Functional skin microenvironments unravelled through single cell and spatial transcriptomic analysis

Bayanne Olabi

BMedSci (Hons), MBChB, MRCP (UK), FHEA, MSc (Med Ed)

Doctorate of Philosophy (PhD)

Biosciences Institute

Newcastle University

Supervisors: Muzlifah Haniffa,

Sarah Teichmann and Ruth Murphy

September 2023

Abstract

Skin tissue comprises distinct microanatomical regions across the epidermis, dermis and hair follicles that are pivotal for its physiological roles, including barrier protection and immunological surveillance. Within these regions, distinct cellular interactions occur between diverse cell types to govern skin structure and function across the lifespan, which can become perturbed in disease. Novel insights have recently been gained into skin pathophysiology from single cell RNA-sequencing (scRNA-seq) of adult healthy, psoriasis and eczema skin. However, single-cell data derived from dissociated tissue loses valuable information about gene expression *in situ*, which informs cellular communication and identity.

The aims of this project were therefore to use spatial transcriptomics to profile adult skin samples from healthy, lesional and non-lesional psoriasis and eczema skin, and leverage a reference scRNA-seq dataset to map cell states across tissue sections. Comparative analyses of spatial data between healthy and disease conditions demonstrated widespread changes in the tissue location of immune and non-immune cells and revealed functional microanatomical cellular niches that become perturbed in psoriasis and eczema.

Based on these findings, the concept of functional skin cellular microenvironments was then explored during skin and hair morphogenesis by annotating and spatially mapping scRNA-seq data from first and second trimester prenatal skin. The analyses revealed distinct cellular microenvironments and intercellular crosstalk supporting skin and hair follicle development. Integrating adult hair follicle and hair-bearing skin organoid datasets contextualised the differences with *in vivo* prenatal skin and provided support for congenital skin and hair disease modelling *in vitro*.

Together, this spatially-resolved atlas of human skin in adulthood and during prenatal skin development has provided unprecedented microanatomical detail and insight into functional skin cell microenvironments. This has future applications for enhancing our understanding of skin and hair disease pathogenesis, identifying novel therapeutic targets, and facilitating tissue engineering for hair follicle and skin regeneration and transplantation.

Acknowledgements

Institute who led the spatial genomic analyses and with whom I worked closely to interpret the data and help develop computational methods to analyse spatial transcriptomic data from skin samples. I would also like to thank Dr Hudaa Gopee, Dr Rachel Botting, Dr Chloe Admane and Dr Ni Huang, with whom I worked closely on prenatal skin data analysis and interpretation, and Dr Simone Webb, Dr Rachel Botting and Dr Emily Stephenson for their valuable feedback on sections of my first thesis draft.

Finally, words cannot express my gratitude to my husband and family, for their unwavering support and kindness.

Candidate declaration

This PhD thesis is my own work and has correctly acknowledged the work of others. No part of the material presented here has been previously submitted for a degree or other qualification in this or any other University or other institute of learning.

The study presented in this thesis represents an interdisciplinary collaborative scientific effort between researchers across multiple institutions and focuses on the research areas to which I had the most intellectual input. The contributions made by individuals towards work presented in this thesis are detailed below:

I recruited patients for this study, collected healthy and psoriasis skin samples and generated Visium spatial transcriptomic data from these samples. Prof Graham Ogg in Oxford provided eczema samples for this study and the corresponding Visium data was generated by the Cellular Generation and Phenotyping (CGaP) team at the Wellcome Sanger Institute whilst I was on maternity leave. I annotated the histological skin tissue regions for the whole dataset and Pavel Mazin carried out the computational analyses of the adult skin spatial transcriptomic data. Droplet-based prenatal skin scRNA-seq datasets were generated by Rachel Botting, Emily Stephenson, Justin Englebert and James Fletcher. Annotation of the prenatal skin dataset was carried out by Rachel Botting, Ni Huang and I using known gene markers and literature searching. I performed clustering and differential gene expression analyses of the data. Integration of the prenatal data with organoid, adult hair graft and adult interfollicular skin scRNA-seq datasets was carried out by Ni Huang. Differential abundance testing, logistic regression modelling, cell-cell interaction analysis and trajectory analysis was also carried out by Ni Huang. Gene set enrichment analysis was carried out by Hudaa Gopee. The prenatal skin spatial analysis was carried out by Pavel Mazin. I provided biological interpretation of the data and results, along with Hudaa Gopee and Rachel Botting.

List of publications

During my PhD studies, I have contributed to several research manuscripts, the details of which are listed below. Articles that have been published to date are presented in Appendices A-E.

1. **Olabi B**, Dubois A, Gopee N, Haniffa M. Defining the skin cellular community using single-cell genomics to advance precision medicine. *Journal of Investigative Dermatology* 2021; 141(2): 255-264. doi: 10.1016/j.jid.2020.05.104. PMID: 32713511.

I co-wrote the first draft of this review article, which involved carrying out literature reviews of relevant research in the field of single cell genomics applied to dermatology, and I designed and created the manuscript figures.

2. Reynolds G, Vegh P, Fletcher J, Poyner EFM, Stephenson E, Goh I, Botting RA, Huang N, **Olabi B**, Dubois A, Dixon D, Green K, Maunder D, Engelbert J, Efremova M, Polański K, Jardine L, Jones C, Ness T, Horsfall D, McGrath J, Carey C, Popescu DM, Webb S, Wang XN, Sayer B, Park JE, Negri VA, Belokhvostova D, Lynch MD, McDonald D, Filby A, Hagai T, Meyer KB, Husain A, Coxhead J, Vento-Tormo R, Behjati S, Lisgo S, Villani AC, Bacardit J, Jones PH, O'Toole EA, Ogg GS, Rajan N, Reynolds NJ, Teichmann SA, Watt FM, Haniffa M. Developmental cell programs are co-opted in inflammatory skin disease. *Science* 2021; 371(6527): eaba6500. doi: 10.1126/science.aba6500. PMID: 33479125.

I contributed to the data interpretation of this single cell study from which my spatial genomic project arose and wrote the first draft of the Research Article Summary for *Science*, which featured in the journal's print copy and involved distilling the key findings of the study. I also helped to design and create several figure schematics for the manuscript, including the graphical abstract for the Research Article Summary. I helped to draft the manuscript and address comments from reviewers.

3. Stephenson E, Reynolds G, Botting RA, Calero-Nieto FJ, Morgan MD, Tuong ZK, Bach K, Sungnak W, Worlock KB, Yoshida M, Kumasaka N, Kania K, Engelbert J, **Olabi B**, Spegarova JS, Wilson NK, Mende N, Jardine L, Gardner LCS, Goh I, Horsfall D, McGrath J, Webb S, Mather MW, Lindeboom RGH, Dann E, Huang N, Polanski K, Prigmore E, Gothe F, Scott J, Payne RP, Baker KF, Hanrath AT, Schim van der Loeff ICD, Barr AS, Sanchez-Gonzalez A, Bergamaschi L, Mescia F, Barnes JL, Kilich E, de Wilton A, Saigal A, Saleh A, Janes SM, Smith CM, Gopee N, Wilson C, Coupland P, Coxhead JM, Kiselev VY, van Dongen S, Bacardit J, King HW; Cambridge Institute of Therapeutic Immunology and Infectious Disease-National Institute of Health Research (CITIID-NIHR) COVID-19 BioResource Collaboration; Rostron AJ, Simpson AJ, Hambleton S, Laurenti E, Lyons PA, Meyer KB, Nikolić MZ, Duncan CJA, Smith KGC, Teichmann SA, Clatworthy MR, Marioni JC, Göttgens B, Haniffa M. Single-cell multi-omics analysis of the immune response in COVID-19. *Nature Medicine* 2021; 27(5): 904-916. doi: 10.1038/s41591-021-01329-2. PMID: 33879890.

I contributed to the data interpretation of this research article and created the final main figure of the article that summarised the key findings from the study. I also helped with drafting the manuscript.

4. Haniffa M, Taylor D, Linnarsson S, Aronow BJ, Bader GD, Barker RA, Camara PG, Camp JG, Chédotal A, Copp A, Etchevers HC, Giacobini P, Göttgens B, Guo G, Hupalowska A, James KR, Kirby E, Kriegstein A, Lundeberg J, Marioni JC, Meyer KB, Niakan KK, Nilsson M, **Olabi B**, Pe'er D, Regev A, Rood J, Rozenblatt-Rosen O, Satija R, Teichmann SA, Treutlein B, Vento-Tormo R, Webb S; Human Cell Atlas Developmental Biological Network. A roadmap for the Human Developmental Cell Atlas. *Nature* 2021; 597(7875): 196-205. doi: 10.1038/s41586-021-03620-1. PMID: 34497388.

I co-wrote the first draft of this Perspective article in *Nature* and contributed to re-drafting and editing the manuscript before submission. I carried out literature-based reviews during the preparation of this manuscript and contributed to the supplementary materials, including the figures and the tables.

5. Jardine L, Webb S, Goh I, Quiroga Londoño M, Reynolds G, Mather M, **Olabi B**, Stephenson E, Botting RA, Horsfall D, Engelbert J, Maunder D, Mende N, Murnane C, Dann E, McGrath J, King H, Kucinski I, Queen R, Carey CD, Shrubsole C, Poyner E, Acres M, Jones C, Ness T, Coulthard R, Elliott N, O'Byrne S, Haltalli MLR, Lawrence JE, Lisgo S, Balogh P, Meyer KB, Prigmore E, Ambridge K, Jain MS, Efremova M, Pickard K, Creasey T, Bacardit J, Henderson D, Coxhead J, Filby A, Hussain R, Dixon D, McDonald D, Popescu DM, Kowalczyk MS, Li B, Ashenberg O, Tabaka M, Dionne D, Tickle TL, Slyper M, Rozenblatt-Rosen O, Regev A, Behjati S, Laurenti E, Wilson NK, Roy A, Göttgens B, Roberts I, Teichmann SA, Haniffa M. Blood and immune development in human fetal bone marrow and Down syndrome. *Nature* 2021; 598(7880): 327-331. doi: 10.1038/s41586-021-03929-x. PMID: 34588693.

I carried out analyses of immunofluorescence microscopy images and contributed to the interpretation of the data generated in this research project. I also compiled the main and supplementary figures for presentation in the manuscript and created several figure schematics.

6. Goh I, Botting RA, Rose A, Webb S, Engelbert J, Gitton Y, Stephenson E, Londoño MQ, Mather M, Mende N, Imaz-Rosshandler I, Horsfall D, Basurto-Lozada D, Chipampe N, Rook V, Mazin P, Vijayabaskar MS, Hannah R, Gambardella L, Green K, Ballereau S, Inoue M, Tuck L, Lorenzi V, Kwakwa K, Alsinet C, **Olabi B**, Miah M, Admane C, Popescu D, Acres M, Dixon D, Coulthard R, Lisgo S, Henderson DJ, Dann E, Suo C, Kinston SJ, Park J, Polanski K, Dongen S, Meyer KB, Bruijn M, Palis J, Behjati S, Laurenti E, Wilson NK, Vento-Tormo R, Chédotal A, Bayraktar O, Roberts I, Jardine L, Göttgens B, Teichmann SA, Haniffa M. Multi-organ functions of yolk sac during human early development. Accepted for publication in *Science,* 2023.

I contributed to drafting the manuscript and interpretation of the single cell data generated in this project.

7. Li R, Strobl J, Poyner E, Gardner L, **Olabi B**, Stephenson E, Coulthard R, Botting R, Prigmore E, Gopee N, Engelbert J, Goh I, Tun W, Reynolds G, Gambardella L,

Admane C, Hu S, Liu F, Ballereau S, Horsfall D, Basurto-Lozada D, Bacon C, Weatherhead S, Wang Y, Bai F, Brunner P, Teichmann S, Haniffa M. A specialised tumour-promoting microenvironment distinguishes cutaneous T cell lymphoma from benign inflammatory skin disease. Under review at *Nature Medicine* 2023.

I carried out analyses of the single cell data, including clustering, annotation, differential gene expression and gene set enrichment analyses, to help reveal biological insights into the pathogenesis of cutaneous T-cell lymphoma and how this compares to healthy skin, psoriasis and eczema. I also helped to draft the manuscript.

8. **Olabi B**, Gopee NH, Huang N, Admane C, Botting RA, Foster AR, Torabi F, Winheim E, Sumanaweera D, Goh I, Miah M, Stephenson E, Tun WM, Moghimi P, Rumney B, He P, Lawrence S, Roberts K, Sidhpura K, Englebert J, Jardine L, Reynolds G, Rose A, Ganier C, Rowe V, Pritchard S, Mulas I, Fletcher J, Popescu D, Poyner E, Dubois A, Filby A, Lisgo S, Barker R, Park J, Vento-Tormo R, Le A, Serdy S, Kim J, Deakin C, Lee J, Nikolova M, Rajan N, Ballereau S, Horsfall D, Basurto Lozada D, O'Toole EA, Treutlein B, Kasper M, Mazin P, Gambardella L, Koehler K, Teichmann S, Haniffa M. A human prenatal skin cell atlas reveals immune cell regulation of skin morphogenesis. Submitted to *Nature*, 2023.

I co-led this interdisciplinary collaborative project and carried out several types of computational analyses of the prenatal skin, adult skin and organoid data, which forms the basis of Chapters 4 and 5 of this thesis. I am co-first author of the study and, with my co-authors, interpreted the data, compiled the figure panels and drafted the manuscript and figure/table legends.

Table of contents

Index of figures

Index of tables

Acronyms and abbreviations

AMP	Antimicrobial peptide
ARCI	Autosomal recessive congenital ichthyosis
ASDC	AXL$^+$SIGLEC6$^+$ dendritic cell
BBKNN	Batch balanced K-nearest neighbours
BMP4	Bone morphogenic protein 4
cDC	Classical dendritic cell
cDNA	Complementary deoxyribonucleic acid
Cq	Quantification cycle
DC	Dendritic cell
DEB	Dystrophic epidermolysis bullosa
DEG	Differentially expressed gene
DEJ	Dermoepidermal junction
EBS	Epidermolysis bullosa simplex
ECM	Extracellular matrix
EKV	Erythrokeratodermia variabilis
ER	Endoplasmic reticulum
ESC	Embryonic stem cell
FACS	Fluorescence-activated cell sorting
FBS	Fetal Bovine Serum
FGF	Fibroblast Growth Factor
FISH	Fluorescence In Situ Hybridisation
GEM	Gel Beads-in-Emulsion
GWAS	Genome-wide associated study
H&E	Haematoxylin And Eosin
Hcl	Hydrochloric acid
HDBR	Human Developmental Biology Resource
HTA	Human Tissue Authority
Ig	Immunoglobulin
IGF	Insulin-like growth factor
IFE	Interfollicular epidermis
IFN	Interferon

IL	Interleukin
ILC	Innate lymphoid cell
iPSC	Induced Pluripotent Stem Cell
JEB	Junctional epidermolysis bullosa
KNN	K-nearest neighbour
KOH	Potassium hydroxide
KRT	Keratin
LE	Lymphatic endothelium
LR	Logistic regression
MRC	Medical Research Council
mRNA	Messenger Ribonucleic Acid
MSC	Mesenchymal stem cell
NK	Natural killer
ng	Nanogram
OCT	Optimal Cutting Temperature
PBMC	Peripheral Blood Mononuclear Cells
PBS	Phosphate-Buffered Saline
PCs	Principal Components
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PCW	Post Conception Weeks
pDC	Plasmacytoid dendritic cell
PIL	Patient Information Leaflet
Poly(A)	Poly adenine
Poly(dT)	Poly deoxythymine
PPAR-δ	Peroxisome proliferator-activated receptor-delta
QC	Quality Control
qPCR	Quantitative Polymerase Chain Reaction
REC	Research Ethics Committee
RNA	Ribonucleic acid
RT	Reverse Transcription
scRNA-seq	Single-cell RNA sequencing
SDS	Sodium dodecyl sulphate

SMC	Smooth muscle cell
smFISH	Single Molecule Fluorescence *In Situ* Hybridisation
snRNA-seq	Single nuclei RNA sequencing
SSC	Saline-Sodium Citrate
TGFβ	Transforming Growth Factor β
Tc	Cytotoxic T cells
Th	T helper cell
TNF	Tumour necrosis factor
Treg	Regulatory T cell
TRM	Tissue Resident Macrophages
TSO	Template Switch Oligo
µl	Microlitre
UMAP	Uniform Manifold Approximation and Projection
UMI	Unique Molecular Identifier
UV	Ultraviolet
VEGF	Vascular endothelial growth factor

Chapter 1: Background and literature overview

1.1. The structure and function of human skin

1.1.1. Skin anatomy

Human skin is a dynamic organ that comprises of multiple cell types organised in spatially distinct skin layers and structures. The epidermis forms the superficial layer of the skin and dermis lies deep to the epidermis, separated by a basement membrane. The interface between the epidermis and dermis is known as the dermoepidermal junction (DEJ). The skin has a neurovascular supply, with nerve endings and blood vessels located within the dermal skin layer. The appendages of the skin include the hair follicles, sebaceous glands and eccrine sweat glands (Figure 1).
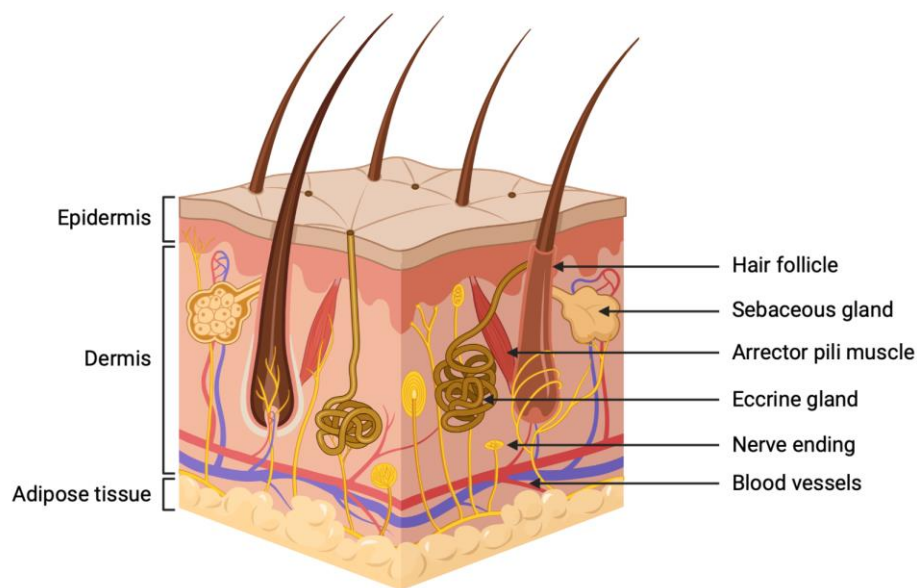
Figure 1: Anatomy of human skin. Schematic representation of human skin in cross-section, showing distinct layers of the skin (epidermis and dermis), bloods vessels and nerve endings located in the dermis and the microanatomical organisation of skin appendages, including hair follicles, sebaceous glands and eccrine glands. Image created with BioRender.com.

The epidermis is a stratified (or multi-layered) epithelium that is composed predominantly of keratinocytes and varies in thickness depending on anatomical site, from less than 0.1mm in eyelid skin to approximately 1.5mm on the palms and soles (Kolarsick, Kolarsick and Goodwin, 2011). The epidermis is arranged into four strata that represent different stages of keratinocyte differentiation. The stratum basale, also known as the basal layer, is made up of mitotically active undifferentiated cells that are continually producing keratinocytes and are attached to the basement membrane by hemidesmosomes (Te Molder, de Pereda and Sonnenberg, 2021). Basal keratinocytes then migrate superficially to form the stratum spinosum, with desmosomes mediating strong cell-cell adhesions between epidermal keratinocytes to maintain tissue integrity (Müller, Hatzfeld and Keil, 2021). Keratinocytes begin to flatten as they differentiate further and migrate superficially to form the stratum granulosum, then become anucleate as they form the uppermost stratum corneum prior to being shed (Norlén, 2006).

In addition to keratinocytes, the epidermis contains other important cell types, including melanocytes and Langerhans cells. Melanocytes are located in the basal layer of the epidermis and produce melanin that is transferred to keratinocytes in melanosomes; this process is responsible for skin pigmentation (Cichorek et al., 2013). Langerhans cells are antigen-presenting cells that are predominantly located in the stratum spinosum (Deckers, Hammad and Hoste, 2018) and determine adaptive immune responses after encountering foreign substances within their local microenvironment (Clayton et al., 2017).

The dermis of the skin is relatively cell sparse compared to the epidermis; the dermal extracellular matrix (ECM) is made up of a fibrous component containing collagen and elastin fibres and an amorphous component containing glycosaminoglycans (Marks and Miller, 2013). Cells that populate the dermis include fibroblasts that produce collagen and other ECM components, endothelial cells and associated mural cells that form the vascular and lymphatic networks, nerve cells and various innate and adaptive immune cells, including macrophages, lymphocytes and mast cells (Nguyen and Soulika, 2019).

The appendages of human skin are specialised structures that originate from the epidermis and include hair follicles and sebaceous glands that together form the pilosebaceous unit in association with arrector pili smooth muscle. The hair shaft is the part of the hair follicle unit that eventually protrudes from the skin and becomes visible, and it is made up of three

concentric layers: the inner medulla, the cortex and the outer cuticle. The medulla is only seen in larger thick hairs. The cortex consists of closely packed cortical cells containing keratin filaments (Harland et al., 2014) and the outer cuticle layer of the hair shaft is a thin layer that functions to protect the cortex (Mirmirani, Huang and Price, 2011). The hair follicle is lined by an inner root sheath, which is composed of terminally differentiated keratinocytes that anchors the hair shaft to the hair follicle and moves with the hair shaft during hair growth, and an outer root sheath that forms a continuous layer with the basal cell layer of the epidermis at the superficial end of the hair follicle and merges distally with the hair bulb (Schneider, Schmidt-Ullrich and Paus, 2009). Hair grows in cycles that involve distinct phases of anagen (active hair growth phase), catagen (transitional phase in which hair growth ceases and the lower cycling portion of the hair follicle regresses) and telogen (where follicles lie dormant in a resting phase) (Alonso and Fuchs, 2006; Hawkshaw et al., 2020).

The skin appendages also include eccrine sweat ducts and glands that arise from the epidermis and are located within the dermis. Larger apocrine sweat glands are found in the axillary and groin regions. The density of other skin appendageal structures also varies depending on anatomical site, with a high density of hair follicles in scalp skin and an absence of hair follicles in palmar and plantar skin, and a high density of eccrine sweat glands in the axillae, palms and soles (Wilke et al., 2007).

1.1.2. Skin physiology

Owing to the complex structure of human skin and diverse cell types present, skin tissue has several functions that are important in protection from the external environment by mechanical and immunological mechanisms, as well in maintaining various homeostatic processes. The stratified epidermal skin layer is impermeable to water and functions as a barrier to external insults, including physical, thermal and chemical trauma. Skin is also protected from ultraviolet (UV) damage by the superficial layers of the epidermis as well as by melanin, which both scatters and absorbs UV radiation to prevent penetration through the epidermis (Brenner and Hearing, 2008).

The skin employs both barrier and immunological mechanisms to protect against different types of microorganisms, including bacteria, viruses and fungi. Upon exposure to and recognition of pathogens, an inflammatory cascade is initiated that involves cytokine production by dendritic cells, recruitment of immune cells including neutrophils and macrophages, and the production of antimicrobial peptides (AMPs) by immune cells and keratinocytes. AMPs include defensins and cathelicidins, and are important effector molecules that function in pathogen defence and skin healing (Coates, Blanchard and MacLeod, 2018). Other antimicrobial skin defences include the production of sebum, which repels pathogens, and the colonisation of commensal bacteria and yeasts on the surface of the skin, which can promote the expression of AMPs (Naik et al., 2015) and enhance the development of cutaneous T cells (Laborel-Préneron et al., 2015; Naik et al., 2012).

Human skin helps to regulate body temperature, whereby dilatation of cutaneous blood vessels and production of sweat allows the loss of heat in warm environments or during exercise, and cutaneous vasoconstriction conserves heat to protect from hypothermia (Romanovsky, 2014). The skin is also a site of vitamin D synthesis following exposure to sunlight. Keratinocytes convert 7-dehydrocholesterol into vitamin D and then converts vitamin D into its active form, which is called 1,25 dihydroxy vitamin D (Bikle, 2012) and has a key role in regulating calcium homeostasis (Khammissa et al., 2018). A further important function of the skin is sensation, with specialised afferent neurones able to distinguish touch, heat, cold, vibration, itch and pain, accounting for an additional protective mechanism from external injury.

1.2. Applications of genomic technologies in skin research

1.2.1. Single-cell RNA sequencing

Our understanding of cellular gene expression and regulation during health and disease is enabled by the ability to profile the transcriptome, or messenger ribonucleic acid (RNA) transcripts, of cells. Early transcriptomic studies of individual cells began with the development of single cell quantitative polymerase chain reaction (qPCR) technology (Kurimoto et al., 2007; Subkhankulova, Gilchrist and Livesey, 2008) and whole

transcriptome analysis using microarrays (Kurimoto et al., 2007; Subkhankulova, Gilchrist and Livesey, 2008). However, these methods were limited by the requirement of *a priori* knowledge about sequences that were under investigation and difficulties in accurately quantifying the expression of lowly and highly expressed transcripts (Kukurba and Montgomery, 2015).

RNA sequencing was then adapted for use on individual cells, with the first report of transcriptomes profiled by single cell RNA sequencing (scRNA-seq) published in 2009 (Tang et al., 2009), only two years after the establishment of bulk RNA sequencing (Kolodziejczyk et al., 2015). The primary drawback of bulk RNA sequencing is that cellular heterogeneity within tissues is masked by measuring the average gene expression across multiple cells. Transcripts cannot be attributed to individual cells, which may lead to specific cell states and rare cell populations being undetected (Li and Wang, 2021). In recent years, scRNA-seq has emerged as a highly robust and sensitive approach for unbiased and high-throughput sequencing of single cell transcriptomes (Zheng et al., 2017) with the ability to dissect cellular heterogeneity in many biological contexts, such as embryonic development (Vento-Tormo et al., 2018; Popescu et al., 2019), immunology and immunotherapy (Szabo et al., 2019; Stubbington et al., 2017), tumour microenvironments and oncogenic processes (Tirosh et al., 2016; Durante et al., 2020) and infections and autoimmune diseases (Reid et al., 2018; Steuerman et al., 2018; Jin et al., 2017). The Human Cell Atlas project has arisen from the potential of scRNA-seq technology, which has become a global collaborative effort to profile the unique 37 trillion cells that make up the human body, helping to transform our understanding of fundamental biological processes and of diagnosing and treating disease (Rozenblatt-Rosen et al., 2017).

The experimental approach of scRNA-seq involves multiple steps that are depicted below (Figure 2).
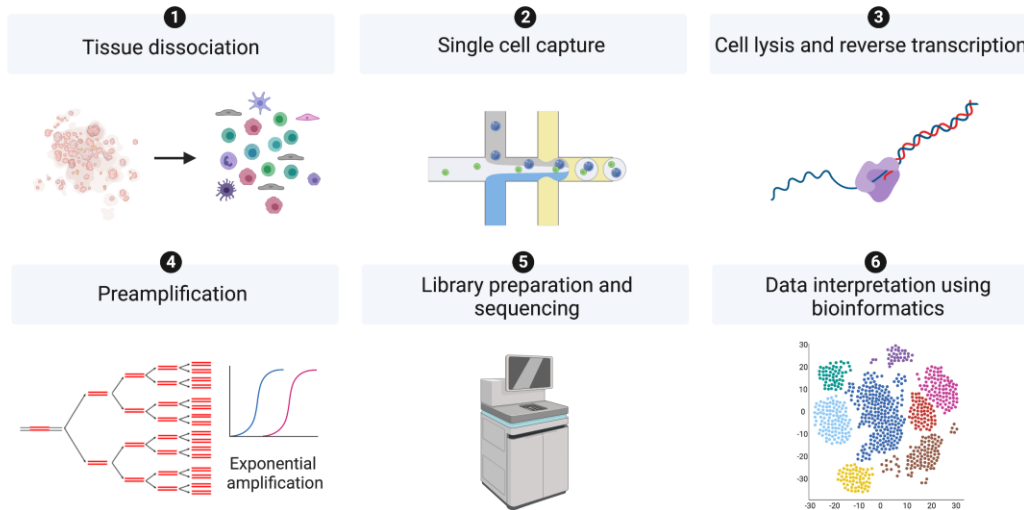
Figure 2: Single cell RNA sequencing experimental workflow. Tissue samples are dissociated to isolate single cells that are then captured, barcoded and lysed. Reverse transcription of mRNA molecules released from cells yields cDNA, which is amplified and sequenced in order to generate single cell data. Image created with BioRender.com.

The first step in the scRNA-seq workflow involves processing tissue samples to yield single cell suspensions, with various mechanical and enzymatic based protocols having been optimised and utilised on different tissue types (Vieira Braga and Miragaia, 2019). Single cells are then isolated, which can be achieved using several methods, including microfluidic- and droplet-based techniques. Following single cell lysis, reverse transcription is performed in order to select for mRNA transcripts by using poly(T) primers and to obtain complementary deoxyribonucleic acid (cDNA) (Kolodziejczyk et al., 2015). The minute quantities of cDNA resulting from this reaction is then amplified, often by polymerase chain reaction (PCR), in order to generate sufficient material for library preparation and sequencing. Computational approaches to examine the sequencing data include quality control and filtering low quality data, clustering and annotation of cell states, and downstream analyses of the dataset such as differential gene expression, differential abundance, inferred cell-cell communication and trajectory inference (Luecken and Theis, 2019) (Figure 2).

1.2.2. Spatial transcriptomics

Although scRNA-seq methodologies have enabled high dimensional analysis of human tissues, data generation relies primarily on tissue digestion and dissociation in order to obtain individual single cells in suspension, which negates valuable spatial information of cells within their microanatomical tissue context. Several technologies for high-dimensional spatially-resolved tissue analysis are available, which rely on targeted RNA hybridisation, such as sequential or multiplexed Fluorescence In Situ Hybridisation (Seq-FISH or MER-FISH) or *in situ* RNA sequencing (Shah et al., 2017; Xia et al., 2019; Gyllborg and Nilsson, 2019). These spatial technologies are able to assess 100-1000 parameters simultaneously at single cell resolution by using combinatorial labelling and sequential imaging, however, they are reliant on having highly-skilled operators and not all are commercially available as robust platforms. These approaches, however, do measure a greater number of parameters than protein-antibody labelling technologies such as imaging mass cytometry, where only up to 100 antigens can be measured (Kuett et al., 2022).

Spatial transcriptomic technologies have been developed in recent years to address the loss of microanatomical context in suspension data, which allows the generation of two-dimensional positional RNA-sequencing data from histological tissue sections (Ståhl et al., 2016). Although the data derived is not of single cell resolution, it is unbiased, enabling the whole transcriptome within the tissue overlying each spot of the capture area to be profiled. Over the last 5 years since the technology was first developed, researchers globally have adopted these methods to interrogate and explore various tissue types and disease states (Williams et al., 2020; Ji et al., 2020; Maniatis, Petrescu and Phatnani, 2021). Furthermore, spatial transcriptomic technologies continue to evolve to be able to profile tissue sections with increasing resolution (Vickovic et al., 2019; Liu et al., 2020; Cho et al., 2021b; Fu et al., 2021; Chen et al., 2021a).

A widely adopted spatial transcriptomic technology to profile tissue sections is called Visium, which has been made commercially available by 10x Genomics and involves several broad experimental processing steps that are depicted below (Figure 3).
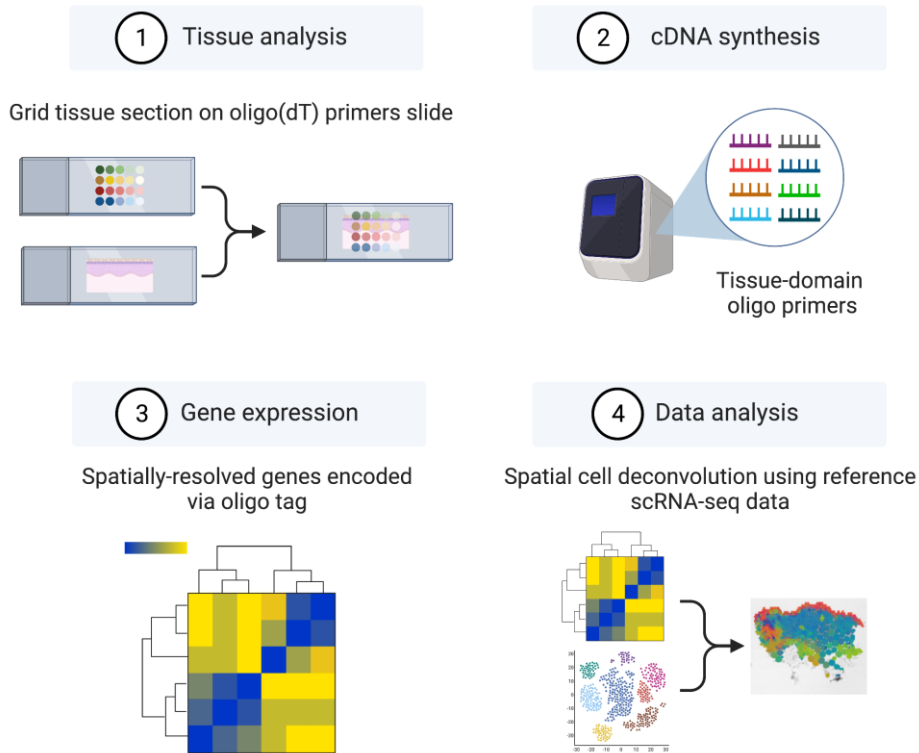
Figure 3: Visium spatial transcriptomic experimental workflow. Tissue sections placed on a slide containing arrayed oligonucleotide primers with positional barcodes are permeabilised to release mRNA. Captured mRNA is synthesised into cDNA, amplified and sequenced to generate spatially-resolved data for downstream analysis. Image created with BioRender.com.

The experimental approach taken during a Visium spatial transcriptomics experiment first involves cryosectioning tissue samples and placing the tissue sections on slides that are engineered with spots containing arrayed poly(dT) oligonucleotide primers with positional barcodes. The Visium spots are approximately 55 μm in size and the distance from the centre of each spot is 100 μm. Tissue sections are then fixed on the slide, stained using haematoxylin and eosin (H&E) and imaged in order to provide histological context for the spatial gene expression data. The same tissue section is then enzymatically permeabilised in order to release mRNA from the tissue sections, which is then captured by the oligonucleotide primers on the slide via the poly(A) tail of mRNA molecules. Reverse transcription then produces spatially barcoded cDNA from captured mRNA, which is then extracted, amplified and used to produce libraries for sequencing. The data generated using this technology therefore allows spatial visualisation of the transcriptome by mapping

spatially-resolved genes back to specific locations on tissue sections (Piñeiro, Houser and Ji, 2022). As Visium data is not of single cell resolution, spatial transcriptomic data can also be combined with scRNA-seq data to map annotated cell states onto tissue sections (Figure 3).

1.3. Investigating skin diseases using genomic technologies

1.3.1. Psoriasis

Psoriasis is a chronic immune-mediated inflammatory skin disease that affects approximately 125 million people worldwide, with an overall prevalence of about 2-3% (Parisi et al., 2013). The clinical manifestations of psoriasis include inflamed thickened red scaly skin plaques, causing skin itching, burning and pain, which can involve the entire body surface (Krueger et al., 2000). This is associated with dystrophy of the nails and joint pain and swelling in a subset of patients (De Rie, Goedkoop and Bos, 2004). The onset of psoriasis is typically between the ages of 15 and 35, however it can affect infants and children, and the chronic natural history of psoriasis necessitates lifelong management and support, which is associated with significant medical resource utilisation, healthcare costs and socioeconomic loss from the impact on patients' livelihood (Fowler et al., 2008). Furthermore, psoriasis has a profound impact on the quality of life of patients (Salman et al., 2018) and is associated with multiple comorbidities, including psoriatic arthritis, cardiovascular disease, type 2 diabetes, anxiety and depression (Gottlieb, Chao and Dann, 2008).

The pathogenesis of psoriasis involves interactions between immune and non-immune cell types that generate inflammatory circuits (Figure 4) and lead to the onset, persistence and progression of the condition in genetically susceptible individuals (Albanesi et al., 2018).
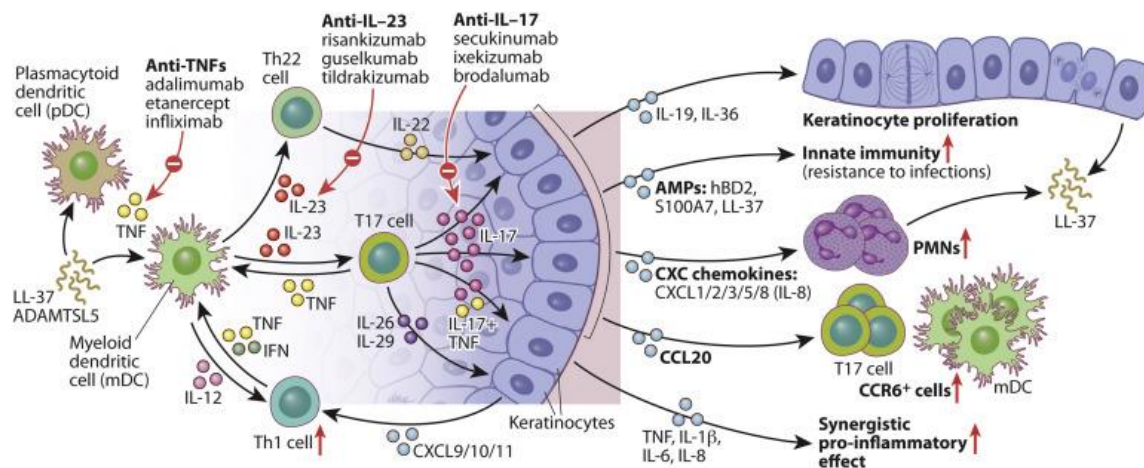
Figure 4: Psoriasis pathogenesis. The schematic shows the immune-mediated effects of T cells and dendritic cells on epidermal keratinocytes, and therapeutic targets for biologic treatments (Hawkes, Chan and Krueger, 2017). TNF = tumour necrosis factor, IL = interleukin, IFN = interferon, Th = T helper cell, AMP = antimicrobial peptide, PMN = polymorphonuclear neutrophil.

Lesions of psoriasis are characterised by dense infiltrates of T cells and dendritic cells that produce pro-inflammatory cytokines, such as interleukin (IL)-17 and IL-23, which cause an induction of keratinocyte proliferation, recruitment of epidermal neutrophils (known as Munro's microabscesses) and dendritic cells, upregulation of the innate immune response via cationic AMPs (e.g. S100A7 and hBD2) and an induced transcription of proinflammatory genes such as *IL6* and *IL8* that, synergistically with *TNF*, sustain the inflammatory processes in psoriasis (Figure 4). Furthermore, increased vascularity within psoriatic plaques is mediated by angiogenic factors (Mahil, Capon and Barker, 2016).

1.3.2. Eczema

The ongoing Global Disease Burden study has identified eczema as one of the most prevalent diseases globally (Hay et al., 2014). Furthermore, a study of disability adjusted life years (DALYs) worldwide revealed that eczema represents a greater disease burden than diabetes or tuberculosis (Hay et al., 2017). In the United Kingdom, eczema affects over 10% of children and approximately 4% of adults, with approximately half of childhood eczema cases persisting into adult life (Leung and Guttman-Yassky, 2017). Robust evidence exists for clinically important associations between eczema and other

allergies and co-morbidities, including asthma, atopic eye disease, allergic rhinitis, food allergies, eosinophilic oesophagitis, neuropsychiatric and cardiovascular disease (Brunner et al., 2017). Furthermore, the disease burden in terms of psychological impact and socioeconomic consequences is substantial, with these allergic disorders costing the NHS over one billion pounds annually (Gupta et al., 2004).

The aetiology of eczema is complex, involving an interplay of genetic and environmental factors (Tanjung et al., 2017; Thomsen et al., 2007). Heterozygote mutation of a single gene called filaggrin (*FLG*), which encodes a skin barrier protein, confers a substantially increased risk of eczema, with an odds ratio of approximately four (Brown and McLean, 2012). Moreover, Genome Wide Association Studies (GWAS) have revealed susceptibility links with innate immunity and T cell function genes for eczema, which are also shared with asthma and allergic rhinitis (Paternoster et al., 2015). These results align with consensus opinion that eczema pathogenesis involves a combination of perturbed skin barrier function, epicutaneous allergen sensitisation and immune dysfunction (Figure 5).
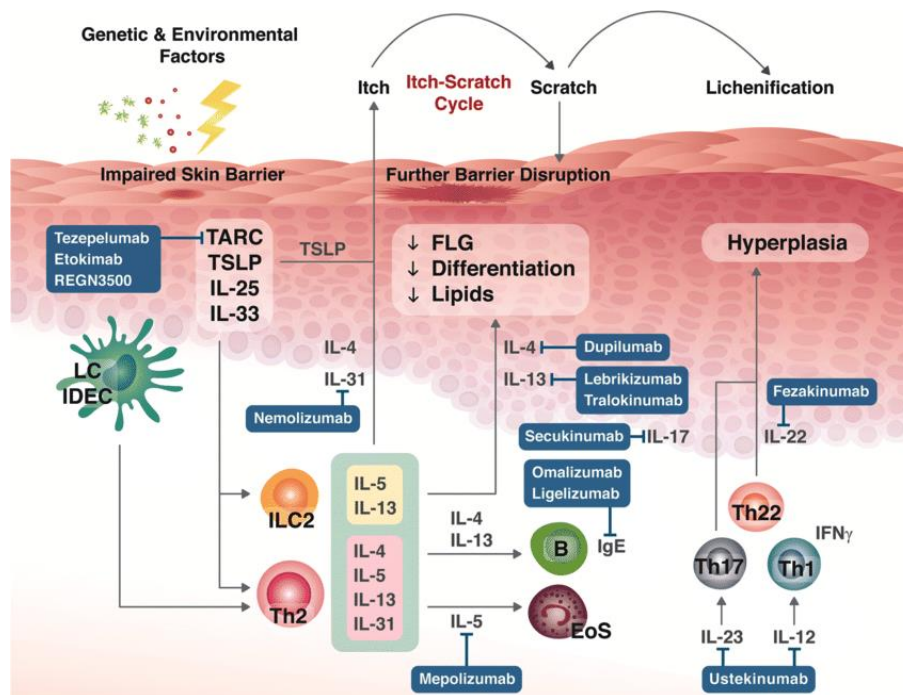
Figure 5: Eczema pathogenesis. Schematic showing the pathogenesis of eczema, including an impaired skin barrier and the immune cells implicated, with therapeutic targets for biologic treatments shown. Adapted from (Chu et al., 2021). B = B cell, EoS = Eosinophil, FLG = Filaggrin, ILC2 = Type II innate lymphoid cell, IDEC = Inflammatory dendritic epidermal cell, IFNγ = Interferon gamma, IL = Interleukin, LC =

Eczema treatment remains challenging because existing options, including topical and systemic steroids and non-specific systemic immunosuppression, do not prevent disease progression or concomitant allergies, are not curative and are not personalised or stratified (Nankervis et al., 2016; Olabi et al., 2020). The adverse safety profile of both topical and systemic treatments also pose a significant barrier to their uptake by patients (Chong and Fonacier, 2016).

1.3.3. Single cell transcriptomic analysis of healthy, psoriasis and eczema skin

Our research group recently published the findings form a systematic analysis of adult healthy trunk skin and lesional and non-lesional skin from patients with psoriasis and eczema using data generated from single-cell RNA-sequencing (scRNA-seq) technology (Reynolds et al., 2021). This dataset included over 500,000 cells and used Fluorescence-activated cell sorting (FACS) to enrich for rare cell states. This study uncovered the cellular and molecular processes occurring within adult skin in health and during disease. Thirty-four cell states were identified in healthy human skin (Figure 6) and two inferred trajectories for keratinocyte differentiation was uncovered. In lesional psoriasis, cytotoxic T cells (Tc) and helper T cells (Th) expressing *IL17A* and *IL17F*, together named Tc17/Th17 cells, were identified. Furthermore, clonally-expanded disease-associated cytotoxic T cells (Tc IL13/IL22 cells) were identified in lesional eczema (Reynolds et al., 2021).
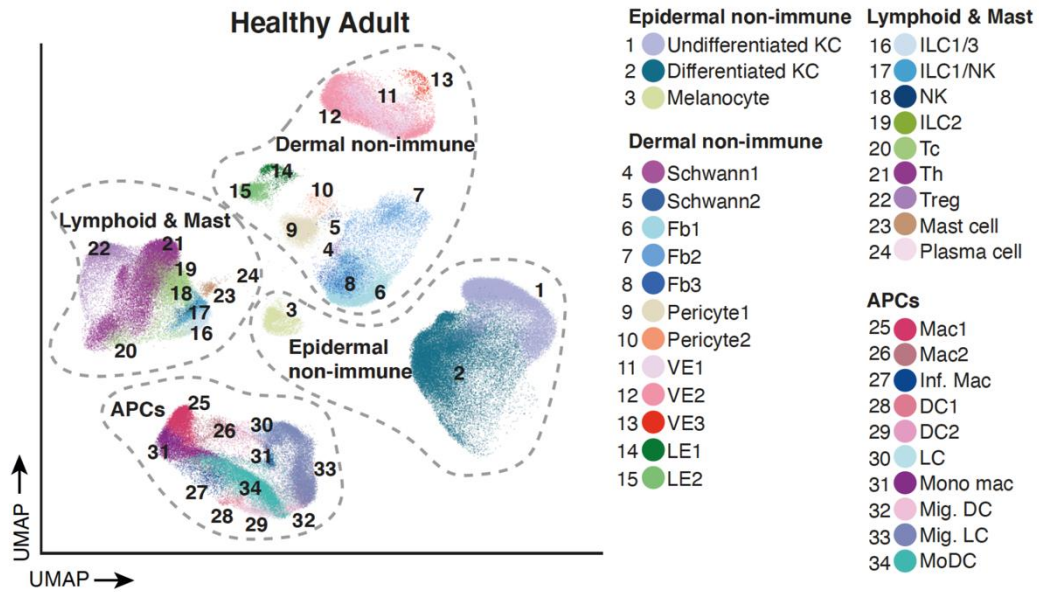
Figure 6: Healthy adult skin cell atlas. UMAP visualisation showing cell types/states in adult healthy skin using scRNA-seq (Reynolds et al., 2021).

The dataset also included scRNA-seq from prenatal skin samples between 7 and 10 post conception weeks (PCW), allowing cell states across development, health and disease to be compared (Reynolds et al., 2021). These analyses revealed the co-optation of prenatal developmental programs in eczema and psoriasis involving a subset of macrophages (Mac2) and vascular endothelial cells (VE3). Furthermore, these cells were expanded in lesional eczema and psoriasis compared with healthy skin and declined following eczema treatment with methotrexate (Figure 7). CellPhoneDB analysis was used to predict interactions between these two cell subsets, identifying that interactions between Mac2 and VE3 are mediated by the *CXCL8* ligand on Mac2 and the *ACKR1* receptor on VE3, and these transcriptional modules are involved in the recruitment of leucocytes into the skin and in angiogenesis (Reynolds et al., 2021). CellPhoneDB analysis was also used to interrogate the interactions between both Mac2 and VE3 with lymphocytes, which demonstrated significant interactions in eczema and psoriasis in comparison to healthy skin (Reynolds et al., 2021).
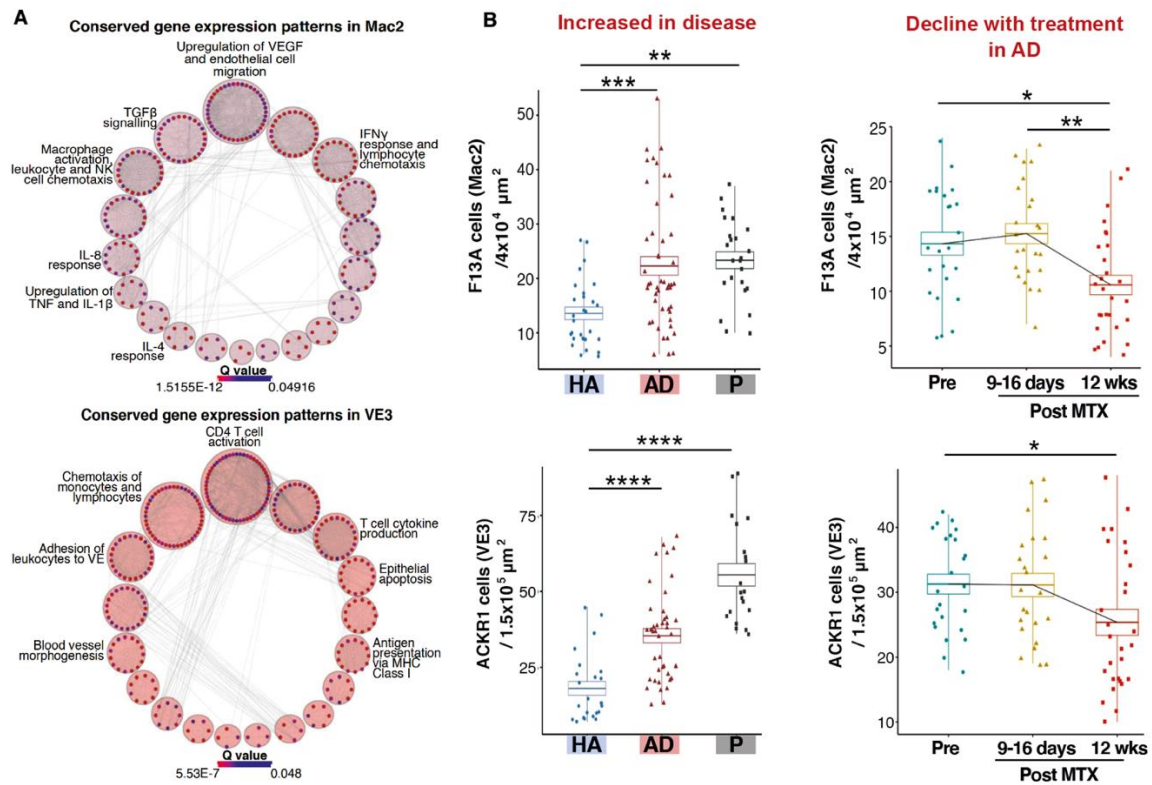
Figure 7: Co-opted Mac2 and VE3 cell states across development and in eczema and psoriasis. Panel A shows network visualisations of conserved pathways (Q value represents gene enrichment score), and panel B shows jitter plots of cell counts in immunostained tissue sections, demonstrating an increase in Mac2 and VE3 in disease and reduction in eczema with methotrexate treatment (Reynolds et al., 2021). HA = healthy adult, AD = atopic dermatitis (syn. eczema), P = psoriasis, MTX = methotrexate.

Spatially contextualising these scRNA-seq data from healthy and diseased skin will provide valuable information related to disease pathogenesis *in situ*. In the context of eczema and psoriasis, the combined approach will allow these cutting-edge technologies to inform dermatopathological microscopic visuospatial examination of cellular disorganisation from skin biopsies.

1.4. Investigating skin and hair follicle development using genomic technologies

De novo human skin and hair follicle formation during prenatal development results from a highly-coordinated series of cellular interactions in spatially distinct tissue regions (Balañá, Charreau and Leirós, 2015). Skin organogenesis begins during embryonic development after gastrulation from two primary germ layers. The epidermis, melanocytes and neuronal cells arise from ectodermal differentiation. The dermis, endothelial and mural cells differentiate from embryonic mesoderm across most anatomical sites, except the facial and cranial skin, where dermal cells arise from ectoderm-derived cranial neural crest cells (Noden and Trainor, 2005; Hu et al., 2018). The epidermis is initially composed of a single layer of ectodermal cells (Coolen et al., 2010). By 4 PCW, two layers can be observed: a basal cell layer and an outer layer known as the periderm that represents the first permeability barrier (King, Balaji and Keswani, 2013). Cells from the periderm are shed into the amniotic fluid during the second trimester when the basal layer begins stratification (Hardman et al., 1999). The process of stratification leads to a complex multi-layered epidermal barrier (Liu, Zhang and Duan, 2013; Damen et al., 2021).

The skin appendages, including hair follicles and sebaceous glands, form in a cephalo-caudal direction during prenatal life (Muller et al., 1991). Hair follicle morphogenesis is initiated by the interaction between epidermal placodes (focal sites of epidermal layer thickening) and dermal condensates (aggregates of dermal fibroblasts). With these interactions, the prenatal hair follicle develops from the epidermal placode that penetrates into the dermis around 11-14 PCW (King, Balaji and Keswani, 2013; Muller et al., 1991). Subsequently, the keratinised hair shaft forms within the centre of the hair follicle, which is surrounded by an epidermal hair sheath and a dermal root sheath (Hu et al., 2018). Sebaceous glands start forming from around 16 PCW, and prenatal hair that protrudes from the skin is observed around 18 PCW (Muller et al., 1991).

There is, however, a paucity of information about the precise cellular composition of human prenatal skin over these developmental periods and whether cells interact in functional microanatomical niches that support skin morphogenesis.

1.5. Applications of organoid models in skin research

In recent years, significant advancements have been made in organoid engineering and its application to biological and clinical research. Organoids are three-dimensional tissue structures cultured *in vitro* that are derived from stem cells (Zhao et al., 2022) and have the ability to recapitulate the cellular heterogeneity, tissue microarchitecture and physiology of human organs (Calà et al., 2023). To date, organoids that mimic different organs have successfully been generated, including brain, retinal, gastrointestinal, cardiac, vascular, kidney, liver and lung organoids (Tang et al., 2022). Furthermore, the first hair-bearing skin organoid was reported in 2020, which was derived from embryonic stem cells (ESC) and induced pluripotent stem cells (iPSC) (Lee et al., 2020).

The skin organoid culture involves treating pluripotent stem cells with factors to promote epidermal induction (including bone morphogenic protein 4 (BMP4) and a transforming growth factor β (TGFβ) inhibitor) and factors to co-induce fibroblasts (including fibroblast growth factor (FGF)). After about 50 days of culture, the epithelium stratifies and becomes multi-layered, analogous to the organisation of prenatal skin with basal, suprabasal and peridermal layers. Epithelial stratification is not seen, however, when epidermal induction is carried out without fibroblast co-induction. Then, over a period of 4-5 months, pigmented hair follicles emerge, with sebaceous glands alongside, representing mature pilosebaceous units (Lee et al., 2020) that are observed later in prenatal skin development during the second trimester.

Understanding the similarities and differences between *in vivo* skin and the *in vitro* skin organoid model has the potential to identify molecular mechanisms that can further enhance skin organoid models in future experimental settings, as well as highlight the potential utility of using skin organoid models for disease modelling.

1.6. Hypothesis

I hypothesise that human skin is comprised of distinct microanatomical cellular niches in healthy adult skin that become perturbed in disease, and that specific cellular niches are also observed during prenatal life that support skin and hair follicle development.

1.7. Aims of this study

The aims of this research study were to:

1. Generate and analyse spatial transcriptomic data on adult healthy, psoriasis and eczema skin and understand the microanatomical cellular organisation of adult skin during health and disease.
2. Comprehensively profile prenatal skin tissue at a single cell resolution across the first and second trimester of gestation and to understand the spatial organisation of cells that support skin development.
3. Dissect the cellular crosstalk underlying hair follicle morphogenesis and to determine the extent to which skin organoids can be used to model skin and hair disorders *in vitro*.

Chapter 2: Materials and methods

2.1. Data generation

To address the aims and objectives of this study, data were generated from adult healthy human skin tissue, adult diseased skin affected by psoriasis and eczema, and prenatal first and second trimester human skin tissue. Skin tissue acquisition, processing and experimental approaches are described below to detail how the data were generated for this project.

2.1.1. Adult skin tissue

Adults were recruited for this research study and consented to provide healthy or diseased skin that was used for spatial transcriptomic processing and analysis, detailed below.

2.1.1.1. Patient recruitment and ethical approval

All research ethics committee and regulatory approval were in place for the collection of research samples at Newcastle and for their storage at the Newcastle Dermatology Biobank (REC reference number: 19/NE/0004) (Appendix G).

Adult healthy skin was sampled from normally discarded surplus skin from skin surgery during defect reconstructions. Multiple anatomical sites were sampled depending on the site of skin surgery, including the face, limbs and trunk. Patients were provided with a patient information leaflet (PIL) relevant to donating normally discarded surplus skin for research (Appendix H) and an associated consent form was reviewed and signed by the patient (Appendix I).

Patients with two common inflammatory skin diseases were recruited for this study: psoriasis and eczema. In Newcastle, I recruited patients with chronic plaque psoriasis who were naïve to biologic treatment and who had not had systemic treatment (such as

methotrexate) for at least 4 weeks prior to donating skin. I had also advised patients not to apply topical steroids to the biopsy site (lower back or abdomen) for at least one week before the biopsy was taken. Each patient had two adjacent 6 mm^2 punch biopsies taken under local anaesthesia from lesional and non-lesional skin. Patients were provided with a PIL relevant to donating skin (that is not normally discarded) for use in research (Appendix J) and an associated consent form was reviewed and signed by the patient prior to recruitment (Appendix K).

Patients with stable, chronic eczema were recruited for this study by Professor Graham Ogg, based in Oxford, who runs specialist eczema research clinics. Eczema patients fulfilled the same eligibility criteria as the psoriasis patients, in that they were naïve to biologic treatment and who have not had systemic treatment for at least 4 weeks are biopsied following a wash out of topical steroid to the biopsy site for 1 week. Similarly, lesional and non-lesional eczema skin was sampled from the trunk under local anaesthesia and immediately frozen and embedded in OCT in the clinic rooms.

2.1.1.2. Sample transport, freezing and OCT embedding

Freshly obtained skin biopsy samples were frozen immediately in order to prevent RNA degradation and avoid crystal formation, which can lead to morphological damage to the skin tissue. In Newcastle, after the adult healthy and psoriasis (lesional and non-lesional) skin was obtained, it was transferred to a vial of pre-cooled HypoThermosol® and transferred from theatre to the lab immediately, located in an adjacent building. A bath of isopentane cooled with dry ice was then prepared and a cryomold was labelled to mark the orientation of the tissue. The cryomold was then half-filled with chilled OCT without introducing bubbles. Using forceps, the skin tissue was placed into the OCT with care to avoid crushing the tissue, and any exposed tissue surface was covered with additional OCT. A light microscope was used to ensure there were no bubbles within the cryomould, especially near the tissue, and if required, a 30G needle was used to carefully remove them. The cryomold was then placed in the isopentane bath and left until the OCT was completely frozen. The cryomould with embedded frozen tissue was then transferred on dry ice and either placed in a sealed container at -80°C for storage or to the cryostat for cryosectioning.

2.1.1.3. Visium spatial transcriptomics experiment

Adult skin was collected from 13 healthy donors, 5 patients with psoriasis and 5 patients with eczema. The frozen, OCT-embedded skin tissue was used to generate spatial transcriptomic data using the Visium Spatial Gene Expression platform by 10x Genomics.

2.1.1.3.1. Cryosectioning and tissue section placement

The OCT tissue block, tissue forceps, razor blades, brushes, specimen stages (chucks) and slides were transferred to a OTF5000 cryostat (Bright Instruments, Bedfordshire, UK) and left to equilibrate to the cryostat chamber temperature for 30 minutes before sectioning. The main cryochamber maintains a stable temperature of -20°C and the specimen temperature was set to -18°C.

Following equilibration, the chuck was filled with OCT and the OCT embedded tissue block was then placed on the chuck with the cutting surface facing away from the chuck. The chuck was then placed on the cryobar inside the cryostat chamber to allow the OCT embedded tissue block to adhere to the chuck. The chuck was then installed onto the specimen head of the cryostat and sectioning commenced until tissue was visible. The temperature of the specimen head was adjusted during sectioning if required: if the sections appeared cracked, the specimen head was too cold and if the sections appeared crumpled, the specimen head was too warm.

The skin tissue was sectioned at 15 µm thickness (Ji et al., 2020) and once the desired section was obtained, it was carefully flattened using cryostat brushes to touch the surrounding OCT. It was then transferred to an equilibrated glass slide using forceps and brushes, and once section placement was adequate, a finger was placed on the underside of the slide for a few seconds to allow the section to adhere.

2.1.1.3.2. Assessment of tissue morphology by H&E staining

Prior to loading the skin tissue sections on to the Visium slides (10x Genomics), morphology was assessed by staining tissue loaded on a glass microscopy slide with haematoxylin and eosin (H&E). First, 800ml of Milli-Q water was dispensed into three

separate beakers and 50ml into a centrifuge tube. The eosin mix was then prepared, which comprised of 100µl eosin Y solution and 900ml tris-acetic acid buffer (0.45M, pH 6.0).

The slide was placed on a clean, non-absorbent surface and 500µl of isopropanol was added to uniformly cover the tissue section. Following incubation for 1 minute at room temperature, the reagent was discarded by holding the slide at an angle above a laboratory wipe. Any excess liquid was removed from the under surface of the slide, without touching the tissue sections, and the slide was then left to air dry for approximately 5 minutes. Haematoxylin 1ml was then added to uniformly cover the tissue section, which was incubated at room temperature for 4 minutes. The reagent was again discarded by holding above a laboratory wipe, then the slide was immersed 5 times in the Milli-Q water in the centrifuge tube. The slide was then immersed 15 times in the water in beaker 1 then 15 times in the water in beaker 2. Any excess liquid was then drained and wiped and the slide was then placed back on the clean non-absorbent work surface. Bluing buffer was next added to uniformly cover the tissue section and left to incubate for 2 minutes, after which the slide was immersed 5 times in beaker 2. After draining and wiping any excess liquid, the eosin mix prepared earlier was added and left to incubate for 1 minute. The slide was then immersed 15 times in beaker 3 and left to air dry.

The glass slide containing the tissue section was then viewed using a Brightfield microscope to assess the morphology of the tissue. Microanatomical landmarks were first identified, to ensure that the tissue being cryosectioned included full thickness skin with good representation of the epidermis and dermis. Furthermore, the overall morphology of the tissue was assessed to ensure the skin was well preserved during the freezing and OCT embedding process (for example, air pockets or uneven freezing that may crack or morphologically damage the tissue).

2.1.1.3.3. Visium tissue optimisation procedure

Prior to proceeding with the generation of spatial gene expression libraries, I carried out a tissue optimisation experiment to ascertain the optimum skin tissue enzymatic permeabilisation time. This must be carried out for each tissue type of interest. Too little tissue digestion results in no mRNA release, whereas over-digestion results in tissue and

mRNA degradation. Therefore, the duration of time that the permeabilisation enzyme is left on the skin tissue to maximise mRNA yield was determined.

The tissue optimisation procedure involved sectioning, fixing and staining tissue on capture areas on a Visium Tissue Optimisation slide (10x Genomics). A proprietary permeabilisation enzyme (10x Genomics) was then applied to each tissue section, which releases mRNA that become fixed to capture area oligonucleotides via poly(dT)-poly(A) binding. The subsequent steps of the experiment involved synthesising fluorescent cDNA, which was imaged to identify the maximum fluorescence signal (representing mRNA release from the tissue) with the lowest signal diffusion (representing mRNA leakage). If in the situation where the signal was the same at two time points, the longer permeabilisation time was taken to be optimal.

2.1.1.3.3.1. Permeabilisation and reverse transcription

Skin tissue sections at 15 µm thickness were mounted onto 7 of the 8 capture areas on the slide, leaving the eighth capture area for the positive mRNA control. Following fixation and staining with H&E, the slide was mounted with a coverslip using 70% glycerol and was imaged using a Zeiss AxioImager microscope (Carl Zeiss Microscopy, Jena, Germany) at 20X using Brightfield imaging.

Next, the permeabilisation enzyme was resuspended and equilibrated at 37°C for 15 minutes and the coverslip was removed. The slide was then placed in a slide cassette, which contains a removable gasket that corresponds to the capture areas on the slides, creating leakproof wells for adding reagents (Figure 8).
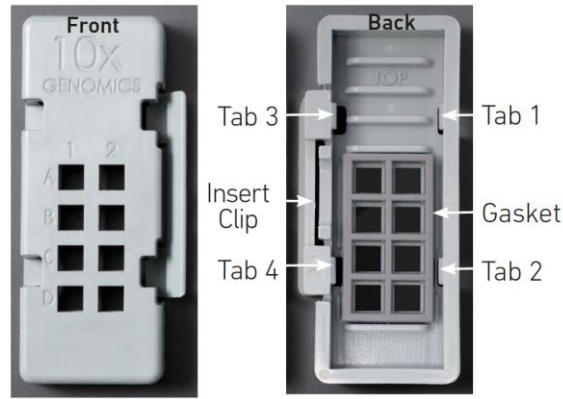
Figure 8: Visium slide cassette

The four tabs and insert clip at the back of the slide cassette ensure that the slide is firmly secure in the cassette and care was taken not to apply excessive force to the slide during insertion and removal from the cassette so as not to break the slide.

Varying permeabilisation time periods were selected based on the optimisation guidelines by Visium and based on optimum times determined by other researchers investigating skin tissue spatial gene expression (Ji et al., 2020). I chose to test the following time periods: 8 minutes, 14 minutes, 20 minutes, 26 minutes, 32 minutes and 38 minutes. The capture area representing the negative control contained a tissue section not exposed to permeabilisation enzyme and the positive control represented reference RNA without any tissue.

First, 2 µl of RNA (1 µg/µl) was added to the centre of the well representing the positive control capture area. Permeabilisation enzyme was then added to the 6 wells containing tissue without introducing bubbles, and the enzyme was not added to the wells representing the positive and negative controls. Uniform coverage of the tissue sections was ensured and tapping the slide cassette gently was done to help ensure this. Furthermore, 100 µl of 0.1X saline-sodium citrate (SSC) buffer (Millipore Sigma) was added to the negative control only. The slide seal was then applied to cover the wells and prevent evaporation during permeabilisation, and the slide cassette was then placed on an Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific), which had been pre-equilibrated to 37°C.

After 8 minutes (which represented the first permeabilisation time period under investigation), the slide cassette was removed from the thermal cycler and the seal was

23

removed. The permeabilisation enzyme was removed from the well in which 8 minutes has been assigned for permeabilisation time, and 100 µl of 0.1X SSC buffer was added. The slide cover was then applied and the cassette was placed immediately in the thermal cycler to continue incubation. After 6 minutes (which represented the interval between the permeabilisation test times), the same process was repeated, with removal of the permeabilisation enzyme from the well that was assigned 14 minutes for permeabilisation, and with addition 100 µl of 0.1X SSC buffer to the same well. This process was repeated until the sixth tissue section has been incubated for 38 minutes in total.

At this stage, all of the wells except the positive control contained tissue sections covered with 0.1X SSC buffer. A proprietary Fluorescent Reverse Transcription (RT) Master Mix (10x Genomics) was then prepared on ice, which contained: 221.8 µl of nuclease-free water, 110 µl of RT Reagent C (which required minimal light exposure), 30.8 µl of the Template Switch Oligo (TSO), 8.8 µl of Reducing Agent B and 68.6 µl of RT Enzyme D. The total volume of this mix was 440 µl, which was the amount required for eight wells on the slide plus an additional 10%. The mix was pipetted and centrifuged briefly to ensure mixing of the reagents.

The 0.1X SSC buffer was then removed from each well and 50 µl of the Fluorescent RT Master Mix was added to each well, ensuring uniform coverage of the slide surface. The slide seal was then replaced and the cassette was placed on an Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific), programmed to the following settings for the reverse transcription reaction:

Table 1: Reverse transcription protocol for fluorescently labelled cDNA synthesis

Step	Temperature	Time (min)
Pre-equilibrate	53ºC	Hold
cDNA synthesis	53ºC	45
Hold	4ºC	Hold

2.1.1.3.3.2. Tissue removal

Following removal of the slide cassette from the thermal cycler and removal of the slide seal, the Fluorescent RT Master Mix was removed from the wells. 100 µl of 0.1X SCC buffer was added to each well then removed again. A proprietary Tissue Removal Mix (10x Genomics) was then prepared at room temperature, which was made up of 539 µl Tissue Removal Buffer and 77 µl Tissue Removal Enzyme, which was mixed by pipetting and briefly centrifuging. 70 µl of the Tissue Removal Mix was added to each well without introducing bubbles and ensuring uniform coverage. The slide seal was then replaced and the cassette was then placed on the Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific), which had been pre-equilibrated to 56°C and set to run for 1 hour at 56°C then hold at 22°C.

At this stage, 45 ml 2X SCC-0.1%SDS was dispensed in a 50 ml centrifuge tube and pre-warmed to 50°C in a water bath. Furthermore, 45 ml 0.2X SCC and 45 ml 0.1X SSC buggers are dispensed in two separate 50 ml centrifuge tubes and maintained at room temperature. Following the incubation on the thermal cycler, the Tissue Removal Mix is removed from the wells and the slide is removed from the cassette.

The slide was then immersed 15 times in the pre-warmed 2X SSC - 0.1% sodium dodecyl sulphate (SDS) solution, then 15 times in 0.2X SSC then 15 times in the 0.1X SSC. The slide was then placed in a 50 ml centrifuge tube in a swinging bucket centrifuge and spun for 30 seconds at 250 rcf. After ensuring that there was no remaining tissue on the slide and that the slide case was covered in aluminium foil to reduce light exposure, the slide was imaged under fluorescent settings.

2.1.1.3.3.3. Fluorescence imaging

The slide was visualised using the Zeiss AxioImager with apotome microscope (Carl Zeiss Microscopy, Jena, Germany) using the fluorescence imaging mode (Zeiss Axiocam 503 monochrome camera module, Zeiss Rhodamine B filter cube, and Colibri 7 LED light source). All eight capture areas were imaged at once without using autoexposure and the imaging parameters were adjusted to ensure that fluorescence spots were clear and in focus.

The maximal fluorescence signal was considered to represent mRNA release, thereby allowing the optimal permeabilisation time for adult skin tissue to be determined.

2.1.1.3.4. Spatial gene expression library preparation

Having optimised the permeabilisation time for adult skin tissue, I proceeded with the spatial gene expression experiment. OCT-embedded frozen adult skin tissue was cryosectioned and placed on one of the 4 capture areas on the Visium Gene Spatial Expression slide (10x Genomics). The gene expression slides are engineered to contain approximately 5,000 barcoded 'spots' within each capture area, unlike the Tissue Optimisation slides. Akin to the scRNA-seq methods where barcoding is used to link single cells to individual beads, the cDNA synthesised from the tissue liberated mRNA in the spatial gene expression experiment is also barcoded to one of the 5,000 spatial spots. This allows us to match up sequencing output with spatial location, building a map of gene expression across the tissue on the slide (Figure 9).
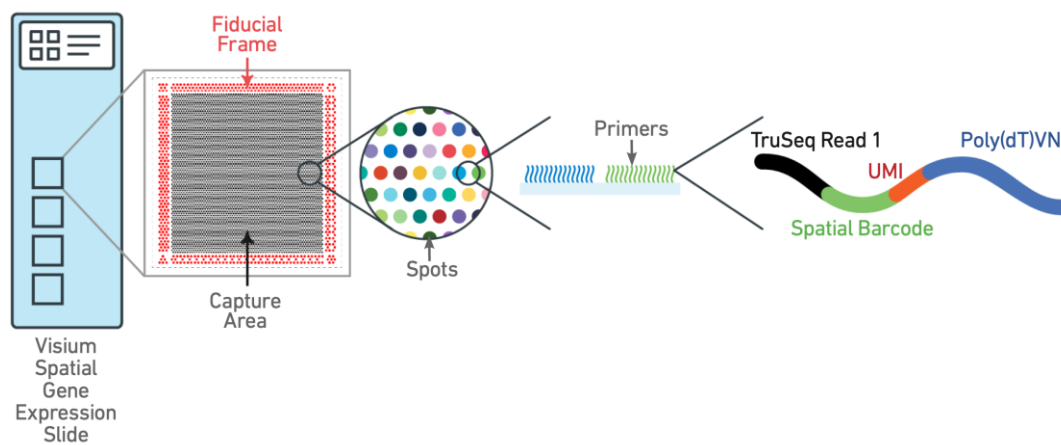
Figure 9: Visium Spatial Gene Expression slide, containing 4 capture areas (6.5 x 6.5mm) each defined by a fiducial frame. Each capture area has ~5,000 spots, each with primers that include an Illumina (Read 1) sequencing primer, a spatial barcode, a unique molecular identifier (UMI) and a poly(dT) sequence that captures poly(A) mRNA. Image from the 10x Genomics Visium protocol.

In designing the experiment, I placed skin tissue from different conditions on the same gene expression slide: for example, two captures areas were used for healthy skin tissue, one capture area for lesional psoriasis skin and 1 capture area for non-lesional psoriasis skin.

This was done to reduce technical artefact in the data when analysing the differences between conditions.

2.1.1.3.4.1. Tissue sectioning and histology imaging

The OCT-embedded adult healthy and diseased skin tissue was cryosectioned at a thickness of 15 µm using an equilibrated OTF5000 cryostat (Bright Instruments). The tissue sections were carefully placed on one of the four capture areas on the Visium Spatial Gene Expression slides (10x Genomics) to avoid tissue folding, crumpling or misplacement (as described in section 2.1.1.3.1). The slide was then stored at -80 °C overnight. The following day, the tissue on the slide was stained with H&E (as described in section 2.1.1.3.1) and then mounted with a coverslip using 70% glycerol. The slide was then imaged using a Zeiss AxioImager with apotome microscope (Carl Zeiss Microscopy) and Brightfield imaging (Zeiss Axiocam 105 48 colour camera module) at 20X magnification. The ZEN blue edition V.3.1 (Carl Zeiss Microscopy) software was then used to acquire the images, adjust the z-plane and light balance, as well as stitching the image tiles to retrieve the overall H&E image file.

2.1.1.3.4.2. Tissue permeabilisation

The slide was then taken back to the lab and loaded into the slide cassette (Figure 8) in order to commence tissue permeabilisation to release the mRNA. 70 µl of pre-warmed permeabilisation enzyme (10x Genomics) was added to each of the 4 wells and the cassette was then placed in an Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific) that had been pre-equilibrated to 37°C. The incubation time at this step corresponded to the previously determined optimum permeabilisation time.

Following this incubation, the slide cassette was removed from the thermal cycler and the slide seal was removed. The permeabilisation enzyme was then removed from the corners of the 4 wells and 100 µl of 0.1X SSC was added to each well.

2.1.1.3.4.3. Reverse transcription

The proprietary Reverse Transcription Master Mix (10x Genomics) was the prepared on ice, which contained: 144.5 µL nuclease-free water, 82.5 µl Reverse Transcription Reagent, 23 µl Template Switch Oligo (TSO), 6.6 µl Reducing Agent B and 51.4 µl Reverse Transcription Enzyme B. This was then pipette mixed and centrifuged briefly.

The 0.1X SSC buffer was then removed from the wells and 75 µl of the Reverse Transcription Master Mix was added to each well. The slide seal was then applied and the slide cassette was placed on an Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific), which had been pre-equilibrated to 53°C and programmed to run for 45 minutes at 53°C then hold at 4°C.

2.1.1.3.4.4. Second strand synthesis and denaturation

The slide cassette was then removed from the thermal cycler and the Reverse Transcription Master Mix was removed from the wells. 75 µL 0.08 M potassium hydroxide (KOH), which was diluted from 8 M stock by mixing 5 µl KOH and 495 µl nuclease-free water, was then added to each well and left to incubate for 5 minutes at room temperature. The KOH was then removed from the wells and 100 µl Buffer EB (Qiagen) was added to each well.

A proprietary mix of second strand synthesis reagents was then prepared on ice, which contained: 305.8 µl Second Strand Reagent, 17.6 µl Second Strand Primer and 6.6 µl Second Strand Enzyme. The mix was then vortexed and centrifuged briefly. The elution buffer was next removed from the wells and 75 µl of the second strand mixture was added to each well. The slide seal was then applied and the slide cassette was placed on an Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific), which had been programmed according to the following settings:

Table 2: Thermal cycler parameters for second strand synthesis

Step	Temperature	Time (min)
Pre-equilibrate	65ºC	Hold
Second strand synthesis	65ºC	15
Hold	4ºC	Hold

Following the incubation, the reagents were removed from the wells and 100 µl of the elution buffer was added to each well. The elution buffer was then removed and 35 µl 0.08 M KOH was then added to each well and left to incubate for 10 minutes for denaturation. During this time, 5 µl Tris-hydrochloric acid (HCl) (1 M, pH 7.0) was added to 4 tubes in an 8 tube strip and a 35 µl sample from each well was added to a corresponding tube containing the Tris-HCl. This was vortexed and centrifuged briefly and placed on ice.

The following image summarises the reactions that take place on the Visium Spatial Gene Expression slide, as described in the steps above:
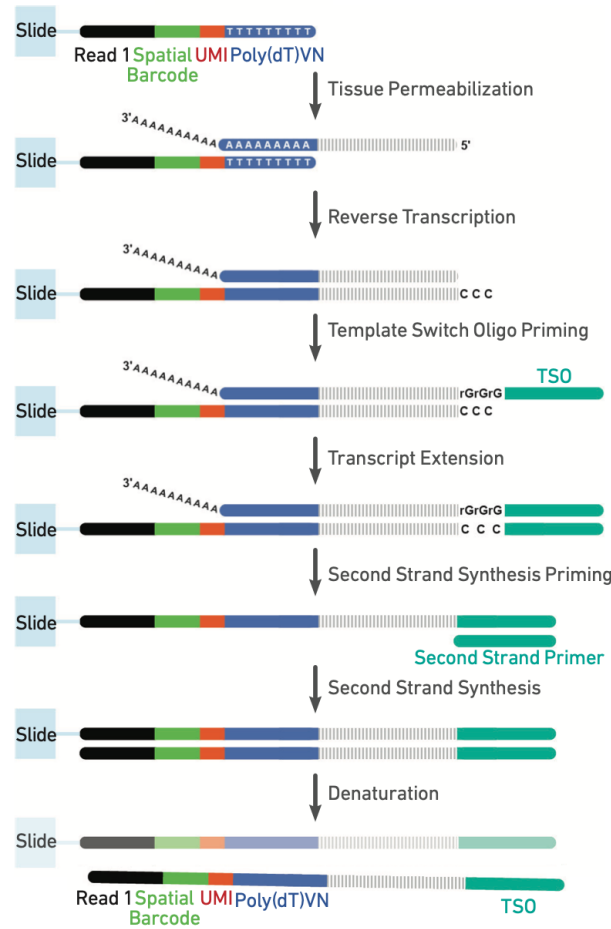
Figure 10: Sample processing steps on the Visium slide, including tissue permeabilisation, reverse transcription, Template Switch Oligo priming, transcript extension, second strand synthesis and denaturation in order to retrieve spatially-barcoded cDNA molecules. Image from the 10x Genomics Visium protocol.

2.1.1.3.4.5. cDNA amplification and quality control

The spatially-barcoded cDNA was then amplified in order to generate sufficient mass for gene expression library construction. The number of PCR cycles was optimised to create enough material for downstream gene expression library construction and, at the same time, minimising resultant undesirable PCR amplification artefacts. In order to determine the optimal number of PCR cycles for cDNA amplification, an optimisation quantitative PCR (qPCR) was carried out. A proprietary quantitative PCR Mix (10x Genomics) was first prepared on ice, which contained 20.4 µl nuclease-free water, 27.5 µl KAPA SYBR FAST qPCR Matster Mix (requiring minimum light exposure by covering the Eppendorf tube with aluminium foil) and 1.7 µl cDNA primers. This mix was then vortexed and centrifuged

briefly. 9 µl of the qPCR Mix was then added to each well in a qPCR plate and 1 µl of the sample that had undergone second strand synthesis was added to each qPCR plate well containing the qPCR Mix. The qPCR plate was then placed on an Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific) programmed according to the following settings:

Table 3: Thermal cycler parameters for qPCR and cycle number determination

Step	Temperature	Time (sec)
1	98ºC	180
2	98ºC	5
4	63ºC	30
5	Repeat from step 2 for 25 cycles	

The optimal number of PCR cycles was then determined using the quantification cycle (Cq) value, which was set along the exponential phase of the amplification plot at approximately 25% of the peak fluorescence value.

Next, the proprietary cDNA Amplification Mix (10x Genomics) was prepared on ice, which contained 220 µl Amp Mix and 66 µl cDNA Primers. 65 µl of the cDNA Amplification Mix was added to the remaining approximately 35 µl sample from the second stand synthesis part of the experiment and pipette mixed and centrifuged briefly. This was then placed in an Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific) for optimised cDNA amplification using the following settings:

Table 4: PCR protocol for cDNA amplification

Step	Temperature	Time (sec)
1	98ºC	180
2	98ºC	15
3	63ºC	20
4	72ºC	60
5	Repeat steps 2–4 for 14* cycles (based on Cq value obtained in last step)	
6	72ºC	60
7	4ºC	Hold

2.1.1.3.4.6. cDNA cleanup

The SPRIselect reagent (Beckman Coulter) was vortexed for resuspension and 60 µl (0.6X) was added to each sample and pipette mixed. This was then incubated at room temperature for 5 minutes and placed on the 10x Genomics Magnetic Separator on the high setting until the solution cleared. The supernatant was then removed and 200 µl 80% ethanol was added to the pellet. After 30 seconds, the ethanol was removed, and the ethanol wash step was repeated. Any remaining ethanol was removed and then the tube strip was left to air dry for 2 minutes and removed from the magnet. 40.5 µl of an elution buffer was then added, followed by pipette mixing and incubation at room temperature for 2 minutes. The tube strip was then placed on the magnetic separator on the low setting until the solution cleared and 40 µl of the sample was transferred to a new tube strip.

2.1.1.3.4.7. cDNA quality control and quantification

1 µl of the sample was then run on a 2100 Agilent Bioanalyzer (Agilent, Santa Clara, CA, USA) High Sensitivity DNA chip. The cDNA concentration (pg/µl) reported by the software was multiplied by the elution volume of the post cDNA amplification reaction clean up sample and divided by 1,000 to obtain the total cDNA yield in ng.

2.1.1.3.4.7. Spatial gene expression library construction

The spatial gene expression library was prepared using 10 µl of the total purified cDNA obtained previously (the remaining cDNA was stored at -20°C and kept for 4 weeks). The 10 µl cDNA sample was transferred to a tube strip and maintained on ice. 25 µl of an elution buffer was added to each sample and 15 µl of a proprietary Fragmentation Mix (containing 22 µl Fragmentation Buffer and 44 µl Fragmentation Enzyme, 10x Genomics) was then added to each sample. This was then pipette mixed and centrifuged briefly, then transferred to an Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific) that was programmed as follows:

Table 5: Thermal cycler incubation protocol for cDNA fragmentation

Step	Temperature	Time (min)
Pre-cool	4°C	Hold
Fragmentation	32°C	5
End Repair and A-tailing	65°C	30
Hold	4°C	Hold

The SPRIselect reagent (0.6X) was resuspended and 30 µl of this was added to each sample and left to incubate for 5 minutes at room temperature. The tube strip was then placed on the magnetic separator at the high setting until the solution cleared and then 75 µl of the supernatant was transferred to a new strip. 10 µl of the SPRIselect reagent (0.8X) was then added to each sample and left to incubate at room temperature for 5 minutes, then placed on the magnet on High until the solution cleared. 80 µl of the supernatant was removed and 125 µl 80% ethanol was added to the pellet. After 30 seconds, the ethanol was removed. The ethanol wash step was repeated then the tube strip was centrifuged briefly and placed on the magnet on Low until the solution cleared. Any remaining ethanol was removed and 50.5 µl elution buffer was added to each sample and left to incubate at room temperature for 2 minutes. It was then placed on the magnet on High until the solution cleared and 50 µl of the sample was transferred to a new tube strip.

An Adaptor Ligation Mix (10x Genomics) was then prepared: 88 µl Ligation Buffer, 44 µl DNA Ligase and 88 µl Adaptor Oligos. 50 µl of this mix was added to the 50 µl sample and pipette mixed and centrifuged briefly. It was then incubated in an Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific) for 15 minutes at 20°C then held at 4°C.

For post ligation cleanup, 80 µl SPRIselect Reagent (0.8X) was added to each sample, pipette mixed and left to incubate at room temperature for 5 minutes. It was then placed on the magnet on High until the solution cleared and the supernatant was removed. 200 µl 80% ethanol was added to the pellet and removed after 30 seconds; this ethanol wash was then repeated. After briefly centrifuging, the tube strip was placed on the magnet on Low, any remaining ethanol was removed and this was left to air dry for 2 minutes. The tube strip was then removed from the magnet and 30.5 µl of the elution buffer was added, pipette mixed and left to incubate at room temperature for 2 minutes. This was then placed on the

magnet on Low until the solution cleared and 30 µl of the sample was transferred to a new strip.

For the sample index PCR procedure, 50 µl of the Amplification Mix (10x Genomics) was added to 30 µl of the sample. 20 µl of an individual Dual Index TT Set (10x Genomics) was also added. After briefly centrifuging, the mixture was incubated in an Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific) using the following PCR protocol

Table 6: PCR protocol for cDNA sample indexing

Step	Temperature	Time (sec)
1	98ºC	45
2	98ºC	20
3	67ºC	30
4	72ºC	20
5	Repeat steps 2–4 for 14 cycles	
6	72ºC	60
7	4ºC	Hold

Post sample index PCR double sided size selection was then carried out according to the manufacturer's instructions. Post library construction quality control (QC) was carried out by running 1 µl of the sample (1:10 dilution) on a 2100 Agilent Bioanalyzer (Agilent, Santa Clara, CA, USA) High Sensitivity DNA chip. The average fragment size was determined from the Bioanalyser trace, which was used as insert size for library quantification.

Finally, the spatial gene expression libraries were pooled and submitted to the Wellcome Genome Campus, Sanger, Cambridge for sequencing using the Illumina Novaseq 6000 (Illumina, CA, USA).

2.1.2. Prenatal skin tissue

2.1.2.1. Tissue acquisition and ethical approval

Prenatal first and second trimester human skin samples were acquired from the Human Developmental Biology Resource (HDBR; http://www.hdbr.org), which is funded by the MRC (Medical Research Council) and Wellcome Trust. The HDBR is licensed by the Human Tissue Authority (HTA; www.hta.gov.uk) and adheres to the guidance set out in the relevant HTA Codes of Practice. Ethical approval and written consent was granted by the Newcastle and North Tyneside NHS Health Authority Joint Ethics Committee (Research Ethics Committee (REC) reference number: 18/NE/0290) (Appendix L).

2.1.2.2. Single-cell RNA sequencing experiment

Prenatal human skin tissue from a total of 15 donors was digested into single cell suspensions for the single-cell RNA sequencing (scRNA-seq) experiment. The ages of the donors ranged from 7 post conception weeks (PCW) to 16 PCW across the first and second trimester of gestation.

2.1.2.2.1. Prenatal skin tissue digestion

Once prenatal skin tissue was transferred from HDBR, it was immediately processed and digested into single cell suspensions. Skin tissue was initially transferred to a sterile 10 mm tissue culture dish and, using a scalpel, sectioned into <1mm segments. The skin tissue was then digested with type IV collagenase, the final concentration of which was 1.6 mg/ml (Worthington) in RPMI (Sigma-Aldrich), and to which 10% heat-inactivated fetal bovine serum (FBS; Gibco) was added. This was kept at 37°C for 30 minutes with intermediate agitation.

A 100 μm cell strainer was then used to pass the digested prenatal skin tissue through, after which the solution was centrifuged at 500g for 5 minutes at 4°C to collect cells in the resultant pellet. At this stage, 1xRBC lysis buffer (eBioscience) was added to the cells for 5 minutes at room temperature. The cells were then washed once with Flow Buffer, which was made up with phosphate-buffered saline (PBS) containing 5% (v/v) FBS and 2 mM

EDTA. The prenatal skin cells were then counted using light microscopy, prior to antibody staining for the next phase of the experiment.

2.1.2.2.2. Flow cytometry and fluorescence-activated cell sorting

The dissociated prenatal skin cells were stained with anti-CD45 antibody (3 microlitres (µL) CD45 BUV395, clone: HI30, BD Biosciences) and left in the dark on ice for 30 minutes. The cells were then passed through a 35 µm filter (Falcon) and DAPI (Sigma-Aldrich) was added at a final concentration of 3 µM. Flow cytometry sorting was carried out using the BD FACS Aria Fusion Flow Cytometer. After gating for single cells, the CD45 positive fraction was sorted from the DAPI-CD45+ gate, and the CD45 negative fraction was sorted from the DAPI-CD45- gate. The CD45 gating was contiguous so that no live cells were lost during sorting. Each population was sorted into chilled PBS within fluorescence-activated cell sorting (FACS) tubes.

2.1.2.2.3. Droplet-based scRNA-seq

2.1.2.2.3.1. GEM generation and barcoding

The live cell suspensions that had undergone sorting were then counted and loaded onto the 10x Genomics Chromium Controller™, where Gel Beads-in-emulsion (GEM) were generated by combining barcoded gel beads, the single cell suspension and a partitioning oil on the Chromium chip (Figure 11). Cells in suspension were added at a low dilution so that a single cell resolution can be attained, whereby over 90% of generated GEMs contain no cells and the remaining GEMs contain a single cell.

The prenatal skin DAPI–CD45+ or DAPI–CD45– FACS- isolated cell concentration that was loaded was calculated to achieve a yield of 10,000 cells per reaction. The Chromium single cell 3' v2 reagent kits from 10x Genomics were used.
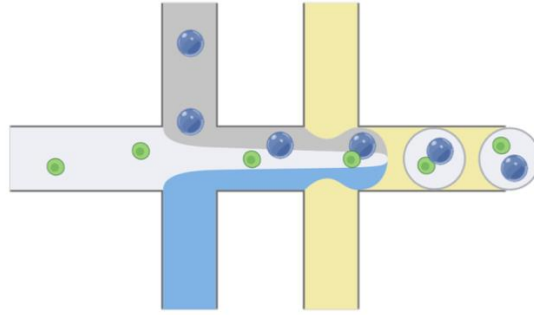
Figure 11: GEM generation, where single gel beads (shown in blue) become encapsulated with single cells (shown in green) in each microfluidic droplet using a partitioning oil. Image created with BioRender.com.

Following GEM generation, the cells were lysed within each microfluidic droplet and the gel beads were dissolved. The gel beads are bound to oligonucleotides that consist of various components: an Illumina sequencing primer, a 10x Genomics barcode that is unique to each bead and therefore each encapsulated cell, a unique molecular identifier (UMI) sequence that is unique to each oligonucleotide and a template switch oligo (TSO).

The single cells within each microfluidic droplet were then lysed, releasing cellular messenger ribonucleic acid (mRNA). A mixture is made than contains the cell lysate, released gel bead oligonucleotides, reverse transcription (RT) reagents and poly deoxythymine (poly(dT)) RT primers. The poly(DT) primers then anneal to the poly adenine (poly(A)) tail of the mRNA molecules and the enzymatic RT reaction then takes place, generating complementary cDNA from the mRNA molecules released from the lysed prenatal skin cells. The TSO sequence allows the cDNA transcript to be extended such that the UMI, 10x barcode and Illumina sequence is transcribed. Therefore, each cDNA molecule is barcoded and associated with each individual cell in each microfluidic droplet.

The RT reaction took place in an Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific) by incubating the reagents at 53 °C for 45 minutes, then 85 °C for 5 minutes.

2.1.2.2.3.2. GEM cleanup

The GEM 'cleanup' process involved removing surplus RT reagents and partitioning oil after the RT reaction was complete. A proprietary 'Recovery Agent' (10x Genomics) was added to each sample at room temperature, and after 2 minutes, a biphasic mixture is seen, containing the Recovery Agent and partitioning oil (that appears pink) and an aqueous (clear) phase that contains the cDNA. The pink Recovery Agent/partitioning oil solution was then removed and discarded, taking care not to remove any of the clear aqueous solution.

Silane magnetic beads (10x Genomics) were then used to purify the barcoded first-strand cDNA in the aqueous phase from the post GEM-RT reaction mixture. The magnetic beads were added to each sample and pipette mixed, then incubated at room temperature for 10 minutes. The solution was then pipette mixed again to resuspend the settled magnetic beads, before a magnetic separator (10x Genomics) was used to fix the cDNA whilst being pipette washed twice with ethanol. The ethanol was then removed, and air drying took place for 1 minute to ensure the residual ethanol evaporated. An Elution solution was then added to separate the magnetic beads from the cDNA using the magnetic separator, then the supernatant containing the cDNA was transferred to a new tube strip to progress with cDNA amplification.

2.1.2.2.3.3. cDNA amplification

The barcoded full-length cDNA was then amplified via polymerase chain reaction (PCR) using a proprietary cDNA amplification mix (10x Genomics) that contains cDNA primers targeting common 3' and 5' ends that were added during the GEM-RT phase of the experiment. The amplification mix was added to each sample and pipette mixed 15 times then centrifuged briefly. The mixture was then incubated in an Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific) with the following settings:

Table 7: PCR protocol for cDNA amplification using the thermal cycler

Step	Temperature	Time (sec)
1	98ºC	180
2	98ºC	15
3	63ºC	20
4	72ºC	60
5	Repeat steps 2–4 for 14 cycles	
6	72ºC	60
7	4ºC	Hold

The number of cycles was optimised to create enough material for downstream gene expression library construction and, at the same time, minimising resultant PCR amplification artefacts.

2.1.2.2.3.4. cDNA cleanup, quality control and quantification

The cDNA cleanup was achieved using a proprietary reagent called SPRIselect (Beckman Coulter), which was added to each sample and pipette mixed 15 times then left to incubate at room temperature for 5 minutes. Magnetic separators were then used to hold the cDNA in situ, after which the cDNA was washed with ethanol twice then left to air dry for 2 minutes to allow the residual ethanol to evaporate. An Elution buffer (Qiagen) was then added, pipette mixed 15 times, then left to incubate for 2 minutes at room temperature. The magnetic separator was used again to retrieve the purified cDNA, and 40 µl from each sample was transferred to a new tube strip.

To assess cDNA quality and to quantify the yield, 1 µl was taken and diluted at a ratio of 1:10. This was then run on a 2100 Agilent Bioanalyzer (Agilent, Santa Clara, CA, USA) according to the manufacturer's instructions. The cDNA yield was calculated by multiplying the cDNA concentration (pg/µl) that is reported by the bioanalyser software by the elution volume of the purified amplified cDNA (40 µl), accounting for the dilution factor, then dividing by 1000 to obtain the total cDNA yield in nanograms (ng). Subsequently, 25% of the total cDNA yield was taken forward to 3' gene expression library preparation.

2.1.2.2.3.5. Gene expression library construction

The amplified full-length cDNA that was generated from the polyadenylated mRNA was next used to construct gene expression libraries. First, the cDNA transcripts were enzymatically fragmented by adding a proprietary Fragmentation Enzyme and Fragmentation Buffer (10x Genomics) to 25% of the cDNA yield and pipette mixing 15 times on ice. The samples were then run on the Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific), which had been pre-cooled to 4ºC, using the following settings:

Table 8: Thermal cycler incubation protocol for cDNA fragmentation

Step	Temperature	Time (min)
Pre-cool	4ºC	Hold
Fragmentation	32ºC	5
End Repair and A-tailing	65ºC	30
Hold	4ºC	Hold

A cleanup step was then carried out following fragmentation to remove any surplus reagents. The cleanup was carried out using SPRIselect (Beckman Coulter), a magnetic separator and ethanol washes, as previously described.

Next, a proprietary Adaptor Ligation Buffer (10x Genomics) was mixed with a DNA ligase and adaptor oligos, added to each sample and pipette mixed. The mixture was then incubated on the Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific) at 20ºC for 15 minutes. This allows an Illumina Read 2 sequence to be appended. A further cleanup step following ligation was carried out. Subsequently, a Dual Index TT Set (10x Genomics) was used to add P5, P7, i5 and i7 sample indices, which were used to deconvolute different samples when they were multiplexed in the sequencing run. The sample indices were added to the samples along with a proprietary Amplification Mix (10x Genomics), and were incubate in the Eppendorf Mastercycler Pro thermal cycler (Thermofisher Scientific) using the following PCR protocol:

Table 9: PCR protocol for cDNA sample indexing using the thermal cycler

Step	Temperature	Time (sec)
1	98ºC	45
2	98ºC	20
3	54ºC	30
4	72ºC	20
5	Repeat steps 2–4 for 14 cycles	
6	72ºC	60
7	4ºC	Hold

Following the PCR run for cDNA indexing, a further cleanup step was carried out using SPRIselect (Beckman Coulter), a magnetic separator and two ethanol washes. The purified cDNA was then analysed on the 2100 Agilent Bioanalyzer (Agilent) in order to determine the average fragment size from the bioanalyser trace, which is then used for library quantification.

The following figure shows the structure of the resultant gene expression library, which contains the P5 and P7 priming sites used in the Illumina sequencer:
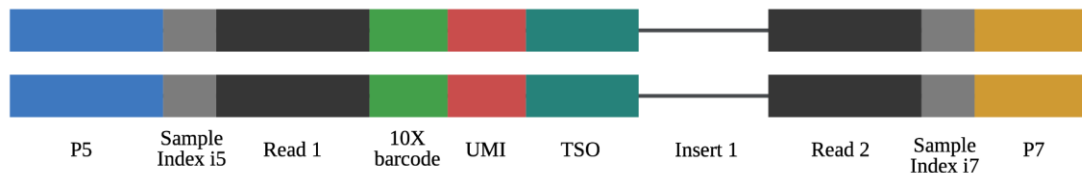
Figure 12: Single cell gene expression dual index library. Image created with BioRender.com and adapted from 10x Genomics.

The prenatal skin gene expression libraries were then pooled and sequenced to achieve a minimum target depth of 20,000 reads per cell using the Illumina Novaseq 6000 (Illumina, CA, USA).

2.1.2.3. Multiplex RNA in situ hybridisation

Human prenatal skin tissue from a 15 PCW embryo was processed to carry out multiplex RNA *in situ* hybridisation using the RNAScope Multiplex Fluorescent Detection Kit v2

(ACDBio, Newcark, California, USA, catalogue number 323100). Thanks are due to Sophie Pritchard and Ilaria Mulas at the Wellcome Sanger Institute in Cambridge for carrying out the multiplex RNA *in situ* hybridisation experiment and to Kenny Roberts at the Sanger for imaging the slides using confocal microscopy and carrying out the post-imaging processing, detailed below.

2.1.2.3.1. Experimental design and procedure

The prenatal skin tissue was frozen and embedded in optimal cutting temperature compound (OCT, Tissue-Tek). 4-plex single molecule fluorescence *in situ* hybridisation (smFISH) was performed according to the manufacturer's instructions using the standard pre-treatment for fresh frozen sections of 10-20 μm and permeabilised with Protease IV at room temperature for 30 minutes.

Human probes against SLC26A7, SHH, NDP and FOXP3 mRNA molecules were used (all from ACDBio catalog probes, Newark, California, USA). Opal dyes (Akoya Biosciences, Marlborough, Massachusetts, USA) were used at a dilution of 1:1,000 for the fluorophore step to develop each channel: Opal 520 Reagent Pack (FP1487001KT), Opal 570 Reagent Pack (FP1488001KT) and Opal 650 Reagent Pack (FP1496001KT) and Atto-425. The following table shows the channels, probes and fluorophores used:

Table 10: RNAScope channels, probes and fluorophores used in the prenatal skin experiment

Channel	Probe	Fluorophore	Colour
1	NDP	Opal 520	Green
2	SHH	Atto 425	Blue
3	FOXP3	Opal 650	Red
4	SLC26A7	Opal 570	Yellow

The slides were then counterstained with DAPI and a cover slip was then placed in order to proceed with imaging using ProLong Gold Antifade Mountant (ThermoFisher, Canoga Park, California, cat. no. P36930).

2.1.2.3.2. Confocal imaging

Confocal imaging was performed on a Perkin Elmer Opera Phenix Plus High-Content Screening System using a 40X (NA 1.1, 0.149 µm/pixel) water-immersion objective with a 2µm z-step. Channels: DAPI (excitation 375 nm, emission 435-480 nm), Atto 425 (ex. 425 nm, em. 463-501 nm), Opal 520 (ex. 488 nm, em. 500-550 nm), Opal 570 (ex. 561 nm, em. 570-630 nm), Opal 650 (ex. 640 nm, em. 650-760 nm).

Confocal image stacks were stitched as two-dimensional maximum intensity projections using proprietary Acapella scripts provided by Perkin Elmer and visualised using OMERO Plus (Glencoe Software).

2.5. Data analysis

2.5.1. Spatial data analysis

2.5.1.1. Alignment, tissue annotation and quality control

Adult skin spatial transcriptomics data was mapped using Spaceranger v.1.3.0 using GRCh38-2020-A reference and prenatal skin data was mapped using Spaceranger v.2.0.1 using GRCh38-1.2.0 reference. Adult skin tissue sections were annotated using the Loupe browser (v 6.1.0; 10x Genomics) in order to define the following regions: epidermis, dermis, pilosebaceous unit, adipose tissue and non-tissue areas. The UMI count was analysed for each tissue section and annotated tissue region. In order to include tissue regions with low UMI counts in downstream analyses, Visium spots that had UMI counts of less than 500 were merged together to form meta-spots. First, all spots were organised into a single level coarse hexagonal metaspot mesh, where each metaspot consists of seven spots (one of a few possible meshes was chosen randomly). Then, spots that belong to the same metaspot and that have total UMI counts below 500 were merged and their gene counts were summed. Resultant counts were assigned to either the central spot in the metaspot if its coverage was below 500 UMIs or to the one of merged spots that had the highest coverage. Original spots that contributed to the merged ones were removed from the object. All spots (or metaspots) that had coverage below 500 UMIs following the abovementioned procedure were removed from the analyses.

2.5.1.2. Spatial cell type deconvolution

The reference scRNA-seq dataset for spatial cell type deconvolution of the adult skin Visium data represented an combination of adult interfollicular healthy, psoriasis and eczema skin data (Reynolds et al., 2021) and adult hair graft data (Takahashi et al., 2020). In order to harmonise non-follicular cell annotations across both datasets, scanpy (v 1.9.1) ingest was used. The reference scRNA-seq for deconvolution analysis of the embryonic limb Visium data (Zhang et al., 2022) included prenatal skin data from samples ≤ 10 PCW. Cell types where less than 20 cells were identified in these prenatal skin samples were excluded from the reference dataset.

To map cell types identified by scRNA-seq in the profiled spatial transcriptomics slides, the Cell2location (v0.1) method was used (Kleshchevnikov et al., 2022). Firstly, a negative binomial regression model was trained to estimate reference transcriptomic profiles for all the cell types profiled with scRNA-seq in the tissue. Lowly expressed genes were excluded using the filtering strategy recommended by developers of the Cell2location package (cell_count_cutoff = 5, cell_percentage_cutoff = 0.03, non_mean_cutoff = 1.12). Training was performed for 250 epochs and reached convergence according to manual inspection. Next, the abundance of cell types in the spatial transcriptomics slides was estimated using reference transcriptomic profiles of different cell types. All slides were analysed jointly. The following cell2location hyperparameters were used: (1) expected cell abundance (N_cells_per_location) = 30; (2) regularisation strength of detection efficiency effect (detection_alpha) = 20. The training was stopped after 50,000 iterations. All other parameters were used at default settings. Cell2location estimates the posterior distribution of cell abundance of every cell type in every spot. Posterior distribution was summarised as 5% quantile, representing the value of cell abundance that the model has high confidence in, and thus incorporating the uncertainty in the estimate used for downstream analysis.

2.5.1.3. Cell distance from dermoepidermal junction analysis

To map adult skin cell states to different tissue depths using the dermoepidermal junction (DEJ) as a reference point, the DEJ was first defined as Visium spots annotated as epidermis contacting spots annotated as dermis. The distance between each spot centre and the centre of the nearest DEJ spot was then calculated in image scale (pixels). The distance was divided by the interspot distance and rounded. Distances for non-epidermal spots were multiplied by minus one. In real scale, the interspot distance is 100 μm.

All spots that were closer than 3 spots to annotated hair follicle regions were excluded from the analyses. Cell2location-predicted cell type abundances were per spot normalized (divided by total abundance). Cell type abundances in spots with the same distance from the DEJ were averaged within each sample and used as independent observations.

2.5.1.4. Tissue microenvironment analysis

To identify microenvironments of co-locating cell types, non-negative matrix factorisation (NMF) was used. The matrix of estimated cell type abundances was first normalised by dividing the matrix by per-spot total abundances. Resulting matrix Xn of dimensions $n \times c$, where n is the total number of spots in the Visium slides and c is the number of cell types in the reference, was decomposed as $Xn = WZ$, where W is a $n \times d$ matrix of latent factor values for each spot and Z is a $d \times c$ matrix representing the fraction of abundance of each cell type attributed to each latent factor. Here the latent factors correspond to tissue microenvironments defined by a set of co-localised cell types. The NMF package for R (Gaujoux and Seoighe, 2010) was used, setting the number of factors d = 10 and using the default algorithm (Brunet et al., 2004). NMF coefficients were normalised by a per-factor maximum. NMF was run 100 times to construct the consensus matrix. The best run was then selected based on lower mean silhouette scores calculated on the consensus matrix; if at least 2 runs had minimal mean silhouette scores, the run with the smallest deviance was selected.

In order to compare spatial microenvironments between adult healthy trunk samples and anatomical site matched disease samples from psoriasis lesional/non-lesional skin and eczema lesional/non-lesional skin, the data was first subsampled prior to running NMF. Four samples from each condition were randomly selected and 50% of the Visium spots for each sample were randomly selected. NMF was then run on per-spot normalised Cell2location predictions 100 times to construct the consensus matrix. Cell type pairs between pairs of conditions were then compared by a proportion test and the difference between consensus matrices.

For cell type abundance correlation analysis in prenatal skin, a per-spot normalised Xn matrix was used. Pearson correlation coefficient was calculated for each pair of cell types (all possible pairs computed) and each sample. For visualisation of correlation analysis, selected cell pairs were plotted, guided by NMF results and by which cell groups and categories formed microenvironments.

2.5.2. scRNA-seq analysis

2.5.2.1. Alignment, quality control, clustering and annotation of prenatal skin dataset

The gene expression data were mapped with Cellranger 2.1.1 and 2.0.2 to an Ensembl 84–based GRCh38 reference (10x Genomics–distributed 1.2.0 version). The Python package emptydrops (https://pypi.org/project/emptydrops/) (v0.0.5) was used to detect cells in each sample. Potential doublets were flagged by Scrublet (v0.2.1) (Wolock, Lopez and Klein, 2019). Low-quality cells were filtered out using the following parameters: minimum number of genes = 100, maximum number of UMIs = 45000, maximum mitochondrial UMI fraction = 0.15. Possible maternal contamination was identified using the Souporcell pipeline (v2.4.0) (Heaton et al., 2020). Data pre-processing was performed using scanpy (v1.4.3) (Wolf, Angerer and Theis, 2018). After pooling data from all samples, genes detected in fewer than 3 cells were removed and data was normalised to 10,000 UMI per cell and log1p transformed.

Highly variable genes were selected based on normalised dispersion (scanpy.pp.highly_variable_genes with flavor = "seurat", min_mean = 0.0125, max_mean = 3, min_dispersion = 0.5). Dimensionality reduction was done by Principal Component Analysis (PCA) and the first 50 principal components (PCs) were used to compute nearest-neighbour graph (scanpy.pp.neighbors with n_neighbors = 15). BBKNN (v1.3.3) (Polański et al., 2020) was used to generate a batch-corrected nearest-neighbour graph considering each donor as a separate batch. Leiden algorithm was used to cluster cells based on the corrected graph with a relatively low resolution (scanpy.tl.leiden with resolution = 0.3) into coarse clusters which were manually annotated into broad lineages using known marker genes.

For each broad lineage, the data was re-processed starting from highly variable gene selection in order to better reveal the finer heterogeneity. At this level, Harmony (v0.1) (Korsunsky et al., 2019) and BBKNN was used in parallel for batch correction (again treating each donor as a separate batch and with batches under 10 cells removed) for every broad lineage. Highly consistent embeddings and clustering were observed between the Harmony and BBKNN approaches. For fibroblasts, the analysis was continued with embeddings and clustering downstream of BBKNN and for all other broad lineages,

Harmony was used. Leiden clusters at the highest-resolution were annotated manually using marker genes identified through literature search and their alignment with differentially expressed genes (DEGs) in each cluster. DEGs were calculated using the sctk (Single Cell analysis Tool Kit) package (https://github.com/Teichlab/sctk), where filtering is carried out then a two-sided Wilcoxon rank-sum test carried out based on pass-filter genes only in a one-vs-all fashion. The sctk package also carries out comparisons between the group of interest (one with highest expression) and the next group (second highly expressed), where the maximum proportion of cells expressing the gene in question in the second most highly expressed group was 0.2. For epidermal annotations, a combined embedding of prenatal skin and skin organoid data (Lee et al., 2020) was created, integrated using the Harmony pipeline, as described above. Harmony corrected PCs were used to compute the batch-corrected nearest neighbourhood graph, and the Leiden algorithm was used to cluster the integrated data, as described above. The sctk package was then used to derive marker genes for derived Leiden clusters. Annotation was carried out on the clusters based on marker genes and refined annotations in the skin organoid data (Lee et al., 2020).

Clusters of doublets were manually flagged and removed by taking into account marker genes and previously calculated scrublet scores. To have a final global visualisation of the atlas, a doublet-free UMAP was generated.

2.5.2.2. Processing, clustering and annotation of skin organoid data

Organoid data was preprocessed, filtered, clustered and annotated separately before integration with prenatal skin data. Briefly, CellRanger (CellRanger 2.1.0 with GRCh38-1.2.0 and CellRanger 3.0.2 with GRCh38-3.0.0) filtered cells from samples of two strains each four time points were pooled and cell QC thresholds for UMI counts, gene counts, percentage of mitochondrial genes and top 50 highly expressed genes were established. The thresholds were: minimum number of genes = 450, maximum number of genes = 5731, minimum number of UMIs = 1063, maximum number of UMIs = 25559, maximum mitochondrial UMI fraction = 0.133, minimum cumulative percentage of counts for 50 most expressed genes in a cell = 23.7%, maximum cumulative percentage of counts for 50 most expressed genes in a cell = 56.6%. Highly variable gene selection, dimensionality reduction and KNN graph construction was done using the same method and parameters as

for prenatal skin. BBKNN was again used to generate a batch-corrected graph, treating combinations of strains and 10x kit versions as batches. Broad lineages were annotated based on known markers. Each broad lineage was then re-processed in the same way as prenatal skin to annotate cell types at a higher resolution.

2.5.2.3. Integration of prenatal skin, adult skin and skin organoid cells

To integrate prenatal skin cells with organoid cells, the datasets were down-sampled for each broad lineage to have roughly balanced cell counts per cell type before integration with Harmony, treating datasets as batches (prenatal skin or organoid) and within dataset batches as covariates (donor for prenatal skin and strain:10x kit version for organoid). Leiden clusters were annotated using known markers.

To integrate prenatal skin, skin organoid and adult skin cells at a broad level, all datasets were down-sampled to have roughly balanced cell counts per broad lineage before integration with Harmony, treating datasets as batches and within-dataset batches as covariates.

2.5.2.4. Differential abundance analysis

Differences in cell abundance associated with gestational age were tested for using Milo (v1.0.0) (Dann et al., 2022). Cells were initially re-embedded into a batch-corrected latent space with a dimension of 20 using the scVI model, as implemented in scvi-tools, considering donor ID as batches. The model was trained using the 5000 most highly variable genes identified using scanpy.pp.highly_variable_genes with parameters: flavor="seurat_v3" and batch="donorID"). Milo then constructed a KNN graph of cells (buildGraph, k=15) based on distances in the latent space and assigned cells to neighbourhoods (makeNhoods, prop=0.05). The number of cells belonging to each sample in each neighbourhood was then counted (countCells) and cell type labels were assigned to neighbourhoods based on majority voting (annotateNhoods) with those labelled as "Mixed" if no single cell type represented more than 70% of the total cells in the neighbourhood. To test for differential abundance across gestational age, the counts in each neighbourhood were modelled using a negative binomial generalised model and a log-

linear model to model the effects of age (post-conception week), with cell sorting method as a categorical co-variable to account for its impact on cell abundance (testNhoods).

2.5.2.5. Cell state and developmental stage predictions using logistic regression

Comparison of cell type correspondence between datasets and probability prediction was carried out using a logistic regression (LR) framework. A model was built using the implementation of the sklearn.linear_model.LogisticRegression module from sklearn package (v1.1.3) (parameters: penalty: L2, solver: saga, regularization strength C=0.1) and trained on the gene expression matrix of the training dataset using all genes that passed QC. The resulting model was used to predict the labels in the target dataset. The correspondence between predicted and original labels in target dataset was computed as Jaccard index and visualised as a heatmap. For comparison of developmental stages, a merged prenatal and adult skin dataset was used as training data and the organoid dataset was used as a target. For comparison of hair follicle cell states, merged prenatal and organoid data was used as training and adult hair follicle data was used as a target.

2.5.2.6. Trajectory analysis

Prenatal skin single-cell trajectories and pseudotime were computed with Monocle3 (https://cole-trapnell-lab.github.io/monocle3) for the epithelial cells and fibroblasts and utilities are available in the 'sctkr' R package (https://github.com/Teichlab/sctkr). Batch-corrected principal components were used as input for Monocle3. The *plot_trajectories()* and *plot_heatmap()* functions in sctkr package were used to respectively generate UMAPs of inferred trajectories (coloured by celltype, PCW and pseudotime) and heatmaps of gene expression along individual trajectories.

2.5.2.7. Cell-cell interaction analysis

The CellPhoneDB (v3.0.0) package (Efremova et al., 2020) was used to infer cell-cell interactions within the prenatal skin scRNA-seq dataset overall and by early/late gestation. In the overall analysis, cell states were randomly subsampled into groups of no more than 200 cells. The subsampled dataset was analysed using the permutation-based method to

establish statistical significance (p-value cut-off = 0.05). For the analysis by early/late gestation, the prenatal skin scRNA-seq dataset was first split into early (≤ 11 PCW) and late (≥ 12 PCW) gestation datasets, which were then subsampled (no more than 200 cells per cell type) and analysed individually (p-value cut-off = 0.05).

A curated list of interactions was plotted for visualisation using ggplot2 (v3.3.6). Circos plots were generated using the Circlize package (v 0.4.15) (Gu et al., 2014) were used for downstream visualisations of selected significant (adjusted p-value <0.05) interactions between co-locating cell types.

2.5.2.8. Gene set enrichment analysis

Gene set enrichment analysis was performed using the implementation of the Enrichr workflow (Chen et al., 2013) in the python package GSEApy (https://gseapy.readthedocs.io/), with Gene Ontology (GO) Biological Process (2021) as the query database. To determine the significantly overexpressed genes for gene set enrichment analysis, differentially expressed genes between ILC3s present at different gestational ages were identified using the Wilcoxon rank-sum test implementation in Scanpy (scanpy.tl.rank_genes_groups, method = "wilcoxon").

Chapter 3: Spatial transcriptomic mapping reveals changes in cellular localisation and altered functional microenvironments in adult psoriasis and eczema skin

3.1. Introduction

Psoriasis and eczema are both highly prevalent inflammatory skin diseases worldwide (Hay et al., 2014; Griffiths et al., 2017) that have a significant impact on patient morbidity and quality of life (Hay et al., 2017). Both conditions are significantly associated with other diseases, such as psoriatic arthritis, diabetes and metabolic syndrome in the case of psoriasis (Gottlieb, Chao and Dann, 2008) and asthma, rhinitis, food allergies and cardiovascular disease in the case of eczema (Brunner et al., 2017). Genetic susceptibility is one aspect of the aetiology of both conditions (Brown and McLean, 2012; Paternoster et al., 2015; Stuart et al., 2010); multiple environmental factors are also implicated in causing disease (Langan et al., 2006; Luschkova et al., 2021; Xie et al., 2021; Vičić et al., 2021), culminating in immune dysfunction and an infiltration of inflammatory cells into the skin that have been therapeutically targeted in disease management. However, treatment approaches to date do not prevent disease progression, are not curative, are not personalised and often rely on non-specific immunosuppression. In the absence of reliable biomarkers to stratify patients and guide treatment options, trial and error prescribing is often carried out, resulting in poor outcomes for patients and a detrimental socioeconomic impact on healthcare services (Strober et al., 2021; Davison et al., 2017; Brown, 2017).

In order to further our understanding of the cellular and molecular mediators of psoriasis and eczema pathogenesis, our research group recently generated a comprehensive single cell atlas of healthy, lesional and non-lesional psoriasis, and lesional and non-lesional eczema skin (Reynolds et al., 2021). This study revealed changes affecting several immune and non-immune cell states in disease, including a significant expansion of cytotoxic and helper T cells expressing interleukin 17 (*IL17A* and *IL17F*) in psoriasis lesional skin, cytotoxic T cells expressing *IL13* and *IL22* in eczema lesional skin and an expansion of

vascular endothelial cell subtype 3 in both psoriasis and eczema lesional and non-lesional skin (Reynolds et al., 2021). Cell-cell communication analyses revealed disease-associated interactions mediated by *ACKR1* expressed on vascular endothelial cell 3 and *CXCL8* on macrophage subtype 2, which is implicated in lymphocyte recruitment and angiogenesis (Reynolds et al., 2021).

Though samples in the scRNA-seq study were enzymatically and mechanically separated into epidermal and dermal tissues prior to single cell profiling to provide a detailed characterisation of the abundance of different cell states across skin compartments (Reynolds et al., 2021), valuable information about the organisation of skin cells *in situ* was still lost during the generation of single cell suspensions. Furthermore, analyses to infer significant cell-cell communication analyses of the scRNA-seq dataset were challenging to systematically validate microanatomically; this was approached by immunostaining of target molecules within focal tissue regions (Reynolds et al., 2021). The aim of my study was therefore to generate a spatial transcriptomic atlas of healthy, psoriasis lesional/non-lesional and eczema lesional/non-lesional skin to further our understanding of the microanatomical changes underpinning disease. The spatial arrangement of cell populations was mapped using reference scRNA-seq data, revealing disease-related perturbations in cellular location and changes in tissue cellular microenvironments across healthy, psoriasis and eczema skin.

3.2. Results

3.2.1. Adult skin sample acquisition

Healthy adult skin samples that are normally discarded during surgical defect reconstruction were collected from 13 patients. Skin samples from various anatomical sites were obtained, including the face, scalp, arm, leg and trunk (Figure 13). Five patients with psoriasis and five patients with eczema were also recruited into the study, where lesional and adjacent non-lesional skin was sampled using punch biopsies (Figure 13). Patients included in the study were naïve to biologic treatment, had not had systemic treatment for at least 4 weeks prior to the biopsies being taken and the biopsy site had not been treated with topical steroids for at least one week prior. Clinical metadata was collected for all donors, including their age and sex (Appendix M).
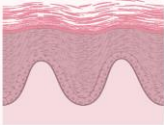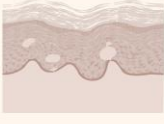
Figure 13: Adult skin sample overview. A summary of the healthy, psoriasis and eczema skin samples collected in this study, including the anatomical sites that the healthy samples were taken from. Image created with BioRender.com.

3.2.2. Tissue optimisation for Visium

A tissue optimisation experiment was carried out to determine the ideal skin tissue digestion time for spatial transcriptomics profiling using Visium. Following the protocol detailed in section 2.1.1.3.3, adult healthy facial skin tissue was sectioned at 15μm thickness onto 7 of the 8 capture areas on the Tissue Optimisation slide (10x Genomics) and H&E staining was initially carried out. Permeabilisation enzyme was subsequently applied to each tissue section for varying periods of time (Figure 14).
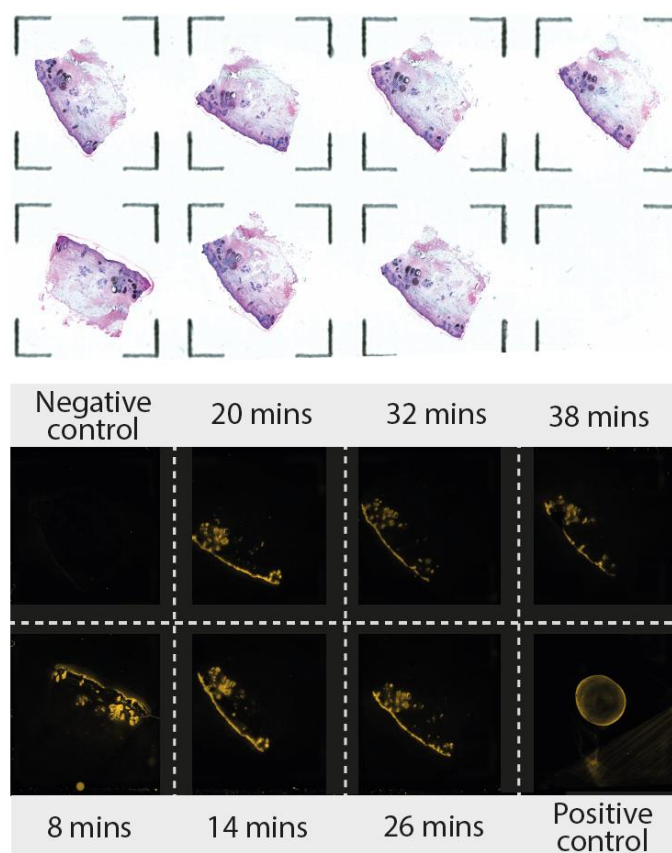
Figure 14: Tissue Optimisation experiment. Adult healthy facial skin was sectioned and stained with H&E (upper panel). Permeabilisation enzyme was then applied to each skin tissue section for varying time periods and the fluorescence signal, which reflects cDNA abundance, was measured across the slide (lower panel).

The negative control, which reflected a tissue section that was not exposed to permeabilisation enzyme, showed minimal fluorescence signal, and therefore was used to help determine the background fluorescence when analysing the image. The positive control, which reflected a sample of reference RNA without any tissue, showed a positive

fluorescence signal, confirming that the reverse transcription reaction was successful (Figure 14).

Focal regions of minimal fluorescence signal were observed across epidermal and dermal regions of the skin tissue following 8 minutes of permeabilisation enzyme application, suggesting insufficient tissue digestion. Beyond 20 minutes of permeabilisation, progressively dimmer fluorescence signals were observed, reflecting RNA degradation with prolonged periods of tissue digestion. The greatest fluorescence intensity with the least signal diffusion was observed following 14 minutes of permeabilisation enzyme application, where epidermal and appendageal structures are well-represented (Figure 14), therefore, this time was chosen for the Visium workflow.

3.2.3. Adult skin sample processing

Adult healthy, psoriasis and eczema samples that were OCT-embedded and frozen were processed to generate spatial transcriptomic sequencing data using the Visium protocol described in Methods section 2.1.1.3.4. I processed 10 healthy samples and the psoriasis samples in Newcastle, whereas the eczema samples and 3 healthy trunk samples were processed by the Cellular Generation and Phenotyping (CGaP) at the Wellcome Sanger Institute whilst I was on maternity leave. The following figure shows the number of tissue sections that sequencing data is available for across the different conditions:

Condition	P728, trunk	P729, face	P730, face	P732, scalp	P734, face and arm	P736, leg	P738, arm	P742, trunk	P746, face	P747, arm	P748, arm	P775, trunk	P777, trunk	P731, trunk	P735, trunk	P737, trunk	P743, trunk	P744, trunk	AD2638, trunk	AD2634, trunk	AD2639, trunk	AD2628, trunk	AD2640, trunk
Healthy	1	2	2	2	4	2 (1)	2	2	2	2	2 (1)	2	4										
Psoriasis lesional														1	2 (1)	2	2	2					
Psoriasis non-lesional														2	2 (1)	2	2	2					
Eczema lesional																			1 (0)	1	2	4	2
Eczema non-lesional																			2	1	2	2	

Figure 15: Spatial transcriptomic data overview. The number of skin tissue sections with spatial transcriptomic data available for each donor and condition is shown. The numbers in brackets denote the data available after filtering.

The experiment was designed with the aim of generating data from two technical replicates for each donor (Figure 15). This was achieved in the majority of cases, however, an exception to this occurred when the psoriasis lesional block from donor P731 was exhausted before the second technical replicate was carried out. Therefore, a single healthy skin section from donor P728 was used in its place on the spatial genomic capture area in order to avoid wasting resources. Furthermore, the H&E images for one experiment that included two healthy (P736 and P748) and a lesional and non-lesional psoriasis sample (P735) was unfortunately corrupted due to a software issue and could not be retrieved. The corresponding sequencing data could not be interpreted fully, therefore the data for these sections were filtered from the downstream analysis, as shown in brackets in Figure 15. The lesional eczema sample from donor AD2638 produced poor quality sequencing data that poorly aligned with the H&E image, and therefore this section was also filtered from the analysable dataset (Figure 15). Four technical replicates were carried out by CGaP for a healthy trunk sample from donor P777 and for a lesional eczema sample from donor AD2628, and any eczema tissue samples where fewer than two technical replicates were carried out represented a lack of remaining tissue in the blocks available for processing.

3.2.4. H&E staining for Visium

To be able to map the spatial transcriptomic sequencing data that was generated to specific microanatomical tissue regions, H&E staining and imaging of each tissue section was carried out as described in section 2.1.1.3.4.1. Representative H&E images for adult healthy, psoriasis and eczema skin are shown below (Figure 16).
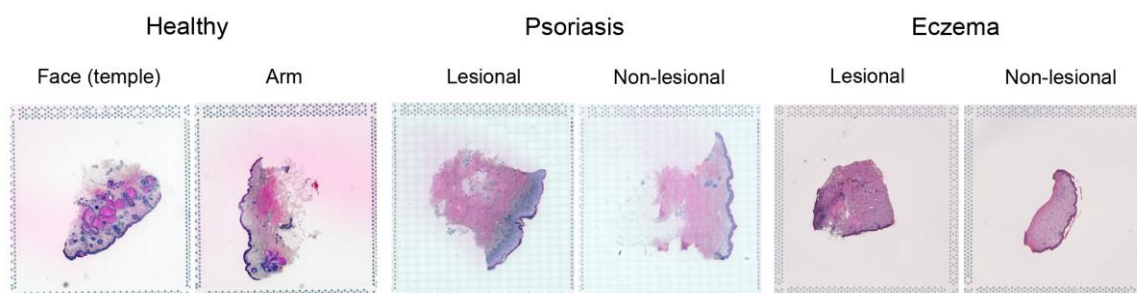
Figure 16: H&E staining of adult healthy, psoriasis and eczema skin tissue. Representative images for each condition are shown, including healthy skin from the face and arm, and lesional and non-lesional skin from patients with psoriasis and eczema.

All skin sections showed a good representation of the epidermis and dermis in each condition, with appendageal structures also observed (Figure 16). The healthy facial skin tissue section exhibited a high density of pilosebaceous units, with sebaceous glands shown as large lobular pink dermal structures and associated hair follicles as smaller round epithelial lined structures mostly seen in the superficial dermis (Figure 16). This finding is in keeping with facial skin having one of the highest concentrations of pilosebaceous units across the body (Fernandez-Flores, 2015). The healthy skin section sampled from the arm showed a pilosebaceous unit, as well as a layer of subcutaneous adipose tissue (Figure 16). The psoriasis lesional tissue section had a markedly thickened epidermal layer and reflects the characteristic elevated, scaly plaques seen clinically (Murphy, Kerr and Grant-Kels, 2007). The non-lesional psoriasis sample taken from clinically uninvolved adjacent skin did not show epidermal hyperplasia. At low magnification, there were no distinguishing features observed in the lesional and non-lesional eczema samples (Figure 16); high magnification histological findings in eczema samples are relatively non-specific and can include spongiosis (intercellular oedema of the epidermis) and an inflammatory cell infiltrate (Houck et al., 2004).

H&E images from each donor in the study, and 63 tissue sections overall, are shown below (Figure 17). Images show characteristic histological features between conditions that have been highlighted using the representative images, including epidermal hyperplasia in all psoriasis lesional samples. The panel of H&E images across the study also demonstrates that each tissue section is unique, reflecting various factors including donor, age, condition, sampling site and sectioning plane. Corresponding metadata for each sample is shown in Appendix M.

Figure 17: H&E staining of adult healthy, psoriasis and eczema tissue. Images from each donor and skin tissue section is shown.

3.2.5. Tissue annotation

The tissue sections were then histologically annotated using a software package called Loupe Browser (V6.1.0; 10x Genomics) in order to facilitate downstream analysis. Based on the H&E images, skin tissue sections were segmented into the following discrete compartments: epidermis, dermis, pilosebaceous units and adipose tissue. I was able to carry out these tissue annotations based on my previous extensive clinical training in dermatopathology. The following figure shows histologically annotated regions within representative tissue sections from healthy, psoriasis and eczema skin:
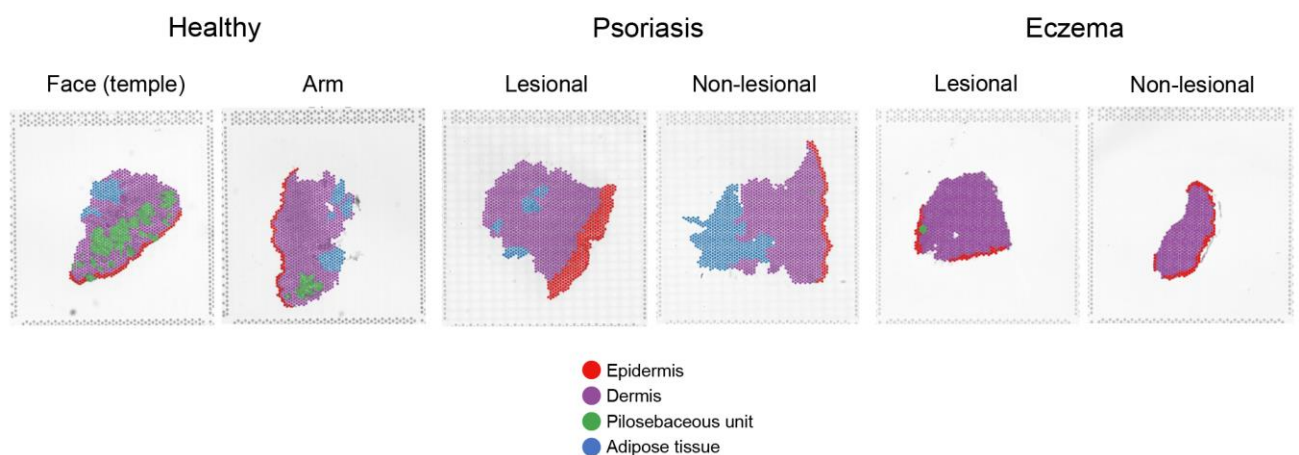
Figure 18: Histological tissue annotation. Representative images for healthy, psoriasis and eczema skin, showing segmentation into epidermal, dermal, pilosebaceous unit and adipose tissue regions.

Each tissue section has good representation of the epidermis and underlying dermis. The distribution of pilosebaceous units and adipose tissues varies between tissue sections, which is dependent on the anatomical site that has been biopsied and the plane in which the individual sample was sectioned.

Annotated sections from each donor and tissue section are shown below (Figure 19, Appendix M), highlighting the variable tissue morphology across the samples.

Figure 19: Histological tissue annotation. Images from each donor and skin tissue section is shown.

3.2.6. Computational quality control

The spatial transcriptomic sequencing data was then analysed in order to determine the quality of the data from each tissue section and determine whether any further filtering was required prior to downstream comparative analyses between healthy and diseased skin. The relative number of transcripts per spot was derived from the UMI (Unique Molecular Identifier) count and was compared across each of the tissue sections processed (Figure 20). Appendix M details the donor information that each tissue section on each slide corresponds to.
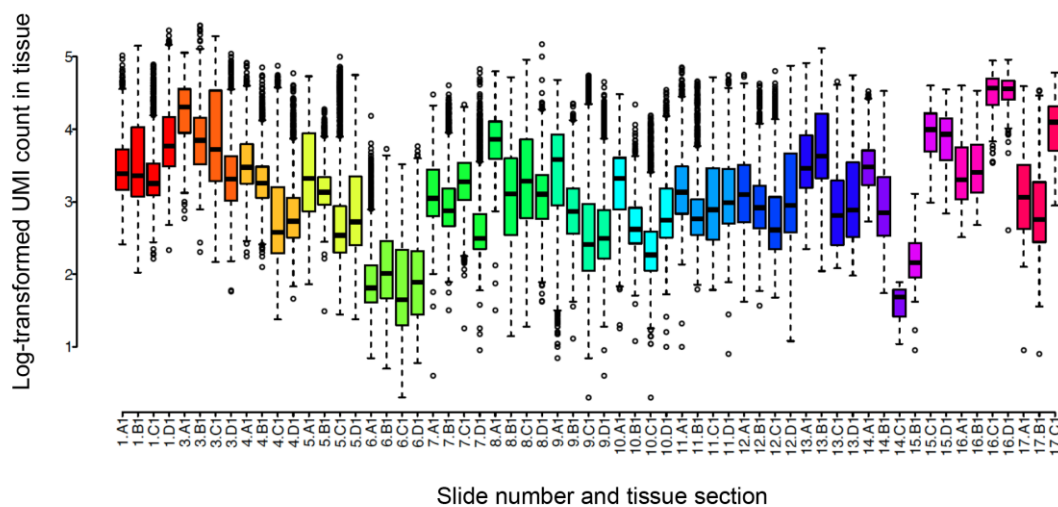
Figure 20: Spatial transcriptomic data quality control. Floating bar chart showing the log-transformed UMI count for each skin tissue section on each slide.

The quality of the data across the separate experiments was good overall, with a high average number of transcripts per spot across most slides and tissue sections (Figure 20). However, the data from slide six showed low UMI counts across all four tissue sections on the slide from two healthy, one psoriasis lesional and one psoriasis non-lesional sample (Appendix M), suggesting that a factor unrelated to the skin tissue samples contributed to the low data quality of this slide. Unfortunately, the technical replicates for these samples were run on Visium slide 2, for which the H&E imaging data was corrupted, as discussed in section 3.2.4. Poor quality data was also observed from two different eczema lesional samples that were processed on two different Visium slides (slide 14 position C1 from donor AD2634 and slide 15 position B1 from donor AD2638 (Appendix M).

The tissue annotations were then used to quantify whether the average number of transcripts per spot varied between epidermal, dermal, pilosebaceous unit and adipose tissue regions between different slides and tissue sections (Figure 21):
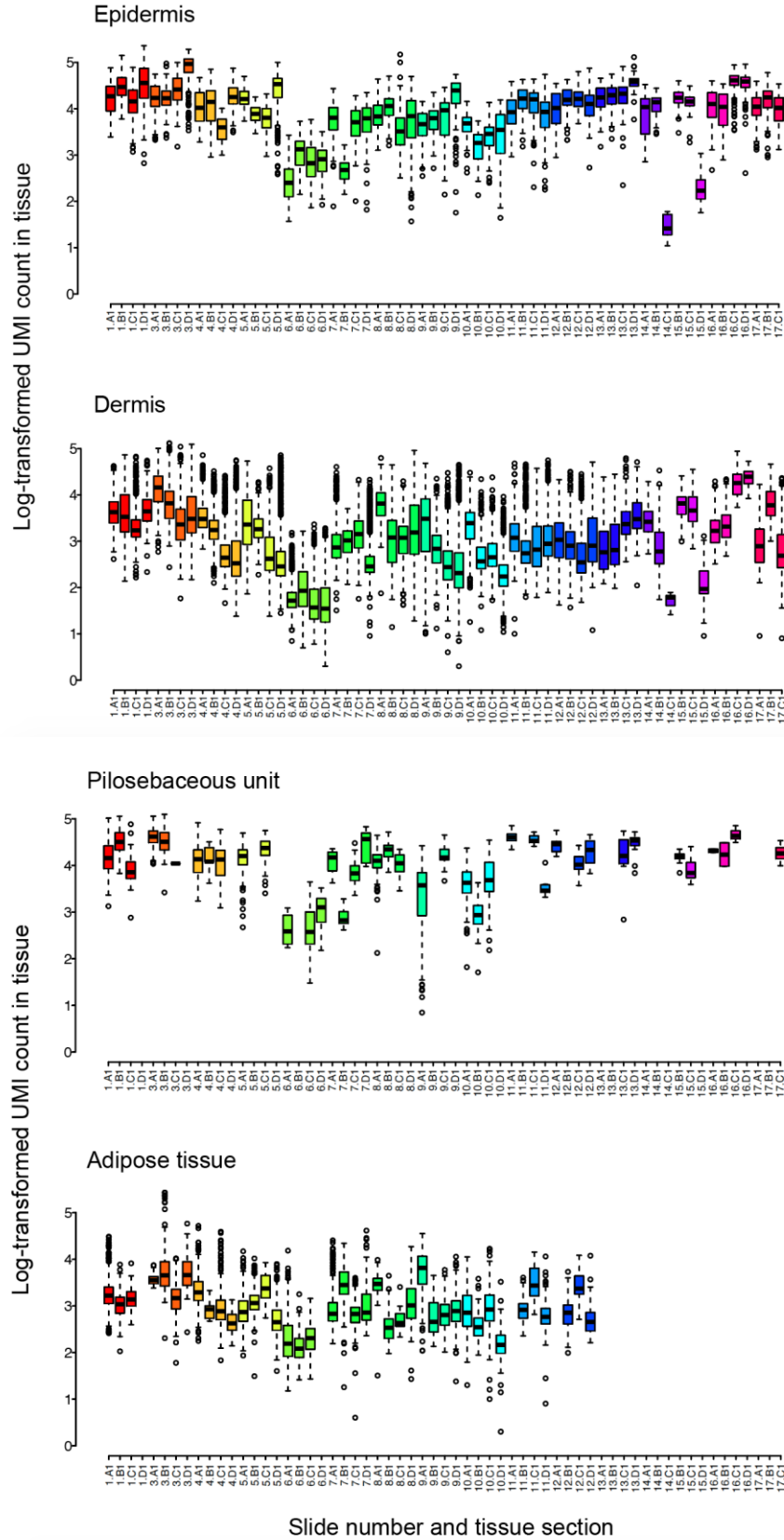
Figure 21: Spatial transcriptomic data quality control. Floating bar charts show the log-transformed UMI count across epidermal, dermal, pilosebaceous unit and adipose tissue regions for each skin tissue section on each slide.

Differences in the log-transformed UMI counts across tissue sections were observed depending on the skin tissue compartment analysed (Figure 21). The spatial transcriptomic data from the epidermal regions was overall of higher quality when compared to the dermal regions. Pilosebaceous and adipose tissue were not observed or annotated in all skin tissue sections. Data available from pilosebaceous tissue varied in quality between samples but were mostly of similar quality to the epidermal data, and data from adipose tissue showed similar log-transformed UMI counts to dermal regions.

Having analysed the spread of log-transformed UMI counts between each sample, the spatial distribution of UMI accounts across tissue sections was visualised next in order to further characterise the quality of the data (Figure 22) and take measures to address low quality regions.
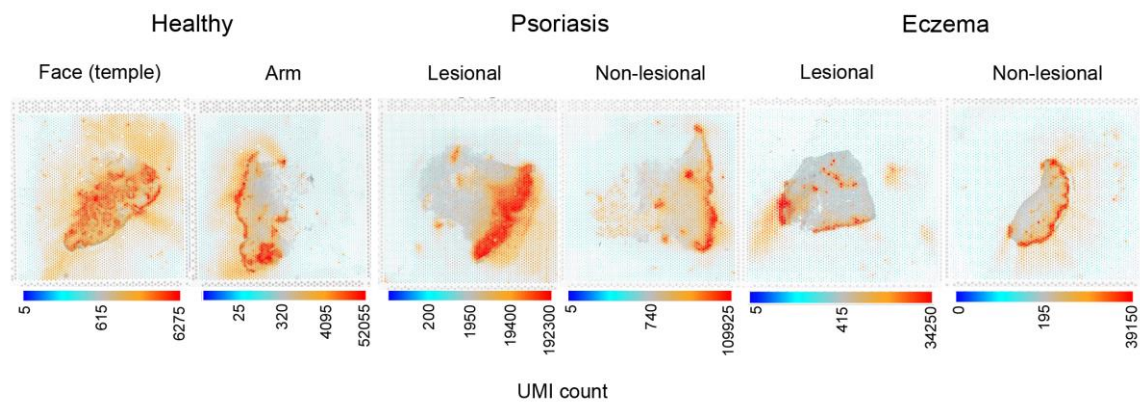
Figure 22: UMI counts across adult healthy, psoriasis and eczema skin tissue sections. Representative images for each condition are shown, highlighting the variation in UMI counts across the tissue.

Visualisation of the UMI count across representative tissue sections highlights the variation between samples and between tissue regions within each sample (Figure 22), which is in accordance with the data shown in Figure 20 and Figure 21. The epidermal and pilosebaceous regions overall showed an approximately 10-fold increase in UMI count compared to the cell sparse dermal and adipose regions (Figure 22). UMI counts across each tissue section in the study is shown below:
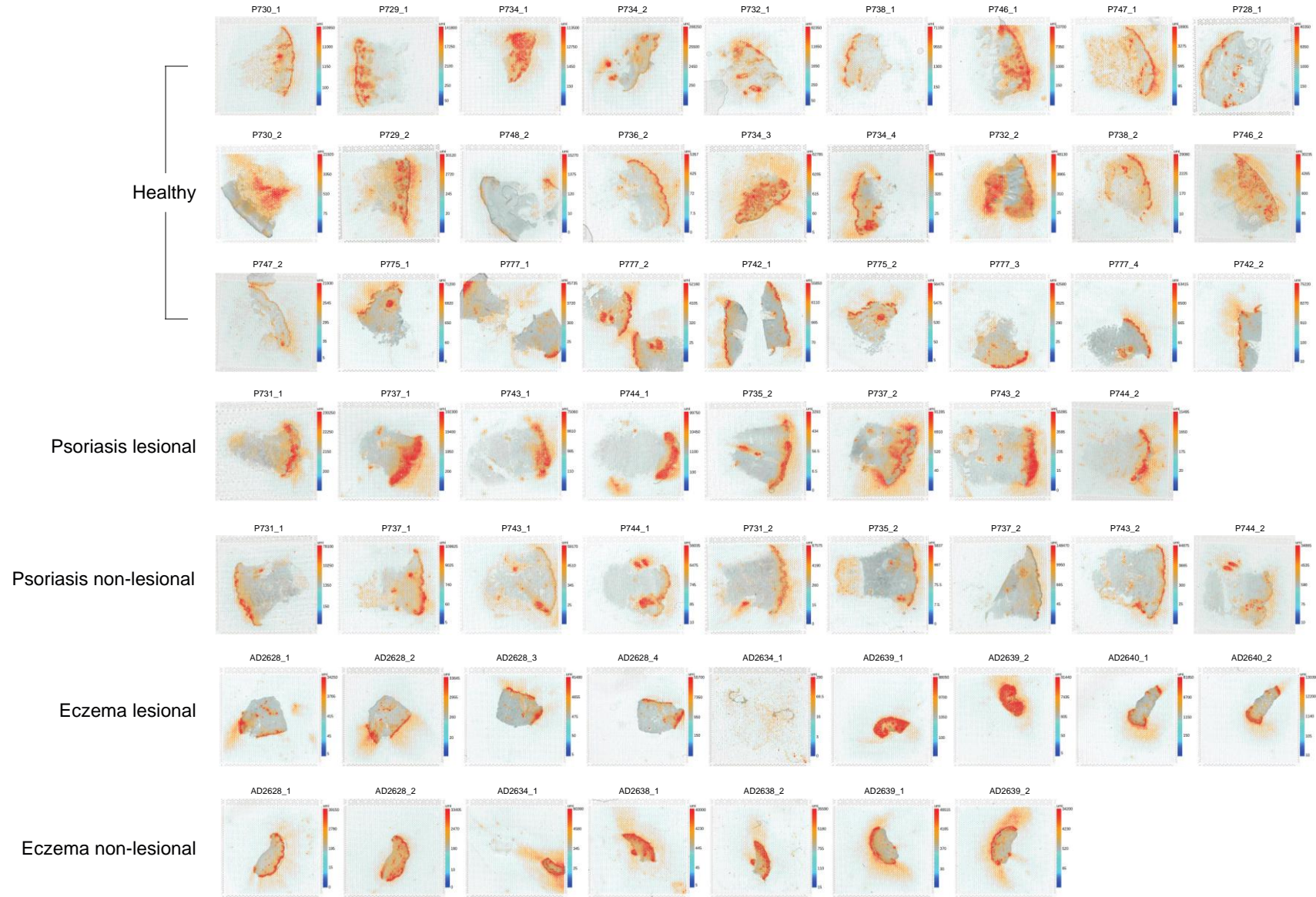
Figure 23: UMI counts across adult healthy, psoriasis and eczema skin tissue. Images from each donor and tissue section is shown.

In order to leverage scRNA-seq data to map cell states onto the spatial transcriptomic data with a high level of confidence, the individual spatial data regions require a sufficient number of transcripts for the analysis. Therefore, the UMI count per spot across each tissue section was first converted from a continuous scale into discrete categories: less than 100, 100 – 499, 500 – 999, 1,000 – 5,000 and above 5,000 (Figure 24). The Visium spots that contained fewer than 500 UMIs were then merged to form meta-spots by grouping up to seven spots together in a hexagonal layout with a central spot, with meta-spots containing between two to seven individual Visium spots. The output of spot merging is shown in Figure 25.

After merging the low coverage Visium spots into meta-spots, the issue of low UMI count spatial data predominantly within the cell sparse dermal regions is markedly improved (Figure 25). Most tissue sections consist of spots and meta-spots greater than 500 UMIs, with the exception of the four sections from experiment using slide six (two healthy samples from P748_2 and P736_2, and psoriasis lesional and non-lesional samples from P735_2), which is consistent with the overall quality control findings discussed above (Figure 20). Poor quality data was also observed for the eczema lesional sample AD2634_1 that would not be rectified using the spot merging procedure (Figure 20). These five samples were therefore excluded from further downstream analysis. Any residual spots that contain fewer than 500 UMIs across the included tissue sections following spot merging were also removed prior to ongoing analysis.
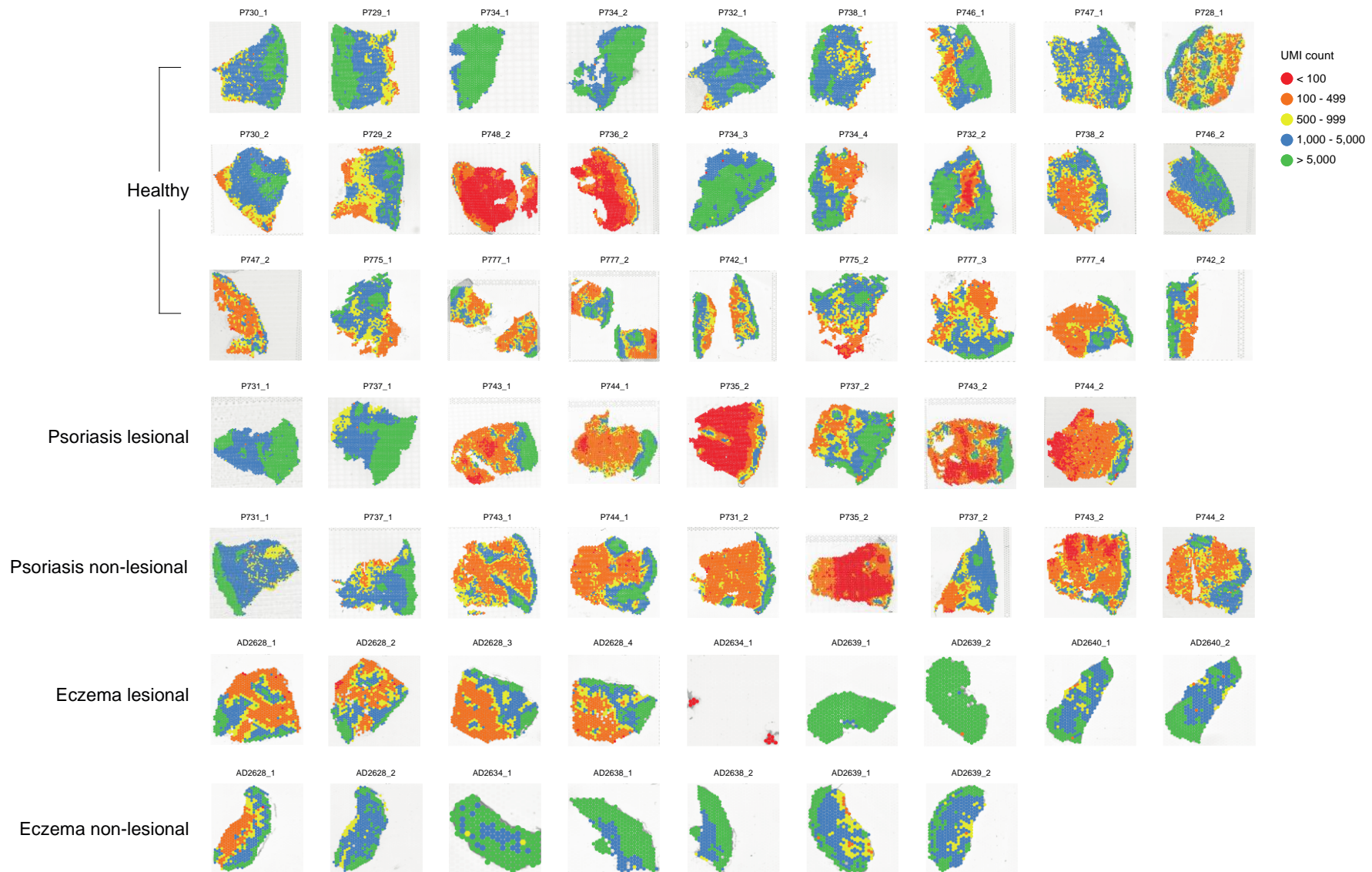
Figure 24: UMI counts across adult healthy, psoriasis and eczema skin tissue sections. Data is shown as a categorical variable prior to spot merging.
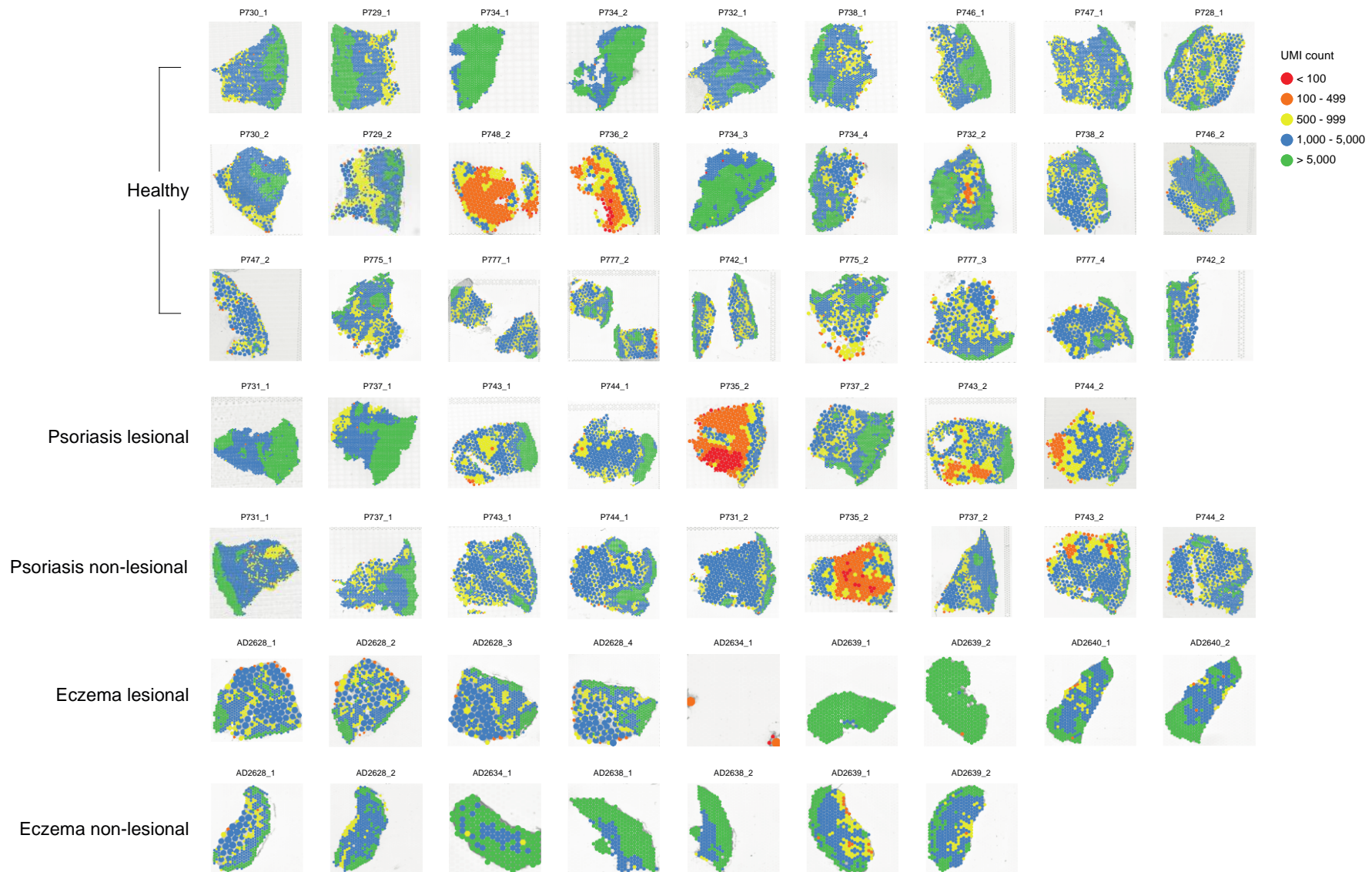
Figure 25: UMI counts across adult healthy, psoriasis and eczema skin tissue sections. Data is shown as a categorical variable following spot merging.

3.2.7. Integration of skin suspension reference datasets

After quality control and filtering of the dataset, the spatial transcriptomic data was next utilised to predict the cell states located within the tissue. A reference scRNA-seq dataset is required that is representative of the cell states expected in the spatial data. The analysis maps the gene expression signatures of specific cell types in the suspension dataset onto specific locations in the spatial dataset.

This spatial transcriptomic study follows on from our previous scRNA-seq skin cell atlas study that profiled skin from healthy donors and lesional/non-lesional skin from patients with psoriasis and eczema (Reynolds et al., 2021), and as such, the inclusion and exclusion criteria of the patients were the same across both phases of recruitment. However, in the scRNA-seq study, only the most superficial 200μm of tissue extending down to the superficial layer of the dermis was profiled, whereas I obtained and processed full thickness skin biopsies extending down to the deep dermis and, in some cases, the subcutaneous adipose tissue layer. Therefore, the scRNA-seq dataset does not include cell states found in deeper skin structures, such as hair follicles and sebaceous glands.

In order to create a reference scRNA-seq dataset that is representative of the spatial data that includes deeper skin structures including the hair follicle and its associated cell states, the scRNA-seq data from the Reynolds *et al.,* 2021 study was integrated with scRNA-seq data from an adult hair graft dataset (Takahashi et al., 2020). The data from this adult hair scRNA-seq study was also utilised in the comparative analyses between prenatal and adult hair follicles in Chapter 4. The following figure shows the output of the integration between both datasets:
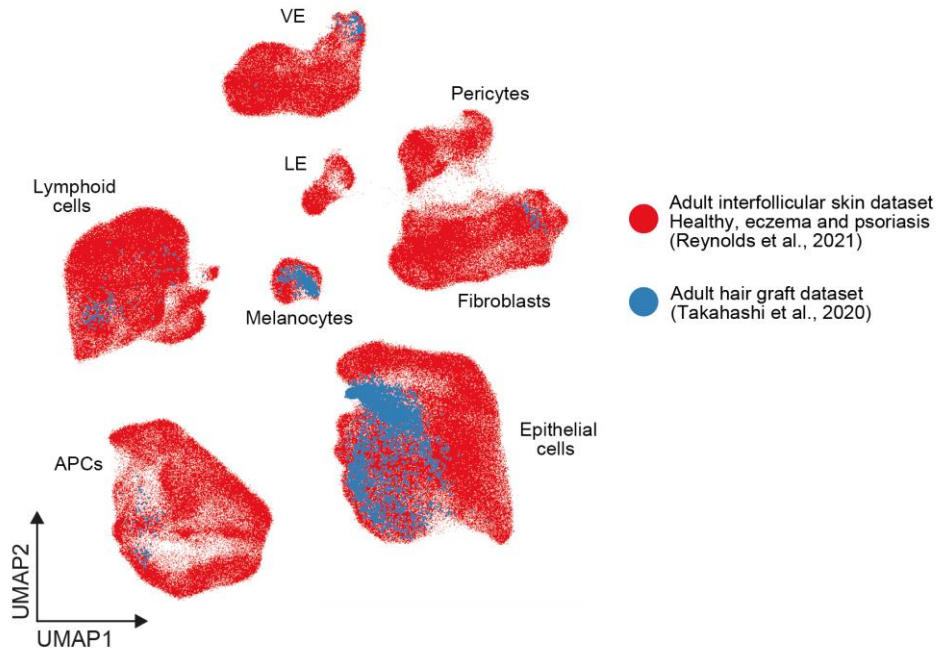
Figure 26: Integration between adult interfollicular and hair follicle scRNA-seq data. The UMAP visualisation shows the broad lineages of cell states from adult interfollicular skin (Reynolds et al., 2021) and adult hair grafts (Takahashi et al., 2020), coloured by dataset.

The integration between scRNA-seq datasets from adult interfollicular skin (Reynolds et al., 2021) and adult hair grafts (Takahashi et al., 2020) showed good mixing of cell states on the combined embedding space (Figure 26). Cell states across both datasets that belong to the same broad lineages clustered together, demonstrating similar transcriptional profiles for these cells across interfollicular skin and hair graft skin samples, without significant technical artefacts between the two studies.

As expected, the hair follicle cell states in the adult hair graft data clustered with the epithelial cells and keratinocytes in the interfollicular skin data (Figure 26). The hair-specialised cell annotations were retained in the final integrated object for spatial cell type deconvolution. However, in order to unify and align the non-follicular cell state annotations in the hair graft dataset with the interfollicular dataset, a label transfer approach was taken (Figure 27):
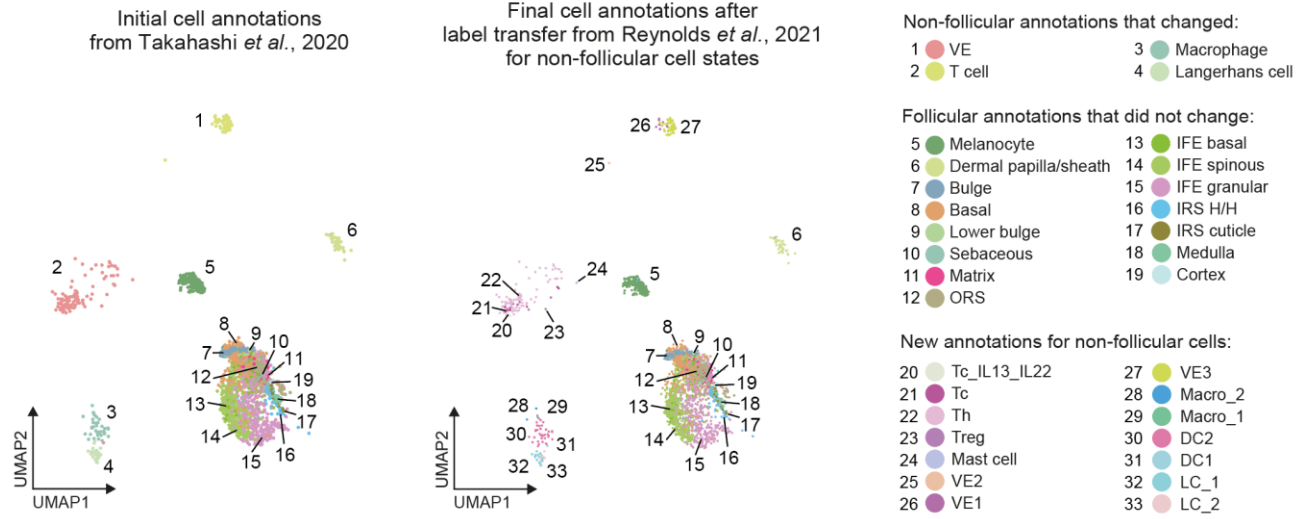
Figure 27: Label transfer of cell annotations between adult hair and interfollicular skin datasets. The original non-follicular cell annotations from the adult hair graft dataset (Takahashi et al., 2020) (left panel) were updated to align with the annotations in the interfollicular skin dataset (Reynolds et al., 2021) (right panel) using a single cell label transfer approach. VE = vascular endothelium, ORS = outer root sheath, IFE = interfollicular epidermis, IRS = inner root sheath, H/H = Henle/Huxley layers, Tc = cytotoxic T cell, IL = interleukin, Th = helper T cell, Treg = regulatory T cell, Macro = macrophage, DC = dendritic cell, LC = Langerhans cell.

The annotations for vascular endothelial cells, T cells, macrophages and Langerhans cells in the adult hair graft dataset (Takahashi et al., 2020) were updated and unified to reflect the higher resolution of cell annotations within the interfollicular skin dataset (Reynolds et al., 2021) (Figure 27). The high resolution hair follicle annotations from the adult hair dataset were retained in the final integrated object.

The final integrated dataset for spatial cell type deconvolution is shown in the following UMAP embedding, with hair follicles labelled as a broad cell type for this visualisation for ease of interpretation:
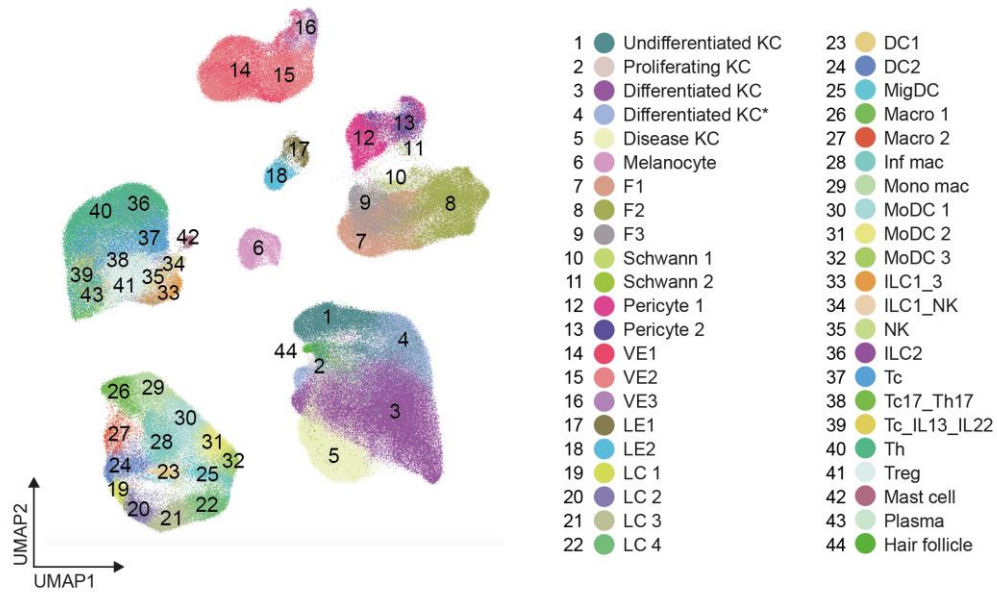
Figure 28: Final integrated dataset for spatial cell type deconvolution. UMAP visualisation shows unified cell annotations in the integrated scRNA-seq dataset between adult interfollicular skin (Reynolds et al., 2021) and adult hair grafts (Takahashi et al., 2020). KC = keratinocyte, F = fibroblast, LE = lymphatic endothelium, MigDC = migratory DC, Inf mac = inflammatory macrophage, Mono mac = monocyte derived macrophage, MoDC = monocyte derived dendritic cell, ILC = innate lymphoid cell, NK = natural killer cell.

3.2.8. Spatial cell type deconvolution

After generating an integrated scRNA-seq dataset that reflects the cell type composition of the spatial transcriptomic data from healthy, psoriasis and eczema skin samples, a spatial cell type deconvolution analysis called Cell2location (Kleshchevnikov et al., 2022) was carried out next in order to map cell states *in situ*. This analysis uses reference cell type signatures from the scRNA-seq data, which correspond to gene expression profiles of annotated cell states, and integrates this with the transcriptomic data from individual spatial locations. This then allows the absolute and relative abundance of each cell type to be defined within each Visium spot.

The results of the Cell2location deconvolution analysis for representative healthy, psoriasis and eczema samples are shown below:
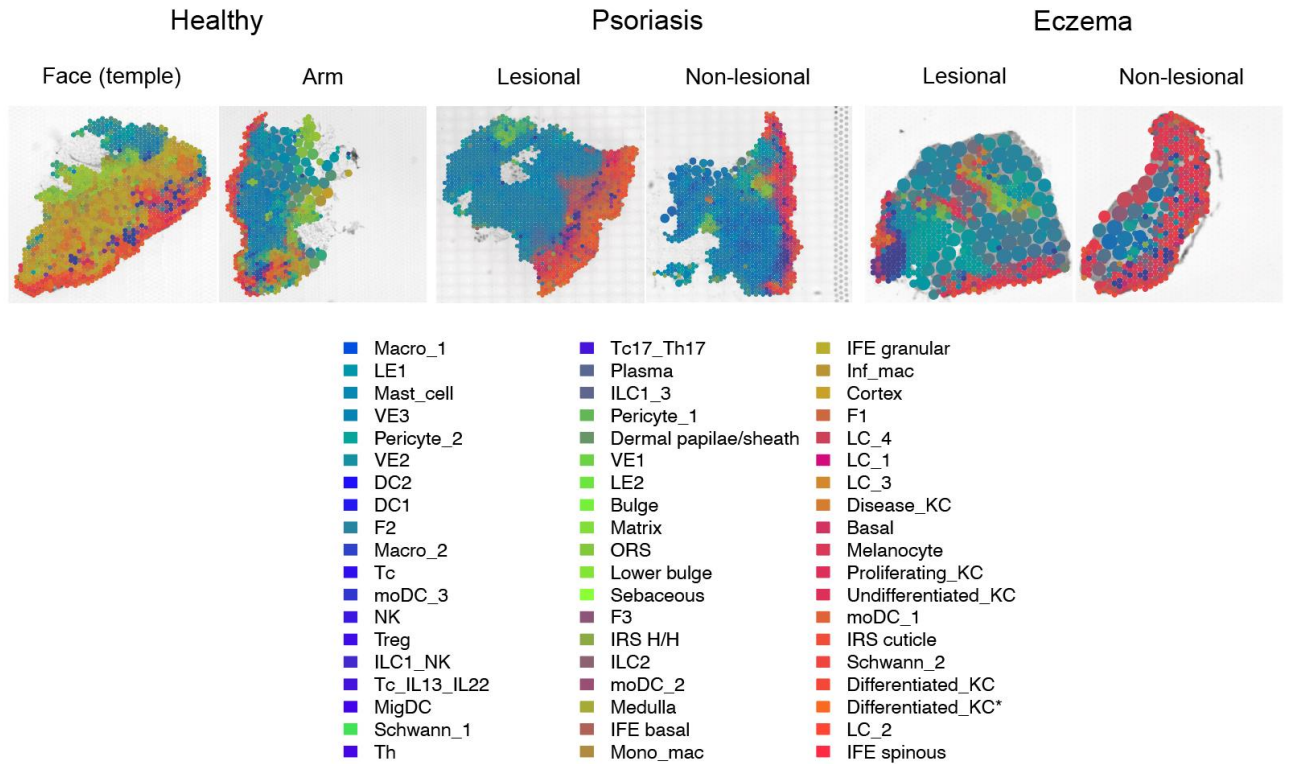
Figure 29: Spatial cell type deconvolution. Cell2location analysis was used to deconvolute single cell states onto spatial locations, with representative images for healthy, psoriasis and eczema skin shown.

The Cell2location analysis mapped single cell states from the integrated dataset onto the Visium spatial data, capturing the diverse cell types within discrete regions of the skin, including keratinocytes in the epidermis (Figure 29). The analysis was extended to include all the of tissue sections in the study, as shown in Figure 30. Due to the large number of reference single cell states mapped to each tissue section, further analyses were carried out next in order to characterise the changes in cell type composition and location between healthy and diseased skin.
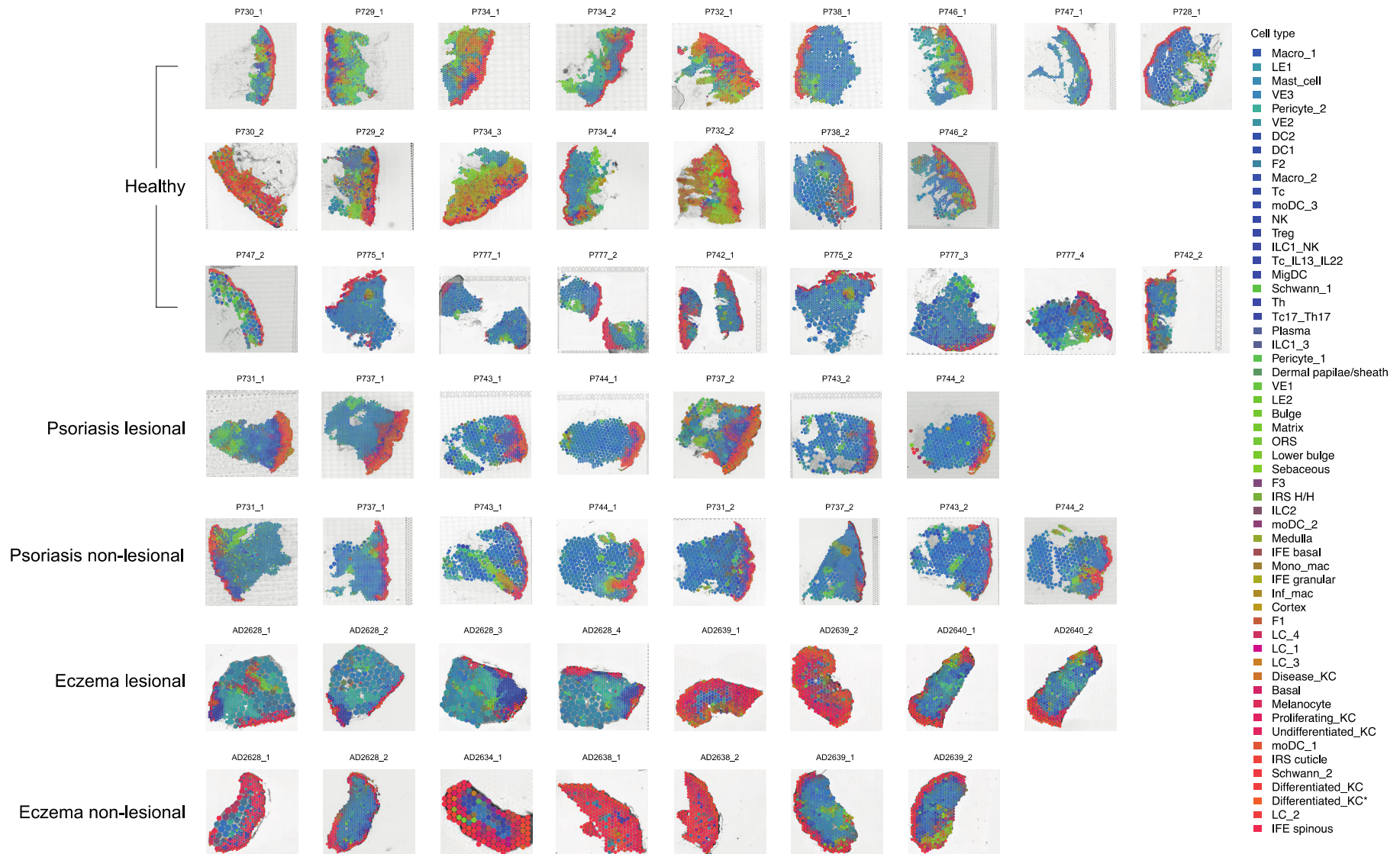
Figure 30: Spatial cell type deconvolution. Cell2location analysis for deconvolution of single cell states onto spatial locations for each donor and skin tissue section is shown.

3.2.9. Differential location of cell types from skin surface between disease conditions

The tissue annotations were next used in order to define the dermoepidermal junction (DEJ), which represents the interface between the epidermal and dermal skin layers. The DEJ was used to help interrogate the spatial transcriptomic data to provide microanatomical context to previously reported single cell findings (Reynolds et al., 2021) and to characterise the changes in cellular tissue locations across healthy and diseased skin. The DEJ was defined as the Visium spots that were annotated as epidermis that are in direct contact with spots annotated as dermis, as visualised below:
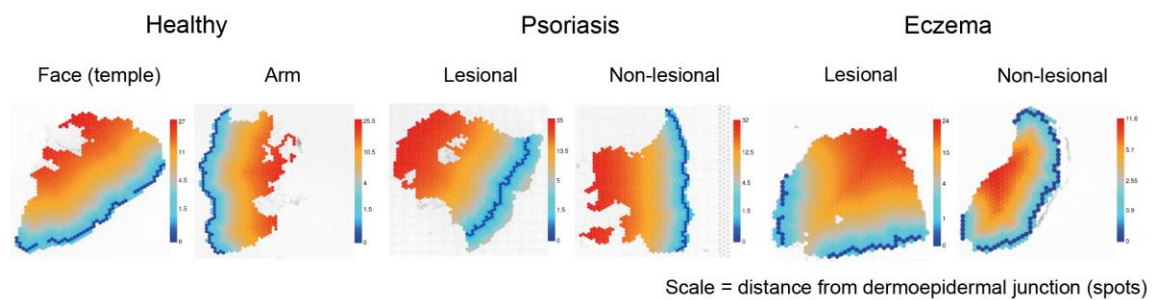
Figure 31: Defining the dermoepidermal junction (DEJ). Representative images for healthy, psoriasis and eczema skin, showing the DEJ as a dark blue line and the distance from the DEJ as a gradual scale.

The epidermis represents a very thin layer of the skin in healthy conditions, and therefore, the DEJ is shown to lie close to the border of the tissue section, similar to that seen in eczema and non-lesional psoriasis. However, in lesional psoriasis, the epidermis is markedly thickened, shown by the DEJ lying further from the tissue border than in healthy skin (Figure 31).

The DEJ and distance from the DEJ across each donor and tissue section in the study is shown below (Figure 32), noting that some eczema samples processed by CGaP were sectioned transversely rather than longitudinally (AD2639_1, AD2639_2 and AD2634_1).

Healthy

Psoriasis lesional

Psoriasis non-lesional

Eczema lesional

Eczema non-lesional

Scale = distance from dermoepidermal junction (spots)
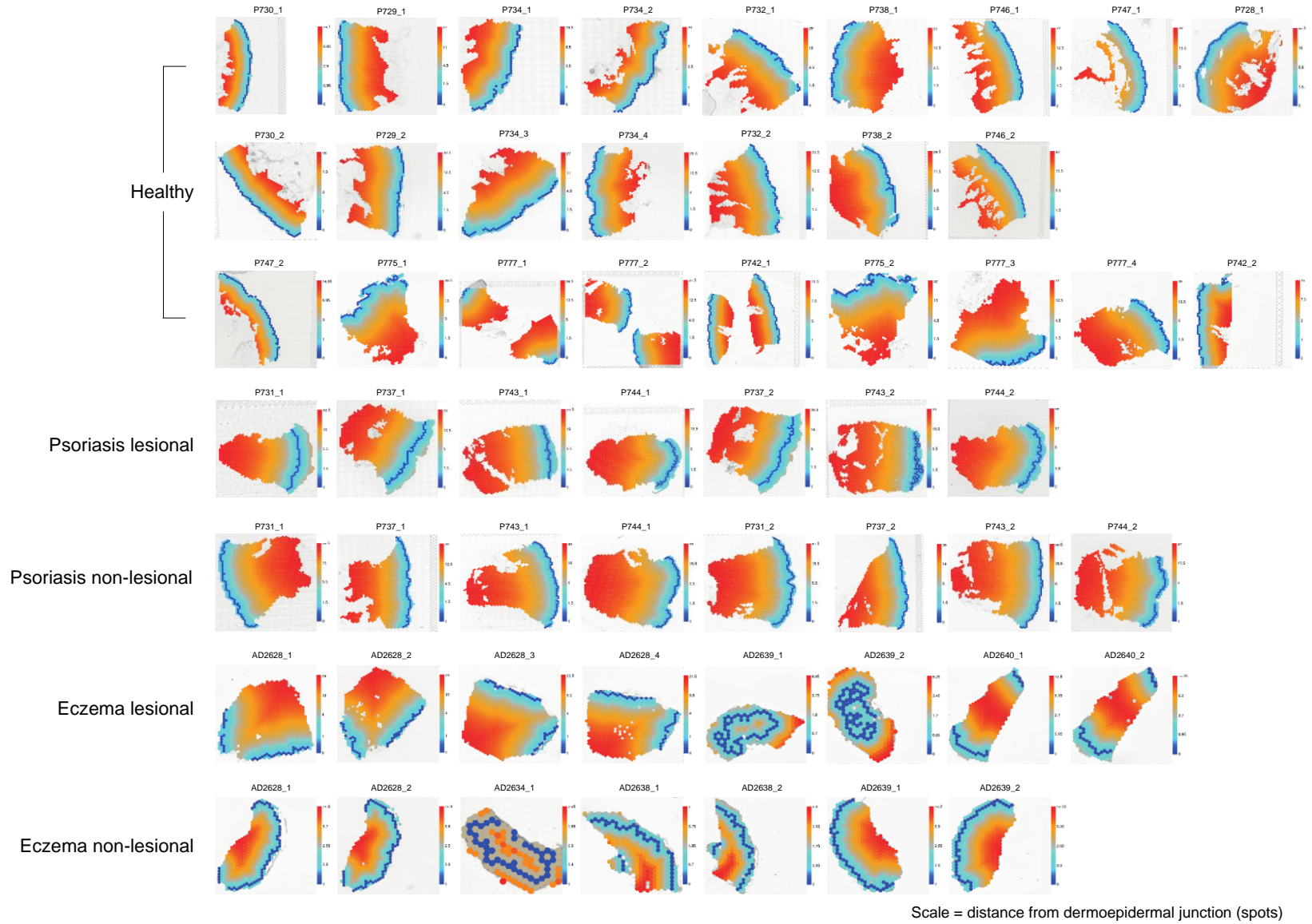
Figure 32: Defining the DEJ. Images from each donor and tissue section is shown.

The distance from the DEJ was then used as a measure to determine the distribution of cell states within the skin from superficial to deep. The relative abundance of each cell within spots at varying distances from the DEJ was first visualised for healthy skin, psoriasis lesional skin and psoriasis non-lesional skin separately (Figure 33). Across all tissue sections in this analysis, the annotated pilosebaceous units and three Visium spots surrounding these structures were excluded so as to reduce the confounding effect of cell types that are not spatially variable within the tissue. Cell states were grouped into epidermal non-immune cells, dermal non-immune cells, lymphoid and mast cells, and antigen presenting cells (Figure 33).

In order to further characterise the changes in cell type distribution and distance from the DEJ between each condition, differential analysis was carried out between pairs of conditions: psoriasis lesional versus healthy skin, psoriasis non-lesional versus healthy skin and psoriasis lesional versus psoriasis non-lesional skin (Figure 34).
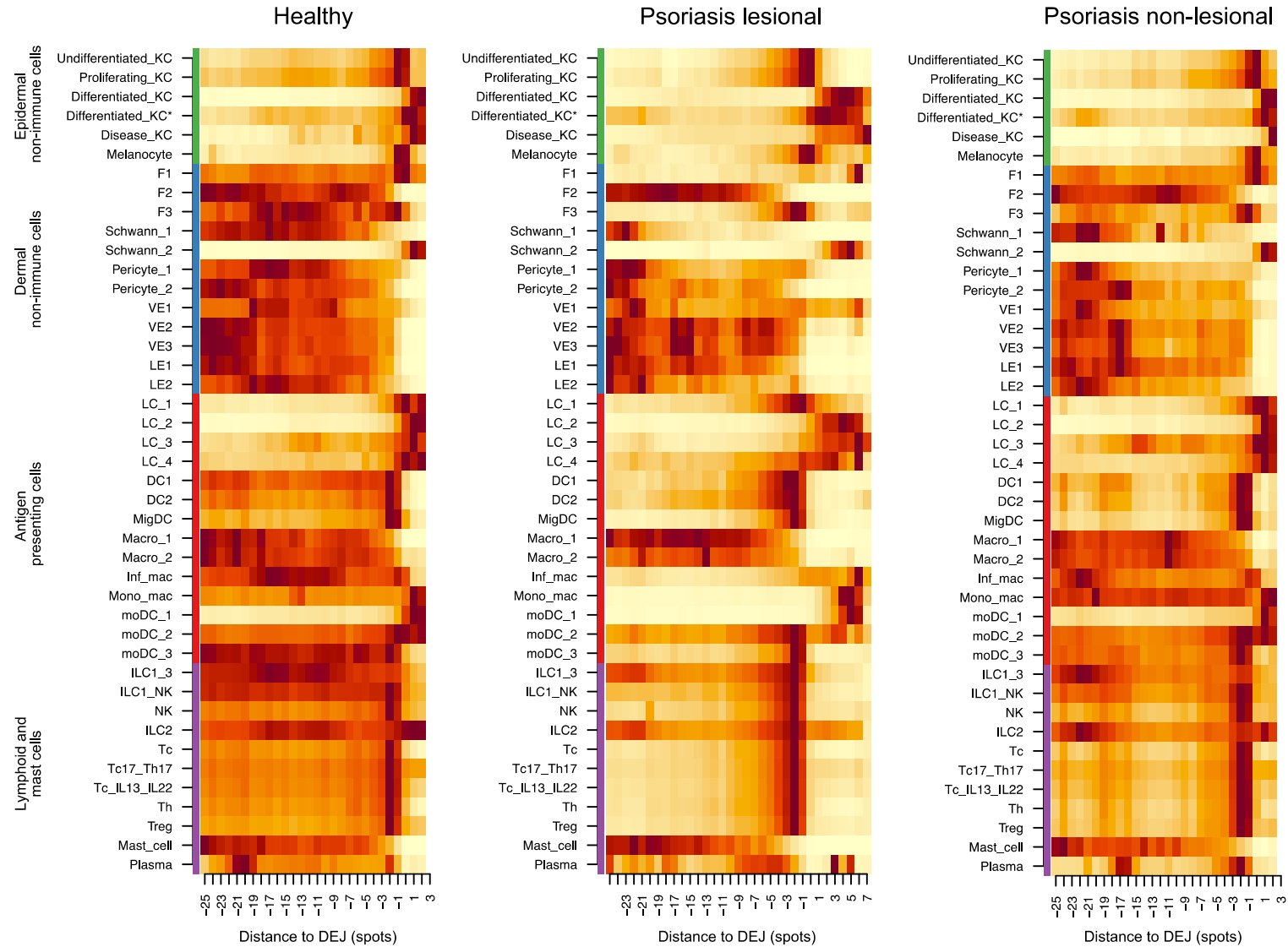
Figure 33: Cell distance from DEJ in healthy and psoriasis skin. Heatmap to show the relative row-normalised abundance of individual cell states within Visium spots located at different distances from the DEJ in healthy, psoriasis lesional and psoriasis non-lesional skin.
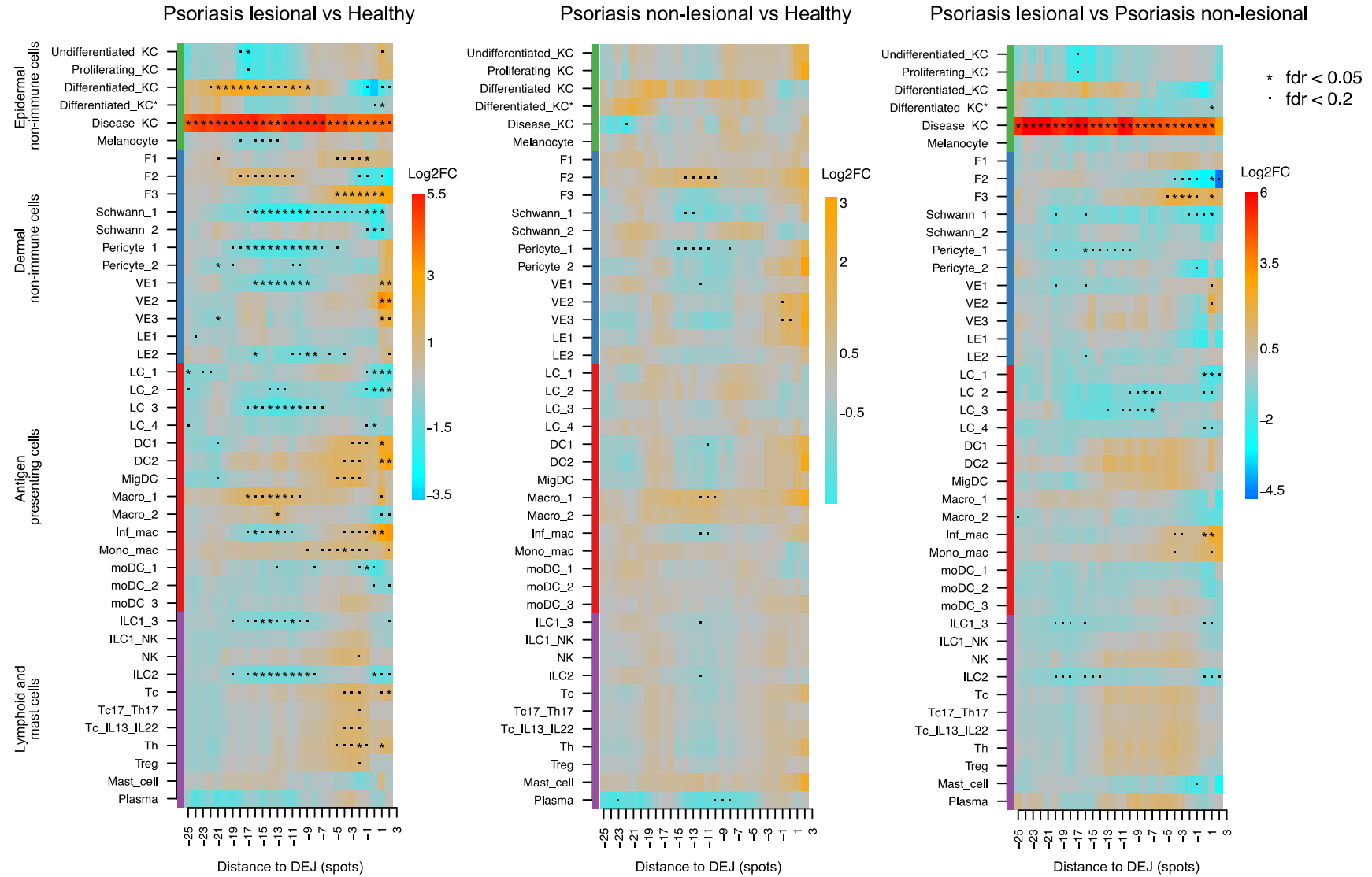
Figure 34: Differences in cell distance from the DEJ across healthy and psoriasis skin. Differential analyses were carried out to determine the changes in cell distance from the DEJ between different conditions. FC = fold change, fdr = false discovery rate.

The comparisons of the relative abundance of cell states at different distances from the DEJ between healthy, psoriasis lesional and psoriasis non-lesional skin revealed several statistically significant differences between these conditions (Figure 34). In comparison to healthy skin, psoriasis lesional skin exhibited a significant enrichment of disease keratinocytes, supporting findings from the scRNA-seq data where disease keratinocytes are mostly represented by psoriasis lesional and eczema lesional datasets. This explains why the enrichment of disease keratinocytes is shown to span the full depth of the skin tissue; an element of RNA leakage occurs across the Visium slide and disease keratinocytes are mostly absent in healthy skin, therefore, comparisons in abundance within deeper skin tissue levels also show a large fold change difference (Figure 34). The remaining cell states in the analysis, however, are well-represented between different conditions, allowing changes in their location and distance from the DEJ to be ascertained.

Differences in the location of fibroblasts were observed between psoriasis lesional and healthy skin, where fibroblast subtypes 1 and 3 were enriched in the superficial dermis of psoriasis lesional skin and fibroblast subtype 2 was enriched in the mid-dermis in psoriasis lesional skin in comparison to healthy skin. Based on analyses of the scRNA-seq dataset used for the spatial cell type deconvolution, fibroblast 2 is known to be the most abundant fibroblast cell type in diseased skin states compared to healthy skin, including psoriasis and eczema lesional and non-lesional skin (Reynolds et al., 2021). Fibroblast 2 highly expresses the chemokine *CXCL12,* which has a role in monocyte, dendritic cell, lymphocyte and neutrophil chemotaxis in psoriasis skin (Abdelaal et al., 2020), and also highly expresses *CCL19,* blockade of which leads to clinical remission of psoriasis lesions with anti-TNF (tumour necrosis factor) treatment (Bosè et al., 2013). Fibroblasts 1 and 3, enriched in the superficial dermis of psoriasis lesional skin compared to healthy skin (Figure 34), are both known to highly express *IL6* (interleukin-6), which interestingly has been shown to stimulate the proliferation of cultured keratinocytes and contribute to the epidermal hyperplasia seen in psoriasis plaques (Grossman et al., 1989). *IL6* also promotes neutrophil differentiation and activation (Bartoccioni et al., 2003; Kaplanski et al., 2003), and the abundance of neutrophils in psoriasis lesions that form intraepidermal collections known as Munro's microabscesses are a histopathological hallmark of the disease (Chiang et al., 2019). However, the reference scRNA-seq dataset did not include neutrophils from adult skin, as neutrophils and other granulocytes are technically challenging to capture to due to

their lower RNA content and their high RNAase levels (Ratnasiri et al., 2023). The spatial deconvolution, therefore, did not map neutrophils to the skin tissue samples.

The vascular endothelial cells also showed an altered microanatomical localisation in psoriasis lesional skin compared to healthy skin (Figure 34), with all three subsets observed to be enriched within the superficial skin regions in disease. This is in accordance with the histological findings in psoriasis lesional skin, where dilated and elongated capillaries are observed in the papillary dermis (Heidenreich, Röcken and Ghoreschi, 2009). Clinically, pinpoint bleeding spots can appear when psoriasis scales are scraped off, known as Auspitz's sign (Nasca et al., 2019), owing to the proximity of the expanded capillary network to the skin surface at sites where the suprapapillary epithelium is thinned. Dermal vascular expansion in psoriasis is a relatively under-researched aspect of the disease (Malecic and Young, 2017), but is known to be mediated by VEGF (vascular endothelial growth factor) produced by keratinocytes, fibroblasts and immune cells, such as mast cells (Detmar et al., 1995; Yan et al., 2018; Wernersson and Pejler, 2014), promoting the recruitment of immune cells to psoriasis skin lesions.

Interestingly, Langerhans cell subtypes 1 and 2 were observed to be enriched in the epidermal regions of healthy skin when compared to psoriasis lesional skin (Figure 34). This suggests a process of Langerhans cell migration away from the epidermis in psoriasis, supported by Figure 33, which shows that Langerhans cell subtypes 1 and 2 are most abundant within the superficial dermis in psoriasis lesional skin. Langerhans cells are tissue-resident antigen presenting cells that are normally distributed in the epidermis and mediate immune responses to inflammation, as well as microorganisms, allergens and ultraviolet radiation (Yan et al., 2020). There have been conflicting reports regarding the localisation of Langerhans cells detected with immunostaining or confocal microscopy in psoriasis skin lesions, with some studies detecting an increase in Langerhans cell density in psoriasis affected epidermis (Fujita et al., 2011; Eaton et al., 2014), some reporting a decrease in density (Bos et al., 1983; Glitzner et al., 2014) and some reporting stable densities compared with unaffected skin (Gommans, van Hezik and van Huystee, 1987; Czernielewski et al., 1985). There is marked interindividual variation in Langerhans cell density within the skin across healthy and psoriasis lesional skin samples, therefore, it is possible that the conflicting results may be a result of underpowered studies.

Psoriasis lesional skin also exhibited a significant enrichment of dendritic cell (DC) subsets 1 and 2 in the epidermis and superficial dermis compared to healthy skin, and migratory DCs in the superficial dermis of psoriasis skin (Figure 34). Macrophages also show altered cell localisation within the tissue samples, with inflammatory macrophages enriched within psoriasis epidermis compared to healthy skin, and monocyte derived macrophages and macrophage subsets 1 and 2 enriched within psoriasis lesional dermis compared to healthy skin (Figure 34). DCs and macrophages are both known to play an important role in the initiation phase of psoriasis inflammation (Kamata and Tada, 2022). Damaged keratinocytes are known to release nucleic acids and antimicrobial peptides that lead to the activation of DCs and macrophages, which in turn then produce tumour necrosis factor (TNF) alpha and interferon (IFN) alpha (Glitzner et al., 2014; Clark and Kupper, 2006). Resident dermal DCs then undergo maturation and differentiation, subsequently producing various cytokines including IL-23 and TNF-α, which promotes the differentiation of naïve T cells into Th1, Th17 and Th22 cells (Lowes et al., 2008; Hu et al., 2021). The underlying pathophysiology is psoriasis is supported by the findings of immune subset enrichment within the superficial regions of lesional skin compared to healthy skin.

The lymphoid cells also showed altered microanatomical location between healthy and psoriasis lesional skin. Innate lymphoid cells in the reference scRNA-seq dataset included a subtype annotated as ILC1/3, reflecting the gene expression signature which shares the characteristics of ILC1s and ILC3s, with plasticity between these lymphoid cell populations having been previously reported (Bernink et al., 2015). ILC2 cells are also present, which highly express *IL7R* and *PTGDR2* (Reynolds et al., 2021). The scRNA-seq analysis did not detect a significant difference in the abundance of ILC1/3 and ILC2 cell states between healthy and diseased skin (Reynolds et al., 2021), whereas the spatial transcriptomic analysis revealed that these cell states are both more significantly enriched in the mid-dermis of healthy skin compared to psoriasis lesional skin (Figure 34). The heatmap analysis of cell distance from the DEJ shows that ILC1/3 and ILC2 in psoriasis lesional skin are most highly enriched in the superficial and deep dermis (Figure 33), suggesting migration of these cells in disease. ILC3s are known to produce IL17 and IL22 (Villanova et al., 2014), which drive psoriasis pathogenesis and lead to keratinocyte activation and proliferation (Bugaut and Aractingi, 2021).

Lymphoid cells that showed a significant enrichment in the superficial dermis of psoriasis lesional skin compared to healthy skin included cytotoxic T cells (Tc), T helper cells (Th), Tc17/Th17 cells, Tc IL13/IL22 cells and Tregs (Figure 34). Tc IL13/IL22 cells are not expected in psoriasis lesional skin, with analysis of the scRNA-seq data showing that Tc IL13/IL22 cells are present in eczema skin and are not present in healthy and psoriasis skin (Reynolds et al., 2021). However, the spatial deconvolution analysis predicts cell type mapping based on transcriptional similarity between the scRNA-seq and spatial datasets, with some Tc IL13/IL22 cells being mapped onto healthy and psoriasis skin tissue sections (Figure 30). The other lymphoid cells have been shown by scRNA-seq to be present in healthy and psoriasis skin, with an expansion of Tc17/Th17 single cells reported in psoriasis lesional skin (Reynolds et al., 2021). The enrichment of lymphoid cells in the superficial dermis of psoriasis lesional skin is in keeping with histopathological findings, where a lymphocytic inflammatory cell infiltrate is observed (Mihu et al., 2021). Tc, Th, Tregs and Tc17/Th17 cells are known to be key components underlying psoriasis pathogenesis due to their aberrant release of cytokines including IL23, IL17 and TNF-α (Liu et al., 2021; Nussbaum, Chen and Ogg, 2021), which are targeted using biologic therapies (Yiu et al., 2019; Egeberg et al., 2018).

Notably, across the rest of the cell states mapped onto the spatial transcriptomic data, there were no significant differences observed in the distribution of several cell types when comparing psoriasis lesional skin and healthy skin, including melanocytes, lymphatic endothelium, mast cells and plasma cells (Figure 34). These cell types are not recognised to be important contributors to the pathogenesis of psoriasis, which supports the findings here.

Psoriasis non-lesional skin samples were also compared with healthy skin samples to determine the differences in cell distance from the DEJ across all mapped cell states (Figure 33 and Figure 34). Biopsies from non-lesional sites were taken from clinically uninvolved skin adjacent to psoriasis plaques, and the H&E staining of these samples did not show marked changes compared to healthy skin (Figure 16 and Figure 17). However, the spatial transcriptomic analysis did reveal several significant differences between psoriasis non-lesional skin and healthy, which were less widespread than the comparison between lesional and healthy skin (Figure 34). Compared to healthy skin, psoriasis non-lesional skin

showed a significant enrichment of fibroblast subtype 2 in the mid-dermis, vascular endothelial cell subtypes 2 and 3 in the superficial dermis and macrophage 1 in the mid dermis. ILC1/3 and ILC2 were also more significantly enriched in the mid-dermis of healthy skin samples compared to the psoriasis non-lesional samples (Figure 34). These changes in cell distance from the DEJ demonstrate the early microanatomical changes in clinically uninvolved skin in patients with psoriasis. Analysis of the scRNA-seq dataset from which the spatial data was mapped did also show a significant increase in the abundance of fibroblast 2 and vascular endothelial cell 2 and 3 in psoriasis non-lesional compared to healthy skin, however, no significant change in abundance was observed between macrophage and ILC subsets (Reynolds et al., 2021). This highlights the additional insights that can be gained from analysing cell states within their tissue microenvironmental context, as opposed to scRNA-seq data alone where isolated cells in suspension are analysed.

The distance of cell states from the DEJ was also compared across psoriasis lesional and psoriasis non-lesional skin sections in order to reveal the cellular microanatomical changes that occur between clinically uninvolved skin and skin from lesional psoriasis inflamed plaques. This analysis showed that disease keratinocytes were significantly enriched in psoriasis lesional skin compared to non-lesional skin, as was observed in the comparison between psoriasis lesional and healthy skin (Figure 34). Furthermore, in comparison to psoriasis non-lesional skin, psoriasis lesional skin showed a significant enrichment of fibroblast 3 in the superficial dermis, vascular endothelial cell 1 and 2 in the superficial dermis, and inflammatory macrophages and monocyte derived macrophages in the epidermis and superficial dermis (Figure 34). Langerhans cell subtype 1 also showed a significant enrichment in the epidermis of psoriasis non-lesional compared to lesional skin, and subtypes 2 and 4 were more enriched in the superficial dermis of psoriasis non-lesional skin, suggesting a process of Langerhans cell migration, as discussed above when examining psoriasis lesional skin findings compared to healthy skin. By comparing psoriasis non-lesional and lesional skin, the microanatomical changes underlying pathogenic processes in development of lesional plaques can be further understood.

The eczema skin samples were next examined using the same approach taken for psoriasis skin samples. The relative abundance of mapped cell states at varying distances from the DEJ was initially visualised for healthy skin, eczema lesional skin and eczema non-lesional

skin separately (Figure 35). Differential analysis on the cell distance from the DEJ was then carried out between pairs of conditions: eczema lesional skin versus healthy skin, eczema non-lesional versus healthy skin and eczema lesional versus eczema non-lesional skin (Figure 36).
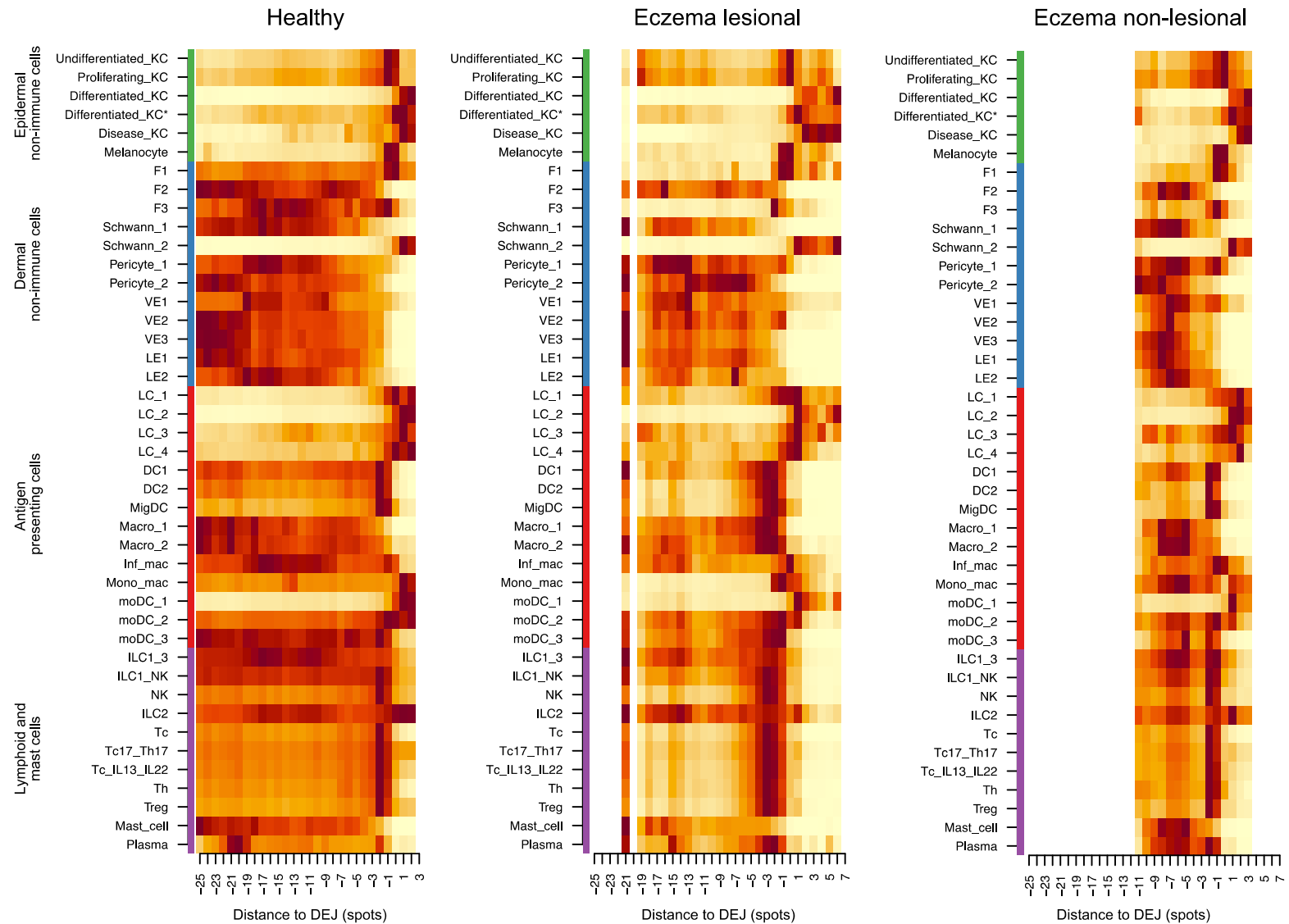
Figure 35: Cell distance from DEJ in healthy and eczema skin. Heatmap to show the relative row-normalised abundance of individual cell states within Visium spots located at different distances from the DEJ in healthy, eczema lesional and eczema non-lesional skin.
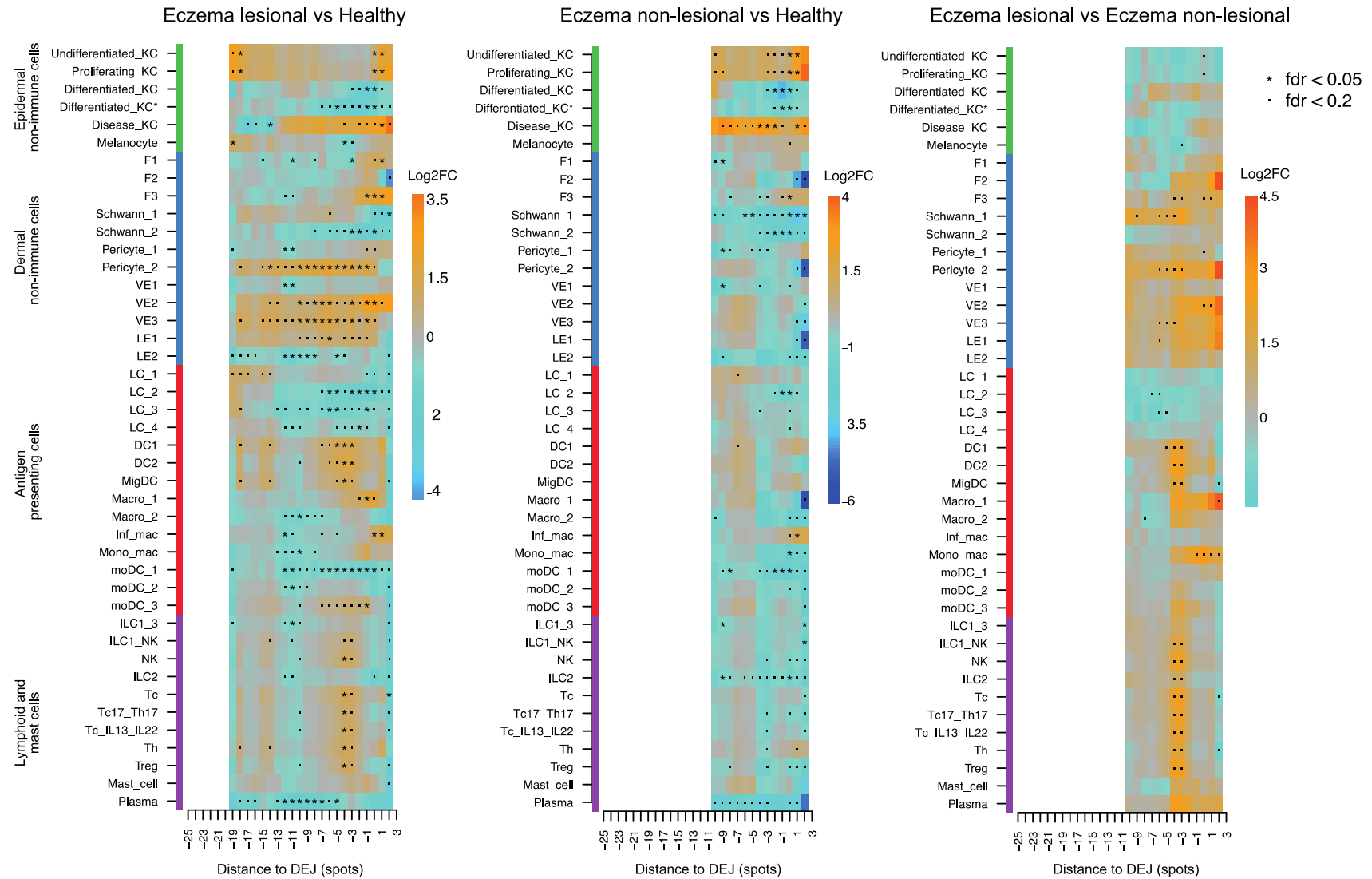
Figure 36: Differences in cell distance from the DEJ across healthy and eczema skin. Differential analyses were carried out to determine the changes in cell distance from the DEJ between different conditions. FC = fold change, fdr = false discovery rate.

Widespread changes across most cell states were observed when comparing cell distance from the DEJ between healthy and eczema lesional skin (Figure 36). The disease keratinocytes were significantly enriched in the epidermis of eczema lesional skin compared to healthy skin. However, an element of RNA leakage was apparent, especially considering that disease keratinocytes are markedly underrepresented in healthy skin, which leads to an increased log fold change between conditions even when the absolute numbers of mapped disease keratinocytes within an eczema dermal spot are low.

Fibroblast subtypes 1 and 3 were observed to be enriched within the superficial skin regions in eczema lesional skin compared to healthy, with no significant differences in cell distance from the DEJ identified for the fibroblast 2 population (Figure 36). This pattern of fibroblast 1 and 3 enrichment in eczema was also observed in psoriasis lesional skin compared to healthy skin (Figure 34). Both subtypes of fibroblasts are known from scRNA-seq analysis to highly express *IL6*, which was discussed in the context of psoriasis pathogenesis above, as well as *MMP2* (matrix metallopeptidase 2) (Reynolds et al., 2021). IL6 is also released as part of the cytokine milieu that drives eczema, with previous reports focusing on the T helper 2 (T_H2) polarised CD4$^+$ T cells that overproduce IL6, IL5, IL4 and IL13 (Toshitani et al., 1993; Werfel et al., 2016). Studies have shown that activated macrophages and monocyte derived DCs also secrete IL6 in atopic individuals (Lee et al., 1992; Su et al., 2020). IL6 exerts its effect by activating IL4-producting CD4+ T cells, inducing the development of Th17 cells from naïve T cells and promoting the acute phase eczema response (Kimura and Kishimoto, 2010; Rincón et al., 1997). IL6 has been targeted therapeutically using a monoclonal antibody called tocilizumab that blocks the IL6 receptor, which has been shown to improve eczema severity, however also promotes bacterial superadded infection of the skin (Navarini, French and Hofbauer, 2011); its use has not be adopted in routine clinical practice (Choong and Tan, 2021).

Fibroblast subsets 1 and 3 enriched in the superficial dermis of eczema lesional skin compared to healthy skin (Figure 36) also highly express *MMP2* (matrix metallopeptidase 2) (Reynolds et al., 2021). MMP2 is known to be produced by fibroblasts in response to TNF-α and mediates proteolytic activity in the skin that contributes to the skin barrier dysfunction observed in eczema (Han et al., 2001; Xu et al., 2014). MMP2 has been shown to be raised in the serum of patients with eczema and the levels correlate with disease

activity scores in moderate and severe eczema (Basałygo et al., 2021). Transepidermal water loss, which is associated with skin barrier dysfunction in eczema, has interestingly been shown to be dramatically reduced with MMP2 inhibition (Li et al., 2010), highlighting the pathogenic role that MMP2-secreting fibroblasts 1 and 3 enriched in the superficial dermis has in eczema lesional skin (Figure 36).

The cell distance from the DEJ comparison between eczema lesional skin and healthy skin also revealed an altered microanatomical localisation of vascular endothelial cells and pericytes (Figure 36). Vascular endothelial cell subtypes 2 and 3 and pericyte subtype 2 showed widespread enrichment in the superficial and mid-dermis of eczema lesional skin. The changes in cell localisation of vascular endothelial cell subtype 1 were less widespread, with an enrichment in the mid-dermis in healthy skin compared to eczema lesional skin (Figure 36). The findings in eczema skin were markedly different to those observed when comparing psoriasis lesional and healthy skin, where all three vascular endothelial cell subsets were enriched in a very localised region close to the skin surface (Figure 34). Endothelial network expansion throughout the dermis in eczema lesional skin reflects clinical signs on examination, including erythema, oedema and white dermographism, a blanching response after firm skin stroking secondary to capillary vasoconstriction (Steinhoff et al., 2006). Previous scRNA-seq analyses have shown that vascular endothelial cell 3 is significantly more abundant in eczema lesional and non-lesional skin compared to healthy skin, whereas the abundance of subtypes 1 and 2 are not significantly different (Reynolds et al., 2021), demonstrating that spatial transcriptomics reveals additional insights into changes in microanatomical localisation that cannot be determined from cell suspension data. Interestingly, vascular endothelial cells 2 and 3 both highly express *ACKR1* (Atypical Chemokine Receptor 1), whereas *ACKR1* is not expressed by vascular endothelial cell 1 (Reynolds et al., 2021). Cell-cell communication analyses have revealed that endothelial cells expressing *ACKR1* interact with macrophage 2 via *CXCL8* (C-X-C Motif Chemokine Ligand 8) to promote lymphocyte recruitment into inflamed skin, and that this signalling program was also identified in developing first trimester prenatal skin, suggesting that developmental cell programs can be co-opted in the pathogenesis of inflammatory skin disease (Reynolds et al., 2021). Therefore, the enrichment of *ACKR1*-expressing vascular endothelial cells 2 and 3 throughout the dermis in eczema lesional skin compared to healthy skin (Figure 36) is likely to drive the disease process by interacting with macrophages to promote lymphocyte recruitment.

The spatial location of lymphatic endothelial (LE) cells in eczema lesional skin was also shown to be perturbed compared to healthy skin: LE subtype 1 showed an enrichment within the superficial and mid-dermal regions in eczema, whereas LE subtype 2 was not enriched in eczema compared to healthy skin (Figure 36). LE1 is characterised by high expression of the chemoattractant *CCL21* (C-C Motif Chemokine Ligand 21), which facilitates DC migration into skin draining lymph nodes (Gupta et al., 2006). LE1 also highly expresses angiogenic factors *CAVIN2* (Caveolae Associated Protein 2) and *CCND1* (Cyclin D1) that promote endothelial network expansion and inflammatory cell recruitment to the skin (Tammela and Alitalo, 2010), therefore supporting the role of LE1 in eczema pathogenesis and highlighting the significance of LE1 enrichment in the superficial and mid-dermis of eczema lesional skin (Figure 36).

Similar to that observed in psoriasis lesional skin, the Langerhans cells in eczema lesional skin also showed altered microanatomical location. In healthy skin, Langerhans cells were enriched within the epidermis, however in eczema lesional skin, Langerhans cells showed high enrichment close to the DEJ within the basal epidermis or within the superficial dermis, with all four Langerhans cell subsets showing a similar pattern of altered localisation (Figure 35 and Figure 36). In healthy skin conditions where the skin barrier is intact, Langerhans cells are predominantly quiescent and mediate tolerance to self antigens (Seneschal et al., 2012). In the context of a dysfunction skin barrier in eczema, however, the proinflammatory environment causes Langerhans cells to promote a Th2 type immune response involving polarisation of naïve T cells to IL-13 producing CD4+ T cells (Elentner et al., 2009). Langerhans cells in eczema pathogenesis also have role in antigen capture, migration to lymph nodes and presentation of antigens to T cells (Dubrac, Schmuth and Ebner, 2010). Our understanding of Langerhans cell function within eczema lesions is supported by the spatial genomic observation that Langerhans cells migrate from the epidermis in healthy skin to co-locate with other inflammatory cells in the superficial dermis in eczema lesional skin (Figure 35 and Figure 36).

The microanatomical cellular location of other antigen presenting cells was also altered when comparing eczema lesional with healthy skin. DC1, DC2, monocyte derived macrophage subtype 3, macrophage 1 and migratory macrophages were all enriched within the superficial dermis in eczema lesional skin compared to healthy skin, and inflammatory

macrophages were enriched within the epidermis of eczema lesional skin compared to healthy skin (Figure 35 and Figure 36). Previous studies aiming to gain insight into the abundance and tissue distribution of DCs and macrophages in eczema lesional skin have used immunohistochemistry and light microscopic evaluation from atopy patch test skin sites modelling acute eczematous inflammation, and found increased macrophage cell numbers in eczematous skin and the presence of heterogeneous populations of dermal DCs and macrophages, including monocyte-derived cells expressing markers for both DC and macrophages (Kiekens et al., 2001). The findings from this spatial transcriptomic study builds on these previous findings by providing the cellular localisation of cell states derived from suspension data at single cell resolution.

Across the lymphoid cell subsets, changes in cellular tissue localisation between eczema lesional and healthy skin was observed within the natural killer (NK) and T cell subsets. NK cells, cytotoxic T cells, Tc17/Th17 cells, Tc IL13/IL22 cells, helper T cells and Tregs were all shown to be significantly enriched in the superficial dermis of eczema lesional skin compared to healthy skin (Figure 36). This is in keeping with the histological features of acutely inflamed eczematous skin, where a superficial dermal lymphocytic inflammatory cell infiltrate is observed (Bieber, 2010; Eckert, 1991). Previous analysis of the scRNA-seq dataset used for spatial cell deconvolution showed that across the different lymphoid cell states, only Tc IL13/IL22 cells were significantly more abundant in eczema lesional skin compared to healthy skin (Reynolds et al., 2021), highlighting the additional insights gained from this spatial transcriptomic approach of mapping cell states that are not differentially abundant between conditions but that show altered location within the tissue microenvironment. Notably, Tc17/Th17 cells were not identified in the eczema samples within the scRNA-seq data (Reynolds et al., 2021), and identifying their presence within eczema tissue sections may reflect the statistical approach taken in the spatial mapping algorithm that predicts transcriptionally similar single cell signatures within the Visium spots (Kleshchevnikov et al., 2022).

The comparisons of cell distance from the DEJ were also carried out between eczema non-lesional and healthy skin. Across the epidermal non-immune cells, dermal non-immune cells, antigen presenting cells and lymphoid cells, few cell states showed localised enrichment in eczema non-lesional versus healthy skin, compared to the changes observed between eczema lesional and healthy skin (Figure 36), reflecting the minimal clinical

changes on skin examination. However, there were some cell states that showed significantly altered microanatomical tissue localisation in eczema non-lesional skin compared to healthy skin, which were also noted in the above comparisons between eczema lesional and healthy skin. For example, disease keratinocytes were shown to be enriched in eczema non-lesional skin, reflecting skin barrier dysfunction and fibroblast 3 cells were enriched within the superficial dermal region of eczema non-lesional skin, which express disease promoting cytokines and MMPs, as discussed above. Furthermore, inflammatory macrophages and helper T cells were enriched within the epidermal region of eczema non-lesional skin compared to healthy skin (Figure 36), suggesting early recruitment of these inflammatory cell states in eczema disease pathogenesis prior to clinical signs developing (Kortekaas Krohn et al., 2022). These microanatomical tissue location findings provide insight into how specific cell states become perturbed in non-lesional skin in eczema compared to healthy unaffected skin.

Eczema lesional and non-lesional skin was also compared in order to understand the cell location differences between clinically unaffected non-lesional skin and acutely inflamed eczematous lesions. Significant differences were predominantly observed in the vascular endothelial cell and immune cell populations. Vascular endothelial cell subtypes 2 and 3 were found to be enriched in the superficial and mid-dermal regions of eczema lesional skin compared to non-lesional skin (Figure 36), which reflects the development of erythema across involved skin clinically. The expansion in the vascular endothelial network also facilitates the recruitment of inflammatory cells in the skin, which is reflected in the enrichment of DC1, DC2, migratory DCs, ILC2s, ILC1/NK cells, NK cells, cytotoxic T cells, Tc17/Th17 cells, Tc IL13/IL22 cells, helper T cells and Tregs in the superficial dermis of eczema lesional skin compared to non-lesional skin, and the enrichment of macrophage 1 cells and monocyte derived macrophages in the epidermis and superficial dermis of eczema lesional compared to eczema non-lesional skin (Figure 36). Many of these cellular tissue location changes were also observed when comparing eczema lesional and healthy skin, reflecting the spatial organisation of cell states in lesional inflamed skin regardless of whether the spatial transcriptomic data is compared with non-lesional or healthy skin, and giving further validity to the findings observed in lesional skin.

93

3.2.10. Tissue cellular microenvironments between disease conditions

The cell states mapped within the tissue sections were next assessed to determine whether specific cell types were more likely to be co-located within microanatomical tissue niches (or microenvironments) using non-negative matrix factorisation analysis, as described in the Methods. Co-location was indicated by a high proportion of two or more cell types sharing a microenvironment.

The samples were analysed to assess differences in the cells that occupy specific microenvironments between pairs of healthy and psoriasis disease conditions initially, including psoriasis lesional versus healthy skin (Figure 37), psoriasis non-lesional versus healthy skin (Figure 38) and psoriasis lesional versus psoriasis non-lesional skin (Figure 39).
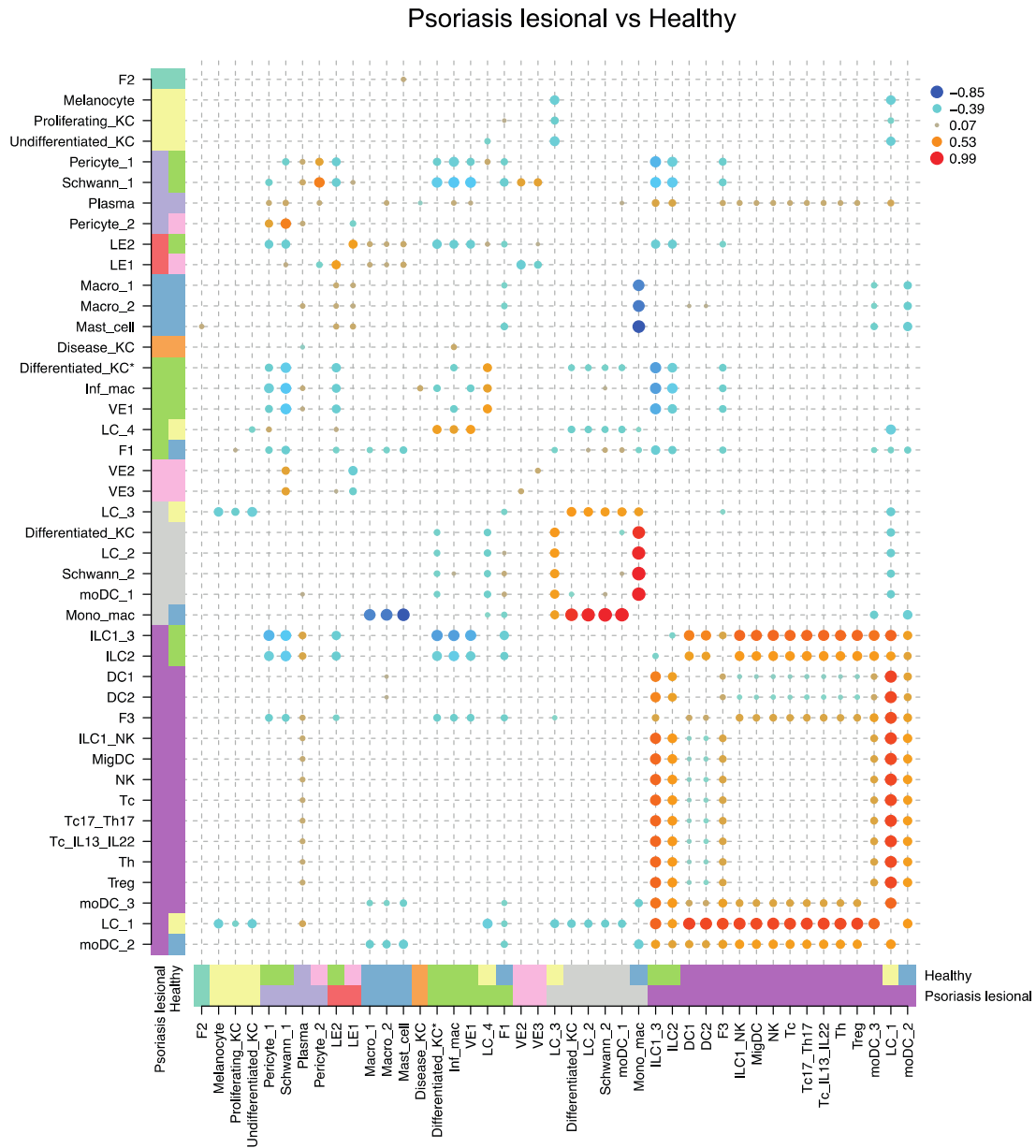
Figure 37: Spatial microenvironments in healthy and psoriasis lesional skin. Cell type to microenvironment assignment is shown by the coloured bars on the x- and y-axes. The differences in the non-negative matrix factorisation derived consensus matrices between healthy and psoriasis lesional skin are shown as a dotplot with significance values.

The comparison in spatial microenvironments between healthy and psoriasis lesional skin revealed several cell types that change their microenvironment between the two conditions. Monocyte derived macrophages are observed to co-locate with macrophage 1, macrophage 2, mast cells, fibroblast 1 and monocyte derived DC 2 in healthy skin. However, in psoriasis lesional skin, monocyte derived macrophages change their microenvironment and instead

95

co-locate with differentiated keratinocytes, Langerhans cell subsets 2 and 3, and monocyte derived DC 1 (Figure 37). This can be explained by monocyte derived macrophages migrating to the epidermis in psoriasis lesional skin, which is supported by the cell distance from the DEJ analysis shown previously (Figure 34). Unexpectedly, Schwann cell subtype 2 was also identified in the monocyte derived macrophage microenvironment along with differentiated keratinocytes and Langerhans cells in psoriasis lesional skin, which may be explained by an element of RNA leakage within the Visium experiment. Schwann cells make up a small proportion of cell types in the data compared with differentiated keratinocytes (Reynolds et al., 2021), therefore, the cell type mapping predictions for keratinocytes are more robust.

Interestingly, Langerhans cell 1 was also observed to change its microenvironment between conditions. In healthy skin, Langerhans cell 1 co-located with undifferentiated keratinocytes, proliferating keratinocytes, melanocytes and Langerhans cell subtypes 3 and 4. However, in psoriasis lesional skin, Langerhans cell 1 localised to a microenvironment comprising of fibroblast 3, several subtypes of antigen presenting cells (including DC1, DC2, migratory DC, monocyte derived DC 2 and 3), as well as all lymphoid cells subsets mapped in the dataset (Figure 37). This finding is supported by the cell distance from the DEJ analysis, where Langerhans cell 1 was observed to be enriched in the epidermis of healthy skin and in the superficial dermis of psoriasis lesional skin (Figure 34), demonstrating a process of Langerhans cell 1 migration away from the epidermis in psoriasis. A previous study using confocal microscopy in psoriasis lesional skin has shown that Langerhans cells in psoriasis can form cellular aggregates with T cells in the dermis close to lymphatic vessels (Martini et al., 2017). The aggregation of Langerhans cells in the dermis has been explained by migrational impairment in psoriasis due to the keratinocyte secrotome in response to IL17 stimulation (Eaton et al., 2018). However, these cellular microenvironment analyses between conditions and previously reported findings on Langerhans cell localisation also raise the question of possible interactions between Langerhans cells, T cells and other inflammatory cells in psoriasis, which has yet to be fully explored.

The ILCs are also observed to change their microenvironment between healthy and psoriasis lesional skin. ILC1/3 and ILC2 co-locate with vascular endothelial cell 1, pericyte 1, Schwann cell 1, inflammatory macrophages and inflammatory differentiated

keratinocytes in healthy skin, whereas they move to the same microenvironment as Langerhans cell 1 in lesional psoriasis, comprising of lymphoid cells, antigen presenting cells and fibroblast 3 (Figure 37). This is also supported by the cell distance from the DEJ analysis, where ILCs were enriched throughout the dermis in healthy skin but localised to the superficial and deep dermal skin layers lesional psoriasis (Figure 34). ILC3s are known to be enriched in the skin and blood of patients with psoriasis and secrete IL17 and IL22 implicated in disease progression (Villanova et al., 2014), however, migration of both ILC subsets to co-locate with other inflammatory cells in psoriasis and the potential intercellular interactions driving pathogenesis require further investigation.

Spatial microenvironments between psoriasis non-lesional and healthy skin were then compared to ascertain the changes in the cells that form microenvironments between the two conditions (Figure 38).
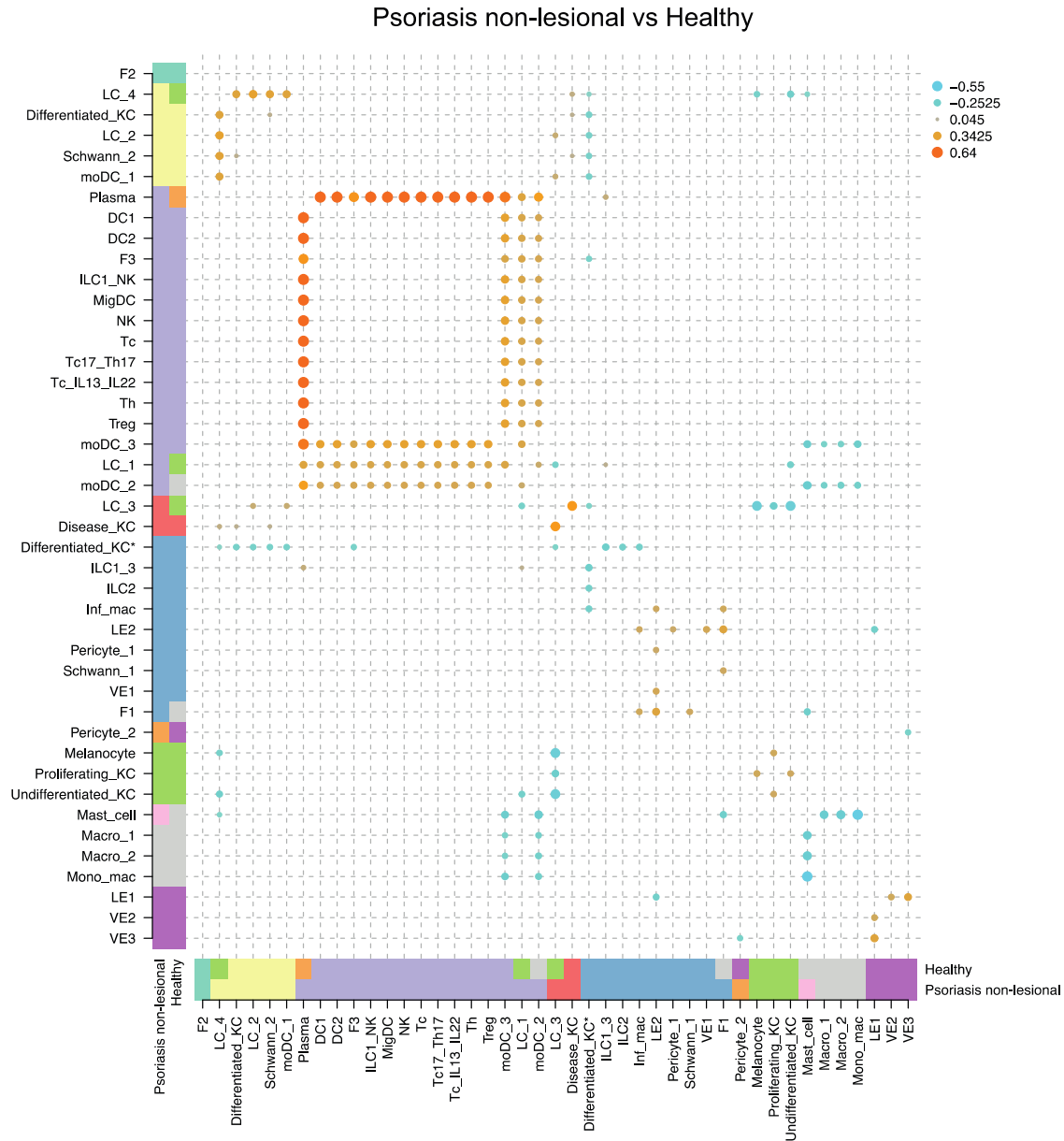
Figure 38: Spatial microenvironments in healthy and psoriasis non-lesional skin. Cell type to microenvironment assignment is shown by the coloured bars on the x- and y-axes. The differences in the non-negative matrix factorisation derived consensus matrices between healthy and psoriasis non-lesional skin are shown as a dotplot with significance values.

The comparison in spatial microenvironments between healthy and psoriasis non-lesional skin demonstrated fewer cell microenvironment changes than was observed between psoriasis lesional and healthy skin, reflecting the clinical manifestations and the cell distance to DEJ analysis (Figure 34 and Figure 38). Plasma cells in healthy skin form a homogenous microenvironment, whereas they change their microenvironment in psoriasis

non-lesional skin to co-locate with lymphoid cells, several subtypes of antigen presenting cells (DC1, DC2, migratory DC, monocyte derived DC 2 and 2, and Langerhans cell 1) and fibroblast 3 cells (Figure 38). The DEJ analysis showed that plasma cells were enriched in the deep dermis in healthy skin, in the deep and superficial dermis in psoriasis non-lesional skin, and in the superficial dermis and epidermis in psoriasis lesional skin (Figure 33), however, statistical significance of these cell location differences was only observed in the comparison between psoriasis non-lesional and healthy skin (Figure 34). Psoriasis is classically considered to be a T cell mediated inflammatory skin disease, however, there has been recent literature providing support for the role of B cells in psoriasis pathogenesis. A recently published scRNA-seq study of peripheral blood mononuclear cells (PBMCs) from healthy donors and patients with psoriasis revealed increased levels of plasma cells in psoriasis that express immunoglobulin (Ig) A or IgG (Liu et al., 2023), and a previous report documented an increase in CD19[+] B cell subsets in skin lesions and the peripheral blood of patients with psoriasis that correlates with disease severity (Lu et al., 2016). These reports and the findings from this study suggest a potential role for B cell activation in psoriasis pathogenesis.

As was observed in the comparison of microenvironments between psoriasis lesional and healthy skin, Langerhans cell 1 and monocyte derived DC 2 in healthy skin are shown to change their microenvironment to co-locate with other lymphoid cells in psoriasis non-lesional skin (Figure 37 and Figure 38). This lends support to the migration of antigen presenting cells toward lymphoid cells in the superficial dermis early in the pathogenesis of psoriasis, prior to clinical skin changes being apparent.

Spatial microenvironments between psoriasis lesional and psoriasis non-lesional skin were also compared in order to determine the microanatomical cellular changes occurring in the development of an inflamed psoriasis plaque (Figure 39).
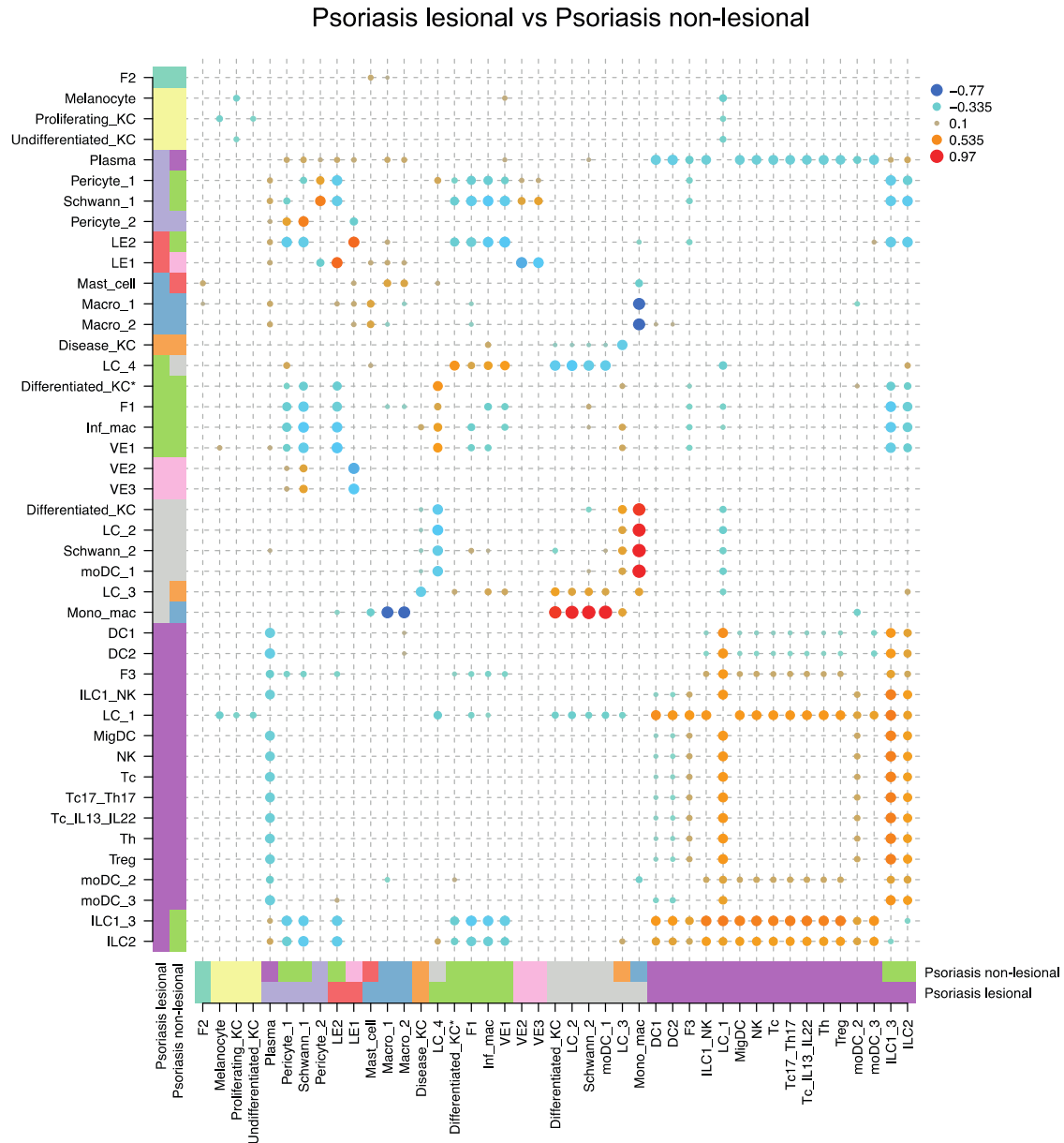
Figure 39: Spatial microenvironments in psoriasis lesional and non-lesional skin. Cell type to microenvironment assignment is shown by the coloured bars on the x- and y-axes. The differences in the non-negative matrix factorisation derived consensus matrices between psoriasis lesional and non-lesional skin are shown as a dotplot with significance values.

The comparison between psoriasis non-lesional and lesional skin microenvironments revealed similar changes to those observed between healthy and psoriasis lesional skin. This includes movement of monocyte derived macrophages in psoriasis non-lesional skin from a dermal microenvironment consisting of macrophages 1 and 2 to an epidermal

microenvironment in psoriasis lesional skin consisting of differentiated keratinocytes and Langerhans cell subtypes 2 and 3 (Figure 39). Furthermore, ILC1/3 and ILC2 were shown to change their microenvironment to co-locate with lymphoid cells in the superficial dermis (Figure 33 and Figure 39). This analysis highlights common cellular microenvironmental changes implicated in the development of psoriasis lesions when compared to both healthy and psoriasis non-lesional skin.

The spatial cell microenvironments in eczema skin were examined next, and differences between healthy and eczema lesional skin (Figure 40), healthy and eczema non-lesional skin (Figure 41) and lesional and non-lesional eczema skin (Figure 42) were examined.
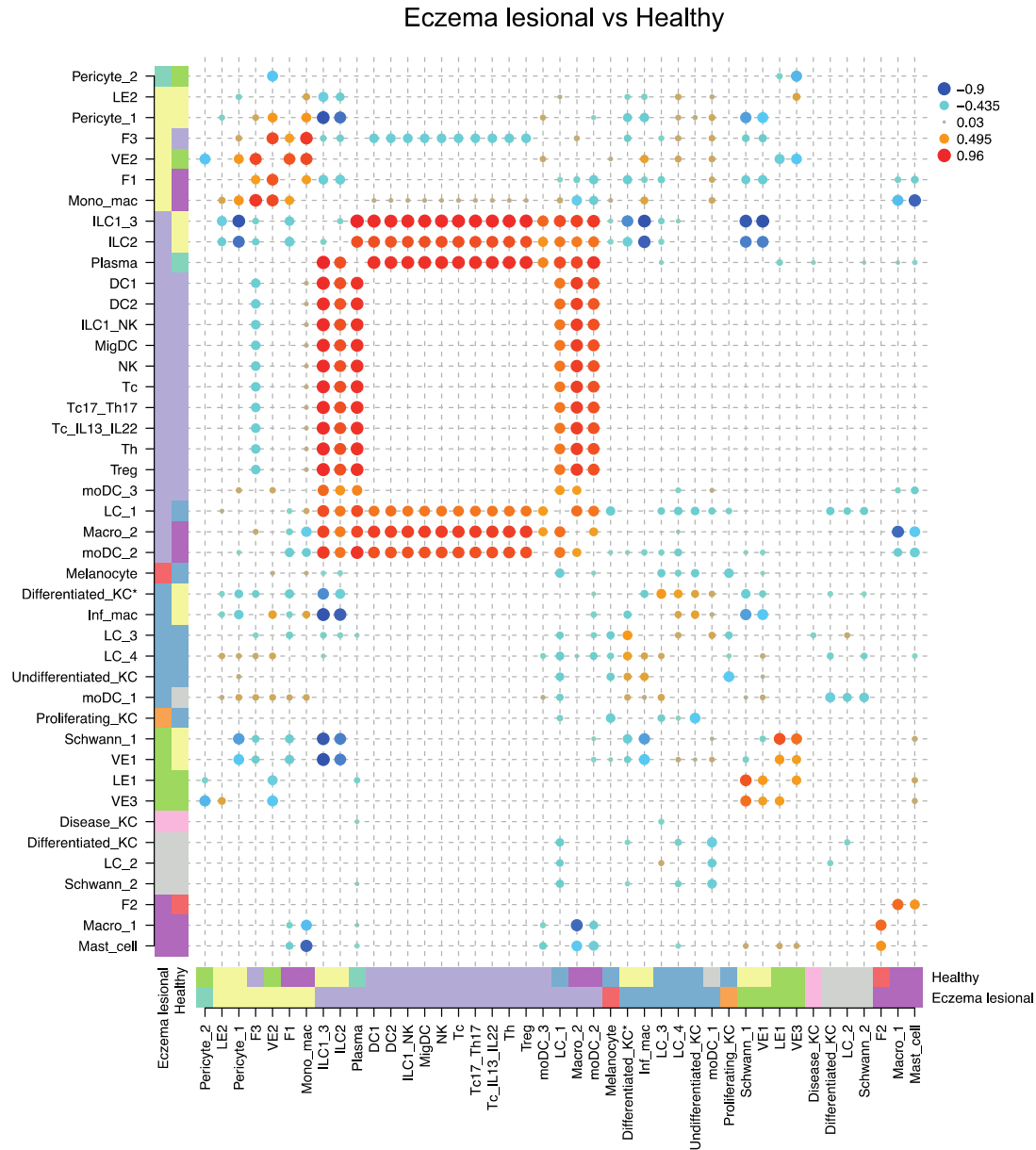
Figure 40: Spatial microenvironments in healthy and eczema lesional skin. Cell type to microenvironment assignment is shown by the coloured bars on the x- and y-axes. The differences in the non-negative matrix factorisation derived consensus matrices between healthy and eczema lesional skin are shown as a dotplot with significance values.

The spatial microenvironment analyses between eczema lesional and healthy skin revealed that the most significant changes in cellular localisation involved the movement of several antigen presenting cell subsets in healthy skin, including ILC1/3, ILC2, Langerhans cell 1, macrophage 2 and monocyte derived DC 2, into a microenvironment in eczema lesional skin composed predominantly of lymphoid cells (Figure 40). It is interesting to note that, despite these significant findings in the spatial microenvironment analysis, ILC1/3, ILC2

and macrophage 2 cell states did not show a significantly differential enrichment in the cell distance from the DEJ analysis between eczema lesional skin compared to healthy skin (Figure 36), demonstrating that each analysis uncovers different types of insights. The cell distance from the DEJ analysis examines the localisation of cell states broadly across the tissue sections from superficial to deep, whereas the spatial micromovement analysis provides information on how individual cells change their colocalisation patterns, even if they remain within the same broad tissue depth level, as in the case of the ILCs and macrophage 2 cells.

There have relatively few studies carried out on the contribution of ILCs to human eczema pathogenesis. Though no significant difference in the abundance of ILCs was observed between eczema and healthy skin in the scRNA-seq data used for the spatial deconvolution analysis (Reynolds et al., 2021), previous studies using flow cytometry have shown an increase in the numbers of ILC2s in lesional eczema skin biopsies that produce type 2 cytokines including IL5 and IL13 (Salimi et al., 2013; Alkon et al., 2022). The localisation of ILCs in eczema skin and their relationship to other cell states is also poorly understood. Immunolabelling techniques have been used to detect ILC subsets within the skin of eczema lesions (Kim et al., 2014a; Brüggen et al., 2016), with a recent study showing the highest enrichment of ILCs in the upper dermis of eczema skin compared to healthy skin (Alkon et al., 2022). However, only by generating transcriptomic data across tissue sections and mapping diverse cell states within the skin can we identify cellular microenvironments between conditions, rather than being limited by the cells and markers that are selected for cell detection with immunolabelling approaches. The findings from the spatial microenvironment highlights that interactions between ILCs and other lymphoid cells in eczema lesional skin are significant and requires further study.

Identification of a significant change in cellular microenvironments between healthy and eczema lesional skin for macrophage 2 cells is of interest in the context of the previous findings from the scRNA-seq analyses of this reference dataset (Reynolds et al., 2021). Macrophage 2 cells were observed to be significantly expanded in eczema lesional and non-lesional skin and psoriasis lesional skin compared to healthy, which was corroborated in the analysis of a bulk RNA sequencing dataset from a larger patient cohort (Reynolds et al., 2021). Macrophage 2 cells highly express *CXCL8*, which encodes IL8 and promotes lymphocyte recruitment (Nedoszytko et al., 2014), which supports the spatial

microenvironment analysis showing that macrophage 2 in healthy skin co-locates with fibroblast 1, monocyte derived macrophages, macrophage 1, monocyte derived DC 2 and mast cells, whereas macrophage 2 in eczema lesional skin shared a microenvironment with all of the mapped lymphoid cells in the dataset (Figure 40). Macrophage 2 was shown to interact with vascular endothelial cell subtype 3 via *CXCL8* and *ACKR1* respectively to promote angiogenesis, which also supports the findings in the cell distance from the DEJ analysis in eczema lesional skin compared to healthy skin discussed above (Figure 36). Interestingly, this interaction was also observed between equivalent cell states in first trimester prenatal skin, suggesting that this cell signalling program is leveraged in both development and disease for lymphocyte recruitment and angiogenesis (Reynolds et al., 2021).

The spatial microenvironment analysis also showed that plasma cells form their own microenvironment in healthy skin, whereas they co-locate with lymphoid cells in eczema lesional skin (Figure 40). In support of this, the DEJ analysis showed plasma cells to be most enriched in the deep dermis in healthy skin, whereas they were most enriched within the superficial dermis in eczema lesional skin, at the same depth as other lymphoid cells (Figure 36). Plasma cells are not known to be a predominant feature of eczematous inflammation, and infiltration of plasma cells on histological analysis may indicate an alternative differential diagnosis (Walsh et al., 2019). However, very few plasma cells were captured from the healthy and eczema skin biopsies (Reynolds et al., 2021) and the cell distance from the DEJ analysis is normalised and therefore masks absolute cell numbers. Plasma cells are known to constitutively reside in healthy skin and have an important role in supporting homeostatic skin barrier integrity by providing adaptive IgM in response to the cutaneous environment (Wilson et al., 2019). Plasma cells have also been shown to accumulate in inflamed skin (Wilson et al., 2019), including in the context of eczema (Cipriani et al., 2014), though very little has been reported on their putative role in eczematous skin inflammation. The findings from the spatial microenvironment analysis suggest that plasma cells co-locate with other inflammatory cells in eczema lesional skin and therefore any potential intercellular interactions present that drive or dampen eczematous inflammation require further investigation.

Spatial microenvironments between eczema non-lesional skin and healthy skin were compared next in order to ascertain the differences in cells that form microenvironments

between clinically unaffected skin in patients with eczema and skin from healthy donors (Figure 41).
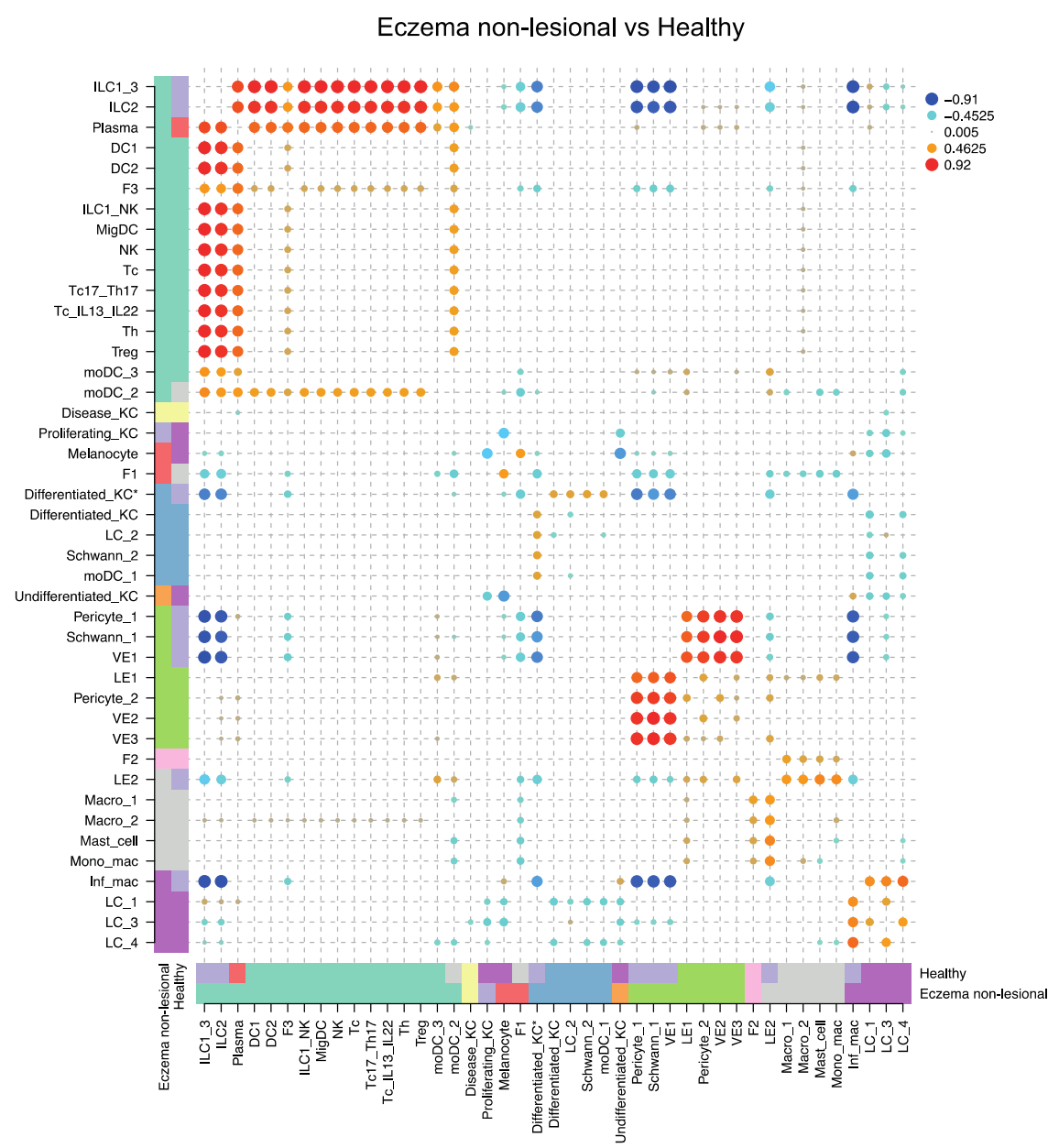
Figure 41: Spatial microenvironments in healthy and eczema non-lesional skin. Cell type to microenvironment assignment is shown by the coloured bars on the x- and y-axes. The differences in the non-negative matrix factorisation derived consensus matrices between healthy and eczema non-lesional skin are shown as a dotplot with significance values.

The comparison in spatial microenvironments between healthy and eczema non-lesional skin demonstrated that ILC1/3, ILC2 and plasma cells change their microenvironment from healthy skin to co-locate with other lymphoid cells, monocyte derived DCs and fibroblast

subtype 3 (Figure 41). The migration of ILCs to a microenvironment comprising of other lymphoid cells reflects the observations made between healthy and eczema lesional skin (Figure 40). Whether ILCs identified in eczema skin samples initiate inflammatory processes or are present as a consequence of inflammation is contentious (Alkon et al., 2022). Observing microanatomical changes in ILCs between healthy and clinically unaffected non-lesional eczema skin samples may suggest that ILCs may be contributing to a process of subclinical inflammation prior to the development of eczema lesions.

Spatial microenvironments between eczema lesional and eczema non-lesional skin were also compared in order to determine the microanatomical cellular changes occurring in the development of inflamed lesions in patients with eczema (Figure 42).
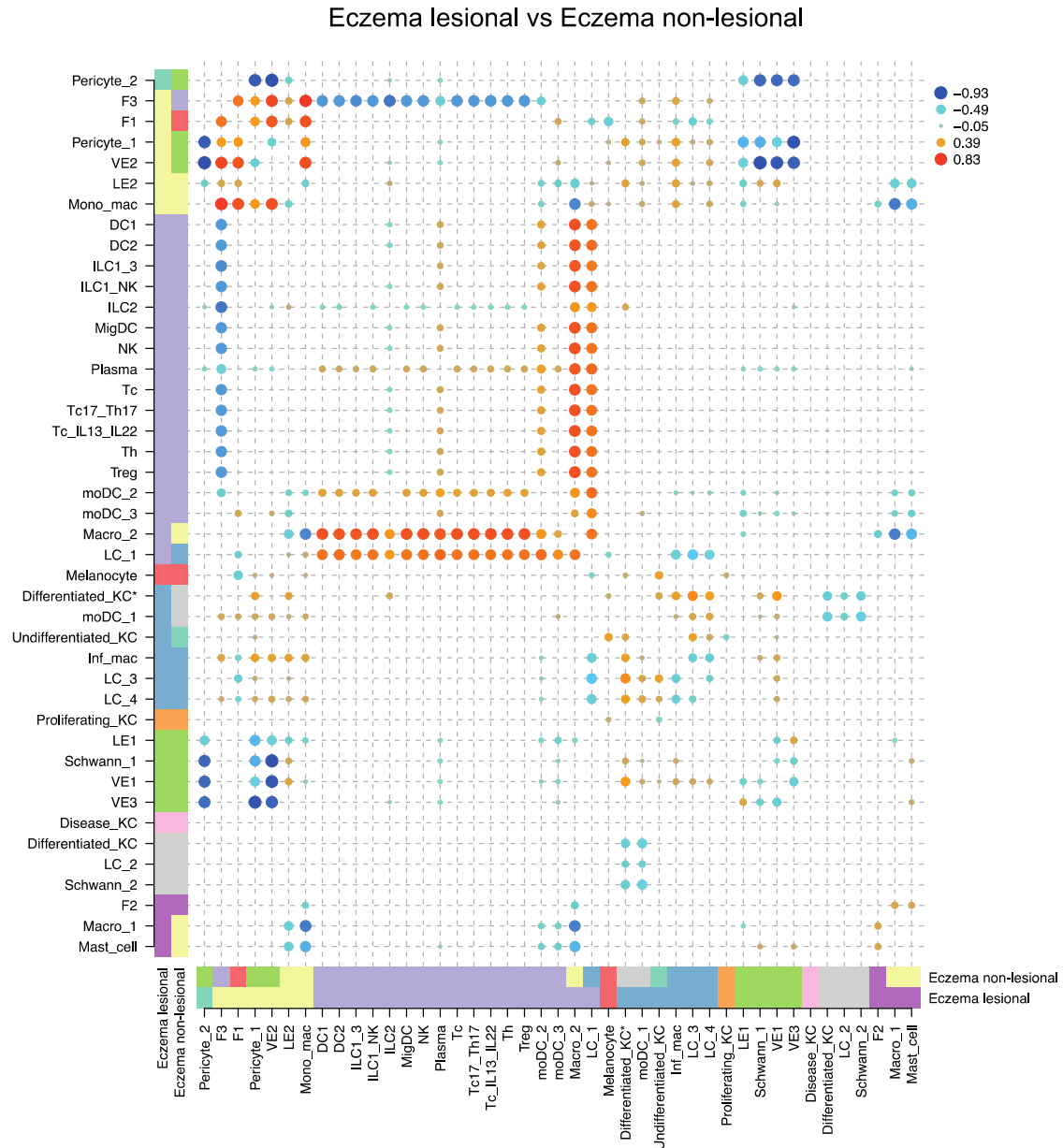
Figure 42: Spatial microenvironments in eczema lesional and non-lesional skin. Cell type to microenvironment assignment is shown by the coloured bars on the x- and y-axes. The differences in the non-negative matrix factorisation derived consensus matrices between eczema lesional and non-lesional skin are shown as a dotplot with significance values.

The comparison in cellular microenvironments between eczema lesional and eczema non-lesional skin showed that Langerhans cell 1 and macrophage 2 change their microenvironment to co-locate with lymphoid cell states and some antigen presenting cells (including DCs and ILCs) in eczema lesional skin, from microenvironments in non-lesional skin that comprised predominantly of other Langerhans cell and macrophage cell subsets

(Figure 42). Of the six cell states that changed their microenvironment between healthy and eczema lesional skin described previously (ILC1/3, ILC2, plasma cells, monocyte derived DC 2, Langerhans cell 1 and macrophage 2) (Figure 40), it is interesting to note that the ILCs, plasma cells and monocyte derived DC2 cells migrated to co-locate with lymphoid cells from healthy to eczema non-lesional skin (Figure 41) and Langerhans cell 1 and macrophage 2 were observed to change between eczema non-lesional and lesional skin (Figure 42), indicating an element of chronicity to the initial changes observed between eczema lesional skin and healthy. The findings across the different comparisons suggest that ILCs, plasma cells and monocyte derived DC2 cells migrate to co-locate with other lymphoid cells early in eczema pathogenesis prior to clinical manifestations becoming apparent, and that Langerhans cell 1 and macrophage 2 migration toward lymphoid cells occurs later when eczema lesions have developed.

As discussed in the previous section on the cell distance from the DEJ analyses, Langerhans cells in eczema have been reported to have a role in antigen capture, migration to skin draining lymph nodes and interacting with T cells for antigen presentation (Dubrac, Schmuth and Ebner, 2010), therefore explaining why Langerhans cells and lymphoid cells are observed within the same microenvironment in eczema lesional skin (Figure 42). Furthermore, macrophage 2 express chemokines that facilitate lymphocyte recruitment to the inflamed skin tissue (Reynolds et al., 2021). A significant intercellular interaction between macrophage 2 and vascular endothelial cell 3 was predicted in the scRNA-seq analyses of the reference dataset, which was confirmed by immunohistochemical analysis of eczema skin samples (Reynolds et al., 2021). However, the results of the spatial deconvolution analysis showed that macrophage 2 are located in the same microenvironment as lymphoid cells and other antigen presenting cells and that vascular endothelial cell 3 is identified in a separate microenvironment with other endothelial cells (Figure 42), which may suggest that the inflammatory cell aggregates are more abundant and significant across the tissue sections compared with the number of macrophages co-locating with the vascular endothelium.

3.3. Discussion

In this Chapter, the generation of a spatial transcriptomic atlas of human healthy adult skin and diseased skin from patients with psoriasis and eczema is presented. Psoriasis and eczema samples were collected from inflamed lesional sites and clinically uninvolved non-lesional sites in order to reveal the cellular microanatomical changes occurring between these different conditions. This study extends from a previous project conducted with our research group that has recently been published, using scRNA-seq to profile around 500,000 cells from healthy skin, lesional and non-lesional psoriasis, and lesional and non-lesional eczema, in order to develop a further understanding of the cellular and molecular mediators underlying these diseases (Reynolds et al., 2021). However, this published dataset was generated from single cell suspensions from dissociated tissue, therefore losing valuable information about spatial expression within a microenvironmental context. This spatial transcriptomic dataset therefore provides complementary insights to the suspension dataset, especially because patients were recruited with the same inclusion and exclusion criteria, allowing robust integration of both datasets for spatial cell deconvolution analyses to determine the location of refined skin cell states that are expected within the tissue samples.

The analyses of the spatial transcriptomic data revealed several insights into the microanatomical changes occurring between healthy, psoriasis and eczema skin, and across lesional and non-lesional disease samples. Some findings were expected based on our understanding of the well-documented histological changes observed in each condition (De Rosa and Mignogna, 2007; Murphy, Kerr and Grant-Kels, 2007; Eckert, 1991). Previous *in situ* hybridisation and immunolabelling studies also provide context to the findings from the spatial transcriptomic data analysis (Alkon et al., 2022; Sankar et al., 2017), however, these approaches involve the detection of a limited number of molecules or cell types of interest based on markers they express, whereas using spatial transcriptomic technology has allowed unbiased transcriptome data generation and mapping of all cell states simultaneously from a reference suspension dataset. This spatial genomic approach has therefore allowed more extensive and systematic interpretation of the microanatomical cellular changes underpinning psoriasis and eczema pathogenesis and has allowed

quantitative comparisons to be made between healthy and disease states and between lesional and non-lesional disease samples.

The analysis of the spatial transcriptomic data from psoriasis lesional samples and comparisons with healthy and non-lesional psoriasis samples revealed microanatomical changes affecting various different cell lineages. The cell distance outputs from the DEJ analysis showed some findings that were expected based on our understanding of the histopathological changes in psoriasis lesional skin, including the expansion of disease keratinocytes secondary to hyperkeratosis, enrichment of vascular endothelial cells in the superficial skin regions reflecting vascular expansion in the papillary dermis with thinning of the suprapapillary epithelium, and enrichment of lymphoid cells in the superficial dermis reflecting the superficial dermal lymphocytic infiltrates observed histologically (Heidenreich, Röcken and Ghoreschi, 2009; Mihu et al., 2021). However, spatially mapping multiple refined cell states that have been characterised in terms of their gene expression profiles (Reynolds et al., 2021) revealed novel insights into how cell subsets are organised *in situ* beyond that which has been achieved using histology, *in situ* hybridisation and immunolabelling approaches. For example, an enrichment in fibroblast subtypes 1 and 3 in psoriasis lesional skin was observed compared to healthy and psoriasis non-lesional skin, and these cell states have been shown to highly express *IL6*, unlike fibroblast subtype 2, which is enriched in the mid-dermis in psoriasis lesional skin. Previous studies have shown that IL6 stimulates the proliferation of keratinocytes and the recruitment of neutrophils that can be identified in the stratum corneum of psoriasis plaques, therefore the localisation of *IL6*-expressing fibroblast subtypes 1 and 3 close to the epidermis provides further insights into their role in psoriasis pathogenesis, particularly as most studies focus on immune cell mediated pathology (Lowes, Bowcock and Krueger, 2007; Grossman et al., 1989). Comparative analysis of tissue cellular microenvironments across conditions also revealed novel insights that have not been possible using previous technologies, owing to the ability to simultaneously and systematically explore the microanatomical location of a large number of cell types profiled by scRNA-seq using spatial cell deconvolution. These analyses revealed that monocyte derived macrophages change their cellular microenvironment in the dermis of healthy skin to co-locate with Langerhans cell subtypes 2 and 3 in the epidermis of psoriasis lesional skin, and that Langerhans cell subtype 1 migrate from the epidermis to the dermis to co-locate with groups of lymphoid cells. Previous literature on the localisation of Langerhans cells in psoriasis has been contentious

110

(Eaton et al., 2014; Glitzner et al., 2014; Fujita et al., 2011), which may in part be attributable to underpowering of these studies. A previous confocal microscopic analysis of psoriasis lesional skin, however, has shown the formation of dermal Langerhans cell aggregates with T cells close to lymphatic vessels (Martini et al., 2017), which supports the spatial genomic findings presented here. However, previous studies have not robustly explored the role of different Langerhans cell subsets in psoriasis and how each transcriptomically distinct cell state is interacting with other co-locating cell types, which would further our understanding of the complex cellular crosstalk underpinning this disease and potentially reveal novel therapeutic targets.

The spatial analyses of psoriasis non-lesional skin samples demonstrated modest cellular microanatomical changes overall, which is in keeping with the clinical and histological features of this condition. Interestingly, however, the cell distance from the DEJ analysis revealed that vascular endothelial cell subtypes 2 and 3 were enriched in the superficial dermis of psoriasis non-lesional skin compared to healthy skin, and of the three vascular endothelial subsets, subtypes 2 and 3 highly express *ACKR1* compared to subtype 1, which in turn was not significantly enriched in psoriasis non-lesional skin. Single cell analyses have shown that *ACKR1* on vascular endothelial cells mediate disease-associated signalling with macrophages to promote angiogenesis and lymphocyte recruitment, and that although this intercellular interaction is not identified in healthy skin, it is present in prenatal skin with a potential contributory role during skin morphogenesis (Reynolds et al., 2021). The enrichment of *ACKR1*-expressing vascular endothelial cells 2 and 3 in the superficial dermis of psoriasis non-lesional skin therefore provides new insights into the early changes that mediate immune cell recruitment in psoriasis, prior to clinical signs becoming apparent.

The analysis of the spatial transcriptomic data generated from eczema lesional and non-lesional samples also revealed several insights in the microanatomical organisation of cell states between conditions. Findings from the cell distance from the DEJ analysis highlighted changes that were expected based on the histological features of eczema lesional skin, including an enrichment of lymphoid and some antigen presenting cells in the superficial dermis compared to healthy and non-lesional eczema skin, reflecting the dermal inflammatory cell infiltrate observed histologically (Bieber, 2010; Eckert, 1991). However, this systematic analysis of cell states that have been characterised using single

cell transcriptomic profiling has provided further insights into the cell subsets that are enriched across different skin regions, showing that DC1, DC2 and migratory DCs specifically are enriched in the superficial dermis of eczema lesional skin compared to healthy and non-lesional skin, in addition to most mapped lymphoid subsets. Furthermore, *ACKR1*-expressing vascular endothelial cell subtypes 2 and 3 were found to be enriched across the superficial and mid-dermis of eczema lesional skin compared to healthy and non-lesional skin, whereas subtype 1 was not significantly enriched in eczema lesional skin. This provides spatial context to the scRNA-seq findings showing the role of *ACKR1* mediating lymphocyte recruitment and angiogenesis in inflammatory skin disease via interactions with macrophages (Reynolds et al., 2021).

The tissue cellular microenvironment analyses comparing healthy, eczema non-lesional and eczema lesional skin revealed the migration of several immune cell states to a microenvironment in eczema that predominantly contains lymphoid cells, including NK cells, cytotoxic and helper T cells and Tregs. Several cell states change their microenvironment to co-locate with lymphoid cells in eczema non-lesional skin, including ILC1/3, ILC2, monocyte derived DC 2 and plasma cells, and when comparing eczema non-lesional with lesional skin, Langerhans cell 1 and macrophage 2 change their microenvironment to co-locate with lymphoid cells in eczema lesional skin. These separate findings are then reflected in the comparison between eczema lesional and healthy skin, where all six cell states (ILC1/3, ILC2, monocyte derived DC 2, plasma cells, Langerhans cell 1 and macrophage 2) move to the lymphoid microenvironment in eczema lesional skin. Interestingly, scRNA-seq analysis of the reference dataset did not show a differential abundance in ILCs between healthy and eczema skin (Reynolds et al., 2021), demonstrating the additional insights into ILC biology during disease using spatial methods. Macrophage 2, however, was shown by scRNA-seq to be more abundant in both eczema lesional and non-lesional skin compared to healthy skin (Reynolds et al., 2021), though the spatial analysis only showed a significant change in the cellular microenvironment of macrophage 2 in lesional skin rather than non-lesional eczema skin. *CXCL8* on macrophage 2 was shown by cell-cell interaction analysis to interact with *ACKR1* on vascular endothelial cells to mediate angiogenesis and lymphocyte recruitment in disease (Reynolds et al., 2021), and though these two cell states were not identified in the same microenvironment, the spatial genomic analyses provide an overall representation of the most significant

microenvironments across mapped cell states between conditions, highlighting intercellular interactions to be explored in future studies on disease pathogenesis.

Prior to generating the spatial transcriptomic data for this project, I carried out various optimisation experiments in order to help ensure that the imaging and sequencing data obtained was of high quality. The research group that I was working in had not previously generated Visium data prior to the commencement of my project, and therefore, I was involved in optimising cryosectioning conditions and adapting the available imaging software tools to capture data from Visium slides. I also carried out an optimisation experiment using adult healthy facial skin to determine the ideal skin tissue digestion, as described in section 3.2.2. The data and guidance from 10x Genomics at the time indicated that tissue optimisation is required for each tissue type of interest, and that carrying out further tissue optimisation experiments to account for donor variables was not required. However, recent experience in our research group has been that the optimal permeabilisation time for skin can vary depending on the age of tissue donors, the anatomical site from which skin was sampled or whether the skin was affected by disease, which reflects the differences in skin structure and cellular organisation with age, site and disease. By using the same permeabilisation time of 14 minutes across all of the tissue sections included in this study, some tissue regions may have been either under- or over-digested, resulting in either insufficient RNA release from the tissue or RNA degradation, respectively. Despite this, the quality of the data obtained for this project is comparable to published spatial transcriptomic datasets (Ji et al., 2019; Castillo et al., 2023) and allowed robust analyses to be carried out after merging Visium spots with low UMI counts that were predominantly in the cell sparse dermal regions.

The Visium spatial gene expression platform was only launched in 2019 and computational approaches to analyse spatial transcriptomic data have been under ongoing development, with Cell2location for spatial cell type deconvolution analysis only being published last year (Kleshchevnikov et al., 2022). Other cellular deconvolution approaches have also been developed for spatial transcriptomic data, which were compared in a benchmarking report published earlier this year, showing that Cell2location was amongst the three best performing methods (Li et al., 2023a). There has been wide variation in the approaches that different research groups have taken to analyse Visium spatial transcriptomic data, with a

common method being to manage the data like single cell data and perform unsupervised clustering of the Visium spots, annotate the clusters and determine differences in gene expression and mapped cell states between clustered tissue regions (Castillo et al., 2023; Long et al., 2023; Ji et al., 2019). However, a major limitation of this method is that the clustering algorithm is not able to delineate biologically meaningful tissue regions and is confounded by the effects of RNA leakage during the Visium experiment. In this case, clusters cross important tissue compartment borders and are often simply annotated numerically with little biological relevance. To overcome this, I worked with Pasha Mazin, a computational biologist at the Wellcome Sanger Institute who led the downstream advanced spatial analysis and to whom I am grateful for his expertise and input, to develop novel computational methods to analyse the skin spatial transcriptomic data. First, I annotated each tissue section to demarcate the following regions: epidermis, dermis, pilosebaceous unit, adipose tissue and regions not containing tissue but that still contained gene expression data secondary to RNA leakage. The adipose and non-tissue regions were excluded from downstream analyses and the epidermal and dermal annotations were used to determine the position of the DEJ, which was then used to determine the location and tissue depth of mapped cell states and allow quantitative comparisons to be made.

For this study, we developed a novel method to address Visium spots which contained low UMI counts, which primarily affected the skin dermis due to the relatively low number of cells present compared to cell dense epidermal and hair follicle regions. Visium spots were merged to form meta-spots based on specified thresholds detailed in the Methods, therefore allowing more robust cell type deconvolution analyses to be performed within these regions. However, a limitation of this approach of merging $50\mu m$ spots into meta-spots is that the area of tissue onto which cells are mapped in the deconvolution increases depending on the number of spots that require merging, therefore lowering the resolution of cell mapping. Other spatial transcriptomic technologies have recently been under development to profile tissue sections at higher resolutions, such as Slide-seq ($10\mu m$) (Rodriques et al., 2019), Seq-scope ($1\mu m$) (Cho et al., 2021a), and Visium HD, which can recover transcript-coupled spatial barcodes at $2\mu m$ resolution (Vickovic et al., 2019) and is due to be commercially available by 10x Genomics over the coming year.

The spatial microenvironment analysis also required a novel computational approach to be developed. Non-negative matrix factorisation has been previously used to determine cellular microenvironments across tissue sections (Suo et al., 2022), however, this approach involves analysing and interpreting individual samples separately in order to assess the cell states that map to distinct spatial microenvironments. An aim of this project was to compare spatial cellular microenvironments between health and disease and across lesional and non-lesional disease states. Therefore, a computational method was developed that compared spatial microenvironments between two samples by subsampling the data and performing the non-negative matrix factorisation analysis in multiple iterations in order to account for biological variability of the data and allow comparisons to be made, details of which are described in the Methods. In the near future, we plan to compile the methods to analyse spatial transcriptomic data developed during the course of this project into a computational package for use by the wider research community and hope to publish and disseminate the computational approaches taken and the application of these methods to help reveal biological insights into healthy, psoriasis and eczema skin.

Future work is also required to systematically investigate the interactions between co-locating cell states within spatial microenvironments in each condition in order to further our understanding of disease pathogenesis and identify potentially novel therapeutic targets. Spatial genomic technologies may also facilitate the stratification of patients based on the distinct cellular and molecular profiles expressed in sampled disease tissue, allowing a personalised medical treatment approach to be explored with patients in the future.

Chapter 4: A comprehensive single-cell and spatial atlas of prenatal skin reveals distinct cellular microenvironments that support development

4.1. Introduction

The spatial transcriptomic analyses of healthy and diseased skin presented in the previous Chapter revealed cellular microenvironments that underly different skin conditions. These findings raised the hypothesis that distinct functional microenvironments may be present within prenatal skin during development to support skin and appendageal morphogenesis, which was therefore investigated using single cell and spatial genomics technologies.

Skin morphogenesis is a dynamic process that begins early in gestation and involves epidermal specification and stratification, development of the dermis and formation of the skin appendages (Hu et al., 2018). Developmental processes are mediated by intercellular interactions between distinct prenatal skin cell states, however, this has historically been difficult to study due to the scarcity of human prenatal skin samples (Haniffa et al., 2021). Furthermore, prenatal skin tissue comprises of small numbers of different cell types that are difficult to detect and investigate using conventional experimental techniques. High-throughput scRNA-seq technology allows unbiased detailed cellular characterisation of individual cell states, including rare cell populations, using a small about of biological material (Kolodziejczyk et al., 2015; Papalexi and Satija, 2018). This technology has been successfully used to study various prenatal tissue organs at a single cell resolution to yield novel biological insights about prenatal development, including the liver, thymus, bone marrow and spleen (Suo et al., 2022; Jardine et al., 2021; Popescu et al., 2019; Park et al., 2020). However, a systematic characterisation of the cellular composition of prenatal skin across the first and second trimester of gestation has yet to be carried out.

Although there have been informative single cell reports profiling skin and hair follicle development in murine models (Mok et al., 2019; Gupta et al., 2019), studies on human

prenatal skin to date have been limited. A recently published scRNA-seq study profiled human prenatal skin between 8 – 15 PCW, however only CD45$^+$ immune cells were isolated and non-immune cells were not captured (Xu et al., 2021). They characterised 13 different immune cell types, including myeloid and lymphoid precursors, monocytes, macrophages, DCs, NK cells, ILCs, T cells, B cells and mast cells (Xu et al., 2021). Their analysis showed that as skin morphology changed during this period of development, so too did the transcriptomic profile of several immune subsets, including macrophages and ILCs. This indicated evolving functional roles of immune cells during prenatal skin development, with genes related to cell proliferation in macrophages more highly expressed earlier in gestation and genes related to focal adhesion and extracellular matrix (ECM) formation enriched later in gestation (Xu et al., 2021), suggesting a role for immune cells supporting skin morphogenesis. A further scRNA-seq study profiled tissue resident macrophages (TRMs) in embryonic skin less than 9 PCW, identifying diverse subsets within prenatal skin (Bian et al., 2020). Both of these prenatal skin scRNA-seq studies, however, did not include prenatal skin non-immune cells or any spatial contextualisation of the data, and therefore, the microanatomical organisation of cells *in situ* could not be addressed.

In this Chapter, I describe a single-cell RNA sequencing study of human prenatal skin across the first and second trimester of gestation, where samples were obtained between 7 and 16 PCW. During this time of time in prenatal development, the architecture of human skin matures, the epidermis stratifies to become a complex multi-layered structure and hair follicles first develop (Whitting et al., 2008). Prenatal cell states are annotated at broad and refined levels to provide a comprehensive overview of the cellular composition of prenatal skin during this developmental period.

The prenatal skin data is also contextualised with scRNA-seq data from an embryonic stem cell (ESC) / induced pluripotent stem cell (iPSC)-derived hair-bearing skin organoid model (Lee et al., 2020) to determine whether the skin organoid recapitulates physiological differentiation during prenatal development at a molecular level across different cell lineages. The comparison between prenatal skin and a skin organoid model allows the potential utility of skin organoids for functional experimentation, including disease modelling, to be evaluated, which will be discussed in Chapter 5.

Prenatal single cell data is then contextualised with spatial transcriptomic data of skin regions within human embryonic limb (Zhang et al., 2022) to reveal the microanatomical skin cellular niches present in prenatal skin. The cell states interacting within cellular microanatomical niches, and their functional roles, are then explored to demonstrate how they may be supporting prenatal skin morphogenesis.

This study is the first to provide a comprehensive single cell characterisation of immune and non-immune cells in human prenatal skin across the first and second trimester of gestation. Furthermore, by contextualising the analysis with spatial transcriptomic data, this study represents the first combined single cell and spatial atlas of prenatal skin during development.

4.2. Results

4.2.1. Prenatal skin sample acquisition

Fifteen prenatal skin samples that were aged between 7 and 16 PCW, spanning the first and second trimester, were obtained from HDBR. Prenatal tissue samples underwent mechanical disruption and enzymatic digestion to yield single cell suspensions. Droplet-based scRNA-seq methods were employed via the 10x Genomics platform to generate single cell transcriptomic sequencing data, as described in section 2.1.2.2.

The following figure illustrates the ages of the prenatal skin samples used in this study, with representation across almost every week during this period of gestation:
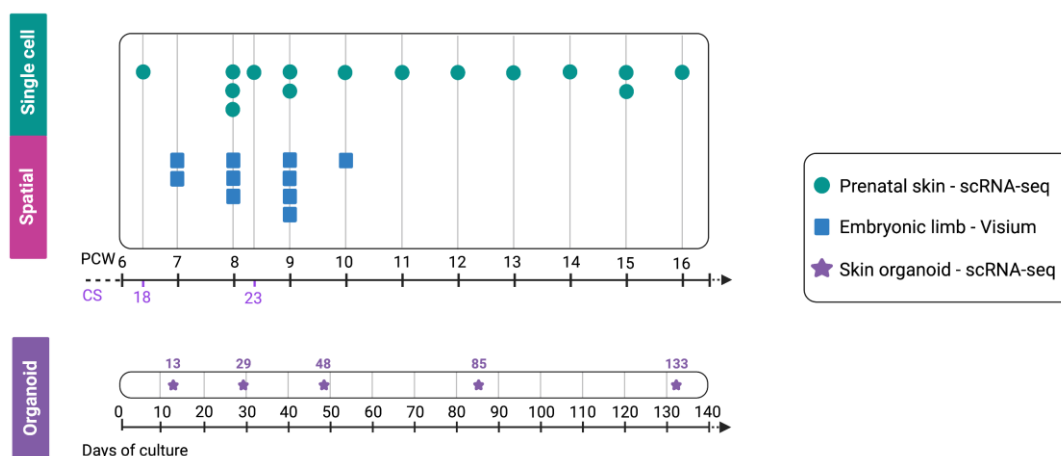
Figure 43: Sample overview. Schematic of prenatal skin scRNA-seq, embryonic limb Visium and skin organoid scRNA-seq data used in this study and the sample ages (in PCW) or days in culture for the respective datasets. Image created with BioRender.com.

The embryonic limb tissues from which spatial transcriptomic data was generated and published (Zhang et al., 2022) included samples ranging from 7 to 10 PCW, with every week during this period represented in the data, as seen in Figure 43 above.

The scRNA-seq data from the ESC/iPSC-derived hair-bearing skin organoid model (Lee et al., 2020) represented multiple time periods of culture between 13 days (~2 weeks) and 133 days (19 weeks or 4-5 months). The data time points are shown in Figure 43 above using a separate scale bar to that of the prenatal scRNA-seq and spatial samples.

4.2.2. Prenatal skin sample processing

After obtaining and digesting the prenatal skin samples into single cell suspensions, cells were FACS-isolated prior to single cell RNA sequencing. The sorting strategy involved removing cellular debris, dead cells and doublets to obtain live, single prenatal skin cells (Figure 44). The cells were then sorted to isolate CD45 positive and negative fractions (Figure 44), before proceeding with the scRNA-seq experiment.
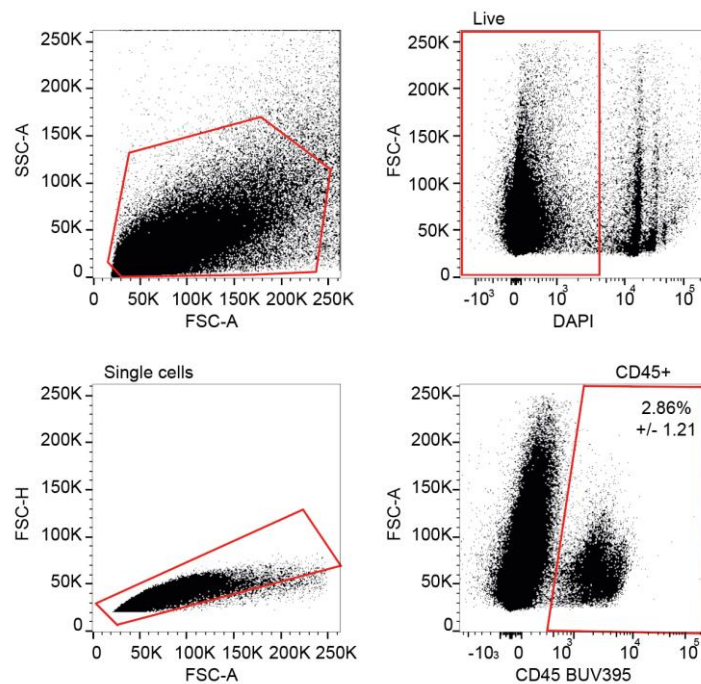
Figure 44: Sorting strategy. Live, single prenatal skin cells were isolated and sorted by CD45+ and CD45- fractions. Representative of n=15, data is shown as mean percentage +/- SD values.

4.2.3. Computational analysis and quality control

The prenatal scRNA-seq data was aligned with CellRanger software (10x Genomics) to a reference human genome (see Methods section 2.5.2.1 for details). Counts matrices were generated that comprised of cells (columns) and genes (rows) for each sample, which were combined (concatenated) to build a single large matrix, prior to pre-processing the data, including quality control (QC).

The total prenatal skin data object comprised of 235,201 cells. The data was then examined to remove poor quality cells using several pre-processing and QC steps, in order to account for any technical noise that may obscure biological signals in the data. The data pre-processing steps were carried out in Scanpy (Wolf, Angerer and Theis, 2018) and included the detection and removal of potential doublets, due to the possibility of more than one cell being trapped in a droplet during encapsulation (see Methods section 2.1.2.2.3.1. This was carried out using the Scrublet (**S**ingle-**c**ell **r**emoval of do**ublet**s) package (Wolock, Lopez and Klein, 2019).

Low quality cells were then filtered out by removing cells that contained fewer than 100 genes and greater than 45,000 UMIs. Furthermore, the mitochondrial UMI fraction was assessed for each cell; if this value is too high, it may represent cellular damage secondary to membrane perforation, cytoplasmic RNA leakage and a large proportion of retained mitochondrial RNA in the cell (Lun, McCarthy and Marioni, 2016). Therefore, cells with a maximum mitochondrial UMI fraction equal to 15% were also filtered out.

Possible maternal contamination was identified using the Souporcell pipeline (Heaton et al., 2020), which is a computational method that clusters cells by genotype using genetic variants identified within the scRNA-seq reads. This resulted in the removal of 49 cells at this stage of the quality control process.

After pooling the data from all the samples, a further quality control step was carried out, whereby genes detected in fewer than 3 cells were removed. The data was then normalised to 10,000 UMI per cell in order to account for the differences in sequencing coverage between cells. Log transformation was subsequently carried out so that differences gene expression is analysed on a relative, rather than an absolute, scale (Lun, 2018).

The quality control metrics for each prenatal skin sample are shown below (Figure 45).
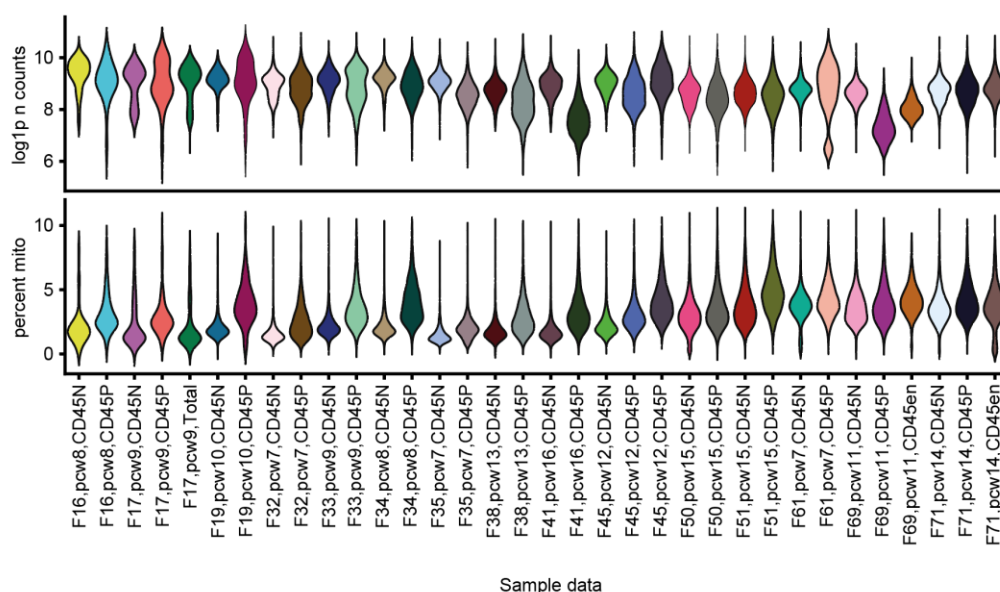
Figure 45: Computational quality control (QC). Violin plots to show the QC metrics across the prenatal skin scRNA-seq data, including the frequency distribution of UMI counts (log-transformed) and percentage of UMI counts in mitochondrial genes per sample fraction (CD45+/-).

Following pre-processing, quality control and doublet removal, 186,582 cells were taken forward for downstream analysis.

The discovery of highly variable genes based on normalised dispersion was then carried out in order to identify the genes that contribute most to cell-to-cell variation within homogenous cell groupings. Next, dimensionality reduction was carried out using by Principal Component Analysis (PCA) to transform the sequencing output into graphs that can be interpreted more easily to gain biological insights into the data. The first 50 principal components (PCs) were used to compute nearest-neighbour graph.

A batch correction approach was then taken to account for any possible technical variation in the scRNA-seq data. The batch balanced K-nearest neighbours (BBKNN) (Polański et al., 2020) algorithm and the Harmony tool (Korsunsky et al., 2019) was used to generate batch-corrected nearest-neighbour graphs, with each donor considered a separate batch. Full details on the computational approaches used for the prenatal scRNA-seq dataset are found in Methods section 2.5.2.1.

4.2.4. Broad prenatal cell annotations

The next step in the analysis was to use the Leiden algorithm to cluster cells and reveal transcriptionally similar groupings based on the batch-corrected graph. A relatively low resolution of 0.3 was used to cluster the cells into broad groupings.

Once the Leiden clusters were derived, the gene expression pattern in each cluster was assessed, allowing the cell clusters to be manually annotated into broad lineages based on known marker genes. From the 186,582 prenatal skin cells that passed QC, clustering and gene expression analysis resulted in 25 broad cell groupings, which are shown in the following Uniform Manifold Approximation and Projection (UMAP) embedding:
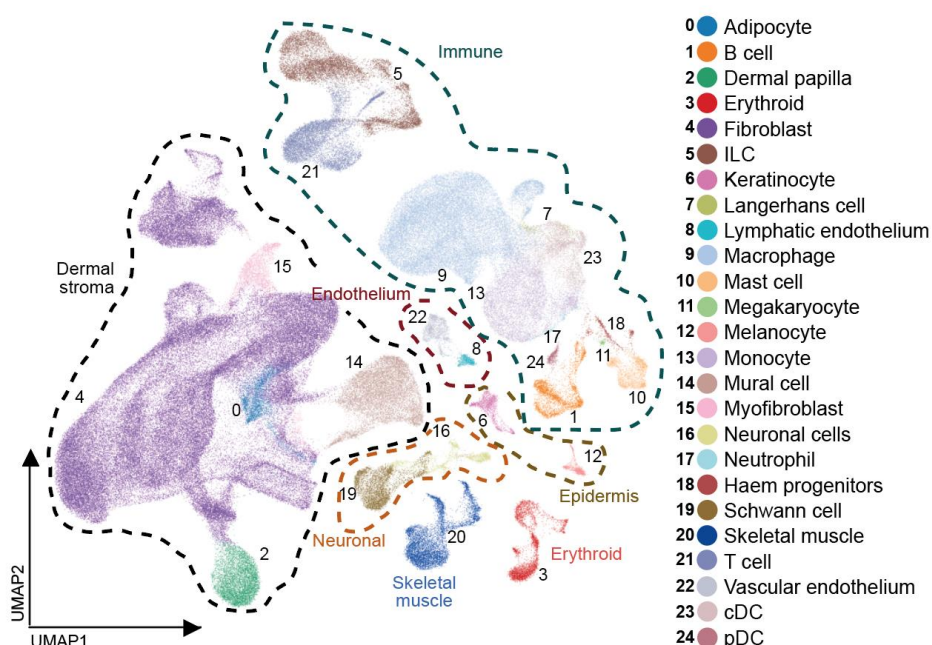
Figure 46: Prenatal broad cell clusters. UMAP visualisation of the prenatal scRNA-seq dataset (7-16 PCW, n = 15, k = 186,582) coloured by broad annotations of cell states. ILC = innate lymphoid cell, Haem = haematopoietic, cDC = classical dendritic cell, pDC = plasmacytoid dendritic cell.

The prenatal skin single cell dataset across first and second trimester of gestation is represented by epidermal cells, dermal stromal cells, endothelial cells, immune cells, neuronal cells, erythroid cells and skeletal muscle cells. These overall cell groupings comprise of 25 broadly annotated cell states, as shown in Figure 46, and marker genes are shown in Figure 47.
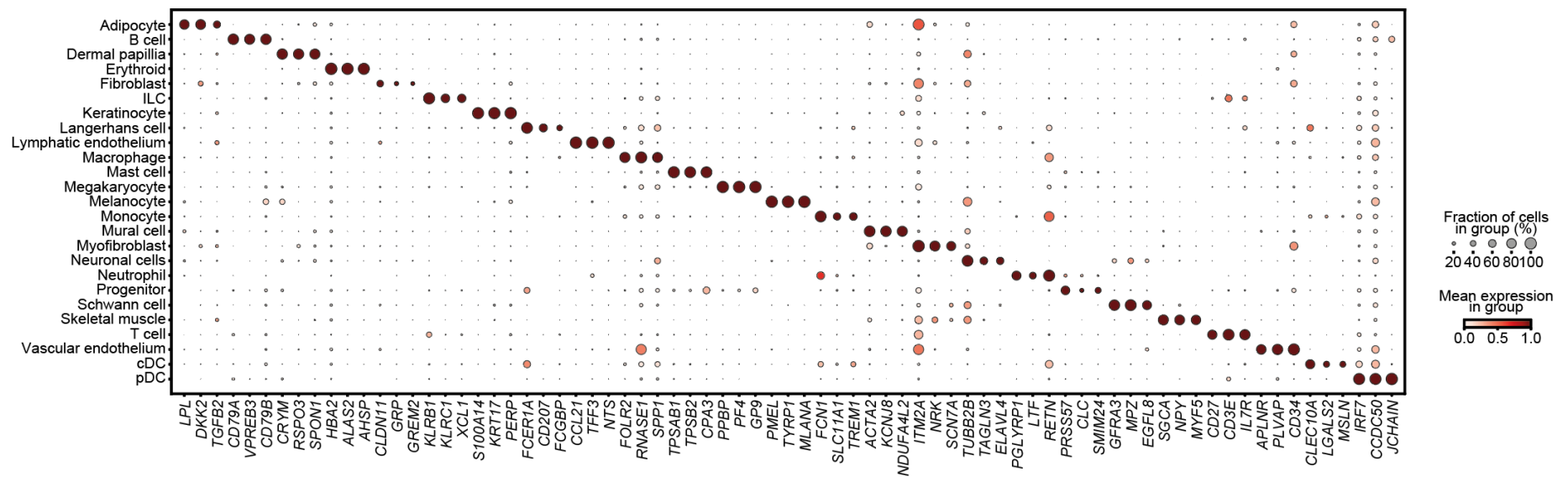
Figure 47: Differentially expressed genes (DEGs) of broad cell states. Dotplot showing DEGs in each broad cluster of the prenatal skin scRNA-seq dataset. The dot colour represents log-transformed, normalised and variance-scaled mean gene expression for each broad cell annotation and the dot size represents the percentage of each cell type expressing the marker gene.

The epidermal component of prenatal skin comprises of two main broad groupings of cells: keratinocytes and melanocytes. The keratinocyte population in prenatal skin is characterised by high expression of keratin 17 (*KRT17*) (Figure 47), which is a type I epithelial keratin that is expressed in the skin and provides mechanical support to keratinocytes to maintain epidermis integrity via desmosome attachments between cells (Windoffer et al., 2011; Porter and Lane, 2003). Developmental skin studies using murine models have identified that expression of *Krt17* first occurs in the single-layered undifferentiated epithelial cell layer, then in the ensuing days, the *Krt17*-expressing cells give rise to the periderm (a transient outer epithelial layer also seen in human development) and gives rise to placodes (which are precursors to ectoderm-derived appendages, including hair follicles and glands) (McGowan and Coulombe, 1998). This supports the key role of *KRT17* in developing skin keratinocytes, as supported by the high expression in the prenatal skin scRNA-seq data (Figure 47). The prenatal skin keratinocytes also highly express *PERP* (p53 apoptosis effector related to *PMP22*), which is also known to be important in stratified epithelial integrity and cell-cell adhesion due to its role in promoting intercellular desmosome assembly (Ihrie et al., 2005).

The melanocytes in first and second trimester prenatal skin highly express classical melanocyte gene markers, including *PMEL* (Premelanosome protein), *TYRP1* (Tyrosinase related protein 1) and *MLANA* (Melan-A). *PMEL* is enriched in melanosomes, which are melanin-forming organelles found in melanocytes, and *MLANA* has a vital role in the stability, trafficking and processing of the protein encoded by *PMEL* (Bissig, Rochin and van Niel, 2016). *TYRP1* encodes an enzyme expressed by melanocytes that plays an important role in the melanin biosynthetic pathway (Sturm, Teasdale and Box, 2001); defects in this gene are known to cause oculocutaneous albinism, which is a rare inherited disorder characterised by a reduction or a lack of melanin pigment in the skin, hair and eyes (Sarangarajan and Boissy, 2001).

The dermal stromal component of the prenatal scRNA-seq dataset comprises predominantly of fibroblasts, which reflects that seen in the skin dermis after birth. Other cell types observed within the prenatal stroma include the dermal papilla (a component of the hair follicle), mural cells (which comprise of pericytes and vascular smooth muscle cells (SMCs) that associate with the endothelial network), myofibroblasts and adipocytes (Figure 46). The prenatal stromal compartment that spans these 5 broad cell groupings were

also subclustered and annotated at a finer resolution into 12 refined cell groupings, including 7 subtypes of fibroblasts; this is discussed later in this chapter in section 4.2.6.2.

Within the endothelial cell compartment, vascular endothelial cells and lymphatic endothelial cells were observed (Figure 46). The endothelial cell compartment was similarly subclustered to reveal 8 refined cell clusters that are discussed in detail in section 4.2.6.3. Overall, however, the vascular endothelial cells were annotated by their high expression of *PLVAP* (Plamsmalemma vehicle-associated protein) and *CD34* (Figure 47). *PLVAP* is an endothelial cell-specific membrane protein that has a role in the formation of the diaphragms that bridge endothelial fenestrae, thereby influencing endothelial cell permeability (Denzer et al., 2023). *CD34* is also a known constitutively expressed marker of vascular endothelial cells that was used whilst broadly annotating the prenatal skin cell states to help distinguish vascular from lymphatic endothelium.

The immune compartment of the prenatal skin scRNA-seq dataset comprised of 12 broad categories of cell states, including myeloid cells (which were annotated at a higher resolution, as described later in the chapter in section 4.2.6.4 and lymphoid cells (which were also annotated further, as described in section 4.2.6.5). The gene expression markers used to identify broad categories of immune cells included *CD3E* for T cells, *CD79A* and *CD79B* for B cells and *CD207* for Langerhans cells (Figure 47) (Valladeau, Dezutter-Dambuyant and Saeland, 2003; Koyama et al., 1997).

The three remaining broad categories that were annotated in the prenatal skin scRNA-seq data were neuronal cells, skeletal muscle cells and erythroid cells (Figure 46). Neuronal cells were characterised by high expression of *TUBB2B*, *TAGLN3* and *ELAVL4* (Figure 47). *TUBB2B* (Tubulin Beta 2B Class IIb) has a functional role in axon guidance within both central and peripheral nervous system axon tracts (Cederquist et al., 2012), and *TAGLN3* (Transgelin 3) is a neurone-specific microtubule associated protein (Ratié et al., 2013). *ELAVL4* (ELAV RNA binding protein 4) is also a marker of neuronal cells that has a role in neurone-specific RNA processing contributing to neural progenitor cell proliferation and neural development (Bronicki and Jasmin, 2013; Beckel-Mitchener et al., 2002).

The skeletal muscle cells are characterised by high expression of *SGCA* (Sarcoglycan alpha) and *MYF5* (Myogenic factor 5) (Figure 47). *SGCA* expression is known to be restricted to skeletal muscle cells and encodes a component of the dystrophin-glycoprotein complex that is essential for muscle fibre stability (Sciandra et al., 2003). Mutations in the *SCGA* gene results in limb-girdle muscular dystrophy, which is characterised by progressive weakness of the muscles (Gonzalez-Quereda et al., 2018). The *MYF5* gene also represents a skeletal muscle cell marker owing to its expression pattern and role in muscle cell fate commitment and skeletal muscle cell differentiation (Zammit, 2017).

The prenatal erythroid cells showed high expression of *HBA2*, *ALAS2* and *AHSP* (Figure 47). *HBA2* (Haemoglobin subunit alpha 2) forms part of the haemoglobin molecule in erythroid cells and *AHSP* (Alpha haemoglobin stabilising protein) is involved in haemoglobin assembly (Krishna Kumar et al., 2010). *ALAS2* (5'-Aminolevulinate Synthase 2) encodes an erythroid specific mitochondrially located enzyme that catalyses the first step in the haem biosynthetic pathway (Krishna Kumar et al., 2010).

Having broadly annotated the prenatal skin scRNA-seq dataset, the cellular composition of skin during this developmental period across the first and second trimester was defined, allowing further downstream analysis for biological insight.

4.2.5. Integration with skin organoid and adult skin datasets

With the broad prenatal skin annotations in place, a comparison of the cellular composition and transcriptional similarity of prenatal skin with the cells that make up the ESC/iPSC-derived hair-bearing skin organoid model was made (Lee et al., 2020), thereby assessing how faithful the model is to *in vivo* skin. This also allows the potential utility that the skin model has for functional experimentation to be determined, such as disease modelling and tissue engineering. The prenatal skin and skin organoid was also integrated with an adult interfollicular skin dataset (Reynolds et al., 2021) in order to contextualise the findings between each condition and reveal and additional or missing cell clusters, or identify any transcriptionally distinct clusters. This can provide more biological insight into how skin changes between prenatal and adult life, and whether the cell states in the skin organoid model resemble prenatal or adult skin cells.

Using the same approach that was taken for the prenatal skin data, the skin organoid scRNA-seq data was pre-processed, filtered, clustered and assigned broad cell annotations. Full details on these analyses are found in the Methods. To integrate the prenatal, organoid and adult skin cells, batch correction was carried out, treating each dataset as batches and treating within-dataset batches as covariates. The combined UMAP embedding was then coloured by dataset, showing the contribution of cells from prenatal skin, organoid and adult skin to each broad lineage, as shown in Figure 48 below:
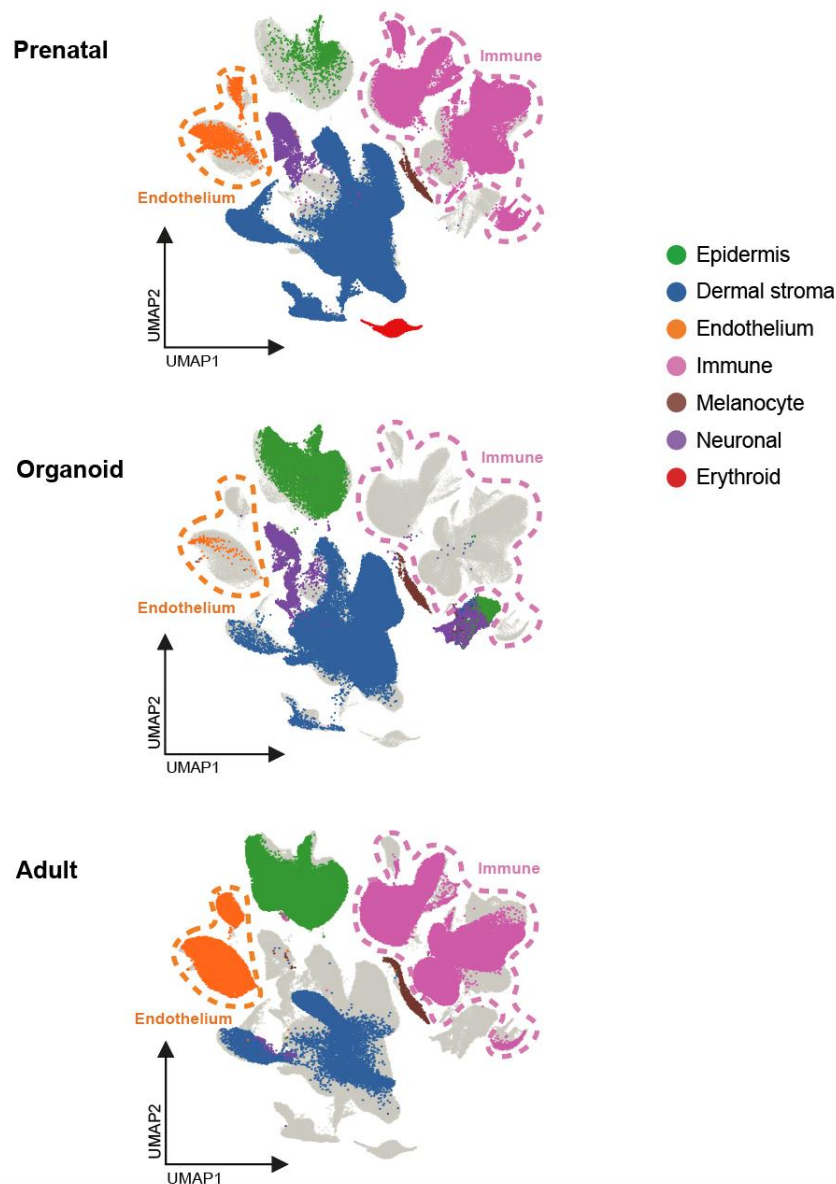
Figure 48: Integration of prenatal (top), organoid (Lee et al., 2020) (middle) and adult interfollicular skin (Reynolds et al., 2021) (bottom), shown as a UMAP projections coloured by broad cell groupings.

The integration between prenatal skin, organoid and adult skin highlighted the differences in broad cell populations between each dataset (Figure 48). Within the epidermal compartment, the prenatal skin data had a smaller representation of cell states compared with the skin organoid and adult skin. Furthermore, the adult skin cell states represented distinct portions of the epidermal embedding space, highlighting transcriptional differences between adult and prenatal/organoid epidermal cells.

The converse is seen in the dermal stromal compartment, where prenatal cell states occupy distinct areas in the combined integrated embedding space, again signifying transcriptionally distinct prenatal stromal cell populations that are absent in adult, but that are mostly represented in the skin organoid (Figure 48). This also demonstrates that within the stromal compartment, the skin organoid cell states are more comparable to prenatal that to adult skin.

Interestingly, within the endothelial compartment, the skin organoid model showed an incomplete representation of cell states compared to that seen in prenatal and adult skin, and within the immune compartment, the organoid model lacked immune cells altogether (Figure 48). These differences would be important to bear in mind when considering the use of this skin organoid model to investigate biological questions that relate to endothelial and immune cell interactions within the skin. As with other compartments, the prenatal and adult skin cell states within the immune compartment occupy different regions within the integrated embedding space, suggesting the present of distinct cell states.

Melanocytes are represented across all three datasets, whereas erythroid cells are only observed within the prenatal skin single cell data. The neuronal cells are only observed in prenatal skin and skin organoid, and are lacking in the adult skin data (Figure 48), which is likely to represent the sampling strategy for adult skin where only the most superficial 200 μm of skin was processed for scRNA-seq (Reynolds et al., 2021). Interestingly, the skin organoid data showed a large representation of neuronal cells in two discrete areas of the UMAP embedding space, which may reflect the experimental culture conditions promoting ectodermal differentiation (Lee et al., 2020).

To assess the transcriptionally distinct cell states in each dataset within the overall lineage groupings highlighted above, the broad level of annotation carried out for the prenatal scRNA-seq data that is shown in Figure 46 was extended to the organoid and adult skin datasets. This enabled distinct populations between prenatal, organoid and adult datasets to be assessed within the same UMAP embedding space:
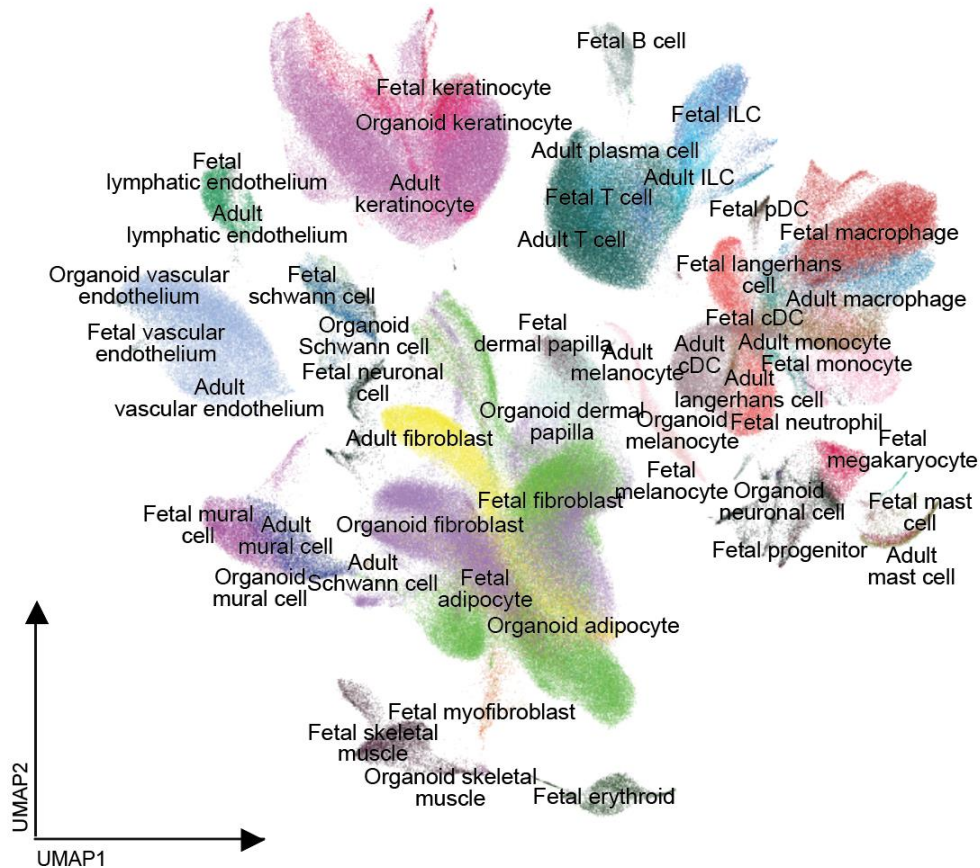
Figure 49: Annotated integration of prenatal, organoid (Lee et al., 2020) and adult interfollicular skin (Reynolds et al., 2021).

This analysis using the broad annotations of cell states within each lineage highlighted the similarities and differences between prenatal skin, organoid and adult skin (Figure 49). For example, cell types that were identified in prenatal skin that were absent in the skin organoid and in adult skin included B cells, megakaryocytes and erythroid cells. Furthermore, most subpopulations of immune cells were observed in both prenatal and adult skin, and were absent in the organoid model, as discussed above. The same was observed for lymphatic endothelial cells, which were observed in prenatal and adult skin, but not in the skin organoid. The dermal papilla cells were identified in the prenatal and skin organoid datasets, but were absent in adult skin, likely reflecting the depth of sampling,

as discussed above. By highlighting the presence and absence of discrete cell clusters in each dataset, the biological differences between skin before birth and in adulthood can be appreciated, and the cell populations that are absent in the skin organoid model are noted.

Within broad cell groupings that span different datasets, the transcriptional similarity between the broad cell states across prenatal skin, skin organoid and adult skin was assessed. This was done by calculating the average distance between cell states in the integrated and batch corrected latent space, which was represented as a heatmap, as shown below in Figure 50.

Overall, this revealed a strong transcriptional correlation of most broad cell states between *in vivo* prenatal, skin organoid and adult skin, reflecting their proximity in the integrated embedding space. However, transcriptional similarity was observed between fibroblasts in prenatal skin, organoid and adult skin, myofibroblasts in prenatal skin and adipocytes in prenatal skin and skin organoid; of note, myofibroblasts were not annotated in the organoid and adult skin datasets and adipocytes were not included in the adult skin dataset because skin was sampled down to the dermal layer of skin (Reynolds et al., 2021). The transcriptional similarity between fibroblasts, prenatal myofibroblasts and prenatal / organoid adipocytes may reflect the early differentiation trajectories of myofibroblasts and adipocytes from fibroblast-like mesenchymal stem cells (MSCs) (Cristancho and Lazar, 2011), explaining their proximity to the large number of fibroblast in the combined integrated embedding space.
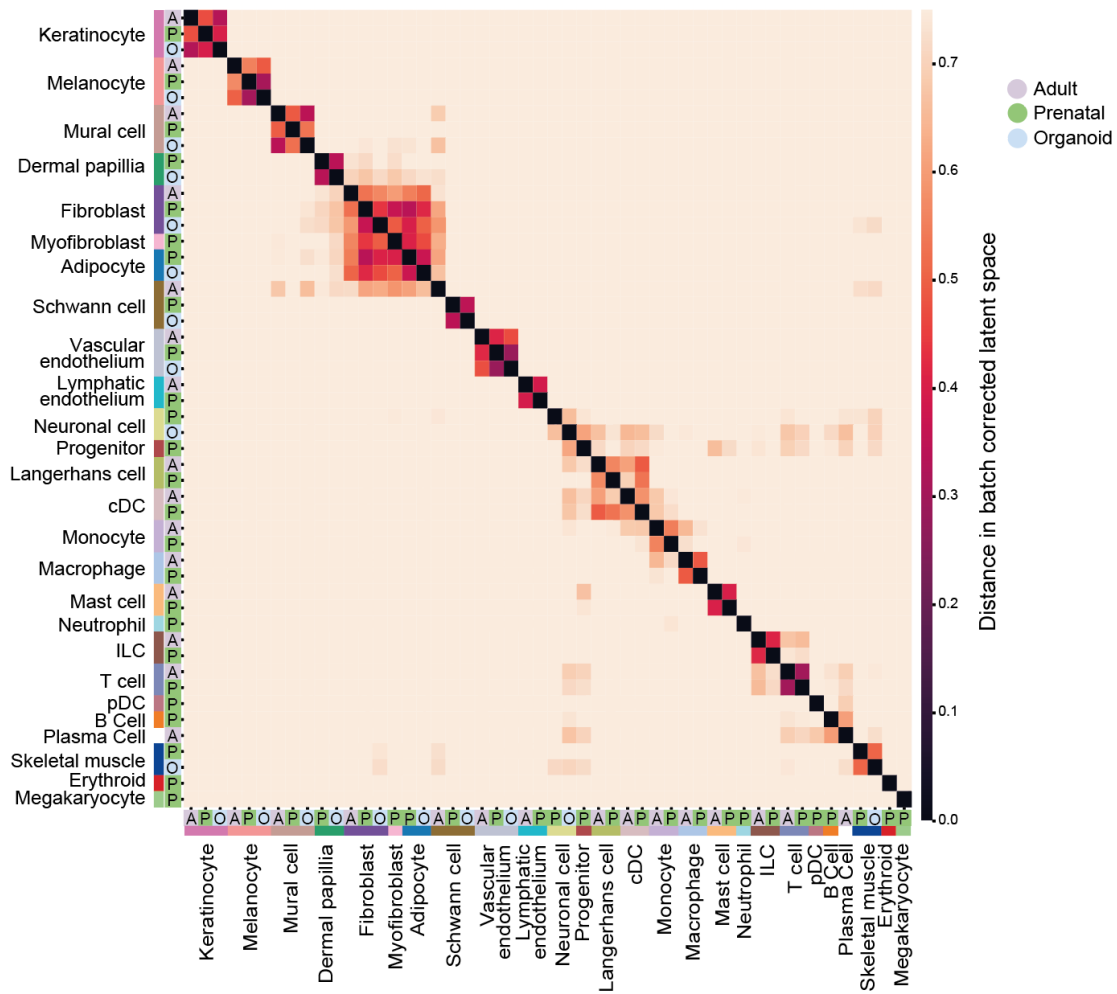
Figure 50: Heatmap showing transcriptional similarity (measured by distance in PC space) between cell states in prenatal skin and skin organoid (Lee et al., 2020) and adult healthy skin (Reynolds et al., 2021).

In order to determine whether the skin organoid model recapitulates physiological differentiation during prenatal development at a molecular level, a logistic regression (LR) framework was used to compare the transcriptional similarity of cells at individual developmental time points, expressed as PCW in the prenatal data and weeks of culture in the organoid data. A model was computationally built and trained on the gene expression matrix of the training dataset, which was the merged prenatal and adult dataset, using all the genes that had passed quality control. The resulting model was used to make predictions in the target organoid dataset. The correspondence between the training and target datasets for each developmental time point was computed as a Jaccard index and visualised as a heatmap (Figure 51). Full details on the analysis approach are found in the Methods.
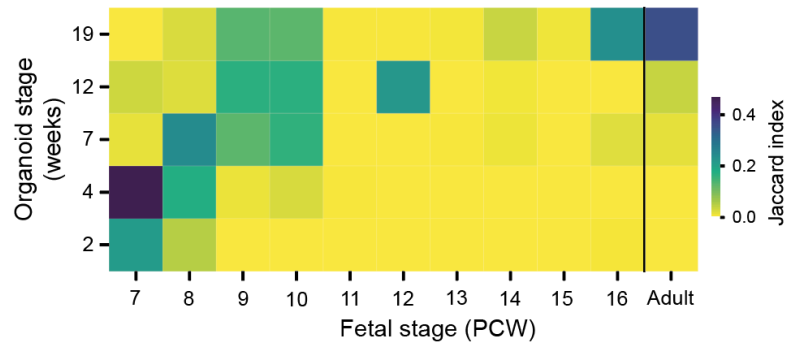
Figure 51: Heatmap showing transcriptional similarity between cell states in prenatal, organoid (Lee et al., 2020) and adult skin (Reynolds et al., 2021). The similarity is measured as the Jaccard index between organoid stages (y-axis, weeks of culture) and per-cell predicted prenatal stages (x-axis, PCW), using a logistic regression model trained on the prenatal/adult skin dataset.

This LR analysis showed that although the differentiation of skin organoids across culture duration resembles prenatal skin development, it progresses at an accelerated tempo. After 4 weeks of culture, skin organoid cell states overall align with prenatal skin at 7 PCW, and after 19 weeks of culture, align with cells in adult skin.

The analysis above included all broad cell types from each dataset, and therefore, molecular changes over time across separate cell lineages was assessed using the same computational approach (Figure 52). This showed that the precise timing of differentiation varies depending on cell lineage, with fibroblasts, endothelial cells and Schwann cells aligning with 7 PCW prenatal skin at 4 weeks of organoid culture, and keratinocytes and melanocytes aligning most closely with 7 PCW skin at 2 weeks of culture.
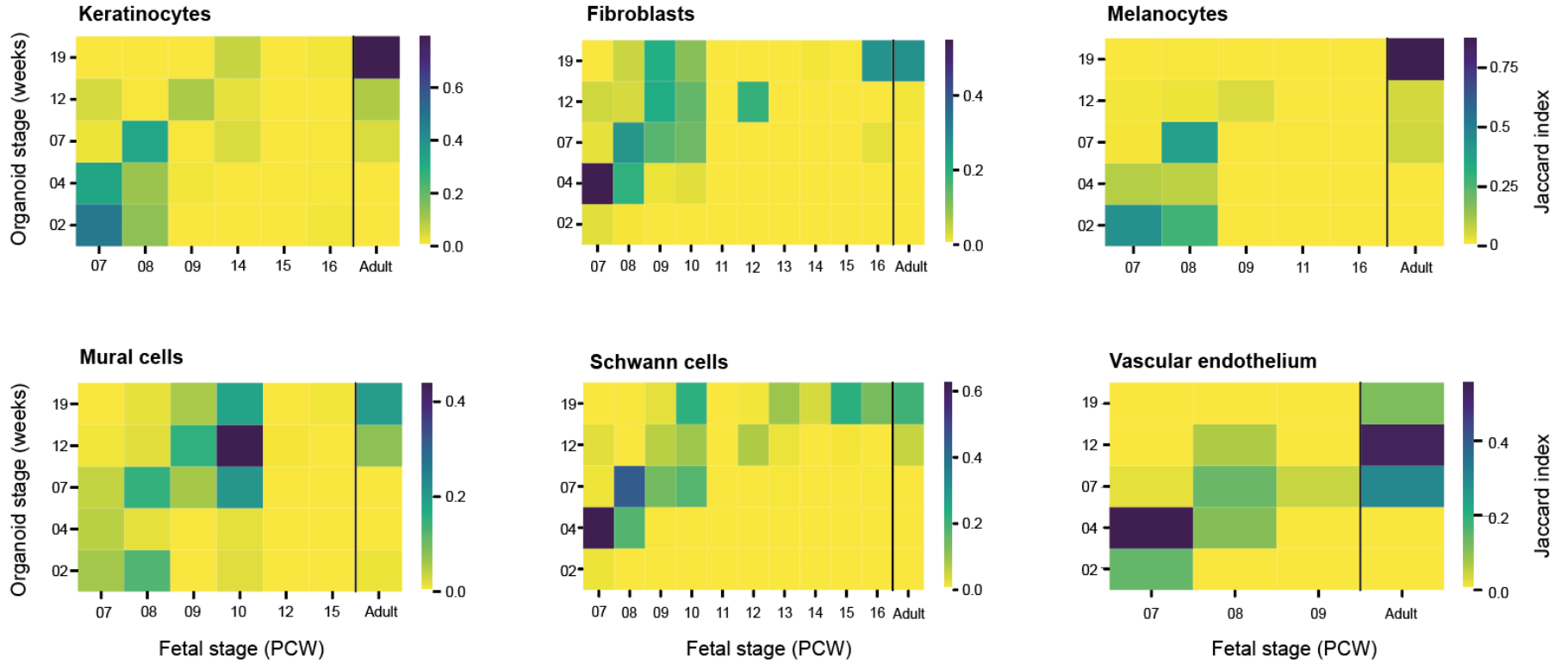
Figure 52: Heatmap showing transcriptional similarity between cell states in prenatal, organoid (Lee et al., 2020) and adult skin (Reynolds et al., 2021) for broad cell categories. The similarity is measured as the Jaccard index between organoid stages (y-axis, weeks of culture) and per-cell predicted prenatal stages (x-axis, PCW), using a logistic regression model trained on the prenatal/adult skin dataset.

4.2.6. Refined prenatal cell annotations

The prenatal skin scRNA-seq data was annotated to a higher resolution to characterise refined cell states within each broad lineage. This was carried out by re-processing the data for each broad lineage, starting from highly variable gene selection to reveal the finer heterogeneity in the dataset. Batch correction was then carried out, treating each donor as a separate batch. Leiden clusters at a high resolution were derived and annotated manually using marker genes identified through literature searches and using differentially expressed genes (DEGs) that were calculated for each cluster. Full details on these analysis steps are found in methods section 2.5.2.1.

4.2.6.1. Epithelial cells

The epithelial cells within the prenatal skin single cell dataset were relatively few in number and failed to represent several portions of the integrated embedding space when compared to the hair-bearing skin organoid (Lee et al., 2020) and adult skin (Reynolds et al., 2021) datasets, as discussed in the previous section and shown in Figure 48. When attempting to use the Leiden clustering algorithm iteratively at high resolutions for the prenatal skin epithelial lineage, biologically relevant clusters and distinct hair follicle cell states did not cluster separately, which was likely due to the relatively small cell numbers in this subset. Therefore, an integrated combined embedding of the prenatal and organoid epithelial cell states was generated using Harmony batch correction (Korsunsky et al., 2019) in order to pull apart epithelial cell states in the nearest neighbourhood graph and allow fine-grained annotation to be carried out. The epithelial cell cluster annotations were based on marker genes and the refined annotations in the organoid scRNA-seq dataset (Lee et al., 2020).

The following figure shows separate UMAP visualisations for the prenatal and organoid epithelial cell states, which were based on the same integrated embedding space:
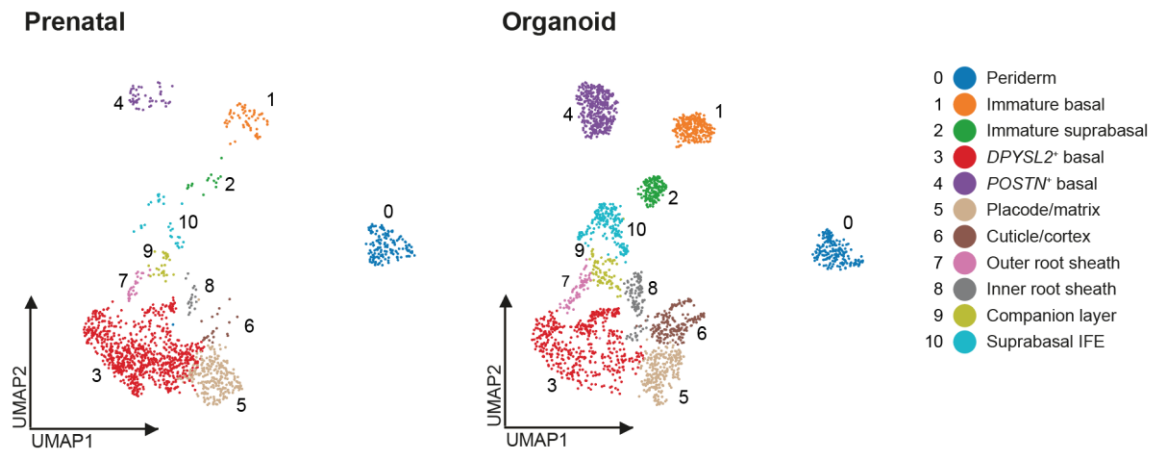
Figure 53: UMAP visualisations of integrated data from prenatal skin and skin organoid (Lee et al., 2020), coloured by refined epithelial cell types. IFE = interfollicular epidermis.

Overall, 11 broad clusters of epithelial cell states were identified and annotated, including periderm cells, interfollicular epidermal cells (immature basal, immature suprabasal, *DPYSL2*[+] basal, *POSTN*[+] basal and suprabasal IFE) and hair follicle specialised epithelial cells (placode/matrix, cuticle/cortex, outer root sheath, inner root sheath and companion layer) (Figure 53). The periderm is a transient outer layer of prenatal skin that is present from about 4 PCW (King, Balaji and Keswani, 2013). It acts as a permeability barrier and is shed into the amniotic fluid following the second trimester when epidermal stratification begins (Lee and Wine, 2019). The cluster containing periderm cells was characterised by high expression of *KRT4*, *KRT8* and *KRT18* (see Appendix N for prenatal skin cell refined cell annotation DEGs), which are commonly expressed in single layer epithelial tissues (Kalabusheva et al., 2023). The periderm cells also highly express *SCEL* (sciellin), which is a known precursor to the cornified envelope of terminally differentiated keratinocytes (Kvedar et al., 1992), and *MUC16* (Mucin 16, Cell Surface Associated), which has a role in forming a protective mucous barrier in the prenatal periderm (Rump et al., 2004). The periderm cells are present in the prenatal skin dataset only until 11 PCW (Figure 54), though previous studies have reported that the periderm is sloughed into the amniotic fluid towards the end of the second trimester (Hardman et al., 1999).
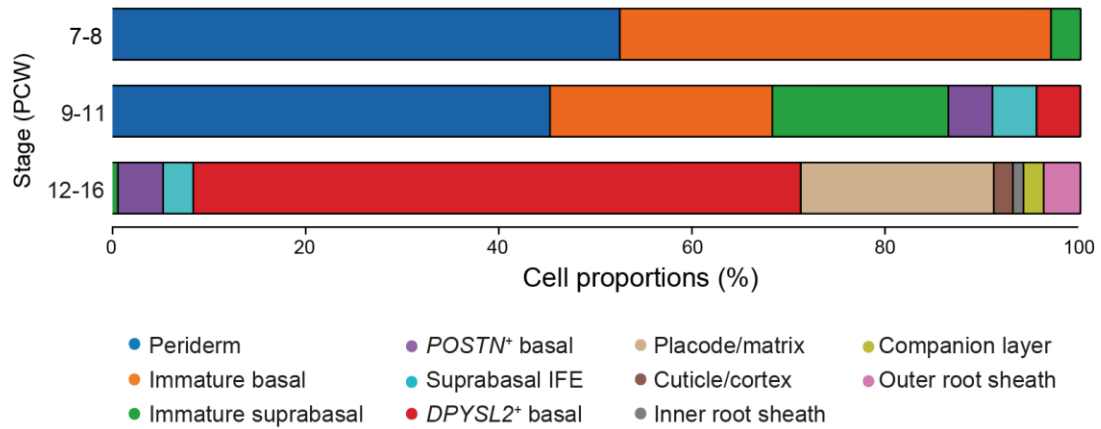
Figure 54: Bar plot showing frequency of epithelial cell states across gestational age in prenatal skin, with colour of bar representing each cell type.

Alongside the periderm, the immature basal and immature suprabasal cell states represented predominant other cell states in prenatal skin during the first trimester (Figure 53 and Figure 54). The immature basal cells expressed *KRT19*, which is typical for simple epithelium (Kalabusheva et al., 2023) and the immature suprabasal cells express *KRT1* and *KRT 10,* which are markers of suprabasal epidermis. The suprabasal cells also express *DSC1* (desmocollin 1), which is a desmosomal component expressed by epithelial cells that mediates cell-cell adhesion (Appendix N) (Nuber et al., 1996).

A discrete cluster of cells present in the prenatal and organoid datasets was annotated as *POSTN*[+] basal cells (Figure 53) owing to its high expression of *POSTN* (periostin), which promotes cellular attachment and adhesion (Gillan et al., 2002). The *POSTN*[+] basal cells also highly express *KRT14* (a marker of basal cells) and *COL17A1* (Appendix N). This expression pattern is similar to a cluster of basal stem cells identified in a recent scRNA-seq study of human newborn skin, where immunofluorescent imaging using of the neonatal epidermis was also carried out and showed that these cells are enriched in specific zones between the rete ridges (Wang et al., 2020b). The *DPYSL2*[+] basal cells, which also express the *KRT14* basal cell marker, then makes up the predominant cell type in prenatal skin beyond 12 PCW (Figure 54). The suprabasal IFE (interfollicular epidermis) cells were present from 9 PCW onwards (Figure 54) and showed high expression of *KRTDAP* (keratinocyte differentiation associated protein) (Appendix N), which is known to be more highly expressed in suprabasal compared to basal epidermal cells (Kabir et al., 2022).

The remaining annotated cell clusters from the combined prenatal and organoid embedding comprised of hair follicle cell states, including the placode/matrix, cuticle/cortex, inner root sheath, companion later and outer root sheath. As discussed above, the prenatal cells were annotated using the hair cell annotations from the organoid dataset, which had a greater representation of hair follicle cell states than the prenatal scRNA-seq data had (Figure 53). The prenatal skin cell DEG analysis, however, did highlight several known markers of these hair follicle cell types, including *SHH* (Sonic Hedgehog Signalling Molecule) and *WNT10A* (Wnt Family Member 10A) for placode/matrix, *KRT75* for the companion layer and *KRT6B* for the outer root sheath (Appendix N) (Kalabusheva et al., 2023; Reddy et al., 2001).

4.2.6.2. Dermal stromal cells

The prenatal dermal stromal cells were annotated using known markers derived from literature search and DEGs calculated between Leiden clusters on a prenatal skin stromal lineage embedding. The following UMAP visualisation shows the annotated stromal cell clusters in the prenatal skin dataset, which is presented in a combined integrated embedding with the organoid stromal cells (Lee et al., 2020), demonstrating transcriptionally similar refined cell states between prenatal skin and skin organoid.
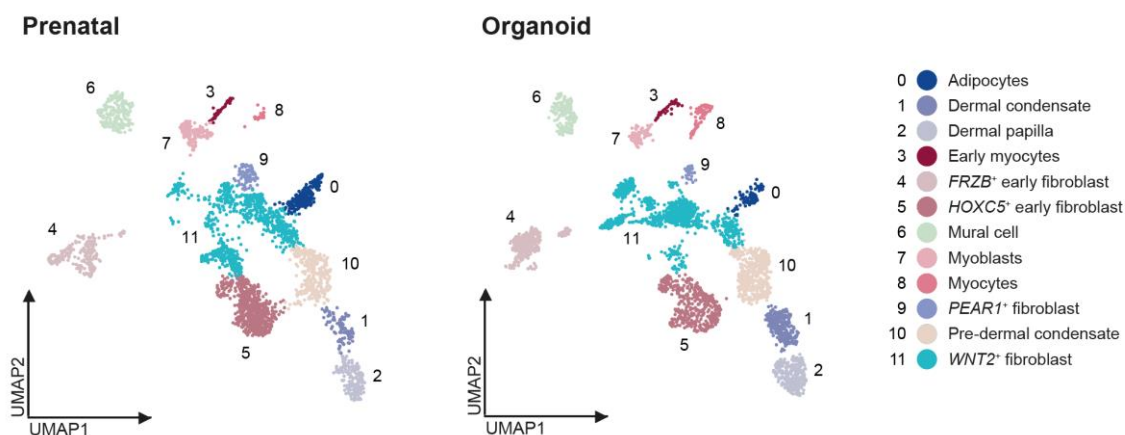
Figure 55: UMAP visualisations of integrated data from prenatal skin and skin organoid (Lee et al., 2020), coloured by refined stromal cell types.

The annotation of the adipocytes, mural cells and muscle cells (myocytes) using known marker genes was discussed earlier in this chapter in section 4.2.4. The rest of the stromal

compartment comprises of fibroblasts, which include those related to the prenatal hair follicle (pre-dermal condensate, dermal condensate, dermal papilla). The hair follicle associated fibroblasts will be discussed in detail in chapter 4, including their differentiation trajectory and the genes they express that mediate their functional roles and differentiation pathway. The non-hair follicle associated fibroblasts include *FRZB*⁺ early fibroblasts, *HOXC5*⁺ early fibroblasts, *WNT2*⁺ fibroblasts and *PEAR1*⁺ fibroblasts, which were annotated based on genes that are highly expressed between the fibroblast clusters (Appendix N). The fibroblasts that were given the 'early' annotation (*FRZB*⁺ early fibroblasts and *HOXC5*⁺ early fibroblasts) were identified in the earlier gestational prenatal skin samples (as shown in Figure 60). The non-hair follicle associated fibroblast subsets and the genes mediating their differentiation trajectory are also discussed in detail in chapter 4.

4.2.6.3. Endothelial cells

The prenatal endothelial cell compartment comprised of vascular and lymphatic endothelium; the distinction between these broad cell types was discussed in detail in section 4.2.4. Within the vascular endothelial compartment, arteriolar, venular and capillary cell states were identified, as shown in the following UMAP visualisation:
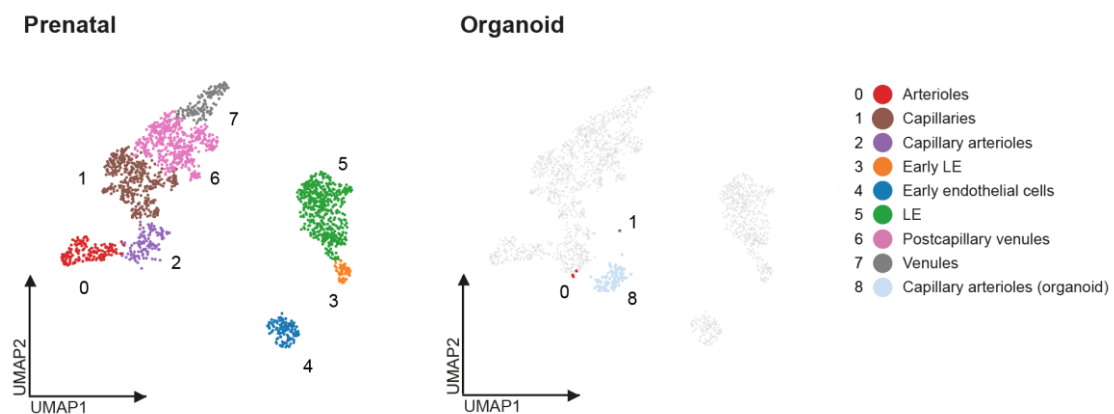
Figure 56: UMAP visualisations of integrated data from prenatal skin and skin organoid (Lee et al., 2020), coloured by refined endothelial cell types. LE = lymphatic endothelium.

By creating a combined integrated embedding between prenatal and organoid endothelial cells and plotting the cell states separately, the paucity of endothelial cells in the skin organoid model (Lee et al., 2020) is highlighted (Figure 56) as discussed above in section 4.2.5.

Within the prenatal vascular endothelial cell grouping, the arteriolar cells are characterised by high expression of *GJA5* (Gap Junction Protein Alpha 4), which is involved in arteriolar morphogenesis (Lu and Wang, 2017) and *FCN3* (Ficolin 3), which is an arteriole-specific marker gene (Voigt et al., 2022) (Appendix N). The prenatal capillary arteriole cells highly express *NETO2* (Neuropilin and Tolloid Like 2), which is involved in vascular patterning and is implicated in infantile proliferative capillary haemangiomas (Calicchio, Collins and Kozakewich, 2009). The capillary arteriole cells also express *PRND* (Prion Like Protein Doppel) and *PRRG3* (Proline Rich And Glia Domain 3), both of which have an important role in endothelial cell development and sprouting angiogenesis (Chen et al., 2020c; Crouch et al., 2023). The capillary cells were annotated based on their high expression of *CD36* and *FABP4*, which are known marker genes (Iso et al., 2013; Son et al., 2018).

The venular cells within the prenatal skin dataset included the venules and postcapillary venules. The venule cells showed high expression of *FBLN2* (Fibulin 2), which encodes an extracellular matrix (ECM) protein. In a recently published scRNA-seq study profiling adult skin and enriching for endothelial cell states, *FBLN2* was shown to be highly expressed in adult skin venule cells in comparison to capillary cells (Li et al., 2021), reflecting that seen in the prenatal data. Furthermore, this study on adult skin demonstrated that *SELE* (Selectin E), which mediates leucocyte adhesion to the endothelium (Lawrence and Springer, 1993), was a highly specific marker for postcapillary venules (Li et al., 2021); *SELE* was also highly and specifically expressed in postcapillary venule cells in the prenatal skin dataset (Appendix N). The postcapillary venules also expressed *IFI27* (Interferon Alpha Inducible Protein 27), which is a known venous endothelial cell marker (Crouch et al., 2023) that is also expressed on capillary cells (Schupp et al., 2021), reflecting this postcapillary segment of the vasculature in prenatal skin.

4.2.6.4. Myeloid cells

The organoid data is not shown alongside in an integrated embedding, as was carried out for the epithelial, stromal and endothelial lineages, because the skin organoid lacked immune cells (Lee et al., 2020); this was discussed above in section 4.2.5.
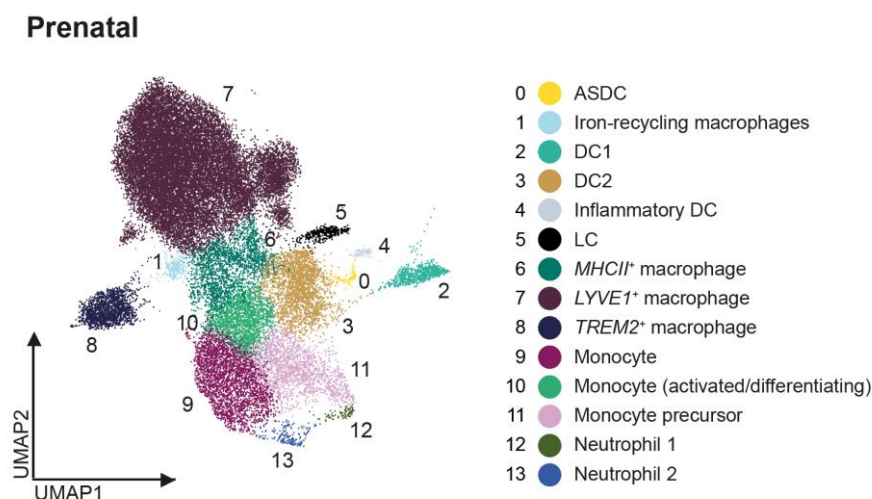
Figure 57: UMAP visualisation of data from prenatal skin, coloured by refined myeloid cell types. DC = dendritic cell, ASDC = AXL⁺SIGLEC6⁺ DC, LC = Langerhans cell.

The prenatal myeloid cell compartment included dendritic cell (DC), Langerhans cells (LC), monocytes, macrophages and neutrophils (Figure 57). The DC subsets annotated in the prenatal skin dataset included DC1, DC2, inflammatory DC and ASDC ($AXL^+SIGLEC6^+$ dendritic cell). The DC1 cell state was characterised by high expression of *CLEC9A* (C-Type Lectin Domain Containing 9A), which is an endocytic receptor specialised for the uptake and processing of material from necrotic cells (Zhang et al., 2012), and is a specific marker gene for the DC1 subset (Villani et al., 2017). The DC1 cells also highly expressed *IDO1* (Indoleamine 2,3-Dioxygenase 1) (Appendix N), which is a tryptophan metabolising enzyme also known to be expressed by DC1 cell states (Chen et al., 2020b). The DC2 cells were annotated based on their high expression of *CLEC10A* (C-Type Lectin Domain Containing 10A), and were also noted to highly express *IL1R2* (Interleukin 1 Receptor Type 2), which has been shown by scRNA-seq of peripheral blood mononuclear cells (PBMCs) to be highly expressed in DC2 subsets (Gao et al., 2021). The inflammatory DCs were characterised by high *LAMP3* (Lysosomal Associated Membrane Protein 3) expression (Appendix N, which is a marker for mature DCs that is upregulated

upon activation (Salaun et al., 2004), and CCL22 (C-C Motif Chemokine Ligand 22), which is a cytokine expressed by DCs upon activation (Vulcano et al., 2001) that mediates cell interaction with regulatory T cells (Tregs) (Rapp et al., 2019). The ASDC cell cluster was annotated owing to its high expression of *SIGLEC6* (Sialic Acid Binding Ig Like Lectin 6) (Appendix N), which is a defining marker gene for ASDCs. *HAMP* (Hepcidin Antimicrobial peptide) was also highly expressed by the prenatal ASDC cells, which has been shown to be a reproducible gene expression marker based on integration of scRNA-seq data from multiple studies (Hao et al., 2023). The prenatal ASDC cell state also highly expressed *LTK* (Leucocyte Receptor Tyrosine Kinase), which represents one of the top most discriminative genes for ASDCs (Villani et al., 2017). The Langerhans cells were annotated based on their expression of *CD207* and *CD1A* (Appendix N), as discussed above in the annotation of broad cell groupings in section 4.2.4.

The prenatal monocyte cell states were annotated as monocyte precursors, monocytes and activated/differentiating monocytes (Figure 57). The monocyte precursors differentially expressed *ZBTB71* (Zing Finger And BTB Domain Containing 7A), which encodes a transcription factor that represses monocyte differentiation (Redondo Monte et al., 2020), and *SERPINB8* (Serpin Family B Member 8), which inhibits monocyte transformation into macrophages (Kappert et al., 2013). The monocyte cell cluster was characterised by high expression of *VNN2* (Vanin 2), which is known to be expressed on mature monocytes in comparison to other lymphocytes (Bornhauser et al., 2020) and *CLEC4E* (C-Type Lectin Domain Family 4 Member E), which is monocyte cell marker with a role in pathogen recognition and phagocytosis (Gren et al., 2015) (Appendix N). The activated / differentiating monocyte cluster was annotated based on high expression of *LYZ* (Lysozyme), which has a role in host defence against bacterial microorganisms, *CD14*, which mediates innate immune responses and *S100A8/9* (S100 Calcium Binding Protein A8/9), which modulates inflammatory responses by inducing cytokine secretion and stimulating leucocyte recruitment (Wang et al., 2018), all of which are recognised markers for this monocyte subset (Kapellos et al., 2019).

The prenatal macrophage subsets annotated within the dataset included iron-recycling macrophages, *MHCII*+ macrophages, *LYVE1*+ macrophages and *TREM2*+ macrophages. These subsets were assigned based on the macrophage classification presented in a recently published integration of prenatal immune cells where scRNA-seq data from nine different

tissues were included across haematopoietic, lymphoid and non-lymphoid peripheral organs (Suo et al., 2022). The iron-recycling macrophages expressed the high levels of *SLC40A1* (Solute Carrier Family 40 Member 1, also known as Ferroportin-1), which encodes a cell membrane iron transporter that balances cellular iron levels (Donovan et al., 2005) and are characterised by high expression of *CD5L* (CD5 Molecule Like), which is a macrophage marker that promotes macrophage survival against apoptosis-inducting stimuli (Miyazaki et al., 1999). The *MHCII⁺* macrophages express the highest levels of *HLA-DRA* and *HLA-DPA1* amongst the macrophage subsets (Suo et al., 2022) and the *LYVE1⁺* macrophages highly express *LYVE1* (Lymphatic Vessel Endothelial Hyaluronan Receptor 1), which has a known role in supporting blood vessel integrity and angiogenesis (Kieu et al., 2022). The *TREM2⁺* macrophages express microglia-associated transcripts, including *P2RY12* (Purinergic Receptor P2Y12) and *TMEM144* (Transmembrane Protein 144) (Appendix N) (van Wageningen et al., 2019).

The prenatal neutrophil cell states clustered into two transcriptionally distinct subsets (Figure 57). The Neutrophil 1 population was characterised by specific expression of *DEFA4* (Defensin Alpha 4, also known as Neutrophil Defensin 4), which encodes an antimicrobial peptide found abundantly in the granules of neutrophils (Basingab et al., 2022), and high expression of *AZU1* (Azurocidin), which encodes a neutrophil granule derived peptide with monocyte chemotactic activity (Almansa et al., 2012). The Neutrophil 2 subpopulation was characterised by specific expression of *LTF* (Lactotransferrin, also known as Neutrophil Lactoferin), which is an iron-binding protein uniquely abundant in neutrophils (Zhao et al., 2018), and high expression of CAMP (Cathelicidin Antimicrobial Peptide, also known as LL37) which encodes an antimicrobial peptide important in innate immune defence (Li et al., 2006).

4.2.6.5. Lymphoid cells

The annotated refined cell states within the lymphoid compartment of the prenatal scRNA-seq dataset is shown in the UMAP visualisation below (Figure 58).
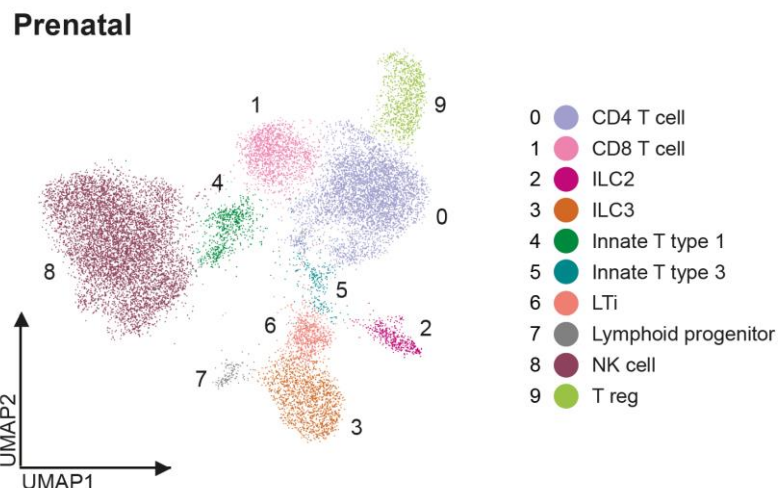
Figure 58: UMAP visualisation of data from prenatal skin, coloured by refined lymphoid cell types. ILC = innate lymphoid cell, LTi = lymphoid tissue inducer cell, NK cell = natural killer cell, Treg = regulatory T cell.

The prenatal skin CD4 and CD8 T cells were annotated based on their expression of *CD4* and *CD8A/B*, respectively, as well as their shared expression of *CD3E* and *IL7R*, as was discussed above in section 4.2.4. The ILC2 cell state was characterised by high expression of *PTGDR2* (Prostaglandin D2 Receptor 2), which is a known gene expression marker (Mazzurana et al., 2021), and *IL9R* (Interleukin 9 Receptor), which is constitutively expressed by ILC2 cells (Mohapatra et al., 2016) (Appendix N). The prenatal ILC3 cell cluster highly expressed *HPN* (Hepsin), which is a serine protease, and *SCN1B* (Sodium Voltage Gated Channel Beta Subunit 1); both genes have been recently shown using scRNA-seq to be highly co-expressed in a subset of innate lymphoid cells (Elmentaite et al., 2021).

The unconventional T cells included the innate T type 1 cells and innate T type 3 cells. Innate T type 1 cells were characterised by high expression of *EOMES* (Eomesodermin), which has a role in the regulation of T cell function (Llaó-Cid et al., 2021) and *TBX21* (T-Box Transcription Factor 21), which encodes a transcription factor important in T cell development (Stolarczyk, Lord and Howard, 2014) (Suo et al., 2022). Innate T type 3 cells demonstrated high expression of *RORC* (RAR Related Orphan Receptor C), which encodes a DNA-binding transcription factor that has a role in cellular differentiation (Capone and Volpe, 2020) and *CCR6* (C-C Motif Chemokine Receptor 6), which encodes a membrane protein involved in immune cell recruitment (Yamazaki et al., 2008) (Suo et al., 2022).

144

The lymphoid tissue inducer (LTi) cells differentially expressed *IL17A* (Interleukin 17A), which is known to be highly expressed by LTi cells even upon genetic deletion of its transcriptional regulator *STAT3* (Signal Transducer and Activator of Transcription 3) (Takatori et al., 2009), and has a functional role in host defence and maintenance of tissue integrity (Qian et al., 2010). The lymphoid progenitor population highly expressed *JCHAIN* (Joining Chain OF Multimeric IgA and IgM), which is known to be expressed by multi-lymphoid progenitors that have the potential to differentiate into multiple different functional cell states (Karamitros et al., 2018), and *MEF2C* (Myocyte Enhancer Factor 2C), which has a crucial role in regulating cell fate in multipotent progenitors (Stehling-Sun et al., 2009). The natural killer (NK) cells were characterised by high expression of *GZMH* (Granzyme H), which is known to be constitutively expressed by NK cells and has a role in the cytotoxic arm of the innate immune response (Krzewski and Coligan, 2012), and *KLRD1* (Killer Cell Lectin Like Receptor D1), which encodes an immune receptor involved in self-nonself discrimination and is also known to be specifically expressed by NK cells (Yang et al., 2019) (Appendix N). The regulatory T cells (Tregs) were characterised by high expression of *FOXP3* (Forkhead Box P3), which is a known specific marker gene and a transcriptional regulator important for the inhibitory function of Tregs (Ono et al., 2007), and CTLA4 (Cytotoxic T-lymphocyte Associated Protein 4), which is expressed on Tregs and negatively regulates Treg homeostasis (Tang et al., 2008) (Appendix N).

4.2.7. Differential abundance testing across gestation using Milo

After assigning refined cell annotations to prenatal skin cell states from samples obtained throughout the first and second trimester of gestation, the abundance of each of these cell populations across gestational time was then assessed. This allows the differential cellular composition of prenatal skin over time to be evaluated, prior to identifying which cell states may be co-locating and interacting during specific developmental windows.

The changes in cellular composition across gestation was calculated using a recently developed computational approach called Milo, which performs differential abundance testing on cell neighbourhoods within a k-nearest neighbour (KNN) graph (Dann et al., 2022). This analysis was initially carried using the broad prenatal skin cell annotations,

allowing temporal changes across overall lineages to be evaluated first. The output of the analysis is shown below as a beeswarm plot in Figure 59.
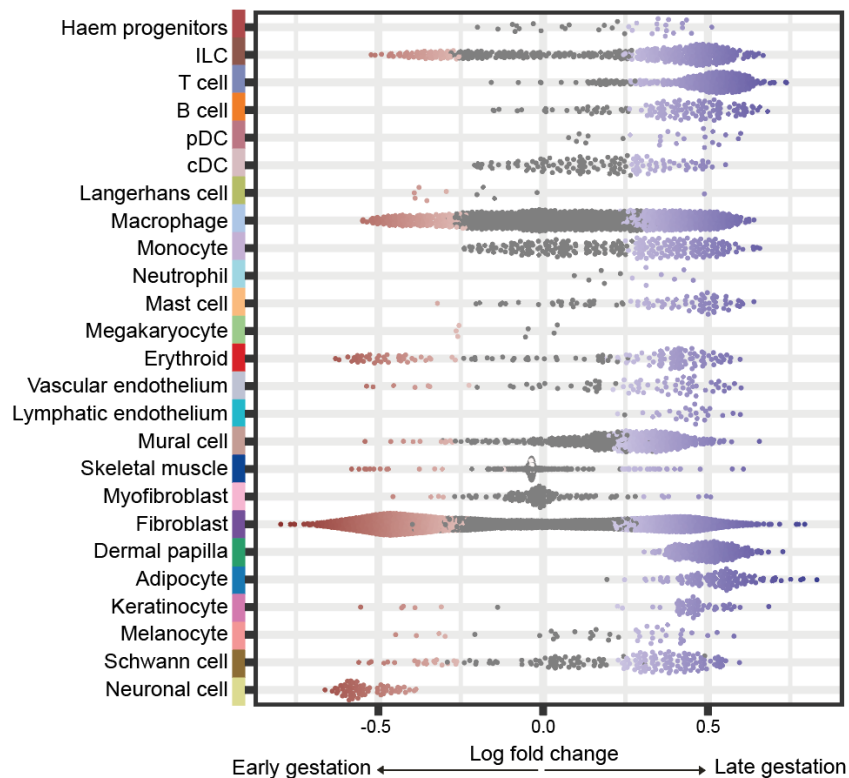
Figure 59: Differential abundance analysis. Beeswarm plot showing log-fold change in abundance between early and late gestation prenatal skin cells in neighbourhoods from different broad cell state clusters. Coloured dots indicate a significant difference in abundance.

The differential abundance analysis showed that the enrichment of some cell populations depended on gestational age. For instance, neuronal cells were observed early in gestation, which are likely to be arising from ectodermal differentiation alongside skin epidermal differentiation; embryonic ectodermal cells that are exposed to Wnt and BMP signalling become epidermis, and in the absence of these two signals become neural cells (Moody et al., 2013). Several prenatal skin cell populations were shown to be abundant in later gestation, including T and B cells (Figure 59), which accompanies prenatal thymus, bone marrow and spleen formation from around 10 PCW (Varga et al., 2011; Jardine et al., 2021; Suo et al., 2022).

Several prenatal skin cell types were present across both early and late gestational stages, including ILCs, macrophages and fibroblasts (Figure 59). Since the Milo analysis was carried out based on neighbourhoods on a KNN graph, the cells abundant across all gestational stages therefore show distinct gene expression profiles between early and late gestation. This either suggests functional evolution of cells during development or dual waves of production.

The differential abundance testing using Milo was next carried out using the refined prenatal skin cell annotations, highlighting the fine-grained differences in cellular composition within each broad cell grouping across gestation. The output of this analysis is shown below in Figure 60.
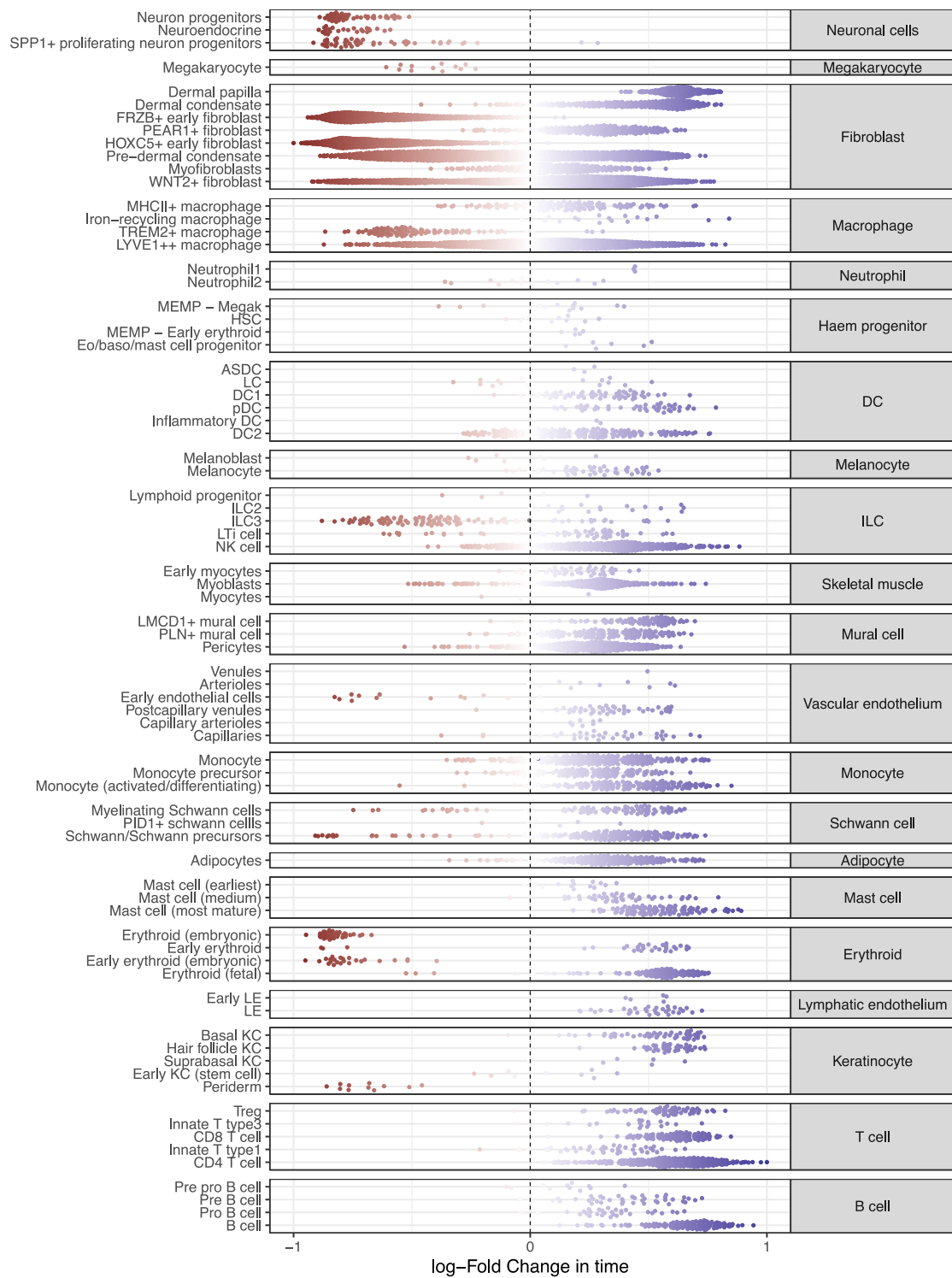
Figure 60: Differential abundance analysis. Beeswarm plot showing log-fold change in abundance between early and late gestation prenatal skin cells in neighbourhoods from different refined cell state clusters. Coloured dots indicate a significant difference in abundance.

Interestingly, within the broad cell groupings above in which cells were enriched across the gestational sampling period during the first and second trimester, some refined subsets showed differential abundance in either early or late gestation. For example, within the fibroblast lineage, the *FRZB*$^+$ early fibroblasts and *HOXC5*$^+$ early fibroblasts were enriched in early gestation, however, the hair follicle associated fibroblasts, such as the dermal condensate cells and dermal papilla cells were enriched in late gestation. This aligns with the onset of hair follicle morphogenesis from the second trimester onwards (de Groot et al., 2021), which will be discussed in more detail in Chapter 4.

Similarly, within the macrophage cell grouping, the *TREM2*$^+$ microglia-like macrophages are enriched earlier in gestation, and the iron-recycling macrophages are enriched later in gestation, suggesting distinct roles for each cell type within specific time periods of prenatal skin morphogenesis. Within the ILC compartment, ILC3s are enriched mostly in early gestation, with some ILC3s present later in gestation, suggesting transcriptional differences and possible functional differences between these cells at different times during skin development, whereas ILC2 are enriched only during late gestation.

Several refined stromal cell states are mostly enriched during late gestation, including subsets of skeletal muscle cells, mural cells, vascular endothelial cells, lymphatic endothelial cells and adipocytes, highlighting the development of more complex skin architecture over gestational time. Across the immune subsets, the majority of cell states are enriched later in gestation, including within the DC, monocyte and mast cell groupings, as well as the T and B cell subsets discussed above, suggesting a diversification of the immune repertoire within the prenatal skin as it develops over gestational time.

4.2.8. Spatial deconvolution reveals prenatal cellular microenvironments

The next line of investigation was to leverage the prenatal skin scRNA-seq data and its refined cell annotations to determine the microanatomical organisation of cells within the developing skin. To do this, a recently published spatial transcriptomic dataset of embryonic lower limb, which included skin region profiling, was used (Zhang et al., 2022).

To map the transcriptomic signatures of the refined single cell states onto the embryonic limb spatial transcriptomic data *in situ*, a computational framework called Cell2location

was used (Kleshchevnikov et al., 2022). This analysis is implemented as a hierarchical Bayesian model and uses the prenatal skin scRNA-seq data to deconvolute mRNA counts in the spatial data. This then allows the estimation of the absolute and relative abundance of each refined prenatal skin cell type at each spatial location in the embryonic limb skin regions.

In order to determine whether specific prenatal skin cell types were more likely to be co-located within microanatomical tissue niches (or microenvironments), a non-negative matrix factorisation (NMF) analysis was then carried out, with co-location indicated by a high proportion of two or more cell types sharing a microenvironment. The output of this analysis is shown below in Figure 61, which reveals that prenatal skin is made up of several microenvironments that comprise of diverse groupings of cell types, including epidermal, dermal, vascular, neuronal and immune cells.
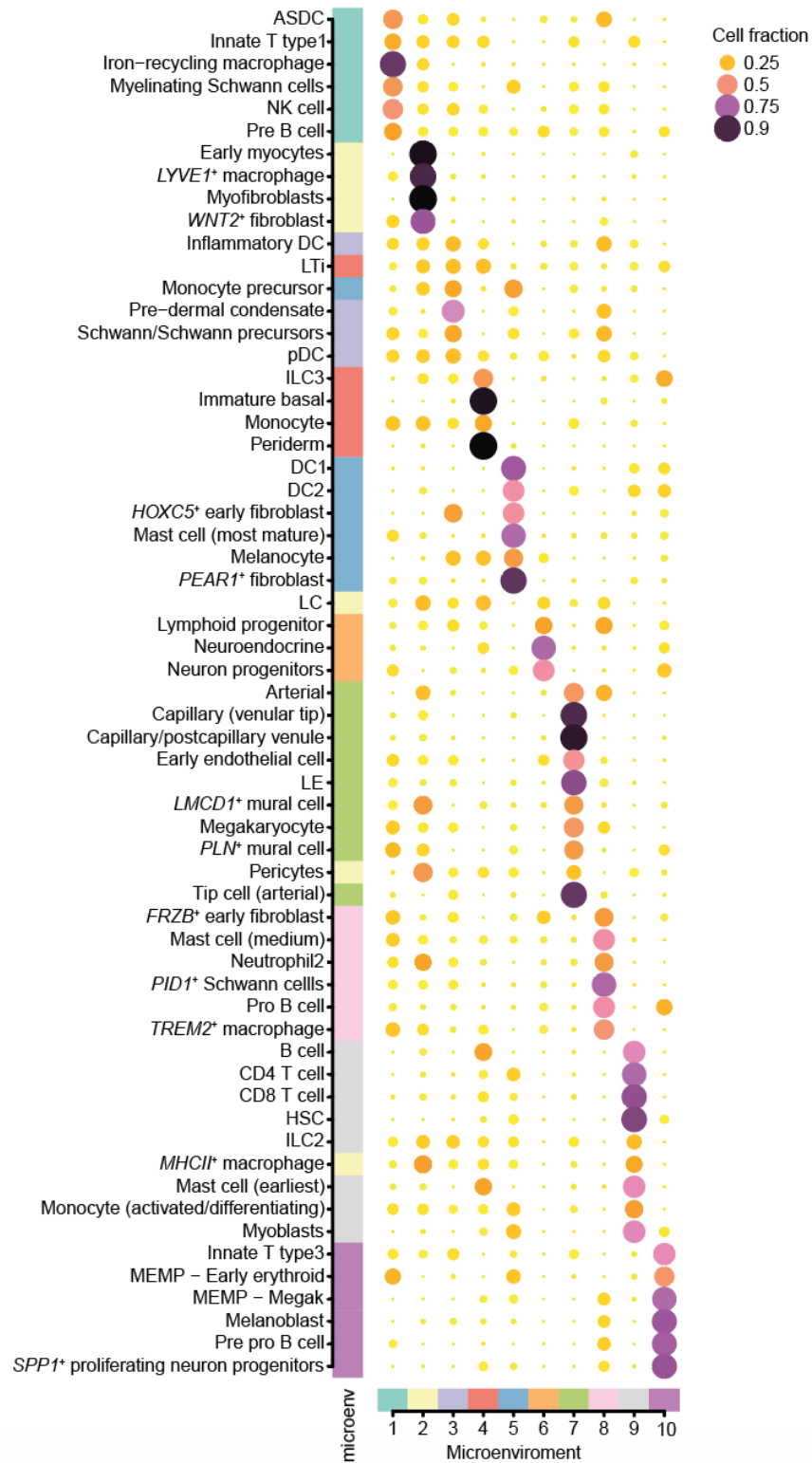
Figure 61: Spatial microenvironments in prenatal skin. The dot plot shows cell type to microenvironment coefficients that are normalised by cell type sums. Cell type to microenvironment assignment is shown by colour.
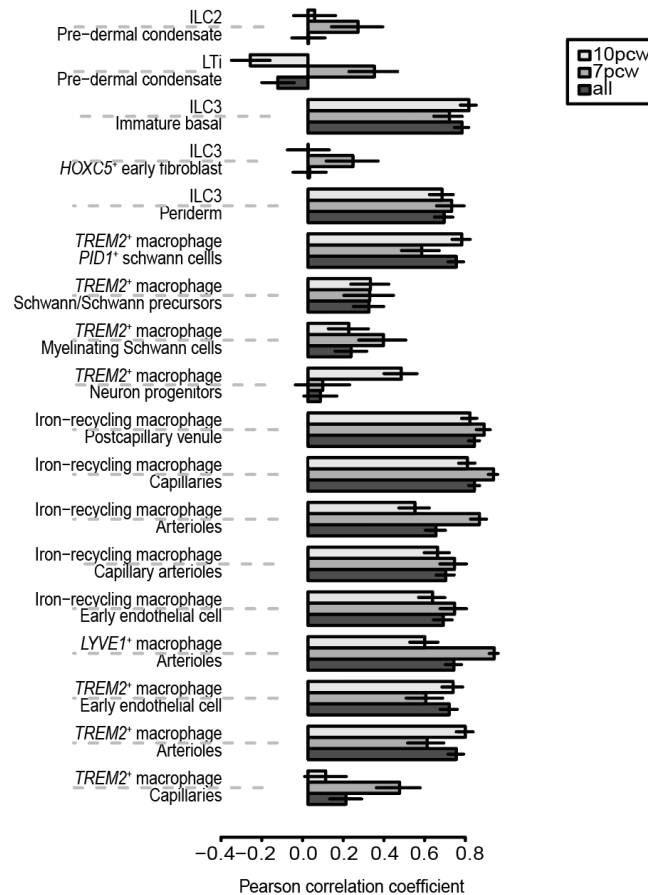
Figure 62: Cell type co-location. Bar plot showing Pearson correlation coefficients for selected cell type pairs, calculated across all spots of Visium samples; either all samples or two time points were used. Error bars show 95% confidence intervals.

Interestingly, the cellular microenvironments uncovered by this analysis relate to specific developmental processes in prenatal skin, including interfollicular epidermal differentiation (where immature epithelial cells, including periderm and immature basal cells, are identified in the same microenvironment as ILC3 cells), the early stages of hair follicle formation (where pre-dermal condensate cells are identified in the same tissue microenvironment as ILC2 cells) and neurovascular genesis (for example, where vascular and neuronal cells show high correlated colocalisation with different macrophage cell subtypes (Figure 61 and Figure 62). This highlights the diverse cell states co-locating during specific time periods across prenatal skin development, and the potential cellular cross talk between non-immune and immune cells supporting skin morphogenesis.

4.2.9. Spatial microenvironments that support prenatal skin development

To investigate the cellular microenvironments in more depth and determine the transcriptional profiles and potential interactions between cells within distinct microanatomical niches, further analysis was carried out. I have focused this analysis on the ILC3 microenvironment referred to in section 4.2.8 due to its relevance to inflammatory skin disease discussed in Chapter 3.

Among the lymphoid cells, ILC3 cells co-localised with $HOXC5^+$ early fibroblasts at 7 PCW (within tissue microenvironment 3) and with immature basal interfollicular epidermal cells at 7 to 10 PCW (within tissue microenvironment 4) (Figure 61 and Figure 62). Their functional role of ILC3s during these time periods was investigated, particularly because differential abundance testing using Milo showed that different neighbourhoods of ILC3 cells were enriched between early and late gestation (Figure 60), suggesting transcriptional heterogeneity across gestational time.

The DEGs between ILC3 cells during specific gestational time periods were therefore calculated (Figure 63), and a gene set enrichment analysis using the derived DEGs was carried out using a reference database to determine the functional categorisation of the upregulated genes (Figure 64).
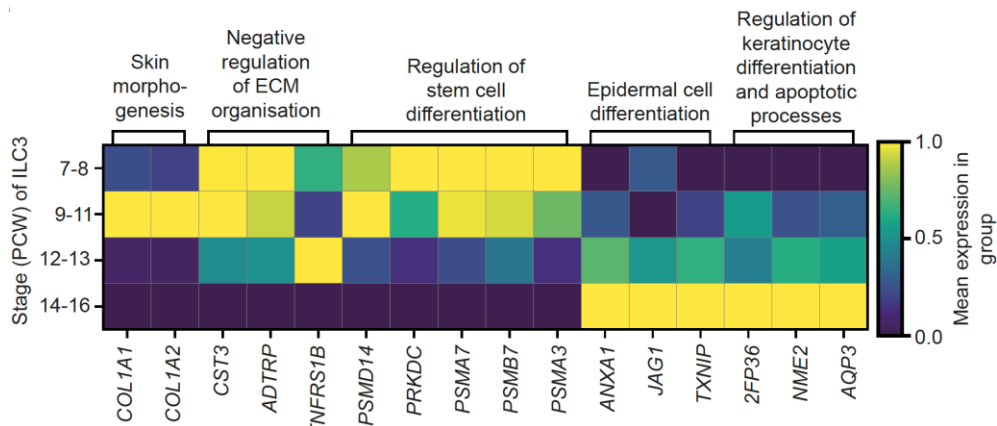
Figure 63: ILC3 DEGs across gestation. Matrix plot showing mean expression (colour) of Milo-generated DEGs by gestational age in the ILC3 population. DEGs are grouped by function.
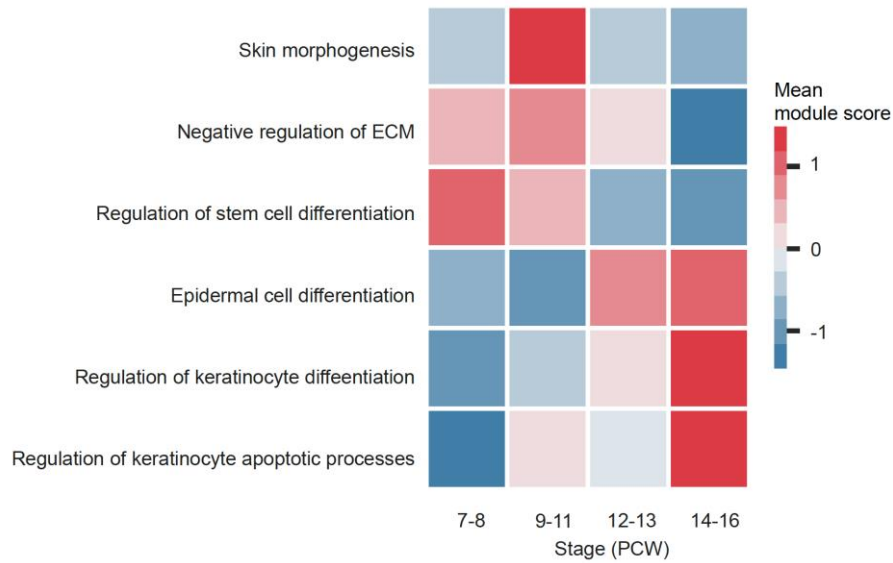
Figure 64: ILC3 functional modules across gestation. Heatmap of z-normalised gene set (GO Biological process) module scores enriched in ILC3 by gestational age (grouped by PCW).

As can be seen from Figure 63 and Figure 64 above, ILC3 cells during early gestation upregulated genes involved in the regulation of ECM organisation (*COL1A1*, *COL1A2*, *CST3*) and stem cell differentiation (*PSMD14*, *PRKDC*, *PSMB7*). In contrast, late gestation ILC3 cells supported epidermal differentiation and upregulated genes implicated in keratinocyte differentiation and apoptosis (*ANXA1*, *JAG1*, *ZFP36*, *AQP3*).

Having characterised the transcriptional profile and the functional roles of ILC3s in prenatal skin across gestational time, a cell-cell interaction analysis was carried out in order to understand how the ILC3 cells are communicating with other cells within the microanatomical environments defined above (section 4.2.8). The cell-cell communication analysis was performed using a computational framework called CellPhoneDB (Efremova et al., 2020), which predicts ligand-receptor interactions between cell states. The output of this analysis is shown as a heatmap below (Figure 65), which revealed that ILC3 cells and *HOXC5*[+] early fibroblasts were interacting through *TGFB1,* which is involved in ECM organisation (Horiguchi, Ota and Rifkin, 2012; Toyoda et al., 2022; Xu et al., 2018), and *TGFBR3*.

154
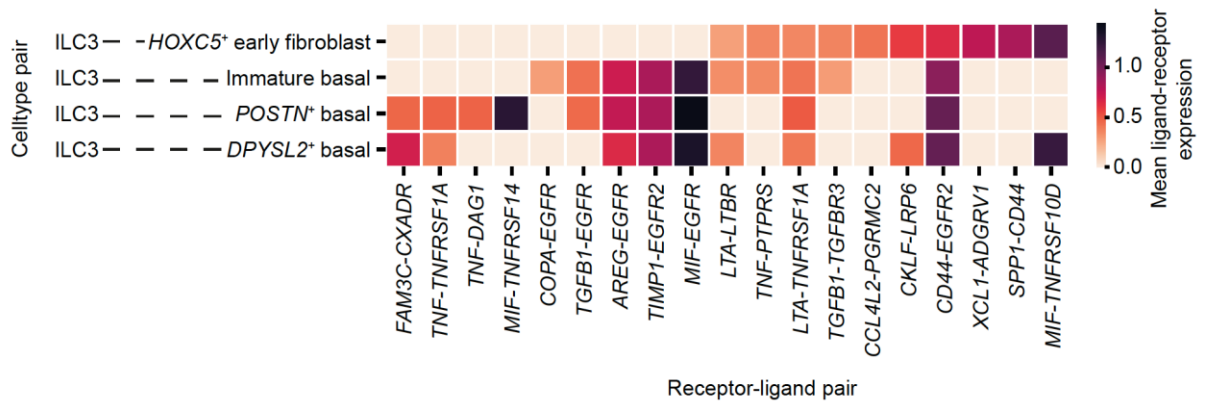
Figure 65: Cell-cell interaction analysis between ILC3s and epidermal cells. Heatmap visualisation of CellphoneDB predicted interactions in the prenatal skin scRNA-seq data. Colour scale represents the mean expression values of each ligand-receptor pair for the corresponding pair of celltypes.

The late gestation ILC3 cells and basal epidermal cells were predicted to interact via *AREG-EGFR*, which are involved in keratinocyte proliferation and differentiation (Chen et al., 2016; Stoll et al., 2016), and *TNF/LTA-TNFRSF1A*, involved in cell survival and apoptosis (Sasaki et al., 2019; Piao et al., 2019; Kumari et al., 2014) (Figure 65). ILC3s are known to contribute to epithelial proliferation and ECM modelling during wound healing in postnatal mouse and human skin, as well as to psoriasis pathogenesis, which is a skin disease characterised by excessive keratinocyte turnover (Ward and Umetsu, 2014; Villanova et al., 2014; Cox, Cruickshank and Saunders, 2021; Li et al., 2016). These findings, therefore, suggest that similar functions are deployed by ILC3 cells during prenatal skin development and morphogenesis.

4.3. Discussion

In this Chapter, the dynamic cellular composition of human prenatal skin during the early stages of skin morphogenesis and *de novo* hair follicle formation during the first and second trimester of gestation is presented. Systematic capture, transcriptomic profiling and annotation of both immune and non-immune skin cell states was carried out, which has not previously been done. Only immune cells in human prenatal skin have been studied by scRNA-seq by other research groups previously (Xu et al., 2021; Bian et al., 2020). This data therefore allows the complex architecture of skin during development to be explored, including how epithelial cell states differentiate and contribute to the hair follicle, the supportive and structural role of the dermal stroma, the development of the endothelial network and the roles that immune cells have during skin morphogenesis *in utero*. The epithelial and stromal lineages and their contribution to hair follicle morphogenesis will be discussed in detail in Chapter 5.

Human prenatal skin was sampled between 7 and 16 PCW, and across this developmental period, 25 broad cell groupings were characterised. Refined cell annotations were then carried out, revealing 82 distinct cell states across various lineages. The differential abundance of the prenatal cell states across the gestational period studied was presented, revealing how the cellular composition dynamically changes over time.

The refined annotation of the epithelial cell states relied on a prior integration of scRNA-seq data between prenatal skin and a hair-bearing skin organoid model (Lee et al., 2020). This was due to the paucity of epithelial cells captured from the prenatal skin samples, which meant that there was insufficient data for the discrete hair follicle cell states to cluster apart without integration from the corresponding cell states in the skin organoid data (Lee et al., 2020). Future work to increase the sample size of the prenatal skin data, collecting more samples from later gestational periods, as well as hair-bearing sites such as scalp skin, would help expand the dataset and allow transcriptomic differences between refined hair follicle cell states to be assessed more easily. Furthermore, other technologies, such as single nuclei RNA sequencing (snRNA-seq), could be employed to process prenatal skin and hair samples, where the mRNA from the nucleus rather than the cytoplasm is profiled (Ding et al., 2020). Support for this approach has recently been shown in a study comparing

scRNA-seq and snRNA-seq data from human adult skin, which showed that snRNA-seq profiling yielded an increased representation of epithelial and stromal cell types compared to scRNA-seq (Eraslan et al., 2022).

By integrating the prenatal skin data with scRNA-seq data from a skin organoid model, the transcriptomic profiles of cell states between *in vivo* and *in vitro* skin and how these change over time was compared. The organoid model lacked immune cells and had a paucity of endothelial cells, however, across the remaining lineages, the cell states present were transcriptionally very similar to both prenatal and adult interfollicular skin cells (Reynolds et al., 2021). Furthermore, the timing of cell differentiation in the organoid model proceeded at an accelerated rate, with organoid cells at 4 weeks of culture aligning closest with prenatal skin at 7 PCW, and after 19 weeks of culture, aligning closest with adult skin cells.

A limitation of this study is that the skin organoid data represented 5 time points of culture: 13 days (~2 weeks), 29 days (~4 weeks), 48 (~7 weeks), 85 (~12 weeks) and 133 days (19 weeks or 4-5 months) (Lee et al., 2020). Therefore, when integrating the data with the prenatal skin data and comparing the transcriptomic profile across time to determine whether physiological *in vivo* differentiation is recapitulated *in vitro*, key molecular changes that may be occurring between data points may have been missed. Future work to build on this would be develop further skin organoids and increase the sampling for scRNA-seq profiling. This work would also lay the foundation for functional experimentation based on the findings from *in vivo* skin analysis, as well as the ability to manipulate the culture protocol to expand the cellular repertoire within the organoid model to better reflect prenatal physiological development, such as addressing the lack of immune cells and endothelial cells in the model. Skin organoid models can also be used for potential disease modelling, which will be discussed in further detail in Chapter 4.

This Chapter also presents a spatial component to the prenatal skin cell atlas, which was carried out using a deconvolution analysis to map the scRNA-seq data to a spatial transcriptomic dataset of embryonic limb that contains regions of prenatal skin (Zhang et al., 2022). This spatial characterisation of human prenatal skin using refined cell states is novel and has not been shown in other studies previously. The analysis revealed distinct microanatomical cellular niches within developing prenatal skin, which is explored to

understand how colocalising cells may be functioning and interacting to support skin development.

The spatial aspect of this prenatal skin cell atlas was, however, limited by the fact that the published spatial transcriptomic study used in the deconvolution analysis only represented samples between 7 and 10 PCW (Zhang et al., 2022). Therefore, the spatial organisation of prenatal cell states later in gestation could not be accurately determined. Furthermore, the spatial data related to embryonic limb, whereas the prenatal samples for scRNA-seq were taken from different anatomical sites, primarily the trunk. Therefore, potential site-specific transcriptomic changes could confound the data or obscure important differences between skin development at different anatomical sites. Future work to generate site-matched spatial transcriptomic data representing the full range of prenatal scRNA-seq data until 16 PCW in the second trimester would be informative in extending our understanding of microanatomical skin cellular organisation beyond 10 PCW.

Overall, this Chapter presented a single cell and spatial atlas of prenatal skin during the first and second trimester of gestation, highlighting broad and refined cell states and how they change in abundance over time, and microanatomical cellular niches present during prenatal skin development. Comparison with a skin organoid model demonstrated the similarities and differences between *in vivo* and *in vitro* skin, which lays the foundational understanding for future functional experimentation, including disease modelling.

Chapter 5: Understanding cellular crosstalk in hair follicle morphogenesis *in vivo* and the support for disease modelling *in vitro*

5.1. Introduction

This Chapter focuses on the human hair follicle and how it develops *de novo* during prenatal life within functional microanatomical cell niches. Previous histological reports detailing the morphological changes that occur during human prenatal skin development have shown that early hair follicle related structures, such as budding of cells from the epidermal basal layer into the underlying mesenchyme, can be visualised at around 14 PCW (Ersch and Stallmach, 1999). The early aggregation of mesenchymal cells beneath the hair bud can also be visualised at this stage, which then develops to form a mature dermal papilla structure by 16 PCW (Breathnach and Smith, 1968; Ersch and Stallmach, 1999). Therefore, this study was designed so that prenatal skin samples that were collected represented the onset of hair follicle morphogenesis and the early development of hair follicle structures.

Though the studies detailing the morphological characteristics of human hair follicles during prenatal life have provided insights into the different stages of hair follicle formation, the cellular and molecular mechanisms underlying this process remain poorly understood. In particular, individual human precursor cell states and the gene expression programs that mediate their differentiation into hair follicle structures that can then be visualised remain unknown. Our understanding of human hair follicle biology has largely been derived from studies on adult hair as it progresses through cycles of anagen (growth), catagen (regression) and telogen (resting) (Alonso and Fuchs, 2006; Hawkshaw et al., 2020). The developmental processes underlying hair follicle formation prenatally have been studied using animal models (Danilenko, Ring and Pierce, 1996; Guo et al., 2020), particularly murine models (Mill et al., 2003; Botchkarev and Fessing, 2005). In recent years, single cell transcriptomics has been employed to study embryonic murine hair development, which has revealed novel insights into transitional cell states during hair

follicle formation and the genes regulating these processes (Sennett et al., 2015; Mok et al., 2019; Saxena, Mok and Rendl, 2019), however, it is unknown as to whether these processes are conserved between species. Therefore, in this study, prenatal skin samples during the gestational period in which hair follicles first develop were profiled using scRNA-seq and analysed to characterise the cellular states, differentiation trajectories and interactions within microanatomical niches involved in human hair follicle morphogenesis. The analyses were then leveraged to gain novel insights into hair follicle related diseases and, using the integrated skin organoid dataset (Lee et al., 2020) discussed in the previous Chapter, explore whether these diseases can be modelled *in vitro*.

5.2. Results

5.2.1. Prenatal skin morphological changes during hair follicle neogenesis

To characterise the prenatal skin and hair follicle morphological features in the samples that were collected, and confirm the formation of hair follicles that are known to develop from around 14 PCW in humans (Ersch and Stallmach, 1999), haematoxylin and eosin (H&E) staining and multipex RNA *in situ* hybridisation using RNAScope was carried out.

5.2.1.1. Histological analysis

Representative prenatal skin samples from different developmental stages were prepared for histological examination using H&E staining, including a sample at 11 PCW prior to the onset of hair follicle formation, a sample at 14 PCW as hair follicles have begun to develop and a sample at 16 PCW to assess the morphology of the skin and hair follicles later in gestation (Figure 66).
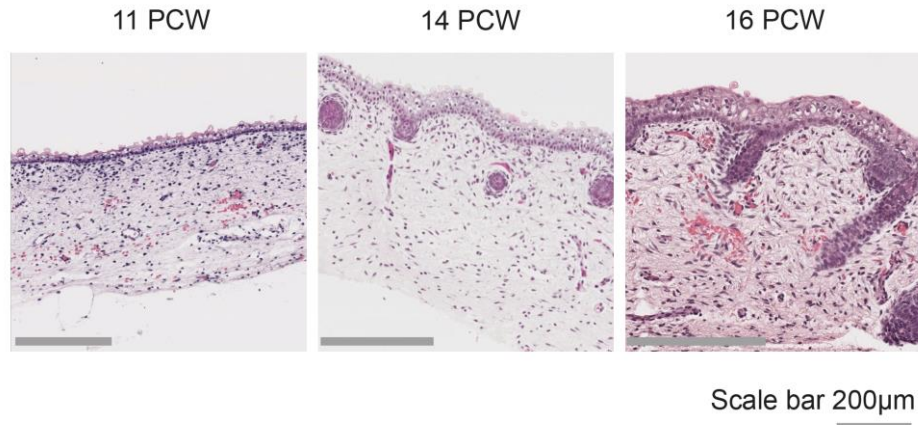
Figure 66: Prenatal histology. Representative H&E stained tissue sections showing different developmental stages of prenatal skin and hair follicle morphogenesis.

The histological images of prenatal skin show the morphological changes underlying skin and hair follicle development between 11 and 16 PCW (Figure 66). Prenatal skin at 11 PCW comprises of a layer of epidermal cells overlying the dermal stroma without skin appendages. At 14 PCW, budding of basal cells (hair placodes/germs) are observed, as well as elongated hair follicles (hair pegs) that are visualised as circular structures in cross

section of the hair follicles. The representative histological image at 16 PCW shows longitudinally sectioned, elongated hair follicle structures, as well as a stratified multi-layered epidermis.

5.2.1.2. Multiplex RNA in situ hybridisation

The developing prenatal hair follicles in later gestation were also visualised with multiplex RNA *in situ* hybridisation using RNAScope in order to spatially characterise the cellular components that make up the prenatal hair follicle and how they are microanatomically organised in relation to one another.

Differential gene expression analysis was used to help determine which specific probes to use in order to visualise each cell type. The hair follicle cell types visualised were the outer root sheath (using a human probe against *SLC26A7*), the matrix (*SHH*) and the dermal papilla (*NDP*). These cells were selected because the outer root sheath is known to encapsulate the hair follicle in prenatal skin, helping to visualise where the outer limit of the hair follicles lie within the prenatal skin samples. Furthermore, the dermal papilla cells are known to lie directly adjacent to the matrix within the hair bulb to regulate hair shaft formation (Sennett and Rendl, 2012), and therefore, confirming their cellular microanatomical proximity within the samples was carried out. Regulatory T cells (Tregs) are known to preferentially localise around the hair follicle in late second trimester (23 PCW) (Dhariwala et al., 2020) and postnatal skin (Sanchez Rodriguez et al., 2014), with a role in facilitating hair follicle stem cell differentiation (Ali et al., 2017). Therefore, their presence earlier in gestation at 15 PCW was investigated using a probe against *FOXP3*.
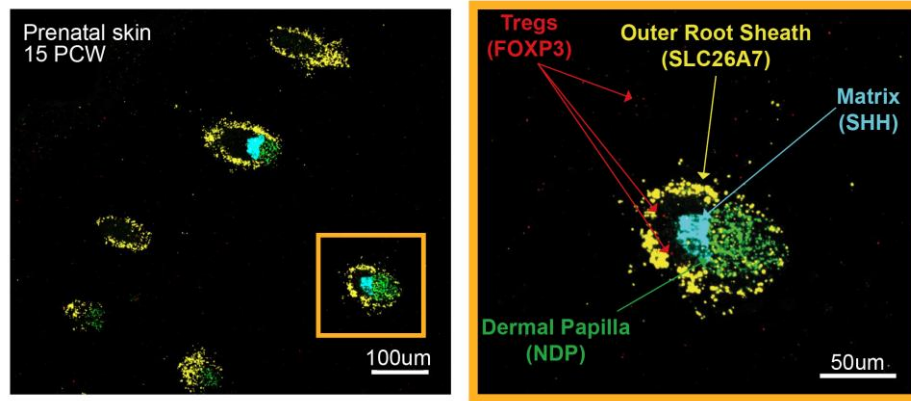
Figure 67: Hair follicle multiplex RNA in situ hybridisation. RNAScope images of prenatal skin from a representative 15 PCW sample, demonstrating outer root sheath (SLC26A7; yellow), matrix (SHH; cyan), dermal papilla (NDP; green) and Tregs (FOXP3; red) cell types. The large-area image is shown on the left (scale bar = 100μm) and the magnified image on the right (scale bar = 50μm).

The low magnification RNAScope image of prenatal skin at 15 PCW showed multiple hair follicles within the sample, demonstrated by the outer root sheath cells that line the hair follicle structures. Some hair follicles were sectioned at a plane in which the hair bulb containing the matrix and dermal papilla cells were visualised, confirming their close proximity at this stage of gestation. Furthermore, *FOXP3*[+] Tregs were identified primarily within and around hair follicles, as well as the interfollicular epidermis, from as early as 15 PCW (Figure 67).

5.2.2. Prenatal epithelial and stromal cell states related to hair follicle formation

In order to delineate early hair follicle development in human prenatal skin, the epithelial and stromal cells that contribute to this process were explored in further depth. As described in the previous Chapter in section 4.2.6, the epithelial and stromal cells were clustered to a high resolution and given refined annotations; the hair follicle cell states were annotated following integration of the prenatal skin data with scRNA-seq data from a hair-bearing skin organoid model (Lee et al., 2020).

Of the epithelial cell states, those that relate to hair follicle formation, including the placode/matrix, outer root sheath, companion layer, inner root sheath and cuticle/cortex,

were identified beyond 12 PCW during the second trimester of gestation (Figure 54). This is in keeping with the findings from the prenatal tissue histology and multiplex RNA *in situ* hybridisation using RNAScope shown above, where hair follicle structures and cell states become apparent within later gestation prenatal skin samples.

The following schematic provides context to the microanatomical organisation of epithelial and stromal cell states that were annotated in the prenatal skin dataset involved in hair follicle formation, which is based on prior knowledge of skin and hair follicle morphological structure during development (Mesler et al., 2017; Schneider, Schmidt-Ullrich and Paus, 2009).
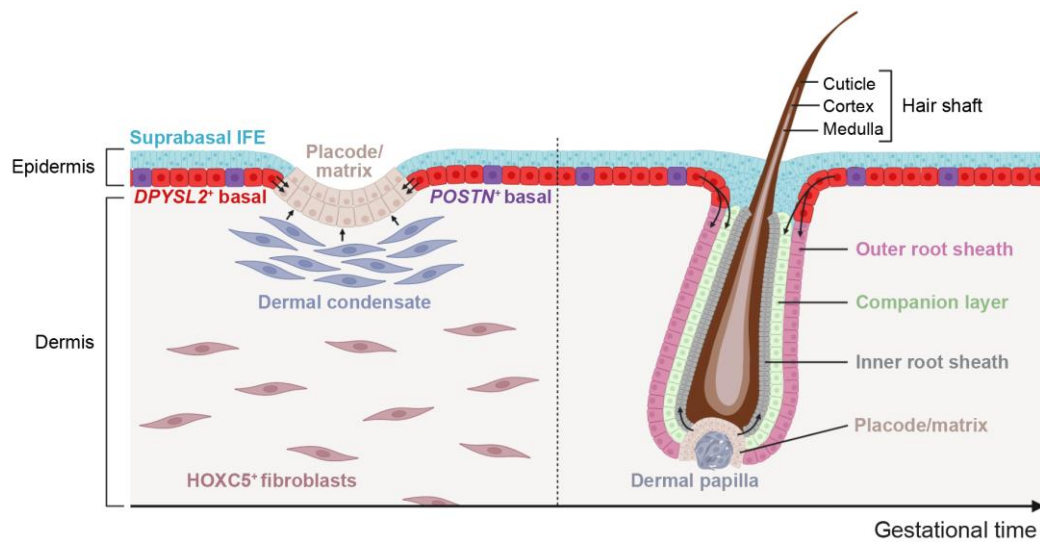
Figure 68: Hair follicle development overview. Schematic to show early (left) and late (right) stages of hair follicle formation, and the epithelial and stromal cell states involved. IFE = interfollicular epidermis. Image created with BioRender.com by Chloe Admane.

The skin comprises of the superficial epidermal layer and the dermis, which lies deep to the epidermis separated by a basement membrane (Figure 68). Within the prenatal single cell dataset, *DPYSL2*+ basal cells made up most of the basal cell states. The *DPYSL2*+ basal cells were first identified between 9-11 PCW, however they were small in number at that stage of gestation. Between 12-16 PCW, the *DPYSL2*+ basal cells made up the majority of epithelial cell states in the dataset (Figure 54), hence why they are represented as the predominant cell type in the basal layer of the epidermis in the schematic above (Figure 68). The *POSTN*+ basal cells made up smaller proportion of basal cells that were

predominantly seen from 9 PCW onwards (Figure 54), with their expression pattern closely resembling a cluster of basal stem cells in human newborn skin scRNA-seq data (Wang et al., 2020b), as discussed in section 4.2.6.1. The *POSTN*$^+$ basal cells are therefore represented within the basal cell layer, interspersed between the *DPYSL2*$^+$ basal cells. Suprabasal interfollicular epidermal (IFE) cells represent differentiated basal cells and lie directly superficial to the basal cell layer (Blanpain and Fuchs, 2009), as shown in Figure 68.

The hair placode represents areas of epithelial thickening at the onset of hair follicle formation, which extend into the underlying dermis to initially form the hair germ, then extends further to form the hair peg (Mesler et al., 2017). The matrix cells were annotated within the same cluster as the placode cells due to the high expression of *SHH* (Sonic Hedgehog Signalling Molecule), which is known to be highly expressed in placode and matrix cells (Michno et al., 2003; Schmidt-Ullrich et al., 2006) (Appendix N). During the early stages of hair follicle formation, the placode is shown as epidermal thickening in the schematic above, and during later stages, the matrix cells are shown at the base of the hair follicle (Figure 68). Matrix cells then persist within the hair follicle, including throughout postnatal life, and they terminally differentiate into cells that form the concentric layers of the hair shaft (Mesler et al., 2017). The epithelial cells that surround the hair shaft in a concentric fashion include the inner root sheath, the companion layer and the outer root sheath (Figure 68).

Within the dermal compartment of the prenatal skin dataset, *HOXC5*$^+$ fibroblasts were observed during the early gestational period in which hair follicles are known to begin forming (Figure 60), which is represented by their presence in the dermis in the schematic in Figure 68. Hair follicle specialised fibroblasts were also observed in the dataset, which play a key role in the mesenchymal to epithelial signals known to initiate hair follicle formation (Rishikaysh et al., 2014; Rendl, Lewis and Fuchs, 2005; Millar, 2002). The dermal condensate cells are fibroblasts that aggregate together at the onset of hair follicle morphogenesis. The dermal condensate is a precursor to the mesenchymal component of the hair follicle, known as the dermal papilla, which regulates hair cycling throughout life (Morgan, 2014).

Having contextualised the morphology and cellular components of the developing hair follicle and the brief roles of the epithelial and dermal cell types involved, analysis of the hair follicle related cells in the prenatal scRNA-seq data is presented next, with the aim of uncovering molecular regulatory mechanisms and cellular interactions involved in hair follicle neogenesis.

5.2.3. Comparison between prenatal and adult hair follicle cell states

De novo human hair follicle formation is a process that uniquely occurs in prenatal skin across the gestational period that was sampled in this study. However, the precise mechanisms underlying hair follicle development have largely been inferred from murine studies (de Groot et al., 2021) or from adult human studies where established hair follicles cycle through distinct phases of hair growth (anagen), regression (catagen) and resting (telogen) (Randall and Botchkareva, 2009; Hsu, Pasolli and Fuchs, 2011).

The hair follicle related cell states present in prenatal skin and their transcriptional profile was therefore compared to adult hair follicle cell states in the anagen growth phase, to provide further understanding as to how prenatal hair morphogenesis differs from hair growth in postnatal life. In a recent scRNA-seq study, human scalp micrografts collected for hair transplantation were sampled, each of which comprised of several hair follicles and hair bulbs, as well as some surrounding interfollicular tissue, such as the epidermis, dermis and sebaceous glands (Takahashi et al., 2020). H&E staining was used to confirm that the hair follicles were in the anagen growth phase, as represented by the presence of the hair bulb; in the telogen resting hair follicle phase, the hair bulb is not present. The tissue was then dissociated into single cells for scRNA-seq (Takahashi et al., 2020). This adult hair follicle dataset was then integrated with the hair follicle related epithelial cell states in the combined prenatal / hair-bearing skin organoid (Lee et al., 2020) data, which was visualised using a UMAP embedding:
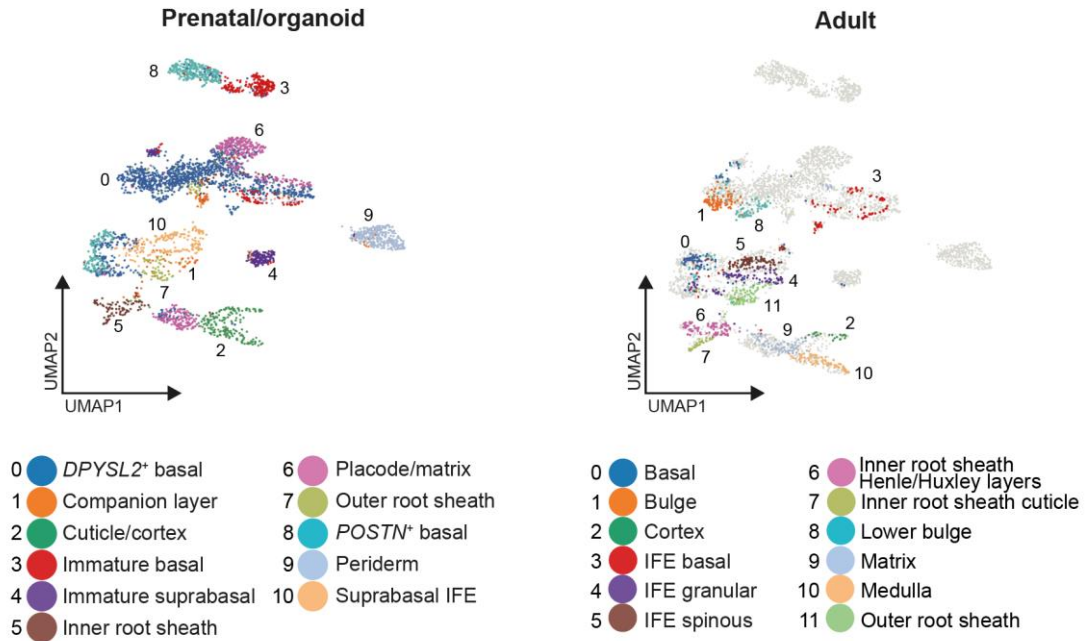
Figure 69: UMAP visualisations showing integrated data from prenatal/skin organoid (left) and adult hair follicles (Takahashi et al., 2020) (right), coloured cell types.

By plotting cell states from the combined prenatal and organoid data alongside cells states from the adult hair follicle data and showing annotations of the cell states (Figure 69), this allows the assessment of transcriptionally similar cell states between prenatal/organoid and adult to be determined, as well as assess whether cell states are present or absent during development or adulthood. The UMAP embedding showed that several cell states present during development were absent in adult hair follicles, including periderm, immature basal, immature suprabasal and *POSTN*⁺ basal cells. Periderm cells are known to be a transient outer epithelial barrier in prenatal life that is shed into the amniotic fluid during the second trimester of gestation. Within the prenatal dataset, immature basal and immature suprabasal cells were only observed in first trimester samples between 7 and 11 PCW and were not observed beyond 12 PCW, as was discussed in the previous Chapter in section 4.2.6.1 and shown in Figure 54. This finding is in accordance with the absence of an equivalent cell state in the adult hair follicle dataset (Figure 69). The *POSTN*⁺ basal cells were also not observed in the adult dataset, despite their presence beyond 12 PCW in the prenatal data (Figure 54). As discussed in section 4.2.6.1 the expression pattern of *POSTN*⁺ basal cells closely resembles that of basal stem cells in human newborn skin (Wang et al., 2020b). Interestingly, this integration analysis between prenatal/organoid and adult cell states therefore indicates that *POSTN*⁺ basal cells do not persist into adult life.

167

The placode/matrix cells in the prenatal/organoid data pulled apart into two transcriptionally distinct clusters following integration with the adult data (Figure 69). One of the two prenatal/organoid clusters transcriptionally aligned with adult hair follicle matrix cells and was observed in close proximity to the cuticle/cortex cells in the prenatal/organoid embedding and to the medulla and cortex cells in the adult hair follicle data embedding. This is in keeping with our understanding of hair follicle biology, where matrix cells persist within the hair follicle throughout life and terminally differentiate into the cells that form the concentric layers of the hair shaft, including the central medulla in large and thick hairs, the cortex and the outer cuticle layer (Mesler et al., 2017). This explains the proximity on the UMAP embedding, and therefore the transcriptional similarity, between the matrix, medulla, cortex and cuticle cell states across development and adulthood (Figure 69). The second cluster of placode/matrix cells on the prenatal/organoid embedding space therefore represents placode cells, which is in keeping with their proximity on the UMAP graph to basal cells; epidermal placodes are focal thickenings of the basal cell layer that are observed in the early stages of hair follicle formation, which are absent in adult established hair follicles (Figure 69).

Interestingly, several cell states closely aligned transcriptionally between the prenatal/organoid and adult hair follicle datasets. For example, the prenatal/organoid suprabasal IFE cells and the adult IFE granular and IFE spinous cells transcriptionally align with each other. The adult IFE granular and IFE spinous cells are the only suprabasal cells annotated in the adult dataset. To contextualise these adult suprabasal annotations, a fully stratified epidermis contains several cell layers that represent the differentiation trajectory of keratinocytes: the basal, spinous, granular and cornified layers, listed from deep to superficial (Moreci and Lechler, 2020). The cornified epidermal layer is made up of thickened, flattened dead cells, hence why they were not captured by scRNA-seq (Takahashi et al., 2020).

Across the prenatal/organoid and adult datasets, the outer and inner root sheath cells also aligned closely, with the adult inner root sheath cells being annotated to represent the several layers present in a developed hair follicle unit. The inner root sheath comprises of the Henle, Huxley and inner root sheath cuticle layers, listed from outer to inner (Schneider, Schmidt-Ullrich and Paus, 2009). Furthermore, the hair shaft cells aligned transcriptionally

across development and adulthood, with the prenatal/organoid cuticle/cortex cells observed in the equivalent UMAP embedding space as the adult medulla and cortex cells. Medulla cells are only present in thicker, larger hair shafts, which is in keeping with the absence of medulla cells in prenatal skin, where small, thin hair are developing at the onset of hair follicle neogenesis *in utero*.

Overall, this integration analysis highlighted development-specific skin and hair follicle cell states, including interfollicular cells and hair placode cells. This highlights the importance of using prenatal skin samples to delineate the molecular mechanisms underlying skin and hair follicle morphogenesis, which cannot be inferred from adult skin and hair follicle data. For the cell states in which equivalent annotations were across development and adulthood, the scRNA-seq data was then used to calculate the transcriptional similarity between cells within the annotated clusters, as shown in the heatmap below:
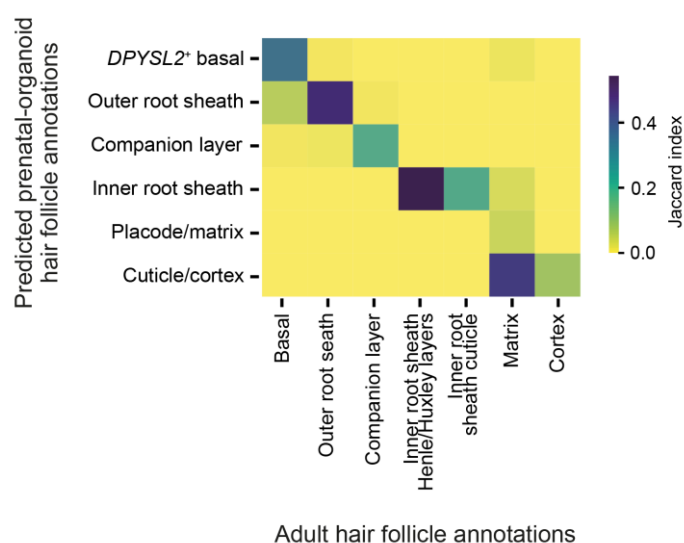
Figure 70: Transcriptional similarity between prenatal/organoid and adult hair follicle cells. Heatmap showing the average proportion (normalised by Jaccard index) of predicted adult epidermal and hair cell states (x-axis) (Takahashi et al., 2020) assigned to integrated prenatal skin and skin organoid epidermal and hair cell states (y-axis).

This comparative analysis between prenatal/organoid and adult basal and hair follicle cell states showed close transcriptional alignment of most cell states, except for placode/matrix cells (Figure 70). To investigate this further, a differential gene expression analysis was

therefore carried out between the combined prenatal/organoid placode/matrix cells and the adult matrix cells. The top DEGs between the annotated cell states in each dataset are shown in Figure 71 below.
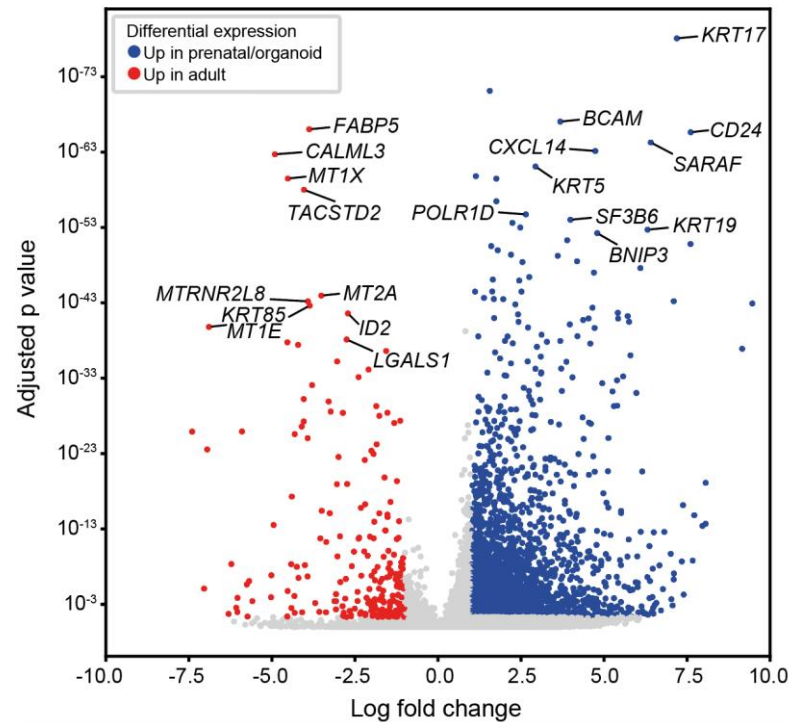
Figure 71: DEGs between prenatal/organoid and adult matrix cells (Takahashi et al., 2020), shown as a volcano plot.

This analysis highlighted several highly expressed genes in adult matrix cells, including *KRT85* (keratin 85), which encodes a protein that forms a structural component of the hair and nails, reflecting the role of matrix cells in hair shaft formation (Mesler et al., 2017). Mutations in *KRT85* are known to cause hair and nail ectodermal dysplasia, which is a congenital disorder characterised by hypotrichosis (where little to no hair growth occurs) and nail dystrophy (which includes abnormalities in the shape, texture and growth of nails) (Shimomura et al., 2010). The adult matrix cells also highly express *LGALS1* (Galectin 1), which is known to play an important role in cell proliferation and differentiation, which is also in keeping with matrix cells self-renewing and terminally differentiating into cells that form the concentric layers of the hair shaft in the anagen growth phase (Chen et al., 2020a). Furthermore, in murine models, LGALS1 was identified to be a key secreted protein that is enriched in embryonic skin and is essential and sufficient, alongside apolipoprotein-A1 (APOA1) and lumican (LUM), to induce new hair follicles by activating Wnt and IGF

(insulin-like growth factor) signalling in hair follicle related cells (Fan et al., 2018). The adult matrix keratinocyte cells in anagen phase also differentially expressed *FABP5* (Fatty acid binding protein 5), which encodes an intracellular carrier protein that delivers retinoic acid to the nuclear receptor PPAR-δ (peroxisome proliferator-activated receptor-delta, which promotes keratinocyte proliferation and survival (Di-Po1 et al., 2003; Morgan, Kannan-Thulasiraman and Noy, 2010). Furthermore, *FABP5* has been shown to be a prominent marker of anagen follicle bulbs, which has a dynamic regulatory role during adult hair follicle cycling, as evidenced by blockage of Wnt/β-catenin signalling leading to an upregulation of FABP5 expression (Collins and Watt, 2008).

When compared with adult matrix cells, prenatal and skin organoid placode/matrix cells showed differential expression of genes involved in immune recruitment and signalling (Figure 71). For example, *CXCL14* (C-X-C Motif Chemokine Ligand 14) was highly expressed, which encodes a cytokine that displays chemotactic activity for immune cells, including monocytes and dendritic cells (Meuter et al., 2007). Depletion of CXCL14 has also been shown to inhibit the accumulation of regulatory T cells (Tregs), whereas supplementation had the opposite effect in a murine model, with CXCL14 promoting interleukin-2 (IL-2)-induced Treg differentiation (Lee et al., 2017). The prenatal and skin organoid placode/matrix cells also highly expressed *CD24* (Cluster of differentiation 24), which encodes a cell signalling molecule that modulates immune responses and has a pivotal role in regulating cell differentiation (Kim et al., 2014b; Ayre et al., 2016). *CD24* has also been shown to be highly expressed in the rapidly dividing bulge cells of the murine hair follicle, which has the growth potential expected from stem cells (Magnaldo and Barrandon, 1996), mirroring the proliferating and differentiation potential of matrix cells during development. The expression of immune recruiting and signalling genes in the developmental placode/matrix cell states compared to the equivalent cells in the adult hair follicle signifies a potential role of immune cells during hair follicle morphogenesis. This is supported by the results of the multiplex RNA *in situ* hybridisation experiment presented in section 5.2.1.2 where *FOXP3*[+] Tregs were shown to localise within and around the developing hair follicle at 15 PCW (Figure 67).

The placode/matrix cells in prenatal skin and skin organoid also differentially expressed *KRT17* (Keratin 17) and *KRT19* (Keratin 19) (Figure 71). *KRT17* is mainly expressed

within appendageal structures, including the hair follicle, nail bed and sebaceous glands (McGowan and Coulombe, 2000). Interestingly, keratin 17 also has a role in hair follicle cycling, as demonstrated by deficits observed in keratin 17 null mice. Alopecia develops in the first week postnatally in keratin 17 null mice, correlating with hair shaft fragility and untimely TNF-α dependent apoptosis in the hair bulb (Tong and Coulombe, 2006), which is where matrix cells are located. *KRT19* is also expressed by placode/matrix cells during development (Figure 71), which is known to be expressed in highly proliferative postnatal human hair follicle bulge stem cells (Rittie et al. 2009), mirroring the biological behaviour across these cell states.

This analysis highlights the transcriptional and biological differences between placode and matrix cells during development and in adulthood during the anagen hair follicle growth phase. The findings suggest a role for immune cell recruitment during development and demonstrates parallels between our current understanding of hair follicle biology in the literature and the possible roles that highly expressed genes in development and adulthood may have.

5.2.4. Trajectory inference of epithelial cell types

To further understand the mechanisms underlying hair follicle neogenesis, the differentiation trajectory between prenatal skin epithelial cell states was investigated next. Fate-mapping and live imaging experiments using murine embryonic models have shown that the invaginating placode differentiates into the matrix, which in turn differentiates into the inner layers of the hair follicle, including the inner root sheath, cuticle and cortex (Saxena, Mok and Rendl, 2019; Morita et al., 2021; Mesler et al., 2017). Previous murine studies have reported that the companion layer arises from the hair matrix (Mesler et al., 2017; Gu and Coulombe, 2007), whereas the outer root sheath has been thought to extend directly from basal epidermal cells (Hardy, 1992; Sequeira and Nicolas, 2012).

In recent years, single cell transcriptomic data has been leveraged to allow the calculation of differentiation trajectories between a group of cell states (Tritschler et al., 2019). A widely-used computational package to carry out trajectory analyses is called Monocle3, which orders cells on a psuedotemporal scale by detecting gradual transcriptional changes

along a cellular differentiation pathway (Trapnell et al., 2014). To date, trajectory inference using scRNA-seq data on human prenatal hair follicle cell states during the gestational period in which hair follicle neogenesis occurs has not been previously reported; previous scRNA-seq studies on human prenatal skin have been restricted to immune cell populations only (McFaline-Figueroa, Trapnell and Cuperus, 2020), as discussed section 4.1. Furthermore, although murine models have provided valuable insights into cell fates during hair follicle morphogenesis, it remains unclear at present how precisely these models recapitulate the cellular and molecular processes underpinning hair follicle development in humans.

An inferred trajectory and pseudotime analysis that included the prenatal epithelial hair follicle cell states was therefore carried out using Monocle3 (McFaline-Figueroa, Trapnell and Cuperus, 2020), which revealed the following differentiation trajectories:
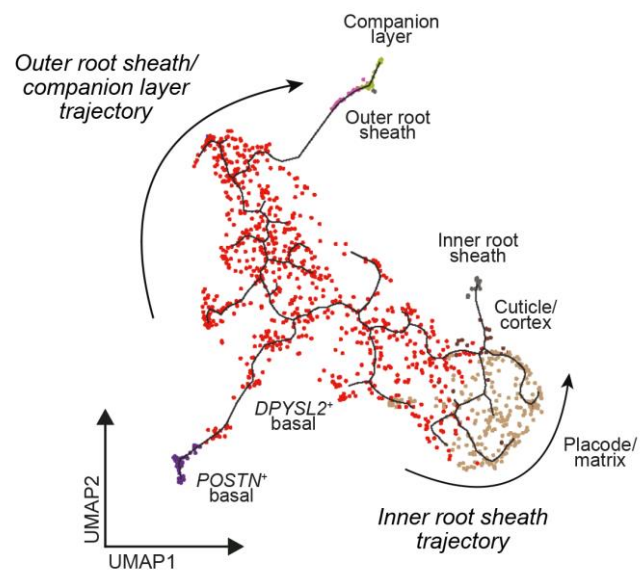
Figure 72: Inferred trajectory of hair follicle epithelial cell differentiation. Pseudotime trajectory analysis using Monocle3 shows epithelial cell states differentiating along outer root sheath / companion layer and inner root sheath trajectories.

This inferred trajectory analysis predicted the differentiation of *POSTN*$^+$ basal cells into *DPYSL2*$^+$ basal cells, prior to bifurcation into two trajectories. Along one trajectory, the *DPYSL2*$^+$ basal cells differentiate into placode/matrix cells then the cuticle/cortex and inner root sheath cells, which will be referred to as the inner root sheath trajectory. Along the

other trajectory, the *DPYSL2⁺* basal cells differentiate into the outer root sheath cells together with the companion layer cells, which will be referred to as the outer root sheath / companion layer trajectory.

In addition to the analysis output being visualised based on cell states along each trajectory above (Figure 72), the graph embedding was visualised based on the Monocle3 derived pseudotime and gestational age (PCW):
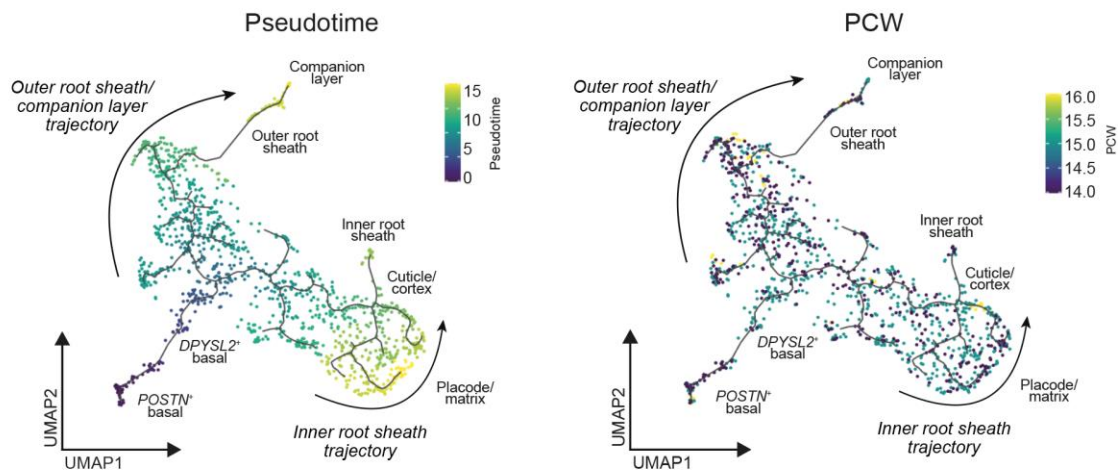
Figure 73: Inferred trajectory of hair follicle epithelial cell differentiation, coloured by pseudotime (left) and gestational age (PCW) (right).

These UMAP visualisations highlight that prenatal epithelial hair follicle cell states that were ordered on a pseudotime scale based on their transcriptional changes along a differentiation trajectory were obtained from prenatal samples of varied gestational ages between 14 and 16 PCW, and that pseudotime does not directly correlate with gestational age (Figure 73). This lends further support to power of single cell transcriptomic analysis to detect subtle molecular changes between individual cell states, allowing differentiation trajectories to be inferred, and that relying on morphological observations of hair follicle development over time alone has limitations.

The findings of the pseudotime trajectory are largely in agreement with what has previously been reported in the literature about skin and hair follicle morphogenesis. *POSTN⁺* basal cells are known to represent a basal cell stem cell population present until at least newborn life, and which are not present in adult skin (Wang et al., 2020b). Furthermore, some basal

cell populations in the epidermis are known to differentiate into the placode at sites where hair follicles eventually develop, whereas other basal populations form either the interfollicular epidermis or outer root sheath surrounding the hair follicle (Schneider, Schmidt-Ullrich and Paus, 2009; Sequeira and Nicolas, 2012). However, several previous studies have suggested that the companion layer arises alongside the inner root sheath from the hair matrix (Mesler et al., 2017; Gu and Coulombe, 2007), as discussed earlier in this section above. This pseudotime trajectory analysis on prenatal human skin showed that the companion layer arises along the same differentiation pathway as the outer root sheath from $DPYSL2^+$ basal cells, independent of the hair matrix (Figure 73). In support of the findings of this analysis, a recent single cell analysis of murine hair follicles also showed greater transcriptional similarity of companion layer cells to outer root sheath cells rather than the inner root sheath cells, which was unexpected based on previous reports in the literature (Joost et al., 2020).

To understand the molecular mediators underlying each differentiation trajectory, genes that were highly expressed along the outer root sheath / companion layer trajectory and the inner root sheath trajectory were derived. Expression of the derived genes across the pseudotime trajectory embedding is shown in Figure 74 below, and expression across distinct epithelial cell states within the trajectories is shown in Figure 75 below.
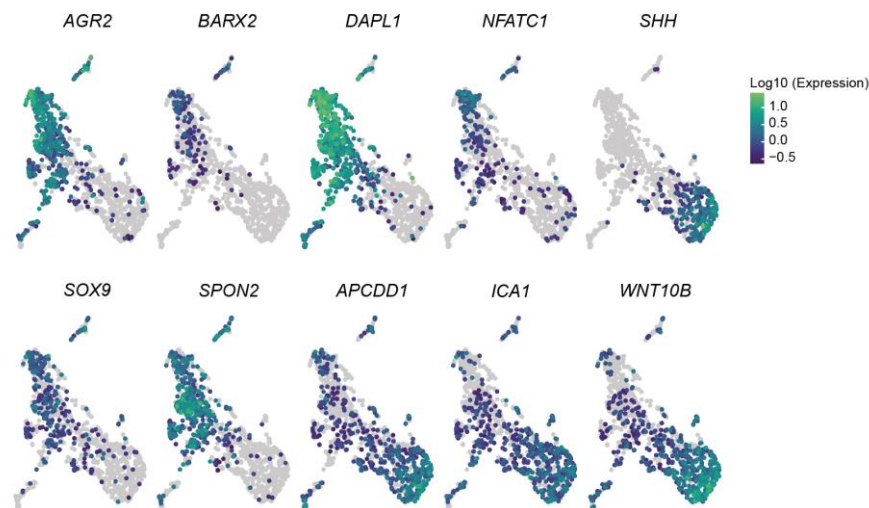
Figure 74: Inferred trajectory of hair follicle epithelial cell differentiation, coloured by gene expression (log-transformed).
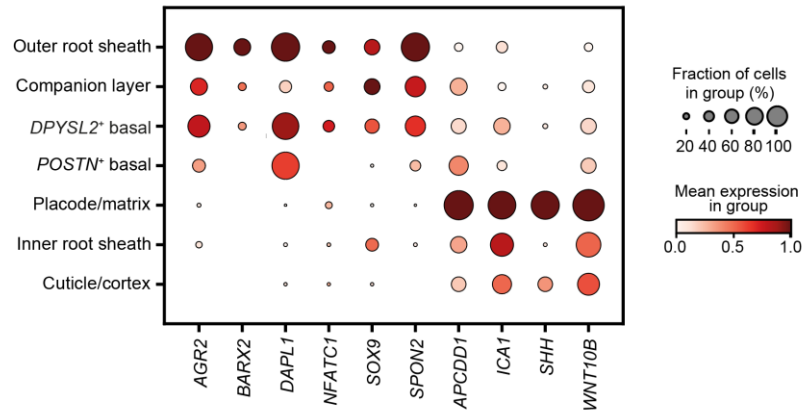
Figure 75: Genes differentially expressed along the outer root sheath / companion layer and inner root sheath trajectories. Dot plot shows expression of variance-scaled, mean expression (dot colour) and percent of expressing cells (dot size).

Along the outer root sheath / companion layer trajectory, genes that are previously reported to be involved in outer root sheath differentiation were identified, including *BARX2* and *SOX9 (*Figure 74 and Figure 75). *BARX2* (BARX Homeobox 2) encodes a transcription factor that has been shown to be expressed in the outer root sheath of adult animal hair follicles, though not in the inner root sheath, hair shaft or dermal hair follicle cells (Sander et al., 2000). During development, it is expressed throughout the embryonic ectoderm but is transiently downregulated at the onset of hair follicle morphogenesis. *BARX2* expression is then observed within epithelial cells of the developing hair follicle except for cells within the placode and matrix that eventually differentiate into the inner root sheath and hair shaft (Sander et al., 2000), which is in accordance with the findings of the human prenatal skin gene expression analysis across the inferred trajectory (Figure 74 and Figure 75). *BARX2* has been shown to have an important role in maintaining keratinising epithelial structures and regulate hair follicle remodelling (Olson et al., 2005; Sander et al., 2000). *SOX9* (SRY-Box Transcription Factor 9) encodes a transcription factor that has been shown to be essential for outer root sheath differentiation in murine models, though is dispensable for the initiation of hair follicle morphogenesis (Vidal et al., 2005).

This inferred trajectory analysis across the prenatal human epithelial hair follicle cell states also revealed new several genes that have not previously been reported to be related to outer root sheath differentiation, including *SPON2* and *DAPL1*. *SPON2* (Spondin 2) encodes the integrin ligand mindin with diverse roles in various tissues, including the

positive regulation of cytokine production and recruitment of lymphocytes, which relates to the possible role of immune cells in hair follicle morphogenesis. *DAPL1* (Death associated protein like 1) has been found to be expressed in epithelial cells directly above the proliferative compartment, including the epidermis and cornea, and has been speculated to have an important role in the early stages of epithelial cell differentiation and apoptosis (Sun et al., 2006; Chen et al., 2021b), which is in keeping with the high expression of *DAPL1* found along the outer root sheath trajectory from basal skin cells (Figure 74 and Figure 75).

Along the inner root sheath differentiation trajectory, *DPYSL2*$^+$ basal cells that were predicted to differentiate into the placode/matrix, cuticle/cortex and inner root sheath showed high expression of known matrix cell markers, including *SHH* and *WNT10B* (Saxena, Mok and Rendl, 2019; Joost et al., 2020), which are regulators of the Wnt/β-catenin signalling pathway. *SPON2* and *AGR2* were observed to be downregulated along the inner root sheath trajectory. In addition to the immune response function of *SPON2* discussed above, *SPON2* has also been reported to have an important role cellular adhesion (Feinstein et al., 1999; Jia, Li and He, 2005). Loss of *AGR2* (Anterior Gradient 2), a molecular chaperone involved in the assembly of cysteine-rich receptors enriched in hair follicles, has been shown to promote cell migration (Martisova et al., 2022; Delom et al., 2020). Together, these findings suggest that reduced cellular adhesion and increased cellular migration properties in *DPYSL2*$^+$ basal cells are involved in matrix specification and dermal invagination.

The analysis above compared broad gene expression differences between the outer root sheath / companion layer trajectory and the inner root sheath trajectory, and therefore, further characterisation of the dynamically expressed genes underpinning each trajectory separately was investigated next. Differentially expressed genes across pseudotime were calculated and the results are shown as a heatmap of gene expression for the outer root sheath trajectory / companion layer in Figure 76 and for the inner root sheath trajectory in Figure 77.
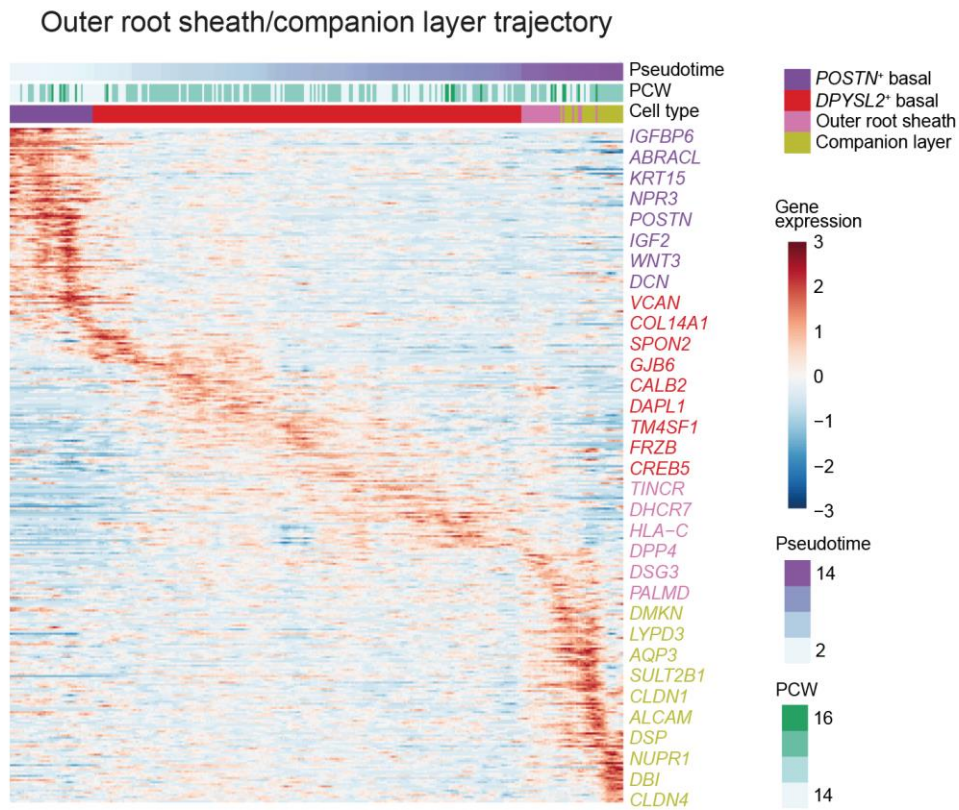
Figure 76: Heatmap showing differentially expressed genes across pseudotime along the outer root sheath/companion layer trajectory.

This analysis highlights genes that are dynamically upregulated during the differentiation of prenatal epithelial cell states from *POSTN*⁺ basal cells to *DPYSL2*⁺ basal cells, then to outer root sheath cells then companion layer cells across Monocle3-inferred pseudotime. Interestingly, the temporal expression of *GJB6* and *DSP* that underpin outer root sheath and companion layer development are also implicated in congenital hair diseases. This will be discussed and explored in further depth in section 5.2.7, which covers pathological processes during skin and hair morphogenesis.
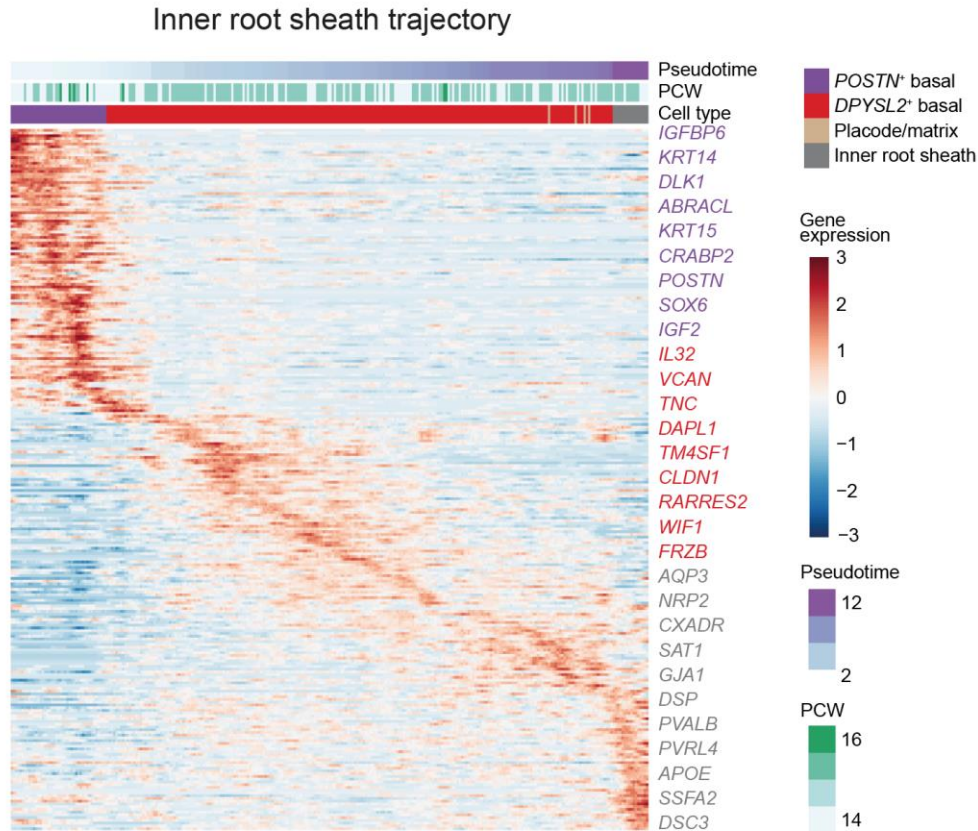
Figure 77: Heatmap showing differentially expressed genes across pseudotime along the inner root sheath trajectory.

The differentially expressed gene analysis across pseudotime for the inner root sheath trajectory highlights genes that are dynamically upregulated during the differentiation of *POSTN*⁺ basal cells to *DPYSL2*⁺ basal cells, then to placode/matrix cells then inner root sheath cells. This analysis revealed that *PVRL4* and *DSC3* are expressed during the later stages of this differentiation pathway, which are genes are also implicated in specific subtypes of congenital hair diseases. This will be also discussed further in section 5.2.7 on skin and hair pathology.

5.2.5. Trajectory inference of fibroblast cell types

Having characterised the differentiation trajectory of epithelial cell states in prenatal skin, the dermal cell types that are involved in crosstalk with epithelial cells during early hair follicle development were investigated next. Within the human prenatal skin single cell dataset, mesenchymal cell states of the developing hair follicle were characterised, and a population of pre-dermal condensate cells was identified. Pre-dermal condensate cells have not previously been reported in human studies but are known to be a transitional fibroblast state involved in murine hair follicle development (Mok et al., 2019; Ma et al., 2022; Abbasi et al., 2021). Pre-dermal condensate cells in embryonic mouse skin are unclustered within the dermis, and aggregate to form the dermal condensate after epidermal placode formation (Mok et al., 2019). The dermal condensate cells form a microanatomical signalling niche with placode cells to regulate ongoing hair follicle development. The hair germ then extends down into the dermis with the dermal condensate cells at the leading edge. After the hair germ becomes an elongated hair peg, dermal condensate cells become encapsulated within the hair shaft-producing bulb region as dermal papilla cells (Saxena, Mok and Rendl, 2019; Grisanti et al., 2013). Notably, the pre-dermal condensate and dermal condensate cells are only present in prenatal skin and not in adult skin in human and mouse (Mok et al., 2019), with self-renewing dermal cells maintaining the dermal papilla during adult hair follicle cycling (Rahmani et al., 2014; Wang et al., 2020a).

To infer the origin of the pre-dermal condensate, dermal condensate and dermal papilla cells in human prenatal skin during hair follicle morphogenesis, inferred trajectory analysis was performed using Monocle3 across the different fibroblast clusters in the prenatal skin dataset. The output of this analysis is shown as a graph embedding below:
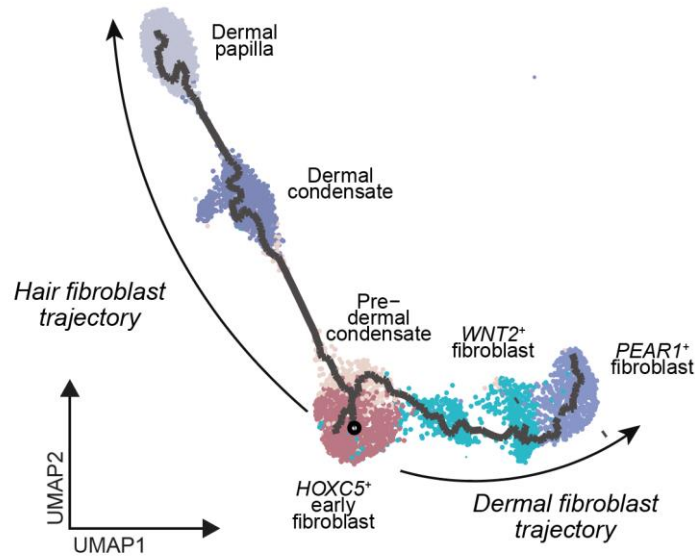
Figure 78: Inferred trajectory of hair follicle fibroblast differentiation. Pseudotime trajectory analysis using Monocle3 shows fibroblast differentiation along the 'hair fibroblast' and 'dermal fibroblast' trajectories.

The inferred trajectory analysis predicted that *HOXC5*[+] early fibroblasts, which are absent after 11 PCW (Figure 60), differentiate along two distinct paths. Along one trajectory, the *HOXC5*[+] early fibroblasts differentiate into pre-dermal condensate cells, then into dermal condensate cells then into dermal papilla cells, which will be referred to as the hair fibroblast trajectory (Figure 78). This finding is in accordance with hair fibroblast differentiation observed in murine studies, discussed above (Mok et al., 2019; Ma et al., 2022; Abbasi et al., 2021).

Along the second trajectory, the *HOXC5*[+] early fibroblasts differentiate into *WNT2*[+] fibroblasts then into *PEAR1*[+] fibroblasts, which is referred to as the dermal fibroblast trajectory here (Figure 78). As this study aims to delineate the cellular and molecular processes underlying prenatal hair follicle morphogenesis, the focus of the interpretation will be on the hair-specialised fibroblasts. Downstream analyses will also focus on the hair fibroblast trajectory and the genes underpinning this differentiation pathway later in this section.

181

In addition to the analysis output being visualised based on cell states along each trajectory above (Figure 78), the graph embedding was visualised based on the Monocle3 derived pseudotime and gestational age (PCW):
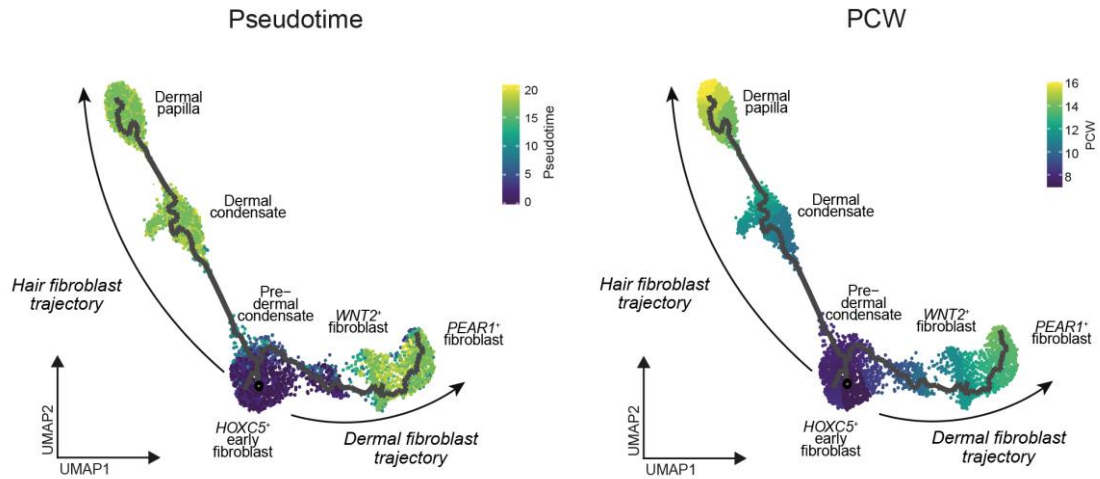
Figure 79: Inferred trajectory of fibroblast differentiation, coloured by pseudotime (left) and gestational age (PCW) (right).

These UMAP visualisations demonstrate the that the prenatal hair follicle fibroblast cell states that are at the distal end of the inferred trajectory are found in mostly later gestational age samples, which is in keeping with the sampling period spanning the stages in which the hair follicle first forms. This gives biological support to this computational approach in which cell transcriptomes are ordered on a pseudotime scale to predict differentiation paths.

Differentially expressed genes across pseudotime were calculated next to determine the dynamic transcriptomic changes unpinning the hair fibroblast trajectory. The results are shown as a heatmap of gene expression in Figure 80.
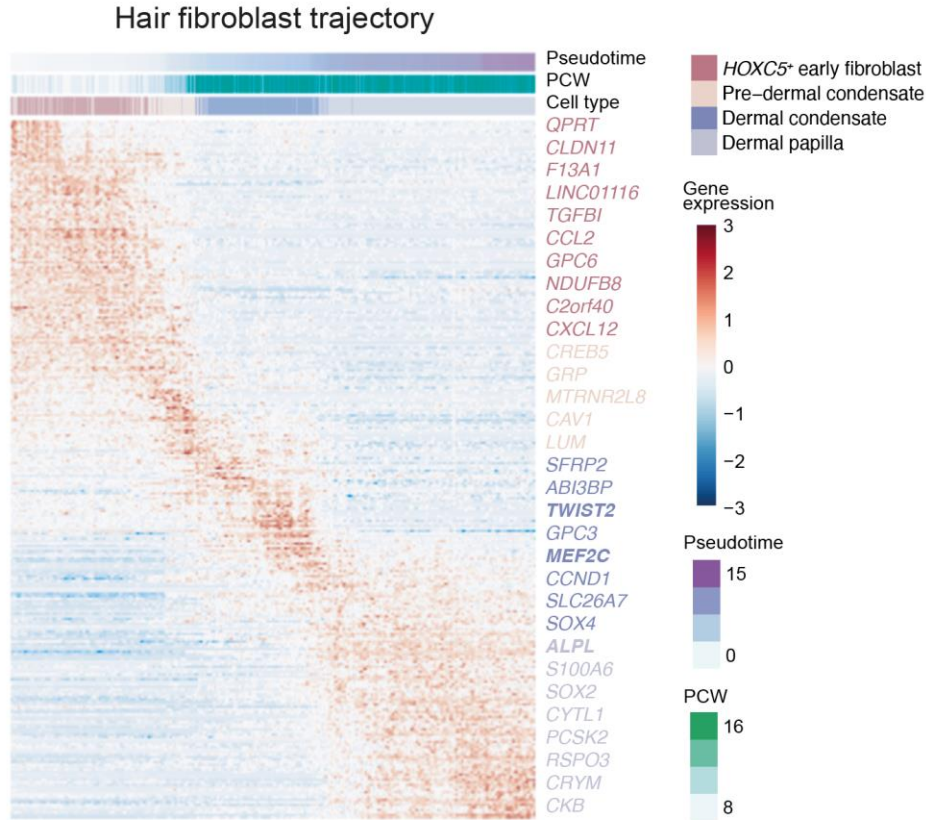
Figure 80: Heatmap showing differentially expressed genes across pseudotime along the hair fibroblast trajectory.

This analysis highlights genes that are dynamically upregulated across the hair fibroblast trajectory and provides further insights into the processes involved during prenatal hair specialised fibroblast differentiation into the dermal papilla. Several notable genes were upregulated by *HOXC5*[+] early fibroblasts as they differentiate into pre-dermal condensate cells, including *TGFB1 (*Transforming Growth Factor Beta 1), which is involved in the reduction of cell-matrix adhesion (Boguslawska et al., 2016; Shelton and Rada, 2009), *CLDN11* (Claudin 11), which has a role in the maintenance of cell-cell contact (Wessells et al., 2009) and *CXCL12* (C-X-C Motif Chemokine Ligand 12), which mediates directed cell migration (Dillenburg-Pilla et al., 2015). Pre-dermal condensate are known from murine studies to migrate to the epidermis at this stage of their differentiation, and these dynamically upregulated genes support the process of collective cell migration (Mok et al., 2019; Li et al., 2019).

As the pre-dermal condensate cells aggregate and differentiate to form the dermal condensate cells, genes implicated in collagen fibril formation and cell adhesion, including

LUM (Lumican) and *ABI3BP* (ABI Family Member 3 Binding Protein) were upregulated (Krishnan et al., 2012; Hsiao et al., 2020; Delfín, DeAguero and McKown, 2019) (Figure 80). Formation of the dermal condensate was characterised by genes such as *RSPO3* (R-Spondin 3) and *SOX2* (SRY-Box Transcription Factor 2) (Figure 80). *SOX2* is known to be expressed from the dermal condensate stage in mouse hair follicles, but its expression is restricted to the dermal papilla in humans (Weber et al., 2020). *RSPO3* and *SOX2* are involved in spatially coordinating the migration and proliferation of the adjacent hair matrix epithelial cells (Saxena, Mok and Rendl, 2019; Hagner et al., 2020; Takeo, Lee and Ito, 2015).

5.2.6. Mesenchymal-epithelial cellular crosstalk during hair follicle neogenesis

Embryonic hair follicle formation is known to rely on the exchange of tightly regulated molecular signals between mesenchymal and epithelial cells (Millar, 2002). However, our understating of these inductive processes have largely been inferred from murine studies (Mok et al., 2019; Saxena, Mok and Rendl, 2019), and to date, single cell transcriptomic data of non-immune as well as immune cells in prenatal skin spanning the period of hair follicle neogenesis has not been previously published. This data was therefore leveraged to interrogate the mesenchymal-epithelial interactions that are implicated in early human hair follicle formation. Cell-cell communication analysis using CellphoneDB (Efremova et al., 2020) was carried out, which predicts ligand-receptor interactions between cell states. Various mesenchymal and epithelial cell populations are only present during specific time periods, as was demonstrated in the differential abundance analysis across gestation in section 4.2.7 and the trajectory analyses in sections 5.2.4 and 5.2.5, which was factored into the analysis. Cell-cell communication between *HOXC5*+ early fibroblasts and pre-dermal condensate cells were both tested against immature basal cells present until 11PCW, and cell-cell communication between dermal condensate cells and dermal papilla cells were both tested against epithelial cells present from 12 PCW (including DPYSL2+ basal cells, POSTN+ basal cells, outer root sheath cells, companion layer cells, inner root sheath cells and cuticle/cortex cells. The output of the analysis is shown as a heatmap below (Figure 81).
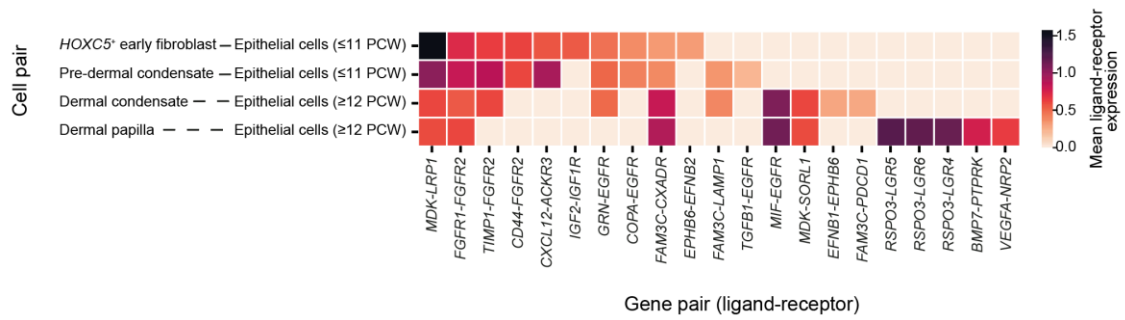
Figure 81: Mesenchymal-epithelial cell interaction analysis. Heatmap visualisation of significant (adjusted p-value <0.05) predicted interactions (CellphoneDB) between hair mesenchymal cell states and early epithelial cells (≤11 PCW; immature basal) or late epithelial cells (≥12 PCW; DPYSL2⁺ basal, POSTN⁺ basal, placode/matrix, outer root sheath, companion layer, inner root sheath, cuticle/cortex) in prenatal skin. Colour scale represents the mean expression values of each ligand-receptor pair for the corresponding pairs of cell types.

This mesenchymal-epithelial cell-cell communication analysis across the gestational period in which human hair follicles first develop revealed several insights into the molecular signals underlying this process. For example, the *CXCL12* ligand on *HOXC5⁺* early fibroblasts and pre-dermal condensate cells was shown to interact with the *ACKR3* (Atypical Chemokine Receptor 3) receptor on immature basal cells until 11 PCW (Figure 81). In order to further characterise the temporal and cellular specificity of signalling via *CXCL12* during hair follicle formation, *CXCL12* gene expression was analysed across the different hair follicle related fibroblast subsets and across gestational age, as shown in the violin plots below in Figure 82.
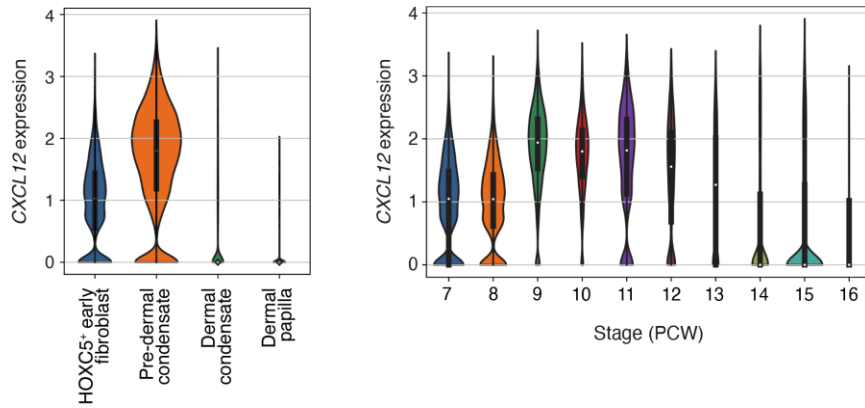
Figure 82: Violin plots showing expression of CXCL12 in hair follicle mesenchymal cells by cell type (left) and gestational age (right). Gene expression values on the y-axis are log-transformed, normalised and scaled.

This analysis showed that *CXCL12* expression by hair follicle dermal cells was highest in *HOXC5+* early fibroblasts and pre-dermal condensate cells, and expression was downregulated after 12 PCW (Figure 82), suggesting that CXCL12 interacts with the ACKR3 receptor on epithelial cells. *CXCL12* is known to have an important role in directed cell migration (Dillenburg-Pilla et al., 2015; Belmadani et al., 2009) and was shown to be dynamically upregulated in the trajectory analysis during pre-dermal condensate differentiation (Figure 80). Collectively, this suggests that CXCL12 interacts with the ACKR3 receptor on epithelial cells and mediates the migration of pre-dermal condensate cells to form the dermal condensate.

Formation of the dermal condensate is known to be accompanied by dermal invagination of the placode/matrix, and the cell-cell communication analysis revealed that *FAM3C* (FAM3 Metabolism Regulating Signalling Molecule C) expressed on dermal condensate cells was predicted to interact with *LAMP1* (Lysosomal Associated Membrane Protein 1) and with *CXDAR* (CXADR Ig-Like Cell Adhesion Molecule) expressed on hair follicle epithelial cells beyond 12 PCW (Figure 81). *FAM3C* encodes a secreted signalling protein that is known to promote cell migration and invasion (Kral et al., 2017; Yin et al., 2018), which is likely to facilitate dermal condensate encapsulation within the developing hair bulb. Furthermore, *EFNB1* (Ephrin B1) on dermal condensate cells were predicted to interact with *EPHB6* (EPH Receptor B6) on hair follicle epithelial cells (Figure 81). *EFNB1* encodes a transmembrane signalling protein that has been shown to have an important role

in epithelial cell migration during development and wound healing (Moore et al., 2004; Nunan et al., 2015), which reflects the developmental processes at this stage of hair follicle morphogenesis.

The cell-cell communication analysis revealed that dermal papilla cells expressing *RSPO3* were predicted to interact with *LGR4/6* (Leucine Rich Repeat Containing G Protein-Coupled Receptor 4/6) on hair follicle epithelial cells beyond 12 PCW (Figure 81). As discussed in section 5.2.5, where inferred trajectory analysis of hair specialised fibroblasts showed an upregulation of *RSPO3* at the late stages of differentiation into dermal papilla cells, *RSPO3* is known to contribute to the proliferation of adjacent hair matrix epithelial cells (Saxena, Mok and Rendl, 2019; Hagner et al., 2020).

The following schematic summarises the results and interpretation of the analyses detailed in this chapter on the developmental cellular processes underlying hair follicle neogenesis, with a focus on the mesenchymal-epithelial interactions described in this section (Figure 83).
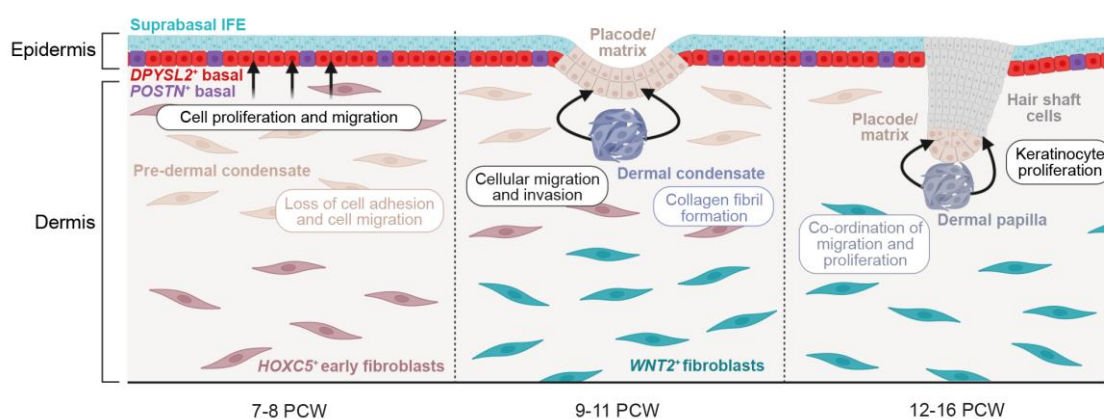
Figure 83: Schematic representation of mesenchymal-epithelial signalling and cellular processes involved during hair formation. Image created with BioRender.com by Chloe Admane.

During the early stages of prenatal skin development, *HOXC5*⁺ early fibroblasts differentiate into pre-dermal condensate cells, which involves an upregulation of genes and cell-cell interactions that mediate fibroblast cell migration to the epidermis and fate specification of epidermal cells to proliferate and form placodes. Subsequently, the dermal

condensate forms, which involves collagen fibril formation, and abuts the invaginating placode, which is reflected in an upregulation of genes and cellular crosstalk involving cellular migration and invasion. Beyond 12 PCW, as the hair peg extends down into the dermis, the dermal papilla forms and becomes encapsulated within the hair bulb, interacting with epithelial matrix cells to regulate keratinocyte proliferation to form the developing hair shaft.

5.2.7. Gene mutations implicated in skin and hair diseases and the support for disease modelling *in vitro*

Having mapped the formation and differentiation of human prenatal skin hair follicles at a single cell level, this information was then leveraged to gain insights into genetic hair disorders. As briefly mentioned in section 5.2.4, genes that were dynamically expressed during the differentiation of prenatal epithelial cell states along the outer root sheath / companion layer trajectory included *GJB6* (Gap Junction Protein Beta 6) and *DSP* (Desmoplakin). *GJB6* was upregulated when epithelial cell states along the trajectory had differentiated into $DPYSL2^+$ basal cells, whereas DSP was upregulated at the distal end of the differentiation pathway when cells had differentiated into companion layer cells surrounding the hair shaft.

Interestingly, *GJB6* and *DSP* are implicated in different subtypes of congenital diseases that affect hair development, which reflects important role of these genes during hair follicle formation prenatally. Mutations in *GJB6* leads to a condition called Clouston syndrome, a rare autosomal dominant genetic skin disease that is also known as hidrotic ectodermal dysplasia (Liu et al., 2015). Ectodermal dysplasias are a heterogeneous group of genetic diseases in which the ectodermal layer of the embryo fails to develop normally, which can affect the skin, hair, nails and sweat glands (Wright et al., 2019). Different subtypes of ectodermal dysplasia are characterised by specific disease manifestations that depend on the gene mutations underlying each disease subtype, the cells that they are expressed in and their expression during earlier or later development, affecting the severity of the disease (Deshmukh and Prashanth, 2012). Individuals affected by Clouston syndrome have pale, fine, sparse hair that grows very slowly, and in some cases, hair loss (alopecia) may occur. The eyelashes and eyebrows are also either absent or short and sparse (Cammarata-Scalisi et al., 2019). Clouston syndrome is also characterised by nail

dystrophy (abnormalities in the growth of the nails) and palmoplantar hyperkeratosis (thickening of the skin of the palms and soles) (Liu et al., 2015). As a type of hidrotic ectodermal dysplasia, Clouston syndrome does not affect the sweat glands, whereas hypohidrotic ectodermal dysplasias affect the ability to produce sweat (Reyes-Reali et al., 2018). The clinical manifestations of Clouston syndrome highlight the important role of *GJB6* in ectodermal development of the skin and its appendages, including the hair follicle.

Desmoplakin is encoded by the *DSP* gene and is a cytoskeletal component of desmosomes that mediates cell-cell adhesion in epithelial structures, including the skin and hair follicle (Brooke, Nitoiu and Kelsell, 2012). Mutations in *DSP* are known to cause skin fragility-woolly hair syndrome, which is a rare autosomal recessive condition characterised by excessive skin blistering secondary to minimal mechanical trauma, and thin woolly hair that is very kinked and difficult to brush (Al-Owain et al., 2011). Some cases are also associated with cardiomyopathy, due to the role of desmoplakin in cardiac myocyte intercellular junctions (Brandão et al., 2023; Al-Owain et al., 2011). Expression of *DSP* along the outer root sheath trajectory of epithelial differentiation of the hair follicle and the disease manifestations caused by mutations of *DSP* highlight its role in skin and hair morphogenesis.

As briefly mentioned in section 5.2.4, genes that were dynamically expressed during the differentiation of prenatal epithelial cell states along the inner root sheath trajectory included *PVRL4* (Nectin Cell Adhesion Molecule 4) and *DSC3* (Desmocollin 3). Both genes were upregulated at the distal end of the differentiation pathway after matrix cells had differentiated into inner root sheath cells. *PVRL4* is implicated in a congenital disorder known as ectodermal dysplasia – syndactyly syndrome, which is characterised by sparse, thin hair affecting the scalp, eyebrows and eyelashes, and alopecia in some cases. Affected individuals also have palmoplantar keratoderma, hypoplastic small and underdeveloped nails, conical-shaped teeth and cutaneous syndactyly, where fingers are webbed or fused together at birth (Raza et al., 2015). Desmocollin 3 is a component of intercellular desmosome junctions (Petrof, Mellerio and McGrath, 2012), and mutations in *DSC3* are known to cause a congenital condition called hypotrichosis and recurrent skin vesicles (Ayub et al., 2009). The scalp hair of affected individuals is sparse and fragile, and eyebrows and eyelashes are absent. Vesicles are small fluid-filled lesions that develop across most of the body of individuals with this disease. The manifestations of these

diseases highlight the critical role that these genes have during skin and hair follicle development.

To gain further insights into the role of these genes implicated in disease during development, the prenatal single cell data was leveraged to determine the level of gene expression across the cell states involved in hair follicle development, which is shown in Figure 84 below. Furthermore, the expression of these genes was determined within cell states of the hair-bearing skin organoid model single cell dataset (Lee et al., 2020) in order to determine whether the gene expression is similar across these cell states and the potential utility of the skin organoids to study disease (Figure 84).
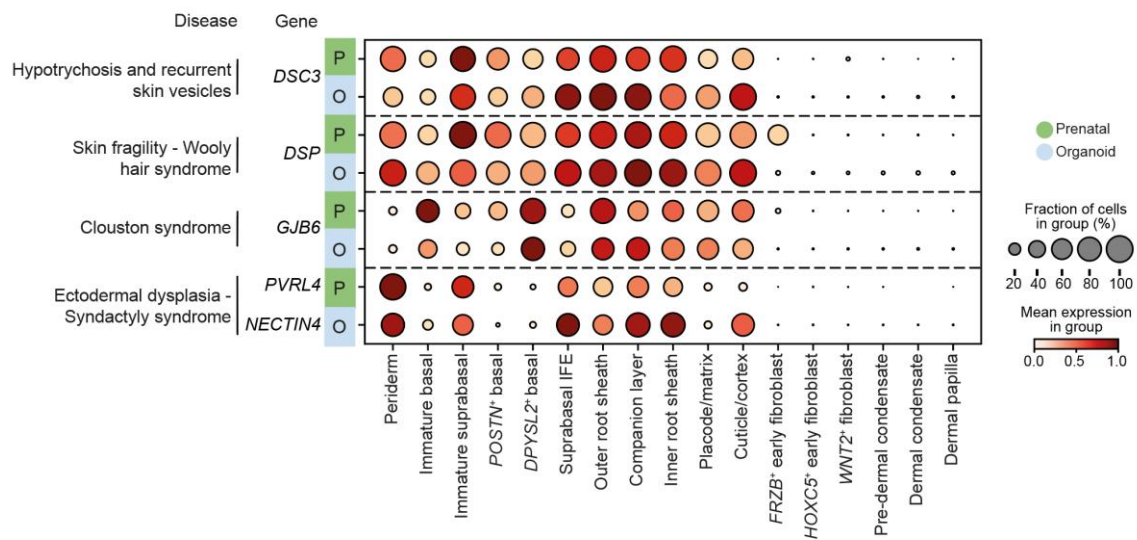
Figure 84: Genetic hair diseases. Dot plot showing variance-scaled expression of genes implicated in genetic hair diseases in prenatal skin (P; green) and skin organoid skin (Lee et al., 2020) (O; blue). Mean expression is represented by dot colour and percentage of expressing cells is represented by dot size.

This analysis showed that the genes implicated in the genetic skin and hair diseases discussed above are expressed within hair follicle related epithelial cell states rather than hair specialised fibroblasts, reflecting their role in epithelial integrity, and which is supported by their expression along the outer root sheath / companion layer and inner root sheath trajectories, as shown in section 5.2.4. Furthermore, there are differences in the level of expression of some genes between the hair follicle epithelial cells, reflecting the heterogeneous clinical manifestations across these disease conditions. Interestingly, the

genes implicated in the congenital skin and hair diseases are similarly expressed in equivalent cell states in the skin organoid model (Figure 84), illustrating the potential value of skin organoids as a disease modelling tool for studying these disorders.

The inferred trajectory and differentially expressed gene expression analysis of hair specialised fibroblasts across pseudotime, which was presented in section 5.2.5, revealed a panel of genes that regulate the differentiation process from *HOXC5*$^+$ early fibroblasts to pre-dermal condensate cells, then to dermal condensate cells, then to dermal papilla cells. Interestingly, several of the genes upregulated during this differentiation pathway are also known, from genome-wide association studies (GWAS), to be associated with androgenetic alopecia, which is a genetically predetermined disorder characterised by an excessive response to androgens that causes patterned hair loss (Lolli et al., 2017). Androgenetic alopecia is sometimes known as male-pattern baldness, and affects up to 80% of the male population, causing bitemporal and vertex hair thinning and loss. However, androgenetic alopecia also affect up to 50% women, causing hair thinning across the frontal and parietal scalp, with retention of the anterior hair line (Piraccini and Alessandrini, 2014; Price, 2003). GWAS studies have reported that the following genes that are expressed along the hair fibroblast trajectory (Figure 80) are associated with androgenetic alopecia: *TWIST2* (Twist Family BHLH Transcription Factor 2), *MEF2C* (Myocyte Enhancer Factor 2C) and *ALPL* (Alkaline Phosphatase) (Pirastu et al., 2017; Marcińska et al., 2015). Together, these genes are known to be involved in androgen receptor sensitivity and Wnt signalling (Kwack et al., 2019; Jacob et al., 2021; Wyce et al., 2010), which supports their role in the pathogenesis of androgenetic alopecia in which the hair follicle miniaturises in response to circulating androgens.

The gene expression of *TWIST2*, *MEF2C* and *ALPL* across the hair follicle related prenatal and organoid single cell populations was determined to gain further insights into their role during development and assess whether the skin organoid cells demonstrate similar expression patterns to *in vivo* skin cells:
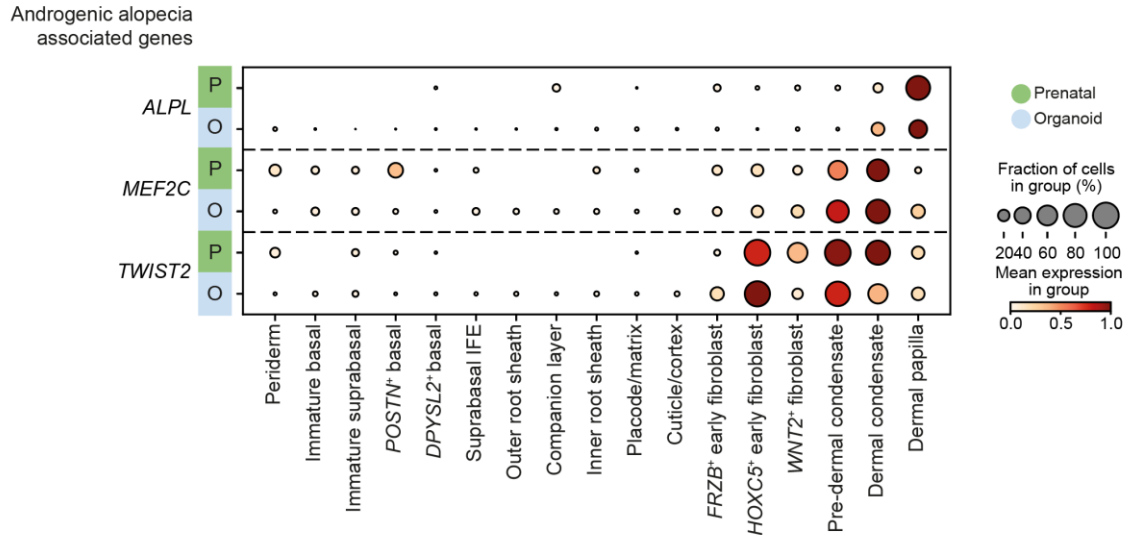
Figure 85: Androgenetic alopecia. Dot plot showing variance-scaled expression of genes implicated in androgenetic alopecia in prenatal skin (P; green) and skin organoid skin (Lee et al., 2020) (O; blue). Mean expression is represented by dot colour and percentage of expressing cells is represented by dot size.

This analysis shows that the expression of genes associated with androgenetic alopecia is restricted to hair follicle specialised fibroblasts rather than epithelial cells. Expression of *ALPL* is highest in the dermal papilla cells, which reflects its upregulation at the distal end of the hair fibroblast differentiation pseudotime trajectory (Figure 80), whereas expression of *MEF2C* and *TWIST2* is higher in pre-dermal condensate cells and dermal condensate cells, reflecting their upregulation at an earlier stage of the pseudotime trajectory. This supports the hypothesis that early mesenchymal differentiation relating to hair follicle formation may be altered *in utero* and contribute towards developing androgenetic alopecia in adult life (Pirastu et al., 2017; Marcińska et al., 2015; Heilmann-Heimbach et al., 2016). As was observed for the genes expressed in epithelial cell states that are implicated in congenital skin and hair diseases, a similar pattern of expression between prenatal and skin organoid cell states was seen across the fibroblast populations for genes implicated in androgenetic alopecia. This provides further support to the utility hair-bearing skin organoid models to investigate diseases that affect both the epithelial and mesenchymal compartments of the hair follicle during development.

In order to extend the notion and value of skin organoids as a model for skin diseases, the expression of genes causing an inherited blistering skin disorder called epidermolysis

bullosa was assessed. Epidermolysis bullosa presents at birth or during early infancy, suggesting *in utero* onset of the disease (Has et al., 2020), and has traditionally been classified into different subtypes depending on the site of formation of blisters within the skin layers: epidermolysis bullosa simplex (EBS) causes blistering within the epidermis, junctional epidermolysis bullosa (JEB) causes blistering at the level of the skin basement membrane, dystrophic epidermolysis bullosa (DEB) causes blistering below the level of the basement membrane and Kindler epidermolysis bullosa can cause intraepidermal, junctional or dermal skin cleavage but is characterised by acral blister formation at the distal extremities (Fine et al., 2008). Mutational analyses in conjunction with transmission electron microscopy and immunofluorescence mapping has led to the development of updated epidermolysis bullosa classification systems in which certain gene mutations can lead to specific subtypes of the disease (Kotalevskaya and Stepanov, 2023; Has et al., 2020). The expression of gene mutations implicated in different subtypes of epidermolysis bullosa was determined across the prenatal and organoid cell populations to further characterise their role during development *in vivo* and *in vitro*:
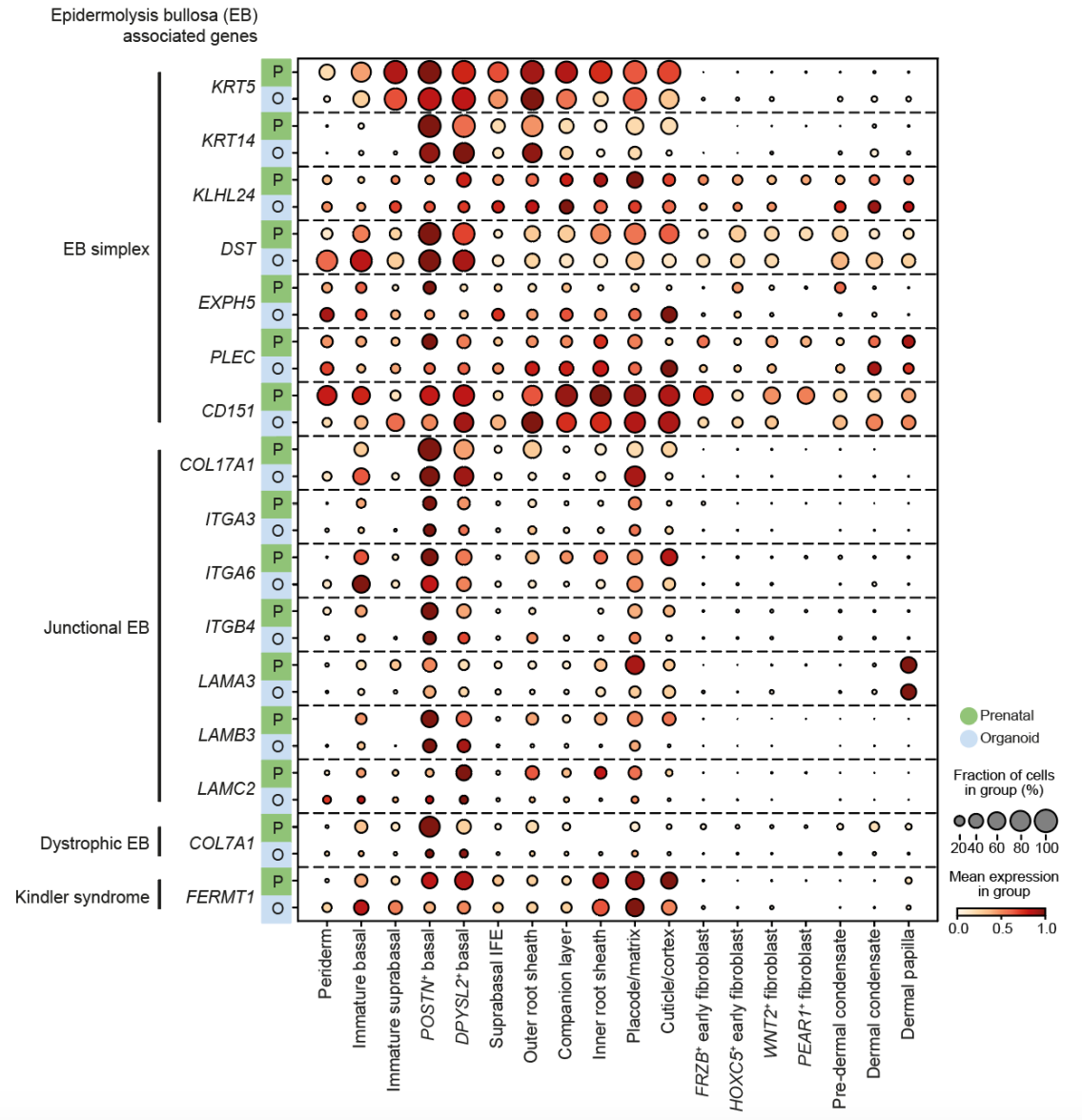
Figure 86: Epidermolysis bullosa. Dot plot showing variance-scaled expression of genes causing Epidermolysis Bullosa in prenatal skin (P; green) and skin organoid skin (Lee et al., 2020) (O; blue). Mean expression is represented by dot colour and percentage of expressing cells is represented by dot size.

This analysis showed a higher expression of genes implicated in EBS (*KRT5*, *KRT14*, *KLHL24*, *DST*, *EXPH5*, *PLEC* and *CD151*), JEB (*COL17A1*, *ITGA3*, *ITGA6*, *ITGB4*, *LAMA3*, *LAMB3* and *LAMC2*), DEB (*COL7A1*) and Kindler epidermolysis bullosa (*FERMT1*) in epithelial cells compared to mesenchymal cells. Interestingly, across the epithelial cell subtypes, the highest expression of most genes (especially genes causing JEB and DEB) was observed in *POSTN*⁺ basal cells (Figure 86). This suggests potential stronger anchorage between *POSTN*⁺ basal epidermal cells and the basement membrane and dermis

194

compared to other epithelial cell states (Wang et al., 2020b). Moreover, gene therapy studies for DEB have identified that fibroblasts expressing *COL7A1* may be a better target for retroviral transfer than keratinocytes (Goto et al., 2006; Varki et al., 2007), and within the dermal compartment, highest prenatal expression of *COL7A1* was observed in the dermal condensate cells, which can help guide future gene therapy studies for epidermolysis bullosa. When comparing the gene expression of epidermolysis bullosa genes between prenatal skin and the skin organoid, a similar pattern of expression was seen for most disease subtypes (Figure 86). As discussed in related to the other skin and hair diseases discussed, this provides further support for the range of conditions that can be modelled and studied using *in vitro* methods.

The prenatal and organoid datasets were also leveraged to better understand the pathogenesis of congenital ichthyoses, a group of heterogeneous monogenic skin diseases that lead to disordered keratinisation, skin scaling and hair loss. Congenital ichthyosis can be associated with ophthalmic complications such as ectropion (outward turning of the eyelid margin and symptoms due inadequate ocular lubrication) and ear complications such as hearing loss (Mazereeuw-Hautier et al., 2019). In severe cases, such as a subtype known as Harlequin ichthyosis caused by a mutation in *ABCA12*, congenital ichthyosis is associated with premature death in neonatal life (Vahlquist, Gånemo and Virtanen, 2008). Congenital ichthyoses are classified based on their clinical presentation and distinguishes between non-syndromic ichthyoses that are confined to the skin (including autosomal recessive congenital ichthyosis (ARCI), epidermolytic ichthyosis and erythrokeratodermia variabilis (EKV)) and syndromic ichthyoses that affect the skin and other organ systems (Mazereeuw-Hautier et al., 2019). The gene expression of mutations implicated in different subtypes of congenital ichthyoses in prenatal and organoid skin is shown below:
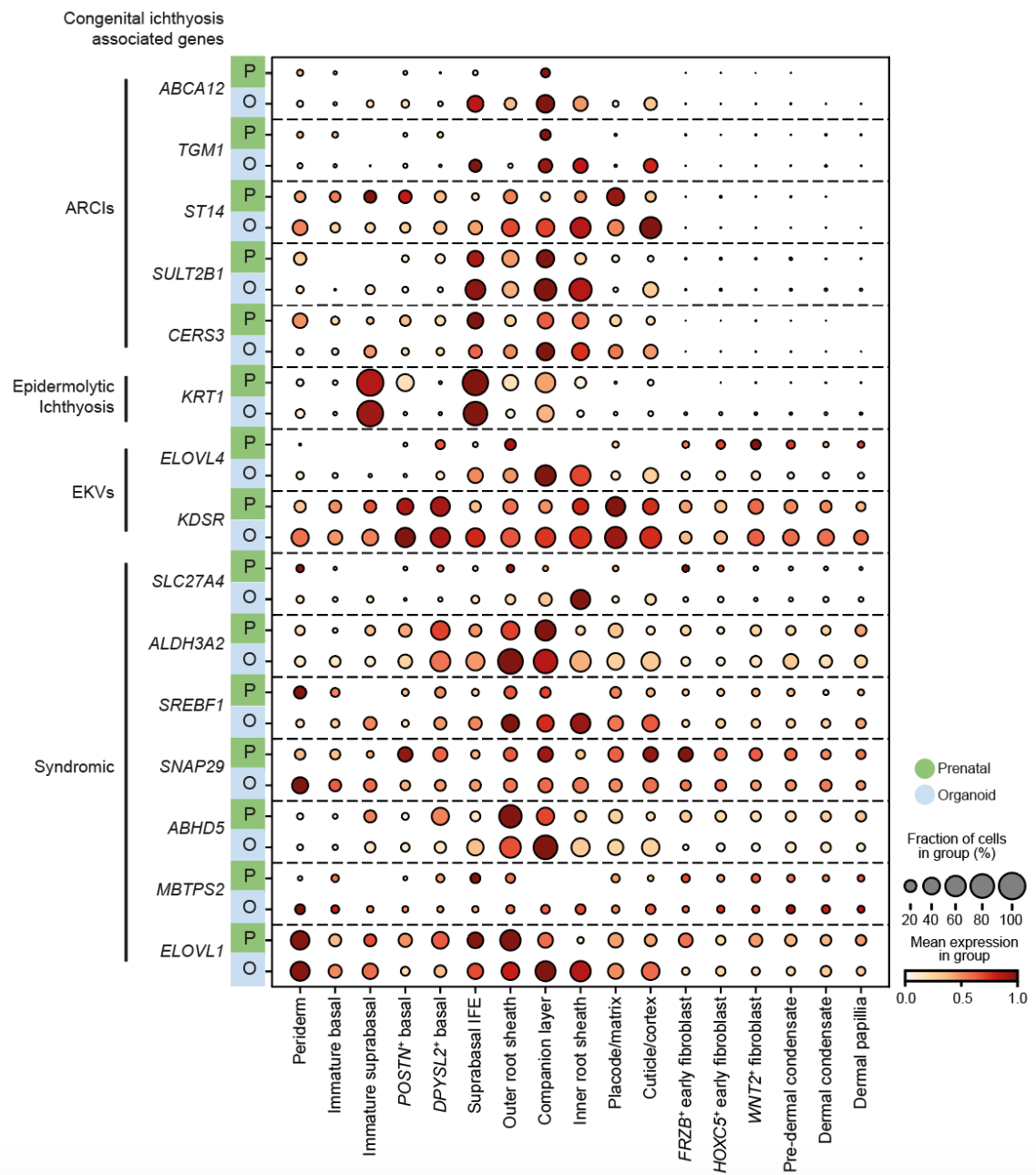
Figure 87: Congenital ichthyosis. Dot plot showing variance-scaled expression of genes causing congenital ichthyoses in prenatal skin (P; green) and skin organoid skin (Lee et al., 2020) (O; blue). Mean expression is represented by dot colour and percentage of expressing cells is represented by dot size. ARCI = autosomal recessive congenital ichthyosis, EKV = erythrokeratodermia variabilis.

This analysis showed that genes mutated in the syndromic forms of congenital ichthyosis that involve other organ systems are expressed highly across both epithelial and mesenchymal cell states, whereas genes causing non-syndromic forms with manifestations predominantly confined to the skin are mostly expressed in epithelial cells (Figure 87).

Furthermore, genes implicated in most congenital ichthyosis subtypes, particularly the ARCIs, showed high expression in the hair follicle specialised epithelial cells. This expression patten highlights the important role that the genes implicated in congenital ichthyoses have in hair follicle development as well as skin morphogenesis, which is reflected by the clinical manifestation of hair loss in affected individuals. Though there were some differences in the expression patterns of disease related genes between prenatal and skin organoid cell states, there is an overall similar pattern of expression across the epithelial and dermal compartments. Therefore, these findings lay the foundations for disease modelling using skin organoid systems and show that any future experiments aiming to use skin organoids to model either congenital ichthyoses or other skin or hair diseases must consider any cellular differences in disease related gene expression between *in vivo* and *in vitro* skin.

5.3. Discussion

In this Chapter, the *de novo* development of human hair follicles during prenatal life is presented. This study represents the first single cell transcriptomic analysis of human prenatal hair follicle specialised epithelial and stromal cells captured during the gestational period in which hair follicles first develop. Human studies of prenatal hair follicle formation have been limited to morphological descriptions (Schirren et al., 1997; Ersch and Stallmach, 1999; Silva et al., 2020, 2021), and previous scRNA-seq studies on human prenatal skin during this early period of gestation have profiled only immune cells, without any skin epithelial and stromal context to the findings (Xu et al., 2021; Bian et al., 2020). Our understanding of the cellular and molecular processes involved have been largely been derived from studies on murine models (Mok et al., 2019; Saxena, Mok and Rendl, 2019), however, it is unknown as to whether the precise mechanisms underlying skin and hair follicle development are conserved between species.

The microanatomical structures within the skin during the gestational period in which samples were collected were assessed using histological analysis and multiplex RNA *in situ* hybridisation, confirming that hair follicles form during the second trimester of gestation after 11 PCW. The epithelial and stromal cell states that contribute to hair follicle formation were then characterised at a single cell level, allowing the integration and comparison with a scRNA-seq dataset of adult hair follicles in the anagen growth phase of the hair growth cycle (Takahashi et al., 2020). Interestingly, this revealed prenatal development specific cell states in human skin and hair follicles, including interfollicular cells and hair placode cells. Hair placode formation represents one of the initial stages of hair follicle morphogenesis, and analysis of the transcriptomic profile of human placode cells, their differentiation and intercellular interactions has provided unique insights into human hair follicle development that cannot be inferred from adult skin data.

The prenatal skin and hair follicle single cell data was then leveraged to predict differentiation trajectories of epithelial and dermal fibroblast cell states and calculate the genes that are dynamically upregulated along these differentiation pathways. Within the epithelial cell compartment, the analysis revealed two distinct differentiation trajectories from basal keratinocytes into the placode/matrix, cuticle/cortex and inner root sheath along

one differentiation pathway and basal keratinocytes into the outer root sheath and companion layer along another. The companion layer is classically considered to arise alongside the inner root sheath from the hair matrix (Sequeira and Nicolas, 2012; Saxena, Mok and Rendl, 2019). However, the findings from the inferred trajectory analysis in human prenatal epithelial cell states indicate that the companion layer arises along the same differentiation pathway as the outer root sheath, independent of the hair matrix. This novel finding has recently been supported by a murine fate-mapping study that showed that the companion layer develops prior to matrix formation and in the absence of matrix cells (Mesler et al., 2017). A recent single cell study of postnatal murine hair follicles also supported the findings in human prenatal skin, which reported that companion layer cells show a greater transcriptional similarity to outer root sheath cells than inner root sheath cells (Joost et al., 2020). These studies highlight the power of single cell transcriptomic data to detect cellular and molecular characteristics of cell states and how they relate to one another along a differentiation path.

The inferred trajectory analysis of mesenchymal cells revealed a differentiation pathway that leads to the development of hair specialised fibroblasts, in which $HOXC5^+$ early fibroblasts differentiate into pre-dermal condensate cells, then into dermal condensate cells, then into dermal papilla cells. $HOXC5^+$ early fibroblasts also differentiate into dermal fibroblasts that are not associated with the hair follicle. This resembles the differentiation of hair specialised and dermal fibroblasts in murine skin, which also originate from a common fibroblast progenitor (Driskell et al., 2013). Previous murine studies have also identified pre-dermal condensate cells to be a transitional fibroblast cell state involved in hair follicle formation (Mok et al., 2019; Ma et al., 2022; Abbasi et al., 2021), and this study represents the first characterisation of pre-dermal condensate cells in human skin. Furthermore, pre-dermal condensate and dermal condensate cells are uniquely found in prenatal developing skin and are not present in adult human and murine skin (Mok et al., 2019), which again highlights the need to sample prenatal skin during this stage of gestation in order to be able to characterise the cellular mechanisms underlying hair follicle formation. By leveraging the single cell transcriptomic data to calculate genes that are dynamically upregulated along the hair fibroblast cell pseudotime trajectory and determine cellular crosstalk between these mesenchymal cells and epithelial cell states, unique insights into the molecular mechanisms regulating early hair follicle morphogenesis are presented.

It is important to note that prenatal skin samples until 16 PCW of gestation were collected and processed in this study, and therefore, the dataset did not capture several mature hair follicle related structures that develop later in gestation. The dermal sheath is a structure of layered connective tissue that surrounds the outer root sheath layer of the hair follicle (Martino, Heitman and Rendl, 2021), and a distinct cluster of dermal sheath cells were not detected in this study. A recently published transcriptomic analysis of human prenatal scalp skin from samples aged 16 and 17 PCW did however identify dermal sheath cells (Weber et al., 2020), which is in keeping with the temporal difference in hair follicle development across different anatomical sites. Furthermore, sebaceous gland and apocrine gland cells, which eventually form part of the hair follicle unit, called the pilosebaceous unit, were not captured in this study, as these structures are known to develop after 16 PCW (Akiyama, Smith and Shimizu, 2000). Eccrine sweat glands also develop from around 20 PCW, and therefore cell states that make up eccrine glands were also not captured in this study (Ersch and Stallmach, 1999)

A further consideration to be made in the interpretation of this analysis is how accurately the hair-bearing skin organoid model (Lee et al., 2020) can model skin and hair diseases. The analyses presented in this chapter demonstrated that genes implicated in congenital hair diseases, androgenetic alopecia, epidermolysis bullosa and congenital ichthyosis are broadly similarly expressed epithelial and dermal cell states within prenatal and organoid skin, lending support to the recapitulation of cellular and molecular manifestations of the disease within the organoid model. However, as discussed in Chapter 3, the skin organoid model lacks immune cells and has a paucity of endothelial cells when compared to *in vivo* prenatal skin. The pathogenesis of the diseases discussed in this Chapter are complex, mediated by the interactions of different cell states, and are known to be involve immune cell infiltration and activation (Zhou et al., 2022; Akiyama, 2022; Huitema et al., 2021; Patzelt et al., 2008). Therefore, future studies would be required to assess how disease gene mutations manifest within a skin organoid model *in vitro*, and whether various aspects of disease pathogenesis and treatment can be studied using this system.

Overall, this Chapter presented a morphological and single cell analysis of human prenatal hair follicle formation spanning the gestational time period in which hair follicles first form. A comparison with adult hair follicle cell states highlighted developmental specific cell

states and pseudotime trajectory analyses revealed novel insights into the differentiation of human hair follicle specialised epithelial cells and fibroblasts during prenatal life. The cellular crosstalk between mesenchymal and epithelial cells was then explored to delineate molecular mechanisms underlying the different stages of hair follicle formation. Lastly, the genes dynamically upregulated during the differentiation of hair follicle cell states that are known to be implicated in various skin and hair diseases were investigated further by determining their expression *in vivo* and *in vitro*, exploring the potential utility of skin organoids to model disease.

Chapter 6: Overall discussion

The development of cutting-edge genomic technologies over the last few years to be able to profile individual human cells and generate a spatial map of cells within their tissue microanatomical context has provided exciting opportunities to gain unparalleled insights into human biology and disease. Indeed, spatially resolved transcriptomics was announced as "Method of the Year 2020" by Nature Methods in 2021 (Marx, 2021), highlighting the impact of this technology for the wider research community.

The combined analysis of single cell transcriptomic data, which provides a detailed characterisation of the molecular properties of individual cells, with spatial transcriptomic Visium data allows the location of cell states within different tissues to be determined. A recent scRNA-seq study conducted within our research group characterised the perturbations occurring in adult psoriasis and eczema lesional and non-lesional skin compared to healthy skin and inferred intercellular interactions mediating disease pathogenesis (Reynolds et al., 2021). However, information about the microanatomical location of individual cell states within skin tissues was lost in the generation of single cell suspensions for profiling. The study presented in Chapter 3 of this thesis was therefore carried out with the aim of addressing the biological question of how adult healthy skin cell states are organised *in situ* and how this becomes perturbed in inflammatory skin disease.

A comprehensive spatial cell atlas of healthy, psoriasis lesional, psoriasis non-lesional, eczema lesional and eczema non-lesional skin was generated and novel computational approaches were developed to analyse the spatial transcriptomic data, allowing comparisons of cell location and cell microenvironments between conditions to be made. The results uncovered perturbations in spatial cell organisation within disease states that involved various different cell lineages, including vascular endothelial cells, antigen presenting cells, and innate and adaptive lymphoid cells. Vascular endothelial cell subtypes 2 and 3, which were shown in the single-cell analyses to highly express *ACKR1* and interact with macrophages via *CXCL8* to mediate angiogenesis and lymphocyte recruitment (Reynolds et al., 2021), were found in the spatial cell atlas to be enriched within the

superficial dermis of psoriasis lesional, psoriasis non-lesional and eczema lesional skin compared to healthy skin. Furthermore, dendritic cells, cytotoxic T cells and T helper cells were shown in healthy skin to be distributed throughout the dermis, however, these cells became significantly enriched within the superficial dermis of psoriasis and eczema lesional skin. Skin cell states were also found to be microanatomically arranged into discrete functional skin microenvironments in adult healthy skin, which became altered in psoriasis and eczema lesional and non-lesional skin. For example, a spatial cellular microenvironment containing adaptive lymphoid cells was identified in healthy skin to which other immune cells migrated to in disease, including ILCs and Langerhans cell subtype 1 in psoriasis and eczema lesional skin. Future work is required to systematically investigate signalling occurring between cells within the same microenvironment to further our understanding of disease pathogenesis and potentially identify novel therapeutic targets. Data generated using Visium is not of single cellular resolution and intercellular interactions are currently inferred from scRNA-seq data, which can predict receptor ligand signalling occurring between two microanatomically distant cell states. Investigating cellular signalling within distinct functional microenvironments could be achieved by spatially profiling tissues at cellular or subcellular resolution. New technologies have recently been developed with the aim of generating highly resolved spatial transcriptomic data, including Seq-scope and Stereo-seq, which both use spatial barcoding to generate unbiased transcriptome profiles at 0.5-0.8μm and 0.22μm resolution, respectively (Cho et al., 2021a; Chen et al., 2022). High-definition (HD) spatial transcriptomics has also recently been developed, which profiles tissues at 2μm resolution (Vickovic et al., 2019); this platform is due to become commercially available as Visium HD over the coming year and leveraging this technology to investigate healthy and diseased skin tissue will allow us to further characterise the molecular mediators underpinning different functional skin microenvironments.

The importance of open access data sharing in biomedical research is becoming increasingly recognised and valued, helping to facilitate interdisciplinary collaborations and accelerate translational research (Armeni et al., 2021). In addition to ensuring that all data from this research study is made publicly available to the wider research community, I plan to work with research software engineers in our group to launch the spatial transcriptomic data in an interactive browsable web atlas portal (Li et al., 2023b). This web

atlas will allow users to visualise H&E-stained skin tissue sections along with corresponding spatial gene expression *in situ* and the locations of mapped cell states. Users will also be able to zoom in to specific tissue regions of interest. The web atlas portal will broaden the reach and utility of the data by removing the technical burden of handling large genomic datasets and allowing investigators, clinicians and students to interact with the data (Horsfall et al., 2023). This approach has potential future implications for clinical practice by revolutionising dermatopathological microscopic visuospatial examination of cellular disorganisation from skin biopsies for diagnostic benefit. Spatial transcriptomic profiling and examination of gene expression patterns in skin tissue from diseased sites also has the potential to inform the development of prognostic and therapeutic biomarkers for disease endotype stratification.

By spatially contextualising the single cell analysis of adult skin samples, distinct functional skin microenvironments underlying healthy and diseased skin states were identified. Furthermore, the scRNA-seq data revealed common functional gene programs between inflammatory skin disease and prenatal skin between 7 and 10 PCW (Reynolds et al., 2021). The study in presented in Chapter 4 of this thesis therefore aimed to investigate whether functional skin cellular microenvironments are also identified in prenatal skin and to further our understanding of the spatial organisation of individual cell states during skin morphogenesis *in utero*. A comprehensive prenatal skin single cell atlas between 7 and 16 PCW was generated, during which skin architecture matures, the epidermis becomes stratified and skin appendages within the dermis begin to form (Hu et al., 2018). This study is the first to profile both immune and non-immune cell states in prenatal skin during this period in development. Microanatomical cellular organisation was then investigated by using a deconvolution analysis to map the scRNA-seq data to spatial transcriptomic data of skin regions within human embryonic limb (Zhang et al., 2022), which revealed distinct functional skin microenvironments present in prenatal skin. These microenvironments were found to relate to specific developmental processes, including interfollicular epidermal differentiation, the early stages of hair follicle formation and neurovascular genesis. The microenvironment related to interfollicular epidermal differentiation was shown to contain immature epidermal cells and ILC3 cells. Further intercellular communication analyses showed that ILC3s in prenatal skin were predicted to interact with basal epidermal cells via *AREG-EGFR* to regulate keratinocyte proliferation and differentiation and *TNF-TNFRSF1A*, involved in cell survival and apoptosis. ILC3s are

known to contribute to excessive keratinocyte turnover and to be implicated in psoriasis pathogenesis. The findings from this study therefore reveals genes programmes deployed by ILC3 cells during prenatal skin morphogenesis that may also be implicated in psoriasis, lending further support to the concept that developmental programs are co-opted in disease (Reynolds et al., 2021).

The skin microanatomical niches comprising of early follicular cell states was investigated next in order to further our understanding of the cellular crosstalk between mesenchymal and epithelial cells during human hair follicle formation, which is presented in Chapter 5 of this thesis. The prenatal skin cell atlas extends to 16 PCW, which covers the gestational period during which hair follicles first develop *de novo*. This study represents the first single cell characterisation of human hair follicle neogenesis; our understanding of the complex processes underlying hair follicle formation have largely been inferred from murine studies (Ge et al., 2020; Saxena, Mok and Rendl, 2019), and human studies to date have been limited to the study of cycling adult hair follicles (Hsu, Pasolli and Fuchs, 2011). The scRNA-seq dataset included placode and pre-dermal condensate cells, which have not previously been characterised in human skin before, and defined the trajectories through which they differentiate into mature follicular cells. This study also provides further insights into the origin of the companion layer, which appears to develop along the same trajectory as the outer root sheath, and not from the hair matrix, as previously reported (Morita et al., 2021; Saxena, Mok and Rendl, 2019). A recent murine fate-mapping study also showed that the companion layer develops prior to matrix formation and in the absence of matrix cells (Mesler et al., 2017). In addition, a recent single cell analysis of mouse hair follicles identified greater transcriptional similarity of companion layer cells to outer root sheath cells rather than inner root sheath cells (Joost et al., 2020), consistent with the findings in human prenatal skin. The prenatal skin cell atlas was also leveraged to further our understanding of the pathogenesis of hair diseases. Interestingly, genes implicated in congenital hair diseases and androgenetic alopecia were found to be expressed along hair follicle differentiation pseudotimes, suggesting that these disorders result from dysfunctional hair follicle development. Furthermore, similar patterns of gene expression implicated in hair diseases and several subtypes of epidermolysis bullosa and congenital ichthyosis were observed in equivalent prenatal and skin organoid cell states, thereby supporting the value of skin organoids as a model to study congenital hair and skin diseases. The prenatal skin dataset has been recently launched on a web atlas portal and includes a

database of skin and hair diseases that I compiled, allowing users to search for the expression of disease-associated genes in prenatal skin (https://developmental.cellatlas.io/fetal-skin; password: fs2023). Observing genes implicated in disease expressed in prenatal skin supports an *in utero* onset for these disorders.

To conclude, these comprehensive spatially-resolved atlases of human skin in adulthood across health and disease, and during prenatal skin and hair follicle development, has provided unprecedented microanatomical detail and insight into functional skin cell microenvironments underpinning skin development and disease. This has future applications for enhancing our understanding of skin and hair disease pathogenesis, predicting the impact of immunotherapies and helping to identify novel diagnostic and prognostic biomarkers and treatment targets. These atlases can also serve as a valuable blueprint to guide faithful *in vitro* skin organoid studies, enabling future skin and hair engineering studies for therapeutic applications, including disease modelling, hair regeneration and skin transplant.

References

Abbasi, S. et al. (2021). Distinct regulatory programs control the latent regenerative potential of dermal fibroblasts during wound healing. *Cell stem cell*, 28 (3), pp.581–583.

Abdelaal, N. H. et al. (2020). Evaluation of the expression of the stromal cell-derived factor-1 alpha (CXCL 12) in psoriatic patients after treatment with Methotrexate. *Journal of cosmetic dermatology*, 19 (1), pp.253–258.

Akiyama, M. (2022). Understanding immune profiles in ichthyosis may lead to novel therapeutic targets. *The Journal of allergy and clinical immunology*, 149 (4), pp.1210–1212.

Akiyama, M., Smith, L. T. and Shimizu, H. (2000). Changing patterns of localization of putative stem cells in developing human hair follicles. *The journal of investigative dermatology*, 114 (2), pp.321–327.

Albanesi, C. et al. (2018). The interplay between keratinocytes and immune cells in the pathogenesis of psoriasis. *Frontiers in immunology*, 9, p.1549.

Ali, N. et al. (2017). Regulatory T cells in skin facilitate epithelial stem cell differentiation. *Cell*, 169 (6), pp.1119-1129.e11.

Alkon, N. et al. (2022). Single-cell analysis reveals innate lymphoid cell lineage infidelity in atopic dermatitis. *The journal of allergy and clinical immunology*, 149 (2), pp.624–639.

Almansa, R. et al. (2012). Critical COPD respiratory illness is linked to increased transcriptomic activity of neutrophil proteases genes. *BMC research notes*, 5 (1), p.401.

Alonso, L. and Fuchs, E. (2006). The hair cycle. *Journal of cell science*, 119 (3), pp.391–393.

Al-Owain, M. et al. (2011). Novel homozygous mutation in DSP causing skin fragility-woolly hair syndrome: report of a large family and review of the desmoplakin-related phenotypes. *Clinical genetics*, 80 (1), pp.50–58.

Armeni, K. et al. (2021). Towards wide-scale adoption of open science practices: The role of open science communities. *Science & public policy*, 48 (5), pp.605–611.

Ayre, D. C. et al. (2016). Analysis of the structure, evolution, and expression of CD24, an important regulator of cell fate. *Gene*, 590 (2), pp.324–337.

Ayub, M. et al. (2009). A homozygous nonsense mutation in the human desmocollin-3 (DSC3) gene underlies hereditary hypotrichosis and recurrent skin vesicles. *The American Journal of Human Genetics*, 85 (4), pp.515–520.

Balañá, M. E., Charreau, H. E. and Leirós, G. J. (2015). Epidermal stem cells and skin tissue engineering in hair follicle regeneration. *World journal of stem cells*, 7 (4), pp.711–727.

Bartoccioni, E. et al. (2003). IL-6, monocyte infiltration and parenchymal cells. *Trends in immunology*, 24 (6), Elsevier BV., pp.299–300; author reply 300-1.

Basałygo, M. et al. (2021). Assessment of serum concentrations of matrix metalloproteinase 1, matrix metalloproteinase 2 and tissue inhibitors of metalloproteinases 1 in atopic dermatitis in correlation with disease severity and epidermal barrier parameters. *Postepy dermatologii i alergologii*, 38 (5), pp.773–779.

Basingab, F. et al. (2022). Alterations in immune-related defensin Alpha 4 (DEFA4) gene expression in health and disease. *International journal of inflammation*, 2022, p.9099136.

Beckel-Mitchener, A. C. et al. (2002). Poly(A) tail length-dependent stabilization of GAP-43 mRNA by the RNA-binding protein HuD. *The journal of biological chemistry*, 277 (31), pp.27996–28002.

Belmadani, A. et al. (2009). The chemokine SDF-1/CXCL12 regulates the migration of melanocyte progenitors in mouse hair follicles. *Differentiation; research in biological diversity*, 77 (4), pp.395–411.

Bernink, J. H. et al. (2015). Interleukin-12 and -23 control plasticity of CD127(+) group 1 and group 3 innate lymphoid cells in the intestinal lamina propria. *Immunity*, 43 (1), pp.146–160.

Bian, Z. et al. (2020). Deciphering human macrophage development at single-cell resolution. *Nature*, 582 (7813), pp.571–576.

Bieber, T. (2010). Atopic dermatitis. *Annals of dermatology*, 22 (2), pp.125–137.

Bikle, D. D. (2012). Vitamin D and the skin: Physiology and pathophysiology. *Reviews in endocrine & metabolic disorders*, 13 (1), pp.3–19.

Bissig, C., Rochin, L. and van Niel, G. (2016). PMEL amyloid fibril formation: The bright steps of pigmentation. *International journal of molecular sciences*, 17 (9), p.1438.

Blanpain, C. and Fuchs, E. (2009). Epidermal homeostasis: a balancing act of stem cells in the skin. *Nature reviews. Molecular cell biology*, 10 (3), pp.207–217.

Boguslawska, J. et al. (2016). Expression of genes involved in cellular adhesion and extracellular matrix remodeling correlates with poor survival of patients with renal cancer. *The journal of urology*, 195 (6), pp.1892–1902.

Bornhauser, B. et al. (2020). The hematopoietic stem cell marker VNN2 is associated with chemoresistance in pediatric B-cell precursor ALL. *Blood advances*, 4 (17), pp.4052–4064.

Bos, J. D. et al. (1983). Immunocompetent cells in psoriasis. *Archives of dermatological research*, 275 (3), pp.181–189.

Bosè, F. et al. (2013). Inhibition of CCR7/CCL19 axis in lesional skin is a critical event for clinical remission induced by TNF blockade in patients with psoriasis. *The American journal of pathology*, 183 (2), pp.413–421.

Botchkarev, V. A. and Fessing, M. Y. (2005). Edar signaling in the control of hair follicle development. *The journal of investigative dermatology. Symposium proceedings*, 10 (3), pp.247–251.

Brandão, M. et al. (2023). Desmoplakin cardiomyopathy: Comprehensive review of an increasingly recognized entity. *Journal of clinical medicine*, 12 (7). [Online]. Available at: doi:10.3390/jcm12072660.

Breathnach, A. S. and Smith, J. (1968). Fine structure of the early hair germ and dermal papilla in the human foetus. *Journal of anatomy*, 102 (Pt 3), pp.511–526.

Brenner, M. and Hearing, V. J. (2008). The protective role of melanin against UV damage in human skin. *Photochemistry and photobiology*, 84 (3), pp.539–549.

Bronicki, L. M. and Jasmin, B. J. (2013). Emerging complexity of the HuD/ELAVl4 gene; implications for neuronal development, function, and dysfunction. *RNA (New York, N.Y.)*, 19 (8), pp.1019–1037.

Brooke, M. A., Nitoiu, D. and Kelsell, D. P. (2012). Cell-cell connectivity: desmosomes and disease. *The Journal of pathology*, 226 (2), pp.158–171.

Brown, S. J. (2017). Evidence-based medicine for atopic eczema: identifying the knowns and unknowns. *The British journal of dermatology*, 176 (4), Wiley., pp.842–844.

Brown, S. J. and McLean, W. H. I. (2012). One remarkable molecule: filaggrin. *Journal of Investigative Dermatology*, 132 (3), pp.751–762.

Brüggen, M.-C. et al. (2016). In situ mapping of innate lymphoid cells in human skin: Evidence for remarkable differences between normal and inflamed skin. *The journal of investigative dermatology*, 136 (12), pp.2396–2405.

Brunet, J.-P. et al. (2004). Metagenes and molecular pattern discovery using matrix factorization. *Proceedings of the National Academy of Sciences of the United States of America*, 101 (12), pp.4164–4169.

Brunner, P. M. et al. (2017). Increasing comorbidities suggest that atopic dermatitis is a systemic disorder. *Journal of Investigative Dermatology*, 137 (1), pp.18–25.

Bugaut, H. and Aractingi, S. (2021). Major role of the IL17/23 axis in psoriasis supports the development of new targeted therapies. *Frontiers in immunology*, 12, p.621956.

Calà, G. et al. (2023). Primary human organoids models: Current progress and key milestones. *Frontiers in bioengineering and biotechnology*, 11, p.1058970.

Calicchio, M. L., Collins, T. and Kozakewich, H. P. (2009). Identification of signaling systems in proliferating and involuting phase infantile hemangiomas by genome-wide transcriptional profiling. *The American journal of pathology*, 174 (5), pp.1638–1649.

Cammarata-Scalisi, F. et al. (2019). Novel clinical features associated with Clouston syndrome. *International journal of dermatology*, 58 (8), pp.e143–e146.

Capone, A. and Volpe, E. (2020). Transcriptional regulators of T helper 17 cell differentiation in health and autoimmune diseases. *Frontiers in immunology*, 11, p.348.

Castillo, R. L. et al. (2023). Spatial transcriptomics stratifies psoriatic disease severity by emergent cellular ecosystems. *Science immunology*, 8 (84), p.eabq7991.

Cederquist, G. Y. et al. (2012). An inherited TUBB2B mutation alters a kinesin-binding site and causes polymicrogyria, CFEOM and axon dysinnervation. *Human molecular genetics*, 21 (26), pp.5484–5499.

Chen, A. et al. (2021a). Large field of view-spatially resolved transcriptomics at nanoscale resolution. *bioRxiv*.

Chen, A. et al. (2022). Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell*, 185 (10), pp.1777-1792.e21.

Chen, C.-L. et al. (2020a). Functional complexity of hair follicle stem cell niche and therapeutic targeting of niche dysfunction for hair regeneration. *Journal of biomedical science*, 27 (1), p.43.

Chen, E. Y. et al. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC bioinformatics*, 14 (1), p.128.

Chen, H.-B. et al. (2021b). DAPL1 is a novel regulator of testosterone production in Leydig cells of mouse testis. *Scientific reports*, 11 (1), p.18532.

Chen, J. et al. (2016). Expression and function of the epidermal growth factor receptor in physiology and disease. *Physiological reviews*, 96 (3), pp.1025–1069.

Chen, Y.-P. et al. (2020b). Single-cell transcriptomics reveals regulators underlying immune cell diversity and immune subtypes associated with prognosis in nasopharyngeal carcinoma. *Cell research*, 30 (11), pp.1024–1042.

Chen, Z. et al. (2020c). The vascular endothelial cell-expressed prion protein doppel promotes angiogenesis and blood-brain barrier development. *Development (Cambridge, England)*, 147 (18), p.dev.193094.

Chiang, C.-C. et al. (2019). Neutrophils in psoriasis. *Frontiers in immunology*, 10, p.2376.

Cho, C.-S. et al. (2021a). Microscopic examination of spatial transcriptome using Seq-Scope. *Cell*, 184 (13), pp.3559-3572.e22.

Cho, C.-S. et al. (2021b). Seq-Scope: Submicrometer-resolution spatial transcriptomics for single cell and subcellular studies. *bioRxiv*.

Chong, M. and Fonacier, L. (2016). Treatment of eczema: corticosteroids and beyond. *Clinical reviews in allergy & immunology*, 51 (3), pp.249–262.

Choong, D. J. and Tan, E. (2021). Does tocilizumab have a role in dermatology? A review of clinical applications, its adverse side effects and practical considerations. *Dermatologic therapy*, 34 (4), p.e14990.

Chu, C.-Y. (2021). Treatments for childhood atopic dermatitis: An update on emerging therapies. *Clinical reviews in allergy & immunology*, 61 (2), pp.114–127.

Cichorek, M. et al. (2013). Skin melanocytes: biology and development. *Postepy dermatologii i alergologii*, 30 (1), pp.30–41.

Cipriani, F. et al. (2014). Autoimmunity in atopic dermatitis: biomarker or simply epiphenomenon? *The journal of dermatology*, 41 (7), pp.569–576.

Clark, R. A. and Kupper, T. S. (2006). Misbehaving macrophages in the pathogenesis of psoriasis. *The journal of clinical investigation*, 116 (8), pp.2084–2087.

Clayton, K. et al. (2017). Langerhans cells-programmed by the epidermis. *Frontiers in immunology*, 8, p.1676.

Coates, M., Blanchard, S. and MacLeod, A. S. (2018). Innate antimicrobial immunity in the skin: A protective barrier against bacteria, viruses, and fungi. *PLoS pathogens*, 14 (12), p.e1007353.

Collins, C. A. and Watt, F. M. (2008). Dynamic regulation of retinoic acid-binding proteins in developing, adult and neoplastic skin reveals roles for beta-catenin and Notch signalling. *Developmental biology*, 324 (1), pp.55–67.

Coolen, N. A. et al. (2010). Comparison between human fetal and adult skin. *Archives of dermatological research*, 302 (1), pp.47–55.

Cox, J. R., Cruickshank, S. M. and Saunders, A. E. (2021). Maintenance of Barrier Tissue Integrity by Unconventional Lymphocytes. *Frontiers in immunology*, 12, p.670471.

Cristancho, A. G. and Lazar, M. A. (2011). Forming functional fat: a growing understanding of adipocyte differentiation. *Nature reviews. Molecular cell biology*, 12 (11), pp.722–734.

Crouch, E. E. et al. (2023). Disentangling brain vasculature in neurogenesis and neurodegeneration using single-cell transcriptomics. *Trends in neurosciences*. [Online]. Available at: doi:10.1016/j.tins.2023.04.007.

Czernielewski, J. et al. (1985). Langerhans' cells in patients with psoriasis: effect of treatment with PUVA, PUVA bath, etretinate and anthralin. *Acta dermato-venereologica*, 65 (2), pp.97–101.

Damen, M. et al. (2021). High proliferation and delamination during skin epidermal stratification. *Nature communications*, 12 (1), p.3227.

Danilenko, D. M., Ring, B. D. and Pierce, G. F. (1996). Growth factors and cytokines in hair follicle development and cycling: recent insights from animal models and the potentials for clinical therapy. *Molecular medicine today*, 2 (11), pp.460–467.

Dann, E. et al. (2022). Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nature biotechnology*, 40 (2), pp.245–253.

Davison, N. J. et al. (2017). Identification of factors that may influence the selection of first-line biological therapy for people with psoriasis: a prospective, multicentre cohort study. *British Journal of Dermatology*, 177 (3), pp.828–836. [Online]. Available at: doi:10.1111/bjd.15551.

De Rie, M. A., Goedkoop, A. Y. and Bos, J. D. (2004). Overview of psoriasis. *Dermatologic therapy*, 17 (5), pp.341–349.

De Rosa, G. and Mignogna, C. (2007). The histopathology of psoriasis. *Reumatismo*, 59 Suppl 1, pp.46–48.

Deckers, J., Hammad, H. and Hoste, E. (2018). Langerhans cells: Sensing the environment in health and disease. *Frontiers in immunology*, 9. [Online]. Available at: doi:10.3389/fimmu.2018.00093.

Delfín, D. A., DeAguero, J. L. and McKown, E. N. (2019). The extracellular matrix protein ABI3BP in cardiovascular health and disease. *Frontiers in cardiovascular medicine*, 6, p.23.

Delom, F. et al. (2020). The anterior gradient-2 interactome. *American journal of physiology. Cell physiology*, 318 (1), pp.C40–C47.

Denzer, L. et al. (2023). The role of PLVAP in endothelial cells. *Cell and tissue research*, 392 (2), pp.393–412.

Deshmukh, S. and Prashanth, S. (2012). Ectodermal dysplasia: a genetic review. *International journal of clinical pediatric dentistry*, 5 (3), pp.197–202.

Detmar, M. et al. (1995). Keratinocyte-derived vascular permeability factor (vascular endothelial growth factor) is a potent mitogen for dermal microvascular endothelial cells. *The journal of investigative dermatology*, 105 (1), pp.44–50.

Dhariwala, M. O. et al. (2020). Developing human skin contains lymphocytes demonstrating a memory signature. *Cell reports. Medicine*, 1 (8), p.100132.

Dillenburg-Pilla, P. et al. (2015). SDF-1/CXCL12 induces directional cell migration and spontaneous metastasis via a CXCR4/Gαi/mTORC1 axis. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 29 (3), pp.1056–1068.

Ding, J. et al. (2020). Author Correction: Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. *Nature biotechnology*, 38 (6), p.756.

Di-Po, N. et al. (2003). The anti-apoptotic role of PPARβ contributes to efficient skin wound healing. *The Journal of steroid biochemistry and molecular biology*, 85 (2–5), pp.257–265.

Donovan, A. et al. (2005). The iron exporter ferroportin/Slc40a1 is essential for iron homeostasis. *Cell metabolism*, 1 (3), pp.191–200.

Driskell, R. R. et al. (2013). Distinct fibroblast lineages determine dermal architecture in skin development and repair. *Nature*, 504 (7479), pp.277–281.

Dubrac, S., Schmuth, M. and Ebner, S. (2010). Atopic dermatitis: the role of Langerhans cells in disease pathogenesis. *Immunology and cell biology*, 88 (4), pp.400–409.

Durante, M. A. et al. (2020). Single-cell analysis reveals new evolutionary complexity in uveal melanoma. *Nature communications*, 11 (1), p.496.

Eaton, L. H. et al. (2014). Guttate psoriasis is associated with an intermediate phenotype of impaired Langerhans cell migration. *The British journal of dermatology*, 171 (2), pp.409–411.

Eaton, L. H. et al. (2018). Impaired Langerhans cell migration in psoriasis is due to an altered keratinocyte phenotype induced by interleukin-17. *The British journal of dermatology*, 178 (6), pp.1364–1372.

Eckert, E. (1991). Histopathological and Immunohistological Aspects of Atopic Eczema. In: *Handbook of Atopic Eczema*. Berlin, Heidelberg: Springer Berlin Heidelberg. pp.127–131.

Efremova, M. et al. (2020). CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nature protocols*, 15 (4), pp.1484–1506.

Egeberg, A. et al. (2018). Safety, efficacy and drug survival of biologics and biosimilars for moderate-to-severe plaque psoriasis. *The British journal of dermatology*, 178 (2), pp.509–519.

Elentner, A. et al. (2009). Langerhans cells are critical in the development of atopic dermatitis-like inflammation and symptoms in mice. *Journal of cellular and molecular medicine*, 13 (8B), pp.2658–2672.

Elmentaite, R. et al. (2021). Cells of the human intestinal tract mapped across space and time. *Nature*, 597 (7875), pp.250–255.

Eraslan, G. et al. (2022). Single-nucleus cross-tissue molecular reference maps toward understanding disease gene function. *Science (New York, N.Y.)*, 376 (6594), p.eabl4290.

Ersch, J. and Stallmach, T. (1999). Assessing gestational age from histology of fetal skin: an autopsy study of 379 fetuses. *Obstetrics and gynecology*, 94 (5 Pt 1), pp.753–757.

Fan, S. M.-Y. et al. (2018). Inducing hair follicle neogenesis with secreted proteins enriched in embryonic skin. *Biomaterials*, 167, pp.121–131.

Feinstein, Y. et al. (1999). F-spondin and mindin: two structurally and functionally related genes expressed in the hippocampus that promote outgrowth of embryonic hippocampal neurons. *Development (Cambridge, England)*, 126 (16), pp.3637–3648.

Fernandez-Flores, A. (2015). Regional variations in the histology of the skin. *The American journal of dermatopathology*, 37 (10), pp.737–754.

Fine, J.-D. et al. (2008). The classification of inherited epidermolysis bullosa (EB): Report of the Third International Consensus Meeting on Diagnosis and Classification of EB. *Journal of the American Academy of Dermatology*, 58 (6), pp.931–950.

Fowler, J. F. et al. (2008). The impact of psoriasis on health care costs and patient work loss. *Journal of the American Academy of Dermatology*, 59 (5), pp.772–780.

Fu, X. et al. (2021). Continuous Polony Gels for Tissue Mapping with High Resolution and RNA Capture Efficiency. *bioRxiv*.

Fujita, H. et al. (2011). Lesional dendritic cells in patients with chronic atopic dermatitis and psoriasis exhibit parallel ability to activate T-cell subsets. *The journal of allergy and clinical immunology*, 128 (3), pp.574-82.e1-12.

Gao, Y. et al. (2021). Single-cell analysis reveals the heterogeneity of monocyte-derived and peripheral type-2 conventional dendritic cells. *The journal of immunology*, 207 (3), pp.837–848. [Accessed 8 June 2023].

Gaujoux, R. and Seoighe, C. (2010). A flexible R package for nonnegative matrix factorization. *BMC bioinformatics*, 11, p.367.

Ge, W. et al. (2020). Single-cell Transcriptome Profiling reveals Dermal and Epithelial cell fate decisions during Embryonic Hair Follicle Development. *Theranostics*, 10 (17), pp.7581–7598.

Gillan, L. et al. (2002). Periostin secreted by epithelial ovarian carcinoma is a ligand for alpha(V)beta(3) and alpha(V)beta(5) integrins and promotes cell motility. *Cancer research*, 62 (18), pp.5358–5364.

Glitzner, E. et al. (2014). Specific roles for dendritic cell subsets during initiation and progression of psoriasis. *EMBO molecular medicine*, 6 (10), pp.1312–1327.

Gommans, J. M., van Hezik, S. J. and van Huystee, B. E. (1987). Flow cytometric quantification of T6-positive cells in psoriatic epidermis after PUVA and methotrexate therapy. *The British journal of dermatology*, 116 (5), pp.661–666.

Gonzalez-Quereda, L. et al. (2018). A new mutation of the SCGA gene is the cause of a late onset mild phenotype limb girdle muscular dystrophy type 2D with axial involvement. *Neuromuscular disorders: NMD*, 28 (8), pp.633–638.

Goto, M. et al. (2006). Fibroblasts show more potential as target cells than keratinocytes in COL7A1 gene therapy of dystrophic epidermolysis bullosa. *The Journal of investigative dermatology*, 126 (4), pp.766–772.

Gottlieb, A. B., Chao, C. and Dann, F. (2008). Psoriasis comorbidities. *Journal of Dermatological Treatment*, 19 (1), pp.5–21.

Gren, S. T. et al. (2015). A single-cell gene-expression profile reveals inter-cellular heterogeneity within human monocyte subsets. *PloS one*, 10 (12), p.e0144351.

Griffiths, C. E. M. et al. (2017). The global state of psoriasis disease epidemiology: a workshop report. *The British journal of dermatology*, 177 (1), pp.e4–e7.

Grisanti, L. et al. (2013). Tbx18 targets dermal condensates for labeling, isolation, and gene ablation during embryonic hair follicle formation. *The journal of investigative dermatology*, 133 (2), pp.344–353.

de Groot, S. C. et al. (2021). Back to the future: From appendage development toward future human hair follicle neogenesis. *Frontiers in cell and developmental biology*, 9, p.661787.

Grossman, R. M. et al. (1989). Interleukin 6 is expressed in high levels in psoriatic skin and stimulates proliferation of cultured human keratinocytes. *Proceedings of the National Academy of Sciences of the United States of America*, 86 (16), pp.6367–6371.

Gu, L.-H. and Coulombe, P. A. (2007). Keratin expression provides novel insight into the morphogenesis and function of the companion layer in hair follicles. *The journal of investigative dermatology*, 127 (5), pp.1061–1073.

Gu, Z. et al. (2014). circlize Implements and enhances circular visualization in R. *Bioinformatics (Oxford, England)*, 30 (19), pp.2811–2812.

Guo, T. et al. (2020). Comparative proteomics reveals genetic mechanisms underlying secondary hair follicle development in fine wool sheep during the fetal stage. *Journal of proteomics*, 223 (103827), p.103827.

Gupta, K. et al. (2019). Single-cell analysis reveals a hair follicle dermal niche molecular differentiation trajectory that begins prior to morphogenesis. *Developmental cell*, 48 (1), pp.17-31.e6.

Gupta, P. et al. (2006). A chronic contact eczema impedes migration of antigen-presenting cells in alopecia areata. *The journal of investigative dermatology*, 126 (7), pp.1559–1573.

Gupta, R. et al. (2004). Burden of allergic disease in the UK: secondary analyses of national databases. *Clinical & Experimental Allergy*, 34 (4), pp.520–526.

Gyllborg, D. and Nilsson, M. (2019). *HybISS: Hybridization-based In Situ Sequencing v1*. [Online]. Available at: doi:10.17504/protocols.io.xy4fpyw.

Hagner, A. et al. (2020). Transcriptional profiling of the adult hair follicle mesenchyme reveals R-spondin as a novel regulator of dermal progenitor function. *iScience*, 23 (4), p.101019.

Han, Y. P. et al. (2001). TNF-alpha stimulates activation of pro-MMP2 in human skin through NF-(kappa)B mediated induction of MT1-MMP. *Journal of cell science*, 114 (Pt 1), pp.131–139.

Haniffa, M. et al. (2021). A roadmap for the Human Developmental Cell Atlas. *Nature*, 597 (7875), pp.196–205.

Hao, Y. et al. (2023). Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nature biotechnology*. [Online]. Available at: doi:10.1038/s41587-023-01767-y.

Hardman, M. J. et al. (1999). Barrier formation in the human fetus is patterned. *The journal of investigative dermatology*, 113 (6), pp.1106–1113.

Hardy, M. H. (1992). The secret life of the hair follicle. *Trends in genetics: TIG*, 8 (2), pp.55–61.

Harland, D. P. et al. (2014). Three-dimensional architecture of macrofibrils in the human scalp hair cortex. *Journal of structural biology*, 185 (3), pp.397–404.

Has, C. et al. (2020). Consensus reclassification of inherited epidermolysis bullosa and other disorders with skin fragility. *The British journal of dermatology*, 183 (4), pp.614–627.

Hawkes, J. E., Chan, T. C. and Krueger, J. G. (2017). Psoriasis pathogenesis and the development of novel targeted immune therapies. *Journal of Allergy and Clinical Immunology*, 140 (3), pp.645–653.

Hawkshaw, N. J. et al. (2020). Deciphering the molecular morphology of the human hair cycle: Wnt signalling during the telogen-anagen transformation. *The British journal of dermatology*, 182 (5), pp.1184–1193.

Hay, R. J. et al. (2014). The global burden of skin disease in 2010: an analysis of the prevalence and impact of skin conditions. *Journal of Investigative Dermatology*, 134 (6), pp.1527–1534.

Hay, S. I. et al. (2017). Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, 390 (10100), pp.1260–1344.

Heaton, H. et al. (2020). Souporcell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. *Nature methods*, 17 (6), pp.615–620.

Heidenreich, R., Röcken, M. and Ghoreschi, K. (2009). Angiogenesis drives psoriasis pathogenesis. *International journal of experimental pathology*, 90 (3), pp.232–248.

Heilmann-Heimbach, S. et al. (2016). Hunting the genes in male-pattern alopecia: how important are they, how close are we and what will they tell us? *Experimental dermatology*, 25 (4), pp.251–257.

Horiguchi, M., Ota, M. and Rifkin, D. B. (2012). Matrix control of transforming growth factor-β function. *Journal of biochemistry*, 152 (4), pp.321–329.

Horsfall, D. et al. (2023). Research software engineering accelerates the translation of biomedical research for health. *Nature medicine*, 29 (6), pp.1313–1316.

Houck, G. et al. (2004). Eczema and the spongiotic dermatoses: a histologic and pathogenic update. *Seminars in cutaneous medicine and surgery*, 23 (1), pp.39–45.

Hsiao, K.-C. et al. (2020). Elevated expression of lumican in lung cancer cells promotes bone metastasis through an autocrine regulatory mechanism. *Cancers*, 12 (1), p.233.

Hsu, Y.-C., Pasolli, H. A. and Fuchs, E. (2011). Dynamics between stem cells, niche, and progeny in the hair follicle. *Cell*, 144 (1), pp.92–105.

Hu, M. S. et al. (2018). Embryonic skin development and repair. *Organogenesis*, 14 (1), pp.46–63.

Hu, P. et al. (2021). The role of helper T cells in psoriasis. *Frontiers in immunology*, 12, p.788940.

Huitema, L. et al. (2021). Immunological mechanisms underlying progression of chronic wounds in recessive dystrophic epidermolysis bullosa. *Experimental dermatology*, 30 (12), pp.1724–1733.

Ihrie, R. A. et al. (2005). Perp is a p63-regulated gene essential for epithelial integrity. *Cell*, 120 (6), pp.843–856.

Iso, T. et al. (2013). Capillary endothelial fatty acid binding proteins 4 and 5 play a critical role in fatty acid uptake in heart and skeletal muscle. *Arteriosclerosis, thrombosis, and vascular biology*, 33 (11), pp.2549–2557.

Jacob, T. et al. (2021). Zebrafish twist2/dermo1 regulates scale shape and scale organization during skin development and regeneration. *Cells & development*, 166 (203684), p.203684.

Jardine, L. et al. (2021). Blood and immune development in human fetal bone marrow and Down syndrome. *Nature*, 598 (7880), pp.327–331.

Ji, A. et al. (2019). 136 Dissecting intratumoral heterogeneity and microenvironment interactions in SCC through single-cell RNA-sequencing. *The journal of investigative dermatology*, 139 (5), p.S24.

Ji, A. L. et al. (2020). Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma. *Cell*, 182 (2), pp.497-514. e22.

Jia, W., Li, H. and He, Y.-W. (2005). The extracellular matrix protein mindin serves as an integrin ligand and is critical for inflammatory cell recruitment. *Blood*, 106 (12), pp.3854–3859.

Jin, Z. et al. (2017). Single-cell gene expression patterns in lupus monocytes independently indicate disease activity, interferon and therapy. *Lupus science & medicine*, 4 (1), p.e000202.

Joost, S. et al. (2020). The Molecular Anatomy of Mouse Skin during Hair Growth and Rest. *Cell stem cell*, 26 (3), pp.441-457.e7.

Kabir, M. F. et al. (2022). Single cell transcriptomic analysis reveals cellular diversity of murine esophageal epithelium. *Nature communications*, 13 (1), p.2167.

Kalabusheva, E. P. et al. (2023). A kaleidoscope of keratin gene expression and the mosaic of its regulatory mechanisms. *International journal of molecular sciences*, 24 (6). [Online]. Available at: doi:10.3390/ijms24065603.

Kamata, M. and Tada, Y. (2022). Dendritic cells and macrophages in the pathogenesis of psoriasis. *Frontiers in immunology*, 13, p.941071.

Kapellos, T. S. et al. (2019). Human monocyte subsets and phenotypes in major chronic inflammatory diseases. *Frontiers in immunology*, 10, p.2035.

Kaplanski, G. et al. (2003). IL-6: a regulator of the transition from neutrophil to monocyte recruitment during inflammation. *Trends in immunology*, 24 (1), pp.25–29.

Kappert, K. et al. (2013). Proprotein convertase subtilisin/kexin type 3 promotes adipose tissue-driven macrophage chemotaxis and is increased in obesity. *PloS one*, 8 (8), p.e70542.

Karamitros, D. et al. (2018). Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells. *Nature immunology*, 19 (1), pp.85–97.

Khammissa, R. A. G. et al. (2018). The biological activities of vitamin D and its receptor in relation to calcium and bone homeostasis, cancer, immune and cardiovascular systems, skin biology, and oral health. *BioMed research international*, 2018, pp.1–9.

Kiekens, R. C. et al. (2001). Heterogeneity within tissue-specific macrophage and dendritic cell populations during cutaneous inflammation in atopic dermatitis. *The British journal of dermatology*, 145 (6), pp.957–965.

Kieu, T. Q. et al. (2022). Kinetics of LYVE-1-positive M2-like macrophages in developing and repairing dental pulp in vivo and their pro-angiogenic activity in vitro. *Scientific reports*, 12 (1), p.5176. [Accessed 9 June 2023].

Kim, B. S. et al. (2014a). Basophils promote innate lymphoid cell responses in inflamed skin. *The journal of immunology*, 193 (7), pp.3717–3725.

Kim, T. S. et al. (2014b). Distinct dendritic cell subsets dictate the fate decision between effector and memory CD8(+) T cell differentiation by a CD24-dependent mechanism. *Immunity*, 40 (3), pp.400–413.

Kimura, A. and Kishimoto, T. (2010). IL-6: regulator of Treg/Th17 balance. *European journal of immunology*, 40 (7), pp.1830–1835.

King, A., Balaji, S. and Keswani, S. G. (2013). Biology and Function of Fetal and Pediatric Skin. *Facial Plastic Surgery Clinics of North America*, 21 (1), pp.1–6. [Online]. Available at: doi:10.1016/j.fsc.2012.10.001.

Kleshchevnikov, V. et al. (2022). Cell2location maps fine-grained cell types in spatial transcriptomics. *Nature biotechnology*, 40 (5), pp.661–671.

Kolarsick, P. A. J., Kolarsick, M. A. and Goodwin, C. (2011). Anatomy and physiology of the skin. *Journal of the Dermatology Nurses' Association*, 3 (4), pp.203–213.

Kolodziejczyk, A. A. et al. (2015). The technology and biology of single-cell RNA sequencing. *Molecular cell*, 58 (4), pp.610–620.

Korsunsky, I. et al. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature methods*, 16 (12), pp.1289–1296.

Kortekaas Krohn, I. et al. (2022). T-cell subsets in the skin and their role in inflammatory skin disorders. *Allergy*, 77 (3), pp.827–842.

Kotalevskaya, Y. Y. and Stepanov, V. A. (2023). Molecular genetic basis of epidermolysis bullosa. *Vavilovskii zhurnal genetiki i selektsii*, 27 (1), pp.18–27.

Koyama, M. et al. (1997). CD79 alpha/CD79 beta heterodimers are expressed on pro-B cell surfaces without associated mu heavy chain. *International immunology*, 9 (11), pp.1767–1772.

Kral, M. et al. (2017). Covalent dimerization of interleukin-like epithelial-to-mesenchymal transition (EMT) inducer (ILEI) facilitates EMT, invasion, and late aspects of metastasis. *The FEBS journal*, 284 (20), pp.3484–3505.

Krishna Kumar, K. et al. (2010). AHSP (α-haemoglobin-stabilizing protein) stabilizes apo-α-haemoglobin in a partially folded state. *The Biochemical journal*, 432 (2), pp.275–282.

Krishnan, A. et al. (2012). Lumican, an extracellular matrix proteoglycan, is a novel requisite for hepatic fibrosis. *Laboratory investigation; a journal of technical methods and pathology*, 92 (12), pp.1712–1725.

Krueger, G. G. et al. (2000). Two considerations for patients with psoriasis and their clinicians:: What defines mild, moderate, and severe psoriasis? What constitutes a clinically significant improvement when treating psoriasis? *Journal of the American Academy of Dermatology*, 43 (2), pp.281–285.

Krzewski, K. and Coligan, J. E. (2012). Human NK cell lytic granules and regulation of their exocytosis. *Frontiers in immunology*, 3, p.335.

Kuett, L. et al. (2022). Three-dimensional imaging mass cytometry for highly multiplexed molecular and cellular mapping of tissues and the tumor microenvironment. *Nature cancer*, 3 (1), pp.122–133.

Kukurba, K. R. and Montgomery, S. B. (2015). RNA sequencing and analysis. *Cold Spring Harbor protocols*, 2015 (11), pp.951–969.

Kumari, S. et al. (2014). Sharpin prevents skin inflammation by inhibiting TNFR1-induced keratinocyte apoptosis. *eLife*, 3. [Online]. Available at: doi:10.7554/eLife.03422.

Kurimoto, K. et al. (2007). Global single-cell cDNA amplification to provide a template for representative high-density oligonucleotide microarray analysis. *Nature protocols*, 2 (3), pp.739–752.

Kvedar, J. C. et al. (1992). Characterization of sciellin, a precursor to the cornified envelope of human keratinocytes. *Differentiation; research in biological diversity*, 49 (3), pp.195–204.

Kwack, M. H. et al. (2019). Overexpression of alkaline phosphatase improves the hair-inductive capacity of cultured human dermal papilla spheres. *Journal of dermatological science*, 95 (3), pp.126–129.

Laborel-Préneron, E. et al. (2015). Effects of the staphylococcus aureus and staphylococcus epidermidis secretomes isolated from the skin Microbiota of atopic children on CD4+ T cell activation. *PloS one*, 10 (10), p.e0141067.

Langan, S. M. et al. (2006). An exploratory prospective observational study of environmental factors exacerbating atopic eczema in children. *The British journal of dermatology*, 154 (5), pp.979–980.

Lawrence, M. B. and Springer, T. A. (1993). Neutrophils roll on E-selectin. *The journal of immunology*, 151 (11), pp.6338–6346.

Lee, C. E. et al. (1992). Interleukin-6 is released in the cutaneous response to allergen challenge in atopic individuals. *The journal of allergy and clinical immunology*, 89 (5), pp.1010–1020.

Lee, H.-T. et al. (2017). A crucial role of CXCL14 for promoting regulatory T cells activation in stroke. *Theranostics*, 7 (4), pp.855–875.

Lee, J. et al. (2020). Hair-bearing human skin generated entirely from pluripotent stem cells. *Nature*, 582 (7812), pp.399–404.

Lee, L. and Wine, K. A. (2019). Embryogenesis of the Skin. *Pediatric Dermatology*, pp.1–35.

Leung, D. Y. M. and Guttman-Yassky, E. (2017). Assessing the current treatment of atopic dermatitis: unmet needs. *Journal of Allergy and Clinical Immunology*, 139 (4), pp.S47–S48.

Li, C. et al. (2010). Development of atopic dermatitis-like skin disease from the chronic loss of epidermal caspase-8. *Proceedings of the National Academy of Sciences of the United States of America*, 107 (51), pp.22249–22254.

Li, C.-F. et al. (2019). Snail-induced claudin-11 prompts collective migration for tumour progression. *Nature cell biology*, 21 (2), pp.251–262.

Li, H. et al. (2023a). A comprehensive benchmarking with practical guidelines for cellular deconvolution of spatial transcriptomics. *Nature communications*, 14 (1), p.1548.

Li, Q. et al. (2021). Single-cell transcriptome profiling reveals vascular endothelial cell heterogeneity in human skin. *Theranostics*, 11 (13), pp.6461–6476.

Li, T. et al. (2023b). WebAtlas pipeline for integrated single cell and spatial transcriptomic data. *bioRxiv*. [Online]. Available at: doi:10.1101/2023.05.19.541329.

Li, X. et al. (2006). Solution structures of human LL-37 fragments and NMR-based identification of a minimal membrane-targeting antimicrobial and anticancer region. *Journal of the American Chemical Society*, 128 (17), pp.5776–5785.

Li, X. and Wang, C.-Y. (2021). From bulk, single-cell to spatial RNA sequencing. *International journal of oral science*, 13 (1), p.36.

Li, Z. et al. (2016). Epidermal Notch1 recruits RORγ(+) group 3 innate lymphoid cells to orchestrate normal skin repair. *Nature communications*, 7, p.11394.

Liu, J. et al. (2021). Single-cell RNA sequencing of psoriatic skin identifies pathogenic Tc17 cell subsets and reveals distinctions between CD8+ T cells in autoimmunity and cancer. *The journal of allergy and clinical immunology*, 147 (6), pp.2370–2380.

Liu, S., Zhang, H. and Duan, E. (2013). Epidermal development in mammals: key regulators, signals from beneath, and stem cells. *International journal of molecular sciences*, 14 (6), pp.10869–10895.

Liu, X. et al. (2023). Single-cell atlas reveals a high selection of IgA1- or IgG1-expressing plasma cells in patients with psoriasis. *Molecular immunology*, 153, pp.85–93.

Liu, Y. et al. (2020). High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. *Cell*, 183 (6), pp.1665-1681. e18.

Liu, Y. T. et al. (2015). Novel mutations in GJB6 and GJB2 in Clouston syndrome. *Clinical and experimental dermatology*, 40 (7), pp.770–773.

Llaó-Cid, L. et al. (2021). EOMES is essential for antitumor activity of CD8+ T cells in chronic lymphocytic leukemia. *Leukemia*, 35 (11), pp.3152–3162.

Lolli, F. et al. (2017). Androgenetic alopecia: a review. *Endocrine*, 57 (1), pp.9–17.

Long, Y. et al. (2023). Spatially informed clustering, integration, and deconvolution of spatial transcriptomics with GraphST. *Nature communications*, 14 (1), p.1155.

Lowes, M. A. et al. (2008). Psoriasis vulgaris lesions contain discrete populations of Th1 and Th17 T cells. *The journal of investigative dermatology*, 128 (5), pp.1207–1211.

Lowes, M. A., Bowcock, A. M. and Krueger, J. G. (2007). Pathogenesis and therapy of psoriasis. *Nature*, 445 (7130), pp.866–873.

Lu, J. et al. (2016). CD19+ B cell subsets in the peripheral blood and skin lesions of psoriasis patients and their correlations with disease severity. *Brazilian Journal of Medical and Biological Research*, 49 (9), p.e5374.

Lu, X.-J. and Wang, H.-T. (2017). Reduced Gja5 expression in arterial endothelial cells impairs arteriogenesis during acute ischemic cardiovascular disease. *Experimental and therapeutic medicine*, 14 (5), pp.4339–4343.

Luecken, M. D. and Theis, F. J. (2019). Current best practices in single-cell RNA-seq analysis: a tutorial. *Molecular systems biology*, 15 (6), p.e8746.

Lun, A. (2018). Overcoming systematic errors caused by log-transformation of normalized single-cell RNA sequencing data. *bioRxiv*, bioRxiv., p.404962. [Online]. Available at: doi:10.1101/404962.

Lun, A. T. L., McCarthy, D. J. and Marioni, J. C. (2016). A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Research*, 5, p.2122.

Luschkova, D. et al. (2021). Atopic eczema is an environmental disease. *Allergologie select*, 5 (01), pp.244–250.

Ma, Y. et al. (2022). Direct reprograming of mouse fibroblasts into dermal papilla cells via small molecules. *International journal of molecular sciences*, 23 (8), p.4213.

Magnaldo, T. and Barrandon, Y. (1996). CD24 (heat stable antigen, nectadrin), a novel keratinocyte differentiation marker, is preferentially expressed in areas of the hair follicle containing the colony-forming cells. *Journal of cell science*, 109 (Pt 13) (13), pp.3035–3045.

Mahil, S. K., Capon, F. and Barker, J. N. (2016). *Update on psoriasis immunopathogenesis and targeted immunotherapy*. Springer.

Malecic, N. and Young, H. S. (2017). Excessive angiogenesis associated with psoriasis as a cause for cardiovascular ischaemia. *Experimental dermatology*, 26 (4), pp.299–304.

Maniatis, S., Petrescu, J. and Phatnani, H. (2021). Spatially resolved transcriptomics and its applications in cancer. *Current Opinion in Genetics & Development*, 66, pp.70–77.

Marcińska, M. et al. (2015). Evaluation of DNA variants associated with androgenetic alopecia and their potential to predict male pattern baldness. *PloS one*, 10 (5), p.e0127852.

Marks, J. G., Jr and Miller, J. J. (2013). Structure and function of the skin. In: *Lookingbill and Marks' Principles of Dermatology*. Elsevier. pp.2–10.

Martini, E. et al. (2017). Dynamic changes in resident and infiltrating epidermal dendritic cells in active and resolved psoriasis. *The journal of investigative dermatology*, 137 (4), pp.865–873.

Martino, P. A., Heitman, N. and Rendl, M. (2021). The dermal sheath: An emerging component of the hair follicle stem cell niche. *Experimental dermatology*, 30 (4), pp.512–521.

Martisova, A. et al. (2022). Identification of AGR2 gene-specific expression patterns associated with epithelial-mesenchymal transition. *International journal of molecular sciences*, 23 (18). [Online]. Available at: doi:10.3390/ijms231810845.

Marx, V. (2021). Method of the Year: spatially resolved transcriptomics. *Nature methods*, 18 (1), pp.9–14.

Mazereeuw-Hautier, J. et al. (2019). Management of congenital ichthyoses: European guidelines of care, part two. *The British journal of dermatology*, 180 (3), pp.484–495.

Mazzurana, L. et al. (2021). Tissue-specific transcriptional imprinting and heterogeneity in human innate lymphoid cells revealed by full-length single-cell RNA-sequencing. *Cell research*, 31 (5), pp.554–568.

McFaline-Figueroa, J. L., Trapnell, C. and Cuperus, J. T. (2020). The promise of single-cell genomics in plants. *Current opinion in plant biology*, 54, pp.114–121.

McGowan, K. M. and Coulombe, P. A. (1998). Onset of keratin 17 expression coincides with the definition of major epithelial lineages during skin development. *The journal of cell biology*, 143 (2), pp.469–486.

McGowan, K. M. and Coulombe, P. A. (2000). Keratin 17 expression in the hard epithelial context of the hair and nail, and its relevance for the pachyonychia congenita phenotype. *The journal of investigative dermatology*, 114 (6), pp.1101–1107.

Mesler, A. L. et al. (2017). Hair follicle terminal differentiation is orchestrated by distinct early and late matrix progenitors. *Cell reports*, 19 (4), pp.809–821.

Meuter, S. et al. (2007). Murine CXCL14 is dispensable for dendritic cell function and localization within peripheral tissues. *Molecular and cellular biology*, 27 (3), pp.983–992.

Michno, K. et al. (2003). Shh expression is required for embryonic hair follicle but not mammary gland development. *Developmental biology*, 264 (1), pp.153–165.

Mihu, C. et al. (2021). Novel concepts in psoriasis: histopathology and markers related to modern treatment approaches. *Revue roumaine de morphologie et embryologie [Romanian journal of morphology and embryology]*, 62 (4), pp.897–906.

Mill, P. et al. (2003). Sonic hedgehog-dependent activation of Gli2 is essential for embryonic hair follicle development. *Genes & development*, 17 (2), pp.282–294.

Millar, S. E. (2002). Molecular mechanisms regulating hair follicle development. *The journal of investigative dermatology*, 118 (2), pp.216–225.

Mirmirani, P., Huang, K. P. and Price, V. H. (2011). A practical, algorithmic approach to diagnosing hair shaft disorders. *International journal of dermatology*, 50 (1), pp.1–12.

Miyazaki, T. et al. (1999). Increased susceptibility of thymocytes to apoptosis in mice lacking AIM, a novel murine macrophage-derived soluble factor belonging to the scavenger receptor cysteine-rich domain superfamily. *The journal of experimental medicine*, 189 (2), pp.413–422.

Mohapatra, A. et al. (2016). Group 2 innate lymphoid cells utilize the IRF4-IL-9 module to coordinate epithelial cell maintenance of lung homeostasis. *Mucosal immunology*, 9 (1), pp.275–286.

Mok, K.-W. et al. (2019). Dermal condensate niche fate specification occurs prior to formation and is placode progenitor dependent. *Developmental cell*, 48 (1), pp.32-48.e5.

Moody, S. A. et al. (2013). On becoming neural: what the embryo can tell us about differentiating neural stem cells. *American journal of stem cells*, 2 (2), pp.74–94.

Moore, K. B. et al. (2004). Morphogenetic movements underlying eye field formation require interactions between the FGF and ephrinB1 signaling pathways. *Developmental cell*, 6 (1), pp.55–67.

Moreci, R. S. and Lechler, T. (2020). Epidermal structure and differentiation. *Current biology: CB*, 30 (4), pp.R144–R149.

Morgan, B. A. (2014). The dermal papilla: an instructive niche for epithelial stem and progenitor cells in development and regeneration of the hair follicle. *Cold Spring Harbor perspectives in medicine*, 4 (7), p.a015180.

Morgan, E., Kannan-Thulasiraman, P. and Noy, N. (2010). Involvement of fatty acid binding protein 5 and PPARβ/δ in prostate cancer cell growth. *PPAR research*, 2010, pp.1–9.

Morita, R. et al. (2021). Tracing the origin of hair follicle stem cells. *Nature*, 594 (7864), pp.547–552.

Müller, L., Hatzfeld, M. and Keil, R. (2021). Desmosomes as signaling hubs in the regulation of cell behavior. *Frontiers in cell and developmental biology*, 9, p.745670.

Muller, M. et al. (1991). Embryology of the hair follicle. *Early human development*, 26 (3), pp.159–166.

Murphy, M., Kerr, P. and Grant-Kels, J. M. (2007). The histopathologic spectrum of psoriasis. *Clinics in dermatology*, 25 (6), pp.524–528.

Naik, S. et al. (2012). Compartmentalized control of skin immunity by resident commensals. *Science (New York, N.Y.)*, 337 (6098), pp.1115–1119.

Naik, S. et al. (2015). Commensal-dendritic-cell interaction specifies a unique protective skin immune signature. *Nature*, 520 (7545), pp.104–108.

Nankervis, H. et al. (2016). *Scoping systematic review of treatments for eczema*.

Nasca, M. R. et al. (2019). Image Gallery: Reproduction of the Auspitz sign by videodermatoscopy, confocal microscopy and horizontal histopathology. *The British journal of dermatology*, 180 (6), p.e178.

Navarini, A. A., French, L. E. and Hofbauer, G. F. L. (2011). Interrupting IL-6-receptor signaling improves atopic dermatitis but associates with bacterial superinfection. *The journal of allergy and clinical immunology*, 128 (5), pp.1128–1130.

Nedoszytko, B. et al. (2014). Chemokines and cytokines network in the pathogenesis of the inflammatory skin diseases: atopic dermatitis, psoriasis and skin mastocytosis. *Postepy dermatologii i alergologii*, 31 (2), pp.84–91.

Nguyen, A. V. and Soulika, A. M. (2019). The dynamics of the skin's immune system. *International journal of molecular sciences*, 20 (8), p.1811.

Noden, D. M. and Trainor, P. A. (2005). Relations and interactions between cranial mesoderm and neural crest populations. *Journal of anatomy*, 207 (5), pp.575–601.

Norlén, L. (2006). Stratum corneum keratin structure, function and formation - a comprehensive review. *International journal of cosmetic science*, 28 (6), pp.397–425.

Nuber, U. A. et al. (1996). Patterns of desmocollin synthesis in human epithelia: immunolocalization of desmocollins 1 and 3 in special epithelia and in cultured cells. *European journal of cell biology*, 71 (1), pp.1–13.

Nunan, R. et al. (2015). Ephrin-BS drive junctional downregulation and actin stress fiber disassembly to enable wound re-epithelialization. *Cell reports*, 13 (7), pp.1380–1395.

Nussbaum, L., Chen, Y. L. and Ogg, G. S. (2021). Role of regulatory T cells in psoriasis pathogenesis and treatment. *The British journal of dermatology*, 184 (1), pp.14–24.

Olabi, B. et al. (2020). What's new in atopic eczema? An analysis of systematic reviews published in 2018. Part 2: systemic therapies. *Clinical and Experimental Dermatology*, 45 (8), pp.980–985.

Olson, L. E. et al. (2005). Barx2 functions through distinct corepressor classes to regulate hair follicle remodeling. *Proceedings of the National Academy of Sciences of the United States of America*, 102 (10), pp.3708–3713.

Ono, M. et al. (2007). Foxp3 controls regulatory T-cell function by interacting with AML1/Runx1. *Nature*, 446 (7136), pp.685–689.

Papalexi, E. and Satija, R. (2018). Single-cell RNA sequencing to explore immune cell heterogeneity. *Nature reviews. Immunology*, 18 (1), pp.35–45.

Parisi, R. et al. (2013). Global epidemiology of psoriasis: a systematic review of incidence and prevalence. *Journal of Investigative Dermatology*, 133 (2), pp.377–385.

Park, J.-E. et al. (2020). A cell atlas of human thymic development defines T cell repertoire formation. *Science*, 367 (6480). [Online]. Available at: doi:10.1126/science.aay3224.

Paternoster, L. et al. (2015). Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. *Nature genetics*, 47 (12), p.1449.

Patzelt, A. et al. (2008). Hair follicles, their disorders and their opportunities. *Drug discovery today. Disease mechanisms*, 5 (2), pp.e173–e181.

Petrof, G., Mellerio, J. E. and McGrath, J. A. (2012). Desmosomal genodermatoses. *The British journal of dermatology*, 166 (1), pp.36–45.

Piao, X. et al. (2019). Blockade of TNF receptor superfamily 1 (TNFR1)-dependent and TNFR1-independent cell death is crucial for normal epidermal differentiation. *The Journal of allergy and clinical immunology*, 143 (1), pp.213-228.e10.

Piñeiro, A. J., Houser, A. E. and Ji, A. L. (2022). Research techniques made simple: Spatial transcriptomics. *The journal of investigative dermatology*, 142 (4), pp.993-1001.e1.

Piraccini, B. M. and Alessandrini, A. (2014). Androgenetic alopecia. *Giornale italiano di dermatologia e venereologia: organo ufficiale, Societa italiana di dermatologia e sifilografia*, 149 (1), pp.15–24.

Pirastu, N. et al. (2017). GWAS for male-pattern baldness identifies 71 susceptibility loci explaining 38% of the risk. *Nature communications*, 8 (1). [Online]. Available at: doi:10.1038/s41467-017-01490-8.

Polański, K. et al. (2020). BBKNN: fast batch alignment of single cell transcriptomes. *Bioinformatics* , 36 (3), pp.964–965.

Popescu, D.-M. et al. (2019). Decoding human fetal liver haematopoiesis. *Nature*, 574 (7778), pp.365–371.

Porter, R. M. and Lane, E. B. (2003). Phenotypes, genotypes and their contribution to understanding keratin function. *Trends in genetics: TIG*, 19 (5), pp.278–285.

Price, V. H. (2003). Androgenetic alopecia in women. *The journal of investigative dermatology. Symposium proceedings*, 8 (1), pp.24–27.

Qian, Y. et al. (2010). IL-17 signaling in host defense and inflammatory diseases. *Cellular & molecular immunology*, 7 (5), pp.328–333.

Rahmani, W. et al. (2014). Hair follicle dermal stem cells regenerate the dermal sheath, repopulate the dermal papilla, and modulate hair type. *Developmental cell*, 31 (5), pp.543–558.

Randall, V. A. and Botchkareva, N. V. (2009). The biology of hair growth. In: *Cosmetics Applications of Laser & Light-Based Systems*. Elsevier. pp.3–35.

Rapp, M. et al. (2019). CCL22 controls immunity by promoting regulatory T cell communication with dendritic cells in lymph nodes. *The journal of experimental medicine*, 216 (5), pp.1170–1181.

Ratié, L. et al. (2013). Novel genes upregulated when NOTCH signalling is disrupted during hypothalamic development. *Neural development*, 8 (1), p.25.

Ratnasiri, K. et al. (2023). Single-cell RNA-seq methods to interrogate virus-host interactions. *Seminars in immunopathology*, 45 (1), pp.71–89.

Raza, S. I. et al. (2015). A novel homozygous nonsense mutation in the PVRL4 gene and expansion of clinical spectrum of EDSS1. *Annals of human genetics*, 79 (2), pp.92–98.

Reddy, S. et al. (2001). Characterization of Wnt gene expression in developing and postnatal hair follicles and identification of Wnt5a as a target of Sonic hedgehog in hair follicle morphogenesis. *Mechanisms of development*, 107 (1–2), pp.69–82.

Redondo Monte, E. et al. (2020). ZBTB7A prevents RUNX1-RUNX1T1-dependent clonal expansion of human hematopoietic stem and progenitor cells. *Oncogene*, 39 (15), pp.3195–3205.

Reid, A. J. et al. (2018). Single-cell RNA-seq reveals hidden transcriptional variation in malaria parasites. *eLife*, 7. [Online]. Available at: doi:10.7554/elife.33105.

Rendl, M., Lewis, L. and Fuchs, E. (2005). Molecular dissection of mesenchymal-epithelial interactions in the hair follicle. *PLoS biology*, 3 (11), p.e331.

Reyes-Reali, J. et al. (2018). Hypohidrotic ectodermal dysplasia: clinical and molecular review. *International journal of dermatology*, 57 (8), pp.965–972.

Reynolds, G. et al. (2021). Developmental cell programs are co-opted in inflammatory skin disease. *Science*, 371 (6527).

Rincón, M. et al. (1997). Interleukin (IL)-6 directs the differentiation of IL-4-producing CD4+ T cells. *The journal of experimental medicine*, 185 (3), pp.461–469.

Rishikaysh, P. et al. (2014). Signaling involved in hair follicle morphogenesis and development. *International journal of molecular sciences*, 15 (1), pp.1647–1670.

Rittié, L. et al. (2009). Hedgehog signaling maintains hair follicle stem cell phenotype in young and aged human skin. *Aging cell*, 8 (6), pp.738–751.

Rodriques, S. G. et al. (2019). Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science (New York, N.Y.)*, 363 (6434), pp.1463–1467.

Romanovsky, A. A. (2014). Skin temperature: its role in thermoregulation. *Acta physiologica (Oxford, England)*, 210 (3), pp.498–507.

Rozenblatt-Rosen, O. et al. (2017). The Human Cell Atlas: from vision to reality. *Nature*, 550 (7677), pp.451–453.

Rump, A. et al. (2004). Binding of ovarian cancer antigen CA125/MUC16 to mesothelin mediates cell adhesion. *The journal of biological chemistry*, 279 (10), pp.9190–9198.

Salaun, B. et al. (2004). CD208/dendritic cell-lysosomal associated membrane protein is a marker of normal and transformed type II pneumocytes. *The American journal of pathology*, 164 (3), pp.861–871.

Salimi, M. et al. (2013). A role for IL-25 and IL-33-driven type-2 innate lymphoid cells in atopic dermatitis. *The journal of experimental medicine*, 210 (13), pp.2939–2950.

Salman, A. et al. (2018). Impact of psoriasis in the quality of life of children, adolescents and their families: a cross-sectional study. *Anais brasileiros de dermatologia*, 93 (6), pp.819–823.

Sanchez Rodriguez, R. et al. (2014). Memory regulatory T cells reside in human skin. *The journal of clinical investigation*, 124 (3), pp.1027–1036.

Sander, G. et al. (2000). Expression of the homeobox gene, Barx2, in wool follicle development. *The journal of investigative dermatology*, 115 (4), pp.753–756.

Sankar, L. et al. (2017). Expression of angiogenic factors in psoriasis vulgaris. *Journal of clinical and diagnostic research: JCDR*, 11 (3), pp.EC23–EC27.

Sarangarajan, R. and Boissy, R. E. (2001). Tyrp1 and oculocutaneous albinism type 3. *Pigment cell research*, 14 (6), pp.437–444.

Sasaki, K. et al. (2019). Modulation of autoimmune pathogenesis by T cell-triggered inflammatory cell death. *Nature communications*, 10 (1), p.3878.

Saxena, N., Mok, K.-W. and Rendl, M. (2019). An updated classification of hair follicle morphogenesis. *Experimental dermatology*, 28 (4), pp.332–344.

Schirren, C. G. et al. (1997). Fetal and adult hair follicle. An immunohistochemical study of anticytokeratin antibodies in formalin-fixed and paraffin-embedded tissue. *The American journal of dermatopathology*, 19 (4), pp.335–340.

Schmidt-Ullrich, R. et al. (2006). NF-kappaB transmits Eda A1/EdaR signalling to activate Shh and cyclin D1 expression, and controls post-initiation hair placode down growth. *Development (Cambridge, England)*, 133 (6), pp.1045–1057.

Schneider, M. R., Schmidt-Ullrich, R. and Paus, R. (2009). The hair follicle as a dynamic miniorgan. *Current biology: CB*, 19 (3), pp.R132-42.

Schupp, J. C. et al. (2021). Integrated single-cell atlas of endothelial cells of the human lung. *Circulation*, 144 (4), pp.286–302.

Sciandra, F. et al. (2003). Dystroglycan and muscular dystrophies related to the dystrophin-glycoprotein complex. *Annali dell'Istituto superiore di sanita*, 39 (2), pp.173–181.

Seneschal, J. et al. (2012). Human epidermal Langerhans cells maintain immune homeostasis in skin by activating skin resident regulatory T cells. *Immunity*, 36 (5), pp.873–884.

Sennett, R. et al. (2015). An integrated transcriptome atlas of embryonic hair follicle progenitors, their niche, and the developing skin. *Developmental cell*, 34 (5), pp.577–591.

Sennett, R. and Rendl, M. (2012). Mesenchymal-epithelial interactions during hair follicle morphogenesis and cycling. *Seminars in cell & developmental biology*, 23 (8), pp.917–927.

Sequeira, I. and Nicolas, J.-F. (2012). Redefining the structure of the hair follicle by 3D clonal analysis. *Journal of cell science*, 125 (20), pp.e1–e1.

Shah, S. et al. (2017). SeqFISH accurately detects transcripts in single cells and reveals robust spatial organization in the hippocampus. *Neuron*, 94 (4), pp.752-758.e1.

Shelton, L. and Rada, J. A. S. (2009). Inhibition of human scleral fibroblast cell attachment to collagen type I by TGFBIp. *Investigative ophthalmology & visual science*, 50 (8), pp.3542–3552.

Shimomura, Y. et al. (2010). Mutations in the keratin 85 (KRT85/hHb5) gene underlie pure hair and nail ectodermal dysplasia. *The journal of investigative dermatology*, 130 (3), pp.892–895.

Silva, L. M. A. et al. (2020). Immunostaining study of cytokeratins in human hair follicle development. *Anais brasileiros de dermatologia*, 95 (3), pp.278–282.

Silva, L. M. A. et al. (2021). Immunoexpression of adhesion molecules during human fetal hair development. *Revista Científica Hospital Santa Izabel*, 5 (3), pp.159–166.

Son, N.-H. et al. (2018). Endothelial cell CD36 optimizes tissue fatty acid uptake. *The journal of clinical investigation*, 128 (10), pp.4329–4342.

Ståhl, P. L. et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, 353 (6294), pp.78–82.

Stehling-Sun, S. et al. (2009). Regulation of lymphoid versus myeloid fate "choice" by the transcription factor Mef2c. *Nature immunology*, 10 (3), pp.289–296.

Steinhoff, M. et al. (2006). Role of vasculature in atopic dermatitis. *The journal of allergy and clinical immunology*, 118 (1), pp.190–197.

Steuerman, Y. et al. (2018). Dissection of influenza infection in vivo by single-cell RNA sequencing. *Cell systems*, 6 (6), pp.679-691.e4.

Stolarczyk, E., Lord, G. M. and Howard, J. K. (2014). The immune cell transcription factor T-bet: A novel metabolic regulator. *Adipocyte*, 3 (1), pp.58–62.

Stoll, S. W. et al. (2016). The EGF receptor ligand amphiregulin controls cell division via FoxM1. *Oncogene*, 35 (16), pp.2075–2086.

Strober, B. et al. (2021). A survey of community dermatologists reveals the unnecessary impact of trial-and-error behavior on the psoriasis biologic treatment paradigm. *Dermatology and therapy*, 11 (5), pp.1851–1860.

Stuart, P. E. et al. (2010). Genome-wide association analysis identifies three psoriasis susceptibility loci. *Nature genetics*, 42 (11), pp.1000–1004.

Stubbington, M. J. T. et al. (2017). Single-cell transcriptomics to explore the immune system in health and disease. *Science (New York, N.Y.)*, 358 (6359), pp.58–63.

Sturm, R. A., Teasdale, R. D. and Box, N. F. (2001). Human pigmentation genes: identification, structure and consequences of polymorphic variation. *Gene*, 277 (1–2), pp.49–62.

Su, H. et al. (2020). Transglutaminase 3 promotes skin inflammation in atopic dermatitis by activating monocyte-derived dendritic cells via DC-SIGN. *The journal of investigative dermatology*, 140 (2), pp.370-379.e8.

Subkhankulova, T., Gilchrist, M. J. and Livesey, F. J. (2008). Modelling and measuring single cell RNA expression levels find considerable transcriptional differences among phenotypically identical cells. *BMC genomics*, 9 (1), p.268.

Sun, L. et al. (2006). EEDA: a protein associated with an early stage of stratified epithelial differentiation. *Journal of cellular physiology*, 206 (1), pp.103–111.

Suo, C. et al. (2022). Mapping the developing human immune system across organs. *Science*, 376 (6597), p.eabo0510.

Szabo, P. A. et al. (2019). Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nature communications*, 10 (1), p.4706.

Takahashi, R. et al. (2020). Defining transcriptional signatures of human hair follicle cell states. *The journal of investigative dermatology*, 140 (4), pp.764-773.e4.

Takatori, H. et al. (2009). Lymphoid tissue inducer-like cells are an innate source of IL-17 and IL-22. *The journal of experimental medicine*, 206 (1), pp.35–41.

Takeo, M., Lee, W. and Ito, M. (2015). Wound healing and skin regeneration. *Cold Spring Harbor perspectives in medicine*, 5 (1), p.a023267.

Tammela, T. and Alitalo, K. (2010). Lymphangiogenesis: Molecular mechanisms and future promise. *Cell*, 140 (4), pp.460–476.

Tang, A. L. et al. (2008). CTLA4 expression is an indicator and regulator of steady-state CD4+ FoxP3+ T cell homeostasis. *The journal of immunology*, 181 (3), pp.1806–1813.

Tang, F. et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature methods*, 6 (5), pp.377–382.

Tang, X.-Y. et al. (2022). Human organoids in basic research and clinical applications. *Signal transduction and targeted therapy*, 7 (1), p.168.

Tanjung, C. et al. (2017). Study protocol to investigate the environmental and genetic aetiology of atopic dermatitis: the Indonesian Prospective Study of Atopic Dermatitis in Infants (ISADI). *BMJ open*, 7 (3).

Te Molder, L., de Pereda, J. M. and Sonnenberg, A. (2021). Regulation of hemidesmosome dynamics and cell signaling by integrin α6β4. *Journal of cell science*, 134 (18). [Online]. Available at: doi:10.1242/jcs.259004.

Thomsen, S. F. et al. (2007). *Importance of genetic factors in the etiology of atopic dermatitis: a twin study*.

Tirosh, I. et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science (New York, N.Y.)*, 352 (6282), pp.189–196.

Tong, X. and Coulombe, P. A. (2006). Keratin 17 modulates hair follicle cycling in a TNFalpha-dependent fashion. *Genes & development*, 20 (10), pp.1353–1364.

Toshitani, A. et al. (1993). Increased interleukin 6 production by T cells derived from patients with atopic dermatitis. *The journal of investigative dermatology*, 100 (3), pp.299–304.

Toyoda, S. et al. (2022). Transforming growth factor β1 signaling links extracellular matrix remodeling to intracellular lipogenesis upon physiological feeding events. *The Journal of biological chemistry*, 298 (4), p.101748.

Trapnell, C. et al. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature biotechnology*, 32 (4), pp.381–386.

Tritschler, S. et al. (2019). Concepts and limitations for learning developmental trajectories from single cell genomics. *Development (Cambridge, England)*, 146 (12), p.dev170506.

Vahlquist, A., Gånemo, A. and Virtanen, M. (2008). Congenital ichthyosis: an overview of current and emerging therapies. *Acta dermato-venereologica*, 88 (1), pp.4–14.

Valladeau, J., Dezutter-Dambuyant, C. and Saeland, S. (2003). Langerin/CD207 sheds light on formation of birbeck granules and their possible function in Langerhans cells. *Immunologic research*, 28 (2), pp.93–107.

Varga, I. et al. (2011). The thymus: picture review of human thymus prenatal development. *Bratislavske lekarske listy*, 112 (7), pp.368–376.

Varki, R. et al. (2007). Epidermolysis bullosa. II. Type VII collagen mutations and phenotype–genotype correlations in the dystrophic subtypes. *Journal of medical genetics*, 44 (3), pp.181–192. [Accessed 27 March 2023].

Vento-Tormo, R. et al. (2018). Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature*, 563 (7731), pp.347–353.

Vičić, M. et al. (2021). Current concepts of psoriasis immunopathogenesis. *International journal of molecular sciences*, 22 (21). [Online]. Available at: doi:10.3390/ijms222111574.

Vickovic, S. et al. (2019). High-definition spatial transcriptomics for in situ tissue profiling. *Nature methods*, 16 (10), pp.987–990.

Vidal, V. P. I. et al. (2005). Sox9 is essential for outer root sheath differentiation and the formation of the hair stem cell compartment. *Current biology: CB*, 15 (15), pp.1340–1351.

Vieira Braga, F. A. and Miragaia, R. J. (2019). Tissue Handling and Dissociation for Single-Cell RNA-Seq. In: Proserpio, V. (Ed). *Single Cell Methods: Sequencing and Proteomics*. New York, NY: Springer New York. pp.9–21.

Villani, A.-C. et al. (2017). Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science (New York, N.Y.)*, 356 (6335). [Online]. Available at: doi:10.1126/science.aah4573.

Villanova, F. et al. (2014). Characterization of innate lymphoid cells in human skin and blood demonstrates increase of NKp44+ ILC3 in psoriasis. *The Journal of investigative dermatology*, 134 (4), pp.984–991.

Voigt, A. P. et al. (2022). Choroidal endothelial and macrophage gene expression in atrophic and neovascular macular degeneration. *Human molecular genetics*, 31 (14), pp.2406–2423.

Vulcano, M. et al. (2001). Dendritic cells as a major source of macrophage-derived chemokine/CCL22 in vitro and in vivo. *European journal of immunology*, 31 (3), pp.812–822.

van Wageningen, T. A. et al. (2019). Regulation of microglial TMEM119 and P2RY12 immunoreactivity in multiple sclerosis white and grey matter lesions is dependent on their inflammatory environment. *Acta neuropathologica communications*, 7 (1), p.206.

Walsh, N. M. et al. (2019). Plasmacytic cutaneous pathology: A review. *Journal of cutaneous pathology*, 46 (9), pp.698–708.

Wang, B. et al. (2020a). Human hair follicle-derived mesenchymal stem cells: Isolation, expansion, and differentiation. *World journal of stem cells*, 12 (6), pp.462–470.

Wang, S. et al. (2018). S100A8/A9 in Inflammation. *Frontiers in immunology*, 9. [Online]. Available at: doi:10.3389/fimmu.2018.01298.

Wang, S. et al. (2020b). Single cell transcriptomics of human epidermis identifies basal stem cell transition states. *Nature communications*, 11 (1), p.4239.

Ward, N. L. and Umetsu, D. T. (2014). A new player on the psoriasis block: IL-17A- and IL-22-producing innate lymphoid cells. *The Journal of investigative dermatology*, 134 (9), pp.2305–2307.

Weber, E. L. et al. (2020). Human fetal scalp dermal papilla enriched genes and the role of R-Spondin-1 in the restoration of hair neogenesis in adult mouse cells. *Frontiers in cell and developmental biology*, 8, p.583434.

Werfel, T. et al. (2016). Cellular and molecular immunologic mechanisms in patients with atopic dermatitis. *The journal of allergy and clinical immunology*, 138 (2), pp.336–349.

Wernersson, S. and Pejler, G. (2014). Mast cell secretory granules: armed for battle. *Nature reviews. Immunology*, 14 (7), pp.478–494.

Wessells, H. et al. (2009). Transcriptional profiling of human cavernosal endothelial cells reveals distinctive cell adhesion phenotype and role for claudin 11 in vascular barrier function. *Physiological genomics*, 39 (2), pp.100–108.

Whitting, D. A. et al. (2008). Biology of the hair follicle. *Hair growth and disorders*, pp.1–22.

Wilke, K. et al. (2007). A short history of sweat gland biology. *International journal of cosmetic science*, 29 (3), pp.169–179.

Williams, S. et al. (2020). *50 Spatially resolved molecular investigation of triple negative breast cancer and its immune microenvironment*. BMJ Specialist Journals.

Wilson, R. P. et al. (2019). IgM plasma cells reside in healthy skin and accumulate with chronic inflammation. *The journal of investigative dermatology*, 139 (12), pp.2477–2487.

Windoffer, R. et al. (2011). Cytoskeleton in motion: the dynamics of keratin intermediate filaments in epithelia. *The journal of cell biology*, 194 (5), pp.669–678.

Wolf, F. A., Angerer, P. and Theis, F. J. (2018). SCANPY: large-scale single-cell gene expression data analysis. *Genome biology*, 19 (1), p.15.

Wolock, S. L., Lopez, R. and Klein, A. M. (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell systems*, 8 (4), pp.281-291.e9.

Wright, J. T. et al. (2019). Ectodermal dysplasias: Classification and organization by phenotype, genotype and molecular pathway. *American journal of medical genetics. Part A*, 179 (3), pp.442–447.

Wyce, A. et al. (2010). The androgen receptor modulates expression of genes with critical roles in muscle development and function. *Endocrine reviews*, 31 (4), pp.604–604.

Xia, C. et al. (2019). Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 116 (39), pp.19490–19499.

Xie, W. et al. (2021). Modifiable lifestyle and environmental factors associated with onset of psoriatic arthritis in patients with psoriasis: A systematic review and meta-analysis of

observational studies. *Journal of the American Academy of Dermatology*, 84 (3), pp.701–711.

Xu, W. et al. (2014). The expression of proinflammatory genes in epidermal keratinocytes is regulated by hydration status. *The journal of investigative dermatology*, 134 (4), pp.1044–1055.

Xu, X. et al. (2018). Transforming growth factor-β in stem cells and tissue homeostasis. *Bone research*, 6, p.2.

Xu, Y. et al. (2021). Single-cell transcriptome analysis reveals the dynamics of human immune cells during early fetal skin development. *Cell reports*, 36 (6), p.109524.

Yamazaki, T. et al. (2008). CCR6 regulates the migration of inflammatory and regulatory T cells. *The journal of immunology*, 181 (12), pp.8391–8401.

Yan, B. et al. (2020). The role of Langerhans cells in epidermal homeostasis and pathogenesis of psoriasis. *Journal of cellular and molecular medicine*, 24 (20), pp.11646–11655.

Yan, B.-X. et al. (2018). Comparative expression of PEDF and VEGF in human epidermal keratinocytes and dermal fibroblasts: from normal skin to psoriasis. *Discovery medicine*, 25 (136), pp.47–56.

Yang, C. et al. (2019). Heterogeneity of human bone marrow and blood natural killer cells defined by single-cell transcriptome. *Nature communications*, 10 (1), p.3931.

Yin, S. et al. (2018). Overexpression of FAM3C protein as a novel biomarker for epithelial-mesenchymal transition and poor outcome in gastric cancer. *International journal of clinical and experimental pathology*, 11 (8), pp.4247–4256.

Yiu, Z. Z. N. et al. (2019). A standardization approach to compare treatment safety and effectiveness outcomes between clinical trials and real-world populations in psoriasis. *The British journal of dermatology*, 181 (6), pp.1265–1271.

Zammit, P. S. (2017). Function of the myogenic regulatory factors Myf5, MyoD, Myogenin and MRF4 in skeletal muscle, satellite cells and regenerative myogenesis. *Seminars in cell & developmental biology*, 72, pp.19–32.

Zhang, B. et al. (2022). A human embryonic limb cell atlas resolved in space and time. *bioRxiv*. [Online]. Available at: doi:10.1101/2022.04.27.489800.

Zhang, J.-G. et al. (2012). The dendritic cell receptor Clec9A binds damaged cells via exposed actin filaments. *Immunity*, 36 (4), pp.646–657.

Zhao, X. et al. (2018). Beneficial role of neutrophils through function of lactoferrin after intracerebral hemorrhage. *Stroke; a journal of cerebral circulation*, 49 (5), pp.1241–1247.

Zhao, Z. et al. (2022). Organoids. *Nature reviews. Methods primers*, 2 (1). [Online]. Available at: doi:10.1038/s43586-022-00174-y.

Zheng, G. X. Y. et al. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature communications*, 8 (1), p.14049.

Zhou, Y. et al. (2022). Investigation of hub genes and immune infiltration in androgenetic alopecia using bioinformatics analysis. *Annals of translational medicine*, 10 (22), p.1226.

Appendices

Appendix A. Co-first author of Journal of Investigative Dermatology Review Article: Defining the Skin Cellular Community Using Single-Cell Genomics to Advance Precision Medicine

Montagna Symposium on the Biology of Skin

Defining the Skin Cellular Community Using Single-Cell Genomics to Advance Precision Medicine

Anna Dubois[1,4], Nusayhah Gopee[1,4], Bayanne Olabi[2,4] and Muzlifah Haniffa[1,3]

Single-cell genomics has revolutionized biological science, enabling high-resolution analysis of human tissues. The ability to demonstrate the role and function of distinct cell types comprising human tissues paves the way for a new understanding of cellular pathways, interactions, and future research directions. The skin, easily accessible and possessing a diverse and complex role in defending us both physically and immunologically from the outside world, lends itself ideally to single-cell genomics analysis. Here, we outline the benefits of single-cell RNA sequencing while also highlighting the challenges in achieving a meaningful result from its use. Key milestones relating to the study of skin in this way are introduced, covering both healthy and diseased states, and we discuss the potential promise of single-cell RNA sequencing to result in tangible medical advances, with a particular focus on precision medicine.

INTRODUCTION

In addition to providing a physical barrier between the organism and the external world, the skin utilizes both the innate and adaptive immune systems to provide protection from external insult. The skin is a complex tissue of different embryological origins and comprising a wide range of cell types. The potential to fully understand this intriguing system in both health and disease is a step closer to being realized with the use of techniques that study tissues at the individual cell level. Early single-cell transcriptome analysis leveraged DNA microarray or qPCR analysis (Jensen and Watt, 2006; Tang et al., 2009). Single-cell RNA sequencing (scRNA-seq), which this article will focus on, has been widely adopted as a robust method to dissect cellular heterogeneity

in many biological contexts. The technology has evolved in the recent years as a highly reliable and sensitive method of transcriptomic analysis, enabling the profiling of an exponentially increased number of cells, with an unprecedented degree of accuracy and reproducibility at a single-cell level (Svensson et al., 2018; Wu et al., 2018). It overcomes the limitations imposed by other single-cell approaches such as FACS, which restricts the analysis to a small number of selected proteins, often biased toward already defined genes (Stubbington et al., 2017). ScRNA-seq also addresses the drawbacks of bulk RNA sequencing of whole tissue, which obscures cellular diversity and rarer small subpopulations of cells through misinterpretation as technical noise and averaging of cell expression (Stubbington et al., 2017; Suvà and Tirosh, 2019; Wu et al., 2018).

ScRNA-seq has enabled high dimensional analysis of human tissues, revealing novel cell states and biological pathways underpinning human development, tissue homeostasis, and disease (Miragaia et al., 2019). ScRNA-seq and other single-cell omics technologies have created a new paradigm to study human tissues at scale, investigating the whole one cell at a time. ScRNA-seq has been leveraged in various fields, contributing to new discoveries and improved understanding, for instance, in areas of embryonic medicine and developmental biology (Li et al., 2018; Popescu et al., 2019; Vento-Tormo et al., 2018), tumor ecosystems and oncogenic processes (Durante et al., 2020; Patel et al., 2014; Puram et al., 2017; Tirosh et al., 2016), immunology and immunotherapy (Jerby-Arnon et al., 2018; Stubbington et al., 2017; Szabo et al., 2019), and infectious and autoimmune diseases (Der et al. 2019a; Gawel et al., 2019; Jin et al., 2017; Jordão et al., 2019; Reid et al., 2018; Steuerman et al., 2018). The Human Cell Atlas (Regev et al., 2017) exemplifies the type of ambitious project that has arisen from the potential of this technique, aiming to identify the molecular properties of each cell that forms a healthy human. The advantage of studying discrete cellular units of any living tissue is that the constituent parts of the whole and its emergent properties become visible. ScRNA-seq, one of the earliest high-throughput single-cell techniques deployed for Human Cell Atlas projects, demonstrated the potential to identify cells and infer their function on the basis of transcriptome profile. Challenges, opportunities, and evidence of real progress, as applied to skin biology and pathophysiology, are highlighted below, and there is much promise for the future.

ScRNA-SEQ PLATFORMS AND DATA GENERATION

High-throughput scRNA-seq data generation primarily relies on tissue dissociation to obtain cells in suspension.

[1]Department of Dermatology and NIHR Newcastle Biomedical Research Centre, Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne, United Kingdom; [2]Department of Dermatology, Lauriston Building, Lauriston Place, Edinburgh, United Kingdom; and [3]Biosciences Institute, Newcastle University, Newcastle upon Tyne, United Kingdom

[4]These authors contributed equally to this work.

Correspondence: Muzlifah Haniffa, Biosciences Institute, Newcastle University, Newcastle upon Tyne, NE2 4HH, United Kingdom. E-mail: m.a. haniffa@ncl.ac.uk

Abbreviations: KC, keratinocyte; scRNA-seq, single-cell RNA sequencing

www.jidonline.org 255

Various mechanically and enzymatically (collagenase and/or trypsin) based protocols have been utilized to yield single cells from different types of tissue (Vieira Braga and Miragaia, 2019). This requires careful optimization because it may affect cellular yield, viability, and transcriptome. Details of the specific tissue dissociation protocols used for skin scRNA-seq profiling are detailed in Table 1. Tissue dissociation enables high-throughput analysis but negates spatial information of cells within their tissue context. Most platforms (e.g., 10× Genomics, Drop-seq, Massively Parallel RNA Single-Cell Sequencing) rely on sequencing short transcripts from the 3′ end of the mRNA (Vegh and Haniffa, 2018). In contrast, Smart seq2 enables full transcripts to be sequenced, providing information on alternatively spliced mRNA molecules. The different platforms capture variable mRNA molecules ranging from hundreds to thousands of transcripts per cell (Stubbington et al., 2017).

In addition, several technologies are currently available for spatially resolved tissue analysis, which rely on targeted RNA hybridization (sequential- or multiplexed-FISH or in situ RNA sequencing) (Gyllborg and Nilsson, 2020; Shah et al., 2017; Xia et al., 2019). These technologies are able to assess >100−1,000 parameters simultaneously at single-cell resolution but are dependent on having a highly skilled operator and not all commercially available. Spatial transcriptomics (Visium) is another new technology that generates two-dimensional positional RNA sequencing data from histological tissue sections (Ståhl et al., 2016). This technique can capture all mRNA within the tissue overlying each spot to be profiled but not at a single-cell level. Combining spatial transcriptomics with single-cell suspension will provide valuable knowledge and information on disease pathogenesis in situ.

Each scRNA-seq protocol involves isolation of single cells, library preparation, and sequencing (Figure 1). Single cells can be captured by FACS isolation into a PCR plate or microfluidic system. Cells can also be captured with barcoded beads and primers into droplets or microwell arrays for cDNA synthesis. Combinatorial indexing (Vitak et al., 2017) or sequential barcoding (Rosenberg et al., 2018) can also be used for scRNA-seq. There has been a rapid acceleration of throughput in single-cell transcriptomics in the past 10 years (Svensson et al., 2018) accompanied by a reduction in analysis cost, which will facilitate its deployment in a clinical setting. Hundreds of thousands of cells can now be routinely sequenced in a single experiment. Alongside the promise of new technology come challenges relating to experimental techniques at all stages (Chen et al., 2019). Although the study of the human skin using scRNA-seq is in its infancy, as costs of experiments fall and clinical utility rises, we can expect to see this technique being used to study both healthy skin and skin disease, leading to a greater understanding of cellular pathways and potential new treatments. Achieving the detail required to gain meaningful outputs from scRNA-seq analysis—be it detecting rare cell types, picking apart differences in a population, or understanding the trajectory of differentiating cells—requires several complex steps to maximize information gained while minimizing inaccuracy.

ScRNA-SEQ DATA ANALYSIS AND INTERPRETATION

The small amount of material from each cell analyzed in scRNA-seq results in data that are noisier than from bulk sequencing. Amplification of the starting material can lead to biases by increasing transcripts that are already abundant. Technical dropouts, the phenomenon by which a gene may be seen in some but not all cells of the same type, mean zero values are observed, as only a small part of the transcript has been captured. Other than true biological variation, the most common cause for this is usually a technical difficulty, such as mRNA degradation during cell lysis or resulting from sampling variation (Qiu, 2020). Adjusting variables such as the sequencing depth or platform used may go some way to correcting for this and improving the sparsity of data—a significant issue in scRNA-seq. The difficulty of analyzing sparse data can be addressed with caution by either using statistical approaches that are able to take this into account and correct for the observed zeros or by attempts to impute missing values to improve results (Andrews and Hemberg, 2019).

Doublets are generated when two cells are inadvertently captured together, resulting in mixed transcriptional profiles. Droplet-encapsulation methods have a higher doublet rate than FACS isolation into wells (Ilicic et al., 2016) but enable higher throughput of cells to be analyzed. Identification and removal of doublets are important during experimentation and analysis, and this can be achieved in various ways. Multiplexing to exploit genetic variation has been used, as have techniques involving oligonucleotide-tagged antibodies against cell-surface proteins (Kang et al., 2018; Stoeckius et al., 2018). Doublets can also be inferred from gene expression profiles, where the expression profile of the doublet is a composite of two-cell states without any unique gene expression profile. Several computation methods have also been devised to identify doublets in scRNA-seq datasets (McGinnis et al., 2019; Wolock et al., 2019).

Single-cell data may be combined from multiple experiments with many differing variables. These differences result in variation or batch effect in the data, which needs to be corrected so as not to confound and invalidate the results. The difficulty lies in preserving key biological differences at the same time as removing those arising from nonbiological factors. Computational algorithm methods, for example, Harmony or batch-balanced k nearest neighbors, can be used to correct for different technologies and nonidentical cell types and can be used across multiple batches and for large datasets (Korsunsky et al., 2019; Polański et al., 2020) with a statistical assessment of batch correction (e.g., kBET statistic) (Park et al., 2018).

As the number of cells being analyzed per experiment increases, there is a need for single-cell data interpretation to keep pace. Gaining useful insights from scRNA-seq experiments requires a rigorous approach and involves several complex steps to make sense of the results. Normalization of the data is required to correct for cell-to-cell variation before analyzing these results with dimensionality reduction techniques. Simplifying visual representation of the data using a lower dimension space and removal of uninformative genes enable algorithms and clustering approaches to be applied to huge data sets. Nonlinear dimensionality reduction, such as

Table 1. Summary of Recent scRNA-seq Studies on Human Skin

Reference	Skin State	Number of Samples	Time Before Tissue Digestion	Tissue Dissociation Method	Cell-Type Enrichment	Transcriptomic Platform	Number of Cells Sequenced
Healthy skin							
Finnegan et al., 2019	Neonatal foreskin	3	Tissue stored at 4 °C overnight before dissociation.	Epidermis mechanically separated from dermis, then incubated in trypsin for 15 min.	FACS to isolate live, single cells	10X Genomics	22,338
Philippeos et al., 2018	Adult female abdominal skin	1	N/A	1 h treatment with dispase II, followed by epidermal peeling. Dermis was then treated with Miltenyi whole-skin dissociation kit overnight.	FACS isolation of lin- cells (CD31-CD45-ECad-) and lin-CD90+ cells from dermis	SmartSeq2	184
Solé-Boldo et al., 2020	Sun-protected site from male Caucasian skin aged 25 and 27 years	2	No longer than 1 h	Miltenyi whole-skin dissociation kit, followed by gentleMACS mechanical dissociation.	Magnetic isolation of live cells	10X Genomics	15,457
	Sun-protected site from male Caucasian skin aged between 53 and 70 years	3					
Tabib et al., 2018	Dorsal mid-forearm skin from males and females of varying ages	6	Unspecified transport time between biopsy and dissociation.	2 h treatment with Miltenyi whole-skin dissociation kit, followed by gentleMACS mechanical dissociation.	No enrichment	10X Genomics	8,522
Takahashi et al., 2020	Follicle-enriched human skin from hair transplant micrografts	2	Tissue stored at 4 °C overnight before dissociation.	30 min incubation with dispase, followed by dissociation with a P1000 pipette and 10 min incubation in trypsin.	FACS high-speed cell sorter	Drop-Seq	5,270 cells
Vorstandlechner et al., 2020	Surplus trunk skin removed during abdominoplasty from healthy female donors aged 30–43 y	3	Unspecified transport time between biopsy and dissociation.	2.5 h enzymatic digestion using Miltenyi whole-skin dissociation kit, followed by gentleMACS mechanical dissociation.	FACS to isolate live, single cells	10X Genomics	5,000
Xue et al., 2020	Dorsal mid-forearm skin from males and females of varying ages	10	N/A	2 h enzymatic digestion using Miltenyi whole-skin dissociation kit, followed by gentleMACS mechanical dissociation.	No enrichment	10X Genomics	27,869
Diseased skin							
Apostolidis et al., 2018	Diffuse cutaneous systemic sclerosis—dorsal mid-forearm skin.	4	N/A	2 h treatment with Miltenyi whole-skin dissociation kit, followed by gentleMACS mechanical dissociation.	FACS into single-cell wells in plates	SmartSeq2	96	HSPG2 and APLNR were identified as two of the top markers of endothelial cell injury in systemic sclerosis.
	Age- and sex-matched healthy controls	4					88
Cheng et al., 2018	Healthy adult scalp epidermis	3	Tissue stored at 4 °C for 1 −2 days before dissociation.	2 h treatment with dispase, followed by epidermal peeling. Epidermis was then trypsin-treated for 15 min.	FACS to isolate live, single cells	10X Genomics	92,889 (2,000 −12,000 cells/ sample)	Stereotyped KC subpopulations comprising human epidermis were characterized. Inflammatory skin states were identified, including S100 activation in the interfollicular epidermis of the normal scalp, enrichment of a CD1C+CD301A+ myeloid DC population in psoriatic epidermis, and IL1βhi CCL3hiCD14+ monocyte-derived macrophages enriched in the foreskin.
	Healthy adult truncal epidermis	3					
	Healthy neonatal foreskin epidermis	3					

(continued)

247

258

Journal of Investigative Dermatology (2021), Volume 141

A *Dubois* et al.
Single-Cell Genomics in Dermatology

Table 1. Continued

Reference	Skin State	Number of Samples	Time Before Tissue Digestion	Tissue Dissociation Method	Cell-Type Enrichment	Transcriptomic Platform	Number of Cells Sequenced
	Psoriasis - truncal epidermis	3					
Der et al. 2019b	Healthy skin of patients with lupus nephritis	17	2 h transport time between biopsy and either immediate dissociation or cryopreservation.	Samples were treated with liberase for 15 min.	No enrichment	Fluidigm C1 800-well	~250 cells/sample	Type I IFN response signatures in KCs distinguished patients with lupus nephritis from healthy control subjects.
	Healthy control skin	3					
Gaydosik et al., 2019	Lesional skin of advanced-stage cutaneous T-cell lymphoma (stage 2b—4a)	5	N/A	2 h treatment with Miltenyi whole-skin dissociation kit, followed by gentleMACS mechanical dissociation.	All skin cells	10X Genomics	30,663	Large intratumor gene expression heterogeneity of T-lymphocytes was found. A 17-gene expression signature was identified in highly proliferating lymphocytes in CTCL samples, with three genes (*PCNA, ATP5C1, NUSAP1*) validated by immunohistochemistry. Tumor-infiltrating CD8+ lymphocytes exhibited heterogeneity of effector and exhaustion programs across samples.
	Healthy control skin	4					14,719
He et al., 2020a	Lesional (4) and nonlesional (5) skin from extremities in moderate —severe AD	5	Biopsies were cryopreserved before dissociation and sequencing.	Samples were sent for external processing, and no details were given.	No enrichment	10X Genomics	10,169 from lesional AD and 6,653 from nonlesional AD	AD lesions were characterized by expanded type 2 and/or type 22 T cells and inflammatory DCs and by a unique inflammatory fibroblast that may interact with immune cells to regulate lymphoid cell organization and type 2 inflammation.
	Healthy-matched controls	7					22,220
Hughes et al., 2019[1]	Healthy skin and various inflammatory conditions: acne, alopecia, GA, leprosy, and psoriasis	9	N/A	No details provided for the dissociation protocol used (Biorxiv submission).	No enrichment; MACS for certain cell types/samples	Seq-well	20,308	Acne: 2 clusters of endothelial cells marked by the expression of *SLC9A3R2*. Alopecia: a subcluster of T cells characterized by expression of *PDE4D*. Both GA and leprosy were characterized by the presence of immature CD8+ CTL and mature CTL effectors containing CD8+ T-CTL, γδ, and NK cells. Psoriasis: a subcluster of DCs (IRF4+ cDC2) that display the elevated expression of *CCL17, CCL22*, and *IL12B*.
Kim et al., 2020	DRESS	1	N/A	3 h treatment with Miltenyi whole-skin dissociation kit during gentleMACS mechanical dissociation.	FACS to isolate live, single cells	10X Genomics	4,676	The Jak-STAT signaling pathway was identified as a possible target for treatment. In addition, they showed that central memory CD4+ T cells were enriched with DNA from human herpesvirus 6b.
	Flank skin from healthy controls	5					13,524
Tirosh et al., 2016	Melanoma with lymphoid tissue metastasis	10	N/A	Samples were treated with collagenase P + DNase I for 10 min	FACS to isolate live, single cells	SmartSeq2	4,645	AXL-expressing cells (drug-resistant program) were identified in all melanoma samples and enriched in those resistant to RAF/MEK inhibition. Heterogeneity in the tumor microenvironment, including fibroblasts and T-cell composition, vary between melanomas.
	Melanoma with distant metastases	8					
	Primary acral melanoma	1					

Abbreviations: AD, atopic dermatitis; CTCL, cutaneous T-cell lymphoma; CTL, cytotoxic T lymphocyte; DC, dendritic cell; DRESS, drug reaction with eosinophilia and systemic symptoms; GA, granuloma annulare; KC, keratinocyte; MACS, magnetic-activated cell sorting; N/A, not applicable; scRNA-seq, single-cell RNA sequencing.
[1]Hughes TK, Wadsworth MH, Gierahn TM, Do T, Weiss D, Andrade PR, et al. Highly efficient, massively-parallel single-cell RNA-Seq reveals cellular states and molecular features of human skin pathology. BioRxiv 2019.
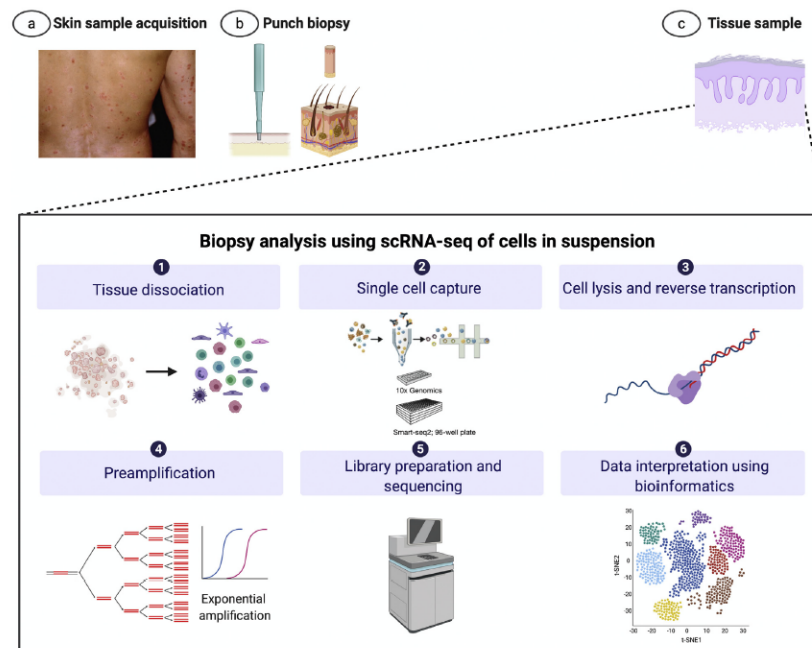
Figure 1. General workflow of scRNA-seq experiments from patient to data. Following sample acquisition of normal or lesional skin with a biopsy procedure, the tissue is processed to generate a single-cell suspension. Various techniques are available to enrich cells of interest, including fluorescence-activated and magnetic cell sorting. Single cells are captured into wells or using microfluidic- and droplet-based platforms. Single cells are barcoded and lysed to extract RNA molecules. The reverse transcription and preamplification steps create amplified and tagged cDNA that is sequenced by next-generation sequencing. Bioinformatic tools are available to deconvolute single-cell transcriptome and analyze and visualize the data. This figure was created with BioRender.com. scRNA-seq, single-cell RNA sequencing; t-SNE, t-distributed stochastic neighbor embedding.

t-distributed stochastic neighbor embedding (van der Maaten and Hinton, 2008) or uniform manifold approximation and projection (Becht et al., 2019), enables high dimensional data structure to be visualized in two dimensions. Distinguishing between bona fide cell states and results caused by artifact remains an issue throughout data interpretation, and careful consideration of these factors is important in ensuring high-quality outputs from the skin samples studied. Establishing cellular heterogeneity within scRNA-seq data is commonly performed using unsupervised clustering methods. There are several algorithms for unsupervised clustering (Andrews and Hemberg, 2018). When using a clustering method, the number of partitions or resolution needs to be specified. Clustering consistency can be assessed by using different algorithms on the same data (Andrews and Hemberg, 2018; Kiselev et al., 2017). Once cells have been sorted into clusters, genes unique to and highly expressed by the clusters (marker genes) are derived. These genes inform the annotation of the cell states with the clusters. Biological validation

of cell-annotated cell states is necessary and critical to assess the accuracy of annotation and provide insights into the biological relevance of identified cell states.

UNDERSTANDING SKIN BIOLOGY IN HEALTH AND DISEASE

Understanding the complex interplay between skin-barrier proteins, immune cells, and inflammatory mediators has been difficult to resolve. ScRNA-seq has expanded our understanding of healthy and diseased skin by allowing high-resolution mapping of different cell types and states, analysis of the impact of differing conditions on gene expression, and elucidation of cellular developmental trajectories and spatial organization to decipher cellular interactions (Griffiths et al., 2018; Mok et al., 2018; Szabo et al., 2019). Table 1 shows a summary of recent scRNA-seq studies on human skin.

Studies on murine keratinocyte (KC) differentiation have demonstrated multiple distinct cell states in the follicular and

249

interfollicular epidermis of healthy telogen and anagen epidermis (Joost et al., 2018; Joost et al., 2016). These include previously undescribed cell states in the hair follicle simultaneously expressing outer and inner bulge gene signatures and a specific cell population lining sebaceous gland openings, which potentially play a critical role in protecting the hair follicle bulge against microorganisms (Joost et al., 2016). Previously undetectable transient cell states were also identified through the reconstruction of the genetic programs involved during hair follicle differentiation, with stem cells sharing a conserved basal gene expression program but spatially distinct transcriptome profile (Joost et al., 2016). More recently, Joost et al. (2020) revealed two distinct cell populations in the outer root sheath with different affinity for interaction with the surrounding stroma. They showed that a common pool of germinative layer cells differentiates into three separate lineages to form the inner root sheath, cortex and/or cuticle, and medulla (Joost et al., 2020).

Human epidermis from healthy scalp and trunk, neonatal foreskin, and psoriatic plaque analysis identified several KC states (Cheng et al., 2018). Apart from basal, spinous, and granular KCs, four additional states termed WntI, mitotic, follicle, and channel clusters were defined. These cell states were characterized by genes relating to antagonizing Wnt signaling, DNA synthesis and cell division, hair follicle and sebaceous glands, and cell–cell communication and ion channels, respectively. Inferred trajectory analysis revealed nonlinear and distinct differentiation pathways for KC cell states (Cheng et al., 2018), which were also observed in hair follicle KCs (Takahashi et al., 2020). Interestingly, inflammatory genes were found to be upregulated in clinically normal scalp epidermis, and psoriatic samples were not only enriched in cells from the mitotic cluster but also the channel cluster (Cheng et al., 2018).

Progress in understanding the contributions of the hair follicle cellular subunits and wound healing has also been made using scRNA-seq and lineage tracing. Two separate sources of stem cells responding to injury have been identified: one population in the basal interfollicular epithelium, which is already primed to interact with the wound stroma, and a second population in the hair follicle, which activates an interfollicular epithelium-like signature within 24 hours of wounding, even before migration to the wound area to enable interaction with the wound environment (Joost et al., 2018; Takahashi et al., 2020).

Deconvoluting the role of fibroblasts, which has lagged behind other cutaneous cellular components, is also being advanced by scRNA-seq. Two major fibroblast populations have been identified in human papillary and reticular dermis: *SFRP2*-expressing fibroblasts, which are small, elongated, and located within collagen bundles, in contrast to *FMO1*-expressing fibroblasts, which are larger and distributed in the interstitium and perivascular areas, in addition to minor subclusters (Tabib et al., 2018). When combined with gene ontology analysis, fibroblast heterogeneity by expression of *SFRP2, FMO1,* and *COL11A1* was associated with regulating signaling pathways, matrix deposition, as well as possibly serving as pluripotent mesenchymal cells (Philippeos et al., 2018; Tabib et al.,

2018). Similar functional and spatial heterogeneity has been demonstrated in murine fibroblast populations (Joost et al., 2020; Salzer et al., 2018) and more comprehensively defined in a recent larger study of human skin (Solé-Boldo et al., 2020). Analysis of full-thickness murine anagen and telogen skin has further shown that dermal fibroblasts exhibit distinct transcriptional states correlating to hair cycle stage (Joost et al., 2020). A potential role of Wnt signaling in determining fibroblast identity has also been uncovered, with papillary cell markers being lost in culture in the absence of Wnt signaling. This may imply a key function of Wnt signaling between basal KCs and papillary dermal fibroblasts (Philippeos et al., 2018). Novel insights into skin aging highlight age-related loss of fibroblast cellular identity with alteration in functional and spatial gene expression signatures, which may be altered by dietary and metabolic interventions (Salzer et al., 2018; Solé-Boldo et al., 2020).

A scRNA-seq study of 19 patients with melanoma (Tirosh et al., 2016) challenged the findings of previous bulk RNA analysis, which classified melanomas as either MITF-high or AXL-high tumors depending on their transcriptional states. They showed that all tumors contained malignant cells of both transcriptional states in varying degrees. An AXL-high genetic profile associates with intrinsic resistance to RAF/MEK inhibitors used for the treatment of melanoma (Tirosh et al., 2016). This program was found to be upregulated in tumor samples resistant to RAF/MEK inhibition, suggesting selective enrichment of AXL-high cells in tumors resistant to RAF/MEK inhibitors following treatment (Tirosh et al., 2016). In addition, an AXL-high transcriptional state was found to be positively correlated with cancer-associated fibroblasts, further shedding light on the interplay of cells within the tumor environment and its role in determining tumor composition and phenotype (Suvà and Tirosh, 2019; Tirosh et al., 2016).

ScRNA-seq was further combined with bulk RNA-sequenced melanoma profiles from the Cancer Genome Atlas (Jerby-Arnon et al., 2018; NIH, 2019), enabling the detection of a cellular program in a subgroup of melanoma cells that confers the ability for immune evasion and resistance to immune-checkpoint inhibitors. This program had prognostic and predictive value when tested in an independent cohort of patients with melanoma. They demonstrated that CDK4/6 inhibition represses the resistance program, increasing melanoma cell response to immune-checkpoint blockade, representing additional possible therapeutic targets in a patient-specific subgroup (Jerby-Arnon et al., 2018). The stratification of subpopulations of patients and even the potential n-of-1 scRNA-seq studies of exceptional responders (Suvà and Tirosh, 2019), leading to personalized cancer treatments including immunotherapy, has been highlighted (Valdes-Mora et al., 2018). A scRNA-seq study of advanced-stage cutaneous T-cell lymphoma discovered a three-gene expression signature (*PCNA, ATP5C1,* and *NUSAP1*) present in highly proliferating malignant lymphocytes, which could represent a novel biomarker to facilitate cutaneous T-cell lymphoma diagnosis (Apostolidis et al., 2018; Gaydosik et al., 2019).

250

Atopic dermatitis skin profiling revealed disease-related *COL6A5*- and *COL18A1*-expressing fibroblasts and expanded *LAMP3*-expressing dendritic cell, dendritic cell 2, resident memory T cells, and IL-13– and/or IL-22–expressing T cells (He et al., 2020a). Genes previously not linked to the development of systemic sclerosis, such as the *APLNR* and *HSPG2* involved in vascular endothelial activation and dysfunction and fibrotic processes, have been shown to be significantly upregulated in systemic sclerosis skin by scRNA-seq (Apostolidis et al., 2018). This provides new insight into the pathways of endothelial cell injury and pathogenesis of this hitherto poorly understood condition and potentially highlight the biomarkers of vascular injury for earlier identification of patients and accurate prognostication.

The field of skin stem-cell biology has also harnessed the power of scRNA-seq. New stem-cell populations and states in murine interfollicular epithelium and infundibulum have been characterized and distinguished using scRNA-seq in combination with differential expression of α6 and β1 integrins (He et al., 2020b). Embryonic single-layered epithelium eliminates unfit cells through apoptosis and/or engulfment in contrast to asymmetric division and/or differentiation by mature epidermal cells (Ellis et al., 2019). Age and the microenvironment have been shown to impact hair regeneration following wounding, as well as rejuvenation of aged stem cells by neonatal dermal cells (Ge et al., 2020).

The wealth of data amassed from scRNA-seq studies on healthy and diseased human skin continues to grow rapidly. There is a pressing need for the research and clinical communities to reconcile the findings on skin cellular heterogeneity from the various studies to derive a consensus for cell-type nomenclature. Batch effect resulting from technical differences—as mentioned earlier in this review—is in part a contributing factor to the observed variation, but thought must also be given to other differences such as age, anatomical site, and ethnicity when drawing comparisons and explaining discrepancies between results (Vorstandlechner et al., 2020). In addition, a detailed cross-species comparison with mouse skin, a widely used animal model for skin biology and disease studies, should also be attempted.

SINGLE-CELL MULTIOMICS AND PERSONALIZED THERAPY

This unprecedented ability to unravel diseases and observe them at the single-cell level, gaining an exceptionally detailed understanding of the relevant biology and disruptions of developmental and regulatory pathways, presents new potential applications and opportunities in this era of postgenomic science. Increasingly, medical interventions are being redefined toward an approach that is tailored to the medical, genetic, and personal data of each individual, hence the term precision medicine (Collins and Varmus, 2015; Mirnezami et al., 2012). This aims for a more predictive, preventive, and personalized approach as well as being participatory with the hope for active participation by patients (NIHRtv 2018); the term P4 Medicine represents the acme of these new perspectives (Fiala et al., 2019; Hood and Friend, 2011).

The impact of precision medicine has been highlighted in oncology, with the development of cancer therapies that target particular molecular abnormalities (Abrams et al., 2014; Garraway et al., 2013). Tumorigenesis, maintenance, and progression arise from aberrant molecular pathways secondary to mutated or altered oncogenes and tumor suppressor genes, and recent years have seen the development and use of these genomics-derived treatments, for example, in breast cancer (trastuzumab [herceptin]), gastrointestinal stromal tumors (imatinib), and lung cancer (erlotinib) (Dienstmann et al., 2013). Furthermore, these technologies are also being used in biomarker identification and prognostication and further parsing of the molecular mechanisms underlying disease phenotype (Chirshev et al., 2019).

In dermatology, systemic immunosuppressants are often required for severe inflammatory and autoimmune dermatoses. However, existing treatments rely on nonspecific immunosuppression by systemic or biologic treatments, are neither curative nor personalized, and in most cases, do not prevent disease progression or concomitant comorbidities. Although multiple biologic treatments that target specific pathogenic molecules and pathways have been developed and licensed for psoriasis in recent years, there is currently a lack of biomarkers to distinguish psoriasis molecular mechanism heterogeneity (endotype) and predict patient response. Consequently, guidance to help clinicians and patients to choose between different treatments is not available, resulting in trial-and-error prescribing (Iskandar et al., 2017), with significant-associated physical, social, and economic costs (Al Sawah et al., 2017). Collaborative research groups have highlighted the need in dermatology to better understand the variations in underlying disease mechanisms, presentation, and treatment response between subpopulations (Griffiths et al., 2015; Reynolds et al., 2016) to develop clinical tests to predict treatment response in individual patients and therefore provide optimal care.

The use of scRNA-seq data to guide precision treatment has recently been described in a case of drug reaction with eosinophilia and systemic symptoms (Kim et al., 2020). Blood and skin scRNA-seq profile from a patient with drug reaction with eosinophilia and systemic symptoms, refractory to conventional treatments with high-dose prednisolone, etanercept, intravenous immunoglobulin, cyclosporin, and mycophenolate mofetil, revealed the upregulation of the Jak–signal transducer and activator of transcription signaling pathway in skin-infiltrating lymphocytes. Tofacitinib, a JAK1 and JAK3 inhibitor treatment, resulted in resolution of skin inflammation. This case highlights the power of single-cell analyses in precision medicine to reveal molecular mechanisms for therapeutics selection and predict patient response. This proof of concept will shape the future of health care (Shalek and Benson, 2017).

CONCLUSION

In recent years, scRNA-seq has allowed us to gain new insights into the individual cells of human tissues and organs in both health and disease. As we continue to add to our ever-increasing level of understanding, a key

251

challenge will be the integration of this information with the aid of other techniques to understand how individual cells interact with one another, creating the behaviors of tissues and systems that we are able to observe. For example, spatial transcriptomics can be used to map the location context of scRNA-seq data, providing valuable knowledge and information on disease pathogenesis in situ. Techniques such as Cellular Indexing of Transcriptomes and Epitopes by sequencing can aid simultaneous corresponding protein detection and sophisticated computational methods, such as CytoTRACE, allow predictive modeling of cell states using scRNA-seq data. Coupling human skin organoid models with CRISPR-based perturbation systems will additionally facilitate mechanistic studies. As these technologies become more widely utilized in research and clinical practice, they can be used to manage subpopulations of patients afflicted with a particular disease. The promise of precision medicine, treating each individual patient in an individual way, will become a step closer to becoming a reality.

ORCIDs
Anna Dubois: http://orcid.org/0000-0002-5248-3687
Nusayhah Gopee: http://orcid.org/0000-0001-8507-5682
Bayanne Olabi: http://orcid.org/0000-0002-4786-7838
Muzlifah Haniffa: http://orcid.org/0000-0002-3927-2084

CONFLICT OF INTEREST
The authors state no conflict of interest.

AUTHOR CONTRIBUTIONS
Conceptualization: AD, NG, BO, MH; Supervision: MH; Writing - Original Draft Preparation: AB, NG, BO, MH; Writing - Review and Editing: AB, NG, BO, MH

REFERENCES

Abrams J, Conley B, Mooney M, Zwiebel J, Chen A, Welch JJ, et al. National Cancer Institute's Precision Medicine Initiatives for the new National Clinical Trials Network. Am Soc Clin Oncol Educ Book 2014;34: 71—6.

Al Sawah S, Foster SA, Goldblum OM, Malatestinic WN, Zhu B, Shi N, et al. Healthcare costs in psoriasis and psoriasis sub-groups over time following psoriasis diagnosis. J Med Econ 2017;20:982—90.

Andrews TS, Hemberg M. False signals induced by single-cell imputation [version 2; peer review: 4 approved]. F1000Research 2019;7:1740.

Andrews TS, Hemberg M. Identifying cell populations with scRNASeq. Mol Aspects Med 2018;59:114—22.

Apostolidis SA, Stifano G, Tabib T, Rice LM, Morse CM, Kahaleh B, et al. Single cell RNA sequencing identifies HSPG2 and APLNR as markers of endothelial cell injury in systemic sclerosis skin. Front Immunol 2018;9: 2191.

Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol 2019;37:38—44.

Chen G, Ning B, Shi T. Single-cell RNA-Seq technologies and related computational data analysis. Front Genet 2019;10:317.

Cheng JB, Sedgewick AJ, Finnegan AI, Harirchian P, Lee J, Kwon S, et al. Transcriptional programming of normal and inflamed human epidermis at single-cell resolution. Cell Rep 2018;25:871—83.

Chirshev E, Oberg KC, Ioffe YJ, Unternaehrer JJ. Let-7 as biomarker, prognostic indicator, and therapy for precision medicine in cancer. Clin Transl Med 2019;8:24.

Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med 2015;372:793—5.

Der E, Suryawanshi H, Buyon J, Tuschl T, Putterman C. Single-cell RNA sequencing for the study of lupus nephritis. Lupus Sci Med 2019a;6: e000329.

Der E, Suryawanshi H, Morozov P, Kustagi M, Goilav B, Ranabothu S, et al. Tubular cell and keratinocyte single-cell transcriptomics applied to lupus nephritis reveal type I IFN and fibrosis relevant pathways. Nat Immunol 2019b;20:915—27.

Dienstmann R, Rodon J, Barretina J, Tabernero J. Genomic medicine frontier in human solid tumors: prospects and challenges. J Clin Oncol 2013;31: 1874—84.

Durante MA, Rodriguez DA, Kurtenbach S, Kuznetsov JN, Sanchez MI, Decatur CL, et al. Single-cell analysis reveals new evolutionary complexity in uveal melanoma. NatCommun 2020;11:496.

Ellis SJ, Gomez NC, Levorse J, Mertz AF, Ge Y, Fuchs E. Distinct modes of cell competition shape mammalian tissue morphogenesis. Nature 2019;569: 497—502.

Fiala C, Taher J, Diamandis EP. P4 Medicine or O4 Medicine? Hippocrates provides the answer. J Appl Lab Med 2019;4:108—19.

Finnegan A, Cho RJ, Luu A, Harirchian P, Lee J, Cheng JB, et al. Single-cell transcriptomics reveals spatial and temporal turnover of keratinocyte differentiation regulators. Front Genet 2019;10:775.

Garraway LA, Verweij J, Ballman KV. Precision oncology: an overview. J Clin Oncol 2013;31:1803—5.

Gawel DR, Serra-Musach J, Lilja S, Aagesen J, Arenas A, Asking B, et al. A validated single-cell-based strategy to identify diagnostic and therapeutic targets in complex diseases. Genome Medicine Med 2019;11:47.

Gaydosik AM, Tabib T, Geskin LJ, Bayan C-ACA, Conway JF, Lafyatis R, et al. Single cell lymphocyte heterogeneity in advanced cutaneous T-cell lymphoma skin tumors. Clin Cancer Res 2019;25:4443—54.

Ge Y, Miao Y, Gur-Cohen S, Gomez N, Yang H, Nikolova M, et al. The aging skin microenvironment dictates stem cell behavior. Proc Natl Acad Sci USA 2020;117:5339—50.

Griffiths CEM, Barnes MR, Burden AD, Nestle FO, Reynolds NJ, Smith CH, et al. Establishing an academic-Industrial Stratified Medicine Consortium: Psoriasis stratification to optimize relevant therapy. J Invest Dermatol 2015;135:2903—7.

Griffiths JA, Scialdone A, Marioni JC. Using single-cell genomics to understand developmental processes and cell fate decisions. Mol Syst Biol 2018;14:e8046.

Gyllborg D, Nilsson M. HybISS: hybridization-based in situ sequencing. https://doi.org/10.17504/protocols.io.xy4fpyw; 2020 (accessed 22 March 2020).

He H, Suryawanshi H, Morozov P, Gay-Mimbrera J, Del Duca E, Kim HJ, et al. Single-cell transcriptome analysis of human skin identifies novel fibroblast subpopulation and enrichment of immune subsets in atopic dermatitis. J Allergy Clin Immunol 2020a;145: 1615—28.

He W, Ye J, Xu H, Lin Y, Zheng Y. Differential expression of α6 and β1 integrins reveals epidermal heterogeneity at single-cell resolution. J Cell Biochem 2020b;121:2664—76.

Hood L, Friend SH. Predictive, personalized, preventive, participatory (P4) cancer medicine (P4) cancer medicine. Nat Rev Clin Oncol 2011;8: 184—7.

Ilicic T, Kim JK, Kolodziejczyk AA, Bagger FO, McCarthy DJ, Marioni JC, et al. Classification of low quality cells from single-cell RNA-seq data. Genome Biol 2016;17:29.

Iskandar IYK, Ashcroft DM, Warren RB, Evans I, McElhone K, Owen CM, et al. Patterns of biologic therapy use in the management of psoriasis: cohort study from the British Association of Dermatologists Biologic Interventions Register (BADBIR). Br J Dermatol 2017;176:1297—307.

Jensen KB, Watt FM. Single-cell expression profiling of human epidermal stem and transit-amplifying cells: Lrig1 is a regulator of stem cell quiescence. Proc Natl Acad Sci USA 2006;103:11958—63.

252

Jerby-Arnon L, Shah P, Cuoco MS, Rodman C, Su M-JMJ, Melms JC, et al. A cancer cell program promotes T cell exclusion and resistance to checkpoint blockade. Cell 2018;175:984−97.e24.

Jin Z, Fan W, Jensen MA, Dorschner JM, Bonadurer GF 3rd, Vsetecka DM, et al. Single-cell gene expression patterns in lupus monocytes independently indicate disease activity, interferon and therapy. Lupus Sci Med 2017;4:e000202.

Joost S, Annusver K, Jacob T, Sun X, Dalessandri T, Sivan U, et al. The molecular anatomy of mouse skin during hair growth and rest. Cell Stem Cell 2020;26:441−57.e7.

Joost S, Jacob T, Sun X, Annusver K, La Manno G, Sur I, et al. Single-cell transcriptomics of traced epidermal and hair follicle stem cells reveals rapid adaptations during wound healing. Cell Rep 2018;25:585−97. e7.

Joost S, Zeisel A, Jacob T, Sun X, La Manno G, Lönnerberg P, et al. Single-cell transcriptomics reveals that differentiation and spatial signatures shape epidermal and hair follicle heterogeneity. Cell Syst 2016;3: 221−37.e9.

Jordão MJC, Sankowski R, Brendecke SM, Sagar, Locatelli G, Tai YH, et al. Single-cell profiling identifies myeloid cell subsets with distinct fates during neuroinflammation. Science 2019;363:eaat7554.

Kang HM, Subramaniam M, Targ S, Nguyen M, Maliskova L, McCarthy E, et al. Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. Nat Biotechnol 2018;36:89−94.

Kim D, Kobayashi T, Voisin B, Jo JHJH, Sakamoto K, Jin SPSP, et al. Targeted therapy guided by single-cell transcriptomic analysis in drug-induced hypersensitivity syndrome: a case report. Nat Med 2020;26: 36−43.

Kiselev VY, Kirschner K, Schaub MT, Andrews T, Yiu A, Chandra T, et al. SC3: consensus clustering of single-cell RNA-seq data. Nat Methods 2017;14: 483−6.

Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 2019;16:1289−96.

Li L, Guo F, Gao Y, Ren Y, Yuan P, Yan L, et al. Publisher correction: single-cell multi-omics sequencing of human early embryos. Nat Cell Biol 2018;20: 1227.

McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. Cell Syst 2019;8:329−37.e4.

Miragaia RJ, Gomes T, Chomka A, Jardine L, Riedel A, Hegazy AN, et al. Single-cell transcriptomics of regulatory T cells reveals trajectories of tissue adaptation. Immunity 2019;50:493−504.e7.

Mirnezami R, Nicholson J, Darzi A. Preparing for precision medicine. N Engl J Med 2012;366:489−91.

Mok KWKW, Saxena N, Rendl M. More than the sum of its parts: single-cell transcriptomics reveals epidermal cell states. Cell Rep 2018;25: 823−4.

National Institutes of health (NIH). The Cancer Genome Atlas. https://www. cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga/ history; 2019. (accessed March 25, 2020).

NIHRtv. I am research, are you?. https://www.youtube.com/watch?v=_X1 TfeZVG8Q; 2018. (accessed March 25, 2020).

Park JE, Polański K, Meyer K, Teichmann SA. Fast batch alignment of single cell transcriptomes unifies multiple mouse cell atlases into an integrated landscape. BioRxiv 2018.

Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science 2014;344:1396−401.

Philippeos C, Telerman SB, Oulès B, Pisco AO, Shaw TJ, Elgueta R, et al. Spatial and single-cell transcriptional profiling identifies functionally distinct human dermal fibroblast subpopulations. J Invest Dermatol 2018;138:811−25.

Polański K, Young MD, Miao Z, Meyer KB, Teichmann SA, Park JE. BBKNN: fast batch alignment of single cell transcriptomes. Bioinformatics 2020;36: 964−5.

Popescu D-MDM, Botting RA, Stephenson E, Green K, Webb S, Jardine L, et al. Decoding human fetal liver haematopoiesis. Nature 2019;574: 365−71.

Puram SV, Tirosh I, Parikh AS, Patel AP, Yizhak K, Gillespie S, et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck. Cancer Cell 2017;171: 1611−24.e24.

Qiu P. Embracing the dropouts in single-cell RNA-seq analysis. Nat Commun 2020;11:1169.

Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, et al. The Human Cell Atlas. Elife 2017;6:6.

Reid AJ, Talman AM, Bennett HM, Gomes AR, Sanders MJ, Illingworth CJR, et al. Single-cell RNA-seq reveals hidden transcriptional variation in malaria parasites. Elife 2018;7:7.

Reynolds NJ, Sinha A, Elias MS, Meggitt SJ. Translating translation into patient benefit for atopic eczema. Br J Dermatol 2016;175(Suppl. 2): 8−12.

Rosenberg AB, Roco CM, Muscat RA, Kuchina A, Sample P, Yao Z, et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. Science 2018;360:176−82.

Salzer MC, Lafzi A, Berenguer-Llergo A, Youssif C, Castellanos A, Solanas G, et al. Identity noise and adipogenic traits characterize dermal fibroblast aging. Cell 2018:1575−90.e22. 90175.

Shah S, Lubeck E, Zhou W, Cai L. seqFISH accurately detects transcripts in single cells and reveals robust spatial organization in the hippocampus. Neuron 2017;94:752−8.e1.

Shalek AK, Benson M. Single-cell analyses to tailor treatments. Sci Transl Med 2017;9:eaan4730.

Solé-Boldo L, Raddatz G, Schütz S, Mallm J-PJP, Rippe K, Lonsdorf AS, et al. Single-cell transcriptomes of the human skin reveal age-related loss of fibroblast priming. Commun Biol.Biol 2020;3:188.

Ståhl PL, Salmén F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science 2016;353:78−82.

Steuerman Y, Cohen M, Peshes-Yaloz N, Valadarsky L, Cohn O, David E, et al. Dissection of influenza infection in vivo by single-cell RNA sequencing. Cell Syst 2018;6:679−91.e4.

Stoeckius M, Zheng S, Houck-Loomis B, Hao S, Yeung BZ, Mauck WM 3rd, et al. Cell hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. Genome Biol.Biol 2018;19: 224.

Stubbington MJT, Rozenblatt-Rosen O, Regev A, Teichmann SA. Single-cell transcriptomics to explore the immune system in health and disease. Science 2017;358:58−63.

Suvà ML, Tirosh I. Single-cell RNA sequencing in cancer: lessons learned and emerging challenges. Mol Cell 2019;75:7−12.

Svensson V, Vento-Tormo R, Teichmann SA. Exponential scaling of single-cell RNA-seq in the past decade. Nat Protoc 2018;13:599−604.

Szabo PA, Levitin HM, Miron M, Snyder ME, Senda T, Yuan J, et al. Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. Nat Commun 2019;10:4706.

Tabib T, Morse C, Wang T, Chen W, Lafyatis R. SFRP2/DPP4 and FMO1/LSP1 define major fibroblast populations in human skin. J Invest Dermatol 2018;138:802−10.

Takahashi R, Grzenda A, Allison TF, Rawnsley J, Balin SJ, Sabri S, et al. Defining transcriptional signatures of human hair follicle cell states. J Invest Dermatol 2020;140:764−73.e4.

Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods 2009;6: 377−82.

Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science 2016;352:189−96.

Valdes-Mora F, Handler K, Law AMK, Salomon R, Oakes SR, Ormandy CJ, et al. Single-cell transcriptomics in cancer immunobiology: the future of precision oncology. Front Immunol 2018;9:2582.

van der Maaten L, Hinton G. Visualizing Data using t-SNE. J Mach Learn Res 2008;9:2579−605.

Vegh P, Haniffa M. The impact of single-cell RNA sequencing on understanding the functional organization of the immune system. Brief Funct Genomics 2018;17:265−72.

253

Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, et al. Single-cell reconstruction of the early maternal-fetal interface in humans. Nature 2018;563:347—53.

Vieira Braga FA, Miragaia RJ. Tissue handling and dissociation for single-cell RNA-seq. In: Proserpio V, editor. Single Cell Methods: Sequencing and Proteomics. New York, NY: Springer New York; 2019. p. 9—21.

Vitak SA, Torkenczy KA, Rosenkrantz JL, Fields AJ, Christiansen L, Wong MH, et al. Sequencing thousands of single-cell genomes with combinatorial indexing. Nat Methods 2017;14:302—8.

Vorstandlechner V, Laggner M, Kalinina P, Haslik W, Radtke C, Shaw L, et al. Deciphering the functional heterogeneity of skin fibroblasts using single-cell RNA sequencing. FASEB J 2020;34:3677—92.

Wolock SL, Lopez R, Klein AM. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. Cell Syst 2019;918: 281—91.e9.

Wu X, Yang B, Udo-Inyang I, Ji S, Ozog D, Zhou L, et al. Research techniques made simple: single-cell RNA sequencing and its applications in dermatology. J Invest Dermatol 2018;138:1004—9.

Xia C, Fan J, Emanuel G, Hao J, Zhuang X. Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression. Proc Natl Acad Sci USA 2019;116: 19490—9.

Xue D, Tabib T, Morse C, Lafyatis R. Transcriptome landscape of myeloid cells in human skin reveals diversity, rare populations and putative DC progenitors. J Dermatol Sci 2020;97:41—9.

254

Appendix B. Co-author of Science Research Article: Developmental cell programs are co-opted in inflammatory skin disease

RESEARCH ARTICLE SUMMARY

SINGLE-CELL GENOMICS

Developmental cell programs are co-opted in inflammatory skin disease

Gary Reynolds*, Peter Vegh*, James Fletcher*, Elizabeth F. M. Poyner*, Emily Stephenson, Issac Goh, Rachel A. Botting, Ni Huang, Bayanne Olabi, Anna Dubois, David Dixon, Kile Green, Daniel Maunder, Justin Engelbert, Mirjana Efremova, Krzysztof Polański, Laura Jardine, Claire Jones, Thomas Ness, Dave Horsfall, Jim McGrath, Christopher Carey, Dorin-Mirel Popescu, Simone Webb, Xiao-nong Wang, Ben Sayer, Jong-Eun Park, Victor A. Negri, Daria Belokhvostova, Magnus D. Lynch, David McDonald, Andrew Filby, Tzachi Hagai, Kerstin B. Meyer, Akhtar Husain, Jonathan Coxhead, Roser Vento-Tormo, Sam Behjati, Steven Lisgo, Alexandra-Chloé Villani, Jaume Bacardit, Philip H. Jones, Edel A. O'Toole, Graham S. Ogg, Neil Rajan, Nick J. Reynolds, Sarah A. Teichmann†, Fiona M. Watt†, Muzlifah Haniffa†

INTRODUCTION: Human skin provides vital protection from water loss and from external insult through structural adaptations and interplay with the innate and adaptive immune systems. The skin develops and functions in an aquatic environment in utero but rapidly adapts to a contrasting set of physical and pathogenic challenges after birth. The changes that take place across this complex multicellular system during development and upon perturbation by immune-mediated inflammatory diseases are poorly understood. A detailed study will facilitate the development of therapeutic interventions for inflammatory skin disease.

RATIONALE: We generated a comprehensive atlas of human skin in early prenatal life, in adulthood, and during inflammatory skin disease by profiling the transcriptomes of more than 500,000 single cells. We analyzed human embryonic skin between 7 and 10 post-conception weeks, healthy adult skin surplus from mammoplasty surgery, and skin biopsies from patients affected by atopic dermatitis (AD) and psoriasis, two common inflammatory skin diseases. Additionally, we performed single-cell T cell receptor analysis to assess T cell clonality in disease. Validation experiments were conducted at the protein level and used mass cytometry, flow cytometry, and immunostaining in situ of skin biopsies from healthy skin and patients with AD and psoriasis, including a cohort of AD patients before and during treatment with oral methotrexate.

RESULTS: Thirty-four cell states were identified in healthy human skin across the collective dataset, with dynamic changes in the nature and abundance of single-cell gene expression profiles identified across embryonic and adult life and upon perturbation during inflammatory skin disease. This resource can be accessed via an interactive browsable web portal, https://developmentcellatlas.ncl.ac.uk/datasets/hca_skin_portal. Analyses revealed that the immune system of first-trimester embryonic skin consists mainly of innate lymphocytes and macrophages. In adult skin, we defined two inferred trajectories for keratinocyte differentiation and the presence of endothelial cells that formed dilated postcapillary venules. We revealed a migratory dendritic cell (DC) signature in healthy adult skin that is conserved in murine DCs. The migratory DC signature was also evident in the developing human thymus and additional disease states. We identified clonally expanded disease-associated cytotoxic T cells (Tc IL13/IL22 cells) in lesional AD and Tc17/T helper 17 (T_H17) cells in lesional psoriasis. We demonstrated the reemergence of prenatal cellular programs mediated by Mac2 macrophages via the chemokine CXCL8 interacting with the venular capillary marker ACKR1 on VE3 vascular endothelial cells in diseased skin. This interaction is implicated in lymphocyte recruitment and angiogenesis. We identified and validated in situ the expansion of Mac2 and VE3 in lesional AD and lesional psoriasis skin, their close apposition in AD and psoriasis tissue, and their reduction in AD skin during methotrexate treatment, which aligned with an improvement in the clinical severity of disease in this patient cohort.

CONCLUSION: Our single-cell atlas of human skin from prenatal life, healthy adults, and AD and psoriasis patients highlights the dynamic nature of cutaneous homeostasis and immunity. Our study provides insights into perturbed and co-opted developmental cellular programs in inflammatory skin disease. These results may provide potential future translational targets to improve the diagnosis and molecular classification of these diseases and to guide treatment strategies. ∎
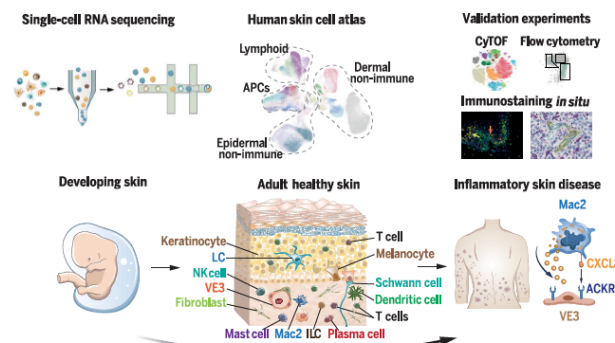
The human skin cell atlas. We used single-cell RNA sequencing to build human skin cell atlases across development, homeostasis, and disease; the analyses were supported by various validation modalities. By delineating cell states in embryonic skin, healthy adult skin, and inflammatory skin diseases (atopic dermatitis and psoriasis), we identified a reemergence of developmental programs in disease mediated by interactions between Mac2 and VE3. APCs, antigen-presenting cells; CyTOF, cytometry by time-of-flight; LC, Langerhans cell; NK, natural killer cell; VE, vascular endothelial cell; Mac, macrophage; ILC, innate lymphoid cell.

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.aba6500

SINGLE-CELL GENOMICS

Developmental cell programs are co-opted in inflammatory skin disease

Gary Reynolds[1]*, Peter Vegh[1]*, James Fletcher[1]*, Elizabeth F. M. Poyner[1,2]*, Emily Stephenson[1], Issac Goh[1], Rachel A. Botting[1], Ni Huang[3], Bayanne Olabi[1,4], Anna Dubois[1,2], David Dixon[1], Kile Green[1], Daniel Maunder[1], Justin Engelbert[1], Mirjana Efremova[3], Krzysztof Polański[3], Laura Jardine[1], Claire Jones[1], Thomas Ness[1], Dave Horsfall[1], Jim McGrath[1], Christopher Carey[1], Dorin-Mirel Popescu[1], Simone Webb[1], Xiao-nong Wang[1], Ben Sayer[1], Jong-Eun Park[3], Victor A. Negri[5], Daria Belokhvostova[5], Magnus D. Lynch[5], David McDonald[1], Andrew Filby[1], Tzachi Hagai[6], Kerstin B. Meyer[3], Akhtar Husain[7], Jonathan Coxhead[1], Roser Vento-Tormo[3], Sam Behjati[3,8], Steven Lisgo[1], Alexandra-Chloé Villani[9,10], Jaume Bacardit[11], Philip H. Jones[3,12], Edel A. O'Toole[13], Graham S. Ogg[14], Neil Rajan[1,2], Nick J. Reynolds[2,15], Sarah A. Teichmann[3,16]†, Fiona M. Watt[5]†, Muzlifah Haniffa[1,2,3]†

The skin confers biophysical and immunological protection through a complex cellular network established early in embryonic development. We profiled the transcriptomes of more than 500,000 single cells from developing human fetal skin, healthy adult skin, and adult skin with atopic dermatitis and psoriasis. We leveraged these datasets to compare cell states across development, homeostasis, and disease. Our analysis revealed an enrichment of innate immune cells in skin during the first trimester and clonal expansion of disease-associated lymphocytes in atopic dermatitis and psoriasis. We uncovered and validated in situ a reemergence of prenatal vascular endothelial cell and macrophage cellular programs in atopic dermatitis and psoriasis lesional skin. These data illustrate the dynamism of cutaneous immunity and provide opportunities for targeting pathological developmental programs in inflammatory skin diseases.

Human skin undergoes major adaptations as it transitions from a relatively pathogen-free aquatic environment in utero to provide mechanical and immunological protection in a nonsterile terrestrial environment. This function requires coordination by specialized cell types that are established during embryonic development. The cellular landscape of prenatal and adult skin, however, remains incompletely defined.

In cancer, developmental cell programs such as angiogenesis, proliferation, and invasion reemerge and lead to interaction between malignant cells and the surrounding stroma (1–4). The current consensus view on inflammatory skin disease pathogenesis is that the interplay between leukocytes and nonleukocytes is involved in disease initiation and progression (5). However, it is unknown whether cell states and gene programs observed in prenatal skin contribute to the pathogenesis of adult-onset inflammatory skin disorders. A detailed understanding of this process may provide a new perspective on inflammatory disease pathogenesis and potentially could identify novel therapeutic targets.

Single-cell genomics, such as RNA sequencing, provides an opportunity to dissect the complex cellular organization of human skin during development and in health and disease at a systems level. Studies of healthy skin to date have primarily focused on adult skin, restricted to specific cell lineages and limited cell numbers (6–10). Large-scale single-cell profiling of human skin should provide a transformative resource to understand aberrations in gene expression resulting from disease.

Deconstructing human skin

In this study, we used single-cell RNA sequencing (scRNA-seq) combined with strategic fluorescence-activated cell sorting (FACS) for comprehensive and deep profiling of healthy and inflamed adult skin. To maximize cell yield and viability from tissue dissociation from previous findings (11–13), we used 200-μm-thick healthy skin samples resulting from mammoplasty, which were separated into epidermis and dermis before dissociation (Fig. 1A and fig. S1A). We adapted our previous FACS gating strategy (12, 14) to isolate various cell fractions (keratinocytes, fibroblasts, and endothelial cells) and immune cells (myeloid and lymphoid cells) to upsample rare cell types for deep cell sampling. We ensured comprehensive capture, minimizing cell loss by placing these FACS gates contiguously (fig. S1B). FACS-isolated cells were characterized with the 10x Genomics platform. We also performed indexed plate-based Smart-seq2 profiling of all epidermal and dermal cells within the CD45⁺HLA-DR⁺ myeloid gate (fig. S1B). To compare cell states in healthy skin with inflammatory disease-induced perturbation, we performed scRNA-seq (10x Genomics) on all CD45⁻ and CD45⁺ cells from lesional and nonlesional skin from patients with atopic dermatitis (AD) and psoriasis (Fig. 1A and fig. S1C).

In total, 528,253 sequenced skin cells (n = 19) passed quality control and doublet exclusion (Fig. 1A). We detected on average ~3000 genes with the 10x Genomics platform and ~6000 genes per cell with Smart-seq2 (15) (fig. S2A). We excluded cells with <200 genes, >20% mitochondrial gene expression, and those identified as doublets (15). To account for biases due to batch effects, we performed data integration of healthy skin samples using BBKNN implementation within Scanpy (16, 17), which showed good sample mixing by UMAP visualization (Fig. 1B and figs. S1, D and E, and S2B). We performed graph-based Leiden clustering and derived differentially expressed genes to annotate the cell clusters, from which 34 cell states were identified (Fig. 1, B and C, and table S1). We were able to identify these cell states by deconvolution analysis using AutoGeneS (18) of adult healthy bulk RNA-seq (fig. S2C). The 34 cell states were discernible even after removal of stress response genes associated with tissue dissociation (19) (fig. S2D). We note the impact of different tissue dissociation protocols on gene expression, as previously reported (20). We found that the same cell states could be identified in a small dataset using the Miltenyi dissociation protocol (7) and through the statistical power and

256

Fig. 1. Deconstructing human skin. (A) Numbers of cells profiled by scRNA-seq and mass cytometry for each condition and schematic of sampling locations. AD, atopic dermatitis. **(B)** UMAP visualization showing all cell states found in the healthy adult scRNA-seq dataset (n = 5). KC, keratinocyte; Fb, fibroblast; VE, vascular endothelium; LE, lymphatic endothelium; ILC, innate lymphoid cell; NK, natural killer cell; Tc, cytotoxic T cell; T_H, T helper cell; T_reg, regulatory T cell; Mac, macrophage; Inf., inflammatory; DC, dendritic cell; LC, Langerhans cell; Mono mac, monocyte-derived macrophage; Mig., migratory; MoDC, monocyte-derived dendritic cell. **(C)** Dot plot showing the expression of discriminatory markers for each cell state in (B). *CD8A*, *CD163*, *CD14*, and *CCR7* were chosen for CyTOF protein validation. **(D)** Top: UMAP visualization of healthy

adult cell states overlaid on developing cell states (n = 7). Bottom: UMAP visualization of developing cell states overlaid on healthy adult cell states. Cells underlaid are shown in gray. **(E)** Median prediction score of developing skin cell states using healthy adult skin as reference derived from the TransferAnchors function in Seurat. Triangles denote developing skin cells; circles denote adult skin cells. **(F)** Bar charts showing the proportions of corresponding cell states found in adult and developing skin. We used generalized linear modeling on a quasi-binomial distribution to compare proportions between developing and adult skin, and found statistically significant changes in vascular endothelium, Schwann cells, and melanocytes ($P = 3.1 \times 10^{-7}$); keratinocytes ($P = 3.1 \times 10^{-4}$); LCs and DCs ($P = 1.4 \times 10^{-4}$); and fibroblasts ($P = 4.0 \times 10^{-4}$). ***$P < 0.001$.

resolution provided by our large dataset to discern rare cell states (fig. S2, E and F). However, our analysis of interfollicular mammoplasty skin sampled to the top layer of the reticular dermis may not have adequately profiled all skin hair follicle and appendageal cells.

We selected genes encoding surface proteins (*CD8A*, *CD163*, *CD14*, and *CCR7*) and additional antigens (where antibodies were commercially available) to derive a cytometry by time-of-flight (CyTOF) panel for protein level validation and frequency assessment of the major cell states on four additional donors (fig. S3, A to E). Gene expression on the relevant cell states was consistent with CyTOF analysis (fig. S3F).

To evaluate the establishment of specific cell states during development and their temporal evolution in adult skin, we compared our adult skin scRNA-seq data with our embryonic/fetal [7 to 10 post-conception weeks (PCW); n = 7] scRNA-seq data (Fig. 1D) (*21*). We used the TransferAnchors function in Seurat to integrate adult and fetal skin cell states (Fig. 1E) (*22*) and calculated the proportional representation of the equivalent cell states in healthy developing and adult skin (Fig. 1F). Our collective dataset of human fetal, adult, and diseased skin cells provides a foundational resource and can be explored using an interactive web portal, https://developmentcellatlas.ncl.ac.uk/datasets/hca_skin_portal.

Transition from innate to adaptive lymphocytes during skin development

Skin T cells consist of three subtypes: cytotoxic (Tc) cells expressing *CD8A/B*, helper (T$_H$) cells expressing *CD4* and *CD40LG*, and regulatory (T$_{reg}$) cells expressing *FOXP3*, *TIGIT*, and *CTLA4* (Fig. 2, A and B) (*23*, *24*). We identified four clusters of innate lymphocytes in adult healthy skin that were CD161(*KLRB1*)$^+$CD3 (*CD3D/CD3G*)$^-$, consisting of ILC1/3, ILC2, ILC1/natural killer (NK), and NK (*KLRD1*$^+$, *GNLY*$^+$, *PRF1*$^+$, *GZMB*$^+$, and *FCGR3A*$^+$) cells (Fig. 2, A and B). ILC1/NK cells have overlapping features of ILC1 and NK cells, as described (*25*, *26*). Plasticity within ILC1 and ILC3 is also recognized, as reflected in our annotation of ILC1/3 (*27*). ILC2 (*IL7R*, *PTGDR2*, and *GATA3*) has the most distinct signature in our data and in existing literature (Fig. 2B) (*26*).

In contrast to adult skin, the fetal skin lymphoid compartment is predominantly populated by ILCs (Fig. 1, D and F) between 7 and 10 PCW, prior to the development of the thymus, bone marrow, and spleen, where T and B lymphocytes differentiate. Fetal NK cells correlate with adult NK cells but express higher levels of *GZMM* and *GZMK* (Figs. 1E and 2B), which suggests that they may be functionally competent (*21*). Fetal skin ILCs (*IL7R*$^+$, *RORC*$^+$, and *KIT*$^+$) resemble adult skin ILC3 (Fig. 2B) (*28*).

To evaluate the impact of epidermal versus dermal microenvironment on T cells, we compared the differentially expressed genes of T cells in the two compartments (Fig. 2C). Epidermal T cells up-regulated the expression of genes associated with skin tissue residency (*RGS1* and *PPP1R15A*) (*29*), effector memory (*CD44* and *ID2*) (*29*, *30*), T cell activation (*TNFRSF18*) (*31*), and inhibition of T cell response (*CD96*, *TSC22D3*, and *DUSP4*) (*32*–*34*), in keeping with previous suggestions that resident memory T cells are poised to mount an effective immune response but express inhibitory molecules to prevent disadvantageous responses to nonpathogenic antigens (*35*). In contrast, dermal T cells express interferon-stimulated genes (*IFITM1*, *IFI6*, and *LY6E*) (*36*) and transcriptionally correlate closer to blood T cells than epidermal T cells do (fig. S4, A and B). Dermal T$_{regs}$ show high mRNA and protein expression of the circulating central memory T cell marker CD62L (*SELL*) (Fig. 2C and fig. S3A).

Disease-associated and clonal T cells

In atopic dermatitis (AD), cytotoxic T cells expressing *IL13*, *IL22*, and *IFNG* (Tc IL13/IL22) are found in both lesional and nonlesional skin but are significantly enriched in lesional skin (P = 0.04, likelihood ratio test). Similarly, in psoriasis, T cells expressing *IL17A*, *IL17F*, *IFNG*, *IL22*, and *IL26* (Tc17/T$_H$17) are found in both lesional and nonlesional skin but are significantly enriched in lesional skin (P < 0.001, likelihood ratio test) (Fig. 2D and fig. S4C). Tc17/T$_H$17 cells are dominant in the epidermis of lesional psoriasis skin, as validated by flow cytometry (Fig. 2D and fig. S4D). These cells express genes characteristic of activated and pathogenic T$_H$17 cells (*KLRB1*, *RBPJ*, and *CXCL13*) (*37*–*39*) (fig. S4E). Tc IL13/IL22 cells are dominant in the dermis of AD lesional skin (Fig. 2D) and express amphiregulin (*AREG*, a member of the epidermal growth factor family) (*40*), skin tissue residency genes (*RGS1*, *NR4A1*, *NR4A2*) (*29*), and effector and activated T cell genes (*ID2*, *PRDM1*, *MAP3K8*, *DUSP2*) (*30*, *41*–*44*) (fig. S4E). To further extend our observation to a larger patient cohort, we used the AutoGeneS tool (*18*) to deconvolute cellular heterogeneity in published AD and psoriasis bulk RNA-seq data (*45*). In accordance with our findings, we observed the presence of a Tc IL13/IL22 signature in lesional AD skin and a Tc17/T$_H$17 signature in lesional psoriasis skin (Fig. 2E).

For both AD and psoriasis, there was a significantly higher proportion of T cell clones shared between nonlesional and lesional skin within donors versus between donors (P < 0.01, two-sample t test) (fig. S4F). In AD and psoriasis lesional skin, disease-associated Tc IL13/22 and Tc17/T$_H$17 cells exhibited significantly higher clonality [P < 0.05 using quasi-

binomial statistics to model proportions, followed by analysis of variance (ANOVA)] and the lowest diversity compared to T$_H$, T$_{reg}$, and Tc cell subsets (Fig. 2F and fig. S4G). Lesional clonal T cells had higher expression of co-stimulation genes (*CD63*, *TNFRSF18*, and *JAML*) and T cell receptor signaling (*EVL*, *LAT*, *LCK*, and *JAK1*) than did nonclonal T cells (Fig. 2G).

Mononuclear phagocytes in adult and developing skin

We observed 14 states of mononuclear phagocytes (MPs) in human skin (Fig. 3, A and B, and fig. S5A) that we annotated by aligning skin and blood MPs using the TransferAnchors function in Seurat (fig. S5B) and expression of MP marker genes (*46*) (fig. S5C). This work revealed the limitations of currently used surface markers and FACS gates to adequately resolve skin MP heterogeneity (fig. S5A).

Two macrophage cell states expressing *CD68* are present in healthy skin. Mac1 shows higher expression of complement transcripts (*C1QB* and *C1QC*) and scavenger receptors (*CD163* and *MARCO*), whereas Mac2 is characterized by the expression of *F13A1* and transcription factors associated with alternative activation and suppression of immune responses (*NR4A1*, *NR4A2*, and *KLF4*) (Fig. 3B and fig. S5C) and, notably, is more closely aligned with fetal macrophages (Fig. 1E).

We observed dendritic cells 1 and 2 (DC1, DC2) and Langerhans cells (LCs) in embryonic skin as early as 7 PCW, prior to bone marrow hematopoiesis, but macrophages were the dominant MP in first-trimester skin (Fig. 1, D to F, and fig. S5D). Interestingly, embryonic/fetal LCs are enriched for macrophage-related genes such as *C1QC*, *FCGR2A*, and *CTSB* (Fig. 3B) and correlate poorly with adult LCs (Fig. 1E). This lends support to a differential origin of prenatal LCs from yolk sac and fetal liver progenitors, as previously reported in mice (*47*), in contrast to the bone marrow–derived hematopoietic stem cell origin of some adult human LCs (*48*).

Migratory dendritic cell signature is conserved across species and augmented in disease states

In the steady state, skin dendritic cells (DCs) undergo a continual process of homeostatic maturation that is required for the induction of tolerance to innocuous environmental antigens (*49*). This is accompanied by their migration to skin draining lymph nodes through lymphatic vessels, a process dependent on CCR7 (*13*). Partition-based approximate graph abstraction (PAGA) analysis revealed three branches of differentiation: LCs, myeloid DCs (DC1 and DC2), and monocyte-derived DCs (moDCs) (Fig. 3C). The clusters at the convergence of these branches [moDC3, LC4, and migratory

258

Fig. 2. Skin innate lymphoid cells, T lymphocytes, and T cell receptor analysis.
(**A**) UMAP visualization of lymphoid cells in healthy adult skin (*n* = 5). ILC, innate lymphoid cell; NK, natural killer cell; Tc, cytotoxic T cell; T_H, T helper cell; T_{reg}, regulatory T cell. (**B**) Dot plot showing expression of marker genes of cell states found in adult skin (circles) shown in (A) and their developmental counterparts (triangles), separated by the dashed line. (**C**) Dot plot showing the differentially expressed genes in T cell subsets between epidermis and dermis in healthy adult skin (*n* = 5). Epi, epidermis. (**D**) Bar charts showing the proportions of lymphoid cell states in healthy and diseased skin. Tc IL13/IL22, cytotoxic T cells expressing *IL13* and *IL22*; Tc17/T_H17, T cells expressing *IL17A* and *IL17F*. Lesional dermal AD skin Tc IL13/IL22 was compared to nonlesional dermal AD skin (modeled counts as negative binomial and analyzed by ANOVA, *P = 0.04); psoriasis epidermal lesional skin Tc17/T_H17 was compared to psoriasis nonlesional epidermis (***P < 0.001). Significance of other cell states was not tested.

HA, healthy adult; AD, atopic dermatitis; P, psoriasis; L, lesional; NL, nonlesional. (**E**) Proportion of cells with Tc IL13/IL22 and Tc17/T_H17 signatures present in bulk RNA-seq data from GSE121212. Modeling the data on a quasibinomial distribution, lesional AD is enriched for Tc IL13/IL22 ($P = 3.4 \times 10^{-2}$) and T_{reg} ($P = 3.2 \times 10^{-2}$) relative to nonlesional skin. Lesional psoriasis is enriched for Tc17/T_H17 ($P = 1.6 \times 10^{-5}$), T_{reg} ($P = 4.6 \times 10^{-7}$), and Tc ($P = 2.5 \times 10^{-4}$) relative to nonlesional. *P* values were calculated using a likelihood ratio test. (**F**) Bar charts showing clonotype size in each T cell subset in lesional AD (left) and psoriasis (right). The color of the bar relates to the size of the clonotype. Modeling the data on a quasibinomial distribution, lesional AD dermis Tc IL13/IL22 cells have a higher proportion of clonotypes with ≥2 cells relative to T_H ($P = 3.3 \times 10^{-2}$), and lesional psoriasis Tc17/T_H17 cells have a higher proportion of clonotypes with ≥2 cells than do T_{regs} ($P = 1.6 \times 10^{-2}$); other comparisons were not significant. *P* values were calculated using a likelihood ratio test. Bars across the top show significance between cell types. (**G**) Dot plot showing the expression of genes that were significantly differentially expressed between nonclonal and clonal T cells in AD (top) and psoriasis (bottom). Hatched and solid colored circles indicate nonclonal and clonal T cells, respectively. *P < 0.05, ***P < 0.001.
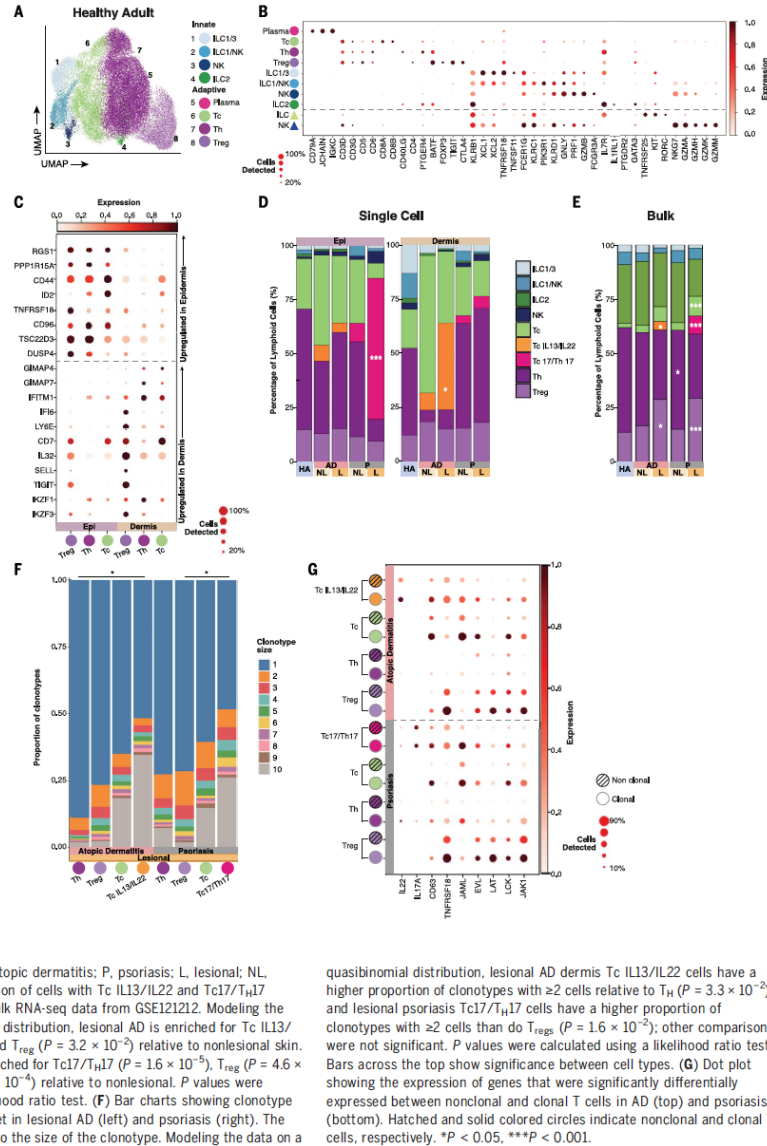
259

Fig. 3. Dermal and epidermal mono-nuclear phagocytes. (A) UMAP visualization of different anti-gen-presenting cell (APC) states found in healthy adult skin (n = 5). See Fig. 1B for abbreviations. (B) Dot plot showing the expression of differentially expressed genes characterizing adult healthy skin cell states (circles) shown in (A) and their developmental counterparts (triangles), separated by the dashed line. (C) Top: Abstracted graph (PAGA) showing connectivity between adult healthy skin DC clusters. The size of the nodes is proportional to cluster size; edge thickness is proportional to the strength of the connection between nodes. Bottom: enrichment of gene signatures for murine splenic Xcr1+DC (DC1) and dermal CD11c+ (DC2) in each node. Av, average. (D) Dot plot of enrichment of gene signature of APC cell types in adult human disease and in the developing thymus. SF, synovial fluid. (E) Bar charts showing the proportions of Mac1 and Mac2 in adult healthy, AD, and psoriasis skin. HA, healthy adult; AD, atopic dermatitis;

P, psoriasis; L, lesional; NL, nonlesional. Mac2 cells are significantly expanded in both lesional AD and psoriasis skin, and Mac1 cells are significantly reduced in both lesional AD and psoriasis skin [Mac1, $P = 5.4 \times 10^{-3}$ (AD lesional), 3.3×10^{-5} (psoriasis lesional); Mac2, $P = 6.3 \times 10^{-3}$ (AD nonlesional), 6.0×10^{-5} (AD lesional), 1.4×10^{-6} (psoriasis lesional)]. P values were calculated using a likelihood ratio test. (F) Proportion of cells with Mac2 signature present in bulk RNA-seq data from GSE121212. We used generalized linear modeling on a quasibinomial distribution to compare proportions of predicted Mac2 cells between healthy and lesional skin and showed statistically significant expansion of Mac2 in lesional AD ($P = 1.7 \times 10^{-7}$). P values were calculated using a likelihood ratio test. (G) Prediction score for alignment using CCA (Seurat) between developing gut, kidney, liver, skin, and thymus

macrophages with Mac1 and Mac2 in healthy adult skin. (H) Network visualizations of pathways conserved between developing skin macrophages and Mac2 in AD and psoriasis. Network nodes are colored by enrichment score ($q < 0.05$) and represent individual enriched gene sets; edges represent genes shared between nodes (intersect ≥10%). (I) Jitter plot displaying the number of positive F13A cells in 4×10^4 µm^2 of healthy adult (n = 7), AD (n = 12), and psoriasis (n = 6) skin (P between healthy adult and AD = 6.6×10^{-4} ; P between healthy adult and psoriasis = 1.2×10^{-3}). (J) Jitter plot displaying the number of positive F13A cells in 4×10^4 µm^2 of AD skin (n = 5) before treatment with methotrexate and 9 to 16 days and 12 weeks post-treatment (P between pretreatment and 12 days post-treatment = 1.5×10^{-2}, P between 9 to 16 days and 12 weeks post-treatment = 2.0×10^{-3}). *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

260

DC (Mig. DC)] express transcripts associated with DC maturation (*CD83, CCR7, LAMP3, CD40,* and *CD86*), immunoregulation (*CD274, IDO1, CD200, PDCD1LG2,* and *SOCS1*) (Fig. 3B and table S1) (*50*), and migration (*FSCN1, PLGRKT, TRAF1, BCL2A1, CFLAR,* and *REL*) (Fig. 3C) as in migratory murine dermal DC2 and splenic DC1 (*49, 51, 52*). Acquisition of this common gene signature is associated with loss of genes conferring subset identity in mice (*53*), which we also observe here for moDC3, LC4, and Mig. DC (Fig. 3B).

Surprisingly, the migratory DC signature is also present across disease states including tonsillitis, ascites (*53, 54*), lung cancer (*50*), and rheumatoid arthritis (Fig. 3D). We previously reported the expression of migratory genes in fetal thymic medullary DCs (*55*), which suggests that developmental gene programs are used in adult tissue homeostasis and augmented in disease.

Fetal macrophage program in AD and psoriasis

We observed an increase in Mac2 in AD and psoriasis skin (Fig. 3E and fig. S5E), which was corroborated in a larger patient bulk RNA-seq dataset (Fig. 3F). Adult healthy skin Mac2 aligned with fetal skin, gut, kidney, liver, and thymus macrophages (Fig. 3G). This led us to hypothesize a cellular program shared between fetal macrophages in AD skin and Mac2 cells in psoriasis skin, which are significantly differentially expressed relative to other skin cells. We derived 91 significantly conserved genes (Seurat FindConservedMarkers using negative binomial test, *P* < 0.05) between analogous macrophage clusters in developing skin and lesional AD and psoriasis skin (table S5). This revealed genes related to stress (*DNAJB1, HSPA1B, HSPA1A, JUN,* and *FOSB*), chemotactic (*CCL4L2, CCL4, CCL3L1,* and *CCL3*), and angiopoietin (*EGR1* and *PTGS2*) signaling. Gene Ontology analysis revealed significantly enriched gene set clusters (hypergeometric test *q* value < 0.05) relating to the regulation of angiogenesis, leukocyte chemotaxis, and transforming growth factor-β (TGF-β) signaling (Fig. 3H and table S6). The role of macrophages in tissue homeostasis and regeneration is gaining recognition (*56*). Our findings add insight into how macrophage programs that support angiogenesis and leukocyte seeding in tissues during fetal development reemerge during AD and psoriasis pathogenesis.

To confirm the role of Mac2 in disease pathogenesis in vivo, we analyzed the abundance of Mac2 in healthy, AD, and psoriasis lesional skin, as well as during AD resolution resulting from systemic treatment with methotrexate (Fig. 3, I and J). We leveraged the marker F13A for Mac2 from our scRNA-seq data (Fig. 3B) and enumerated F13A-expressing Mac2 cells by immunohistochemistry. This revealed a significant increase in Mac2 in AD and

psoriasis lesional skin relative to healthy skin, as well as a decline in Mac2 12 weeks after commencement of methotrexate treatment, in parallel with a reduction in patients' clinical Eczema Area and Severity Index (EASI) score (Fig. 3J and fig. S5F).

In both AD and psoriasis, LC1 has the highest enrichment of cell cycle genes (fig. S5, G and H). To validate our findings, we examined LC proliferation in healthy, AD, and psoriasis epidermis. We found that Ki67+ Langerin+ cells increased in AD and psoriasis lesional skin (fig. S5I), consistent with previous findings (*57, 58*). Using FACS index data, we determined that LC1 is enriched within the Langerin+CD1a^lo^CD11c^lo^ gate distinct from Langerin+CD1a^hi^ LCs (fig. S5A). In contrast, epidermal HLA-DR+CD1a−ang−CD11c+CD1c+ cells are predominantly moDCs and correspond with non–LC-like epidermal cells potent at stimulating T cell proliferation, proinflammatory cytokine production, and transmission of HIV to CD4+ T cells (*59*).

Keratinocyte differentiation in healthy and diseased skin

We characterized four groups of keratinocytes: undifferentiated, proliferating, differentiated, and inflammatory differentiated cells (differentiated KC*) (Fig. 4A). Undifferentiated keratinocytes transcribe basal epidermal proteins (*KRT5* and *KRT14*) and are abundant in the CD49f^hi^ FACS gate (Fig. 1C). Proliferating keratinocytes (*CDK1+* and *PCNA+*) have lower expression of suprabasal cell transcripts (e.g., *KRT1, KRT10*) that characterize differentiated keratinocytes (Fig. 4B). Inflammatory differentiated keratinocytes coexpress lower levels of undifferentiated (*TP63* and *ITGA6*) and differentiated (*KRT1* and *KRT10*) transcripts but additionally express *ICAM1, TNF,* and *CCL20* (Fig. 4B). The gene expression patterns of these keratinocyte subgroups are in agreement with their spatial arrangement in the epidermis (Fig. 4C) (Human Protein Atlas, www.proteinatlas.org) and with a human epidermal scRNA-seq dataset (*60*) (fig. S6A). First-trimester human epidermis, comprising "basal" undifferentiated keratinocyte progenitors overlaid by the periderm, expresses keratin genes and proteins of simple epithelium (keratins 8, 18, and 19) (Fig. 4B and fig. S6B) (*61*).

Force-directed graph (FDG) and PAGA analyses revealed dual inferred differentiation trajectories from the stem cell gene (*TP63, PPP3CA,* and *CAV1/2*)–enriched basal keratinocytes into terminally differentiated keratinocytes expressing *CNFN, FLG,* and *IVL* (Fig. 4D, fig. S6C, and table S1) (*61, 62*). One arm shows high levels of lamellar body (LB)–related transcripts such as *ABCA12, CKAP4,* and *CLIP1* that characterize late epidermal differentiation, and the other arm expresses lower levels of LB-related transcripts (*63*) (Fig. 4E). *IRF1*

and *SOX9*, transcription factors that are differentially expressed in the two pathways (fig. S6D), mark distinct cells by immunofluorescence analysis of healthy skin (Fig. 4F and fig. S6E). The statistically significant differentially expressed genes identified with Monocle (*64*) across keratinocyte differentiation (fig. S6C) recapitulate previous reports in human and mouse (*65*).

Notably, keratinocytes expressing LB-related transcripts coexpress genes associated with autosomal recessive congenital ichthyosis, such as *ABCA12, NIPAL4, SLC27A4,* and *TGM1* (Fig. 4E). However, analysis of fetal keratinocytes showed little to no expression of these congenital ichthyosis-related genes, which suggests that disease onset at the molecular level begins only after 10 PCW (Fig. 4E). This is in keeping with the absence of a granular layer in first-trimester fetal epidermis, where *LOR, FLG, IVL,* and genes required for lamellar body production are expressed (*65, 66*). Inflammatory differentiated keratinocytes express higher levels of genes associated with inflammatory ichthyoses and severe atopy, such as *NIPAL4* and *SPINK5* (Fig. 4E, cluster 8) (*67*).

Both AD and psoriasis lesional skin were enriched for differentiated keratinocytes (Fig. 4G), as supported by deconvolution of bulk RNA-seq data from an extended patient cohort (Fig. 4H). Differentially expressed gene analysis revealed lower expression of stem cell and basal keratinocyte genes (*CAV1/2, KRT14,* and *DUSP10*) but higher expression of commitment genes (*FOS, JUNB, CDKN1A, MAFB*) in undifferentiated lesional psoriasis and AD keratinocytes (fig. S6F). These observations agree with previous reports (*68, 69*) and suggest rapid transition and differentiation of keratinocytes in AD and psoriasis lesional skin. Lesional differentiated keratinocytes additionally expressed inflammatory transcripts including alarmins (*S100A7, S100A8, S100A9*), serpins (*SERPINB4, SERPINB13*), and interferon response genes (*IFI27, IFITM1*) (fig. S6F). The proportion of inflammatory differentiated keratinocytes, resembling previously described CCL20-expressing keratinocytes in murine inflammatory skin disease induced by subcutaneous interleukin-17 (IL-17) and tumor necrosis factor (TNF-α) injection (*70*), are expanded in psoriasis skin (Fig. 4G).

Stromal cell heterogeneity

We next interrogated the heterogeneity within fibroblasts, vascular and lymphatic endothelial cells, and Schwann cells (Fig. 5, A and B). Fibroblasts dominated the nonimmune cell population in developing skin; the proportional representation of melanocytes, Schwann cells, and lymphatic and vascular endothelial cells was increased in adult skin (Fig. 1F).

Three fibroblast subsets expressing extracellular matrix (ECM)–related genes such as

261
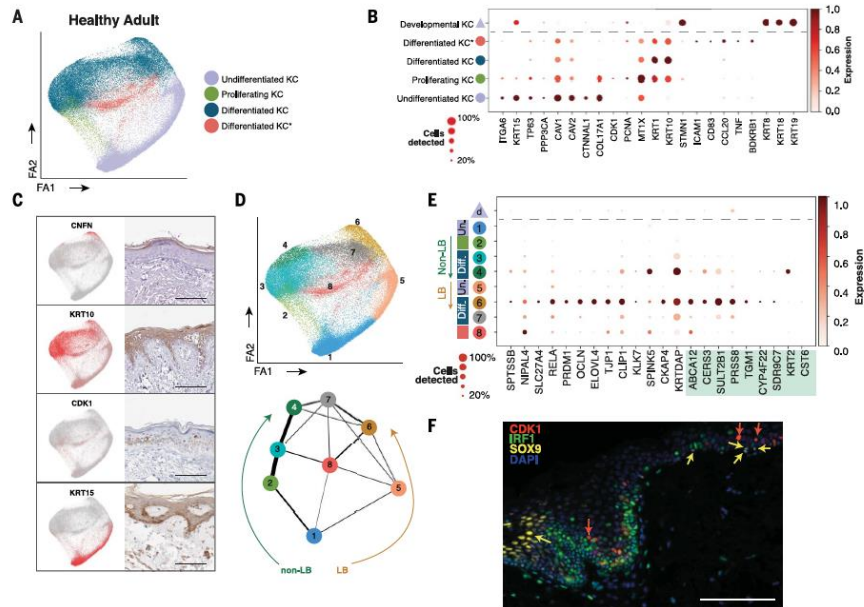
Fig. 4. Keratinocyte cell states in health, AD, and psoriasis.
(**A**) Force-directed graph (FDG) visualization of the different keratinocyte cell states found in healthy adult skin ($n = 5$). KC, keratinocyte. The asterisk indicates the cell state with inflammatory markers. (**B**) Dot plot showing the expression of differentially expressed genes characterizing keratinocyte states in healthy adult skin (circles) shown in (A) and developmental keratinocytes (triangle), separated by the dashed line. (**C**) FDG feature plots showing gene expression of healthy adult skin keratinocyte states shown in (A), together with images of these markers in situ, from the Human Protein Atlas. Scale bars, 100 μm. (**D**) Top: FDG in (A) annotated by Leiden clustering of eight groups: undifferentiated KC (clusters 1 and 5), proliferating (cluster 2), differentiated KC (clusters 3, 4, 6, and 7), and differentiated KC* (cluster 8). Bottom: PAGA showing the relative connectivity between the keratinocyte clusters. Arrows indicate the two differentiation pathways of basal keratinocytes to suprabasal. LB, lamellar body. (**E**) Dot plot of genes related to lamellar body production and ichthyosis (green box) expressed by healthy adult keratinocyte states (circles) shown in (D), as well as fetal keratinocytes (triangle). Un., undifferentiated; Diff., differentiated. (**F**) Immunofluorescence staining of healthy adult skin for CDK1 (red), IRF1 (green), SOX9 (yellow), and DAPI (blue). Red and yellow arrows indicate CDK1+ and SOX9+ cells, respectively, in suprabasal layers. Image is representative of $n = 3$ donors. Scale bar, 100 μm. (**G**) Bar charts showing the proportions of the keratinocyte cell states in healthy and diseased skin. Undifferentiated KCs, $P = 6.1 \times 10^{-4}$ (AD lesional), 3.0×10^{-5} (psoriasis lesional); differentiated KCs, $P = 6.5 \times 10^{-16}$ (AD lesional), 6.5×10^{-7} (psoriasis nonlesional), 8.8×10^{-20} (psoriasis lesional); differentiated

KCs*, $P = 2.3 \times 10^{-2}$ (psoriasis nonlesional), 1.5×10^{-4} (psoriasis lesional); proliferating KCs, $P = 6.5 \times 10^{-10}$ (AD nonlesional), 8.1×10^{-14} (psoriasis lesional). Populations are compared to those in healthy adults. P values were calculated using a likelihood ratio test. (**H**) Percentage of cells with undifferentiated, differentiated, and proliferating keratinocyte signatures present in bulk RNA-seq data from GSE121212. Generalized linear modeling on a quasibinomial distribution was used to compare proportions of predicted keratinocyte subsets between healthy, nonlesional, and lesional skin and showed statistically significant expansion of differentiated keratinocytes in nonlesional AD ($P = 1.1 \times 10^{-3}$), lesional AD ($P = 9.6 \times 10^{-10}$), nonlesional psoriasis ($P = 2.1 \times 10^{-5}$), and lesional psoriasis ($P = 2.0 \times 10^{-16}$). *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

Fig. 5. Stromal and endothelial cells. (**A**) UMAP visualization of the nonimmune, non-keratinocyte cell states found in healthy adult skin (*n* = 5). Fb, fibroblast; LE, lymphatic endothelium; VE, vascular endothelium. (**B**) Dot plot showing the expression of differentially expressed genes characterizing adult healthy skin cell states (circles) shown in (**A**) and their developmental counterparts (triangles), separated by the dashed line. (**C**) 3D reconstruction of Z-stacked images of whole-mount immuno-fluorescence staining of dermis for CD31 (PECAM1, red), γ-synuclein (SNCG, green), and DRAQ5 (blue). White cube outlined in central image represents 40 μm × 40 μm × 40 μm. (**D**) Bar charts showing the proportions of VE in healthy adult and diseased skin [VE1, *P* = 5.5 × 10⁻⁶ (psoriasis nonlesional), 5.0 × 10⁻¹⁵ (psoriasis lesional); VE2, *P* = 1.9 × 10⁻² (psoriasis nonlesional); VE3, *P* = 1.3 × 10⁻² (AD nonlesional), 1.5 × 10⁻⁴ (AD

lesional), 1.3 × 10⁻³ (psoriasis nonlesional), 6.2 × 10⁻⁹ (psoriasis lesional)]. *P* values were calculated using a likelihood ratio test. (**E**) Proportion of cells with VE3 signature present in bulk RNA-seq data from GSE121212. The proportion of VE3 increased in both lesional AD (*P* = 9.1 × 10⁻⁴) and psoriasis (*P* = 8.2 × 10⁻⁴). *P* values were calculated using a likelihood ratio test. (**F**) Prediction score for alignment using CCA (Seurat) between developing gut, kidney, liver, skin, and thymus VE with VE1, VE2, and VE3 in healthy adult skin. (**G**) Network visualization of pathways conserved between developing skin VE and VE3 in AD and psoriasis. Network nodes are colored by enrichment score (*q* < 0.05) and represent individual enriched gene sets; edges represent genes shared between nodes (intersect ≥10%). (**H**) Jitter plot displaying the number of positive ACKR1 cells in 1.5 × 10⁵ μm² of healthy adult (*n* = 6), AD (*n* = 12), and psoriasis (*n* = 6) skin (ANOVA *P* value between healthy
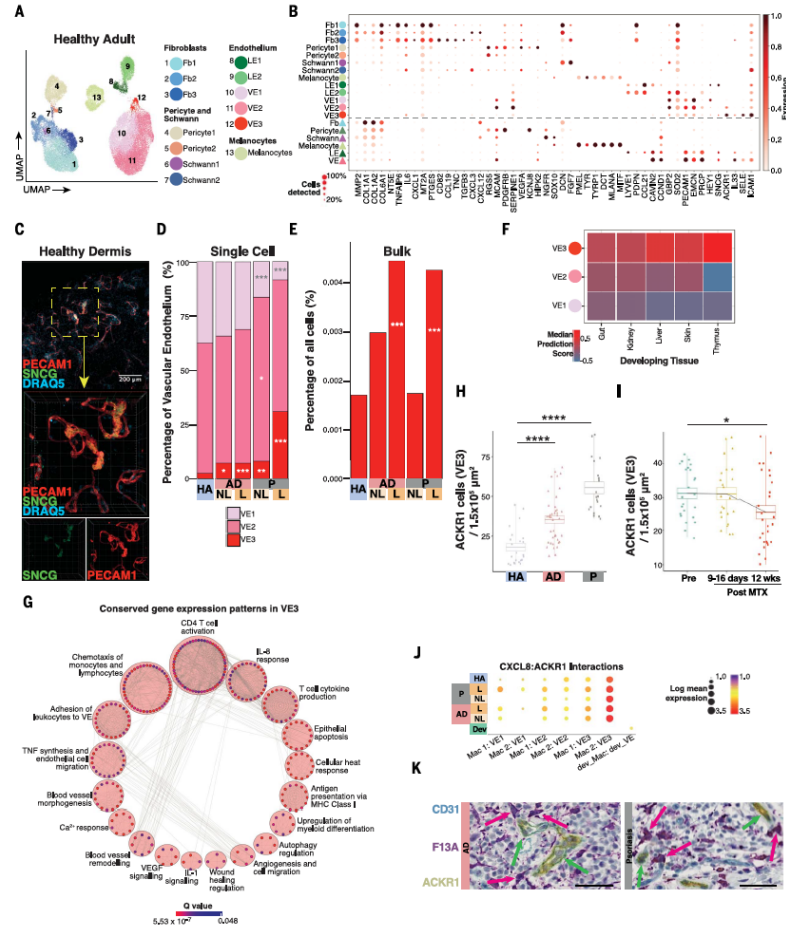
adult and AD, 3.0 × 10⁻⁵; between healthy adult and psoriasis, 2.4 × 10⁻¹³). (**I**) Jitter plot displaying the number of positive ACKR1 cells in 1.5 × 10⁵ μm² of AD skin (*n* = 5) before treatment with methotrexate and 9 to 16 days and 12 weeks post-treatment (ANOVA *P* value between pretreatment and 12 weeks post-treatment, 5.0 × 10⁻²). (**J**) Interactions between macrophage and vascular endothelium subsets predicted by CellPhoneDB. Color and size indicate log₂ mean expression, averaged across the two clusters. Dev, developing skin. (**K**) Immuno-histochemical staining of AD (left) and psoriasis (right) skin for F13A (purple), ACKR1 (yellow), and CD31 (teal), showing the close proximity of Mac2 and VE3. Pink arrows point to F13A-positive macrophages; green arrows point to CD31/ACKR1-positive vascular endothelial cells. Scale bars, 20 μm. Representative images from *n* = 4 for AD and *n* = 6 for psoriasis are shown. *P < 0.05, **P < 0.01, ***P < 0.001.

263

MMP2, COL1A1, COL1A2, and NT5E (encodes CD73) are present in healthy human skin, dominated by Fb1 fibroblasts with minor populations of Fb2 and Fb3 fibroblasts (Fig. 5, A and B). Dermal fibroblast heterogeneity encompassing structural and immunomodulatory subtypes has been previously reported at the single-cell level (6, 8). Interestingly, fibroblasts in fetal skin express more genes relating to Fb2 adult fibroblasts, including COL1A1, COL1A2, and COL6A1, which suggests that they are functionally specialized toward ECM remodeling and maintenance (Fig. 5B). Furthermore, in AD and psoriasis lesional and nonlesional skin, Fb2 fibroblasts are significantly enriched relative to healthy skin and have up-regulated expression of the chemokines CXCL12 and CCL19, in keeping with recent reports (71) (fig. S7, A and B).

Specialized vascular endothelium mediates leukocyte trafficking

Endothelial cells in the healthy adult dermis constitute the vascular endothelium (PECAM1, EMCN) and lymphatic endothelium (LYVE1, PDPN) (Fig. 5, A and B, and table S1). There are two subclusters of lymphatic endothelial cells defined by the differential expression of CCL21 and PDPN, which are higher in LE1 and LE2, respectively (Fig. 5, A and B), the latter resembling PDPN+ collecting lymphatic vessels in human dermis (14). Notably, LE1 cells express higher levels of the chemoattractant CCL21, which mediates DC migration into skin draining lymph nodes, as well as angiogenesis factors CAVIN2 and CCND1, further supporting their function as initial afferent lymphatics (72) (Fig. 5B and table S1).

Three distinct states of PECAM1 (CD31)–expressing vascular endothelial cells (VE1, VE2, VE3) are present in adult dermis. VE3, which forms ~2% of endothelial cells, is characterized by γ-synuclein (SNCG) and high expression of the venular capillary marker ACKR1 (73, 74) (Fig. 5B). In addition, VE3 cells coexpress inflammatory cytokines, chemokines, and leukocyte adhesion molecules including IL6, IL33, SELE, and ICAM1 (Fig. 5B), similar to lymph node high endothelial venules that mediate leukocyte entry (75, 76). We performed whole-mount immunostaining of healthy dermis and identified SNCG+PECAM1+(VE3) distended vascular structures in the superficial dermis (Fig. 5C), which suggests that these cells may be postcapillary venular cells regulating leukocyte adhesion and migration.

Co-optation of developmental gene programs in AD and psoriasis

We observed significant expansion of VE3 in AD and psoriasis lesional skin (Fig. 5D and fig. S7B) that was also evident in the broader patient cohort bulk RNA-seq data (Fig. 5E). Fetal skin VE cells aligned transcriptionally with adult skin VE3 and also expressed genes involved in leukocyte adhesion and trafficking (Figs. 1E and 5B). The expansion of VE3 in inflamed skin and transcriptome alignment of fetal gut, kidney, liver, skin, and thymus VE and of adult skin VE3 (Fig. 5F) led us to hypothesize that developmental VE gene programs are involved in AD and psoriasis pathogenesis, similar to our earlier observation with Mac2. We derived 112 genes that were conserved (Seurat FindConservedMarkers, negative binomial test, P < 0.05) between fetal skin VE and AD and psoriasis VE3. This identified gene sets related to stress (DNAJB1, HSPA6, HSPB1, HSPH1, HSP90AA1), IL-6 (SOCS3), and angiopoietin (EGR1) signaling, similar to Mac2 (tables S5 and S6). Gene Ontology analysis identified significantly enriched gene set clusters (hypergeometric test, q value < 0.05), similar to Mac2 (Figs. 3H and 5G), relating to leukocyte adhesion, T cell activation, and IL-8 response as conserved gene modules in developing skin VE and VE3 in AD and psoriasis, respectively (Fig. 5G and table S6).

To validate the pathogenic role of VE3 in inflammatory skin disease in vivo, we used ACKR1 as a marker for VE3 (based on the scRNA-seq data of healthy, AD, and psoriasis skin) (Fig. 5B and fig. S7C) and compared the abundance of VE3 in healthy, AD, and psoriasis skin before and after treatment with oral methotrexate. This showed a significantly higher (P < 0.05, ANOVA) frequency of VE3 in AD and psoriasis skin and a reduction in VE3 in the skin of AD patients after treatment, in line with clinical response and reduction in EASI score, similar to Mac2 (Fig. 5, H and I). Flow cytometry analysis confirmed the expansion of VE3 in lesional psoriasis skin (fig. S7, D and E).

Mac2 and VE3 are the only skin cell states significantly enriched for these leukocyte migration gene programs; hence, we investigated whether they were interacting with each other or other immune cells to coordinate this function. To assess cell-cell interactions in healthy, AD, and psoriasis skin, we interrogated the CellPhoneDB receptor-ligand database, which predicted a significant enrichment for ACKR1 on VE3 to interact with CXCL8 (IL-8) on Mac2 (Fig. 5J and fig. S7F). We confirmed this cell-cell interaction in situ, demonstrating the close apposition of VE3 and perivascular Mac2 in AD and psoriasis skin (Fig. 5K). CellPhoneDB analysis also predicted enhanced interaction between both VE3 and Mac2 with lymphocytes in AD and psoriasis skin relative to healthy skin, supporting a role for these cells in lymphocyte recruitment into inflamed skin (fig. S7G).

Discussion

The importance of developmental programs in carcinogenesis and metastasis of both childhood and adult-onset tumors is well established (1, 2, 77). The impact of developmental programs in adult-onset neurodegenerative disorders is also emerging (78). Our findings support a broader use of prenatal cellular programs, not only in inflammatory skin disease but also potentially in other immune-mediated inflammatory disorders. Lymphocyte seeding into the developing skin is reliant on the structural network provided by the vasculature and, as our data suggest, also through endothelial cell interactions with macrophages (21), which are the most abundant skin-resident immune cells during embryonic development. We postulate that this interplay is co-opted to recruit immune cells in inflammatory skin disease. The molecular regulation of conserved gene modules during development may be distinct from that observed in disease. Dissecting the precise interplay of known angiogenic triggers, such as hypoxia, Wnt, STAT3, and β-catenin signaling, will pave the way toward a mechanistic understanding of VE3 expansion. Establishing the intrinsic and tissue-extrinsic factors that drive the Mac2 state acquisition in disease may innovate anti-inflammatory strategies. In addition to prenatal endothelial cell and macrophage gene programs, the skin fibroblast (F2) cell program is also augmented in AD and psoriasis, as well as activation of the fetal thymic medullary DC state in inflammation of several adult tissues and cancer (50, 53–55).

Our human skin atlas provides a road map for targeting pathological programs in inflammatory skin diseases and is a foundational resource for the dynamic cutaneous cellular topology that evolves during fetal development, adulthood, and inflammation.

REFERENCES AND NOTES

1. S. Jessa et al., Stalled developmental programs at the root of pediatric brain tumors. Nat. Genet. 51, 1702–1713 (2019). doi: 10.1038/s41588-019-0531-7; pmid: 31768071
2. I. Tirosh et al., Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. Nature 539, 309–313 (2016). doi: 10.1038/nature20123; pmid: 27806376
3. R. Francis et al., Gastrointestinal transcription factors drive lineage-specific developmental programs in organ specification and cancer. Sci. Adv. 5, eaax8898 (2019). doi: 10.1126/sciadv.aax8898; pmid: 31844668
4. M. M. Pomerantz et al., Prostate cancer reactivates developmental epigenomic programs during metastatic progression. Nat. Genet. 52, 790–799 (2020). doi: 10.1038/s41588-020-0664-8; pmid: 32690948
5. E. Guttman-Yassky, J. G. Krueger, Atopic dermatitis and psoriasis: Two different immune diseases or one spectrum? Curr. Opin. Immunol. 48, 68–73 (2017). doi: 10.1016/j.coi.2017.08.008; pmid: 28869867
6. C. Philippeos et al., Spatial and Single-Cell Transcriptional Profiling Identifies Functionally Distinct Human Dermal Fibroblast Subpopulations. J. Invest. Dermatol. 138, 811–825 (2018). doi: 10.1016/j.jid.2018.01.016; pmid: 29381249
7. L. Solé-Boldo et al., Single-cell transcriptomes of the human skin reveal age-related loss of fibroblast priming. Commun. Biol. 3, 188 (2020). doi: 10.1038/s42003-020-0922-4; pmid: 32327715
8. T. Tabib, C. Morse, T. Wang, W. Chen, R. Lafyatis, SFRP2/DPP4 and FMO1/LSP1 Define Major Fibroblast Populations in Human Skin. J. Invest. Dermatol. 138, 802–810 (2018). doi: 10.1016/j.jid.2017.09.045; pmid: 29080679
9. V. Vorstandlechner et al., Deciphering the functional heterogeneity of skin fibroblasts using single-cell RNA sequencing. FASEB J. 34, 3677–3692 (2020). doi: 10.1096/fj.201902001RR; pmid: 31930613

10. D. Xue, T. Tabib, C. Morse, R. Lafyatis, Transcriptome landscape of myeloid cells in human skin reveals diversity, rare populations and putative DC progenitors. *J. Dermatol. Sci.* **97**, 41–49 (2020). doi: 10.1016/j.jdermsci.2019.11.012; pmid: 31836271

11. M. Haniffa et al., Differential rates of replacement of human dermal dendritic cells and macrophages during hematopoietic stem cell transplantation. *J. Exp. Med.* **206**, 371–385 (2009). doi: 10.1084/jem.20081633; pmid: 19171766

12. M. Haniffa et al., Human tissues contain CD141^hi cross-presenting dendritic cells with functional homology to mouse CD103^+ nonlymphoid dendritic cells. *Immunity* **37**, 60–73 (2012). doi: 10.1016/j.immuni.2012.04.012; pmid: 22795876

13. N. McGovern et al., Human dermal CD14^+ cells are a transient population of monocyte-derived macrophages. *Immunity* **41**, 465–477 (2014). doi: 10.1016/j.immuni.2014.08.006; pmid: 25200712

14. X.-N. Wang et al., A three-dimensional atlas of human dermal leukocytes, lymphatics, and blood vessels. *J. Invest. Dermatol.* **134**, 965–974 (2014). doi: 10.1038/jid.2013.481; pmid: 24352044

15. See supplementary materials.

16. K. Polański et al., BBKNN: Fast batch alignment of single cell transcriptomes. *Bioinformatics* **36**, 964–965 (2020). pmid: 31400197

17. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018). doi: 10.1186/s13059-017-1382-0; pmid: 29409532

18. H. Aliee, F. Theis, AutoGeneS: Automatic gene selection using multi-objective optimization for RNA-seq deconvolution. bioRxiv [preprint]. 23 February 2020;pmid: 940650

19. S. C. van den Brink et al., Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat. Methods* **14**, 935–936 (2017). doi: 10.1038/nmeth.4437; pmid: 28960196

20. R. A. Botting et al., Phenotypic and functional consequences of different isolation protocols on skin mononuclear phagocytes. *J. Leukoc. Biol.* **101**, 1393–1403 (2017). doi: 10.1189/jlb.4A1116-496R; pmid: 28270408

21. D.-M. Popescu et al., Decoding human fetal liver haematopoiesis. *Nature* **574**, 365–371 (2019). doi: 10.1038/s41586-019-1652-y; pmid: 31597962

22. T. Stuart et al., Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888–1902.e21 (2019). doi: 10.1016/j.cell.2019.05.031; pmid: 31178118

23. R. A. Clark, Resident memory T cells in human health and disease. *Sci. Transl. Med.* **7**, 269rv1 (2015). doi: 10.1126/scitranslmed.3010641; pmid: 25568072

24. B. V. Kumar, T. J. Connors, D. L. Farber, Human T Cell Development, Localization, and Function throughout Life. *Immunity* **48**, 202–213 (2018). doi: 10.1016/j.immuni.2018.01.007; pmid: 29466753

25. H. Spits, J. H. Bernink, L. Lanier, NK cells and type 1 innate lymphoid cells: Partners in host defense. *Nat. Immunol.* **17**, 758–764 (2016). doi: 10.1038/ni.3482; pmid: 27328005

26. M. L. Robinette et al., Transcriptional programs define molecular characteristics of innate lymphoid cell classes and subsets. *Nat. Immunol.* **16**, 306–317 (2015). doi: 10.1038/ni.3094; pmid: 25621825

27. J. H. Bernink et al., Interleukin-12 and -23 Control Plasticity of CD127^+ Group 1 and Group 3 Innate Lymphoid Cells in the Intestinal Lamina Propria. *Immunity* **43**, 146–160 (2015). doi: 10.1016/j.immuni.2015.06.019; pmid: 26187413

28. E. Montaldo, K. Juelke, C. Romagnani, Group 3 innate lymphoid cells (ILC3s): Origin, differentiation, and plasticity in humans and mice. *Eur. J. Immunol.* **45**, 2171–2182 (2015). doi: 10.1002/eji.201545598; pmid: 26031799

29. J. Li, M. Olshansky, F. R. Carbone, J. Z. Ma, Transcriptional Analysis of T Cells Resident in Human Skin. *PLOS ONE* **11**, e0148351 (2016). doi: 10.1371/journal.pone.0148351; pmid: 26824609

30. K. D. Omilusik et al., Sustained Id2 regulation of E proteins is required for terminal differentiation of effector CD8^+ T cells. *J. Exp. Med.* **215**, 773–783 (2018). doi: 10.1084/jem.20171584; pmid: 29440362

31. E. M. Esparza, R. H. Arch, Glucocorticoid-induced TNF receptor, a costimulatory receptor on naive and activated T cells, uses TNF receptor-associated factor 2 in a novel fashion as an inhibitor of NF-κB activation. *J. Immunol.* **174**, 7875–7882 (2005). doi: 10.4049/jimmunol.174.12.7875; pmid: 15944293

32. K. Stanko et al., CD96 expression determines the inflammatory potential of IL-9-producing Th9 cells. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E2940–E2949 (2018). doi: 10.1073/pnas.1708329115; pmid: 29531070

33. C.-Y. Huang et al., DUSP4 deficiency enhances CD25 expression and CD4^+ T-cell proliferation without impeding T-cell development. *Eur. J. Immunol.* **42**, 476–488 (2012). doi: 10.1002/eji.201041295; pmid: 22101742

34. E. Ayroldi et al., Modulation of T-cell activation by the glucocorticoid-induced leucine zipper factor via inhibition of nuclear factor κB. *Blood* **98**, 743–753 (2001). doi: 10.1182/blood.V98.3.743; pmid: 11468175

35. B. V. Kumar et al., Human Tissue-Resident Memory T Cells Are Defined by Core Transcriptional and Functional Signatures in Lymphoid and Mucosal Sites. *Cell Rep.* **20**, 2921–2934 (2017). doi: 10.1016/j.celrep.2017.08.078; pmid: 28930685

36. J. W. Schoggins, C. M. Rice, Interferon-stimulated genes and their antiviral effector functions. *Curr. Opin. Virol.* **1**, 519–525 (2011). doi: 10.1016/j.coviro.2011.10.008; pmid: 22328912

37. R. Ramesh et al., Pro-inflammatory human Th17 cells selectively express P-glycoprotein and are refractory to glucocorticoids. *J. Exp. Med.* **211**, 89–104 (2014). doi: 10.1084/jem.20130301; pmid: 24395888

38. G. Meyer zu Horste et al., RBPJ Controls Development of Pathogenic Th17 Cells by Regulating IL-23 Receptor Expression. *Cell Rep.* **16**, 392–404 (2016). doi: 10.1016/j.celrep.2016.05.088; pmid: 27346359

39. R. Takagi et al., B cell chemoattractant CXCL13 is preferentially expressed by human Th17 cell clones. *J. Immunol.* **181**, 186–189 (2008). doi: 10.4049/jimmunol.181.1.186; pmid: 18566383

40. T. B. Rojahn et al., Single-cell transcriptomics combined with interstitial fluid proteomics defines cell type-specific immune regulation in atopic dermatitis. *J. Allergy Clin. Immunol.* **146**, 1056–1069 (2020). doi: 10.1016/j.jaci.2020.03.041; pmid: 32344053

41. S. L. Nutt, K. A. Fairfax, A. Kallies, BLIMP1 guides the fate of effector B and T cells. *Nat. Rev. Immunol.* **7**, 923–927 (2007). doi: 10.1038/nri2204; pmid: 17965637

42. F. Z. Chowdhury, L. D. Estrada, S. Murray, J. Forman, J. D. Farrar, Pharmacological inhibition of TPL2/MAP3K8 blocks human cytotoxic T lymphocyte effector functions. *PLOS ONE* **9**, e92187 (2014). doi: 10.1371/journal.pone.0092187; pmid: 24642963

43. W. Wei et al., Dual-specificity phosphatases 2: Surprising positive effect at the molecular level and a potential biomarker of diseases. *Genes Immun.* **14**, 1–6 (2013). doi: 10.1038/gene.2012.54; pmid: 23190643

44. R. M. Welsh, Blimp hovers over T cell immunity. *Immunity* **31**, 178–180 (2009). doi: 10.1016/j.immuni.2009.08.005; pmid: 19699168

45. L. C. Tsoi et al., Atopic Dermatitis Is an IL-13-Dominant Disease with Greater Molecular Heterogeneity Compared to Psoriasis. *J. Invest. Dermatol.* **139**, 1480–1489 (2019). doi: 10.1016/j.jid.2018.12.018; pmid: 30641038

46. A.-C. Villani et al., Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, aah4573 (2017). doi: 10.1126/science.aah4573; pmid: 28428369

47. G. Hoeffel et al., Adult Langerhans cells derive predominantly from embryonic fetal liver monocytes with a minor contribution of yolk sac-derived macrophages. *J. Exp. Med.* **209**, 1167–1181 (2012). doi: 10.1084/jem.20120340; pmid: 22565823

48. M. P. Collin et al., The fate of human Langerhans cells in hematopoietic stem cell transplantation. *J. Exp. Med.* **203**, 27–33 (2006). doi: 10.1084/jem.20051787; pmid: 16390938

49. L. Ardouin et al., Broad and Largely Concordant Molecular Changes Characterize Tolerogenic and Immunogenic Dendritic Cell Maturation in Thymus and Periphery. *Immunity* **45**, 305–318 (2016). doi: 10.1016/j.immuni.2016.07.019; pmid: 27533013

50. B. Maier et al., A conserved dendritic-cell regulatory program limits antitumour immunity. *Nature* **580**, 257–262 (2020). doi: 10.1038/s41586-020-2134-y; pmid: 32269339

51. M. Baratin et al., T Cell Zone Resident Macrophages Silently Dispose of Apoptotic Cells in the Lymph Node. *Immunity* **47**, 349–362.e5 (2017). doi: 10.1016/j.immuni.2017.07.019; pmid: 28801233

52. S. Tamoutounour et al., Origins and functional specialization of macrophages and of conventional and monocyte-derived dendritic cells in mouse skin. *Immunity* **39**, 925–938 (2013). doi: 10.1016/j.immuni.2013.10.004; pmid: 24184057

53. T.-P. Vu Manh, N. Bertho, A. Hosmalin, I. Schwartz-Cornil, M. Dalod, Investigating Evolutionary Conservation of Dendritic Cell Subset Identity and Functions. *Front. Immunol.* **6**, 260 (2015). doi: 10.3389/fimmu.2015.00260; pmid: 26082777

54. T.-L. Tang-Huau et al., Human in vivo-generated monocyte-derived dendritic cells and macrophages cross-present

antigens through a vacuolar pathway. *Nat. Commun.* **9**, 2570 (2018). doi: 10.1038/s41467-018-04985-0; pmid: 29967419

55. J.-E. Park et al., A cell atlas of human thymic development defines T cell repertoire formation. *Science* **367**, eaay3224 (2020). doi: 10.1126/science.aay3224; pmid: 32079746

56. T. A. Wynn, A. Chawla, J. W. Pollard, Macrophage biology in development, homeostasis and disease. *Nature* **496**, 445–455 (2013). doi: 10.1038/nature12034; pmid: 23619691

57. L. Chorro et al., Langerhans cell (LC) proliferation mediates neonatal development, homeostasis, and inflammation-associated expansion of the epidermal LC network. *J. Exp. Med.* **206**, 3089–3100 (2009). doi: 10.1084/jem.20091586; pmid: 19995948

58. D. Terhorst et al., Dynamics and Transcriptomics of Skin Dendritic Cells and Macrophages in an Imiquimod-Induced, Biphasic Mouse Model of Psoriasis. *J. Immunol.* **195**, 4953–4961 (2015). doi: 10.4049/jimmunol.1500551; pmid: 26466959

59. K. M. Bertram et al., Identification of HIV transmitting CD11c^+ human epidermal dendritic cells. *Nat. Commun.* **10**, 2759 (2019). doi: 10.1038/s41467-019-10697-w; pmid: 31227717

60. J. B. Cheng et al., Transcriptional Programming of Normal and Inflamed Human Epidermis at Single-Cell Resolution. *Cell Rep.* **25**, 871–883 (2018). doi: 10.1016/j.celrep.2018.09.006; pmid: 30355494

61. M. B. Omary, N.-O. Ku, P. Strnad, S. Hanada, Toward unraveling the complexity of simple epithelial keratins in human disease. *J. Clin. Invest.* **119**, 1794–1805 (2009). doi: 10.1172/JCI37762; pmid: 19587454

62. A. Mishra et al., A protein phosphatase network controls the temporal and spatial dynamics of differentiation commitment in human epidermis. *eLife* **6**, e27356 (2017). doi: 10.7554/eLife.27356; pmid: 29043977

63. K. R. Feingold, Lamellar bodies: The key to cutaneous barrier function. *J. Invest. Dermatol.* **132**, 1951–1953 (2012). doi: 10.1038/jid.2012.177; pmid: 22797297

64. C. Trapnell et al., The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014). doi: 10.1038/nbt.2859; pmid: 24658644

65. S. Joost et al., Single-Cell Transcriptomics Reveals that Differentiation and Spatial Signatures Shape Epidermal and Hair Follicle Heterogeneity. *Cell Syst.* **3**, 221–237.e9 (2016). doi: 10.1016/j.cels.2016.08.010; pmid: 27641957

66. S.-C. Lee et al., Expression of differentiation markers during fetal skin development in humans: Immunohistochemical studies on the precursor proteins forming the cornified cell envelope. *J. Invest. Dermatol.* **112**, 882–886 (1999). doi: 10.1046/j.1523-1747.1999.00602.x; pmid: 10383733

67. M. Seidl-Philipp et al., Spectrum of ichthyoses in an Austrian ichthyosis cohort from 2004 to 2017. *J. Dtsch. Dermatol. Ges.* **18**, 17–25 (2020). doi: 10.1111/ddg.13968; pmid: 31642606

68. H. B. Schonthaler et al., S100A8-S100A9 protein complex mediates psoriasis by regulating the expression of complement factor C3. *Immunity* **39**, 1171–1181 (2013). doi: 10.1016/j.immuni.2013.11.011; pmid: 24332034

69. U. Sivaprasad et al., SERPINB3/B4 contributes to early inflammation and barrier dysfunction in an experimental murine model of atopic dermatitis. *J. Invest. Dermatol.* **135**, 160–169 (2015). doi: 10.1038/jid.2014.353; pmid: 25111616

70. E. G. Harper et al., Th17 cytokines stimulate CCL20 expression in keratinocytes in vitro and in vivo: Implications for psoriasis pathogenesis. *J. Invest. Dermatol.* **129**, 2175–2183 (2009). doi: 10.1038/jid.2009.65; pmid: 19295614

71. H. He et al., Single-cell transcriptome analysis of human skin identifies novel fibroblast subpopulation and enrichment of immune subsets in atopic dermatitis. *J. Allergy Clin. Immunol.* **145**, 1615–1628 (2020). doi: 10.1016/j.jaci.2020.01.042; pmid: 32035984

72. T. Tammela, K. Alitalo, Lymphangiogenesis: Molecular mechanisms and future promise. *Cell* **140**, 460–476 (2010). doi: 10.1016/j.cell.2010.01.045; pmid: 20178740

73. J. Goveia et al., An Integrated Gene Expression Landscape Profiling Approach to Identify Lung Tumor Endothelial Cell Heterogeneity and Angiogenic Candidates. *Cancer Cell* **37**, 421 (2020). doi: 10.1016/j.ccell.2020.03.002; pmid: 32183954

74. A. Thiriot et al., Differential DARC/ACKR1 expression distinguishes venular from non-venular endothelial cells in murine tissues. *BMC Biol.* **15**, 45 (2017). doi: 10.1186/s12915-017-0381-7; pmid: 28526034

75. J. Pollheimer et al., Interleukin-33 drives a proinflammatory endothelial activation that selectively targets nonquiescent

265

cells. *Arterioscler. Thromb. Vasc. Biol.* **33**, e47–e55 (2013). doi: 10.1161/ATVBAHA.112.253427; pmid: 23162017

76. K. Veerman, C. Tardiveau, F. Martins, J. Coudert, J.-P. Girard, Single-Cell Analysis Reveals Heterogeneity of High Endothelial Venules and Different Regulation of Genes Controlling Lymphocyte Entry to Lymph Nodes. *Cell Rep.* **26**, 3116–3131.e5 (2019). doi: 10.1016/j.celrep.2019.02.042; pmid: 30865898

77. H. S. Phillips *et al.*, Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell* **9**, 157–173 (2006). doi: 10.1016/j.ccr.2006.02.019; pmid: 16530701

78. M. Barnat *et al.*, Huntington's disease alters human neurodevelopment. *Science* **369**, 787–793 (2020). doi: 10.1126/science.aax3338; pmid: 32675289

79. Zenodo doi:10.5281/zenodo.4249674.

SUPPLEMENTARY MATERIALS

Appendix C. Co-author of Nature Medicine Research Article: Single-cell multi-omics analysis of the immune response in COVID-19.

OPEN

Single-cell multi-omics analysis of the immune response in COVID-19

Emily Stephenson[1,43], Gary Reynolds[1,43], Rachel A. Botting[1,43], Fernando J. Calero-Nieto[2,43], Michael D. Morgan[3,4,43], Zewen Kelvin Tuong[5,6,43], Karsten Bach[3,4,43], Waradon Sungnak[6,43], Kaylee B. Worlock[7], Masahiro Yoshida[7], Natsuhiko Kumasaka[6], Katarzyna Kania[4], Justin Engelbert[1], Bayanne Olabi[1], Jarmila Stremenova Spegarova[8], Nicola K. Wilson[2], Nicole Mende[2], Laura Jardine[1], Louis C. S. Gardner[1], Issac Goh[1], Dave Horsfall[1], Jim McGrath[1], Simone Webb[1], Michael W. Mather[1], Rik G. H. Lindeboom[6], Emma Dann[6], Ni Huang[6], Krzysztof Polanski[6], Elena Prigmore[6], Florian Gothe[8,24], Jonathan Scott[8], Rebecca P. Payne[8], Kenneth F. Baker[8,9], Aidan T. Hanrath[8,10], Ina C. D. Schim van der Loeff[8], Andrew S. Barr[10], Amada Sanchez-Gonzalez[10], Laura Bergamaschi[11,12], Federica Mescia[11,12], Josephine L. Barnes[7], Eliz Kilich[13], Angus de Wilton[13], Anita Saigal[14], Aarash Saleh[14], Sam M. Janes[7,13], Claire M. Smith[15], Nusayhah Gopee[1,16], Caroline Wilson[1,17], Paul Coupland[4], Jonathan M. Coxhead[1], Vladimir Yu Kiselev[6], Stijn van Dongen[6], Jaume Bacardit[18], Hamish W. King[6,19], Cambridge Institute of Therapeutic Immunology and Infectious Disease-National Institute of Health Research (CITIID-NIHR) COVID-19 BioResource Collaboration*, Anthony J. Rostron[8,20], A. John Simpson[8], Sophie Hambleton[8], Elisa Laurenti[2], Paul A. Lyons[11,12], Kerstin B. Meyer[6,44], Marko Z. Nikolić[7,13,44], Christopher J. A. Duncan[8,10,44], Kenneth G. C. Smith[11,12,44], Sarah A. Teichmann[6,21,44 ✉], Menna R. Clatworthy[5,6,12,22,23,44 ✉], John C. Marioni[3,4,6,44 ✉], Berthold Göttgens[2,44 ✉] and Muzlifah Haniffa[1,6,9,16,44 ✉]

Analysis of human blood immune cells provides insights into the coordinated response to viral infections such as severe acute respiratory syndrome coronavirus 2, which causes coronavirus disease 2019 (COVID-19). We performed single-cell transcriptome, surface proteome and T and B lymphocyte antigen receptor analyses of over 780,000 peripheral blood mononuclear cells from a cross-sectional cohort of 130 patients with varying severities of COVID-19. We identified expansion of nonclassical monocytes expressing complement transcripts (*CD16+C1QA/B/C+*) that sequester platelets and were predicted to replenish the alveolar macrophage pool in COVID-19. Early, uncommitted CD34+ hematopoietic stem/progenitor cells were primed toward megakaryopoiesis, accompanied by expanded megakaryocyte-committed progenitors and increased platelet activation. Clonally expanded CD8+ T cells and an increased ratio of CD8+ effector T cells to effector memory T cells characterized severe disease, while circulating follicular helper T cells accompanied mild disease. We observed a relative loss of IgA2 in symptomatic disease despite an overall expansion of plasmablasts and plasma cells. Our study highlights the coordinated immune response that contributes to COVID-19 pathogenesis and reveals discrete cellular components that can be targeted for therapy.

The COVID-19 global pandemic has caused >120 million infections and 2.6 million deaths (as of 17 March 2021)[1,2]. Symptoms vary in severity and include acute respiratory distress syndrome, thrombosis and organ failure[3]. COVID-19 is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a single-stranded RNA betacoronavirus that enters host cells through receptors such as angiotensin-converting enzyme 2 (ACE2) and neuropilin (NRP1), which are expressed widely, including in nasal epithelium[4–6].

Several studies have highlighted a complex network of peripheral blood immune responses in COVID-19 infection[7,8]. A reduction in

T cells with disease severity and reduced interferon (IFN)-γ production by lymphocytes have been reported[9]. However, an expansion of highly cytotoxic effector T cell subsets in moderate to severe disease[10,11] and higher expression of exhaustion markers programmed cell death protein 1 and Tim-3 on CD8+ T cells have been described in patients receiving intensive care therapy[5]. In severe cases, classical monocytes have been shown to display a type 1 IFN inflammatory signature[6]; however, low levels of IFNα coupled with a reduction in plasmacytoid dendritic cells (DCs) have been reported in patients with critical disease[12]. Emergency and dysregulated myelopoiesis, and expanded activated megakaryocytes have also

267

been reported[13–15]. Proliferating plasmablasts and extrafollicular B cell activation are present in critically ill patients, but despite high levels of SARS-CoV-2-specific antibodies and antibody-secreting cells, many of these patients do not recover[6,15]. To better understand the coordinated systemic immune response in individuals with asymptomatic and symptomatic COVID-19, we performed combined single-cell transcriptome, cell-surface protein and lymphocyte antigen receptor repertoire analysis of peripheral blood of a cross-sectional patient cohort and integrated results across three UK medical centers.

Results

Altered cellular profiles across COVID-19 severities. We generated single-cell combined transcriptome and surface proteome data from peripheral blood mononuclear cells (PBMCs) from individuals with asymptomatic, mild, moderate, severe and critical[16] COVID-19 across three UK centers in Newcastle, Cambridge and London (Fig. 1a, Extended Data Fig. 1a and Supplementary Table 1). Controls included healthy volunteers, individuals with non-COVID-19 severe respiratory illness and healthy volunteers administered with intravenous lipopolysaccharide (IV-LPS) as a surrogate for an acute systemic inflammatory response (Fig. 1a and Supplementary Table 2). We sequenced 1,141,860 cells from 143 samples. Following computational doublet removal, 781,123 cells passed quality control (QC; minimum of 200 genes and <10% mitochondrial reads per cell; Extended Data Fig. 1b). Data were integrated using Harmony[17] with good mixing of cells by the kBET statistic calculated for each cluster across samples (rejection rate improved from 0.62 to 0.36 following integration, $P < 2.1 \times 10^{-8}$ by Wilcoxon paired signed-rank test; Extended Data Fig. 1c,d).

Following Leiden clustering, cells were manually annotated based on the RNA expression of known marker genes supported by surface protein expression of markers employed in flow cytometry to discriminate subpopulations (Extended Data Fig. 1e). We defined 18 cell subsets (Fig. 1b), with an additional 27 cell states identified following subclustering (Figs. 1b, 2a, 3a,b and 4a,b and Supplementary Table 3). Our annotation was further validated using Azimuth, whereby more than 50% of cells were mapped and matched to a unique cluster in 32/33 of the clusters defined in the Azimuth PBMC dataset (Methods; proliferating CD8+ T cells mapped across two clusters). Clusters unique to our data included proliferating monocytes, innate lymphoid cell subpopulations and isotype-specific plasma cells (Extended Data Fig. 1f).

We observed a relative expansion of proliferating lymphocytes, proliferating monocytes, platelets and mobilized hematopoietic stem and progenitor cells (HSPCs) with worsening disease. Plasmablasts and B cells were expanded in severe and critical disease (Fig. 1c and Extended Data Fig. 2a). These changes matched trends in clinical blood cell counts (Extended Data Fig. 2b and Supplementary Table 4). To assess the broader impacts of patient characteristics and clinical metadata on the altered proportion of cell types/states, we used a Poisson linear mixed model (Methods), which predicted the COVID-19 swab result (Bonferroni-corrected logistic regression (BF-corrected LR), $P = 1.1 \times 10^{-3}$; Methods), disease severity at blood sampling (BF-corrected LR, $P = 8.9 \times 10^{-8}$) and center (contributed by increased red blood cells (RBCs) and reduced monocytes in the Cambridge patient cohort; (BF-corrected LR, $P = 2.0 \times 10^{-142}$) as the main contributing factors among seven different clinical/technical factors (Extended Data Fig. 2c). PBMC composition varied depending on symptom duration, with increased relative frequency of plasmacytoid dendritic cells (pDCs), natural killer (NK) cells, CD14+ and CD16+ monocytes (false discovery rate (FDR), 10%) and decreased relative frequencies of B cells, regulatory T (T_{reg}) cells, RBCs, platelets and CD4+ T cells, with a longer symptomatic interval before hospital admission (Extended Data Fig. 2d). These changes may be due to a subset

of individuals in the critically ill category who reported a longer time since symptom onset, consistent with a protracted course of infection in critical disease (Extended Data Fig. 2e,f). However, concordant changes in immune cell composition were observed when excluding patients with either the longest symptom durations (>24 d) or critical disease (Extended Data Fig. 2g), indicating that disease severity changes were not driven by symptom duration. Cell abundance results were also in agreement when performing a leave-one-out analysis (Extended Data Fig. 2h).

We observed expression of type I/III interferon response genes in monocytes, DCs and HSPCs across the spectrum of COVID-19 severity, but not in individuals challenged with IV-LPS, in keeping with the importance of type I and III interferons in innate immune responses to viral infection (Fig. 1d). Type I/III interferon response-related genes were recently implicated in genome-wide association studies (GWAS) for COVID-19 susceptibility[18,19]. *IFNAR2* was both upregulated in individuals with COVID-19 compared to healthy controls in most circulating cell types and highly expressed by plasmablasts, monocytes and DCs (Extended Data Fig. 2i).

Multiplexed analysis of 45 proteins in serum showed two contrasting profiles between mild/moderate and severe/critical patients. CCL4, CXCL10, interleukin (IL)-7 and IL-1α were associated with severe and critical disease, suggesting an augmented drive for monocyte and NK lymphocyte recruitment, as well as support for T cell activity/pathology (Extended Data Fig. 2j and Supplementary Table 5).

We used Cydar[20] to characterize the immune landscape changes with disease severity based on surface protein expression by dividing cells into phenotypic hyperspheres. We quantified the number of cells from each severity group within the hyperspheres, identifying 608 hyperspheres that differed significantly in abundance with increasing severity (spatial FDR < 0.05; Fig. 1e). Differentially abundant hyperspheres were present in all major immune compartments. Notably, we found an increase in B cells (CD19+/CD20+), plasma cells (CD38+) and HSPCs (CD34+), as well as remodeling of the myeloid compartment[13] (Fig. 1e).

Mononuclear phagocytes and HSPC changes. Transcriptome and surface proteome analysis of blood mononuclear phagocytes (MPs) identified known DC subsets (pDC, ASDC ($AXL^+SIGLEC6^+$ DC), DC1, DC2 and DC3) and several monocyte states (Fig. 2a,b). Three CD14+ monocyte states were present (proliferating, classical CD14+ and activated CD83+) in addition to two CD16+ monocyte states (nonclassical CD16+ and C1QA/B/C+CD16+; Fig. 2a,b). Proliferating monocytes and DCs expressing *MKI67* and *TOP2A* were increased with disease severity (Fig. 2a,b). In contrast, the frequencies of DC2 and DC3 were reduced. Proliferating monocytes, previously reported by flow cytometry analysis of blood from patients with COVID-19[14], transcriptionally resembled CD14+ monocytes and was the only population to change significantly with symptom duration. (Fig. 2a,b and Extended Data Fig. 3a). Proliferating DCs resembled the DC2 subset (Fig. 2a,b). Rare *C1QA/B/C*-expressing CD16+ monocytes were the only source of C1 complement components (Fig. 2b and Extended Data Fig. 3b).

We previously demonstrated egress of blood DCs and monocytes to the alveolar space with rapid acquisition of a lung molecular profile following human inhalational LPS challenge[21]. To investigate the relationship between circulating and lung alveolar MPs in COVID-19, we compared the transcriptome profile of blood DCs and monocytes with their bronchoalveolar lavage (BAL) counterparts using recently published data (GSE145926)[22] (Extended Data Fig. 3d). Partition-based graph abstraction (PAGA) suggested transcriptional similarity between circulating CD14+ monocytes and BAL macrophages in health, aligning with recent data demonstrating that BAL macrophages can arise from circulating CD14+ monocytes
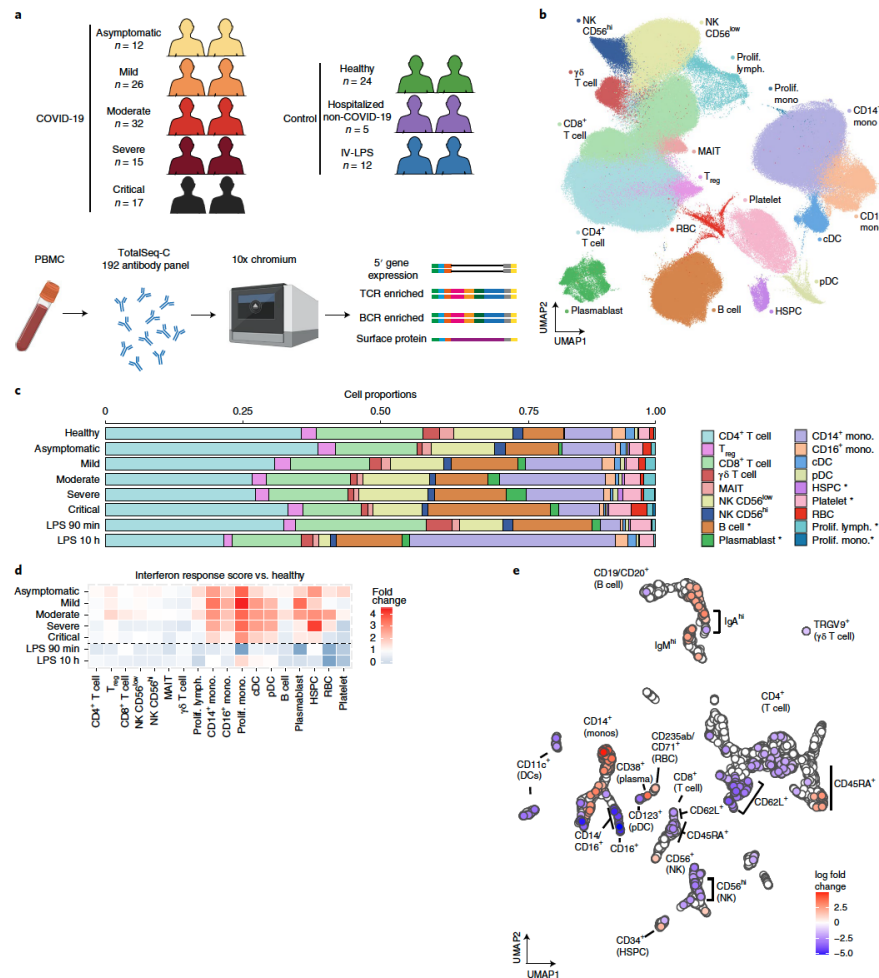
Fig. 1 | Single-cell multi-omics analysis of PBMCs from individuals with COVID-19 and controls. a, Overview of the participants included and the samples and data collected. Figure was created using BioRender.com. **b**, UMAP visualization of all 781,123 cells after QC. Leiden clusters based on 5′ gene expression shown and colored by cell type. Lymph, lymphocyte; mono, monocyte; prolif, proliferating. **c**, Bar plot of the proportion of cell types shown in **b**, separated by condition and COVID-19 severity status. Hypothesis testing was performed using quasi-likelihood F-test comparing healthy controls to individuals with COVID-19 for linear trends across disease severity groups (healthy > asymptomatic > mild > moderate > severe > critical). Differentially abundant cell types were determined using a 10% FDR and are marked with an asterisk. **d**, Enrichment of interferon response of each cell state separated by severity. IFN response was calculated using a published gene list (GO:0034340) **e**, UMAP computed using batch-corrected mean staining intensities of 188 antibodies for 4,241 hyperspheres. Each hypersphere represents an area in the 188-dimensional space and is colored by significant (spatial FDR < 0.05) severity-associated changes in abundance of cells within that space.

(Fig. 2c)[23]. In COVID-19, there was greater transcriptional similarity between BAL macrophages and circulating C1QA/B/C+CD16+ monocytes (Fig. 2c), suggesting a differential origin of alveolar macrophages in healthy donors and individuals with COVID-19. Both BAL macrophages and C1QA/B/C+CD16+ monocytes express *FCGR3A* and *C1QA/B/C* and type I interferon response genes (Fig. 2a). Myeloid hyperinflammation causing lung and peripheral tissue damage via secretion of inflammatory cytokines such as IL-6 and tumor necrosis factor (TNF) in COVID-19 has been reported and in our analysis were primarily expressed by tissue rather than blood MPs (Fig. 2c). Genes differentially expressed by blood monocytes identified *S100A8*, previously reported in COVID-19 as contributing to the cytokine storm in severe infection[24]. BAL macrophages expressed leukocyte-recruiting chemokines including *CCL2, CCL4, CCL7* and *CCL8* (Fig. 2d).

Tissue DCs respond to local inflammation and pathogen challenge by migrating to the draining lymph node to activate naïve T cells. BAL macrophages contain a population of mature, migratory DCs that express *CCR7* and *LAMP3* but downregulate DC-specific markers, such as *CD1C* and *CLEC9A* (Extended Data Fig. 3c). These migratory DCs express *IL10* in health, but *TNF* and the common IL-12 and IL-23 subunit *IL12B* in COVID-19, suggesting altered capacity for T cell polarization (Fig. 2e). In peripheral blood, C1QA/B/C+CD16+ monocytes expressed the highest amount of type 1 IFN response genes compared to all myeloid cells (Fig. 2f and Supplementary Tables 6 and 7). We detected minimal TNF-mediated or IL-6-mediated JAK–STAT signaling activation in circulating monocytes and DCs, but this was upregulated by COVID-19 BAL MPs (Fig. 2f, Supplementary Tables 6 and 7).

Coagulation abnormalities and monocyte–platelet aggregates have been previously reported in COVID-19 (refs. [25,26]), leading us to investigate predicted receptor–ligand interactions between monocytes and platelets using the CellPhoneDB repository. The expression levels of *SIRPA:CD47, FPR1:ANXA1, FPR2:APP* between monocytes:platelets were highest in the C1QA/B/C+CD16+ monocytes (Fig. 2g). Using protein data, we identified ICAM1 interactions on platelets with CD11a/b/c/CD18 primarily on C1QA/B/C+CD16+ monocytes and CD16+ monocytes (Extended Data Fig. 3d), accompanied by increased expression of surface proteins indicative of platelet activation (Fig. 2h).

Our large dataset of 781,123 PBMCs allowed us to interrogate 3,297 CD34+ HSPCs. Leiden clustering and uniform manifold and projection (UMAP) visualization showed a cloud-like representation, consistent with a stem/progenitor cell landscape previously described for bone marrow HSPCs[27] (Fig. 2i and Extended Data

Fig. 3e). Absence of CD38 mRNA and protein expression marks the most immature cells within the CD34+ compartment, while expression of markers such as *GATA1, MPO* and *PF4* characterizes distinct erythroid, myeloid and megakaryocytic progenitor populations, respectively (Fig. 2i). Accordingly, we were able to annotate six transcriptional clusters as CD34+CD38− HSPCs, CD34+CD38+ early progenitor HSPCs and CD34+CD38+ erythroid, megakaryocytic and myeloid progenitors, as well as a small population distinguished by the expression of cell cycle (S phase) genes (Fig. 2i). Megakaryocyte progenitors were absent in healthy and asymptomatic individuals but made up ~5% of CD34+ cells in all symptomatic individuals (Fig. 2j). As peripheral blood is not a site for hematopoiesis[28], this finding likely reflects COVID-19-mediated perturbation of normal homeostatic functioning of the bone marrow HSPC compartment.

In light of our earlier observations of platelet activation and enhanced C1QA/B/C+CD16+ monocyte–platelet interactions (Fig. 2g,h), the appearance of CD34+ megakaryocyte progenitors was of particular interest, as it suggested a rebalancing of the stem/progenitor cell compartment. To further explore this hypothesis, we generated differential gene expression signatures between the megakaryocyte, myeloid and erythroid progenitor clusters to interrogate potential early activation or priming in the most immature HSPC clusters (Extended Data Fig. 3f). We observed enrichment of the megakaryocyte progenitor signature in the CD38+ HSPC populations in moderate COVID-19 compared to the healthy condition (Fig. 2k and Supplementary Table 8), but no enrichment of erythroid or myeloid signatures in either CD38− or CD38+ HSPCs (Extended Data Fig. 3g and Supplementary Table 8). Our earlier observation of increased platelet activation within the context of normal platelet counts (Fig. 2h and Extended Data Fig. 2b) is thus consistent with a model whereby a rebalancing of the HSPC compartment toward megakaryopoiesis may be compensating for peripheral platelet consumption in COVID-19.

T lymphocytes and T cell receptor changes. Fine-resolution clustering of mRNA profiles revealed 11 initial clusters of CD4+ T cells, CD8+ T cells and innate-like T cells including γδ T cells, NK T cells and mucosal-associated invariant T (MAIT) cells (Fig. 3a and Extended Data Fig. 4). Annotations were refined further using RNA expression of effector cytokines and surface protein expression (Fig. 3a–c).

Cellular composition of the T cell compartment varied between healthy and infected groups (Fig. 3d). Based on their relative proportions and differential abundance testing (FDR 10%), we found activated CD4+ T cells expressing *IL22*, circulating follicular helper

Fig. 2 | Expansion of complement-expressing nonclassical monocytes and megakaryocyte-primed progenitor cells and increased platelet activation with COVID-19 disease severity. a, Dot plots of gene (left) and surface protein (right) expression for myeloid populations. **b**, Bar plot of the proportion of myeloid populations from the Newcastle and London sites. Hypothesis testing was performed using a quasi-likelihood *F*-test comparing healthy controls to individuals with COVID-19. Differentially abundant cell types were determined using a 10% FDR and are marked with an asterisk. **c**, PAGA graph representing connectivity between clusters defined in **a** for healthy (top left) and COVID-19 (bottom left) monocytes and BAL macrophages (mac). Expression of *IL6* (top right) and *TNF* (bottom right) in each cluster along the predicted path for COVID-19 monocytes. **d**, Expression of differentially expressed cytokines between CD83+CD14+ monocytes and BAL macrophages shown by cells ordered by pseudotime calculated for cells from **c**. **e**, Dot plot of gene expression of DC-derived T cell polarizing cytokines in peripheral blood DC2 cells and mature BAL DCs. **f**, Heat map displaying gene-set enrichment scores for type 1/3 IFN response, TNF response and JAK–STAT signatures in the myeloid populations. Asterisks indicate significance compared to healthy controls. Absolute values and other comparisons are provided in Supplementary Table 7. **g**, Heat map of predicted ligand–receptor interactions between platelets and monocyte subsets, using RNA data. **h**, Dot plot of significant differentially expressed genes between samples from healthy donors and individuals with COVID-19 filtered for platelet activation markers. **i**, UMAP representation of HSPCs (top) and dot plot of gene expression markers used to annotate clusters (bottom). MK, megakaryocyte; prog, progenitor. **j**, Bar chart depicting the proportion of progenitors. **k**, Box plots displaying the enrichment of a megakaryocyte progenitor signature in CD34+CD38+ HSPCs (right) and CD34+CD38− (left), averaged per donor scores. Comparisons were made by an analysis of variance (ANOVA) with pairwise comparisons using Tukey's test. Asterisks above bars indicate significance and are colored by the severity for which they were compared to. Absolute values are provided in Supplementary Table 8. Boxes denote the interquartile range (IQR), and the median is shown as horizontal bars. Whiskers extend to 1.5 times the IQR, and outliers are shown as individual points (*P* values: CD38-negative cells in healthy versus LPS group (90 min), 0.3×10^{-3}; CD38-positive cells in healthy versus moderate group, 0.7×10^{-3}).

270

T (T_{FH}) cells, type 1 helper T (T_H1) cells, CD8+ effector memory T (T_{EM}) cells and MAIT cells relatively enriched in individuals with asymptomatic and mild infection, with NKT, proliferating CD8+ and CD4+, and CD8+ terminal effector T (T_{TE}) cells enriched in individuals with more severe infection (Fig. 3e and Extended Data Fig. 5a,b). Treating disease severity as an ordinal variable (Methods), multiple cell populations displayed nonlinear differences across disease severity (proliferating CD4+ and CD8+ T cells, CD8+ T_{TE}, CD4+

T_H1, CD4+ T_H17, CD4+ central memory T (T_{CM}) and IL-22+CD4+ T cells), illustrating the complex compositional changes to peripheral T cells that occur with COVID-19 (Fig. 3e and Extended Data Fig. 5b). IL-22-expressing CD4+ T cells seen in asymptomatic and mild disease could be associated with tissue-protective responses that may restrict immunopathology (Fig. 3e) as previously shown for IL-22 in influenza A virus infection[29] and lower viral load in COVID-19 patients' lungs[30]. Proliferating CD4+ and CD8+ T cells

271
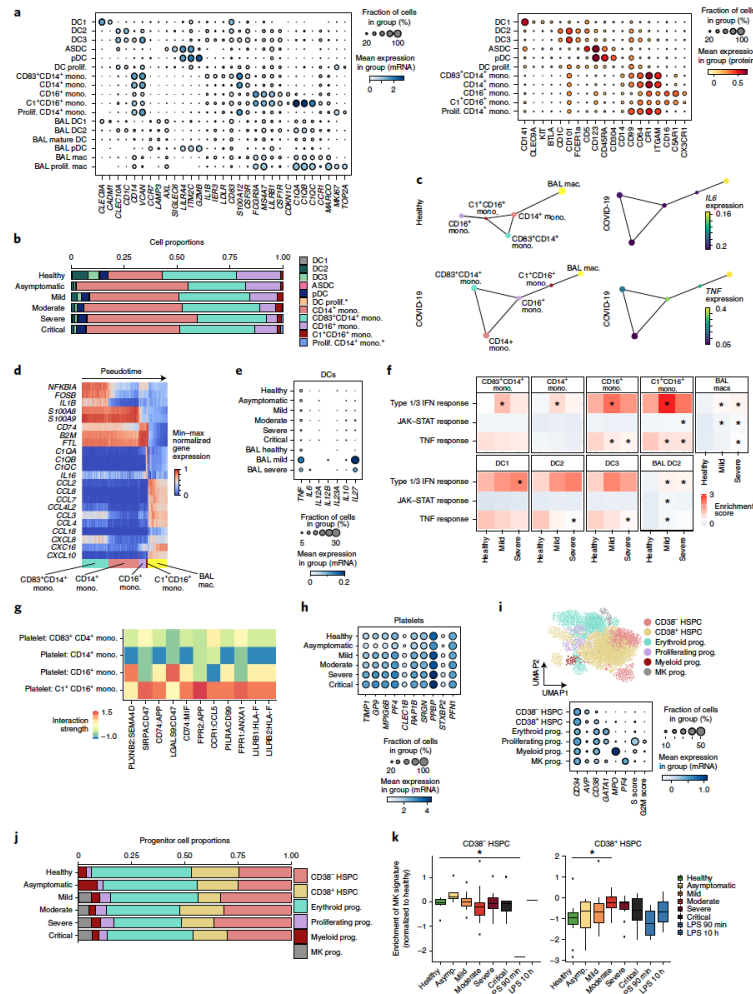
coexpressed exhaustion marker genes *LAG3* and *TOX* (Extended Data Fig. 5c), in keeping with previous studies of patients with severe COVID-19 (ref. [5]). In contrast to disease severity, CD4[+] T_H1, CD4[+] T_H2, CD4[+]IL-22[+] and CD4[+] T_{CM} cells were enriched among individuals with longer symptom duration, while effector populations with a cytotoxic phenotype (CD8[+] T_{TE}, CD8[+] T_{EM}, MAIT and NK T cells) were enriched in individuals with shorter symptom duration (Extended Data Fig. 5d).

Differential gene expression analysis across disease severity (FDR 1%) and gene-set enrichment analysis (GSEA) identified pathways associated with inflammation and T cell activation across multiple subsets, including IL-2–STAT5 signaling, mTORC1 signaling, inflammatory response, IFNγ response, and IL-6-JAK–STAT3 signaling (Extended Data Fig. 5e). The increased activation and cytotoxic phenotype in T cells from individuals with COVID-19 was functionally validated by flow cytometry analysis of PBMCs stimulated ex vivo with SARS-CoV-2 peptide showing upregulation of CD137 and CD107α (Extended Data Fig. 5f).

T cell receptor (TCR) clonality analysis showed that effector CD8[+] T cells were the most clonally expanded (odds ratio (OR) (95% confidence interval (CI)) 1.81 (1.58–2.10)), $P = 2.49 \times 10^{-15}$) and their relative proportion increased with disease severity (Fig. 3f,g and Extended Data Fig. 5g,h and Supplementary Tables 6 and 7). Conversely, the relative proportion of clonally expanded effector memory CD8[+] T_{EM} cells decreased in individuals with more severe disease (OR (95% CI) 0.87 (0.72–1.04), $P = 0.26$; Fig. 3f,g and Supplementary Tables 9 and 10). These clonal alterations were primarily driven by severity rather than differences in symptom duration for more severely ill patients, as CD8[+] T_{EM} clones expanded in individuals who had longer symptom duration, in line with a more developed infection trajectory in these individuals (per day, OR (95% CI) 1.02 (1.01–1.03), $P = 2.66 \times 10^{-10}$). The ratio of effector CD8[+] T cells to CD8[+] T_{EM} cells (TE:EM ratio) correlated with disease severity (linear model β 2.97, $P = 2.92 \times 10^{-18}$; Fig. 3g and Supplementary Tables 9 and 10), suggesting that CD8[+] T cell differentiation outcome may contribute to both antiviral protection and immunopathology, as previously reported in animal models[31], although bystander T cell activation cannot be excluded.

B lymphocytes and B cell receptor changes. Re-clustering of B cells and plasma cells identified nine clusters that were annotated according to canonical marker expression (Fig. 4a,b), and previously published transcriptional signatures (Extended Data Fig. 6a). This included immature, naïve, switched and non-switched memory B cells, and a cluster of cells that enriched for markers previously described in exhausted memory B cells[32,33] (Fig. 4a,b and Extended Data Fig. 6a). We also found a large population of plasmablasts with negative expression of CD19 and CD20, with high expression of the proliferation marker *MKI67*, consistent with previous

reports on severe SARS-CoV-2 infection[15,24], as well as IgM[+], IgG[+] and IgA[+] plasma cells (Fig. 4a,b). In individuals with symptomatic COVID-19, there was a significant expansion of plasmablasts and plasma cells (Fig. 4c and Extended Data Fig. 6b). The magnitude of this expansion increased from mild to moderate disease but was attenuated in severe to critical disease. This observation persisted even after accounting for days from symptom onset (Extended Data Fig. 6b). IgA[+] cells were decreased in individuals with symptomatic COVID-19 due to a significant decrease of the IgA2 subclass (Fig. 4d and Extended Data Fig. 6b,c), suggestive of the maintenance of an effective mucosal humoral response in asymptomatic individuals. In parallel, we observed the greatest expansion of circulating follicular helper T (cT_{FH}) cells in asymptomatic individuals and a strong positive correlation between cT_{FH} cells and plasma cells in individuals with asymptomatic/mild disease that was lost with increasing disease severity (Figs. 3e and 4e and Extended Data Fig. 5a,b). Together, this suggests the presence of coordinated T cell and B cell responses in asymptomatic and mild disease, generating effective antiviral humoral immunity that becomes uncoupled in severe and critical disease. This is consistent with previous findings relating to the requirement of T_{FH} cells for optimal antibody responses and high-quality neutralizing antibodies in viral infection[34].

GSEA analysis identified interferon alpha response and interferon gamma response pathway genes enriched in all B cell subsets in individuals with COVID-19, and this was more marked in those with asymptomatic or mild disease, and attenuated in severe and critical disease (Fig. 4f and Extended Data Fig. 6d). The magnitude of type 1 interferon transcriptional response in B cells mirrored serum IFNα levels, which were highest in individuals with mild disease (Extended Data Fig. 2j), suggesting that the low expression of IFN response genes in B cells in severe or critical disease does not reflect an inability of B cells to respond to IFNα, but rather attenuation of IFNα. This may be because the initial antiviral response has waned in patients with severe or critical disease or because these individuals fail to sustain adequate IFNα production by myeloid cells and pDCs following symptom onset as previously reported[7]. Longitudinal sampling would be required to distinguish these two possibilities.

In asymptomatic individuals, TNF signaling via nuclear factor kappa B (NF-κB) pathway genes was enriched in immature, naïve and switched memory B cells, but decreased in immature B cells and plasma cells in critical and severe disease (Fig. 4f and Extended Data Fig. 6d). Assessment of the leading-edge genes in this pathway demonstrated their markedly higher expression in all B cell and plasmablast/cell subsets in asymptomatic individuals with COVID-19 compared with those with symptomatic disease (Fig. 4g and Extended Data Fig. 6e). TNF was barely detectable in COVID-19 serum samples and highest in individuals with moderate disease (Extended Data Fig. 2j), suggesting that another

---

Fig. 3 | Compositional and clonal analyses of T lymphocytes illustrate the expansion of effector subsets. a, UMAP visualization of 309,617 T cells based on gene expression shown and colored by cell type. Insets show the two-dimensional kernel density estimates of select T cell types in UMAP space. **b,** Dot plots of gene (top) and surface protein (bottom) expression for populations shown in **a. c,** Dot plots of gene expression of cytokine genes for populations shown in **a. d,** Box plots of cell type proportions that are differentially abundant between healthy donors and individuals with COVID-19. Boxes denote the IQR, and the median is shown as horizontal bars. Whiskers extend to 1.5 times the IQR and outliers are shown as individual points ($n = 24$ healthy, $n = 86$ COVID-19 biologically independent samples). **e,** Box plots of the proportion of cell types shown in **a.** Only cell types showing trends of changes by severity status are shown. Boxes denote IQR with median shown as horizontal bars. Whiskers extend to 1.5 times the IQR, and outliers are shown as individual points ($n = 9$ asymptomatic, $n = 23$ mild, $n = 30$ moderate, $n = 13$ severe, $n = 10$ critical biologically independent samples). **f,** Bar plots show the frequency of clonal T cells. Expanded clones denote clonotypes observed more than once. Asterisks indicate significance after multiple-testing correction (logistic regression using two-sided *t*-test with Benjamini–Hochberg FDR correction; CD4[+] T_{CM} adjusted $P = 0.119$, CD4[+] T_{EM} adjusted $P = 0.472$, CD4+IL-22[+] adjusted $P = 0.01$, CD4[+] prolif. adjusted $P = 0.993$, CD4[+] T_H1 adjusted $P = 0.993$, CD4[+] T_{FH} adjusted $P = 0.109$, T_{reg} adjusted $P = 0.993$, CD8[+] prolif. adjusted $P = 0.016$, CD8[+] T_{TE} adjusted $P = 2.49 \times 10^{-15}$, CD8[+] T_{EM} adjusted $P = 0.259$). **g,** Box plots of the proportion of clonally expanded CD8[+] T_{EM} cells (left), effector CD8[+] T cells (middle) and the ratio of effector CD8[+] T cells to CD8[+] T_{EM} cells (right). Boxes denote the IQR, and the median is shown as horizontal bars. Whiskers extend to 1.5 times the IQR, and outliers are shown as individual points. Legend is as in **e.**

cytokine, for example IL-6, or stimulus may be responsible for NF-κB activation in asymptomatic individuals with COVID-19.

Hypoxia pathway genes were enriched in immature and naïve B cells only in asymptomatic individuals (Fig. 4f and Extended Data Fig. 6d). Since these individuals are unlikely to be hypoxic (given their lack of symptoms), we postulated that this signature may reflect another hypoxia inducible factor-activating stimulus, which includes B cell receptor (BCR) cross-linking[35]. We assessed the expression of genes associated with BCR activation, such as CD79A and CD79B, and downstream kinases such as BTK in B cell subsets. Overall, BCR activation-associated genes were most highly expressed in B cells in healthy controls, followed by asymptomatic

273
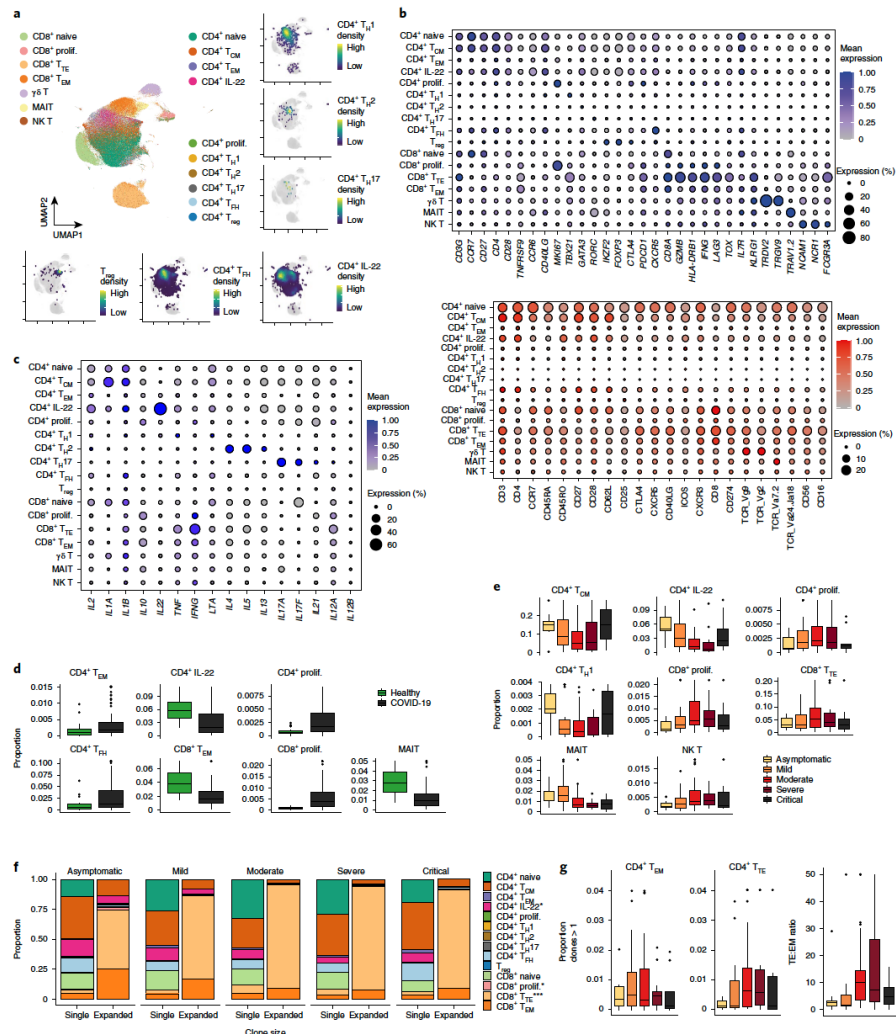
individuals with COVID-19, with lower expression observed in all symptomatic COVID-19 groups (Fig. 4g and Extended Data Fig. 6e). BCR activation threshold is also modulated by immune tyrosine inhibitory motif-containing receptors that recruit phosphatases, increasing the activation threshold of B cells[36]. BCR inhibitory gene expression was limited, but *CD22* was detectable across B cell subsets in asymptomatic COVID-19, while *FCGR2B*, *CD72* and *PTPN6* expression was evident in B cells in severe COVID-19 (Fig. 4g and Extended Data Fig. 6e). Together, this analysis suggests that B cells in asymptomatic individuals with COVID-19 and those with mild disease have a more pronounced response to interferons, increased NF-κB activation and a higher expression of genes associated with BCR activation signaling, suggesting a potential for greater BCR activation. Longitudinal analysis of patient samples will be required to establish if this is due to avid responses early in disease that prevent progression to a more severe phenotype or the immune response in early disease.

Following activation, B cells differentiate into antibody-producing plasma cells, accompanied by a progressive increase in oxidative metabolism[37,38]. We observed differences in metabolic gene pathway expression in plasmablasts and plasma cells between disease severity categories, with enrichment of oxidative phosphorylation pathway genes in all disease groups, and a relative increase in glycolysis pathway genes in asymptomatic patient plasmablasts when compared to symptomatic disease groups (Fig. 4f and Extended Data Fig. 6e).

We next assessed BCR clonality using 'dandelion', a single-cell BCR-sequencing analysis package (Methods), and found substantially more clonal expansion in symptomatic individuals with COVID-19 (Fig. 4h and Extended Data Fig. 7). Expanded clonotypes were found across all major cell types with larger clonotypes present primarily in plasmablast/plasma cell clusters (Extended Data Fig. 8a,b). Within the expanded clonotypes, there was some evidence of class switching within symptomatic COVID-19 groups but not in the asymptomatic/healthy individuals (Extended Data Fig. 8c). Unlike other large-scale single-cell RNA-sequencing (scRNA-seq) studies with BCR profiling[15,24], there was no obvious bias of immunoglobulin heavy-chain variable (*IGHV*) gene usage in general (Extended Data Fig. 9a). Disaggregating the *IGHV* gene usage data to individual gender groups showed that only *IGHV1-46* was significantly increased in women with critical COVID-19 relative to healthy controls (Extended Data Fig. 9a). Some related BCRs were present in different individuals, with more incidence of V and J gene usage and related amino acid sequences of heavy-chain and light-chain CDR3s observed in individuals with severe or critical disease, and in individuals from one center (Newcastle; Fig. 4i), which could arise due to local variants of the virus driving expansion

of specific B cell clones. We note that none of these related BCRs were found to be expanded in the individuals, which was expected as only a relatively small number of B cells per individual were sampled. It would have been unlikely to find exactly matching heavy-chain and light-chain sequences across different individuals (even when allowing for somatic hypermutation variation), given the expected low coverage that arises from a small number of single cells (relative to bulk BCR sequencing). Finally, we observed disproportionate distribution in clonotype size, whether considering expanded or all clonotypes, and increased BCR mutations between men and women with COVID-19, with greater levels of both in women compared with men (Fig. 4j and Extended Data Fig. 9b). These differences in clonal expansion of B cells are consistent with a role in previous reports of worse outcomes in COVID-19 in men[39,40].

We summarize the immunological cellular and molecular profiles observed in our study distinguishing features between asymptomatic/mild disease from severe/critical disease (Fig. 5). Future longitudinal studies may enable us to distinguish if the distinct responses in asymptomatic and milder disease prevent progression to severe phenotypes.

Discussion

Our cross-sectional multi-omics PBMC survey revealed several new insights into COVID-19 pathogenesis. Firstly, peripheral blood monocytes and DCs exhibit an interferon response to infection. We identified CD1QA/B/C+CD16+ monocytes, coexpressing receptors and ligands for interactions with platelets, that are predicted to replenish alveolar macrophages in COVID-19. Secondly, altered hematopoiesis is evident in the peripheral circulation with megakaryocyte-primed gene expression in the earliest CD34+CD38+ HSPCs, and expanded megakaryocyte progenitors in the response to COVID-19. We reveal a balance in protective versus immunopathogenic adaptive immune responses in COVID-19 patients. Previous studies have reported expanded proliferative CD4+ and CD8+ T cells with disease severity[24], but a reduction in γδ T cells[24,41], consistent with our study. In addition, we observed enrichment of T_H1 cells in asymptomatic donors, consistent with previously reported IFNγ and IL-2 antigen-specific T cells in asymptomatic individuals[42]. We report expansion of CD8+ effector T cells, which likely include antigen-specific short-lived effector cells that could lead to uncontrolled inflammation and immunopathology, expanding on previous reports[43–46].

The expansion of plasmablasts and plasma cells is less evident in critical and severe disease than in moderate and mild disease, in contrast to previous studies that reported the diminished plasmablast expansion in convalescent stages and not within active disease[15]. This response is paralleled by the T_{FH} profile in individuals

Fig. 4 | Single-cell analysis of B lymphocytes and BCR repertoire reveal plasmablast expansion and clonality differences between genders. a, UMAP visualization of 74,019 cells in the B cell lineage identified from gene expression data. **b**, Dot plots of gene (top) and surface protein (bottom) expression for populations shown in **a**. **c**, Bar plot of the mean proportion of cell types shown in **a**. **d**, Proportion of total IgA and IgA2 in plasmablast and plasma cells based on BCR data. Kruskal–Wallis test with Benjamini–Hochberg correction. **e**, Coordinated changes between T_{FH} and B cells assessed by differential correlation analysis (empirical $P \leq 0.1$). Shown is the Pearson correlation ($\pm$ bootstrap s.e.m.) between T_{FH} proportions and plasmablast or plasma cell (combined); only significant trends are shown. **f**, GSEA of MSigDB hallmark signatures in naive B cells, switched memory B cells and plasmablasts for asymptomatic/symptomatic COVID-19 versus healthy groups. Size of circles indicate (absolute) normalized enrichment score (NES). GSEA (permutation) nominal $P < 0.05$ and FDR < 0.25 denoted by non-gray colored dots. EMT, epithelial–mesenchymal transition; UV, ultraviolet. **g**, Dot plots of genes related to TNF signaling and BCR signaling in naive B cells, switched memory B cells and plasmablasts. Size of circles indicates the percentage of cells expressing the gene, and color gradient corresponds to increasing mean expression value. **h**, Scatterplot of clonotype size by node closeness centrality gini indices with marginal histograms indicating the distribution. Each dot represents an individual. **i**, BCR overlap incidence plot. Nodes correspond to individual donors colored by (inner ring) severity and (outer ring) site from which samples were collected. Edges indicate if at least one cell from each individual displayed an identical combination of heavy and light-chain V and J gene usage with CDR3 similarity allowance ($\geq 85\%$). **j**, Clonotype size (left) and node closeness centrality gini indices (right) separated by gender. Mann–Whitney U test with Benjamini–Hochberg correction between the gender groups within each severity status. Color of adjusted P values indicates the gender group with the higher mean value. The box portion of the box plots extends from the 25th to 75th percentiles, whiskers extend from the smallest to largest values, and the middle line corresponds to the median. NS, not significant.

with COVID-19 and is consistent with postmortem observations showing a lack of germinal centers in lymph nodes and spleen in individuals with fatal COVID-19 and a decrease in T$_{FH}$ cells[44].

Our data revealed a significant decrease in IgA2 in symptomatic COVID-19 compared to asymptomatic donors, suggesting that maintenance of a robust mucosal humoral immune response

275
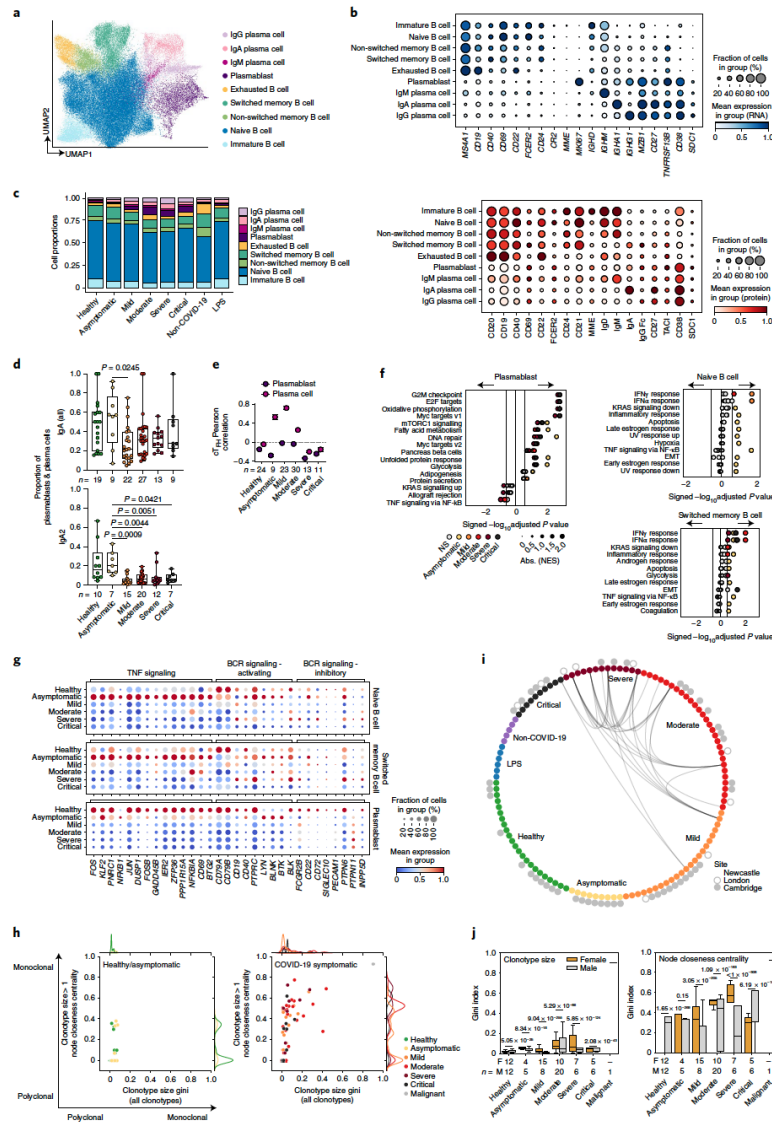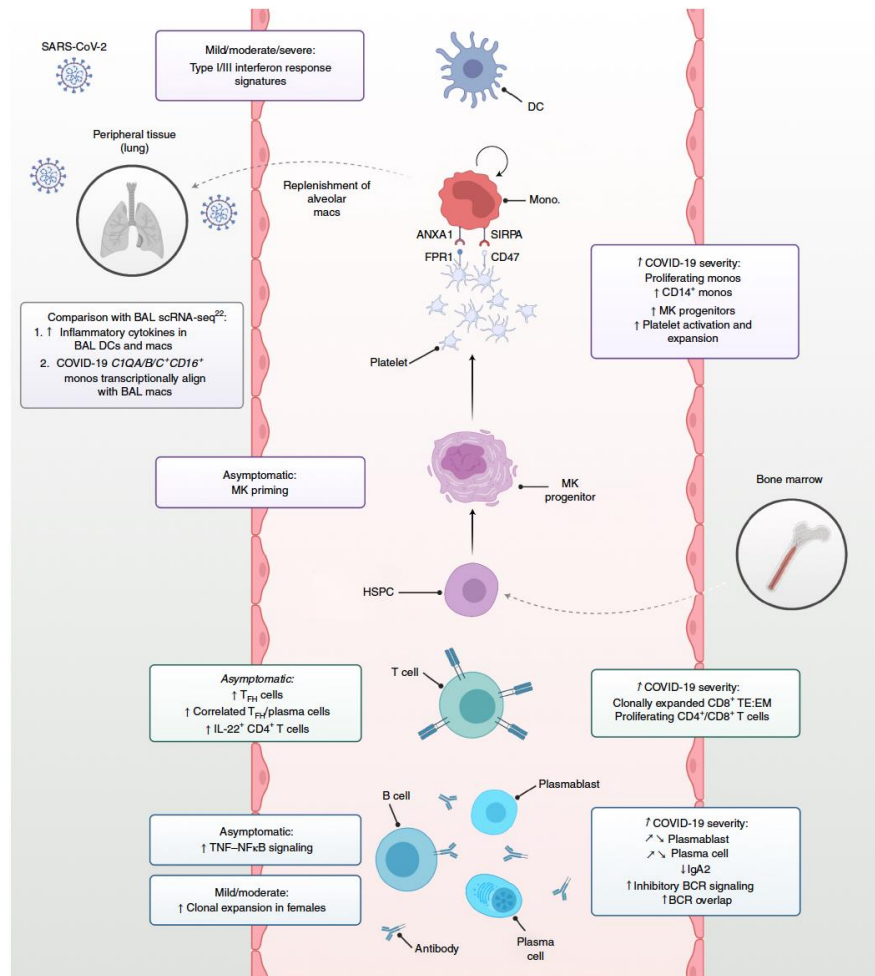
Fig. 5 | Integrated framework of the peripheral immune response in COVID-19. Schematic illustration of study highlights. Created with BioRender.com.

may influence the fate of individuals infected with SARS-CoV-2. We observed a diminished IFNα response in the B cell compartments of individuals with critical and severe disease, further emphasizing a crucial role of these responses in outcomes, as previously reported in patients with COVID-19 who had type I IFN antibodies[47]. Our data also suggest differential BCR clonality and mutation frequencies between gender groups, which may contribute to the differing clinical outcomes observed between men and women

with COVID-19[48]. Our study provides a valuable resource, exploitable for translational studies, and a template for future integrative meta-analysis of single-cell multi-omics datasets from individuals with COVID-19 worldwide.

Online content
Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information,

References

1. Bedford, J. et al. COVID-19: towards controlling of a pandemic. Lancet 395, 1015–1018 (2020).
2. John Hopkins University. COVID-19 map. Johns Hopkins Coronavirus Resource Center. https://coronavirus.jhu.edu/map.html (2020).
3. Chen, N. et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet 395, 507–513 (2020).
4. Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell 181, 271–280 (2020).
5. Diao, B. et al. Reduction and functional exhaustion of T cells in patients with coronavirus disease 2019 (COVID-19). Front. Immunol. 11, 827 (2020).
6. Lee, J. S. et al. Immunophenotyping of COVID-19 and influenza highlights the role of type I interferons in development of severe COVID-19. Sci. Immunol. 5, eabd1554 (2020).
7. Arunachalam, P. S. et al. Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. Science 369, 1210–1220 (2020).
8. Woodruff, M. C. et al. Extrafollicular B cell responses correlate with neutralizing antibodies and morbidity in COVID-19. Nat. Immunol. 21, 1506–1516 (2020).
9. Chen, G. et al. Clinical and immunological features of severe and moderate coronavirus disease 2019. J. Clin. Invest. 130, 2620–2629 (2020).
10. Zhang, J.-Y. et al. Single-cell landscape of immunological responses in patients with COVID-19. Nat. Immunol. 21, 1107–1118 (2020).
11. Liu, C. et al. Time-resolved systems immunology reveals a late juncture linked to fatal COVID-19. Cell https://doi.org/10.1016/j.cell.2021.02.018 (2021).
12. Hadjadj, J. et al. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. Science 369, 718–724 (2020).
13. Schulte-Schrepping, J. et al. Severe COVID-19 is marked by a dysregulated myeloid cell compartment. Cell 182, 1419–1440 (2020).
14. Mann, E. R. et al. Longitudinal immune profiling reveals key myeloid signatures associated with COVID-19. Sci. Immunol. 5, eabd6197 (2020).
15. Bernardes, J. P. et al. Longitudinal multi-omics analyses identify responses of megakaryocytes, erythroid cells and plasmablasts as hallmarks of severe COVID-19. Immunity 53, 1296–1314 (2020).
16. World Health Organization. Clinical management of COVID-19. https://www.who.int/publications/i/item/clinical-management-of-covid-19 (2021).
17. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat. Methods 16, 1289–1296 (2019).
18. Severe COVID-19 GWAS Group et al. Genome-wide association study of severe COVID-19 with respiratory failure. N. Engl. J. Med. 383, 1522–1534 (2020).
19. Pairo-Castineira, E. et al. Genetic mechanisms of critical illness in COVID-19. Nature https://doi.org/10.1038/s41586-020-03065-y (2020).
20. Lun, A. T. L., Richard, A. C. & Marioni, J. C. Testing for differential abundance in mass cytometry data. Nat. Methods 14, 707–709 (2017).
21. Jardine, L. et al. Lipopolysaccharide inhalation recruits monocytes and dendritic cell subsets to the alveolar airspace. Nat. Commun. 10, 1999 (2019).
22. Liao, M. et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. Nat. Med. 26, 842–844 (2020).
23. Evren, E. et al. Distinct developmental pathways from blood monocytes generate human lung macrophage diversity. Immunity https://doi.org/10.1016/j.immuni.2020.12.003 (2020).
24. Ren, X. et al. COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. Cell https://doi.org/10.1016/j.cell.2021.01.053 (2021).
25. Levi, M., Thachil, J., Iba, T. & Levy, J. H. Coagulation abnormalities and thrombosis in patients with COVID-19. Lancet Haematol. 7, e438–e440 (2020).
26. Hottz, E. D. et al. Platelet activation and platelet-monocyte aggregate formation trigger tissue factor expression in patients with severe COVID-19. Blood 136, 1330–1341 (2020).
27. Velten, L. et al. Human haematopoietic stem cell lineage commitment is a continuous process. Nat. Cell Biol. 19, 271–281 (2017).
28. Mende, N. et al. Quantitative and molecular differences distinguish adult human medullary and extramedullary haematopoietic stem and progenitor cell landscapes. Preprint at BioRxiv https://doi.org/10.1101/2020.01.26.919753 (2020).
29. Ivanov, S. et al. Interleukin-22 reduces lung inflammation during influenza A virus infection and protects against secondary bacterial infection. J. Virol. 87, 6911–6924 (2013).
30. Desai, N. et al. Temporal and spatial heterogeneity of host response to SARS-CoV-2 pulmonary infection. Nat. Commun. 11, 6319 (2020).
31. Joshi, N. S. et al. Inflammation directs memory precursor and short-lived effector CD8+ T cell fates via the graded expression of T-bet transcription factor. Immunity 27, 281–295 (2007).
32. Monaco, G. et al. RNA-seq signatures normalized by mRNA abundance allow absolute deconvolution of human immune cell types. Cell Rep. 26, 1627–1640 (2019).
33. Adlowitz, D. G. et al. Expansion of activated peripheral blood memory B cells in rheumatoid arthritis, impact of B cell depletion therapy, and biomarkers of response. PLoS ONE 10, e0128269 (2015).
34. Crotty, S. T. T follicular helper cell biology: a decade of discovery and diseases. Immunity 50, 1132–1148 (2019).
35. Meng, X. et al. Hypoxia-inducible factor-1α is a critical transcription factor for IL-10-producing B cells in autoimmune disease. Nat. Commun. 9, 251 (2018).
36. Pritchard, N. R. & Smith, K. G. C. B cell inhibitory receptors and autoimmunity. Immunology 108, 263–273 (2003).
37. Doughty, C. A. et al. Antigen receptor-mediated changes in glucose metabolism in B lymphocytes: role of phosphatidylinositol 3-kinase signaling in the glycolytic control of growth. Blood 107, 4458–4465 (2006).
38. Boothby, M. & Rickert, R. C. Metabolic regulation of the immune humoral response. Immunity 46, 743–755 (2017).
39. Guan, W.-J. et al. Clinical characteristics of coronavirus disease 2019 in China. N. Engl. J. Med. 382, 1708–1720 (2020).
40. Onder, G., Rezza, G. & Brusaferro, S. Case-fatality rate and characteristics of patients dying in relation to COVID-19 in Italy. J. Am. Med. Assoc. 323, 1775–1776 (2020).
41. Parrot, T. et al. MAIT cell activation and dynamics associated with COVID-19 disease severity. Sci. Immunol. 5, eabe1670 (2020).
42. Le Bert, N. et al. Highly functional virus-specific cellular immune response in asymptomatic SARS-CoV-2 infection. J. Exp. Med. 218, e20202617 (2021).
43. Rydyznski Moderbacher, C. et al. Antigen-specific adaptive immunity to SARS-CoV-2 in acute COVID-19 and associations with age and disease severity. Cell 183, 996–1012 (2020).
44. Kaneko, N. et al. Loss of Bcl-6-expressing T follicular helper cells and germinal centers in COVID-19. Cell 183, 143–157 (2020).
45. Maloy, K. J. et al. CD4+ T cell subsets during virus infection. Protective capacity depends on effector cytokine secretion and on migratory capability. J. Exp. Med. 191, 2159–2170 (2000).
46. Sahin, U. et al. COVID-19 vaccine BNT162b1 elicits human antibody and TH1 T cell responses. Nature 586, 594–599 (2020).
47. Bastard, P. et al. Autoantibodies against type I IFNs in patients with life-threatening COVID-19. Science 370, eabd4585 (2020).
48. Takahashi, T. et al. Sex differences in immune responses that underlie COVID-19 disease outcomes. Nature 588, 315–320 (2020).

277

Appendix D. Co-author of Nature Article: A roadmap for the Human Developmental Cell Atlas.

Perspective

A roadmap for the Human Developmental Cell Atlas

Muzlifah Haniffa[1,2,3,45 ✉], Deanne Taylor[4,5,45], Sten Linnarsson[6,45], Bruce J. Aronow[7], Gary D. Bader[8], Roger A. Barker[9,10], Pablo G. Camara[11], J. Gray Camp[12], Alain Chédotal[13], Andrew Copp[14], Heather C. Etchevers[15], Paolo Giacobini[16], Berthold Göttgens[9,17], Guoji Guo[18], Ania Hupalowska[19], Kylie R. James[2], Emily Kirby[20], Arnold Kriegstein[21], Joakim Lundeberg[22], John C. Marioni[23], Kerstin B. Meyer[2], Kathy K. Niakan[24,25], Mats Nilsson[26], Bayanne Olabi[1], Dana Pe'er[27], Aviv Regev[19,28,44], Jennifer Rood[19], Orit Rozenblatt-Rosen[19,44], Rahul Satija[29], Sarah A. Teichmann[2,30], Barbara Treutlein[31], Roser Vento-Tormo[2], Simone Webb[1] & Human Cell Atlas Developmental Biological Network*

The Human Developmental Cell Atlas (HDCA) initiative, which is part of the Human Cell Atlas, aims to create a comprehensive reference map of cells during development. This will be critical to understanding normal organogenesis, the effect of mutations, environmental factors and infectious agents on human development, congenital and childhood disorders, and the cellular basis of ageing, cancer and regenerative medicine. Here we outline the HDCA initiative and the challenges of mapping and modelling human development using state-of-the-art technologies to create a reference atlas across gestation. Similar to the Human Genome Project, the HDCA will integrate the output from a growing community of scientists who are mapping human development into a unified atlas. We describe the early milestones that have been achieved and the use of human stem-cell-derived cultures, organoids and animal models to inform the HDCA, especially for prenatal tissues that are hard to acquire. Finally, we provide a roadmap towards a complete atlas of human development.

Most modern developmental biology research has historically focused on model organisms. Owing to practical challenges, human development—from a fertilized ovum to a fully formed fetus at birth—has remained a poorly understood 'black box'. The implications of a human developmental cell atlas for understanding human development are far-reaching, as many congenital disorders and childhood cancers may originate during susceptible windows of development[1-3]. The clinical relevance of the atlas extends into adulthood for ageing, cancer and applications in regenerative medicine and stem cell therapies[4-6]. Furthermore, embryonic and fetal stem cells[7,8] and developmental trajectories provide an essential reference and guide for engineering human stem-cell-derived models[9-13], organoids[14] and cellular therapies.

Human development begins with a fertilized oocyte that divides and differentiates through preimplantation, embryonic and fetal stages (Fig. 1). Early studies began with morphometric and qualitative assessments of human embryos, leading to development of the Carnegie staging system[15] (Fig. 1). Advances in imaging, cytometry and genomics technologies have provided further insights into the complex spatiotemporal changes during organogenesis[16]. Recent progress in single-cell profiling technologies has revolutionized our ability to study human development at an unprecedented resolution[17]. Leveraging these advances to build a comprehensive atlas of human development (from the fertilized oocyte to birth) at cellular resolution is an ambitious endeavour that is similar in scale to the Human Genome Project,

[1]Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK. [2]Wellcome Sanger Institute, Hinxton, UK. [3]Department of Dermatology and NIHR Newcastle Biomedical Research Centre, Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [4]Department of Biomedical and Health Informatics (DBHi), The Children's Hospital of Philadelphia, Philadelphia, PA, USA. [5]Department of Pediatrics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA. [6]Division of Molecular Neurobiology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden. [7]Division of Developmental Biology and Biomedical Informatics, Cincinnati Children's Hospital Medical Centre, Cincinnati, OH, USA. [8]The Donnelly Centre, University of Toronto, Toronto, Ontario, Canada. [9]Wellcome and MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK. [10]Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK. [11]Department of Genetics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA. [12]Institute of Molecular and Clinical Ophthalmology Basel (IOB), University of Basel, Basel, Switzerland. [13]INSERM, CNRS, Institut de la Vision, Sorbonne Université, Paris, France. [14]Developmental Biology and Cancer Programme, UCL Great Ormond Street Institute of Child Health, London, UK. [15]MMG, INSERM, U1251, Aix Marseille Université, Marseille, France. [16]Laboratory of Development and Plasticity of the Neuroendocrine Brain, Inserm, CHU Lille, Lille Neuroscience and Cognition, UMR-S 1172, Université Lille, Lille, France. [17]Department of Haematology, University of Cambridge, Cambridge, UK. [18]Center for Stem Cell and Regenerative Medicine, Zhejiang University School of Medicine, Hangzhou, China. [19]Klarman Cell Observatory, Broad Institute of Harvard and MIT, Cambridge, MA, USA. [20]Centre of Genomics and Policy, McGill University, Montreal, Quebec, Canada. [21]Department of Neurology, University of California San Francisco (UCSF), San Francisco, CA, USA. [22]Science for Life Laboratory, KTH Royal Institute of Technology, Solna, Sweden. [23]Cancer Research Institute UK Cambridge Institute, University of Cambridge, Cambridge, UK. [24]Francis Crick Institute, London, UK. [25]Centre for Trophoblast Research, Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK. [26]Science for Life Laboratory, Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden. [27]Computational and Systems Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [28]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA. [29]New York Genome Center, New York University, New York, NY, USA. [30]Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge, UK. [31]Department of Biosystems Science and Engineering, Eidgenössische Technische Hochschule (ETH) Zurich, Basel, Switzerland. [44]Present address: Genentech, South San Francisco, CA, USA. [45]These authors jointly supervised this work: Muzlifah Haniffa, Deanne Taylor, Sten Linnarsson. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: m.a.haniffa@ncl.ac.uk

278
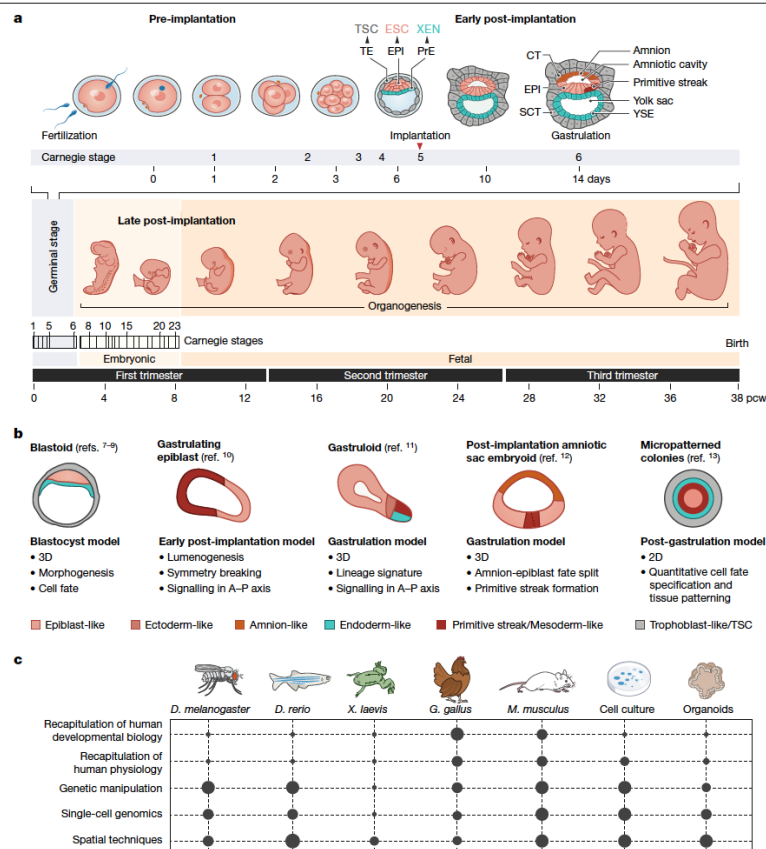
Fig. 1 | Human embryo development and model systems. a, Timeline of human development from fertilization to birth. CT, cytotrophoblast; ESC, embryonic stem cell; EPI, epiblast; PrE, primitive endoderm; pcw, post-conception weeks; SCT, syncytiotrophoblast; TE, trophectoderm; TSC, trophoblast stem cell; XEN, extraembryonic endoderm; YSE, yolk sac endoderm. **b**, Models derived from human stem cells, and associated studies.

In vitro model systems to study early embryonic development. A–P, anterior–posterior. **c**, Experimental model systems to study development, including *Drosophila melanogaster, Danio rerio, Xenopus laevis, Gallus gallus, Mus musculus*, cell culture and organoids, and their amenability to facilitating various aspects of scientific study.

which required multidisciplinary scientific expertise from disparate fields working together collaboratively. Such a community has arisen from the grassroots assembly of global researchers who are working as part of the Human Cell Atlas (HCA)[18] initiative. As with the Human Genome Project, the HCA will be a foundational scientific resource, composed of diverse data types and available freely through browsable and searchable web portals that visualize cells across anatomical space and developmental time.

The HDCA is a strategic focus of HCA[19], and is pursued by scientists from individual laboratories as well as large national and international research consortia (Supplementary Table 1); the HDCA is open to all who adhere to its mission and open science values[20]. The HDCA aims for equity, inclusivity and diversity both in terms of scientific participation

and the representation of human tissue samples. We encourage any interested researcher to become a member, participate, register their study and contribute their data and publications to the HDCA and HCA[21].

Building a developmental cell atlas

The successful construction of a HDCA poses substantial scientific challenges in terms of experimental measurement technologies, computational analysis and visualization algorithms (Fig. 2). In particular, the dynamic nature of gestation creates challenges for designing a sampling strategy, especially to capture transient morphological changes in the first eight weeks. A major endeavour for the HDCA will

Perspective
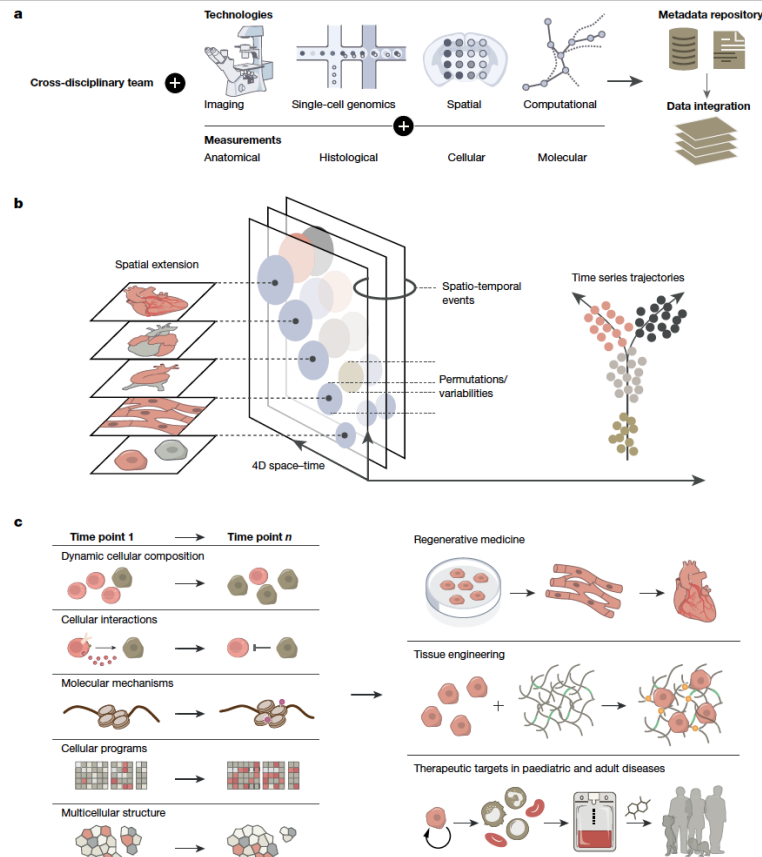
Fig. 2 | Building the HDCA. a, 'How to build an atlas' modules, including an interdisciplinary team (biological disciplines, clinical specialities, computational scientists and technology engineers), multimodal technologies and integration of data across platforms. **b,** Key features of the HDCA. Single-cell measurements across three-dimensional space (alongside a fourth dimension of time) allow for the capture of dynamic developmental processes, including cell proliferation, migration and regulation. Spatial extension captures cell proliferation (changes in organ size and shape), and the sensing and control of morphogenesis; spatio-temporal events include cell differentiation and replication, and changes in cellular composition and gene expression. Permutations and variabilities outline the spatial determinants of cell differentiation, mechanical adhesive feedback and heterocellular signalling by ligands. Time series trajectories reveal fate-biased migrating progenitors, lineage specifications and cell migration. **c,** Utility and applications of the HDCA: cellular and molecular biological insights into the coordination of organ development across the whole embryo (left) are applied to advance regenerative medicine, tissue engineering and therapeutic strategies (right).

be to develop the conceptual and computational framework to capture development with respect to cellular and morphological changes. The HDCA, through coordination with the HCA Organoid Network[22], will incorporate data from in vitro culture model and organoid systems[23] to cautiously infer development between seven days and four weeks after conception (a period in which samples are difficult to obtain) (Fig. 1b, c).

The successful delivery of the HDCA will leverage the Human Genome Project-initiated restructuring of how large science projects are funded, conducted, coordinated and shared (based on the Fort Lauderdale Principles[24]) that forms the basis for the HCA, its committees (for example, computation and ethics) and 'Biological Networks'[20]. This organizational framework has enabled researchers to form large-scale coordinated collaborations across technologies and biological disciplines: developmental biology, embryology, genetics and model systems, computational biology, clinical specialities (including in vitro fertilization), clinical genetics and pathology, as well as coordination with funders. Partnerships with allied biological networks, including organoid and paediatric atlas projects, will facilitate clinical applications.
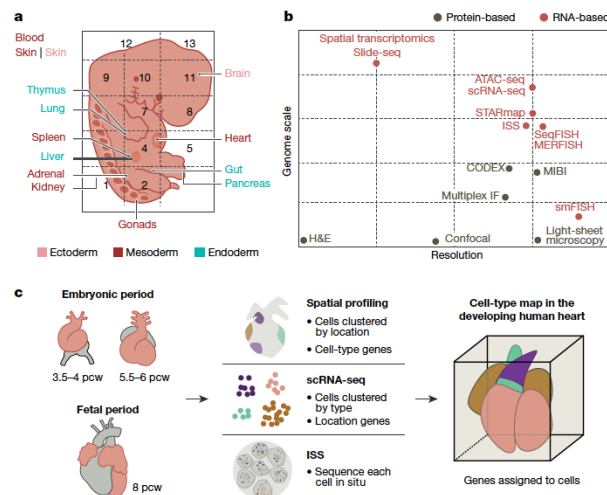
Fig. 3 | Multi-omics profiling and data integration. a, Organ or anatomical unit profiling of a prenatal embryo derived from multiple germ layers. **b**, Single-cell atlas technologies by relative resolution and genome scale. CODEX, co-detection by indexing; H&E, haematoxylin and eosin staining; IF, immunofluorescence; ISS, in situ sequencing; MERFISH, multiplexed error-robust FISH; MIBI, multiplexed ion beam imaging; seqFISH, sequential FISH; smFISH, single-molecule FISH; STARmap, spatially resolved transcript annotation readout mapping. **c**, Integration of datasets from different technologies (for example, spatial transcriptomics, scRNA-seq and targeted ISS) to profile organs or whole embryos.

Ethics, resources and data sharing

Accessing human developmental samples is constrained by general and geographically specific ethical and legal challenges. These include issues relating to donation, access to and research use of legally defined developing human tissue material, regulatory approvals processes and cultural sensitivities. Research on human embryos and fetuses is supported within European and national regulations, such as the UK National Research Ethics Service (NRES) and the French Agence de Biomédecine. In the UK, studies on preimplantation human embryos (up to 14 days after conception) are governed by the Human Fertilisation and Embryology Authority and a research ethics committee (such as NRES). However, in the USA, research on donated human embryonic and fetal materials has increasingly been restricted over the past two decades, despite the existence of similar regulatory oversight.

Nonetheless, resources to support research in human development (such as the UK's Human Developmental Biology Resource (HDBR)[25]) provide material to researchers. Recipients of HDBR tissue who are not based in the UK require their own project-specific ethics approval, before receipt of material. The HDBR provides embryonic and fetal samples from 4 to 20 weeks after conception with karyotype information and, increasingly, with anonymized maternal DNA and clinical history. Material from fetuses with prenatally diagnosed disorders is also available. The French Human Developmental Cell Atlas (HuDeCA) (https://hudeca.genouest.org) has recently been established, and aspires to constitute a comprehensive European resource of human embryonic or early fetal samples.

The international sharing of genomic sequencing and clinical data derived from prenatal or paediatric tissue samples is subject to data protection regulation that considers live versus deceased status, consent regarding research data use and confidentiality. Data from living donors are shared under appropriate access controls. The HCA Ethics Working Group is developing tools, guidance notes[26], consent-form templates and sampling information for embryonic, fetal and paediatric tissue material, and international data-sharing guidance for the HDCA.

Mapping development across space and time

Development is intricately orchestrated in three spatial dimensions and gestation time. Human embryogenesis cannot be easily assessed at high resolution in vivo[27]. Time-lapse studies are limited to in vitro pre-implantation embryos. The application of high-throughput genomics technologies to dissociated cells and tissue sections in situ is beginning to provide data of unprecedented resolution (Fig. 3, Table 1).

Cellular and molecular heterogeneity

Single-cell molecular profiles based on RNA, chromatin accessibility, methylation or select protein signatures have enabled a more nuanced definition of cell types and states. The data underpinning such definitions are increasingly derived from single-cell RNA sequencing (scRNA-seq), barcoded antibodies and accessible chromatin sequencing of dissociated cells[28,29]. Resolving cell types and trajectories at high granularity is aided by full-length scRNA-seq, but is primarily performed by profiling large numbers of cells. Cell-type definition is currently guided by existing knowledge from model organisms and adult cellular profiles, which may not faithfully reflect prenatal cell types, transient cell types that are present only during development and transitional states of differentiation.

To overcome these challenges, many time points need to be profiled and defined cell states need to be mapped back into their 3D space over time and functionally characterized. High levels of multiplexing can attain this level of granularity at an affordable cost for a complete human developmental cell atlas[30,31]. Molecular profiles, morphology,

Perspective

Table 1 | Publications registered with the HDCA

Organ	Stage	Main highlights	Publications
Brain	First and second trimester	Specific brain regions studied, including prefrontal cortex and neocortex; tracing of developmental trajectories of cells; characterization of mechanisms that underlie neuron generation and circuit formation	Refs. [59,60,63]
Gut	First and second trimester, organoids	Profile of transcriptomes of cycling epithelial precursor cells; evaluation of the effect of mesenchymal cells on LGR5 stem cells; comparison of transcriptomes of ex vivo tissues and in vitro fetal organoids, and of transcriptome profiles from paediatric Crohn's disease epithelium with matched healthy controls	Refs. [58,60,64,124]
Heart	First and second trimester	Identification of unique gene profiles that correspond to distinct anatomical regions in each developmental stage; integration of scRNA-seq and spatial data; generation of a web resource of the human developing heart	Refs. [57,60,62,119]
Liver and fetal haematopoiesis	First and second trimester	Identification of the repertoire of human blood and immune cells, and of differentiation trajectories from haematopoietic stem cells and multipotent progenitors; evaluation of the effect of tissue microenvironment on blood and immune-cell development	Refs. [54,60]
Kidney	First trimester	Identification of both known and unknown transcription factors associated with nephron development; characterization of myeloid and lymphoid populations present during fetal development	Refs. [54,60,61,65]
Placenta	First trimester	Characterization of cellular organization of the decidua and placenta; identification of perivascular and stromal cellular subsets; development of a repository of ligand–receptor complexes, and of a statistical tool to predict the cell-type specificity of cell–cell communication via receptor–ligand interactions	Ref. [56]
Thymus	First and second-trimester, paediatric	Identification of more than 50 cell states, novel subpopulations of thymic fibroblasts and epithelial cells, and a cellular network of the thymic niche for T cell development	Ref. [78]
Skin	First trimester	Identification of physiological erythropoiesis; enrichment of innate immune cells; co-option of developmental programs identified in adult inflammatory skin diseases	Refs. [54,114]
Multi-organ	First and second trimester	Integrated analyses of transcriptomes and chromatin accessibility from several fetal organs, including brain, heart, lung, gut, kidney, adrenal glands, stomach, pancreas, spleen, gonads, muscle, eye and skin	Refs. [60,85,118]

There are 48 researchers from 13 countries currently registered with the HDCA. Developmental datasets are contributed to public repositories, including the HCA Data Coordination Portal. Further information on HCA publications is available at https://www.humancellatlas.org/publications.

functional assessment and other features can reflect the multifaceted state of a cell. For example, the transcriptome reflects the present and potential future of a cell; protein expression captures the immediate past and present state of a cell; chromatin profiles reveal its invariant type and potential for future differentiation; and ontogeny reveals its history.

The field of developmental biology has traditionally drawn on ontogenic relationships to define cell types, but this is challenging in humans for whom information is captured as snapshots across gestation. CRISPR scarring is applicable only in stem cells, organoid systems and short-term explants[32,33]. The tracking of somatic mutations is the only available technology to definitively determine ontogeny, but is limited by its current lack of scalability[34,35]. Recent methods that rely on the simultaneous measurement of mitochondrial DNA and RNA, transcriptome and open chromatin may overcome this challenge[36,37]. We anticipate the field moving towards a consensus cell ontology that integrates multimodal single-cell profiling data as well as legacy knowledge of embryonic cell-type definitions augmented by information from diverse animal models.

Mapping cells in 2D and 3D

Spatial genomics methods to measure RNA in tissue sections typically offer a trade-off: high-resolution (single-cell and subcellular) methods that typically measure hundreds of transcripts or whole transcriptome profiles at a multicellular level[38,39]. This trade-off can be mitigated by integration with single-cell profiles from dissociated cells, expanding the genomic coverage by predicting the spatial expression of unmeasured genes or enhancing resolution by deconvolution of multicellular measurements. Tissue clearing methods to render organs transparent[40], combined with whole-mount protein immunostaining and RNA single-molecule fluorescence in situ hybridization (FISH)[41,42], can now provide 3D molecular profiling at cellular or subcellular resolution

using light-sheet microscopy[43–45]. Increasing multiplex capacity and use of artificial intelligence and machine learning algorithms to overcome data analytical challenges have successfully been deployed to image whole-organismal vasculature following tissue clearing[46,47].

Biophysical methods and live imaging

Mounting evidence from *Drosophila* and other models shows that mechanical forces have a key role in development processes and tissue morphogenesis[48]. Surface tension and pressure can be measured in single cells of preimplantation mouse embryos[49]. Adapting these technologies to human preimplantation embryos and stem-cell-based embryo models[50] can build a spatiotemporal mechanical atlas.

Positional landmarks in development

A standard coordinate system for locations in the human body (a common coordinate framework (CCF)) is crucial for the HCA and HDCA[51]. Two types of systems are useful: absolute (similar to postcode or zip code addresses) and relative (similar to a landmark-based address system). CCF anatomical 'postcodes' enable the integration of multimodal datasets of different spatial and longitudinal resolution. The Allen Mouse Brain Reference Atlas version 3 provides a CCF of 3D anatomical features and local features grouped in a hierarchy to facilitate multilevel analysis of the mouse brain. Efforts are currently underway to establish CCFs for adult human organs within the Human Biomolecular Atlas Program of the National Institutes of Health (NIH). The HDCA will need to develop a CCF that incorporates space and time, as well as cell movement and patterns during organogenesis on the basis of existing macro-level 3D coordinates for human embryos (such as the HDBR atlas (http://hdbratlas.org/) and the Transparent Human Embryo (https://transparent-human-embryo.com/)).
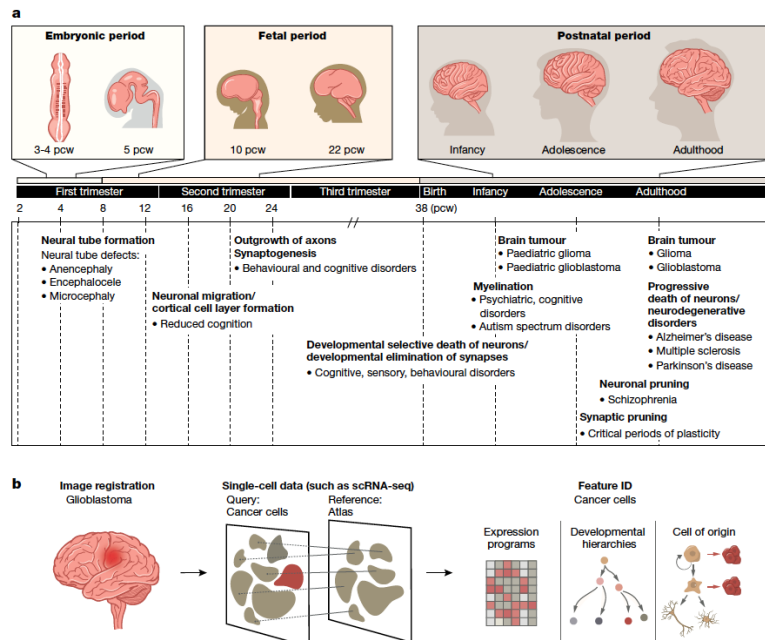
Fig. 4 | Clinical relevance and applications of the HDCA. a, Diseases related to brain development, across lifespan. A timeline of brain development across human life, with examples of diseases with onset at different gestational stages and ages. **b**, Disease state compared to developmental atlas, showing how a single-cell atlas with temporal and spatial information can be used as a reference to understand disease states.

Computation and data visualization

Among the key algorithmic challenges to integrating data into a developmental atlas are (1) mapping cells with more intermediate states compared to adult counterparts; (2) inferring time orderings and lineage relations, including branching lineages and multiple paths converging on the same outcome; (3) inferring spatial movement of cells; (4) building a temporal series of CCFs, each as a probabilistic model for a time window as well as a model for their morphing along space and time[52]; (5) mapping across modalities and time points (for example, chromatin states in one time window to RNA and protein levels of another); and (6) regulatory and molecular network inference within and across cells. Theories and insights from multiple fields will be required to model the mechanisms that underpin tissue formation and growth. It is likely that additional emergent properties of cells and their ecosystems will be discovered using interdisciplinary approaches. These will need new vocabularies, ontologies and modelling approaches to be understood. The HDCA community must also apply FAIR (findability, accessibility, interoperability and reusability) principles to help to ensure reproducibility and data accessibility[53].

Computational integration of multi-omics data for visualizations similar to that of Google Maps, such as the Open Microscopy Environment (https://www.openmicroscopy.org/), will enable zooming to the single-cell level from a large-volume tissue view. Additional complexity will combine visualizations from imaging and sequencing data. A sophisticated abstraction of raw data and integration across modalities, anchored by a developmental CCF, will be essential. Links to clinical relevance and applications will enhance the utility of the atlas.

Emerging cell atlases of human development

The advantages of whole tissue or organ profiling compared to lineage-centric analysis include comprehensive cellular analysis and the discovery of emergent biological properties. For example, the developing liver functions as a haematopoietic organ during early gestation until the middle of the second trimester, before it functionally transitions into a metabolic organ similar to the adult liver[54]. To meet the high demand for erythropoiesis during development, the human skin and adrenal glands can also support erythrocyte maturation during the first trimester[54,55].

In contrast to our terrestrial postnatal life, the human embryo and fetus exist in an aquatic environment: our lung, gut and skin are exposed to amniotic fluid. In contrast to the postnatal lung, the developing lung does not perform oxygen transfer or receive the same volume of blood through the pulmonary veins. The effect of these physiological factors on individual tissues and the role of the placenta and maternal decidua in supporting human embryogenesis and fetal life are emerging[56,57].

Current organ atlases of brain, gut, heart, liver, kidney, placenta, thymus and skin (Table 1) underscore the importance of studying human samples and have revealed unique aspects of human development that are not conserved with animal model systems[58–61]. These include

Perspective

timelines of development during gestation, cell-type markers and the expression pattern of transcription factors between mouse and human organs[62,63].

The specification of functional tissue niches occurs during both prenatal and postnatal life. Studies of the fetal gut have highlighted the importance of interactions between the epithelial and mesenchymal compartments in allowing the formation of villi, and have identified fetal gut transcription factors that are aberrantly activated in individuals with paediatric Crohn's disease[64]. Comparison between the developing and adult kidney have demonstrated the establishment of a dedicated spatial zonation pattern that protects against uropathogenic bacterial challenges postnatally[61,65]. Single-cell transcriptomics of germ cells during development have provided important insights into the main pathways that control their differentiation[66,67], with ongoing studies focused on exploring the regulatory mechanisms of sex determination (https://hugodeca-project.eu).

Early developmental studies of the brain have focused on human and primate cortical development[68-70]. The developing human and rodent midbrain, which contains the clinically relevant dopaminergic cell groups that are lost in individuals with Parkinson's disease, has also extensively been studied[63,71,72], as have the developing mouse spinal cord and cerebellum[73,74], the hypothalamic arcuate nucleus and the diencephalon[75].

Atlases of distributed systems (such as the immune system) have been initiated, detailing haematopoietic organs such as the yolk sac[76,77] and liver[54], lymphoid tissues such as thymus (in which T cells differentiate)[78] and non-lymphoid tissues such as skin and kidney, in which immune cells reside. These studies have revealed an intrinsic change in the differentiation potential of haematopoietic stem progenitor cells with gestational time, together with the importance of the local tissue microenvironment for blood and immune-cell development.

Model organisms and culture systems

Our understanding of human development has largely been inferred from studies on animal model systems that are not always conserved across species[79] (Fig. 1). Two recent studies contrast the kinetics of development between human and mouse, highlighting the need for caution in interpreting heterospecific graft studies and findings from nonprimate preclinical models[80,81]. However, the feasibility of perturbation and in-depth mechanistic studies using animal models and culture systems provide a valuable scaffold and complement the HDCA, particularly for the immediate weeks after implantation during which human samples are inaccessible.

Single-cell molecular profiling has transformed many aspects of developmental biology research across all major model organisms[82-86], providing mechanistic insights into fundamental biological processes (including the early specification of germ layers and diversification of early cardiovascular cells)[29,87]. Comparative biology has the potential to make major contributions to cell ontology. The availability of parallel human and model species data will support expanded cross-species analyses. Computational analysis can align cells and inferred lineages across species to extrapolate findings from nonprimate models and help to optimize animal models of normal and pathological human development. From a computational perspective, it will be important to develop tools for better annotation of 3′ and 5′ untranslated regions of animal model data, as most scRNA-seq technologies capture only these regions. The development of computational tools that can robustly map developmental trajectories across species and that can account for different developmental kinetics between cell types within and between species will be required. Comparative studies of human and mouse preimplantation and gastrulation embryos have revealed conserved and divergent transcriptional programs. For example, *Klf2* expression in mouse embryo-fated epiblast progenitor cells is not observed in humans; by contrast, *KLF17* is enriched in human, but not mouse, epiblast[88].

The self-organization of human embryonic tissue can be captured from the earliest moments in vitro[50,89], and extended to gastrulation, anterior–posterior embryonic patterning and the early phases of somitogenesis[11]. The recent human gastrulation embryo dataset will be informative as a benchmark to further refine in vitro directed differentiation of human cells, including gastruloid models[11]. Other processes during organogenesis can also be monitored, including the clock control of somite segmentation[90,91], boundary formations during hepato-biliary–pancreatic organ budding[92] and patterning of the neural tube. Protocols are now established to mimic the development of diverse human tissues that exhibit morphologies and physiological functionalities of developing human tissues. These organoid systems include hair-bearing skin[93]; the small intestine with a crypt–villus axis[94]; region-specific[95] and multiregion[96] brain tissue that models neurogenesis, neural migration and synapse formation; multilayered neural retina with photoreception responses[97]; and arterio-venous specification during blood vessel development[98].

A comprehensive reference atlas of the cell types and states that are present during human development will be critical to benchmark stem-cell-derived organoids. Such roadmap comparisons will highlight similarities[69] and deficiencies[99], and define strategies for improving organoids for disease modelling. In the future, high-fidelity human stem-cell-derived human organoids and single-cell multi-omic modalities will be powerful tools to understand the mechanisms that control human organogenesis.

Clinical relevance and applications

The interaction of genotype and environment that leads to phenotype underlies developmental disorders. A range of childhood and adult disorders have their origins in prenatal life (Fig. 4). These include structural birth defects[100], neurodevelopmental disorders (including schizophrenia)[101], childhood cancers[2,65], inborn errors of immunity[102], infertility and differences of sex development[103], as well as many paediatric disorders[104]. Thousands of rare genetic diseases can each present a spectrum of perturbed developmental sequelae at birth, and sometimes differ widely in medical presentation even when classified as the same disease or condition[105]. As examples, Down syndrome (trisomy 21)[106] and 22q11.2 deletion syndrome[107] separately present substantial risks for schizophrenia, Alzheimer's disease and hypothyroidism starting in adolescence[108]. Identifying the aetiology of developmental disorders and the effects of maternal genotype, paternal age and other external risk factors (such as diet, alcohol, toxins, endocrine disruptors and pathogens) has been hampered by our limited understanding of normal development in humans.

Development atlases are also revealing the pathogenesis of childhood cancers (Fig. 4). Paediatric and adult brain tumours in their early stages often present impaired developmental programs within tumour cells[109,110]. Comparing the expression profile of tumour cells with the HDCA can identify the cancer cell of origin and its oncogenic pathways. For example, a single-cell atlas of the developing mouse cerebellum has been used to investigate subtypes of human medulloblastoma (a paediatric brain tumour)[2,111], and cell states during nephrogenesis revealed the developmental cellular origin of Wilms' tumour[65]. High-resolution mapping of developing immune cells will inform the molecular basis and extent of disease phenotypes of childhood leukaemias and primary immunodeficiencies.

Many adult cancers also recapitulate a dysregulated version of human developmental programs[112]. The acquisition of early developmental molecular programs is characteristic of malignant pathology, and is a previously unrecognized hallmark of immunological disease and the cancer immune environment[113,114]. HDCA data have also facilitated our understanding of the differential susceptibility of adult and prenatal cells to SARS-CoV-2 through examination of viral entry receptor and protease expression in a wide range of organs[115].

Cell and tissue engineering for clinical therapies and regenerative medicine are areas with considerable potential for the direct utility of the HDCA. Cell therapies derived from human pluripotent stem cells are now entering early clinical trials for the treatment of Parkinson's disease[116], using protocols that were refined on the basis of developmental studies of midbrain dopaminergic neurons[72]. Similar approaches are being followed to develop a range of other stem cell products for human trials[117]. Haematopoietic stem cell transplantation is an established and widely used treatment for many haematological, and increasingly non-haematological, disorders. Leveraging the potency factors of fetal haematopoietic stem cells could have a substantial benefit for patients who receive transplants of haematopoietic stem cells.

Towards a whole embryo atlas

The initial HCA white paper emphasized 12 distinct organ systems within the human body and highlighted the importance of a developmental cell atlas. Integrated multi-organ analyses will provide insights into the tissue microenvironment that shapes resident epithelial, stroma and immune cells and the cellular heterogeneity of innervating blood vessels, lymphatics and peripheral nerves. Eventually, this may illuminate system-level lineage development and cell fate decisions across an entire organism. The datasets from profiling based on human developmental organs have been critical in interpreting recent multi-organ developmental atlases[55,118].

There are several large-scale organ-based studies being undertaken by HDCA researchers. These include the NIH 'Brain Research through Advancing Innovative Neurotechnologies' (BRAIN) initiative (including the BRAIN Initiative Cell Census Network (BICCN) consortium) focusing on the developing human cortex; the Swedish Human Cell Atlas consortium performing large-scale scRNA-seq, assay for transposase-accessible chromatin using sequencing (ATAC-seq) and spatial-omic analyses of the developing human brain, heart[119] and lung during the first trimester; the French HuDeCA consortium mapping eight first-trimester human organs using 3D imaging and scRNA-seq; the European Union (EU) Horizon 2020-funded developing brain (Braintime) and gonad (HUGODECA) projects; the NIH Developmental Genotype-Tissue Expression (dGTEx)[120] project; and UK consortia funded by the Wellcome Trust and Medical Research Council. The logical next step will be to coordinate these efforts and extend the current approach to contextualize the development of different cell lineages across all organs.

However, multi-organ approaches do not permit the analysis of distributed tissue networks as a continuum from a single donor sample. Whole-embryo analysis has been limited to very early preimplantation samples[88,121,122] and one gastrulation-stage embryo[123]. Multi-omics suspension and spatial-genomics profiling of anatomically dissected units from whole human embryos at six to seven weeks after conception are being undertaken by the HDCA researchers based in the UK. We anticipate a first whole human embryo profiling within the next two years. On the basis of existing HDCA data and the rapid changes during early development, we propose a minimum of three replicates for each biologically relevant gestation period (for example, each week from six weeks after conception). All such data produced and shared by the global research community (formally registered with the HCA or not) contributes to the HDCA. Defining a universal organizing framework for these data will enable them to be unified into a complete atlas that will be a transformative resource for the research and clinical communities.

1. Behjati, S., Lindsay, S., Teichmann, S. A. & Haniffa, M. Mapping human development at single-cell resolution. *Development* **145**, dev152561 (2018).
2. Vladoiu, M. C. et al. Childhood cerebellar tumours mirror conserved fetal transcriptional programs. *Nature* **572**, 67–73 (2019).
3. Velmeshev, D. et al. Single-cell genomics identifies cell type-specific molecular changes in autism. *Science* **364**, 685–689 (2019).
4. Gulsuner, S. et al. Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell* **154**, 518–529 (2013).
5. Simmons, R. A. Developmental origins of adult disease. *Pediatr. Clin. North Am.* **56**, 449–466 (2009).
6. Laughney, A. M. et al. Regenerative lineages and immune-mediated pruning in lung cancer metastasis. *Nat. Med.* **26**, 259–269 (2020).
7. Sozen, B., Jorgensen, V., Zhu, M., Cui, T. & Zernicka-Goetz, M. Reconstructing human early embryogenesis in vitro with pluripotent stem cells. Preprint at https://doi.org/10.1101/2021.03.12.435175 (2021).
8. Yu, L. et al. Blastocyst-like structures generated from human pluripotent stem cells. *Nature* **591**, 620–626 (2021).
9. Liu, X. et al. Modelling human blastocysts by reprogramming fibroblasts into iBlastoids. *Nature* **591**, 627–632 (2021).
10. Simunovic, M. et al. A 3D model of a human epiblast reveals BMP4-driven symmetry breaking. *Nat. Cell Biol.* **21**, 900–910 (2019).
11. Moris, N. et al. An in vitro model of early anteroposterior organization during human development. *Nature* **582**, 410–415 (2020).
12. Shao, Y. et al. A pluripotent stem cell-based model for post-implantation human amniotic sac development. *Nat. Commun.* **8**, 208 (2017).
13. Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D. & Brivanlou, A. H. A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nat. Methods* **11**, 847–854 (2014).
14. Camp, J. G., Wollny, D. & Treutlein, B. Single-cell genomics to guide human stem cell and tissue engineering. *Nat. Methods* **15**, 661–667 (2018). **This review highlights the potential utility of single-cell genomics to optimize cell and tissue engineering, with a focus on emerging methodologies that can guide this process (such as transcription factor combinatorics, spatial reconstruction, CRISPR–Cas9 screens and lineage-coupled transcriptomics).**
15. Morgan, L. *Icons of Life: A Cultural History of Human Embryos* (Univ. California Press, 2009).
16. Blonder, L. X. Morphogenesis: The cellular and molecular processes of developmental anatomy. By Jonathan Bard. xi 313 pp. New York: Cambridge University Press. 1990. $37.95. (paper). *Am. J. Hum. Biol.* **5**, 245–246 (1993).
17. Aldridge, S. & Teichmann, S. A. Single cell transcriptomics comes of age. *Nat. Commun.* **11**, 4307 (2020).
18. Human Cell Atlas. Home, https://www.humancellatlas.org/ (2021).
19. Regev, A. et al. Science forum: the human cell atlas. *eLife* **6**, e27041 (2017).
20. The HCA Consortium. *The Human Cell Atlas White Paper*, https://www.humancellatlas.org/wp-content/uploads/2019/11/HCA_WhitePaper_18Oct2017-copyright.pdf (2019).
21. Human Cell Atlas. Register, https://www.humancellatlas.org/register (2021).
22. Bock, C. et al. The organoid cell atlas. *Nat. Biotechnol.* **39**, 13–17 (2021).
23. Subbaraman, N. Lab-grown structures mimic human embryo's earliest stage yet. *Nature* **591**, 510–511 (2021).
24. Wellcome Trust. *Sharing Data from Large-scale Biological Research Projects: A System of Tripartite Responsibility*, http://www.genome.gov/Pages/Research/WellcomeReport0303.pdf (Wellcome Trust, 2003).
25. Gerrelli, D., Lisgo, S., Copp, A. J. & Lindsay, S. Enabling research with human embryonic and fetal tissue resources. *Development* **142**, 3073–3076 (2015). **The HDBR is a biobank that collects and distributes material for research from human embryos (from 4 weeks after conception) and fetuses (up to 22 weeks after conception);** https://www.hdbr.org/ **shows the range of facilities offered by the HDBR and provides access for prospective users.**
26. Human Cell Atlas. Ethics, https://www.humancellatlas.org/ethics/ (2021).
27. Huang, Q. et al. Intravital imaging of mouse embryos. *Science* **368**, 181–186 (2020).
28. Mereu, E. et al. Benchmarking single-cell RNA-sequencing protocols for cell atlas projects. *Nat. Biotechnol.* **38**, 747–755 (2020).
29. Argelaguet, R. et al. Multi-omics profiling of mouse gastrulation at single-cell resolution. *Nature* **576**, 487–491 (2019).
30. Cao, J. et al. Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **357**, 661–667 (2017).
31. McGinnis, C. S. et al. MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat. Methods* **16**, 619–626 (2019).
32. Fujii, M., Clevers, H. & Sato, T. Modeling human digestive diseases with CRISPR–Cas9-modified organoids. *Gastroenterology* **156**, 562–576 (2019).
33. Artegiani, B. et al. Fast and efficient generation of knock-in human organoids using homology-independent CRISPR–Cas9 precision genome editing. *Nat. Cell Biol.* **22**, 321–331 (2020).
34. Lee-Six, H. et al. Population dynamics of normal human blood inferred from somatic mutations. *Nature* **561**, 473–478 (2018).
35. D'Gama, A. M. & Walsh, C. A. Somatic mosaicism and neurodevelopmental disease. *Nat. Neurosci.* **21**, 1504–1514 (2018).
36. Ludwig, L. S. et al. Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. *Cell* **176**, 1325–1339 (2019).
37. Lareau, C. A. et al. Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling. *Nat. Biotechnol.* **39**, 451–461 (2021).
38. Ståhl, P. L. et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (2016).
39. Wang, X. et al. Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* **361**, eaat5691 (2018).
40. Ueda, H. R. et al. Tissue clearing and its applications in neuroscience. *Nat. Rev. Neurosci.* **21**, 61–79 (2020).
41. Yang, B. et al. Single-cell phenotyping within transparent intact tissue through whole-body clearing. *Cell* **158**, 945–958 (2014).
42. Sylwestrak, E. L., Rajasethupathy, P., Wright, M. A., Jaffe, A. & Deisseroth, K. Multiplexed intact-tissue transcriptional analysis at cellular resolution. *Cell* **164**, 792–804 (2016).
43. Casoni, F. et al. Development of the neurons controlling fertility in humans: new insights from 3D imaging and transparent fetal brains. *Development* **143**, 3969–3981 (2016).

285

Perspective

44. Belle, M. et al. Tridimensional visualization and analysis of early human development. *Cell* **169**, 161–173 (2017).
    A 3D map of first-trimester human development by tissue clearing and light-sheet imaging, providing high-resolution images of the developing cardiopulmonary, vascular, peripheral nervous, muscular and urogenital systems, as well as insights into complex processes such as skin innervation and the differential vascularization of male and female genital systems.
45. Zhao, S. et al. Cellular and molecular probing of intact human organs. *Cell* **180**, 796–812 (2020).
46. Todorov, M. I. et al. Machine learning analysis of whole mouse brain vasculature. *Nat. Methods* **17**, 442–449 (2020).
47. Kirst, C. et al. Mapping the fine-scale organization and plasticity of the brain vasculature. *Cell* **180**, 780–795 (2020).
48. Gracia, M. et al. Mechanical impact of epithelial-mesenchymal transition on epithelial morphogenesis in *Drosophila*. *Nat. Commun.* **10**, 2951 (2019).
49. Dumortier, J. G. et al. Hydraulic fracturing and active coarsening position the lumen of the mouse blastocyst. *Science* **365**, 465–468 (2019).
50. Shahbazi, M. N., Siggia, E. D. & Zernicka-Goetz, M. Self-organization of stem cells into embryos: a window on early mammalian development. *Science* **364**, 948–951 (2019).
51. Rood, J. E. et al. Toward a common coordinate framework for the human body. *Cell* **179**, 1455–1467 (2019).
52. Bonneel, N. *Optimal Transport for Computer Graphics and Temporal Coherence of Image Processing Algorithms*. PhD thesis, Sorbonne Univ. (2018).
53. Wilkinson, M. D. et al. The FAIR guiding principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).
54. Popescu, D.-M. et al. Decoding human fetal liver haematopoiesis. *Nature* **574**, 365–371 (2019).
    A detailed single-cell characterization of fetal blood and immune-cell development, revealing inferred differentiation trajectories from haematopoietic stem cells and gestation-specific potential for the differentiation of haematopoietic stem cells.
55. Cao, J. et al. A human cell atlas of fetal gene expression. *Science* **370**, eaba7721 (2020).
    One of a set of two studies focusing on integrating single-cell gene expression[55] and chromatin accessibility[118] from 15 first- and second-trimester human organs.
56. Vento-Tormo, R. et al. Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**, 347–353 (2018).
    A detailed scRNA-seq analysis of first-trimester decidua and placenta, highlighting the cell–cell interactions that take place at the maternal–fetal interface during human development using a receptor–ligand database (CellPhoneDB).
57. Suryawanshi, H. et al. A single-cell survey of the human first-trimester placenta and decidua. *Sci. Adv.* **4**, eaau4788 (2018).
58. Holloway, E. M. et al. Mapping development of the human intestinal niche at single-cell resolution. *Cell Stem Cell* **28**, 568–580 (2021).
59. Pollen, A. A. et al. Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat. Biotechnol.* **32**, 1053–1058 (2014).
60. Han, X. et al. Construction of a human cell landscape at single-cell level. *Nature* **581**, 303–309 (2020).
    A single-cell gene expression study of multiple organs during first- and second-trimester human development, with comparative analyses between human and mouse to identify conserved genetic networks.
61. Stewart, B. J. et al. Spatiotemporal immune zonation of the human kidney. *Science* **365**, 1461–1466 (2019).
62. Cui, Y. et al. Single-cell transcriptome analysis maps the developmental track of the human heart. *Cell Rep.* **26**, 1934–1950 (2019).
63. La Manno, G. et al. Molecular diversity of midbrain development in mouse, human, and stem cells. *Cell* **167**, 566–580 (2016).
64. Elmentaite, R., Ross, A., James, K. R., Ortmann, D. & Gomes, T. Single-cell sequencing of developing human gut reveals transcriptional links to childhood Crohn's disease. *Dev. Cell* **55**, 771–783 (2020).
65. Young, M. D. et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* **361**, 594–599 (2018).
    Comparative single-cell analyses of fetal, paediatric and adult kidneys and of Wilms' tumours, demonstrating the origin of Wilms' tumour as aberrant nephron development.
66. Vértesy, Á. et al. Parental haplotype-specific single-cell transcriptomics reveal incomplete epigenetic reprogramming in human female germ cells. *Nat. Commun.* **9**, 1873 (2018).
67. Li, L. et al. Single-cell RNA-seq analysis maps development of human germline cells and gonadal niche interactions. *Cell Stem Cell* **20**, 858–873 (2017).
68. Nowakowski, T. J. et al. Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358**, 1318–1323 (2017).
69. Camp, J. G. et al. Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. *Proc. Natl Acad. Sci. USA* **112**, 15672–15677 (2015).
70. Lu, Y. et al. Single-cell analysis of human retina identifies evolutionarily conserved and species-specific mechanisms controlling development. *Dev. Cell* **53**, 473–491 (2020).
71. Tiklová, K. et al. Single-cell RNA sequencing reveals midbrain dopamine neuron diversity emerging during mouse brain development. *Nat. Commun.* **10**, 581 (2019).
72. Kee, N. et al. Single-cell analysis reveals a close relationship between differentiating dopamine and subthalamic nucleus neuronal lineages. *Cell Stem Cell* **20**, 29–40 (2017).
73. Rosenberg, A. B., Roco, C. M., Muscat, R. A. & Kuchina, A. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* **360**, 176–182 (2018).
74. Carter, R. A. et al. A single-cell transcriptional atlas of the developing murine cerebellum. *Curr. Biol.* **28**, 2910–2920 (2018).
75. Huisman, C. et al. Single cell transcriptome analysis of developing arcuate nucleus neurons uncovers their key developmental regulators. *Nat. Commun.* **10**, 3696 (2019).
76. Utz, S. G. et al. Early fate defines microglia and non-parenchymal brain macrophage development. *Cell* **181**, 557–573 (2020).
77. Ginhoux, F. & Jung, S. Monocytes and macrophages: developmental pathways and tissue homeostasis. *Nat. Rev. Immunol.* **14**, 392–404 (2014).
78. Park, J.-E. et al. A cell atlas of human thymic development defines T cell repertoire formation. *Science* **367**, eaay3224 (2020).
79. Rossant, J. & Tam, P. P. L. New insights into early human development: lessons for stem cell derivation and differentiation. *Cell Stem Cell* **20**, 18–28 (2017).
80. Rayon, T. et al. Species-specific pace of development is associated with differences in protein stability. *Science* **369**, eaba7667 (2020).
81. Matsuda, M. et al. Species-specific segmentation clock periods are due to differential biochemical reaction speeds. *Science* **369**, 1450–1455 (2020).
82. Pijuan-Sala, B. et al. A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature* **566**, 490–495 (2019).
    A densely sampled time-course analysis covering mouse gastrulation and early organogenesis provides an scRNA-seq reference atlas, which is then exploited to provide insights into early blood and endothelial development through parallel analysis of mouse chimeras that lack the key regulator TAL1 (also known as SCL).
83. Wagner, D. E. et al. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**, 981–987 (2018).
84. Briggs, J. A. et al. The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science* **360**, eaar5780 (2018).
85. Cao, J. et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
86. Cusanovich, D. A. et al. The *cis*-regulatory dynamics of embryonic development at single-cell resolution. *Nature* **555**, 538–542 (2018).
87. Lescroart, F. et al. Defining the earliest step of cardiovascular lineage segregation by single-cell RNA-seq. *Science* **359**, 1177–1181 (2018).
88. Blakeley, P. et al. Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development* **142**, 3151–3165 (2015).
89. Deglincerti, A. et al. Self-organization of the in vitro attached human embryo. *Nature* **533**, 251–254 (2016).
90. Matsuda, M. et al. Recapitulating the human segmentation clock with pluripotent stem cells. *Nature* **580**, 124–129 (2020).
91. Diaz-Cuadros, M. et al. In vitro characterization of the human segmentation clock. *Nature* **580**, 113–118 (2020).
92. Koike, H. et al. Modelling human hepato-biliary-pancreatic organogenesis from the foregut-midgut boundary. *Nature* **574**, 112–116 (2019).
93. Lee, J. et al. Hair-bearing human skin generated entirely from pluripotent stem cells. *Nature* **582**, 399–404 (2020).
94. Spence, J. R. et al. Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. *Nature* **470**, 105–109 (2011).
95. Marton, R. M. & Paşca, S. P. Organoid and assembloid technologies for investigating cellular crosstalk in human brain development and disease. *Trends Cell Biol.* **30**, 133–143 (2020).
96. Lancaster, M. A. et al. Cerebral organoids model human brain development and microcephaly. *Nature* **501**, 373–379 (2013).
97. Quadrato, G. et al. Cell diversity and network dynamics in photosensitive human brain organoids. *Nature* **545**, 48–53 (2017).
98. Wimmer, R. A. et al. Human blood vessel organoids as a model of diabetic vasculopathy. *Nature* **565**, 505–510 (2019).
99. Bhaduri, A. et al. Cell stress in cortical organoids impairs molecular subtype specification. *Nature* **578**, 142–148 (2020).
100. Homsy, J. et al. De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science* **350**, 1262–1266 (2015).
101. Barnat, M. et al. Huntington's disease alters human neurodevelopment. *Science* **369**, 787–793 (2020).
102. Zhang, S.-Y. et al. Human inborn errors of immunity to infection affecting cells other than leukocytes: from the immune system to the whole organism. *Curr. Opin. Immunol.* **59**, 88–100 (2019).
103. Croft, B. et al. Human sex reversal is caused by duplication or deletion of core enhancers upstream of SOX9. *Nat. Commun.* **9**, 5319 (2018).
104. Taylor, D. M. et al. The pediatric cell atlas: defining the growth phase of human development at single-cell resolution. *Dev. Cell* **49**, 10–29 (2019).
105. Haendel, M. et al. How many rare diseases are there? *Nat. Rev. Drug Discov.* **19**, 77–78 (2020).
106. Ly, A. et al. DSCAM is a netrin receptor that collaborates with DCC in mediating turning responses to netrin-1. *Cell* **133**, 1241–1254 (2008).
107. Yamagishi, H. & Srivastava, D. Unraveling the genetic and developmental mysteries of 22q11 deletion syndrome. *Trends Mol. Med.* **9**, 383–389 (2003).
108. Biswas, A. B. & Furniss, F. Cognitive phenotype and psychiatric disorder in 22q11.2 deletion syndrome: a review. *Res. Dev. Disabil.* **53-54**, 242–257 (2016).
109. Jessa, S. et al. Stalled developmental programs at the root of pediatric brain tumors. *Nat. Genet.* **51**, 1702–1713 (2019).
110. Tirosh, I. et al. Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature* **539**, 309–313 (2016).
111. Hovestadt, V. et al. Medulloblastomics revisited: biological and clinical insights from thousands of patients. *Nat. Rev. Cancer* **20**, 42–56 (2020).
112. Phillips, H. S. et al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell* **9**, 157–173 (2006).
113. Sharma, A. et al. Onco-fetal reprogramming of endothelial cells drives immunosuppressive macrophages in hepatocellular carcinoma. *Cell* **183**, 377–394 (2020).

286

114. Reynolds, G. et al. Developmental cell programs are co-opted in inflammatory skin disease. *Science* **371**, eaba6500 (2021).
   Comparative analyses of fetal skin with healthy and diseased adult skin, revealing the co-option of developmental cell programs in two common inflammatory skin conditions (atopic dermatitis and psoriasis).
115. Sungnak, W. et al. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat. Med.* **26**, 681–687 (2020).
116. Barker, R. A., Parmar, M., Studer, L. & Takahashi, J. Human trials of stem cell-derived dopamine neurons for Parkinson's disease: dawn of a new era. *Cell Stem Cell* **21**, 569–573 (2017).
117. Takahashi, J. Preparing for first human trial of induced pluripotent stem cell-derived cells for Parkinson's disease: an interview with Jun Takahashi. *Regen. Med.* **14**, 93–95 (2019).
118. Domcke, S. et al. A human cell atlas of fetal chromatin accessibility. *Science* **370**, eaba7612 (2020).
119. Asp, M. et al. A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart. *Cell* **179**, 1647–1660 (2019).
   A temporal and 3D spatial map of the developing human heart from the first trimester by using a combination of transcriptome-wide scRNA-seq and spatial transcriptomics methods with cellular validation by in situ sequencing.
120. NHGRI & NICHD. Developmental Genotype-Tissue Expression (dGTEx). https://www.genome.gov/Funded-Programs-Projects/Developmental-Genotype-Tissue-Expression (2020).
121. Yan, L. et al. Single-cell RNA-seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1131–1139 (2013).
   A comprehensive scRNA-seq analysis of human oocytes to blastocyst-stage embryos that has widely been used to investigate lineage-associated gene expression and as a comparative analysis to human pluripotent stem cell lines.
122. Petropoulos, S. et al. Single-cell RNA-seq reveals lineage and X chromosome dynamics in human preimplantation embryos. *Cell* **165**, 1012–1026 (2016).
123. Tyser, R. C. V. et al. A spatially resolved single cell atlas of human gastrulation. Preprint at https://doi.org/10.1101/2020.07.21.213512 (2020).
124. Elmentaite, R. et al. Cells of the human intestinal tract mapped across space and time. *Nature*, https://doi.org/10.1038/s41586-021-03852-1 (2021).

Human Cell Atlas Developmental Biological Network

Pascal Barbry[22], Omer Bayraktar[2], Sam Behjati[2], Andreas Bosio[33], Bruno Canque[34], Frédéric Chalmel[35], Yorick Gitton[13], Deborah Henderson[1], Anne Jorgensen[36], Steven Lisgo[1], Jinyue Liu[37], Emma Lundberg[22], Jean-Léon Maître[38], Séverine Mazaud-Guittot[35], Elizabeth Robertson[39], Antoine Rolland[35], Raphael Scharfmann[40], Michèle Souyri[41], Erik Sundström[42] & Matthias Zilbauer[42,43]

[32]Institut de Pharmacologie Moléculaire et Cellulaire, UMR7275, CNRS/UNS, Université Côte d'Azur, Valbonne, France. [33]Miltenyi Biotec B.V. & Co. KG, Bergisch Gladbach, Germany. [34]Laboratoire Développement du Système Immunitaire, Ecole Pratique des Hautes Etudes, INSERM U976, Institut de Recherche Saint Louis, Centre Hayem, Hôpital Saint Louis 1, Paris, France. [35]Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail) UMR_S 1085, Université Rennes, Rennes, France. [36]University Department of Growth and Reproduction, EDMaRC, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark. [37]Genome Institute of Singapore, Singapore, Singapore. [38]Institut Curie, Paris, France. [39]Sir William Dunn School of Pathology, University of Oxford, Oxford, UK. [40]U1016 INSERM Institut Cochin, Groupe Hospitalier Cochin Port Royal, Paris, France. [41]INSERM UMRS 1131, Institut de Recherche Saint Louis, Paris, France. [42]Division of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden. [43]University of Cambridge, Cambridge, UK.

287

Appendix E. Co-author of Nature Research Article: Blood and immune development in human fetal bone marrow and Down syndrome

Article

Blood and immune development in human fetal bone marrow and Down syndrome

Laura Jardine[1,2,21], Simone Webb[1,21], Issac Goh[1], Mariana Quiroga Londoño[3], Gary Reynolds[1], Michael Mather[1], Bayanne Olabi[1], Emily Stephenson[1], Rachel A. Botting[1], Dave Horsfall[1], Justin Engelbert[1], Daniel Maunder[1], Nicole Mende[3], Caitlin Murnane[4], Emma Dann[5], Jim McGrath[1], Hamish King[6], Iwo Kucinski[3], Rachel Queen[1], Christopher D. Carey[7], Caroline Shrubsole[2], Elizabeth Poyner[1], Meghan Acres[1], Claire Jones[8], Thomas Ness[8], Rowen Coulthard[8], Natalina Elliott[4], Sorcha O'Byrne[4], Myriam L. R. Haltalli[3], John E. Lawrence[5], Steven Lisgo[1], Petra Balogh[5], Kerstin B. Meyer[5], Elena Prigmore[5], Kirsty Ambridge[5], Mika Sarkin Jain[5], Mirjana Efremova[9], Keir Pickard[2], Thomas Creasey[2,10], Jaume Bacardit[11], Deborah Henderson[1], Jonathan Coxhead[1], Andrew Filby[1], Rafiqul Hussain[1], David Dixon[1], David McDonald[1], Dorin-Mirel Popescu[1], Monika S. Kowalczyk[12], Bo Li[12], Orr Ashenberg[12,13], Marcin Tabaka[12], Danielle Dionne[12], Timothy L. Tickle[12,14], Michal Slyper[12], Orit Rozenblatt-Rosen[12], Aviv Regev[12], Sam Behjati[5,15], Elisa Laurenti[3], Nicola K. Wilson[3,22], Anindita Roy[4,16,17,18,22], Berthold Göttgens[3,22], Irene Roberts[4,16,17,18,22], Sarah A. Teichmann[5,19,22] & Muzlifah Haniffa[1,5,20,22 ✉]

Haematopoiesis in the bone marrow (BM) maintains blood and immune cell production throughout postnatal life. Haematopoiesis first emerges in human BM at 11–12 weeks after conception[1,2], yet almost nothing is known about how fetal BM (FBM) evolves to meet the highly specialized needs of the fetus and newborn. Here we detail the development of FBM, including stroma, using multi-omic assessment of mRNA and multiplexed protein epitope expression. We find that the full blood and immune cell repertoire is established in FBM in a short time window of 6–7 weeks early in the second trimester. FBM promotes rapid and extensive diversification of myeloid cells, with granulocytes, eosinophils and dendritic cell subsets emerging for the first time. The substantial expansion of B lymphocytes in FBM contrasts with fetal liver at the same gestational age. Haematopoietic progenitors from fetal liver, FBM and cord blood exhibit transcriptional and functional differences that contribute to tissue-specific identity and cellular diversification. Endothelial cell types form distinct vascular structures that we show are regionally compartmentalized within FBM. Finally, we reveal selective disruption of B lymphocyte, erythroid and myeloid development owing to a cell-intrinsic differentiation bias as well as extrinsic regulation through an altered microenvironment in Down syndrome (trisomy 21).

Human BM is established as the site of lifelong blood and immune cell production from 11–12 post conception weeks (PCW)[1,2]. By this time, fetal liver has initiated an immune repertoire, with further differentiation supported by the spleen and thymus[3,4]. The priorities of fetal haematopoiesis are to generate erythrocytes for oxygen transport, platelets for haemostasis, macrophages for tissue remodelling and an immune system that is poised to respond to insult without risking tissue damage. Longer-term haematopoiesis depends on a finite pool of haematopoietic stem cells (HSCs) that are supported by their niche. Perturbations of haematopoiesis in utero can have far-reaching implications, including KMT2A fusions or Down-syndrome-associated mutations in GATA1 that lead to an increased risk of childhood leukaemia[5,6]. To our knowledge, no systematic examination of FBM development or human bone marrow stroma at any time point has so far been achieved.

In this study, we use single-cell multi-omics to investigate the composition of disomic and trisomy-21 human FBM as haematopoiesis emerges and develops during the early second trimester. We perform multi-omics profiling of fetal liver and cord blood cells to compare tissue-specific differentiation landscapes. We validate: (1) newly emerging cell states in FBM by fluorescence-activated cell sorting (FACS)-based prospective isolation for single-cell RNA sequencing (scRNA-seq) and morphology assessment; (2) the regional distribution of subsets of FBM endothelial cells by multiplex immunofluorescence imaging; and (3) the differentiation potential of HSCs using single-cell clonogenic differentiation assays. Drawing on existing scRNA-seq data from yolk sac, fetal liver, cord blood and adult BM (ABM), we show in humans how a complex multilineage blood and immune system is assembled in FBM within a matter of weeks.

A list of affiliations appears at the end of the paper.

288
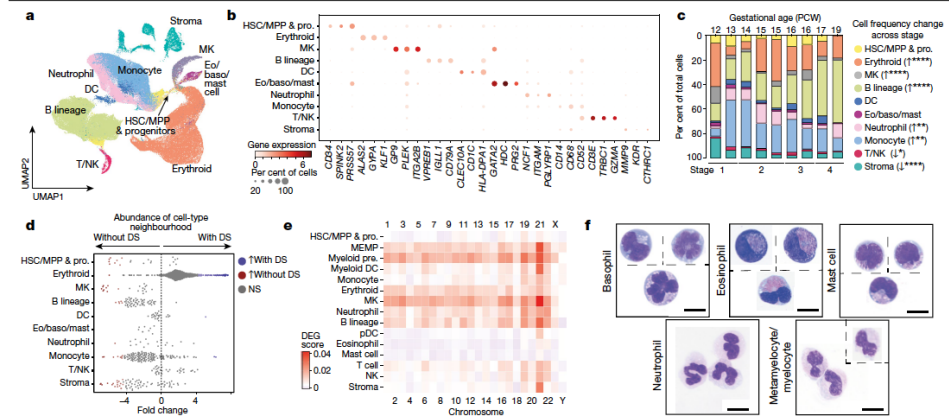
Fig. 1 | A single-cell atlas of human FBM. a, Uniform manifold approximation and projection (UMAP) of FBM scRNA-seq data ($n = 9$, $k = 103,228$, 12–19 PCW) by broad categories (Supplementary Table 7). Baso, basophil; eo, eosinophil; MK, megakaryocyte. **b**, Gene expression dot plot of cell-state-defining genes for broad categories in FBM scRNA-seq data. Dot colour indicates log-transformed, normalized and scaled gene expression value. Dot size indicates the percentage of cells in each category expressing a given gene. **c**, Frequency of broad cell categories in FBM scRNA-seq data. $n = 9$ biologically independent samples are grouped into four developmental stages to facilitate statistical comparison over gestational stage. P values from a quasibinomial regression model (subject to one-sided ANOVA; with correction for sort gates; computed at 95% confidence intervals and adjusted for multiple testing using Bonferroni correction) are shown in parentheses; *$P < 0.05$, **$P < 0.01$, ****$P < 0.0001$ (Supplementary Tables 18,19). **d**, Beeswarm plot of the log-transformed fold change in abundance between cells in equivalent broad categories in FBM scRNA-seq datasets from fetuses with Down syndrome (DS)

($n = 4$, $k = 16,743$) and age-matched fetuses without Down syndrome ($n = 2$, $k = 9,717$) from biologically independent samples (Supplementary Tables 7, 20). Coloured dots indicate a significant difference in abundance (P value adjusted for multiple testing with a false discovery rate of greater than 10%) estimated with a two-sided quasi-likelihood test (null hypothesis: no difference in abundance between the two conditions). NS, not significant. **e**, Heat map visualizing the number of differentially expressed genes (DEGs) between equivalent cell states in FBM scRNA-seq datasets from fetuses with Down syndrome (DS) versus fetuses without Down syndrome. Data are scRNA-seq per chromosome, including correction for number of genes per chromosome (two-sided Wilcoxon rank-sum statistical test with Benjamini–Hochberg procedure for multiple testing correction; Supplementary Table 21). **f**, Cytospin images of FBM eosinophils, basophils, mast cells and neutrophils ($n = 2$ biologically independent samples; both 17 PCW, performed as two independent experiments) sorted according to the gating strategy in Extended Data Fig. 1 and stained with Giemsa. Images were concatenated as shown by the dotted lines. Scale bars, 10 μm.

A single-cell atlas of human FBM

We generated mRNA, T cell receptor and B cell receptor (TCR and BCR, respectively) and CITE-seq data from single FBM cells (enriched for CD45+ or CD45− cells) and CITE-seq data from CD34+-selected cells after mechanical disruption of fetal femur ($n = 9$; 12–19 PCW). We generated single-cell mRNA profiles of FBM cells from fetuses with Down syndrome ($n = 4$) and CITE-seq data from CD34+ fetal liver and cord blood cells. Reference scRNA-seq datasets (yolk sac and fetal liver[4]; cord blood and ABM (https://data.humancellatlas.org/)) were used to investigate haematopoietic development (Extended Data Fig. 1, Supplementary Tables 1–5). Our data are available for exploration (https://fbm.cellatlas.io/).

From 115,993 FBM scRNA-seq cells, 103,228 passed quality control, revealing 64 transcriptionally distinct cell states that were manually grouped into 10 compartments (Fig. 1a, b, Extended Data Fig. 1, Supplementary Tables 1, 6–10). We constructed a continuous decision tree for supervised learning of cell-state-discriminative protein combinations in FBM CITE-seq data (Extended Data Fig. 2, Supplementary Tables 11–14). Between 12 and 19 PCW, the ratio of blood and immune cells to stromal cells expanded from 5:1 to 18:1. B cell lymphopoiesis expanded, but the total proportions of myeloid cells remained consistent (Fig. 1c). Compared with age-matched disomic BM (12–13 PCW), megakaryocyte and B lineages were diminished in Down syndrome (specifically pre pro-B and immature B cells, consistent with fetal liver data in Down syndrome[7]), whereas erythroid cells were significantly more abundant

(mid–late erythroid cells) and more enriched in cell-cycle genes (Fig. 1d, Extended Data Fig. 7). FBM from fetuses with Down syndrome exhibited genome-wide transcriptional differences in addition to increased expression of chromosome-21 genes (Fig. 1e). Megakaryocyte–erythroid–mast cell progenitor (MEMP), megakaryocyte and B-lineage cells from fetuses with Down syndrome overexpressed chromosome 21 transcription factors that have documented roles in haematopoiesis, including *U2AF1* (MEMP), *U2AF1* and *ETS2* (megakaryocyte) and *ETS2* (B lineage) (Extended Data Fig. 7).

Granulocytes first emerge in FBM

Neutrophils, eosinophils and basophils were not detected in age-matched fetal liver[4]. We validated their presence in FBM by morphology and prospective FACS isolation for scRNA-seq (Smart-seq2) (Fig. 1f, Extended Data Fig. 1, Supplementary Table 15). Compared with yolk sac and fetal liver, FBM myeloid cells were significantly expanded in FBM (Fig. 2a). Detailed clustering revealed 18 monocyte, dendritic cell (DC), neutrophil and macrophage states from committed precursors to terminally differentiated cells (Extended Data Fig. 3). In force-directed graph (FDG) embedding, monocyte and neutrophil signatures diverged at the granulocyte and monocyte progenitor (GMP) stage, consistent with mouse data[8] (Extended Data Fig. 3). FBM GMPs expressed higher levels of *CEBPA* (neutrophil specification), relative to *SPI1* (monocyte specification)[9], than did fetal liver GMPs (Extended Data Fig. 3). Across neutrophil differentiation (inferred by Monocle 3), genes associated with leukaemia-risk congenital neutropenias (*SBDS*, *HAX1* and *G6PC3*)
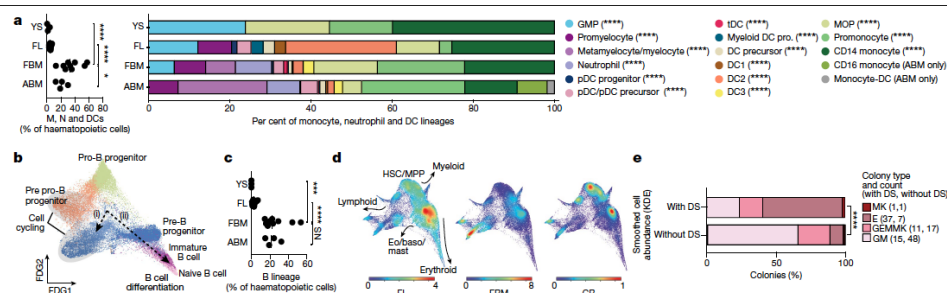
Fig. 2 | Myeloid diversification, B-lineage expansion and tissue-specific properties of HSC/MPPs. a, Left, frequency of monocytes (M), neutrophils (N) and DCs in yolk sac (YS) ($n = 3$), fetal liver ($n = 14$), FBM ($n = 9$) and ABM ($n = 4$) scRNA-seq datasets (n refers to biologically independent samples; line, median; *$P < 0.05$,****$P < 0.0001$ from one-way ANOVA with Tukey's multiple comparison). Right, mean proportions of cell states within monocyte, neutrophil and DC lineages for above datasets (****$P < 0.0001$ from quasibinomial regression model subject to one-sided ANOVA with 95% confidence intervals and Bonferroni correction; Supplementary Tables 19, 22). MOP, monocyte progenitor. **b**, FDG of B-lineage cell states ($k = 28,583$) in FBM scRNA-seq data. Grey ellipses highlight cycling cells. Dashed arrows denote (i) cycling pre-B cell and (ii) B cell differentiation branches. **c**, Frequency of B-lineage cells in yolk sac, fetal liver, FBM and ABM scRNA-seq datasets

(replicates, measures of centre and statistical tests as per **a**; ***$P < 0.001$,****$P < 0.0001$; Supplementary Tables 19, 23). **d**, FDG visualization of CD34$^+$ fetal liver (left), FBM (centre) and cord blood (right) cells on a CITE-seq gene expression landscape (Extended Data Fig. 5b). Colour indicates relative cell abundance per tissue by kernel density estimation (KDE). **e**, Colonies produced by HSC/MPPs (grown on methylcellulose; 17–21 PCW) from fetuses with Down syndrome ($n = 2$, $k = 246$, $k^* = 64$) and age-matched fetuses without Down syndrome ($n = 3$, $k = 365$, $k^* = 73$) (n refers to biologically independent samples, k indicates plated cells and k^* indicates wells producing colonies). The bar graph shows proportions by colony type, with colony numbers provided in parentheses. Statistical differences between fetuses with Down syndrome and the expected distribution based on fetuses without Down syndrome were tested by chi-squared test with ****$P < 10^{-15}$, two-sided.

were expressed in early progenitors, whereas those without recognized leukaemia risk (*AP3B1* and *CXCR4*) were expressed in terminal differentiation stages (Extended Data Fig. 3).

DC subsets diversify in FBM

Plasmacytoid (pDC), transitional DC (tDC) and DC3 emerged during FBM haematopoiesis; however, non-classical CD16$^+$ monocytes and monocyte-DCs were not detected (Fig. 2a). Signatures of DC1 and pDC—but not DC2 and DC3[10]—were conserved between fetal and adult peripheral blood subsets. FDG embedding revealed that the tDC transcriptional state was intermediate between DC2 and pDC, as in adult blood[10]. iRegulon analysis showed that the transcription factors that drive FBM pDC and tDC differentiation are shared (Extended Data Fig. 3).

Mature natural killer and T cells in FBM

We identified natural killer (NK) cells, natural killer T (NKT)-like cells and innate lymphoid cell (ILC) precursors in FBM (Extended Data Fig. 3). FBM NK cells were enriched for NK cytotoxicity genes, relative to yolk sac and fetal liver (Extended Data Fig. 3). In contrast to ABM, FBM contained few T lymphocytes (naive only; Fig. 1a. Supplementary Table 5). As thymic lymphopoiesis is established before FBM is colonized[3], these were single-positive CD4, CD8 and T regulatory cells expressing productive *TRA* and *TRB* (Extended Data Fig. 4).

Expanded B cell lymphopoiesis in FBM

We observed two bursts of proliferative activity during B cell development (in pre pro-B and pre-B progenitors; Fig. 2b). Heavy chain rearrangement was productive from the pre-B progenitor stage and heavy and light chain from the immature B cell stage. The emerging B cell repertoire was diverse, with a small number of shared clonotypes detected (Extended Data Fig. 4). The frequency of B-lineage cells was 10-fold higher in FBM than in fetal liver, and markedly skewed towards the earlier cell states, compared to ABM (Fig. 2c).

Differentiation trajectories predicted by Monocle 3 branched at the pre-B progenitor stage into 'cycling' and 'B cell differentiation' paths.

Apoptosis genes were most enriched in the non-cycling pro-B and pre-B cell stages, in keeping with the programmed death of cells that fail to undergo successful heavy chain recombination and integration into the pre-B receptor (Extended Data Fig. 4).

Small deletions and translocations in a limited set of genes causing B-cell acute lymphoblastic leukaemia (B-ALL)[11], which commonly presents in infancy and childhood, were highly expressed in early B-lineage progenitors in FBM, but expression was less marked in equivalent ABM stages (Extended Data Fig. 4).

Tissue-specific properties of HSC/MPPs

CD34$^+$-selected CITE-seq data from fetal liver, FBM and cord blood allowed us to examine the unique features of FBM progenitors (Extended Data Fig. 5, Supplementary Tables 1, 16, 17). Erythroid precursors dominated fetal liver, whereas lymphoid precursors were most prevalent in FBM. Cord blood was enriched in HSCs and multipotent progenitors (hereafter, HSC/MPPs), common lymphoid progenitors and the earliest erythroid precursors (Fig. 2d). Cell-cycle gene enrichment was lower in cord blood than in fetal tissue HSC/MPPs (consistent with previous work[12]), whereas FBM and fetal liver HSC/MPPs showed similar enrichment of cycling genes (Extended Data Fig. 5). Differentially expressed proteins in HSC/MPPs revealed tissue-specific patterns of adhesion molecules (CD49a and CD146 in fetal liver; integrin β7 in BM), growth factor receptors (EGFR in fetal liver) and molecules associated with HSC activation and recirculation (CD69 and CD31 in cord blood) (Extended Data Fig. 5).

We used direction of transition (DoT) analysis to investigate tissue-specific HSC/MPP differentiation bias. Fetal liver was biased towards erythroid fate and away from lympho-myeloid fate, whereas FBM was biased towards neutrophil and B-lineage fate (Extended Data Fig. 6). We assessed differentiation potential in vitro via single-cell clonal cultures of paired fetal liver and FBM HSC/MPPs. Myeloid colonies arose frequently from both sources of HSC/MPPs, but myeloid-restricted colonies were typical of FBM, supporting the myeloid bias of

Article

FBM HSC/MPPs from DoT analysis and myeloid cell diversity in FBM (Fig. 2a, Extended Data Fig. 6).

Erythroid bias of HSC/MPPs in Down syndrome

FBM HSC/MPPs from fetuses with Down syndrome produced significantly more erythroid colonies and fewer myeloid colonies on methylcellulose compared to HSC/MPPs from age-matched fetuses without Down syndrome (Fig. 2e, Extended Data Fig. 7). Across erythroid differentiation pseudotime in Down syndrome (Monocle 3), an increase in the expression of cell-cycle genes (CCND3 and MKI67) and elevated, sustained expression of the glycolysis gene PKLR (Extended Data Fig. 7) suggested that rapid proliferation and metabolic adaptations compound the erythroid dominance.

Transcription factors with well-defined roles in early haematopoietic programming (SPI1 and FLI1)[13,14] were expressed at lower levels in HSC/MPPs and MEMPs from fetuses with Down syndrome than in those from fetuses without Down syndrome, and PySCENIC inferred downregulation of corresponding regulons. Megakaryocytes from fetuses with Down syndrome expressed lower levels of FLI1, a driver of megakaryocyte differentiation[15], in keeping with recent data showing FLI1 promoter silencing in Down syndrome[16]. Regulons for chromosome-21-encoded GABPA, which is implicated in the differentiation and maintenance of HSC/MPPs[17], were overrepresented in Down syndrome (Extended Data Fig. 7).

Most myeloid lineages in Down syndrome overexpressed TNF (consistent with the higher levels of circulating TNF in Down syndrome[18]) and TNF signalling pathway genes were overrepresented in myeloid, erythroid, NK and stromal cells from fetuses with Down syndrome (Extended Data Fig. 7). CellPhoneDB analysis predicted statistically significant receptor–ligand interactions involving proteins of the TNF family between FBM HSC/MPPs from fetuses with Down syndrome and mature myeloid cells (Extended Data Fig. 7).

Stromal cell heterogeneity in FBM

We identified 19 stromal cell states in FBM, which closely correlate with postnatal mouse BM stroma[19] (Extended Data Fig. 8). Two dominant endothelial cell clusters expressed KDR (VEGFR2) but with differential expression of CD34. One cluster expressed the characteristic sinusoidal endothelial cell genes CTSL, STAB2 and SELENOP (also known as SEPP1) and the other expressed high levels of VIM and CD34, associated with non-sinusoidal endothelial cells in mice[19]. The non-sinusoidal 'tip endothelial cells' expressed canonical markers for cells at the tips of growing vascular structures—PDGFB, UNC5B and DLL4 (ref. [20]) (Fig. 3a, Extended Data Fig. 8).

Regional partitioning of endothelial cell subsets

Using multiplex immunofluorescence microscopy, we identified CD34hiVEGFR2lo branching vessels adjacent to the epiphyseal cartilage (metaphysis). VEGFR2hiCD34lo cells formed convoluted structures in more-distal regions (diaphysis). Thicker-walled CD34hi vessels, co-localizing with CXCL12+ cells, were present in both metaphysis and diaphysis (Fig. 3b, Extended Data Fig. 9). This regional compartmentalization was reminiscent of the L-type (sinusoids) and H-type (metaphyseal arterioles) vessels that have been described in mouse BM, with distinct roles in supporting haematopoiesis[21,22]. Mouse sinusoidal and arteriolar endothelial cell genes were enriched in our FBM sinusoid and tip endothelial cells, respectively (Extended Data Fig. 8). The frequency of CD34+CD117+ HSC/MPPs and progenitors relative to cellular density was similar in metaphyseal and diaphyseal areas (P = 0.431, Wald test) (Extended Data Fig. 9).

Endothelial cells in tissues and in Down syndrome

FBM sinusoidal endothelial cells had significantly higher expression of SELE, VCAM1 and ICAM2 and concordant higher surface protein expression than analogous fetal liver sinusoidal endothelial cells. FBM
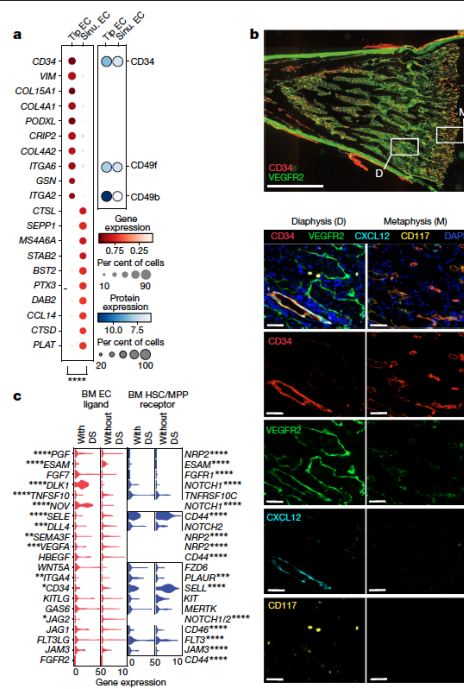
Fig. 3 | Stromal cell heterogeneity in FBM. a, Left, dot plot of the top 10 genes (by P value) differentially expressed between tip endothelial cells (ECs) (k = 362) and sinusoidal endothelial cells (k = 550) in FBM scRNA-seq data (n = 9) (Methods and interpretation as in Fig. 1b; two-sided Wilcoxon rank-sum test with Benjamini–Hochberg correction; ****P < 0.0001, Supplementary Table 24). Right, dot plot of equivalent protein expression (where antibody present in CITE-seq panel) on tip and sinusoidal endothelial cells in the FBM (total) CITE-seq dataset (n = 3). Dot colour indicates DSB-normalized protein expression. Dot size indicates the percentage of cells in each category expressing a given protein. **b,** Top, longitudinal section of fetal femur with multiplex immunofluorescence staining, showing CD34 (red) and VEGFR2 (green) channels to demonstrate regional differences in BM vasculature between metaphysis (M) and diaphysis (D). Scale bar, 2 mm. White boxes mark regions of interest (ROIs) shown below. Bottom, ROIs with all channels and single channels to demonstrate patterns of co-expression in CD34hiVEGFRlo metaphyseal vessels, VEGFR2hiCD34hi diaphyseal sinusoids and CD34hiCXCL12-associated arterioles. Scale bars, 50 μm. Representative images from n = 4 biologically independent FBM samples (14–15 PCW), with staining performed in two independent experiments. **c,** Violin plots of gene expression in HSC/MPPs and pooled endothelial cells from FBM scRNA-seq datasets from fetuses with Down syndrome (n = 4; k = 105 HSC/MPPs; k = 111 endothelial cells) and fetuses without Down syndrome (n = 9; k = 92 HSC/MPPs; k = 938 endothelial cells). Genes shown have a significant receptor–ligand interaction in FBM without Down syndrome predicted by CellPhoneDB analysis (detailed in Extended Data Fig. 10a, b). Significance in expression difference between fetuses with Down syndrome and fetuses without Down syndrome calculated by two-sided Wilcoxon rank-sum test with Benjamini–Hochberg correction; *P < 0.05,**P < 0.01,***P < 0.001,****P < 0.0001 (Supplementary Table 25).

291

sinusoidal endothelial cells expressed more *THBS1*, which may facilitate HSC/MPP retention[23] and matrix metalloproteinases, which are associated with mature cell egress[24]. FBM sinusoidal endothelial cells also expressed more *CCL14*, which is implicated in myeloid progenitor proliferation[25] (Extended Data Fig. 8).

CellphoneDB predicted statistically significant haematopoiesis-supportive interactions between FBM HSC/MPPs and stromal cells including FLT3–FLT3L and KIT–KITLG (confirmed at the protein level). CellPhoneDB analysis also predicted that HSC/MPPs signal to the tip (capillary metaphyseal) and proliferating endothelial cells and to the osteochondral precursors through ANGPT2, DLK1, EFNA1 and FGF7 (Extended Data Fig. 10).

The expression of NOTCH ligands *NOV* (*CCN3*) and *DLK1*—predicted by CellPhoneDB to mediate endothelial cell and HSC/MPP interactions (Extended Data Fig. 10)—was significantly higher in endothelium from fetuses with Down syndrome than in endothelium from fetuses without Down syndrome, and the expression of *NOTCH1* was increased in HSC/MPPs in Down syndrome (Fig. 3c). NOTCH signalling has a critical role in the emergence of HSC/MPPs as well as fetal HSC maintenance and response to pro-inflammatory signals, including TNF[26]. Probing for inflammatory programs in stroma from fetuses with Down syndrome, we found activation of multiple inflammatory pathways, including TNF pathways, in macrophages and osteoclasts from fetuses with Down syndrome versus those without Down syndrome (Extended Data Fig. 7). Type I interferon, interferon-γ (IFNγ) and other inflammatory cytokine (interleukin (IL)-1, IL-6, IL-7 and IL-12) response pathways were overexpressed in endothelial cells and osteochondral cells in Down syndrome (Extended Data Fig. 8). Our collective findings reveal an altered stromal environment in Down syndrome.

Discussion

Survival of the fetus depends on the successful initiation of haematopoiesis in several organs across gestation. We reveal the complete establishment of haematopoiesis in the FBM within the first few weeks of the second trimester and identify the BM as a key site of neutrophil emergence, myeloid diversification and B lymphoid selection. We identify a unique intrinsic molecular profile of FBM HSC/MPPs, and an intrinsic bias of BM stem and progenitor cells in Down syndrome that is underpinned by genome-wide transcriptional changes. A better understanding of human developmental haematopoiesis has the potential to inform regenerative and transplantation therapies; for example, through co-opting developmental programs to accelerate the reconstitution of haematopoietic stem cell transplants, and manipulating the lineage bias of differentiating progenitors to address specific deficiencies or for cellular therapy. For such endeavours to be successful, an initial phase of discovery science is critical. It is in this context that our study provides a comprehensive analysis of human FBM haematopoiesis to address a major previous knowledge gap.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-03929-x.

1. O'Byrne, S. et al. Discovery of a CD10-negative B-progenitor in human fetal life identifies unique ontogeny-related developmental programs. *Blood* **134**, 1059–1071 (2019).
2. Charbord, P., Tavian, M., Humeau, L. & Péault, B. Early ontogeny of the human marrow from long bones: an immunohistochemical study of hematopoiesis and its microenvironment. *Blood* **87**, 4109–4119 (1996).
3. Park, J.-E. et al. A cell atlas of human thymic development defines T cell repertoire formation. *Science* **367**, eaay3224 (2020).
4. Popescu, D.-M. et al. Decoding human fetal liver haematopoiesis. *Nature* **574**, 365–371 (2019).
5. Wiemels, J. L. et al. Prenatal origin of acute lymphoblastic leukaemia in children. *Lancet* **354**, 1499–1503 (1999).
6. Muntean, A. G., Ge, Y., Taub, J. W. & Crispino, J. D. Transcription factor GATA-1 and Down syndrome leukemogenesis. *Leuk. Lymphoma* **47**, 986–997 (2006).
7. Roy, A. et al. Perturbation of fetal liver hematopoietic stem and progenitor cell development by trisomy 21. *Proc. Natl Acad. Sci. USA* **109**, 17579–17584 (2012).
8. Olsson, A. et al. Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Exp. Hematol.* **44**, S24 (2016).
9. Dahl, R. et al. Regulation of macrophage and neutrophil cell fates by the PU.1:C/EBPα ratio and granulocyte colony-stimulating factor. *Nat. Immunol.* **4**, 1029–1036 (2003).
10. Villani, A.-C. et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017).
11. Mullighan, C. G. et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758–764 (2007).
12. Summers, Y. J., Heyworth, C. M., de Wynter, E. A., Chang, J. & Testa, N. G. Cord blood G_0 CD34$^+$ cells have a thousand-fold higher capacity for generating progenitors in vitro than G_1 CD34$^+$ cells. *Stem Cells* **19**, 505–513 (2001).
13. Pimanda, J. E. et al. Gata2, Fli1, and Scl form a recursively wired gene-regulatory circuit during early hematopoietic development. *Proc. Natl Acad. Sci. USA* **104**, 17692–17697 (2007).
14. Iwasaki, H. et al. Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood* **106**, 1590–1600 (2005).
15. Palii, C. G. et al. Single-cell proteomics reveal that quantitative changes in co-expressed lineage-specific transcription factors determine cell fate. *Cell Stem Cell* **24**, 812–820 (2019).
16. Muskens, I. S. et al. The genome-wide impact of trisomy 21 on DNA methylation and its implications for hematopoiesis. *Nat. Commun.* **12**, 821 (2021).
17. Yu, S. et al. GABP controls a critical transcription regulatory module that is essential for maintenance and differentiation of hematopoietic stem/progenitor cells. *Blood* **117**, 2166–2178 (2011).
18. Sullivan, K. D. et al. Trisomy 21 causes changes in the circulating proteome indicative of chronic autoinflammation. *Sci. Rep.* **7**, 14818 (2017).
19. Baccin, C. et al. Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. *Nat. Cell Biol.* **22**, 38–48 (2020).
20. Suchting, S. et al. The Notch ligand Delta-like 4 negatively regulates endothelial tip cell formation and vessel branching. *Proc. Natl Acad. Sci. USA* **104**, 3225–3230 (2007).
21. Kusumbe, A. P., Ramasamy, S. K. & Adams, R. H. Coupling of angiogenesis and osteogenesis by a specific vessel subtype in bone. *Nature* **507**, 323–328 (2014).
22. Itkin, T. et al. Distinct bone marrow blood vessels differentially regulate haematopoiesis. *Nature* **532**, 323–328 (2016).
23. Long, M. W., Briddell, R., Walter, A. W., Bruno, E. & Hoffman, R. Human hematopoietic stem cell adherence to cytokines and matrix molecules. *J. Clin. Invest.* **90**, 251–255 (1992).
24. Lane, W. J. et al. Stromal-derived factor 1–induced megakaryocyte migration and platelet production is dependent on matrix metalloproteinases. *Blood* **96**, 4152–4159 (2000).
25. Schulz-Knappe, P. et al. HCC-1, a novel chemokine from human plasma. *J. Exp. Med.* **183**, 295–299 (1996).
26. Butko, E., Pouget, C. & Traver, D. Complex regulation of HSC emergence by the Notch signaling pathway. *Dev. Biol.* **409**, 129–138 (2016).
27. Mulè, M. P., Martins, A. J. & Tsang, J. S. Normalizing and denoising protein expression data from droplet-based single cell profiling. Preprint at https://doi.org/10.1101/2020.02.24.96360 (2021).

[1]Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK. [2]Haematology Department, Freeman Hospital, Newcastle-upon-Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [3]Department of Haematology, Wellcome–MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK. [4]Department of Paediatrics, University of Oxford, Oxford, UK. [5]Wellcome Sanger Institute, Hinxton, UK. [6]Centre for Immunobiology, Blizard Institute, Queen Mary University of London, London, UK. [7]Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, UK. [8]NovoPath, Department of Pathology, Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [9]Barts Cancer Institute, Queen Mary University of London, London, UK. [10]Wolfson Childhood Cancer Centre, Northern Institute for Cancer Research, Newcastle University, Newcastle upon Tyne, UK. [11]School of Computing, Newcastle University, Newcastle upon Tyne, UK. [12]Klarman Cell Observatory, Broad Institute of Harvard and MIT, Cambridge, MA, USA. [13]Genentech, South San Francisco, CA, USA. [14]Data Sciences Platform, Broad Institute of Harvard and MIT, Cambridge, MA, USA. [15]Department of Paediatrics, University of Cambridge, Cambridge, UK. [16]MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, University of Oxford, Oxford, UK. [17]Department of Paediatrics, Children's Hospital, John Radcliffe Hospital, and MRC WIMM, University of Oxford, Oxford, UK. [18]BRC Blood Theme, National Institute for Health Research (NIHR) Oxford Biomedical Research Centre, Oxford, UK. [19]Theory of Condensed Matter Group, Cavendish Laboratory, University of Cambridge, Cambridge, UK. [20]Department of Dermatology and NIHR Newcastle Biomedical Research Centre, Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK. [21]These authors contributed equally: Laura Jardine, Simone Webb. [22]These authors jointly supervised this work: Nicola K. Wilson, Anindita Roy, Berthold Göttgens, Irene Roberts, Sarah A. Teichmann, Muzlifah Haniffa. ✉e-mail: m.a.haniffa@ncl.ac.uk

292

Appendix F. Co-Principal Investigator of multi-centre research study funded by the Chan Zuckerberg Initiative for $1.7 million: Single-cell multi-omic and spatial cell atlas of pediatric skin.

←   BACK TO PROJECT LIST

PROJECT

Single-Cell Multi-omic and Spatial Cell Atlas of Pediatric Skin

AWARD   Pediatric Networks

PROJECT SUMMARY

INVESTIGATORS

Project Summary

Healthy human skin is contingent on physiological processes established during early life. However, the current limited knowledge of pediatric skin biology restricts understanding of disease pathogenesis and the innovation of therapeutic strategies.

This project aims to use high-throughput single-cell and spatial methods to comprehensively define key milestones in human skin maturation, including adaptation to a non-aquatic environment in neonatal life, pre- and post-adrenarche, and puberty, when recognized changes occur in health and disease. Additionally, the proposed study aims to contextualize pediatric skin data with adult healthy skin and inflammatory skin disease datasets previously generated by members of this research team.

A comprehensive reference of skin biology across human lifespan will empower the wider scientific community to gain insights into processes and time points involved in maturation, aging and disease pathogenesis. By learning more about how skin develops over time, this team hopes to provide insights relevant to pediatric inflammatory and infectious skin conditions, genetic skin diseases, and skin disorders specific to historically marginalized communities. Overall, this study will establish an important, accessible repository of pediatric skin samples, which will be shared freely to enhance skin research and clinical translation.

Investigators

Co-Principal Investigators

Muzlifah Haniffa, MD, PhD
WELLCOME SANGER INSTITUTE
UNITED KINGDOM
ORCID
0000-0002-3927-2084

Dave Horsfall, MSc
NEWCASTLE UNIVERSITY
UNITED KINGDOM
ORCID
0000-0002-8086-812X

Edel O'Toole, MB, BCh, PhD, FRCP
QUEEN MARY UNIVERSITY OF LONDON
UNITED KINGDOM
ORCID
0000-0002-4084-4836

Bayanne Olabi, MBChB, MRCP, MSc
NEWCASTLE UNIVERSITY
UNITED KINGDOM
ORCID
0000-0002-4786-7838

Neil Rajan, MBBS, PhD
NEWCASTLE UNIVERSITY
UNITED KINGDOM
ORCID
0000-0002-5850-5680

Inês Sequeira, MSc, PhD
QUEEN MARY UNIVERSITY LONDON
UNITED KINGDOM
ORCID
0000-0003-3342-665X

Mirjana Efremova, MSc, PhD
QUEEN MARY UNIVERSITY OF LONDON
UNITED KINGDOM
ORCID
0000-0002-8107-9974

A Future For Everyone

Follow Us

Appendix G. Dermatology Biobank REC approval letter

NHS
Health Research Authority

North East – Newcastle & North Tyneside 1 Research Ethics Committee

NHSBT Newcastle Blood Donor Centre
Holland Drive
Newcastle upon Tyne
NE2 4NQ

Telephone: 0207 1048084

14 February 2019

Professor Penny Lovat
The Medical School
Newcastle University
Framlington Place
Newcastle upon Tyne
NE2 4HH

Dear Professor Lovat

Title of the Research Tissue Bank:	**Dermatology Biobank Renewal March 2019**
REC reference:	**19/NE/0004**
Designated Individual:	**Dr Christopher Morris**
IRAS project ID:	**256840**

Thank you for your letter dated 18 January 2019, responding to the Committee's request for further information on the above research tissue bank and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

We plan to publish your research summary wording for the Research Tissue Bank on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all Research Tissue Banks that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact hra.studyregistration@nhs.net outlining the reasons for your request.Under very limited circumstances (e.g. for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the Research Tissue Bank.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a **Favourable** ethical opinion of the above research tissue bank on the basis described in the application form and supporting documentation as revised.

The Committee has also confirmed that the favourable ethical opinion applies to all research projects conducted in the UK using tissue or data supplied by the tissue bank, provided that the release of tissue or data complies with the attached conditions. It will not be necessary for these researchers to make project-based applications for ethical approval. They will be deemed to have ethical approval from this committee. You should provide the researcher with a copy of this letter as confirmation of this. The Committee should be notified of all projects receiving tissue and data from this tissue bank by means of an annual report.

This application was for the renewal of a Research Tissue Bank application. The previous REC Reference number for this application was 08/H0906/95+5.

A Research Ethics Committee established by the Health Research Authority

Duration of ethical opinion

The favourable opinion has been renewed for five years from the end of the previous five year period provided that you comply with the standard conditions of ethical approval for Research Tissue Banks set out in the attached document. You are advised to study the conditions carefully. The opinion may be renewed for a further period of up to five years on receipt of a fresh application. It is suggested that the fresh application is made 3-6 months before the 5 years expires, to ensure continuous approval for the research tissue bank.

Research Tissue Bank Renewals

The previous five year period was scheduled to run from 05/03/2014 to 05/03/2019. This Research Tissue Bank may be renewed for further periods of five years at a time by following the process described in the above paragraph.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Covering letter on headed paper [Cover Letter from PI_Prof Lovat]	1	27 November 2018
Human Tissue Authority licence [HTA Licence]		01 July 2016
Other [Cv Prof Lovat]		01 November 2018
Other [DBB Form 02 Specimen Request]	1	08 November 2018
Other [DBB RA 01 Freezer]	5	01 November 2018
Other [DBB RA 2 Transport]	5	01 November 2018
Other [DBB RA 03 Cryostat]	4	01 November 2018
Other [DBB RA 04 Microtome]	4	01 November 2018
Other [DBB RA 06 Derm Premises]	2	01 November 2018
Other [DBB Form 01 Withdrawal of consent]	2	06 December 2018
Other [DBB Form 03 User Agreement]	2	06 December 2018
Other [DBB RA 05 HTA premises]	2	06 December 2018
Other [DBB SOP 01 Management]	4	06 December 2018
Other [DBB SOP 02 Consent]	5	06 December 2018
Other [DBB SOP 03 Disposal]	4	06 December 2018
Other [DBB SOP 04 Induction]	5	06 December 2018
Other [DBB SOP 05 freezer failure]	5	06 December 2018
Other [DBB SOP 06 Tissue Processing]	5	06 December 2018
Other [DBB SOP 07 Tissue transfer in]	4	06 December 2018
Other [Annual progress report march 2018]	16.03.18	06 December 2018
Other [Response letter re 19/NE/0004]	1	28 January 2019
Participant consent form [Patient Consent 01 Surplus tissue]	10	28 January 2019
Participant consent form [Patient Consent 02 New Tissue]	10	28 January 2019
Participant consent form [Patient Consent 04 Adults normally discarded tissue]	10	28 January 2019
Participant consent form [patient consent 05 previously donated tissue]	5	28 January 2019
Participant consent form [Patient Consent 6 Guardians for child surplus tissue]	10	28 January 2019
Participant consent form [Patient consent form 7 Guardians new child tissue]	1	28 January 2019
Participant consent form [Patient consent form 8 Minors for surplus tissue]	1	28 January 2019
Participant consent form [Patient Consent form 9 Minors for new tissue]	1	28 January 2019

Participant information sheet (PIS) [Patient information sheet 2i]	10	28 January 2019
Participant information sheet (PIS) [Patient information sheet 1i]	10	28 January 2019
Participant information sheet (PIS) [Patient Info 4 Adult normally discarded]	10	28 January 2019
Participant information sheet (PIS) [patient Info 5i previously donated tissue]	5	28 January 2019
Participant information sheet (PIS) [Patient Info 06 Guardians for child surplus tissue]	10	28 January 2019
Participant information sheet (PIS) [Patient information sheet 7i Guardian for new child tissue]	1	28 January 2019
Participant information sheet (PIS) [Patient information sheet 8i- minor surplus tissue]	1	28 January 2019
Participant information sheet (PIS) [Patient information sheet 9i minor new tissue]	1	28 January 2019
Protocol for management of the tissue bank [Dermatology Biobank Protocol Version 9]	9	28 January 2019
REC Application Form [RTB_Form_03122018]		03 December 2018
Relative consent form [Patient Consent 3 Relative]	10	28 January 2019
Relative information sheet [Patient Info 3 Relative]	10	28 January 2019
Summary of research programme(s) [Summary of Derm Biobank Research projects]	2	28 January 2019

Licence from the Human Tissue Authority

Thank you for providing a copy of the above licence.

Research governance

Under the UK Policy Framework for Health and Social Care Research there is no requirement for NHS research permission for the establishment of research tissue banks in the NHS. Applications to NHS R&D offices through IRAS are not required as all NHS organisations are expected to have included management review in the process of establishing the research tissue bank.

Research permission is also not required by collaborators at tissue collection centres (TCCs) who provide tissue or data under the terms of a supply agreement between the organisation and the research tissue bank. TCCs are not research sites for the purposes of the RGF.

Research tissue bank managers are advised to provide R&D offices at all TCCs with a copy of the REC application for information, together with a copy of the favourable opinion letter when available. All TCCs should be listed in Part C of the REC application.

NHS researchers undertaking specific research projects using tissue or data supplied by a research tissue bank must apply for permission to R&D offices at all organisations where the research is conducted, whether or not the research tissue bank has ethical approval.

Site-specific assessment (SSA) is not a requirement for ethical review of research tissue banks.

Registration of Research Tissue Banks

It is a condition of the ethical approval that all Research Tissue Banks are registered on the UK Clinical Research Collaboration (UKCRC) Tissue Directory. The Research Tissue Bank should be registered no later than 6 weeks after the date of this favourable ethical opinion letter or 6 weeks after the Research Tissue Bank holds tissue with the intention to provide for research purposes. Please use the following link to register the Research Tissue Bank on the UKCRC Directory: https://directory.biobankinguk.org/Register/Biobank Registration is defined as having added details of the types of tissue samples held in the tissue bank.

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment or annual progress report form. We will monitor the registration details as part of the annual progress reporting process.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached standard conditions give detailed guidance on reporting requirements for research tissue banks with a favourable opinion, including:

- Notifying substantial amendments
- Submitting Annual Progress reports.

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website: http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at https://www.hra.nhs.uk/planning-and-improving-research/learning/

19/NE/0004	Please quote this number on all correspondence

Yours sincerely

pp

G. Mayer

**Mr Paddy Stevenson
Chair**

E-mail: nrescommittee.northeast-newcastleandnorthtyneside1@nhs.net

Enclosures:	Standard approval conditions SL-AC3
Copy to:	Dr Christopher Morris – Designated Individuals, Newcastle University

Appendix H. Patient information leaflet for donating normally discarded surplus skin for research

The Newcastle upon Tyne Hospitals **NHS**
NHS Foundation Trust

Dermatology Biobank
William Leech Building, M2.121
Translational and Clinical Research Institute,
Newcastle University.

Royal Victoria Infirmary
Queen Victoria Road
Newcastle upon Tyne
NE1 4LP

Tel: 0191 233 6161
Fax: 0191 201 0155
www.newcastle-hospitals.nhs.uk

Consent for the collection, storage and use of adult normally discarded <u>surplus</u> human tissues for research <u>Patient information sheet</u>

Introduction
We understand that you previously had or will be having a surgical procedure in which specimens of tissue removed during surgery are generally discarded. We would like to invite you to contribute this tissue to the Dermatology Biobank, and allow us to store and use this surplus tissue in future medical research.

What will happen if I agree?
With your agreement, the tissue that was removed taken during the procedure will be stored in the Dermatological Sciences Laboratory in a tissue bank and potentially used in future medical research.

With your approval, information from your medical records will be taken, made anonymous and kept safe on a computer in Dermatological Sciences, Newcastle University. A medical doctor or a member of the regulatory authorities or the NHS Trust, where it is relevant to you taking part in this study and who may be a member of the research team will look at your medical records to get the relevant information. Any information needed for research purposes will be made anonymous before it is given to the researcher and the samples will be assigned an anonymous code. The researchers will not be able to find out your name or any personal details from the information that they receive. The Dermatological Sciences Laboratory will be responsible for the safekeeping of any samples donated and for protecting the confidentiality of your personal information, in line with the general data protection regulatory guidance outlined by the Health Research Authority.

You will have the opportunity to discuss with a doctor issues relating to the use of your samples for research purposes. He or she will answer any questions you may have. If you decide that you want the samples to be stored in the tissue bank and used for research purposes, you will be asked to sign a consent form confirming your decision and stating that you have read and understood this sheet. When you sign the form you will give the ownership of the tissue to Dermatological Sciences, Newcastle University. They will then store it for an indefinite period of time and will be able to decide how it should be used for research. They will also have the right to dispose of unused stored material in an appropriate and ethical manner following normal procedures. If at any time you change your mind and do not wish stored tissue to be used for research, any such tissue (unless already used up prior to your request) will be destroyed by incineration, data will be deleted if easily accessible; your medical treatment or legal rights will not be affected.

Medical research and why this programme is important
If you agree, the tissue that we collect will be used as part of research projects set up or approved by the Dermatological Sciences laboratory at Newcastle University and the Newcastle Hospitals NHS Trust. These research projects may help us learn more about the causes of skin/melanocyte derived diseases and how they can be treated and prevented. Some of these research programmes could lead to the development of new products or processes, for example a new treatment for eczema, these may be developed commercially for the improvement of patient care, in which case there would be no financial benefit to you. Some of this research may involve an assessment of genetic material (DNA and/or RNA) to help us understand the genetic basis of health and skin/melanocyte derived disease.

The research will not involve human embryos, although some of our research involves the administration of samples into rodent (mice) or fish. This is only done when it is essential to further our understanding of the way in which a disease develops or responds to treatment. These experiments are performed according to the strict guidelines set up by the Government involving the minimum of distress to the rodent/fish used. In addition some of our research involves collaboration with commercial partners. If you do not wish to allow your sample to be used in this way, then you can still donate your sample without the need to consent to research involving the use of animals or in collaboration with commercial partners.

The tissue bank acts as a custodian of the samples it holds and releases them only to individuals or organizations that have an acceptable scientific background and work to high ethical standards. Any research must be properly approved. Some of the sample or material extracted from it may be used by researchers outside of Newcastle University or the UK and may include transfer of samples to research centres abroad and or to other approved tissue banks.

What are the possible risks of taking part?
There are no extra risks involved in participating in this study.

Are there any possible benefits?
In the future, the results of these research projects might identify patients more likely to develop skin disease or respond or have side effects to a particular treatment. The research will not affect your previous, current or future treatment, therefore you will not benefit personally from taking part in the research. A cost recovery system is operated to pay for the continued biobank running costs covering tissue processing, storage and release to authorized studies and users only.

Do I have to take part?
No. Your participation is this study is entirely voluntary. The research projects have been designed taking into account published National and European guidance on the use of human tissue in research. All medical research will be approved by an independent research ethics committee to ensure that the purpose of the study is adequate and that your welfare and confidentiality are protected.

If you do not want your tissue to be stored in the Tissue bank please tell us and do NOT sign the consent form. If you do not sign this form, the tissue will still be sent to the pathology laboratory to undergo those tests that are necessary for your care but it will not be used for research purposes. All unused tissue from your procedure will be disposed of using normal hospital methods. We will respect your decision and it will not in any way affect the treatment that you receive. Equally if you have already consented to the storage of your tissue in the Dermatology biobank, you are free to change your mind at any time, without giving any reason. This would not affect the standard of care you receive.

The Dermatology biobank has been approved by NRES Committee Newcastle and North Tyneside 1 who ensure that all the work carried out meets the expected national ethical standards

If you have any concerns or other questions about taking part, please contact one of the study principle investigators, Professor Penny Lovat (tel: 0191 208 7170) or Professor Nick Reynolds (tel:0191 208 8936) or contact Dr Tom Ewen, The Dermatology Biobank Manager (Tel: 0191 208 5644). You can also seek advice on the study from the Health Research Authority (HRA, www.hra.nhs.uk; telephone 0207 108 8118). Similarly, the Patient Advice and Liaison Service (PALS, Freephone 0800 032 0202; Text; 01670 511098; email: northoftynepals@nhct.nhs.uk) can provide confidential advice on the study if you wish to have any further details.

300

Appendix I. Consent form for donating normally discarded surplus skin for research

The Newcastle upon Tyne Hospitals NHS
NHS Foundation Trust

Royal Victoria Infirmary
Queen Victoria Road
Newcastle upon Tyne
NE1 4LP

Dermatology Biobank
William Leech Building, M2.121
Translational and Clinical Research Institute,
Newcastle University.

Tel: 0191 233 6161
Fax: 0191 201 0155
www.newcastle-hospitals.nhs.uk

Consent Form for the collection, storage and use of <u>Adult Normally Discarded</u> surplus human tissues for research

Patient Name: …………………….………….. NHS Number: ………………………

Date of Birth: ……………………….

I agree that the tissue referred to below may be used for research, including genetic (DNA and/or RNA) studies. I also agree that this tissue becomes the property of Newcastle University to be used at its discretion (including dispersement to approved tissue banks) in properly approved research programs, which may lead to the development of new patient treatments for the benefit of patient care.

Please initial the boxes if you are agree with the corresponding points:

1. I have read and understood the attached information sheet 4i (version 12, dated 22nd March 2021) on this project and have been given a copy to keep. I have been able to ask questions and I understand why the research is being done and any risks involved.

2. I consent to the use of tissue, taken as a part of my diagnostic investigation or treatment, for research purposes. I understand that I am free to withdraw my approval for use of the samples at any time without my medical treatment or legal rights being affected.

3. I agree that the samples and the information gathered from them can be stored by Dermatological Sciences in a Tissue Bank and used for future research. I understand that relevant sections of my medical notes and /or data collected during the study may be looked at by individuals from regulatory authorities or the NHS Trust, where it is relevant to my taking part in this study. I give permission for these individuals to have access to my records.

4. I give permission for someone in the research team to look at my medical records. I understand that this information will be kept confidential, including from any laboratory staff analyzing the samples. I understand that this information, in an anonymous form may be passed to persons outside the

White copy-original   Yellow copy-patient   Pink copy-medical notes

university in connection with research and may be published in any research findings.

5. I understand that I will not benefit financially if this research leads to the development of a new treatment or medical test. ☐

6. I understand that I will not be told the results of any tests carried out on the samples. Summary research study results can be fed back by the research team if requested; such feedback will not change the medical care I receive. ☐

7. I give my permission for my sample to be used in experiments using rodents (mice) and fish. ☐

8. I give my permission for my samples and anonymized clinical information to be sent to centres outside the UK for research purposes. ☐

9. I give permission from my sample to be used by commercial partners (e.g. pharmaceutical companies) for research purposes. ☐

10. I know how to contact the research team if I need to. ☐

Name…………………..…………… Signature ……..………… Date…………………...

I have explained the request for tissue for research purposes and answered such questions as the patient has asked.

Health professional……………………… Signature………………… Date………………..

A copy of this consent form must be filed in the patient clinical notes and a copy given to the patient. The original consent form must be sent to Prof Penny Lovat / Prof N Reynolds at Dermatological Sciences (Newcastle University) where it will be filed.

To be completed by Clinical staff:

Associated Study Name: ……………………………………………………………………………

Principal Investigator: ……………………………………………………………………………

Sample Study Code: ……………………………………………………………………………

Sample Description: ……………………………………………………………………………

Sample Site: ……………………………………………………………………………
To be completed by Biobank staff:

Biobank Number:………………………………………… …………………………………

Date of Data input ……………………………………………………………………………

Appendix J. Patient information leaflet for donating skin for research

The Newcastle upon Tyne Hospitals **NHS**
NHS Foundation Trust

Dermatology Biobank
William Leech Building, M2.121
Translational and Clinical Research Institute,
Newcastle University.

Royal Victoria Infirmary
Queen Victoria Road
Newcastle upon Tyne
NE1 4LP

Tel: 0191 233 6161
Fax: 0191 201 0155
www.newcastle-hospitals.nhs.uk

Consent for the collection, storage and use of human tissues for research
Patient information sheet

Introduction
We would like to invite you to participate in medical research, by allowing us to take a biopsy (sample of tissue) and/or blood test. The sample will be stored in the Dermatology Biobank and used in future research studies, which may help patients with skin or melanocyte derived disease.

What will happen if I take part?
We would like you to consider having a blood test or small tissue biopsy taken purely for medical research. With your consent these could be used for research to help us learn more about the causes of diseases and how they can be treated and prevented.

In some cases, your doctor may already be planning a routine biopsy or blood test; in this case we would like you to consider allowing extra tissue to be removed from the operation site during your routine procedure or extra blood to be taken, purely for research. Such extra samples will only be taken if you give your consent and if their removal is unlikely to cause you any harmful effects now or in the future.

If you give approval, we will collect your biopsy sample from the operating theatre. The tissue taken will be processed and/or stored in the Dermatological Sciences Laboratory in a tissue bank and potentially used in future medical research.

If your samples are stored, with your approval, information about your health will be taken from your medical records, made anonymous and kept safe on a computer in Dermatological Sciences, Newcastle University. This will help us understand what your illness was like and relate what we find in the laboratory to what happens to patients. A medical doctor or a member of the regulatory authorities or the NHS Trust, where it is relevant to you taking part in this study and who may be a member of the research team will look at your medical records to get the relevant information about your health. Any information about your health and diagnosis needed for research purposes will be made anonymous before it is given to the researchers and your samples will be assigned an anonymous code. The researchers will not be able to find out your name or any personal details about you from the information that they receive. The Dermatological Sciences Laboratory will be responsible for the safekeeping of any samples you donate and for protecting the confidentiality of your personal information, in line with the general data protection regulatory guidance outlined by the Health Research Authority.

You will have the opportunity to discuss with a doctor issues relating to the biopsy procedure and the use of your samples for research purposes. He or she will answer any questions you may have. If you decide that you want to donate samples for medical research and storage in the tissue bank, you will be asked to sign a consent form confirming your decision and stating that you have read and understood this sheet. When you sign the form you will give the ownership of the tissue to Dermatological sciences, Newcastle University. They will then store it for an indefinite period of time and will be able to decide how it should be used for research. They will also have the right to dispose of unused stored material in an appropriate and ethical manner following normal procedures. If at any time you change your mind and do not wish stored tissue to be used for research, any such tissue (unless already used prior to your request) will be destroyed by incineration, data will be deleted if easily accessible; your medical treatment or legal rights will not be affected.

Medical research and why this programme is important
If you agree, the tissue that we collect will be used as part of research projects set up or approved by the Dermatological Sciences laboratory at Newcastle University and the Newcastle Hospitals NHS Trust. These research projects may help us learn more about the causes of skin/melanocyte derived diseases and how they can be treated and prevented. Some of these research programmes could lead to the development of new products or processes, for example a new treatment for eczema, these may be developed commercially for the improvement of patient care, in which case there would be no financial benefit to you. Some of this research may involve an assessment of genetic material (DNA and/or RNA) to help us understand the genetic basis of health and skin/melanocyte derived disease.

The research will not involve human embryos, although some of our research involves the administration of samples into rodent (mice) or fish. This is only done when it is essential to further our understanding of the way in which a disease develops or responds to treatment. These experiments are performed according to the strict guidelines set up by the Government involving the minimum of distress to the rodent/fish used. In addition some of our research involves collaboration with commercial partners. If you do not wish to allow your sample to be used in this way, then you can still donate your sample without the need to consent to research involving the use of animals or in collaboration with commercial partners. The tissue bank acts as a custodian of the samples it holds and releases them only to individuals or organizations that have an acceptable scientific background and work to high ethical standards. Any research must be properly approved. Some of the sample or material extracted from it may be used by researchers outside of Newcastle University or the UK and may include transfer of samples to research centres abroad and or to other approved tissue banks

What are the possible risks of taking part?
There are no extra risks involved in participating in this study.

Are there any possible benefits?
In the future, the results of these research projects might identify patients more likely to develop skin disease or respond or have side effects to a particular treatment. The research will not affect your previous, current or future treatment, therefore you will not benefit personally from taking part in the research. A cost recovery system is operated to pay for the continued biobank running costs covering tissue processing, storage and release to authorized studies and users only.

Do I have to take part?
No. Your participation is this study is entirely voluntary. The research projects have been designed taking into account published National and European guidance on the use of human tissue in research. All medical research will be approved by an independent research ethics committee to ensure that the purpose of the study is adequate and that your welfare and confidentiality are protected.

If you do not want your tissue to be stored in the Tissue bank please tell us and do NOT sign the consent form. If you do not sign this form, the tissue will still be sent to the pathology laboratory to undergo those tests that are necessary for your care but it will not be used for research purposes. All unused tissue from your procedure will be disposed of using normal hospital methods. We will respect your decision and it will not in any way affect the treatment that you receive. Equally if you have already consented to the storage of your tissue in the Dermatology biobank, you are free to change your mind at any time, without giving any reason. This would not affect the standard of care you receive.

The Dermatology biobank has been approved by NRES Committee Newcastle and North Tyneside 1 who ensure that all the work carried out meets the expected national ethical standards.

If you have any concerns or other questions about taking part, please contact one of the study principle investigators, Professor Penny Lovat (tel: 0191 208 7170) or Professor Nick Reynolds (tel:0191 208 8936) or contact Dr Tom Ewen, The Dermatology Biobank Manager (Tel: 0191 208 5644). You can also seek advice on the study from the Health Research Authority (HRA, www.hra.nhs.uk; telephone 0207 108 8118). Similarly, the Patient Advice and Liaison Service (PALS, Freephone 0800 032 0202; Text; 01670 511098; email: northoftynepals@nhct.nhs.uk) can provide confidential advice on the study if you wish to have any further details.

Appendix K. Consent form for donating skin for research

The Newcastle upon Tyne Hospitals **NHS**
NHS Foundation Trust

Royal Victoria Infirmary
Queen Victoria Road
Newcastle upon Tyne
NE1 4LP

Dermatology Biobank
William Leech Building, M2.121
Translational and Clinical Research Institute,
Newcastle University

Tel: 0191 233 6161
Fax: 0191 201 0155
www.newcastle-hospitals.nhs.uk

Patient Consent Form for the collection, storage and use of <u>human tissues</u> for research

Patient Name: …………………………….…   NHS Number : …………….……….

Date of Birth : ………………………..

I agree that the tissue referred to below may be used for research, including genetic (DNA and/or RNA) studies. I also agree that this tissue becomes the property of Newcastle University to be used at its discretion (including dispersement to approved tissue banks) in properly approved research programs, which may lead to the development of new patient treatments for the benefit of patient care.

Please initial the boxes if you are agree with the corresponding points:

1. I have read and understood the attached patient information sheet 2i (version 12, dated 22nd March 2021) on this project and have been given a copy to keep. I have been able to ask questions and I understand why the research is being done and any risks involved.

2. I give consent for blood and/or tissue samples to be taken for medical research purposes, as specified below*. I understand that I am free to withdraw my approval for use of the samples at any time without my medical treatment or legal rights being affected.

3. I agree that the samples I have given and the information gathered about me can be stored by Dermatological sciences in a Tissue Bank and used for future research. I understand that relevant sections of my medical notes and /or data collected during the study may be looked at by individuals from regulatory authorities or the NHS Trust, where it is relevant to my taking part in this study. I give permission for these individuals to have access to my records.

4. I give permission for someone in the research team to look at my medical records to get information about my medical condition and health. I understand that this information will be kept confidential, including from any laboratory staff analyzing the samples. I understand that this information, in an anonymous form may be passed to persons outside the university in connection with research and may be published in any research findings.

5. I understand that I will not benefit financially if this research leads to the development of a new treatment or medical test.

   ☐

6. I understand that I will not be told the results of any tests carried out on my samples. However summary research study results can be fed back via my GP if requested, such feedback will not change my medical care.

   ☐

7. I give my permission for my sample to be used in experiments using rodents (mice) and fish.

   ☐

8. I give my permission for my samples and anonymized clinical information to be sent to centres outside the UK for research purposes.

   ☐

9. I give permission from my sample to be used by commercial partners (e.g. pharmaceutical companies) for research purposes.

   ☐

10. I know how to contact the research team if I need to.

   ☐

Patient's name……………………… Signature…………………..…….. Date……………..

I have explained the request for tissue for research purposes and answered such questions as the patient has asked.

Health professional………………………...Signature…………………..Date……………...

A copy of this consent form must be filed in the patient clinical notes and a copy given to the patient. The original consent form must be sent to Prof Penny Lovat / Prof N Reynolds at Dermatological Sciences (Newcastle University) where it will be filed.

To be completed by Clinical staff:

Associated Study Name: ……………………………………………………………………………

Principal Investigator: ……………………………………………………………………………

Sample Study Code: ……………………………………………………………………………

Sample Description: ……………………………………………………………………………

Sample Site: ……………………………………………………………………………
To be completed by Biobank staff:

Biobank Number:……………………………………………………………………………

Date of Data input ……………………………………………………………………………

Patient consent form 2
Version 10 – 22/03/2021

2

Appendix L. HDBR REC approval letter

NHS
Health Research Authority
North East – Newcastle & North Tyneside 1 Research Ethics Committee

NHSBT Newcastle Blood Donor Centre
Holland Drive
Newcastle upon Tyne
NE2 4NQ

10 December 2018

Telephone: 0207 1048084

Professor Stephen C Robson
Institute of Cellular Medicine
3rd Floor, Leazes Wing
Medical School
Newcastle University
Newcastle upon Tyne
NE2 4HH

Dear Professor Robson

Title of the Research Tissue Bank:	**The Human Developmental Biology Resource**
REC reference:	**18/NE/0290**
Designated Individual:	**Dr Chris Morris**
IRAS project ID:	**250012**

Thank you for your letter of 8 November 2018, responding to the Committee's request for further information on the above research tissue bank and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Vice Chair.

We plan to publish your research summary wording for the Research Tissue Bank on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all Research Tissue Banks that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact hra.studyregistration@nhs.net outlining the reasons for your request.Under very limited circumstances (e.g. for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the Research Tissue Bank.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a **Favourable** ethical opinion of the above research tissue bank on the basis described in the application form and supporting documentation as revised.

Recommendation: The Committee recommend that it could be highlighted in the information documents that possible use of the donated tissue with animal sourced material may occur and if a participant has any moral objections to that then they should not consider donation (so that consent can be full and competent).

The Committee has also confirmed that the favourable ethical opinion applies to all research projects conducted in the UK using tissue or data supplied by the tissue bank, provided that the release of tissue or data complies with the attached conditions. It will not be necessary

for these researchers to make project-based applications for ethical approval. They will be deemed to have ethical approval from this committee. You should provide the researcher with a copy of this letter as confirmation of this. The Committee should be notified of all projects receiving tissue and data from this tissue bank by means of an annual report.

This application was for the renewal of a Research Tissue Bank application. The previous REC Reference number for this application was 08/H0906/21+5.

Duration of ethical opinion

The favourable opinion has been renewed for five years from the end of the previous five year period provided that you comply with the standard conditions of ethical approval for Research Tissue Banks set out in the attached document. You are advised to study the conditions carefully. The opinion may be renewed for a further period of up to five years on receipt of a fresh application. It is suggested that the fresh application is made 3-6 months before the 5 years expires, to ensure continuous approval for the research tissue bank.

Research Tissue Bank Renewals

The previous five year period ran from 04/09/2013 to 03/09/2018. This Research Tissue Bank may be renewed for further periods of five years at a time by following the process described in the above paragraph.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Covering letter on headed paper		16 September 2018
Human Tissue Authority licence		06 July 2016
Other [HDBR Tissue Access Policy]	1	03 August 2018
Other [Protocol for collection of consented material]	3	20 July 2016
Other [HDBR Sample Sign out Form]	2	01 February 2018
Other [HDBR Anonymous Participant Data Questionnaire]	1	03 August 2018
Other [Latest HDBR Annual Report 2016/17]	N/A	26 July 2017
Other [Chief Investigator CV]		
Other [Response letter REC]	N/A	08 November 2018
Other [UCH HTA License]	N/A	01 January 2013
Other [Revised Plain English Research Summary]	1	08 November 2018
Participant consent form	5	27 July 2018
Participant information sheet (PIS)	6.1	08 November 2018
Protocol for management of the tissue bank [Protocol for Recruitment to HDBR]	3	20 July 2016
REC Application Form [RTB_Form_17082018]		17 August 2018
Summary of research programme(s) [HDBR Terms of Reference]	1	03 August 2018

Licence from the Human Tissue Authority

Thank you for providing a copy of the above licence.

Research governance

Under the UK Policy Framework for Health and Social Care Research there is no requirement for NHS research permission for the establishment of research tissue banks in

the NHS. Applications to NHS R&D offices through IRAS are not required as all NHS organisations are expected to have included management review in the process of establishing the research tissue bank.

Research permission is also not required by collaborators at tissue collection centres (TCCs) who provide tissue or data under the terms of a supply agreement between the organisation and the research tissue bank. TCCs are not research sites for the purposes of the RGF.

Research tissue bank managers are advised to provide R&D offices at all TCCs with a copy of the REC application for information, together with a copy of the favourable opinion letter when available. All TCCs should be listed in Part C of the REC application.

NHS researchers undertaking specific research projects using tissue or data supplied by a research tissue bank must apply for permission to R&D offices at all organisations where the research is conducted, whether or not the research tissue bank has ethical approval.

Site-specific assessment (SSA) is not a requirement for ethical review of research tissue banks.

Registration of Research Tissue Banks

It is a condition of the ethical approval that all Research Tissue Banks are registered on the UK Clinical Research Collaboration (UKCRC) Tissue Directory. The Research Tissue Bank should be registered no later than 6 weeks after the date of this favourable ethical opinion letter or 6 weeks after the Research Tissue Bank holds tissue with the intention to provide for research purposes. Please use the following link to register the Research Tissue Bank on the UKCRC Directory: https://directory.biobankinguk.org/Register/Biobank Registration is defined as having added details of the types of tissue samples held in the tissue bank.

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment or annual progress report form. We will monitor the registration details as part of the annual progress reporting process.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.
After ethical review

Reporting requirements

The attached standard conditions give detailed guidance on reporting requirements for research tissue banks with a favourable opinion, including:

- Notifying substantial amendments
- Submitting Annual Progress reports.

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known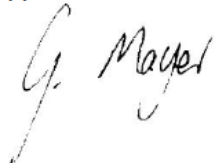 please use the feedback form available on the HRA website: http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at
https://www.hra.nhs.uk/planning-and-improving-research/learning/

18/NE/0290	Please quote this number on all correspondence

Yours sincerely
pp

[signature]

Dr Mike Bone
Vice Chair

E-mail: nrescommittee.northeast-newcastleandnorthtyneside1@nhs.net

Enclosures: RTB Standard approval conditions SL-AC3

Copy to: Dr Chris Morris – Designated Individual, Newcastle University

CONDITIONS OF ETHICAL APPROVAL

Research Ethics Committee:	North East - Newcastle & North Tyneside 1 Research Ethics Committee
Research Tissue Bank:	The Human Developmental Biology Resource
REC reference number:	18/NE/0290
Name of applicant:	Prof Stephen C Robson
Date of approval:	10 December 2018
IRAS project ID:	250012

Ethical approval is given to the Research Tissue Bank ("the Bank") by the Research Ethics Committee ("the Committee") subject to the following conditions.

1. Further communications with the Committee

1.1 Further communications with the Committee are the personal responsibility of the applicant.

2. Duration of approval

2.1 Approval is given for a period of 5 years, which may be renewed on consideration of a new application by the Committee, taking account of developments in legislation, policy and guidance in the interim. New applications should include relevant changes of policy or practice made by the Bank since the original approval together with any proposed new developments.

3. Licensing

3.1 A copy of the Licence from the Human Tissue Authority (HTA) should be provided when available (if not already submitted).

3.2 The Committee should be notified if the Authority renews the licence, varies the licensing conditions or revokes the Licence, or of any change of Designated Individual. If the Licence is revoked, ethical approval would be terminated.

4.   Generic ethical approval for projects receiving tissue

4.1   Samples of human tissue or other biological material may be supplied and used in research projects to be conducted in accordance with the following conditions.

   4.1.1   The research project should be within the fields of medical or biomedical research described in the approved application form.

   4.1.2   The Bank should be satisfied that the research has been subject to scientific critique, is appropriately designed in relation to its objectives and (with the exception of student research below doctoral level) is likely to add something useful to existing knowledge.

   4.1.3   Where tissue samples have been donated with informed consent for use in future research ("broad consent"), the Bank should be satisfied that the use of the samples complies with the terms of the donor consent.

   4.1.4   All samples and any associated clinical information must be non-identifiable to the researcher at the point of release (i.e. anonymised or linked anonymised).

   4.1.5   Samples will not be released to any project requiring further data or tissue from donors or involving any other research procedures.   Any contact with donors must be confined to ethically approved arrangements for the feedback of clinically significant information.

   4.1.6   A supply agreement must be in place with the researcher to ensure storage, use and disposal of the samples in accordance with the HTA Codes of Practice, the terms of the ethical approval and any other conditions required by the Bank.

4.2   A research project in the UK using tissue provided by a Bank in accordance with these conditions will be considered to have ethical approval from the Committee under the terms of this approval.   In England, Wales and Northern Ireland this means that the researcher will not require a licence from the Human Tissue Authority for storage of the tissue for use in relation to this project.

4.3   The Bank may require any researcher to seek specific ethical approval for their project.   Such applications should normally be made to the Committee and booked via the Central Booking System

4.4   A Notice of Substantial Amendment should be submitted to seek the Committee's agreement to change the conditions of generic approval.

5. Records

5.1   The Bank should maintain a record of all research projects to which tissue has been supplied. The record should contain at least the full title of the project, a summary of its purpose, the name of the Chief Investigator, the sponsor, the location of the research, the date on which the project was approved by the Bank, details of the tissue released and any relevant reference numbers.

5.2   The Committee may request access to these records at any time.

6. Annual reports

6.1 An annual report should be provided to the Committee listing all projects for which tissue has been released in the previous year. The list should give the full title of each project, the name of the Chief Investigator, the sponsor, the location of the research and the date of approval by the Bank. The report is due on the anniversary of the date on which ethical approval for the Bank was given.

6.2 The Committee may request additional reports on the management of the Bank at any time.

7. Substantial amendments

7.1 Substantial amendments should be notified to the Committee and ethical approval sought before implementing the amendment.   A substantial amendment generally means any significant change to the arrangements for the management of the Bank as described in the application to the Committee and supporting documentation.

7.2 A Notice of Substantial Amendment should be generated by accessing the original application form on the Integrated Research Application System (IRAS).

7.3 The following changes should always be notified as substantial amendments:

   7.3.1   Any significant change to the policy for use of the tissue in research, including changes to the types of research to be undertaken or supported by the Bank.

   7.3.2   Any significant change to the types of biological material to be collected and stored, or the circumstances of collection.

   7.3.3   Any significant change to informed consent arrangements, including new/modified information sheets and consent forms.

   7.3.4   A change to the conditions of generic approval

   7.3.5   Any other significant change to the governance of the RTB.

8. Serious Adverse Events

8.1 The Committee should be notified as soon as possible of any serious adverse event or reaction, any serious breach of security or confidentiality, or any other incident that could undermine public confidence in the ethical management of the tissue.   The criteria for notifying the Committee will be the same as those for notifying the Human Tissue Authority in the case of research tissue banks in England, Wales and Northern Ireland.

9. Other information to be notified

9.1 The Committee should be notified of any change in the contact details for the applicant

or where the applicant hands over responsibility for communication with the Committee to another person at the establishment.

10. Closure of the Bank

10.1   Any plans to close the Bank should be notified to the Committee as early as possible and at least two months before closure.   The Committee should be informed what arrangements are to be made for disposal of the tissue or transfer to another research tissue bank.

10.2   Where tissue is transferred to another research tissue bank, the ethical approval for the Bank is not transferable. Where the second bank is ethically approved, it should notify the responsible Research Ethics Committee. The terms of its own ethical approval would apply to any tissue it receives.

11. Breaches of approval conditions

11.1 The Committee should be notified as soon as possible of any breach of these approval conditions.

11.2 Where serious breaches occur, the Committee may review its ethical approval and may, exceptionally, suspend or terminate the approval.

Appendix M. Metadata table for adult healthy, psoriasis and eczema skin samples

Donor	Replicate	Condition	Anatomical site	Sex	Age	Slide number	Capture area
P730	P730_1	Healthy	Face - temple	F	74	1	A1
P729	P729_1	Healthy	Face - temple	M	52	1	B1
P748	P748_1	Healthy	Forearm	M	77	2	A1
P736	P736_1	Healthy	Thigh	F	60	2	B1
P734	P734_1	Healthy	Face - temple	M	58	3	A1
P734	P734_2	Healthy	Forearm	M	58	3	B1
P732	P732_1	Healthy	Scalp	F	47	4	A1
P738	P738_1	Healthy	Arm	F	61	4	B1
P746	P746_1	Healthy	Face - cheek	F	33	5	A1
P747	P747_1	Healthy	Forearm	M	82	5	B1
P728	P728_1	Healthy	Sternum	F	41	7	B1
P730	P730_2	Healthy	Face - temple	F	74	7	C1
P729	P729_2	Healthy	Face - temple	M	52	7	A1
P748	P748_2	Healthy	Forearm	M	77	6	A1
P736	P736_2	Healthy	Thigh	F	60	6	B1
P734	P734_3	Healthy	Face - temple	M	58	8	A1
P734	P734_4	Healthy	Forearm	M	58	8	B1
P732	P732_2	Healthy	Scalp	F	47	9	A1
P738	P738_2	Healthy	Arm	F	61	9	B1
P746	P746_2	Healthy	Face - cheek	F	33	10	A1
P747	P747_2	Healthy	Forearm	M	82	10	B1
P775	P775_1	Healthy	Abdomen - right	M	56	11	A1
P777	P777_1	Healthy	Back - right lower	F	62	11	B1
P777	P777_2	Healthy	Back - right lower	F	62	11	C1
P742	P742_1	Healthy	Back - upper	F	70	11	D1
P775	P775_2	Healthy	abdomen - right	F	56	12	A1
P777	P777_3	Healthy	Back - right lower	F	62	12	B1
P777	P777_4	Healthy	Back - right lower	F	62	12	C1
P742	P742_2	Healthy	Back - upper	F	70	12	D1
P731	P731_1	Psoriasis lesional	Lower back	F	18	1	D1
P735	P735_1	Psoriasis lesional	Lower back	M	31	2	C1
P737	P737_1	Psoriasis lesional	Lower back	M	58	3	C1
P743	P743_1	Psoriasis lesional	Lower back	F	71	4	C1
P744	P744_1	Psoriasis lesional	Lower back	M	21	5	C1
P735	P735_2	Psoriasis lesional	Lower back	M	31	6	C1
P737	P737_2	Psoriasis lesional	Lower back	M	58	8	C1
P743	P743_2	Psoriasis lesional	Lower back	F	71	9	C1
P744	P744_2	Psoriasis lesional	Lower back	M	21	10	C1
P731	P731_1	Psoriasis non-lesional	Lower back	F	18	1	C1

P735	P735_1	Psoriasis non-lesional	Lower back	M	31	2	D1
P737	P737_1	Psoriasis non-lesional	Lower back	M	58	3	D1
P743	P743_1	Psoriasis non-lesional	Lower back	F	71	4	D1
P744	P744_1	Psoriasis non-lesional	Lower back	M	21	5	D1
P731	P731_2	Psoriasis non-lesional	Lower back	F	18	7	D1
P735	P735_2	Psoriasis non-lesional	Lower back	M	31	6	D1
P737	P737_2	Psoriasis non-lesional	Lower back	M	58	8	D1
P743	P743_2	Psoriasis non-lesional	Lower back	F	71	9	D1
P744	P744_2	Psoriasis non-lesional	Lower back	M	21	10	D1
AD2628	AD2628_1	Eczema lesional	Lower back	M	60	17	B1
AD2628	AD2628_2	Eczema lesional	Lower back	M	60	14	B1
AD2628	AD2628_3	Eczema lesional	Lower back	M	60	13	C1
AD2628	AD2628_4	Eczema lesional	Lower back	M	60	13	D1
AD2634	AD2634_1	Eczema lesional	Lower back	M	35	14	C1
AD2638	AD2638_1	Eczema lesional	Lower back	F	42	15	B1
AD2639	AD2639_1	Eczema lesional	Lower back	F	23	16	C1
AD2639	AD2639_2	Eczema lesional	Lower back	F	23	16	D1
AD2640	AD2640_1	Eczema lesional	Lower back	M	24	13	A1
AD2640	AD2640_2	Eczema lesional	Lower back	M	24	13	B1
AD2628	AD2628_1	Eczema non-lesional	Lower back	M	60	17	A1
AD2628	AD2628_2	Eczema non-lesional	Lower back	M	60	14	A1
AD2634	AD2634_1	Eczema non-lesional	Lower back	M	35	17	C1
AD2638	AD2638_1	Eczema non-lesional	Lower back	F	42	15	C1
AD2638	AD2638_2	Eczema non-lesional	Lower back	F	42	15	D1
AD2639	AD2639_1	Eczema non-lesional	Lower back	F	23	16	A1
AD2639	AD2639_2	Eczema non-lesional	Lower back	F	23	16	B1

Appendix N. Prenatal skin refined cell annotation differentially expressed genes

	genes	cluster	logfoldchanges	pvals	pvals_adj
0	KRT4	Periderm	12.443899	1.49E-96	1.28E-93
1	IGFL2	Periderm	12.620115	1.82E-85	6.26E-83
2	UPK2	Periderm	9.870495	3.25E-66	5.08E-64
3	GABRP	Periderm	11.289986	1.58E-69	3.88E-67
4	SPRR3	Periderm	11.941079	1.68E-52	2.22E-50
5	SCEL	Periderm	12.125658	4.41E-83	1.26E-80
6	KRT18	Periderm	8.956959	2.27E-92	1.30E-89
7	C2orf54	Periderm	10.988039	4.69E-47	5.77E-45
8	KRT8	Periderm	9.938104	2.42E-100	4.16E-97
9	ELF3	Periderm	10.551799	1.89E-69	4.07E-67
10	IL36RN	Periderm	12.581477	2.34E-33	1.91E-31
11	NDRG2	Periderm	5.243946	9.40E-92	4.04E-89
12	AGR3	Periderm	11.771323	3.93E-37	3.38E-35
13	GPRC5A	Periderm	5.3127484	2.58E-38	2.46E-36
14	MUC16	Periderm	11.721866	4.41E-29	3.04E-27
15	VTCN1	Periderm	12.200967	2.93E-23	1.74E-21
16	AC006262.5	Periderm	12.567321	5.74E-42	6.58E-40
17	SPRR2F	Periderm	9.748786	2.51E-26	1.66E-24
18	C1orf116	Periderm	10.403912	9.71E-30	7.26E-28
19	PCP4	Periderm	6.081897	1.57E-53	2.25E-51
20	PWWP2B	Periderm	5.1951632	9.73E-40	1.05E-37
21	RP11-379F4.4	Periderm	7.454037	9.77E-10	2.62E-08
22	ATP6V1B1	Periderm	8.085769	5.35E-14	2.19E-12
23	PRR15L	Periderm	11.967394	1.71E-24	1.09E-22
24	MARVELD3	Periderm	7.257426	2.81E-29	2.02E-27
25	KCNG1	Periderm	6.2050447	8.31E-21	4.61E-19
26	KB-1562D12.1	Periderm	9.430672	8.55E-24	5.25E-22
27	NEURL1	Periderm	3.6093817	5.05E-15	2.29E-13
28	LRP2	Periderm	9.685073	5.00E-08	1.15E-06
29	UPK1B	Periderm	8.877734	5.60E-17	2.83E-15
30	RP11-123B3.2	Periderm	11.892307	2.18E-08	5.21E-07
31	MUC1	Periderm	7.3156195	7.39E-15	3.10E-13
32	FAM3D	Periderm	10.118874	1.45E-32	1.13E-30
33	MTCL1	Periderm	6.0221815	6.22E-10	1.75E-08
34	WWC1	Periderm	6.5948277	1.27E-11	4.22E-10
35	INSL4	Periderm			
36	MUC22	Periderm			

37	DIRAS3	Periderm		4.8949327	6.15E-15	2.71E-13
38	TJP3	Periderm		8.383868	1.05E-08	2.59E-07
39	TMEM61	Periderm				
40	EVPL	Periderm		8.764293	1.04E-37	9.38E-36
41	NAALADL2	Periderm		2.637306	2.65E-12	8.95E-11
42	TMPRSS2	Periderm		9.108264	2.04E-06	3.81E-05
43	RP11-195B3.1	Periderm				
44	CXCL17	Periderm		10.823533	9.78E-07	1.93E-05
45	GPR87	Periderm		8.694902	5.44E-11	1.61E-09
46	RP11-416N2.4	Periderm		5.8686714	2.08E-05	0.000347749
47	PDE10A	Periderm		4.343164	9.68E-07	1.93E-05
48	AC006262.10	Periderm				
49	SLC15A2	Periderm		4.646712	3.77E-08	8.76E-07
50	AMOT	Periderm		2.667089	1.58E-12	5.56E-11
51	PRSS16	Periderm		9.163827	7.84E-06	0.000141942
52	TMEM125	Periderm		8.57647	1.14E-07	2.55E-06
53	HIC2	Periderm		3.1245623	1.57E-11	5.01E-10
54	KLHL36	Periderm		2.7032712	1.38E-11	4.47E-10
55	PGLYRP4	Periderm		9.852481	9.94E-07	1.93E-05
56	F3	Periderm		3.5077903	9.89E-16	4.86E-14
57	PARD6B	Periderm		5.8848243	9.64E-13	3.53E-11
58	CNGA1	Periderm		6.927306	3.15E-05	0.000511336
59	OCLN	Periderm		5.0452266	8.67E-11	2.49E-09
60	CLDN9	Periderm		8.625488	8.18E-10	2.27E-08
61	ERP27	Periderm		6.292387	1.84E-05	0.000312651
62	MUC20	Periderm				
63	TMCO4	Periderm		4.0873404	6.93E-07	1.44E-05
64	PXDNL	Periderm				
65	CAPS	Periderm		4.312723	1.29E-05	0.000226359
66	GABRA3	Periderm				
67	SH3TC2	Periderm				
68	KLK1	Periderm		5.157	2.90E-08	6.84E-07
69	FBP1	Periderm		5.1314883	3.84E-15	1.79E-13
70	CLDN3	Periderm		9.47084	1.66E-15	7.94E-14
71	CHRD	Periderm		2.3548176	3.61E-06	6.68E-05
72	PLEKHG4B	Periderm		3.9266944	6.46E-09	1.61E-07
73	CCDC64B	Periderm		7.2642055	5.33E-09	1.37E-07
74	NEBL	Periderm		4.075698	4.36E-13	1.63E-11
75	TEPP	Periderm				
76	CCDC183	Periderm		5.6231155	2.19E-05	0.000361949

77	MUC4	Periderm			
78	SLC16A5	Periderm	3.6512783	1.73E-06	3.31E-05
79	TOM1L2	Periderm	2.3643749	1.10E-08	2.66E-07
80	AMFR	Periderm	2.2979639	4.36E-06	7.98E-05
81	BEAN1	Periderm			
82	TMPRSS11E	Periderm	9.622306	1.47E-05	0.000254653
83	RHOBTB2	Periderm			
84	TGFA	Periderm	5.034206	9.63E-07	1.93E-05
85	SLC28A3	Periderm	7.750571	2.52E-07	5.41E-06
86	SLC2A1-AS1	Periderm	5.4908686	4.10E-07	8.70E-06
87	MGAT3	Periderm	4.0366616	1.60E-05	0.000275331
88	TMEM241	Periderm	3.067219	9.09E-07	1.86E-05
89	RARRES1	Periderm			
90	TMEM191C	Periderm			
91	ARHGAP8	Periderm	6.0338964	1.94E-05	0.000326818
92	SERHL2	Periderm			
93	RP11-532F12.5	Periderm	6.182914	3.73E-11	1.13E-09
94	SYT17	Periderm	4.48285	1.79E-07	3.95E-06
95	GNG4	Periderm	6.377764	6.88E-21	3.95E-19
96	MIR210HG	Periderm	4.007667	2.03E-12	7.00E-11
97	DEGS2	Periderm	6.357942	9.69E-06	0.000173697
98	CAB39L	Periderm	3.5027406	5.67E-09	1.43E-07
99	BRINP1	Immature basal	6.635342	2.28E-10	9.82E-08
100	CFAP57	Immature basal	7.6186414	0.0104581	0.829170083
101	RASSF6	Immature basal	6.663823	0.033815768	0.999924195
102	ACOT11	Immature basal	4.475705	0.000616448	0.070685998
103	LNX1	Immature basal	4.928556	0.000425396	0.052262996
104	IRX3	Immature suprabasal	4.711161	3.36E-07	0.000288741
105	LYPD3	Immature suprabasal	7.0004835	3.58E-11	6.15E-08
106	RP11-774O3.3	Immature suprabasal	5.936124	0.084203614	0.999684354
107	LPIN3	Immature suprabasal	5.7418313	0.030897838	0.999684354
108	KIFC2	Immature suprabasal	5.71668	0.087136062	0.999684354
109	ZNF814	Immature suprabasal	4.4141974	0.041419872	0.999684354
110	ZNF358	Immature suprabasal	3.329505	0.059813109	0.999684354
111	SGSM2	Immature suprabasal	4.20413	0.045147158	0.999684354
112	CHP2	Immature suprabasal	6.649369	0.028285793	0.999684354
113	RBBP8NL	Immature suprabasal	10.327675	0.074094786	0.999684354
114	NRARP	Immature suprabasal	5.636678	0.083982018	0.999684354
115	MTMR1	Immature suprabasal	4.54498	0.097716213	0.999684354
116	LINC00923	Immature suprabasal	8.325464	0.373305607	0.999684354

117	RP11-115D19.1	Immature suprabasal	10.664323	0.007358647	0.999684354
118	SERPINB13	Immature suprabasal	10.663053	0.180206103	0.999684354
119	ZNF438	Immature suprabasal	3.2829442	0.132506415	0.999684354
120	ZBTB6	Immature suprabasal	3.7250485	0.245374327	0.999684354
121	ADNP2	Immature suprabasal	3.0906699	0.148880976	0.999684354
122	CGN	Immature suprabasal	8.266206	0.182300637	0.999684354
123	ASTN2	Immature suprabasal	4.198622	0.105730939	0.999684354
124	PDE8B	Immature suprabasal	4.5116925	0.036736692	0.999684354
125	SZT2	Immature suprabasal	3.3461306	0.270281873	0.999684354
126	ZNF99	Immature suprabasal	5.8130994	0.385578928	0.999684354
127	ACPP	Immature suprabasal	6.1885667	0.082103315	0.999684354
128	ZNF641	Immature suprabasal	3.4868536	0.242041721	0.999684354
129	ITPRIPL2	Immature suprabasal	2.8393686	0.168816961	0.999684354
130	RP11-420A23.1	Immature suprabasal	6.6249676	0.377354025	0.999684354
131	TBX6	Immature suprabasal	5.1206675	0.199402899	0.999684354
132	GRM5-AS1	Immature suprabasal	5.9689817	0.383006928	0.999684354
133	FAM200A	Immature suprabasal	3.2815297	0.054635334	0.999684354
134	PIK3R2	Immature suprabasal	5.1405063	0.394660386	0.999684354
135	AC004231.2	Immature suprabasal	7.993656	0.373282441	0.999684354
136	DSC1	Immature suprabasal	10.049288	0.025608199	0.999684354
137	KRT14	POSTN+ basal	11.183341	3.15E-29	2.71E-26
138	COL17A1	POSTN+ basal	9.709164	1.84E-30	3.17E-27
139	EDN2	POSTN+ basal	9.364129	4.32E-21	1.50E-18
140	WNT3	POSTN+ basal	9.010021	4.36E-21	1.50E-18
141	GRIK1-AS1	POSTN+ basal	10.134525	5.74E-12	1.41E-09
142	LMO1	POSTN+ basal	9.680367	9.51E-10	1.82E-07
143	SYT8	POSTN+ basal	8.297426	4.71E-06	0.000505853
144	PTCHD4	POSTN+ basal	9.317444	8.78E-08	1.37E-05
145	UPK3B	POSTN+ basal	9.202848	4.52E-06	0.000505853
146	TP53AIP1	POSTN+ basal	8.749475	4.63E-06	0.000505853
147	LAMB4	POSTN+ basal	8.0767765	0.00037607	0.030801902
148	TNS4	POSTN+ basal	8.588221	1.34E-06	0.000191416
149	WNT16	POSTN+ basal	4.439509	4.51E-05	0.004308334
150	TECRL	POSTN+ basal	10.723921	0.000354083	0.030451113
151	COL8A1	POSTN+ basal	4.8791714	3.24E-05	0.003277673
152	MROH2A	POSTN+ basal	11.671583	0.201834199	0.999754189
153	RHBDL1	POSTN+ basal	5.985862	0.000448517	0.035065858
154	TMEM139	POSTN+ basal	6.6961017	0.043500104	0.999754189
155	WFDC5	POSTN+ basal	11.038086	0.01071556	0.458498611
156	PLA2R1	POSTN+ basal	3.6746929	0.006024484	0.320800451

157	CA9	POSTN+ basal	7.2121696	0.005517571	0.316340711
158	GRIK1	POSTN+ basal	5.4666433	0.006154892	0.320800451
159	FAM46B	POSTN+ basal	6.6794286	0.002461464	0.146115367
160	PRAP1	POSTN+ basal	7.5753326	0.075540686	0.999754189
161	TP73	POSTN+ basal	7.2550178	0.002398072	0.146115367
162	MIR4500HG	POSTN+ basal	5.9954157	0.078493045	0.999754189
163	ZNF8	POSTN+ basal	3.1086001	0.045393646	0.999754189
164	CDCP1	POSTN+ basal	5.0653515	0.000588842	0.044035137
165	GRIK4	POSTN+ basal	5.1752124	0.026215678	0.980238387
166	KCNJ5	POSTN+ basal	4.532326	0.028850015	0.999754189
167	DNER	POSTN+ basal	4.028272	0.00949098	0.429591732
168	AC005082.12	POSTN+ basal	3.438786	0.006498591	0.321410575
169	IGFL1	POSTN+ basal	10.517024	0.021636225	0.86544902
170	LRRC7	POSTN+ basal	4.258289	0.008227279	0.382457272
171	OTX1	POSTN+ basal	7.583703	0.022354562	0.873860144
172	FAM132A	POSTN+ basal	3.5102599	0.01070407	0.458498611
173	ETV4	DPYSL2+ basal	5.3727827	2.14E-55	2.83E-53
174	KRT6A	Suprabasal IFE	12.290916	1.92E-09	1.65E-06
175	KRTDAP	Suprabasal IFE	12.497265	3.79E-13	6.51E-10
176	AADACL2	Suprabasal IFE	8.6131735	0.004595011	0.376353314
177	KLK11	Suprabasal IFE	10.416666	0.004481022	0.376353314
178	SERPINB3	Suprabasal IFE	13.008033	0.11398298	0.999553592
179	CDKN2B	Suprabasal IFE	7.549345	0.00056493	0.097167992
180	THEM5	Suprabasal IFE	7.9017344	0.001653608	0.179134259
181	CHRNA2	Suprabasal IFE	11.577281	0.206144012	0.999553592
182	ANXA8L1	Suprabasal IFE	7.126025	0.060253742	0.999553592
183	CTD-2015H6.3	Suprabasal IFE	5.2263117	0.006258772	0.468047302
184	ZBTB7C	Suprabasal IFE	5.759904	0.030315752	0.999553592
185	PLLP	Suprabasal IFE	5.349284	0.014413663	0.892394058
186	PLD1	Suprabasal IFE	2.7524517	0.019934135	0.999553592
187	PCDH11X	Outer root sheath	6.08219	0.002172579	0.24912234
188	LIPG	Outer root sheath	6.0826683	0.088121629	0.999592492
189	RP11-128M1.1	Outer root sheath	5.327609	0.18113532	0.999592492
190	KRT6B	Outer root sheath	11.488702	0.166128606	0.999592492
191	RBM34	Outer root sheath	2.8846939	0.006548868	0.592844884
192	RP5-1050E16.2	Outer root sheath	6.7352524	0.085910324	0.999592492
193	HR	Outer root sheath	7.998205	0.001911427	0.243515186
194	PPP1R14C	Outer root sheath	3.4016469	0.027792462	0.999592492
195	TNK1	Outer root sheath	6.4242544	0.086883953	0.999592492
196	SYNPR	Outer root sheath	9.5519	0.083675665	0.999592492

197	FUT10	Outer root sheath	2.224069	0.049034588	0.999592492
198	PHF8	Outer root sheath	2.3247817	0.096810996	0.999592492
199	TRAPPC11	Outer root sheath	1.6749569	0.086777917	0.999592492
200	ZNF555	Outer root sheath	2.6516619	0.127825505	0.999592492
201	ZBTB21	Outer root sheath	1.7005777	0.079022215	0.999592492
202	POLI	Outer root sheath	1.6966475	0.089042428	0.999592492
203	HYDIN	Outer root sheath	4.133974	0.098393587	0.999592492
204	ZNF615	Outer root sheath	2.018049	0.162399791	0.999592492
205	C1orf106	Outer root sheath	5.0507164	0.018449199	0.999592492
206	ALOX15	Outer root sheath	5.1357455	0.17754717	0.999592492
207	ZNF136	Outer root sheath	1.7155558	0.076420837	0.999592492
208	ADGRA3	Outer root sheath	1.9966899	0.062004105	0.999592492
209	EXOC3	Outer root sheath	1.1549442	0.129817663	0.999592492
210	ZNF350	Outer root sheath	2.2712567	0.050641782	0.999592492
211	SLC25A14	Outer root sheath	1.9602681	0.060367153	0.999592492
212	BARX2	Outer root sheath	7.8520107	0.000567569	0.122027367
213	WBSCR27	Outer root sheath	2.6591516	0.13253469	0.999592492
214	SRSF12	Outer root sheath	2.9448364	0.070814889	0.999592492
215	UNC5B-AS1	Outer root sheath	6.474258	0.086953788	0.999592492
216	SPDEF	Companion layer	10.131556	0.096148805	0.999576503
217	LINC00675	Companion layer	10.609614	0.045738922	0.999576503
218	MUC6	Companion layer	10.586391	0.019752074	0.999576503
219	RPS6KA6	Companion layer	5.3733444	0.053055593	0.999576503
220	EPHX3	Companion layer	6.0954604	0.000320614	0.063043198
221	KRT75	Companion layer	12.526175	0.095742235	0.999576503
222	KCNK7	Companion layer	7.532109	0.020388607	0.999576503
223	CEL	Companion layer	7.486692	0.320756183	0.999576503
224	SMIM22	Companion layer	9.558874	0.046018365	0.999576503
225	MYO5B	Companion layer	6.7364063	0.047956269	0.999576503
226	FAM3B	Companion layer	6.909851	0.003038832	0.435565949
227	ACADSB	Companion layer	3.0539782	0.021468782	0.999576503
228	EPPK1	Companion layer	7.906335	0.097347221	0.999576503
229	CDC42BPG	Companion layer	6.500451	0.099954406	0.999576503
230	KLK14	Companion layer	8.675966	0.318397991	0.999576503
231	SLC6A11	Companion layer	6.2745047	0.101538538	0.999576503
232	TRPV6	Companion layer	9.236455	0.183227447	0.999576503
233	ADGRF4	Companion layer	8.993343	0.096324698	0.999576503
234	ZNF737	Companion layer	3.2393143	0.018444352	0.999576503
235	NECAB1	Companion layer	5.4233904	0.104350042	0.999576503
236	MYH14	Companion layer	6.8065734	0.000981836	0.168875779

237	NIPAL1	Companion layer	6.60174	0.008456378	0.969664661
238	FBXO2	Companion layer	4.46534	0.204558138	0.999576503
239	ZBTB24	Companion layer	2.712391	0.022871325	0.999576503
240	PCAT6	Companion layer	2.7576149	0.011919745	0.999576503
241	CCL27	Companion layer	5.9903126	0.049347134	0.999576503
242	KRT79	Companion layer	9.715588	0.317824506	0.999576503
243	TGM1	Companion layer	6.1280746	0.188511343	0.999576503
244	PDIK1L	Companion layer	2.7691107	0.085645498	0.999576503
245	DOCK3	Companion layer	8.302869	0.184192086	0.999576503
246	DIRAS1	Companion layer	3.8444266	0.063501298	0.999576503
247	ADPRM	Companion layer	2.211226	0.039341956	0.999576503
248	RP5-1159O4.1	Companion layer	3.6606522	0.207734224	0.999576503
249	SHH	Placode/matrix	11.996095	1.36E-106	2.34E-103
250	RHCG	Placode/matrix	6.8416953	5.62E-25	5.37E-23
251	CKMT2	Placode/matrix	7.1574445	4.66E-27	4.72E-25
252	PRKCG	Placode/matrix	6.9191985	3.46E-35	6.61E-33
253	TRPM3	Placode/matrix	5.9557695	1.82E-34	2.84E-32
254	LINGO1	Placode/matrix	3.7592468	2.43E-31	3.48E-29
255	WNT10A	Placode/matrix	6.868418	1.03E-49	2.53E-47
256	LINC00882	Placode/matrix	4.74487	2.98E-19	2.23E-17
257	BNC1	Placode/matrix	5.231955	3.70E-16	1.99E-14
258	FOXI3	Placode/matrix	8.573352	1.94E-12	8.34E-11
259	HCRTR1	Placode/matrix	8.691096	6.24E-07	1.49E-05
260	NOTUM	Placode/matrix	5.202044	9.16E-20	7.16E-18
261	CILP2	Placode/matrix	5.978341	8.67E-07	2.01E-05
262	SHROOM3	Placode/matrix	3.1104133	1.41E-20	1.16E-18
263	TLE2	Placode/matrix	3.1378155	1.30E-14	6.40E-13
264	CPA6	Placode/matrix	8.046429	1.93E-18	1.28E-16
265	PRDM5	Placode/matrix	2.1778307	2.43E-09	7.21E-08
266	NFE2L3	Placode/matrix	5.1005373	1.68E-16	9.31E-15
267	WNK2	Placode/matrix	3.9156203	3.62E-28	3.89E-26
268	PDZD2	Placode/matrix	2.7804046	2.15E-14	1.03E-12
269	ATP6V1C2	Placode/matrix	4.863823	4.67E-07	1.13E-05
270	ANKRD6	Placode/matrix	3.3939214	3.83E-14	1.78E-12
271	RDH16	Placode/matrix	7.9472194	2.03E-12	8.51E-11
272	AP001258.4	Placode/matrix	3.710443	4.62E-11	1.63E-09
273	GNAL	Placode/matrix	2.873853	1.66E-07	4.31E-06
274	MDGA1	Placode/matrix	4.229711	3.94E-09	1.15E-07
275	LOR	Placode/matrix	8.999555	7.60E-25	6.88E-23
276	SAMD12	Placode/matrix	5.6578794	1.18E-09	3.63E-08

277	STON1	Placode/matrix	2.1746337	5.25E-09	1.51E-07
278	KCTD18	Placode/matrix	1.8118894	3.54E-08	9.53E-07
279	RP11-245D16.4	Placode/matrix	4.204214	1.81E-09	5.47E-08
280	ADAMTSL2	Placode/matrix	4.8862953	6.44E-11	2.22E-09
281	SAMD11	Placode/matrix	4.577732	2.91E-12	1.16E-10
282	HUNK	Placode/matrix	2.201372	2.12E-12	8.67E-11
283	FJX1	Placode/matrix	3.4379036	1.73E-12	7.63E-11
284	HOXC13	Placode/matrix	9.153885	1.66E-18	1.14E-16
285	AMER1	Placode/matrix	3.230423	2.14E-05	0.000466288
286	PHYHIPL	Placode/matrix	4.4254794	6.10E-15	3.18E-13
287	ADAMTS17	Placode/matrix	4.275629	7.25E-19	5.19E-17
288	CPNE7	Placode/matrix	6.6334286	2.42E-07	6.22E-06
289	SLC25A48	Placode/matrix	4.224962	6.67E-11	2.23E-09
290	RNF183	Inner root sheath	9.309183	0.221631816	0.999654225
291	EPC2	Inner root sheath	2.5421717	0.117191652	0.999654225
292	CEP162	Inner root sheath	2.5697315	0.111926529	0.999654225
293	RP11-316M20.1	Inner root sheath	9.843094	0.414823324	0.999654225
294	RP11-479O9.4	Inner root sheath	4.1808457	0.248184348	0.999654225
295	AC034243.1	Inner root sheath	5.4975243	0.233192338	0.999654225
296	LMTK3	Inner root sheath	6.5462985	0.107649735	0.999654225
297	DLG3	Inner root sheath	3.7232125	0.136833509	0.999654225
298	SLC38A9	Inner root sheath	2.354401	0.111678429	0.999654225
299	TMEFF1	Inner root sheath	3.6865878	0.262097407	0.999654225
300	SCUBE2	Inner root sheath	3.6065676	0.027496531	0.999654225
301	LINC01152	Inner root sheath	5.5490274	0.23059172	0.999654225
302	COX18	Inner root sheath	2.8252962	0.081323295	0.999654225
303	AC092835.2	Inner root sheath	4.6106596	0.241276066	0.999654225
304	FANK1	Inner root sheath	5.043655	0.114495323	0.999654225
305	UGT3A2	Inner root sheath	3.8754513	0.024565359	0.999654225
306	ZNF790	Inner root sheath	2.972501	0.153403839	0.999654225
307	C1RL	Inner root sheath	2.5497782	0.097766856	0.999654225
308	FAM160B2	Inner root sheath	2.1263735	0.121401387	0.999654225
309	UNKL	Inner root sheath	2.574853	0.089040285	0.999654225
310	NFXL1	Inner root sheath	2.680377	0.03905792	0.999654225
311	MARVELD2	Inner root sheath	6.9348974	0.043133019	0.999654225
312	PAQR5	Inner root sheath	6.1728387	0.044643482	0.999654225
313	RP11-140K17.3	Cuticle/cortex	3.1126003	0.022114763	0.999626517
314	LINC00680	Cuticle/cortex	2.6354089	0.028228662	0.999626517
315	RP11-60A24.3	Cuticle/cortex	8.879263	0.258045808	0.999626517
316	ZFYVE9	Cuticle/cortex	2.8976824	0.059001926	0.999626517

317	BBS9	Cuticle/cortex	2.2514913	0.079608983	0.999626517
318	MINCR	Cuticle/cortex	2.6920602	0.027741207	0.999626517
319	ZNF549	Cuticle/cortex	3.421411	0.173789606	0.999626517
320	ZXDC	Cuticle/cortex	4.0210266	0.01412308	0.999626517
321	BRMS1L	Cuticle/cortex	1.9544355	0.098134183	0.999626517
322	USP30	Cuticle/cortex	2.534201	0.1272897	0.999626517
323	PPIP5K1	Cuticle/cortex	3.7618806	0.161917748	0.999626517
324	SLCO4A1-AS1	Melanocyte	12.043873	6.30E-218	7.23E-216
325	GJB1	Melanocyte	8.420761	9.25E-125	7.58E-123
326	QPCT	Melanocyte	6.342037	0	0
327	MC1R	Melanocyte	6.9257717	3.58E-130	3.08E-128
328	MCOLN3	Melanocyte	7.0108123	7.92E-151	7.17E-149
329	TYR	Melanocyte	11.225169	1.83E-236	2.25E-234
330	DCT	Melanocyte	12.985922	0	0
331	SLC6A17	Melanocyte	9.886178	3.29E-89	2.26E-87
332	GPR143	Melanocyte	8.205478	9.13E-294	1.43E-291
333	CAPN3	Melanocyte	9.17275	0	0
334	RETSAT	Melanocyte	4.206188	3.74E-89	2.47E-87
335	SLC24A5	Melanocyte	10.53669	3.02E-279	4.33E-277
336	PAMR1	Melanocyte	4.7290874	8.51E-59	4.31E-57
337	NSG1	Melanocyte	6.696851	0	0
338	RP4-529N6.1	Melanocyte	10.6294	2.47E-28	8.16E-27
339	C3orf79	Melanocyte	10.855184	2.19E-241	2.89E-239
340	TPCN2	Melanocyte	3.6413534	1.33E-65	7.17E-64
341	ZNF749	Melanocyte	4.046079	6.86E-40	2.88E-38
342	TYRP1	Melanocyte	12.573613	0	0
343	ROPN1	Melanocyte	9.09343	5.90E-12	1.25E-10
344	FMN1	Melanocyte	5.48432	9.57E-168	9.14E-166
345	CYP2U1	Melanocyte	3.9732797	3.09E-30	1.09E-28
346	DSTYK	Melanocyte	3.9088356	1.35E-109	9.65E-108
347	GAPDHS	Melanocyte	8.656339	6.80E-15	1.65E-13
348	RP11-1055B8.3	Melanocyte	7.7923813	8.58E-27	2.73E-25
349	CA14	Melanocyte	7.60927	2.63E-180	2.83E-178
350	SLC6A15	Melanocyte	7.392114	6.43E-121	5.02E-119
351	ALX1	Melanocyte	7.0734634	4.75E-79	2.92E-77
352	LL22NC03-N95F10.1	Melanocyte	7.7590322	5.30E-23	1.57E-21
353	FAM69C	Melanocyte	4.274584	3.66E-15	9.13E-14
354	CDKL2	Melanocyte	5.692367	5.77E-19	1.63E-17
355	FBXO32	Melanocyte	4.573367	1.28E-68	7.37E-67

356	NRCAM	Melanocyte	3.8509011	5.94E-18	1.62E-16
357	BAIAP2L2	Melanocyte	6.164767	2.74E-11	5.60E-10
358	NPM2	Melanocyte	5.922723	3.06E-25	9.40E-24
359	DIP2C	Melanocyte	2.3255606	6.85E-32	2.56E-30
360	RAB3B	Melanocyte	4.430366	1.45E-17	3.89E-16
361	GALNTL6	Melanocyte	5.3713536	8.94E-50	4.39E-48
362	LINC00518	Melanocyte	9.779316	2.59E-42	1.14E-40
363	IDI2-AS1	Melanocyte	7.4161425	7.13E-32	2.61E-30
364	PRKCE	Melanocyte	2.4247432	4.16E-19	1.19E-17
365	PMEL	Melanocyte	11.773891	0	0
366	MCHR1	Melanoblast	8.783223	2.85E-35	2.34E-33
367	RP11-615I2.2	Melanoblast	6.7648845	2.16E-48	2.32E-46
368	ROPN1B	Melanoblast	8.127593	2.13E-20	1.22E-18
369	LNP1	Melanoblast	5.286263	6.53E-45	6.61E-43
370	MLANA	Melanoblast	11.136707	2.58E-125	4.43E-122
371	KCNJ13	Melanoblast	8.891594	1.07E-15	5.42E-14
372	SLC45A2	Melanoblast	8.953339	7.23E-29	5.65E-27
373	S100A1	Melanoblast	4.8861017	1.54E-25	1.10E-23
374	CHST9	Melanoblast	5.797201	7.52E-22	4.46E-20
375	BIRC7	Melanoblast	5.7565775	4.04E-42	3.86E-40
376	LINC00462	Melanoblast	7.9914756	6.68E-22	4.10E-20
377	MLIP	Melanoblast	5.509444	9.15E-07	2.62E-05
378	TMEM215	Melanoblast	4.537467	1.64E-07	5.24E-06
379	GAPLINC	Melanoblast	3.9049802	3.72E-15	1.83E-13
380	CHCHD6	Melanoblast	4.995898	1.06E-106	2.60E-104
381	STK32A	Melanoblast	6.0629835	1.91E-14	8.89E-13
382	RTTN	Melanoblast	4.089126	3.03E-16	1.58E-14
383	ABCB5	Melanoblast	7.3827934	2.39E-22	1.52E-20
384	FTCDNL1	Melanoblast	4.5229774	6.95E-05	0.001616113
385	HSF4	Melanoblast	3.9716268	1.06E-06	2.94E-05
386	PNMA6A	Melanoblast	2.5432134	0.000651757	0.014190146
387	ST20	Melanoblast	2.4535794	9.74E-08	3.22E-06
388	MITF	Melanoblast	5.7852964	9.88E-116	3.40E-113
389	LGI3	Melanoblast	7.8440614	0.00028728	0.006417171
390	FAM53B	Melanoblast	2.3123689	2.68E-07	8.39E-06
391	GPR19	Melanoblast	5.091654	1.10E-05	0.000281639
392	NAT14	Melanoblast	1.9476153	1.53E-07	4.95E-06
393	CABLES1	Melanoblast	3.5531795	2.38E-24	1.64E-22
394	CCDC140	Melanoblast	6.6146183	2.40E-13	1.03E-11
395	NLGN1	Melanoblast	2.9820738	2.47E-12	1.02E-10

396	GYG2	Melanoblast	3.287917	1.17E-06	3.20E-05
397	LPL	Adipocytes	6.723894	0	0
398	ADAMTS19	Adipocytes	3.7155523	1.64E-259	1.41E-256
399	PQLC2L	Adipocytes	4.879482	1.70E-155	4.18E-153
400	FGF10	Adipocytes	3.5135021	1.01E-146	2.17E-144
401	PRICKLE2	Adipocytes	2.534813	1.11E-108	1.47E-106
402	PTPRT	Adipocytes	4.435348	2.88E-116	4.96E-114
403	HRASLS5	Adipocytes	4.639753	5.00E-47	3.74E-45
404	RP11-392O17.1	Adipocytes	4.5866704	3.19E-72	3.23E-70
405	CITED1	Adipocytes	4.280206	3.74E-52	3.07E-50
406	AADAC	Adipocytes	5.719056	3.07E-172	1.05E-169
407	SLC6A13	FRZB+ early fibroblast	11.471655	0	0
408	SLC1A3	FRZB+ early fibroblast	5.0368733	0	0
409	LRRC17	FRZB+ early fibroblast	3.9496737	0	0
410	SLC6A1	FRZB+ early fibroblast	8.533931	0	0
411	C16orf89	FRZB+ early fibroblast	8.62283	0	0
412	EPHA7	FRZB+ early fibroblast	5.623212	0	0
413	SLC7A11	FRZB+ early fibroblast	7.03022	0	0
414	ARHGAP20	FRZB+ early fibroblast	4.0266857	0	0
415	CYP1B1	FRZB+ early fibroblast	6.9909396	0	0
416	STRA6	FRZB+ early fibroblast	5.976237	0	0
417	COL13A1	FRZB+ early fibroblast	4.313859	0	0
418	RIMS1	FRZB+ early fibroblast	7.925822	0	0
419	RANBP3L	FRZB+ early fibroblast	7.5557947	0	0
420	LAMA1	FRZB+ early fibroblast	7.069982	0	0
421	LINC01139	FRZB+ early fibroblast	4.642066	0	0
422	KLC4	FRZB+ early fibroblast	3.2901027	0	0
423	SNED1	FRZB+ early fibroblast	6.5423117	0	0
424	KCNK2	FRZB+ early fibroblast	4.1698875	0	0
425	CXCL6	FRZB+ early fibroblast	6.2396874	0	0
426	NKX6-1	FRZB+ early fibroblast	8.488359	2.06E-176	6.69E-175
427	DACH2	FRZB+ early fibroblast	5.817186	0	0
428	SERPIND1	FRZB+ early fibroblast	8.21448	3.71E-167	1.18E-165
429	SLC22A6	FRZB+ early fibroblast	10.124752	8.49E-167	2.65E-165
430	ZIC1	FRZB+ early fibroblast	4.7879214	0	0
431	LUZP2	FRZB+ early fibroblast	4.2390356	1.52E-286	5.95E-285
432	INSRR	FRZB+ early fibroblast	8.655241	1.04E-129	2.79E-128
433	DKKL1	FRZB+ early fibroblast	3.509274	0	0
434	PCBP3	FRZB+ early fibroblast	4.1587195	1.92E-216	6.89E-215
435	SLC6A20	FRZB+ early fibroblast	7.9687667	1.77E-104	4.23E-103

436	EPHA3	FRZB+ early fibroblast	4.594174	0	0
437	SLC6A12	FRZB+ early fibroblast	6.88468	7.11E-165	2.14E-163
438	RDH10	FRZB+ early fibroblast	3.2194858	0	0
439	RP11-834C11.4	HOXC5+ early fibroblast	4.27087	0	0
440	ISM1	HOXC5+ early fibroblast	4.777561	0	0
441	HAND1	HOXC5+ early fibroblast	6.4379816	0	0
442	MYO3B	HOXC5+ early fibroblast	4.6762342	0	0
443	DPY19L1	HOXC5+ early fibroblast	3.1691134	0	0
444	ADGRV1	HOXC5+ early fibroblast	4.518062	0	0
445	HOXC5	HOXC5+ early fibroblast	5.2288475	0	0
446	RP4-799D16.1	HOXC5+ early fibroblast	6.9945016	0	0
447	HOXB3	HOXC5+ early fibroblast	3.709301	0	0
448	IRX5	HOXC5+ early fibroblast	3.5723877	0	0
449	C1QTNF7	HOXC5+ early fibroblast	2.7214336	0	0
450	NBL1	HOXC5+ early fibroblast	2.4382265	0	0
451	IRX1	HOXC5+ early fibroblast	2.8611274	0	0
452	MICU3	HOXC5+ early fibroblast	1.8797499	2.74E-279	1.85E-277
453	ITGA8	HOXC5+ early fibroblast	3.0125685	0	0
454	PAX1	HOXC5+ early fibroblast	2.867488	3.37E-178	1.57E-176
455	CACNA1G	HOXC5+ early fibroblast	2.090393	2.53E-288	1.89E-286
456	ATRNL1	HOXC5+ early fibroblast	2.7333574	9.80E-292	7.66E-290
457	LINC01305	HOXC5+ early fibroblast	3.181901	2.80E-279	1.85E-277
458	VIPR2	HOXC5+ early fibroblast	2.5469134	0	0
459	HOXC4	HOXC5+ early fibroblast	2.8359675	5.82E-194	3.13E-192
460	WNT2	WNT2+ fibroblast	6.2262545	0	0
461	COPZ2	WNT2+ fibroblast	2.6785483	0	0
462	SFRP4	WNT2+ fibroblast	4.343009	0	0
463	ACKR4	WNT2+ fibroblast	3.7746124	0	0
464	CILP	WNT2+ fibroblast	4.918565	0	0
465	ADAMTS2	WNT2+ fibroblast	2.363721	0	0
466	P3H3	WNT2+ fibroblast	2.1270008	0	0
467	CPB1	WNT2+ fibroblast	4.868908	0	0
468	HTRA3	WNT2+ fibroblast	2.8553362	0	0
469	SCARA5	WNT2+ fibroblast	3.6503012	0	0
470	WBSCR17	WNT2+ fibroblast	4.00864	0	0
471	PI16	PEAR1+ fibroblast	3.6043487	0	0
472	ABCA10	PEAR1+ fibroblast	3.7314963	6.92E-127	5.96E-124
473	GRP	Pre-dermal condensate	4.5573907	0	0
474	CORIN	Pre-dermal condensate	3.4277306	0	0
475	ADCY1	Pre-dermal condensate	2.9057283	2.84E-306	3.05E-304

476	CUBN	Pre-dermal condensate	3.7782712	0	0
477	SULT1B1	Pre-dermal condensate	3.8507204	0	0
478	IRF9	Pre-dermal condensate	2.2922766	0	0
479	PRR16	Pre-dermal condensate	1.9292406	3.06E-293	2.92E-291
480	SLC26A7	Dermal condensate	8.380465	0	0
481	ASIP	Dermal condensate	3.5913036	0	0
482	CPZ	Dermal condensate	4.510842	0	0
483	TMEM171	Dermal condensate	5.836547	0	0
484	CFHR1	Dermal condensate	5.9561896	1.66E-236	3.17E-234
485	RP11-60A8.1	Dermal condensate	5.9056497	0	0
486	NGEF	Dermal condensate	4.79619	4.93E-117	7.70E-115
487	RP11-286B14.1	Dermal condensate	5.5204086	1.33E-122	2.29E-120
488	CRYM	Dermal papilla	9.047038	0	0
489	RSPO3	Dermal papilla	6.894936	0	0
490	PAPPA2	Dermal papilla	5.994298	0	0
491	SPON1	Dermal papilla	5.224045	0	0
492	NPS	Dermal papilla	13.207621	0	0
493	NDP	Dermal papilla	7.836783	0	0
494	HHIP	Dermal papilla	6.5602727	0	0
495	KCNN2	Dermal papilla	6.4301476	0	0
496	ALPL	Dermal papilla	7.303497	0	0
497	RP11-138M12.1	Dermal papilla	8.801816	1.14E-143	8.14E-142
498	PPEF1	Dermal papilla	6.832318	2.32E-206	1.90E-204
499	ARL9	Dermal papilla	6.3415017	2.50E-222	2.27E-220
500	KIF26B	Dermal papilla	4.0568275	0	0
501	CDKL5	Dermal papilla	3.4521499	3.52E-260	4.04E-258
502	ALX4	Dermal papilla	5.8673506	0	0
503	RP11-680F8.1	Dermal papilla	5.537302	2.04E-262	2.51E-260
504	F2RL2	Dermal papilla	3.7136223	1.57E-204	1.23E-202
505	KCNK4	Dermal papilla	8.2718	4.27E-99	2.53E-97
506	LRRTM2	Dermal papilla	5.674571	1.31E-137	8.64E-136
507	DAND5	Dermal papilla	7.097033	1.59E-68	7.58E-67
508	ANO2	Dermal papilla	5.445151	1.24E-89	6.48E-88
509	CNFN	Dermal papilla	4.435768	5.07E-260	5.45E-258
510	RP11-735G4.1	Dermal papilla	7.597431	8.28E-57	3.56E-55
511	DIO3OS	Dermal papilla	3.262175	2.20E-107	1.40E-105
512	TMEFF2	Myofibroblasts	4.452043	0	0
513	ADAMTSL3	Myofibroblasts	3.9598486	0	0
514	SCN7A	Myofibroblasts	6.244546	0	0
515	ADH1B	Myofibroblasts	4.634799	4.27E-254	9.19E-252

516	NEGR1	Myofibroblasts	3.9789371	1.55E-251	2.97E-249
517	KCNMB2	Myofibroblasts	4.543457	2.10E-134	2.13E-132
518	ALDH1A3	Myofibroblasts	3.728417	3.60E-160	4.76E-158
519	SRPX2	Myofibroblasts	2.4888177	1.11E-150	1.19E-148
520	AGTR2	Myofibroblasts	4.848382	6.49E-103	5.31E-101
521	GSC	Myofibroblasts	4.1433434	4.91E-58	3.01E-56
522	TLL2	Myofibroblasts	5.986955	2.17E-58	1.38E-56
523	ANOS1	Myofibroblasts	3.303324	7.72E-156	8.86E-154
524	KERA	Myofibroblasts	3.17145	5.86E-78	4.20E-76
525	CNTNAP4	Myofibroblasts	3.0778816	1.57E-39	7.97E-38
526	C1QTNF3	Myoblasts	6.8257957	0	0
527	NPY	Myoblasts	9.455095	0	0
528	MYF5	Myoblasts	11.14899	0	0
529	PAX7	Myoblasts	9.631301	0	0
530	MYF6	Myoblasts	6.9552846	0	0
531	CALCR	Myoblasts	9.397291	3.36E-177	3.85E-175
532	TACR3	Myoblasts	7.027777	3.32E-67	2.28E-65
533	NTN4	Myoblasts	5.8203154	2.33E-234	3.64E-232
534	FGF16	Myoblasts	4.8831596	3.84E-60	2.36E-58
535	MYOG	Early myocytes	12.226757	0	0
536	RP1-302G2.5	Early myocytes	9.028112	3.41E-266	1.17E-263
537	SIRT2	Early myocytes	4.090881	2.32E-293	1.33E-290
538	FITM1	Early myocytes	8.571268	0	0
539	RGR	Early myocytes	10.103986	5.57E-153	8.71E-151
540	SCRIB	Early myocytes	5.8726025	5.60E-270	2.41E-267
541	TMEM8C	Early myocytes	11.963775	2.66E-246	7.63E-244
542	RP11-161M6.2	Early myocytes	5.1528244	2.49E-174	4.76E-172
543	EDC4	Early myocytes	4.777489	3.23E-100	2.78E-98
544	MYOD1	Early myocytes	7.8869486	1.74E-160	3.00E-158
545	KRT31	Early myocytes	10.306025	4.08E-78	2.92E-76
546	XXbac-B33L19.12	Early myocytes	8.670017	2.74E-139	3.37E-137
547	ENPP6	Early myocytes	6.1066394	1.75E-75	1.20E-73
548	CHRND	Early myocytes	8.537736	2.28E-147	3.02E-145
549	SOX8	Early myocytes	6.029268	4.90E-103	4.43E-101
550	FNDC5	Early myocytes	4.598112	2.87E-120	3.09E-118
551	RBM24	Early myocytes	6.3629017	1.75E-127	2.00E-125
552	SBK2	Early myocytes	8.27031	7.90E-39	3.58E-37
553	CTD-2545M3.8	Early myocytes	7.108774	2.79E-92	2.18E-90
554	DOK7	Early myocytes	7.393765	1.01E-50	5.77E-49

555	LINC00930	Early myocytes	9.611089	5.57E-37	2.46E-35
556	DMPK	Early myocytes	4.2337627	3.62E-111	3.46E-109
557	KRT39	Early myocytes	13.71028	9.73E-22	2.74E-20
558	VGLL2	Early myocytes	7.868106	8.53E-65	5.43E-63
559	GNB3	Early myocytes	5.396553	5.05E-25	1.64E-23
560	KRT40	Early myocytes	11.096172	7.13E-46	3.96E-44
561	OTOF	Early myocytes	6.9320703	1.05E-41	5.31E-40
562	NBPF20	Early myocytes	3.6130028	4.72E-27	1.66E-25
563	NEU4	Early myocytes	5.172804	9.56E-40	4.57E-38
564	MTERF2	Early myocytes	2.3285437	4.16E-33	1.70E-31
565	CELA2B	Early myocytes	10.504261	1.88E-11	3.85E-10
566	KCNF1	Early myocytes	6.5414286	5.06E-10	1.01E-08
567	MAMSTR	Early myocytes	2.8159733	8.65E-43	4.51E-41
568	CASP9	Early myocytes	3.1952784	4.12E-27	1.48E-25
569	MAP6D1	Early myocytes	6.212143	1.36E-22	3.96E-21
570	RP11-766F14.2	Early myocytes	7.603944	4.00E-13	9.17E-12
571	LMF1-AS1	Early myocytes	3.8460162	1.92E-12	4.19E-11
572	SLC25A34	Early myocytes	4.6889696	1.40E-13	3.29E-12
573	NEB	Early myocytes	7.339848	1.38E-182	2.96E-180
574	P2RX6	Early myocytes	5.714809	3.79E-15	9.44E-14
575	HS6ST2	Early myocytes	4.222504	7.46E-32	2.98E-30
576	CHRNG	Early myocytes	8.237532	5.07E-06	8.22E-05
577	ADGRA1	Early myocytes	7.6381874	6.79E-13	1.52E-11
578	STC2	Early myocytes	4.229174	1.25E-25	4.13E-24
579	GSTT2B	Early myocytes	3.0382395	4.78E-21	1.33E-19
580	KRTAP3-2	Early myocytes	13.210102	3.37E-06	5.58E-05
581	TRIM72	Early myocytes	5.295547	2.30E-15	5.83E-14
582	FAM212B	Early myocytes	3.3613362	5.87E-28	2.15E-26
583	PC	Early myocytes	3.6201394	1.02E-14	2.46E-13
584	NXPE3	Early myocytes	3.2692468	4.26E-31	1.59E-29
585	BVES	Early myocytes	4.3415523	1.17E-22	3.47E-21
586	SIX4	Early myocytes	4.875696	8.55E-18	2.23E-16
587	MYO16	Early myocytes	3.6424968	5.01E-12	1.06E-10
588	ESYT3	Early myocytes	5.372191	6.04E-15	1.49E-13
589	ZNF556	Early myocytes	5.3085337	8.69E-12	1.80E-10
590	RYR1	Early myocytes	3.6500237	1.45E-34	6.10E-33
591	MYLPF	Myocytes	10.337247	1.12E-33	1.25E-30
592	TNNC2	Myocytes	10.247019	2.21E-32	1.27E-29
593	CKM	Myocytes	9.762779	9.66E-29	2.77E-26
594	ACTA1	Myocytes	10.303062	1.46E-33	1.25E-30

595	MYL1	Myocytes	9.823687	2.82E-25	6.07E-23
596	TNNT3	Myocytes	7.613084	4.65E-27	1.14E-24
597	TNNI1	Myocytes	9.1856	4.87E-23	9.30E-21
598	TNNI2	Myocytes	8.414598	5.83E-31	2.51E-28
599	MYOZ2	Myocytes	9.65351	3.11E-18	4.86E-16
600	SLN	Myocytes	9.191896	3.05E-17	4.03E-15
601	COX6A2	Myocytes	9.34774	2.68E-17	3.85E-15
602	TNNC1	Myocytes	7.843628	4.97E-16	5.70E-14
603	MYL5	Myocytes	5.5391927	1.90E-11	1.82E-09
604	DES	Myocytes	8.186448	1.15E-29	3.97E-27
605	IL17B	Myocytes	7.8706913	1.69E-11	1.71E-09
606	ENO3	Myocytes	7.181029	1.36E-20	2.34E-18
607	SMPX	Myocytes	10.168578	6.30E-11	5.70E-09
608	APOBEC2	Myocytes	9.63795	6.35E-09	4.37E-07
609	CAV3	Myocytes	7.903769	3.80E-10	3.11E-08
610	MYH3	Myocytes	8.904871	3.53E-10	3.04E-08
611	MYBPH	Myocytes	9.754949	1.50E-09	1.10E-07
612	MYBPC1	Myocytes	9.588177	4.34E-06	0.000233368
613	CSRP3	Myocytes	9.329485	3.87E-07	2.30E-05
614	NMRK2	Myocytes	9.19525	1.33E-06	7.63E-05
615	HSPB3	Myocytes	7.515379	8.06E-09	5.33E-07
616	PGAM2	Myocytes	8.326417	4.57E-06	0.000238247
617	TCAP	Myocytes	7.9586225	4.37E-05	0.00201221
618	LDB3	Myocytes	7.1706595	1.59E-05	0.00080578
619	CACNG1	Myocytes	9.960569	4.23E-06	0.000233368
620	HFE2	Myocytes	9.853982	3.86E-05	0.001844866
621	MB	Myocytes	9.297532	0.000286613	0.011203961
622	ACTN2	Myocytes	7.9357734	2.70E-16	3.32E-14
623	SRPK3	Myocytes	7.4986157	0.00012015	0.004920439
624	MYL3	Myocytes	4.1950364	0.010325275	0.236792985
625	LINC01497	Myocytes	7.098419	0.00012506	0.005002414
626	MYL2	Myocytes	8.704901	0.016333274	0.364847165
627	MYOZ1	Myocytes	5.49881	0.002640458	0.081746788
628	SMYD1	Myocytes	9.354034	0.000704642	0.0257869
629	CASQ1	Myocytes	8.3550415	0.053649373	0.997122433
630	LINC01405	Myocytes	8.823363	0.029952216	0.620696531
631	CALML6	Myocytes	7.648813	0.000117019	0.00490911
632	CA3	Myocytes	7.916143	1.42E-07	8.72E-06
633	LINC00202-1	Myocytes	8.624578	0.00785141	0.199571599
634	TRDN	Myocytes	7.9553204	0.00800607	0.199571599

635	RP11-299L17.3	Myocytes	6.4883766	0.004246177	0.118758346
636	PRR32	Myocytes	9.105126	0.007798427	0.199571599
637	ANKRD1	Myocytes	8.025108	0.053914178	0.997122433
638	KLHL40	Myocytes	8.547795	0.003738992	0.107184428
639	SH3BGR	Myocytes	4.2235107	0.007856532	0.199571599
640	MIR133A1HG	Myocytes	9.003932	0.015556291	0.352063429
641	MRLN	Myocytes	8.851955	1.54E-09	1.10E-07
642	TMEM38A	Myocytes	6.4719167	0.004280824	0.118758346
643	ARHGAP36	Myocytes	9.076952	0.147174472	0.999417922
644	SYNPO2L	Myocytes	7.090438	0.016617669	0.366440916
645	LANCL1-AS1	Myocytes	9.63207	0.05281355	0.997122433
646	LMOD2	Myocytes	10.283176	0.146118944	0.999417922
647	CACNB1	Myocytes	4.169814	0.007076918	0.192966927
648	MYH8	Myocytes	10.141639	0.052733875	0.997122433
649	MURC	Myocytes	5.185056	0.110019102	0.999417922
650	RP11-446H18.5	Myocytes	8.435473	0.090587663	0.999417922
651	MAP3K7CL	Myocytes	5.2848477	0.000556284	0.02080019
652	ATP1B4	Myocytes	9.546837	0.146517747	0.999417922
653	DGCR6	Myocytes	3.690488	0.002365283	0.076760135
654	LMOD3	Myocytes	5.483297	0.034315338	0.694380952
655	EEF1A2	Myocytes			
656	ITGB1BP2	Myocytes	6.5966024	0.093543977	0.999417922
657	RP11-358H18.3	Myocytes	8.644925	0.146695743	0.999417922
658	C1orf105	Myocytes	7.7013535	0.008004035	0.199571599
659	PLCL1	Pericytes	4.0661883	9.79E-297	2.40E-294
660	CNN1	LMCD1+ mural cell	4.9479027	1.75E-117	4.31E-115
661	ACAN	LMCD1+ mural cell	5.4348793	9.86E-110	1.88E-107
662	ENOX1	LMCD1+ mural cell	2.9853654	1.01E-56	1.08E-54
663	TGFB3	LMCD1+ mural cell	3.0880616	6.01E-68	8.61E-66
664	COL25A1	LMCD1+ mural cell	4.353694	9.24E-110	1.88E-107
665	ITGA11	LMCD1+ mural cell	3.7546089	2.75E-63	3.37E-61
666	PLPPR4	LMCD1+ mural cell	4.806455	2.92E-29	2.09E-27
667	IGFN1	LMCD1+ mural cell	4.1578183	1.58E-43	1.43E-41
668	RGS5	PLN+ mural cell	7.386362	0	0
669	OR51E1	PLN+ mural cell	10.267225	7.67E-265	2.64E-262
670	PLN	PLN+ mural cell	8.930392	3.03E-181	5.21E-179
671	KCNAB1	PLN+ mural cell	4.6367655	1.51E-226	4.32E-224
672	AMPH	PLN+ mural cell	3.9967682	4.78E-182	9.14E-180
673	TBX2-AS1	PLN+ mural cell	4.989215	6.40E-175	9.18E-173
674	TBX2	PLN+ mural cell	5.0922956	2.47E-180	3.86E-178

675	MOCS1	PLN+ mural cell	4.3858204	6.61E-114	6.69E-112
676	CASQ2	PLN+ mural cell	8.864628	2.87E-196	6.18E-194
677	KCNMB1	PLN+ mural cell	4.386083	1.78E-86	1.45E-84
678	NRIP2	PLN+ mural cell	5.6888866	4.51E-39	2.28E-37
679	CNNM2	PLN+ mural cell	3.2661245	1.20E-55	7.35E-54
680	OR51E2	PLN+ mural cell	10.123513	1.53E-26	5.48E-25
681	MED14	PLN+ mural cell	2.1427264	1.03E-56	6.57E-55
682	CTNNA3	PLN+ mural cell	5.280877	3.46E-93	2.97E-91
683	MYH11	PLN+ mural cell	6.6174655	0	0
684	PLCE1	PLN+ mural cell	2.8613224	6.87E-81	5.13E-79
685	SLC51A	PLN+ mural cell	5.2945313	2.46E-34	1.15E-32
686	HOXB-AS1	PLN+ mural cell	3.141791	2.46E-93	2.23E-91
687	MAMDC2	PLN+ mural cell	4.4212265	8.60E-67	5.92E-65
688	LINC00672	PLN+ mural cell	4.381844	7.72E-131	9.48E-129
689	MYOCD	PLN+ mural cell	5.9570785	3.22E-18	8.78E-17
690	GPR20	PLN+ mural cell	4.0559826	1.52E-19	4.28E-18
691	RP11-865I6.2	PLN+ mural cell	5.019987	1.05E-46	5.48E-45
692	EFNB3	Neuroendocrine	5.264022	2.10E-91	1.80E-88
693	SOST	Neuroendocrine	8.605207	1.14E-84	6.56E-82
694	LMX1A	Neuroendocrine	7.922091	1.99E-67	5.71E-65
695	WNT1	Neuroendocrine	11.24843	7.75E-57	1.67E-54
696	PCSK1N	Neuroendocrine	9.210689	1.81E-136	3.12E-133
697	WSCD2	Neuroendocrine	6.435518	9.71E-46	1.28E-43
698	MSX1	Neuroendocrine	5.5403724	7.90E-59	1.94E-56
699	NEFL	Neuroendocrine	7.8948894	7.14E-54	1.12E-51
700	RFX4	Neuroendocrine	9.11778	2.21E-69	7.62E-67
701	WT1	Neuroendocrine	9.731387	8.58E-27	5.47E-25
702	KCNK1	Neuroendocrine	6.21101	7.93E-31	7.18E-29
703	CLEC18B	Neuroendocrine	10.647538	2.35E-30	1.93E-28
704	GDPD2	Neuroendocrine	9.31269	8.35E-31	7.18E-29
705	LINC00472	Neuroendocrine	6.2755327	9.09E-55	1.74E-52
706	SLIT1	Neuroendocrine	7.2524104	5.20E-38	6.39E-36
707	RSPO1	Neuroendocrine	4.2338777	1.43E-35	1.45E-33
708	GDF7	Neuroendocrine	9.365177	7.16E-17	3.16E-15
709	CFAP126	Neuroendocrine	11.444835	8.33E-16	3.58E-14
710	FAM181A	Neuroendocrine	9.051712	3.23E-19	1.55E-17
711	GRIN2A	Neuroendocrine	9.289579	2.11E-11	5.94E-10
712	HOXA5	Neuroendocrine	5.467428	2.26E-27	1.49E-25
713	C5orf49	Neuroendocrine	5.4957385	5.27E-14	1.78E-12
714	RP11-159K7.2	Neuroendocrine	8.505207	6.61E-10	1.56E-08

715	CRB2	Neuroendocrine	9.615373	4.71E-15	1.84E-13
716	MIR9-3HG	Neuroendocrine	5.447925	2.44E-12	7.23E-11
717	LRRTM1	Neuroendocrine	4.6563854	2.35E-12	7.09E-11
718	ERBB4	Neuroendocrine	7.7896857	1.37E-14	4.91E-13
719	NELL1	Neuroendocrine	6.435572	2.81E-10	7.00E-09
720	DNALI1	Neuroendocrine	4.1016564	3.38E-30	2.64E-28
721	C6orf118	Neuroendocrine	6.610608	6.40E-08	1.27E-06
722	RIPPLY3	Neuroendocrine	6.1914277	1.29E-07	2.52E-06
723	MASP1	Neuroendocrine	3.6381764	4.74E-20	2.40E-18
724	NUAK2	Neuroendocrine	5.292714	6.12E-30	4.58E-28
725	NEFM	Neuroendocrine	6.613922	2.81E-36	3.02E-34
726	PAX6	Neuroendocrine	7.0116744	3.43E-09	7.47E-08
727	GABRA5	Neuroendocrine	6.27784	3.94E-07	6.84E-06
728	FOXJ1	Neuroendocrine	5.3264637	5.72E-24	3.39E-22
729	SERTM1	Neuroendocrine	6.261303	3.78E-07	6.63E-06
730	C22orf15	Neuroendocrine	5.1287203	6.49E-10	1.55E-08
731	C1orf194	Neuroendocrine	9.609568	1.46E-07	2.81E-06
732	RP11-649A16.1	Neuroendocrine	5.3416624	2.54E-13	8.24E-12
733	MORN5	Neuroendocrine			
734	WT1-AS	Neuroendocrine	9.003304	2.35E-09	5.24E-08
735	CCDC74B	Neuroendocrine	4.244421	5.32E-15	2.03E-13
736	GRM8	Neuroendocrine			
737	CLEC18A	Neuroendocrine	6.5426526	5.91E-07	1.01E-05
738	ENKUR	Neuroendocrine	6.897196	8.71E-10	2.00E-08
739	RASL10B	Neuroendocrine	5.103434	1.06E-06	1.73E-05
740	LRRN2	Neuroendocrine	5.32202	9.18E-11	2.43E-09
741	MSI1	Neuroendocrine	5.6047354	6.54E-20	3.21E-18
742	LINC00907	Neuroendocrine			
743	PTPRO	Neuroendocrine	4.3805947	3.55E-14	1.22E-12
744	DPP10	Neuroendocrine	5.508889	1.01E-08	2.12E-07
745	LRRIQ1	Neuroendocrine	3.8533332	1.01E-11	2.90E-10
746	ILDR2	Neuroendocrine	6.061857	8.05E-11	2.16E-09
747	MPP5	Neuroendocrine	2.5852349	1.14E-12	3.50E-11
748	ASTN1	Neuroendocrine			
749	RP11-96L14.8	Neuroendocrine			
750	RP11-387A1.5	Neuroendocrine			
751	PPIL6	Neuroendocrine	4.3034196	3.09E-09	6.81E-08
752	CCDC74A	Neuroendocrine	2.8900182	1.68E-07	3.17E-06
753	ALDH1A2	Neuroendocrine	3.9790092	2.23E-11	6.18E-10
754	RNF182	Neuroendocrine	6.303894	1.11E-14	4.13E-13

755	KIRREL2	Neuroendocrine	5.8744826	2.51E-07	4.54E-06
756	BAIAP3	Neuroendocrine			
757	LRP1B	Neuroendocrine	5.2490816	1.46E-13	4.83E-12
758	CCDC181	Neuroendocrine	3.4908235	2.67E-08	5.46E-07
759	IQCG	Neuroendocrine			
760	MAPK15	Neuroendocrine	6.6524887	6.13E-08	1.23E-06
761	MIMT1	Neuroendocrine	4.999209	6.22E-07	1.05E-05
762	FSIP1	Neuroendocrine			
763	ZCCHC18	Neuroendocrine			
764	SDK2	Neuroendocrine	3.7518017	4.41E-13	1.41E-11
765	CTNNA2	Neuroendocrine	6.7735653	1.36E-10	3.48E-09
766	HOXA-AS2	Neuroendocrine	4.5592866	1.32E-10	3.44E-09
767	LINC00461	Neuroendocrine	7.363261	1.51E-25	9.30E-24
768	MPPED2	Neuroendocrine	2.6852906	6.08E-13	1.90E-11
769	RND2	Neuroendocrine	4.61416	2.36E-15	9.43E-14
770	TNFRSF11B	Neuroendocrine	5.21648	5.68E-17	2.57E-15
771	POU3F4	Neuroendocrine			
772	GRIA1	Neuroendocrine			
773	RP11-384F7.2	Neuroendocrine	6.242899	5.68E-10	1.38E-08
774	TOX3	Neuroendocrine	4.8088017	2.41E-07	4.46E-06
775	VSTM2L	Neuroendocrine	4.400357	6.88E-07	1.15E-05
776	STMN2	Neuron progenitors	9.8962755	1.47E-258	3.61E-256
777	PPP1R17	Neuron progenitors	10.951153	4.84E-266	1.39E-263
778	TAGLN3	Neuron progenitors	9.825979	0	0
779	NEUROD1	Neuron progenitors	11.698586	1.90E-303	6.53E-301
780	DCX	Neuron progenitors	7.573279	0	0
781	RTN1	Neuron progenitors	6.254555	1.53E-165	1.75E-163
782	ELAVL4	Neuron progenitors	8.331352	0	0
783	ISL1	Neuron progenitors	10.032765	2.09E-151	2.25E-149
784	ELAVL3	Neuron progenitors	9.805101	1.13E-219	2.16E-217
785	MLLT11	Neuron progenitors	4.21388	0	0
786	NHLH1	Neuron progenitors	9.957776	2.88E-177	3.53E-175
787	SRRM4	Neuron progenitors	10.295021	1.77E-195	2.76E-193
788	NEUROD2	Neuron progenitors	10.960141	2.85E-108	1.96E-106
789	NEUROD6	Neuron progenitors	10.311495	9.59E-49	2.46E-47
790	RAB3A	Neuron progenitors	3.9812376	9.57E-134	9.14E-132
791	RPRM	Neuron progenitors	5.554044	1.65E-100	1.02E-98
792	INA	Neuron progenitors	9.781502	3.29E-94	1.77E-92
793	GNG3	Neuron progenitors	8.3380165	1.39E-118	1.09E-116
794	NHLH2	Neuron progenitors	10.683877	2.03E-94	1.13E-92

795	HPCA	Neuron progenitors	6.1146073	3.62E-121	3.11E-119
796	GAP43	Neuron progenitors	5.1330805	4.59E-222	9.86E-220
797	KIF5C	Neuron progenitors	6.988046	9.46E-98	5.61E-96
798	ATP1A3	Neuron progenitors	8.894436	4.16E-117	3.11E-115
799	TLX3	Neuron progenitors	11.835845	1.59E-65	6.23E-64
800	PRPH	Neuron progenitors	8.969972	2.14E-53	6.14E-52
801	FAM57B	Neuron progenitors	8.650248	1.07E-92	5.58E-91
802	BTBD17	Neuron progenitors	8.684577	3.71E-103	2.36E-101
803	ELAVL2	Neuron progenitors	6.5343847	9.39E-92	4.75E-90
804	OLFM1	Neuron progenitors	4.6768365	3.53E-107	2.33E-105
805	RBFOX1	Neuron progenitors	6.5003405	6.01E-131	5.44E-129
806	ATCAY	Neuron progenitors	8.268711	1.68E-63	6.14E-62
807	THSD7B	Neuron progenitors	7.0579286	1.06E-71	4.55E-70
808	STMN4	Neuron progenitors	10.323136	9.12E-97	5.23E-95
809	SUSD2	Neuron progenitors	7.666305	4.42E-69	1.81E-67
810	KIF1A	Neuron progenitors	6.7267756	6.84E-109	4.90E-107
811	GDAP1L1	Neuron progenitors	8.420856	6.03E-85	2.96E-83
812	CHRNA3	Neuron progenitors	8.740883	2.51E-57	8.00E-56
813	SSTR2	Neuron progenitors	6.03716	6.87E-179	9.09E-177
814	CELF3	Neuron progenitors	8.554238	3.96E-64	1.51E-62
815	NEUROD4	Neuron progenitors	10.513488	4.15E-58	1.35E-56
816	GNAO1	Neuron progenitors	5.7434726	3.33E-56	1.02E-54
817	ACTL6B	Neuron progenitors	10.778753	8.87E-51	2.42E-49
818	MIAT	Neuron progenitors	4.150545	5.11E-76	2.31E-74
819	NPPA	Neuron progenitors	9.193224	1.01E-68	4.05E-67
820	CHGB	Neuron progenitors	6.7316737	2.32E-51	6.55E-50
821	KLHL35	Neuron progenitors	5.2939863	6.29E-34	1.35E-32
822	HMP19	Neuron progenitors	10.406486	7.05E-40	1.71E-38
823	PLPPR1	Neuron progenitors	5.2406235	1.53E-60	5.27E-59
824	CELF4	Neuron progenitors	5.1802197	1.26E-61	4.44E-60
825	INSM1	Neuron progenitors	8.9330635	8.66E-44	2.19E-42
826	EYA2	Neuron progenitors	6.5867963	5.77E-80	2.68E-78
827	KLHDC8A	Neuron progenitors	6.678169	3.06E-51	8.48E-50
828	CNTN2	Neuron progenitors	6.6861086	4.00E-57	1.25E-55
829	TLX2	Neuron progenitors	10.772753	2.65E-24	5.00E-23
830	CAMKV	Neuron progenitors	8.262839	4.36E-22	7.65E-21
831	DRGX	Neuron progenitors	8.667975	5.64E-28	1.13E-26
832	AC091878.1	Neuron progenitors	6.925561	4.07E-39	9.58E-38
833	SEZ6	Neuron progenitors	6.6636066	1.03E-26	2.01E-25
834	CDH7	Neuron progenitors	4.3886714	9.17E-55	2.77E-53

835	CELF5	Neuron progenitors	7.323848	4.85E-23	8.79E-22
836	SNAP25	Neuron progenitors	4.01512	8.89E-54	2.64E-52
837	TTC9B	Neuron progenitors	8.983475	2.50E-27	4.94E-26
838	MYT1L	Neuron progenitors			
839	CHST8	Neuron progenitors	8.465033	1.28E-53	3.73E-52
840	DLL3	Neuron progenitors	4.2283463	5.43E-50	1.46E-48
841	REM2	Neuron progenitors	6.8017955	6.85E-39	1.59E-37
842	DLG4	Neuron progenitors	3.1966858	3.27E-60	1.10E-58
843	P2RX3	Neuron progenitors	6.4458	3.53E-24	6.61E-23
844	SYP	Neuron progenitors	5.32047	2.04E-120	1.67E-118
845	TBR1	Neuron progenitors			
846	SMIM18	Neuron progenitors			
847	CEND1	Neuron progenitors	7.202552	5.62E-24	1.04E-22
848	POU4F1	Neuron progenitors	6.426981	3.55E-59	1.18E-57
849	MAST1	Neuron progenitors	6.557081	8.57E-35	1.87E-33
850	NEUROG1	Neuron progenitors	8.934353	1.75E-73	7.71E-72
851	SYT4	Neuron progenitors	7.466494	1.78E-26	3.44E-25
852	PAK7	Neuron progenitors	5.133634	6.01E-22	1.04E-20
853	ATP6V1G2	Neuron progenitors	4.184562	3.21E-28	6.65E-27
854	LINC01551	Neuron progenitors	7.7901535	1.01E-41	2.48E-40
855	CLVS1	Neuron progenitors	7.8398128	5.75E-49	1.50E-47
856	SCRT2	Neuron progenitors			
857	RAB26	Neuron progenitors	7.7090387	1.34E-22	2.41E-21
858	LRRC53	Neuron progenitors			
859	KCNJ6	Neuron progenitors			
860	PHF21B	Neuron progenitors	5.8920546	3.65E-21	6.22E-20
861	FOXG1	Neuron progenitors	8.83988	4.73E-33	1.00E-31
862	NCAN	Neuron progenitors	8.072658	4.30E-20	7.12E-19
863	GDAP1	Neuron progenitors	3.805459	5.92E-36	1.32E-34
864	SPTBN4	Neuron progenitors			
865	CTD-2314G24.2	Neuron progenitors			
866	XKR7	Neuron progenitors			
867	Sep-03	Neuron progenitors	6.081374	1.42E-36	3.26E-35
868	HP09025	Neuron progenitors			
869	SLC17A7	Neuron progenitors			
870	AC004158.3	Neuron progenitors			
871	RCOR2	Neuron progenitors	4.55248	5.11E-36	1.16E-34
872	CACNA1A	Neuron progenitors			
873	KIF5A	Neuron progenitors			
874	NTN5	Neuron progenitors	7.240624	1.13E-63	4.22E-62

875	CHRNB4	Neuron progenitors	5.9722013	1.45E-20	2.42E-19
876	RPS6KL1	Neuron progenitors			
877	FRMPD1	Neuron progenitors			
878	DUSP26	Neuron progenitors			
879	FAIM2	Neuron progenitors			
880	IGDCC3	Neuron progenitors			
881	JAKMIP2	Neuron progenitors			
882	B3GAT1	Neuron progenitors			
883	NFASC	Neuron progenitors			
884	LINC00599	Neuron progenitors			
885	FEZF2	Neuron progenitors			
886	DPYSL5	Neuron progenitors			
887	HCN1	Neuron progenitors			
888	NOL4	Neuron progenitors			
889	CDK5R1	Neuron progenitors			
890	KLC2	Neuron progenitors			
891	PLCD4	Neuron progenitors			
892	SLC17A6	Neuron progenitors			
893	MAP6	Neuron progenitors			
894	PTCHD2	Neuron progenitors			
895	NPTX1	Neuron progenitors			
896	MAPK8IP1	Neuron progenitors			
897	SDK1	Neuron progenitors			
898	DCC	Neuron progenitors			
899	FMN2	Neuron progenitors			
900	PEX5L	Neuron progenitors			
901	ASNS	Neuron progenitors			
902	LHX9	Neuron progenitors			
903	CCNJL	Neuron progenitors			
904	ASPHD1	Neuron progenitors			
905	SVOP	Neuron progenitors			
906	APC2	Neuron progenitors			
907	RP11-247C2.2	Neuron progenitors			
908	LRRC49	Neuron progenitors			
909	SBK1	Neuron progenitors			
910	KCNH6	Neuron progenitors			
911	SRRM3	Neuron progenitors			
912	TIAM2	Neuron progenitors			
913	SYT5	Neuron progenitors			
914	BRSK2	Neuron progenitors			

915	SERP2	Neuron progenitors			
916	KCNB2	Neuron progenitors			
917	CADPS	Neuron progenitors			
918	JPH4	Neuron progenitors			
919	TMEM35	Neuron progenitors			
920	SH3GL2	Neuron progenitors			
921	RNF165	Neuron progenitors			
922	TMEM169	Neuron progenitors			
923	PRDM8	Neuron progenitors			
924	NEUROG2	Neuron progenitors			
925	DNAJB5	Neuron progenitors			
926	ISLR2	Neuron progenitors			
927	CECR2	Neuron progenitors			
928	TMEM59L	Neuron progenitors			
929	RIMS3	Neuron progenitors			
930	PNMA2	Neuron progenitors			
931	DPF1	Neuron progenitors			
932	C1orf61	Neuron progenitors			
933	PRDM12	Neuron progenitors			
934	TRHDE	Neuron progenitors			
935	DRAXIN	Neuron progenitors			
936	ABCC2	SPP1+ proliferating neuron progenitors	8.882806	4.70E-74	1.35E-71
937	GJC3	SPP1+ proliferating neuron progenitors	6.6883497	2.98E-34	2.84E-32
938	COL9A1	SPP1+ proliferating neuron progenitors	4.488566	2.34E-60	4.03E-58
939	FAM64A	SPP1+ proliferating neuron progenitors	3.198473	2.11E-41	2.59E-39
940	TKTL1	SPP1+ proliferating neuron progenitors	5.1046824	1.13E-69	2.42E-67
941	PRSS56	SPP1+ proliferating neuron progenitors	6.5629387	1.80E-15	1.03E-13
942	ADGRG2	SPP1+ proliferating neuron progenitors	3.495936	8.92E-10	3.34E-08
943	SCRG1	PID1+ schwann cellls	7.5873876	2.22E-70	3.83E-67
944	KCNC2	PID1+ schwann cellls	11.535546	1.52E-11	3.27E-09
945	LINC00237	PID1+ schwann cellls	10.498134	2.89E-06	0.00022569
946	NKX2-2	PID1+ schwann cellls	11.715125	7.17E-07	6.86E-05
947	ITIH6	PID1+ schwann cellls	10.319773	1.07E-05	0.000765676

340

948	RP11-472N13.3	PID1+ schwann cellls	5.7586203	2.93E-15	8.40E-13
949	POU3F1	PID1+ schwann cellls	7.6949167	5.49E-11	9.44E-09
950	ADAMTSL1	PID1+ schwann cellls	3.0983114	2.56E-09	3.39E-07
951	ATP10B	PID1+ schwann cellls	8.667722	1.83E-08	2.25E-06
952	GINS3	PID1+ schwann cellls	3.643901	1.40E-06	0.000114867
953	COL20A1	PID1+ schwann cellls	7.387785	6.62E-67	5.70E-64
954	RBMS3-AS3	PID1+ schwann cellls	5.3298297	1.96E-10	3.06E-08
955	TIMP4	PID1+ schwann cellls	3.6599178	2.99E-05	0.002053819
956	AZGP1	PID1+ schwann cellls	7.008661	9.37E-07	8.48E-05
957	GFRA1	Schwann/Schwann precursors	5.650927	0	0
958	DGKB	Schwann/Schwann precursors	7.194529	0	0
959	RIMS2	Schwann/Schwann precursors	5.5077634	1.98E-120	1.55E-118
960	RP11-420N3.2	Schwann/Schwann precursors	5.5005407	1.90E-131	1.82E-129
961	PKHD1	Schwann/Schwann precursors	8.124764	6.27E-35	2.84E-33
962	FP325317.1	Schwann/Schwann precursors	8.129267	2.24E-125	1.91E-123
963	IL1RAPL2	Schwann/Schwann precursors	5.762308	4.88E-51	2.89E-49
964	COL28A1	Schwann/Schwann precursors	7.2824597	9.14E-127	8.27E-125
965	ART3	Schwann/Schwann precursors	6.7515616	1.45E-136	1.66E-134
966	SEMA3B	Myelinating Schwann cells	6.12836	0	0
967	GFRA3	Myelinating Schwann cells	8.634487	0	0
968	XKR4	Myelinating Schwann cells	7.475747	5.14E-246	1.77E-243
969	DHH	Myelinating Schwann cells	8.713778	3.95E-129	8.50E-127
970	SORCS1	Myelinating Schwann cells	5.972452	1.08E-64	8.42E-63
971	AATK	Myelinating Schwann cells	5.301558	6.95E-93	9.19E-91
972	NTRK3	Myelinating Schwann cells	4.11876	7.90E-97	1.23E-94
973	LGI1	Myelinating Schwann cells	4.116931	3.55E-26	1.39E-24
974	CHL1	Myelinating Schwann cells	7.0613003	1.07E-86	1.31E-84
975	MAG	Myelinating Schwann cells	6.4723854	3.46E-31	1.80E-29
976	AC108142.1	Myelinating Schwann cells	5.2271028	1.11E-17	3.88E-16
977	TMPRSS5	Myelinating Schwann cells	6.1796904	3.30E-47	2.10E-45
978	CLDN19	Myelinating Schwann cells	6.8370275	2.19E-41	1.22E-39
979	ADGRB3	Myelinating Schwann cells	3.7499418	6.54E-29	2.96E-27
980	FSTL5	Myelinating Schwann cells	4.08517	1.24E-85	1.42E-83
981	SLC5A7	Myelinating Schwann cells	7.402215	1.03E-15	3.46E-14
982	CLSTN2	Myelinating Schwann cells	5.342893	8.75E-31	4.42E-29
983	HSPA12A	Myelinating Schwann cells	3.814108	3.48E-69	3.33E-67
984	IFIT2	Myelinating Schwann cells	3.2289913	3.01E-51	1.99E-49
985	DLX1	Myelinating Schwann cells	7.3504086	2.89E-30	1.42E-28
986	PPP1R1C	Myelinating Schwann cells	6.670739	5.95E-65	4.88E-63
987	GRB14	Myelinating Schwann cells	5.9382644	3.43E-27	1.44E-25

988	CCDC184	Myelinating Schwann cells	3.8278823	2.89E-95	4.14E-93
989	RP11-357H14.17	Myelinating Schwann cells	7.6721263	1.88E-27	8.07E-26
990	LINC01314	Myelinating Schwann cells	5.977144	4.35E-23	1.66E-21
991	RP11-598F7.3	ASDC	5.909327	9.66E-13	2.77E-10
992	BDKRB2	ASDC	4.9578834	0.00010716	0.007679784
993	HAMP	ASDC	7.6330476	5.87E-06	0.000531378
994	SLC41A2	ASDC	4.5264044	2.48E-06	0.000237231
995	LTK	ASDC	6.413557	0.000593445	0.036454459
996	MIR4432HG	ASDC	8.5879345	0.017851334	0.61408589
997	CDH23	ASDC	3.4348419	0.00091285	0.050648441
998	SLC4A3	ASDC	3.1412525	0.008479265	0.331462162
999	PPM1J	ASDC	5.3641706	1.36E-09	2.12E-07
1000	GRAMD4	ASDC	2.4471948	3.77E-05	0.002946537
1001	SIGLEC6	ASDC	6.735953	2.11E-08	2.79E-06
1002	RRM2B	ASDC	1.9697793	0.000384505	0.026453917
1003	CLEC9A	DC1	10.867453	0	0
1004	BATF3	DC1	6.3713055	4.63E-249	2.65E-246
1005	CLNK	DC1	7.3586397	6.13E-151	2.64E-148
1006	XCR1	DC1	10.534615	3.97E-43	3.25E-41
1007	IDO1	DC1	10.726732	3.73E-266	3.20E-263
1008	TAP2	DC1	3.7403233	2.59E-61	3.71E-59
1009	AIM2	DC1	7.4784	4.26E-19	1.74E-17
1010	PPY	DC1	10.644698	5.59E-14	1.85E-12
1011	HLA-DOB	DC1	6.9728537	2.10E-89	5.16E-87
1012	CCSER1	DC1	5.499637	1.50E-70	2.87E-68
1013	BTLA	DC1	5.3950496	6.75E-68	1.16E-65
1014	PLEKHM3	DC1	3.8767958	9.34E-22	4.59E-20
1015	AC096772.6	DC1	3.5906603	3.66E-18	1.43E-16
1016	SLC46A3	DC1	3.942808	1.47E-37	1.10E-35
1017	SLAMF8	DC1	6.2692766	1.83E-56	2.10E-54
1018	VMO1	DC1	4.1771455	5.35E-35	3.54E-33
1019	RP11-798K3.3	DC1	6.5036855	2.19E-78	4.71E-76
1020	FAM135A	DC1	2.2485926	3.52E-36	2.42E-34
1021	BCL6	DC1	2.4347525	1.39E-39	1.09E-37
1022	SERPINF2	DC1	5.361456	2.01E-46	1.73E-44
1023	TRERF1	DC1	1.4485062	7.47E-16	2.73E-14
1024	PPM1M	DC1	2.101836	1.73E-34	1.06E-32
1025	CHN2	DC1	2.6486096	7.96E-29	4.03E-27
1026	LAMP3	Inflammatory DC	11.607614	6.35E-30	1.09E-26
1027	CCL22	Inflammatory DC	10.685211	3.61E-20	2.07E-17

1028	TVP23A	Inflammatory DC	5.6982365	8.22E-20	3.53E-17
1029	NCCRP1	Inflammatory DC	10.666358	1.24E-18	3.55E-16
1030	EBI3	Inflammatory DC	7.10456	1.09E-16	2.67E-14
1031	AOC1	Inflammatory DC	12.640075	1.21E-07	1.30E-05
1032	GPR157	Inflammatory DC	6.9619	7.87E-13	1.69E-10
1033	IL15	Inflammatory DC	6.609557	3.51E-12	6.72E-10
1034	SLCO5A1	Inflammatory DC	8.449662	4.14E-10	5.79E-08
1035	CCL19	Inflammatory DC	9.81733	1.45E-08	1.79E-06
1036	CLLU1OS	Inflammatory DC	10.318002	0.000868851	0.036449343
1037	HMSD	Inflammatory DC	6.216898	7.28E-08	8.34E-06
1038	ARHGAP22	Inflammatory DC	6.6907406	5.17E-07	4.94E-05
1039	CD1B	Inflammatory DC	8.984627	0.000208761	0.010880852
1040	TFPI2	Inflammatory DC	3.7633893	2.19E-06	0.000164131
1041	LRRK1	Inflammatory DC	3.6972272	2.14E-10	3.35E-08
1042	MIR155HG	Inflammatory DC	4.5580177	4.38E-10	5.79E-08
1043	CCL17	Inflammatory DC	10.602584	0.000424759	0.020294054
1044	LINC01539	Inflammatory DC	9.955095	0.018783534	0.598290355
1045	LY75	Inflammatory DC	5.3513217	7.74E-07	6.65E-05
1046	MAP3K14	Inflammatory DC	3.963899	1.14E-06	8.92E-05
1047	CD80	Inflammatory DC	7.2537994	0.003498178	0.130801427
1048	CLEC4G	Inflammatory DC	5.4524055	0.022612622	0.694530519
1049	CD274	Inflammatory DC	5.8098893	9.96E-06	0.000658895
1050	BCL2L14	Inflammatory DC	7.7762547	0.006395111	0.229158141
1051	PDCD1LG2	Inflammatory DC	6.586241	0.003618567	0.132424149
1052	MREG	Inflammatory DC	5.263753	0.000138009	0.007657291
1053	CTA-384D8.35	Inflammatory DC	6.817494	0.011513318	0.396058125
1054	AC079767.4	Inflammatory DC	6.829719	0.001924776	0.075241242
1055	ANKRD55	Inflammatory DC	5.2382183	0.000144905	0.007788659
1056	CXCL9	Inflammatory DC	8.706409	0.050394192	0.9998238
1057	UBD	Inflammatory DC	8.746961	0.050379594	0.9998238
1058	CXCL10	Inflammatory DC	5.7767887	0.012485797	0.421089609
1059	TRAF3	Inflammatory DC	2.91121	5.89E-05	0.003378208
1060	SBNO2	Inflammatory DC	2.9369836	0.000774795	0.033316203
1061	SMPD3	pDC	9.6409645	8.62E-199	3.71E-196
1062	LINC00996	pDC	6.1632514	1.54E-153	3.79E-151
1063	PTCRA	pDC	11.286246	3.12E-233	1.79E-230
1064	DERL3	pDC	6.431002	3.39E-146	7.30E-144
1065	IRF7	pDC	5.9626455	0	0
1066	LILRA4	pDC	9.374616	1.13E-171	3.88E-169
1067	GZMB	pDC	7.4458804	7.58E-299	6.52E-296

343

1068	CLEC4C	pDC	11.117163	6.59E-163	1.89E-160
1069	MYBL2	pDC	5.417145	2.17E-85	3.73E-83
1070	ZFAT	pDC	5.932146	2.08E-72	2.98E-70
1071	AC104024.1	pDC	9.199553	6.43E-36	5.82E-34
1072	SLC15A4	pDC	4.2948003	1.27E-77	1.98E-75
1073	RP11-73G16.2	pDC	12.639896	1.16E-25	8.00E-24
1074	CPLX1	pDC	6.145637	3.52E-38	3.56E-36
1075	PACSIN1	pDC	8.045663	2.68E-37	2.56E-35
1076	NLRP7	pDC	12.307331	5.77E-10	2.16E-08
1077	AC011893.3	pDC	8.591529	1.99E-32	1.56E-30
1078	KCTD5	pDC	2.558793	5.16E-30	3.70E-28
1079	PHEX	pDC	6.497609	4.92E-23	3.02E-21
1080	PARP10	pDC	3.5493464	9.43E-33	7.73E-31
1081	RP11-117D22.2	pDC	7.2918324	1.27E-17	6.60E-16
1082	TLR9	pDC	9.185045	5.49E-09	2.01E-07
1083	IGLJ2	pDC	9.781215	4.55E-14	2.06E-12
1084	DRD4	pDC	7.188687	1.97E-11	8.45E-10
1085	NEK8	pDC	4.0248413	2.83E-15	1.35E-13
1086	TTC39A	pDC	7.1404943	1.03E-22	6.08E-21
1087	LRRC26	pDC	9.47263	3.74E-11	1.50E-09
1088	ADGRG5	pDC	5.2463036	1.26E-22	7.25E-21
1089	RP11-542M13.3	pDC	6.1595583	4.51E-11	1.76E-09
1090	RP4-647C14.2	pDC	4.1436267	1.33E-08	4.41E-07
1091	CUX2	pDC	6.751465	3.62E-11	1.48E-09
1092	CD207	LC	12.509914	2.01E-98	3.46E-95
1093	FCGBP	LC	5.697465	1.18E-36	5.06E-34
1094	CD1A	LC	8.475122	9.10E-15	1.04E-12
1095	RP11-597D13.8	LC	7.78333	3.58E-48	2.06E-45
1096	TCHH	LC	5.4297323	2.42E-16	3.20E-14
1097	LACC1	LC	3.986921	1.18E-22	2.89E-20
1098	GNGT2	LC	4.235744	3.60E-20	6.89E-18
1099	ZBTB7A	Monocyte precursor	2.2907977	1.06E-64	9.62E-63
1100	SERPINB8	Monocyte precursor	3.80276	6.63E-131	1.14E-128
1101	AQP9	Monocyte	7.6475916	0	0
1102	VNN2	Monocyte	6.6335034	0	0
1103	CCL20	Monocyte	6.282831	0	0
1104	CLEC4E	Monocyte	6.755689	0	0
1105	MIR3945HG	Monocyte	5.341787	6.86E-156	7.37E-154
1106	SPATA13	Monocyte	4.106592	4.56E-141	4.13E-139
1107	DNAAF1	Monocyte	5.761776	1.00E-214	1.73E-212

344

1108	IRG1	Monocyte	7.9202423	2.49E-52	1.16E-50
1109	LIPN	Monocyte	6.453432	8.05E-72	4.62E-70
1110	PROK2	Monocyte	6.052801	1.06E-150	1.01E-148
1111	VNN1	Monocyte	5.168205	3.42E-66	1.84E-64
1112	STEAP4	Monocyte	4.03008	4.65E-140	4.00E-138
1113	FCGR3B	Monocyte	6.5437603	1.70E-71	9.41E-70
1114	PLA2G7	Iron-recycling macrophage	5.645552	3.40E-101	5.85E-98
1115	TREM2	Iron-recycling macrophage	6.7460093	1.62E-90	1.39E-87
1116	CD5L	Iron-recycling macrophage	6.638904	3.65E-09	1.74E-07
1117	SDSL	Iron-recycling macrophage	3.884111	3.61E-19	5.18E-17
1118	FAM20C	Iron-recycling macrophage	2.3892982	6.11E-10	3.50E-08
1119	SIGLEC7	Iron-recycling macrophage	3.7093816	4.10E-18	5.04E-16
1120	CMKLR1	Iron-recycling macrophage	2.9508636	8.93E-21	1.92E-18
1121	SLC37A2	Iron-recycling macrophage	3.461943	5.27E-23	1.29E-20
1122	FPR3	Iron-recycling macrophage	5.630993	2.76E-11	1.69E-09
1123	RAB42	Iron-recycling macrophage	2.3628151	3.02E-06	0.000110436
1124	RP11-733O18.1	Iron-recycling macrophage	3.840296	1.17E-09	6.27E-08
1125	SIGLEC9	Iron-recycling macrophage	3.5424316	2.80E-17	3.01E-15
1126	CD180	Iron-recycling macrophage	3.2497935	2.99E-13	2.45E-11
1127	OTOA	Iron-recycling macrophage	3.6025887	2.22E-08	9.77E-07
1128	DAB2	LYVE1++ macrophage	4.900291	0	0
1129	SIGLEC1	LYVE1++ macrophage	7.2378864	0	0
1130	CD209	LYVE1++ macrophage	7.3622336	0	0
1131	CCL13	LYVE1++ macrophage	7.3843455	0	0
1132	C2	LYVE1++ macrophage	6.718223	0	0
1133	TIMD4	LYVE1++ macrophage	6.868403	0	0
1134	MMP19	MHCII+ macrophage	2.68129	1.33E-89	2.08E-87
1135	C3	TREM2+ macrophage	7.606228	0	0
1136	RP11-480C22.1	TREM2+ macrophage	10.787864	0	0
1137	P2RY12	TREM2+ macrophage	7.69721	5.84E-280	2.51E-277
1138	BHLHE41	TREM2+ macrophage	5.0434036	1.44E-241	4.94E-239
1139	FCGR1B	TREM2+ macrophage	5.096401	1.96E-132	5.61E-130
1140	SYT6	TREM2+ macrophage	5.8961825	1.72E-66	1.74E-64
1141	TMEM144	TREM2+ macrophage	4.5871735	1.43E-88	2.73E-86
1142	CSF1	TREM2+ macrophage	3.260051	3.05E-81	4.77E-79
1143	RP11-552D4.1	TREM2+ macrophage	6.888534	1.67E-42	1.11E-40
1144	SUCNR1	TREM2+ macrophage	7.2800646	1.08E-44	7.44E-43
1145	ADORA3	TREM2+ macrophage	5.6324315	1.35E-76	1.93E-74
1146	SALL1	TREM2+ macrophage	6.2987275	2.00E-29	1.11E-27
1147	TMEM52B	TREM2+ macrophage	5.885583	5.92E-45	4.24E-43

1148	LINC01094	TREM2+ macrophage	5.107903	5.44E-45	4.07E-43
1149	SIGLEC8	TREM2+ macrophage	8.166105	1.64E-21	7.84E-20
1150	TMIGD3	TREM2+ macrophage	4.224584	3.01E-67	3.45E-65
1151	DEFA4	Neutrophil1	15.352081	5.14E-58	5.10E-55
1152	AZU1	Neutrophil1	13.182836	2.07E-55	1.19E-52
1153	MPO	Neutrophil1	13.04753	5.93E-58	5.10E-55
1154	PRTN3	Neutrophil1	13.1273	1.29E-44	5.54E-42
1155	ELANE	Neutrophil1	13.577745	4.21E-35	1.03E-32
1156	DEFA3	Neutrophil1	15.000713	1.93E-42	6.65E-40
1157	DEFA1	Neutrophil1	13.618472	4.16E-18	7.15E-16
1158	DEFA1B	Neutrophil1	13.718551	2.07E-17	2.97E-15
1159	RNASE3	Neutrophil1	9.024401	1.33E-16	1.63E-14
1160	SLPI	Neutrophil1	9.452569	1.21E-16	1.61E-14
1161	CEACAM6	Neutrophil1	11.394709	3.12E-11	2.98E-09
1162	CCDC88B	Neutrophil1	5.4294314	6.08E-18	9.51E-16
1163	BPI	Neutrophil1	6.829222	7.16E-26	1.37E-23
1164	FUT4	Neutrophil1	4.922105	6.28E-06	0.00036676
1165	FAM95C	Neutrophil1	10.942098	0.026947004	0.891323991
1166	SLC2A5	Neutrophil1	6.589313	0.001123743	0.053689934
1167	MROH6	Neutrophil1	4.611239	0.006589861	0.26987048
1168	SERPINB10	Neutrophil1	8.231854	4.28E-09	3.68E-07
1169	CEBPE	Neutrophil1	7.3098197	2.08E-06	0.000127869
1170	PGLYRP1	Neutrophil2	12.064688	1.99E-101	1.71E-98
1171	LTF	Neutrophil2	13.938079	3.04E-64	1.74E-61
1172	CAMP	Neutrophil2	12.107831	5.45E-37	2.34E-34
1173	S100A8	Neutrophil2	12.017343	6.84E-115	1.18E-111
1174	CHI3L1	Neutrophil2	10.040782	8.92E-23	1.53E-20
1175	CD177	Neutrophil2	11.046282	7.96E-14	8.56E-12
1176	FOLR3	Neutrophil2	7.795285	1.86E-36	6.41E-34
1177	CRISP3	Neutrophil2	10.444011	1.35E-15	1.78E-13
1178	CHIT1	Neutrophil2	10.592263	2.13E-13	2.03E-11
1179	CYP4F3	Neutrophil2	8.599696	0.000151414	0.006024956
1180	MMP8	Neutrophil2	11.904649	0.018232459	0.435553194
1181	CEACAM8	Neutrophil2	11.812595	2.78E-14	3.41E-12
1182	ARG1	Neutrophil2	8.535819	6.68E-30	1.92E-27
1183	OLFM4	Neutrophil2	7.0941815	0.003325713	0.102146902
1184	ORM1	Neutrophil2	7.038842	3.76E-13	3.23E-11
1185	ADGRG3	Neutrophil2	6.863625	6.63E-14	7.60E-12
1186	PTH2	Lymphoid progenitor	7.8540654	0.000314529	0.049180845
1187	IL17A	LTi	7.965111	2.80E-15	3.22E-13

1188	PTGDR2	ILC2		7.7617545	8.84E-72	7.60E-69
1189	IL9R	ILC2		9.473588	1.34E-30	3.84E-28
1190	FXYD7	ILC2		5.107498	2.94E-20	4.60E-18
1191	AC002331.1	ILC2		10.232541	7.79E-05	0.003722972
1192	FSTL4	ILC2		6.7699747	5.15E-07	4.02E-05
1193	HPN	ILC3		9.377386	2.20E-211	1.89E-208
1194	LINC00299	ILC3		8.656792	1.22E-274	2.10E-271
1195	SCN1B	ILC3		5.936243	4.72E-165	2.03E-162
1196	PDZK1	ILC3		8.350523	1.02E-59	8.38E-58
1197	KRT81	ILC3		6.835972	8.66E-72	9.92E-70
1198	EYS	ILC3		7.2167997	1.27E-48	9.93E-47
1199	KRT86	ILC3		6.8656554	4.17E-197	2.39E-194
1200	BCAS1	ILC3		5.6912847	9.03E-101	1.94E-98
1201	GGT1	ILC3		5.0731654	2.21E-60	1.90E-58
1202	SLC4A10	ILC3		5.62661	1.65E-62	1.49E-60
1203	CXCR5	ILC3		6.6515126	2.94E-147	1.01E-144
1204	TTC39C-AS1	ILC3		7.714494	1.18E-86	1.69E-84
1205	COL4A4	ILC3		6.657636	3.61E-34	2.22E-32
1206	TRGC2	ILC3		5.285743	2.55E-41	1.76E-39
1207	SERPINA11	ILC3		6.9396663	4.39E-14	1.51E-12
1208	IPCEF1	ILC3		3.7694588	1.95E-69	1.97E-67
1209	GZMK	Innate T type1		7.381604	0	0
1210	TRDV2	Innate T type1		8.99884	2.96E-76	1.02E-73
1211	IFNG-AS1	Innate T type1		7.262675	4.49E-17	3.22E-15
1212	TRGV8	Innate T type1		7.414204	1.89E-08	7.95E-07
1213	RP5-1028K7.2	Innate T type1		5.5532603	3.76E-12	2.31E-10
1214	GPR25	Innate T type3		6.990213	5.87E-20	1.68E-17
1215	GZMH	NK cell		9.5908165	0	0
1216	GNLY	NK cell		8.464121	0	0
1217	CMC1	NK cell		5.4888654	0	0
1218	KLRD1	NK cell		8.986831	0	0
1219	KLRF1	NK cell		8.509406	0	0
1220	GZMA	NK cell		7.65959	0	0
1221	NKG7	NK cell		9.4140625	0	0
1222	IFNG	NK cell		8.666936	0	0
1223	S1PR5	NK cell		7.4763265	0	0
1224	CD160	NK cell		8.911394	5.88E-221	4.22E-219
1225	SH2D1B	NK cell		7.7320223	0	0
1226	NCR1	NK cell		7.719786	7.21E-176	4.13E-174
1227	KIR2DL4	NK cell		9.341735	3.24E-96	1.50E-94

1228	KIR2DL1	NK cell	9.835931	1.50E-82	6.29E-81
1229	IL18RAP	NK cell	6.590556	0	0
1230	FASLG	NK cell	7.8470745	1.57E-263	1.28E-261
1231	KIR2DL3	NK cell	9.3781805	4.23E-101	2.08E-99
1232	KIR3DL1	NK cell	9.371177	9.27E-55	2.85E-53
1233	EOMES	NK cell	7.1674957	1.74E-104	8.80E-103
1234	ERVH48-1	CD4 T cell	6.9512396	2.03E-76	1.58E-74
1235	CD8B	CD8 T cell	8.864881	0	0
1236	RP11-291B21.2	CD8 T cell	9.813023	0	0
1237	PASK	CD8 T cell	5.891054	0	0
1238	CHI3L2	CD8 T cell	6.0316787	0	0
1239	REG4	CD8 T cell	7.077896	5.44E-15	2.34E-13
1240	CTLA4	Treg	9.004111	4.48E-241	7.71E-238
1241	FOXP3	Treg	10.286537	5.73E-139	1.97E-136
1242	TIGIT	Treg	6.401906	4.66E-136	1.33E-133
1243	IL2RA	Treg	8.10019	4.56E-165	1.96E-162
1244	RP11-1399P15.1	Treg	10.016526	1.69E-20	1.12E-18
1245	FAS	Treg	3.4699621	1.03E-41	1.11E-39
1246	GNG8	Treg	5.203077	9.56E-28	7.83E-26
1247	LAG3	Treg	4.3252454	1.55E-11	7.62E-10
1248	LAIR2	Treg	5.816273	7.26E-65	1.14E-62
1249	RP5-887A10.1	Pre B cell	8.54369	1.05E-130	1.80E-127
1250	CCDC191	Pre B cell	5.1948724	7.58E-38	1.19E-35
1251	TCL1B	Pre B cell	7.967291	6.87E-21	6.95E-19
1252	RP11-1070N10.3	Pre B cell	7.1888685	1.34E-29	1.77E-27
1253	BEST3	Pre B cell	8.90387	3.26E-25	3.50E-23
1254	CFAP73	Pre B cell	8.327322	4.46E-08	2.79E-06
1255	UMODL1	Pre pro B cell	8.577388	7.71E-11	9.47E-09
1256	TNFRSF13B	Pre pro B cell	7.68184	6.11E-12	9.56E-10
1257	SCN3A	Pre pro B cell	6.6270137	3.52E-15	8.65E-13
1258	ENAM	Pre pro B cell	8.52941	9.57E-08	8.45E-06
1259	KB-208E9.1	Pre pro B cell	8.157725	9.83E-08	8.45E-06
1260	SLC8A1-AS1	Pre pro B cell	7.2047434	2.09E-29	1.20E-26
1261	BAHCC1	Pre pro B cell	3.3291929	1.11E-10	1.19E-08
1262	DNTT	Pro B cell	11.113144	6.80E-126	5.85E-123
1263	VPREB1	Pro B cell	10.90359	1.35E-223	2.32E-220
1264	LINC01013	Pro B cell	9.621903	9.22E-69	3.96E-66
1265	RP11-301G19.1	Pro B cell	10.246528	1.04E-24	1.38E-22
1266	LCN6	Pro B cell	10.869826	3.94E-25	6.16E-23
1267	LINC00426	Pro B cell	8.232993	2.96E-37	8.49E-35

348

1268	ALDH5A1	Pro B cell	4.3928223	2.45E-21	2.81E-19
1269	MS4A1	B cell	11.6786995	0	0
1270	SP140	B cell	5.5451937	5.92E-242	5.09E-239
1271	PARP15	B cell	4.980935	1.41E-151	2.43E-149
1272	FCRL5	B cell	9.973266	2.64E-43	1.68E-41
1273	FCRL1	B cell	9.380558	3.68E-129	4.52E-127
1274	RP11-231C14.7	B cell	6.350166	2.84E-142	4.45E-140
1275	TNFRSF13C	B cell	5.1954107	2.28E-115	2.61E-113
1276	FCER2	B cell	4.917365	1.93E-44	1.33E-42
1277	CLEC17A	B cell	7.708571	8.97E-31	4.98E-29
1278	PNOC	B cell	6.883121	3.14E-38	1.93E-36
1279	RP11-325F22.2	B cell	4.1634445	4.58E-99	4.92E-97
1280	FCRL2	B cell	7.728107	1.38E-16	5.06E-15
1281	RP11-164H13.1	B cell	8.0755625	1.03E-53	8.09E-52
1282	SPINK2	HSC	9.894599	3.66E-61	6.30E-58
1283	AP001171.1	HSC	7.840521	1.39E-15	7.96E-13
1284	TRH	HSC	4.5291324	2.42E-16	2.08E-13
1285	CHST13	HSC	4.7016225	9.57E-06	0.00137181
1286	RP11-299J3.8	HSC	2.6472287	1.15E-05	0.001408012
1287	CLDN15	HSC	3.158422	5.13E-05	0.005118427
1288	UBR5-AS1	HSC	4.237116	1.44E-06	0.000226451
1289	HTR1F	HSC	5.179932	0.000779238	0.044676307
1290	PRSS2	HSC	10.396776	0.000598626	0.036772761
1291	CLDN10	HSC	5.6408687	7.99E-09	1.72E-06
1292	NR6A1	HSC	3.3803432	2.71E-05	0.002912343
1293	CCDC171	HSC	2.5108223	0.000121924	0.009419165
1294	NOG	HSC	3.6243124	0.000314256	0.021620834
1295	MT1F	Erythroid (embryonic)	8.121728	0	0
1296	MT1H	Erythroid (embryonic)	11.559282	0	0
1297	MT1G	Erythroid (embryonic)	10.936268	0	0
1298	SLC39A8	Erythroid (embryonic)	5.912409	0	0
1299	MT1X	Erythroid (embryonic)	5.346427	0	0
1300	HBE1	Erythroid (embryonic)	16.178913	0	0
1301	MT1E	Erythroid (embryonic)	6.987057	0	0
1302	RP5-1198O20.4	Erythroid (embryonic)	6.1726727	4.49E-64	3.09E-62
1303	PRRT1	Erythroid (embryonic)	4.237276	1.51E-31	7.02E-30
1304	NHLRC4	Erythroid (embryonic)	5.2261615	5.72E-22	2.24E-20
1305	PKLR	Early erythroid (embryonic)	9.458901	6.37E-164	9.13E-162
1306	RP11-470P21.2	Early erythroid (embryonic)	9.47033	1.33E-65	1.15E-63
1307	RP11-797H7.5	Early erythroid (embryonic)	8.402389	1.14E-127	1.40E-125

1308	CTSE	Early erythroid (embryonic)	9.165415	6.83E-207	2.94E-204
1309	TF	Early erythroid (embryonic)	7.1948094	7.03E-101	8.07E-99
1310	ST6GALNAC1	Early erythroid (embryonic)	7.144578	1.43E-66	1.30E-64
1311	SLC25A21	Early erythroid (embryonic)	5.099903	1.58E-25	7.53E-24
1312	NT5DC4	Early erythroid (embryonic)	8.471015	9.06E-31	5.03E-29
1313	SLC30A10	Early erythroid (embryonic)	6.806208	1.70E-36	1.01E-34
1314	SPESP1	Early erythroid (embryonic)	2.8637652	1.18E-07	4.30E-06
1315	ABCB10	Early erythroid (embryonic)	3.306749	8.67E-16	3.73E-14
1316	GATA5	Early erythroid (embryonic)	8.15584	7.70E-39	4.91E-37
1317	TMCC2	Early erythroid	9.477422	1.15E-162	1.24E-160
1318	EIF5	Early erythroid	3.7753313	1.51E-187	1.86E-185
1319	XPO7	Early erythroid	5.960391	4.31E-125	3.90E-123
1320	PIM1	Early erythroid	6.6465144	6.13E-233	2.64E-230
1321	CAT	Early erythroid	4.750076	5.70E-212	8.91E-210
1322	PNP	Early erythroid	6.1306863	1.27E-170	1.46E-168
1323	TSPO2	Early erythroid	9.995054	3.85E-85	3.16E-83
1324	EPB42	Early erythroid	8.890583	7.90E-232	2.72E-229
1325	IBA57	Early erythroid	5.255339	2.09E-38	1.33E-36
1326	YPEL4	Early erythroid	7.803597	7.94E-76	6.21E-74
1327	RSAD2	Early erythroid	6.6686196	1.51E-22	6.84E-21
1328	ACSL6	Early erythroid	7.246163	6.89E-32	4.23E-30
1329	TRIM10	Early erythroid	8.783466	3.52E-17	1.41E-15
1330	ART4	Early erythroid	8.840714	9.84E-19	4.13E-17
1331	TRIM58	Early erythroid	8.751879	3.38E-219	7.27E-217
1332	CRB1	Early erythroid	8.464662	4.17E-10	1.41E-08
1333	NCEH1	Early erythroid	5.167127	4.74E-17	1.85E-15
1334	TBCEL	Early erythroid	4.9835224	8.88E-24	4.24E-22
1335	DYRK3	Early erythroid	5.36092	1.05E-23	4.90E-22
1336	FECH	Early erythroid	7.295683	7.88E-255	6.77E-252
1337	TMEM86B	Early erythroid	6.083541	8.93E-25	4.39E-23
1338	BSG	Early erythroid	4.6415215	4.82E-244	2.76E-241
1339	SLC14A1	Early erythroid	7.312956	7.32E-51	5.25E-49
1340	DCAF12	Erythroid (fetal)	7.873213	0	0
1341	HBB	Erythroid (fetal)	15.293702	0	0
1342	SNCA	Erythroid (fetal)	8.998622	0	0
1343	AC104389.1	Erythroid (fetal)	8.684248	0	0
1344	FKBP8	Erythroid (fetal)	4.7481437	0	0
1345	STRADB	Erythroid (fetal)	7.74948	0	0
1346	GABARAPL2	Erythroid (fetal)	4.68941	0	0
1347	BNIP3L	Erythroid (fetal)	5.7522936	0	0

1348	MKRN1	Erythroid (fetal)	4.9372716	0	0
1349	PHOSPHO1	Erythroid (fetal)	9.79329	0	0
1350	IFIT1B	Erythroid (fetal)	9.76659	6.86E-76	3.93E-74
1351	FAM178B	MEMP - Early erythroid	10.714776	2.84E-57	4.89E-54
1352	PVT1	MEMP - Early erythroid	5.369701	3.65E-39	3.14E-36
1353	PNMT	MEMP - Early erythroid	7.3194475	1.13E-28	4.87E-26
1354	SRM	MEMP - Early erythroid	1.7778957	3.17E-13	3.90E-11
1355	BOLA3-AS1	MEMP - Early erythroid	2.753205	2.50E-12	2.87E-10
1356	SLC27A2	MEMP - Early erythroid	6.692016	8.32E-14	1.19E-11
1357	DNAH8	MEMP - Early erythroid			
1358	NMNAT3	MEMP - Early erythroid	4.7326393	6.92E-21	1.98E-18
1359	GAL	MEMP - Early erythroid	3.7502172	5.38E-10	4.02E-08
1360	ATAD3A	MEMP - Early erythroid	2.2368069	7.69E-10	5.51E-08
1361	RP11-98O2.1	MEMP - Early erythroid			
1362	COX10	MEMP - Early erythroid	1.896247	1.95E-08	1.05E-06
1363	ECE2	MEMP - Early erythroid	2.28153	2.21E-10	1.90E-08
1364	MRPS17	MEMP - Early erythroid	2.2767947	1.97E-09	1.25E-07
1365	POLR1B	MEMP - Early erythroid	2.6531916	8.83E-07	3.53E-05
1366	SYCP1	MEMP - Early erythroid			
1367	AC002467.7	MEMP - Early erythroid	3.1220767	1.91E-06	6.86E-05
1368	AMHR2	MEMP - Early erythroid	7.116643	9.48E-16	1.81E-13
1369	SLX1B	MEMP - Early erythroid	4.501797	0.00034669	0.007184415
1370	RHEBL1	MEMP - Early erythroid	3.5236213	1.58E-05	0.000468744
1371	DDX55	MEMP - Early erythroid	1.5716976	1.27E-06	4.85E-05
1372	GDPD1	MEMP - Early erythroid	4.0913405	6.25E-09	3.71E-07
1373	PDSS1	MEMP - Early erythroid	2.7495384	5.26E-07	2.20E-05
1374	PDCD2L	MEMP - Early erythroid	2.4488325	1.40E-05	0.000428904
1375	ZNF639	MEMP - Early erythroid	1.537416	1.27E-06	4.85E-05
1376	CCDC169	MEMP - Early erythroid			
1377	LARP1B	MEMP - Early erythroid	1.4943618	8.67E-06	0.00027623
1378	TRIM6	MEMP - Early erythroid	4.421252	0.002589711	0.044543037
1379	MRPL30	MEMP - Early erythroid	1.4041439	1.90E-06	6.86E-05
1380	CTC-338M12.5	MEMP - Early erythroid			
1381	CEP152	MEMP - Early erythroid	2.4728336	3.83E-08	1.99E-06
1382	RFT1	MEMP - Early erythroid	1.3580681	3.32E-06	0.000111878
1383	MIF-AS1	MEMP - Early erythroid	2.578641	5.56E-05	0.001447907
1384	CRISP2	MEMP - Early erythroid	8.573316	0.001864448	0.033404697
1385	PFAS	MEMP - Early erythroid	1.9020844	0.000445601	0.009124207
1386	FBXO4	MEMP - Early erythroid	1.4330695	2.68E-06	9.23E-05
1387	WDR4	MEMP - Early erythroid	2.4221396	2.47E-05	0.000688367

1388	MAK16	MEMP - Early erythroid	1.8213499	1.22E-07	6.01E-06
1389	ANKLE1	MEMP - Early erythroid	4.089981	0.000218855	0.004888718
1390	POLE2	MEMP - Early erythroid	2.7860508	1.48E-07	6.86E-06
1391	MARS2	MEMP - Early erythroid	2.7696476	0.001423902	0.02638301
1392	SLX1A	MEMP - Early erythroid			
1393	PRSS21	MEMP - Early erythroid	4.7664194	1.49E-11	1.50E-09
1394	NAT10	MEMP - Early erythroid	1.8008904	1.63E-06	6.08E-05
1395	MTO1	MEMP - Early erythroid	1.2472458	9.91E-05	0.00243421
1396	ZNF768	MEMP - Early erythroid	1.6405009	0.001464589	0.026516764
1397	C9orf40	MEMP - Early erythroid	2.1371675	0.000972896	0.018802026
1398	ERI1	MEMP - Early erythroid	1.2633963	0.000254645	0.005544176
1399	HPDL	MEMP - Early erythroid	4.220856	4.23E-08	2.14E-06
1400	NUP62CL	MEMP - Early erythroid	4.365681	4.39E-05	0.001180372
1401	TCEB3-AS1	MEMP - Early erythroid			
1402	DPH2	MEMP - Early erythroid	1.7717191	5.67E-05	0.00145681
1403	COQ3	MEMP - Early erythroid	1.899382	4.96E-06	0.00016396
1404	MRM1	MEMP - Early erythroid	3.8944674	1.41E-07	6.73E-06
1405	UBL7-AS1	MEMP - Early erythroid	1.5534049	0.000590869	0.011920682
1406	TRMT61A	MEMP - Early erythroid	1.3692249	0.00024252	0.005347868
1407	FBXO45	MEMP - Early erythroid	2.1655803	9.51E-05	0.002369755
1408	LINC01150	MEMP - Early erythroid	4.12021	0.000769968	0.015049378
1409	CCDC51	MEMP - Early erythroid	1.5977621	4.89E-05	0.001293567
1410	C5orf63	MEMP - Early erythroid	2.126044	8.98E-05	0.002270895
1411	ZNF77	MEMP - Early erythroid	2.2524436	0.00227303	0.039893995
1412	PXN-AS1	MEMP - Early erythroid	1.7004887	0.000113519	0.002654604
1413	RNASEH1-AS1	MEMP - Early erythroid	1.5958788	1.42E-05	0.000428962
1414	CH507-9B2.3	MEMP - Early erythroid	2.4502168	1.23E-05	0.000384256
1415	GATB	MEMP - Early erythroid	1.7379624	2.69E-07	1.22E-05
1416	HBD	MEMP - Megak	10.209379	2.38E-126	4.10E-123
1417	SPINK4	MEMP - Megak	8.463351	1.07E-37	1.23E-35
1418	GDF11	MEMP - Megak	2.9126244	2.48E-22	1.77E-20
1419	AF196970.3	MEMP - Megak			
1420	KCNK5	MEMP - Megak	5.022985	6.78E-15	3.15E-13
1421	MIR3142HG	MEMP - Megak	5.389055	2.76E-16	1.40E-14
1422	ENTPD1-AS1	MEMP - Megak	2.343953	2.09E-17	1.12E-15
1423	XRCC2	MEMP - Megak	3.154367	2.55E-11	8.77E-10
1424	C11orf98	MEMP - Megak	2.5399966	3.12E-14	1.34E-12
1425	STAC	MEMP - Megak			
1426	WRN	MEMP - Megak	1.5084637	8.42E-12	3.02E-10
1427	NBPF26	MEMP - Megak	1.5822277	1.90E-07	4.41E-06

1428	AP4S1	MEMP - Megak	2.1252282	7.56E-10	2.28E-08
1429	CENPP	MEMP - Megak	2.058586	3.95E-11	1.33E-09
1430	RAD54L	MEMP - Megak	2.5556035	5.26E-08	1.34E-06
1431	RP13-726E6.2	MEMP - Megak	5.5604773	1.18E-12	4.63E-11
1432	PPBP	Megakaryocyte	14.0542345	8.02E-111	4.60E-108
1433	PF4	Megakaryocyte	13.659931	1.96E-113	3.38E-110
1434	GP9	Megakaryocyte	11.914752	3.79E-112	3.26E-109
1435	GP1BA	Megakaryocyte	10.909913	6.45E-89	7.92E-87
1436	TREML1	Megakaryocyte	12.255163	2.42E-83	1.90E-81
1437	NRGN	Megakaryocyte	8.332227	8.88E-104	3.05E-101
1438	HPSE	Megakaryocyte	7.7135077	5.25E-80	3.76E-78
1439	ILK	Megakaryocyte	4.0199623	8.26E-95	1.58E-92
1440	ZYX	Megakaryocyte	4.1373763	1.66E-87	1.68E-85
1441	TUBB1	Megakaryocyte	9.721976	2.07E-87	1.97E-85
1442	CMTM5	Megakaryocyte	8.706047	1.21E-102	3.48E-100
1443	LAT	Megakaryocyte	6.8278775	2.73E-93	4.27E-91
1444	CLEC1B	Megakaryocyte	12.213498	2.23E-98	4.80E-96
1445	TLN1	Megakaryocyte	4.3663344	2.56E-88	2.94E-86
1446	C6orf25	Megakaryocyte	10.560176	2.37E-104	1.02E-101
1447	RSU1	Megakaryocyte	3.9481626	6.36E-86	5.47E-84
1448	RGS18	Megakaryocyte	7.5390325	3.13E-93	4.48E-91
1449	FERMT3	Megakaryocyte	6.314248	4.42E-102	1.09E-99
1450	TMEM40	Megakaryocyte	11.601661	7.39E-78	5.08E-76
1451	GP6	Megakaryocyte	10.829315	4.17E-56	2.56E-54
1452	CCND3	Megakaryocyte	5.292612	4.24E-93	5.61E-91
1453	TUBA4A	Megakaryocyte	6.3070703	1.55E-84	1.27E-82
1454	SPX	Megakaryocyte	8.248187	5.22E-64	3.45E-62
1455	TNNI3	Megakaryocyte	9.727106	1.81E-42	8.64E-41
1456	ITGB3	Megakaryocyte	8.15733	4.85E-62	3.09E-60
1457	MAX	Megakaryocyte	4.3807864	1.13E-86	1.03E-84
1458	ITGA2B	Megakaryocyte	9.823692	1.38E-94	2.38E-92
1459	PF4V1	Megakaryocyte	9.078819	1.85E-50	1.06E-48
1460	ACRBP	Megakaryocyte	6.6324167	9.02E-36	3.69E-34
1461	DCLRE1A	Megakaryocyte	6.42619	1.49E-43	7.32E-42
1462	ALOX12	Megakaryocyte	10.561812	4.87E-81	3.64E-79
1463	DGKI	Megakaryocyte	7.7921863	2.14E-29	7.52E-28
1464	RP11-401P9.5	Megakaryocyte	7.4524584	3.03E-23	8.83E-22
1465	CD226	Megakaryocyte	6.1728597	9.61E-50	5.33E-48
1466	EHD3	Megakaryocyte	5.6193547	8.94E-37	3.94E-35
1467	F2RL3	Megakaryocyte	8.6088705	6.08E-44	3.07E-42

353

1468	SAMD14	Megakaryocyte	5.1431355	1.12E-48	6.04E-47
1469	TUBA8	Megakaryocyte	9.075804	1.13E-41	5.26E-40
1470	GNAZ	Megakaryocyte	6.66615	2.51E-32	9.57E-31
1471	RP11-672A2.6	Megakaryocyte	7.9475436	2.39E-23	7.20E-22
1472	AP001189.4	Megakaryocyte	8.980323	1.37E-19	3.47E-18
1473	ABCC3	Megakaryocyte	8.341001	2.25E-23	6.92E-22
1474	RP11-879F14.2	Megakaryocyte	6.5317526	2.13E-28	7.18E-27
1475	GP5	Megakaryocyte	7.540501	5.24E-17	1.20E-15
1476	SYTL4	Megakaryocyte	5.6507773	3.55E-30	1.27E-28
1477	GSAP	Megakaryocyte	4.841238	1.08E-35	4.31E-34
1478	LINC00989	Megakaryocyte	6.9827576	1.83E-32	7.15E-31
1479	BZRAP1-AS1	Megakaryocyte	5.427374	4.47E-36	1.92E-34
1480	LY6G6F	Megakaryocyte	13.321681	3.37E-16	7.42E-15
1481	DNAAF3	Megakaryocyte	7.7141356	2.62E-23	7.77E-22
1482	ZGLP1	Megakaryocyte	5.3502946	2.42E-22	6.50E-21
1483	EGF	Megakaryocyte	10.148702	3.02E-15	6.34E-14
1484	AC113404.1	Megakaryocyte	5.517482	4.30E-32	1.61E-30
1485	ELOVL7	Megakaryocyte	7.4058414	2.66E-31	9.75E-30
1486	DGKG	Megakaryocyte	6.5961337	3.17E-24	9.92E-23
1487	SLC35D3	Megakaryocyte	10.823749	6.52E-14	1.32E-12
1488	TMSB4Y	Megakaryocyte	6.8437366	2.49E-20	6.60E-19
1489	SRC	Megakaryocyte	4.193851	5.49E-36	2.30E-34
1490	HRAT92	Megakaryocyte	7.6443768	3.50E-14	7.17E-13
1491	RP11-81H14.2	Megakaryocyte	7.08472	4.38E-15	9.07E-14
1492	ADCY6	Megakaryocyte	3.6248124	2.72E-29	9.35E-28
1493	GAS2L1	Megakaryocyte	2.8266847	2.66E-20	6.93E-19
1494	SIAE	Megakaryocyte	4.65651	3.34E-23	9.57E-22
1495	MOB3C	Megakaryocyte	3.5272555	1.02E-22	2.83E-21
1496	RP11-672A2.4	Megakaryocyte			
1497	AP003068.23	Megakaryocyte	6.2995667	3.62E-18	8.52E-17
1498	MFAP3L	Megakaryocyte	6.3379583	1.13E-27	3.73E-26
1499	WDR44	Megakaryocyte	2.999032	2.15E-19	5.36E-18
1500	CFAP161	Megakaryocyte	9.435371	3.35E-09	5.99E-08
1501	RP11-930O11.1	Megakaryocyte			
1502	HIST1H3H	Megakaryocyte	4.241618	1.10E-17	2.55E-16
1503	BEND2	Megakaryocyte			
1504	CASS4	Megakaryocyte	4.819186	1.51E-22	4.11E-21
1505	DENND2C	Megakaryocyte	4.380255	8.64E-17	1.96E-15
1506	EFHC2	Megakaryocyte	5.8687973	7.82E-23	2.21E-21
1507	RP11-367G6.3	Megakaryocyte			

1508	CFAP45	Megakaryocyte	7.0398107	6.25E-11	1.18E-09
1509	PTGES3L	Megakaryocyte	7.27459	1.68E-07	2.84E-06
1510	HGD	Megakaryocyte	9.895655	1.18E-19	3.02E-18
1511	KCNQ4	Megakaryocyte	6.5543876	4.17E-08	7.17E-07
1512	EXOC3L4	Megakaryocyte			
1513	SCFD2	Megakaryocyte	2.6138632	5.33E-16	1.14E-14
1514	ZNF185	Megakaryocyte	4.2587337	5.07E-11	9.80E-10
1515	AC114752.3	Megakaryocyte			
1516	SENCR	Megakaryocyte	3.6601527	9.23E-19	2.21E-17
1517	CLCN4	Megakaryocyte	3.7783146	9.76E-16	2.07E-14
1518	ABHD11-AS1	Megakaryocyte	7.2642064	1.60E-07	2.73E-06
1519	GYS2	Megakaryocyte			
1520	ABHD11	Megakaryocyte	2.6711166	5.22E-16	1.14E-14
1521	PRUNE	Megakaryocyte	2.095831	5.36E-11	1.02E-09
1522	LINC00504	Megakaryocyte			
1523	XXbac-BPG32J3.19	Megakaryocyte			
1524	LIPH	Megakaryocyte			
1525	C15orf53	Megakaryocyte	8.808073	3.41E-08	5.92E-07
1526	GLOD5	Megakaryocyte			
1527	XYLT2	Megakaryocyte	3.2646177	8.68E-19	2.10E-17
1528	HEXIM2	Megakaryocyte	2.309031	8.27E-10	1.51E-08
1529	LEFTY1	Megakaryocyte			
1530	METTL22	Megakaryocyte	2.242397	6.10E-09	1.08E-07
1531	SPDYC	Megakaryocyte			
1532	NCKIPSD	Megakaryocyte			
1533	ITPKA	Megakaryocyte			
1534	USP20	Megakaryocyte			
1535	MLH3	Megakaryocyte			
1536	HTR2A	Megakaryocyte			
1537	PLCH1	Megakaryocyte			
1538	RYR2	Megakaryocyte			
1539	RP3-329E20.2	Megakaryocyte			
1540	MYEOV	Megakaryocyte			
1541	AC015969.3	Megakaryocyte			
1542	RSPH14	Megakaryocyte			
1543	PDZK1IP1	Megakaryocyte			
1544	AC137932.6	Megakaryocyte			
1545	RP11-672A2.5	Megakaryocyte			
1546	CMTM2	Megakaryocyte			

1547	TRAPPC3L	Megakaryocyte
1548	SLC25A15	Megakaryocyte
1549	MYOM1	Megakaryocyte
1550	NT5M	Megakaryocyte
1551	FAM63A	Megakaryocyte
1552	POTEE	Megakaryocyte
1553	HRASLS	Megakaryocyte
1554	INHBA-AS1	Megakaryocyte
1555	CA13	Megakaryocyte
1556	RP11-1398P2.1	Megakaryocyte
1557	LINC01003	Megakaryocyte
1558	POTEF	Megakaryocyte
1559	TUBAL3	Megakaryocyte
1560	PROSER2	Megakaryocyte
1561	INPP5B	Megakaryocyte
1562	TEC	Megakaryocyte
1563	SEC14L5	Megakaryocyte
1564	PCYOX1L	Megakaryocyte
1565	ACCS	Megakaryocyte
1566	TTC7B	Megakaryocyte
1567	MED12L	Megakaryocyte
1568	TNFSF4	Megakaryocyte
1569	CTB-55O6.8	Megakaryocyte
1570	RP11-888D10.3	Megakaryocyte
1571	CXCL5	Megakaryocyte
1572	MPL	Megakaryocyte
1573	BORCS6	Megakaryocyte
1574	PTPRN	Megakaryocyte
1575	LRP12	Megakaryocyte
1576	CATSPER1	Megakaryocyte
1577	LINC01137	Megakaryocyte
1578	ZBTB11-AS1	Megakaryocyte
1579	GCLM	Megakaryocyte
1580	FAM73B	Megakaryocyte
1581	CDHR1	Megakaryocyte
1582	TMEM104	Megakaryocyte
1583	CTD-3214H19.6	Megakaryocyte
1584	RP11-446N19.1	Megakaryocyte
1585	CTNS	Megakaryocyte
1586	RP11-556E13.1	Megakaryocyte

1587	LSMEM1	Megakaryocyte			
1588	ZNF778	Megakaryocyte			
1589	LINC01001	Megakaryocyte			
1590	PEX3	Megakaryocyte			
1591	ZNF175	Megakaryocyte			
1592	AXIN1	Megakaryocyte			
1593	CHADL	Megakaryocyte			
1594	PANX1	Megakaryocyte			
1595	FUT8-AS1	Megakaryocyte			
1596	RP11-395G23.3	Megakaryocyte			
1597	INAFM2	Megakaryocyte			
1598	APOBEC3F	Megakaryocyte			
1599	ZG16B	Megakaryocyte			
1600	CLC	Eo/baso/mast cell progenitor	15.153702	1.57E-85	2.71E-82
1601	ATP10D	Eo/baso/mast cell progenitor	5.8443913	5.66E-34	2.43E-31
1602	HCAR3	Eo/baso/mast cell progenitor	7.7753925	4.68E-40	2.69E-37
1603	SAMSN1	Eo/baso/mast cell progenitor	6.274184	4.04E-66	3.47E-63
1604	CCNA1	Eo/baso/mast cell progenitor	8.274321	2.72E-08	3.12E-06
1605	CR1	Eo/baso/mast cell progenitor	5.2911367	3.93E-09	5.63E-07
1606	CCR3	Eo/baso/mast cell progenitor	8.970312	0.00205832	0.116054226
1607	IL4	Eo/baso/mast cell progenitor	10.1209	1.18E-07	1.19E-05
1608	CYP11A1	Eo/baso/mast cell progenitor	5.68971	0.002091675	0.116054226
1609	RP11-462G2.1	Eo/baso/mast cell progenitor	9.963682	0.003267888	0.160593372
1610	EPX	Mast cell (earliest)	11.190589	4.09E-26	1.55E-23
1611	IL13	Mast cell (earliest)	8.715403	1.01E-27	5.76E-25
1612	TNFRSF9	Mast cell (earliest)	6.016863	2.08E-14	3.98E-12
1613	NRIP3	Mast cell (earliest)	3.632653	1.51E-08	1.44E-06
1614	CDK15	Mast cell (earliest)	6.821331	9.18E-11	1.05E-08
1615	CCDC71L	Mast cell (earliest)	3.717658	1.68E-06	0.000137597
1616	AC004381.6	Mast cell (earliest)	3.9229467	1.89E-09	1.91E-07
1617	SLC38A11	Mast cell (medium)	6.305733	3.46E-26	3.13E-24
1618	CMA1	Mast cell (medium)	8.483405	5.23E-61	1.29E-58
1619	PLA2G2A	Mast cell (medium)	6.2974267	4.76E-34	5.85E-32
1620	NDST2	Mast cell (medium)	4.9148073	3.16E-34	4.19E-32
1621	RP11-557H15.4	Mast cell (medium)	7.610438	4.21E-52	8.04E-50
1622	TPSG1	Mast cell (medium)	8.768275	1.21E-10	4.75E-09
1623	MAOB	Mast cell (medium)	6.7470636	6.17E-70	1.77E-67
1624	MIR202HG	Mast cell (medium)	8.673044	6.75E-22	5.05E-20
1625	LAX1	Mast cell (medium)	5.2844996	6.81E-50	1.17E-47
1626	GDF10	Early LE	6.542229	7.17E-26	1.76E-23

1627	SV2C	Early LE	9.155369	6.47E-23	1.36E-20
1628	TCTN3	Early LE	3.9209301	2.07E-30	5.95E-28
1629	RADIL	Early LE	6.7712703	1.33E-33	4.58E-31
1630	SBSPON	Early LE	5.71437	9.96E-34	4.28E-31
1631	SCG3	Early LE	9.136127	7.13E-23	1.36E-20
1632	RP11-520P18.5	Early LE	27.391434	0.002637093	0.058151285
1633	PROX1	Early LE	7.9311676	1.47E-42	8.40E-40
1634	FOXC2	Early LE	5.7907405	1.70E-15	2.25E-13
1635	BHMT	Early LE	9.992049	3.05E-13	3.28E-11
1636	ABCA4	Early LE	7.5122	1.58E-10	1.51E-08
1637	SLC41A1	Early LE	3.888286	1.95E-13	2.23E-11
1638	NPY1R	Early LE	5.802308	1.30E-08	8.63E-07
1639	MSMP	Early LE	8.499233	7.39E-09	5.30E-07
1640	TFF2	Early LE			
1641	BHMT2	Early LE	5.143732	1.67E-09	1.25E-07
1642	AC073254.1	Early LE	5.994912	0.001463308	0.03647667
1643	PCLO	Early LE	5.810251	4.82E-08	3.07E-06
1644	C11orf87	Early LE	3.574652	0.006878502	0.13444345
1645	PRR34-AS1	Early LE	4.5602393	5.59E-06	0.000267288
1646	GCOM1	Early LE	7.334697	0.010413981	0.184660282
1647	APTR	Early LE	2.0448604	5.42E-06	0.000266148
1648	PM20D1	Early LE			
1649	CCDC141	Early LE	6.4960504	7.75E-10	6.67E-08
1650	GPER1	Early LE	3.731354	0.002766649	0.060235902
1651	NRG3	Early LE	5.2925673	0.011915844	0.202923283
1652	NTN1	Early LE	2.6402144	0.000220639	0.007744886
1653	ROPN1L	Early LE			
1654	IQCD	Early LE			
1655	NTS	Early LE	11.377788	4.67E-44	4.44E-41
1656	AC008063.2	Early LE			
1657	ZNF678	Early LE	2.2653077	0.000459251	0.013858109
1658	RAI2	Early LE	3.4486785	0.000394288	0.012110279
1659	HSD3B7	Early LE	2.6987479	3.12E-05	0.0012764
1660	IGHMBP2	Early LE	2.2722073	0.000137708	0.005039532
1661	TBC1D19	Early LE	1.9092246	0.002072777	0.047714673
1662	STRADA	Early LE	2.5398538	0.001067626	0.028692443
1663	HIVEP2	Early LE	2.411641	0.002300586	0.052065888
1664	GHDC	Early LE	2.44813	0.000282315	0.009338123
1665	TNFAIP8L3	Early LE	5.8033967	1.23E-06	7.05E-05
1666	VSIG10	Early LE	2.3709795	0.007672244	0.143437598

1667	RP11-251M1.1	Early LE	5.545965	0.000343428	0.010938822
1668	GRM3	Early LE	4.034133	0.007624948	0.143437598
1669	LINC00636	Early LE	7.8363314	2.55E-06	0.000132866
1670	ABLIM2	Early LE			
1671	ABCC10	Early LE			
1672	DYRK1B	Early LE	2.1624742	0.000519305	0.015400076
1673	LRP11	Early LE	2.288066	0.001004987	0.027880292
1674	PCDHB11	Early LE			
1675	ALKBH6	Early LE	1.7681053	0.008506633	0.157326983
1676	MYZAP	Early LE	6.0645404	9.17E-10	7.51E-08
1677	CELSR1	Early LE	6.7981157	6.84E-10	6.19E-08
1678	RARA-AS1	Early LE	2.5594876	0.005345322	0.112121399
1679	ZBTB46	Early LE	3.6847315	0.000614892	0.017925666
1680	SYNM	Early LE	5.4308367	2.76E-05	0.001188448
1681	CARD6	Early LE	3.936775	0.001191008	0.031515898
1682	SPATA6L	Early LE	5.145594	0.006239649	0.124792971
1683	ZSCAN2	Early LE	2.3208382	0.002336141	0.052183924
1684	RP11-483L5.1	Early LE	4.733398	0.001754212	0.042231632
1685	KIF17	Early LE	4.095448	4.57E-05	0.001829375
1686	ATXN1L	Early LE	1.1860789	0.013136666	0.219369561
1687	SGIP1	Early LE	5.2865343	2.85E-05	0.001195481
1688	SLC24A1	Early LE	3.7753928	0.001500703	0.03687442
1689	NALCN	Early LE	4.414167	0.00096746	0.027279207
1690	NIPAL2	Early LE	3.1272771	0.000125579	0.004799923
1691	STAB2	LE	13.551779	0.00E+00	3.20E-306
1692	IL7	LE	10.211235	1.50E-102	5.17E-100
1693	SCNN1B	LE	10.906203	3.25E-53	3.73E-51
1694	GPR182	LE	12.402483	8.97E-49	9.64E-47
1695	ELMOD1	LE	5.9476013	1.56E-36	1.12E-34
1696	RP11-322E11.5	LE	9.294467	5.45E-66	9.38E-64
1697	NNMT	LE	4.9140167	8.73E-76	2.14E-73
1698	C6orf141	LE	8.659012	2.25E-61	3.51E-59
1699	LINC01558	LE	8.937398	1.03E-82	2.94E-80
1700	GPR1	LE	5.543996	6.21E-14	1.48E-12
1701	SLC30A3	LE	8.017138	1.41E-34	8.95E-33
1702	BCL2L10	LE	7.969268	2.94E-16	8.02E-15
1703	NFIA-AS2	LE	6.412098	1.10E-37	9.04E-36
1704	PRR5L	LE	4.2250304	1.33E-29	6.56E-28
1705	ZP2	LE	10.043701	5.13E-32	2.85E-30
1706	PLIN5	LE	6.8770056	4.43E-35	2.93E-33

1707	P4HA3	LE	3.8428628	3.27E-17	9.38E-16
1708	LRCH2	LE	2.7015288	2.19E-16	6.07E-15
1709	PIEZO2	LE	2.3454216	7.26E-37	5.43E-35
1710	WASF3	LE	3.227173	2.87E-21	1.03E-19
1711	TGFBRAP1	LE	3.3973646	2.70E-28	1.22E-26
1712	HOXD10	LE	8.2465725	5.19E-33	2.97E-31
1713	OMG	LE	8.22318	1.10E-17	3.27E-16
1714	KCNIP1	LE	4.2105174	9.16E-21	3.15E-19
1715	GALNT15	LE	5.4460673	6.93E-21	2.43E-19
1716	UBAP1L	LE	5.2864733	8.89E-18	2.68E-16
1717	KLKB1	LE			
1718	TNFRSF11A	LE	4.3629365	2.68E-53	3.29E-51
1719	GIMAP5	LE	4.27506	2.54E-30	1.32E-28
1720	SLC35E4	LE	3.1290302	2.40E-11	5.09E-10
1721	TFF3	LE	12.020079	0	0
1722	RP11-544L8__B.4	LE	5.356959	6.92E-16	1.80E-14
1723	RSPH4A	LE	5.4239545	4.10E-11	8.60E-10
1724	STOX2	LE	2.6335454	8.12E-23	3.33E-21
1725	SLC26A4	LE	5.748854	2.02E-08	3.81E-07
1726	CRTAC1	LE	5.7417526	3.12E-08	5.83E-07
1727	PGM5-AS1-1	LE	4.825012	2.42E-22	9.30E-21
1728	TAC3	LE	4.799044	1.90E-13	4.35E-12
1729	ACKR2	LE	6.8842916	1.50E-28	7.19E-27
1730	ZDHHC14	LE	3.360037	2.43E-22	9.30E-21
1731	DTX4	LE	3.1244316	7.56E-22	2.83E-20
1732	SLC38A5	Early endothelial cells	8.195626	6.79E-87	1.17E-83
1733	MYLK2	Early endothelial cells	8.688176	1.24E-25	2.14E-23
1734	C1orf64	Early endothelial cells	10.499416	2.08E-26	4.47E-24
1735	SLC38A3	Early endothelial cells	5.967941	3.09E-51	2.65E-48
1736	SPINK8	Early endothelial cells	12.4607315	3.95E-27	9.70E-25
1737	XIRP2	Early endothelial cells	7.7817507	9.27E-19	1.06E-16
1738	PPARD	Early endothelial cells	3.6972904	2.01E-33	8.65E-31
1739	QRFPR	Early endothelial cells	9.036257	1.06E-23	1.66E-21
1740	FLVCR2	Early endothelial cells	6.684845	2.94E-29	1.01E-26
1741	UCP1	Early endothelial cells	10.017733	3.28E-13	2.35E-11
1742	FOXF2	Early endothelial cells	5.8664684	1.21E-14	1.10E-12
1743	HAPLN1	Early endothelial cells	4.900504	3.93E-19	4.83E-17
1744	FOXL2NB	Early endothelial cells	7.787049	0.000281723	0.006730058
1745	PAQR6	Early endothelial cells	5.286531	4.50E-23	6.44E-21

1746	CA5A	Early endothelial cells	6.153009	5.53E-05	0.001442349
1747	ERVMER34-1	Early endothelial cells	4.629756	0.002024148	0.041446846
1748	ADORA2A	Early endothelial cells	5.1478434	4.81E-10	2.36E-08
1749	ZBTB2	Early endothelial cells	2.1418488	4.69E-07	1.61E-05
1750	SLC25A44	Early endothelial cells	2.5618742	6.07E-08	2.37E-06
1751	PORCN	Early endothelial cells	2.7856355	3.27E-07	1.15E-05
1752	MATN3	Early endothelial cells	4.43668	8.75E-16	8.85E-14
1753	FAM110D	Early endothelial cells	6.982435	1.83E-06	6.04E-05
1754	ZNF366	Early endothelial cells	6.700243	3.44E-09	1.60E-07
1755	KIAA1161	Early endothelial cells	6.278667	1.48E-11	9.11E-10
1756	AC017002.2	Early endothelial cells	4.1159353	6.22E-05	0.001596934
1757	NKAIN1	Early endothelial cells	4.484429	2.64E-07	9.45E-06
1758	ABHD17C	Early endothelial cells	4.917722	7.97E-07	2.69E-05
1759	ABCB1	Early endothelial cells	4.854517	6.22E-13	4.12E-11
1760	PITPNA	Early endothelial cells	2.4306269	1.33E-08	5.87E-07
1761	B3GALT1	Early endothelial cells			
1762	EXOC2	Early endothelial cells	2.654748	4.27E-11	2.45E-09
1763	KCNC4	Early endothelial cells	5.4301667	3.38E-05	0.000922957
1764	MYRIP	Early endothelial cells	7.1237864	1.63E-13	1.27E-11
1765	GLYCTK	Early endothelial cells	2.7338207	1.71E-07	6.26E-06
1766	RGS17	Early endothelial cells	3.5403097	9.97E-10	4.77E-08
1767	SDCBP2	Early endothelial cells	2.4190266	0.001098632	0.023620577
1768	EPHX4	Early endothelial cells	6.473838	1.95E-06	6.22E-05
1769	FOXL2	Early endothelial cells			
1770	TOR1B	Early endothelial cells	1.5198246	1.02E-05	0.000296396
1771	PLEKHF1	Early endothelial cells	3.2237906	9.90E-08	3.70E-06
1772	DCLK2	Early endothelial cells	3.4322536	4.14E-08	1.73E-06
1773	CHST2	Early endothelial cells	1.952452	7.51E-08	2.87E-06
1774	KLF11	Early endothelial cells	2.0785902	1.91E-06	6.21E-05
1775	SHROOM1	Early endothelial cells	4.1083655	1.00E-10	5.37E-09
1776	PANK4	Early endothelial cells	1.8891857	6.39E-05	0.001616378
1777	IGLON5	Early endothelial cells	3.0523956	2.19E-05	0.00060774
1778	DCBLD1	Early endothelial cells	3.9589157	2.19E-08	9.42E-07
1779	CHST1	Capillary arterioles	7.9215446	8.85E-36	5.07E-33
1780	NETO2	Capillary arterioles	4.702522	1.85E-24	5.29E-22
1781	PRND	Capillary arterioles	10.694583	1.64E-86	2.83E-83
1782	PRRG3	Capillary arterioles	3.7570074	3.39E-17	6.48E-15
1783	CA4	Capillary arterioles	4.5944743	7.19E-06	0.000353346
1784	CREG2	Capillary arterioles	8.239782	3.49E-08	2.59E-06
1785	RP1-16A9.1	Capillary arterioles	8.586419	9.42E-30	4.05E-27

1786	LY86-AS1	Capillary arterioles	6.063576	2.18E-08	1.70E-06
1787	WSCD1	Capillary arterioles	5.8307996	1.43E-11	1.64E-09
1788	CHRM3	Capillary arterioles	7.0676107	1.17E-13	1.68E-11
1789	LOXHD1	Capillary arterioles	7.433704	0.001061189	0.033800821
1790	TCFL5	Capillary arterioles	3.7150857	0.000601413	0.02060492
1791	LTBP2	Capillary arterioles	4.0878954	1.49E-09	1.43E-07
1792	AC022182.3	Capillary arterioles	6.5193534	0.000211789	0.008884823
1793	SCN11A	Capillary arterioles	6.8877263	1.77E-08	1.45E-06
1794	RP11-768F21.1	Capillary arterioles	6.354223	0.000676233	0.022367722
1795	PITPNM2	Capillary arterioles	3.8143613	1.44E-05	0.000686571
1796	EFCAB14	Capillary arterioles	4.9066772	1.98E-77	1.70E-74
1797	NOX4	Capillary arterioles	4.7321305	6.79E-12	8.34E-10
1798	NOS2	Capillary arterioles	7.9240837	0.00061096	0.02060492
1799	LIMD1	Capillary arterioles	2.1982605	1.27E-08	1.09E-06
1800	ANKRD40	Capillary arterioles	1.5691677	0.000297726	0.011909025
1801	KIAA0040	Capillary arterioles	3.4445748	1.09E-10	1.10E-08
1802	CASC8	Capillary arterioles	6.158759	0.004923552	0.122732015
1803	TMEM233	Capillary arterioles	5.348594	5.04E-11	5.42E-09
1804	CDC34	Capillary arterioles	1.8988698	4.99E-06	0.000252653
1805	MAPK11	Capillary arterioles	5.048877	2.64E-09	2.39E-07
1806	MXRA7	Capillary arterioles	2.237872	2.75E-06	0.00014771
1807	GJA5	Arterioles	8.603976	3.29E-34	1.42E-31
1808	VEGFC	Arterioles	5.005253	4.41E-21	9.48E-19
1809	PIK3C2B	Arterioles	3.8401318	1.45E-13	2.49E-11
1810	WWC2	Arterioles	2.771025	3.45E-07	2.37E-05
1811	CCDC178	Arterioles	6.107472	2.75E-09	2.36E-07
1812	FCN3	Arterioles	8.614598	0.000136962	0.00654374
1813	DAB2IP	Arterioles	3.4196978	9.66E-09	7.91E-07
1814	PRSS51	Arterioles	11.36725	0.000602892	0.026589083
1815	SMAD7	Arterioles	4.0553007	1.19E-11	1.28E-09
1816	IMPG2	Arterioles	6.657661	0.004187179	0.156564073
1817	GPR4	Arterioles	6.594894	4.08E-10	3.69E-08
1818	AVPR2	Capillaries	8.079218	6.00E-15	7.37E-13
1819	SELE	Postcapillary venules	10.867261	1.20E-202	2.07E-199
1820	GIPC2	Postcapillary venules	6.4238772	1.70E-43	4.19E-41
1821	AOC2	Postcapillary venules	7.056947	7.25E-15	5.94E-13
1822	TACR1	Postcapillary venules	6.827874	1.88E-07	6.74E-06
1823	C2CD4B	Postcapillary venules	7.886147	2.01E-38	3.46E-36
1824	FAM184A	Postcapillary venules	3.9586868	4.59E-18	4.16E-16
1825	IFI27	Postcapillary venules	6.4433875	3.70E-148	3.18E-145

1826	FBLN2	Venules	5.2530994	4.99E-32	8.58E-29
1827	PDIA5	Venules	5.0467205	4.16E-29	3.58E-26
1828	RAMP3	Venules	7.6692367	1.46E-26	8.38E-24
1829	AKR1C1	Venules	7.080758	9.13E-21	3.93E-18
1830	HTR2B	Venules	9.234	2.24E-05	0.002409836
1831	LYPD2	Venules	12.155155	2.01E-09	4.33E-07
1832	CLRN3	Venules	10.728618	2.08E-06	0.000289784
1833	AKR1C3	Venules	7.1477304	2.82E-09	5.39E-07
1834	FRMPD4	Venules	8.023091	0.00619109	0.306116508
1835	MEDAG	Venules	4.288007	2.82E-05	0.00285423
1836	GDF3	Venules	9.54783	0.000477202	0.034199451
1837	GATA6-AS1	Venules	6.475415	0.013592414	0.615235589
1838	SMAD9	Venules	4.057689	1.20E-05	0.001472572
1839	BMPER	Venules	5.167112	1.50E-07	2.34E-05
1840	RAPGEF4	Venules	6.2292843	8.13E-05	0.007446846
1841	ATOH8	Venules	2.689458	0.003234403	0.17945722
1842	HSPB8	Venules	5.6114726	0.000647081	0.042806898
1843	LIN7A	Venules	2.5459445	0.002776077	0.159161739
1844	MEOX2-AS1	Venules	5.3551297	0.00736926	0.342570994
1845	DDAH1	Venules	2.8765793	0.006229115	0.306116508
1846	ACER2	Venules	5.804265	0.049162003	0.999435306
1847	SMURF2	Venules	1.70339	0.004679461	0.243899177