Mapping the Alloimmune Effector Cells in Graft Versus Host Disease

Yu Kin Jason Lam

Doctor of Philosophy

April 2023

Translational and Clinical Research Institute

Newcastle University

Supervisor: Professor Matthew Collin

Abstract

Graft versus host disease (GVHD) is a life-threatening complication of haematopoietic stem cell transplantation. It is conventionally described as a process mediated by donor immune cells attacking healthy cells of the transplant recipient, causing inflammation in different organs including the skin. However, the cell types involved in the process and their immune mechanisms remain poorly understood.

To define the immune landscape of GVHD, I collected paired blood samples and skin biopsies from patients with GVHD (n = 5) at intervals post transplantation (Day 19 – 261) and compared them with transplant controls without GVHD (n = 2) and healthy donors (paired n = 2, skin unpaired n = 2). Single-cell RNA sequencing was performed on peripheral blood mononuclear cells and skin, which was enzymatically split into dermis and epidermis. Using the 10x Genomics platform with 5' chemistry, a total of 173,168 transcriptomes were obtained and their genotypes (donor or recipient origin) were deconvoluted. T cell receptors were also sequenced to investigate the dynamics of T cell activation across tissues.

The cells from each anatomical compartment were defined and compared between conditions. The cellular composition of GVHD-affected tissues was vastly altered, with pronounced infiltration of the skin by donor derived cells from both myeloid and lymphoid lineages. Detailed myeloid cell characterisation identified macrophage and dendritic cell subsets, including resident macrophages and DC3. Inflammatory macrophages and dendritic cells were significantly expanded in GVHD. GVHD-specific features were also evident, which included the presence of plasmacytoid dendritic cells, as well as the expression of an activation signature among dendritic cells. Within the lymphoid compartment, GVHD featured the infiltration of predominately donor-derived, clonally expanded T cells from the blood into the dermis and epidermis. Clonal expansion of recipient-derived T cell was also observed in one case, indicating a potential host versus graft reaction. This study provides a high-resolution analysis of the complex cross-tissue immune reaction in GVHD.

Declaration

This thesis was composed by myself, that the work contained herein is my own with the following exceptions:

Sequencing of libraries generated from experiments were performed by members of staff from the Newcastle Genomics Core Facility. Pre-processing of raw sequencing data, and variant identification from whole exome sequence data were performed by members of staff from the Newcastle Bioinformatics Support Unit.

Clinical details of patients were collected by Dr Callum Wright.

Histological examination of skin biopsies was performed by Dr James Sampson.

Flow cytometry panel was designed by other members of the Haematopoiesis and Immunity Lab.

Acknowledgements

List of Abbreviations

ADP	Adenosine diphosphate
ALL	Acute lymphoblastic leukaemia
AML	Acute myeloid leukaemia
ATP	Adenosine triphosphate
CCL	C-C motif chemokine ligand
CCR	C-C motif chemokine receptor
CD	Cluster of differentiation
cDC1	Classical DC 1
cDC2	Classical DC 2
CITE-seq	Cellular indexing of transcriptomes and epitopes by sequencing
CLL	Chronic lymphocytic leukaemia
CLR	C-type lectin receptors
CM	Central memory
CML	Chronic myeloid leukaemia
CMML	Chronic myelomonocytic leukaemia
CMV	Cytomegalovirus
CSF-1	Colony stimulating factor 1
CXCR	C-X-C motif chemokine receptor
Cy	Cyclophosphamide
DAMP	Danger-associated molecular pattern
DC	Dendritic cell
DLI	Donor lymphocyte infusion
EBV	Epstein-Barr virus
EDTA	Ethylenediaminetetraacetic acid
EM	Effector memory
FCS	Foetal calf serum
G-CSF	Granulocyte colony-stimulating factor
GI	Gastrointestinal
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GVHD	Graft versus host disease

GVL	Graft versus leukaemia
Haplo	Haploidentical
HLA	Human leukocyte antigen
HSC	Haematopoietic stem cell
HSCT	Haematopoietic stem cell transplantation
IFN	Interferon
Ig	Immunoglobulin
IL	Interleukin
ILC	Innate lymphoid cell
LC	Langerhans cell
LE	Lymphatic endothelial cell
LPS	Lipopolysaccharides
MAIT	Mucosal-associated invariant T cell
M-CSF	Macrophage colony-stimulating factor
MDS	Myelodysplastic syndrome
MHC	Major histocompatibility complex
migDC	Migratory dendritic cell
miHA	Minor histocompatibility antigen
MM	Multiple myeloma
mregDC	Mature dendritic cell enriched in immunoregulatory molecules
NK	Natural killer cell
PAMP	Pathogen-associated molecular pattern
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate-buffered saline
PCA	Principal component analysis
pDC	Plasmacytoid dendritic cell
PRR	Pattern recognition receptor
PT	Post-transplant
Rec.mac	Recruited macrophage
Res.mac	Resident macrophage
RIC	Reduced-intensity conditioning
scRNA-seq	Single-cell RNA sequencing

Sib	Sibling
Super.mac	Super-activated macrophage
TBI	Total body irradiation
TCR	T cell receptor
TGF-β	Transforming growth factor beta
Th1	T helper cell type 1
Th2	T helper cell type 2
Th17	T helper cell type 17
TLR	Toll-like receptor
TNF	Tumour necrosis factor
Treg	Regulatory T cell
Trm	Tissue-resident memory T cell
t-SNE	t-distributed stochastic neighbour embedding
UMAP	Uniform Manifold Approximation and Projection
UMI	Unique molecular identifier
VE	Vascular endothelial cell

List of Tables and Figures

Chapter 4

Appendix

Table of Contents

Chapter 1. Introduction

1.1 Haematopoietic stem cell transplantation

1.1.1 Background

Haematopoietic stem cell transplantation (HSCT) is a curative therapy for haematological malignancies, congenital haematological disorders and immunological diseases (Snowden *et al.*, 2022). The goal of HSCT is to re-establish normal haematopoiesis and, in the case of haematological malignancy, to generate curative donor alloimmunity. There are two main types of HSCT, namely allogeneic (stem cells are harvested from another individual) and autologous (stem cells are harvested from the recipient) HSCT. The indications of these two types of HSCT are distinctive. Autologous HSCT is out of the scope of this thesis so herein only allogeneic HSCT will be discussed.

In adults, the majority of allogeneic HSCTs are performed for the treatment of acute myeloid leukaemia (AML), accounting for more than one-third of all allogeneic HSCTs (Snowden *et al.*, 2022). Other indications include acute lymphoblastic leukaemia (ALL), myelodysplastic syndrome (MDS), lymphomas, and to a lesser extent, chronic myeloid leukaemia (CML), chronic lymphocytic leukaemia (CLL) and multiple myeloma (MM) (Snowden *et al.*, 2022).

Before the transplant procedure, a genetically compatible stem cell source must be identified. Allogeneic HSCT requires stem cells from a healthy donor with compatible human leukocyte antigens (HLA). The HLA system is the major histocompatibility complex (MHC) in humans which codes for a set of proteins responsible for the regulation of immune response (Bjorkman *et al.*, 1987). The HLA system is especially relevant in the context of transplantation immunology, as these proteins differentiate cells between self and non-self. The degree of HLA compatibility is associated with post-transplant complications such as graft versus host disease (GVHD) (Kawase *et al.*, 2007). The role of HLA genes in GVHD will be further elaborated in subsequent sections.

Common source of stem cells used in HSCT includes bone marrow, peripheral blood stem cells and cord blood from an HLA-matched or haploidentical donor (Cheuk, 2013). Being less invasive than bone marrow harvest, granulocyte colony-stimulating factor (G-CSF) mobilized

peripheral blood stem cells are most commonly used in modern HSCT (Gratwohl *et al.*, 2013). The source of stem cells affects the transplant outcome. For example, comparing to the transplantation of bone marrow, transplantation of peripheral blood stem cells is associated with lower rates of graft failure but higher incidence of GVHD (Anasetti *et al.*, 2012).

To prepare a patient for HSCT, conditioning regimens are administered which serves a dual purpose: to eliminate tumour cells and suppress the immune system of the host to promote donor stem cell engraftment. In the past, conditioning regimens are myeloablative and involve high doses of chemotherapy and radiation to fully eradicate the host immune cells. However, myeloablative conditioning regimens are associated with significant toxicity and high non-relapse mortality (Song *et al.*, 2021). As the benefits of donor immune cells against malignant host cells were being recognized, reduced-intensity conditioning (RIC) regimens have been developed. These non-myeloablative regimens are more tolerable and safer, therefore opening the opportunity for patients with older age and comorbidities to receive HSCT, without compromising the overall survival rate (Song *et al.*, 2021). Once the conditioning regimen is complete, the patient is infused with stem cells, and prophylaxis is administered to minimize the risk of complications including infection and GVHD.

1.1.2 Immune reconstitution

Effective immune reconstitution contributes to the success of HSCT by protection against opportunistic infection and clearance of residual malignant disease. After stem cell infusion, different cell types follow a predictable pattern of recovery, with the pace being influenced by a multitude of factors (Figure 1.1). In general, the recovery of innate immune cells precedes adaptive immune cells. Recipients remain pancytopenic for approximately 14 days post-transplantation (Stern *et al.*, 2018). Neutrophils reconstitute first, followed by monocytes and dendritic cells (DCs) around the first month post-transplantation (Stern *et al.*, 2018). NK cells are the first lymphocyte to recover, while T cells and B cells can remain deficient throughout the first year post-transplantation (Stern *et al.*, 2018).

Figure 1.1
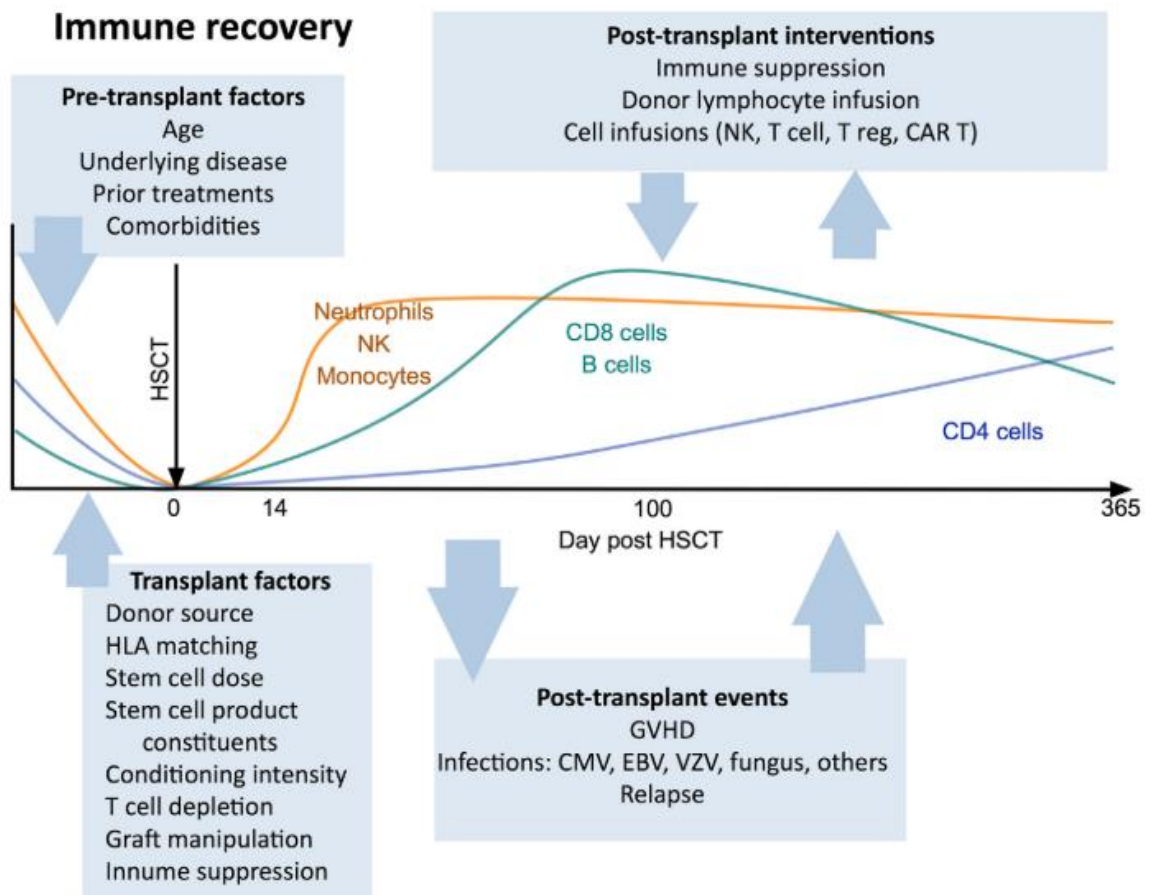
Figure 1.1 Pattern of immune reconstitution. Major immune cell subsets follow a predictable pattern of recovery. This is influenced by the underlying condition of the patient, transplant protocol, and post-transplant events such as infections and GVHD. CMV = cytomegalovirus; EBV = Epstein-Barr virus; VZV = varicella-zoster virus; CAR T = chimeric antigen receptor T cells. Image from Stern *et al.,* 2018.

The initial reconstitution of T cells relies on the oligoclonal expansion of peripheral mature donor T cells (Simons, Cavazzana and André, 2019). However, the recovery of a fully immune-competent, diverse T cell repertoire requires bone marrow-derived T cell progenitors to undergo selection processes in the thymus (Krenger and Holländer, 2008; Simons, Cavazzana and André, 2019). This process is further prolonged in patients with GVHD-induced thymic injury (Krenger and Holländer, 2008; Simons, Cavazzana and André, 2019). GVHD is also associated with impaired reconstitution and function of B cells (Storek *et al.*, 2001), NK cells (Bunting *et al.*, 2017), and DCs (Reddy *et al.*, 2004; Markey *et al.*, 2012). The impairment of

antigen presentation functions of DCs during GVHD causes further immune dysregulation by defective regulatory T cell (Treg) generation (Leveque-El Mouttie *et al.*, 2016).

1.1.3 Graft versus leukaemia

An important component of successful HSCT relies on donor immune cells to eliminate residual malignant disease. This immune-mediated response is known as graft-versus-leukaemia (GVL), and it is critical to minimizing relapse (Kanda *et al.*, 2012). However, the alloimmune response may also be capable of recognizing epitopes on healthy tissues, causing inflammation and tissue damage. This phenomenon is termed GVHD. Approximately 20% of transplant recipients die from GVHD (Styczyński *et al.*, 2020). Other causes of death include relapse and infection, which are also linked to GVHD and its treatments. Although low grade GVHD is associated with better overall survival compared with the absence of GVHD, potentially owing to the beneficial GVL effects, severe GVHD is associated with higher mortality rates (Kanda *et al.*, 2012).

Therefore, much effort has been devoted to exploring HSCT strategies to maximize the GVL effects while minimizing GVHD. Common approaches to induce remission in patients with relapse include lower dose and early withdrawal of immunosuppression, and donor lymphocyte infusion, yet methods are associated with increased risk of GVHD (Kolb *et al.*, 1995; Brandenburg, Gottlieb and Bradstock, 1998). More recently, numerous T and NK cell-based adoptive cell therapies have been tested to enhance the GVL effect while limiting GVHD (Blazar, Hill and Murphy, 2020). Although some have demonstrated promising results, concerns of GVHD risk still remain as molecular targets on leukaemia cells are usually non-specific alloantigens (Blazar, Hill and Murphy, 2020).

1.1.4 Graft versus host disease

GVHD presents in acute or chronic forms. Acute GVHD mainly damages the skin, gastrointestinal (GI) tract and liver, while chronic GVHD has more diverse manifestations and often resemble autoimmune disease (Shlomchik, 2007). Formerly, acute and chronic GVHD were classified by onset within and after 100 days of transplantation, respectively. As it has

been recognized that the expanding transplant practices, such as the use of RIC regimens and donor lymphocyte infusions (DLI), alter the classical onset of acute and chronic manifestations (e.g., symptoms of acute GVHD can appear after 100 days of transplantation and present simultaneously with symptoms of chronic GVHD), the classification system of GVHD has been refined, which is based on the clinical and histopathological features instead of time of symptomatic onset post-transplantation (Table 1.1) (Filipovich *et al.*, 2005). Yet, both forms of GVHD heavily restrict the therapeutic potential of allogeneic HSCT.

Table 1.1

Categories of acute and chronic graft-versus-host disease

Category	Time of symptoms after HSCT or DLI	Presence of acute GVHD features	Presence of chronic GVHD features
Acute GVHD			
Classic acute GVHD	≤100 days	Yes	No
Persistent, recurrent, or late-onset acute GVHD	>100 days	Yes	No
Chronic GVHD			
Classic chronic GVHD	No time limit	No	Yes
Overlap syndrome	No time limit	Yes	Yes

Table 1.1 Refined classification system of GVHD. Table from Strong *et al.*, 2018.

1.2 GVHD pathophysiology

1.2.1 Classical perspective

More than half a century ago, Billingham outlined three critical requirements for the development of GVHD (Billingham, 1966). Firstly, the graft must contain immunologically competent cells. Secondly, the host must express tissue antigens that are not present in the donor. Lastly, the recipient must have a compromised immune system so that the graft is not being rejected. Subsequent experimental work in animals established that donor T cells are

necessary for GVHD (Korngold and Sprent, 1978), with the highly polymorphic HLA proteins being the most relevant tissue antigens (Petersdorf *et al.*, 1995).

Nonetheless, T cells may not be sufficient in GVHD pathology, as GVHD-associated tissue damage can still occur in recipients receiving T cells which are defective in cytotoxic pathways (Baker *et al.*, 1996). This observation has led to the investigation of other potential cellular players involved in GVHD. Macrophages, in particular, have gained attention for being the main producer of inflammatory cytokines that are important mediators of GVHD (Holler *et al.*, 1990; Nestel *et al.*, 1992). Clinically, patients with GVHD can have minimal T cell involvement, but macrophages are almost always present (Nishiwaki *et al.*, 2009). Therefore, GVHD is likely to be mediated not just by T cells, but also by other cell types.

Based on the findings from animal models and clinical data, Ferrara and Deeg have summarized the mechanism of GVHD induction into a three-step process (Ferrara and Deeg, 1991; Ferrara *et al.*, 2009) (Figure 1.2). In the first step, antigen-presenting cells (APCs) are activated. In the second step, donor T cells are activated, and in the third and final step, organ damage occurs.

Figure 1.2
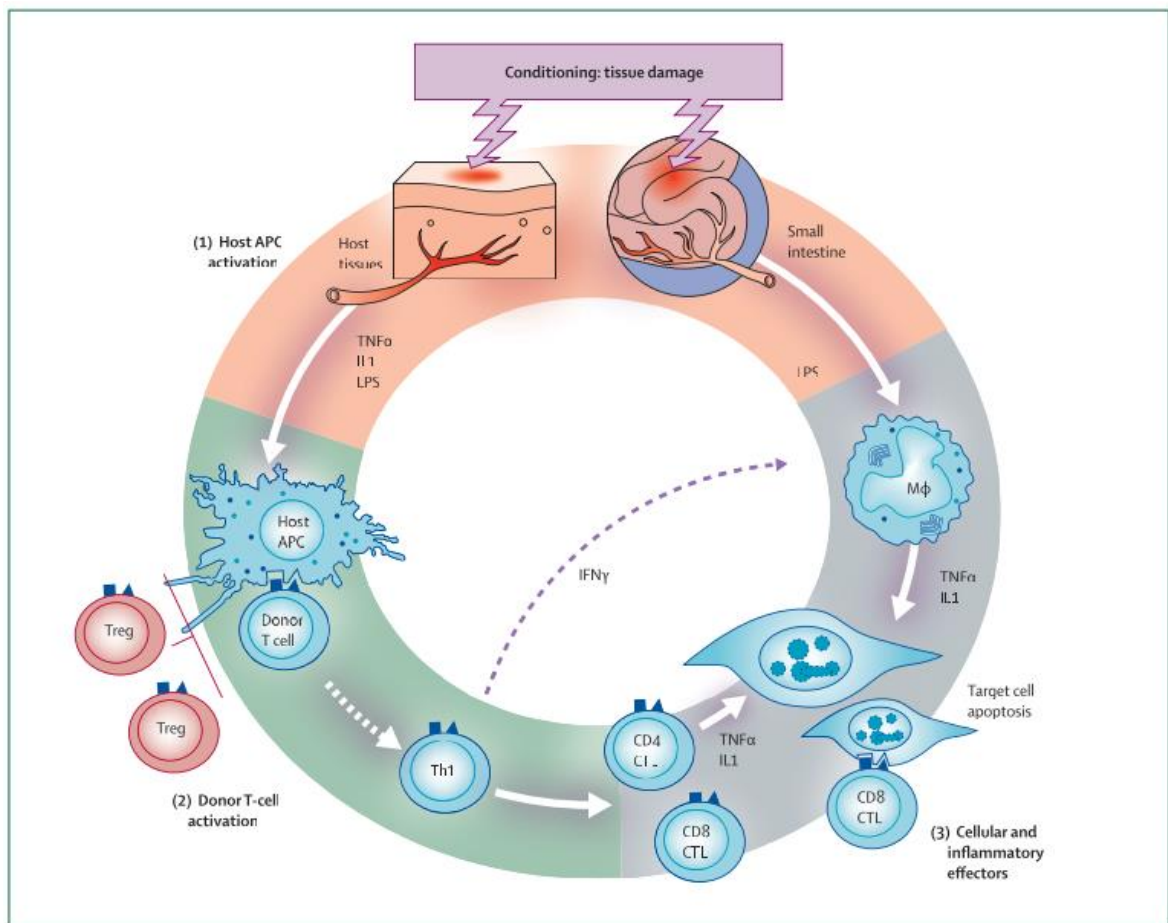
Figure 1.2 Classical pathophysiology of acute GVHD. Three stage model of GVHD pathogenesis: (1) Conditioning damages host tissue and leads to the activation of host APC. (2) Donor T cells are activated by host APC. (3) Cytotoxic lymphocytes and inflammatory cytokines cause host tissue apoptosis. TNFα = tumour necrosis factor alpha; IL1 = interleukin 1; LPS = lipopolysaccharide; IFNγ = interferon gamma; Treg = regulatory T cell; Th1 = T helper cell type 1; CTL = cytotoxic T lymphocyte. Image from Ferrara *et al.*, 2009.

Initially, host APCs are activated as a result of the underlying disease and the transplant conditioning regimen. Conditioning plays a critical role in successful HSCT, however it also collaterally damages host tissues. The destruction of host tissue leads to an excessive production of pro-inflammatory cytokines, including tumour necrosis factor (TNF) and interleukins 1 and 6 (IL-1 and IL-6) (Xun *et al.*, 1994; Hill *et al.*, 1997). Furthermore, conditioning compromises the integrity of the GI tract, leading to the release of bacterial inflammatory stimuli such as lipopolysaccharides (LPS) and other pathogen/danger-associated molecular patterns (PAMPs/DAMPs) into systemic circulation (Hill *et al.*, 1997). These molecules are recognised by Toll-like receptors (TLR) expressed on APCs (Iwasaki and Medzhitov, 2004).

The next step of GVHD pathogenesis involves the priming, expansion, and differentiation of donor alloreactive T cells. Apart from the involvement of the "cytokine storm" generated from the first phase, cognate interaction between donor T cells and host APCs presenting alloantigen is required for T cell activation. In the case of HLA-matched HSCT, peptide recognition by donor T cell is analogous to pathogen-derived antigens. Donor T cells can only recognize HLA-bound peptides that are present in the host but not in the donor. GVHD development after receiving HLA-identical graft is considered to be mainly caused by disparities between donor and host proteins as a result of genetic polymorphism (Chao, 2004). These proteins are known as minor histocompatibility antigens (miHAs). Genes that are only expressed in the host but not the donor (Murata, Warren and Riddell, 2003), as well as the polymorphisms that interfere with peptide processing or post-translational modification (Brickner *et al.*, 2001; Yadav *et al.*, 2003), can also give rise to miHAs. The potency of different miAHs in inducing GVHD varies and shows a hierarchical immunodominance (Perreault *et al.*, 1996). Since the expression pattern of miHAs can differ between tissues (de Bueger *et al.*, 1992), it may dictate the organs targeted by GVHD as direct cell-cell interaction between donor T cell and its target is required for CD8+ T cell-mediated cytotoxicity (Shlomchik, 2007). However, in the case of HLA-mismatched HSCT, both HLA molecule and its associated peptide contribute to donor T cell activation. Therefore, the identity of peptides that promote GVHD are more indiscriminative. In terms of the type of T cells involved, CD4+ T cells and CD8+ T cells are involved in MHC class II and MHC class I mismatches, respectively, and in the case of HLA-matched/miHA-mismatched both T cell subsets are involved (Korngold and Sprent, 1982, 1985). Th1 and its cytokines (IFN-γ, IL-2 and TNF-α) are strongly implicated in gut GVHD (Hill *et al.*, 1997; Yi *et al.*, 2009).

The third and final step of GVHD pathogenesis describes target tissue damage mediated by effector cells and soluble factors. Cytotoxicity has been mainly attributed to T cells. Expression of chemokine receptor CCR5 on T cells promotes homing from lymphoid tissues to target organs (Wysocki *et al.*, 2005). In CD8-mediated GVHD, target tissue injury requires cognate interaction between T cells and the target cells (Matte-Martone *et al.*, 2008). Tissue apoptosis is mediated via pathways including perforin/granzyme and Fas/Fas ligand (Baker *et al.*, 1996; Braun *et al.*, 1996). Soluble factors including TNF-α, IL-1, IFN-γ and nitric oxide, primarily as a result of CD4-mediated GVHD, inflict direct tissue damage independent of any cognate interactions (Piguet *et al.*, 1987; Garside *et al.*, 1992; Teshima *et al.*, 2002). IFN-γ also mediates the effector response of macrophages by priming the production of TNF-α in response to low levels of LPS (Nestel *et al.*, 1992).

1.2.2 Current insights

Ferrara's model of GVHD pathophysiology provided a framework for the key events of GVHD induction. However, some important aspects have yet to be addressed: What are the APCs? Are they donor or recipient in origin? Does antigen presentation only occur in secondary lymphoid tissue? And if so, why does GVHD predominately affect only a subset of tissues?

It is now known that both donor and host APCs contribute to GVHD, although with spatial and temporal differences. In MHC class I-dependent GVHD, alloantigens are mainly presented by host haematopoietic APCs (Shlomchik *et al.*, 1999; Toubai *et al.*, 2012). It is less clear on which APCs are involved in MHC class II-dependent GVHD. Host hematopoietic APCs are sufficient to induce MHC class II-dependent GVHD (Sprent *et al.*, 1986; Teshima *et al.*, 2002). However, experiments ablating individual host haematopoietic APC subsets have failed to pinpoint a single APC subset that is mandatory for inducing GVHD, suggesting redundancy in APC functions (Li *et al.*, 2012). Host macrophages and B cells attenuate GVHD through SIRPα-CD47 interaction-dependent phagocytosis of donor T cells and IL-10 production, respectively (Rowe *et al.*, 2006; Hashimoto *et al.*, 2011). Intriguingly, host DCs were shown to be protective against GVHD (Koyama *et al.*, 2012).

The role of host non-haematopoietic APCs in inducing GVHD has also gained increasing attention. Host non-haematopoietic APCs are fully capable of inducing alloreactive T cell-

mediated GVHD, primarily through the MHC class II-dependent pathway (Koyama *et al.*, 2012; Li *et al.*, 2012). Recently, it has been demonstrated that conditioning therapy amplifies the microbiota-induced MHC class II expression on intestinal epithelial cells in the ileum (Koyama *et al.*, 2019) (Figure 1.3). Donor T cells can also be found shortly after HSCT (Koyama *et al.*, 2012). The recruitment of naïve T cells to the colon is potentially mediated by the CCL19-CCR7 axis (Radulovic *et al.*, 2013). Gut microbiota is likely to be integral to GVHD induction, as GVHD is largely absent in germ-free mice (Van Bekkum *et al.*, 1974). Antibiotic-mediated intestinal decontamination in patients also demonstrated protective effects against GVHD (Vossen *et al.*, 2014). Since tissue microbiota-derived PAMP signals are localized, and T cells can be activated outside secondary lymphoid tissue by non-haematopoietic APCs, these findings suggest that GVHD induction may be a tissue-localized event.

Figure 1.3
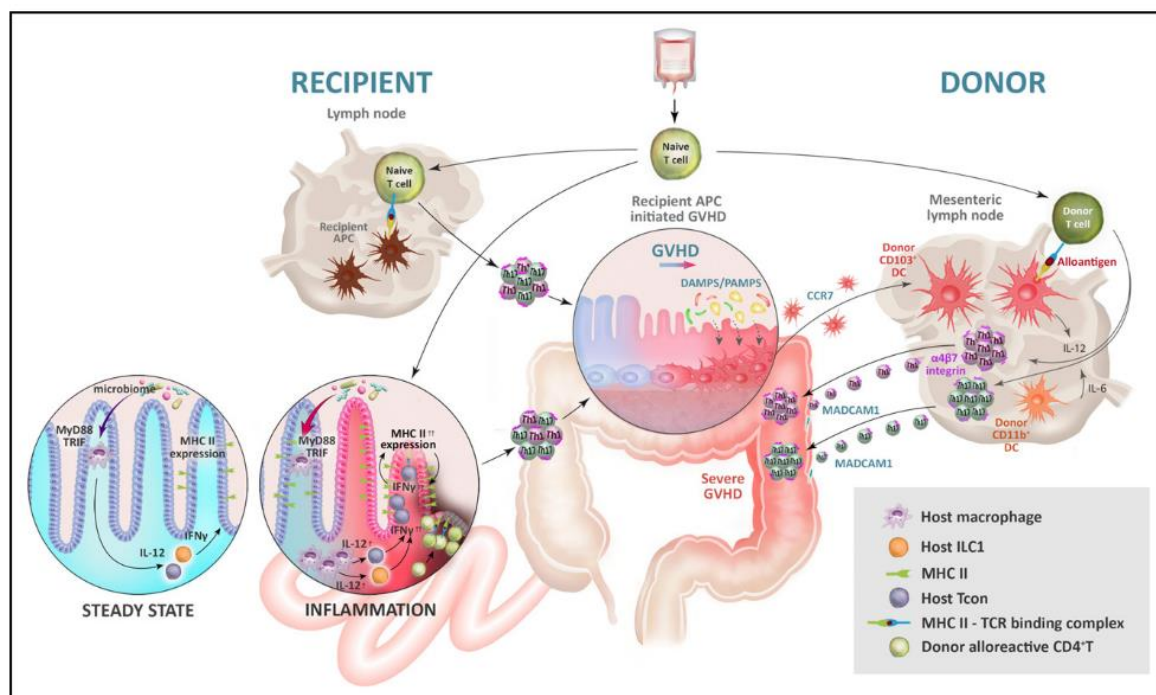
Figure 1.3 Host non-haematopoietic APCs participate in GVHD induction. Microbiota-driven cytokine signals augment antigen presentation by host haematopoietic APCs (e.g., DCs) and non-haematopoietic APCs (e.g., intestinal epithelial cells) after conditioning, leading to pathogenic differentiation of donor naïve T cells to Th1/Th17 subtype and initiate GVHD. Donor DCs expand at a later stage and migrate to the mesenteric lymph node under the influence of CCR7 to drive pathogenic differentiation of new donor T cells. Expression of α4β7 integrin on T cells then facilitate their migration into the gut and perpetuate severe GVHD. Image from Koyama and Hill, 2019.

On the other hand, donor APCs are poor inducers of GVHD but can subsequently exacerbate and maintain GVHD (Matte *et al.*, 2004; Koyama *et al.*, 2012). Donor APCs are activated by PAMPs and DAMPs in the GI tract after initial tissue damage. DCs migrate to the mesenteric lymph nodes, where they activate donor T cells and imprint the gut-homing integrin α4β7 (Koyama *et al.*, 2015). T cells then acquire an effector phenotype and migrate to the GI tract to mediate tissue damage. Donor T cells also produce GM-CSF which drives donor DC expansion and alloantigen presentation in the colon and mesenteric lymph nodes (Gartlan *et al.*, 2019).

1.2.3 Prevention of GVHD

Apart from minimizing risk factors during the donor selection process, further interventions are often necessary. This is achieved mostly via the manipulation of the T cell population. T cell-depleted grafts are routinely used for HSCT (Ho and Soiffer, 2001). Common GVHD prophylaxis strategies include the use of T cell-targeting antibodies, such as Alemtuzumab and anti-thymocyte globulin (ATG) (Ali *et al.*, 2017). Other frequently used pharmacological agents, including tacrolimus and cyclosporine, target T cell activation pathway via the inhibition of calcineurin (Flanagan *et al.*, 1991). Calcineurin is a calcium and calmodulin dependent serine/threonine protein phosphatase which activates the transcription factor NFAT. This in turn upregulates the expression of IL-2, a cytokine with the capability of stimulating the proliferation of T cells (Shaw *et al.*, 1988). Calcineurin inhibitors are frequently being used in combination with other immunosuppressants with alternative mechanism of action such as mycophenolate mofetil and Sirolimus (Sandmaier *et al.*, 2019). Post-transplant cyclophosphamide has shown to be highly effective in haploidentical transplantation, and is being studied for the use in HLA-matched transplantation (Williams *et al.*, 2020).

However, the advantage of GVHD prevention is usually offset by other complications, such as higher rates of graft failure, relapse of malignancy and infections (Ho and Soiffer, 2001). The use of calcineurin inhibitors also adversely inhibit the generation of GVHD-suppressing Tregs, which is an IL-2 dependent process (Whitehouse *et al.*, 2017). Novel GVHD pathophysiology-based approaches such as blocking the chemotaxis of effector cells to target organs, protecting

the gut from conditioning injury, and adoptive transfer of tolerogenic cell populations, are being tested in preclinical and clinical settings (Zeiser and Blazar, 2017).

1.3 Immune cell involvement in cutaneous GVHD

The skin plays various important physiological functions such as temperature regulation, prevention of excessive water loss, and synthesis of vitamin D. As an interface separating the body and the environment, the skin also forms a crucial part of host immunity, protecting against external pathogens. Human skin is composed of two main layers, the epidermis and dermis. The epidermis is the outermost layer of the skin, which keratinocytes continuously proliferate and cornify to form a protective barrier. Melanocytes and Langerhans cells (LCs) are also among the major cell types that can be found in the epidermis. Whereas the dermis is the inner layer of skin which is rich in connective tissue, blood vessels, nerves, hair follicles and other structures. The main cell populations that reside in the dermis are fibroblasts, macrophages, DCs, endothelial cells, pericytes, T cells, innate lymphoid cells (ILCs), and mast cells.

As an immune cell-rich organ, the skin is also very commonly affected by GVHD. The presence of maculopapular rash is often the first sign of GVHD. This process is thought to be mediated by donor alloreactive T cells and other recruited myeloid cells. However, the identity and origin of these cells are poorly understood. In this section, the characteristics, development, and function of the major immune cell types that are involved in GVHD pathogenesis will be reviewed.

1.3.1 Langerhans cells

LCs are the primary APCs in the epidermis. They are strategically positioned at the interface with the environment, acting as "immune sentinels" by sampling the environment within the keratinocyte junctions (Kubo *et al.*, 2009). Upon encountering antigens, LCs migrate to skin-draining lymph nodes to activate T cells (Johnston, Halliday and King, 2000). LCs also play an

important role in maintaining tissue homeostasis and immune tolerance (Hatakeyama *et al.*, 2017).

LCs are characterised by the expression of C-type lectin langerin (CD207) and MHC class I-like glycoprotein CD1a. Langerin is a transmembrane protein which is internalized into LC-specific cytoplasmatic organelles named Birbeck granules upon binding to antigen (Valladeau *et al.*, 2000). CD1a specializes in presenting lipid antigens to T cells (de Jong and Ogg, 2021). LCs express high levels of FcεR1 and CD39. FcεR1 is the high-affinity receptor for the Fc region of IgE. This receptor is also expressed in other granulocytes such as mast cells, basophils and eosinophils, and plays a major role in mediating allergic responses (Galli and Tsai, 2012). In the context of antigen presentation, FcεR1 improves presentation efficiency of IgE-bound antigens (Shin and Greer, 2015). The outcome of FcεR1-mediated antigen presentation can be either immunogenic or tolerogenic, potentially determined by the crosslinking status of FcεR1 during engagement with antigens (Shin and Greer, 2015). CD39 is an ATPase that is expressed on the surface of LCs. It hydrolyses extracellular ATP and ADP, which might contribute to the protective role of LCs against inflammation (Mizumoto *et al.*, 2002). As professional APCs, LCs also express high levels of MHC class II molecules. The expression of junctional proteins including EpCAM, occludins and claudins allows LCs to integrate into the keratinocyte layer (Tang *et al.*, 1993; Kubo *et al.*, 2009).

The development of LCs is unique among other DCs as they are seeded in the skin before birth. LCs are established from the yolk sac and foetal liver haematopoiesis, and are capable of self-renewal throughout adulthood without depending on the bone marrow (Hoeffel et al., 2012). During inflammation, LCs can be replenished by bone marrow derived progenitors potentially through a transient population of classical monocytes, followed by CD1c+ blood DCs to restores stable self-renewing LCs (Collin and Milne, 2016). Although lineage tracing experiments showed that the development of LCs is closely related to both DCs and macrophages (Wu et al., 2016), LCs transcriptionally resemble classical DCs (cDCs) and acquire DC-like morphology and functionality (Carpentier et al., 2016).

The role of LCs in GVHD induction has received significant attention. In humans, host-derived LCs survive conditioning and may be present up to 100 days post-HSCT depending on the intensity of conditioning and prior GVHD (Collin *et al.*, 2006; Haniffa *et al.*, 2009; Mielcarek *et al.*, 2014). The fact that patients with GATA2 mutation, which causes complete loss of DCs but preservation of LCs, still develop GVHD after HSCT suggests that host-derived LCs may

participate in instigating GVHD (Cuellar-Rodriguez *et al.*, 2011). However, in such scenario host macrophages and other non-professional APCs may also contribute to donor T cell activation. During immune reconstitution, donor engraftment in the LC compartment results in mixed host-donor chimerism. The overlapping timeframe of donor LC engraftment and acute GVHD represents a possible role of donor LC in GVHD pathology. Although degree of LC chimerism was not shown to correlate with clinical or histological skin GVHD, donor LC chimerism was associated with increased skin inflammation (Mielcarek *et al.*, 2014). The role of donor LCs in GVHD remains largely unexplored.

Pre-clinical models allow manipulation of LCs to better understand their roles in GVHD. In mouse, host LCs are sufficient in causing cutaneous GVHD (Merad et al., 2004), but depletion of either host or donor LCs does not abrogate GVHD when other host APCs are present (Li et al., 2011). This observation suggests that host LCs may contribute to, but are not essential for the GVHD pathology. However, another group has demonstrated that host LCs are required for cutaneous GVHD, primarily through locally licensing primed alloreactive CD8 effector cells to induce epithelial damage (Bennett et al., 2011; Santos E Sousa et al., 2018).

1.3.2 Macrophages

Macrophages are tissue cells which form a component of the innate immune system, as well as play an important role in tissue homeostasis (Shapouri-Moghaddam *et al.*, 2018). They are widely distributed in lymphoid and nearly all non-lymphoid organs, such as marginal zone macrophages in the spleen, alveolar macrophages in the lung, osteoclasts in the bone, Kupffer cells in the liver, as well as in immune-privileged sites such as microglial cells in the central nervous system (Shapouri-Moghaddam *et al.*, 2018). Macrophages represent a highly heterogenous population of cells with specialized functions depending on the anatomical location. This includes clearance of dead cells and debris (Boada-Romero *et al.*, 2020), tissue repair and remodelling (Wynn and Vannella, 2016), and host defence (Weiss and Schaible, 2015). Macrophage biology is also linked to a broad spectrum of diseases including infections, metabolic disorders, autoimmune diseases, and cancer (Shapouri-Moghaddam *et al.*, 2018).

The homeostasis of tissue macrophages was first described to be dependent on circulating monocytes in the mononuclear phagocyte system in 1968 by Van Furth (van Furth and Cohn, 1968). Indeed, monocytes are rapidly recruited to inflamed tissues and can acquire a wide range of macrophage and DC functionality (Collin and Bigley, 2018). To classify these monocyte-derived cells, a dichotomy activation model has been established which divides macrophages into M1 (pro-inflammatory) and M2 (anti-inflammatory) based on the type of stimulation and contrasting functional characteristics *in vitro* (Mills *et al.*, 2000). Subsequent study of human macrophage activation with 28 stimulation conditions demonstrated that the M1/M2 paradigm was not sufficient to explain the diverse macrophage states (Xue *et al.*, 2014). In reality, macrophages are exposed to a wide range of stimuli simultaneously. With high-resolution single cell analysis, finer levels of macrophage heterogeneity can be profiled. Different types of macrophages have been identified across multiple tissues in healthy and disease states (Mulder *et al.*, 2021). This study also highlighted the inability of the M1/M2 paradigm to capture the macrophage diversity, especially macrophages related to cancers (tumour-associated macrophages) and macrophages with embryonic origin.

For long, tissue macrophages were thought to solely arise from the differentiation of circulating blood monocytes. Now it is known that the self-renewal of embryonic-derived macrophages is the predominant mechanism of macrophage population maintenance in the steady state in most tissues (Hashimoto *et al.*, 2013; Ginhoux and Guilliams, 2016). This circumstance is further demonstrated by the fact that patients with monocytopenia has largely unaffected tissue macrophage population (Bigley *et al.*, 2011). In the skin, however, recruitment of monocytes is integral to the homeostasis of dermal macrophages (McGovern *et al.*, 2014; Ginhoux and Guilliams, 2016). Dermal "long-term" resident macrophages can be distinguished from monocyte-derived macrophages by immunostaining of intracellular FXIIIa (Zaba *et al.*, 2007). In flow cytometry, resident macrophages are autofluorescent in the FITC channel due to the presence of dense cytoplasmic melanin granules within their cytoplasm (Haniffa *et al.*, 2009). Additional markers such as *HES1*, *FOLR2*, *LYVE1* have been used in single-cell RNA sequencing (scRNA-seq) to identify resident macrophages (Mulder *et al.*, 2021). Nonetheless, resident macrophages and monocyte-derived macrophages share similar functional profiles. Both macrophage populations do not migrate to lymph nodes upon stimulation (lack of CCR7 expression), are poor stimulators of naïve T cells but potent inducers

of memory T cells, and are capable of producing pro-inflammatory cytokines (Haniffa *et al.*, 2009; McGovern *et al.*, 2014).

The involvement of macrophages in GVHD has been described as a type of effector cell which secrete pro-inflammatory cytokines to exacerbate tissue injury (Hill and Ferrara, 2000). Manipulation of the macrophage CSF-1 (M-CSF) axis in murine models have offered insights to the roles of host tissue macrophages and donor bone marrow-derived macrophages in GVHD. Pre-transplant expansion of host macrophages conferred protection against GVHD (Hashimoto *et al.*, 2011), while depletion of host macrophages exacerbates GVHD (MacDonald *et al.*, 2010). On the other hand, post-transplant depletion of donor macrophages ameliorates GVHD (Alexander *et al.*, 2014). Host macrophages may provide protection against GVHD through phagocytosis of alloreactive T cells (Hashimoto *et al.*, 2011), and producing transforming growth factor beta (TGF-β) and promoting the expansion of Tregs (D'Aveni *et al.*, 2015). The effector mechanism of donor macrophages may be mediated by GM-CSF-producing donor T cells (Tugues *et al.*, 2018).

In humans, the content of myeloid cell infiltrates bearing macrophage markers correlates with GVHD clinical severity (Nishiwaki *et al.*, 2009). However, these cells could potentially originate from either the host or the donor, as host dermal resident macrophages survive conditioning, leading to mixed host-donor chimerism for over 1 year post-HSCT (Haniffa *et al.*, 2009). Recently, it has been demonstrated that the infiltrates are donor monocyte-derived macrophages which can activate allogeneic T cells and has the potential to mediate cutaneous GVHD (Jardine *et al.*, 2020).

1.3.3 Dendritic cells

DCs are professional APCs that play a vital role in immunity. They bridge the innate and adaptive immune system by presenting antigens to T cells (Banchereau *et al.*, 2000). Dendritic cells express a wide range of pattern recognition receptors (PRRs) such as C-type lectin receptors (CLRs), Fc receptors and Toll-like receptors (TLRs) which bind to various antigens through PAMPs and DAMPs (Wang *et al.*, 2020). Antigen engagement with its corresponding receptor results in it being endocytosed, processed, and presented via highly diverse peptide-binding MHC molecules. At the same time dendritic cells migrate to secondary lymphoid

organs to interact with T cells. Together with co-stimulatory molecules and cytokines, T cells are activated and polarized to specific subsets catered to eliminate the antigen (Patente, Pelgrom and Everts, 2019). DCs can also exert tolerogenic functions by deletion of self-reactive T cells and promoting expansion of Tregs (Steinman, Hawiger and Nussenzweig, 2003).

DCs are a class of haematopoietic stem cell (HSC)-derived cells which can be divided into three major subsets: plasmacytoid dendritic cells (pDCs), cDC1 and cDC2 (Collin and Bigley, 2018). The development of each DC subset requires a combination of critical transcription factors such as GATA2, PU.1, IRF4 and IRF8, and their deficiency can lead to the loss of certain DC subsets (Dickinson et al., 2011; Hambleton et al., 2011). Different types of DC express a unique suite of PRRs and are catered to elicit specific types of immune responses.

pDCs are the major producer of type I interferon. In humans, they are commonly categorised by the expression of CD303 (CLEC4C/BDCA-2) and CD304 (Neuropilin/BDCA-4). They respond to viral infections by sensing their RNA and DNA through endosomal TLR7 and TLR9, respectively, which ultimately leads to production of type I interferons and pro-inflammatory cytokines such as TNF-α and IL-6 (Bao and Liu, 2013).

On the other hand, cDC1s are superior in cross-presenting antigens via MHC class I to activate CD8+ T cells (Haniffa *et al.*, 2012). They can also polarize CD4 T cells towards Th1 phenotype through IL-12 (Jongbloed *et al.*, 2010). cDC1s are characterised by the expression of CD141 (BDCA-3), CLEC9A, CADM1, BTLA and XCR1. Expressing TLR3, 9, 10, cDC1s specialise in responding to viral and intracellular antigens (Collin and Bigley, 2018). cDC1s are also major producers of type I and III interferons (Lauterbach *et al.*, 2010). Meanwhile, cDC2s are potent in priming naïve CD4+ T cells for Th2 or Th17 polarization (Schlitzer *et al.*, 2013; Williams *et al.*, 2013). Example of markers that are commonly used to identify cDC2s include CD1c, CD2, FcεR1. cDC2s express a wide range of receptors including lectins and TLRs, and are capable of mediating infections of bacteria, fungi and parasites (Collin and Bigley, 2018). cDC2 represents the largest DC population in human blood and dermis (Alcántara-Hernández *et al.*, 2017).

It has recently been discovered that the cDC2 population is heterogenous and can be subdivided into a "DC-like" DC2 population and a "monocyte-like" inflammatory DC3 population (Villani et al., 2017). DC3 can be distinguished from DC2 by the expression of CD14, CD163, and a lower expression of CD5 (Villani et al., 2017; Yin et al., 2017). The development of DC3 is more closely related to monocyte rather than DC2, however, DC3 are not monocyte-

derived but rather arises through an independent pathway of haematopoiesis (Cytlak *et al.*, 2020). DC3 shares characteristics of both DC2 and monocytes, such as being potent at priming naïve T cells and secreting inflammatory cytokines, respectively (Bourdely et al., 2020). This population is shown to be expanded in patients with systemic lupus erythematosus (Dutertre et al., 2019) and psoriasis (Nakamizo et al., 2021).

In GVHD settings DCs are activated by PAMPs and DAMPs as a result of host tissue injuries, and subsequently present antigens to prime alloreactive T cells (Ferrara *et al.*, 2009; Koyama and Hill, 2019). Selective depletion of recipient and donor-derived DC subsets in murine HSCT models has demonstrated that multiple DC subsets can induce GVHD (Yu et al., 2019). This is further complicated by the presence of non-DC APCs (e.g., B cells and macrophages) and non-haematopoietic cells that have antigen-presenting capabilities (e.g., intestinal epithelial cells). The loss of tolerogenic functions of DCs also contributes to GVHD. In mice, the depletion of pDCs has led to increased GVHD severity (Banovic et al., 2009). Compromised antigen presentation function of cDCs could also promote GVHD by limiting Treg expansion and function (Leveque-El Mouttie et al., 2016). Clinical studies have indicated that impaired DC reconstitution was correlated to increased risk of GVHD (Reddy et al., 2004). As such, manipulating the DC population in the graft to limit GVHD may be an attractive strategy. However, it remains a challenge to determine which DC subsets are optimal for preventing GVHD, and to obtain sufficient amount and quality of the desired DC population, which likely requires *ex vivo* generation and expansion procedures.

1.3.4 T cells

T cells form the core part of the adaptive immunity and are responsible for defending against pathogens and cancers, as well as promoting inflammation and autoimmune diseases. T cells originate from HSCs in the bone marrow and migrate to the thymus for maturation before entering the circulation (Kumar, Connors and Farber, 2018). During development, the genes encoding for T cell receptor (TCR) undergo rearrangement, giving rise to a huge diversity of unique TCRs and thus their ability to recognize a myriad of antigens (Krangel, 2009).

Since inappropriate activation of T cells can be devastating to the host, this process is tightly regulated by multiple signals to confirm that T cells only activate at the appropriate space and

time. T cell activation requires three signals from APCs (Gutcher and Becher, 2007), with the first signal being the presentation of antigen through MHC class II molecule to the antigen-specific TCR. The second signal involves engagement of the cell-surface co-stimulatory molecules such as CD28 on T cells interacting with CD80 and CD86 on APCs. The third signal is mediated by cytokines produced by APCs to polarize the differentiation of T cells into an effector subtype that mediates the appropriate immune response. Activation of T cells also leads to formation of antigen-specific memory T cells which are capable of mediating a rapid immune response should the host re-encounters the same antigen (Farber, Yudanin and Restifo, 2014).

The classical GVHD model states that T cells are necessary for its pathogenesis (Korngold and Sprent, 1978). Donor naïve T cells are activated by host APC in secondary lymphoid tissues (Anderson *et al.*, 2003; Ferrara *et al.*, 2009). T cells then acquire an effector phenotype and mediate tissue damage via cognate interactions and soluble factors (Ferrara *et al.*, 2009). GVHD was proposed to be mediated by Th1 cells (Hill *et al.*, 1997), but Th2 and Th17 cells also contribute to GVHD pathology depending on the type of tissue involved (Yi *et al.*, 2009). On the other hand, Tregs have the potential to suppress GVHD (Hoffmann *et al.*, 2002). The roles of innate and nonconventional T cells have also been studied in experimental and clinical GVHD (Hill *et al.*, 2021). It is important to note that T cells are also responsible for the elimination of malignant cells and thereby exerting the beneficial GVL effect (Blazar, Hill and Murphy, 2020). Therefore, it is of great interest to understand the mechanistic action of conditioning regimens and post-transplant immunosuppression towards different T cell subsets, in order to not only remove the T cells promoting GVHD, but also preserve the cells that alleviate GVHD and maximize the effect of GVL.

Apart from donor T cells, the role of host T cells in mediating GVHD is starting to gain traction. It was once assumed that conditioning regimens completely abrogates host T cell immunity. However, the discovery of tissue-resident memory T cells (Trm) and the observation of T cell-rich organs being the major sites affected by GVHD suggest that the role of these recipient T cells have been overlooked (Park and Kupper, 2015). Trm are non-circulating lymphocytes that express CD69, and they maintain tissue residency by upregulation of cell adhesion molecules (e.g., ITGAE (CD103) and ITGA1 (CD49a)) and tissue-homing chemokine receptor (e.g., CXCR6), while downregulating egress cues (e.g., SELL (CD62L)) (Kumar et al., 2017). Indeed, studies have shown that conditioning regimens deplete T cells in the circulation but not in the tissues

(Divito et al., 2020; Strobl et al., 2020). In the skin, host Trm can be activated by donor-derived monocyte and promote GVHD through production of pro-inflammatory cytokines (Divito et al., 2020) (Figure 1.4). There is also evidence that the activation of host Trm precedes GVHD, suggesting the involvement of host Trm in early mechanisms of GVHD development (Strobl et al., 2020). Recently, it has been suggested that host Trm from the skin of patients affected by GVHD may recirculate and promote inflammation in other organs (Strobl *et al.*, 2021). Overall, this may represent a host versus graft reaction and potentially be responsible for clinical low-grade acute GVHD, while more severe GVHD may require the participation of donor T cells (Jardine *et al.*, 2020).

Figure 1.4
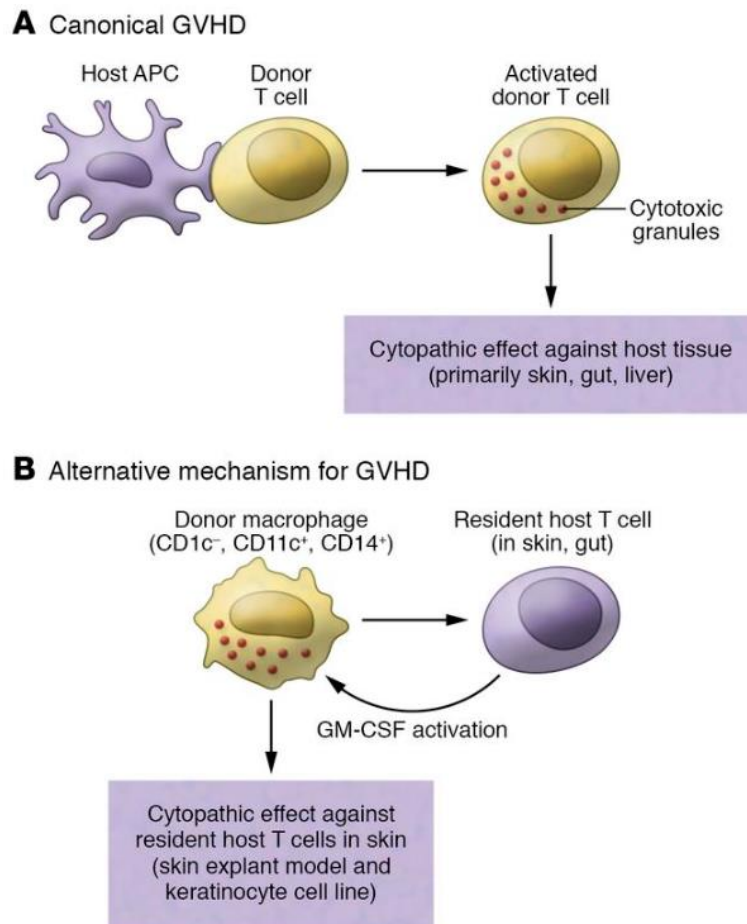
Figure 1.4 Alternative GVHD model. (A) In canonical GVHD, donor T cells mediate host tissue damage after being activated by host APCs. (B) In the alternative model, donor CD14+ CD11c+ macrophages activate resident host T cells. These macrophages may also mediate a direct cytopathic effect against host cells. The inclusion of host effector T cells suggests a host-versus-graft component in GVHD. Image from Young, 2020.

1.3.5 Innate lymphoid cells

Innate lymphoid cells (ILCs) are a family of bone marrow-derived lymphocytes that lacks antigen-specific receptors (Eberl *et al.*, 2015). It encompasses NK cells, ILC1, ILC2 and ILC3, and they play a key role in host defence and tissue homeostasis. In humans, NK cells can be subdivided into CD16+ CD56dim and CD16- CD56bright and can be found in peripheral blood and tissues (Poli *et al.*, 2009), while ILC1-3 are mainly tissue-resident (Eberl *et al.*, 2015). ILCs are being considered the innate counterpart of T cells due to the similarity of the suite of cytokine they produce. NK cells mirror cytotoxic T cell by the ability of producing cytolytic products such as perforin and granzymes, while ILC1-3 mirror the cytokine profiles of Th1, Th2 and Th17 cells respectively (Eberl *et al.*, 2015).

In transplant settings, NK cells are the first lymphoid cell to be reconstituted (Stern *et al.*, 2018). However, it remains controversial whether they promote or protect against GVHD. Donor NK cells were seen to infiltrate GVHD-affected organs (Shah *et al.*, 2015) and might mediate GVHD via production of pro-inflammatory cytokines such as IFN-γ (Cooley *et al.*, 2005). On the other hand, donor NK cells could suppress GVHD by either directly depleting alloreactive T cells (Olson et al., 2010; Sheng et al., 2020), or indirectly by depleting APCs (Meinhardt et al., 2015). Clinical transplantation suggested that NK cell alloreactivity might be associated with reduced relapse and GVHD, at least in mismatched settings (Ruggeri *et al.*, 2002).

Meanwhile, ILCs reconstitute slower than NK cells and their reconstitution pattern resembles T cells (Piperoglou *et al.*, 2022). The presence of circulating ILCs is associated with reduced GVHD (Marius Munneke *et al.*, 2014; Kroeze *et al.*, 2022). Donor ILC2s alleviate GVHD by producing cytokines to limit the development of inflammatory Th1 and Th17 cells (Bruce *et al.*, 2017), whereas host ILC3s exert their protective effects by mediating intestinal homeostasis via the production IL22 (Hanash *et al.*, 2012). However, it is recently demonstrated that ILC1s promote gut GVHD by producing IFN-γ, which subsequently upregulate MHC expression on intestinal epithelial cells and prime donor T cells (Koyama *et al.*, 2019).

1.4 Single-cell RNA sequencing

1.4.1 Advantage over other RNA profiling techniques

RNA profiling has been an indispensable tool for defining a cell's identity and its biological activities. Early RNA studies were performed using northern blots and quantitative polymerase chain reactions, which only measure low number of genes (Kukurba and Montgomery, 2015). Hybridization-based microarray technologies and sequence-based approaches were then developed. However, these methods have several limitations, such as high background levels due to cross-hybridization and the inability to distinguish isoforms, respectively (Kukurba and Montgomery, 2015). The development of next-generation sequencing and genome-wide RNA-sequencing have since revolutionized the field (Wang, Gerstein and Snyder, 2009). Although it has led to many novel discoveries, readout of the bulk population sequencing is the gene expression averaged across all the cells. The variability of gene expression across a seemingly homogenous population of cells has gained increasing appreciation. A prime example is the study of tumour cell heterogeneity has led to the discovery of immune escape and drug resistance mechanisms (Zhang *et al.*, 2021).

The initial concept of scRNA-seq based on a next-generation sequencing platform was published in 2009, which sequenced a single mouse blastomere (Tang *et al.*, 2009). Recent advances in technology have tremendously scaled up the throughput and reduced the cost, enabling the study of tens of thousands of single cell transcriptomes (Jovic *et al.*, 2022). scRNA-seq allowed unbiased identification of cells by eliminating the need to define cells with a handful of established markers. This confers numerous advantages towards understanding the underlying biology, such as discerning closely related cell types by trajectory inference (Trapnell *et al.*, 2014). The field of scRNA-seq is still rapidly evolving. This technology has been combined with approaches that study different modalities, including chromatin accessibility, protein expression and the spatial location of each cell within their native environment (Efremova and Teichmann, 2020). A major limitation of scRNA-seq, however, is that the cost per experiment remains high, limiting the number of biological samples profiled. This is expected to improve in the near future.

To conduct a scRNA-seq experiment, deciding which single cell platform to use is critical to the success of the experiment, and this depends on the research questions to be answered. Mainstream single cell platforms can be summarized into two categories: plate-based (e.g., SMART-seq2) and droplet-based (e.g., 10x Genomics Chromium). Plate-based methods are recognized for the number of features obtained per cell, while droplet-based methods can achieve a higher throughput at the cost of sequencing depth. If the study involves investigating splice variant or isoforms, or detection of lowly expressed genes, then plate-based methods are preferred as they provide information of full-length transcripts and offers higher sensitivity, whereas if the goal of study is to identify cellular heterogeneity within a tissue, droplet-based methods triumph as maximizing number of cells sampled is key (Haque *et al.*, 2017). In recent years, droplet-based methods, which are much less labour intensive have become dominant. The slight reduction in transcriptomic detail is almost completely offset by a higher cell count, which is superior at overcoming the problem of transcript drop-out, common to all methods.

The human skin has been extensively studied with scRNA-seq technologies in the past few years, primarily through droplet-based methods (Theocharidis *et al.*, 2022). Since scRNA-seq only gained popularity recently, only a few groups have used this technique to study GVHD effector cells. Tkachev and colleagues deployed a rhesus macaques MHC-mismatch HSCT model and identified that pathogenic donor CD8+ T cells infiltrating the GI tract simultaneously acquire a cytotoxic and Trm phenotype (Tkachev *et al.*, 2021). Engel and colleagues used a mouse MHC-mismatch HSCT model and discovered that donor DCs in the mesenteric lymph node may prime alloreactive CD4+ T cells into a quiescent state which the cells are clonally expanded but lack effector functions (Engel *et al.*, 2020). Piper and colleagues used a mouse MHC-mismatched model and identified an IFN-γ− subset within the CD4+ GM-CSF+ T cell population which may be responsible for inducing pathologic damage in the colon (Piper *et al.*, 2022). Lastly, Strobl and colleagues described a circulating skin-derived T cell population in patients with GVHD which may reseed and propagate inflammation in other organs (Strobl *et al.*, 2021). To the best of my knowledge, there are no publications on profiling the myeloid cells in human cutaneous GVHD using scRNA-seq.

1.4.2 Droplet-based platform

The principle of scRNA-seq relies on preserving the information of cellular origin for each transcript. Most protocols utilize unique nucleotide sequences, with each sequence being incorporated into all captured transcripts from an individual cell, acting as a "cellular barcode". Some protocols also use additional nucleotide sequences, called unique molecular identifiers (UMIs), for transcript quantification. UMIs are a pre-determined set of nucleotides, acting as "molecular tags". They are incorporated into all captured transcripts before any PCR amplification steps so that copies of the same transcript arise from amplification can be identified. The end point of a scRNA-seq experiment is to sequence the barcoded cDNA library generated from the transcripts, so that the gene expression profile of each cell within the sample can be characterized.

In terms of experiment workflow, using the protocol from 10x Genomics (the platform of choice in this study) as an example, single cell suspension is first loaded onto a microfluidic device along with the reagents required for reverse transcription and the gel beads, which contain the cellular barcodes and UMIs. The Chromium controller then partitions the cells into droplets, with each droplet containing a single cell, multiple copies of a single cellular barcode, a set of UMIs and the reagents. Cells within the droplet are then lysed and polyadenylated RNA molecules are captured. This is followed by reverse transcription within each droplet, resulting in cDNA molecules which are barcoded and UMI-tagged. The cDNA molecules are then pooled and amplified. To generate cDNA libraries for sequencing, cDNA molecules are fragmented, and adaptor sequences are added so that the resultant libraries are compatible with next generation sequencing. Index sequences, which act as "sample barcodes", are also incorporated to allow multiple samples to be sequenced in a single sequencing run. Specific regions on the cDNA, such as T cell receptor sequences, can be separately amplified to generate additional libraries for sequencing (Borcherding *et al.*, 2021).

1.4.3 Data analysis

Pre-processing and quality control

The goal of analysing single-cell data is to uncover biologically meaningful signals from billions of sequencing reads. These reads originate from tens of thousands of cells across multiple samples, therefore they have to be assigned back to their sample and cell of origin. This process is known as demultiplexing. Multiple raw data processing pipelines are available, such as Cell Ranger (Zheng *et al.*, 2017), inDrop (Klein *et al.*, 2015), SEQC (Azizi *et al.*, 2018), and zUMIs (Parekh *et al.*, 2018), for performing quality control of reads, demultiplexing, genome alignment and quantification. This results in count matrices with dimension equal to the number of cell barcodes times the number of genes detected, which contains the number of copies of each transcript detected by every cell. Before conducting any downstream analysis, this list of cells must be filtered to remove poor-quality cells and doublets, which are the result of two (or more) cells being incorporated into a single droplet during the partitioning stage.

To determine the viability of cells, three covariates are commonly inspected: number of genes detected per cell, number of UMIs detected per cell (count depth), and fraction of counts originating from mitochondrial genes per cell (Luecken and Theis, 2019). Cells with few detected genes, low count depth and high percentage of mitochondrial genes may indicate a ruptured cell membrane, as cytosolic mRNA content is lost and only the mRNA located within the mitochondria is preserved. In contrast, cells with abnormally high number of detected genes may indicate doublets. Of note, these covariates should be jointly considered when determining a suitable threshold, as there may be a biological interpretation to why their value fall outside the expected range. One example would be cells with a higher fraction of mitochondrial genes may be involved in highly metabolically active processes (Luecken and Theis, 2019). Genes that are not expressed in more than a few cells should also be removed as they are not informative of the cellular heterogeneity (Luecken and Theis, 2019). Further quality control procedures can also be implemented, such as correcting the count data for ambient RNA (Young and Behjati, 2020). Quality control very often is an iterative process, where optimal thresholds can only be determined based on downstream analysis performance (Luecken and Theis, 2019).

There are also more sophisticated doublet detection tools available to further improve the quality of data. The method of doublet detection generally falls within two categories: programmes such as Scrublet (Wolock, Lopez and Klein, 2019) and DoubletFinder (McGinnis, Murrow and Gartner, 2019) detect doublets by identifying cells that contain transcripts originating from two different cell types, whereas Demuxlet (Kang *et al.*, 2018) and Vireo (Huang, McCarthy and Stegle, 2019) harness genetic variations between individuals to identify cells that contain two sets of genotypes. The method of doublet detection by genotype is only applicable if the sample contains cells from two or more individuals. This can be achieved by mixing single-cell suspensions from different individuals, or in the scenario of allogeneic HSCT, a natural *in vivo* mixing of donor and recipient cells.

Exploring and interpreting data

With each cell expressing thousands of genes, the resulting dataset is enormous and multi-dimensional. The challenge of analysing single-cell data is to make meaningful comparison between groups of cells that express distinct sets of genes. This is achieved by reducing the dimensions, simplifying the dataset through only concerning the most informative genes, followed by using clustering algorithms to group similar cells together. The details of each step will be discussed below.

After removing low quality cells, the data must first be normalized to account for the difference in count depth arising from technical effects (Hafemeister and Satija, 2019). Similarly, gene counts can be scaled so that all genes are weighted equally (zero mean and unit variance) for downstream analysis (Butler *et al.*, 2018). However, this step removes some biological information and therefore is not implemented in all pipelines (Street *et al.*, 2018).

The first step of dimensional reduction involves feature selection. This step highlights the biological signal by selecting a subset of genes that drive the most variability in the data for downstream analysis (Brennecke *et al.*, 2013). To further reduce the complexity of data, principal component analysis (PCA) is commonly performed. This summarises the dataset with top principal components, preserving as much variability as possible with fewer variables. Finally, clustering algorithms are used to group cells with similar PCA data into distinct clusters in an unbiased manner. Clustering results are visualized in a two- to three-dimensional space. Common dimensionality reduction methods for visualization include t-distributed stochastic

neighbour embedding (t-SNE) (van der Maaten and Hinton, 2008) and Uniform Manifold Approximation and Projection (UMAP) (McInnes, Healy and Melville, 2018). While both t-SNE and UMAP are non-linear, graph-based methods for dimensionality reduction, UMAP is more efficient and better at capturing the global structure, therefore it is generally recommended over t-SNE (Luecken and Theis, 2019).

These clusters are then given biological interpretations by annotating them with cell type labels. Clusters can be defined by either manually inspecting their expression profile or using automated annotation software, or a combination of both. To manually define clusters, differential expression testing can be performed between two groups of cells (comparing cells in one cluster to all other cells in the dataset) to identify marker genes. Alternatively, expression values of literature-derived marker genes can be directly visualized. On the other hand, automated annotation software, such as CellTypist (Domínguez Conde *et al.*, 2022) and DISCO (Li *et al.*, 2022), predicts the identity of cells by harnessing the gene expression data from a large quantity of literature across different tissues and physiological conditions. Though convenient, these are only estimations, as very often the exact cell identity (e.g., a disease-specific cell state) won't be available in these reference atlases. Therefore, manual annotation remains an irreplaceable approach when exploring finer details within the dataset. Once the identity of cells is determined, further analysis can be performed, such as studying the transitions between cell identities (trajectory inference).

In many cases, multiple scRNA-seq experiments are carried out with the goal of comparing the cellular composition and gene expression differences between samples obtained across various experimental conditions (e.g., comparing cells in control group to those that are stimulated with a cytokine, or comparing clinical samples obtained from patients to those obtained from healthy subjects). In these situations, running a single integrated analysis can be advantageous. Data integration works by correcting the batch effect between experiments and align shared cell populations across datasets (Stuart *et al.*, 2019). Comparative analysis can then be easily performed on a single dataset.

1.5 Hypotheses and aims

1.5.1 Hypotheses

This is a data-led study. Hypotheses will be formulated and discussed at the end of the thesis.

1.5.2 Overall aims

The aim of this project is to use single-cell RNA sequencing to define the effector cells in cutaneous GVHD. By characterising the immune populations in the blood and the skin, I seek to delineate specific changes between healthy and disease state and explore their developmental relationship between tissues, in terms of identity, functional status and T cell receptor usage. I also aim to define the relative contribution of host and donor cells in cutaneous GVHD.

Chapter 2. Materials and methods

2.1 Generation of single cell suspension

2.1.1 Skin

Fresh skin sample from healthy controls (*n* = 4) was either obtained from surplus of breast reconstruction surgery (*n* = 2) or by shave biopsy (*n* = 2). Fresh GVHD-affected skin sample (presented as a rash) (*n* = 5) and transplant control skin sample (absence of GVHD 100 days post-HSCT) (*n* = 2) were obtained by shave biopsy. The size of skin samples obtained from shave biopsy was approximately 0.5 cm^2. A small section of biopsy from patients with GVHD was used for histopathology examination (except patient GVH5) (Figure 2.1). To obtain size-matched sample from healthy surplus skin, it was first cut into approximately 1.5 cm wide strips in cold PBS. Whole skin layer of 200 μm thickness was then obtained using a dermatome fitted with a Pilling WecPrep blade and a .008 gauge Goulian guard (Teleflex, US). Skin sample of 0.5 cm^2 in size was obtained using a scalpel.

Figure 2.1
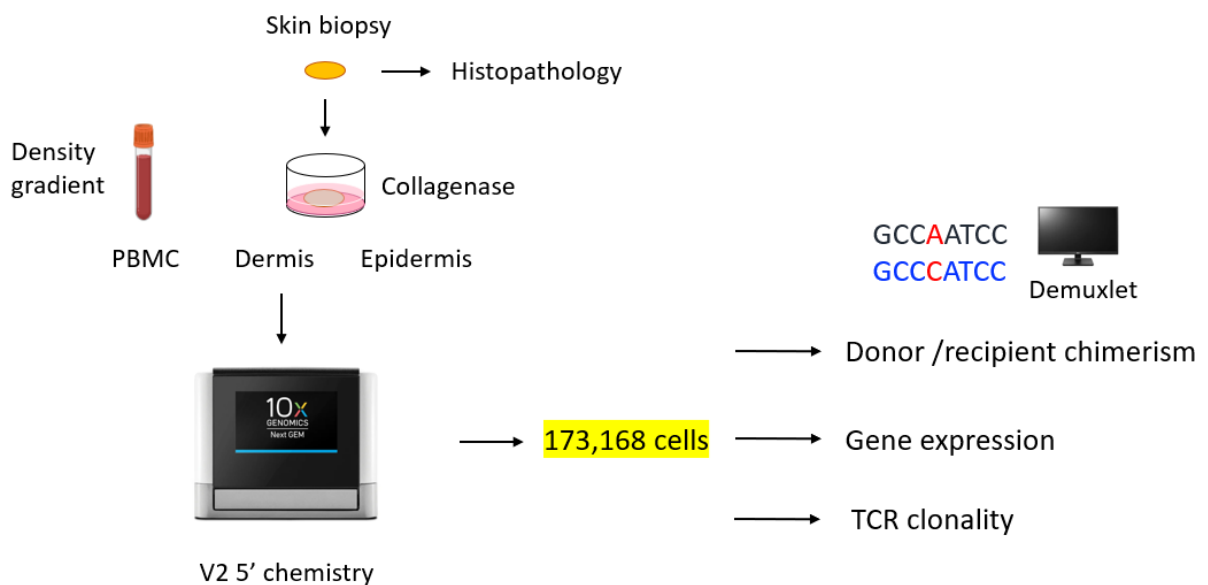
Figure 2.1 Outline of the experimental procedure.

The samples were first digested in 1 mL of 1 U/mL Dispase II (Gibco, #17105041) in RPMI (Sigma-Aldrich, #R0883) in a 24-well plate in an incubator (37 °C, 5% CO_2) for 75 minutes to separate the epidermal and dermal compartments. The two compartments were pulled apart using two pairs of tweezers. The tissues were then rinsed in separate wells filled with PBS to avoid cross-contamination. To obtain single-cell suspensions, epidermis and dermis were then digested separately in 1 mL of 1.6 mg/mL Worthington's Collagenase Type 4 (LS004188. CODE: CLS4) in RF10 medium (RPMI with 10% FCS (Sigma-Aldrich, #F9665)) in an incubator (37 °C, 5% CO_2) for approximately 9 hours (or until the dermis was totally digested). The resultant suspensions were washed with PBS at 500 g for 5 minutes twice before proceeding to cell counting. Suspensions were then adjusted to appropriate concentrations with PBS and kept at 4 °C until loading onto the single-cell platform.

2.1.2 Blood

Blood samples from healthy donors and patients were collected in EDTA tubes. Samples from patients with GVHD and transplant controls were obtained in the same instance of shave biopsy. Whole blood samples were first diluted 1:1 in PBS. Peripheral blood mononuclear cells (PBMCs) were then isolated by density gradient using Lymphoprep (STEMCELL Technologies, #07851/07861). PBMCs were washed with PBS at 500 g for 5 minutes twice. Red cell lysis buffer (Biolegend, #420302) was used to remove excessive erythrocytes from the samples. Suspensions were then adjusted to appropriate concentrations with PBS and kept at 4 °C until loading onto the single-cell platform.

2.2 Single-cell RNA sequencing

Approximately 20,000 cells from each sample were loaded onto separate channels of the Chromium chip K (10x Genomics, US) with the aim of recovering 10,000 live single cells. cDNA was generated from barcoded single cells and indexed libraries (gene expression and T cell receptor) were prepared according to manufacturer's protocol (v1 – Chromium Single Cell V(D)J Reagent Kits User Guide, CG000086 Rev H; v2 – Chromium Next GEM Single Cell 5' Reagent Kits v2 (Dual Index), CG000331 Rev A). The size distribution of fragments and

concentration of each library were measured by Tapestation (Agilent Technologies, US) and Qubit (Thermo Fisher Scientific, US), respectively. Libraries were pooled prior to sequencing. Illumina NovaSeq 6000 sequencing system was used with the configuration specified in the respective protocols to achieve a minimum depth of 50,000 raw reads per cell. Sequencing was performed by members of staff from the Newcastle Genomics Core Facility.

2.3 Whole exome sequencing

Transplant donor DNA material was obtained from recipient blood neutrophils post-HSCT (donor-derived). Recipient DNA material was obtained from skin fibroblasts (remain recipient derived post-HSCT). To obtain sufficient fibroblasts for DNA extraction, dermal digest was used as a source of fibroblasts for culture. Cells were cultured in RF20 (RPMI with 20% FCS) in a 6-well plate in an incubator (37 °C, 5% CO_2) until sufficient quantity was obtained. All cells were stored as a pellet at - 80 °C until extraction.

DNA was extracted using QIAamp DNA Micro Kit (Qiagen, #56304). The quality and concentration of DNA were measured by nanodrop (Thermo Fisher Scientific, US) and Qubit respectively prior to whole exome sequencing. Whole exome sequencing was performed by members of staff from the Newcastle Genomics Core Facility.

2.4 Data analysis

2.4.1 Sample demultiplexing and read alignment

This step was done by the Newcastle Bioinformatics Support Unit. Cell Ranger v4.0 (10x Genomics, US) was used to allocate sequencing reads back to their sample of origin by library indices. The reads were aligned to the human genome GRCh38 and feature-barcode matrices were generated for each sample.

2.4.2 Deconvolution of cell origin

Demuxlet (Kang *et al.*, 2018) (https://github.com/statgen/demuxlet) was used to assign each cell to either donor or recipient in origin by genotype. BAM files and lists of barcodes for input were generated using Cell Ranger v4.0. VCF files containing individual genotypes were obtained by performing variant calling on individual whole exome sequence datasets. Variant calling was performed by members of staff from the Newcastle Bioinformatics Support Unit. The VCF files of each donor-recipient pair were then combined, and chromosome orders were rearranged according to the software requirement.

2.4.3 Quality control and doublet removal

Quality control was performed on each sample individually prior to downstream analysis. Seurat (v4.1.1) (Butler *et al.*, 2018) (https://satijalab.org/seurat/) in R (v4.2.1) was used for the analysis of scRNA-seq data. Genes that were expressed in fewer than 3 cells were removed. Genes mapped to variable, diversity, and joining regions of immunoglobulin and T cell receptors were also removed. Appropriate filters for the total number of genes and molecules detected in each cell, as well as the percentage of mitochondrial transcripts, were also applied to each sample accordingly. DoubletFinder (v2.0.3) (McGinnis, Murrow and Gartner, 2019) (https://github.com/chris-mcginnis-ucsf/DoubletFinder) was used to simulate and identify doublets in each dataset. Droplets that were identified as doublet by DoubletFinder (gene expression approach), or identified as "ambiguous" or doublet by demuxlet (genotype approach), were subsequently removed.

2.4.4 Data integration and cluster analysis

Each dataset was first individually analysed. Data normalization, variance stabilization and feature selection (*n* = 3000) were performed using the *SCTransform* function. Principle components were calculated using the *RunPCA* function. Datasets were visualized by plotting the coordinates calculated with the *RunUMAP* function using the top 30 dimensions. Cells

were clustered with the *FindNeighbors* (top 30 dimensions) and *FindClusters* function. The *FindAllMarkers* function was used to obtain differentially expressed genes of each cluster for cluster annotation. A Wilcoxon Rank Sum test was run on genes expressed in at least 10% of cells in either population, and had a 0.1-fold difference in expression (in log2 scale).

For sample integration, *SCTransform* was run on each dataset, and variations due to the difference in total number of genes and molecules detected in each cell, as well as the abundance of mitochondrial transcripts, were regressed out. Residual keratinocytes and melanocytes were removed from the dermis samples before integration. Datasets were integrated by running the *FindIntegrationAnchors* and *IntegrateData* function, using the top 3000 most variable genes and top 30 dimensions. PCs, UMAP coordinates, and clusters of the integrated dataset were then computed as described above. To perform detailed analysis of myeloid and lymphoid subpopulations, the relevant cells were subsetted and the sample integration process was repeated. The statistical significance of the differences in cell type proportions between conditions was examined by performing a chi-squared test in SPSS (v28.0.1.1).

2.4.5 Comparison of differentially expressed genes between blood CD14+ monocytes of patients affected by GVHD and transplant controls

FindMarkers function was used on the CD14+ monocytes of the integrated PBMC object to obtain differentially expressed genes between the two conditions. A Wilcoxon Rank Sum test was run on genes expressed in at least 10% of cells in either population, and had a 0.1-fold difference in expression (in log2 scale).

2.4.6 Reference mapping

The dermal myeloid reference dataset was generated by integrating all the myeloid cells from all dermis samples. Multiple query datasets were obtained: 1. myeloid cells from the blood and epidermis of all subjects in this study; 2. APCs from healthy donors (Reynolds *et al.*, 2021); 3. lesional and non-lesional myeloid cells in the skin (Nakamizo *et al.*, 2021).

To project query cells onto the reference dataset, *SCTransform* was first run on the query dataset. Anchors between the reference and query dataset were identified with the *FindTransferAnchors* function, specifying normalization method to SCT and name of reduction to PCA. Query cells were then projected onto the reference UMAP with the *MapQuery* function, using PCA from the reference for neighbour finding and UMAP as reduction model.

The blood T cell reference dataset was generated by subsetting the CD4 and CD8 T cells from the integrated PBMC dataset generated from healthy controls. Query dataset was obtained by subsetting the CD4 and CD8 T cells from all the anatomical compartments of patients affected by GVHD and transplant controls. Query cells were projected onto the reference dataset as described above.

2.4.7 DC Module scoring

Module score of myeloid cells was calculating using the *AddModuleScore* function. DC2 and DC3 modules were constructed using genes labelled "Higher in CD1C_A (cluster DC2) / Lower in CD1C_B (cluster DC3)" and "Lower in CD1C_A (cluster DC2) / Higher in CD1C_B (cluster DC3)" from supplementary table 3 (CD1C subsets) of another study (Villani *et al.*, 2017), respectively. The statistical significance of the difference in module score between DC2 and DC3 was examined by performing a Welch two sample t-test in R.

2.4.8 T cell receptor clonotype analysis

scRepertoire (v1.7.2) (Borcherding, Bormann and Kraus, 2020) (https://github.com/ncborcherding/scRepertoire) was used to perform T cell receptor clonotype analysis. Contig outputs from CellRanger were consolidated into a single list of clonotypes associated with single-cell barcode with the *combineTCR* function. A clonotype was defined by a combination of nucleotide sequence and VDJC gene sequence (CTstrict option). Clonotype calls from poor quality cells were then filtered by attaching the clonotypic information to the already quality-controlled Seurat objects using the *combineExpression* function.

To investigate the clonal expansion of T cells across anatomical compartments, alluvial plots were constructed with datasets from the blood, dermis, and epidermis of each patient, using the *compareClonotypes* function.

2.5 Ethics

All human samples were obtained with informed consent according to the protocols approved by the following: Improving Haematopoietic Stem Cell Transplantation Outcome, Newcastle and North Tyneside Research Ethics Committee 2 (reference 14/ NE/1136); or Newcastle Biobank, Newcastle and North Tyneside Research Ethics Committee 1 (reference 17/NE/0361).

Chapter 3. Profiling the peripheral blood mononuclear cells

3.1 Introduction

Monocytes and macrophages have been described as a type of accessory cell that mediates GVHD pathology by producing inflammatory cytokines in response to LPS (Nestel *et al.*, 1992). Indeed, monocytes are among the myeloid cell infiltrates in early acute GVHD of the gut (Zeiser and Blazar, 2017). Knockout of the P2Y2, a receptor for ATP (a type of DAMP), on recipient monocytes reduces GVHD lethality (Klämbt *et al.*, 2015). Activation of peripheral blood monocytes has been reported in patients affected by GVHD (Reinhardt *et al.*, 2014; Reinhardt-Heller *et al.*, 2017, 2018).

In the skin, two macrophage populations exist: transient CD14+ monocyte-derived macrophages and tissue-resident macrophages (McGovern *et al.*, 2014). Recently it has been demonstrated that donor inflammatory monocyte-derived macrophages infiltrate the dermis of patients affected by acute cutaneous GVHD and may contribute to pathogenesis (Jardine *et al.*, 2020). This study also suggested that circulating CD14+ monocytes are primed for extravasation and cytotoxic functions. Therefore, understanding the characteristics of these monocytes may contribute to new therapeutic approaches of GVHD.

Chapter aims:

1. Define the PBMC populations in transplant controls and patients affected by GVHD

2. Define the origin (donor/recipient) of each cell

3. Investigate the compositional changes of PBMC between controls and patients affected by GVHD

4. Compare the gene expression of CD14+ monocytes between transplant controls and patients affected by GVHD

3.2 Results

3.2.1 Patient details

In this study, a total of seven patients were recruited. Within the patient cohort, four patients were clinically diagnosed with acute GVHD, and one patient was diagnosed with chronic GVHD (Table 3.1). Details of the chemotherapy regimens were listed in Appendix table 1.

The samples from patients with GVHD (the term "GVH samples" was used interchangeably hereafter) were obtained between day 19 to day 261 post transplantation. Samples from the two transplant controls were obtained during bone marrow biopsy at approximately day 100 post transplantation. The important clinical events were summarized in the swimmer plot (Figure 3.1). All patients were negative for cytomegalovirus (CMV) by blood PCR test at the time of skin biopsy. Histopathology images of the skin biopsy were also obtained and analysed (displayed in the next chapter).

Figure 3.1
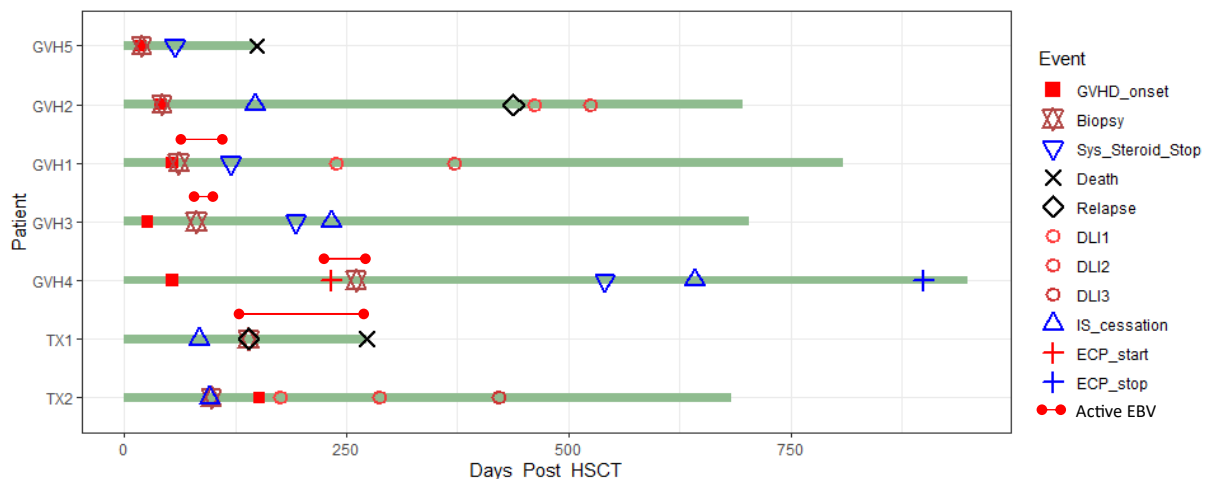
Figure 3.1 Summary of the key clinical events of each patient. Patients affected by GVHD were ordered by the timeframe between HSCT and sample being taken (biopsy), with GVH5 (top) being the sample taken at the earliest timepoint. Sys_Steroid = systemic steroid; DLI = donor lymphocyte infusion; IS = immunosuppression; ECP = extracorporeal photopheresis; EBV = Epstein-Barr virus.

Table 3.1 Summary of the patient details from whom the samples were collected. (Clinical data gathered by Dr Callum Wright)

	GVH5	GVH2	GVH1	GVH3	GVH4	TX1	TX2
Day of sample collection (post-HSCT)	19	42	61	81	261	140	98
Day of onset (post-HSCT)	19	41	54	26	54	NA	NA (152)
Age at transplant	30	60	68	32	58	48	61
Sex	Male	Male	Male	Male	Female	Female	Male
Indication	Hodgkin Lymphoma	AML	AML	ALL	AML (CR2)	AML (CR2)	AML (evolved from CMML)
Treatment protocol	RIC-Haplo	RIC-Haplo	FMA30	Cy-TBI RIC ALL	FMA30	RIC-Haplo	FMA30
Donor	Haplo	Haplo	MUD	12/12 Sib	MUD	Haplo	12/12 Sib
Conditioning intensity	RIC	RIC	RIC	RIC	RIC	RIC	RIC
TBI	YES	YES	NO	YES	NO	YES	NO
Prophylaxis	PT-Cy	PT-Cy	Alemtuzumab	Alemtuzumab	Alemtuzumab	PT-Cy	Alemtuzumab
CMV status (Donor/Recipient)	(+/-)	(-/+)	(-/+)	(+/+)	(unknown/+)	(+/+)	(-/+)
GVHD grade at biopsy	II (skin stage 3)	II (skin stage 3)	I (skin stage 2)	II (skin stage 3)	Flare, Chronic (Max grade 1)	NA	NA
CMV activation	NO	NO	NO	Prior	Prior	Prior	No
EBV at time of biopsy	0	0	< 1000 IU/mL	435000 IU/mL	< 1000 IU/mL	8670 IU/mL	0
Immunosuppression at time of biopsy	Tacrolimus	Tacrolimus + 24 hours of topical steroid	Ciclosporin + topical steroid	Ciclosporin + topical steroids	Prednisolone + ECP + Ciclosporin + Rituximab	Tacrolimus + mycophenolate mofetil	Ciclosporin
Peripheral blood chimerism at time of biopsy	CD15: 100% CD3: 100%	CD15: 100% CD3: 100%	100% (whole blood)	CD15: 100% CD3: 62%	CD15: 100% CD3: 100%	CD15: 100% CD3: 93% Marrow: 87%	CD15: 100% CD3: 57%
Transplant outcome	Non-relapse mortality (Infection)	Alive at last follow-up	Alive at last follow-up	Alive at last follow-up	Alive at last follow-up	Died from relapsed AML	Subsequently developed GVHD

3.2.2 Overview of PBMCs from patients with GVHD and transplant controls

To obtain an overview of the PBMC compartment, all blood samples from patients with GVHD, together with the samples obtained from transplant controls and healthy controls were integrated (Figure 3.2). Clusters were annotated according to the expression of marker genes of the major cell types (Figure 3.3; Table 3.2).

Figure 3.2
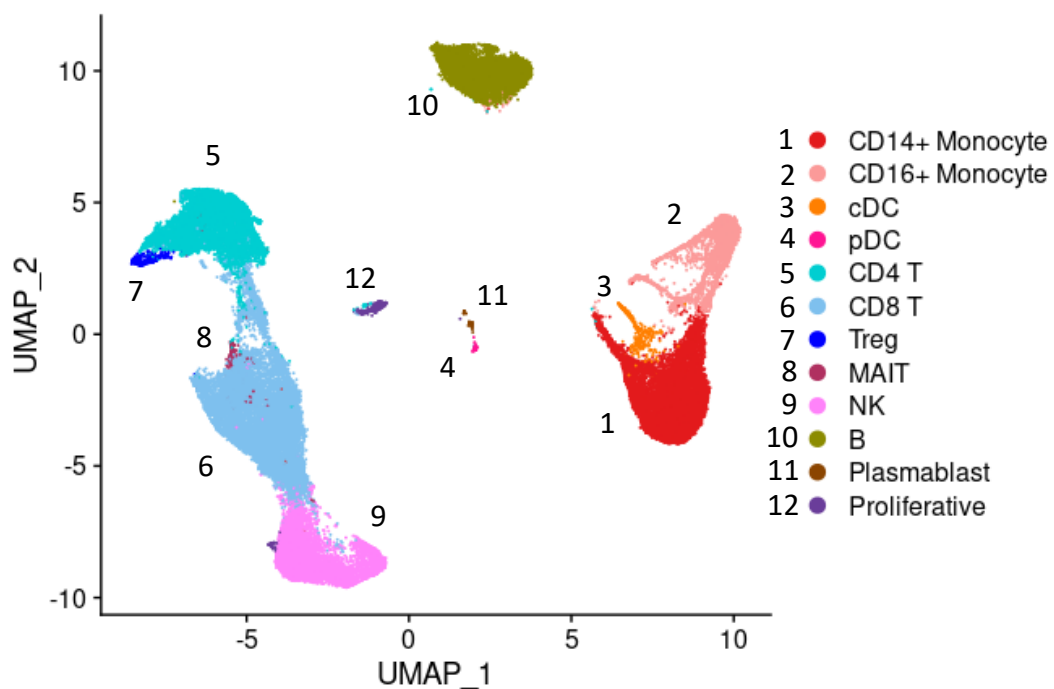
Figure 3.2 UMAP visualization of PBMCs from all samples. Samples including healthy donors (n = 3), transplant controls (n = 2) and GVHD (n = 5). Each dot represents a cell. Clusters were manually annotated by inspecting expression of canonical marker genes. cDC = classical dendritic cell; pDC = plasmacytoid dendritic cell; Treg = regulatory T cell; MAIT = mucosal-associated invariant T cell; NK = natural killer cell.

Figure 3.3
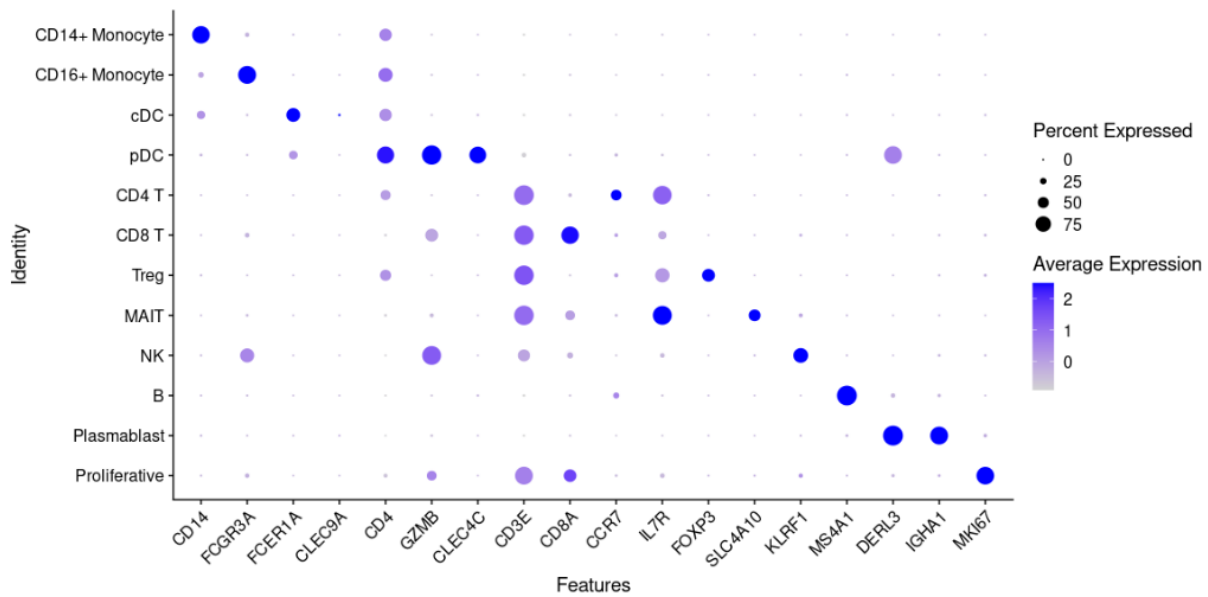
Figure 3.3 Dot plot showing the expression of marker genes of each PBMC population defined in Figure 3.2. The size of each dot corresponds to the percentage of cells within the cluster expressing the gene. The intensity of the dot corresponds to the average level of expression of the gene across all cells within the cluster.

Table 3.2

Cluster	Marker genes
CD14+ Monocyte	*CD14*
CD16+ Monocyte	*FCGR3A*
cDC	*FCER1A, CLEC9A*
pDC	*CD4, GZMB, CLEC4C*
CD4 T	*CD3E, CD4, CCR7, IL7R*
CD8 T	*CD3E, CD8A*
Treg	*CD3E, FOXP3*
MAIT	*CD3E, IL7R, SLC4A10*
NK	*FCGR3A, GZMB, KLRF1*
B	*MS4A1*
Plasmablast	*DERL3, IGHA1*
Proliferative	*MKI67*

Table 3.2 List of representative marker genes used in Figure 3.3 for cluster identification.

The cellular composition of the GVHD-affected PBMC samples was compared with healthy and transplant controls (Figure 3.4). Within the myeloid compartment, the most striking difference between GVHD-affected samples and controls was the enrichment of monocytes in GVHD-affected samples that were collected at an early timepoint post HSCT. This difference was especially profound with the CD16+ monocytes. The proportion of monocytes in GVHD-affected samples collected at a later timepoint was comparable to controls. On the other hand, the proportion of lymphoid cells in the early timepoint samples were much lower than the late timepoint samples and control samples. In terms of the type of T cells, later timepoint GVHD-affected samples and transplant controls contained more CD8 T cells than CD4 T cells, whereas healthy controls contained more CD4 T cells than CD8 T cells. B cells were present in two of the five GVHD-affected samples (GVH1 and GVH3). Moreover, there were also more proliferative cells in GVHD-affected samples compared to the controls.

Figure 3.4
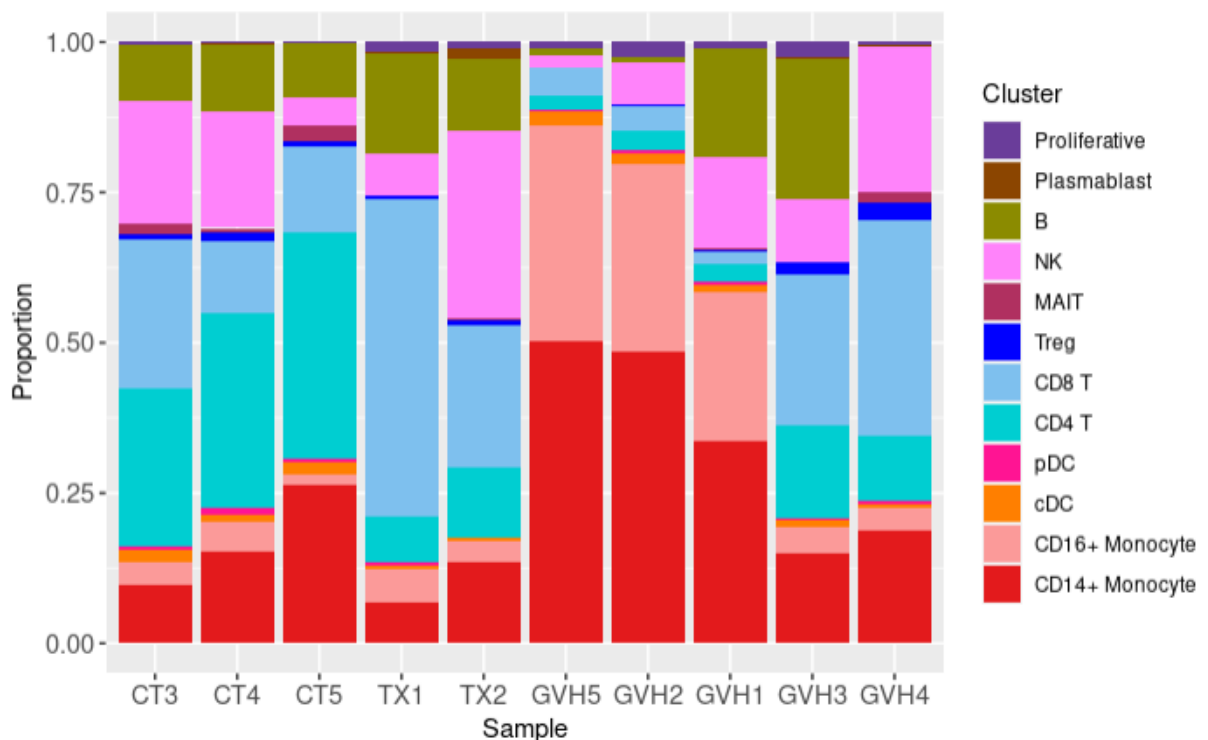
Figure 3.4 Bar chart showing the proportion of each PBMC population by sample. Samples from patients affected by GVHD were ordered by the timeframe between HSCT and sample being taken, with GVH5 (left) being the sample taken at the earliest timepoint. Dataset CT5 was downloaded from 10x Genomics (Human PBMC from a Healthy Donor, 10k cells (v2)). CT = healthy donor; TX = transplant control; GVH = graft versus host disease.

3.2.3 Individual PBMC samples

Since the manifestation of GVHD could be diverse, the PBMC sample of each patient was then analysed separately. The samples from transplant controls were clustered (Figure 3.5, panel A) and the origin of each cell was shown (Figure 3.5, panel B). T cell subpopulations could be identified by using additional markers: *CCR7* for naïve/central memory (CM), *KLRB1* for CD4 effector memory (EM) and *CCL5* for CD8 EM. CD8 EM was the largest T cell population in both samples. Intriguingly, in one of the samples (TX1), a cluster of cells expressing haematopoietic progenitor markers could be identified (Table 3.3), which was labelled "AML blast".

Figure 3.5
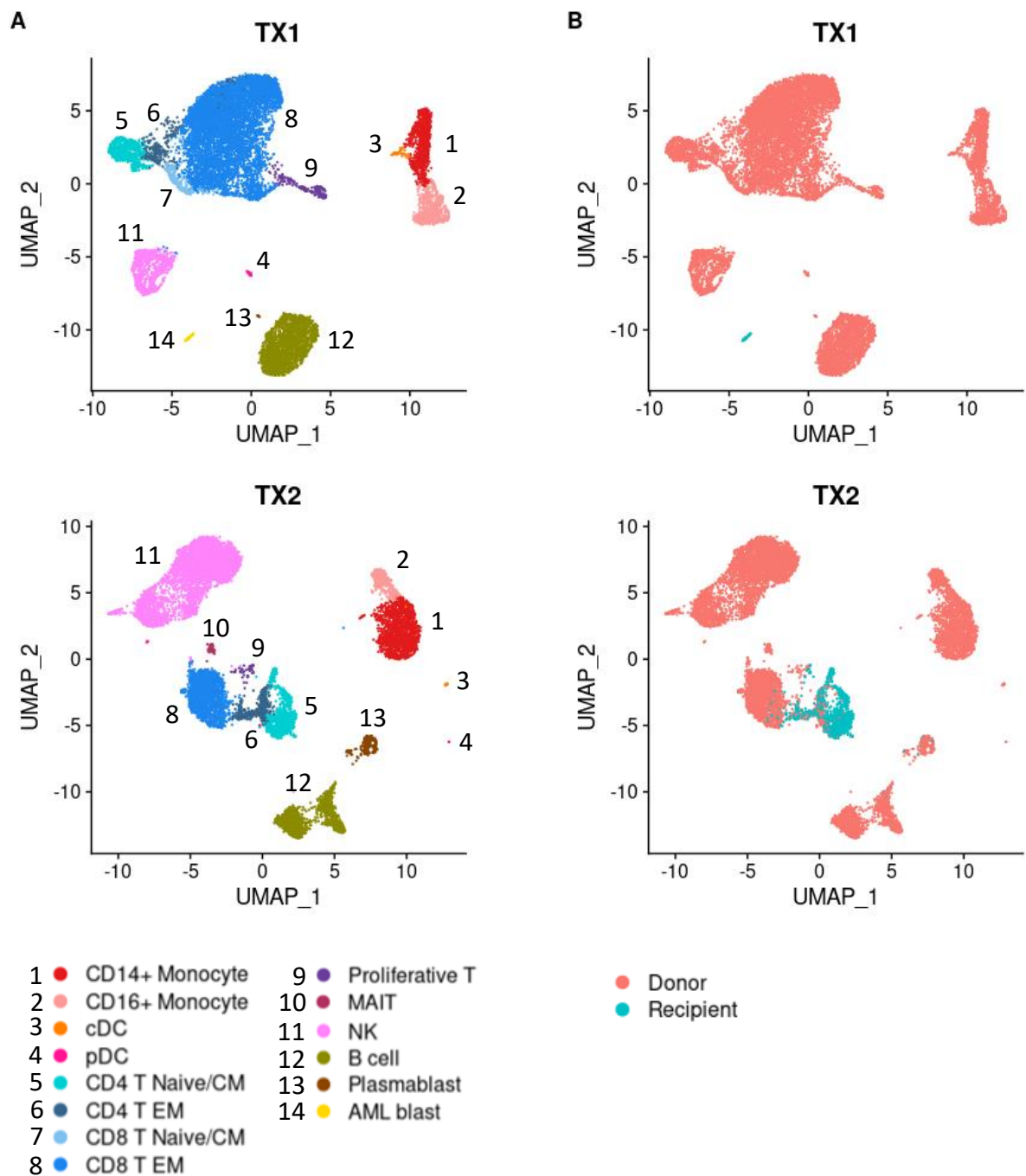
Figure 3.5 UMAP visualization of each PBMC sample from transplant controls. (A) Clusters were manually annotated by inspecting expression of canonical marker genes. (B) Each cell was annotated by its origin, either derived from the patient (recipient) or from the graft that the patient received (donor).

Table 3.3

p_val	avg_log2FC	pct.1	pct.2	p_val_adj	gene
0	3.091907	0.985	0.012	0	*PRSS57*
0	2.105391	0.971	0.008	0	*EGFL7*
0	2.068955	0.971	0.037	0	*FAM30A*
0	2.012772	0.897	0	0	*MYCN*
0	1.946316	0.926	0.002	0	*GATA2*
0	1.796355	0.926	0	0	*XIST*
0	1.751574	0.956	0.003	0	*CRNDE*
0	1.700635	0.971	0.01	0	*BAHCC1*
0	1.687692	0.956	0.002	0	*HOXA9*
0	1.686433	0.912	0.023	0	*FSCN1*
0	1.616884	0.926	0.035	0	*AC084033.3*
0	1.482371	0.779	0.007	0	*LAPTM4B*
0	1.466646	0.838	0.001	0	*SMIM24*
0	1.420787	0.912	0.001	0	*TFPI*
0	1.389162	0.868	0.027	0	*ATP8B4*

Table 3.3 List of top differentially expressed genes of the AML blast cluster from the PBMC of TX1. The average fold change was calculated between this cluster and all other clusters, presented in log2 scale (avg_log2FC). pct.1 = percentage of cells in the cluster expressing the gene. pct.2 = percentage of cells in all the other clusters where the gene was detected. p_val = p value; p_val_adj = adjusted p value.

In terms of cell origin, cells from both the myeloid and lymphoid compartments of TX1 were entirely donor derived. The AML blasts, in contrast, were recipient derived. Whereas in the other transplant control sample (TX2), the myeloid compartment was entirely donor derived but in the CD4 T cell compartment most cells were recipient derived, contrary to the CD8 T cell compartment in which the majority of cells were donor derived.

Next, the samples from patients affected by GVHD were clustered (Figure 3.6, panel A) and the origin of each cell was shown (Figure 3.6, panel B). Samples were ordered from top to bottom by the timeframe between HSCT and the sample being taken. Subsets of T cell did not form separate clusters, potentially owing to small number of T cells. However, a cluster of gamma-delta T cells expressing CD3E, TRDC and lack CD4 and CD8A expression could be identified in sample GVH3. Moreover, B cells were only present in two out of five samples (GVH1 and GVH3).

Regarding their origin, cells from both the myeloid and lymphoid compartments were mostly donor derived for all the samples except for sample GVH3. In this sample, the Treg population and gamma-delta T cell population were mostly recipient derived. The proliferative T cell population and CD8 T cell population also contained a considerable proportion of recipient-derived cells.

Figure 3.6
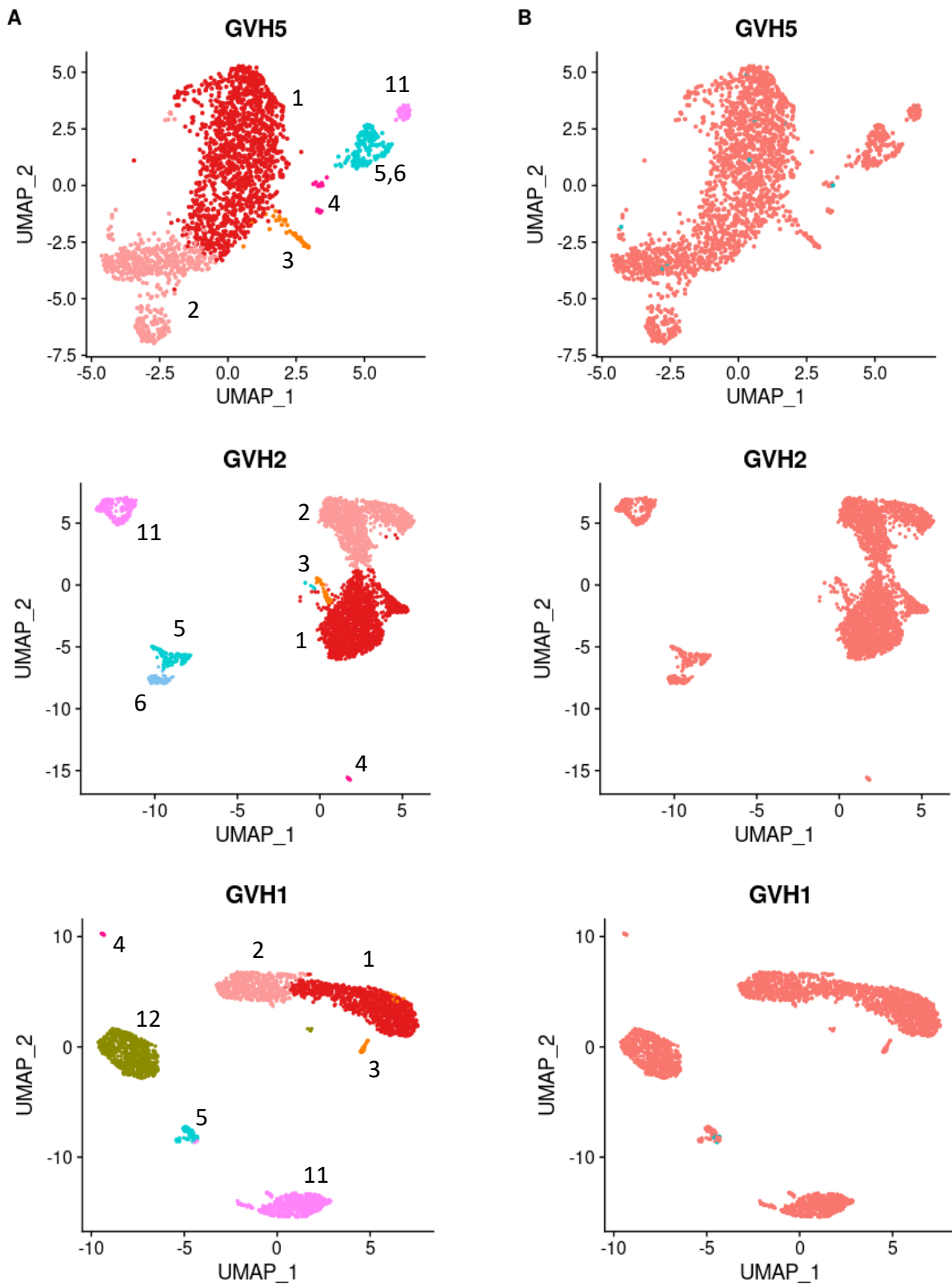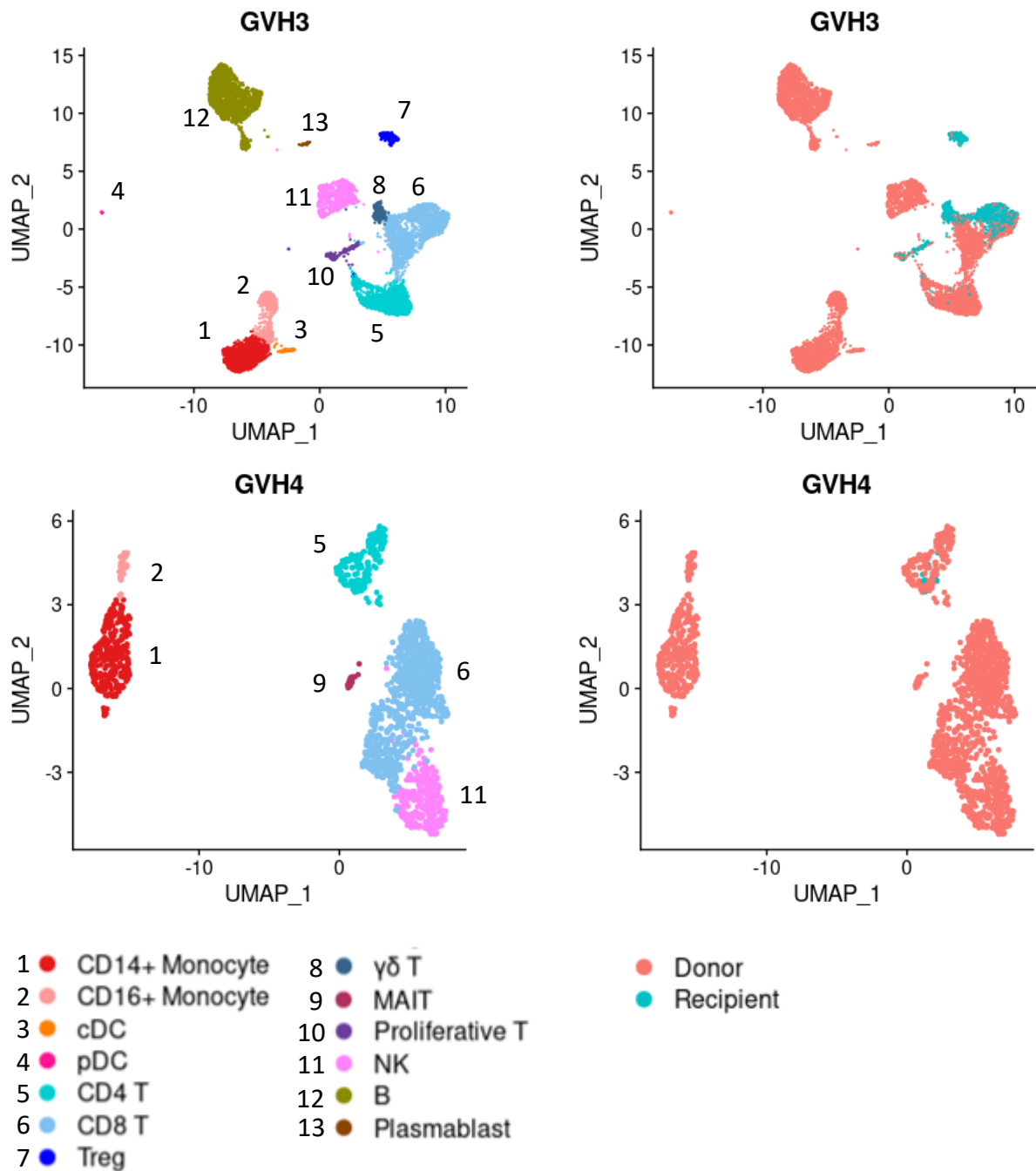
Figure 3.6 UMAP visualization of each PBMC sample from patients affected by GVHD. (A) Clusters were manually annotated by inspecting expression of canonical marker genes. Samples were ordered by the timeframe between HSCT and sample being taken, with GVH5 (top) being the sample taken at the earliest timepoint. (B) Each cell was annotated by its origin, either derived from the patient (recipient) or from the graft that the patient received (donor).

3.2.4 Monocyte priming

As priming of blood CD14+ monocytes potentially represents an important event in GVHD pathology, the gene expression of CD14+ monocytes from patients affected by GVHD was compared to those from transplant controls (Figure 3.7). Adhesive glycoprotein *THBS1*, IFN-response genes *IFITM1, IFITM3, MT2A, IFI6, ISG15*, and pro-inflammatory transcription factor *KLF6* were upregulated by the monocytes from patients with GVHD. On the other hand, immunomodulatory genes such as *IL10RA* and *NRROS* were downregulated. Interestingly, genes that are related to TNF-α signalling (*TNFRSF1B, TNFAIP2*) were also downregulated.

Figure 3.7
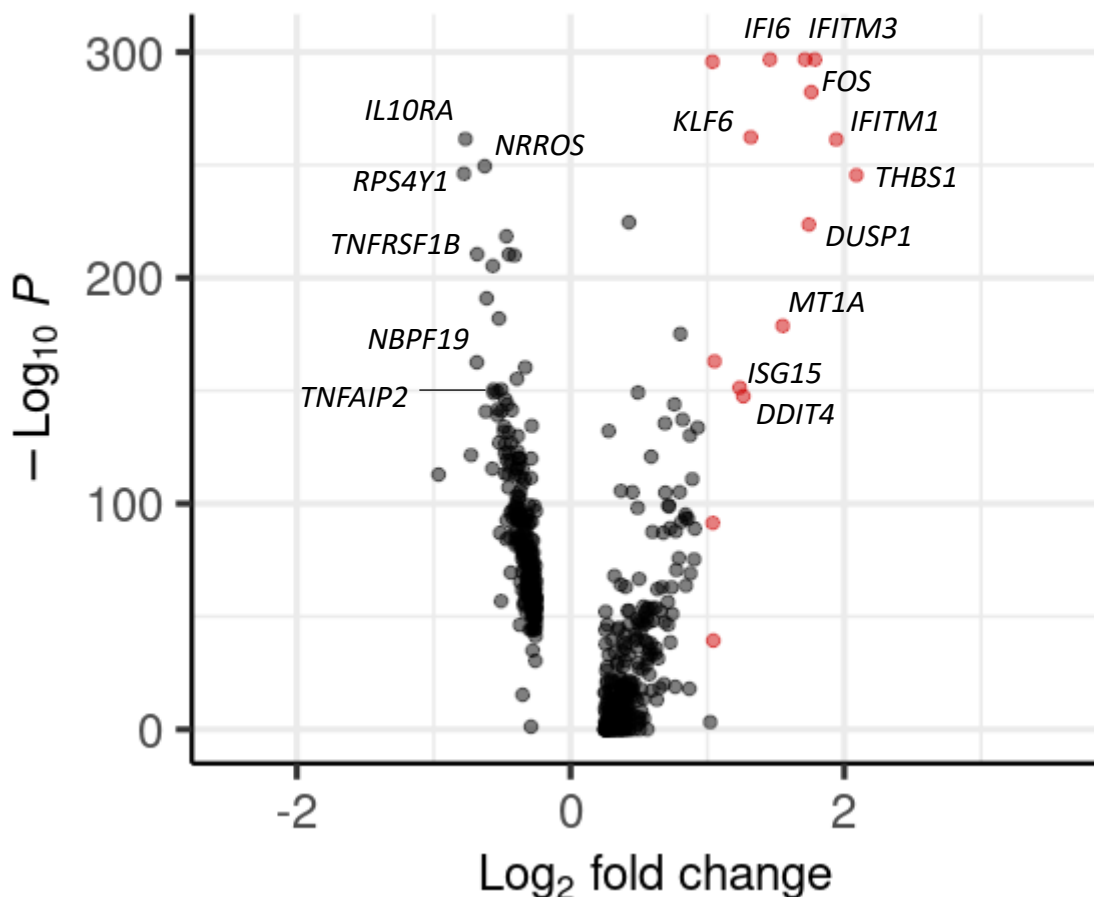
Figure 3.7 Volcano plot showing the top differentially expressed genes between the blood CD14+ monocytes of patients affected by GVHD and transplant controls. Points in red showed genes with absolute log2 fold change larger than 1.

3.3 Chapter discussion

The pathology of GVHD requires the orchestration of various haematopoietic and non-hematopoietic cells of donor and recipient origin, however these cells were not well characterized, especially in clinical settings. Here, the PBMC from patients with GVHD at various timepoints post HSCT, ranging from early (Day 19) to late (Day 261), were sampled along with transplant controls at approximately day 100 post transplantation. Results showed that the clinical variations of the patients receiving transplant were reflected by scRNA-seq data, such as the interval between HSCT and sampling, their EBV status, and the prophylaxis the patients received. There was an inverse relationship between the timepoint of the sample and the proportion of monocytes in the blood, with the early timepoint samples having a higher proportion of monocytes compared to the late timepoint samples. An inversion of CD4 T cell to CD8 T cell ratio was also observed in HSCT recipients compared to healthy controls. These patterns exemplified immune reconstitution after HSCT (Stern et al., 2018). Moreover, the presence of B cells correlated with EBV reactivation. In one of the samples (GVH1), a population of B cells was detectable by scRNA-seq even when the EBV level was very low (< 1000 IU/mL). This patient subsequently (after a week) had high levels of EBV (392000 IL/mL). B cells were not detected in one of the EBV-positive patients (GVH4), likely due to the effect of rituximab, an immunosuppressive medication that targets B cells in the blood. Additionally, the ratio of NK cells to T cells may indicate the type of prophylaxis the patients received. Comparing the two transplant controls, the control who received alemtuzumab (TX2), which preferentially depletes T cells (Roex *et al.*, 2021), had a higher proportion of NK cells and lower proportion of T cells in the blood.

Comparison of cellular composition between conditions showed that monocytes, especially the CD16+ population, were enriched in GVHD-affected samples collected at an early timepoint compared to those collected at a later timepoint and controls. The pathogenic roles of intermediate monocytes (CD14+ CD16+) has been documented in patients with GVHD and it is correlated to the activation of a subset of pro-inflammatory Th17 cell (Reinhardt-Heller *et al.*, 2017). Although in the current study the CD14+ CD16+ monocyte population was not specifically inspected, the relative abundance of monocytes together with low proportions of lymphocytes in these patients suggest that monocytes are important mediators of GVHD. The origin of each cell, whether donor or recipient derived, was determined and as anticipated, all

monocytes were donor derived as a result of haematopoiesis of the graft. Partial donor chimerism was observed in two transplant recipient (GVH3 and TX2), potentially due to incomplete ablation of host immunity during conditioning or persistence of Trm (Divito *et al.*, 2020; Strobl *et al.*, 2021), which will be discussed in subsequent chapters.

Recently, it has been shown that circulating CD14+ monocytes in patients with GVHD are primed for extravasation and cytotoxic functions (Jardine *et al.*, 2020). In the current study, the gene expression of CD14+ monocytes from patients with GVHD was compared to those from transplant controls. Results showed evidence of monocyte activation, however genes related to cytotoxicity were not included among the top differentially expressed genes, potentially due to sample and technical variations. The downregulation of genes related to TNF-α signalling was also intriguing, as during GVHD, macrophages were primed to produce large amounts of TNF-α (Nestel *et al.*, 1992). This finding may indicate a negative-feedback loop of TNF-α signalling.

A limitation of this study is that ideal transplant control samples (patient recovers from HSCT without relapse and developing GVHD) are challenging to collect, as transplant controls that are free from GVHD and relapse at the time of sampling can still subsequently develop GVHD and/or relapse. One approach to address this issue would be to collect samples from multiple transplant controls at day 100 for long-term storage and select the most suitable samples for scRNA-seq at a later date based on the conditions of the patients. Nonetheless, sample storage and retrieval may inadvertently introduce variations to the samples such as affecting the proportion of cell types and their RNA expression. On the other hand, examining samples from patients before disease manifest (relapse and/or developing GVHD) may also provide valuable insights. Therefore, I opted to perform scRNA-seq with fresh samples at day 100, from transplant controls who at the time were free from relapse and GVHD.

Out of the two transplant controls recruited, one control (TX1) was subsequently diagnosed with relapsing AML from a clinical bone marrow biopsy sample (taken on the same day as the scRNA-seq was carried out), while the other control (TX2) subsequently developed GVHD (approximately 1.5 months after scRNA-seq was carried out). Analysis of cellular composition of the blood revealed that, apart from the common myeloid and lymphoid cells that were found in other GVH and healthy control samples, AML-blasts could also be identified from the transplant control who had relapsed. Unlike most other cells which were derived from the graft, these cells were recipient derived and expressed genes of haematopoietic progenitor

50

cells. The detection and characterization of measurable residual diseases in AML has been a popular area of research, as understanding how cancer cells survive chemotherapy is crucial for the development of more effective treatments (Robinson et al., 2022). Further characterization of AML blasts in the blood with single-cell technologies may shed new light on AML biology and potentially lead to the development of less invasive diagnostics and better patient stratification.

Chapter 4. Profiling the dermal immune cells

4.1 Introduction

Skin is the most frequently affected organ by GVHD (Ferrara *et al.*, 2009). Clinically, cutaneous acute GVHD usually presents as a pruritic maculopapular rash, while chronic GVHD involves various lichenoid and sclerodermatous changes (Santos e Sousa, Bennett and Chakraverty, 2018). Skin biopsies of patients with acute cutaneous GVHD are characterized by vacuolar interface dermatitis with extensive leukocyte infiltration, and in its chronic form, characterized by features such as hyperkeratosis, acanthosis and basal cell necrosis (Santos e Sousa, Bennett and Chakraverty, 2018).

The role of donor naïve T cells in GVHD has been well-established, however the role of myeloid cells has gained significant interest in recent years. Cutaneous GVHD can be mediated with minimal T cell involvement, but macrophages are almost always present and very abundant (Nishiwaki *et al.*, 2009). However, it is not known whether these cells are donor or host derived. Host resident macrophages survive up to one year post-HSCT and are potent inducers of memory T cells (Haniffa *et al.*, 2009). Murine studies manipulating the CSF-1 axis subsequently showed that host macrophages confer protective effect against GVHD while donor-derived macrophages exacerbates GVHD (MacDonald *et al.*, 2010; Hashimoto *et al.*, 2011; Alexander *et al.*, 2014). Recently, it has been demonstrated that in humans, donor monocyte-derived macrophages infiltrate the dermis and potentially mediate GVHD pathology (Jardine *et al.*, 2020). Meanwhile, DCs are also present in the human skin but their roles in GVHD are elusive.

With scRNA-seq become more widely available, multiple cell types can be characterized simultaneously and at a higher resolution. Myeloid cells in human skin have been characterized with scRNA-seq technologies and are implicated in a range of immune disorders (Xue *et al.*, 2020; Nakamizo *et al.*, 2021; Reynolds *et al.*, 2021). A subset of cDC2, namely DC3, is also found to be enriched in lesional skin of patients with psoriasis (Nakamizo *et al.*, 2021).

Chapter aims:

1. Define the dermal cell populations in transplant controls and patients affected by GVHD

2. Define the origin (donor/recipient) of each cell

3. Investigate the compositional changes of dermal cells between controls and patients affected by GVHD

4. Define the dermal myeloid cell subpopulations
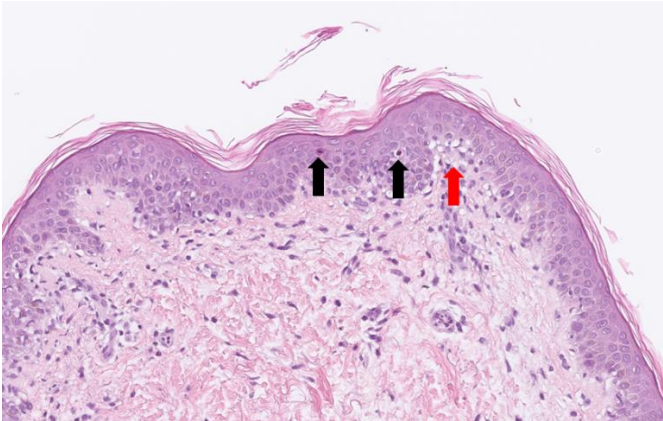
4.2 Results
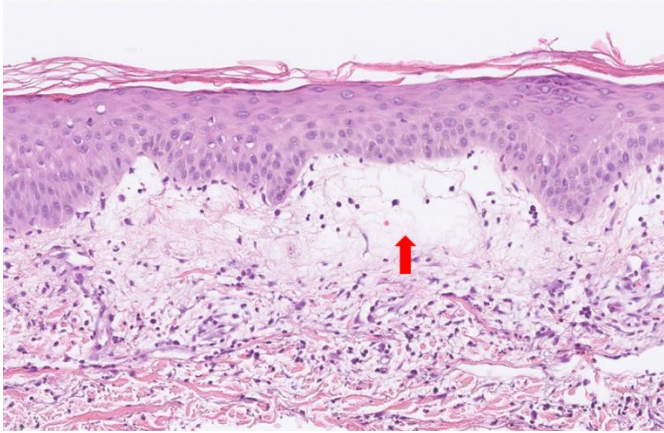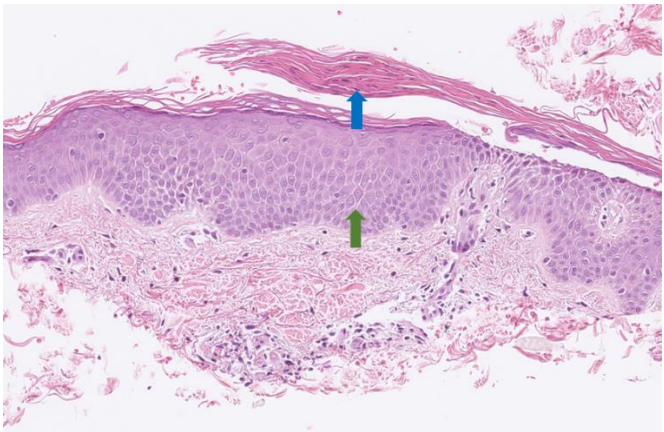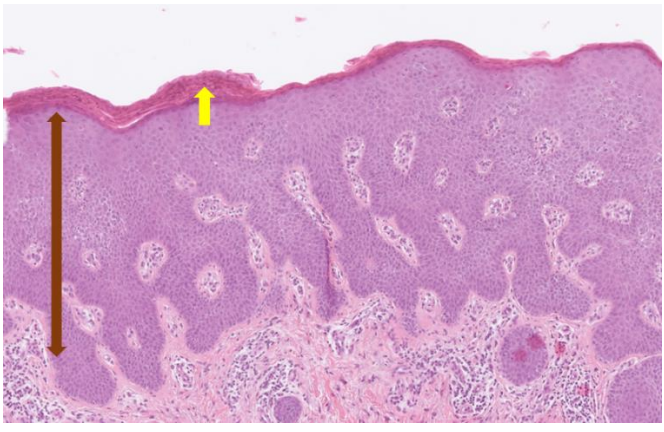
4.2.1 Histopathology of skin biopsies from patients with GVHD

The skin biopsies from four of the five patients with GVHD were examined by an independent pathologist (Table 4.1). Interestingly, despite all the patients presenting with clinical GVHD features, one skin biopsy (GVH3) lacked typical histological features of GVHD, instead demonstrating features of spongiotic dermatitis. One skin biopsy (GVH4) demonstrated histological features of chronic GVHD, in keeping with the clinical presentation of this patient.

Table 4.1

(Images of the pathology slides were provided by Dr James Sampson. Labelling of images was assisted by Dr James Sampson and Dr Callum Wright. Description of the images were extracted from the clinical pathology reports with the assistance by Dr James Sampson.)

Sample	Description
(A) GVH2	Mild lymphocytic inflammation is present at the dermo-epidermal junction which is associated with basal cell hydropic degeneration and occasional dyskeratotic cells, many above the basal layer. The number of eosinophils is not significantly increased. There is no overlying parakeratosis, in keeping with an acute process. The appearances are those of a vacuolar interface dermatitis with features fully in keeping with GVHD of grade II. Support the diagnosis of GVHD (grade III).

(B) GVH1	There is basal cell hydropic degeneration with areas of clefting, associated with oedema and moderate amounts of lymphocytic infiltrates within dermis. Scattered colloid bodies are seen within the epithelium as well as intraepithelial lymphocytes. Eosinophils are present but are not prominent. A dPAS stain for fungi is negative. The appearances are those of a florid vacuolar interface dermatitis with features that support the diagnosis of GVHD (grade III).
(C) GVH3	There is focal parakeratosis. There is mild spongiotic changes with intra cellular oedema and focal exocytosis of lymphocytes. There is no evidence of acute GVHD. The appearances are in keeping with spongiotic dermatitis, consistent with eczema.
(D) GVH4	Sections show crosscut pieces of skin with psoriasiform hyperplasia and confluent parakeratosis; aggregates of neutrophil polymorphs are present in the stratum corneum. There is a perivascular inflammatory cell infiltrate comprising lymphocytes and plasma cells within the papillary dermis. There is no interface change. The appearances are in keeping with chronic graft versus host disease.

Table 4.1 H&E-stained images of skin biopsies from patients with GVHD. Sample from patients affected by GVHD were ordered by the timeframe between HSCT and sample being taken. Coloured arrows represent some key features: black – apoptotic keratinocytes; red – vacuolar interface; blue – parakeratosis; green – spongiosis; brown – acanthosis; yellow – presence of neutrophils.

4.2.2 Overview of dermal cells from patients with GVHD and transplant controls

To obtain an overview of the dermal compartment, all dermis samples from patients with GVHD, together with the samples obtained from transplant controls and healthy controls were integrated (Figure 4.1). Clusters were annotated according to the expression of marker genes of the major cell types (Figure 4.2, Table 4.2).

Figure 4.1
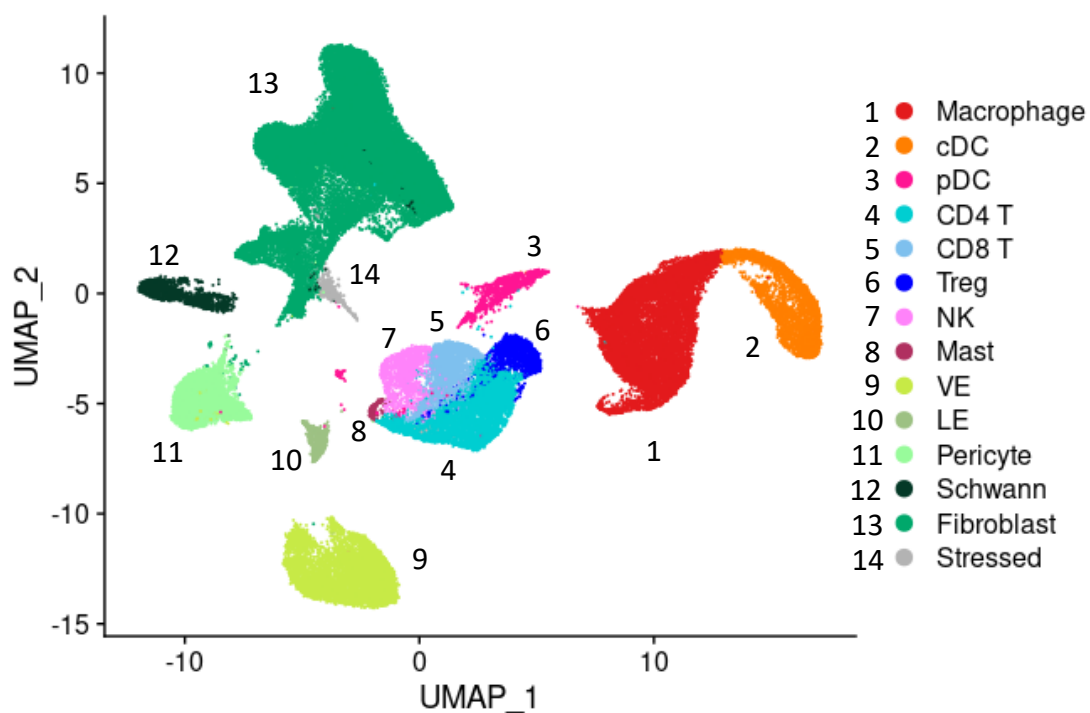
Figure 4.1 UMAP visualization of dermal cells from all samples. Samples including healthy donors (*n* = 4), transplant controls (*n* = 2) and GVHD (*n* = 5). Clusters were manually annotated by inspecting expression of canonical marker genes. cDC = classical dendritic cell; pDC = plasmacytoid dendritic cell; Treg = regulatory T cell; NK = natural killer cell; VE = vascular endothelial cell, LE = lymphatic endothelial cell.

Figure 4.2
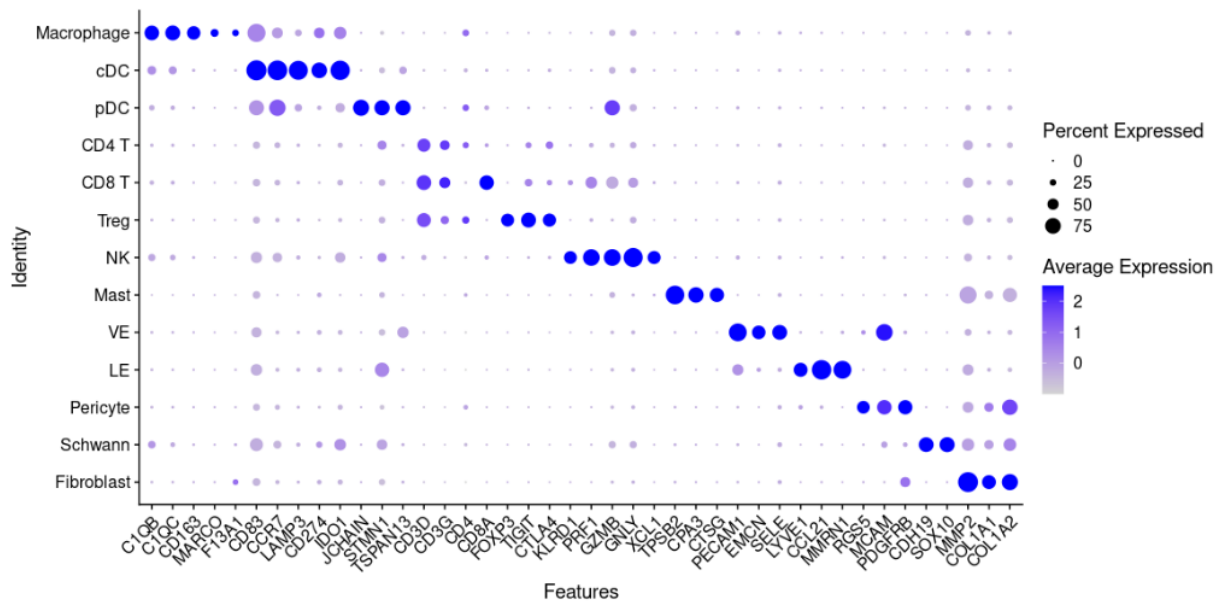
Figure 4.2 Dot plot showing the expression of marker genes of each dermal cell population defined in Figure 4.1. The size of each dot corresponds to the percentage of cells within the cluster expressing the gene. The intensity of the dot corresponds to the average level of expression of the gene across all cells within the cluster.

Table 4.2

Cluster	Marker genes
Macrophage	*C1QB, C1QC, CD163, MARCO, F13A1*
cDC	*CD83, CCR7, LAMP3, CD274, IDO1*
pDC	*JCHAIN, STMN1, TSPAN13*
CD4 T	*CD3D, CD3G, CD4*
CD8 T	*CD3D, CD3G, CD8*
Treg	*CD3D, CD3G, FOXP3, TIGIT, CTLA4*
NK	*KLRD1, PRF1, GZMB, GNLY, XCL1*
Mast	*TPSB2, CPA3, CTSG*
VE	*PECAM1, EMCN, SELE*
LE	*LYVE1, CCL21, MMRN1*
Pericyte	*RGS5, MCAM, PDGFRB*
Schwann	*CDH19, SOX10*
Fibroblast	*MMP2, COL1A1, COL1A2*

Table 4.2 List of representative marker genes used in Figure 4.2 for cluster identification.

The cellular composition of the GVHD-affected dermis samples was compared with healthy and transplant controls (Figure 4.3). A chi-square test was performed and there was a significant association between the cell types and the condition of the samples, χ^2 (80) = 17357, p < .001 (Appendix table 2). There was a notable enrichment of immune cells in GVHD-affected dermis compared to healthy and transplant controls, and the type of enriched immune cells appeared to be dependent on the timeframe between HSCT and sample being taken. The enrichment of myeloid cells, including macrophages, pDCs and cDCs, were more pronounced in early timepoint samples, whereas the enrichment of lymphoid cells, including CD4 T cells, CD8 T cells and Tregs, were evident in late timepoint samples. In these samples, Treg constituted a relatively large proportion of the immune cell population compared to other samples. Additionally, NK cells were enriched in all GVHD-affected dermis regardless of timepoint.

Figure 4.3
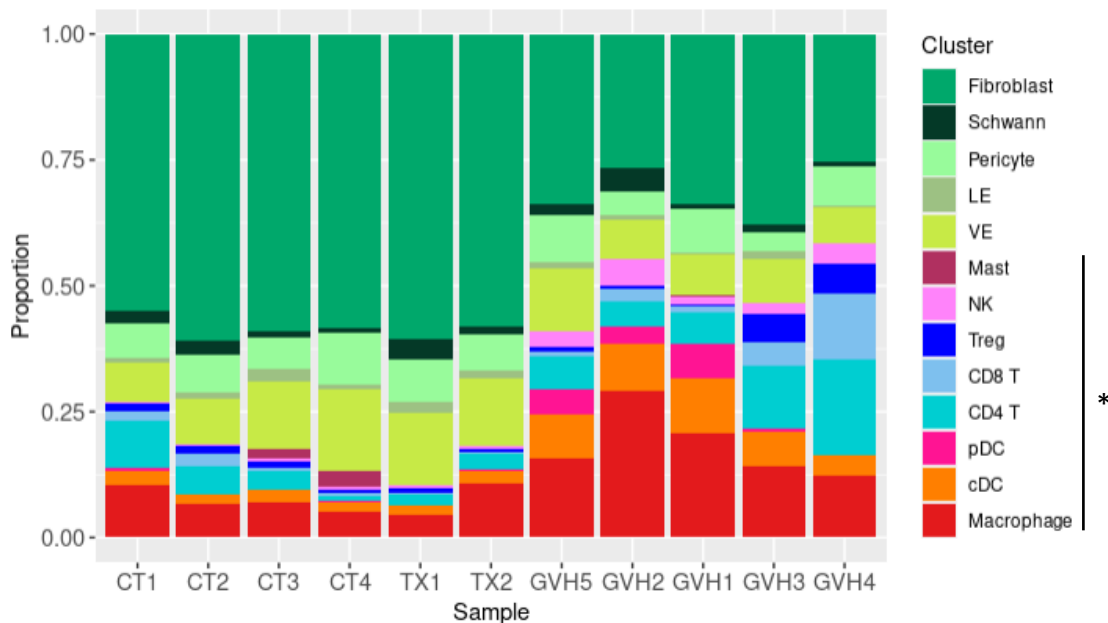
Figure 4.3 Bar chart showing the proportion of each dermal cell population by sample. Sample from patients affected by GVHD were ordered by the timeframe between HSCT and sample being taken, with GVH5 (left) being the sample taken at the earliest timepoint. (* p < 0.05 GVH samples compared to controls). CT = healthy donor; TX = transplant control; GVH = graft versus host disease.

4.2.3 Individual dermis samples

Next, the dermis sample of each patient was analysed separately. The samples from transplant controls were clustered (Figure 4.4, panel A) and the origin of each cell was shown (Figure 4.4, panel B). The major dermal immune and stromal cell populations could be identified. Most immune cells were donor derived. Macrophages and T cells of recipient origin could be detected.

Figure 4.4
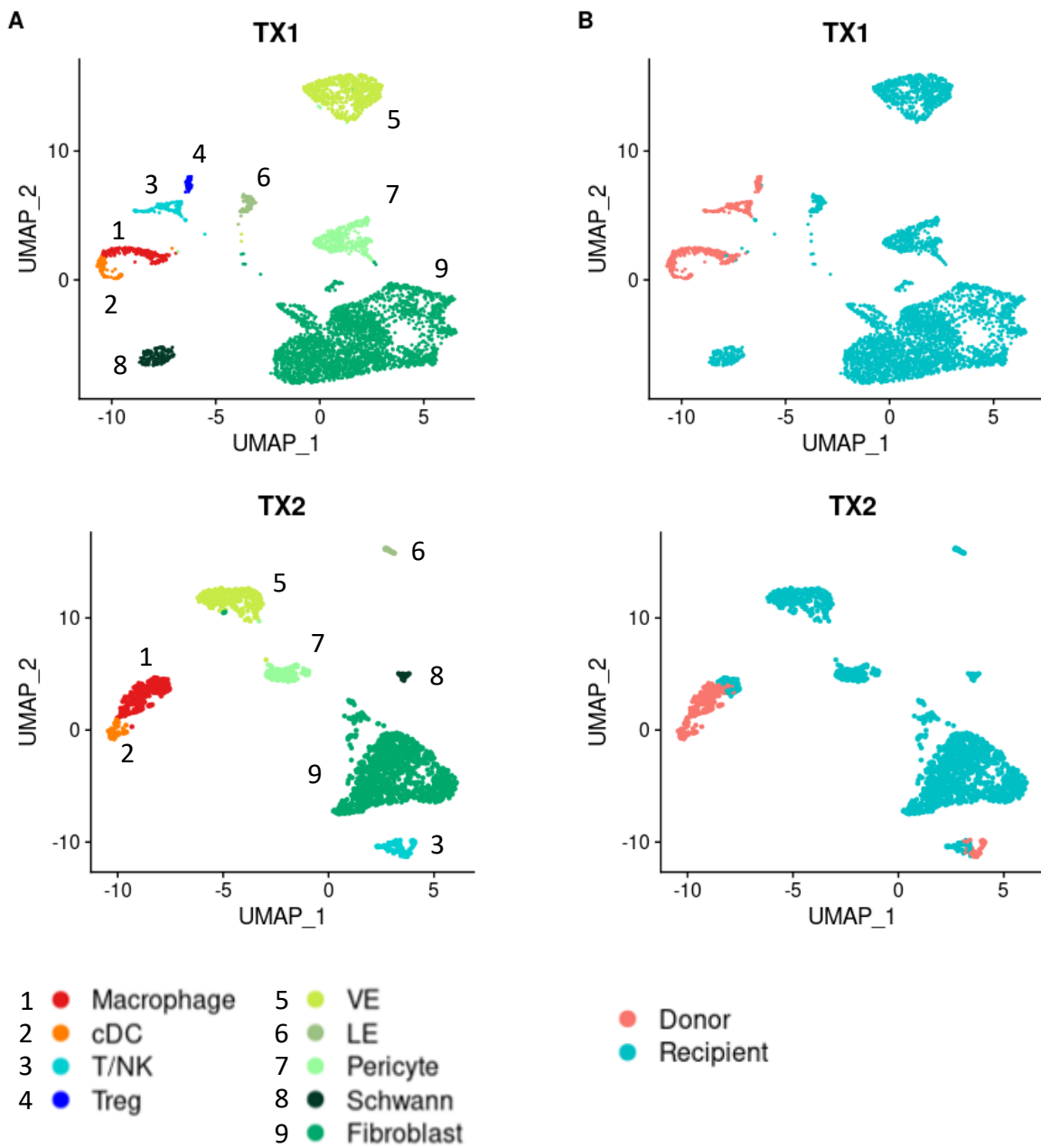
Figure 4.4 UMAP visualization of each dermis sample from transplant controls. (A) Clusters were manually annotated by inspecting expression of canonical marker genes. (B) Each cell was annotated by its origin, either derived from the patient (recipient) or from the graft that the patient received (donor).

The samples from patients affected by GVHD were then clustered (Figure 4.5, panel A) and the origin of each cell was shown (Figure 4.5, panel B). Samples were ordered from top to bottom by the timeframe between HSCT and the sample being taken. Within the myeloid compartment, pDCs were present in samples that were collected at an early timepoint post HSCT (GVH5, GVH2 and GVH1). Regarding the origin of cells, most cells were donor derived with the exception of a macrophage population, termed resident macrophage. These macrophages were recipient derived in early timepoint samples, and were donor derived in samples collected at a later timepoint. Resident macrophages could be identified by the expression of a specific suite of marker genes. The details of myeloid cell populations would be discussed in subsequent sections. On the other hand, most lymphoid cells were donor derived. Interestingly, the presence of recipient-derived T cells could also be detected to various extents. A small population of donor-derived migratory Langerhans cells expressing *CD1A* and *CD207* was also present in sample GVH3.

Figure 4.5
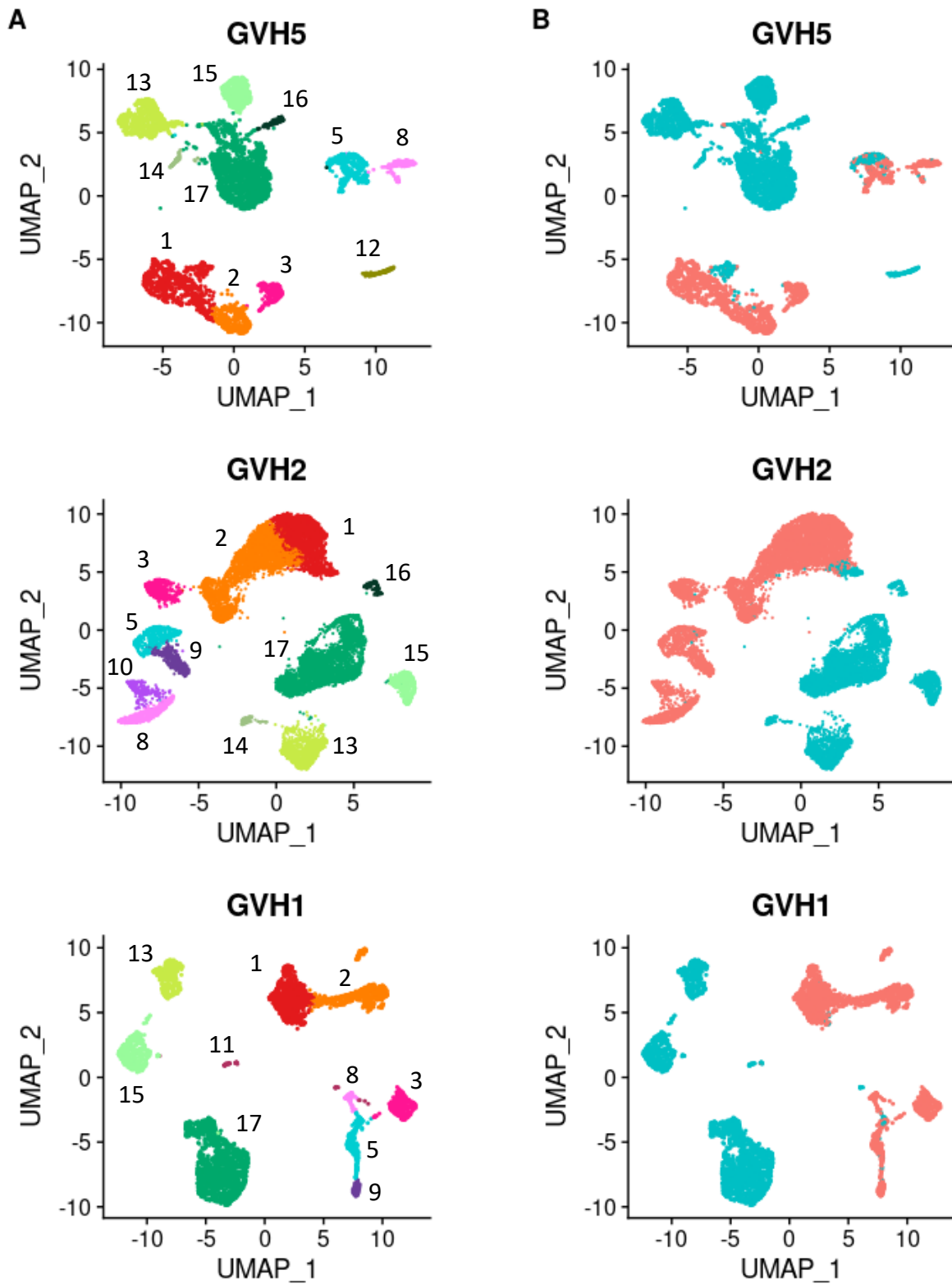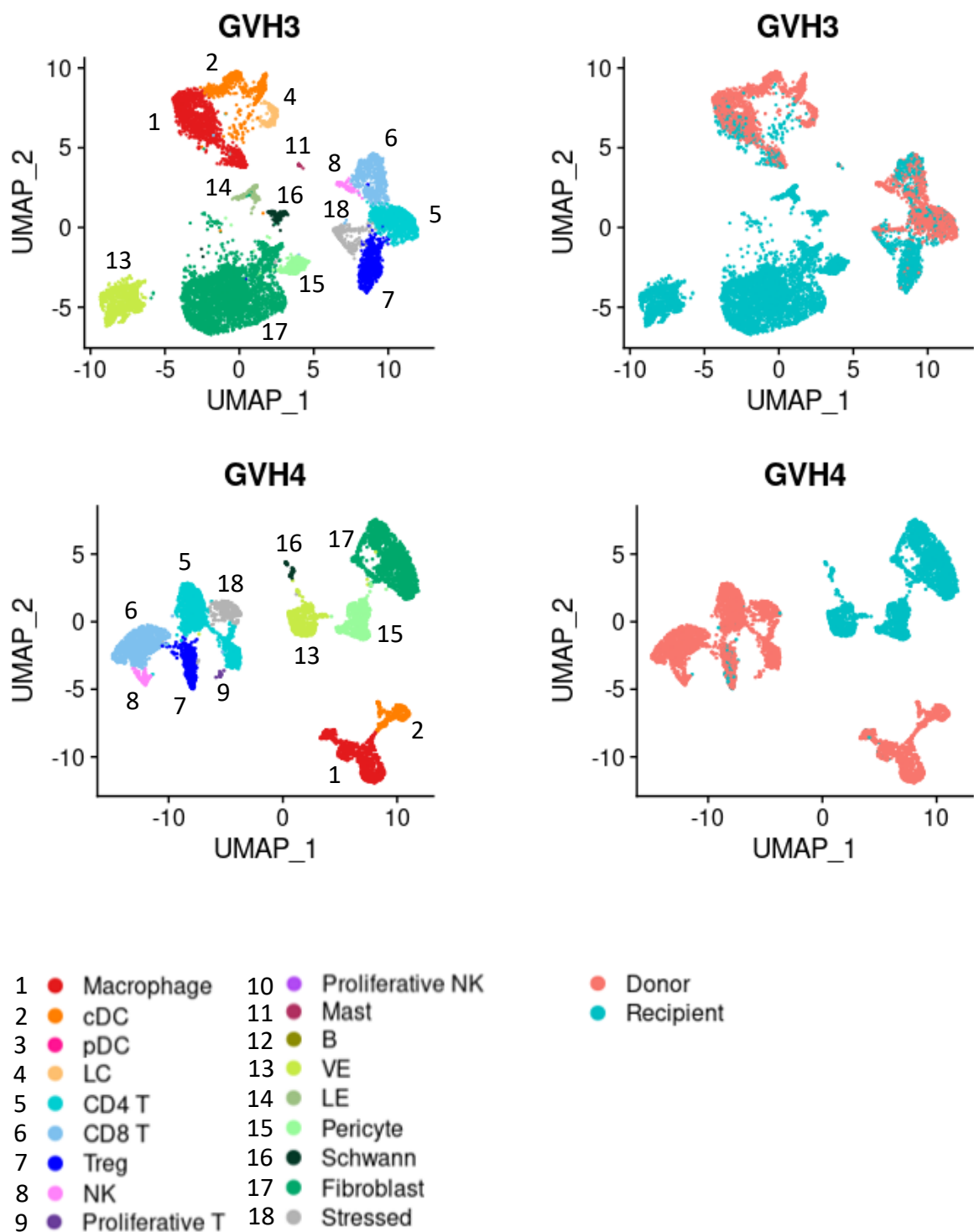
Figure 4.5 UMAP visualization of each dermis sample from patients affected by GVHD. (A) Clusters were manually annotated by inspecting expression of canonical marker genes. Samples were ordered by the timeframe between HSCT and sample being taken, with GVH5 (top) being the sample taken at the earliest timepoint. (B) Each cell was annotated by its origin, either derived from the patient (recipient) or from the graft that the patient received (donor).

4.2.4 Characterization of dermal myeloid cells

To study the myeloid populations in detail, myeloid cells were subsetted and re-analysed. Clustering resulted in six macrophage populations and five DC populations (Figure 4.6, panel A). The origin of cells was also displayed (Figure 4.6, panel B). Clusters were annotated according to the expression of canonical marker genes (Figure 4.7, Table 4.3).

Figure 4.6
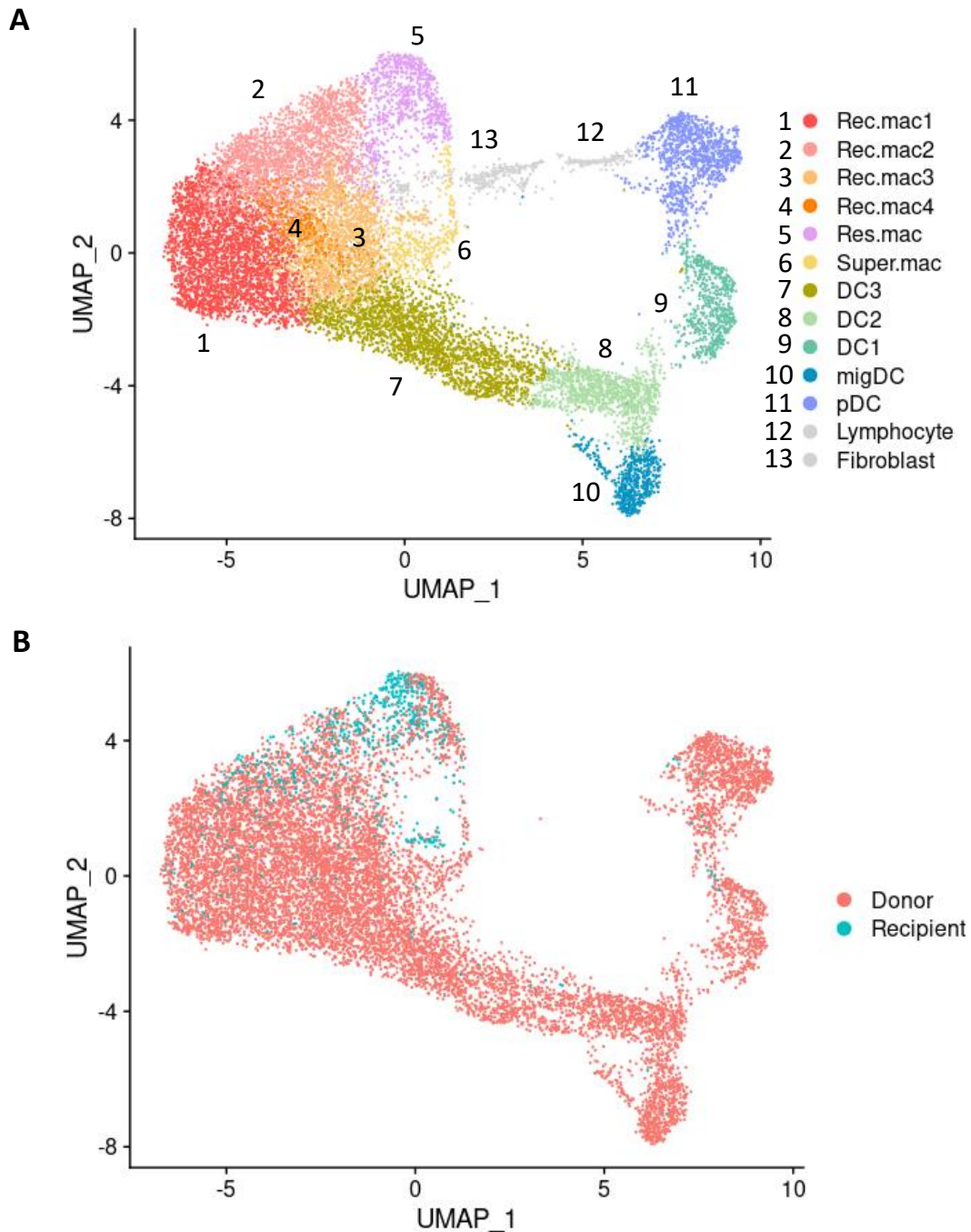
Figure 4.6 UMAP visualization of dermal myeloid cells from all samples. Samples including healthy donors (*n* = 4), transplant controls (*n* = 2) and GVHD (*n* = 5). (A) Clusters were manually annotated by inspecting expression of canonical marker genes. (B) Each cell from transplant recipients was annotated by its origin, either derived from the patient (recipient) or from the graft that the patient received (donor). Rec.mac = recruited macrophage; Resi.mac = resident macrophage; Super.mac = super-activated macrophage; DC = dendritic cell; migDC = migratory dendritic cell; pDC = plasmacytoid dendritic cell.

Figure 4.7
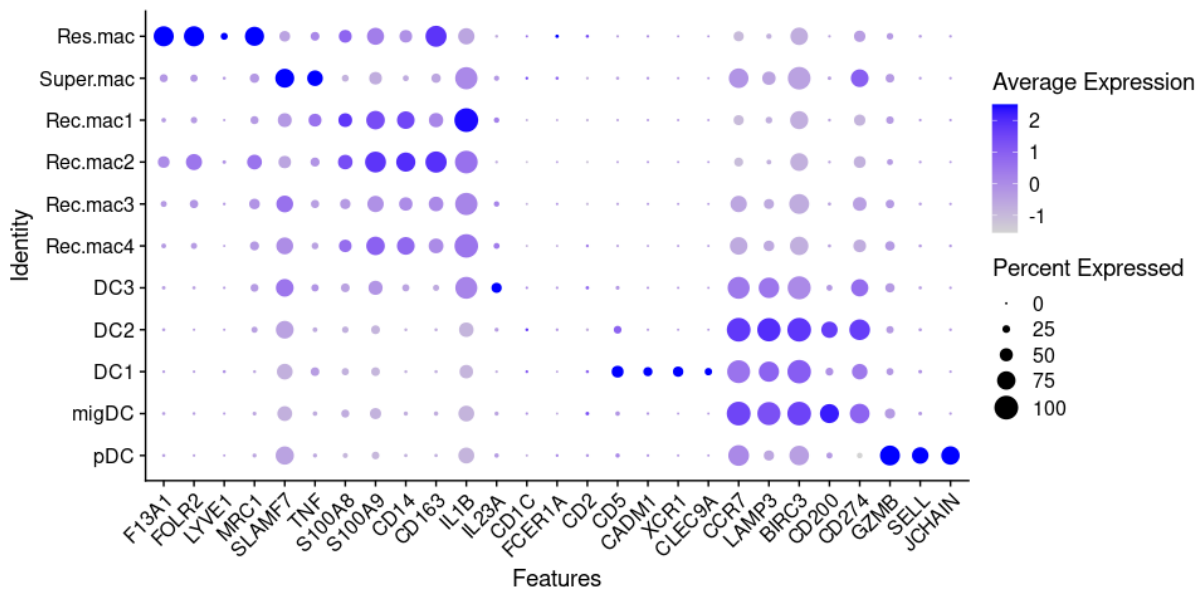
Figure 4.7 Dot plot showing the expression of marker genes of each dermal myeloid cell population defined in Figure 4.6. The size of each dot corresponds to the percentage of cells within in the cluster expressing the gene. The intensity of the dot corresponds to the average level of expression of the gene across all cells within the cluster.

Table 4.3

Cluster	Marker genes
Resident macrophage	*F13A1, FOLR2, LYVE1, MRC1*
Super-activated macrophage	*SLAMF7, TNF*
Recruited macrophage	*S100A8, S100A9, CD14, CD163, IL1B*
DC3	*CD14, CD163, IL1B, IL23A*
DC2	*CD1C, FCER1A, CD2, CD5*
DC1	*CADM1, XCR1, CLEC9A*
migDC	*CCR7, LAMP3, BIRC3, CD200, CD274*
pDC	*GZMB, SELL, JCHAIN*

Table 4.3 List of representative marker genes used in Figure 4.7 for cluster identification.

66

Differentially expressed genes of each macrophage cluster were computed to explore their functional profiles (Tables 4.4 – 4.9). Resident macrophages expressed tissue residency markers *F13A1, FOLR2, LYVE1 and MRC1* (Table 4.4, Figure 4.7). This population also expressed genes with immune functions, which were upregulated along with other trophic-related genes in GVHD (Figure 4.8). Recruited macrophage 2 was related to resident macrophages, expressing immunomodulatory gene *TGFBI* and lower levels of resident markers (Table 4.5, Figure 4.7).

There were also four other inflammatory macrophage populations. Super-activated macrophages expressed *SLAMF7* and *TNF* which were capable of amplifying inflammation through an autocrine signalling loop (Table 4.6, Figure 4.7) (Simmons et al., 2022). Recruited macrophage 1 expressed genes related to inflammation and immune cell recruitment, such as *IL1B*, *CCL3* and *CXCL3* (Table 4.7, Figure 4.7). Recruited macrophage 3 expressed *TXN, REL* and *HLA* (Table 4.8). Recruited macrophage 4 expressed genes of the metallothionein family including *MT1G*, *MT1E* and *MT1M* (Table 4.9).

Table 4.4

	p_val	avg_log2FC	pct.1	pct.2	p_val_adj
SELENOP	0	2.990123	0.689	0.136	0
F13A1	0	2.892736	0.827	0.221	0
FGL2	0	2.178674	0.778	0.229	0
MS4A6A	0	2.201383	0.758	0.164	0
CTSC	0	2.812796	0.845	0.347	0
CCL13	0	2.947776	0.735	0.102	0
HSPB1	2.40E-295	3.71268	0.882	0.49	6.86E-291
SCN9A	8.38E-289	0.832465	0.294	0.018	2.39E-284
FOLR2	4.45E-287	2.052239	0.83	0.349	1.27E-282
VSIG4	8.14E-281	1.097078	0.391	0.045	2.33E-276
CCL18	3.91E-249	3.1432	0.849	0.397	1.12E-244
HSP90AA1	7.59E-246	2.518584	0.969	0.891	2.17E-241
EMB	3.57E-245	1.341179	0.528	0.113	1.02E-240
PLTP	1.07E-241	2.247949	0.9	0.544	3.05E-237
SLC40A1	7.99E-234	1.359103	0.293	0.028	2.28E-229
LGMN	3.97E-231	1.81525	0.853	0.469	1.13E-226
CST3	4.78E-230	2.055583	0.994	0.932	1.37E-225
HSPA6	2.16E-220	3.923801	0.534	0.137	6.16E-216
MS4A4A	5.71E-217	1.627404	0.653	0.22	1.63E-212
BLVRB	4.08E-213	1.567552	0.751	0.33	1.16E-208

Table 4.4 List of top differentially expressed genes of resident macrophages. The average fold change was calculated between this cluster and all other macrophage clusters, presented in log2 scale (avg_log2FC). pct.1 = percentage of cells in the cluster expressing the gene; pct.2 = percentage of cells in all the other clusters where the gene was detected; p_val = p value; p_val_adj = adjusted p value.

Figure 4.8
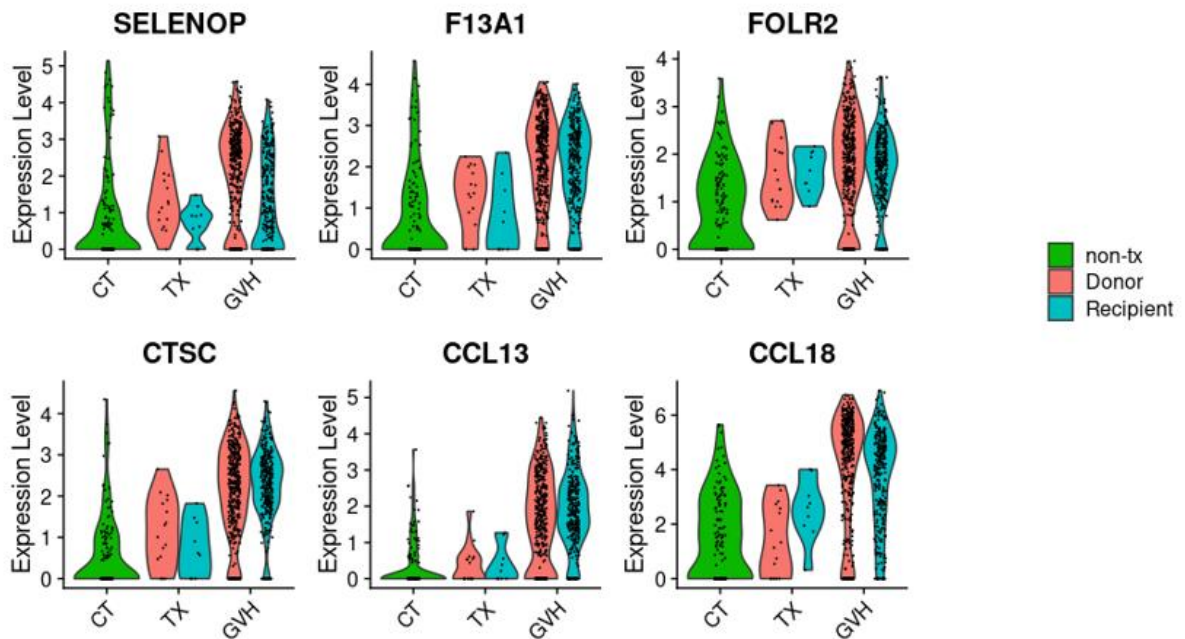
Figure 4.8 Violin plots comparing the expression of genes related to trophic and immune function by resident macrophages of each condition. Some of the top differentially expressed genes were shown. CT = healthy donor; TX = transplant control; GVH = graft versus host disease; non-tx = non-transplant cells (neither donor nor recipient).

Table 4.5

	p_val	avg_log2FC	pct.1	pct.2	p_val_adj
TGFBI	0	1.532967	0.888	0.558	0
FOLR2	0	1.086738	0.642	0.222	0
CD163	0	1.36669	0.881	0.562	0
RNASE1	0	1.778164	0.873	0.444	0
PLTP	0	1.43769	0.823	0.423	0
FTL	0	0.825596	1	1	0
LILRB5	2.60E-283	0.967825	0.496	0.122	7.41E-279
CST3	6.50E-264	1.054117	0.991	0.906	1.86E-259
CCL18	8.54E-264	2.727643	0.649	0.287	2.44E-259
NPC2	2.37E-237	0.698666	0.973	0.885	6.77E-233
PLD3	1.63E-231	0.955703	0.843	0.553	4.66E-227
CTSB	2.79E-208	0.760618	0.994	0.926	7.96E-204
F13A1	6.20E-200	0.860964	0.434	0.129	1.77E-195
LGMN	5.53E-186	0.908497	0.677	0.379	1.58E-181
ABCA1	3.63E-175	0.775967	0.931	0.81	1.04E-170
MSR1	2.81E-171	0.63536	0.364	0.098	8.03E-167
MARCO	3.64E-170	1.873525	0.552	0.267	1.04E-165
C1QC	2.57E-168	1.032668	0.873	0.619	7.33E-164
CTSL	7.02E-166	0.87751	0.997	0.96	2.00E-161
SELENOP	4.62E-165	0.73373	0.3	0.065	1.32E-160

Table 4.5 List of top differentially expressed genes of recruited macrophage 2. The average fold change was calculated between this cluster and all other macrophage clusters, presented in log2 scale (avg_log2FC). pct.1 = percentage of cells in the cluster expressing the gene; pct.2 = percentage of cells in all the other clusters where the gene was detected; p_val = p value; p_val_adj = adjusted p value.

Table 4.6

	p_val	avg_log2FC	pct.1	pct.2	p_val_adj
NEURL3	2.71E-201	0.9276585	0.379	0.034	7.74E-197
ABTB2	1.39E-158	0.6778048	0.496	0.079	3.98E-154
IRF8	2.47E-144	1.3350458	0.472	0.079	7.05E-140
RHOF	4.94E-129	1.1003507	0.694	0.203	1.41E-124
MIR155HG	7.25E-129	2.2957929	0.892	0.413	2.07E-124
TRAF4	2.69E-127	0.8907790	0.453	0.083	7.69E-123
DUSP2	1.55E-120	2.8027251	0.935	0.559	4.43E-116
SNHG15	1.92E-118	2.0579185	0.810	0.344	5.49E-114
GADD45B	2.05E-114	2.2147002	0.978	0.730	5.86E-110
TNFRSF9	1.99E-112	1.1021337	0.621	0.176	5.69E-108
STAT5A	4.23E-104	1.3101217	0.696	0.266	1.21E-99
H3F3B	1.48E-101	0.8844308	1.000	0.986	4.23E-97
CKB	2.96E-98	1.9505303	0.463	0.112	8.45E-94
REL	2.97E-98	1.4382950	0.986	0.841	8.50E-94
UBC	1.91E-97	1.0196759	0.997	0.994	5.46E-93
HES4	4.24E-97	1.0240382	0.520	0.136	1.21E-92
CD40	8.66E-97	1.7062784	0.859	0.514	2.47E-92
ZBTB10	1.05E-94	1.1507251	0.737	0.301	2.99E-90
OASL	2.29E-94	1.3742373	0.572	0.175	6.54E-90
TRAF1	4.50E-94	1.2562048	0.954	0.736	1.29E-89

Table 4.6 List of top differentially expressed genes of super-activated macrophage. The average fold change was calculated between this cluster and all other macrophage clusters, presented in log2 scale (avg_log2FC). pct.1 = percentage of cells in the cluster expressing the gene; pct.2 = percentage of cells in all the other clusters where the gene was detected; p_val = p value; p_val_adj = adjusted p value.

Table 4.7

	p_val	avg_log2FC	pct.1	pct.2	p_val_adj
IL1B	0	1.288701	0.999	0.945	0
CCL3	0	1.40199	0.994	0.896	0
CCL3L1	0	1.798061	0.867	0.523	0
IL1A	2.27E-307	1.883531	0.721	0.356	6.49E-303
CXCL3	9.25E-297	1.292046	0.997	0.972	2.64E-292
CXCL2	6.87E-281	1.166721	0.996	0.978	1.96E-276
CXCL8	9.06E-242	0.966219	1	0.995	2.59E-237
PTGS2	1.38E-236	0.962266	0.973	0.835	3.95E-232
CCL4	4.11E-208	1.240022	0.917	0.769	1.17E-203
PLAUR	1.80E-204	0.619135	1	0.991	5.14E-200
CCL20	3.37E-203	1.604292	0.805	0.578	9.64E-199
SERPINB2	4.10E-181	2.093183	0.747	0.519	1.17E-176
CLEC5A	1.74E-172	0.901999	0.296	0.061	4.96E-168
CXCL1	8.43E-167	1.219179	0.945	0.847	2.41E-162
CXCL5	2.75E-166	1.527325	0.866	0.705	7.86E-162
F3	1.05E-161	1.515828	0.496	0.225	3.01E-157
IL6	1.81E-156	1.165096	0.824	0.647	5.16E-152
CCL4L2	8.13E-152	1.638099	0.61	0.348	2.32E-147
RAC2	5.86E-145	0.722213	0.602	0.346	1.67E-140
INHBA	1.63E-127	0.826064	0.903	0.759	4.66E-123

Table 4.7 List of top differentially expressed genes of recruited macrophage 1. The average fold change was calculated between this cluster and all other macrophage clusters, presented in log2 scale (avg_log2FC). pct.1 = percentage of cells in the cluster expressing the gene; pct.2 = percentage of cells in all the other clusters where the gene was detected; p_val = p value; p_val_adj = adjusted p value.

Table 4.8

	p_val	avg_log2FC	pct.1	pct.2	p_val_adj
CCR7	7.15E-156	1.483876	0.655	0.366	2.04E-151
REL	1.42E-107	0.698479	0.925	0.818	4.06E-103
TXN	3.69E-104	0.639756	0.999	0.993	1.06E-99
PNRC1	9.44E-98	0.407199	1	0.996	2.69E-93
HLA-A	1.18E-97	0.478884	0.998	0.994	3.38E-93
BTG1	3.30E-90	0.480143	0.996	0.984	9.43E-86
IDO1	1.08E-83	1.031813	0.686	0.506	3.09E-79
EEF1A1	5.58E-81	0.357205	1	0.999	1.59E-76
CXCR4	7.53E-81	0.689622	0.655	0.438	2.15E-76
LITAF	1.24E-80	0.58812	0.92	0.845	3.53E-76
HLA-DPB1	6.28E-75	0.499662	0.968	0.908	1.79E-70
WARS	1.28E-74	0.710955	0.686	0.51	3.66E-70
BNIP3L	1.42E-74	0.695665	0.95	0.899	4.05E-70
TTYH2	2.96E-72	0.33195	0.273	0.104	8.45E-68
ISG20	2.70E-70	0.594396	0.775	0.614	7.70E-66
FAM49A	1.81E-69	0.438692	0.64	0.444	5.16E-65
PTPN1	3.30E-68	0.482118	0.863	0.781	9.42E-64
TSPAN33	1.12E-66	0.472407	0.482	0.281	3.20E-62
PPA1	3.81E-63	0.572053	0.508	0.322	1.09E-58
TNFSF13B	2.51E-62	0.61851	0.551	0.354	7.17E-58

Table 4.8 List of top differentially expressed genes of recruited macrophage 3. The average fold change was calculated between this cluster and all other macrophage clusters, presented in log2 scale (avg_log2FC). pct.1 = percentage of cells in the cluster expressing the gene; pct.2 = percentage of cells in all the other clusters where the gene was detected; p_val = p value; p_val_adj = adjusted p value.

Table 4.9

	p_val	avg_log2FC	pct.1	pct.2	p_val_adj
MT1G	2.98E-157	2.592289	0.969	0.588	8.52E-153
MT1E	6.44E-154	2.072155	0.978	0.641	1.84E-149
MT1M	3.56E-150	2.317523	0.909	0.442	1.02E-145
MT1H	1.32E-137	2.715028	0.668	0.19	3.76E-133
MT1X	7.26E-132	1.843063	0.995	0.764	2.07E-127
MT1F	1.53E-112	1.649965	0.933	0.615	4.36E-108
MT2A	5.61E-102	1.226545	0.998	0.978	1.60E-97
MT1A	2.65E-86	1.799881	0.599	0.208	7.57E-82
TMSB10	2.74E-45	0.651378	1	0.993	7.81E-41
CCR7	1.93E-41	0.831637	0.709	0.418	5.50E-37
SLC1A2	3.82E-40	0.576578	0.286	0.089	1.09E-35
ADAM19	6.16E-39	0.800428	0.565	0.284	1.76E-34
C15orf48	9.20E-39	0.562638	0.995	0.987	2.63E-34
ALDH2	7.50E-38	0.604069	0.555	0.295	2.14E-33
IDO1	2.64E-37	0.930558	0.762	0.536	7.53E-33
SLC39A8	5.33E-34	0.531554	0.969	0.845	1.52E-29
GPX4	1.42E-33	0.457039	0.962	0.915	4.05E-29
TMSB4X	5.11E-33	0.628875	0.995	0.985	1.46E-28
GRINA	3.70E-31	0.503073	0.93	0.898	1.06E-26
B2M	4.76E-31	0.358509	1	1	1.36E-26

Table 4.9 List of top differentially expressed genes of recruited macrophage 4. The average fold change was calculated between this cluster and all other macrophage clusters, presented in log2 scale (avg_log2FC). pct.1 = percentage of cells in the cluster expressing the gene; pct.2 = percentage of cells in all the other clusters where the gene was detected; p_val = p value; p_val_adj = adjusted p value.

DCs included a pDC population expressing *GZMB*, *SELL* and *JCHAIN*, a DC1 population expressing *CADM1*, *XCR1* and *CLEC9A*, a DC2 population expressing *CD5* and lack DC1 markers, and a DC3 population expressing *IL1B*, *IL23A* along with some *CD14* and *CD163* expression (Nakamizo *et al.*, 2021). There was also a GVHD-specific DC cluster expressing high level of genes related to maturation and immunoregulation (*CCR7*, *LAMP3*, *BIRC3*, *CD200*, *CD274*) which featured migDC.

The cellular composition of myeloid subpopulations between GVHD-affect dermis and controls was then compared (Figure 4.9). A chi-square test was performed and there was a significant association between the myeloid subpopulations and the condition of the samples $\chi2$ (110) = 9451, p < .001 (Appendix table 3). There was an enrichment of most myeloid subpopulations (except DC1, Rec.mac2, super.mac and Resi.mac) in most GVHD-affect dermis samples compared to healthy and transplant controls. This phenomenon was more pronounced in GVHD-affected samples collected at an early timepoint post transplantation. The proportion of resident macrophage in sample GVH3 was the highest and showed a significant difference compared to all other samples.

Figure 4.9
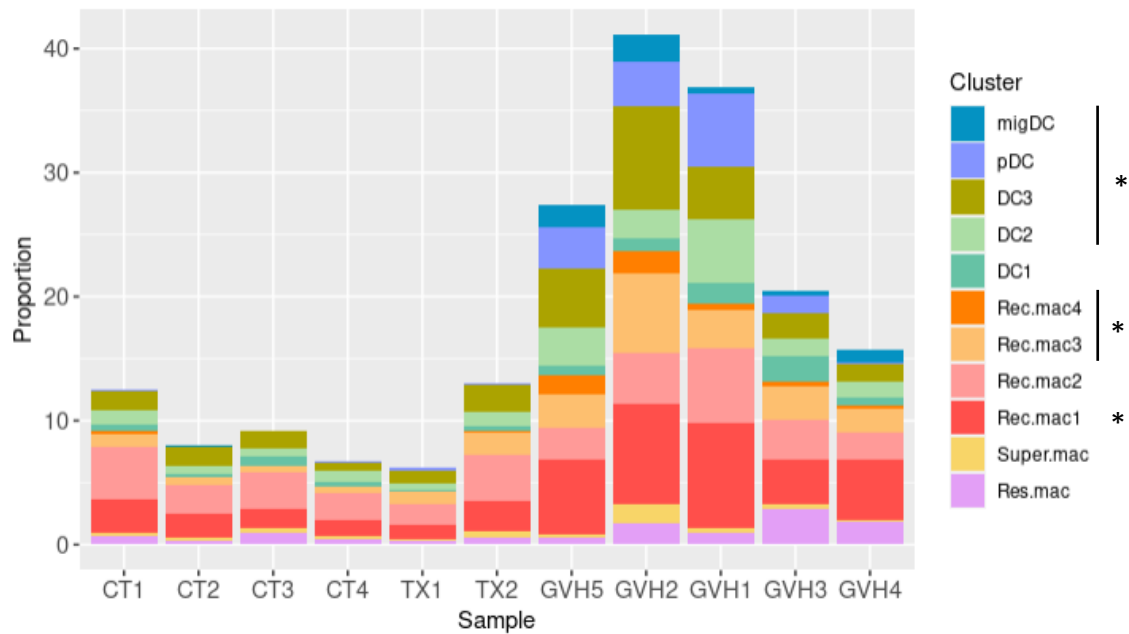
Figure 4.9 Bar chart showing the proportion of each dermal myeloid cell subset by sample, normalized to the total number of cells of individual samples. Sample from patients affected by GVHD were ordered by the timeframe between HSCT and sample being taken, with GVH5 (left) being the sample taken at the earliest timepoint. (* $p < 0.05$ GVH samples compared to controls). CT = healthy donor; TX = transplant control; GVH = graft versus host disease.

4.2.5 Validation of myeloid subset annotations

To cross-validate the annotations from the previous section, the dermal myeloid cells were constructed into a reference, and cells from other published human skin single-cell datasets (Nakamizo *et al.,* 2021; Reynolds *et al.,* 2021) were projected onto the reference dataset (Figure 4.10). The majority of cells from the resident macrophage clusters "Macro_1" and "Macro_2" from Reynolds et al., and "MARCO" from Nakamizo et al. were projected onto the macrophage clusters expressing residency markers, whereas the monocyte cluster "Mono" from Reynolds et al. and inflammatory macrophage cluster "CCR1" from Nakamizo et al. were projected onto the macrophage cluster with inflammatory properties (Recruited macrophage 1). Regarding the DC aspect, the DC annotations across the DC3, DC2, DC1 and migDC spectra in the reference dataset were aligned to inflammatory monocytes "inf_mono", monocyte-derived DC clusters "moDC1", "moDC2" and "moDC3" and migratory DCs "MigDC" from Reynolds et al., and to DC3, DC2, DC1 and mregDC from Nakamizo et al.. A cluster of DC (DC2) from Reynolds et al. and a cluster of DC3 (DC3_3) from Nakamizo et al. were projected onto the resident macrophage cluster. The DC1 clusters from Reynolds et al. and Nakamizo et al. were projected onto the pDC cluster and DC1 cluster, respectively.

Figure 4.10
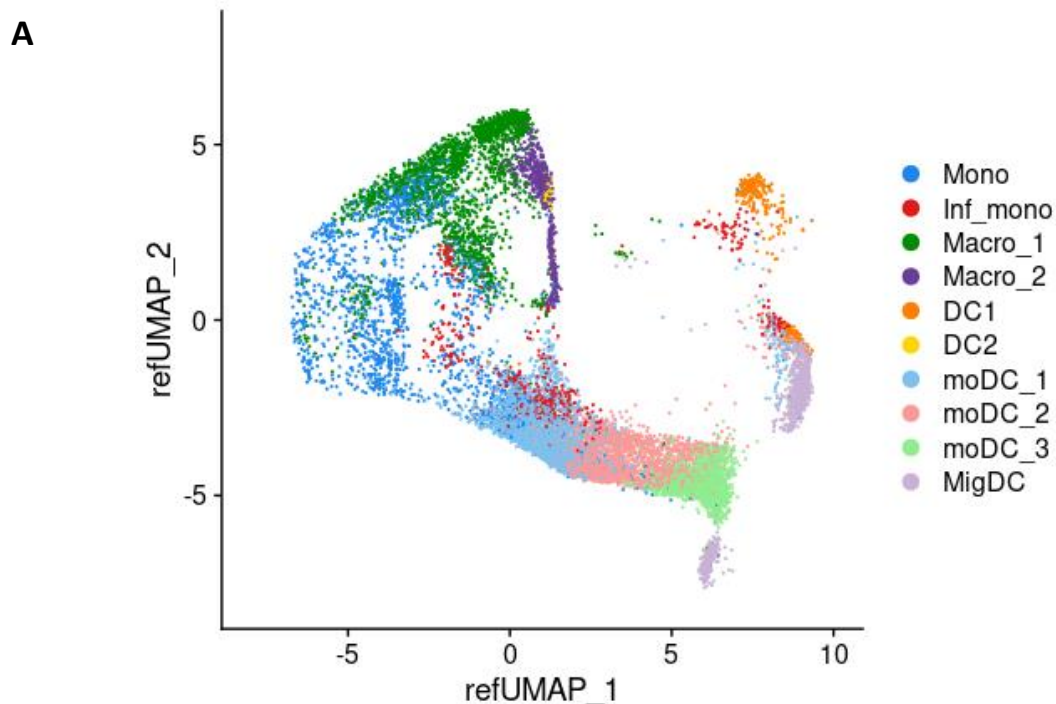
(Figure continued next page)
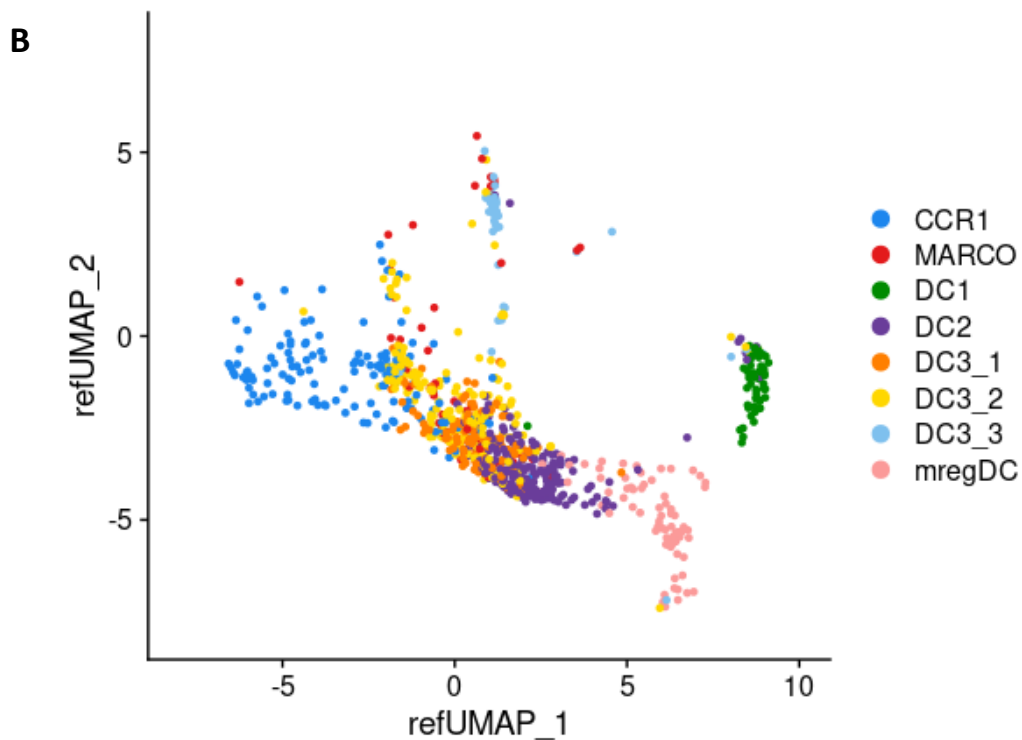
Figure 4.10 Myeloid cell datasets from Reynolds *et al.* and Nakamizo *et al.* (query) projected onto the dermal myeloid cells from Figure 4.6 (reference). Annotations from the authors of the dataset were displayed. (A) Dataset from Reynolds *et al.,* 2021. APCs from the skin of healthy controls were shown. Cells with predicted score lower than 0.7, and cell types with fewer than 10 cells, were removed. (B) Dataset from Nakamizo *et al.,* 2021. Myeloid cells from lesional and non-lesional skin were shown. Cells with predicted score lower than 0.6, cell types with fewer than 20 cells, and dying cells, were removed. Mono = monocyte; Inf_mono = inflammatory monocyte; Macro = macrophage; DC = dendritic cell; moDC = monocyte-derived dendritic cell; MigDC = migratory dendritic cell; CCR1 = CCR1+ macrophage, MARCO = MARCO+ macrophage; mregDC = mature DCs enriched in immunoregulatory molecules.

It is also of interest that whether the genes that distinguish between DC2 and DC3 in the blood (Villani *et al.*, 2017) were useful in classifying DC2 and DC3 in the skin (Figure 4.11). Although there wasn't a distinct separation between DC2 and DC3, a gradient of module scores transitioning between DC3 to DC2 could be seen. There was a significant difference in the means of module scores between DC2 and DC3 for their respective gene sets (Figure 4.12). It was also evident that migDC had lower DC2 signatures and macrophages had DC3 signatures.

Figure 4.11
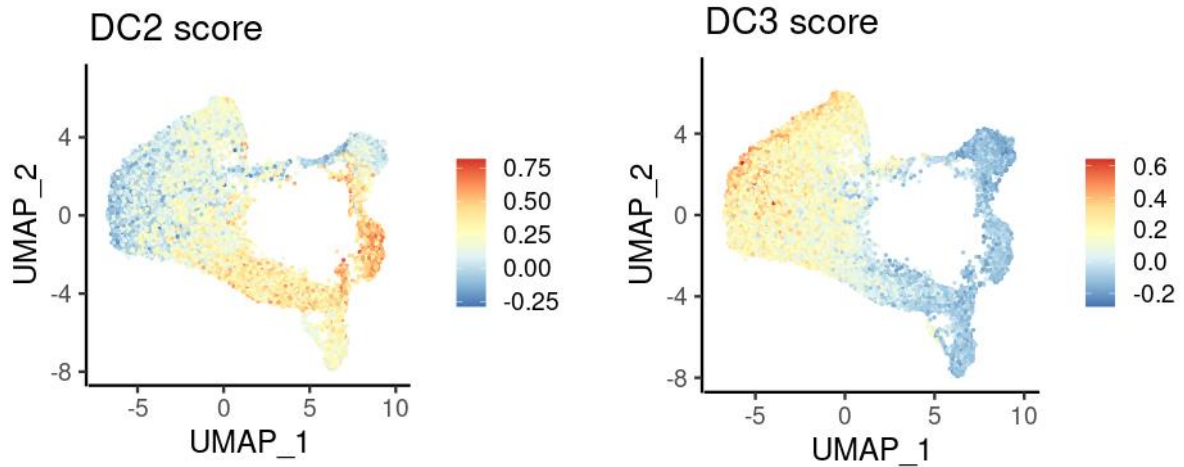
Figure 4.11 Scoring of myeloid cells by DC2 and DC3 gene modules. Modules were constructed based on the genes in Villani *et al.,* 2017 - supplementary table 3 (CD1C subsets).

Figure 4.12
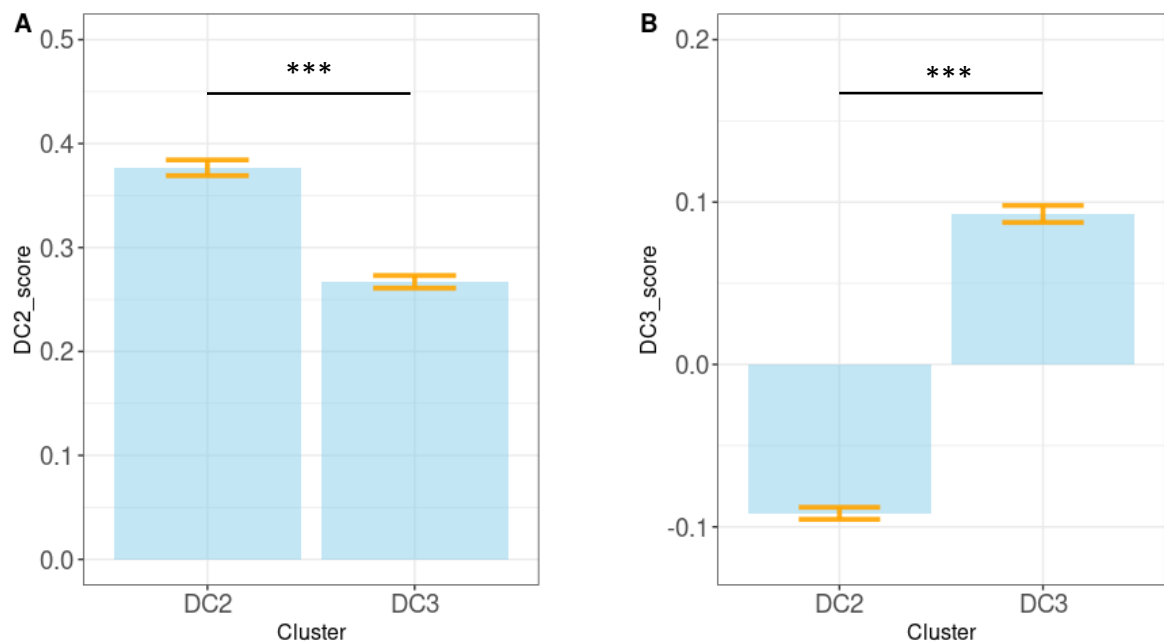
Figure 4.12 Bar charts comparing the mean score of each module between dermal DC2 and DC3. Welch two sample t-test was used due to unequal variance between the comparing groups. (A) DC2 module score: DC2 (M = 0.377, SD = 0.126) and DC3 (M = 0.267, SD = 0.142); t(2508) = 22.7, p < 0.001; (B) DC3 module score: DC2 (M = -0.0917, SD = 0.0624) and DC3 (M = 0.0927, SD = 0.123); t(3330) = -57.3, p = < 0.001.

4.3 Chapter discussion

A common feature of cutaneous GVHD is the infiltration of the skin by immune cells. Through scRNA-seq, I characterized the stromal and immune cell populations in GVHD-affected dermis. Results demonstrated that while there was an enrichment of macrophages in the dermis of patients with GVHD, the enrichment of T cells was only observed in samples collected at a later timepoint. This finding corroborated previous results obtained by flow cytometry (Jardine *et al.*, 2020). However, scRNA-seq results from this study showed that myeloid infiltrates included both macrophages and cDCs, in contrast to the flow cytometry findings in which macrophage was the only myeloid cell type that was enriched. This discrepancy was likely due to variations in experimental conditions and patient cohorts. The roles of T cells were well-established in GVHD, yet patients with an early GVHD onset were often lymphopenic and immunosuppressed. The enrichment of myeloid cells in these patients highlighted their potential involvement in GVHD pathology. Surprisingly, Tregs were enriched in the GVH samples collected at later timepoint. Given the protective roles of Tregs against GVHD, these cells might be inactivated or dysregulated in patients affected by GVHD.

Additionally, pDCs and NK cells were also present in the dermis of patients with GVHD, with the former enriched in the samples collected at an early timepoint. These cells are not present in the skin under physiological conditions and are even thought to have protective roles against GVHD (Olson *et al.*, 2010; Sheng *et al.*, 2020; Tian *et al.*, 2021). Nonetheless, the pathogenic roles of pDCs and NK cells were also being described in GVHD and other skin conditions such as psoriasis due to their capability of producing the relevant inflammatory and cytotoxic mediators (Cooley *et al.*, 2005; Nestle *et al.*, 2005). Further investigation of these cells is required to discern their roles in GVHD. On the other hand, donor-derived LCs were present in the dermis of one of the patients with GVHD (GVH3). LCs normally reside in the epidermis, and their appearance in the dermis may be narrating the process of carrying antigen from the epidermis towards the draining lymph node for antigen presentation through the lymphatics. The role of LCs in GVHD has been described as a type of host-derived APC that triggers the differentiation of pathogenic effector T cells in situ (Santos E Sousa *et al.*, 2018). The observation in the current study reveals a potential mechanism of donor-derived LCs in mediating cutaneous GVHD.

Meanwhile, donor monocyte-derived macrophages have been suggested to be potential mediators of GVHD pathology due to their capability of activating T cells and exert direct cytotoxic effect (Jardine *et al.*, 2020). In the current study, dermal macrophages populations were classified into various "resident" and "recruited" populations. Resident macrophages express tissue residency markers *F13A1*, *FOLR2*, and also *LYVE1* which suggests they may be derived from embryonic precursors (Chakarov *et al.*, 2019; Mulder *et al.*, 2021). This population of macrophages self-propagate independently from circulating monocytes and HSC progenitors (Hashimoto *et al.*, 2013). Current data showed that in transplant settings the resident macrophage population consisted of cells from both the donor and recipient, implying the contribution of bone marrow-derived progenitors to this population of cells (Haniffa *et al.*, 2009). Lymphocyte chemotactic factors such as *CCL13* and *CCL18* were upregulated by resident macrophages affected by GVHD compared to controls, which suggest they may play a potential role in mediating GVHD. However, it is uncertain whether these effects exerted by resident macrophages are protective or exacerbates GVHD. Another closely related macrophage population, which I termed "Recruited macrophage 2", express *TGFBI* which codes for an extracellular matrix protein, and lower levels of resident macrophage markers. This population consisted mostly of donor derived cells and possibly represents an intermediate state between resident macrophage and its progenitor. The proportion of these two populations of macrophage were not changed in GVHD comparing to healthy and transplant controls, potentially implying that the process of maintaining the resident macrophage population was not affected by GVHD. However, further investigation with larger sample size is required to elucidate the relationship between resident macrophages and GVHD.

Conversely, other donor-derived, pro-inflammatory "recruited" macrophage populations were also identified. "Recruited macrophage 1" was the most abundant population and highly expresses inflammatory mediators. "Recruited macrophage 3" expressed *REL* and may drive immune and inflammation processes via NF-KB pathway. "Recruited macrophage 4" expressed genes of the metallothionein family. Metallothioneins are small metal-binding proteins and play an important role in maintaining metal homeostasis and responding to cellular stress. These proteins also emerged to be involved in various immune responses. It has been documented that the expression of metallothioneins were upregulated by GM-CSF activated macrophages and were required for optimal antimicrobial effector functions (Subramanian Vignesh *et al.*, 2013). Nonetheless, metallothioneins also possess various immunomodulatory

properties (Dai *et al.*, 2021). Lastly, "super-activated" macrophages were capable of amplifying inflammation through an autocrine signalling loop (Simmons *et al.*, 2022). However, this population was not enriched in GVHD and therefore likely to be irrelevant in this context.

In general, these recruited macrophage populations were enriched in the dermis of patients with GVHD. They also expressed monocyte-associated antigens (*S100A8, S100A9*), consistent with recent emigration of monocytes from the blood (Reinhardt *et al.*, 2014). Since there was no evidence of dermal macrophages expressing proliferative markers, it is unlikely that the enrichment of dermal macrophages was due to local proliferation of pre-existing macrophages.

Regarding classical DCs, cDC2s (including DC2 and DC3) were enriched in the dermis of patients with GVHD. The inflammatory roles of DC3 have been described recently and these cells are found to be selectively expanded in various conditions such as autoimmune diseases and cancer (Dutertre *et al.*, 2019; Bourdely *et al.*, 2020; Nakamizo *et al.*, 2021). In the skin, DC3 can be distinguished from DC2 by the expression of *IL1B*, *IL23A*, and the lack of *CD5* expression (Nakamizo *et al.*, 2021). These cells also express monocyte/macrophage-related markers *CD14* and *CD163* but at a lower level comparing to dermal macrophages. Additionally, I cross-validated the cell type annotations by harnessing other publicly available scRNA-seq datasets (Nakamizo *et al.*, 2021; Reynolds *et al.*, 2021). It is worth mentioning that the nomenclature of monocyte-derived DC populations ("moDC") used by Reynolds *et al*. was based on the expression of monocyte-related traits by dendritic cells but not indicative of a monocyte origin. Besides, the projection of a small number of DCs onto the resident macrophage cluster was likely driven by the expression of genes related to stress (HSP genes). Nonetheless, DC3 constitutes part of the myeloid infiltrate and potentially contributes to GVHD pathophysiology.

Another subset of DC was only present in the dermis of patients with GVHD but not in transplant and healthy controls, which I termed migDC. migDCs express transcripts associated with DC maturation, immunoregulation and migration (Maier *et al.*, 2020; Nakamizo *et al.*, 2021; Reynolds *et al.*, 2021). The term "migDC" was used here instead of "mregDC" is due to the fact that the expression of immunoregulatory genes is integral to the maturation and migration process of DCs. I showed that this migDC programme is relevant in GVHD and it will be interesting to further explore its roles in GVHD. Interestingly, other DCs also expressed this migDC programme, which might explain the challenge encountered when defining the boundaries between DC subsets, as during DC maturation different types of DCs express shared activation and migration modules (Bosteels *et al.*, 2020). MigDCs also lacked DC1- and

DC2-specific markers detectable by scRNA-seq (Maier *et al.*, 2020). Therefore, experimental approaches incorporating protein expression profiling such as CITE-seq, or projection of data onto high-quality references, will be beneficial in such context.

In summary, T cells are likely not the sole effector cell population in GVHD, as other immune cells are also present and sometimes even more abundant. It is anticipated that GVHD pathogenesis is a balancing act of various pro-inflammatory and immunomodulatory cell populations of donor and host origin.

Chapter 5. Profiling the epidermal immune cells

5.1 Introduction

Apoptosis of keratinocytes and epithelial stem cells are features of cutaneous GVHD (Gilliam *et al.*, 1996; Zhan *et al.*, 2012). As the main haematopoietic APC in the epidermis, much focus has been on the potential role of LC in driving cutaneous GVHD pathology. Similar to macrophages, host LCs survive HSCT but reach complete donor chimerism at a higher rate (Collin *et al.*, 2006; Haniffa *et al.*, 2009; Mielcarek *et al.*, 2014). The contribution of LCs to GVHD pathology, however, remains obscure as preclinical models demonstrated conflicting results. In one study, cutaneous GVHD remined intact in the absence of either host or donor LCs (Li *et al.*, 2011). On the contrary, another group showed that host LCs trigger the differentiation of pathogenic effector T cells in situ (Santos E Sousa *et al.*, 2018). Donor LC may contribute to GVHD through promoting cutaneous inflammation (Mielcarek *et al.*, 2014).

The discovery of Trm has also shed new light to GVHD pathogenesis. GVHD has long been considered to be mediated entirely by donor T cells (Ferrara *et al.*, 2009). It is now known that host Trm survive HSCT and may potentially mediate GVHD (Divito *et al.*, 2020; Strobl *et al.*, 2020) (see section 1.3.4). This has led to the hypothesis that clinical low-grade acute GVHD is a local host versus graft reaction, while donor T cells are responsible for more severe GVHD (Jardine *et al.*, 2020).

Chapter aims:

1. Define the epidermal cell populations in transplant controls and patients affected by GVHD

2. Define the origin (donor/recipient) of each cell

3. Investigate the compositional changes of dermal cells between controls and patients affected by GVHD

4. Compare epidermal myeloid cells to their dermal counterparts

5. Compare the engraftment rate of myeloid cell populations

5.2 Results

5.2.1 Overview of epidermal cells from patients with GVHD and transplant controls

Lastly, the epidermis compartment was characterized. All epidermis samples from patients with GVHD, together with the samples obtained from transplant controls and healthy controls were integrated (Figure 5.1). Clusters were annotated according to the expression of marker genes of the major cell types (Figure 5.2; Table 5.1). In addition to the myeloid and lymphoid cells that were already described in the previous chapters, keratinocytes and melanocytes were also present in the epidermis. These cells could be identified by the expression of genes responsible for the synthesis of keratins (e.g., *KRT5*, *KRT14*) and melanosomes (e.g., *PMEL*, *MLANA*), respectively. A cluster of proliferative cells was also present, which consisted of CD4 T cells, CD8 T cells and NK cells from GVHD-affected samples, and keratinocytes exclusively from healthy and transplant control samples.

Figure 5.1
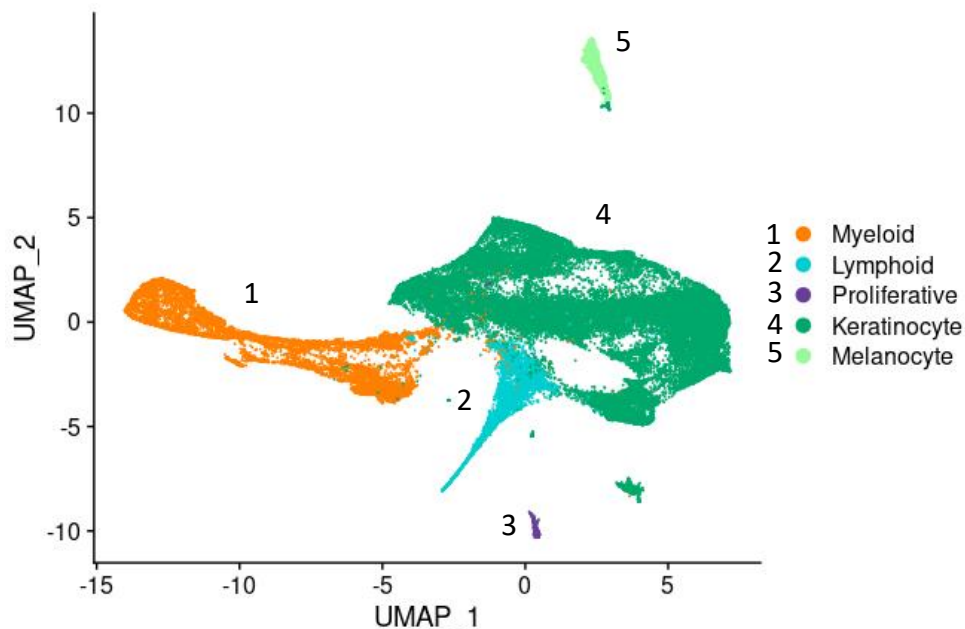
Figure 5.1 UMAP visualization of epidermal cells from all samples. Samples including healthy donors (*n* = 2), transplant controls (*n* = 2) and GVHD (*n* = 5). Clusters were manually annotated by inspecting expression of canonical marker genes.

Figure 5.2
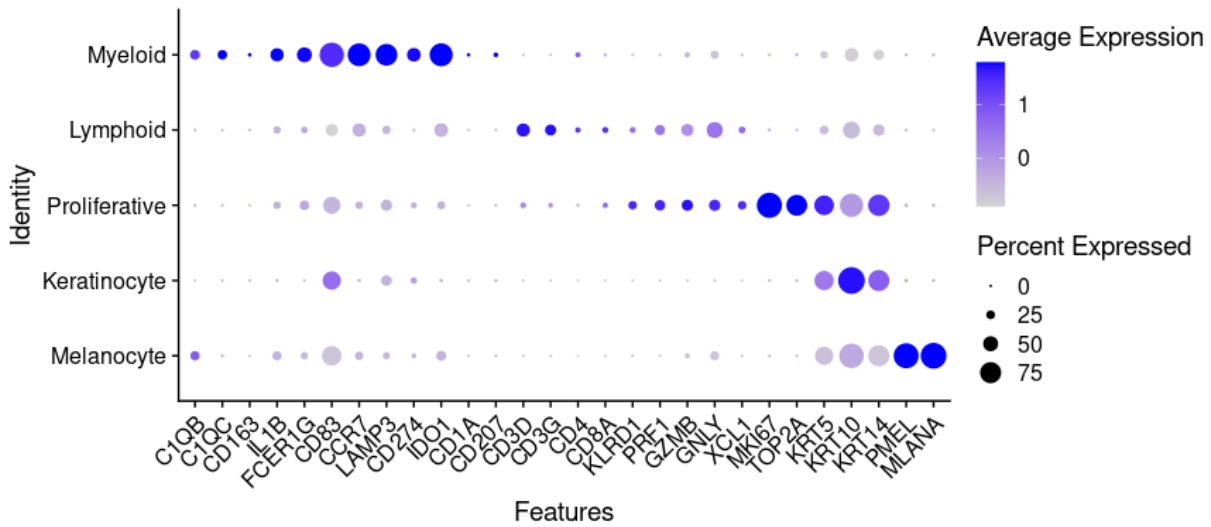
Figure 5.2 Dot plot showing the expression of marker genes of each epidermal cell population defined in Figure 5.1. The size of each dot corresponds to the percentage of cells within the cluster expressing the gene. The intensity of the dot corresponds to the average level of expression of the gene across all cells within the cluster.

Table 5.1

Cluster	Marker genes	
Myeloid	*C1QB, C1QC, CD163, IL1B,*	(Macrophage)
	FCER1G, CD83, CCR7, LAMP3, CD274, IDO1,	(DC)
	CD1A, CD207	(LC)
Lymphoid	*CD3D, CD3G, CD4, CD8A,*	(T)
	KLRD1, PRF1, GZMB, GNLY, XCL1	(NK)
Proliferative	*MKI67, TOP2A*	
Keratinocyte	*KRT5, KRT10, KRT14*	
Melanocyte	*PMEL, MLANA*	

Table 5.1 List of representative marker genes used in Figure 5.2 for cluster identification.

Most immune cells originated from GVHD-affected samples, while most stromal cells originated from healthy and transplant control samples (Figure 5.3). Sample GVH3 was an exception, of which the cellular composition was more similar to controls. Despite testing various integration parameters, a small number of keratinocytes from healthy control samples were still clustered together with lymphocytes, leading to a slight over-estimation of

lymphocyte proportion in both healthy control samples. Different subsets of myeloid and lymphoid cells were also not well separated. Therefore, analysing each sample individually or further clustering of immune cells would be beneficial and allow a more detailed analysis of the immune cell populations.

Figure 5.3
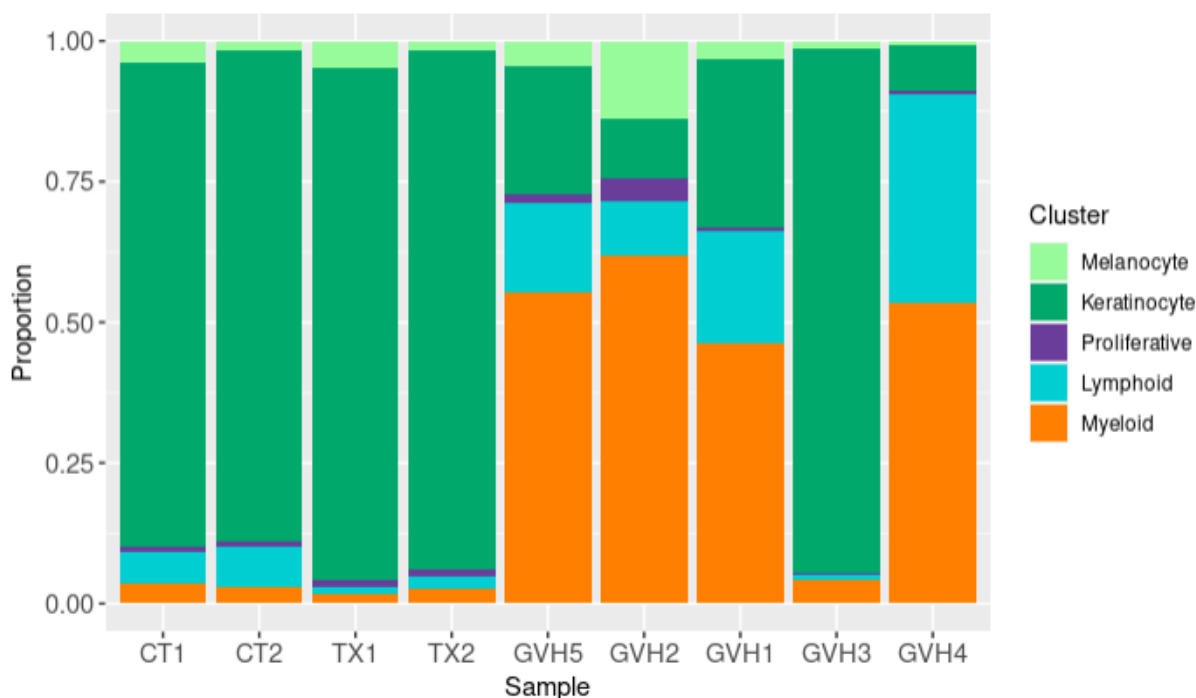
Figure 5.3 Bar chart showing the proportion of each epidermal cell population by sample. Sample from patients affected by GVHD were ordered by the timeframe between HSCT and sample being taken, with GVH5 (left) being the sample taken at the earliest timepoint. CT = healthy donor; TX = transplant control; GVH = graft versus host disease.

5.2.2 Individual epidermis samples

The epidermis sample of each patient was analysed separately. The samples from transplant controls were clustered (Figure 5.4, panel A) and the origin of each cell was shown (Figure 5.4, panel B). Little evidence of immune cell infiltration was seen. A small population of LC expressing *CD1A* and *CD207* was present, which was mostly donor derived. A small population of T cells was also present which consisted of both donor derived and recipient derived cells.

Figure 5.4
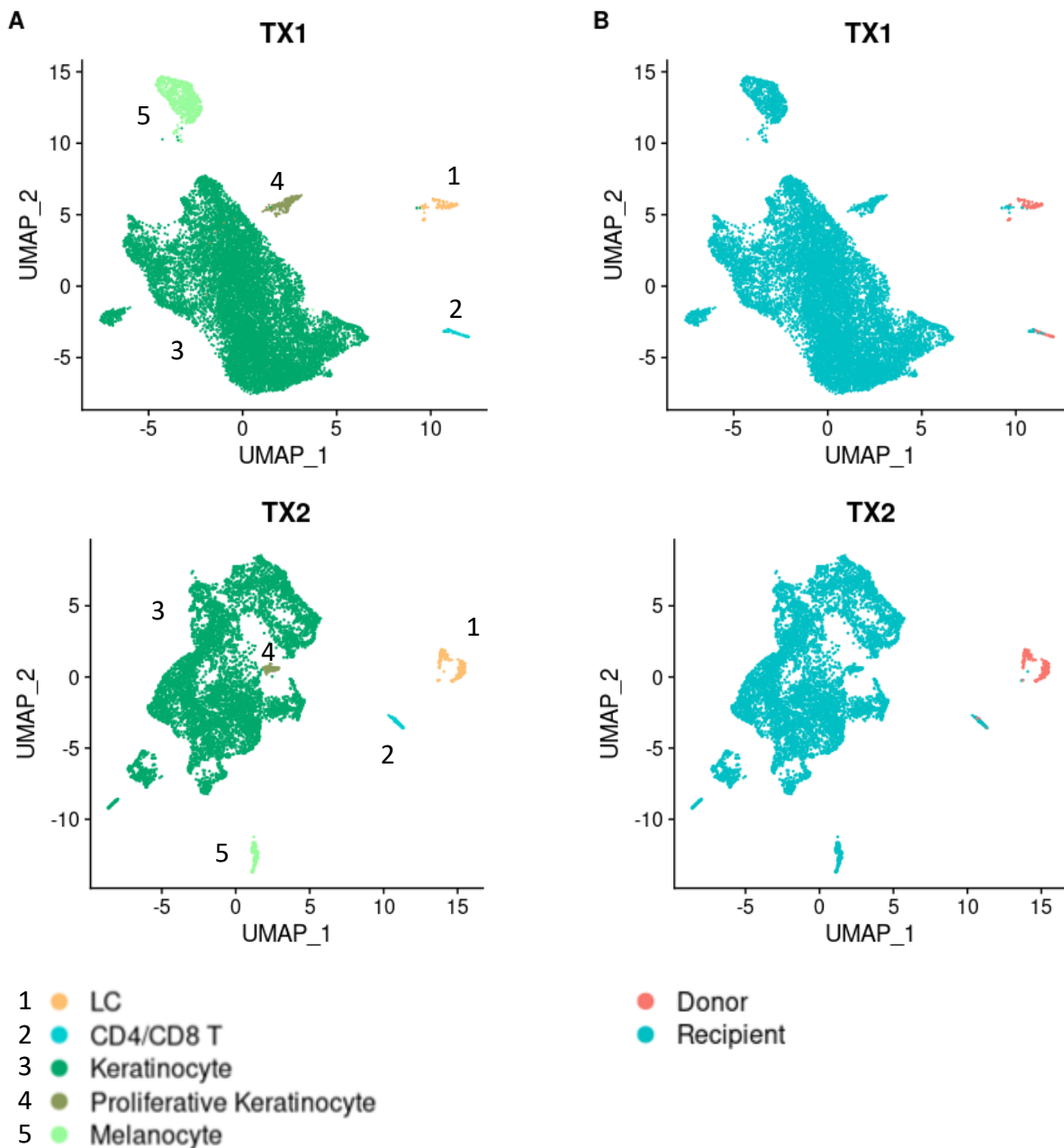
Figure 5.4 UMAP visualization of each epidermis sample from transplant controls. (A) Clusters were manually annotated by inspecting expression of canonical marker genes. (B) Each cell was annotated by its origin, either derived from the patient (recipient) or from the graft that the patient received (donor). LC = Langerhans cell.

The samples from patients affected by GVHD were then clustered (Figure 5.5, panel A) and the origin of each cell was shown (Figure 5.5, panel B). Samples were ordered from top to bottom by the timeframe between HSCT and the sample being taken. In general, there was a profound infiltration of donor-derived immune cells in GVHD-affected epidermis samples except for sample GVH3. Infiltrating myeloid cells included macrophages and DCs. Consistent with the dermis samples, pDCs could only be found in the epidermis samples collected at an early timepoint post transplantation (GVH5, GVH2 and GVH1). LCs could also be found to form a separate cluster in some samples (GVH5, GVH1, GVH3). This cluster contained mostly donor derived cells but also contained traces amount of recipient derived cells. There were no indications of the presence of resident macrophages.

Within the lymphoid compartment, infiltrates include CD4 T cells, CD8 T cells and NK cells. The majority of these cells were donor derived and some also express proliferative markers. Interestingly, a large population of recipient-derived T cells was present in the sample collected at the earliest timepoint (GVH5). Recipient T cells were also evident in sample GVH1 and GVH3. A cluster of Treg was present in the sample collected at the latest timepoint (GVH4) which contained cells derived from both the donor and the recipient.

Figure 5.5
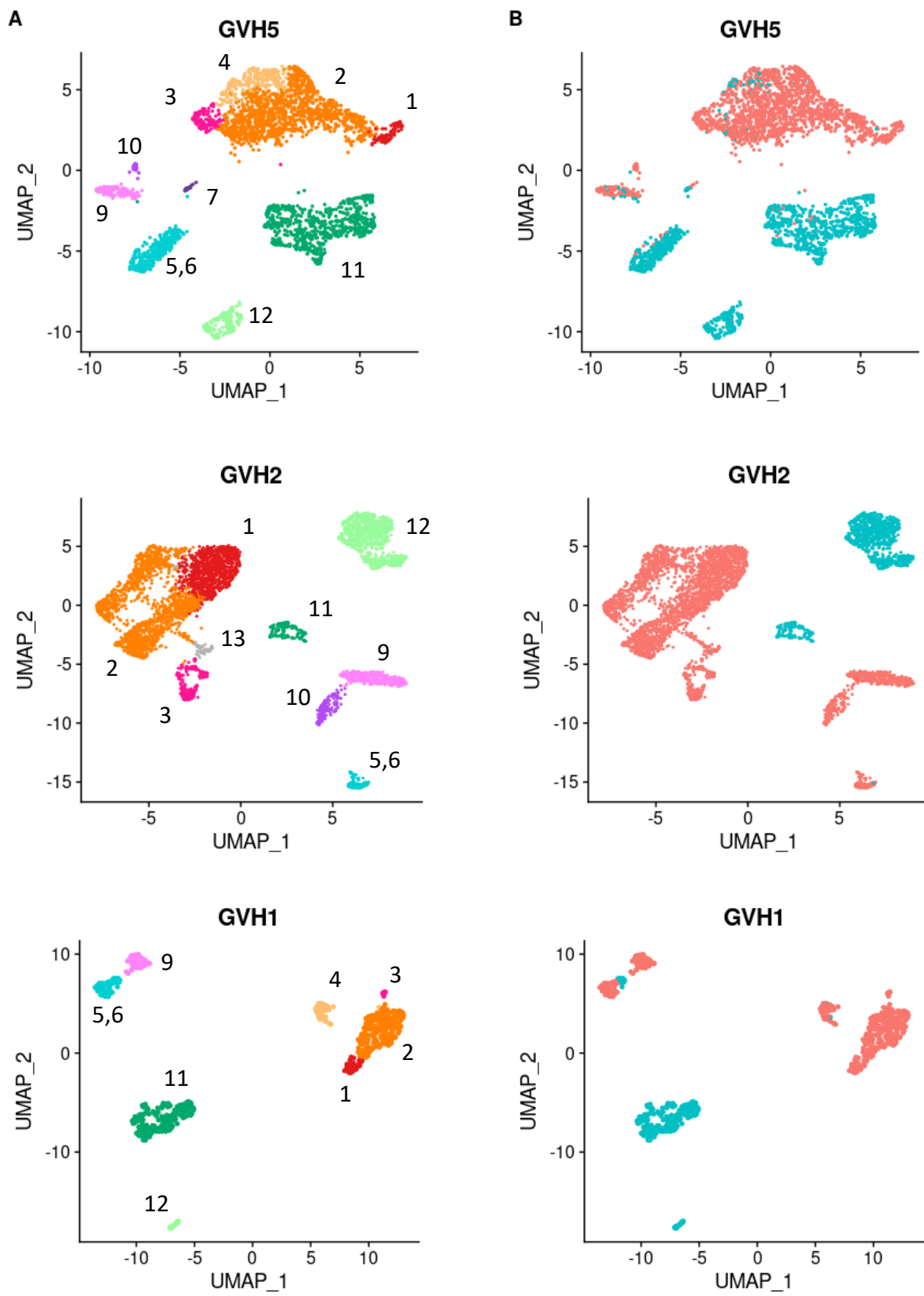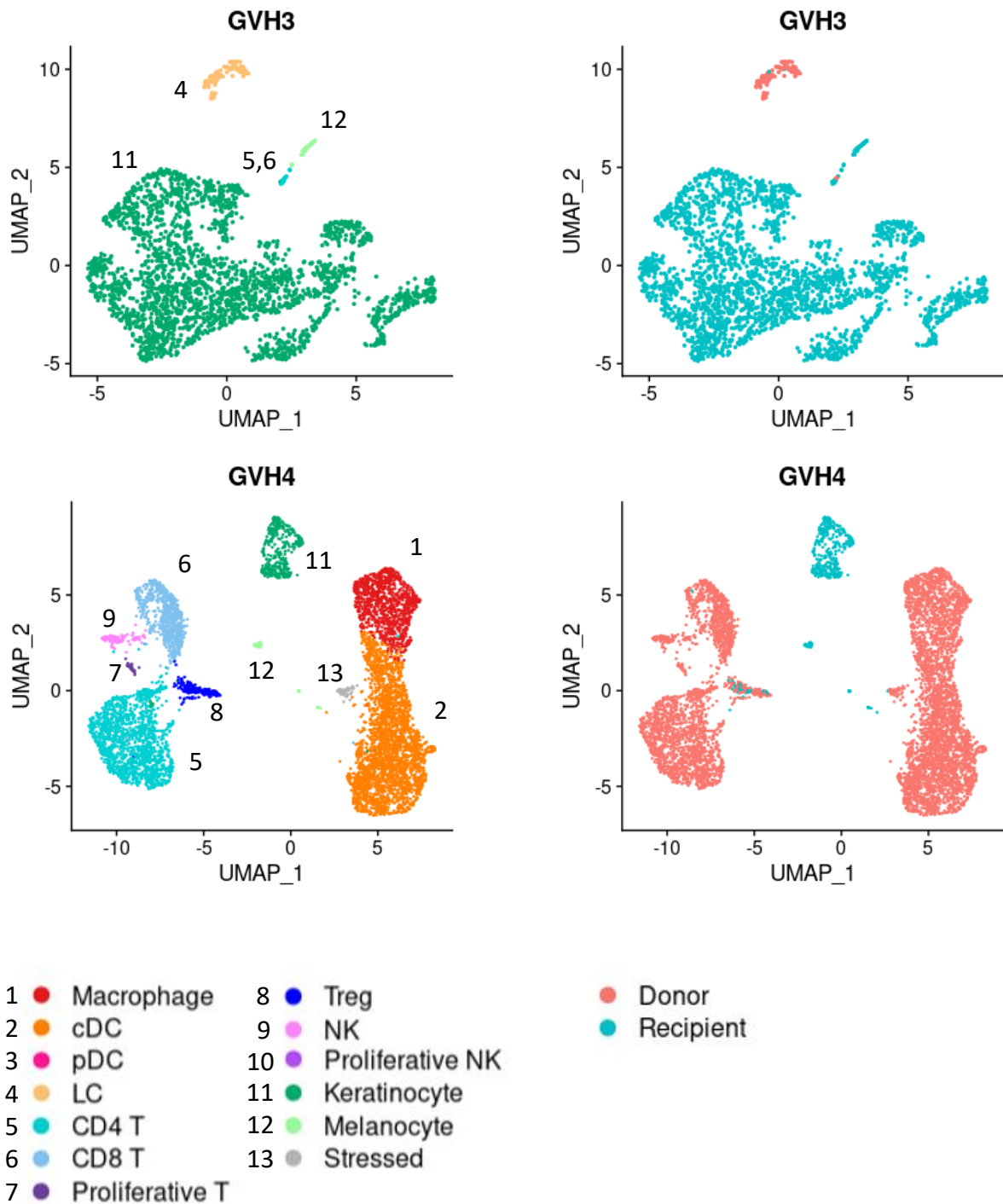
Figure 5.5 UMAP visualization of each epidermis sample from patients affected by GVHD. (A) Clusters were manually annotated by inspecting expression of canonical marker genes. Samples were ordered by the timeframe between HSCT and sample being taken, with GVH5 (top) being the sample taken at the earliest timepoint. (B) Each cell was annotated by its origin, either derived from the patient (recipient) or from the graft that the patient received (donor).

5.2.3 Characterization of epidermal immune cells

In order to analyse the immune cells at a higher resolution, immune cells from individual samples were subsetted and re-integrated (Figure 5.6, panel A). As the proportion of immune cells in control samples was low, most cells selected for re-integration were originated from GVHD-affected samples (Figure 5.6, panel B).

Figure 5.6
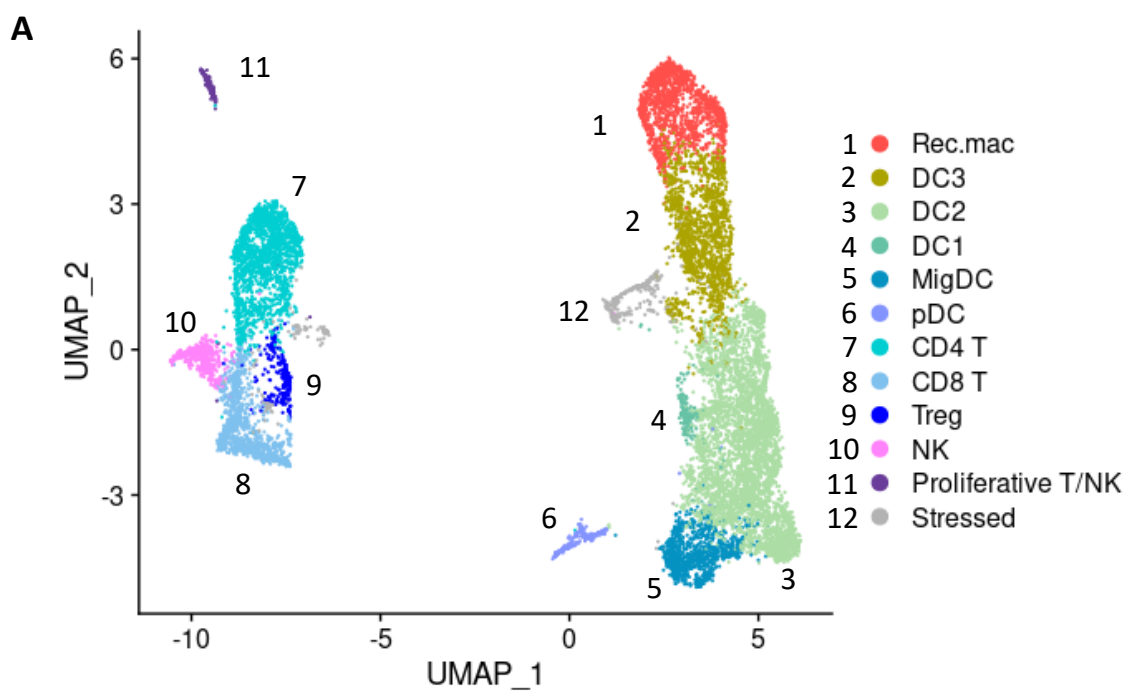
(Figure continued next page)
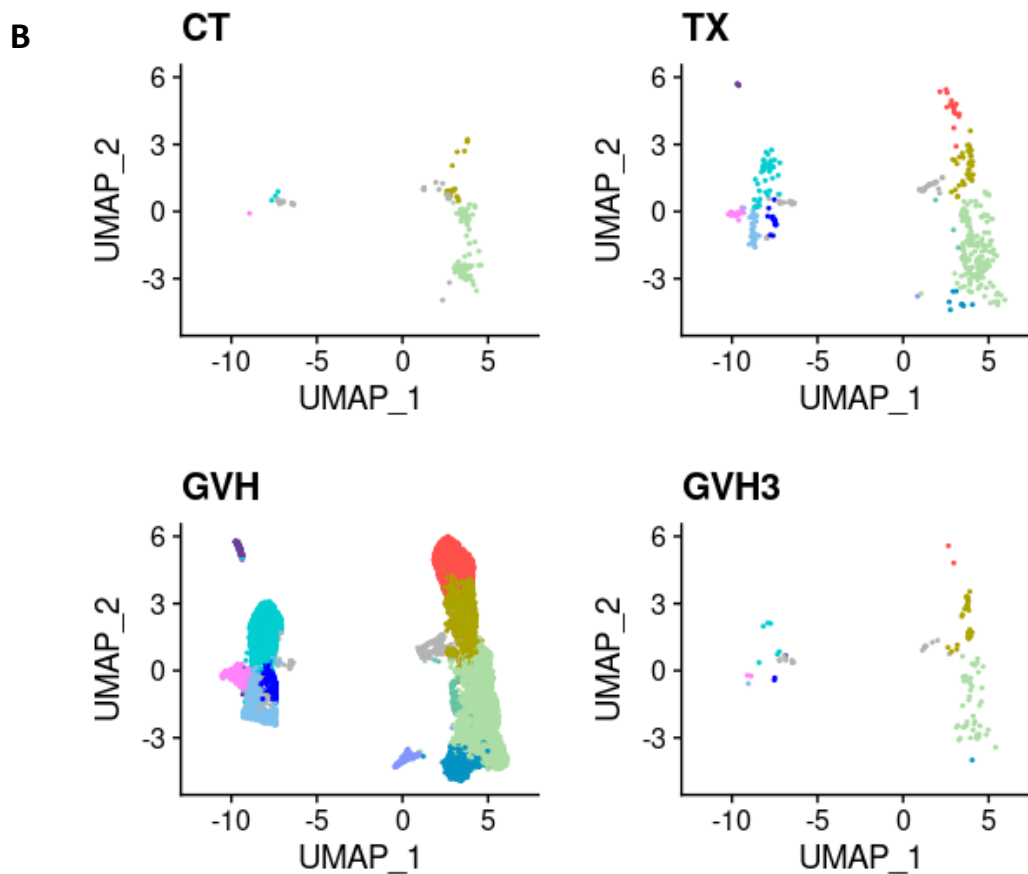
Figure 5.6 UMAP visualization of epidermal immune cells from all samples. Samples including healthy donors (*n* = 2), transplant controls (*n* = 2) and GVHD (*n* = 5). (A) Lymphoid cell clusters were manually annotated by inspecting expression of canonical marker genes. Myeloid cell clusters were annotated by inspecting expression of canonical marker genes and projecting onto dermal myeloid cells. (B) Contribution of cells from each condition was shown.

To annotate the clusters, the expression of canonical marker genes for each cluster was first investigated (Figure 5.7, panel A and B; genes were listed previously in Table 4.2 and 4.3). The identity of most clusters could be clearly determined. The macrophage cluster was being annotated as "recruited" instead of "resident" due to the expression of markers related to inflammation and lack expression of genes related to tissue residency. The distinction between DC2 and migDC, however, was not apparent as the expression of the DC2 marker genes was low. Furthermore, both clusters were expressing migDC signature genes at similar levels. LCs also did not form a separate cluster and mainly reside within the DC1 and DC2 clusters as inferred by the expression of marker genes *CD1A, CD207*.

Figure 5.7

A

B
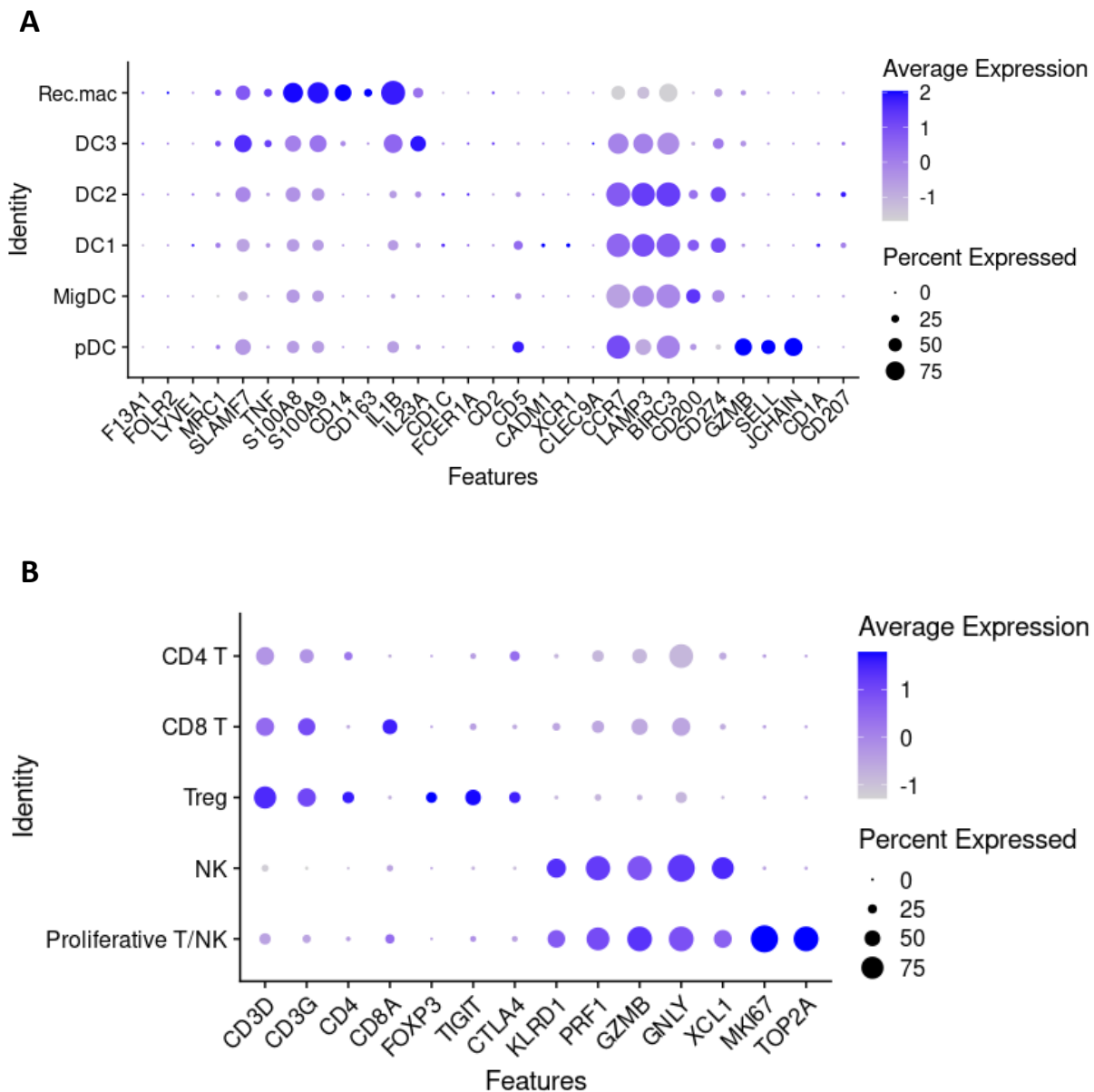
Figure 5.7 Dot plot showing the expression of marker genes of each epidermal immune population defined in Figure 5.6. (A) Myeloid populations, and (B) lymphoid populations. The size of each dot corresponds to the percentage of cells within in the cluster expressing the gene. The intensity of the dot corresponds to the average level of expression of the gene across all cells within the cluster.

To further classify the epidermal myeloid cells, the myeloid cells from Figure 5.6 were projected onto the dermal myeloid cell reference dataset from the previous chapter (Figure 5.8). The cells from the epidermal compartment were mapped to their corresponding dermal

counterparts. The recruited macrophages from the epidermis were projected almost exclusively to the "recruited macrophage 1" cluster of the dermis.

Figure 5.8
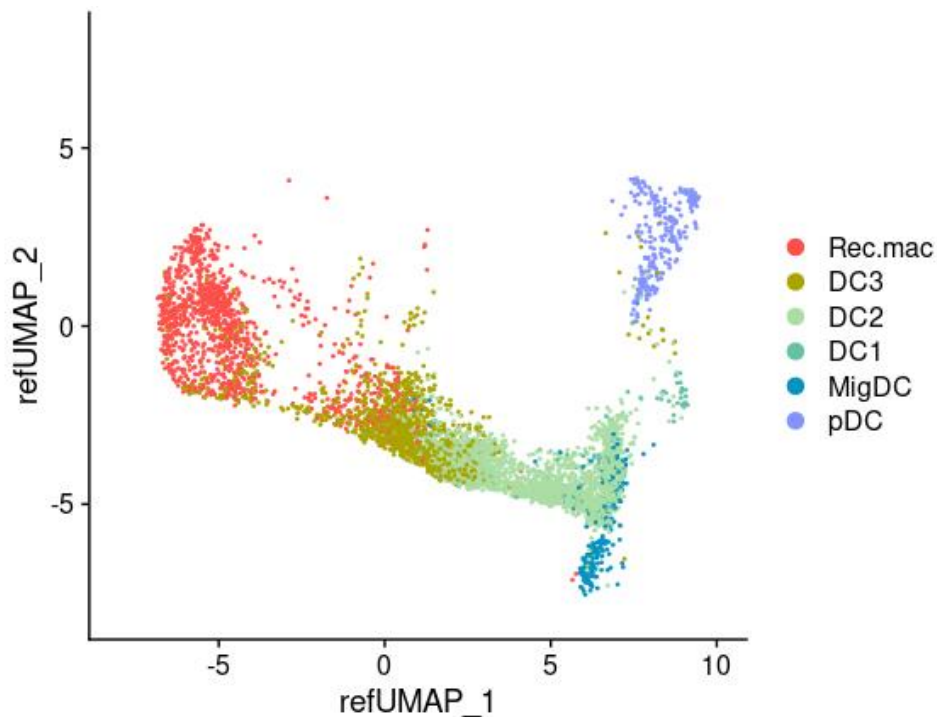
Figure 5.8 Epidermal myeloid cells (query) projected onto the dermal myeloid cells from Figure 4.6 (reference). Cells with predicted score lower than 0.7 were removed.

The proportion of myeloid and lymphoid cells were compared between the samples (Figure 5.9, panel A and B). There was a remarkable enrichment of immune cells in GVHD-affected epidermis compared to controls, except for sample GVH3. The majority of myeloid cell infiltrates were DC2 and DC3. Similar to the dermis, pDCs were only evident in early timepoint GVHD-affected epidermis, with the proportion being the highest in the sample obtained at the earliest timepoint. The involvement of migDC was also evident in GVHD-affected epidermis. Within the macrophage compartment, most cells were pro-inflammatory recruited macrophage 1.

Regarding lymphoid cells, the proportion of CD4 T cell to CD8 T cell in GVHD-affected epidermis was roughly similar. The presence of NK cell and Treg were also evident. In the sample collected at the latest timepoint (GVH4), the lymphoid compartment was dominated by T cells.

Figure 5.9

A
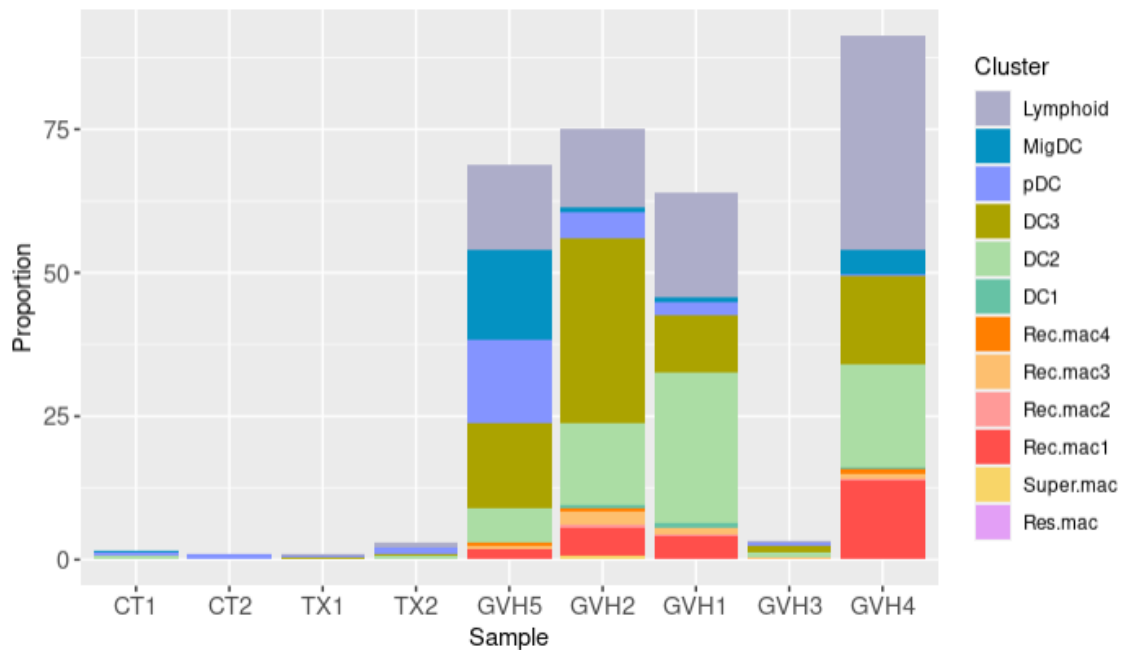
B
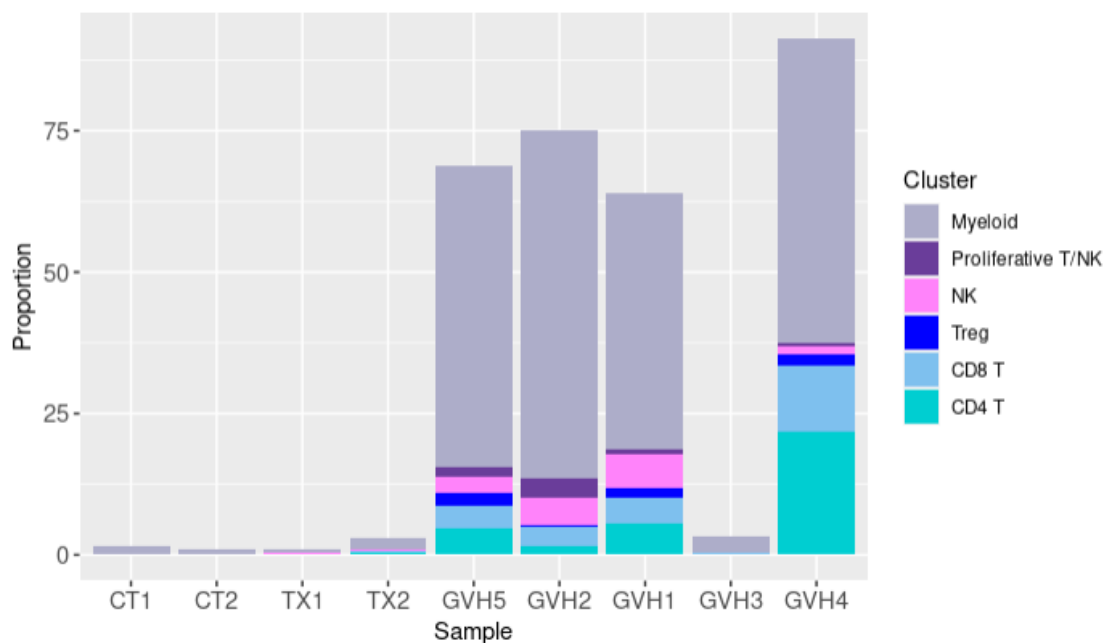
Figure 5.9 Bar charts showing the proportion of each epidermal immune cell subset by sample, normalized to the total number of cells of individual samples. (A) Myeloid populations, and (B) lymphoid populations. Sample from patients affected by GVHD were ordered by the timeframe between HSCT and sample being taken, with GVH5 (left) being the sample taken at the earliest timepoint. CT = healthy donor; TX = transplant control; GVH = graft versus host disease.

5.2.4 Reconstitution of myeloid cells

As the identity and origin of the myeloid cells were determined, the reconstitution dynamics of each type of myeloid cells were then compared (Figure 5.10). Since LCs did not cluster separately in some samples, positive expression of *CD207* within the myeloid cell clusters was used to estimate the number of LCs.

In the epidermis, there was complete donor chimerism within the macrophage and DC compartment. Recipient cells were only present among LCs in small proportions and could be found in the earliest timepoint sample (GVH5) and one of the transplant controls (TX1). In the dermis, however, recipient cells were more prominent. The majority of resident macrophages in samples collected at an early timepoint (GVH5, GVH2) and in one of the transplant controls (TX2) were recipient cells. Recipient dermal DCs could also be seen in one sample (GVH3). With regard to the rate of engraftment, the epidermal LC compartment reached complete donor chimerism earlier than the dermal resident macrophage compartment.

Figure 5.10

Figure 5.10 Donor chimerism of each myeloid cell population in the skin.
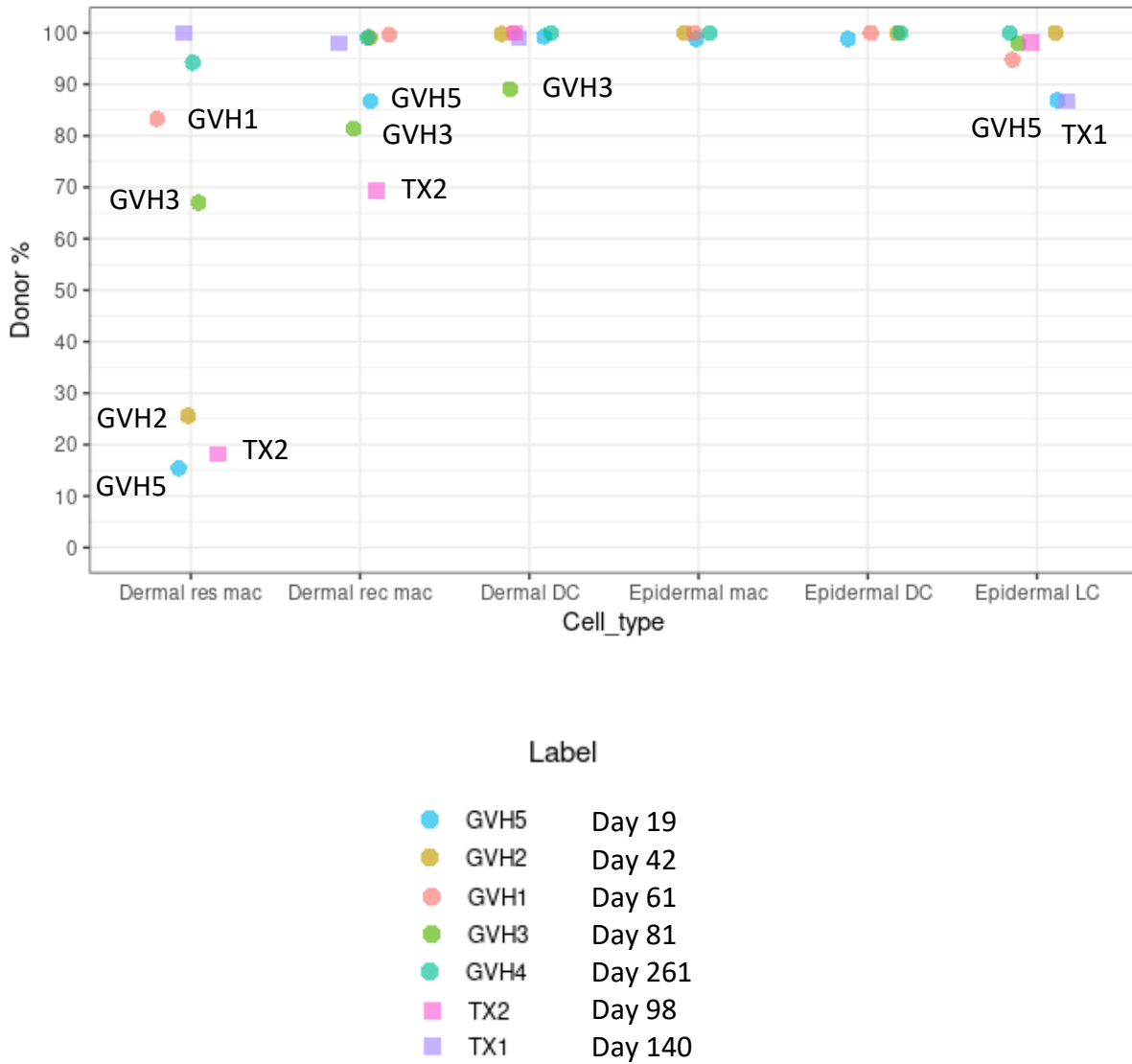
Figure 5.10 Donor chimerism of each myeloid cell population in the skin. Macrophages and DCs were not present in the epidermis of transplant controls and sample GVH3. Datapoints below 90% were labelled on the figure.

5.3 Chapter discussion

Similar to the dermis, cells from both the myeloid and lymphoid lineage were enriched in the epidermis of patients with GVHD. pDCs were also only enriched in the epidermis samples that were collected at an early timepoint. In the epidermis, however, myeloid infiltrates were dominated by DCs.

The presence of immune cells in the epidermis was unlikely due to contaminations of cells from the dermis during sample processing, since cells that were normally present in the dermis (e.g., fibroblast) could not be found in the epidermis samples. Instead, the vast majority of these cells were projected onto the "Recruited macrophage 1" cluster, suggesting a pro-inflammatory role by infiltrated macrophages in the epidermis. These macrophages may also have specific traits that allow access to the epidermis compared to other macrophage subsets. On the other hand, a cluster of proliferative keratinocytes was only present in the epidermis of transplant and healthy controls but not in patients with GVHD. This was possibly due to the pathogenic effector cells in the epidermis killing the epidermal stem cells, leading to an impartment of skin homeostasis (Takahashi et al., 2018).

An interesting observation among the epidermis samples was that in one of the GVHD-affected samples (GVH3), the infiltration of immune cells was minimal. From the previous chapters I established that large populations of recipient-derived T cells, including Tregs, were present in the blood and dermis of that patient. These features were atypical for GVHD, which according to the classical model, pathophysiology was driven by donor-derived effector T cells. Histology of the skin also lacked common signs of GVHD. This may be consistent with the heterogeneous skin manifestations of patients clinically diagnosed with GVHD. The difference may also be reflected in the blood. Indeed, different forms of cutaneous manifestations have been reported, including an eczema-like GVHD (Chan, Wood and Mesbah Ardakani, 2020). Although the T cells in the epidermis and dermis of other skin biopsies were mainly donor derived, recipient T cells and Tregs were also present to various extent. In one epidermis sample (GVH5) almost all T cells were recipient derived. In the next chapter, I will be studying the dynamics of T cells across tissues and compare between different types of GVHD manifestation.

With regard to the aspect of myeloid cells, HSCT offers fascinating insights to their ontogeny. Similar to dermal resident macrophages, LCs are also known for their self-renewed properties

and bone marrow dependency after HSCT (Collin *et al.*, 2006; Haniffa *et al.*, 2009; Mielcarek *et al.*, 2014). Through scRNA-seq, I demonstrated that epidermal LCs survive HSCT and were replaced by donor bone marrow-derived progenitor cells at a higher rate than dermal resident macrophages, which supports previous finding. Current results also suggested that GVHD promotes engraftment of LCs, as one of the transplant controls (TX1) had the lowest proportion of donor-derived LCs despite at 140 days post-HSCT.

Nonetheless, there are limitations to using scRNA-seq to determine the rate of replacement of cells in GVHD in general. Studies involving scRNA-seq techniques usually have a very small sample size due to cost limitations. This made it challenging to take into account the natural biological variations between sample donors. Moreover, the cells of interest might not always form a separate cluster, which complicates the quantification process. In this dataset, LCs in samples affected by GVHD did not always form a cluster. This might be due to low number of cells and the transcriptional changes of LCs after activation resembling other activated DCs. Incorporating protein labelling may address the problem, but it is not always possible especially if the starting material for scRNA-seq experiment is limited.

In conclusion, through scRNA-seq I demonstrated that in GVHD there was a profound infiltration of the epidermis by immune cells. The degree of infiltration and the cell types involved may potentially be indicative of the underlying type of GVHD. I also showed the differential replacement rate of skin DCs, macrophages and LCs by donor derived cells, which suggested their distinct ontogeny and dependency of bone marrow-derived progenitors upon significant perturbations.

Chapter 6. Cross-tissue analysis

6.1 Introduction

Reduced T cell receptor diversity after HSCT is associated with higher risk of GVHD and relapse (Yew *et al.*, 2015). Pinpointing the pathogenic clones responsible for GVHD, however, is challenging as it involves clonal expansion of rare donor clones, and the dominant clones were unique to each recipient (Wu *et al.*, 2021). Regardless, studying personal T cell clonotypes may provide valuable insights towards GVHD pathology, especially with the recent discovery of host Trm after HSCT (Divito *et al.*, 2020; Strobl *et al.*, 2020).

On the other hand, expansion of host macrophages and inhibition of donor-derived macrophages confers protection towards GVHD in mice (MacDonald *et al.*, 2010; Hashimoto *et al.*, 2011; Alexander *et al.*, 2014). Although donor CD14+ monocyte-derived macrophages have been shown to potentially mediate GVHD (Jardine *et al.*, 2020), I have demonstrated in the previous chapter that donor-derived cells were also responsible for replenishing macrophages bearing tissue resident signatures which have potential immunomodulatory properties. In steady states, the human skin consists of a transient, CD14+ monocyte-derived macrophage population (McGovern *et al.*, 2014). However, the mechanism of long-term tissue-resident macrophages replacement, especially in HSCT settings, remained undefined.

Chapter aims:

1. Investigate overlapping of clonotypes between the blood, dermis and epidermis

2. Define the origin (donor/recipient) and properties of the expanded clonotypes

3. Establish a connection between blood progenitors and skin macrophages

6.2 Results

6.2.1 Overview of clone size distribution across tissues

Since clonal expansion of T cells may represent a response to alloantigens, the clone size distribution of the expanded T cells between GVHD-affected tissues and transplant controls was compared as a proxy of the degree of graft versus host response (Figure 6.1). Larger clones could be found in GVHD-affected dermis and epidermis compared to transplant controls. They were also present at a higher frequency.

Figure 6.1
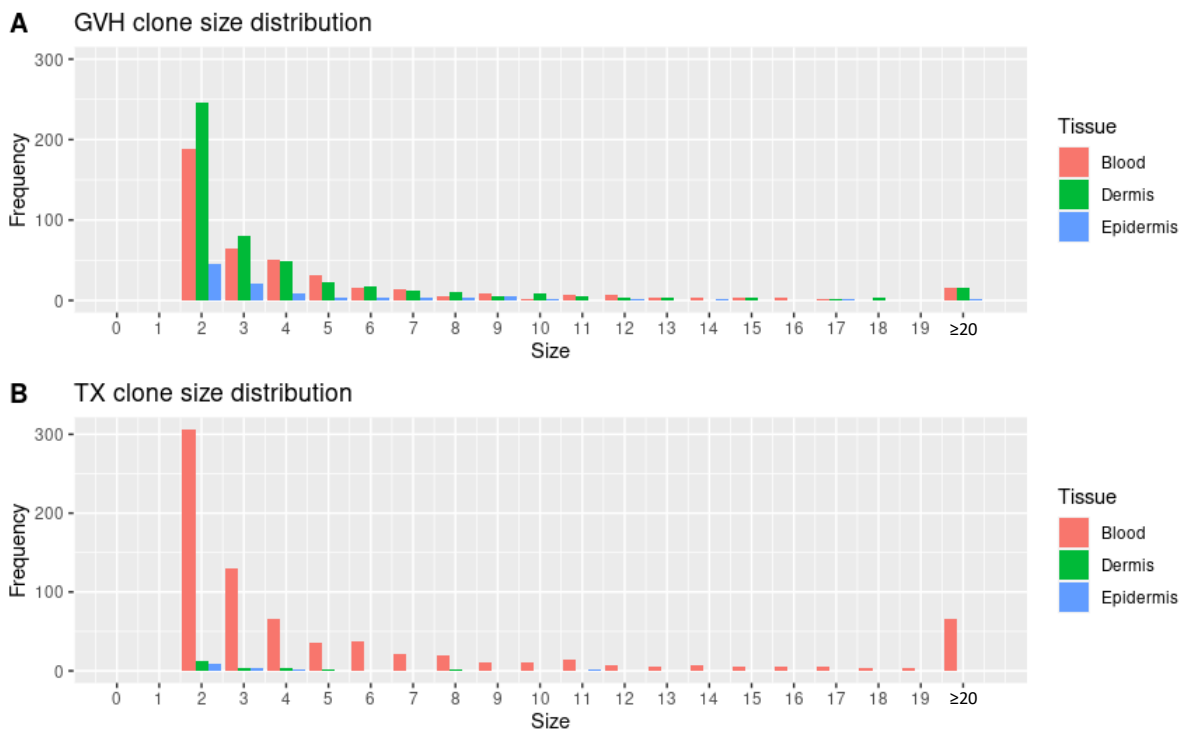
Figure 6.1 Bar charts showing the distribution of clone sizes across tissues. (A) In GVHD (n = 5), and (B) transplant controls (n = 2). Each bar represents the cumulative frequency of all clonotypes with the specified clone size across all samples of a condition. Frequency of clonotypes with clone size equal to one (not expanded within the sampled tissue) was not shown.

6.2.2 Overlapping of clonotypes between tissue compartments

In order to understand the selection and expansion of T cell clonotypes across tissues, the distribution of TCR sequences between blood, dermis and epidermis was compared. There were many unique TCR sequences in each compartment, so I focused on clones identified by the same TCR sequence appearing at least twice in any compartment. I started either in the blood, mapping clones forward into the dermis and epidermis (Figure 6.2, panel A), or in the epidermis, mapping back to the dermis and blood (Figure 6.2, panel B). Bidirectional mapping from the dermis did not reveal any additional information and was not displayed. The origin of the clones, whether they were derived from the donor or the recipient, was also determined.

In transplant controls, there were small number of recipient T cell clones in the epidermis (Figure 6.2, panel A). These clones were found in lower proportions in the dermis but very rarely in the blood. On the other hand, massive number of donor T cell clones were found in the blood (Figure 6.2, panel B). These cells had low access to the dermis and almost no access to the epidermis. There were more common clonotypes between the blood and dermis of sample TX2 compared to TX1. Patient TX2 subsequently developed GVHD (approximately 40 days after sampling).

Figure 6.2
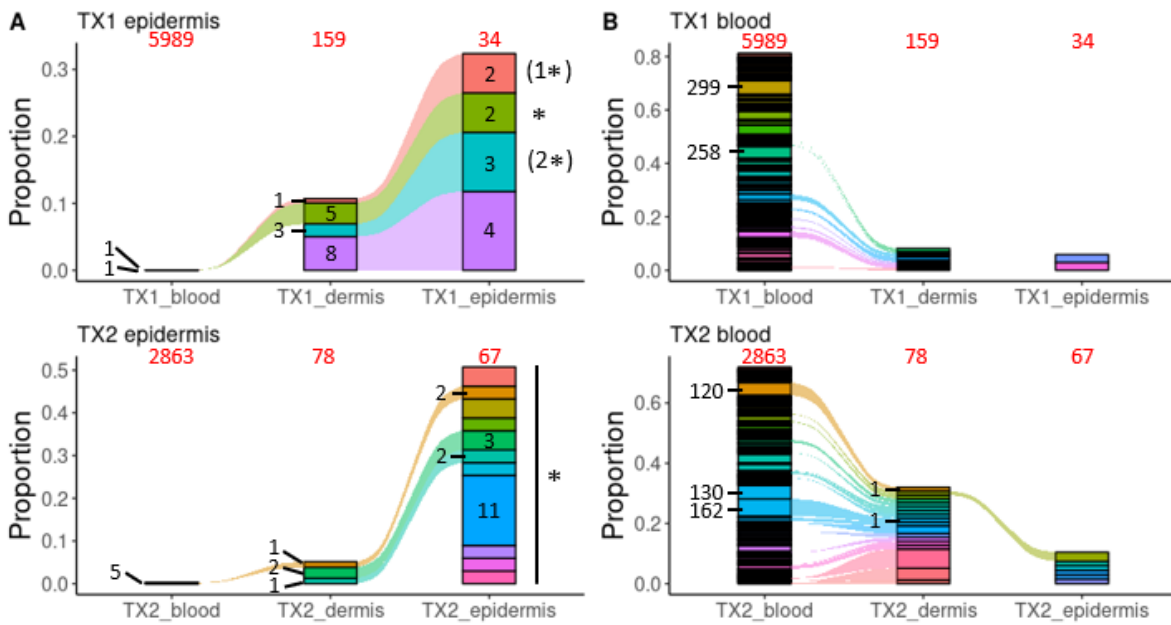
Figure 6.2 Alluvial plots showing the clonotype relationship between tissues of each transplant control. Clonotypes with clone size larger than one in the (A) epidermis and (B) blood compartment were displayed. The asterisk (*) next to the clonotype indicated clones derived from the recipient. Number at the top of each bar displayed the total number of T cells in each sample. Number within each box displayed the clone size.

In the case of GVHD, large number of donor T cell clones were present in the epidermis (Figure 6.3, panel A). These clonotypes were found in high proportions in the dermis and were also well represented in the blood. On the other hand, donor T cell clones in the blood had free access to the dermis and the epidermis (Figure 6.3, panel B). Additionally, specific features could be seen in some samples. In the earliest timepoint epidermis sample (GVH5), all expanded clones were recipient derived and they were only present in the skin. Moreover, the latest timepoint epidermis sample (GVH4) contained oligoclonal donor T cells, with a small number of clonotypes occupying the majority of the T cell repertoire.

Conversely, dataset GVH3 was dissimilar to the others. Clonal expansion in the epidermis was minimal. The largest clones in the blood were recipient derived, and they did not enter the epidermis.

Figure 6.3
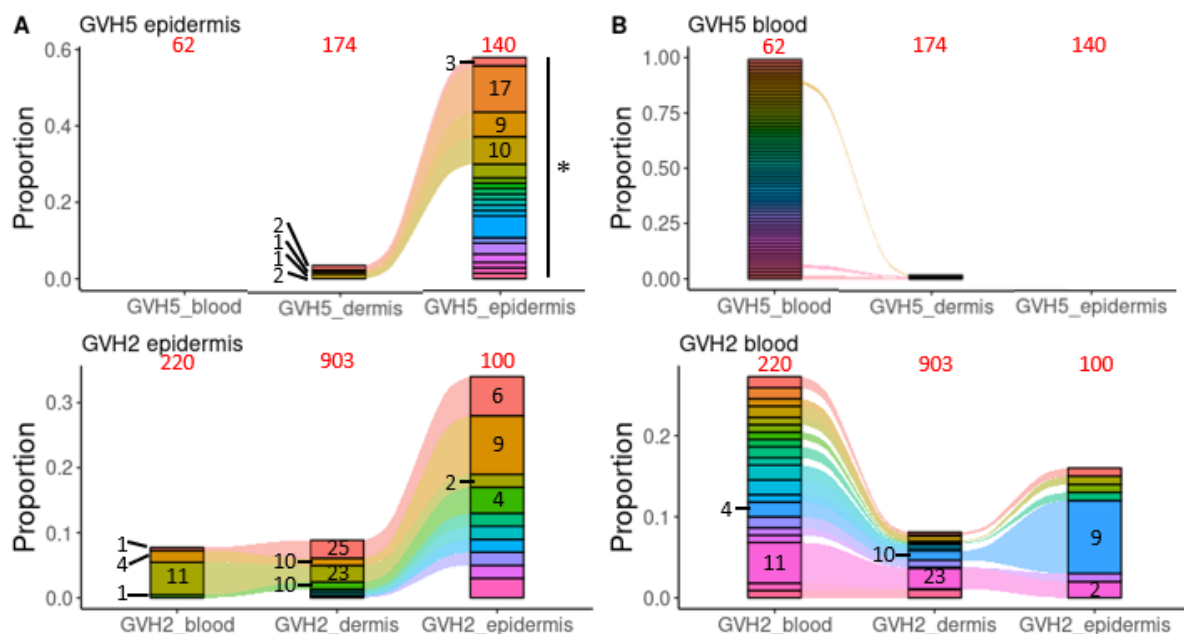
(Figure continued next page)
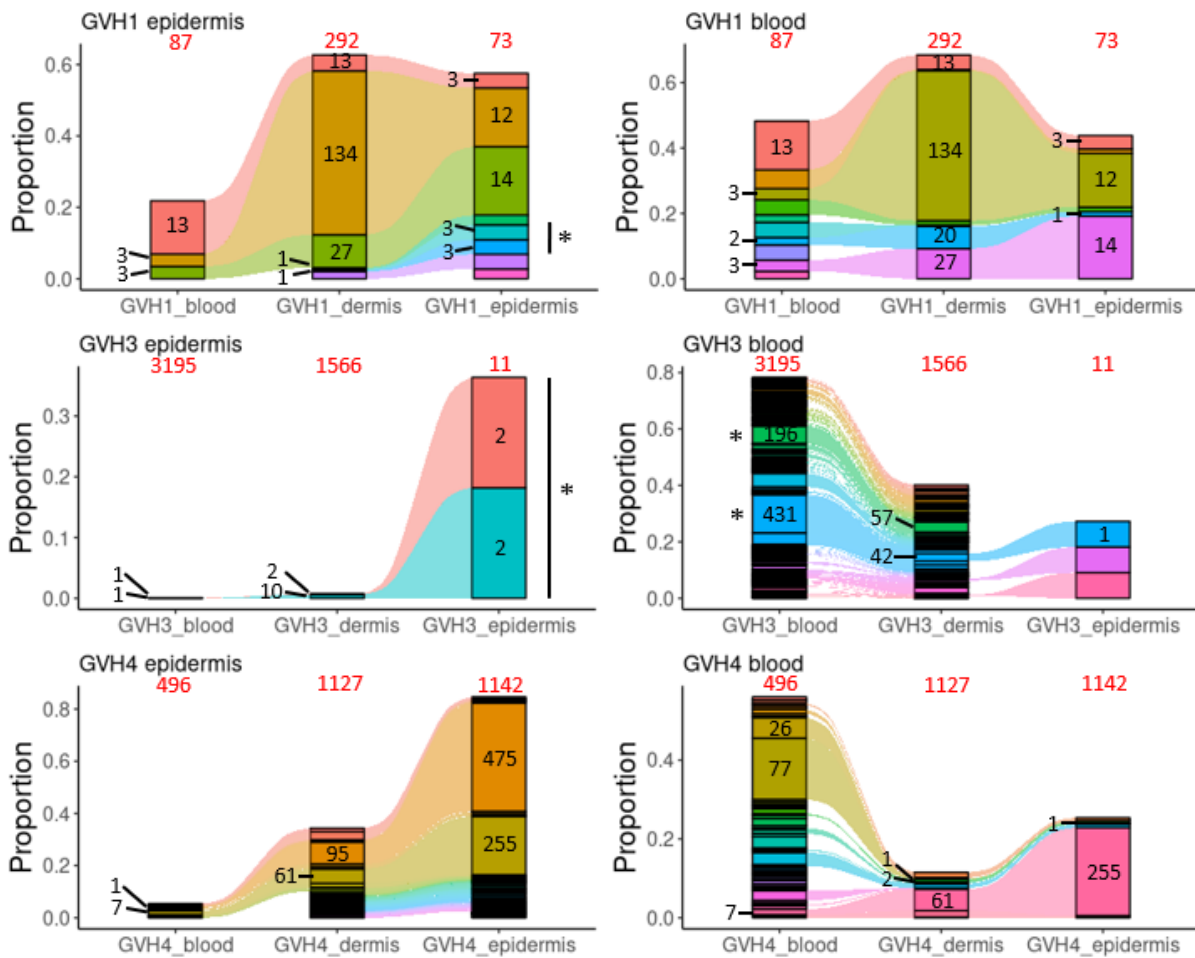
Figure 6.3 Alluvial plots showing the clonotype relationship between tissues of each patient affected by GVHD. Samples were ordered by the timeframe between HSCT and sample being taken, with GVH5 (top) being the sample taken at the earliest timepoint. Clonotypes with clone size larger than one in the (A) epidermis and (B) blood compartment were displayed, with the exception of sample GVH5 which all clonotypes from the blood compartment (clone size of all clonotypes equal to one) were displayed. The asterisk (*) next to the clonotype indicated clones derived from the recipient. Number at the top of each bar displayed the total number of T cells in each sample. Number within each box displayed the clone size.

6.2.3 Characterization of the expanded clonotypes

After studying the clonal relationship of the expanded clonotypes across tissues, these clonotypes from each sample were then characterized in more detail. A reference UMAP was constructed by using the T cells from the blood of healthy controls, and the T cells from the patient samples were projected onto the reference (Figure 6.4). Most cells from the transplant samples were projected onto either CD4 or CD8 effector memory T cell populations.

Figure 6.4
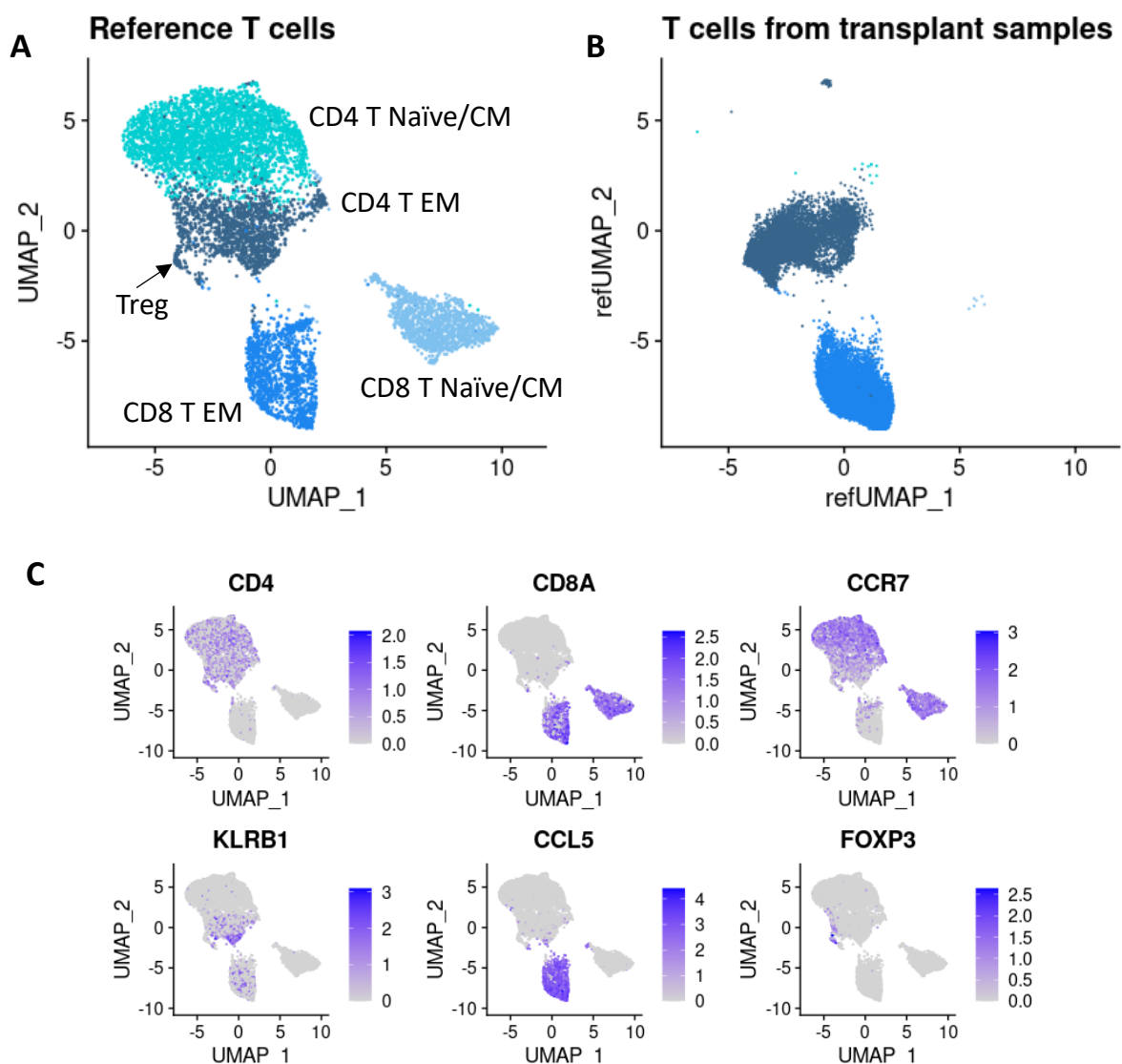
Figure 6.4 T cells from transplant controls and patients affected by GVHD (query) projected onto the T cells from the blood of healthy controls (reference). (A) UMAP of reference T cells. (B) Cells from transplant controls and patients affected by GVHD. (C) Expression of representative marker genes. Cells with predicted score lower than 0.8 were removed.

The two most expanded clonotypes from the blood and the epidermis of each sample were then highlighted. In transplant controls, the most expanded clonotypes in the blood were CD8 effector memory T cells (Figure 6.5, panel A). Expanded clonotypes in the epidermis included both CD4 and CD8 effector memory T cells, albeit in lower numbers. Expansion of Treg within the epidermis was also evident in one of the samples (TX2). In the case of GVHD, expansion of CD4 and CD8 effector memory T cells were observed in both the blood and the epidermis (Figure 6.5, panel B). An extensive involvement of CD4 effector memory T cells was observed in the epidermis of the latest timepoint sample (GVH4). Moreover, expansion of Treg in the epidermis was evident in the atypical sample (GVH3).

Figure 6.5
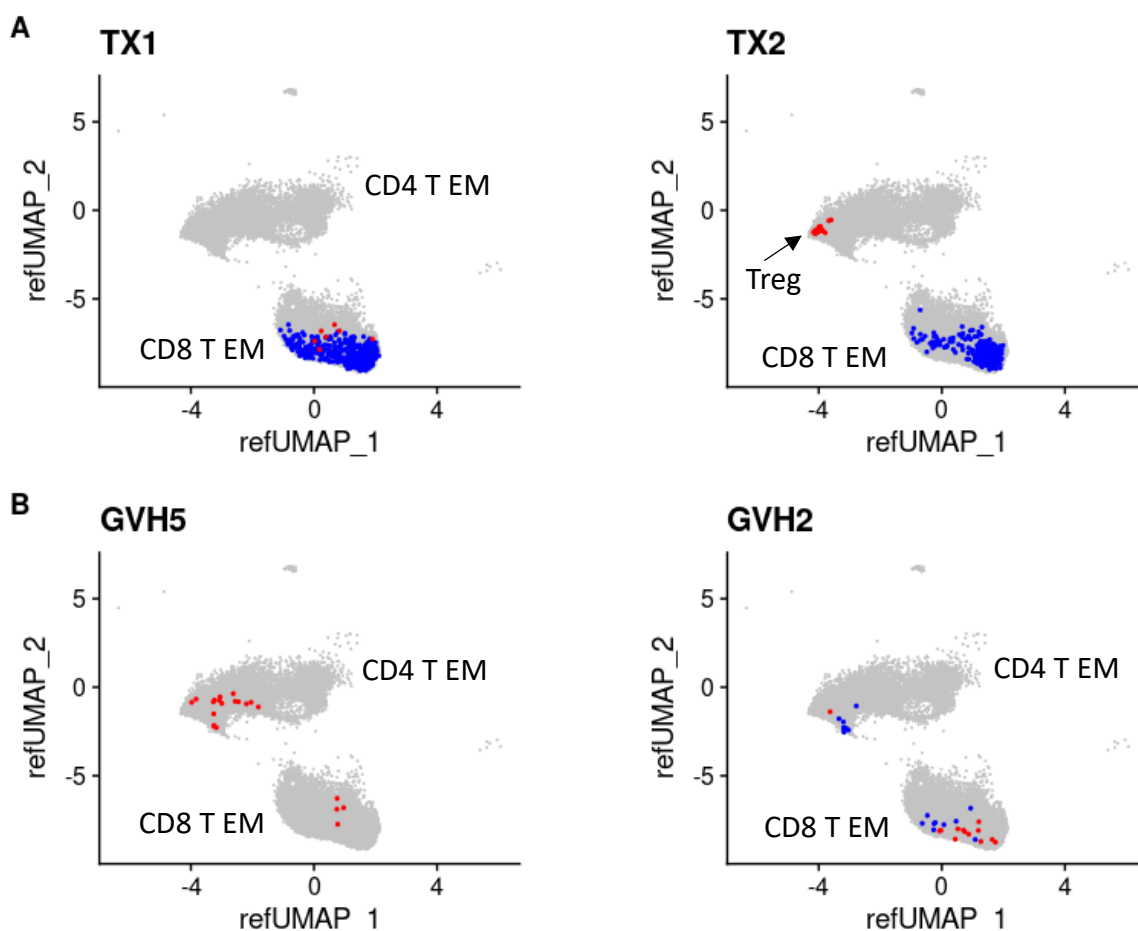
(Figure continued next page)
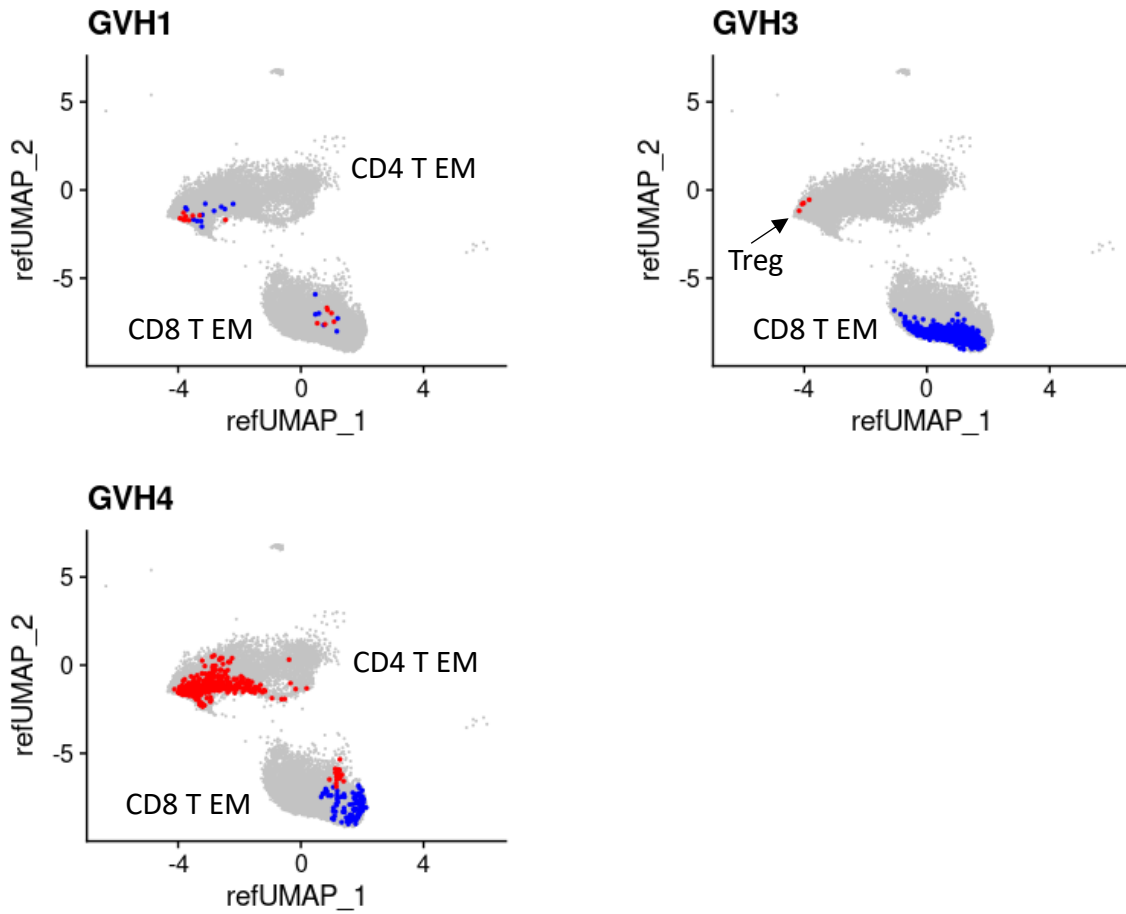
Figure 6.5 Characterization of T cells from transplant controls and patients affected by GVHD. The two most expanded clonotypes in the epidermis (red) and blood (blue) of each sample were highlighted. (A) Transplant controls, and (B) patients affected by GVHD. Sample from patients affected by GVHD were ordered by the timeframe between HSCT and sample being taken, with GVH5 being the sample taken at the earliest timepoint.

6.2.4 Connection between blood progenitors and skin macrophages

I next explored the relationship between the myeloid cells in the blood and the skin. The myeloid cells from the blood were projected onto the dermal myeloid cells from Figure 4.6. pDC and cDC from the blood were projected onto the pDC and DC1 cluster of the dermis, respectively (Figure 6.6). The majority of CD14+ monocytes from the blood were projected onto the "recruited macrophage 1" cluster of the dermis, with some projected onto the DC3 cluster. Surprisingly, the majority of CD16+ monocytes from the blood were projected onto the resident macrophage cluster of the dermis. There were also some CD16+ monocytes projected onto various recruited macrophage clusters.

Figure 6.6
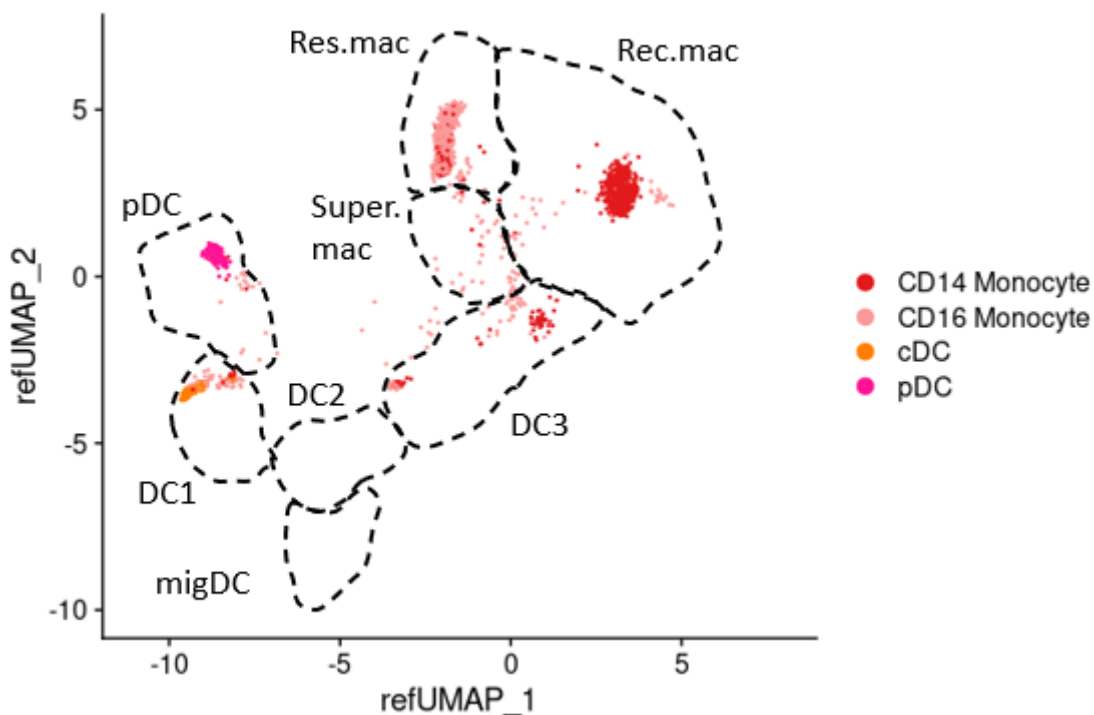
Figure 6.6 Blood myeloid cells from all samples (query) projected onto the dermal myeloid cells from Figure 4.6 (reference). Contaminating fibroblasts and lymphocytes from the dermal myeloid dataset were removed before the blood myeloid cells were projected onto the dataset. Blood myeloid cells with predicted score lower than 0.6 were removed.

To further investigate the "resident macrophage-like" CD16+ monocytes, the gene expression between these cells and the other CD16+ monocytes that were projected onto the recruited macrophage clusters were compared (Table 6.1, Figure 6.7). The top differentially expressed genes of these resident macrophage-like monocytes included genes that were also expressed by dermal resident macrophages, such as those involved in the complement system (*C1QA, C1QB, C1QC*).

Table 6.1

	p_val	avg_log2FC	pct.1	pct.2	p_val_adj
C1QA	2.382499e-31	2.3899368	0.644	0.219	5.802100e-27
C1QB	3.593907e-14	2.1753154	0.355	0.095	8.752242e-10
C1QC	4.602803e-13	1.7806872	0.278	0.045	1.120921e-08
EGR1	1.643619e-13	1.5218873	0.327	0.085	4.002704e-09
HLA-DRB5	1.170947e-40	1.1815863	0.982	0.647	2.851607e-36
IFI27	1.012416e-04	1.1524074	0.187	0.080	1.000000e+00
MNDA	5.270647e-29	1.0429257	0.709	0.234	1.283561e-24
FGL2	1.823309e-44	1.0059806	0.968	0.562	4.440303e-40
APOBEC3A	1.935275e-38	1.0011959	0.929	0.453	4.712976e-34
IFITM3	3.915857e-57	0.9560934	1.000	0.990	9.536287e-53
ANXA1	3.998534e-16	0.9430956	0.937	0.701	9.737629e-12
EPSTI1	2.516253e-49	0.9419400	0.864	0.259	6.127831e-45
ZFP36L2	7.081779e-33	0.9282726	0.951	0.577	1.724626e-28
TMEM176B	1.315570e-18	0.9269356	0.606	0.274	3.203807e-14
HLA-DQA1	4.335809e-32	0.9228858	0.928	0.507	1.055899e-27
CD52	3.002188e-34	0.9134721	0.986	0.701	7.311228e-30
VAMP5	7.198150e-38	0.8935401	0.982	0.597	1.752966e-33
CTSC	2.638257e-39	0.8904658	0.961	0.567	6.424947e-35
LAP3	5.494492e-40	0.8834487	0.941	0.458	1.338074e-35
IFITM1	5.426134e-27	0.8651038	0.979	0.851	1.321426e-22

Table 6.1 List of top differentially expressed genes of blood CD16+ monocytes that were projected onto the resident macrophage cluster compared to other CD16+ monocytes that were projected onto the recruited macrophage clusters. The average fold change was presented in log2 scale (avg_log2FC). pct.1 = percentage of cells in the cluster expressing the gene; pct.2 = percentage of cells in all the other clusters where the gene was detected; p_val = p value; p_val_adj = adjusted p value.

Figure 6.7
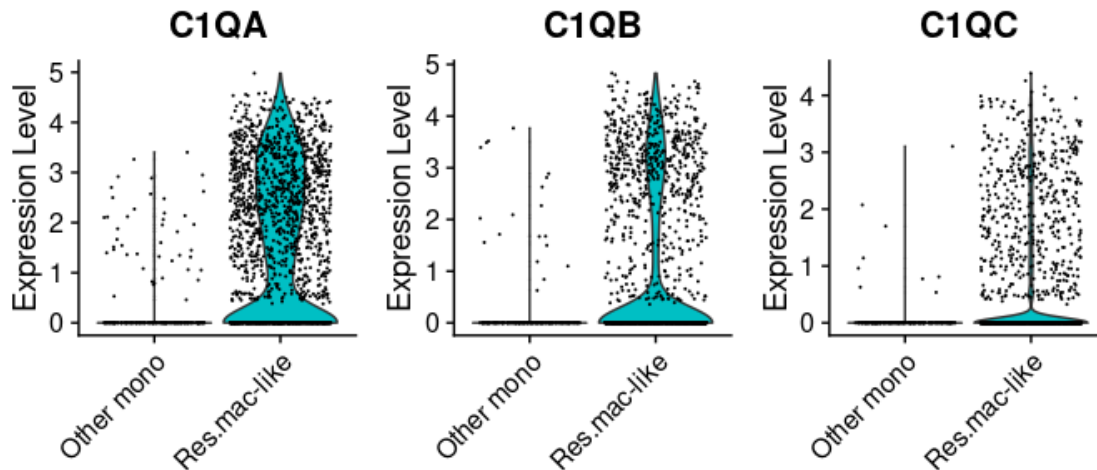
Figure 6.7 Expression of genes from the C1q family by CD16+ monocytes that were projected onto the resident macrophage cluster (Res.mac-like) compared to other CD16+ monocytes (Other mono) projected onto the recruited macrophage clusters.

Since the blood CD16+ monocytes that were mapped to the resident macrophage cluster might potentially be the progenitors of the donor-derived dermal resident macrophages, I sought to further characterize this population. I utilized a flow cytometry panel previously established for distinguishing myeloid subsets (Cytlak et al., 2019) and investigated whether this population could be discerned from other CD16+ monocytes by phenotype in the blood (Figure 6.8).

Using the single-cell dataset, I first compared the gene expression of the markers in the flow cytometry panel between the CD16+ resident macrophage-like monocytes and other CD16+ monocytes (Figure 6.9). Most markers had similar expression between the groups except IL3RA (CD123) and HLA-DRA. Flow cytometry of the blood sample collected at the earliest timepoint (GVH5) showed that although the CD16+ monocyte population could be identified, this population appeared homogenous with the provided markers (Figure 6.10).

Figure 6.8
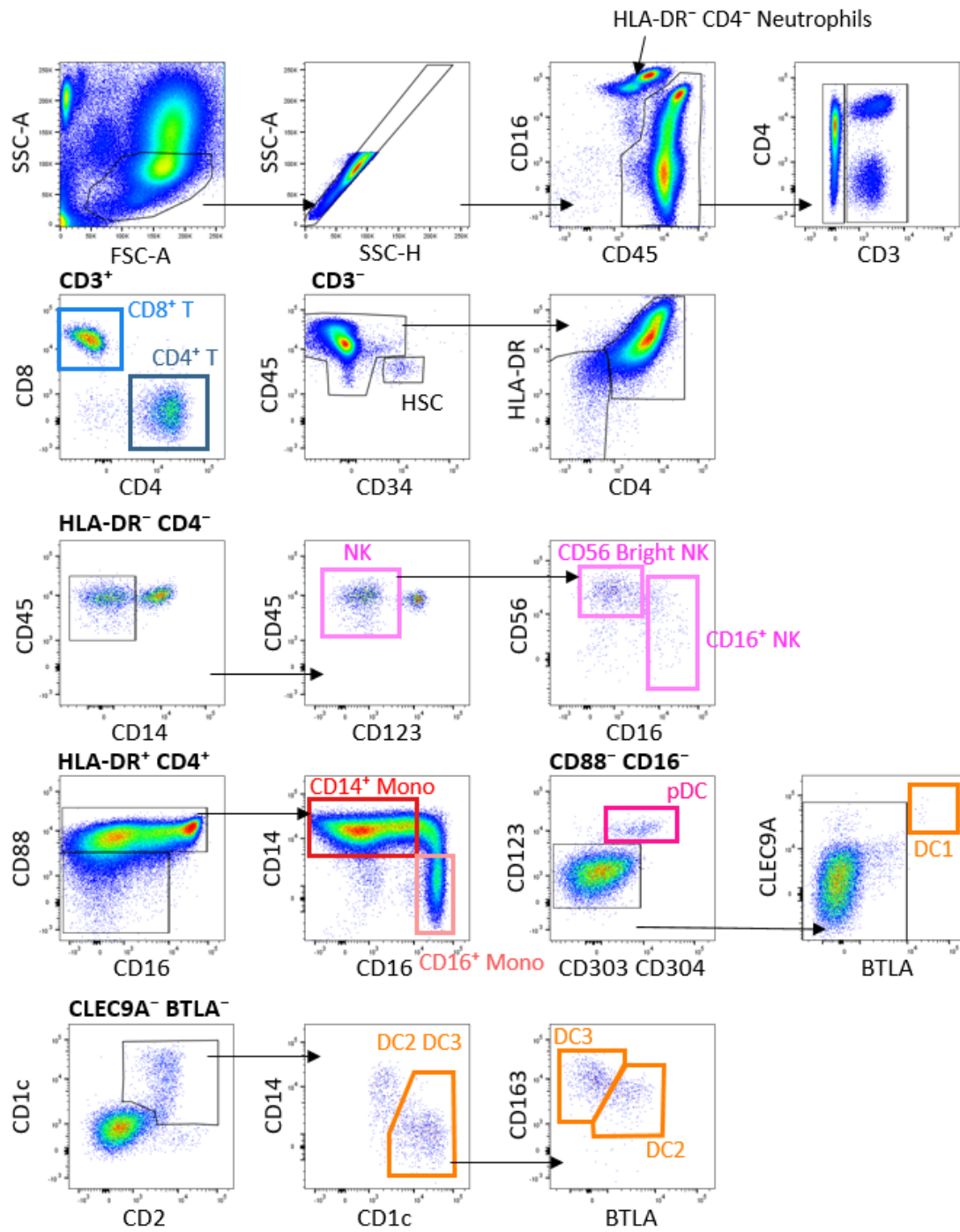
Figure 6.8 Flow cytometry gating of one early timepoint GVH sample (GVH5).

Figure 6.9
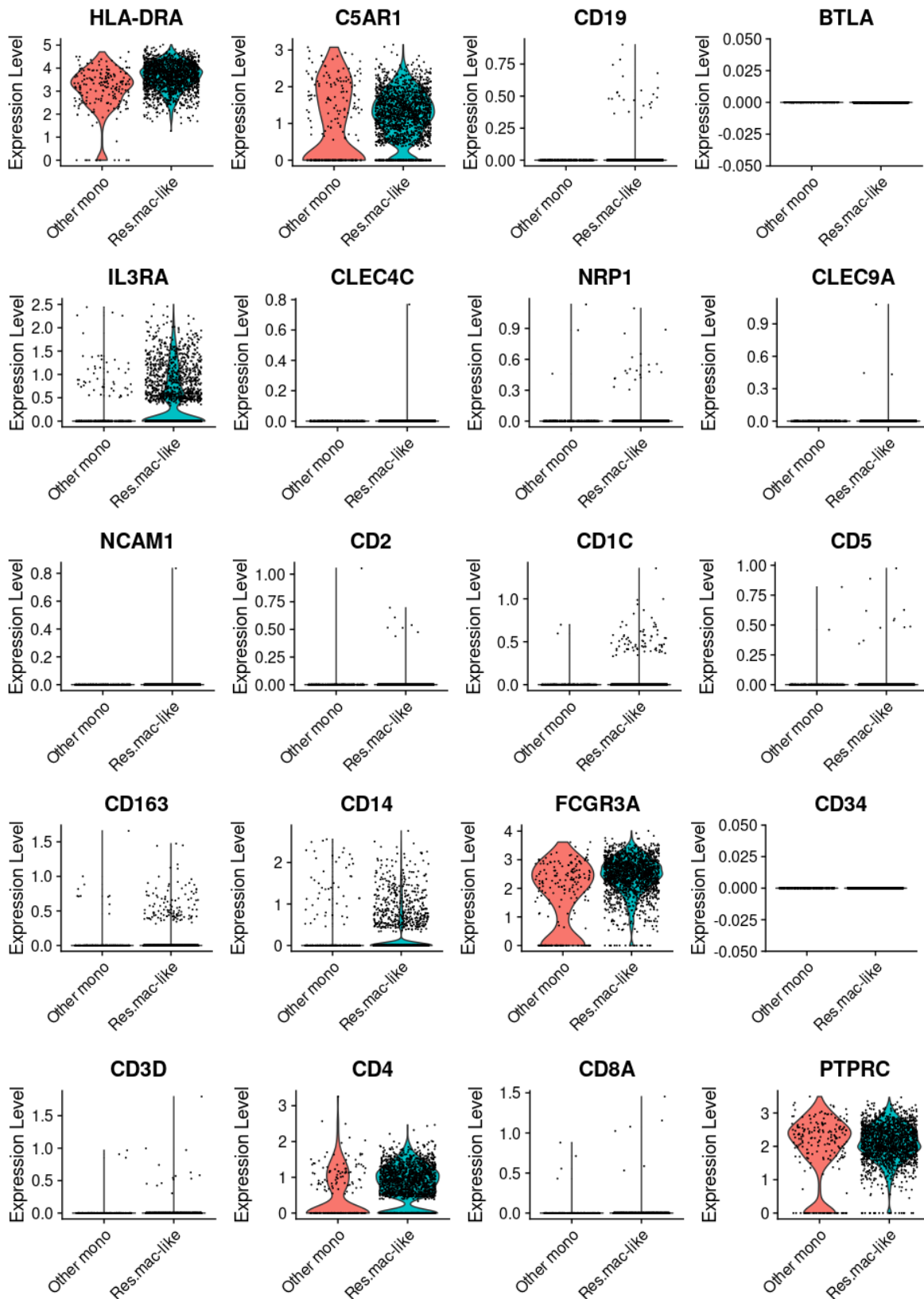
Figure 6.9 Violin plots comparing the expression of marker genes in the flow cytometry panel between the resident macrophage-like CD16+ monocytes and other CD16+ monocytes. Res.mac-like = resident macrophage-like CD16+ monocytes.

Figure 6.10
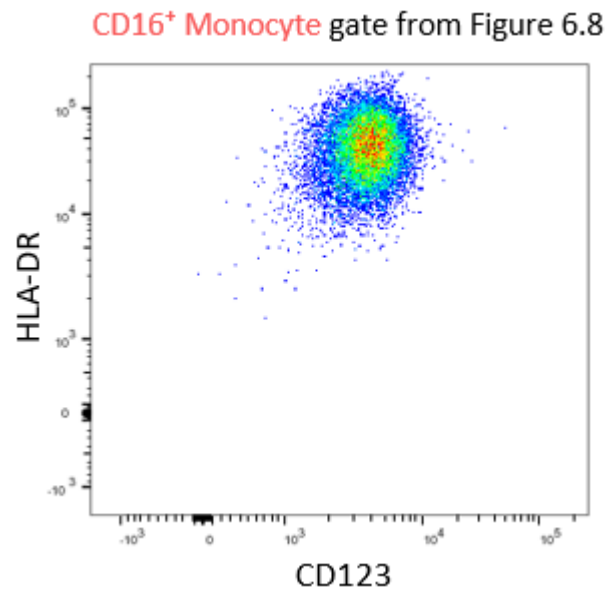
Figure 6.10 Expression of HLA-DR and CD123 (IL3RA) within the CD16+ monocyte population from Figure 6.8.

6.3 Chapter discussion

Donor T cell-mediated damage to the epidermis is a hallmark of cutaneous GVHD (Santos e Sousa, Bennett and Chakraverty, 2018). Through single-cell TCR profiling, I discovered that cutaneous GVHD features clonal overlapping of expanded, donor-derived CD4 and CD8 effector T cells across tissues (blood, dermis and epidermis). Although there were expanded clones in the blood of transplant controls, only a limited proportion had access to the dermis, and even fewer to the epidermis. In addition, the skin from transplant controls retained a higher percentage of host resident T cells. Nonetheless, variations existed between individuals regarding the type of T cells involved and their origin, which will be discussed below. After collating the clonotype patterns and the findings from previous chapters, I grouped the types of GVHD into two main categories: typical and atypical.

The typical category includes patient GVH2, GVH1 and GVH4. This group of patients present classical features of GVHD including infiltration of the dermis and epidermis by clonally expanded, donor-derived T cells from the blood. Few Tregs were present in these samples, with no evidence of clonal expansion, hence they might act as bystanders or be dysregulated. Additionally, the dermis and epidermis were infiltrated by donor-derived myeloid cells including macrophages and DCs. The rate of donor macrophage and LC engraftment also fit the pace of GVHD as previously described (Haniffa *et al.*, 2009). Together with the presence of classical features in histology, patient GVH2 and GVH1 represent typical cases of acute GVHD, while patient GVH4 represent a typical case of chronic GVHD. Comparing to other GVHD samples, the dermis and epidermis of GVH4 contained the highest proportion of T cells. The expansion of donor T cells in the epidermis was oligoclonal. Moreover, the extensive involvement of CD4 T cells in the epidermis agrees with the classical description of chronic GVHD (Yamashita *et al.*, 2004).

Conversely, GVHD can sometimes manifest in an atypical manner. Although infiltration of myeloid cells could be seen in the skin of patient GVH5, the enrichment of T cells in the epidermis was due to local clonal expansion of recipient cells. Of note, patient GVH5 received prior checkpoint inhibitor therapy and it is well-documented that this treatment is associated with increased risk of GVHD (Ijaz *et al.*, 2019). Here I showed that the expansion of recipient T cells may be contributing to this phenomenon.

116

Additionally, patient GVH3 displayed a lack of immune cell enrichment in the epidermis and the expansion of recipient-derived Tregs was evident. Some recipient-derived DCs also persisted in the dermis. Moreover, the histology of skin biopsy showed lack of typical GVHD features but resembled eczema. Interestingly, the most expanded clonotypes in the blood of this patient were recipient derived. The observation of recipient-derived T cells in the blood is unexpected as chemotherapy and conditioning regimen used in HSCT is thought to eliminate all T cells in the blood. Recently, it has been suggested that recipient-derived T cells from the skin can recirculate into the blood and this process may prime GVHD in other tissues (Strobl *et al.*, 2021). In our dataset, however, it is difficult to conclude whether this recirculation process occurred, as recipient-derived T cells were always present in the blood in all follow-up sampling (data not shown). Therefore, the possibility of some recipient T cells remained in the circulation before HSCT cannot be excluded. Nonetheless, this further highlights the potential involvement of recipient-derived T cells in GVHD.

A caveat of this part of this part of study was that the transplant controls were not optimal. Transplant control TX1 subsequently died due to relapse. It is possible that the donor cells were not fully capable of causing a graft versus host reaction. Transplant control TX2 developed GVHD approximately 1.5 months after sampling. Results suggested that increased overlapping of donor clonotypes between the blood and dermis may be a feature of pre-GVHD onset.

Finally, the developmental relationship between myeloid cells from the blood and skin was also investigated. Myeloid cells from the blood were projected onto the dermal myeloid cell reference dataset. Most CD14+ monocytes were projected onto the "recruited macrophage 1" cluster, which corroborated previous finding that blood CD14+ monocytes and skin macrophages are closely-related populations (McGovern *et al.*, 2014; Jardine *et al.*, 2020). Surprisingly, CD16+ monocytes were projected onto the resident macrophage population. Under physiological conditions, the resident macrophage population is known to propagate independently from circulating monocytes (Hashimoto *et al.*, 2013). Although this population is replaced by donor bone marrow HSC-derived cells after HSCT, the developmental pathway between HSCs and resident macrophages remained elusive. I showed that CD16+ monocytes that were projected onto resident macrophages highly express genes of the complement system (*C1QA*, *C1QB*, *C1QC*). Since these genes are also highly expressed by dermal resident macrophages, this suggests a potentially connection between CD16+ monocytes and dermal resident macrophages. Nonclassical (CD16+) monocytes play an important role in the

maintenance of the vasculature and are implicated in inflammation and chronic diseases (Narasimhan *et al.*, 2019). Although under homeostatic conditions they do not readily differentiate into macrophages (Hanna *et al.*, 2011), they can be recruited to the site of injury and differentiate into alternatively activated, wound healing macrophages (Olingy *et al.*, 2017). Recently, it has also been demonstrated with a humanized mouse model that a subset of lung macrophages with high C1Q expression is developed from circulating CD16+ monocytes (Evren *et al.*, 2021). This model, together with the results in this chapter suggest a potential developmental relationship between CD16+ monocytes and dermal resident macrophages, at least during immune reconstitution after HSCT.

Chapter 7. Concluding discussions

7.1 Summary of findings

Understanding the immune mechanism of GVHD is pivotal to the development of better treatments. This thesis is the first myeloid-focused human GVHD study using droplet-based high throughput scRNA-seq, providing a comprehensive view of the cellular heterogeneity of the skin and blood from patients with GVHD at the time of disease onset. Results showed that the cellular composition of GVHD-affected tissues was vastly altered compared to the steady state. By harnessing SNPs of each donor-recipient pair, these changes could be attributed mostly to donor inflammatory cells. The expansion of recipient T cells was also detected in one case, indicating a potential host versus graft reaction. Moreover, through mapping the myeloid cells between the blood and skin, the mechanism of long-term tissue-resident macrophages replacement was investigated. Based on these findings, two hypotheses are formulated and will be discussed in the next section.

Although much insight was gained from this study, it was not without limitations. Despite the skin being the most common organ affected by GVHD, skin-limited GVHD is not associated with the most adverse outcomes. This study will also benefit from having a larger cohort size, particularly transplant controls who are GVHD- and relapse-free long term. Investigating longitudinal samples of each patient will provide unique insight towards the pathophysiology of GVHD. Moreover, transcriptome is only one layer of machinery that governs cellular function. The measurement of other modalities such as surface protein expression and spatial information at a single-cell resolution will complement, validate and expand the understanding towards GVHD.

7.2 Formulated hypotheses

Based on the generated data, two main hypotheses were formulated:

<u>Hypothesis 1: Both myeloid and lymphoid cells play a key role in local immune reactions at the site of GVHD</u>

The role of T cells in GVHD pathophysiology is irrefutable. Recently, myeloid cells have also been shown to be important GVHD mediators. In this thesis, I defined the cellular composition of the blood, the dermis and the epidermis, and how they are altered in patients affected by GVHD. Results support the notion that both myeloid and lymphoid cells are critical in mediating local GVHD reactions.

In typical cases of cutaneous GVHD, T cells constituted part of the epidermal immune infiltrates (Figure 5.9). These cells were donor derived CD4 and CD8 effector memory T cells and were clonally expanded (Figure 6.5). The same clones could also be found across the dermis and PBMC (Figure 6.3). The role of host Trm in mediating GVHD has been suggested (Divito *et al.*, 2020; Strobl *et al.*, 2021). Evidence presented in this thesis demonstrated that host Trm were only expanded in checkpoint inhibitor therapy-mediated cutaneous GVHD, although the potential role of these cells in mediating GVHD in other tissues cannot be excluded. I also described another atypical case of cutaneous (eczematoid) GVHD of which T cells and other immune infiltrates were absent in the epidermis.

Similarly, myeloid cells were prominent in GVHD-affected dermis and epidermis, especially early after HSCT when the patients were lymphopenic (Figure 4.3, Figure 5.3). I have investigated the heterogeneity among the myeloid cells and revealed that donor-derived pro-inflammatory macrophage populations were enriched in the skin of patients affected by GVHD (Figure 4.9, Figure 5.9). I also identified a population of resident macrophages with the slowest donor engraftment rate among all myeloid cell subsets (Figure 5.10). Additionally, donor-derived DCs were enriched in the skin of patients affected by GVHD (Figure 4.9, Figure 5.9). This population of cells was the main myeloid infiltrate in the epidermis. Results showed that the newly defined DC subset, DC3, were among the upregulated DC populations, suggesting these cells may potentially contribute to GVHD pathophysiology. MigDC and pDC were only present in the skin of patients affected by GVHD. Further studies may provide valuable insight towards their roles in mediating GVHD. In terms of LCs, their roles were more challenging to define owing to low abundance. Host LCs were rapidly replaced by donor cells. A potential role

of donor-derived LCs was to carry antigens and present them in the lymph node, however this was only suggested by the atypical GVHD sample (Figure 4.5).

Furthermore, I showed that other immune cell populations such as NK cells and, in some cases, Tregs, were also enriched in the skin of patients affected by GVHD (Figure 4.3, Figure 5.9). It is likely that GVHD pathogenesis is a balancing act between various pro-inflammatory and immunomodulatory myeloid and lymphoid cell populations.

Hypothesis 2: Classical (CD14+) monocytes and nonclassical (CD16+) monocytes represent two distinct pathways of tissue macrophage development after HSCT

Two populations of macrophage exist in the dermis – a transient/recruited macrophage population and a resident macrophage population (reviewed in section 1.3.2). After HSCT, host dermal macrophages are gradually replaced by donor cells. This process is dependent on the donor bone marrow, however the exact mechanism remains elusive.

Previously it has been demonstrated that in GVHD settings, blood CD14+ monocytes are primed and can differentiate into dermal macrophages (Jardine *et al.*, 2020). I observed that CD14+ monocytes were enriched in the blood of the patients affected by GVHD (Figure 3.4) and were activated (Figure 3.7). The CD14+ monocytes were projected onto the recruited macrophage cluster (Figure 6.6), which supports the previous findings that CD14+ monocytes and recruited macrophages are two closely related populations. However, this leaves the replacement of resident macrophages unaccounted for.

Reconstitution of different myeloid cell subsets revealed that resident macrophages were replaced by donor bone marrow-derived cells (Figure 5.10), yet their blood progenitors were yet to be defined. I showed that donor-derived CD16+ monocytes were also enriched in the blood of the patients affected by GVHD. The majority of CD16+ monocytes were projected onto the resident macrophage cluster (Figure 6.6) and expressed some resident macrophage markers (Figure 6.7, Table 6.1), suggesting a "resident macrophage-like" characteristic of CD16+ monocytes (discussed in section 6.3). Although the direct proof of developmental relationship requires genetic manipulation of animal models, the evidence presented in this thesis suggests that blood CD16+ monocytes are blood progenitors of dermal resident macrophages, at least in HSCT settings. Donor-derived CD16+ monocytes may also be

connected to a resident macrophage-like subset (Recruited macrophage 2) which express immunomodulatory gene TGFBI (Figure 4.6, Table 4.5). Therefore, CD14+ monocyte and CD16+ monocyte may represent two distinct pathways of tissue macrophage development after HSCT, potentially playing opposing roles in GVHD pathogenesis.

7.3 Future directions

There are still many aspects remaining unexplored within the current dataset. With the cellular composition being defined, further analysis can be performed, such as delineating the functional differences of immune cells between samples collected from patients with GVHD and controls. Studying receptor-ligand interactions may also yield fascinating insights as the interplay between APCs, effector cells and their targets are intrinsic to GVHD pathogenesis. Additionally, the antigen-presenting properties of stromal cells can also be explored. With regard to the developmental potential of circulating monocytes, a trajectory can be constructed with the dermal recruited and resident macrophages to decipher the transcriptional programmes during this dynamic event.

Future experiments can incorporate the study of spatial information, such as using the 10x Genomics Visium platform. High-resolution, multichannel microscopy can be used to verify novel markers discovered by scRNA-seq, and study cell-cell interactions that may be regarded as doublets and removed during scRNA-seq analysis. GVHD-affected gut samples can also be collected to survey the early events of pathogenesis. These may eventually be translated into clinical improvements in GHVD diagnosis and potential for early detection of relapsed underlying disease.

References

Alcántara-Hernández, M. *et al.* (2017) 'High-Dimensional Phenotypic Mapping of Human Dendritic Cells Reveals Interindividual Variation and Tissue Specialization', *Immunity*, 47(6), pp. 1037-1050.e6. doi: 10.1016/j.immuni.2017.11.001.

Alexander, K. A. *et al.* (2014) 'CSF-1-dependant donor-derived macrophages mediate chronic graft-versus-host disease', *Journal of Clinical Investigation*, 124(10), pp. 4266–4280. doi: 10.1172/JCI75935.

Ali, R. *et al.* (2017) 'The role of anti-thymocyte globulin or alemtuzumab-based serotherapy in the prophylaxis and management of graft-versus-host disease', *Biomedicines*, 5(4). doi: 10.3390/biomedicines5040067.

Anasetti, C. *et al.* (2012) 'Peripheral-Blood Stem Cells versus Bone Marrow from Unrelated Donors', *New England Journal of Medicine*, 367(16), pp. 1487–1496. doi: 10.1056/nejmoa1203517.

Anderson, B. E. *et al.* (2003) 'Memory CD4+ T cells do not induce graft-versus-host disease', *Journal of Clinical Investigation*, 112(1), pp. 101–108. doi: 10.1172/JCI17601.

Azizi, E. *et al.* (2018) 'Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment', *Cell*. Elsevier Inc., 174(5), pp. 1293-1308.e36. doi: 10.1016/j.cell.2018.05.060.

Baker, M. B. *et al.* (1996) 'The role of cell-mediated cytotoxicity in acute GVHD after MHC-matched allogeneic bone marrow transplantation in mice', *Journal of Experimental Medicine*, 183(6), pp. 2645–2656. doi: 10.1084/jem.183.6.2645.

Banchereau, J. *et al.* (2000) 'Immunobiology of Dendritic Cells', *Annual Review of Immunology*, 18(1), pp. 767–811. doi: 10.1146/annurev.immunol.18.1.767.

Bao, M. and Liu, Y. J. (2013) 'Regulation of TLR7/9 signaling in plasmacytoid dendritic cells', *Protein and Cell*, 4(1), pp. 40–52. doi: 10.1007/s13238-012-2104-8.

Van Bekkum, D. W. *et al.* (1974) 'Mitigation of secondary disease of allogeneic mouse radiation chimeras by modification of the intestinal microflora', *Journal of the National Cancer Institute*, 52(2), pp. 401–404. doi: 10.1093/jnci/52.2.401.

Bigley, V. *et al.* (2011) 'The human syndrome of dendritic cell, monocyte, B and NK lymphoid deficiency', *Journal of Experimental Medicine*, 208(2), pp. 227–234. doi: 10.1084/jem.20101459.

Billingham, R. E. (1966) 'The biology of graft-versus-host reactions.', *Harvey lectures*, 62, pp. 21–78. Available at: http://www.ncbi.nlm.nih.gov/pubmed/4875305.

Bjorkman, P. J. *et al.* (1987) 'The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens', *Nature*, 329(6139), pp. 512–518. doi: 10.1038/329512a0.

Blazar, B. R., Hill, G. R. and Murphy, W. J. (2020) 'Dissecting the biology of allogeneic HSCT to enhance the GvT effect whilst minimizing GvHD', *Nature Reviews Clinical Oncology*. Springer US, 17(8), pp. 475–492. doi: 10.1038/s41571-020-0356-4.

Boada-Romero, E. *et al.* (2020) 'The clearance of dead cells by efferocytosis', *Nature Reviews Molecular Cell Biology*. Springer US, 21(7), pp. 398–414. doi: 10.1038/s41580-020-0232-1.

Borcherding, N. *et al.* (2021) 'Mapping the immune environment in clear cell renal carcinoma by single-cell genomics', *Communications Biology*. Springer US, 4(1), pp. 1–11. doi: 10.1038/s42003-020-01625-6.

Borcherding, N., Bormann, N. L. and Kraus, G. (2020) 'scRepertoire: An R-based toolkit for single-cell immune receptor analysis', *F1000Research*, 9, pp. 1–17. doi: 10.12688/f1000research.22139.2.

Bosteels, C. *et al.* (2020) 'Inflammatory Type 2 cDCs Acquire Features of cDC1s and Macrophages to Orchestrate Immunity to Respiratory Virus Infection', *Immunity*. Elsevier Inc., 52(6), pp. 1039-1056.e9. doi: 10.1016/j.immuni.2020.04.005.

Bourdely, P. *et al.* (2020) 'Transcriptional and Functional Analysis of CD1c+ Human Dendritic Cells Identifies a CD163+ Subset Priming CD8+CD103+ T Cells', *Immunity*, 53(2), pp. 335-352.e8. doi: 10.1016/j.immuni.2020.06.002.

Brandenburg, U., Gottlieb, D. and Bradstock, K. (1998) 'Antileukemic effects of rapid cyclosporin withdrawal in patients with relapsed chronic myeloid leukemia after allogeneic bone marrow transplantation', *Leukemia and Lymphoma*, 31(5–6), pp. 545–550. doi: 10.3109/10428199809057613.

Braun, M. Y. *et al.* (1996) 'Cytotoxic T cells deficient in both functional Fas ligand and perforin

show residual cytolytic activity yet lose their capacity to induce lethal acute graft-versus-host disease', *Journal of Experimental Medicine*, 183(2), pp. 657–661. doi: 10.1084/jem.183.2.657.

Brennecke, P. *et al.* (2013) 'Accounting for technical noise in single-cell RNA-seq experiments', *Nature Methods*, 10(11), pp. 1093–1098. doi: 10.1038/nmeth.2645.

Brickner, A. G. *et al.* (2001) 'The Immunogenicity of a New Human Minor Histocompatibility Antigen Results from Differential Antigen Processing', *The Journal of Experimental Medicine*, 193(2), pp. 195–205. doi: 10.1084/jem.193.2.195.

Bruce, D. W. *et al.* (2017) 'Type 2 innate lymphoid cells treat and prevent acute gastrointestinal graft-versus-host disease', *Journal of Clinical Investigation*, 127(5), pp. 1813–1825. doi: 10.1172/JCI91816.

de Bueger, M. *et al.* (1992) 'Tissue distribution of human minor histocompatibility antigens. Ubiquitous versus restricted tissue distribution indicates heterogeneity among human cytotoxic T lymphocyte-defined non-MHC antigens.', *The Journal of Immunology*, 149(5), pp. 1788–1794. doi: 10.4049/jimmunol.149.5.1788.

Bunting, M. D. *et al.* (2017) 'GVHD prevents NK-cell-dependent leukemia and virus-specific innate immunity', *Blood*, 129(5), pp. 630–642. doi: 10.1182/blood-2016-08-734020.

Butler, A. *et al.* (2018) 'Integrating single-cell transcriptomic data across different conditions, technologies, and species', *Nature Biotechnology*, 36(5), pp. 411–420. doi: 10.1038/nbt.4096.

Chakarov, S. *et al.* (2019) 'Two distinct interstitial macrophage populations coexist across tissues in specific subtissular niches', *Science*, 363(6432). doi: 10.1126/science.aau0964.

Chan, S. L., Wood, B. A. and Mesbah Ardakani, N. (2020) 'Eczematous graft-vs-host disease: A report of three cases and review of the literature', *Journal of Cutaneous Pathology*, 47(11), pp. 1085–1095. doi: 10.1111/cup.13836.

Chao, N. J. (2004) 'Minors come of age: Minor histocompatibility antigens and graft-versus-host disease', *Biology of Blood and Marrow Transplantation*, 10(4), pp. 215–223. doi: 10.1016/j.bbmt.2003.10.003.

Cheuk, D. K. (2013) 'Optimal stem cell source for allogeneic stem cell transplantation for hematological malignancies', *World Journal of Transplantation*, 3(4), p. 99. doi:

10.5500/wjt.v3.i4.99.

Collin, M. and Bigley, V. (2018) 'Human dendritic cell subsets: an update', *Immunology*, 154(1), pp. 3–20. doi: 10.1111/imm.12888.

Collin, M. and Milne, P. (2016) 'Langerhans cell origin and regulation', *Current Opinion in Hematology*, 23(1), pp. 28–35. doi: 10.1097/MOH.0000000000000202.

Collin, M. P. *et al.* (2006) 'The fate of human Langerhans cells in hematopoietic stem cell transplantation', *Journal of Experimental Medicine*, 203(1), pp. 27–33. doi: 10.1084/jem.20051787.

Cooley, S. *et al.* (2005) 'KIR reconstitution is altered by T cells in the graft and correlates with clinical outcomes after unrelated donor transplantation', *Blood*, 106(13), pp. 4370–4376. doi: 10.1182/blood-2005-04-1644.

Cuellar-Rodriguez, J. *et al.* (2011) 'Successful allogeneic hematopoietic stem cell transplantation for GATA2 deficiency', *Blood*, 118(13), pp. 3715–3720. doi: 10.1182/blood-2011-06-365049.

Cytlak, U. *et al.* (2020) 'Differential IRF8 Transcription Factor Requirement Defines Two Pathways of Dendritic Cell Development in Humans', *Immunity*. Elsevier Inc., 53(2), pp. 353-370.e8. doi: 10.1016/j.immuni.2020.07.003.

D'Aveni, M. *et al.* (2015) 'G-CSF mobilizes CD34+ regulatory monocytes that inhibit graft-versus-host disease', *Science Translational Medicine*, 7(281). doi: 10.1126/scitranslmed.3010435.

Dai, H. *et al.* (2021) 'Metallothionein 1: A New Spotlight on Inflammatory Diseases', *Frontiers in Immunology*, 12(November), pp. 13–21. doi: 10.3389/fimmu.2021.739918.

Divito, S. J. *et al.* (2020) 'Peripheral host T cells survive hematopoietic stem cell transplantation and promote graft-versus-host disease', *Journal of Clinical Investigation*, 130(9), pp. 4624–4636. doi: 10.1172/JCI129965.

Domínguez Conde, C. *et al.* (2022) 'Cross-tissue immune cell analysis reveals tissue-specific features in humans', *Science*, 376(6594). doi: 10.1126/science.abl5197.

Dutertre, C. A. *et al.* (2019) 'Single-Cell Analysis of Human Mononuclear Phagocytes Reveals Subset-Defining Markers and Identifies Circulating Inflammatory Dendritic Cells', *Immunity*.

Elsevier Inc., 51(3), pp. 573-589.e8. doi: 10.1016/j.immuni.2019.08.008.

Eberl, G. *et al.* (2015) 'Innate lymphoid cells: A new paradigm in immunology', *Science*, 348(6237). doi: 10.1126/science.aaa6566.

Efremova, M. and Teichmann, S. A. (2020) 'Computational methods for single-cell omics across modalities', *Nature Methods*, 17(1), pp. 14–17. doi: 10.1038/s41592-019-0692-4.

Engel, J. A. *et al.* (2020) 'Single-cell transcriptomics of alloreactive CD4+ T cells over time reveals divergent fates during gut graft-versus-host disease', *JCI Insight*, 5(13), pp. 1–19. doi: 10.1172/jci.insight.137990.

Evren, E. *et al.* (2021) 'Distinct developmental pathways from blood monocytes generate human lung macrophage diversity', *Immunity*, 54(2), pp. 259-275.e7. doi: 10.1016/j.immuni.2020.12.003.

Farber, D. L., Yudanin, N. A. and Restifo, N. P. (2014) 'Human memory T cells: generation, compartmentalization and homeostasis', *Nature Reviews Immunology*, 14(1), pp. 24–35. doi: 10.1038/nri3567.

Ferrara, J. L. *et al.* (2009) 'Graft-versus-host disease', *The Lancet*. Elsevier Ltd, 373(9674), pp. 1550–1561. doi: 10.1016/S0140-6736(09)60237-3.

Ferrara, J. L. and Deeg, H. J. (1991) 'Graft-versus-host disease.', *The New England journal of medicine*, 324(10), pp. 667–74. doi: 10.1056/NEJM199103073241005.

Filipovich, A. H. *et al.* (2005) 'National Institutes of Health Consensus Development Project on criteria for clinical trials in chronic graft-versus-host disease: I. diagnosis and staging working group report', *Biology of Blood and Marrow Transplantation*, 11(12), pp. 945–956. doi: 10.1016/j.bbmt.2005.09.004.

Flanagan, W. M. *et al.* (1991) 'Nuclear association of a T-cell transcription factor blocked by FK-506 and cyclosporin A', *Nature*, 352(6338), pp. 803–807. doi: 10.1038/352803a0.

van Furth, R. and Cohn, Z. A. (1968) 'THE ORIGIN AND KINETICS OF MONONUCLEAR PHAGOCYTES', *Journal of Experimental Medicine*, 128(3), pp. 415–435. doi: 10.1084/jem.128.3.415.

Galli, S. J. and Tsai, M. (2012) 'IgE and mast cells in allergic disease', *Nature Medicine*, 18(5), pp. 693–704. doi: 10.1038/nm.2755.

Garside, P. *et al.* (1992) 'Nitric oxide mediates intestinal pathology in graft-vs.-host disease', *European Journal of Immunology*, 22(8), pp. 2141–2145. doi: 10.1002/eji.1830220827.

Gartlan, K. H. *et al.* (2019) 'Donor T-cell–derived GM-CSF drives alloantigen presentation by dendritic cells in the gastrointestinal tract', *Blood Advances*, 3(19), pp. 2859–2865. doi: 10.1182/bloodadvances.2019000053.

Gilliam, A. C. *et al.* (1996) 'Apoptosis is the predominant form of epithelial target cell injury in acute experimental graft-versus-host disease', *Journal of Investigative Dermatology*. Elsevier Masson SAS, 107(3), pp. 377–383. doi: 10.1111/1523-1747.ep12363361.

Ginhoux, F. and Guilliams, M. (2016) 'Tissue-Resident Macrophage Ontogeny and Homeostasis', *Immunity*. Elsevier Inc., 44(3), pp. 439–449. doi: 10.1016/j.immuni.2016.02.024.

Gratwohl, A. *et al.* (2013) 'Quantitative and qualitative differences in use and trends of hematopoietic stem cell transplantation: A Global Observational Study', *Haematologica*, 98(8), pp. 1282–1290. doi: 10.3324/haematol.2012.076349.

Gutcher, I. and Becher, B. (2007) 'APC-derived cytokines and T cell polarization in autoimmune inflammation', *Journal of Clinical Investigation*, 117(5), pp. 1119–1127. doi: 10.1172/JCI31720.

Hafemeister, C. and Satija, R. (2019) 'Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression', *Genome Biology*. Genome Biology, 20(1), pp. 1–15. doi: 10.1186/s13059-019-1874-1.

Hanash, A. M. *et al.* (2012) 'Interleukin-22 Protects Intestinal Stem Cells from Immune-Mediated Tissue Damage and Regulates Sensitivity to Graft versus Host Disease', *Immunity*, 37(2), pp. 339–350. doi: 10.1016/j.immuni.2012.05.028.

Haniffa, M. *et al.* (2009) 'Differential rates of replacement of human dermal dendritic cells and macrophages during hematopoietic stem cell transplantation', *Journal of Experimental Medicine*, 206(2), pp. 371–385. doi: 10.1084/jem.20081633.

Haniffa, M. *et al.* (2012) 'Human Tissues Contain CD141 hi Cross-Presenting Dendritic Cells with Functional Homology to Mouse CD103 + Nonlymphoid Dendritic Cells', *Immunity*, 37(1), pp. 60–73. doi: 10.1016/j.immuni.2012.04.012.

Hanna, R. N. *et al.* (2011) 'The transcription factor NR4A1 (Nur77) controls bone marrow differentiation and the survival of Ly6C- monocytes', *Nature Immunology*, 12(8), pp. 778–785. doi: 10.1038/ni.2063.

Haque, A. *et al.* (2017) 'A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications', *Genome Medicine*. Genome Medicine, 9(1), pp. 1–12. doi: 10.1186/s13073-017-0467-4.

Hashimoto, D. *et al.* (2011) 'Pretransplant CSF-1 therapy expands recipient macrophages and ameliorates GVHD after allogeneic hematopoietic cell transplantation', *Journal of Experimental Medicine*, 208(5), pp. 1069–1082. doi: 10.1084/jem.20101709.

Hashimoto, D. *et al.* (2013) 'Tissue-resident macrophages self-maintain locally throughout adult life with minimal contribution from circulating monocytes', *Immunity*, 38(4), pp. 792–804. doi: 10.1016/j.immuni.2013.04.004.

Hatakeyama, M. *et al.* (2017) 'Anti-Inflammatory Role of Langerhans Cells and Apoptotic Keratinocytes in Ultraviolet-B–Induced Cutaneous Inflammation', *The Journal of Immunology*, 199(8), pp. 2937–2947. doi: 10.4049/jimmunol.1601681.

Hill, G. and Ferrara, J. (2000) 'The primacy of the gastrointestinal tract as a targer organ of acute graft-versus-host disease: rationale for the use of cytokine shields in allogeneic bone marrow transplantation', *Blood*, 95(9), pp. 2754–2759.

Hill, G. R. *et al.* (1997) 'Total body irradiation and acute graft-versus-host disease: the role of gastrointestinal damage and inflammatory cytokines', *Blood*, 90(8), pp. 3204–13. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9376604.

Hill, G. R. *et al.* (2021) 'Current Concepts and Advances in Graft-Versus-Host Disease Immunology', *Annual Review of Immunology*, 39(1), pp. 19–49. doi: 10.1146/annurev-immunol-102119-073227.

Ho, V. T. and Soiffer, R. J. (2001) 'The history and future of T-cell depletion as graft-versus-host disease prophylaxis for allogeneic hematopoietic stem cell transplantation Review article The history and future of T-cell depletion as graft-versus-host disease prophylaxis for allogeneic he', *Blood*, 98(12), pp. 3192–3204. doi: 10.1182/blood.V98.12.3192.

Hoffmann, P. *et al.* (2002) 'Donor-type CD4+CD25+ regulatory T cells suppress lethal acute

graft-versus-host disease after allogeneic bone marrow transplantation', *Journal of Experimental Medicine*, 196(3), pp. 389–399. doi: 10.1084/jem.20020399.

Holler, E. *et al.* (1990) 'Increased serum levels of tumor necrosis factor alpha precede major complications of bone marrow transplantation', *Blood*, 75(4), pp. 1011–6. doi: 10.1182/blood.v75.4.1011.1011.

Huang, Y., McCarthy, D. J. and Stegle, O. (2019) 'Vireo: Bayesian demultiplexing of pooled single-cell RNA-seq data without genotype reference', *Genome Biology*. Genome Biology, 20(1), pp. 1–12. doi: 10.1186/s13059-019-1865-2.

Ijaz, A. *et al.* (2019) 'Significant Risk of Graft-versus-Host Disease with Exposure to Checkpoint Inhibitors before and after Allogeneic Transplantation', *Biology of Blood and Marrow Transplantation*. Elsevier Inc., 25(1), pp. 94–99. doi: 10.1016/j.bbmt.2018.08.028.

Iwasaki, A. and Medzhitov, R. (2004) 'Toll-like receptor control of the adaptive immune responses', *Nature Immunology*, 5(10), pp. 987–995. doi: 10.1038/ni1112.

Jardine, L. *et al.* (2020) 'Donor monocyte-derived macrophages promote human acute graft-versus-host disease', *Journal of Clinical Investigation*, 130(9), pp. 4574–4586. doi: 10.1172/JCI133909.

Johnston, L. J., Halliday, G. M. and King, N. J. C. (2000) 'Langerhans cells migrate to local lymph nodes following cutaneous infection with an arbovirus', *Journal of Investigative Dermatology*. Elsevier Masson SAS, 114(3), pp. 560–568. doi: 10.1046/j.1523-1747.2000.00904.x.

de Jong, A. and Ogg, G. (2021) 'CD1a function in human skin disease', *Molecular Immunology*, 130, pp. 14–19. doi: 10.1016/j.molimm.2020.12.006.

Jongbloed, S. L. *et al.* (2010) 'Human CD141+ (BDCA-3)+ dendritic cells (DCs) represent a unique myeloid DC subset that cross-presents necrotic cell antigens', *Journal of Experimental Medicine*, 207(6), pp. 1247–1260. doi: 10.1084/jem.20092140.

Jovic, D. *et al.* (2022) 'Single-cell RNA sequencing technologies and applications: A brief overview', *Clinical and Translational Medicine*, 12(3). doi: 10.1002/ctm2.694.

Kanda, J. *et al.* (2012) 'Impact of graft-versus-host disease on outcomes after allogeneic hematopoietic cell transplantation for adult T-cell leukemia: A retrospective cohort study',

Blood, 119(9), pp. 2141–2148. doi: 10.1182/blood-2011-07-368233.

Kang, H. M. *et al.* (2018) 'Multiplexed droplet single-cell RNA-sequencing using natural genetic variation', *Nature Biotechnology*, 36(1), pp. 89–94. doi: 10.1038/nbt.4042.

Kawase, T. *et al.* (2007) 'High-risk HLA allele mismatch combinations responsible for severe acute graft-versus-host disease and implication for its molecular mechanism', *Blood*, 110(7), pp. 2235–2241. doi: 10.1182/blood-2007-02-072405.

Klämbt, V. *et al.* (2015) 'A Novel Function for P2Y2 in Myeloid Recipient–Derived Cells during Graft-versus-Host Disease', *The Journal of Immunology*, 195(12), pp. 5795–5804. doi: 10.4049/jimmunol.1501357.

Klein, A. M. *et al.* (2015) 'Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells', *Cell*. Elsevier Inc., 161(5), pp. 1187–1201. doi: 10.1016/j.cell.2015.04.044.

Kolb, H. J. *et al.* (1995) 'Graft-versus-leukemia effect of donor lymphocyte transfusions in marrow grafted patients', *Blood*, 86(5), pp. 2041–2050. doi: 10.1182/blood.v86.5.2041.bloodjournal8652041.

Korngold, R. and Sprent, J. (1978) 'Lethal graft-versus-host disease after bone marrow transplantation across minor histocompatibility barriers in mice: Prevention by removing mature T cells from marrow*', *Journal of Experimental Medicine*, 148(6), pp. 1687–1698. doi: 10.1084/jem.148.6.1687.

Korngold, R. and Sprent, J. (1982) 'Features of T cells causing H-2-restricted lethal graft-vs.-host disease across minor histocompatibility barriers', *Journal of Experimental Medicine*, 155(3), pp. 872–883. doi: 10.1084/jem.155.3.872.

Korngold, R. and Sprent, J. (1985) 'Surface markers of T cells causing lethal graft-vs-host disease to class I vs class II H-2 differences.', *The Journal of Immunology*, 135(5), pp. 3004–3010. doi: 10.4049/jimmunol.135.5.3004.

Koyama, M. *et al.* (2012) 'Recipient nonhematopoietic antigen-presenting cells are sufficient to induce lethal acute graft-versus-host disease', *Nature Medicine*, 18(1), pp. 135–142. doi: 10.1038/nm.2597.

Koyama, M. *et al.* (2015) 'Donor colonic CD103+ dendritic cells determine the severity of

acute graft-versus-host disease', *Journal of Experimental Medicine*, 212(8), pp. 1303–1321. doi: 10.1084/jem.20150329.

Koyama, M. *et al.* (2019) 'MHC Class II Antigen Presentation by the Intestinal Epithelium Initiates Graft-versus-Host Disease and Is Influenced by the Microbiota', *Immunity*, 51(5), pp. 885-898.e7. doi: 10.1016/j.immuni.2019.08.011.

Koyama, M. and Hill, G. R. (2019) 'The primacy of gastrointestinal tract antigen-presenting cells in lethal graft-versus-host disease', *Blood*, 134(24), pp. 2139–2148. doi: 10.1182/blood.2019000823.

Krangel, M. S. (2009) 'Mechanics of T cell receptor gene rearrangement', *Current Opinion in Immunology*, 21(2), pp. 133–139. doi: 10.1016/j.coi.2009.03.009.

Krenger, W. and Holländer, G. A. (2008) 'The immunopathology of thymic GVHD', *Seminars in Immunopathology*, 30(4), pp. 439–456. doi: 10.1007/s00281-008-0131-6.

Kroeze, A. *et al.* (2022) 'Presence of innate lymphoid cells in allogeneic hematopoietic grafts correlates with reduced graft-versus-host disease', *Cytotherapy*, 24(3), pp. 302–310. doi: 10.1016/j.jcyt.2021.10.011.

Kubo, A. *et al.* (2009) 'External antigen uptake by Langerhans cells with reorganization of epidermal tight junction barriers', *Journal of Experimental Medicine*, 206(13), pp. 2937–2946. doi: 10.1084/jem.20091527.

Kukurba, K. R. and Montgomery, S. B. (2015) 'RNA Sequencing and Analysis', *Cold Spring Harbor Protocols*, 2015(11), p. pdb.top084970. doi: 10.1101/pdb.top084970.

Kumar, B. V, Connors, T. J. and Farber, D. L. (2018) 'Human T Cell Development, Localization, and Function throughout Life', *Immunity*, 48(2), pp. 202–213. doi: 10.1016/j.immuni.2018.01.007.

Lauterbach, H. *et al.* (2010) 'Mouse CD8α+ DCs and human BDCA3+ DCs are major producers of IFN-λ in response to poly IC', *Journal of Experimental Medicine*, 207(12), pp. 2703–2717. doi: 10.1084/jem.20092720.

Leveque-El Mouttie, L. *et al.* (2016) 'Corruption of dendritic cell antigen presentation during acute GVHD leads to regulatory T-cell failure and chronic GVHD', *Blood*, 128(6), pp. 794–804. doi: 10.1182/blood-2015-11-680876.

Li, H. *et al.* (2011) 'Langerhans cells are not required for graft-versus-host disease', *Blood*, 117(2), pp. 697–707. doi: 10.1182/blood-2010-07-299073.

Li, H. *et al.* (2012) 'Profound Depletion of Host Conventional Dendritic Cells, Plasmacytoid Dendritic Cells, and B Cells Does Not Prevent Graft-versus-Host Disease Induction', *The Journal of Immunology*, 188(8), pp. 3804–3811. doi: 10.4049/jimmunol.1102795.

Li, M. *et al.* (2022) 'DISCO: A database of Deeply Integrated human Single-Cell Omics data', *Nucleic Acids Research*. Oxford University Press, 50(D1), pp. D596–D602. doi: 10.1093/nar/gkab1020.

Luecken, M. D. and Theis, F. J. (2019) 'Current best practices in single-cell RNA-seq analysis: a tutorial', *Molecular Systems Biology*, 15(6). doi: 10.15252/msb.20188746.

van der Maaten, L. and Hinton, G. (2008) 'Visualizing Data using t-SNE', *Journal of Machine Learning Research*, 9(nov), pp. 2579–2605. doi: 10.1007/s10479-011-0841-3.

MacDonald, K. P. A. *et al.* (2010) 'An antibody against the colony-stimulating factor 1 receptor depletes the resident subset of monocytes and tissue- and tumor-associated macrophages but does not inhibit inflammation', *Blood*, 116(19), pp. 3955–3963. doi: 10.1182/blood-2010-02-266296.

Maier, B. *et al.* (2020) 'A conserved dendritic-cell regulatory program limits antitumour immunity', *Nature*, 580(7802), pp. 257–262. doi: 10.1038/s41586-020-2134-y.

Marius Munneke, J. *et al.* (2014) 'Activated innate lymphoid cells are associated with a reduced susceptibility to graft-versus-host disease', *Blood*, 124(5), pp. 812–821. doi: 10.1182/blood-2013-11-536888.

Markey, K. A. *et al.* (2012) 'Immune insufficiency during GVHD is due to defective antigen presentation within dendritic cell subsets', *Blood*, 119(24), pp. 5918–5930. doi: 10.1182/blood-2011-12-398164.

Matte-Martone, C. *et al.* (2008) 'CD8+ but not CD4+ T cells require cognate interactions with target tissues to mediate GVHD across only minor H antigens, whereas both CD4+ and CD8+ T cells require direct leukemic contact to mediate GVL', *Blood*, 111(7), pp. 3885–3892. doi: 10.1182/blood-2007-11-125294.The.

Matte, C. C. *et al.* (2004) 'Donor APCs are required for maximal GVHD but not for GVL',

Nature medicine, 10(9), pp. 987–992. doi: 10.1038/nm1089.

McGinnis, C. S., Murrow, L. M. and Gartner, Z. J. (2019) 'DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors', *Cell Systems*. Elsevier Inc., 8(4), pp. 329-337.e4. doi: 10.1016/j.cels.2019.03.003.

McGovern, N. *et al.* (2014) 'Human dermal CD14+ cells are a transient population of monocyte-derived macrophages', *Immunity*, 41(3), pp. 465–477. doi: 10.1016/j.immuni.2014.08.006.

McInnes, L., Healy, J. and Melville, J. (2018) 'UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction'. Available at: http://arxiv.org/abs/1802.03426.

Mielcarek, M. *et al.* (2014) 'Langerhans cell homeostasis and turnover after nonmyeloablative and myeloablative allogeneic hematopoietic cell transplantation', *Transplantation*, 98(5), pp. 563–568. doi: 10.1097/TP.0000000000000097.

Mills, C. D. *et al.* (2000) 'M-1/M-2 Macrophages and the Th1/Th2 Paradigm', *The Journal of Immunology*, 164(12), pp. 6166–6173. doi: 10.4049/jimmunol.164.12.6166.

Mizumoto, N. *et al.* (2002) 'CD39 is the dominant Langerhans cell-associated ecto-NTPDase: Modulatory roles in inflammation and immune responsiveness', *Nature Medicine*, 8(4), pp. 358–365. doi: 10.1038/nm0402-358.

Mulder, K. *et al.* (2021) 'Cross-tissue single-cell landscape of human monocytes and macrophages in health and disease', *Immunity*, 54(8), pp. 1883-1900.e5. doi: 10.1016/j.immuni.2021.07.007.

Murata, M., Warren, E. H. and Riddell, S. R. (2003) 'A Human Minor Histocompatibility Antigen Resulting from Differential Expression due to a Gene Deletion', *The Journal of Experimental Medicine*, 197(10), pp. 1279–1289. doi: 10.1084/jem.20030044.

Nakamizo, S. *et al.* (2021) 'Single-cell analysis of human skin identifies CD14+ type 3 dendritic cells co-producing IL1B and IL23A in psoriasis', *Journal of Experimental Medicine*, 218(9). doi: 10.1084/jem.20202345.

Narasimhan, P. B. *et al.* (2019) 'Nonclassical Monocytes in Health and Disease', *Annual Review of Immunology*, 37, pp. 439–456. doi: 10.1146/annurev-immunol-042617-053119.

Nestel, F. P. *et al.* (1992) 'Macrophage priming and lipopolysaccharide-triggered release of

tumor necrosis factor α during graft-versus-host disease', *Journal of Experimental Medicine*, 175(2), pp. 405–413. doi: 10.1084/jem.175.2.405.

Nestle, F. O. *et al.* (2005) 'Plasmacytoid predendritic cells initiate psoriasis through interferon-α production', *Journal of Experimental Medicine*, 202(1), pp. 135–143. doi: 10.1084/jem.20050500.

Nishiwaki, S. *et al.* (2009) 'Impact of macrophage infiltration of skin lesions on survival after allogeneic stem cell transplantation: A clue to refractory graft-versus-host disease', *Blood*, 114(14), pp. 3113–3116. doi: 10.1182/blood-2009-03-209635.

Olingy, C. E. *et al.* (2017) 'Non-classical monocytes are biased progenitors of wound healing macrophages during soft tissue injury', *Scientific Reports*. Springer US, 7(1), pp. 1–16. doi: 10.1038/s41598-017-00477-1.

Olson, J. A. *et al.* (2010) 'NK cells mediate reduction of GVHD by inhibiting activated, alloreactive T cells while retaining GVT effects', *Blood*, 115(21), pp. 4293–4301. doi: 10.1182/blood-2009-05-222190.

Parekh, S. *et al.* (2018) 'zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs', *GigaScience*. Oxford University Press, 7(6), pp. 1–9. doi: 10.1093/gigascience/giy059.

Park, C. O. and Kupper, T. S. (2015) 'The emerging role of resident memory T cells in protective immunity and inflammatory disease', *Nature Medicine*, 21(7), pp. 688–697. doi: 10.1038/nm.3883.

Patente, T. A., Pelgrom, L. R. and Everts, B. (2019) 'Dendritic cells are what they eat: how their metabolism shapes T helper cell polarization', *Current Opinion in Immunology*. Elsevier Ltd, 58, pp. 16–23. doi: 10.1016/j.coi.2019.02.003.

Perreault, C. *et al.* (1996) 'Identification of an immunodominant mouse minor histocompatibility antigen (MiHA): T cell response to a single dominant MiHA causes graft-versus-host disease', *Journal of Clinical Investigation*, 98(3), pp. 622–628. doi: 10.1172/JCI118832.

Petersdorf, E. W. *et al.* (1995) 'The significance of HLA-DRB1 matching on clinical outcome after HLA-A, B, DR identical unrelated donor marrow transplantation', *Blood*. American

Society of Hematology, 86(4), pp. 1606–1613. doi: 10.1182/blood.v86.4.1606.bloodjournal8641606.

Piguet, P. F. *et al.* (1987) 'Tumor necrosis factor/cachectin is an effector of skin and gut lesions of the acute phase of graft-vs.-host disease', *Journal of Experimental Medicine*, 166(5), pp. 1280–1289. doi: 10.1084/jem.166.5.1280.

Piper, C. *et al.* (2022) 'Single-cell immune profiling reveals a developmentally distinct CD41 GM-CSF1 T-cell lineage that induces GI tract GVHD', *Blood Advances*, 6(9), pp. 2791–2804. doi: 10.1182/bloodadvances.2021006084.

Piperoglou, C. *et al.* (2022) 'Innate lymphoid cell recovery and occurrence of GvHD after hematopoietic stem cell transplantation', *Journal of Leukocyte Biology*, 111(1), pp. 161–172. doi: 10.1002/JLB.5A1019-522RR.

Poli, A. *et al.* (2009) 'CD56bright natural killer (NK) cells: An important NK cell subset', *Immunology*, 126(4), pp. 458–465. doi: 10.1111/j.1365-2567.2008.03027.x.

Radulovic, K. *et al.* (2013) 'The Early Activation Marker CD69 Regulates the Expression of Chemokines and CD4 T Cell Accumulation in Intestine', *PLoS ONE*, 8(6). doi: 10.1371/journal.pone.0065413.

Reddy, V. *et al.* (2004) 'Low dendritic cell count after allogeneic hematopoietic stem cell transplantation predicts relapse, death, and acute graft-versus-host disease', *Blood*, 103(11), pp. 4330–4335. doi: 10.1182/blood-2003-09-3325.

Reinhardt-Heller, K. *et al.* (2017) 'Increase of Intermediate Monocytes in Graft-versus-Host Disease: Correlation with MDR1+Th17.1 Levels and the Effect of Prednisolone and $1\alpha,25$-Dihydroxyvitamin D3', *Biology of Blood and Marrow Transplantation*. Elsevier Inc., 23(12), pp. 2057–2064. doi: 10.1016/j.bbmt.2017.08.008.

Reinhardt-Heller, K. *et al.* (2018) 'Characterization of monocyte subtypes regarding their phenotype and development in the context of graft-versus-host disease', *Transplant Immunology*, 50(October 2018), pp. 48–54. doi: 10.1016/j.trim.2018.06.004.

Reinhardt, K. *et al.* (2014) 'Monocyte-Induced Development of Th17 Cells and the Release of S100 Proteins Are Involved in the Pathogenesis of Graft-versus-Host Disease', *The Journal of Immunology*, 193(7), pp. 3355–3365. doi: 10.4049/jimmunol.1400983.

Reynolds, G. *et al.* (2021) 'Developmental cell programs are co-opted in inflammatory', *Science*, 371(6527), p. eaba6500. doi: 10.1126/science.aba6500.

Roex, M. C. J. *et al.* (2021) 'Effect of alemtuzumab-based T-cell depletion on graft compositional change in vitro and immune reconstitution early after allogeneic stem cell transplantation', *Cytotherapy*. Elsevier Inc., 23(1), pp. 46–56. doi: 10.1016/j.jcyt.2020.08.003.

Rowe, V. *et al.* (2006) 'Host B cells produce IL-10 following TBI and attenuate acute GVHD after allogeneic bone marrow transplantation', *Blood*, 108(7), pp. 2485–2492. doi: 10.1182/blood-2006-04-016063.

Ruggeri, L. *et al.* (2002) 'Effectiveness of Donor Natural Killer Cell Alloreactivity in Mismatched Hematopoietic Transplants', *Science*, 295(March), pp. 2097–2100.

Sandmaier, B. M. *et al.* (2019) 'Addition of sirolimus to standard cyclosporine plus mycophenolate mofetil-based graft-versus-host disease prophylaxis for patients after unrelated non-myeloablative haemopoietic stem cell transplantation: a multicentre, randomised, phase 3 trial', *The Lancet Haematology*, 6(8), pp. e409–e418. doi: 10.1016/S2352-3026(19)30088-2.

Santos E Sousa, P. *et al.* (2018) 'Peripheral tissues reprogram CD8+ T cells for pathogenicity during graft-versus-host disease', *JCI insight*, 3(5), pp. 1–19. doi: 10.1172/jci.insight.97011.

Santos e Sousa, P., Bennett, C. L. and Chakraverty, R. (2018) 'Unraveling the mechanisms of cutaneous graft-versus-host disease', *Frontiers in Immunology*, 9(MAY). doi: 10.3389/fimmu.2018.00963.

Schlitzer, A. *et al.* (2013) 'IRF4 Transcription Factor-Dependent CD11b+ Dendritic Cells in Human and Mouse Control Mucosal IL-17 Cytokine Responses', *Immunity*. Elsevier Inc., 38(5), pp. 970–983. doi: 10.1016/j.immuni.2013.04.011.

Shah, N. N. *et al.* (2015) 'Acute GVHD in patients receiving IL-15/4-1BBL activated NK cells following T-cell-depleted stem cell transplantation', *Blood*, 125(5), pp. 784–792. doi: 10.1182/blood-2014-07-592881.

Shapouri-Moghaddam, A. *et al.* (2018) 'Macrophage plasticity, polarization, and function in health and disease', *Journal of Cellular Physiology*, 233(9), pp. 6425–6440. doi: 10.1002/jcp.26429.

Shaw, J.-P. *et al.* (1988) 'Identification of a Putative Regulator of Early T Cell Activation Genes', *Science*, 241(4862), pp. 202–205. doi: 10.1126/science.3260404.

Sheng, L. *et al.* (2020) 'Cytotoxicity of Donor Natural Killer Cells to Allo-Reactive T Cells Are Related With Acute Graft-vs.-Host-Disease Following Allogeneic Stem Cell Transplantation', *Frontiers in Immunology*, 11(July), pp. 1–13. doi: 10.3389/fimmu.2020.01534.

Shin, J. S. and Greer, A. M. (2015) 'The role of FcεRI expressed in dendritic cells and monocytes', *Cellular and Molecular Life Sciences*. Springer Basel, 72(12), pp. 2349–2360. doi: 10.1007/s00018-015-1870-x.

Shlomchik, W. D. *et al.* (1999) 'Prevention of Graft Versus Host Disease by Inactivation of Host Antigen-Presenting Cells', *Science*, 285(July), pp. 412–415.

Shlomchik, W. D. (2007) 'Graft-versus-host disease', *Nature Reviews Immunology*, 7, pp. 340–352. doi: 10.1201/9781315160498.

Simmons, D. P. *et al.* (2022) 'SLAMF7 engagement superactivates macrophages in acute and chronic inflammation', *Science Immunology*, 7(68), pp. 139–148. doi: 10.1126/sciimmunol.abf2846.

Simons, L., Cavazzana, M. and André, I. (2019) 'Concise Review: Boosting T-Cell Reconstitution Following Allogeneic Transplantation—Current Concepts and Future Perspectives', *Stem Cells Translational Medicine*, 8(7), pp. 650–657. doi: 10.1002/sctm.18-0248.

Snowden, J. A. *et al.* (2022) 'Indications for haematopoietic cell transplantation for haematological diseases, solid tumours and immune disorders: current practice in Europe, 2022', *Bone Marrow Transplantation*. Springer US, 57(8), pp. 1217–1239. doi: 10.1038/s41409-022-01691-w.

Song, Y. *et al.* (2021) 'Reduced Intensity Conditioning Followed by Allogeneic Hematopoietic Stem Cell Transplantation Is a Good Choice for Acute Myeloid Leukemia and Myelodysplastic Syndrome: A Meta-Analysis of Randomized Controlled Trials', *Frontiers in Oncology*, 11(October), pp. 1–11. doi: 10.3389/fonc.2021.708727.

Sprent, J. *et al.* (1986) 'Properties of purified T cell subsets: II. in vivo responses to class I VS. class II H-2 differences', *Journal of Experimental Medicine*, 163(4), pp. 998–1011. doi:

10.1084/jem.163.4.998.

Steinman, R. M., Hawiger, D. and Nussenzweig, M. C. (2003) 'Tolerogenic dendritic cells', *Annual Review of Immunology*, 21, pp. 685–711. doi: 10.1146/annurev.immunol.21.120601.141040.

Stern, L. *et al.* (2018) 'Mass Cytometry for the assessment of immune reconstitution after hematopoietic stem cell transplantation', *Frontiers in Immunology*, 9(JUL). doi: 10.3389/fimmu.2018.01672.

Storek, J. *et al.* (2001) 'Factors influencing B lymphopoiesis after allogeneic hematopoietic cell transplantation', *Blood*, 98(2), pp. 489–491. doi: 10.1182/blood.V98.2.489.

Street, K. *et al.* (2018) 'Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics.', *BMC genomics*. BMC Genomics, 19(1), p. 477. doi: 10.1186/s12864-018-4772-0.

Strobl, J. *et al.* (2020) 'Long-term skin-resident memory T cells proliferate in situ and are involved in human graft-versus-host disease', *Science Translational Medicine*, 12(570), pp. 1–18. doi: 10.1126/SCITRANSLMED.ABB7028.

Strobl, J. *et al.* (2021) 'Human resident memory t cells exit the skin and mediate systemic th2-driven inflammation', *Journal of Experimental Medicine*, 218(11). doi: 10.1084/jem.20210417.

Strong, K. *et al.* (2018) 'Cutaneous Graft-Versus-Host Disease : Diagnosis and Treatment', *American Journal of Clinical Dermatology*. Springer International Publishing, 19, pp. 33–50. doi: 10.1007/s40257-017-0306-9.

Stuart, T. *et al.* (2019) 'Comprehensive Integration of Single-Cell Data', *Cell*. Elsevier Inc., 177(7), pp. 1888-1902.e21. doi: 10.1016/j.cell.2019.05.031.

Styczyński, J. *et al.* (2020) 'Death after hematopoietic stem cell transplantation: changes over calendar year time, infections and associated factors', *Bone Marrow Transplantation*, 55(1), pp. 126–136. doi: 10.1038/s41409-019-0624-z.

Subramanian Vignesh, K. *et al.* (2013) 'Granulocyte macrophage-colony stimulating factor induced Zn sequestration enhances macrophage superoxide and limits intracellular pathogen survival', *Immunity*. Elsevier Inc., 39(4), pp. 697–710. doi: 10.1016/j.immuni.2013.09.006.

Tang, A. *et al.* (1993) 'Adhesion of epidermal Langerhans cells to keratinocytes mediated by E-cadherin', *Nature*, 361(6407), pp. 82–85. doi: 10.1038/361082a0.

Tang, F. *et al.* (2009) 'mRNA-Seq whole-transcriptome analysis of a single cell', *Nature Methods*, 6(5), pp. 377–382. doi: 10.1038/nmeth.1315.

Teshima, T. *et al.* (2002) 'Acute graft-versus-host disease does not require alloantigen expression on host epithelium.', *Nature medicine*, 8(6), pp. 575–81. doi: 10.1038/nm0602-575.

Theocharidis, G. *et al.* (2022) 'Single-cell transcriptomics in human skin research: available technologies, technical considerations and disease applications', *Experimental Dermatology*, 31(5), pp. 655–673. doi: 10.1111/exd.14547.

Tian, Y. *et al.* (2021) 'Graft-versus-host disease depletes plasmacytoid dendritic cell progenitors to impair tolerance induction', *Journal of Clinical Investigation*, 131(1). doi: 10.1172/JCI136774.

Tkachev, V. *et al.* (2021) 'Spatiotemporal single-cell profiling reveals that invasive and tissue-resident memory donor CD8+ T cells drive gastrointestinal acute graft-versus-host disease', *Science Translational Medicine*, 13(576). doi: 10.1126/SCITRANSLMED.ABC0227.

Toubai, T. *et al.* (2012) 'Induction of acute GVHD by sex-mismatched H-Y antigens in the absence of functional radiosensitive host hematopoietic-derived antigen-presenting cells', *Blood*, 119(16), pp. 3844–3853. doi: 10.1182/blood-2011-10-384057.

Trapnell, C. *et al.* (2014) 'The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells', *Nature Biotechnology*. Nature Publishing Group, 32(4), pp. 381–386. doi: 10.1038/nbt.2859.

Tugues, S. *et al.* (2018) 'Graft-versus-host disease, but not graft-versus-leukemia immunity, is mediated by GM-CSF–licensed myeloid cells', *Science Translational Medicine*, 10(469), p. eaat8410. doi: 10.1126/scitranslmed.aat8410.

Valladeau, J. *et al.* (2000) 'Langerin, a Novel C-Type Lectin Specific to Langerhans Cells, Is an Endocytic Receptor that Induces the Formation of Birbeck Granules cross-priming, a mechanism by which exogenous anti-gen is alternatively routed into the MHC class I pathway for presentat', *Immunity*, 12, pp. 71–81.

Villani, A.-C. *et al.* (2017) 'Single-cell RNA-Seq reveals new types of human blood dendritic cells, monocytes, and progenitors', *Science*, 356(6335), p. eaah4573. doi: 10.1097/TP.0000000000001890.

Vossen, J. M. *et al.* (2014) 'Complete suppression of the gut microbiome prevents acute graft-versus-host disease following allogeneic bone marrow transplantation', *PLoS ONE*, 9(9). doi: 10.1371/journal.pone.0105706.

Wang, Y. *et al.* (2020) 'Dendritic cell biology and its role in tumor immunotherapy', *Journal of Hematology and Oncology*, 13(1). doi: 10.1186/s13045-020-00939-6.

Wang, Z., Gerstein, M. and Snyder, M. (2009) 'RNA-Seq: a revolutionary tool for transcriptomics', *Nature Reviews Genetics*, 10(1), pp. 57–63. doi: 10.1038/nrg2484.

Weiss, G. and Schaible, U. E. (2015) 'Macrophage defense mechanisms against intracellular bacteria', *Immunological Reviews*, 264(1), pp. 182–203. doi: 10.1111/imr.12266.

Whitehouse, G. *et al.* (2017) 'IL-2 therapy restores regulatory T-cell dysfunction induced by calcineurin inhibitors', *Proceedings of the National Academy of Sciences of the United States of America*, 114(27), pp. 7083–7088. doi: 10.1073/pnas.1620835114.

Williams, J. W. *et al.* (2013) 'Transcription factor IRF4 drives dendritic cells to promote Th2 differentiation', *Nature Communications*. Nature Publishing Group, 4(May). doi: 10.1038/ncomms3990.

Williams, L. *et al.* (2020) 'Post-transplantation Cyclophosphamide: From HLA-Haploidentical to Matched-Related and Matched-Unrelated Donor Blood and Marrow Transplantation', *Frontiers in Immunology*, 11(April), pp. 1–7. doi: 10.3389/fimmu.2020.00636.

Wolock, S. L., Lopez, R. and Klein, A. M. (2019) 'Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data', *Cell Systems*. Elsevier Inc., 8(4), pp. 281-291.e9. doi: 10.1016/j.cels.2018.11.005.

Wu, Y. *et al.* (2021) 'Donor T-Cell Repertoire Profiling in Recipient Lymphoid and Parenchyma Organs Reveals GVHD Pathogenesis at Clonal Levels After Bone Marrow Transplantation in Mice', *Frontiers in Immunology*, 12(December), pp. 1–13. doi: 10.3389/fimmu.2021.778996.

Wynn, T. A. and Vannella, K. M. (2016) 'Macrophages in Tissue Repair, Regeneration, and Fibrosis', *Immunity*. Elsevier Inc., 44(3), pp. 450–462. doi: 10.1016/j.immuni.2016.02.015.

Wysocki, C. A. *et al.* (2005) 'Leukocyte migration and graft-versus-host disease', *Blood*, 105(11), pp. 4191–4199. doi: 10.1182/blood-2004-12-4726.

Xue, D. *et al.* (2020) 'Transcriptome landscape of myeloid cells in human skin reveals diversity, rare populations and putative DC progenitors', *Journal of Dermatological Science*. Japanese Society for Investigative Dermatology, 97(1), pp. 41–49. doi: 10.1016/j.jdermsci.2019.11.012.

Xue, J. *et al.* (2014) 'Transcriptome-Based Network Analysis Reveals a Spectrum Model of Human Macrophage Activation', *Immunity*. Elsevier Inc., 40(2), pp. 274–288. doi: 10.1016/j.immuni.2014.01.006.

Xun, C. Q. *et al.* (1994) 'Effect of total body irradiation, busulfan-cyclophosphamide, or cyclophosphamide conditioning on inflammatory cytokine release and development of acute and chronic graft-versus-host disease in H-2- incompatible transplanted SCID mice', *Blood*, 83(8), pp. 2360–2367. doi: 10.1182/blood.v83.8.2360.bloodjournal8382360.

Yadav, R. *et al.* (2003) 'The H4b Minor Histocompatibility Antigen Is Caused by a Combination of Genetically Determined and Posttranslational Modifications', *The Journal of Immunology*, 170(10), pp. 5133–5142. doi: 10.4049/jimmunol.170.10.5133.

Yamashita, K. *et al.* (2004) 'Severe chronic graft-versus-host disease is characterized by a preponderance of CD4+ effector memory cells relative to central memory cells', *Blood*, 103(10), pp. 3986–3988. doi: 10.1182/blood-2003-09-3286.

Yew, P. Y. *et al.* (2015) 'Quantitative characterization of T-cell repertoire in allogeneic hematopoietic stem cell transplant recipients', *Bone Marrow Transplantation*. Nature Publishing Group, 50(9), pp. 1227–1234. doi: 10.1038/bmt.2015.133.

Yi, T. *et al.* (2009) 'Reciprocal differentiation and tissue-specific pathogenesis of Th1, Th2, and Th17 cells in graft-versus-host disease', *Blood*, 114(14), pp. 3101–3112. doi: 10.1182/blood-2009-05-219402.

Young, J. W. (2020) 'Alternative mechanisms that mediate graft-versus-host disease in allogeneic hematopoietic cell transplants', *Journal of Clinical Investigation*, 130(9), pp. 4532–4535. doi: 10.1172/JCI140064.

Young, M. D. and Behjati, S. (2020) 'SoupX removes ambient RNA contamination from

droplet-based single-cell RNA sequencing data', *GigaScience*, 9(12), pp. 1–10. doi: 10.1093/gigascience/giaa151.

Zaba, L. C. *et al.* (2007) 'Normal human dermis contains distinct populations of CD11c +BDCA-1+ dendritic cells and CD163+FXIIIA + macrophages', *Journal of Clinical Investigation*, 117(9), pp. 2517–2525. doi: 10.1172/JCI32282.

Zeiser, R. and Blazar, B. R. (2017) 'Acute Graft-versus-Host Disease Biology, Prevention and Therapy', *The New England Journal of Medicine*, 377(22), pp. 2167–2179. doi: 10.1002/cncr.27633.Percutaneous.

Zhan, Q. *et al.* (2012) 'Graft-versus-host disease-related cytokine-driven apoptosis depends on p73 in cytokeratin 15-positive target cells.', *Biology of blood and marrow transplantation : journal of the American Society for Blood and Marrow Transplantation*, 18(6), pp. 841–51. doi: 10.1016/j.bbmt.2012.02.004.

Zhang, Y. *et al.* (2021) 'Single-cell RNA sequencing in cancer research', *Journal of Experimental and Clinical Cancer Research*, 40(1), pp. 1–17. doi: 10.1186/s13046-021-01874-1.

Zheng, G. X. Y. *et al.* (2017) 'Massively parallel digital transcriptional profiling of single cells', *Nature Communications*. Nature Publishing Group, 8, pp. 1–12. doi: 10.1038/ncomms14049.

Appendix

Appendix table 1

	Chemotherapy regimens
GVH5	Prior chemotherapy (in chronological order): ABVD, DHAP, IVE, LEAM auto, Brentuximab vedotin, Nivolumab Fludarabine 30 mg/m^2 on days -6, -5, -4, -3, -2. Cyclophosphamide 14.5 mg/kg on days -6 and -5. Total body irradiation 200 cGy was performed on day -1. Post-transplant cyclophosphamide was administered at 60-72 hours and at 84-96 hours following stem cell infusion at a dose of 50 mg/kg. GVHD prophylaxis after transplantation was performed with tacrolimus with weaning by day 100 if no GVHD.
GVH2	Prior chemotherapy: CPX Fludarabine 30 mg/m^2 on days -6, -5, -4, -3, -2. Cyclophosphamide 14.5 mg/kg on days -6 and -5. Total body irradiation 200 cGy was performed on day -1. Post-transplant cyclophosphamide was administered at 60-72 hours and at 84-96 hours following stem cell infusion at a dose of 50 mg/kg. GVHD prophylaxis after transplantation was performed with tacrolimus with weaning by day 100 if no GVHD.
GVH1	Prior chemotherapy: CPX Fludarabine 30 mg/m^2 on days -6, -5, -4, -3, -2. Alemtuzumab 30 mg, melphalan 140 mg/m^2 and ciclosporin 2.5 mg/kg on day -1. GVHD prophylaxis after transplantation was performed with ciclosporin with weaning by day 100 if no GVHD.
GVH3	Prior chemotherapy: UKALL14 protocol, FLAG-IDA Cyclophosphamide 50 mg/kg on days -6 and -5. Total body irradiation 2 Gy x 2 (7 hour gap) was performed on days -3 and -2. Alemtuzumab 30 mg and ciclosporin 2.5 mg/kg on day -1. GVHD prophylaxis after transplantation was performed with ciclosporin with weaning by day 100 if no GVHD.

(table continued on next page)

GVH4	Prior chemotherapy: DA Fludarabine 30 mg/m2 on days -6, -5, -4, -3, -2. Alemtuzumab 30 mg, melphalan 140 mg/m2 and ciclosporin 2.5 mg/kg on day -1. GVHD prophylaxis after transplantation was performed with ciclosporin with weaning by day 100 if no GVHD.
TX1	Fludarabine 30 mg/m^2 on days -6, -5, -4, -3, -2. Cyclophosphamide 14.5 mg/kg on days -6 and -5. Total body irradiation 200 cGy was performed on day -1. Post-transplant cyclophosphamide was administered at 60-72 hours and at 84-96 hours following stem cell infusion at a dose of 50 mg/kg. GVHD prophylaxis after transplantation was performed with tacrolimus and mycophenolate mofetil (first 35 days) with weaning by day 100 if no GVHD.
TX2	Fludarabine 30 mg/m^2 on days -6, -5, -4, -3, -2. Alemtuzumab 30 mg, melphalan 140 mg/m^2 and ciclosporin 2.5 mg/kg on day -1. GVHD prophylaxis after transplantation was performed with ciclosporin with weaning by day 100 if no GVHD.

Appendix table 1 Details of chemotherapy regimens received by each patient.

Appendix table 2

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
cell_type * sample	67745	100.0%	0	0.0%	67745	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Monte Carlo Sig. (2-sided)			Monte Carlo Sig. (1-sided)		
				Significance	99% Confidence Interval		Significance	99% Confidence Interval	
					Lower Bound	Upper Bound		Lower Bound	Upper Bound
Pearson Chi-Square	17357.250[a]	80	.000	.000[b]	.000	<.001			
Likelihood Ratio	15579.768	80	.000	.000[b]	.000	<.001			
Fisher-Freeman-Halton Exact Test	15511.942			.000[b]	.000	<.001			
Linear-by-Linear Association	989.349[c]	1	<.001	.000[b]	.000	<.001	.000[b]	.000	<.001
N of Valid Cases	67745								

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 5.73.

b. Based on 10000 sampled tables with starting seed 624387341.

c. The standardized statistic is -31.454.

cell_type * sample Crosstabulation

			CT1	CT2	CT3	CT4	GVH5	GVH2	GVH1	GVH3	GVH4	TX1	TX2	Total
cell_type	Mast	Count	5a, b, c, d, e, f, g, h, i	0f, g, h, i	38j	121j	0d, e, h, i	4c, e, g, i	10b, k	29k	5a, b, c, d, e, f, g, h, i, k	0a, c, d, e, f, g, h, i	0a, b, c, d, e, f, g, h, i, k	212
		% within sample	0.1%	0.0%	2.1%	3.1%	0.0%	0.0%	0.3%	0.3%	0.1%	0.0%	0.0%	0.3%
	NK	Count	39a	24a	7a	24a	154b	724c	61d, e	190e	254b	32a	14a, d	1523
		% within sample	0.4%	0.3%	0.4%	0.6%	3.2%	5.4%	1.5%	1.9%	4.0%	0.7%	0.7%	2.2%
	Treg	Count	144a	111a, b	22a, b, c	22c	45a, b, c	75c	20c	553d	358d	42a, b, c	11b, c	1403
		% within sample	1.5%	1.5%	1.2%	0.6%	0.9%	0.6%	0.5%	5.6%	5.6%	0.9%	0.6%	2.1%
	CD8 T cell	Count	179a, b, c	196c	16b, d	24d	36d	334a, c	40d	454e	819f	26d	10d	2134
		% within sample	1.9%	2.6%	0.9%	0.6%	0.7%	2.5%	1.0%	4.6%	12.9%	0.5%	0.5%	3.2%
	CD4 T cell	Count	873a	419b, c, d	65e, f	37g	315d	654c, f	251b, d	1190h	1156i	99j	62e, j	5121
		% within sample	9.3%	5.5%	3.5%	1.0%	6.5%	4.9%	6.4%	12.1%	18.2%	2.1%	3.2%	7.6%
	pDC	Count	42a, b, c	6d	2a, b, c, d	4c, d	250e	480f	265e	58b	9a, c, d	7a, c, d	1a, b, c, d	1124
		% within sample	0.4%	0.1%	0.1%	0.1%	5.2%	3.6%	6.7%	0.6%	0.1%	0.1%	0.1%	1.7%
	cDC	Count	261a	143b	47a, b, c	80a, b	413d	1238d, e	425e	654f	252c	85b	53a, b, c	3651
		% within sample	2.8%	1.9%	2.6%	2.1%	8.6%	9.2%	10.8%	6.7%	4.0%	1.8%	2.7%	5.4%
	Macrophage	Count	991a	496b	126b	194b, c	756d	3905e	813f	1371d	755a	213c	209a	9829
		% within sample	10.5%	6.6%	6.9%	5.0%	15.7%	29.0%	20.6%	14.0%	11.9%	4.5%	10.7%	14.5%
	Stromal	Count	6862a	6155b	1508b	3355c	2845d	6058e	2053f	5303f	2735e	4273g	1601b	42748
		% within sample	73.0%	81.5%	82.4%	86.9%	59.1%	45.0%	52.1%	54.1%	43.1%	89.4%	81.6%	63.1%
Total		Count	9396	7550	1831	3861	4814	13472	3938	9802	6343	4777	1961	67745
		% within sample	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of sample categories whose column proportions do not differ significantly from each other at the .05 level.

Appendix table 2 Chi-square test of dermal cell populations across samples.

Appendix table 3

Crosstabs

Case Processing Summary

| | Cases | | | | | |
| | Valid | | Missing | | Total | |
	N	Percent	N	Percent	N	Percent
cell_type * sample	67745	100.0%	0	0.0%	67745	100.0%

Chi-Square Tests

| | | | Asymptotic Significance (2-sided) | Monte Carlo Sig. (2-sided) | | | Monte Carlo Sig. (1-sided) | | |
| | | | | | 99% Confidence Interval | | | 99% Confidence Interval | |
	Value	df		Significance	Lower Bound	Upper Bound	Significance	Lower Bound	Upper Bound
Pearson Chi-Square	9451.895[a]	110	.000	.000[b]	.000	<.001			
Likelihood Ratio	9174.936	110	.000	.000[b]	.000	<.001			
Fisher-Freeman-Halton Exact Test	9081.544			.000[b]	.000	<.001			
Linear-by-Linear Association	233.387[c]	1	<.001	.000[b]	.000	<.001	.000[b]	.000	<.001
N of Valid Cases	67745								

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 9.97.

b. Based on 10000 sampled tables with starting seed 2000000.

c. The standardized statistic is -15.277.

148

cell_type * sample Crosstabulation

cell_type			CT1	CT2	CT3	CT4	sample GVH5	GVH2	GVH1	GVH3	GVH4	TX1	TX2	Total
cell_type	pDC	Count	12a, b	0b	0a, b	4a, b	159c	490c	236d	139e	5a, b	12a	2a, b	1059
		% within sample	0.1%	0.0%	0.0%	0.1%	3.3%	3.6%	6.0%	1.4%	0.1%	0.3%	0.1%	1.6%
	mregDC	Count	9a	6a	0a, b, c, d	0a	86e, f	300f	17c, d	40b, d	67e	2a	1a, b, c, d	528
		% within sample	0.1%	0.1%	0.0%	0.0%	1.8%	2.2%	0.4%	0.4%	1.1%	0.0%	0.1%	0.8%
	DC1	Count	55a	12b	14a, c, d, e	15a, b	38a, e	136c, d, e	65d, f	204f	46a, c, e	8b	6a, b, c, e	599
		% within sample	0.6%	0.2%	0.8%	0.4%	0.8%	1.0%	1.7%	2.1%	0.7%	0.2%	0.3%	0.9%
	DC2	Count	99a, b	55b, c	10a, b, c	33a, b, c	148d	324d	201e	137a	76a, b	19c	23a, b	1125
		% within sample	1.1%	0.7%	0.5%	0.9%	3.1%	2.4%	5.1%	1.4%	1.2%	0.4%	1.2%	1.7%
	DC3	Count	146a, b, c	114a, b, c	27a, b, c, d	26d	226e	1112f	166e	203c	93a, b, c	52b, d	42a, c	2207
		% within sample	1.6%	1.5%	1.5%	0.7%	4.7%	8.3%	4.2%	2.1%	1.5%	1.1%	2.1%	3.3%
	Rec.mac1	Count	253a	139b, c	29a, b, c	48c	292d	1091e	335e	350f	313d	54c	49a, b, f	2953
		% within sample	2.7%	1.8%	1.6%	1.2%	6.1%	8.1%	8.5%	3.6%	4.9%	1.1%	2.5%	4.4%
	Rec.mac2	Count	389a	177b, c, d, e	54a, d, e, f, g, h, i	87b, c, d, e, h, i	120b, c, d, e, g, i	566a	236j	315f, g, h, i	132c, e	78c	72a, b, d, f, g, h, i	2226
		% within sample	4.1%	2.3%	2.9%	2.3%	2.5%	4.2%	6.0%	3.2%	2.1%	1.6%	3.7%	3.3%
	Rec.mac3	Count	103a, b	51b, c	10b, c	18c	129d, e	850f	120e	265d, e	122d	53a, b, c	36a, d, e	1757
		% within sample	1.1%	0.7%	0.5%	0.5%	2.7%	6.3%	3.0%	2.7%	1.9%	1.1%	1.8%	2.6%
	Rec.mac4	Count	19a, b, c, d, e, f, g, h	2g, h	0a, b, c, d, e, f, g, h, i, j, k, l	0e, f, h	78m	246m	22i, j, k, l	30c, d, k, l	16b, d, f, j, l	0a, e, g, h	3a, b, c, d, e, f, g, h, i, j, k, l	416
		% within sample	0.2%	0.0%	0.0%	0.0%	1.6%	1.8%	0.6%	0.3%	0.3%	0.0%	0.2%	0.6%
	Super.mac	Count	33a	20a, b	7a, b	10a, b	14a, b	197c	17a	50a	4b	7a, b	10a	369
		% within sample	0.4%	0.3%	0.4%	0.3%	0.3%	1.5%	0.4%	0.5%	0.1%	0.1%	0.5%	0.5%
	Resi.mac	Count	62a, b, c, d, e	25e	17c, d, f, g	17a, b, c, d, e	26a, b, c, d, e	234g	36b, d	276h	121f, g	14a, e	11a, b, c, d, e	839
		% within sample	0.7%	0.3%	0.9%	0.4%	0.5%	1.7%	0.9%	2.8%	1.9%	0.3%	0.6%	1.2%
	non_myeloid	Count	8216a	6949b, c	1663c	3603b, d	3498e	7926f	2487g	7793h	5348i	4478d	1706a, i	53667
		% within sample	87.4%	92.0%	90.8%	93.3%	72.7%	58.8%	63.2%	79.5%	84.3%	93.7%	87.0%	79.2%
Total		Count	9396	7550	1831	3861	4814	13472	3938	9802	6343	4777	1961	67745
		% within sample	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of sample categories whose column proportions do not differ significantly from each other at the .05 level.

Appendix table 3 Chi-square test of dermal myeloid cell subsets across samples.