Uncovering the molecular mechanisms behind mycetoma

Jonathan Chapman

Thesis submitted for the degree of Doctor of Philosophy

Newcastle University Biosciences Institute, Faculty of Medical Sciences, Newcastle University

September 2021

Abstract

Mycetoma is a chronic, painless, inflammatory condition, caused by either invading fungi or bacteria. It is one of twenty neglected tropical diseases formally recognised by the World Health Organisation. Following inoculation of the causative organisms into the subcutaneous tissue of a host, they organise into structures called grains. These in turn initiate the formation of granulomas and the development of a large, tumour-like mass. This lesion growth is reported to be painless by the majority of patients. Additionally, through unknown mechanisms, the pathogens appear to be able to persist in the host and evade their immune response.

This thesis focuses on actinomycetoma, which is exclusively caused by bacteria of the phylum Actinobacteria. Such bacteria are a major source of specialised metabolites, such as antibiotics, antitumour compounds and immunosuppressants. The central hypothesis of the thesis is that bacterial pathogens produce one or more specialised metabolites that mediate painless lesion development and pathogen persistence. A key aim was therefore to isolate and characterise any such compounds, which may also have therapeutic potential.

A new host-pathogen interaction assay was developed and applied to study virulence mechanisms of the actinomycetoma pathogen *Streptomyces sudanensis*. RNAseq, cytokine ELISAs, an NF-κB activity assay and microscopy were deployed to observe how murine macrophages and *S. sudanensis* interact. A unique immune profile was observed to be induced within the macrophages, characterised by a mixture of pro- and anti-inflammatory features. Multiple potential virulence factors were also identified within *S. sudanensis*. Additionally, human tissue culture cells were shown to undergo pyroptosis when interacting with the pathogen. Two related compounds that appear to be responsible for this activity were isolated from the bacteria and identified as 2,5-diketopiperazines.

This work provides novel insights into how mycetoma pathogens interact with the immune system and of the molecular mechanisms underlying this disease.

To my family

Acknowledgements

I would like to thank my brilliant supervisors, Prof. Jeff Errington, Dr Bernhard Kepplinger and Dr Katarzyna Mickiewicz. Their guidance, wisdom and support both in the lab and beyond was always insightful, considerate and comforting. I feel immensely privileged to have been guided by not just one but three great scientists.

I need to give thanks to Dr Nick Allenby for collaborating on bacterial genome sequencing and for assembling the genome of *Streptomyces sudanensis*. I also acknowledge Dr Olga Chrobak for providing training in RNA extraction and sample handling. I need to thank John Casement of the Bioinformatics Support Unit for his helpful guidance on performing RNAseq analysis, as well as for preparing the *S. sudanensis* transcriptome files. I would also like to thank Dr Andrew Watson for his excellent advice and many useful discussions regarding my RNAseq data.

I would like to thank Dr Yousef Dashti for working so patiently with me on the chemical purification process and for performing the NMR analysis. Additionally, I would like to thank Prof. Matthias Trost for donating THP-1 cells to the lab and Dr Abeer Dannoura for providing training and advice in how to culture these cells. Special thanks also go to Fran Davison for all of her help with tissue culture and more, and for somehow managing to keep the lab running smoothly. I also want to give a big thank you to all members of the Errington Group and everyone on the fourth floor in the CBCB for creating such a welcoming and supportive lab environment.

I especially want to thank my parents for all of their love and support through the years and without whom I would not be where I am today. I owe them so much.

Finally, I would like to thank my partner Akvile, who has supported me through some of the best and worst days of this project. Even when I was plagued by self-doubt, she always believed in me and inspired me to push on. Aš negalėjau to užbaigti be tavęs.

Table of Contents

Table of Figures

Table of Tables

Abbreviations

^{1}H-NMR – proton nuclear magnetic resonance

2-ME – 2-mercaptoethanol

aa – amino acids

ACN – acetonitrile

ACP-domain – acyl carrier protein domain

A-domain – adenylation domain

AMP – adenosine monophosphate

antiSMASH – antibiotics and secondary metabolite analysis shell

AT-domain – acyl transferase domain

ATP – adenosine triphosphate

BCDH – branched-chain α-keto acid dehydrogenase

BCR – B cell receptor

BGC – biosynthetic gene cluster

BLASTP – Protein Basic Local Alignment Search Tool

cDNA – complementary DNA

C-domain – condensation domain

CDPS – cyclodipeptide synthase

CDS – protein coding DNA sequence

CL – cutaneous leishmaniasis

CO_2 – carbon dioxide

DAMP – danger associated molecular pattern

DC – dendritic cell

DMEM – Dulbecco's Modified Eagle's Medium

DMSO – dimethyl sulfoxide

DNA – deoxyribose nucleic acid

ECM – extracellular matrix

E-domain – epimersation domain

EDTA – Ethylenediaminetetraacetic acid

ELISA – enzyme-linked immunosorbent assay

ESI – electrospray ionisation

FBS – foetal bovine serum

HGA – homogentisic acid

HPLC – high performance liquid chromatography

HTH – helix turn helix

Ig – immunoglobulin

IL – interleukin

ILC – innate lymphoid cell

KS-domain – ketosynthase domain

LC-MS – liquid chromatography-mass spectrometry

LL – lepromatous

LPS – lipopolysaccharide

M – macrophages

M+C – macrophages + *S. coelicolor*

M+DC – macrophages + dead *S. coelicolor*

M+DS – macrophages + dead *S. sudanensis*

M+S – macrophages + *S. sudanensis*

MGC – multinucleate giant cell

MHC – major histocompatibility complex

MMP – matrix metalloproteinase

mRNA – messenger RNA

MZ – marginal zone

NGS – next generation sequencing

NK cell – natural killer cell

NO – nitric oxide

NRPS – non-ribosomal peptide synthetase

OD – optical density

ONT – Oxford Nanopore Technologies

PacBio – Pacific Biosciences

PAMP – pathogen associated molecular pattern

PC – principal component

PCA – principal component analysis

PCP-domain – peptidyl carrier protein domain

PCR – polymerase chain reaction

PKS – polyketide synthase

PRR – pattern recognition receptor

PTM – post-translational modification

RiPP – ribosomally synthesised and post-translationally modified peptide

RNA – ribose nucleic acid

RNAseq – RNA sequencing

ROS – reactive oxygen species

RPMI media – Roswell Park Memorial Institute media

rRNA – ribosomal RNA

RT-PCR – real time polymerase chain reaction

S – *S. sudanensis* in isolation

S+M – *S. sudanensis* + macrophages

SEAP – secreted alkaline phosphatase

SN – supernatant

SNP – single nucleotide polymorphism

TB – tuberculosis

TBE – tris borate EDTA

TCR – T cell receptor

TCS – two component system

TE-domain – thioesterase domain

TF – transcription factor

TGF – transforming growth factor

Th1, 2 or 17 – T helper cell type 1, 2 or 17

TLR – toll-like receptor

TNF – tumour necrotic factor

TOF – time of flight

Treg – regulatory T cell

tRNA – transfer RNA

TT – tuberculoid

VDJ segments – Variable, Diversity and Joining segments

VEGF – vascular endothelial growth factor

VL – visceral leishmaniasis

WHO – World Health Organization

WT – wild type

Chapter 1. Introduction

1.1 Mycetoma - an overview

1.1.1 A neglected inflammatory disease

Mycetoma is a chronic, granulomatous, inflammatory condition, which is progressive in its nature (van de Sande, 2013, van de Sande et al., 2014, Fahal et al., 2015, Cardenas-de la Garza et al., 2020). First described clinically in 1842, it is confined for the most part to tropical regions, as shown by Figure 1.1. The countries with the highest prevalence are Mauritania, Sudan, Mexico, Senegal, Niger and Somalia (van de Sande, 2013, van de Sande et al., 2018).

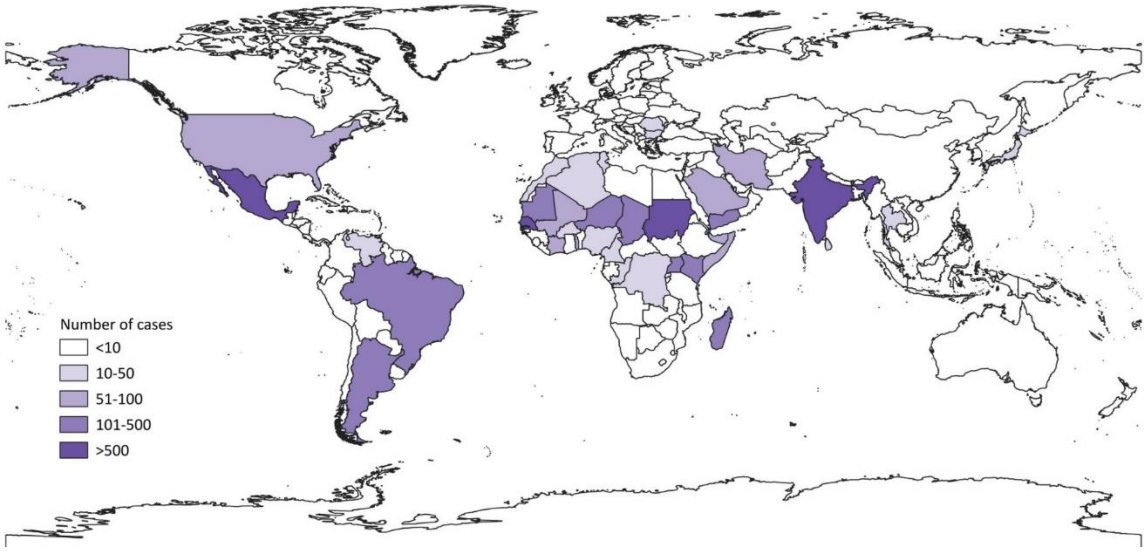


Figure 1.1 Map of the world indicating the number of mycetoma cases reported within each country. Image taken from Figure 3 of Emery and Denning (2020).

Mycetoma is formally listed by the World Health Organisation (WHO) as one of twenty neglected tropical diseases in the world, reflective of its incidence mainly amongst individuals of low socio-economic status (Fahal, 2017). It can be caused by either fungi, in which case it is termed eumycetoma, or by bacteria, referred to as actinomycetoma (Arenas et al., 2017). A meta-analysis of all published cases, conducted by van de Sande in 2013, estimated that actinomycetoma accounted for just over 50% of all mycetoma cases globally and it is this bacterial form of the disease which is the focus of this study.

1.1.2 The causative agents of actinomycetoma

There are at least 58 species (bacterial and fungal) thought to cause mycetoma, although the true number of causative agents is hard to ascertain, as historically neither form of mycetoma has been a reportable disease in clinics (van de Sande, 2013, Zijlstra et al., 2016, Gilquin et al., 2016, Gueneau et al., 2020). In terms of actinomycetoma, the most common causative species, starting with the most prevalent, are *Actinomadura madurae*, *Streptomyces somaliensis*, *Actinomadura pelletieri*, *Nocardia brasiliensis* and *Nocardia asteroides* (van de Sande, 2013). They, along with the other less common bacterial agents, are all members of the order Actinomycetales and are collectively referred to as actinomycetes.

1.1.3 Pathology of actinomycetoma

The pathology of mycetoma is well characterised. Once the causative agent has been inoculated into the subcutaneous tissue of an individual, it triggers the formation of structures called grains, comprised of the bacterial or fungal mycelium embedded in a hard cement-like material composed of agglomerated proteins, lipids, melanins and metal ions (Ibrahim et al., 2013, Sheehan et al., 2020). These become surrounded by neutrophils and macrophages as an inflammatory response is triggered, visible in Figure 1.2(a-c) (van de Sande, 2013, Arenas et al., 2017).

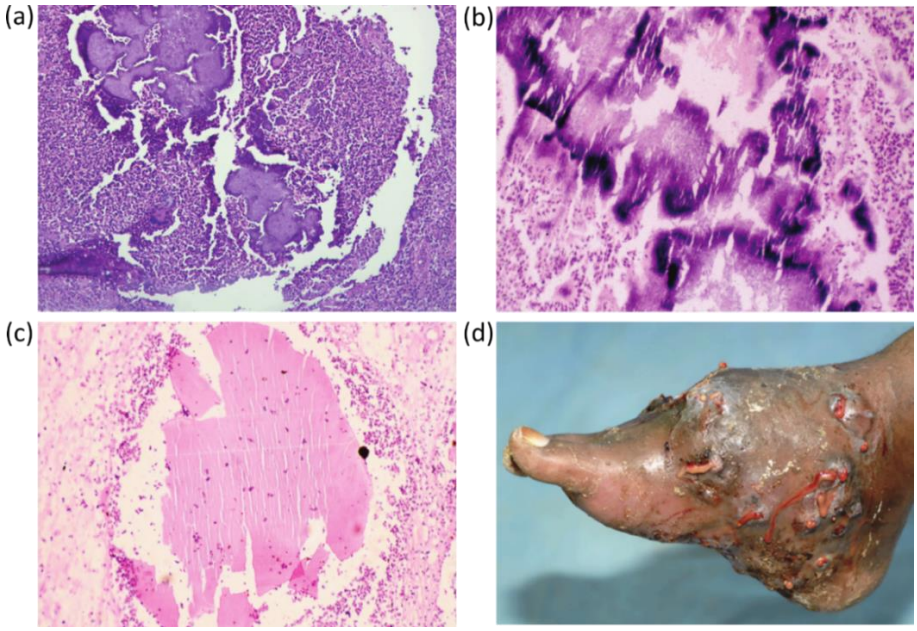


Figure 1.2 (a) Actinomycetoma grains formed by *A. pelletieri*, surrounded by neutrophils and macrophages. Image from Fahal et al. (2015). **(b)** Actinomycetoma grains formed by *A. madurae*. Image from Arenas et al. (2017). **(c)** Actinomycetoma grains formed by *S. somaliensis*. Image from Arenas et al. (2017). **(d)** Foot of patient, featuring a large mycetoma growth. Image from (Fahal et al., 2015).

Recruited immune cells then form a stratified, multicellular structure around the grain, called a ‘granuloma’, represented in Figure 1.3, around which a fibrotic cap then develops. This is a key feature of the pathology of the disease.

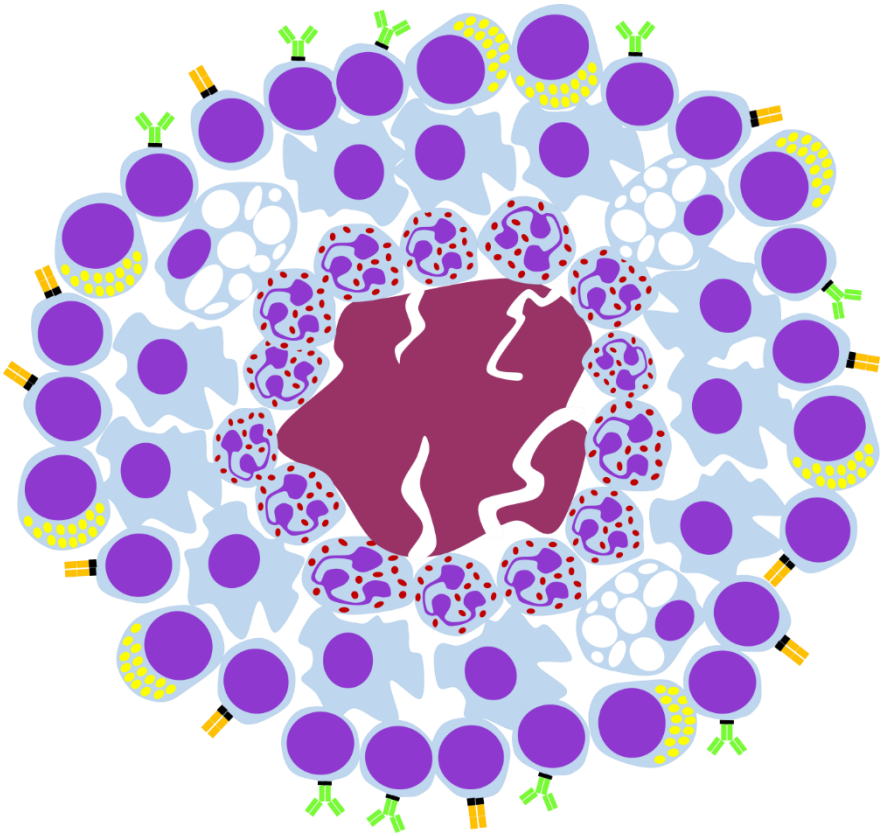


Figure 1.3 Diagrammatic example of a mycetoma granuloma. The mycetoma grain is surrounded by immune cells, which have formed the layered granuloma structure.

The most common site in the body for this initial infection is the foot, followed by the hand, but mycetoma of the head and the torso have also been documented (Fahal et al., 2015). From the grain-containing granulomas, nodules and then a larger subcutaneous mass will develop, visible as tumour-like swelling and causing major deformity to the infected tissue, as in Figure 1.2(d). Abscesses form, leading to the secretion of grain-containing pus and eventual scarring. If the condition is left untreated at this point, the bone will be invaded and surgery becomes the only treatment option, with limb amputations being common for individuals in the late stages of the disease (Fahal et al., 2015, Cardenas-de la Garza et al., 2020). In the most severe cases, the condition can spread further to the internal organs, with involvement of the lungs, bladder and spinal cord all being reported (Fahal et al., 2015).

1.1.4 Knowledge gaps

There are multiple aspects of mycetoma which require further scientific investigation. For example, how the causative agents are inoculated into individuals is unknown (van de Sande et al., 2014). Some mycetoma pathogens have been successfully cultured from soil, but the potential mechanism of transition from soil to humans remains poorly understood (Zijlstra et al., 2016). With no known animal reservoir or vector, the current prevailing theories regard trauma to patients' skin caused by plant thorns as the most likely inoculation route (van de Sande et al., 2014, Zijlstra et al., 2016). Disease prevention remains a challenge as long as the infection route remains unknown.

Individuals in the early phase of the disease do not show any obvious external symptoms, meaning diagnosing the condition early has proven to be difficult (van de Sande et al., 2014). When symptoms do manifest, proper diagnosis requires complex histopathology and imaging techniques (Fahal et al., 2015, Siddig et al., 2019a). Such methods are very rapid and highly accurate, but also require surgical biopsies and expensive equipment. They are therefore not readily available in the poorer and more remote areas where mycetoma occurs. Simpler and cheaper culturing techniques to identify the causative species, which use grains from patients, also run into problems as many of the bacterial species are difficult to grow in labs, meaning such methods take roughly four times longer to yield a diagnostic result than histological examinations (Siddig et al., 2019a). Pathogens cultured from grains are also very easily incorrectly identified, leading to incorrect drug choice for treatment and patients suffering as a result (Fahal et al., 2015). Other diagnostic alternatives are PCR-based techniques, which have been shown to reliably identify pathogenic species (van de Sande et al., 2014). They are a promising area for the development of diagnostic tools, with the first multiplex RT-PCR assay that is able to accurately identify multiple eumycetoma pathogens having been recently developed (Arastehfar et al., 2020). But, as with other techniques, their current relative high cost makes them inapplicable in mycetoma endemic areas.

Clinicians have emphasised the importance of developing accurate serological tests for use in the field, which would allow for immediate and low cost identification and characterisation of mycetoma infections (Hay et al., 2019). Antibody tests exist for *Nocardia* spp. but their specificity for mycetoma infections versus other nocardial diseases has not been studied and therefore their validity could be limited (Salinas Carmona et al., 1993, Cardenas-de la Garza et

al., 2020). Recently however, *Actinomadura* species have tested positive in these *Nocardia* serological tests, but the underlying molecular cause of these positive results has not yet been investigated (Cardenas-de la Garza et al., 2020). Ultimately, novel markers common across the majority of actinomycetoma infections need to be identified to allow for swift and accurate diagnosis in the field.

It has been observed that the course of the disease is dependent on what the infective organism is, with actinomycetoma being more aggressive and damaging than eumycetoma (Zijlstra et al., 2016, Emmanuel et al., 2018). This effect can also be seen between the different bacterial species, where the disease progresses faster with some than with others. There are no suggestions for specific potential virulence factors which could be responsible for a more aggressive mycetoma spread. Speculation has attributed differences in disease progression to the diverse chemistry abilities of mycetoma pathogens and whether they produce more "substances that destroy tissue" (Vera-Cabrera et al., 2012). Related to this effect is the observation from patient studies in Sudan that the duration of the disease can vary from a few months in some individuals, all the way up to 60 years in others (Fahal et al., 2015). Questions as to how the infection is able to evade, or at the very least modulate, the immune system for such lengths of time and why this effect varies between individuals, have not been answered.

The same patient study highlighted another curious aspect of the disease; 70.5% of patients experienced no pain, despite the presence of a subcutaneous mass. Any pain that was experienced was intermittent and attributed to secondary bacterial infections of the affected site, as opposed to the primary mycetoma agent itself. As with the disease duration effect, the molecular mechanisms behind this lack of pain and whether it is attributable to the causative bacteria have not been resolved. Additionally, how bacteria form grains within tissue and what molecular mechanisms initiate this process has not been explored (van de Sande et al., 2014).

Taken together, these gaps in knowledge serve to highlight that there are myriad questions relating to the underlying molecular biology of mycetoma which need to be addressed.

1.2 The human immune system

To set the pathology of mycetoma in its proper context, the composition and function of the human immune system first needs to be examined.

1.2.1 An overview

The human immune system functions to distinguish foreign agents, such as pathogenic microbes, and abnormal cells, such as cancerous growths, from the healthy self. It eliminates anything recognised as non-self and repairs tissue damage. The system is comprised of a highly complex network of cells, tissues and organs, highlighted in Figure 1.4. These all act in concert to safeguard the body.

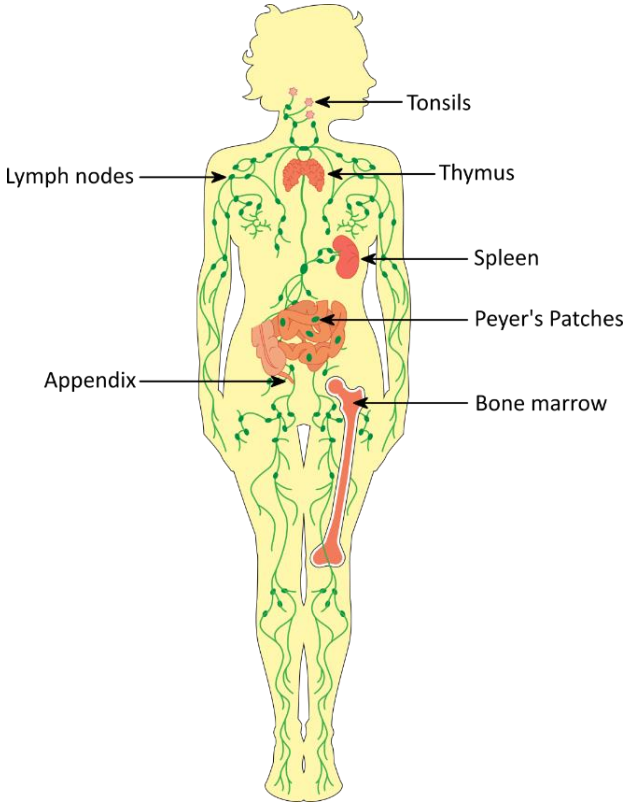


Figure 1.4 Tissues and organs of the human immune system. Image adapted from Nagwa (n.d.).

Immune activities are mediated by a wide variety of soluble proteins and white blood cells, which freely circulate through nearly all tissues. White blood cells have a wide variety of functions and are classically grouped into distinct two arms of the immune system, namely innate and adaptive immunity (Akira et al., 2006). The innate arm constitutes a germline-

encoded first line of defence, while the adaptive arm involves highly specific recognition of pathogens and the generation of immunological memory.

Innate immunity is mediated by the complement cascade of soluble proteins and the following cell types: phagocytic macrophages and dendritic cells (DCs); granulocytic neutrophils, basophils, eosinophils and mast cells; and innate lymphoid cells (ILCs) (Akira et al., 2006, Iwasaki and Medzhitov, 2015). All can be seen in Figure 1.5(a-g). Adaptive immunity is mediated by only two cell types: T and B lymphocytes. These are shown in Figure 1.5(h-i). While these two arms are defined separately, they do not exist in isolation, with each reinforcing and regulating the other. The primary mechanisms of this regulation are the release of cytokines, the chemical messengers of the immune system, and presentation of foreign antigen to other immune cells to trigger activation. Innate and adaptive immunity combine to mount a cohesive immune response.

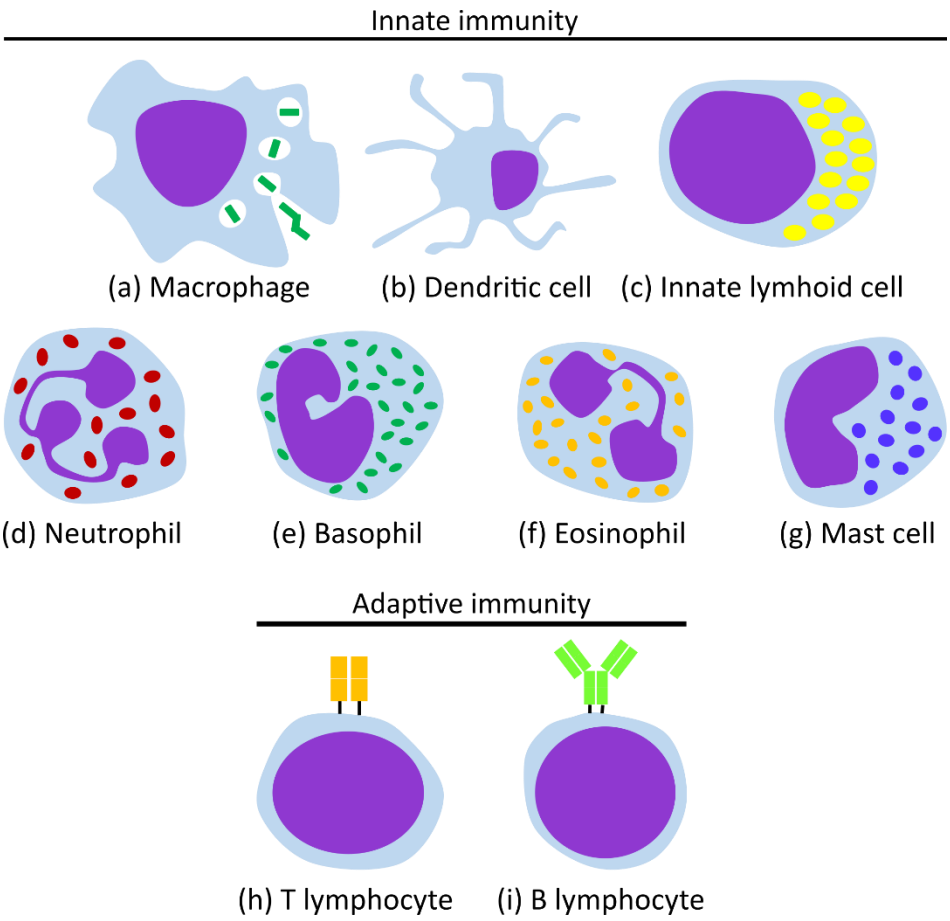


Figure 1.5 Cell types of the innate and adaptive branches of the human immune system. **(a)** Macrophage, carrying out phagocytosis of a pathogen. **(b)** Dendritic cell. **(c)** Innate lymphoid cell, represented here by a natural killer cell. **(d)** Neutrophil. **(e)** Basophil**. (f)** Eosinophil. **(g)** Mast cell. **(h)** T lymphocyte, displaying its defining T cell receptor. **(i)** B lymphocyte, displaying its defining B cell receptor.

The microbial pathogens that the immune system can respond to include species of bacteria, fungi, viruses and parasites. Such pathogens contain evolutionarily conserved components, generally essential to their survival, which are collectively known as pathogen associated molecular patterns (PAMPs) (Akira et al., 2006, Iwasaki and Medzhitov, 2015, Fitzgerald and Kagan, 2020). For bacterial species, PAMPs may include lipopolysaccharide (LPS), lipoproteins or peptidoglycan. Additionally, pathogen activities and attacks against the host, mediated by their virulence factors, result in detectable conserved effects, known as danger associated molecular patterns (DAMPs) (Iwasaki and Medzhitov, 2015, Fitzgerald and Kagan, 2020). Examples of DAMPs include peptide fragments, indicating protease activity, or debris from lysed host cells, indicating the activity of pore-forming toxins. Determination of non-self from self by the immune system is enabled by the detection of PAMPs and DAMPs. Following recognition of a foreign threat, an immune response is elicited. The following sections will focus on the response to pathogens, with particular emphasis on bacteria.

1.2.2 Detection of PAMPs and DAMPs

Threat detection begins with pattern recognition receptors (PRRs) and the complement system (Iwasaki and Medzhitov, 2015, Reis et al., 2019, Fitzgerald and Kagan, 2020). PRRs are expressed on the surfaces and within the endosomes of all host cells, i.e., not only immune cells but also other cell types, such as epithelial cells. They detect pathogens via the recognition of PAMPs and DAMPs. Classes of PRRs include Toll-like receptors (TLRs), nucleotide-binding oligomerisation domain-like receptors, RIG-I-like receptors, C-type lectin receptors and AIM-2-like receptors (Iwasaki and Medzhitov, 2015).

Ligand binding to PRRs triggers signalling cascades that lead to the activation of immunity related transcription factors and expression of response genes. Taking the example of TLRs, ligand recognition leads to the activation of an intracellular signalling cascade, as shown in Figure 1.6 (Fitzgerald and Kagan, 2020). This ultimately leads to the translocation into the nucleus of NF-κB and AP-1, key mediators of the inflammatory immune response. These transcription factors bind to DNA promoter sequences, triggering expression of cytokines, chemokines and numerous other immune response genes, the products of which mediate activation and upregulation of an immune response.

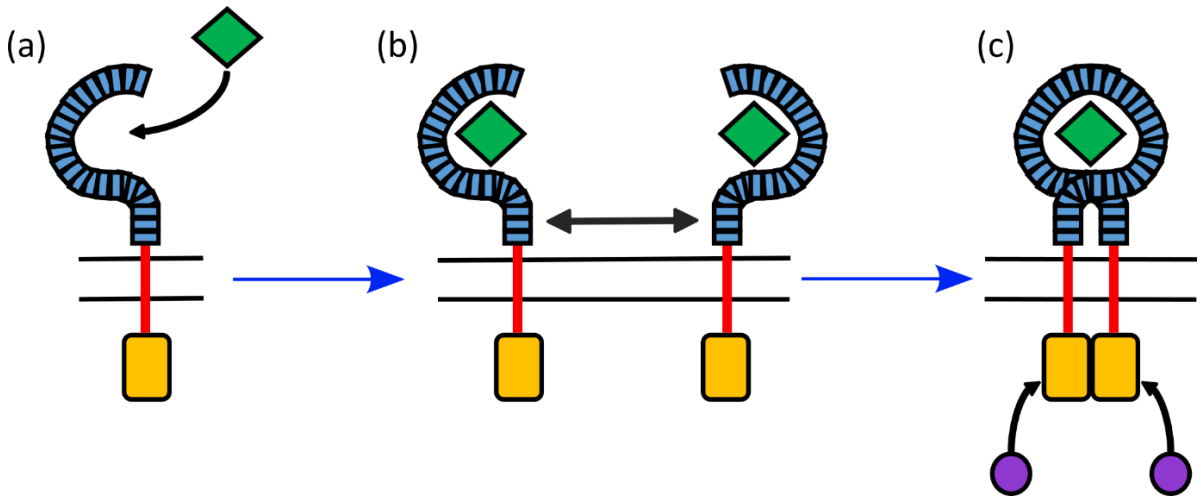


Figure 1.6 Structure of a TLR and activation of its signalling pathway. TLRs consist of an extracellular leucine-rich repeat domain (blue) and an intracellular TIR domain (orange). **(a)** A PAMP (green) is bound by the TLR. **(b)** TLR undergoes dimerization. **(c)** Adaptor proteins (purple) bind to the receptors' TIR domains, initiating a signalling cascade.

Proteins of the complement system can also recognise PAMPs and DAMPs (Reis et al., 2019). Recognition triggers a proteolytic cascade of the proteins, producing peptide fragments with various functions. Some are bound by complement receptors on the surfaces of immune cells, resulting in activation of those cells. In the case of innate immunity, this amplifies the immune response by enhancing NF-κB and AP-1 activation. Other complement peptides coat pathogens to enhance their uptake by phagocytosis, trigger immune cell migration by acting as chemoattractants, regulate cell adherence or assemble into a membrane attack complex, which then inserts into the membranes of pathogens and triggers lysis.

1.2.3 Macrophages and dendritic cells as the first line of defence

Activation of the above innate signalling systems brings tissue-resident macrophages and DCs into action (Akira et al., 2006, Iwasaki and Medzhitov, 2015). Macrophages differentiate from circulating monocytes in the blood, after which they patrol the peripheral tissues of the body (Mosser and Edwards, 2008, Iwasaki and Medzhitov, 2015). Differing cocktails of PAMPs, DAMPs, cytokines and chemokines trigger activation of these cells and determine their polarisation to a particular phenotype. At one end of this polarisation scale are classically activated, or M1, macrophages (Mosser and Edwards, 2008). They arise in response to the detection of bacterial, fungal or viral components, exposure to apoptotic cells and upon encountering pro-inflammatory cytokines, such as tumour necrotic factor (TNF) α or interferon (IFN) γ, produced by other host cells (Mosser and Edwards, 2008, O'Neill and

Pearce, 2016, Parisi et al., 2018). M1 macrophages have an enhanced ability to kill microbes. This is mediated by increased phagocytosis and the upregulation of bactericidal reactive oxygen species (ROS) and nitric oxide (NO) production. They also themselves secrete large quantities of pro-inflammatory cytokines, such as interleukin (IL) 1β, TNFα and IL6, promoting activation, differentiation and recruitment of other pro-inflammatory immune cells involved in cell-mediated immunity.

At the other end of the polarisation spectrum are alternatively activated, or M2, macrophages. They arise in response to anti-inflammatory signals, such as IL10 and IL4, which are produced in the later stages of the pro-inflammatory response by lymphocytes and in response to infection by parasites (Mosser and Edwards, 2008, O'Neill and Pearce, 2016). They can be further divided into wound-healing and regulatory phenotypes (Mosser and Edwards, 2008). The main feature of wound-healing cells is their upregulation of the conversion of arginine to ornithine, which in turn can be converted into collagen and polyamines. Collagen contributes to the production of a new extracellular matrix (ECM) and thus mediates tissue repair and wound healing. Polyamines suppress the clonal expansion of T and B lymphocytes, thus constraining adaptive immunity. Regulatory macrophages meanwhile produce high levels of anti-inflammatory cytokines, such as IL10, the key mediator of the anti-inflammatory response, and transforming growth factor (TGF) β. M2 macrophages are essential for supressing the immune system to prevent a runaway immune response that would ultimately kill the host. It should be emphasised that macrophage polarisation is not binary, but rather a sliding scale with M1 and M2 at either end and a range of mixed phenotypes in between (Parisi et al., 2018).

DCs also continuously patrol the peripheral tissues, but their role is to act as the main antigen presenting cells of the immune system (Banchereau and Steinman, 1998, Bieber and Autenrieth, 2020). DCs are continually taking up particles from their environment by phagocytosis, degrading them and presenting the resulting peptide epitopes on their surfaces in complexes with major histocompatibility complexes (MHC), a process detailed in Figure 1.7 (O'Neill and Pearce, 2016).

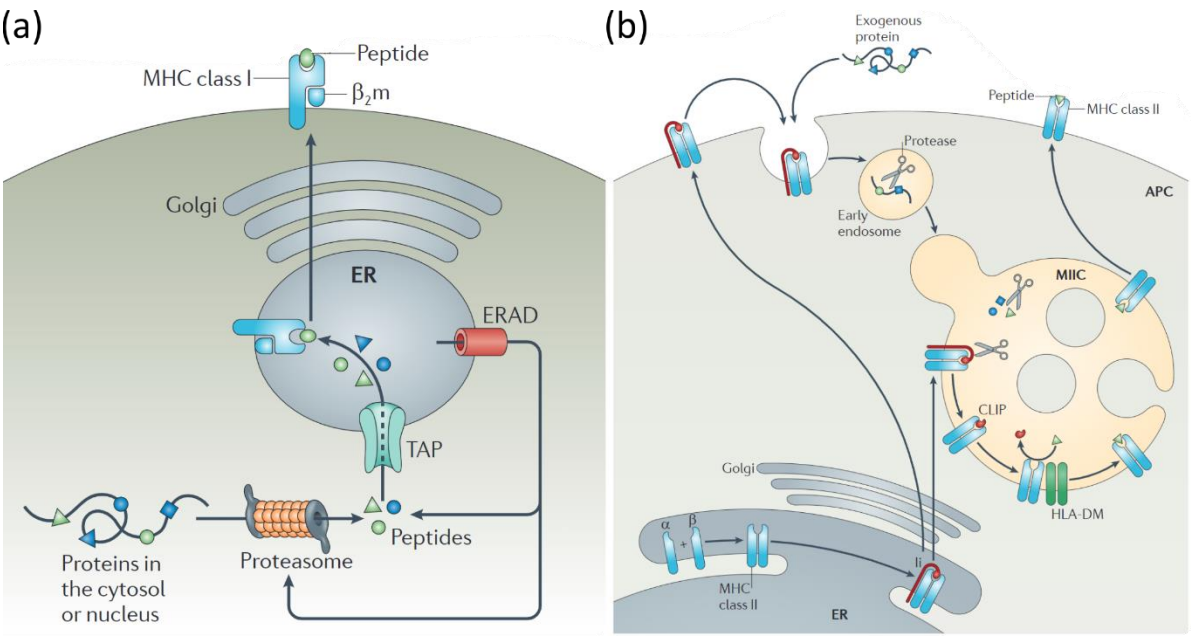


Figure 1.7 Pathways for the presentation of antigens at the surface of cells. **(a)** Presentation of intracellular antigens with MHC I. Proteins are degraded by the proteasome, transported into the endoplasmic reticulum (ER) via transporter associated with antigen presentation (TAP) and complexed with MHC I. The complex is transported to the plasma membrane via the Golgi. **(b)** Presentation of extracellular antigens with MHC II. MHCII is assembled with the invariant chain (Ii) in the ER and transported via the Golgi to an MHC II compartment (MIIC). The Ii is cleaved and, via the chaperone HLA-DM, replaced with antigen peptide, derived from endocytosed and proteolytically cleaved exogenous protein. The MHC II-antigen complex is then transported to the plasma membrane. Adapted from figures 1 and 3 of Neefjes et al. (2011).

Through this process they are able to present “self” antigens, indicating to other immune cells the absence of pathogens, and also enabling the tolerisation of lymphocytes, preventing autoimmune reactions (Banchereau and Steinman, 1998). When microbes or PAMPs are encountered, following phagocytosis, the foreign epitopes are presented with type II MHCs on the DC surface, indicating to other cells the presence of an extracellular infection (see Figure 1.7(b)). If a foreign antigen is detected within the DC cytoplasm, it is instead presented on type I MHC, signalling an intracellular infection (see Figure 1.7(a)). Encounters with pathogens trigger DC activation via PRRs. The DCs are induced to boost antigen presenting capacity through increased display of MHC at the plasma membrane, upregulate expression of lymphocyte co-stimulatory molecules and release pro-inflammatory cytokines and chemokines (Banchereau and Steinman, 1998, Bieber and Autenrieth, 2020). Activated DCs will also migrate to lymphoid tissues in order to present antigen to naïve cells T and B lymphocytes, engaging adaptive immunity in the fight against the infection.

1.2.4 Neutrophils are recruited following macrophage activation

During a bacterial infection, cytokines and chemokines released by the activated macrophages and DCs recruit a milieu of circulating innate leukocytes to the site of infection. These include monocytes, which differentiate into additional macrophages and DCs, and the granulocytic neutrophils (Iwasaki and Medzhitov, 2015). Neutrophils are the most abundant leukocytes, accounting for 70% of all circulating white blood cells in humans (Papayannopoulos, 2018, Rosales, 2020). Like macrophages, they are highly phenotypically heterogeneous, however their specific subclasses have not yet been formally categorised (Rosales, 2020). Once at the site of infection, neutrophils carry out their core microbicidal function. This microbial killing is mediated through a variety of methods, namely: phagocytosis of pathogens, release of toxins from cytoplasmic granules via degranulation, release of ROS and formation of neutrophil extracellular traps (Iwasaki and Medzhitov, 2015, Papayannopoulos, 2018).

Neutrophils can also cooperate and coordinate their attacks against larger groups of pathogens, which would normally overwhelm individual cells, through a poorly characterised process known as 'swarming' (Kienle and Lammermann, 2016, Alex et al., 2020). Additionally, they release their own cocktail of pro-inflammatory cytokines to help maintain and enhance immune cell recruitment and activation to fight the infection (Rosales, 2020). Neutrophils are some of the most biologically well-equipped and potentially destructive cells of the immune system.

1.2.5 The role of innate lymphoid cells in shaping the immune response

Also activated by the cytokine milieu released by macrophages and DCs are ILCs (Iwasaki and Medzhitov, 2015). These cells have shared phenotypic features and reside in peripheral tissues, in particular barrier surfaces like mucosal membranes, as opposed to in lymph nodes or the blood (Spits and Cupedo, 2012, Spits et al., 2013, Klose and Artis, 2016). They lack PRRs and can only be activated by cytokines (Klose and Artis, 2016). ILCs are grouped into three distinct functional categories (Spits et al., 2013). Group 1 are characterised by their high production of the pro-inflammatory cytokine IFNγ and play a key role in anti-viral immunity. Natural killer (NK) cells are members of this group and are responsible for detecting infected or stressed host cells and lysing them (Iwasaki and Medzhitov, 2015, Klose and Artis, 2016). They are the only cytotoxic ILCs. Group 2 ILCs are involved in immune responses to extracellular parasites and in allergy and characteristically produce the cytokines IL4, IL5 and

IL13 (Spits et al., 2013, Klose and Artis, 2016). Group 3 produce IL17 and IL22 and play a role in the formation and repair of lymphoid tissue, as well as in defence against extracellular bacteria. ILCs ultimately serve to regulate and tailor immune responses to suit the specific threat encountered.

1.2.6 Activation and roles of T and B cells

As described above, the second arm of the immune system is adaptive immunity, which is mediated by T and B lymphocytes. T lymphocytes take their name from the fact that they mature in the thymus, after which they move to peripheral tissues (Kumar et al., 2018). In the periphery, they can be found in three forms: regulatory T cells (Tregs), naïve T cells and memory T cells. Tregs are defined by their expression of the transcription factor FOXP3 and maintain immune self-tolerance and homeostasis by suppressing lymphocyte activity (Kitagawa and Sakaguchi, 2017). Naïve T cells are inactive and remain in the lymph nodes, awaiting stimulation by foreign antigens complexed with MHC I or II and costimulatory molecules, all presented on the surfaces of DCs and macrophages (Kumar et al., 2018). Detection of antigens is mediated via their characteristic T cell receptor (TCR) (Alcover et al., 2018). Following activation, the naïve cells will undergo clonal expansion and differentiate into a variety of effector subtypes, which are split into two camps: those expressing the surface receptor CD8 (CD8+) and those expressing CD4 (CD4+) (Zhu et al., 2010).

CD8+ T cells are cytotoxic, directly killing infected or stressed host cells and therefore playing a key role in cell mediated immunity. CD4+ cells are known as T 'helper' cells and play a similar role to ILCs, in that they also regulate and tailor the immune response to best suit the specific pathogenic threat encountered. They themselves are subdivided into three groups, namely Th1, Th2 and Th17, the functionalities of which are mirrored by the ILC groups (Ziegler, 2016). Th1 cells produce IFNγ, TNFα and IL2 and so are involved in promoting the inflammatory response and encouraging cell mediated immunity, for instance by increasing and maintaining macrophage activation (Zhu et al., 2010, Ziegler, 2016). Th2 cells mainly produce IL4, IL5, IL10 and IL13, so suppress inflammation and help to combat extracellular parasitic infections. They also encourage humoral immunity, through boosting antibody production by B lymphocytes and therefore are important for a strong antibody response (Mosmann and Sad, 1996). Finally, Th17 cells are characterised by their production of IL17A and IL17F and boosting tissue

inflammation (Ziegler, 2016). Their key function is in promoting clearance of extracellular bacterial and fungal infections (Patel and Kuchroo, 2015).

Following activation and differentiation, T cells migrate from the lymph nodes to the site of infection and proceed with their defined roles (Kumar et al., 2018). Upon clearance of an infection, effector T cells undergo apoptosis, leaving the final class of T cells in their wake, the memory T cells. These cells are long-lived and recognise the specific antigen/pathogen that was previously cleared. Should the same pathogen infect the host again, it can be recognised by the memory cells, triggering an adaptive immune response at a much faster rate than before. It is through this class of cell that the adaptive immune system generates immunological memory and long-term immunity.

B lymphocytes take their name from the fact that they mature in the bone marrow and are distinguished as the producers of antibody and by their expression of the B cell receptor (BCR) (Banchereau and Steinman, 1998, Treanor, 2012). They mediate humoral immunity. Like T cells, once matured they exist in a naïve state and can be subdivided into distinct classes, namely B1a, B1b, B2 and marginal zone (MZ) cells (D'Souza and Bhattacharya, 2019). B1a and B1b cells play a role in the early humoral response and in fact have innate immunity characteristics, as they can only produce antibodies with affinity to evolutionarily conserved antigens (Haas et al., 2005, Prieto and Felippe, 2017). B2 and MZ cells on the other hand can recognise previously unencountered antigens and are therefore truly adaptive, coming into play in the later stages of an infection. The key distinction between them is that B2 cells circulate throughout the body, while MZ cells are localised exclusively in the spleen.

Binding of antigen ligands to the BCR, alongside recognition of costimulatory molecules and cytokines supplied by Th2 cells, triggers activation, differentiation and clonal expansion of naïve B cells (Treanor, 2012, D'Souza and Bhattacharya, 2019). Activation can also be triggered by TLRs. Differentiation is either into memory B cells, which provide lasting immunity to a pathogen in the same manner as memory T cells, or into plasma cells, which devote the vast majority of their cellular resources to the production of antibodies (D'Souza and Bhattacharya, 2019).

1.2.7 Antibodies – the arsenal of B cells

Antibodies are extracellular proteins, comprised of light and heavy peptide chains, which themselves are comprised of multiple immunoglobulin domains. In humans, there are 5 classes of antibody: IgA, IgD, IgE, IgG and IgM (Stavnezer, 1996). Each class has a different functional role determined by their structures, which can be seen in Figure 1.8.

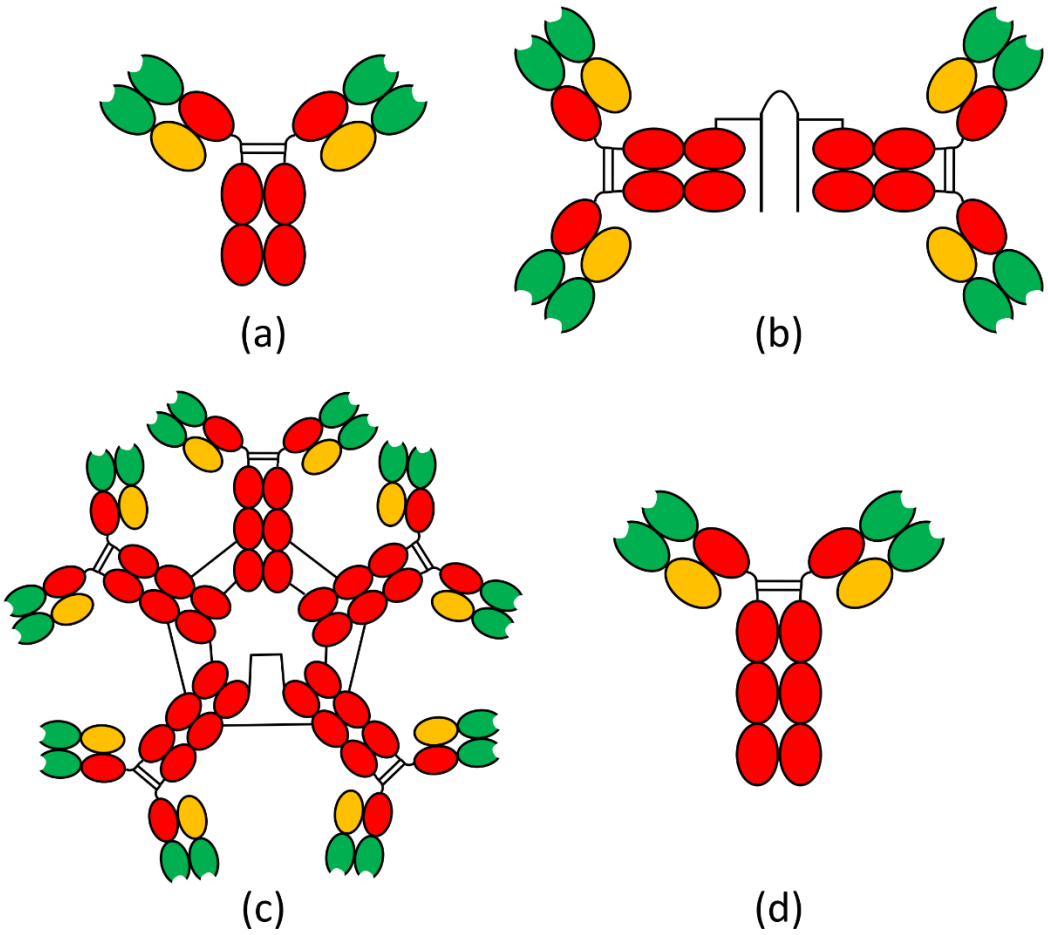


Figure 1.8 Structures of the 5 classes of antibodies. Each oval represents an immunoglobulin domain; the heavy chain is coloured red; the light chain is coloured orange; variable regions are coloured green. **(a)** IgG and IgD. IgD is membrane bound, acting as the BCR. IgG is secreted. **(b)** IgA, a secreted dimer involved in mucosal immunity. **(c)** IgM, a secreted pentamer. **(d)** IgE, usually membrane bound and plays role in allergy and parasite defence.

B cells have the ability to undergo class switching, whereby they transition from producing one class of antibody to another (Stavnezer, 1996). Class switching is promoted by T helper cells and is dependent on the nature of the pathogenic threat encountered, allowing the most functionally effective antibody class in a given scenario to be synthesised. B2 cells have the ability to class switch to all isotypes, while B1a and B1b cells are more limited, primarily producing IgM and being able to class switch only to IgA (Prieto and Felippe, 2017).

Antibodies have multiple modes of action (Lu et al., 2018). They can bind directly to antigens on the surfaces of pathogens, preventing them from exerting their virulent effects in a process called neutralisation. In the case of the multimeric IgA and IgM, multiple pathogens can be bound by the same antibody, clumping the invaders together and making them a more attractive target for phagocytes and physically preventing their dissemination in the host. This is known as agglutination. Phagocytes express surface receptors that recognise antibodies and so coating a pathogen with antibodies makes it more easily detectable and promotes phagocytosis. This is termed opsonisation. NK cells also have receptors for antibodies and are triggered to release their cytotoxic cargo upon receptor activation, so antibodies can boost cytotoxicity. Antibodies can also be bound by complement proteins, triggering the cascade and its downstream effector functions. Finally, antibodies can bind soluble factors secreted by pathogens to aid virulence, causing them to precipitate out of solution and rendering them ineffective.

1.2.8 Generating adaptive immunity – somatic recombination and hypermutation

T and B lymphocytes can modify their TCRs, BCRs and antibodies to specifically recognise newly encountered antigens. These fundamental processes are what make this branch of the immune system "adaptive" and are critical to its function. In the case of the TCR, this adaptation is solely mediated by somatic DNA recombination of the variable (V), diversity (D) and joining (J) segments of the genes that encode the receptor peptide chains (Hozumi and Tonegawa, 1976, Oettinger et al., 1990). Recombination of different V, D and J segments generates an entirely new amino acid sequence and therefore a new tertiary structure for the binding region of the receptor peptide chain. Thus, through essentially trial and error, a T cell can be generated that is able to recognise an antigen that was previously unknown to the immune system.

B cells also use VDJ recombination to generate structural diversity for the immunoglobulin chains of their BCRs and antibodies. But they also have another genetic tool at their disposal, somatic hypermutation (Di Noia and Neuberger, 2007). This process involves introducing point mutations in the immunoglobulin heavy and light chain genes, resulting in changes to their amino acid sequences and protein structures. With this process working alongside VDJ recombination, B cells have an even greater capacity for generating structural diversity and stronger binding affinity for novel antigens than T cells.

Moving from this broad perspective of the immune system, the way in which an actinomycetoma infection interacts with many of the above-described processes will now be examined.

1.3 Mycetoma and the immune system

1.3.1 Granulomas and disease

A key feature of mycetoma pathology is the formation of granulomas in the host tissues. A granuloma was authoritatively defined by Dolph Adams in a 1976 review as "A compact collection of mature mononuclear phagocytes, which is not necessarily accompanied by accessory features, such as necrosis" (Adams, 1976). Research in the decades since this definition has shown that of the mononuclear phagocytes, macrophages are the central architects of these cellular structures (Ramakrishnan, 2012). These cells organise into granulomas in the presence of foreign bodies or persistent antigens, to physically wall-off and contain threats which they have been unable to destroy. Granuloma-inducing antigens are derived from various species of parasites, viruses, fungi or bacteria (Pagan and Ramakrishnan, 2018). In the case of nearly all infectious granulomatous diseases, the identity of the inciting antigen is unknown. Following granuloma formation by macrophages, other immune cell types can also be recruited, including T and B lymphocytes, NK cells and DCs (Pagan and Ramakrishnan, 2018, Wilson et al., 2019). If the stimulating antigen is still not cleared by this response, then granuloma formation becomes excessive and leads to severe pathologies in the host, its protective function subverted (Pagan and Ramakrishnan, 2018).

Depending on the infectious agent present, macrophages can differentiate into various cell types as the granuloma matures. They can undergo a mesenchymal-epithelial transition into large, elongated cells, which pack together with interdigitated membranes (Adams, 1974, Pagan and Ramakrishnan, 2018). Such macrophages are called 'epithelioid' cells. They resemble epithelial cells and express canonical epithelial markers, such as E-cadherin (Cronan et al., 2016). The presence of epithelioid cells provides rigid structure to granulomas. It has been shown that disruption of epithelial markers in macrophages results in the development of disorganised granulomas, which allow greater infiltration of neutrophils and lymphocytes (Cronan et al., 2016, Kauffman et al., 2018). Greater infiltration of these cells correlates with

increased pathogen killing and host survival. Epithelioid granulomas can therefore promote pathogen persistence. Their development process is shown in Figure 1.9. Infectious granulomas are typically epithelioid, raising the question of whether there is a common initiating stimulus shared amongst pathogens (Pagan and Ramakrishnan, 2018). Macrophages can also fuse membranes with their neighbours, becoming multinucleate giant cells (MGCs), or oxidise intracellular lipids to become foam cells.

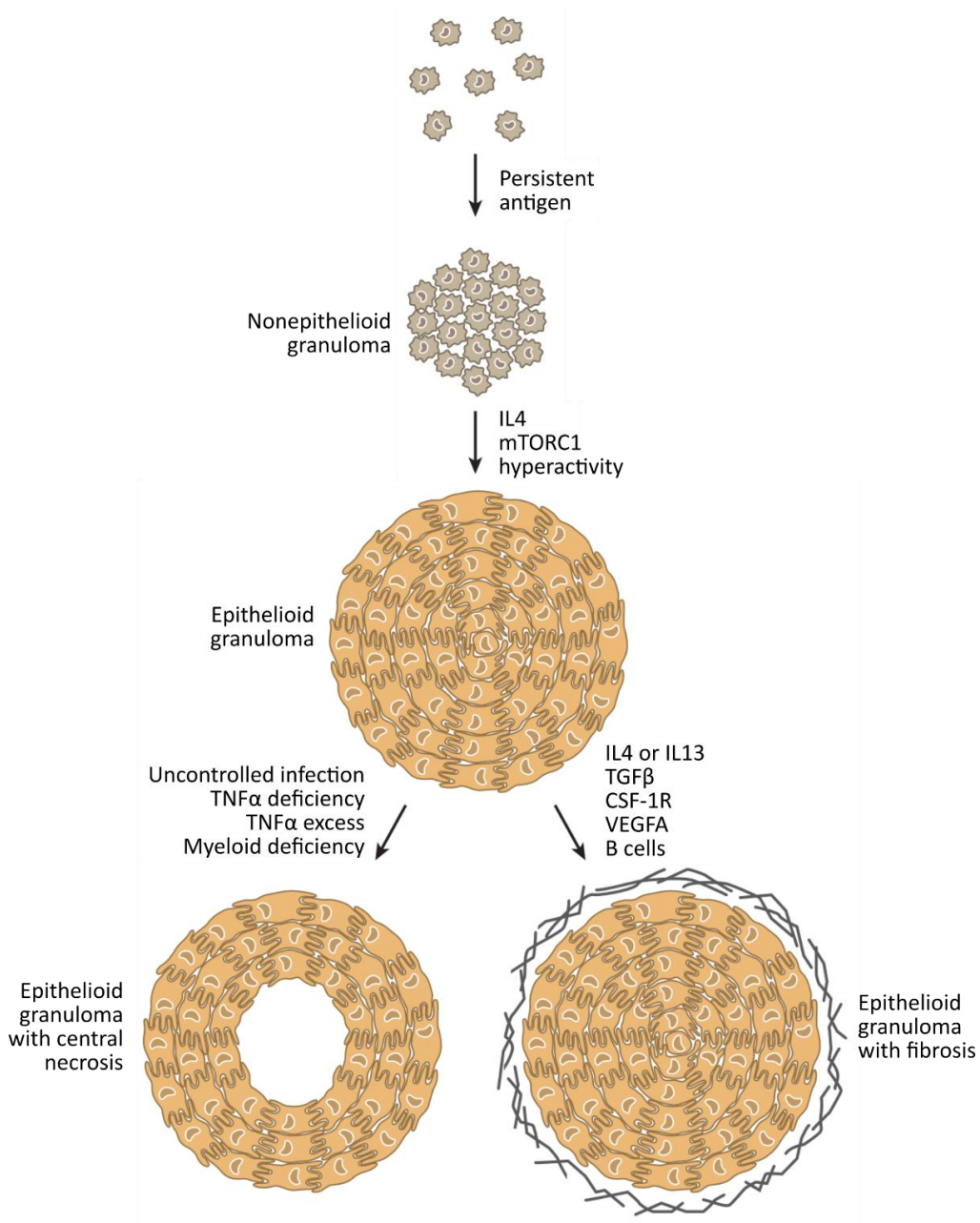


Figure 1.9 Summary of epithelioid granuloma development, including stimulants proposed to trigger each stage. Persistent antigen triggers macrophages to form into a granuloma. Epithelioid differentiation is then thought to be triggered either by IL4 or chronic mTORC1 signalling. Depending on the stimuli present, granulomas can then progress either to a necrosis or fibrosis. Taken from Figure 1 of Pagan and Ramakrishnan (2018).

Other common structural changes in mature granulomas are necrosis and fibrosis, seen as diverging pathways in Figure 1.9. Necrosis is confined to the core of the granuloma and is most common during infections by intracellular pathogens. Fibrosis is an excessive production of connective tissue, which builds up around the perimeter of granulomas, forming a fibrotic cap. The presence of either of these features is an indicator of a poor health outcome for a patient. As with the initiation of granuloma formation and epithelioid differentiation, it is unknown why some pathogens induce these effects while others do not.

1.3.2 The mycetoma grain and granuloma

As described above, granuloma development in mycetoma follows the formation of a grain by the invading pathogen. A study of fungal grain formation during a eumycetoma infection in the model organism *Galleria mellonella* found this process to consist of four steps (Sheehan et al., 2020). First, the host and pathogen recognise one another. This induces the pathogen to secrete enzymes that cross-link ECM proteins to the fungal mycelium and each other, while immune cells form a sheath around the mycelium. The immune cells then begin to produce ROS and antimicrobial peptides, triggering a stress response in the fungal cells. This stress response leads to the initiation of melanisation by the fungi, forming an encapsulated grain. Granuloma formation by recruited immune cells then follows.

Study of the granulomas formed during eumycetoma development from *Madurella mycetomatis* infection led to the observation of the presence of three types of host tissue reaction, represented in Figure 1.10 (Fahal et al., 1995).

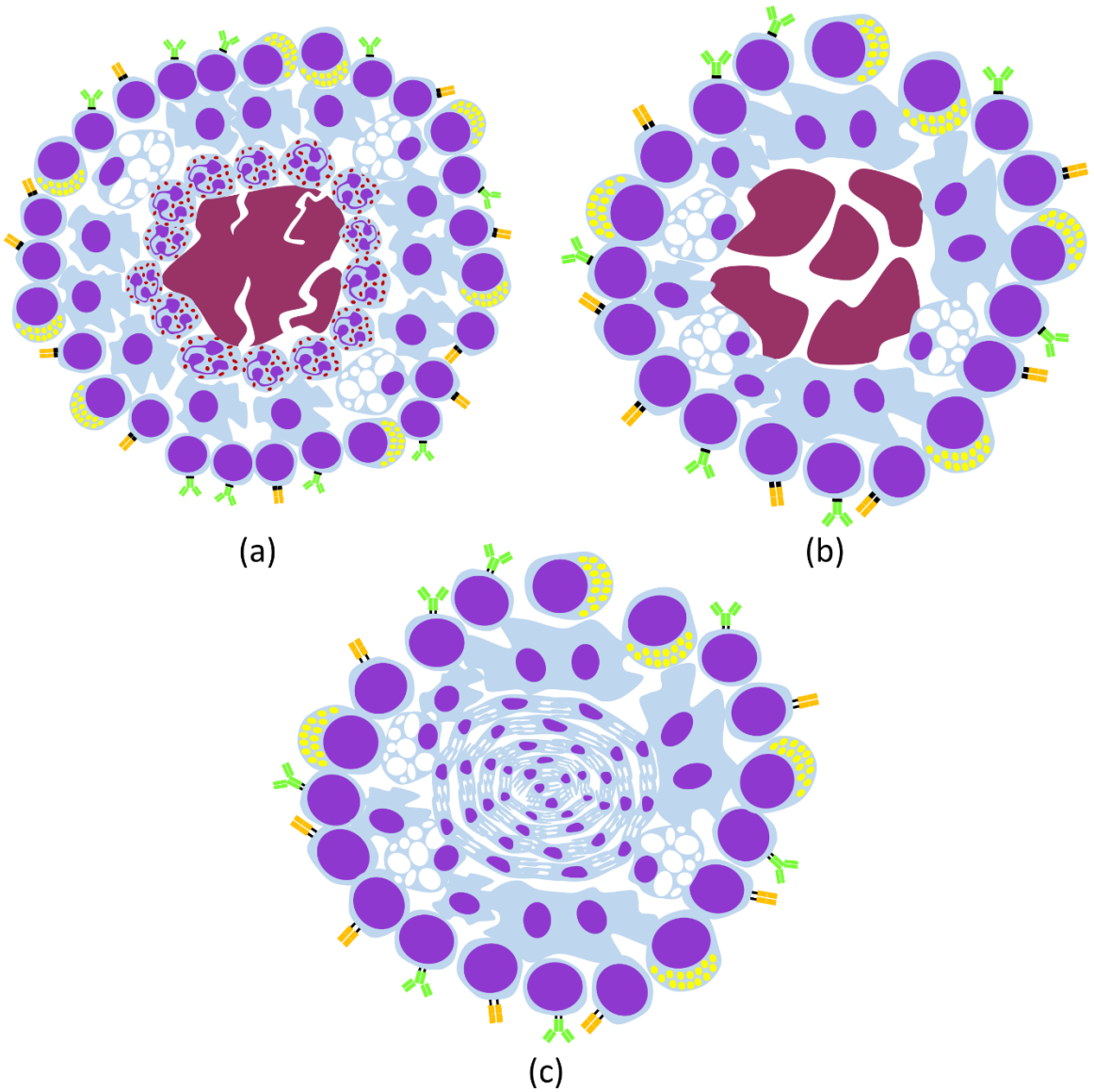


Figure 1.10 Types of granuloma observed during mycetoma infections. **(a)** Type 1: 3 zones of cells surround the central bacterial/fungal grain. The first consists of neutrophils, which start to fragment the grain. The second contains of macrophages and foam cells. The third is comprised of lymphocytes (T, B and NK cells). **(b)** Type 2: Neutrophil zone is replaced by a larger zone of macrophages, which also includes MGCs. These cells clear the fragmented grain, any pathogens and dead neutrophils. Lymphocytes are still present. **(c)** Type 3: The grain has been cleared, leaving an epithelioid granuloma structure. This is comprised of a core of epithelioid macrophages. MGCs and foam cells are still present. The outer layer of lymphocytes also persists.

Type I granuloma tissue is comprised of neutrophils immediately surrounding the grain in a central layer, encircled by a second layer of macrophages with vacuolated cytoplasm and lymphocytes, as seen in Figure 1.10(a). The grain at the centre may be partially fragmented, due to the activity of the neutrophils. In type II tissue the inner neutrophil layer is not present, replaced instead with macrophages and MGCs, shown in Figure 1.10(b). It was observed that

macrophages in this tissue type appear to be involved in clearing the fragmented grain and any dead neutrophils. Type III tissue has a similar organisation to type II, except the grain at the centre has been cleared, with a core of epithelioid macrophages and MGCs taking its place, shown in Figure 1.10(c).

These tissue types were later observed surrounding grains of *S. somaliensis* in an actinomycetoma infection (El Hassan et al., 2001). Very rarely would one type be seen in isolation within an infection site, with combinations of the two being most common, and it is speculated that type I in fact gives way to type II as grain fragmentation progresses. It was noted in the case of eumycetoma that while this response by immune cells did successfully destroy existing grains, new ones would still form, indicating the failure of this immune response to completely eliminate the invading organism (Fahal et al., 1995). Pathogen persistence may in fact be favoured by this response. In eumycetoma, some giant cells within type II tissue were seen to contain viable *M. mycetomatis* hyphae, while foam cells were observed around the granulomas of an actinomycetoma infection (Fahal et al., 1995, Salinas-Carmona et al., 2012).

1.3.3 Genetic factors affecting the immune response to mycetoma

Three patient studies examined the role genetic factors might play in increasing an individual's susceptibility to eumycetoma. Owing to the presence of neutrophils at the centre of mycetoma granulomas, the first study looked into single nucleotide polymorphisms (SNPs) that impacted neutrophil function (van de Sande et al., 2007). Key differences in allele occurrence were found between eumycetoma patients and a healthy control group; namely for the genes *CR1*, *CXCL8*, *CXCR2*, *TSP4* and *NOS2*. *CR1* encodes complement receptor 1 on neutrophils and binds to pathogens coated in complement proteins, allowing the neutrophil to 'present' such pathogens to a macrophage for phagocytosis. Eumycetoma patients were found to have SNPs that would reduce the efficacy of this receptor for binding complement, which would therefore limit neutrophil functionality.

CXCL8 encodes IL8, *CXCR2* encodes its receptor and *TSP4* produces thrombospondin-4, an initiator of IL8 production. IL8 is a chemokine for neutrophils that recruits the cells to sites of infection or tissue damage. In the mycetoma patients, alleles associated with increased IL8 production were found for all three genes, meaning an excessive number of neutrophils would

be drawn to sites of damage. It is postulated that the activity of these additional neutrophils could result in further damage to tissues and subsequent progression of mycetoma.

The final gene flagged in this study, *NOS2*, encodes NO Synthase Type 2 in neutrophils. NO production is used as a method of killing bacteria and within the mycetoma patients *NOS2* was expressed at lower levels than in the control group. Lower expression of *NOS2* lowers NO production, which logically results in neutrophils becoming less competent at killing bacteria.

The second genetic study looked at genes involved in the production of the chemokine CCL5 and the cytokine IL10 (Mhmoud et al., 2013). CCL5 is secreted by macrophages and promotes the formation of granulomas by attracting T lymphocytes, DCs, NK cells, mast cells and basophils to the site of infection. In the case of mycetoma patients, two alleles for CCL5 were identified which were not present in the healthy control group and thus were deemed to be linked to increased susceptibility for the disease. The alleles appeared to contribute to increasing CCL5 expression in the mycetoma patients.

As described above, IL10 is an anti-inflammatory cytokine. It suppresses the production of NO, ROS and cell mediated immunity, promoting humoral immunity instead (Cyktor and Turner, 2011). Within mycetoma patients, an allele linked to increased expression of this cytokine was found (Mhmoud et al., 2013). Increased IL10 production by immune cells could create an environment of less direct killing of bacteria and the host cells infected by them, due to a down regulation of the cell mediated immune response.

The third study examined expression of matrix metalloproteinases (MMPs) in eumycetoma infections (Geneugelijk et al., 2014). MMPs are proteolytic enzymes expressed in response to infection or injury and which play roles in tissue remodelling, wound healing, activation or inhibition of cytokines and promoting the migration of immune cells through tissues (Nissinen and Kahari, 2014, Tomlin and Piccinini, 2018). Excessive MMP activity has been shown to contribute to the pathology of inflammatory diseases. Within the context of mycetoma, these enzymes were speculated to be responsible for the accumulation of collagen that leads to the granulomas undergoing fibrosis (Geneugelijk et al., 2014). The study found that not only was expression of MMPs raised in eumycetoma, but that the disease was associated with a particular allele in *TIMP-1*. This gene encodes an inhibitor of MMP activity and the allele associated with the disease was one that reduces its expression (Ali et al., 2020). The presence

of the allele was therefore speculated to contribute to raised MMP activity and the development of fibrosis in mycetoma lesions.

1.3.4 Cytokine profile of mycetoma

Assaying for cytokines produced in *N. brasiliensis*-induced actinomycetoma revealed that pro-inflammatory cytokines are strongly expressed in the early stages of the disease (Solis-Soto et al., 2008). Infecting keratinocytes with *A. madurae* yielded a similar set of observations (Santiago-Tellez et al., 2019). Furthermore, *in vivo* cytokine detection in patients with eumycetoma produced the same result, with raised levels of IL1β, IL2, IL12, TNFα and IFNγ all being detected (Nasr et al., 2016, Abushouk et al., 2019). These interleukins are indicative of a pro-inflammatory response and promote cell mediated immunity (Mosmann and Sad, 1996, Zhu et al., 2010, Abushouk et al., 2019). Additionally, IL17A and MMP9 have been found to be expressed in mycetoma granulomas (Siddig et al., 2019b). IL17A promotes production of pro-inflammatory cytokines by immune cells, while MMP9 degrades the extracellular matrix and proteolytically activates chemokines to aid cell migration. Expression levels of both correlated with increased disease duration in patients.

Also detected in both eumycetoma and actinomycetoma infections were the cytokines IL4, IL5, IL6, IL10, IL13, IL35 and IL37 (El Hassan et al., 2001, Mendez-Tovar et al., 2004, Nasr et al., 2016, Abushouk et al., 2019). These cytokines all induce an anti-inflammatory response, as well as promoting antibody production to enable humoral immunity (Mosmann and Sad, 1996, Zhu et al., 2010).

Interestingly, in the actinomycetoma studies, IL10 and IL4 were the cytokines found to be at the highest levels within lesions, indicating the mounting of a strong anti-inflammatory response by the immune system (El Hassan et al., 2001, Salinas-Carmona et al., 2012). A patient study also found IFNγ, a key effector of inflammation, to be significantly lower in actinomycetoma patients when compared to a control group, as well as IL10 and IL4 to be significantly higher (Mendez-Tovar et al., 2004, Zhu et al., 2010). IL35 and IL37 levels were also found to be positively correlated with lesion size and disease duration in patients, while IL1β and IL12 levels showed a negative correlation (Abushouk et al., 2019). The dampening of the pro-inflammatory response would inhibit cell mediated immunity, allowing the causative mycetoma bacteria to survive the immune response and persist within the host, thus exacerbating the disease (Mendez-Tovar et al., 2004, Salinas-Carmona et al., 2012).

1.3.5 The role of adaptive immunity in the host response to mycetoma

The adaptive immune response to mycetoma has been examined over the last three decades (Rico et al., 1982, Deem et al., 1983, Salinas-Carmona and Perez-Rivera, 2004). Initial investigations showed that T cell deficient mice had increased susceptibility towards *N. brasiliensis*-induced actinomycetoma, while B cell deficient mice were able to control the infection via a T cell response (Rico et al., 1982). Additionally, dosing of T cell deficient and normal mice with anti-*Nocardia* serum gave no protection against mycetoma and instead appeared to cause the symptoms to severely worsen. The humoral immune response was further implicated in the development of the disease when analysis of human actinomycetoma lesions, caused by *S. somaliensis*, showed the presence of the antibodies IgG and IgM on the surface of grains at the centre of granulomas (El Hassan et al., 2001).

Another mouse study found that T cells previously challenged with *N. asteroides* would kill these bacteria *in vitro*, but were not as effective against closely related species, such as *N. brasiliensis* (Deem et al., 1983). This confirmed that T cells were able to become immunologically specific for the causative agents of mycetoma and could successfully eliminate them. These findings seemingly point to cell mediated immunity being the key method for controlling and subduing mycetoma, while humoral immunity could be facilitating the disease. It therefore highlights again how the anti-inflammatory cytokine profile benefits mycetoma pathogens.

This assumption is challenged by later studies, which found that, rather than humoral immunity not protecting against mycetoma, it is in fact specifically IgG that provides no protection (Salinas-Carmona and Perez-Rivera, 2004, Gonzalez-Suarez et al., 2009). Evidence showed that IgM serum administered to mice was able to defend against an experimental *N. brasiliensis* actinomycetoma, whereas IgG serum gave no such effect. Despite it being produced in the early stages of an infection and being thought to be a short-lived response, IgM has also been shown to provide long term immunity against intracellular infection by other bacteria (Racine et al., 2011).

The basis for its effectiveness could lie in its structure, which is shown in Figure 1.8(c). IgM is a pentamer, while IgG is a monomer, meaning it is significantly more efficient at agglutination of pathogens and is able to bind multiple antigens on pathogens simultaneously (Ehrenstein and Notley, 2010). More efficient agglutination would lead to a greater enhancement of

bactericidal killing by phagocytes and prevent the bacteria from spreading further into the host's tissues. The gradual transition from an IgM to an IgG response in the tissue around a grain has been suggested to be the source of the delay in the development of mycetoma in patients following the initial infection (Zijlstra et al., 2016).

1.3.6 Recognition of the mycetoma by the innate immune system

Expression of PRRs during actinomycetoma development was examined in mice, revealing an increase in TLR2 expression and a decrease in TLR4 expression (Millan-Chiu et al., 2011). In the early stages of infection, TLR2 was highly expressed by neutrophils and macrophages in contact with the causative bacteria, while TLR4 was expressed by mast cells. But going into the late stage, TLR4 expression became downregulated overall and TLR2 became upregulated in foam cells and fibroblasts around the edges of granulomas.

TLR4 is known to recognise LPS, a cell wall component of Gram-negatives, hence its downregulation in the case of mycetoma, which is caused by Gram-positive bacteria (Akira et al., 2006). TLR2 however, can recognise various components of the Gram-positive cell wall, such as lipoproteins and peptidoglycan. So, its upregulation in this case could allow the immune system to better respond to the bacterial infection. Cell wall components have been implicated in triggering the inflammatory response towards *N. brasiliensis*, with removal of cell wall-associated lipids by washing bacteria in solvent preventing both inflammation and the development of actinomycetoma in mice (Trevino-Villarreal et al., 2012).

It should be noted that as well as triggering a pro-inflammatory response, TLR2 signalling has also been shown to induce the expression of IL10 (Cyktor and Turner, 2011). This would further shift the immune system towards an anti-inflammatory response, aiding pathogen persistence and mycetoma development. The host's attempts to respond to the bacteria by upregulating TLR2 would then have a runaway effect of promoting mycetoma progression even further.

While precise molecular mechanisms of disruption have yet to be elucidated, it is suggested by the existing body of work that the development of mycetoma may depend upon the causative organism disrupting the balance between different arms of the immune system. This could result in a switch from a pro-inflammatory to an anti-inflammatory response, allowing survival of the pathogen.

1.4 Wider actinobacterial disease

It is possible that actinomycetoma causative agents may share virulence mechanisms with other, more widely studied actinobacterial pathogens, responsible for diseases with similar pathologies. Such diseases are described in detail below.

1.4.1 Nocardiosis

Nocardiosis is an umbrella term for disease caused by infections with *Nocardia* spp. of bacteria (Wilson, 2012). These same bacterial species are actinomycetoma pathogens and *Nocardia*-induced mycetoma is also described as chronic subcutaneous nocardiosis in the literature (Ambrosioni et al., 2010). The summary of actinomycetoma above therefore also applies to subcutaneous nocardiosis.

Where there is divergence from other mycetoma pathogens is that 50-70% of *Nocardia* infections are pulmonary, rather than subcutaneous, due to inhalation into the lungs being a common route of infection (Ambrosioni et al., 2010, Wilson, 2012). The clinical profile of pulmonary nocardiosis is highly similar to that of tuberculosis (TB), presenting with pneumonia, necrotic abscesses, lesions and granulomas (Ambrosioni et al., 2010, Fujita et al., 2016). The granulomas have been observed to be comprised of macrophages, lymphocytes and various granulocytes (Fujita et al., 2016). More specific observations of the molecular pathology of the disease and the nature of its granuloma response are non-existent in the literature, due to lack of study.

Of relevance to both nocardiosis and actinomycetoma is that at least one *Nocardia* spp., namely *N. terpenica* (formerly *N. brasiliensis* IFM 0406), is known to produce an immunosuppressive compound named brasilicardin A (Shigemori et al., 1998, Komaki et al., 1999, Komatsu et al., 2005, Hayashi et al., 2008, Schwarz et al., 2018). This compound exerts its effect by blocking progression through the cell cycle in lymphocytes and inhibiting amino acid uptake (Usui et al., 2006). Brasilicardin A production has not been identified in other species and so it appears to be a novel product of only *N. terpenica* (Zijlstra et al., 2016). The immunosuppressive activity of this compound would no doubt contribute to the establishment and progression of *N. terpenica* infections, whether they be pulmonary nocardiosis or actinomycetoma. However, the compound's lack of distribution across the genus and the broader Actinobacteria phylum suggests it is not the central mediator of either of these diseases.

1.4.2 Actinomycosis

As with nocardiosis, actinomycosis is an umbrella term, encapsulating all infections by species of *Actinomyces* bacteria (Kononen and Wade, 2015). Its causative agents, such as *A. israelii* are listed as actinomycetoma pathogens (van de Sande, 2013). The description of actinomycosis also matches that of actinomycetoma, namely that it is a chronic, granulomatous infection, caused when bacteria are inoculated into deep tissues via trauma (Kononen and Wade, 2015). Histological images also show the formation of bacterial grains, as in actinomycetoma. Therefore, this disease is considered here to be the same as actinomycetoma.

1.4.3 Tuberculosis

According to the WHO, 10 million people fell ill with TB in 2019, with 1.4 million dying, putting this disease in the top ten causes of death globally. Resulting from an infection of the lungs by the Actinobacteria species *Mycobacterium tuberculosis*, the disease presents with classic symptoms of a chronic cough and haemoptysis, among others (Zumla et al., 2013). Despite its high mortality rate, *M. tuberculosis* is contained as an asymptomatic, latent infection in 90% of infected individuals.

Once in the lungs, *M. tuberculosis* first infects alveolar macrophages and triggers a pro-inflammatory response (Wilson et al., 2019). Additional immune cells are recruited to the infection site, leading to epithelioid granuloma formation (Peyron et al., 2008). Studies have shown that this process is manipulated by *M. tuberculosis*. The bacteria secrete a protein, ESAT-6, which induces expression of MMP9 and IL8 in epithelial cells adjacent to infected macrophages (Volkman et al., 2010, Boggaram et al., 2013). These proteins are thought to promote granuloma formation in this disease, IL8 being a chemokine that draws cells to the site and MMP9 degrading the ECM and activating other chemokines. ESAT-6 has also been shown to initially induce a pro-inflammatory phenotype in resting or anti-inflammatory macrophages in early stages of infection, again promoting granuloma formation (Refai et al., 2018).

In line with this activity, TB granulomas have been found to have pro-inflammatory cores, which are also usually necrotic (Marakalala et al., 2016). Two mechanisms underpin this necrosis. Firstly, *M. tuberculosis* triggers excessive production of TNFα and ROS, resulting in rapid necroptosis of host cells (Roca and Ramakrishnan, 2013). Secondly, the bacteria produce

necrosis-inducing factors, such as the toxin CpnT or the type VII secretion system ESX-1 (Danilchanka et al., 2014, Beckwith et al., 2020). Necrosis favours the bacteria as death of infected macrophages releases them into the extracellular space (Roca and Ramakrishnan, 2013). This is significant as the tissue layers surrounding the necrotic core of the granuloma have been shown to have an anti-inflammatory environment (Marakalala et al., 2016, Beckwith et al., 2020, Cronan et al., 2021). Such an environment is growth permissive for the bacteria and allows transmission of the infection to new cells.

The anti-inflammatory outer zone partly results from the presence of eosinophils and T cells producing the anti-inflammatory cytokines IL4 and IL13 (Cronan et al., 2021). Studies have also shown that *M. tuberculosis* antigens are able to induce expression of IL10, via TLR2 signalling (Cyktor and Turner, 2011). Indeed, blocking IL10 activity in CBA/J mice actually improved survival during infection with *M. tuberculosis* (Beamer et al., 2008). Additionally, infected macrophages have been observed to switch from an initial pro-inflammatory phenotype, induced within 8 hrs, to an anti-inflammatory phenotype after 24 to 48 hrs (Shi et al., 2019). This so-called 'biphasic' response has been attributed to the activity of ESAT-6, which boosts IL10 production (Refai et al., 2018).

Recruitment of cells to a granuloma also favours TB latency. *M. tuberculosis* can induce macrophages to differentiate into foam cells, which are commonly found in granulomas, as described above (Peyron et al., 2008). The bacteria produce mycolic acids, which induce the differentiation into foam cells following phagocytosis. Foam cells are unable to digest the engulfed bacteria or carry out further phagocytosis. *M. tuberculosis* can then lie dormant within these cells, safe from the immune response and with a nutrient source in the form of the triglyceride-rich organelles of the foam cells.

These features exemplify how *M. tuberculosis* has evolved numerous mechanisms which enable it to manipulate host immunity. It promotes granuloma formation to facilitate its own cell-to-cell transmission and latency. It then dysregulates the transition to an anti-inflammatory state to maintain its replicative niche and own persistence. The granuloma response therefore appears to benefit the pathogen. However, adding further complexity is the fact that complete inhibition of granuloma formation in mice has been shown to significantly enhance bacterial growth and lead to negative outcomes for the host (Cronan et al., 2021).

1.4.4 Leprosy

Caused by infections of the obligate intracellular pathogens *Mycobacterium leprae* and *Mycobacterium lepromatosis*, leprosy presents with symptoms of cutaneous lesions and peripheral neuropathy (Mi et al., 2020, Scollard et al., 2015, Serrano-Coll et al., 2018). It is a spectrum of disease, defined by two extremes, tuberculoid (TT) and lepromatous (LL), with a range of phenotypes in between. Regardless of the disease phenotype, macrophages and myelin-producing Schwann cells are specifically targeted for invasion by the bacteria (Serrano-Coll et al., 2018).

TT leprosy, also called 'paucibacillary', is characterised by a strong cell mediated immune response, the formation of epithelioid granulomas and low bacterial load (Fonseca et al., 2017, Pagan and Ramakrishnan, 2018). The granulomas restrict bacterial spread and replication and have a strong pro-inflammatory profile, with raised levels of TNFα, IFNγ, IL2 and MMPs 2-9 (Scollard et al., 2015, Fonseca et al., 2017, Pagan and Ramakrishnan, 2018, Serrano-Coll et al., 2018, Mi et al., 2020). This response promotes immune cell recruitment and direct killing of bacteria. It is thought that neuropathy in this form of the disease is not directly caused by the pathogens, but instead by the immune response (Serrano-Coll et al., 2018). By disrupting the tissues in which they form, granulomas inhibit normal biological function, which in the case of leprosy infections localised to peripheral nerves inevitably causes nerve damage (Fonseca et al., 2017). Additionally, the pro-inflammatory environment triggered by TT leprosy will cause the death of host cells colonised by the pathogens. For a leprosy infection, this means the death of Schwann cells, leading to nerve demyelination, the first stage of neuropathy.

LL leprosy, also known as 'multibacillary', features a strong humoral immune response, little to no granuloma formation and high bacterial load (Fonseca et al., 2017, Pagan and Ramakrishnan, 2018). The pathogens can more freely replicate and disseminate, with immune cells polarised to an anti-inflammatory profile, induced and sustained by high levels of IL4, IL5 and IL10 produced by infected macrophages (Scollard et al., 2015, Fonseca et al., 2017, Mi et al., 2020). These macrophages also differentiate into foam cells, reducing their bactericidal abilities further. In the LL form, neuropathy appears to be directly caused by the bacterial infection of the nerve, due to the higher bacterial load (Fonseca et al., 2017). *M. leprae* produces phenolic glycolipid, which stimulates excessive NO production in macrophages (Madigan et al., 2017). High levels of NO in the tissue environment damages the mitochondria

of nerve axons and initiates demyelination. Additionally, the bacteria can induce lipid droplet formation in Schwann cells, diverting intracellular lipids away from the myelin biosynthetic pathway (Mietto et al., 2020). Reduced myelin production then further contributes to demyelination of the nerve. Schwann cell de-differentiation has also been proposed as another method of demyelination, triggered by the pathogens binding to the host plasma membrane receptor ERBB2 (Serrano-Coll et al., 2018).

Cytokine profiles clearly determine the disease phenotype of leprosy. However, the molecular trigger or triggers produced by leprosy pathogens which determine the set of cytokines expressed, and in turn the disease phenotype, remain unknown.

1.4.5 Buruli ulcer

Buruli ulcer is a neglected tropical disease, occurring in Asia, South America, Africa and Australia and caused by a subcutaneous infection with the soil-dweller *Mycobacterium ulcerans* (Silva et al., 2009, Adusumilli et al., 2005, Van der Werf et al., 1999). The disease initially develops as a firm, painless nodule, before then progressing to ulceration, involving severe necrosis of cutaneous and subcutaneous tissues and the sloughing of skin (George et al., 1999, Van der Werf et al., 1999). Development of the necrotic lesions is slow, attributed to the slow growth rate of the pathogen, and the infection can persist for anywhere up to a decade, showing immunosuppressive character (George et al., 1999, Silva et al., 2009, Adusumilli et al., 2005). Additionally, despite the extensive necrosis, patients report the lesions to be painless (George et al., 1999).

At the cellular level, Buruli ulcer lesions are characterised by the presence of extracellular colonies of *M. ulcerans* within a necrotic core, surrounded by a ‘belt’ of immune cells, which form a barrier between necrotic and healthy tissues (George et al., 1999, Ruf et al., 2017). The necrotic core itself expands beyond the infection site, which was suggested to be consistent with the free diffusion of a specialised metabolite (George et al., 1999). Further investigation led to the discovery of the toxin mycolactone, produced by *M. ulcerans*.

With its biosynthetic genes encoded on a 174 kb giant plasmid, mycolactone has been shown in numerous studies to have cytotoxic, immunosuppressive and analgesic activities and thus accounts for the defining characteristics of Buruli ulcer pathogenesis (George et al., 1999, Stinear et al., 2004, Adusumilli et al., 2005, Silva et al., 2009, Marion et al., 2014, Song et al.,

2017, Demangel and High, 2018, Foulon et al., 2020). Cytotoxicity has been demonstrated against multiple cell types and is mediated by mycolactone disrupting the cell membrane and increasing its permeability (Adusumilli et al., 2005, Silva et al., 2009, Foulon et al., 2020). A recent study showed this membrane disruption induces production of ROS, which in turn activates intracellular sensors called inflammasomes (Foulon et al., 2020). Once activated, inflammasomes trigger an inflammatory form of cell death termed pyroptosis.

Buruli ulcer is not a granulomatous disease. However, mycolactone negative *M. ulcerans* mutants have been shown to produce an intracellular, pro-inflammatory infection which induces granuloma formation (Adusumilli et al., 2005).

1.5 Other granulomatous diseases

Beyond actinobacterial diseases, there are parasite-induced tropical diseases which also feature granuloma formation as part of their pathology. While the fundamental biology of these protozoan and multicellular parasites differs from that of the bacterial actinomycetoma pathogens, it is possible that they share a mechanism for the induction of granuloma formation. These diseases are examined in detail below.

1.5.1 Leishmaniasis

Leishmaniasis is caused by an infection of various *Leishmania* spp., which are protozoan parasites (Kaye and Beattie, 2016, Rossi and Fasel, 2018, Giorgio et al., 2020). It has a wide spectrum of clinical manifestations depending on the causative species. A common observation across infections with all species is the formation of granulomas (Kaye and Beattie, 2016). The spectrum of disease that comprises this condition can be broadly classified into two groups: cutaneous leishmaniasis (CL) and visceral leishmaniasis (VL).

CL is characterised by the appearance of ulcerating skin lesions (Rossi and Fasel, 2018). It is generally a self-healing condition (Giorgio et al., 2020). During CL infections, two forms of granulomas have been described (Kaye and Beattie, 2016). The first is comprised of epithelioid cells, pro-inflammatory macrophages, has no neutrophil infiltration, no necrosis and shows clear signs of parasite killing. The second features neutrophil infiltration, a necrotic core and widespread macrophage lysis. Various lymphocytes are present within both types. The mechanisms that determine the form a granuloma will take are unknown.

VL is the deadlier form of the disease and is characterised by dissemination of parasites to the spleen and the liver (Rossi and Fasel, 2018). VL lesions are not self-healing (Giorgio et al., 2020). The location of a VL infection determines whether a granuloma response is induced, with granulomas observed only in the liver and not in the spleen. It is not clear why this is the case. Patients with a robust hepatic granuloma response are often asymptomatic, while mouse models have shown that parasites are able to easily persist in the spleen in the absence of granulomas (Kaye and Beattie, 2016, Giorgio et al., 2020). The establishment of granulomas in either form of leishmaniasis begins first with macrophages engulfing the invading parasites (Kaye and Beattie, 2016). The infected macrophages then fuse either with each other or with non-infected cells to form MGCs, initiating granuloma formation (Kaye and Beattie, 2016). Additional immune cells are then recruited to the site via the release of chemokines.

Studies have found that the anti-inflammatory cytokines IL4 and IL13 are required for the formation of a well-organised leishmanial granuloma. Additionally, depletion studies in mice revealed that the loss of CD8+ T cells impaired granuloma formation, pointing to a role for adaptive immunity in this process (Stern et al., 1988). Once fully formed, the granulomas become dominated by pro-inflammatory cytokine and cell profiles, promoting cell mediated immunity and the death of infected cells (Kaye and Beattie, 2016). Evidence from dog disease models shows that a strong granuloma response results in a lower parasite burden within the host and it is considered that leishmaniasis granulomas are generally protective to the host, rather than damaging (Kaye and Beattie, 2016, Giorgio et al., 2020).

1.5.2 Schistosomiasis

Schistosomiasis is caused by an infection of various *Schistosoma* spp., which are parasitic flat worms (Wilson et al., 2019). Once in the subcutaneous tissues of the host, the worms migrate to the liver and the gut, where they lay eggs. By an unknown mechanism, the immature eggs are immunologically inert, avoiding the recruitment of immune cells to their location (Takaki et al., 2021). Once matured however, parasitic antigens are secreted through the eggshell, attracting macrophages and triggering the formation of an epithelioid granuloma around the egg. Studies have shown that when eggs are in the gut tissues, granuloma formation promotes their extrusion into the gut lumen and thus their excretion in faeces (Hams et al., 2013, Takaki et al., 2021). It has therefore been speculated that *Schistosoma* eggs have evolved specific antigens to promote the granuloma response, in order to facilitate parasite dissemination.

The schistosomiasis granuloma has a predominantly anti-inflammatory environment, with recruited immune cells showing potent upregulation of IL4, IL5, IL13 and IgE (Hams et al., 2013). Development of fibrosis is also very common, as the prolonged anti-inflammatory response triggers excessive wound healing (Hams et al., 2013, Pagan and Ramakrishnan, 2018). This is harmful to the host, due to the growth of fibrotic lesions around granulomas that eventually cause internal bleeding and death (Hams et al., 2013). However, if no granuloma forms, then toxins secreted from the egg disseminate through the host issues and exert their own pathological effects. This results in a more acute and lethal disease and reduced survival in animal models (Hams et al., 2013, Pagan and Ramakrishnan, 2018). Granulomas therefore have both protective and deleterious effects in schistosomiasis.

1.6 Comparisons between mycetoma and other diseases of interest

Comparisons with the above diseases could prove beneficial to understanding mycetoma development. These comparisons are directly made in Table 1.1. Epithelioid differentiation is a commonly shared feature of these diseases, along with the induction of fibrosis around the granuloma. This highlights the possibility that study of mycetoma could give insight into important pathological molecular mechanisms conserved across multiple global diseases. Actinomycetoma stands apart from other bacterial diseases of TB and leprosy in being an extracellular rather than intracellular infection that triggers granuloma formation. It is also interesting to note that in both mycetoma and schistosomiasis, granulomas build around large bodies embedded in the host tissues, namely the grain in the case of mycetoma and the parasite egg for schistosomiasis. Such bodies are absent in the other diseases. It could therefore be speculated that the mechanisms of granuloma formation in mycetoma and schistosomiasis have the most in common with each other, rather than with the other listed diseases. They may be closer to the original foreign body response that granuloma formation evolved as.

Additionally, while it has been found that mycetoma initially induces a pro-inflammatory response, its granuloma becomes dominated by anti-inflammatory cytokines. This again matches with schistosomiasis. For the other diseases, the cytokine environment is more mixed or is pro-inflammatory. Finally, mycetoma stands out with leprosy and Buruli ulcer in having analgesia as part of its pathology. Leprosy and Buruli ulcer pathogens produce this effect via differing mechanisms, despite all being *Mycobacterium* spp. This demonstrates that there are

diverse evolutionary pathways for bacteria to develop such mechanisms, even within the same genus. This is worth considering when trying to identify conserved virulence systems within mycetoma pathogens.

Disease	Type of causative agent	Intra- or extracellular infection	Painless development	Granulomatous	Epitheiloid granuloma	Granuloma protective or harmful	Fibrosis	Necrosis	Cytokine environment
Mycetoma	Bacteria/ Fungi	Extracellular	Yes	Yes, surrounding grains	Yes	Likely both	Yes, around granuloma	No	Initial pro-inflammatory, switching to anti-inflammatory
Nocardiosis (pulmonary)	Bacteria	Unclear	No	Yes	Yes	Unclear	Yes, around granuloma	Yes, at granuloma core	Unclear
Actinomycosis	Bacteria	Extracellular	No	Yes, surrounding grains	Yes	Unclear	Yes, around granuloma	No	Unclear
Tuberculosis	Bacteria	Intracellular	No	Yes	Yes	Both	Yes, around granuloma	Yes, at granuloma core	Pro-inflammatory granuloma core, but anti-inflammatory outer layers
Leprosy	Bacteria	Intracellular	Yes, via neuropathy	Yes, but only TT form, not LL.	Yes	Both	Yes, around granuloma	Yes, but not within the granuloma	TT form is pro-inflammatory, LL form is anti-inflammatory
Buruli ulcer	Bacteria	Extracellular	Yes, via mycolactone activity	No	n/a	n/a	No	Yes, via mycolactone activity	Triggers pro-inflammatory cell death, but suppresses immunity
Leishmaniasis	Protozoan parasite	Intracellular	No	Yes	Yes	Protective	No	Some granulomas are necrotic, others are not	Granulomas dominated by pro-inflammatory cytokines
Schistosomiasis	Parasitic flatworm	Extracellular	No	Yes, around worm eggs	Yes	Both	Yes, around granuloma	No	Anti-inflammatory within granuloma, tirggers excessive wound healing

Table 1.1 Comparison of wider actinobacterial and granulomatous diseases to mycetoma.

1.7 Natural Products

It has been well reported that microbial pathogens have evolved to secrete products which can modulate aspects of the host response, such as immune pathways or wider metabolic processes, in order to promote their own survival (Rahman and McFadden, 2011, Asrat et al., 2015, Friedrich et al., 2017, Bussi and Gutierrez, 2019, Konig et al., 2021). Actinobacterial pathogens are no different in this regard, as shown by *N. terpenica*, described above, which produces the immunosuppressant brasilicardin A (Shigemori et al., 1998). This is also demonstrated in Buruli ulcer, where secreted mycolactone mediates much of the disease pathology (George et al., 1999). As it is possible that the pathogenic Actinobacteria responsible for actinomycetoma are able to disrupt the immune response, it may be that this effect is achieved through the production of a compound by the bacteria. It could therefore be useful to explore the repertoire of molecules that actinomycetes are known to produce.

1.7.1 Actinomycetes and natural products

Actinomycetes are Gram-positive, filamentous bacteria that have been found in both soil and aquatic environments and also as pathogens and commensals within plant and animal species (Barka et al., 2016). Their life cycle involves multiple morphological changes, with initial growth from spores into a vegetative mycelium, followed by the development of aerial hyphae, which will then produce more spores for distribution.

Actinomycetes became of significant interest to the scientific community when, in the 1940s, the antibiotic streptomycin was discovered by Waksman and colleagues, who isolated it from *Streptomyces griseus* (Schatz et al., 1944). Penicillin had already demonstrated the utility of natural products derived from microbes to the world, but the actinomycete screening approach used by Waksman and his group led to a shift in industry that facilitated the discovery of multiple new products, the majority of which were sourced from actinomycetes (Katz and Baltz, 2016, De Simeis and Serra, 2021).

In this context, the phrase ‘natural products’ refers to specialised metabolites. These are molecules produced by an organism, granting it an advantage within its ecological niche, but which are not essential to its survival when being grown under laboratory conditions (Katz and Baltz, 2016). They are very chemically diverse molecules (van der Heul et al., 2018). There are wide ranges of molecular weights, numbers of stereo-specific centres and differing levels of complexity. The list of applications is long and varied also, with uses as antibiotic, anti-

parasitic, anti-cancer and immunosuppressive drugs all well known (Hoskisson and van Wezel, 2019).

Production of individual specialised metabolites within actinomycetes is regulated by genes arranged in clusters within the chromosome and on plasmids, termed biosynthetic gene clusters (BGCs) (Rudd and Hopwood, 1980, Arakawa, 2010). The earliest evidence for the existence of BGCs came in 1979, during the study of the biosynthesis of actinorhodin within *Streptomyces coelicolor*, when it was observed that mutations to each biosynthetic gene could be mapped between the same closely linked loci on the chromosome (Rudd and Hopwood, 1979). At the turn of the century, with the advancement of sequencing technology, the whole genome sequence for *S. coelicolor* was published and with it came the revelation that there were 18 previously unknown BGCs present within the genome, which appeared to be transcriptionally silent (Bentley et al., 2002).

Genome sequencing of more species over time revealed that, generally, actinomycetes had the ability to produce ten times as many specialised metabolites as was previously estimated, with genomes containing as many as 20-50 BGCs (Barka et al., 2016, Katz and Baltz, 2016, Hoskisson and Seipke, 2020). The majority of these clusters do not appear to be expressed under standard laboratory conditions, hence the underestimation of the true repertoire of these organisms. For these transcriptionally silent BGCs, and even for many that have been shown to be expressed, their true functions remain unknown and their products are 'cryptic' (Hoskisson and Seipke, 2020). This means that, despite actinomycetes being exploited for nearly the last 80 years, there is still great potential in terms of novel natural products to be unlocked from within the genomes of these bacteria (Kim, 2021).

1.7.2 Non-ribosomal peptide synthetases

Non-ribosomal peptide synthetases (NRPSs) are large enzyme assemblies responsible for the biosynthesis of certain peptides (Jaremko et al., 2020). Such peptides often contain amino acids from outside the repertoire of the 20 proteinogenic amino acids used in the assembly of proteins during translation at the ribosome. These assemblies were first discovered in the 1960s, when it was found that dosing *Bacillus* cells with ribosome inhibitors and RNases did not inhibit the synthesis of certain peptides, thus indicating the existence of peptide synthesis distinct from the ribosome (Mach et al., 1963, Dell et al., 2021).

NRPSs are modular in nature, with each module incorporating one amino acid into a polypeptide chain (Finking and Marahiel, 2004). The modules are themselves made up of domains, which are classified by the particular enzymatic function they carry out within the module. There are three types of domain that must be present within a module for it to have minimum functionality; the adenylation (A)-domain, the peptidyl carrier protein (PCP)-domain and the condensation (C)-domain (Finking and Marahiel, 2004, Felnagle et al., 2008, Payne et al., 2016).

The A-domain has selectivity for certain amino acids and upon binding will convert it to an aminoacyl AMP adenylate, in a reaction that consumes an ATP and binds the monomer to the domain (Finking and Marahiel, 2004, Payne et al., 2016). The now activated monomer can then be accepted by a neighbouring PCP-domain, which binds it via a 4' phosphopantetheine cofactor. The C-domain is responsible for peptide bond formation and catalyses this process between one amino acyl group downstream of it in the module and another amino acyl group upstream of it. A dipeptide is then formed, which can then participate in a condensation reaction at the next C domain, thus building up a polypeptide. This scheme is shown in Figure 1.11.

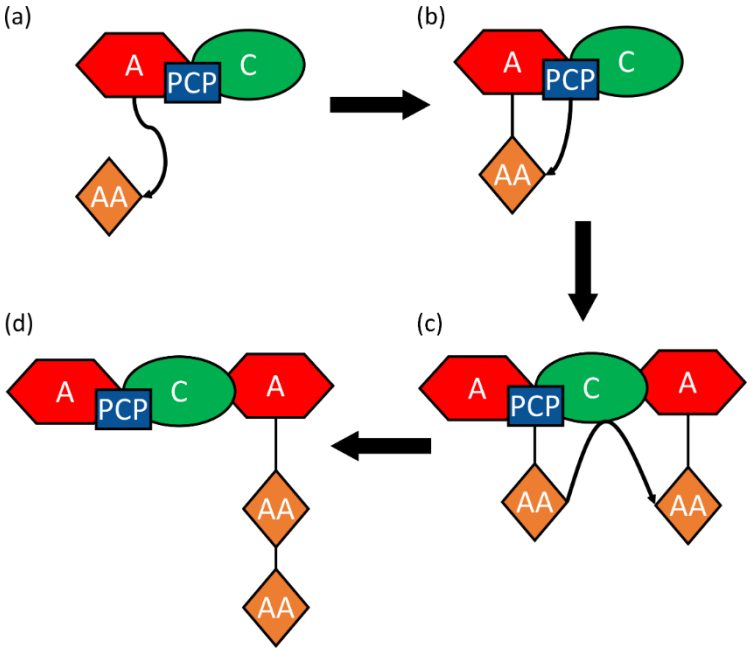


Figure 1.11 Reaction scheme for an NRPS module assembling amino acids into a polypeptide chain. **(a)** An A-domain recognises and binds an amino acid. **(b)** The PCP-domain accepts the activated amino acid from the A-domain. **(c)** The PCP-domain carries the substrate across to a C-domain, which catalyses a condensation reaction between the downstream substrate and an upstream substrate on an A-domain in a neighbouring module. **(d)** A dipeptide is formed and the process repeats to build it up to a polypeptide.

Aside from these three core domains, there are also additional “tailoring domains”, for making modifications to the polypeptide, such as glycosylation and methylation (Payne et al., 2016). The two most abundant of these domain types are the epimersation (E)- and thioesterase (TE)-domains. An E-domain will change the stereochemistry of any of the residues present in the polypeptide, while the TE-domain is normally present in the final module of an NRPS assembly and will release the chain from the assembly, finishing its synthesis. A key feature of NRPSs is that they are not restricted to using just twenty amino acids, as ribosomes are, but instead have access to at least 300 different substrates, meaning wide structural diversity can be produced (Felnagle et al., 2008). Specialised metabolites synthesised by these systems include the antibiotics vancomycin, shown in Figure 1.12(a), capreomycin and all β-lactams; the bleomycin family of anti-cancer drugs; and the immunosuppressant cyclosporin A, visible in Figure 1.12(b).

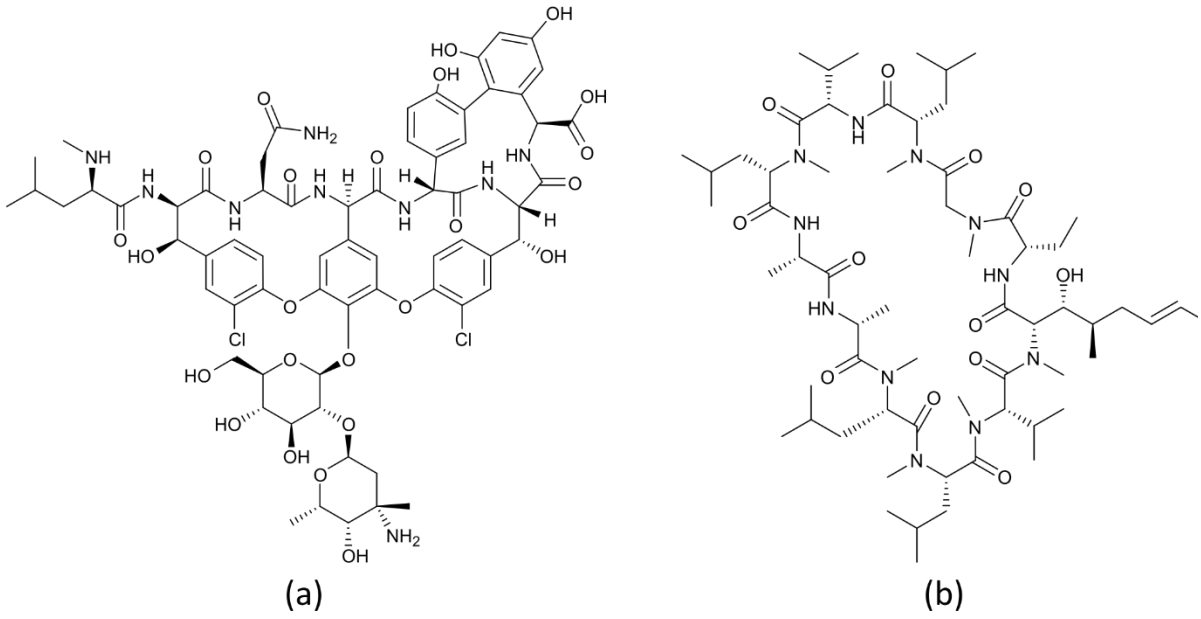


Figure 1.12 Example structures of two compounds synthesised by NRPS systems. **(a)** Vancomycin, an antibiotic. **(b)** Cyclosporin A, an immunosuppressant. Their structures serve to highlight the great structural complexity that can be achieved by NRPSs.

1.7.3 Polyketide synthases

Polyketide synthases (PKS) are very large protein assemblies, comprised of modules, which in turn are made up of domains (Smith et al., 2021). In this regard, they are very structurally similar to NRPS assemblies. Evidence of their existence as massive assemblies came in 1990,

when genetic analysis of the erythromycin gene cluster revealed a gene for a synthase encoding one large polypeptide (Cortes et al., 1990).

As in an NRPS, each PKS module has three core domains which are needed for minimum functionality; an acyl transferase (AT)-domain, a ketosynthase (KS)-domain and an acyl carrier protein (ACP)-domain (Weissman, 2015). The AT-domain binds and activates an acyl-CoA substrate and is what gives the assembly its selectivity. The activated acyl group is then transferred to the ACP-domain, via formation of a thioether link (Hertweck, 2009). Once bound to the ACP-domain, a neighbouring KS-domain will then catalyse a condensation reaction between the original substrate and a second substrate on an upstream ACP-domain (Weissman, 2015). A polyketide chain is thus formed. This scheme is shown in Figure 1.13.

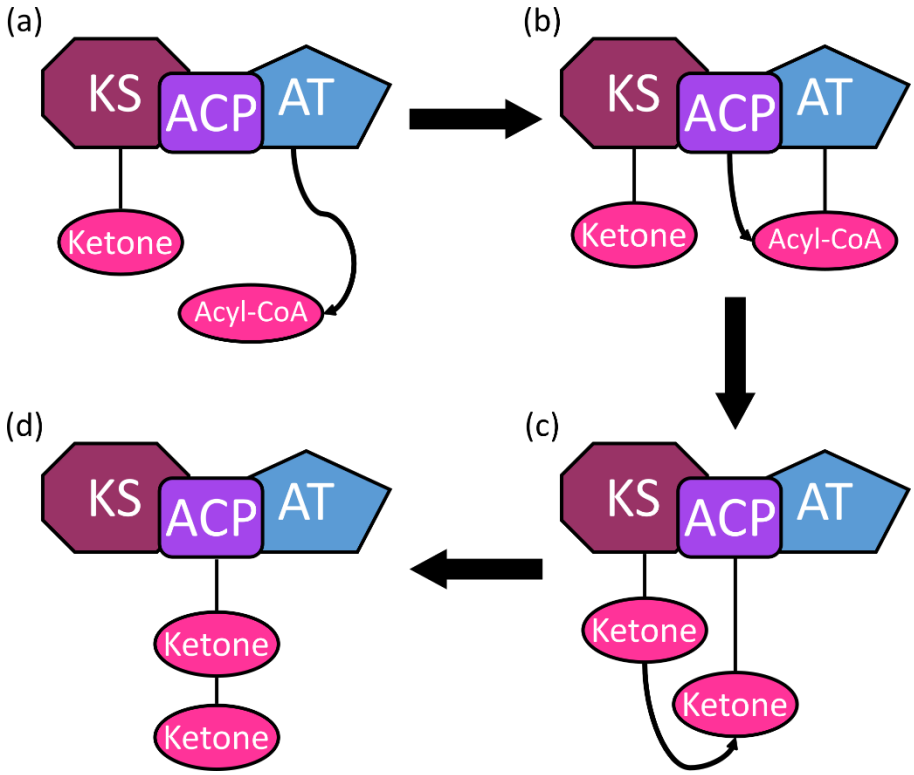
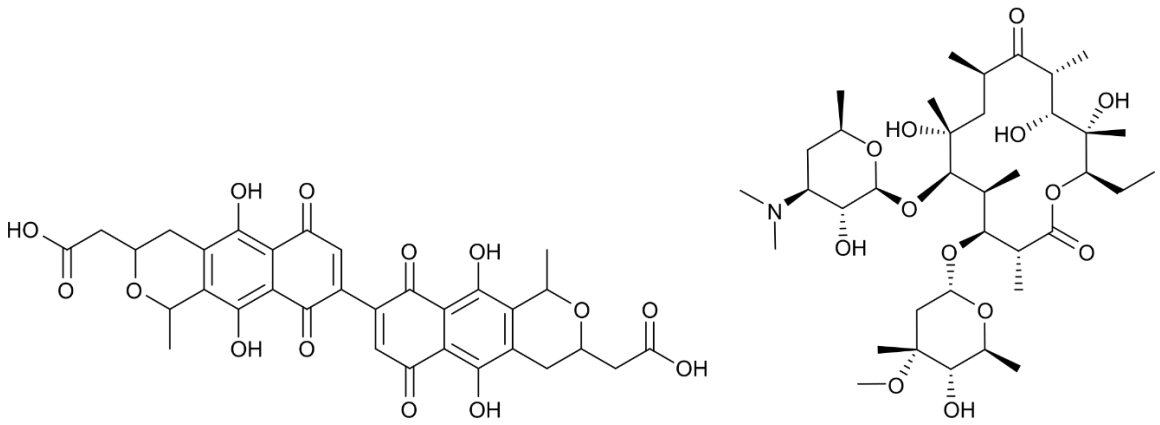


Figure 1.13 Reaction scheme for a PKS module assembling ketones into a polyketide chain. **(a)** AT-domain binds and activates an acyl-CoA substrate. **(b)** Acyl group (a ketone) is transferred to an ACP-domain. **(c)** KS-domain within the module, already loaded with a ketone, catalyses a reaction between its substrate and the second ketone bound to the ACP-domain. **(d)** A polyketide is formed and moves to the next module for further elongation.

The keto groups on the chain can be modified by additional modules, such as ketoreductase, dehydratase or enoyl reductase. These are analogous to the tailoring domains of an NRPS. In a further similarity with NRPSs, the final polyketide chain is released from the PKS by the activity of a TE domain. Once a polyketide skeleton has been released it undergoes cyclisation and other additional modifications, such as glycosylation, forming the final compound structure (Hertweck, 2009).

PKSs are classified into three types, based on their constituent parts (Hertweck, 2009, Shimizu et al., 2017). Type I systems are comprised of subunits covalently bound together, which themselves are made up of the different domains described above, with the essential KS-, ACP- and AT-domains always present. Type II PKSs are formed from loosely associated subunits, each of which only contain KS- and ACP-domains. Each subunit therefore only has a single function. A type III PKS is a homodimer, with only a KS-domain in its subunits, which carries out all synthesis functions.

The polyketide products synthesised by these assemblies include a diverse range of compounds, such as erythromycin (antibiotic) (Fig. 10(b)), actinorhodin (antibiotic) (Fig. 10(a)), avermectin (anti-parasitic), rapamycin (immunosuppressant) and amphotericin (anti-fungal) (Hertweck, 2015).



(a) (b)

Figure 1.14 Example structures of two compounds synthesised by PKS systems. **(a)** Actinorhodin, which is synthesised by a type II PKS. **(b)** Erythromycin, which is synthesised by a type I PKS. Both are antibiotics.

1.7.4 Ribosomally synthesised and post-translationally modified peptides

Abbreviated as RiPPs, this group of natural products contains 22 sub-families, e.g. the lanthipeptides, classified based on their structures and biosynthesis routes and is therefore an umbrella term for an extremely large and diverse set of compounds (Ortega and van der Donk, 2016). This diversity is derived from extensive post-translational modifications (PTMs), which grant these compounds chemical properties not available to unaltered peptides and are a defining trait of this class (Arnison et al., 2013). In addition to making compounds more chemically versatile, PTMs also make these peptides more resistant to proteolysis and decrease the entropy of a compound binding to its target (Hetrick and van der Donk, 2017).

While the specifics of the process differ between each sub-family, there is a general biosynthetic pathway that is common across all RiPPs, which is termed 'post-ribosomal peptide synthesis' (Arnison et al., 2013). This process is summarised in Figure 1.15.

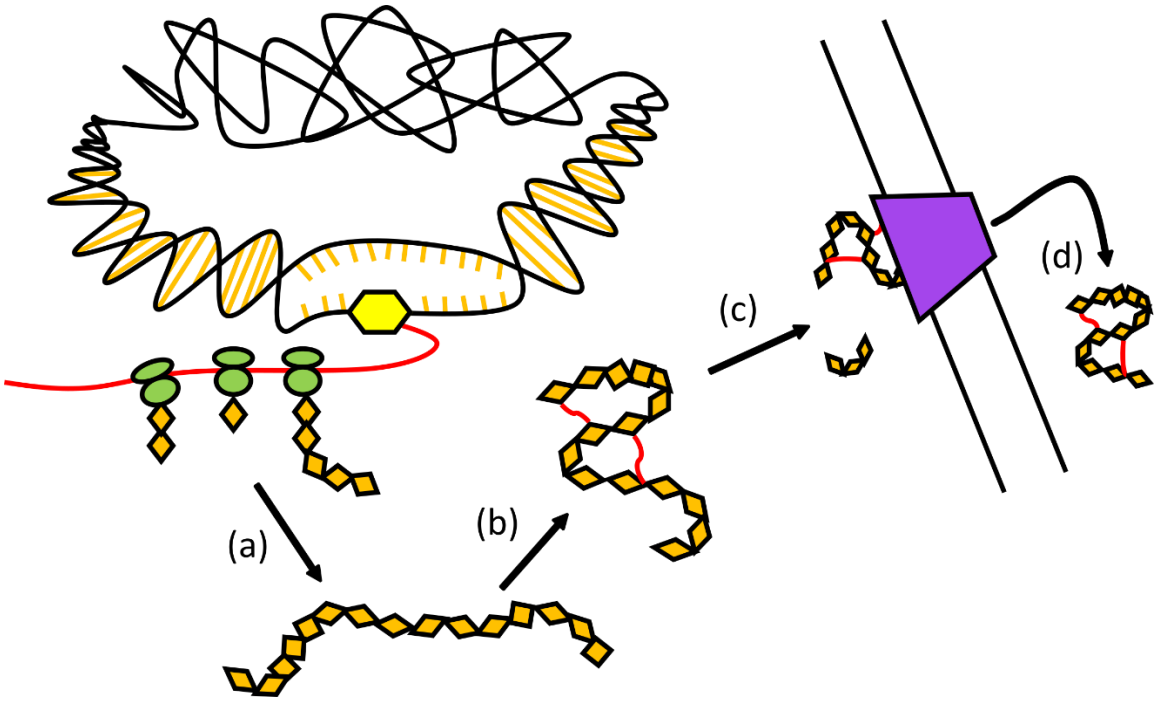


Figure 1.15 Biosynthesis of a RiPP by a bacterial cell. **(a)** The BGC is transcribed by RNA polymerase from the bacterial DNA and translated by ribosomes to produce a nascent polypeptide chain. **(b)** The polypeptide undergoes significant post-translational modifications, for example the formation of thioester linkages to produce lanthionine. **(c)** The modified peptide has its leader sequence cleaved and is transported across the plasma membrane by an ABC transporter. **(d)** The peptide is now fully active.

Compounds are initially synthesised by the ribosomes as a precursor peptide, which can be divided into a leader peptide and a core peptide (Hubrich et al., 2021). The core peptide is the location of all of the PTM sites and so is the section of the precursor that will actually become the final RiPP compound (Arnison et al., 2013, Ortega and van der Donk, 2016). The leader peptide is located at the N-terminus of the core peptide, although some rare examples do exist where it is located at the C-terminus, and has various roles in biosynthesis. It is recognised by the biosynthetic enzymes responsible for making PTMs, allowing them to bind the core peptide and begin modification. The leader peptide must then be cleaved to produce the final RiPP structure and render the peptide active. For a large number of RiPPs, such cleavage has been shown to be directly linked to export of the compound from the cell via an ATP-binding cassette transporter, which has associated peptidase activity (Ortega and van der Donk, 2016).

Export only occurs once all PTMs have been made. These modifications can include glycosylation, thioether cross-link formation, amide bond formation, dehydration and cyclisation to name a few (Truman, 2016). Each contributes to the final physical structure of a RiPP. Gene clusters encoding different RiPPs can show homology to one another, owing to the fact that in many cases the same enzymes are used to make certain sets of PTMs (Arnison et al., 2013). This makes characterisation of novel RiPPs via *in silico* genome analysis much simpler. Examples of RiPPs produced by *Streptomyces* spp. include the antibiotics duramycin and cinnamycin, the structure of which is shown in Figure 1.16. (Knerr and van der Donk, 2012).

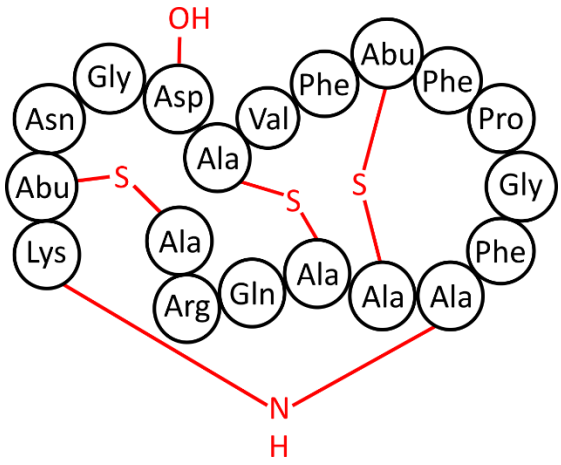


Figure 1.16 Structure of the RiPP cinnamycin, shown as its constituent amino acids. PTMs are highlighted in red. The tertiary structure of cinnamycin is held together by these modifications, highlighting the importance of their addition during RiPP biosynthesis.

1.7.5 Isoprenoids

The isoprenoids comprise the largest group of compounds within the natural world and are produced by both eukaryotes and prokaryotes alike (Dairi, 2005, Daum et al., 2009, Malico et al., 2020). They have been put to a wide variety of uses as antibiotics, anti-cancer drugs, insecticides, food flavourings and plant hormones. All consist of the same basic units of an isopentenyl diphosphate and a dimethylallyl diphosphate joined together via a condensation reaction to form a polyprenyl chain, as shown in Figure 1.17 (Daum et al., 2009). The chain then undergoes cyclisation by cyclase enzymes, forming the final compound structure.

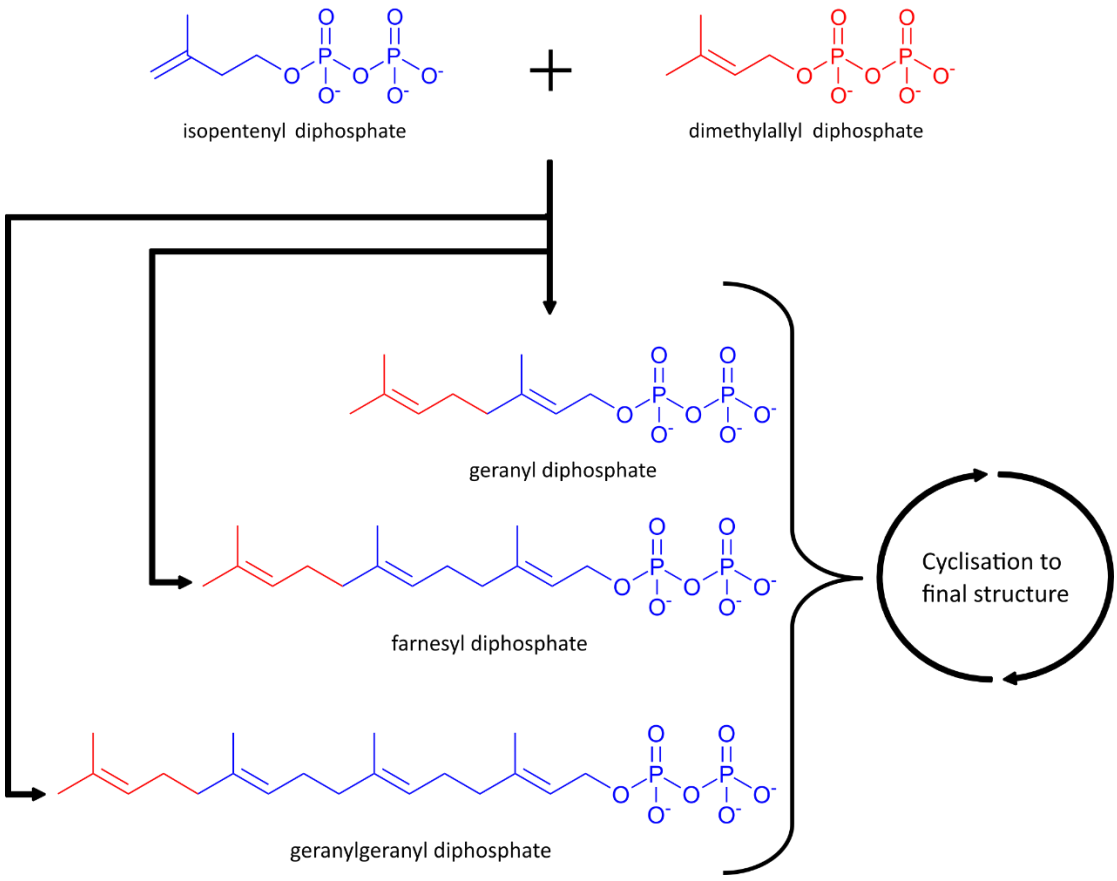


Figure 1.17 Precursors involved in the biosynthesis of isoprenoids. Isopentenyl diphosphate (highlighted in blue) and its isomer dimethylallyl disphosphate (highlighted in red) are the starting units of all isoprenoids. They can be bonded together to form one of 3 polyprenyl chains; geranyl diphosphate, farnesyl diphosphate or geranylgeranyl diphosphate. These chains then either undergo cyclisation and other modifications to form the final compound structure or can themselves be bonded together and then cyclised, to produce even larger compounds.

Historically, plants were the primary source for these compounds, with bacteria not generally considered to be prolific producers (Smanski et al., 2012). A steady stream of useful isoprenoids were obtained from actinomycetes over the years however, such as novobiocin in 1956, an antibiotic which features an isoprenoid side chain, and brasilicardin A (Komaki et al., 1999, Durr et al., 2006). The structure of brasilicardin A is shown in Figure 1.18. The advances in genome sequencing over the last two decades and subsequent identification of gene clusters revealed that actinomycetes have a greater capacity for synthesising isoprenoids than was previously estimated (Smanski et al., 2012). They in fact produce quite a sizeable range compared to other prokaryotes. With the rise of antimicrobial resistance, researchers are now turning to this under-utilised isoprenoid repertoire of actinomycetes as a source of desperately needed novel antibiotics.

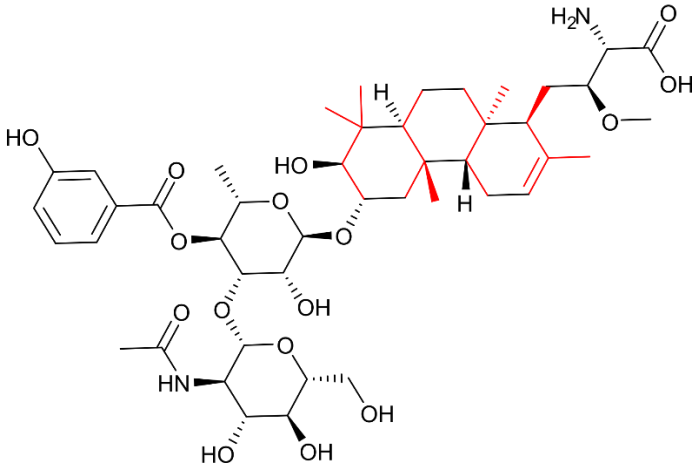


Figure 1.18 Structure of the immunosuppressive compound, brasilicardin A. The cyclised isoprenoid core is highlighted in red.

1.7.6 Aminoglycosides

The well-known antibiotic streptomycin, structure shown in Figure 1.19(a), was the first of this class of natural product to be discovered, during the early 1940s (Yu et al., 2017). These compounds have therefore been in clinical use for a significant period of time. The other antibiotics that later joined streptomycin include kanamycin (Figure 1.19(b)), gentamicin, tobramycin and apramycin (Kudo and Eguchi, 2009).

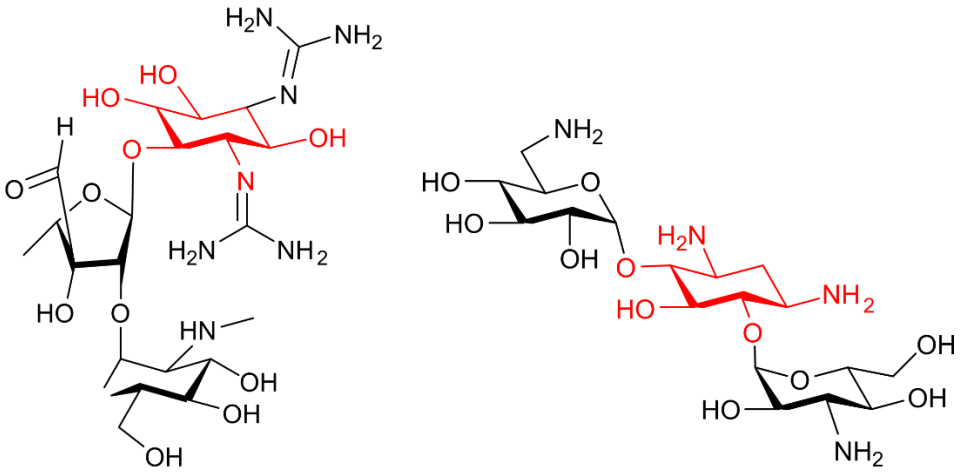


Figure 1.19 Example structures of **(a)** streptomycin and **(b)** kanamycin, two aminoglycoside antibiotics. Highlighted in red are their different aminocyclitol groups, which form the cores of these compounds. Biosynthesis of aminoglycosides begin with these core molecules.

All of these compounds are derived, for the most part, from sugars and despite being very structurally diverse, they do share some general physical features (Kudo and Eguchi, 2016). The core of each compound is comprised of an aminocyclitol (see Fig. 15), a derivative of inositol, which is linked via glycosidic bonds to an aminosugar (Becker and Cooper, 2013, Park et al., 2017). The rest of the structure will then usually contain two or more additional amino groups and several hydroxyl groups. The sharing of general structural features is highlighted by the fact that all of these antibiotics have the same mode of action, in that they all target the bacterial ribosome (Kudo and Eguchi, 2016). Unique features are added onto the general structure during biosynthesis however, creating a diverse set of compounds, which are actually sub-divided into smaller classes based on these features and their biosynthesis (Becker and Cooper, 2013).

In terms of biosynthesis, this is performed by a series of individual enzymes, each encoded within the biosynthetic gene cluster of a compound (Katz and Baltz, 2016). As some structural features are shared across all of the compounds, there are naturally several genes that are conserved across the gene clusters of different aminoglycosides, usually for the generation of common intermediates (Kudo and Eguchi, 2016). This is similar to the biosynthetic enzyme genes of RiPPs and makes the identification of aminoglycoside gene clusters simpler, as genes can be identified via homology searches.

The process of biosynthesis of an aminoglycoside, while being different and specific for each molecule, shares a common homology across the class (Kudo and Eguchi, 2009). A hydroxyl group undergoes oxidation/transamination to an amino group, followed by attachment of glucosamine via glycosylation/deacetylation. A ribose group is added via phophoribosylation/dephosphorylation, followed by glycosylation. A deoxysugar group is then formed by dehydration/reduction and the whole compound then undergoes methylation and epimerisation. This general process serves to highlight the number of enzymes required within a gene cluster to synthesise an aminoglycoside.

The majority of natural products can be grouped into one of the classes described above. This is helpful in enabling the identification of key functional groups and predicting modes of action of newly discovered compounds, as well as predicting the synthesis of compounds from the genomes of organisms. But, there are many examples of natural products that do not fit into one class in particular, for example the antibiotic daptomycin is synthesised by an NRPS-PKS hybrid (Fischbach and Walsh, 2006, Skellam, 2021). Classification can therefore prove to be restrictive and place limits on analysis.

1.8 Aims

Based on the known gaps in knowledge regarding mycetoma, this study had the following overarching aims:

- Investigate the nature of actinomycetoma pathogens' interactions with the immune system.
- Identify potential virulence factors involved in causing the disease state to arise.
- Isolate novel natural products from actinomycetoma pathogens which may play a role in pathogenesis or have therapeutic potential.

Chapter 2. Materials and Methods

2.1 General methods

2.1.1 Chemicals

All chemicals used were sourced from Sigma-Aldrich (now Merck), unless stated otherwise.

2.1.2 Bacterial strains

The wild type (WT) *S. sudanensis* strain used was obtained from the Leibniz Institute DSMZ - German Collection of Microorganisms and Cell Cultures GmbH (see Table 2.1). The non-pathogenic control strains *Streptomyces albus* J1074 and *Streptomyces coelicolor* M1152 (see Table 2.1) were originally sourced from the lab of Andreas Bechthold, University of Freiberg, DE, and the John Innes Centre in Norwich, UK respectively. Strains were maintained on GYEA media (10 g/L glucose, 10 g/L yeast extract, 14 g/L agar), incubated at 30°C. Starter cultures of each strain were set up by adding a loop-full of mycelia to 15 mL GYE media (10 g/L glucose, 10 g/L yeast extract) in a shake flask and incubated for one to three days at 30°C, on a three storey shelf shaker. Glass beads were added to liquid cultures to break up the mycelia. Starter cultures were used as inocula, at a concentration of 5%, to set up larger volume cultures in shake flasks.

Strain	Genetic background	Source
S. sudanensis DSM41923	WT	DSM strain collection
S. albus J1074	WT	Bechthold Lab
S. coelicolor M1152	*Δact Δred Δcpk Δcda rpoB*:C1298T	John Innes Centre

Table 2.1 Bacterial strains used in this study.

2.1.3 Mammalian cell line and tissue culture

RAW-blue cells, a murine macrophage reporter cell line, were obtained from InvivoGen. This cell line is derived from RAW 264.7 murine macrophages and features an NF-κB inducible secreted embryonic alkaline phosphatase (SEAP) reporter integrated into the chromosome (Invivogen, n.d.). The cell line was maintained in Dulbecco's Modified Eagle's Medium (DMEM), with 5% foetal bovine serum (FBS) and 0.1% 2-mercaptoethanol (2-ME), as 1 mL cultures in two separate Corning Costar 24-well plates. Incubation was at 37.5°C and 5% CO_2,

in a Sanyo CO_2 incubator and subculturing was performed every three to four days. 10 mL macrophage cultures were set up by adding 9 mL DMEM + 5% FBS + 0.1% 2-ME and 1 mL cell culture to T25 Nunc EasyFlasks, with incubation for three days at 37.5°C and 5% CO_2.

THP-1 cells, a human monocyte cell line, were generously donated by Professor Matthias Trost, Newcastle University. The cell line was maintained in Roswell Park Memorial Institute (RPMI) 1640 media, supplemented with 5% FBS, 0.1% 2-ME and 2 g/L sodium bicarbonate, as 10 mL cultures in two separate T25 Nunc EasyFlasks. Incubation was at 37.5°C and 5% CO_2, in a Sanyo CO_2 incubator and subculturing was performed every three to four days.

2.1.4 Extraction of high molecular weight genomic DNA from S. sudanensis

15 mL of bacterial cell culture was harvested, centrifuged at 1774 xg for 10 minutes and the supernatant removed. The bacterial biomass was then lysed using a micropestle and resuspended in 5 mL SET buffer (75 mM NaCl, 25 mM EDTA pH 8.0, 20 mM Tris-HCl pH 7.5). This was followed by incubation with 1.5 mg/mL lysozyme for 1.5 hours, at 37°C. 10 µL of 10 mg/mL RNase was then added, followed by incubation at room temperature for 1 minute. Pronase was then added to a concentration of 0.5 mg/mL and SDS to a concentration of 1%, with further incubation at 37°C for 2 hours. Next, 2 mL of 5M NaCl and 5 mL of chloroform was added, followed by incubation at room temperature for 30 minutes, with continuous inversion. Samples were then centrifuged at 4542 xg for 15 minutes, the supernatant harvested and 0.6 vol of isopropanol added to each. After inverting and waiting 3 minutes, the genomic DNA was spooled onto a glass Pasteur pipette, air dried and dissolved in 200 µL sterile water. Samples were then analysed using a NanoDrop ND-1000 spectrophotometer.

2.1.5 Pulsed field gel electrophoresis of extracted gDNA

150 mL of a 0.8% agarose gel was made, using 0.5X tris borate EDTA (TBE). All DNA samples were diluted to a concentration of 50 ng/µL and 100 µL of each was mixed with 100 µL 2% agarose gel and set into plugs. The plugs were placed into the 0.8% gel once it had set. The gel was run at 180 V, for 20 hours, with intervals set to start at 10 seconds and increase up to 70 seconds. Staining solution was prepared by adding 2 mL SYBR Gold to 200 mL 1X TBE. The gel was incubated in staining solution for 20 minutes and then agitated in the solution for a further 20 minutes. It was then visualised.

2.1.6 Testing the use of dialysis tubing for the interaction assay

S. sudanensis was cultured as described in section 2.1.2, then centrifuged at 3270 xg, the old media poured off and the cell pellet resuspended in DMEM to an OD_{600} of 0.1. 2.5 mL of this culture was then transferred into a piece of dialysis tubing (MWCO= 3000kDa), with both of its ends then knotted to seal the tubing. A 10 mL RAW-blue cell culture was prepared as described in section 2.1.3. After incubation, the old culture medium was removed and replaced with 5 mL fresh medium. Macrophages were then suspended into the fresh medium by forcefully tapping the flask on the lab bench, after which a cell count was performed using a haemocytometer. The macrophage culture was then diluted to give a cell density of 1 million cells/mL. 2 mL of macrophage culture was then added to three wells of a Corning Costar six well tissue culture plate. The plate was subsequently incubated for 30 minutes at 37.5°C and 5% CO_2, to allow the macrophages to attach to the surfaces of the wells. The dialysis tubing containing bacteria was then placed into one of the plate wells. Sealed dialysis tubing containing 2.5 mL fresh DMEM was placed in another well, to act as a control. The plate was then incubated at 37.5°C and 5% CO_2 for 72 hrs, with images taken of the wells every 24 hrs.

2.1.7 Testing the use of Transwell tissue culture inserts for the interaction assay

A 10 mL RAW-blue cell culture was prepared as described in section 2.1.3. After incubation, the old culture medium was removed and replaced with 5 mL fresh DMEM. Macrophages were then suspended into the fresh medium by forcefully tapping the flask on the lab bench, after which a cell count was performed using a haemocytometer. The macrophage culture was then diluted to give a cell density of 1 million cells/mL. 2 mL of macrophage culture was then added to wells of a Corning Costar six well tissue culture plate. The plate was subsequently incubated for 30 minutes at 37.5°C and 5% CO_2, to allow the macrophages to attach to the surfaces of the wells. Corning Transwell tissue culture inserts, with pore sizes of 0.4 µm, were then placed into each of the wells. *S. sudanensis* was cultured as described in section 2.1.2, then centrifuged at 3270 xg, the old media poured off and the cell pellet resuspended in DMEM to an OD_{600} of 0.1. 2 mL of this culture was added to inserts in four of the wells; two of the wells contained macrophages and two contained only DMEM. 2 mL DMEM was added to the inserts of the remaining two macrophage-only wells. The plate was then incubated for 72 hours, at 37.5°C and 5% CO_2.

2.1.8 Testing structural integrity of agarose gels for use in the indirect interaction assay

Gels were prepared using deionised water and agarose concentrations of 0.25%, 0.3% and 0.5%. While still melted, 0.5 mL of each gel was pipetted into Transwell inserts and allowed to set. 1 mL of DMEM was then added to the inserts, which were then placed into empty wells on a Corning Costar six well plate and observed for leakage of DMEM through the gels and the insert, into the wells.

2.1.9 Testing agarose gel within Transwell inserts for the interaction assay

S. sudanensis was cultured as described in section 2.1.2, then centrifuged at 3270 xg, the old media poured off and the cell pellet resuspended in DMEM to an OD_{600} of 0.1. A 0.5% agarose gel was prepared using deionised water. While still melted, 0.5 mL of the gel was pipetted into two Corning Transwell inserts and allowed to set. 2 mL DMEM was added to two wells of a Corning Costar six well plate. The inserts containing agarose gel were placed into the two wells and 1 mL of the *S. sudanensis* culture in DMEM was added to one of the inserts. 1 mL CO_2. During incubation, the wells under the inserts were observed for signs of bacterial growth. At the conclusion of the incubation period, a sample of the DMEM from the well beneath the insert was streaked out on a nutrient agar plate and incubated at 37°C for six days. The same was also done for the bacterial culture within the insert, as a positive control.

2.1.10 Testing non-pathogenic control strains for us within the interaction assay

S. coelicolor and *S. albus* were cultured at described in section 2.1.2, then centrifuged at 3270 xg, the old media poured off and the cell pellet resuspended in DMEM to an OD_{600} of 0.1. A 10 mL RAW-blue cell culture was prepared as described in section 2.1.3. After incubation, the old culture medium was removed and replaced with 5 mL fresh medium. Macrophages were then suspended into the fresh medium by forcefully tapping the flask on the lab bench, after which a cell count was performed using a haemocytometer. The macrophage culture was then diluted to give a cell density of 0.5 million cells/mL. 2 mL of macrophage culture was then added to three wells of a Corning Costar six well tissue culture plate. The plate was subsequently incubated for 30 minutes at 37.5°C and 5% CO_2, to allow the macrophages to attach to the surfaces of the wells. A 0.5% agarose gel was prepared using deionised water. While still melted, 0.5 mL of the gel was pipetted into two Corning Transwell inserts and allowed to set. The Transwell inserts were then placed into the three wells containing macrophages. 1 mL fresh DMEM was then added to an insert, as a negative control. 1 mL of

the *S. coelicolor* and *S. albus* were added to the remaining two inserts. The plate was then incubated for 48 hrs, at 37.5°C and 5% CO_2.

2.1.11 Validation of heat killing method for bacteria

S. coelicolor and *S. sudanensis* were cultured at described in section 2.1.2. A sample of each culture was streaked out on nutrient agar plates. The remaining volumes of each culture were then autoclaved in a Prestige Medical Classic benchtop autoclave. A sample of each autoclaved culture was then streaked out on nutrient agar plates. The plates of living and autoclaved bacteria were incubated at 37°C for seven days. This temperature was selected to replicate the temperature bacteria would be incubate at during indirect interaction assay trials.

2.1.12 Indirect interaction assay between bacteria and murine macrophages

A 10 mL culture of RAW-blue macrophages was set up in DMEM + 5% FBS + 0.1% 2-ME media, in a T25 Nunc EasyFlask and incubated at 37.5°C and 5% CO_2, for three days. After incubation, the old culture medium was removed and replaced with 5 mL fresh medium. Macrophages were then suspended into the fresh medium by forcefully tapping the flask on the lab bench, after which a cell count was performed using a haemocytometer. The macrophage culture was then diluted to give a cell density of 1 million cells/mL. 2 mL of macrophage culture was then added to five wells of a Corning Costar six well tissue culture plate, with 2 mL of fresh medium added to the last well. The plate was subsequently incubated for 30 minutes at 37.5°C and 5% CO_2, to allow the macrophages to attach to the surfaces of the wells.

In parallel to this, 15 mL liquid cultures of *S. sudanensis* and *S. coelicolor* were set up in GYE media and incubated on a shaker at 30°C, for three days. At the end of incubation, the OD_{600} of the two cultures was measured using a Thermo Scientific Genesys 20 spectrophotomer. 2 mL of each culture was then transferred to microfuge tubes and the remaining volumes in the flasks were autoclaved in a Prestige Medical Classic benchtop autoclave, to heat kill the bacteria. 100 µL of the autoclaved cultures and of the still living 2 mL cultures in the microfuge tubes was streaked out onto nutrient agar plates, which were then incubated at 30°C, to check for successful heat killing of bacteria and for contamination. 2 mL of each autoclaved culture was then transferred to two microfuge tubes. The aliquots of living and dead bacterial cultures were then centrifuged at 20238 xg, the supernatant removed and the pellets resuspended in the same volume of the macrophage culture medium. The centrifugation and resuspension

steps were then repeated, to wash the bacterial cells. After washing, the living cultures were diluted to an OD_{600} of 0.1 and the dead cultures to an OD_{600} of 2.

Agarose was then prepared, at a concentration of 0.5% in ultrapure water. While melted, 0.5 mL of agarose was pipetted into six Corning Transwell polyester membranes, with 0.4 µm pores, cell culture inserts and allowed to set. Once cooled, the inserts were added to the six wells in the tissue culture plate that had been seeded with macrophages earlier. 1 mL of fresh macrophage medium was added to the first insert, as well as 1 mL of each of the diluted living and dead bacterial cultures to the next four. To the final insert, which was placed in a well containing only media and no macrophages, 1 mL of living *S. sudanensis* was added. This set up six conditions: macrophages in isolation (M), macrophages + *S. sudanensis* (M+S), macrophages + dead *S. sudanensis* (M+DS), macrophages + *S. coelicolor* (M+C), macrophages + dead *S. coelicolor* (M+DC), *S. sudanensis* in isolation (S). The assay plate was then incubated for 48 hrs at 37°C and 5% CO_2.

2.1.13 Extraction of bacterial and macrophage RNA

Well inserts from the assay plate were removed and set aside in empty petri dishes. 1.5 mL of macrophage culture supernatant was harvested from each well of the plate and stored in microfuge tubes at -20°C. The two living *S. sudanensis* cultures, in the conditions M+S and S, were harvested from their respective inserts and centrifuged for 1 minute at 18407 xg. The supernatant was removed and 250 µL of RNAlater stabilisation reagent added, followed by a 5 minute incubation. The samples were then centrifuged for 1 minute at 6010 xg and the supernatant removed. Pellets were resuspended in 200 µL of 50 mg/mL lysozyme and incubated at 37°C, for 4 minutes.

The macrophages in the wells of the assay plate were dislodged into the remaining 0.5 mL of media using cell scrapers, transferred to microfuge tubes and centrifuged at 18407 xg, for 1 minute. The supernatants were removed and the cells resuspended in 250 µL of RNAlater. Both the bacteria and macrophage samples were then centrifuged at 6010 xg, for 3 minutes, their supernatants removed and 350 µL RLT buffer from the Qiagen RNeasy Mini Kit added. From this point on, the RNeasy Mini Kit was used to isolate RNA, following the manufacturer's protocol. The presence of intact RNA in the samples was then checked for using gel electrophoresis (100 V, 30 minutes), with a 0.8% agarose gel. Sample purity was assessed using a NanoDrop ND-1000 spectrophotometer.

2.1.14 Preparation of cDNA libraries for RNA sequencing

Pure mRNA was obtained by using the Illumina Ribo-Zero rRNA Removal Kit on all macrophage and bacterial RNA samples, with the protocol followed as specified by the manufacturer. Following this procedure, samples were purified using the Zymo Clean and Concentrate – 5 kit. Preparation of a cDNA library from each sample was then carried out using the NEBNext Ultra II RNA Library Prep Kit for Illumina, with the protocol followed as specified by the manufacturer. Quality of the cDNA libraries was then assessed using an Agilent 2100 Bioanalyzer, with samples loaded onto a DNA 1000 chip.

2.1.15 Activity assays with THP-1 monocytes

At various points throughout the project, bacterial culture samples, purified extracts or compounds needed to be tested for activity against THP-1 monocytes. The following procedure was always used, unless otherwise stated. A 10 mL THP-1 culture was set up (see section 2.1.3). Following incubation, the culture was transferred to a 50 mL Falcon tube and centrifuged at 233 xg to pellet the cells. The old culture media was removed and then replaced with a volume of fresh media to give 250,000 cells/mL. Cells were then seeded into however many wells were needed, in either Corning Costar 24-well plates or Falcon 96-well plates. For the 24-well plates, 1 mL of culture was added to each well, while for the 96-well plates 200 µl was added. A range of dilutions of the samples being tested were then set up across the macrophage wells, with one well always left as a "0%" control. The assay plate would then be incubated at 37°C and 5% CO_2, for a maximum of 72 hrs, with images of the wells captured every 24 hrs.

2.1.16 Bacterial culture supernatants activity assay with THP-1 cells

S. sudanensis and *S. coelicolor* were cultured in 25 mL GYE as described in section 2.1.2. To harvest supernatant (SN), cultures were poured into 50 mL Falcon tubes and centrifuged at 3270 xg for 10 minutes. The SNs were collected and bacterial pellets discarded. SNs were then sterilised by passing through 0.2 µm pore Sartorious syringe filters. SNs were then added at a range of concentrations to THP-1 cell cultures, set up in a 96-well plate as described in section 2.1.9.

2.1.17 NLRP3 activation and inhibition assay with THP-1 cells

A 20 mM stock of MCC950 was prepared by dissolving 10 mg of compound in 1.235 mL DMSO. A 6.7 mM stock of nigericin was prepared by dissolving 10 mg of compound in 2 mL 100%

ethanol. *S. sudanensis* culture SN was prepared as detailed in section 2.1.10. THP-1 cells were set up in a 96-well plate as described in section 2.1.9. Cells in four of the wells were first primed with LPS, which was added to a concentration of 1 μg/mL and the plate incubated at 37°C and 5% CO_2 for 3 hrs. After incubation, the media containing LPS was removed from the wells and replaced with fresh RPMI. The following conditions were then set up in different wells: cells in isolation, '+ LPS priming', '+ LPS priming + nigericin', '+ LPS priming + MCC950', '+ LPS priming + nigericin + MCC950', '+ nigericin', '+ MCC950', '+ nigericin + MCC950', '+ *S. sudanensis* SN', '+ *S. sudanensis* SN + MCC950'. Nigericin and MCC950 were each added to a concentration of 10 μM to the appropriate wells. Bacterial SN was added to concentrations of 30%, 40% and 50% to the appropriate wells. The assay plate was then incubated at 37°C and 5% CO_2, for 72 hrs, with images of the wells captured every at 1, 6, 24, 48 and 72 hrs.

2.1.18 Aqueous and organic phases of S. sudanensis SN activity assay with THP-1 cells

For preparation of the aqueous and organic phases of culture SN, see section 2.4.2. A small sample of the organic phase was dried under reduced pressure using a Buchi R-100 Rotavapor. The dried sample was then resuspended in 1 mL sterile ultrapure water. A sample of the aqueous phase was sterilised by passing through a 0.2 μm pore Sartorious syringe filter. The samples were then added to concentrations of 30, 40 and 50% to THP-1 cells set up in a 96-well plate, as described in section 2.1.9. The assay plate was then incubated at 37°C and 5% CO_2, for 48 hrs, with images of the wells captured every 24 hrs.

2.1.19 High performance liquid chromatography fractions activity assay with THP-1 cells

Collected high performance liquid chromatography (HPLC) fractions were evaporated to dryness in a SP Scientific Genevac Series 3 HT Evaporator. Fractions were then resuspended in 120 μL sterile ultrapure water. THP-1 cells were set up in a 96-well plate as described in section 2.1.9. The HPLC fractions were added to a concentration of 20% to each well. The assay plate was then incubated at 37°C and 5% CO_2, for 72 hrs, with images of the wells captured every 24 hrs.

2.1.20 Synthetic cyclic dipeptides activity assay with THP-1 cells

Cyclo(-Leu-Pro) and cyclo(-Val-Pro) were purchased from Bachem. 1 M stock solutions of each were made using DMSO and further diluted to give additional stocks of 200 mM, 20 mM, 10 mM, 2 mM, 0.2 mM. THP-1 cells were set up in a 96-well plate as described in section 2.1.9. Cells in ten of the wells were primed with LPS, which was added to a concentration of 1 μg/mL

and the plate incubated at 37°C and 5% CO_2 for 3 hrs. After incubation, the media containing LPS was removed from the wells and replaced with fresh RPMI. The following conditions were then set up in different wells: cells in isolation, '+ LPS priming', '+ LPS priming + DMSO (0.5%)', '+ DMSO (0.5%)', '+ LPS priming + DMSO (1%)', '+ DMSO (1%)', '+ LPS priming + cyclo(-Leu-Pro)', '+ cyclo(-Leu-Pro)', '+ LPS priming + cyclo(-Val-Pro)', '+ cyclo(-Val-Pro)'. The cyclic dipeptides were added to concentrations of 10 mM, 5 mM, 1 mM, 100 uM, 50 uM, 10 uM, 1 uM to the appropriate wells. The assay plate was then incubated at 37°C and 5% CO_2, for 72 hrs, with images of the wells captured every 24 hrs.

2.2 Bioinformatical methods

2.2.1 De novo genome sequencing

Extracted *S. sudanensis* genomic DNA was submitted to the NU-OMICS facility at Northumbria University for Illumina sequencing, using a MiSeq sequencer. Additionally, the DNA was sequenced on an Oxford Nanopore Technologies MinION sequencer by Demuris Ltd. The sequencing library was prepared in the lab using the Oxford Nanopore Rapid Barcoding Kit before being submitted for sequencing. Assembly of the MinION and Illumina reads and annotation of the genome was then performed by Dr Nick Allenby of Demuris Ltd. The program Canu was used for assembly, while Prokka was used for annotation (Seemann, 2014, Koren et al., 2017).

2.2.2 Identification of biosynthetic gene clusters

The presence of biosynthetic gene clusters within the assembled *S. sudanensis* genome were predicted using an online analysis tool, the antibiotics and secondary metabolite analysis shell (antiSMASH), version 4.2.0 (Blin et al., 2017). Sequences were submitted to antiSMASH for analysis using the default settings of the software.

2.2.3 Differential expression analysis of RNAseq data

Barcoded RNAseq libraries, prepared as described in section 2.1.8, were pooled and submitted for sequencing to Edinburgh Genomics. Illumina sequencing of the libraries was performed on a single lane of a NovaSeq machine. The quality of the sequencing data was assessed using FastQC, version 0.11.8 (Babraham Bioinformatics). The mouse transcriptome GRCm38.p6

(release M22) was downloaded from the GENCODE online database. The *S. sudanensis* transcriptome was created by extracting protein coding sequences from the bacterial genome, which was performed by John Casement of the Bioinformatics Support Unit, Newcastle University. RNAseq reads were mapped to the appropriate transcriptome using the program Salmon, version 0.8.2 (Patro et al., 2017). Additional mapping of *S. sudanensis* RNAseq reads to the full *S. sudanensis* genome was carried out using BowTie2, version 2.3.2 (Langmead and Salzberg, 2012). The R package tximport, version 1.2, was used to import the mapping data from Salmon into R, version 3.6.1 (Soneson et al., 2015). Differential expression analysis was performed on the RNAseq data using the R package DESeq2, version 1.32 (Love et al., 2014).

2.2.4 Protein alignments and homology searches

The amino acid sequences of proteins encoded by genes of interest were compared to others using the NCBI Protein Basic Local Alignment Search Tool (BLASTP) to search for homology (Altschul et al., 1990). Precise alignments of two proteins was performed using the EMBL-EBI Clustal Omega multiple sequence alignment program, version 1.2.4 (Madeira et al., 2019).

2.3 Immunological methods

2.3.1 QUANTI-Blue assay

QUANTI-Blue is a dye produced by InvivoGen that changes from violet to blue due to the activity of alkaline phosphatase. It was used to quantify the amount of SEAP produced by RAW-blue macrophages during the indirect interaction assay with bacteria. As SEAP expression is induced by activated NF-κB, the levels of SEAP measured correspond to the levels of NF-κB activation in the macrophages. QUANTI-Blue reagent was prepared according to the manufacturer's instructions and 180 µL was added to wells of a Falcon 96-well plate. The macrophage culture supernatants harvested from 3 repeats of the indirect interaction assay were thawed at room temperature and 20 µL of each was added to the wells containing QUANTI-Blue. Wells were set up with three technical repeats for each sample. 20 µL ultrapure water was used as a blank. The plate was incubated at 37°C, for 1 hr. The OD_{625} of each of the wells was then measured, using a BMG SPECTROstar Nano plate reader. Raw absorbance values were then blank corrected.

2.3.2 Enzyme-linked immunosorbent assays

Macrophage culture supernatants harvested from six repeats of the indirect interaction assay were assayed for the presence of four different cytokines using enzyme-linked immunosorbent assays (ELISAs). The following kits from ThermoFisher Scientific were used: IL10 Mouse Uncoated ELISA Kit, IL1β Mouse Uncoated ELISA Kit, TGFβ1 Human/Mouse Uncoated ELISA Kit, TNFα Mouse Uncoated ELISA Kit. The protocols were followed as described by the manufacturer. Clean DMEM + 5% FBS + 0.1% 2-ME was used a blank for each assay. Upon conclusion of the kit protocols, the OD_{425} of the plate wells were measured using a BMG CLARIOstar plate reader. Raw absorbance values were then blank corrected.

2.4 Chemical extraction and purification methods

2.4.1 Preparation of aqueous extract from liquid bacterial culture

S. sudanensis was cultured in 3L of GYE media for 96 hrs, using conditions described in section 2.1.2. The bacterial biomass was then removed by centrifugation at 11325 xg and the SN collected.

2.4.2 Liquid phase organic solvent extraction

The aqueous *S. sudanensis* culture SN was mixed with an equal volume of ethyl acetate, within a separation funnel. Ethyl acetate has medium polarity, making it ideal for extracting a wide variety of both polar and non-polar compounds from the aqueous phase. The two phases were then decanted into separate flasks and the process repeated, with the two organic phases pooled together. The organic and aqueous phases were then tested for activity against THP-1 cells, as described in section 2.1.12. The active organic phases was then evaporated to dryness under reduced pressure using a Buchi R-100 Rotavapor.

2.4.3 HPLC fractionation

An Agilent 1260 Infinity II Preparative HPLC system, with C18 reversed phase column and fraction collector, was used to perform preparative HPLC on the active organic phase in collaboration with Dr Yousef Dashti, Newcastle University. The dried organic phase was dissolved in 2 mL 100% methanol. The entirety of the sample was loaded onto a cotton roll, which was left to dry out before being placed into the HPLC column. The column was run at a

flow rate of 0.2 mL/min, for a duration of 60 minutes. A gradient of 5-100% acetonitrile (ACN), with a 5 minute ACN wash at the end, was used. Fractions were collected every minute. Half of each of the fractions were then evaporated *in vacuo* using a Genevac HT Series 3 centrifugal evaporator and tested for activity against THP-1 cells, as described in section 2.1.13. The remaining volumes of each fraction were reserved for proton nuclear magnetic resonance (^{1}H-NMR) analysis.

2.4.4 Liquid chromatography-mass spectrometry

HPLC fractions found to be active were analysed by liquid chromatography-mass spectrometry (LC-MS) using an Agilent 1260 Infinity II HPLC system with C18 reversed phase column and automated multisampler module, coupled to a Bruker microTOF mass spectrometer. This was done in collaboration with Dr Yousef Dashti, Newcastle University. 50 µL of each fraction in water was mixed with 200 µL methanol and transferred to a HPLC vial. Samples were loaded onto the HPLC column by the automated injector and were run with a flow rate of 0.2 mL/min, for a duration of 60 minutes. A gradient of 5-100% acetonitrile (ACN) was used.

2.4.5 ^{1}H-NMR analysis

^{1}H-NMR analysis of active HPLC fractions was performed by Dr Yousef Dashti, Newcastle University. Spectra were recorded using a Bruker Avance III 700 MHz spectrometer equipped with a TCI cryoprobe at 25°C. Chemical shifts were referenced to the solvent peaks of DMSO-d_6 at δ_H 2.50.

2.5 Data visualisation and statistical methods

2.5.1 Creation of data plots

The plot of the genome of *Streptomyces sudanensis* was created using Artemis, version 18.1.0. All other data plots were created in R, version 3.6.1. Bar and box plots were generated using the package ggplot2, version 3.3.5. Heatmaps were created using the heatmap.2 function of the package gplots, version 3.1.1. Dendrograms were created using the hclust function of the stats package, version 3.6.2. Principal component analysis (PCA) plots were generated using the plotPCA function in the package DESeq2, version 1.32. Dispersion plots were generated

using the plotDispEsts function within DESeq2. Venn diagrams were created using the online software Venny, version 2.1.

2.5.2 Statistical thresholds applied within RNAseq differential expression analysis

The R package DESeq2, version 1.32, was used for differential expression analysis of RNAseq data from RAW 264.7 macrophages and *S. sudanensis*. By default, DESeq2 outputs a results table containing log2 fold changes, p-values and adjusted p-values for each gene in a data set. Adjusted p-values were derived using the Benjamini-Hochberg method. The thresholds set for differential expression for both data sets were a log2 fold change of +/- 1 and an adjusted p-value of < 0.05. Results tables were parsed using these two thresholds.

2.5.3 Statistical analyses of RAW macrophage cell length to width ratios

Data for each condition were plotted as histograms using the hist() function of R, version 3.6.1, to check their distribution. Only data in the control condition M were normally distributed. The non-parametric Kruskal-Wallis test was therefore used to compare the effect of each condition on cell length to width ratios. A post-hoc Dunn test with Bonferroni correction was then used for pairwise comparisons between conditions.

2.5.4 Statistical analyses of cytokine ELISA results

For all ELISAs, mean absorbances were calculated for each condition. An Analysis of variance (ANOVA) test was performed to compare the effects of both living and dead *S. sudanensis* and *S. coelicolor* on cytokine release from RAW 264.7 macrophages during the indirect interaction assay. Tukey's Honestly Significant Difference post-hoc test was accordingly performed on all three datasets to give pairwise comparisons between conditions.

Chapter 3. Characterisation of the macrophage response to *Streptomyces sudanensis*

3.1 Introduction

One of the core aims of this study was to interrogate how mammalian immune cells respond to actinomycetoma pathogens. It was decided that the optimal approach to this question would be to design an assay method wherein the bacteria and immune cells are co-cultured. This would allow the observation of biological effects induced within either population.

Macrophages (specifically, the murine macrophage cell line RAW-blue) were selected as the immune cell type for use in the assay. As described in Chapter 1, these cells act as the front line of innate immunity and are therefore amongst the first immune cells to encounter mycetoma infections (Akira et al., 2006, Iwasaki and Medzhitov, 2015). They are key to initiating and defining the early immune response to invading pathogens. Additionally, as detailed in Chapter 1, macrophages are implicated in the initiation and formation of granulomas, one of the defining features of mycetoma pathogenesis (Ramakrishnan, 2012).

Several methods were chosen for the assessment of biological effects induced by macrophages and bacteria interacting during the assay and will be described below. These methods included: an NF-κB activity assay, enzyme-linked immunosorbent assays (ELISAs) to detect and quantify the release of certain cytokines from the macrophages, transcriptome analysis by RNA sequencing (RNAseq) and phenotypic assessment of the macrophages by microscopy.

RAW-blue is an NF-κB reporter cell line, derived from RAW 264.7, which contains a chromosomally integrated secreted alkaline phosphatase (SEAP) reporter, inducible by the transcription factor NF-κB (Invivogen, n.d.). As described in Chapter 1, this transcription factor upregulates the expression of pro-inflammatory genes in response to pathogen invasion or tissue damage (Fitzgerald and Kagan, 2020). Therefore, measuring its activity in response to a given stimulus allows the extent of immune activation to be observed. The amount of SEAP produced by RAW-blue cells can be easily quantified in a colourimetric assay, in turn allowing quantification of the scale of NF-κB activity. This cell line also expresses a wide range of cytokines and pattern recognition receptors (PRRs), such as Toll-like and nucleotide-binding

oligomerisation domain-like receptors, and so is ideally suited for studying macrophage detection of, and activation by, pathogens.

ELISAs are colourimetric assays that detect target proteins through the use of enzymes conjugated to antibodies that are specific to the protein. The type of ELISA chosen for cytokine detection was a sandwich ELISA, an overview of which is shown in Figure 3.1.

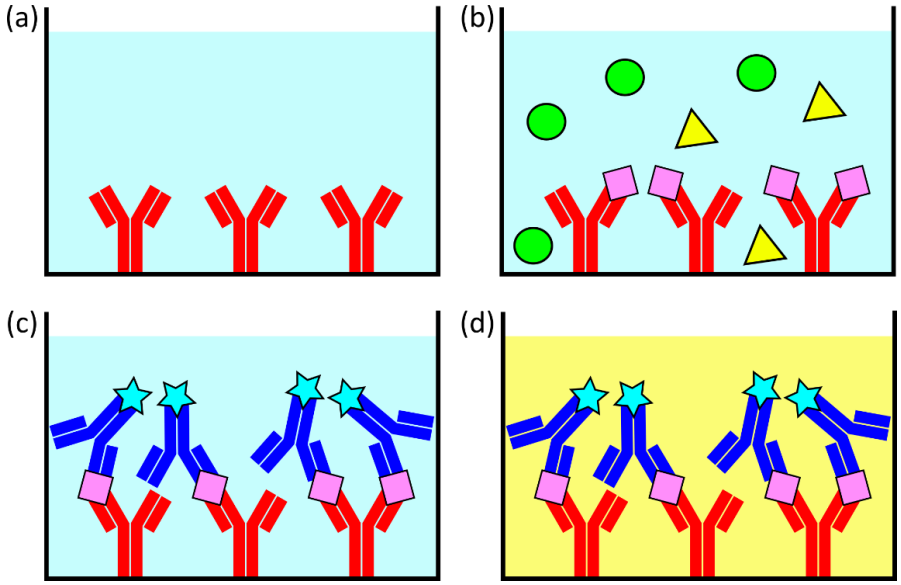


Figure 3.1 Summary of a sandwich ELISA. **(a)** A well is coated with capture antibodies, specific to the target protein. **(b)** Sample is added to the well, containing a mixture of different proteins. The target protein is bound by the capture antibodies. The well is washed to remove unbound proteins. **(c)** Detection antibodies, also specific for the target protein, are added and bind to the target. The detection antibodies are conjugated to an enzyme (shown as a star). The well is washed to remove unbound detection antibody. **(d)** The substrate for the conjugated enzyme is added to the well. The enzyme cleaves the substrate, forming a coloured product. The measured light absorbance of the well is proportional to the level of enzyme activity and therefore to the levels of target protein bound by the antibodies. Target protein quantity can thus be calculated.

RNAseq is a powerful technique, enabled by next generation sequencing (NGS) methods. From its development in 1977, Sanger sequencing was dominant (Metzker, 2010). Shown in Figure 3.2, this is a sequencing by synthesis method that uses chain terminating, fluorescently labelled nucleotides to determine DNA sequences. It is highly accurate but low throughput, as only one DNA fragment can be sequenced at a time. This makes it a very slow and inefficient technique for sequencing large libraries of DNA fragments.

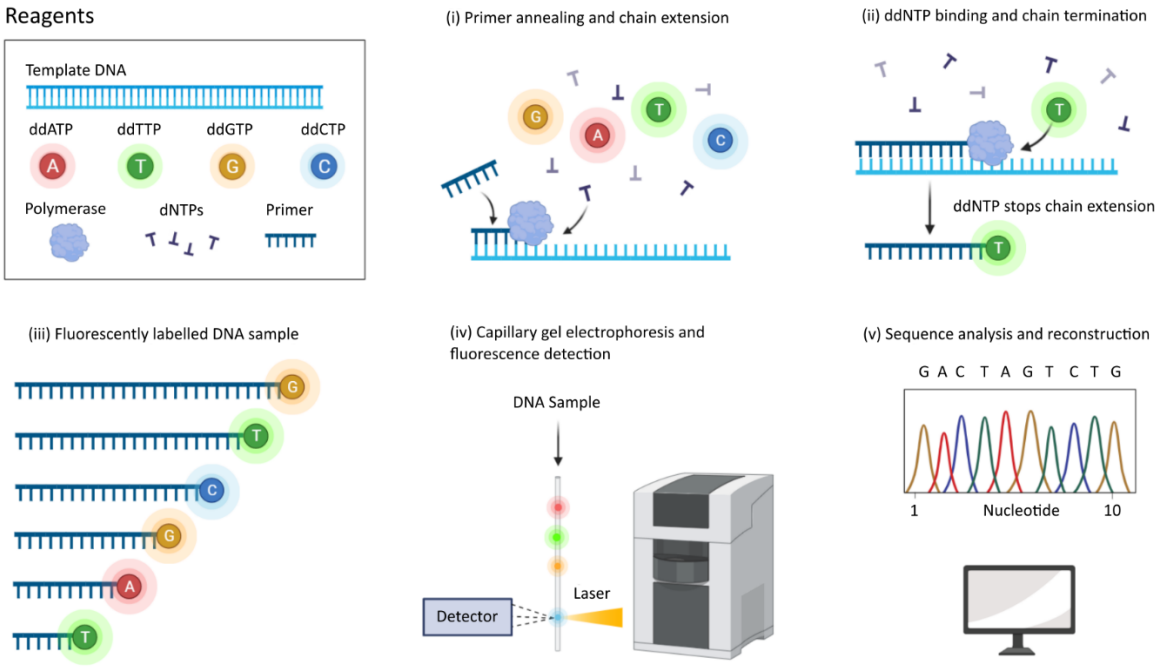


Figure 3.2 Sanger DNA sequencing method. Image adapted from BioRender (2020).

From 2005 onwards, new sequencing methods were introduced that massively parallelised the sequencing by synthesis process, termed NGS methods (McCombie et al., 2019). Of the initial NGS technologies, the dominant one that remains regularly in use to this day is that of Illumina Inc., whose method is illustrated in Figure 3.3. DNA is first fragmented and then adapter sequences ligated to the fragments, creating a sequencing library (Metzker, 2010). The library is then loaded onto a flow cell, which is itself coated with adapters. The library fragments anneal to their complimentary adapters in the flow cell and undergo rounds of bridge amplification, as shown in Figure 3.3(a). Sequencing by synthesis is then performed for all library fragments simultaneously, via the mechanism shown in Figure 3.3(b). This method makes DNA sequencing high throughput, allowing for whole genomes to be sequenced in one run, and increases the depth of sequencing coverage.

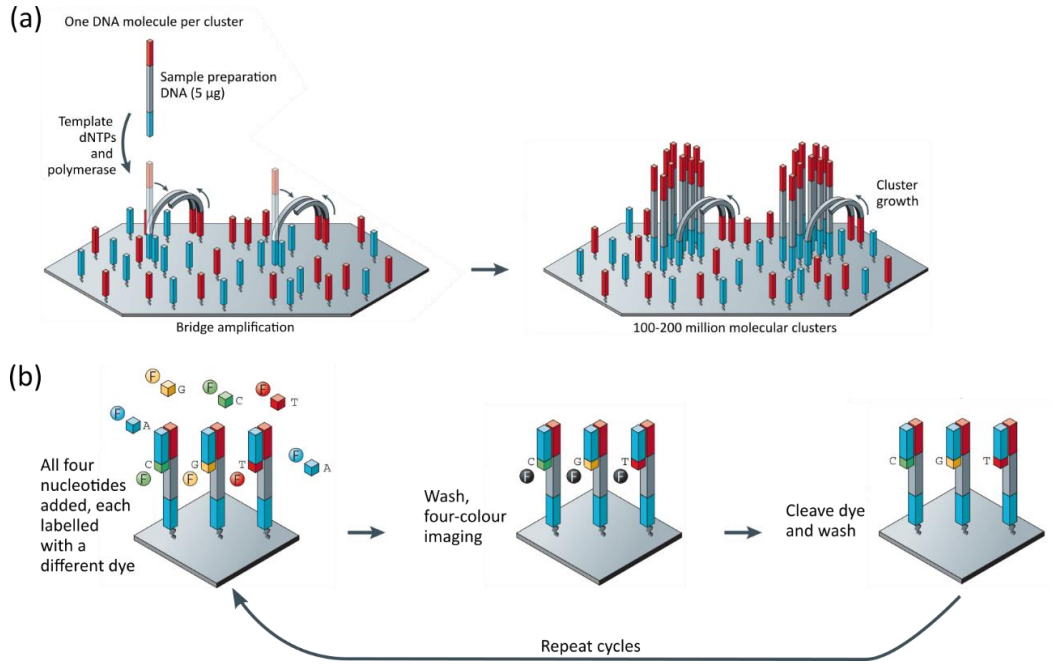


Figure 3.3 Illumina DNA sequencing. **(a)** Sequencing library fragments undergo bridge amplification within the Illumina flow cell, creating clusters of complimentary forward and reverse strands and increasing sequencing coverage of the library. **(b)** Sequencing by synthesis is then performed for all DNA fragments, using fluorescently labelled, chain terminating nucleotides. Live four-colour imaging of the flow cell shows which nucleotides have been incorporated within each fragment. The labels are cleaved and the process is repeated to sequence the entire fragment. Adapted from Figures 1 and 2 of Metzker (2010).

NGS methods have progressed further still over the last two decades, with the development of the so-called third generation technologies: single-molecule real-time sequencing and nanopore sequencing (McCombie et al., 2019). These are produced by Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT) respectively. Their key innovation is the ability to produce reads that range from thousands to hundreds of thousands of base pairs in length (van Dijk et al., 2018, McCombie et al., 2019). This is orders of magnitude greater than those produced by Illumina sequencing, which outputs on a scale of only hundreds of base pairs. PacBio and ONT employ differing methods, both of which are shown in Figure 3.4. The PacBio system (Figure 3.4(a)) still uses sequencing by synthesis, but is dependent on the use of a zero mode waveguide well, which is designed to provide the smallest possible volume for light detection (Rhoads and Au, 2015). This allows the sequencer to visualise the incorporation of individual nucleotides during continuous strand synthesis by DNA polymerase. The ONT system (Figure 3.4(b)) instead relies on DNA strands being fed through protein nanopores,

which are embedded in a membrane (Stoddart et al., 2009). It then measures changes in the electrical current across the membrane, caused by translocation of the DNA.

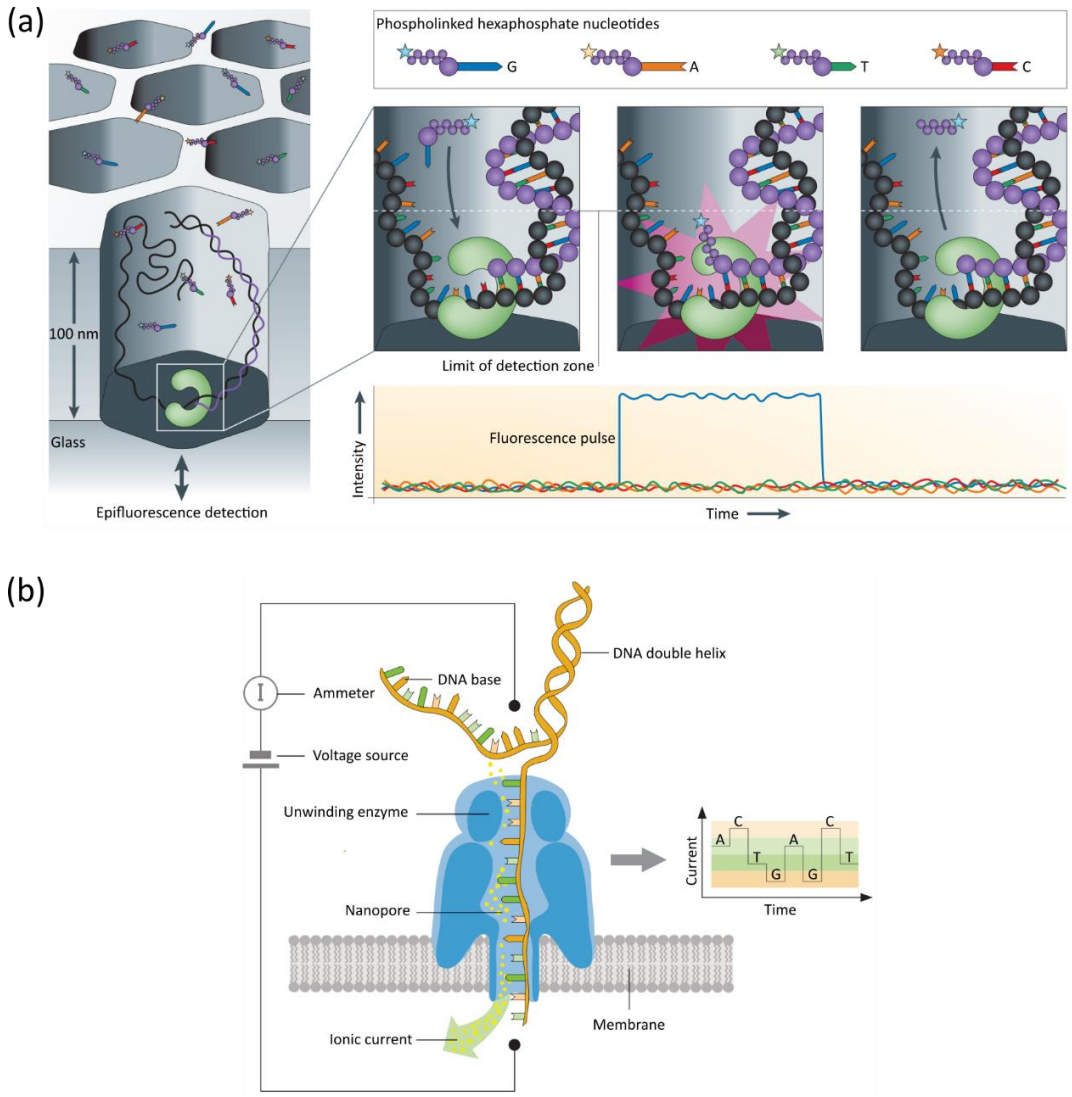


Figure 3.4 (a) PacBio single-molecule real-time sequencing. Libraries are loaded onto a cell, which contains roughly 150,000 zero mode waveguides. A single DNA polymerase is fixed to the bottom of each one. The polymerase synthesises a new strand from the sequencing library template. It incorporates fluorescently labelled nucleotides which produce light pulses as they are held by the enzyme. Light pulses are used to identify each nucleotide and are continuously recorded across the whole cell, allowing multiple library fragments to be sequenced simultaneously. The series of light pulses from each zero mode waveguide are then analysed to produce the nucleotide sequence. Adapted from Figure 4 of Metzker (2010). **(b)** ONT nanopore sequencing. Libraries are loaded onto a flow cell, which contains a phospholipid membrane, studded with protein pores. A current is applied across the membrane. DNA fragments are bound by the pore proteins, unwound and pulled through the pore. As nucleotides pass through the pore, they generate a unique change in the current, due to altering the flow of ions. Changes in the current are measured and used to identify each nucleotide and thus build up a DNA sequence. This process occurs simultaneously across the flow cell membrane. Adapted from Figure 2 of Gopfrich and Judge (2018).

NGS innovations not only increased the speed and affordability of genomic DNA sequencing, but their high throughput nature also meant that sequencing a cell's transcriptome became viable and RNAseq methods were accordingly developed (Mortazavi et al., 2008, Nagalakshmi et al., 2008). For transcriptomic analysis, RNA is extracted from cells, the mRNA isolated and a cDNA library generated for sequencing. The sequencing data can then be mapped to a reference genome or transcriptome. This allows for the identification of the mRNA transcripts present in a cell at a given moment, providing a snapshot of which genes were being expressed. Transcripts can also be quantified, giving a direct measure of gene expression levels.

It was decided that RNAseq would be performed on the macrophages and the bacteria, both in the presence and absence of one another. Differential expression analysis could then be performed, whereby expression levels of genes under different conditions are compared. This would allow for the identification of macrophage genes that were up- or downregulated in the presence of the bacteria and *vice versa*. Thus, it would be possible to observe how the different cell populations respond to one another at the transcriptional level. This could uncover immune pathways activated by the pathogen, as well as host cell systems that are disrupted, providing insight into the pathogenesis of mycetoma.

As described in Chapter 1, Actinobacteria are major producers of specialised metabolites (Katz and Baltz, 2016). It is therefore possible that such compounds may have a role in mycetoma pathology. The physical design of the co-culture assay method was therefore considered, with the aim of identifying specialised metabolites produced by actinomycetoma pathogens, which may act as virulence factors. Such metabolites would usually be secreted from the bacterial cells and then naturally diffuse through their culture medium. Therefore, the bacteria and macrophages need not be in direct physical contact with each other for the effects of any specialised metabolites to be observed. Additionally, if the cultures were mixed together, the PRRs of the macrophages would detect standard bacterial cell wall pathogen associated molecular patterns (PAMPs), e.g. peptidoglycan and lipoproteins. This would trigger well-characterised immune responses that aren't the main focus of this study. A permeable barrier separating the bacteria and macrophages would allow for metabolite diffusion into the macrophage culture, while minimising macrophage exposure to common PAMPs.

With this core aspect of the assay design decided, it was thereafter referred to as the 'indirect interaction assay'. Method optimisation was undertaken to determine what permeable barrier to use and other appropriate experimental conditions.

3.2 Design and optimisation of a novel interaction assay between *S. sudanensis* and macrophages

3.2.1 Testing dialysis tubing as a barrier between bacteria and macrophage cultures

Dialysis tubing was trialled first for use as a barrier between the two cell cultures. This tubing is a semi-permeable membrane that will only allow molecules of a certain size to diffuse through it. The molecules able to diffuse can be controlled by selecting tubing with specific pore sizes. The size chosen for this trial was 3000 kDa, which should allow for diffusion of specialised metabolites while crucially blocking the passage of bacteria through the membrane. The interaction assay using dialysis tubing was set up as described in Chapter 2 and the images taken during the incubation period are shown in Figure 3.5.

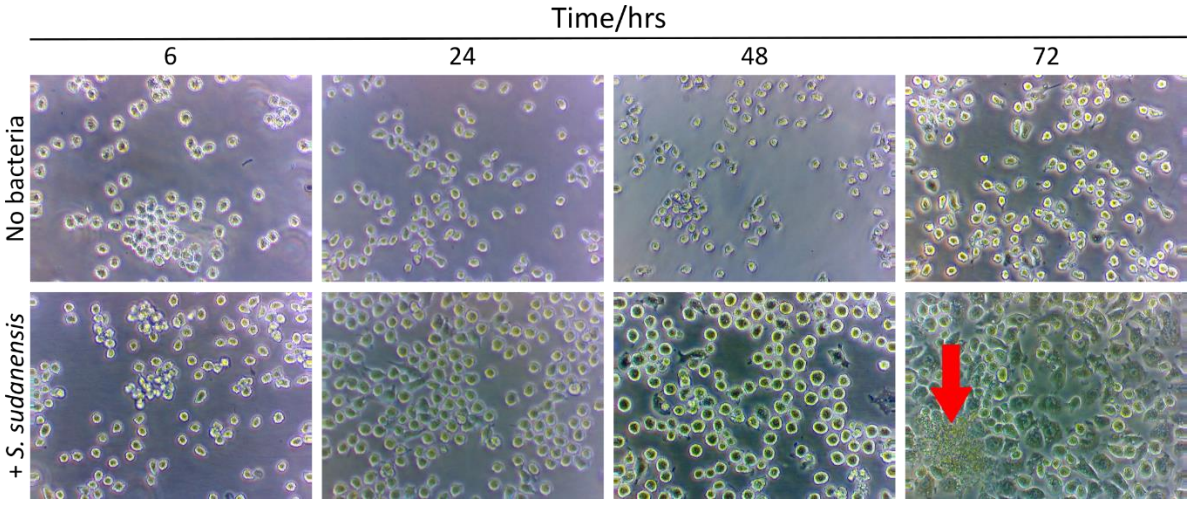


Figure 3.5 Images of RAW-Blue macrophages taken over the course of 72 hrs, during the trialling of dialysis tubing as a barrier between macrophage and bacterial cultures. Sealed tubing was placed into wells with the macrophages. Images in the top row are from a well where fresh DMEM was inside the sealed tubing. Those of the bottom row are from a well where *S. sudanensis* culture was inside the tubing. The red arrow indicates a growing bacterial mycelium. Images taken at 200X magnification.

The images of the 'No bacteria' control show the macrophages remained in a resting state for at least the first 48 hrs of the incubation period. The cells remained a constant size, their cytoplasm remained clear and no pseudopodia were visible. It therefore appeared that the dialysis tubing is itself not immunogenic. However, by 72 hrs cell morphology became more irregular with some pseudopodia having formed, indicating a low level of macrophage activation. It is likely that after incubating for 72 hrs, media nutrients would have been greatly depleted, while waste products excreted from the cells would have accumulated. These conditions will likely have caused some cells to die and subsequently release danger associated molecular patterns (DAMPs). If so, the released signals might then have activated other macrophages in the well, explaining the appearance of pseudopodia at this time point.

In the '+ *S. sudanensis*' condition, the macrophages at 24 hrs were clearly enlarged compared to at 6 hrs. The cells were also enlarged compared to those in the 'No bacteria' condition at 24 hrs. At 48 hrs, this enlargement was more evident, along with the presence of a small number of fully activated cells, indicated by the presence of multiple pseudopodia and an irregular cell morphology. By 72 hrs, the well was dominated by strongly activated macrophages. Most cells were enlarged further and had developed a "starry" morphology, with multiple pseudopodia spread outwards. High numbers of vesicles were also visible in their cytoplasm, indicating the cells had carried out phagocytosis.

Also shown in the image at 72 hrs is a growing mycelium of *S. sudanensis*, highlighted by the red arrow. The density and size of the mycelium suggests that the bacteria were likely present in the macrophage culture at earlier time points, which may explain the activation seen from 24 hrs onwards. It was therefore impossible to conclude whether the effects induced in the macrophages were due to specialised metabolite production or direct interaction with the bacteria. It is most likely that the bacteria leaked from the dialysis tubing through one of the sealed ends, which were tied off only by knotting. Multiple attempts to fully block bacterial escape by sealing the tubing in different ways were unsuccessful. Therefore, it was decided to abandon the use of dialysis tubing for the assay.

3.2.2 Using 3D tissue culture well inserts as a barrier between bacteria and macrophage cultures

Transwell permeable supports were next trialled as a barrier between cultures. These are plastic inserts with porous bases that are placed into multiwell plates. The design of the assay using the Transwell system is shown in Figure 3.6(a).

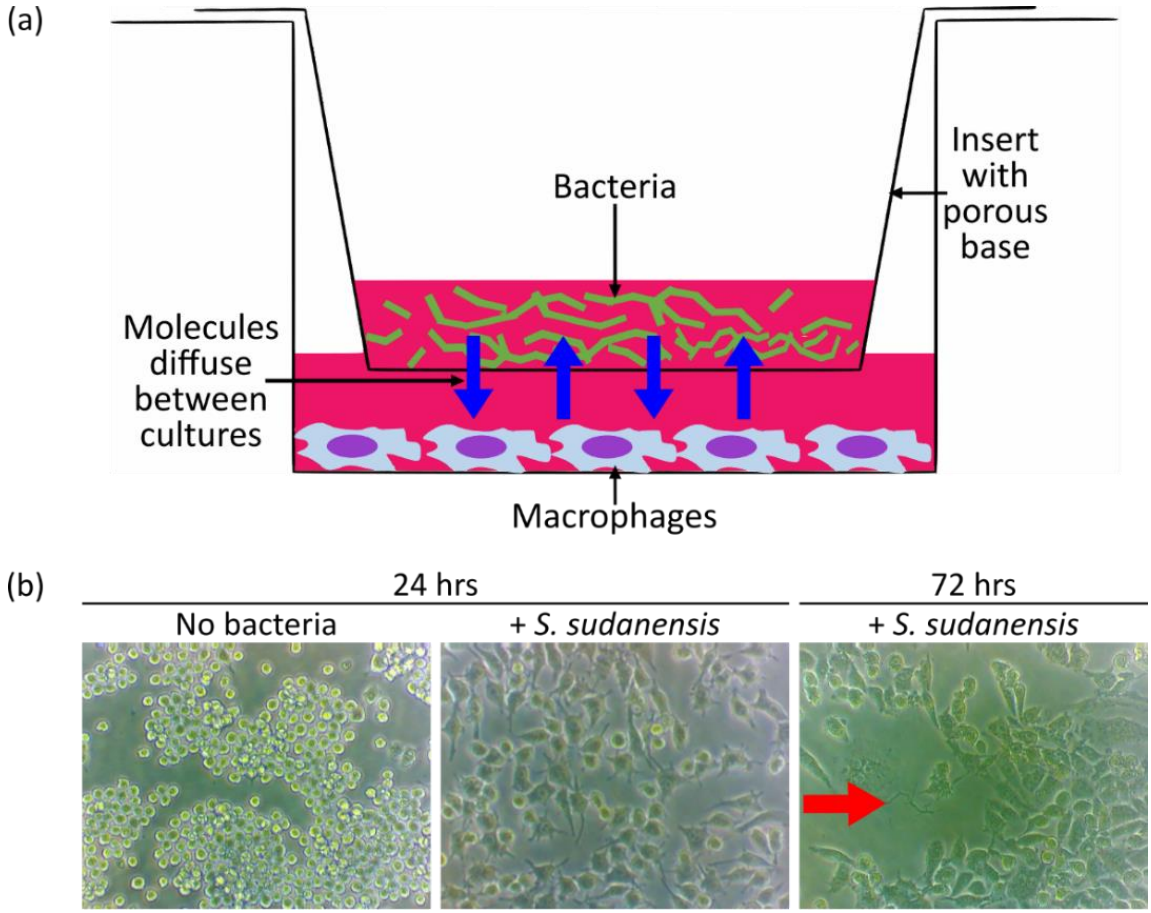


Figure 3.6 (a) Diagram of interaction assay method using Transwell permeable insert. **(b)** Images of macrophages incubated underneath a Transwell insert. The insert either contained fresh DMEM or *S. sudanensis*. The red arrow indicates a growing bacterial mycelium. Images taken at 200X magnification.

The smallest pore size available for the Transwell inserts was 0.4 µm, which should have been sufficiently small to block transmission of Streptomyces hyphae, which are typically 0.8 µm or more. Images taken during the trial of this method are shown in Figure 3.6(b). By 24 hrs, there was a clear contrast between the 'No bacteria' control and '+ *S. sudanensis*'. With the bacteria in the insert, the macrophages were visibly enlarged and had developed a "starry"

morphology, with multiple pseudopodia reaching outwards. It was noted that this activation state had been reached in a much shorter incubation time than during the dialysis tubing trial. This could be due to the larger pores of the Transwell insert allowing for easier diffusion of bacterially derived molecules. Macrophages in the negative control well, on the other hand, had maintained their size and regular, round morphology. The Transwell inserts themselves therefore showed no immunogenicity.

The trial was run up to 72 hrs, at which point a bacterial mycelium was observed in the '+ *S. sudanensis*' well (red arrow in Figure 3.6(b)). This suggested that *S. sudanensis* was able to pass through the 0.4 µm pores of the insert. As with the dialysis tubing trial, this meant it was impossible to distinguish clearly between effects induced by potential specialised metabolites and those triggered by close contact with bacterial cells.

3.2.3 Optimisation of agarose gel density within Transwell inserts

When cultured on solid agar plates some streptomycetes are known to be able to penetrate the surface of the agar. However, we noted that *S. sudanensis* did not appear to do this. We therefore reasoned that a layer of agarose might prevent *S. sudanensis* from passing through the pores of the transwell insert. The modified experimental design is shown in Figure 3.7(a).

The potential of the agarose to hinder the diffusion of specialised metabolites was considered. Different agarose densities were trialled to find the lowest density of agarose that would retain barrier activity for the bacteria.

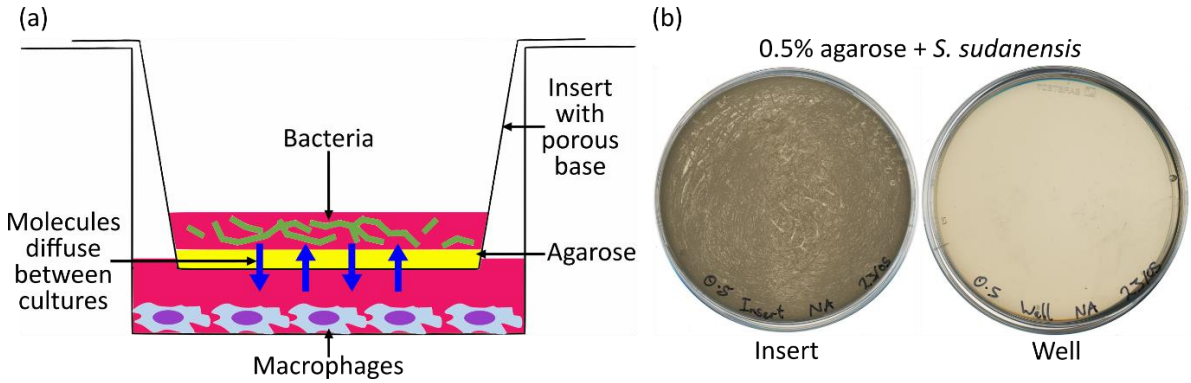


Figure 3.7 (a) Modified design of the Transwell interaction assay. **(b)** Nutrient agar plates, streaked with 100 µL *S. sudanensis* culture from the insert (left) and DMEM from the well beneath the insert (right). The 'Insert' plate was incubated for 24 hrs before imaging. The 'Well' plate was incubated for 6 days. Both were incubated at 37°C.

Gels with agarose concentrations of 0.25%, 0.3% and 0.5% were tested for structural integrity within Transwell inserts using the method detailed in Chapter 2. It was observed that 0.25% and 0.3% gels did not block the flow of culture media through the insert pores, with media droplets visible by eye on the underside of the insert after four minutes. By 30 minutes these droplets had fallen to the base of their wells, with new ones still forming. At the same time point, no droplets were visible under the 0.5% agarose insert or on the base of its well. It was thus concluded that the densities of the 0.25% and 0.3% agarose gels were too low to use in the assay. 0.5% would therefore be the lowest agarose concentration used in the rest of the trial.

Next, the ability of the 0.5% agarose gel to block *S. sudanensis* contamination of the well was assessed, using the assay set up shown in Figure 3.7(a). No bacterial growth was observed in the well under the insert for four days. At the conclusion of the incubation period, a final check for bacterial contamination of the well was carried out. A sample of the media from the well beneath the insert was streaked out on a nutrient agar plate. The same was also done for the bacterial culture within the insert, as a positive control. Images of the two plates following incubation are shown in Figure 3.7(b). The plate streaked from the insert grew a dense bacterial lawn, whereas the plate streaked from the well showed no signs of growth, even after six days of incubation. It was therefore concluded that a layer of 0.5% agarose gel successfully blocked bacterial growth through the Transwell insert pores. Accordingly, this density of gel was selected for use in the interaction assay.

3.2.4 Selection of a non-pathogenic Streptomyces control strain

To identify features of the macrophage response to *S. sudanensis* that are unique to the pathogen, a control interaction condition was needed. Model *Streptomyces* strains that are thought to be non-pathogenic were selected for this control. Such organisms would have the same fundamental biology as *S. sudanensis*, while theoretically lacking any mycetoma-related virulence factors. Therefore, any features of the macrophage response shared between the pathogen and non-pathogen could be attributed to immune reactions to common *Streptomyces* PAMPs. On the other hand, any features unique to the pathogen could be attributed to specific virulence mechanisms.

The other reason for using non-pathogenic model strains was that they are genetically tractable. Should a specialised metabolite of interest be found in *S. sudanensis*, the

biosynthetic gene cluster (BGC) responsible for its production could be heterologously expressed in the non-pathogenic host. If that strain then gained virulence, the role of the metabolite in mycetoma pathogenesis would be confirmed.

The two model organisms initially selected were *S. albus* J1074 and *S. coelicolor* M1152. Both are standard models for *Streptomyces* spp. and so have been very well studied and extensively genetically characterised (Baltz, 2010). These strains are also widely used for heterologous expression of BGCs (Gomez-Escribano and Bibb, 2014, Kallifidas et al., 2018). *S. albus* J1074 has an unusually small genome for a *Streptomyces*, being only 6.8 Mb compared to an average of roughly 9 Mb (Zaburannyi et al., 2014). This bacterium also has a faster growth rate than many other *Streptomyces*.

S. coelicolor M1152 is a derivative of the strain M145, itself derived from the wild-type strain A3(2), and is distinguished by carrying the following mutations: *Δact Δred Δcpk Δcda rpoB[C1298T]* (Gomez-Escribano and Bibb, 2011). These mutations have knocked out four BGCs, reducing the presence of competing carbon and nitrogen sinks and making the strain devoid of antibiotic activity. Meanwhile, the point mutation in *rpoB* serves to increase specialised metabolite production. This strain is therefore highly optimised for heterologous expression. Additionally, the strain being devoid of production of its most abundant antibiotics increases the appeal of using it in the interaction assay. Lack of production means the strain would be secreting fewer metabolites that could interact with the macrophages and obscure the results of the assay. The suitability of these two strains as non-pathogenic controls was tested. The bacteria were set up in the Transwell system with macrophages, as pictured in Figure 3.7(a). Images of the macrophages at the end of the assay incubation period are shown in Figure 3.8.

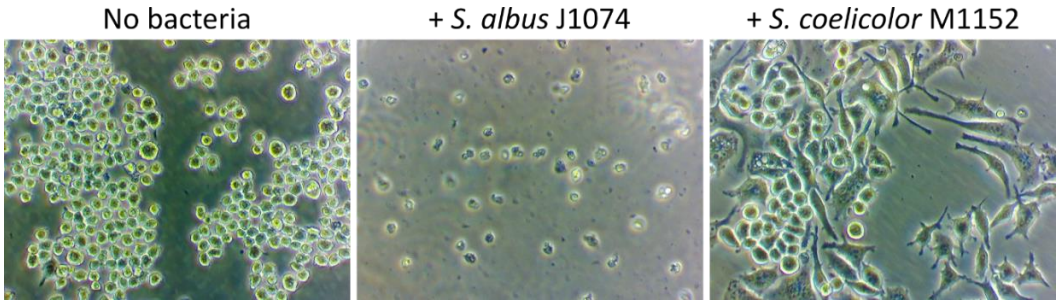


Figure 3.8 Macrophages, at 0.5 million cells per mL, following incubation at 37°C and 5% CO_2, in the Transwell assay system, with no bacteria in the insert, *S. albus* J1074 in the insert and *S. coelicolor* M1152 in the insert, respectively. Images taken at 48 hrs and at 200X magnification.

Within the 'No bacteria' control, the macrophages remained an approximately regular size and maintained their round morphology, as expected. With *S. albus* J1074 in the insert however, the number of macrophages in the culture massively declined and the cells that remained became small and shrivelled. This was a clear sign of cell death. Small particles were also visible throughout the well, which were likely the remains of dead macrophages. Those exposed to *S. coelicolor* M1152 on the other hand, did not die but instead underwent strong activation. All cells were enlarged and others had adopted a "starry" morphology, with multiple pseudopodia. Many had also carried out phagocytosis, evidenced by the numerous vesicles in their cytoplasm.

From these results, it was clear that *S. albus* J1074 produces a cytotoxic compound and was therefore not suitable for use as a non-pathogenic control. A search of the literature following this trial revealed that this strain is known to produce antimycins (Olano et al., 2014). These are compounds that inhibit mitochondrial electron transport chains, which could explain the high level of toxicity to macrophages. The observation of cell death did, however, confirm that secreted compounds could diffuse through the 0.5% agarose gel on the insert base. *S. coelicolor* M1152 showed no such toxicity but proved to be immunogenic. The macrophage activation seen with this strain in the insert could have been indicative of a standard immune response to the detection of PAMPs, which may have diffused through the agarose gel and insert pores. It was decided to use *S. coelicolor* M1152 as the non-pathogenic control strain in the assay.

3.2.5 Validation of method for heat-killing bacteria for use as control conditions

A final control condition of placing dead bacteria in the Transwell insert was designed for the interaction assay. For obvious reasons, dead bacteria do not synthesise specialised metabolites. Therefore, if *S. sudanensis* does produce a compound that can affect macrophages, any effects due to the compound's activity would only be seen in the presence of living bacteria. Differences between macrophage responses to dead and live *S. sudanensis* could be attributed to the bacteria secreting compounds as part of their living metabolism. Use of dead bacteria would also provide a further control for immune reactions triggered by standard *Streptomyces* PAMPs diffusing into the macrophage culture.

Autoclaving was selected as the method for killing bacteria, both for its ease of use and reliable sterilisation ability. Additionally, the use of chemical agents for killing would risk

contamination of the macrophage cultures during the assay. There was a concern that autoclaving could cause the bacterial cells to lose structural integrity and fully break down. For this control to work at its best, whole dead cells would be needed. Their gradual decomposition during the assay period would more accurately replicate the similar process taking place in the live culture, where some cells will die and others will shed cell wall components as they grow. To address this, autoclaving of *S. sudanensis* and the non-pathogen *S. coelicolor* M1152 was trialled. The autoclaved cultures were examined by microscopy and compared to living cultures. Samples of each culture type were also streaked out on plates, to test the efficacy of heat-killing by the autoclave. The results of this trial can be seen in Figure 3.9.

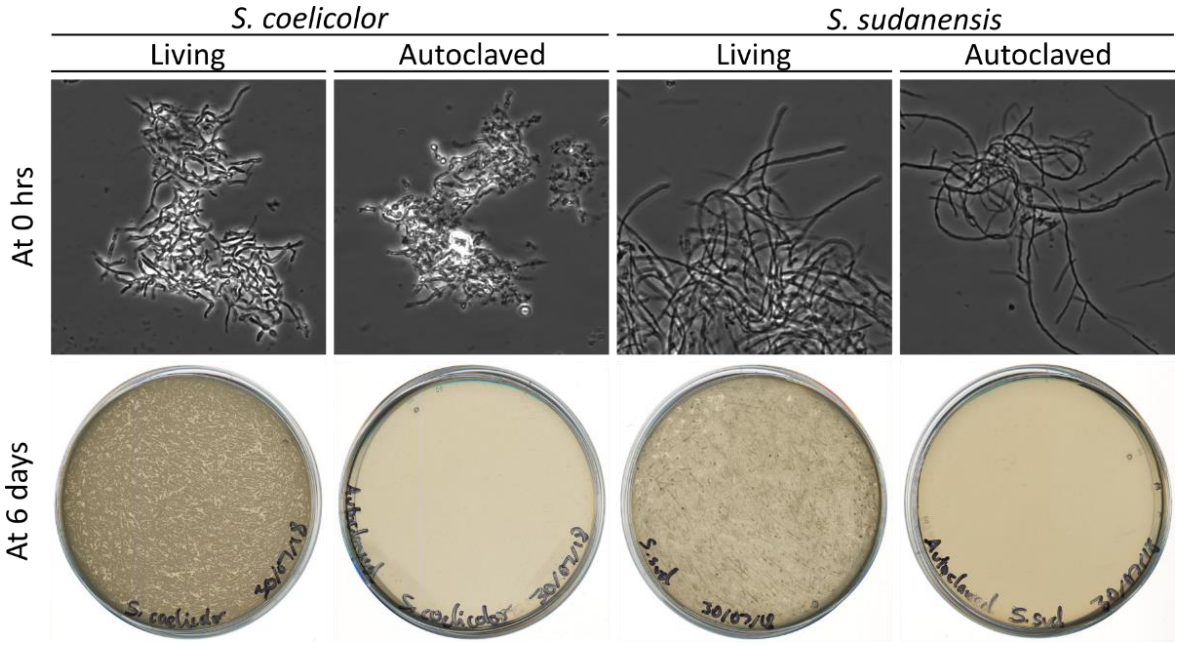


Figure 3.9 Comparison of living and autoclaved *S. coelicolor* M1152 and *S. sudanensis* cultures. Microscopy images taken at 1000X magnification.

In the microscopy images of *S. coelicolor*, cells that had undergone autoclaving appeared shrivelled and clumped together. Some rough mycelial morphology was maintained, with individual chains of dead cells visible. Nascent spore chains were also visible, indicating that some cells had undergone sporulation during autoclaving. However, the agar plates beneath these images suggest that sporulation was unsuccessful, as no bacterial growth was observed after incubation for six days. This result confirmed that autoclaving is highly effective at killing *S. coelicolor*. The microscope images of living and autoclaved *S. sudanensis* appeared nearly identical in terms of cell morphology. The mycelia mainly retained their structure following

autoclaving, without the shrivelling or agglomeration seen with *S. coelicolor*. The agar plates confirmed that the bacteria had been killed, as no growth was observed on the plate streaked from the autoclaved culture. Based on these results, autoclaving was chosen for heat-killing of *S. sudanensis*.

3.3 Visual observations of RAW-blue macrophages during the indirect interaction assay

With the indirect interaction assay protocol optimised, full trials were carried out. For each trial the macrophages were placed under five conditions: Macrophages in isolation (M), macrophages + *S. sudanensis* (M+S), macrophages + dead *S. sudanensis* (M+DS), macrophages + *S. coelicolor* M1152 (M+C) and macrophages + dead *S. coelicolor* M1152 (M+DC). Typical microscopy images of the macrophage cultures at the end of the assay period are shown in Figure 3.10.

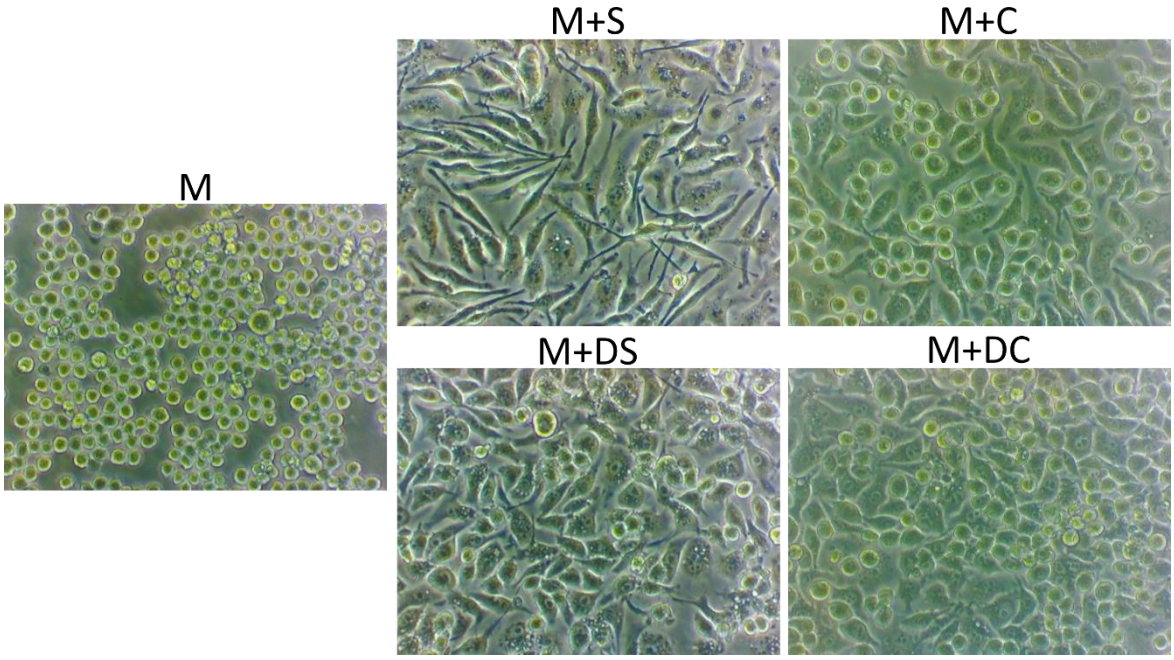


Figure 3.10 Images of RAW-blue macrophages after 48 hrs incubation within the indirect interaction assay. M: macrophages in isolation. M+S: macrophages + *S. sudanensis.* M+C: macrophages + *S. coelicolor* M1152. M+DS: macrophages + dead *S. sudanensis*. M+DC: macrophages + dead *S. coelicolor* M1152. Macrophage density was 1 million cells per mL. Images taken at 200X magnification and are representative of three biological repeats.

In the M condition, the macrophages had retained a rounded morphology, with largely uniform diameter, indicating that they remained in a resting state. The M+C, M+DC and M+DS conditions all appeared similar. The macrophages underwent enlargement, developed multiple pseudopodia, and an irregular cell shape. High numbers of vesicles were also present

in their cytoplasm, indicating phagocytic activity. These were all signs that the cells had undergone immune activation. Activation appeared to be stronger in M+DS and M+DC, compared to M+C.

Macrophages in M+S culture were visually distinct from all other conditions. While vesicles were visible within the cells, showing they too had carried out phagocytosis, their morphology was different in that nearly all cells had undergone significant elongation, resulting in a long and thin cell shape. This difference in morphology was also assessed quantitatively by comparing the length to width ratios of cells. The greater the ratio, the thinner the cell. The lengths and widths of macrophages within each of the images in Figure 3.10 were measured using the line tool of the image analysis software FIJI. A box plot of the resulting data is shown in Figure 3.11.

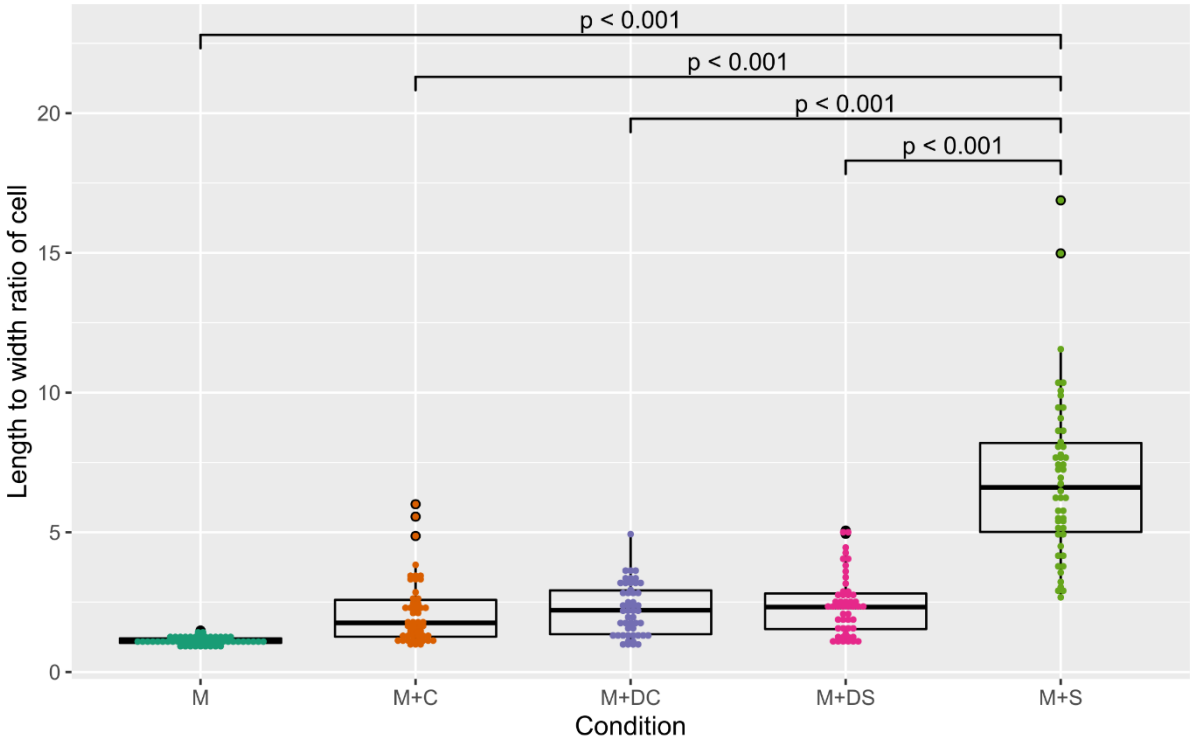


Figure 3.11 Box plot of length to width ratios of macrophages under the five different conditions of the indirect interaction assay. Lengths and widths measured for 50 cells in each of the images within Figure 3.10. Kruskal-Wallis test output: Chi squared = 157.85, df = 4, p-value < 0.001. M: macrophages in isolation. M+S: macrophages + *S. sudanensis.* M+C: macrophages + *S. coelicolor* M1152. M+DS: macrophages + dead *S. sudanensis*. M+DC: macrophages + dead *S. coelicolor* M1152.

The plot confirms, firstly, that the macrophages were activated, as indicated by the heterogeneous cell shapes, by exposure to bacterial components. This is indicated by the wide y-axis spread of data points for the four challenge conditions. Secondly, these data confirm the visual observation that in M+S the cells became significantly thinner compared to other conditions. The median length to width ratio for M+S was about 3- to 5-fold greater than for any of the other conditions. On average, a macrophage in M+S was over six times longer than it was wide. Even the lowest ratio for a cell in M+S was still greater than the median ratios for all other conditions. The visual and quantitative differences in morphology between the M+S and M+DS conditions were particularly striking. Dead *S. sudanensis* did not induce the same morphology as the living bacteria. The living mycetoma pathogen induced an entirely unique activation phenotype in the macrophages.

3.4 Overview of RAW-blue macrophage RNAseq dataset

RNA was harvested from the macrophages, as detailed in Chapter 2, across the five already listed conditions at the 48 hr time point. This was done for three biological repeats, with samples numbered accordingly. RNAseq libraries were then prepared as described in Chapter 2. Illumina sequencing of the libraries was performed by Edinburgh Genomics. The quality of the resulting sequencing data was then assessed using FastQC, which showed no significant problems.

The sequencing reads were mapped to the GENCODE M22 release of the *Mus musculus* transcriptome, using the mapping software Salmon. The number of reads that mapped to the transcriptome are listed in Table 3.1. The mapping rates ranged from 40.86% up to 70.51%.

Sample	Reads	Mapped reads	Mapped (%)
M_1	127,510,153	59,884,239	46.96
M_2	19,786,380	8,514,730	43.03
M_3	18,124,108	9,499,585	52.41
M+C_1	14,587,062	8,248,903	56.55
M+C_2	19,335,370	12,717,256	65.77
M+C_3	205,041,977	107,394,227	52.38
M+DC_1	12,983,242	5,911,534	45.53
M+DC_2	44,933,168	20,419,406	45.44
M+DC_3	25,698,685	12,107,259	47.11
M+DS_1	14,226,683	8,052,989	56.60
M+DS_2	24,908,678	13,779,254	55.32
M+DS_3	23,828,247	11,620,161	48.77
M+S_1	28,408,338	20,031,729	70.51
M+S_2	27,857,175	11,383,711	40.86
M+S_3	39,058,771	19,420,831	49.72

Table 3.1 Number of RAW-blue RNAseq reads obtained for each sample taken from the indirect interaction assay and their mapping rates to the *M. musculus* transcriptome.

The mapping rates were lower than expected, with eight of the samples having less than 50% of reads mapping. However, all but one of the samples had over 8,000,000 reads successfully mapped, which was considered high enough for differential expression analysis (Liu et al., 2014). Tximport was used to import data produced by Salmon into DESeq2, an R-based software package for differential expression analysis. DeSeq2 first carries out statistical normalisation of datasets, based on read library size and per gene average transcript lengths. The read counts for each gene are corrected based on the size of the read library. This is then combined with normalisation based on average transcript lengths. Longer transcripts are more likely to produce a greater number of reads and thus would otherwise be overrepresented in the data.

Following normalisation, the dispersion estimates for each gene were checked. Dispersion refers to the variance of the count data across the three repeats for each gene (Love et al., 2014). DeSeq2 can plot the dispersion and mean counts for each gene, allowing the relationship between the two to be visualised for the whole dataset. This acts as a diagnostic tool to check for irregularities in the data. The plot constructed from the macrophage data is shown in Figure 3.12. The dispersion should decrease as the mean counts increase and that is indeed the trend shown by the plot.

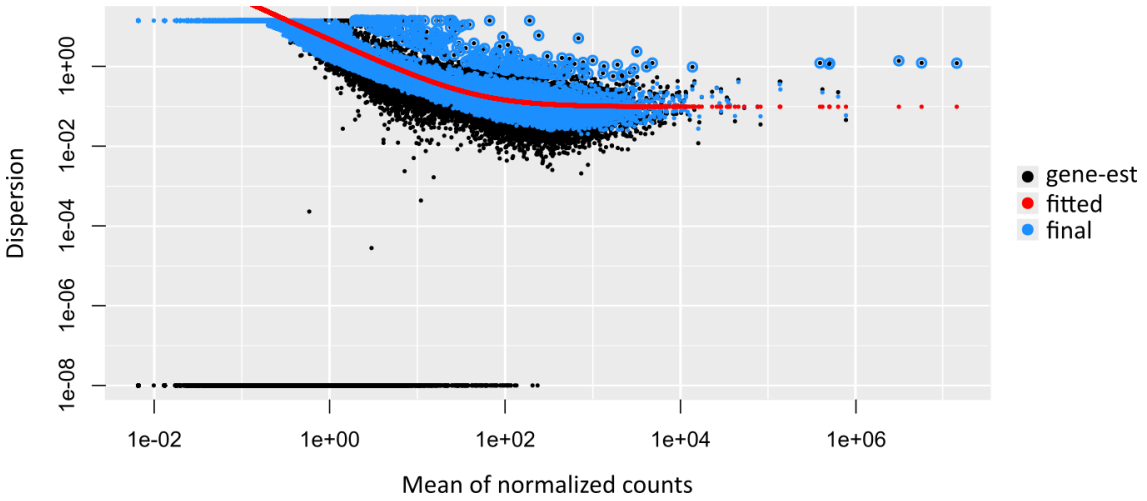


Figure 3.12 Plot produced by DESeq2 showing the relationship between dispersion and mean counts for each gene in the RAW-blue RNAseq dataset.

The data was next checked for outliers among the biological repeats. Firstly, principal component analysis (PCA) was performed and the resulting plot can be seen in Figure 3.13.

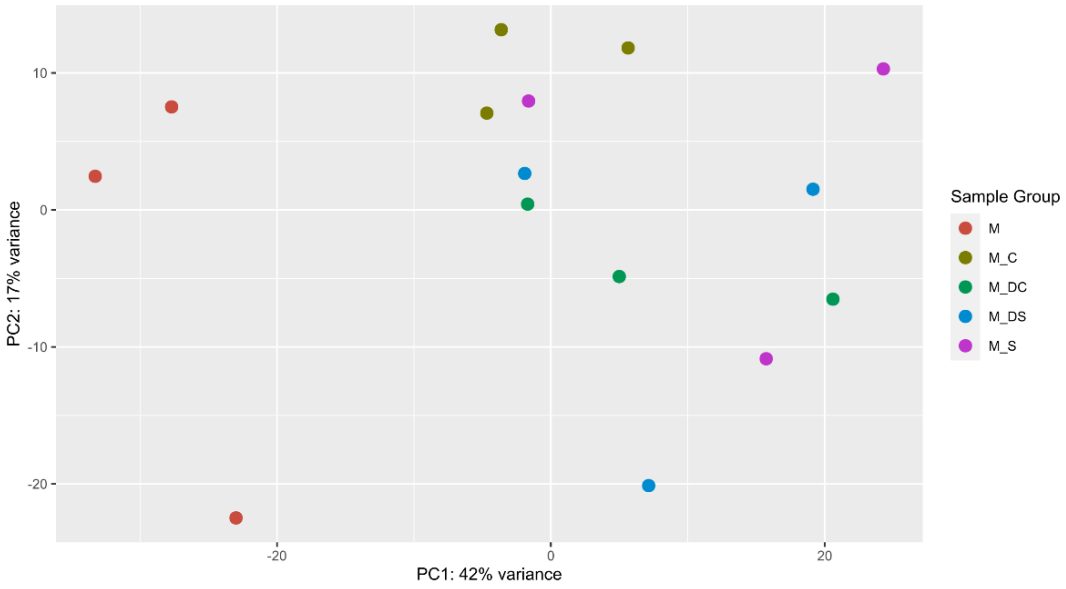


Figure 3.13 Plot of PCA of the RAW-blue RNAseq dataset, featuring all biological repeats.

Given that there were five different conditions in this experiment, a complex picture of the sources of variation between samples was to be expected. With the presence and absence of bacteria, the presence of *S. sudanensis* versus *S. coelicolor* and the effects of dead versus living cells all being compared, it was unlikely that any one principal component (PC) would account for the majority of the variance. This complexity also made interpretation of the PCA plot more challenging.

Indeed, the largest source of variance between samples (42%) was PC1. The negative control M condition samples were clustered at one end of PC1, with the remaining '+ bacteria' conditions spread from the midpoint of the axis across to the other end. Therefore, the largest proportion of the variance appeared to be mostly attributable to whether bacteria, dead or alive, were present. The grouping of the four '+ bacteria' conditions into a large cluster demonstrates that the macrophage transcriptional responses to various bacterial components were more similar to each other than to the macrophage resting state. There was also some intra-condition variation accounted for in PC1, with M+S, M+DS and M+DC all showing distribution of samples along this axis.

PC2 is the next most significant source of variance, accounting for 17%. It is likely reflective of intra-sample variation. One of the M samples is positioned at the opposite end of PC2, away from the remaining two. Additionally, the plot shows that M+S and M+DS are both partly distributed along PC2. It was also possible to pick out some inter-condition variance in PC2, with M+C clustered above the M+DC samples. This indicates variance attributable to whether *S. coelicolor* was alive or dead.

In terms of searching for outliers, a cautious approach was taken, as removal of any individual samples would reduce the number of biological repeats and thus the statistical power of the RNAseq data analysis. It was clear from the plot that sample M_2 was an outlier within its condition. No obvious outliers were identified among the other four conditions, with most of the samples within each condition being roughly equidistant from each other. Sample M_2 was accordingly removed from the dataset. The DESeq2 normalisation and statistical analyses described above were then repeated for the remaining samples.

Differentially expressed genes were identified by using DESeq2 to compare expression levels within each of the four bacterial challenge conditions against macrophages in isolation. The thresholds used for differential expression were a fold change in expression greater than +/-2 and an adjusted p-value of less than 0.05. Fold changes of +2 and -2 were chosen as the objective was to study the largest transcriptional changes induced. A doubling or halving of transcription seemed to be an appropriate starting point. The number of differentially expressed genes under each of the conditions is shown in Table 3.2.

	No. of differentially expressed genes		
Condition	**Upregulated**	**Downregulated**	**Total**
M+C	398	195	593
M+DC	538	251	789
M+DS	529	300	829
M+S	569	450	1019

Table 3.2 Numbers of differentially expressed genes found when comparing each of the listed conditions to macrophages in isolation.

The M+S condition had the largest total number of differentially expressed genes (1019), of which 56% were upregulated and the rest downregulated. It also had the largest individual numbers of both up- and downregulated genes. M+C had the smallest amount of transcriptional change, with 593 differentially expressed genes, of which 67% were upregulated. M+DS and M+DC had similar numbers of genes undergo differential expression, with M+DS having only 40 more. For each of them, 64% and 68% of genes respectively were upregulated. For all four conditions, most of the differential expression involved upregulation, rather than repression. Overall, macrophages exposed to living *S. sudanensis* showed the greatest amount of change in transcriptional activity.

Next, the numbers of genes that were uniquely up- or downregulated under each condition were quantified (Figure 3.14). Examination of both Venn diagrams revealed that M+S had a total of 398 uniquely differentially expressed genes, with 180 uniquely upregulated and 218 uniquely downregulated. M+C had a total of 125 such genes, of which 63 were upregulated and 62 downregulated. For M+DS, the total of unique genes was 137, with 62 and 75 up- and downregulated respectively. Finally, M+DC had 65 uniquely upregulated genes and 64 downregulated, giving a total of 129. The living pathogen therefore induced the greatest quantity of unique transcriptional activity.

The responses to living and dead bacteria were then compared. M+S and M+DS shared 491 differentially expressed genes, corresponding to 59% of M+DS genes and 48% of M+S genes. M+C and M+DC shared 358, which came to 60% of M+C genes and 45% of those in M+DC. These high proportions of shared genes indicated that living and dead bacteria induced a high amount of identical transcriptional activity in the macrophages. This supported the relevance of the use of dead cells as controls. Of further note in Figure 3.14(a) is that there were 208

upregulated genes shared by all conditions. This is the largest single group of genes in that Venn diagram and suggests a core transcriptional response shared by the macrophages across all challenge conditions.

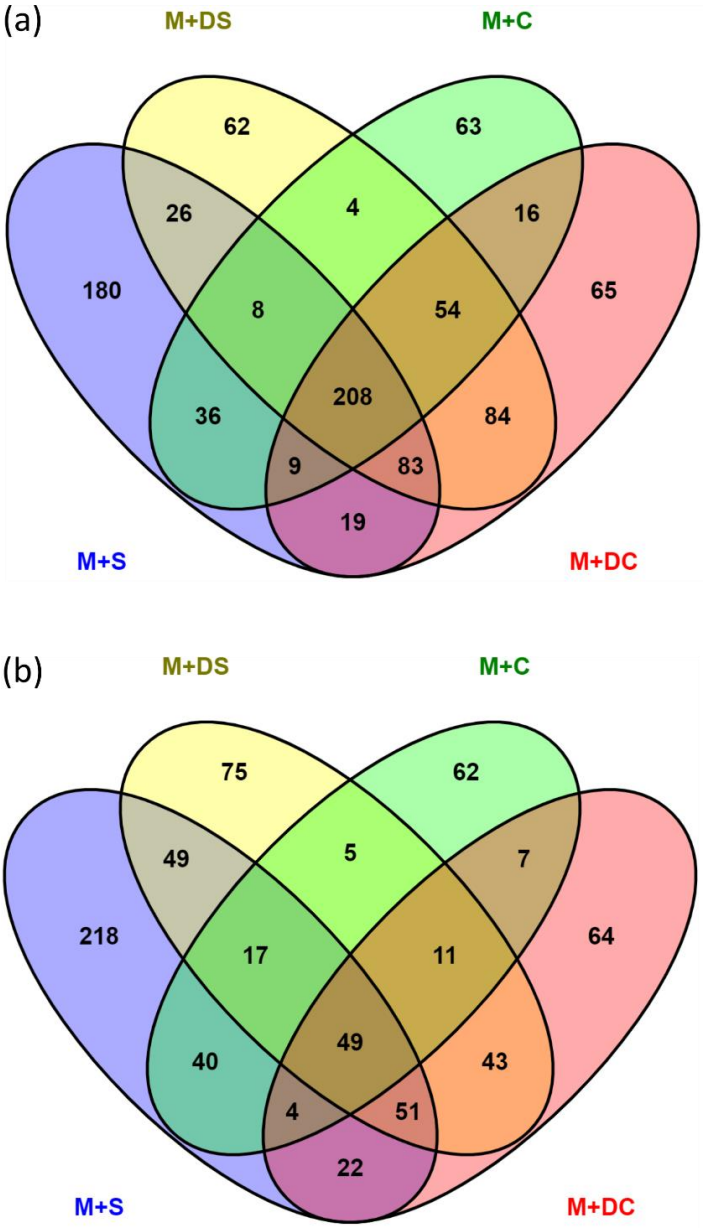


Figure 3.14 Venn diagrams showing numbers of differentially expressed RAW-blue genes shared between or unique to the four bacterial challenge conditions of the indirect interaction assay. **(a)** Upregulated genes **(b)** Downregulated genes.

The numbers from the Venn diagrams were also used to calculate the proportion of genes that were uniquely differentially expressed. This allowed for a better comparison of the scale of unique transcriptional activity between samples. Table 3.3 contains the manually calculated proportions.

	Differentially expressed genes that are unqiue (%)		
Condition	**Upregulated**	**Downregulated**	**Total**
M+C	15.8	31.8	21.1
M+DC	12.1	25.5	16.3
M+DS	11.7	25.0	16.5
M+S	31.6	48.4	39.1

Table 3.3 Proportion of RAW-blue genes within each challenge condition that were uniquely differentially expressed, shown as a percentage of each condition's total number of differentially expressed genes.

The M+S condition once again came out top, with 39% of its differentially expressed genes being unique. Despite having the lowest total number of differentially expressed genes, M+C has the second highest proportion of unique genes. So, not only did living *S. sudanensis* induce the greatest total amount of transcriptional change in macrophages, but it also induced the greatest amount of unique change, in terms of both quantity and proportion.

3.5 Comparison of changes to pro-inflammatory gene expression induced in RAW-blue macrophages during the indirect interaction assay

With the differentially expressed genes identified within all four challenge conditions, the data was manually examined for genes of interest. Genes involved in the pro-inflammatory, M1 macrophage response, based on their annotations in the UniProt database, were analysed first (UniProt, 2021). A heat map of fold changes for a selection of these genes is shown in Figure 3.15.

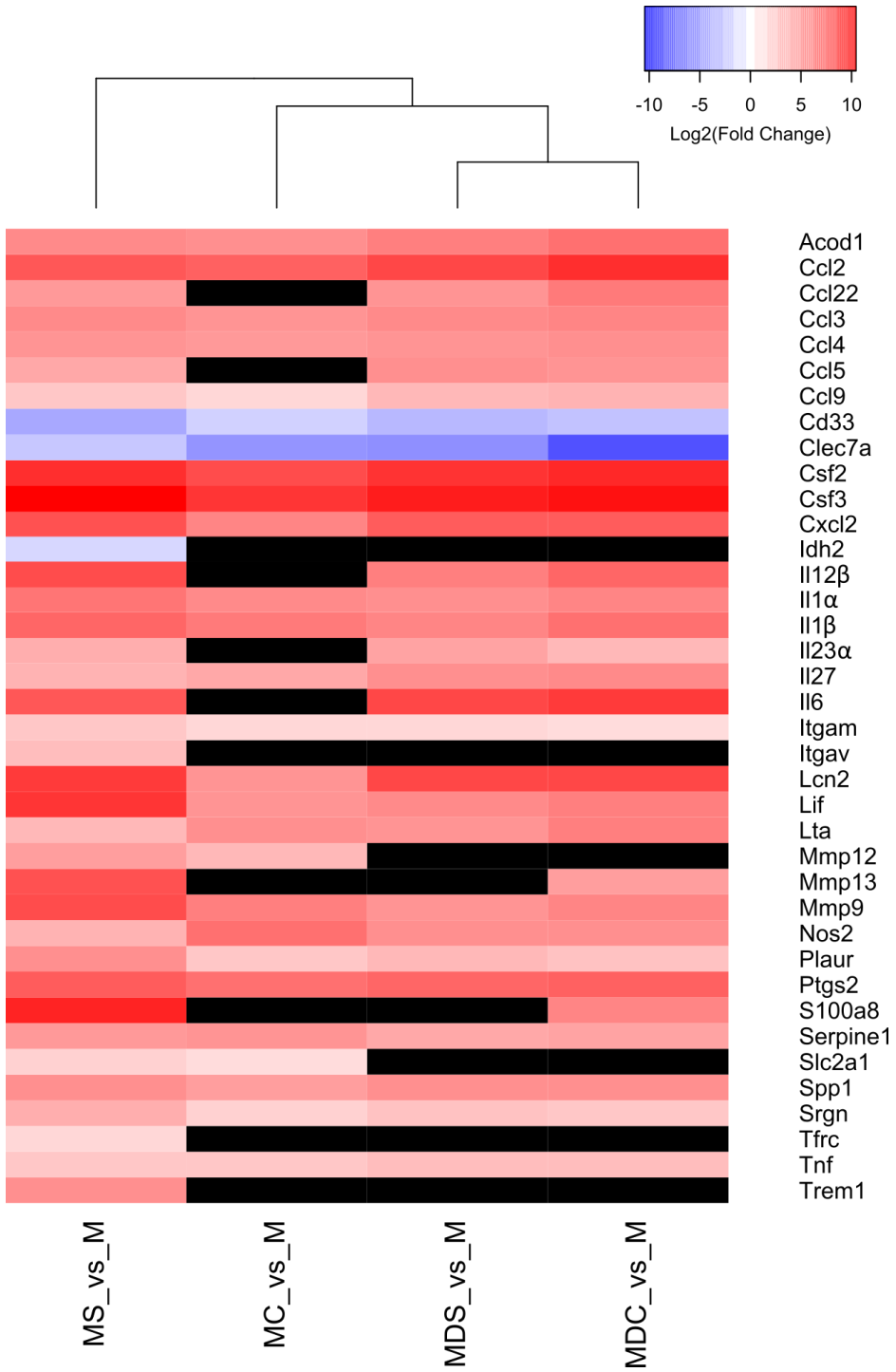


Figure 3.15 Heat map of the transformed fold changes in expression of pro-inflammatory associated genes during the indirect interaction assay, across all four challenge conditions. Samples are clustered by overall similarity, shown by the dendrogram at the top of the heat map. Black shading indicates genes that did not meet the set thresholds for fold change or adjusted p-value under a specific condition and were therefore not considered to be differentially expressed.

Descriptions of the functions for each of the above genes are detailed in Table 3.4.

Gene	Function of gene product	UniProt entry code
Acod1	Aconitate decatboxylase. Converts aconitate to itaconate.	P54987
Ccl2	Chemokine that acts through CCR2. Chemotactic for monocytes and basophils.	P10148
Ccl22	Chemokine that is chemotactic for T-lymphocytes.	O88430
Ccl3	Monokine with potent inflammatory activity and chemotactic for eosinophils.	P10855
Ccl4	Monokine with inflammatory and chemotactic activities.	P14097
Ccl5	Chemokine that attracts monocytes, eosinophils and T-helper cells.	P30882
Ccl9	Monokine with inflammatory and chemotactic activities.	P51670
Cd33	Cell surface antigen that maintains immune cells in a resting state.	Q63994
Clec7a	C-type lectin that acts as a PRR for bacterial and fungal cell wall components. Acts via TLR2.	Q6QLQ4
Csf2	Granulocyte-macrophage colony-stimulating factor. Promotes growth and differentiation of immune cells.	P01587
Csf3	Granulocyte colony-stimulating factor. Induces production, differentiation and activation of granulocytes.	P09920
Cxcl2	Chemokine that attracts granulocytes.	P10889
Idh2	Isocitrate dehydrogenase. Activity is part of the Krebs cycle.	P54071
Il12 β	Cytokine that associates with IL23A, to form IL23. Induces additional pro-inflammatory production.	P43432
Il1 α	Cytokine produces by activated macrophages. Inflammatory cytokine through pyrogenic activity.	P01582
Il1 β	Highly potent inflammatory cytokine, a major inducer of the pro-inflammatory response.	P10749
Il23 α	Cytokine that associates with IL12, to form IL23. Induces additional pro-inflammatory production.	Q9EQ14
Il27	Cytokine that can have pro- or anti-inflammatory activity, depending on its environmental context.	Q8K316
Il6	Cytokine with pleiotropic immune effects. Induces acute phase inflammatory response to infections.	P08505
Itgam	Integrin involved in adhesive interactions between immune cells and mediates uptake of complement-coated pathogens.	P05555
Itgav	Integrin involved in adhesive interactions between immune cells and a receptor for multiple signals, including SPP1, osteopontin.	P43406
Lcn2	Lipocalin, a siderophore that sequesters iron from bacterial siderophores.	P11672
Lif	Member of the IL6 family of cytokines. Has pleiotropic effects in response to injury and inflammation.	P09056
Lta	Lymphotoxin alpha. Inflammatory cytokine that can induce death of infected host cells.	P09225
Mmp12	Matrix metalloproteinase that hydrolyses elastin. Involved in tissue remodelling in response to injury.	P34960
Mmp13	Matrix metalloproteinase that degrades multiple ECM proteins, such as collagen. Involved in tissue remodelling and wound healing.	P33435
Mmp9	Matrix metalloproteinase that degrades ECM proteins and promotes leukocyte migration.	P41245
Nos2	Inducible nitric oxide synthase. Produces NO, an antibacterial molecule, and enhances inflammatory cytokine production.	P29477
Plaur	Receptor for urokinase plasmin activator. Promotes plasmin formation. Upregulated in various inflammatory disorders.	P35356
Ptgs2	Prostaglandin synthase, also know as COX-2. Synthesises pro-inflammatory lipids.	Q05769
S100a8	One of 2 subunits of the iron, zinc and calcium chelator calprotectin. Calprotectin has pro-inflammatory and antibacterial effects.	P27005
Serpine1	Inhibitor of PLAUR and so inhibits plasmin activation and blood clotting. Inhibits wound healing.	P22777
Slc2a1	Glucose transporter. Upregulation associated with M1 macrophage phenotype.	P17809
Spp1	Osteopontin. Acts as cytokine which promotes IL-12b and IFN-g production.	P10923
Srgn	Serglycin. Involved in formation of secretory granules in immune cells. Specifically regulates secretion of TNF.	P13609
Tfrc	Transferrin receptor. Allows cells to uptake extracellular iron via receptor mediated endocytosis.	Q62351
Tnf	Potent pro-inflammatory cytokine, secreted by macrophages and causes fever. Stimulates IL-1B secretion and further inflammation.	P06804
Trem1	Cell surface receptor that amplifies inflammatory responses. Has role in chronic inflammatory diseases.	Q9JKE2

Table 3.4 Function annotations for genes listed in Figure 3.15, obtained from the UniProt database (UniProt, 2021).

The heat map shows that key pro-inflammatory genes, such as *Tnfα, Il1α*, *Il1β* and *Nos2*, were upregulated across all four challenge conditions. In fact, the pro-inflammatory colony stimulating factor genes *Csf2* and *Csf3* were within the top twelve most highly upregulated genes for each condition. Interestingly, the dendrogram in Figure 3.15 indicates that, for this set of genes, the M+S condition was the most dissimilar from the others. The two dead bacteria conditions clustered together and were therefore the most similar, with M+C then clustered next, followed lastly by M+S. Also of note is that for 20 of the 37 genes listed in the heat map, their magnitudes of fold change were greatest under M+S. When M+S was then compared to M+C only, i.e. when the two conditions with living bacteria were compared, this number rose further to 32 out of 37, or 87% of the genes.

Nine interleukin genes were differentially expressed in the M+S dataset, of which six are shown in Figure 3.15. Interleukins are a major class of cytokines and, as described in Table 3.4, these six are drivers of the pro-inflammatory response. Three of these genes underwent the greatest upregulation in M+S, namely *Il1α*, *Il1β* and *Il12β*. *Il23α* was most upregulated in M+DS, but then followed by M+S, meaning *S. sudanensis* triggered stronger expression than *S. coelicolor*. Upregulation of five of the six interleukins was greater in M+S than in M+C. In fact, M+C didn't even induce differential expression of *Il6*. The genes *Csf2* and *Csf3*, again encoding cytokines, were also more highly upregulated in M+S than in M+C.

A pro-inflammatory cytokine gene notable for its highly extreme upregulation in M+S was *Lif*, encoding a member of the IL6 family. While its expression was raised in all conditions, it was several orders of magnitude higher in M+S, with a fold change of +265. This made it the eighth most highly upregulated gene in the M+S dataset. For comparison, *Lif* fold changes in the other three conditions ranged from +21 to +36.

There are six pro-inflammatory chemokines listed in Figure 3.15. *Ccl2, Ccl3, Ccl4* and *Ccl9* experienced strong upregulation across all conditions. *Ccl5* and *Ccl22*, however, were not differentially expressed in M+C. Upregulation of these genes in both M+DS and M+DC was higher than in M+S. However, M+S did show greater upregulation across all six chemokines when compared to M+C.

As described in Chapter 1, matrix metalloproteinases (MMPs) are proteolytic enzymes which have their expression induced by pro-inflammatory conditions (Nissinen and Kahari, 2014,

Tomlin and Piccinini, 2018). There are three MMPs in Figure 3.15 and their specific functions are described in Table 3.4. All three were most strongly upregulated in M+S. *Mmp9* and *Mmp13* underwent particularly strong upregulation, with fold changes of +138 and +122 respectively. Fold changes under the other conditions went no higher than +34.

Upregulation of siderophores can also be seen in the dataset. *Lcn2*, encoding lipocalin, was upregulated in all conditions, with M+S once again showing the greatest increase and M+C the lowest. *S100a8* was also upregulated, but only in M+S and M+DC (Figure 3.15). The magnitude of the fold change in M+S, however, was much greater than in M+DC: +493 compared to +32. In fact, this was the fifth most highly upregulated gene in the M+S data set. As described in Table 3.4, this metal chelator also has pro-inflammatory activity. Furthermore, as seen in the heat map, the siderophore receptor *Tfrc* was uniquely upregulated in M+S.

The heat map also gives evidence that cells in all conditions underwent metabolic changes associated with the M1 macrophage activation state. Specifically, there are four genes directly involved in metabolic processes (Table 3.4). Upregulation of *Acod1* is observed in nearly all pro-inflammatory leukocytes and so is a reliable marker for inflammation (O'Neill and Artyomov, 2019). It was upregulated in all experimental conditions within this dataset.

Pro-inflammatory macrophages undergo the Warburg shift, a metabolic change whereby aerobic glycolysis is upregulated for the rapid production of ATP, while the Krebs cycle is repressed (Wilson et al., 2019). Upregulation of *Slc2a1* and downregulation of *Idh2* are two markers of the Warburg shift and both occurred in the M+S condition. Only *Slc2a1* upregulation was seen in M+C, while neither gene was differentially expressed in M+DS or M+DC. *Ptgs2* is another metabolic gene of interest, particularly as a target for non-steroidal anti-inflammatory drugs. Once again, it's upregulation was greatest in M+S, followed by M+DC, M+DS and then M+C.

In addition to *Idh2*, *Cd33* is another gene whose downregulation promotes inflammatory activity. As the product of this gene maintains immune cells in a resting state, its downregulation presumably leads to greater activation. It was downregulated in all four conditions, but crucially, most strongly in M+S. *Clec7a* is the only remaining downregulated gene not discussed and is of note because it was least downregulated in M+S. Being a PRR that

triggers an inflammatory response, the comparison of fold changes suggested that cells in M+S remained more sensitive to inflammatory stimuli.

There was at least one key aspect of the M+S pro-inflammatory response that was weaker than the other conditions, namely the upregulation of *Nos2*. Expression increased by only 8-fold in M+S, compared to +24, +22 and +52 in M+DS, M+DC and M+C respectively. Production of nitric oxide by the protein product of this gene is a standard macrophage weapon used to kill invading bacteria (MacMicking et al., 1997). Thus, there were clear differences between the transcriptional responses of proinflammatory genes to living and dead *S. sudanensis* and particularly between pathogen and non-pathogen.

3.6 Other transcriptional changes induced in RAW-blue macrophages by *S. sudanensis*

3.6.1 Unique changes in transcriptional regulation

Among the genes uniquely differentially expressed under M+S were two that form a transcription factor (TF), both of which were upregulated. A third TF gene was upregulated in both M+S and M+DS only and so was still uniquely induced by *S. sudanensis* (Table 3.5).

Gene	Fold Change	Adjusted p-value
Jun	2.1	0.0029
Fosl2	2.2	0.0041
Cebpb	1.5	0.0026

Table 3.5 Differential expression data for three transcription factor genes, uniquely upregulated in the M+S condition.

The protein products of *Jun* and *Fosl2* dimerise to form an AP-1 TF (Karin et al., 1997). As described in Chapter 1, AP-1s are a family of TFs that activate expression of immune genes in response to infection (Fitzgerald and Kagan, 2020). The related genes *Junb* and *Fosl1* also encode proteins that can dimerise to form their own AP-1 TF and were upregulated in all four assay challenge conditions, again showing there was a common response to exposure to bacteria. But *Jun* and *Fosl2* upregulation being confined to M+S implies a TF not found in the other three conditions may have been expressed in response to the living pathogen.

Cebpb encodes a TF reported to be heavily involved in various immune processes, particularly inflammation (Akira et al., 1990, Uematsu et al., 2007). Binding sites for this TF have been found upstream of multiple key pro-inflammatory genes, including *Tnfα*, *Il1α*, *Il6* and *Nos2*. It is therefore interesting to note that both *S. sudanensis* conditions induced its upregulation, while the *S. coelicolor* controls did not. Again, this points to the potential presence of a unique TF within the macrophages under these two conditions.

It was also noted that the gene *Batf* was upregulated in M+C, M+DC and M+DS, but not differentially expressed in M+S. This gene encodes a protein that dimerises with JUNB to form yet another AP-1 TF (Williams et al., 2001). So, in this instance, cells exposed to M+S may be lacking transcription regulatory activity seen in the other four conditions.

3.6.2 Unique upregulation of anti-inflammatory wound healing

As detailed in Chapter 1, wound healing is associated with the M2 macrophage phenotype and forms part of the anti-inflammatory immune response (Mosser and Edwards, 2008). Genes involved in this response generally do not contribute to the killing of bacterial pathogens. Figure 3.16 shows a heat map of genes involved in the promotion of anti-inflammatory wound healing, five out of six of which were only differentially expressed in M+S. Basic functions of the proteins encoded by these genes are detailed in Table 3.6.

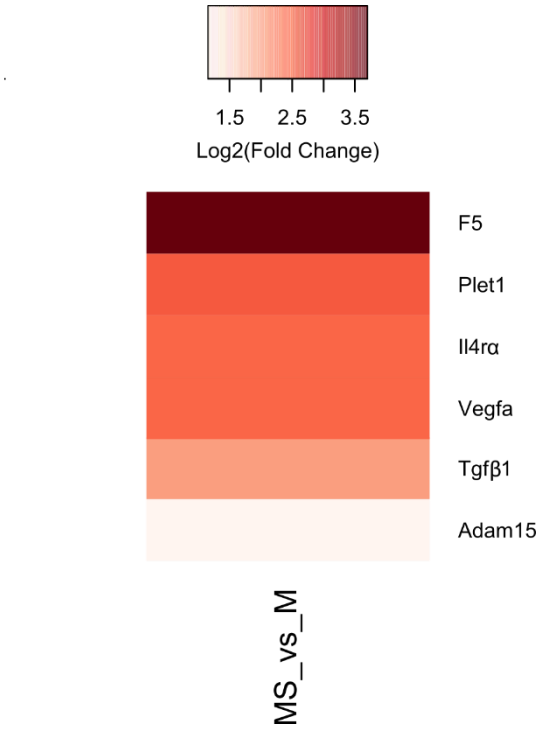


Figure 3.16 Heat map of transformed fold changes for genes involved in the promotion of anti-inflammatory wound healing, under the M+S condition.

Gene	Function of gene product	UniProt entry code
F5	Central regulator of haemostasis. Cofactor for Factor Xa and so facilitates proteolytic activation of thrombin.	O88783
Plet1	Modulates adhesion of cells to extracellular matrix proteins and promotes wound healing.	Q8VEN2
Il4r α	Receptor for IL4, an anti-inflammatory cytokine that induces the M2 macrophage activation phenotype.	P16382
Vegfa	Growth factor that induces angiogenesis and endothelial cell growth. Permeabilises blood vessels and promotes macrophage and neutrophil migration.	Q00731
Tgf β 1	Regulates multiple cell types. Suppresses macrophages at high concentrations, but is chemotactic at low levels. Promotes bone remodelling and wound healing.	P04202
Adam15	Membrane-bound metalloproteinase that cleaves collagen, gelatin and proteolytically processes adhesion molecules. Active in wound healing processes.	O88839

Table 3.6 Function annotations for genes listed in Figure 3.7, obtained from the UniProt database (UniProt, 2021).

Il4rα underwent upregulation in M+DC as well as M+S. However, it's fold change under that condition was +3.8, compared to +5.5 for M+S. Furthermore, the adjusted p-value for the change in M+DC was on the edge of significance, being 0.04, while for M+S it was 0.004, a factor of ten smaller. So, the upregulation of *Il4rα* was both far more biologically and statistically powerful during challenge with living *S. sudanensis*, hence why it was included within this group of genes.

3.6.3 Suppression of the mevalonate pathway

Uniquely downregulated within the M+S dataset were a group of genes encoding enzymes which catalyse sequential steps of the mevalonate pathway. A heat map of these genes is visible, alongside their places in the pathway, in Figure 3.17. As Figure 3.17(b) shows, the pathway is responsible for cholesterol and isoprenoid biosynthesis through its sterol and non-sterol branches respectively (Gruenbacher and Thurnher, 2017). Mammalian cell membranes are comprised of approximately 25% cholesterol, which maintains their fluidity across varying temperatures (Ikonen, 2008). Cholesterol also contributes to the formation of lipid rafts within membranes, which are important for cell signalling. Finally, it also acts as a precursor molecule for the synthesis other sterols, including steroid hormones. Isoprenoids are commonly used for the post-translational modification of proteins (Wang and Casey, 2016). The mevalonate pathway is therefore highly important for maintaining cell homeostasis.

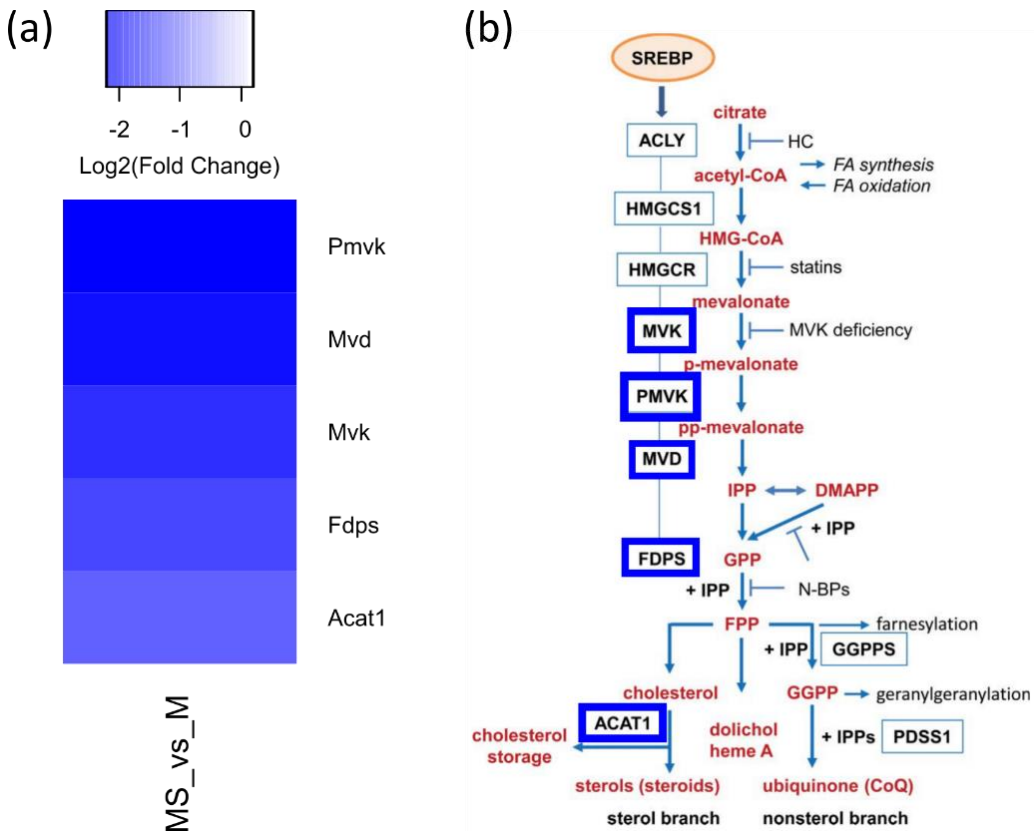


Figure 3.17 (a) Heat map of transformed fold changes for genes involved in the mevalonate pathway, under the M+S condition. **(b)** The mevalonate pathway. Bold blue boxes indicates proteins whose genes were uniquely downregulated in M+S. Pathway Figure adapted from Gruenbacher and Thurnher (2017).

3.6.4 Priming of the NLRP3 inflammasome

Upregulation of the gene *Nlrp3* was observed only in the M+S and M+DS conditions, so uniquely in the presence of *S. sudanensis*. Under M+S, the gene underwent a +3.8-fold change in expression, with an associated adjusted p-value <0.001. Under M+DS, the fold change was +2.4, with an associated adjusted p-value of 0.047. Of immediate note is that the fold change in M+S, so in the presence of living bacteria, was larger than that in M+DS. The adjusted p-value for the fold change under M+DS was also at the edge of the significance threshold set for differential expression. The same measure for the M+S fold change was many times smaller, indicating a much stronger statistical trend across all three biological repeats for this upregulation.

Nlrp3 encodes a PRR, which acts as the receptor domain of the NLRP3 inflammasome (Silvis et al., 2021). NLRP3 oligomerises with the proteins ASC and pro-caspase-1, forming a multimeric protein assembly (Figure 3.18). The assembled inflammasome acts as an intracellular sensor for the presence of pathogens and their activities.

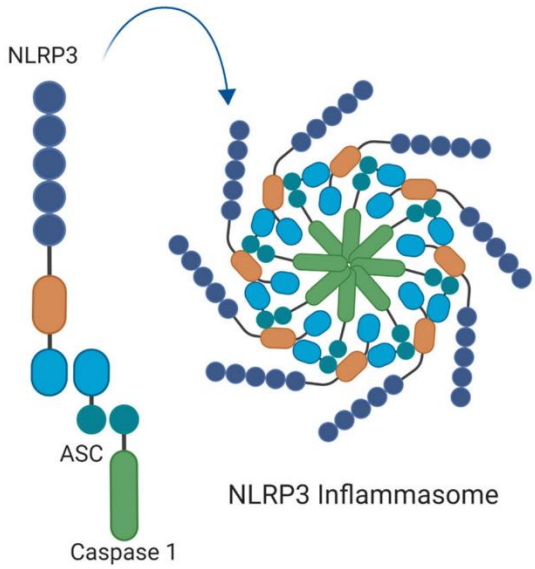


Figure 3.18 Structure of the NLRP3 inflammasome. Figure taken from Silvis et al. (2021).

Following detection of an activation signal, conformational shape changes in the inflammasome lead to the proteolytic activation of caspase-1 (Kelley et al., 2019). This caspase then cleaves and activates the inflammatory cytokines pro-IL1β and pro-IL18. Gasdermin D is also cleaved into an active form by caspase-1 and assembles in the cell membrane to form pores. These pores trigger a programmed form of cell death called pyroptosis. This is an inflammatory cell death, releasing pro-inflammatory cytokines and host-derived DAMPs. This greatly enhances the pro-inflammatory response in the local environment, promoting pathogen killing, but also causes host tissue damage.

It is currently thought that the NLRP3 inflammasome is regulated by a two-signal system (Figure 3.19) (Kelley et al., 2019). Signal one can come in the form of cytokines or PAMPs and triggers intracellular signalling cascades that ultimately lead to the NF-κB-mediated expression of *Nlrp3*. Once the NLRP3 protein is present at a minimum threshold concentration, it oligomerises with the constitutively expressed ASC and pro-caspase-1 to form the complete inflammasome. This whole process is known as inflammasome priming. The primed inflammasome then awaits detection of signal two, known as inflammasome activation

(Figure 3.19). This signal can come in numerous forms, the best characterised of which is potassium ion efflux. Once signal two has been provided, the inflammasome then initiates pyroptosis and activates the pro-inflammatory cytokines.

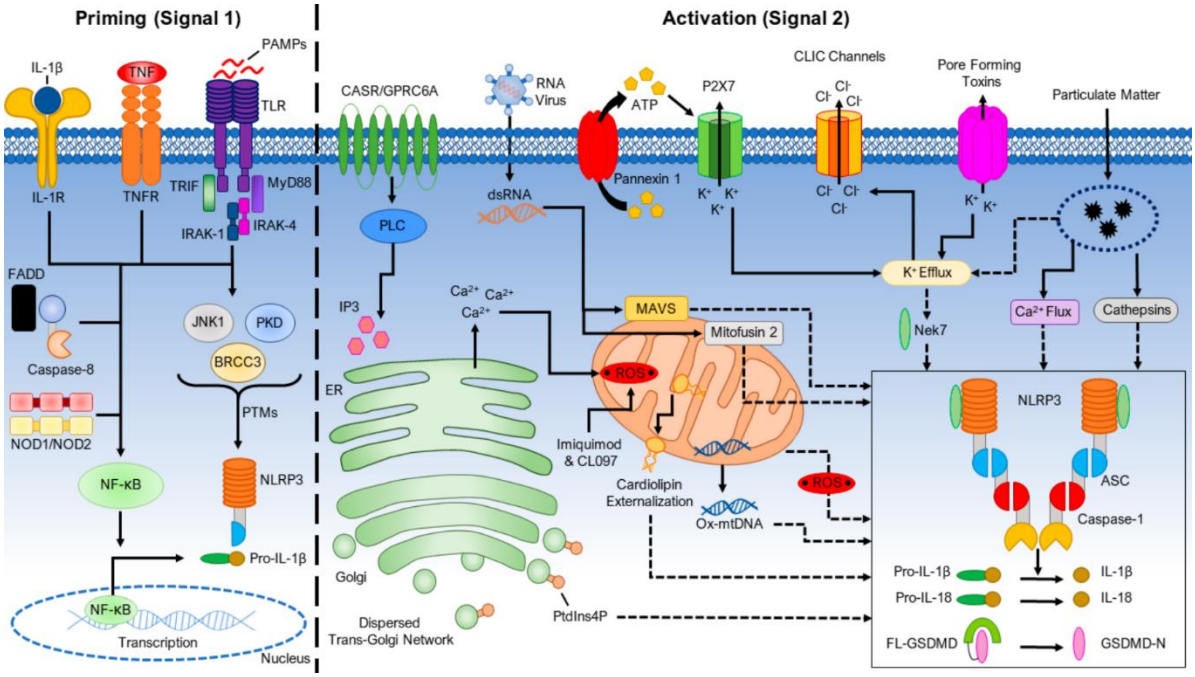


Figure 3.19 The two-signal system for priming and activation of the NLRP3 inflammasome. Figure from Kelley et al. (2019).

With upregulation of *Nlrp3* detected in the RNAseq data, it can be stated that both living and dead *S. sudanensis* induced priming of the NLRP3 inflammasome. Living and dead *S. coelicolor* on the other hand, failed to do this. It is important to note at this point that the gene encoding the inflammasome structural protein ASC is naturally transcriptionally silent in the RAW-blue cell line (Pelegrin et al., 2008). This means RAW-blue cells are unable to assemble an inflammasome and therefore cannot undergo pyroptosis or cleave IL1β into its active form.

3.7 Cytokine production by RAW-blue macrophages during the indirect interaction assay

In addition to RNA, macrophage culture supernatants (SNs) were also harvested from each condition during the interaction assay, and multiple ELISAs were used to detect and quantify cytokine release.

3.7.1 Tumour necrotic factor α

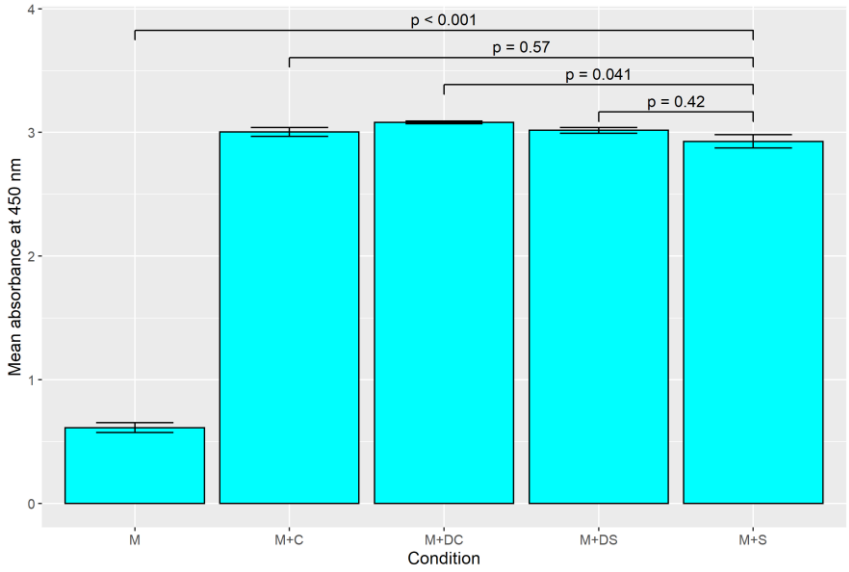


Figure 3.20 Results of ELISA measuring quantities of TNFα released by RAW-blue macrophages at 48 hrs, during the indirect interaction assay. Data consists of six biological repeats. Errors bars show the standard error of the mean. Displayed p-values were calculated using Tukey's Honestly Significant Difference pairwise comparison test, following an ANOVA test on the whole dataset.

Figure 3.20 shows that supernatants of macrophages in their resting state, condition M, did not contain significant quantities of the pro-inflammatory mediator TNFα. Interpolation from the standard curve predicted the concentration to be 9 pg/ml. This was expected, as no inflammatory stimuli should have been present within that well. TNFα accumulation across each of the four challenge conditions was relatively uniform and substantially greater than in the M condition. The absorbances shown in Figure 3.20 convert to concentrations ranging from 633 pg/ml for M+S, up to 675 pg/ml for M+DC. ANOVA testing of the dataset gave an F-value of 904.4, with an associated p-value <0.001. Post hoc pairwise comparisons to M+S further demonstrated the likelihood that this condition triggered similar TNFα release to the other challenge conditions, as shown by the high p-values in Figure 3.20.

3.7.2 Transforming growth factor β

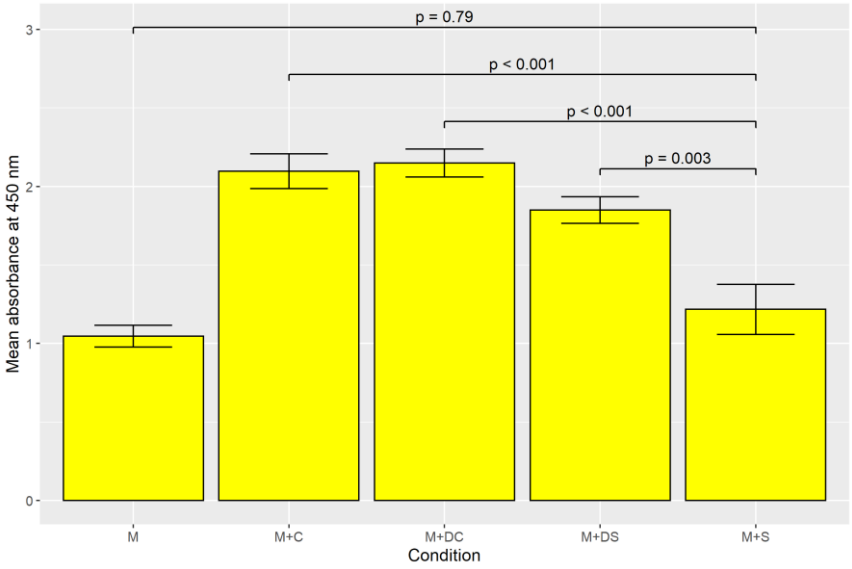


Figure 3.21 Results of ELISA measuring quantities of TGFβ released by RAW-blue macrophages at 48 hrs, during the indirect interaction assay. Data consists of six biological repeats. Errors bars show the standard error of the mean. Displayed p-values were calculated using Tukey's Honestly Significant Difference pairwise comparison test, following an ANOVA test on the whole dataset.

From Figure 3.21, it can be seen that TGFβ releases in M and M+S were very similar, with interpolation from the standard curve giving predicted concentrations of 221 pg/ml and 278 pg/ml, respectively. The next highest TGFβ release occurred in M+DS, at 520 pg/ml. Dead pathogen therefore had a stronger effect than the living cells. The two conditions featuring *S. coelicolor* produced the highest levels of the cytokine. ANOVA testing gave an F-value of 22.47, with an associated p-value <0.001. As shown in Figure 3.21, pairwise post hoc testing produced a very high p-value for M-M+S, while p-values for comparisons to the other challenge conditions were extremely low.

3.7.3 IL10

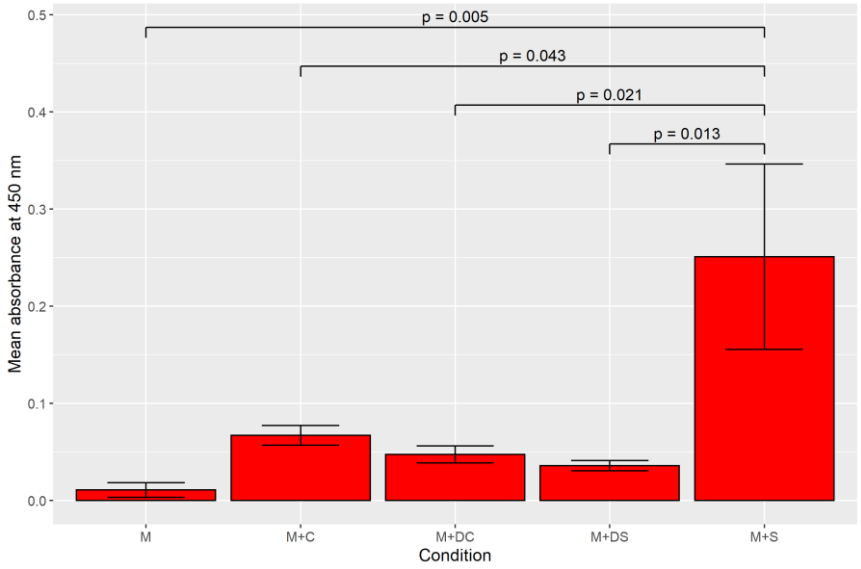


Figure 3.22 Results of ELISA measuring quantities of IL10 released by RAW-blue macrophages at 48 hrs, during the indirect interaction assay. Data consists of six biological repeats. Errors bars show the standard error of the mean. Displayed p-values were calculated using Tukey's Honestly Significant Difference pairwise comparison test, following an ANOVA test on the whole dataset.

Figure 3.22 shows that only the M+S condition induced significant release of the anti-inflammatory cytokine IL10. The absorbance shown for M+S converted to a concentration of 622 pg/ml, while the measurements for M, M+DC and M+DS were below the sensitivity limit of the ELISA. An ANOVA test gave an F-value of 4.958 for these data, with an associated p-value of 0.004. Pairwise comparisons to M+S gave the lowest p-values following post hoc testing, as shown in figure 3.22. P-values from other comparisons were all greater than 0.88.

3.8 NF-κB activity

As described above, the RAW-blue cell line contains a SEAP that is inducible by NF-κB. SEAP levels in culture SNs harvested during the indirect interaction assay were quantified using the QUANTI-Blue assay to measure the activity of this transcription factor within each interaction assay condition. Results are plotted in Figure 3.23.

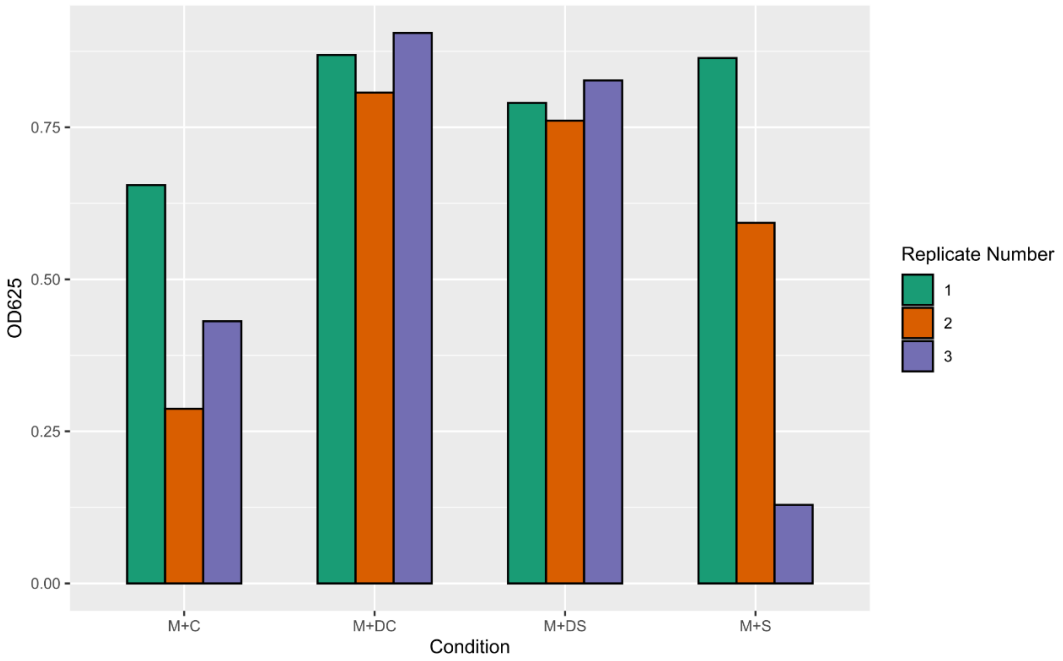


Figure 3.23 Results of QUANTI-Blue assay to detect SEAP production from RAW-blue cells during the indirect interaction assay, at 48 hrs. OD625 measured for macrophages in isolation was used as a blank. Sample key: M+C – macrophages + *S. coelicolor*, M+DC - macrophages + dead *S. coelicolor*, M+DS – macrophages + dead *S. sudanensis*, M+S – macrophages + *S. sudanensis*.

The plot shows that NF-κB activity was induced in all four conditions, across all three biological replicates. Conditions containing dead bacteria triggered consistently high levels of NF-κB activity, with only minimal variation between the three repeats. For M+S and M+C, however, NF-κB activity was not consistent across all replicates. This variation was greatest for M+S. M+S did, however, induce greater activity than M+C in two of the three repeats.

3.9 The response of human THP-1 monocytes to *S. sudanensis*

As shown in 3.6.4, *S. sudanensis* was found to prime the NLRP3 inflammasome. But as RAW-blue cells are unable to assemble an inflammasome, it remained unknown as to whether full inflammasome activation and subsequent pyroptosis would also be induced by the pathogen. It was therefore decided to use a different cell line, the human monocytic cell line, THP-1 (Tsuchiya et al., 1980), to investigate mycetoma pathogen-induced inflammasome activation. These cells can be differentiated into macrophage-like cells, for which they are well established models (Chanput et al., 2014).

3.9.1. Visual observations of THP-1 cells following challenge with S. sudanensis culture supernatant

As living bacteria induced a stronger priming effect than dead bacteria, it was speculated that the cause of inflammasome priming might be a metabolite produced by live bacteria. Additionally, the stimulus that induced priming must have diffused through the Transwell insert during the assay and so it was anticipated that the metabolite should be present in the SN of *S. sudanensis* cultures. It was therefore decided to use a simplified challenge assay method, whereby sterile bacterial culture SNs would be added directly to THP-1 cells. Images of THP-1 cells challenged with culture SNs are shown in Figure 3.24.

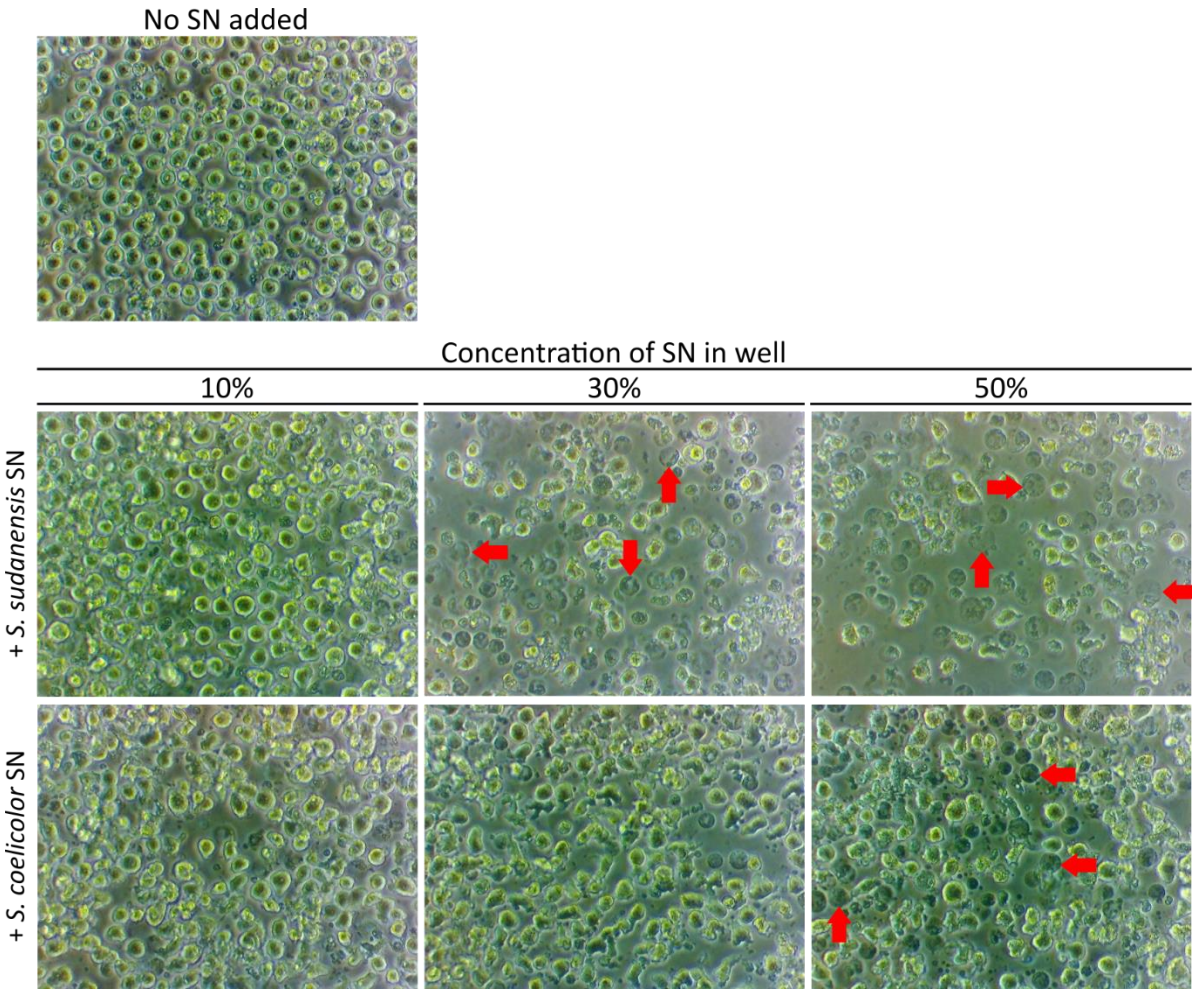


Figure 3.24 Images taken at 48hrs of THP-1 cells challenged with sterile SNs from *S. sudanensis* and *S. coelicolor* M1152 cultures. Images taken at 200X magnification and representative of three biological repeats. Red arrows indicate examples of the swollen, perfectly round and faded cell death phenotype of interest.

As shown in Figure 3.24, both sets of SNs induced enlargement and irregular morphologies in cells compared to the negative control, from a concentration of 10% upwards. But at a concentration of 30%, the *S. sudanensis* SN induced the formation of many perfectly round, swollen cells, which also seemingly lacked nuclei and other organelles, based on their phase pale appearance. This phenomenon was even more pronounced at 50% SN concentration, where this phenotype has become dominant in the well. Since they lack a nucleus, these cells were almost certainly dead. In contrast, the *S. coelicolor* SN did not have this effect at 30% concentration. At 50%, multiple perfectly round cells were visible, however they were darkened, rather than faded, and they were not the dominant cell type in the well. It was concluded that *S. sudanensis* is a prolific producer of a molecule that induces a specific form of cell death.

To check whether components of the bacterial culture media, GYE, were inducing the observed THP-1 phenotype, cells were challenged with sterile GYE media alongside *S. sudanensis* SN. Images from this assay trial are shown in Figure 3.25.

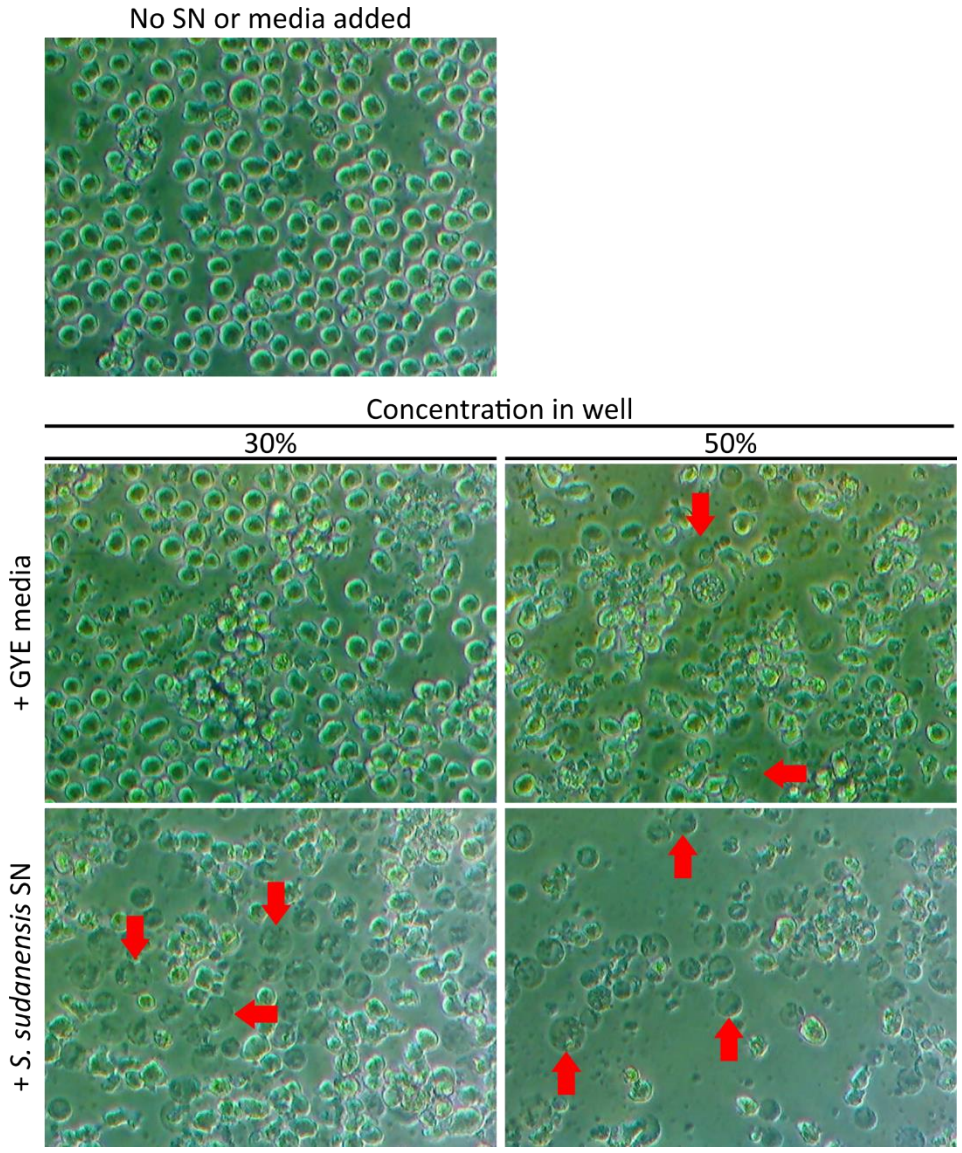


Figure 3.25 Images taken at 48hrs of THP-1 cells challenged with sterile SN from *S. sudanensis* culture and sterile GYE culture medium. Images taken at 100X magnification and representative of three biological repeats. Red arrows indicate examples of the swollen, perfectly round and faded cell death phenotype of interest.

The figure shows again that the *S. sudanensis* SN induced the same swollen, "ghost-like" cell death at 30% concentration and upwards. GYE did not induce this effect at 30% but did show some toxicity. Shrivelled cells are visible in the image in Figure 3.25, alongside small black debris, probably the remains of dead cells. At 50% concentration, the toxicity of the GYE was even more pronounced. Also visible at this concentration were a small minority of cells which

appeared to have the same phenotype as that induced by the *S. sudanensis* SN. They are visible in 50% GYE image in Figure 3.25. The induction of the phenotype was visually not as strong as with the pathogen SN, as the swollen dead cells were not dominant in the well. This difference in strength is very clear between the 50% images in Figure 3.25.

Therefore, components of GYE media do have the ability to induce this unusual form of cell death, but only at high concentrations. This more than likely explains the appearance of the perfectly round, darkened cells in the 50% *S. coelicolor* SN well, seen in Figure 3.24. From these observations, it was concluded that the effect induced by *S. sudanensis* SN is largely independent of the GYE media but may be enhanced by it at higher concentrations.

3.9.2 Induction of NLRP3-mediated pyroptosis in THP-1 cells using lipopolysaccharide and nigericin and comparison with the phenotype induced by S. sudanensis culture supernatant

Based on observations in the literature, it was speculated that the unusual form of cell death induced by *S. sudanensis* SN could be pyroptosis, mediated by the NLRP3 inflammasome (Davis et al., 2019). Assays were devised to check this. The first involved inducing NLRP3-mediated pyroptosis, using lipopolysaccharide (LPS) to prime the inflammasome and the potassium ionophore nigericin, shown in Figure 3.26(a), to activate it. This a widely used and well characterised method of NLRP3 activation (Mariathasan et al., 2006, Bauernfeind et al., 2009, Wang et al., 2013b, Coll et al., 2015, Coll et al., 2019). THP-1s were also challenged with *S. sudanensis* SN, so that the phenotypes induced in both conditions could be directly compared. The results of the assay are shown in Figure 3.26(b).

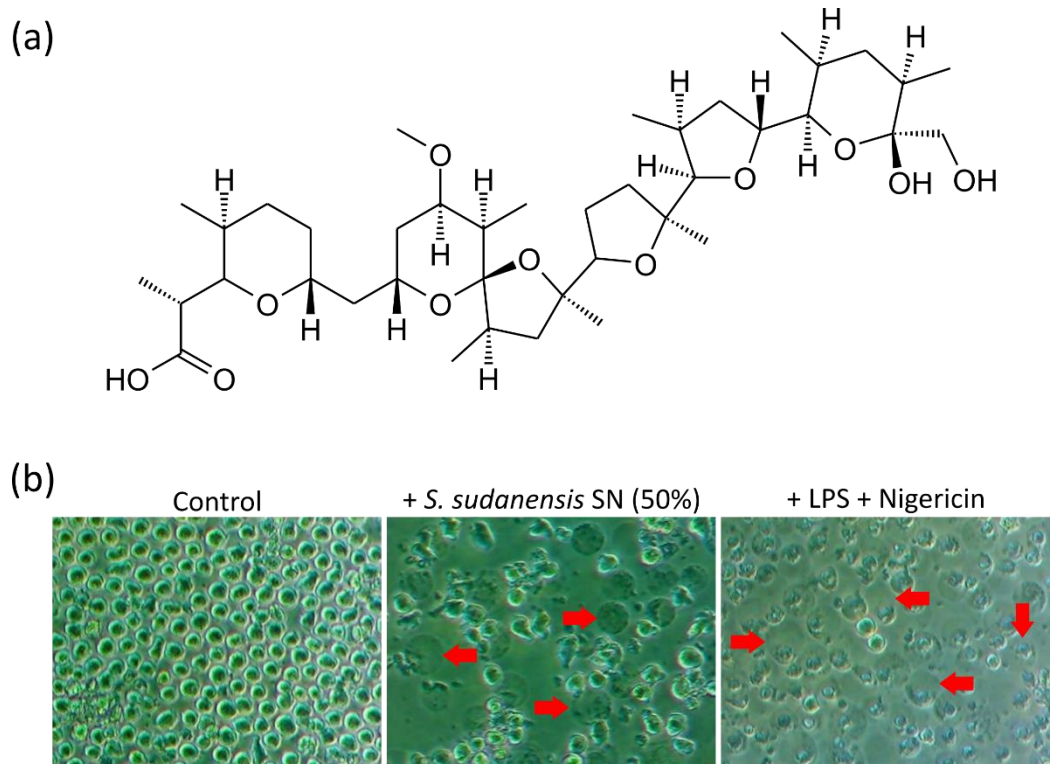


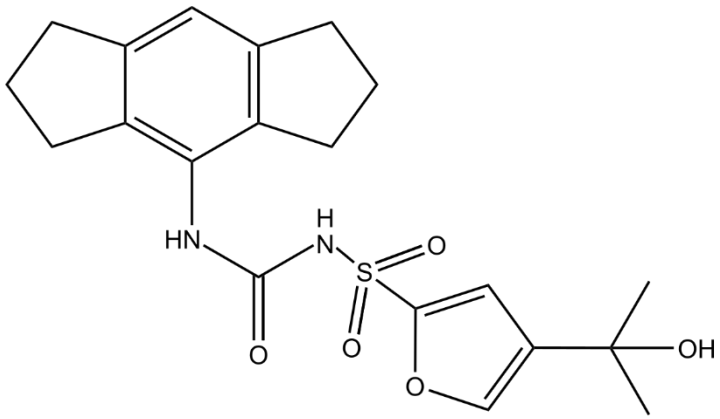
Figure 3.26 (a) Structure of nigericin, a potassium ionophore. **(b)** Images taken at 6 hrs of THP-1 cells challenged with sterile SN from *S. sudanensis* culture and a combination of LPS and nigericin. Images taken at 100X magnification. Red arrows indicate examples of the swollen, perfectly round and faded cell death phenotype of interest.

The images in Figure 3.26(b) show again the swollen form of cell death induced by *S. sudanensis* SN. Examination of the image of the '+ LPS + Nigericin' condition revealed a phenotype similar to that of SN addition and resemble those described by Davis et al. (2019), who identified this phenotype as pyroptosis. The visual similarity between the two conditions supports the idea that the bacterial SN also induces pyroptosis.

3.9.3 Effect of the NLRP3 inflammasome inhibitor MCC950 on the phenotype induced by S. sudanensis culture supernatant

The NLRP3 inhibitor MCC950 (Figure 3.27(a)) binds directly to the NLRP3 protein, preventing it from oligomerising and thus inhibiting inflammasome activity (Coll et al., 2015, Coll et al., 2019). If the *S. sudanensis* SN promoted inflammasome activation, it should be inhibited by MCC950. This was tested and a time course of images from this assay are shown in Figure 3.27(b).

(a)



(b)

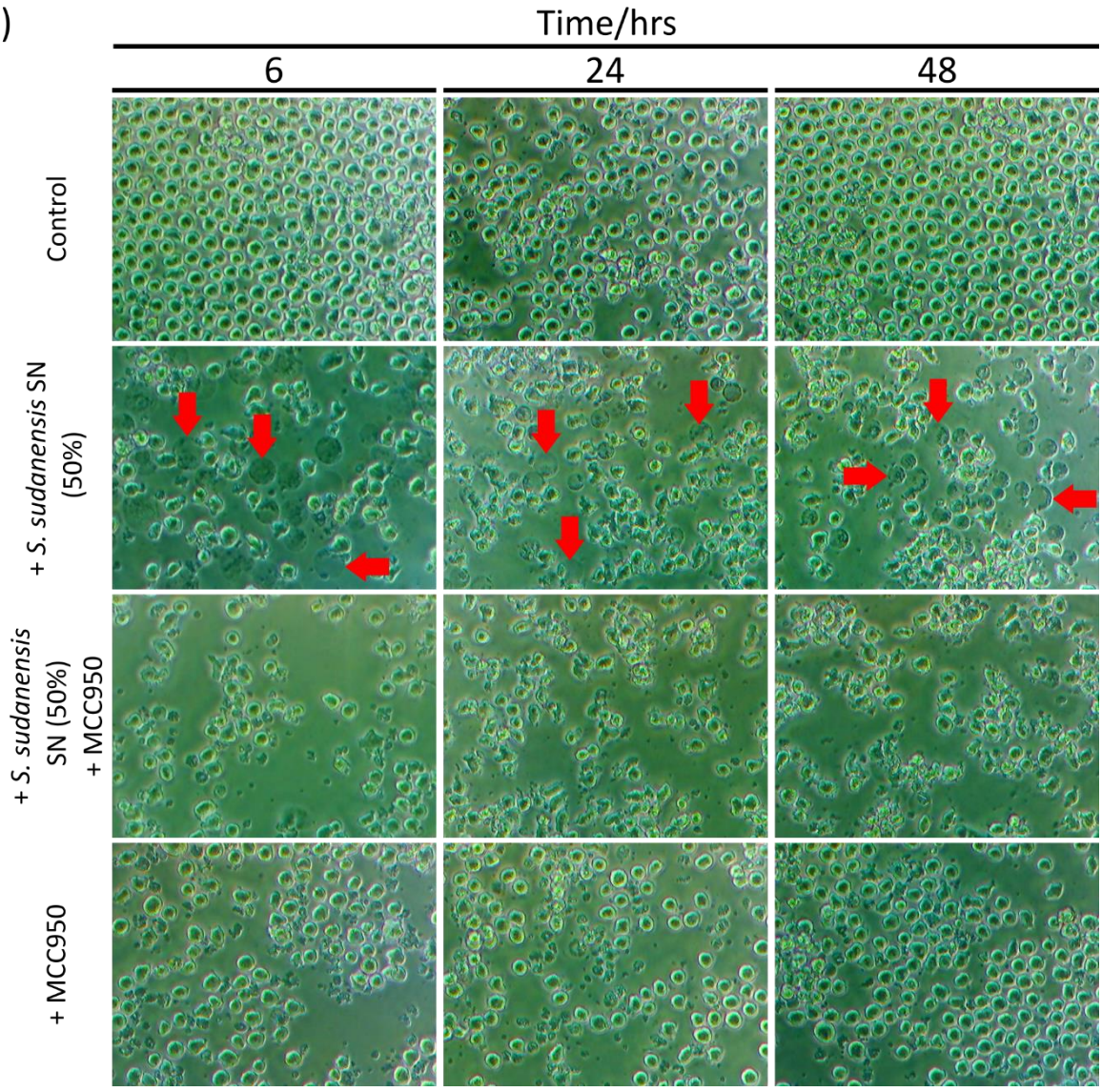


Figure 3.27 (a) Structure of MCC950, an inhibitor of the NLRP3 inflammasome. **(b)** Images of THP-1 cells challenged with sterile SN from *S. sudanensis* culture, with and without the addition of MCC950. Images taken at 100X magnification. Red arrows indicate examples of the swollen, perfectly round and faded cell death phenotype of interest.

MCC950 by itself was mildly cytotoxic, as a small number of shrivelled cells were visible in the well, as seen in the images for 6 and 24 hrs in Figure 3.27(b). But this was a weak effect and so was unlikely to have interfered with the assay result. The *S. sudanensis* SN induced the expected effect in the THP-1 cells, with multiple swollen dead cells visible at all three time points. When MCC950 was added, in addition to the bacterial SN, this cell death phenotype was completely blocked for the entire incubation period (compare the images from the two conditions in Figure 3.27(b)). As MCC950 is thought to be a specific inhibitor of NLRP3, the complete blockage of SN-induced swollen cell death indicates that this phenotype is most likely mediated by NLRP3. It therefore appears that *S. sudanensis* produces one or more compounds that have pyroptotic activity.

3.10 Summary and conclusions

A custom host-pathogen interaction assay was designed to fulfil two aims. Firstly, to broadly investigate how RAW-blue macrophages and the pathogen *S. sudanensis* respond to one another. Secondly, to detect the activities of any specialised metabolites potentially produced by the bacteria. The designed method involved seeding the macrophages into a tissue culture well and placing a porous Transwell insert above them. A layer of 0.5% agarose was then set on the base of the insert and the bacterial culture loaded on top of it. The agarose was found to block bacterial growth through the insert pores.

Separation using a permeable barrier was intended to allow the diffusion of secreted molecules and compounds between pathogen and host cultures, whilst blocking the movement of whole cells. So, if the pathogen did produce a metabolite that can affect immune cells, the metabolite would be able to diffuse through the well insert and into the macrophage culture. Its effects could then be observed. At the same time, blocking the passage of whole cells would prevent contact mediated immune reactions that would obscure the assay result.

Multiple control conditions were developed for the assay. A non-pathogenic control strain was selected: the model strain *S. coelicolor* M1152. Dead cultures of both bacteria were also used, to attempt to control for physical components of the cell breaking down and diffusing into the macrophage culture. Cultures were heat-killed by autoclaving.

When the assay was run for 48 hrs, it was observed that living *S. sudanensis* induced a unique morphology in the macrophages. The cells became extremely elongated and narrow. Dead *S. sudanensis* and both *S. coelicolor* conditions all induced a different morphology, with the cells adopting irregular, starry shapes.

RNA was harvested from macrophages at the 48hr time point of the assay and RNAseq was performed to enable differential expression analysis of genes. Living *S. sudanensis* was found to have induced differential expression in the greatest number of genes, with living *S. coelicolor* inducing the lowest number. Living *S. sudanensis* also induced the greatest amount of unique transcriptional changes.

Analysis and comparison of the datasets revealed a high number of pro-inflammatory markers to be upregulated in all four conditions. A majority of these markers showed the greatest upregulation in response to living *S. sudanensis*. Several unique transcriptional changes of interest were also observed in macrophages exposed to the living pathogen. Firstly, two transcription factors were found to be upregulated, indicating potential mediators of the unique transcriptional regulation that occurred in macrophages under this condition. Secondly, key genes involved in the anti-inflammatory immune response, including two pleiotropic cytokines, were upregulated in the M+S condition. Thirdly, five genes whose products catalyse steps of the mevalonate pathway were heavily downregulated. Finally, the NLRP3 inflammasome, responsible for initiating pyroptosis, was primed by both living and dead *S. sudanensis*, with the living cells inducing the stronger priming effect.

SNs were also harvested from the interaction assay at 48hrs. ELISAs were performed to quantify the release of three cytokines from macrophages during the assay. Levels of TNFα were found to be raised in all four challenge conditions, compared to the control condition M. The quantities of this cytokine were similar for all four conditions. Increases in TGFβ levels were also produced under M+C, M+DC and M+DS, while for M+S the level of this cytokine remained similar to that of M. No accurately measurable accumulation of IL10 was found in M, M+DC or M+DS. Only minimal release, just above the sensitivity limit of the assay, was measured for M+C. Living *S. sudanensis* on the other hand, induced very high levels of IL10 release from macrophages. A QUANTI-Blue assay was also performed on the SNs, to measure NF-κB activity. This revealed that while dead bacteria gave consistently high activity levels, the

activity measured during exposure to both living strains was highly variable from one replicate to another.

NLRP3 inflammasome activation by *S. sudanensis* was investigated further using the human monocytic cell line THP-1. When these cells were challenged with sterile SNs from *S. sudanensis* and *S. coelicolor* cultures, it was found that the pathogen SN induced a pyroptosis-like form of cell death. To determine whether this cell death was NLRP3-mediated pyroptosis, the death phenotype was directly compared to that triggered by nigericin, a known stimulator of NLRP3. A high visual similarity was observed between them. Furthermore, when the NLRP3 inhibitor MCC950 was added to THP-1 cultures in addition to *S. sudanensis* SN, the unusual death phenotype was blocked, confirming its likely assignment as NLRP3-mediated pyroptosis.

In conclusion, the indirect interaction assay demonstrated that *S. sudanensis* produces one or more molecules that stimulate a specific response in RAW-blue macrophages. RNAseq provided important insights into the nature of the macrophage response to this pathogen. It also showed that numerous aspects of this immune response were specific to the live pathogen and not induced by either dead pathogen cells or a non-pathogen. Finally, evidence was obtained that *S. sudanensis* produces one or more compounds that have pyroptotic activity via activation of the NLRP3 inflammasome.

Chapter 4. Characterisation of the response of the actinomycetoma pathogen *Streptomyces sudanensis* to macrophages

4.1 Introduction

As described in Chapter 3, the development of next generation sequencing technologies has increased the throughput, read lengths and depth of coverage of DNA sequencing, while also significantly enhancing the speed and affordability of the process (McCombie et al., 2019). The number of organisms whose genomes have been fully sequenced has resultantly increased significantly, with nearly 350,000 bacterial genomes now publicly available (Land et al., 2015, NCBI, n.d.-b). As described in Chapter 1, within the field of specialised metabolite research, this has led to the discovery that actinomycetes have the ability to produce ten times as many specialised metabolites as was previously estimated (Barka et al., 2016, Katz and Baltz, 2016). Bioinformatics-driven 'genome mining' has since emerged as a routine method for the discovery of novel specialised metabolites (Ward and Allenby, 2018). Utilising comparative genomics with known producers of specialised metabolites, the presence of biosynthetic gene clusters (BGCs) can be identified in newly sequenced organisms. The widely used antibiotics and Secondary Metabolite Analysis Shell (antiSMASH) is a typical example of a computational tool that deploys this analytical method (Blin et al., 2021). Predicted BGCs can then theoretically be targeted by experimental approaches (Challis, 2014). It was decided to apply this approach to *S. sudanensis*, to identify candidate BGCs that may produce virulence factors.

No genome assemblies for *S. sudanensis* were publicly available, meaning *de novo* sequencing of the organism was necessary. Actinobacterial genomes are extremely rich in guanine and cytosine (Barka et al., 2016). Additionally, they contain long, repetitive sequence regions, often within BGCs (Lee et al., 2020). These factors create a challenge when sequencing using technologies that produce relatively short reads hundreds of base pairs in length, such as Illumina sequencing. The repetitive regions exceed the lengths of the reads produced, so assembling a high quality, low contig number genome sequence is challenging.

Oxford Nanopore Technologies (ONT) and Pacific Biosciences (PacBio) sequencing technologies produce reads that range from thousands to hundreds of thousands of base pairs in length and therefore avoid this problem (van Dijk et al., 2018, McCombie et al., 2019). ONT sequencing was chosen for use for *S. sudanensis*, due to its lower cost per genome compared

to PacBio sequencing. ONT does, however, have a lower base calling accuracy than Illumina, with a 15% error rate compared to a range of 0.1 to 2% for Illumina's instruments (Illumina, 2017, Pfeiffer et al., 2018, van Dijk et al., 2018). A common workaround for this is to sequence using both technologies. The longer ONT reads can be used to create an assembly; the shorter, but more accurate, Illumina reads are then used to correct errors (Gomez-Escribano et al., 2015, Wick et al., 2017, Lee et al., 2020). It was therefore decided to sequence the *S. sudanensis* genome using Illumina technology as well.

Sequencing of *S. sudanensis* would also enable RNA sequencing (RNAseq) to be performed on the bacteria at the end of the indirect interaction assay incubation period, as was done with the macrophages. As described in Chapter 3, the assay was developed to interrogate how *S. sudanensis* interacts with mammalian macrophages, without being in direct physical contact with them. Monitoring changes in bacterial gene transcription following exposure to macrophages could provide insights into how *S. sudanensis* adapts to the presence of immune cells. It could also highlight potential virulence factors, by generating a list of genes that are transcriptionally upregulated only in the presence of macrophages. Furthermore, expression of predicted BGCs during the assay could be examined.

4.2 Genome sequencing of *S. sudanensis* and identification of biosynthetic gene clusters

4.2.1 De novo genome sequencing of S. sudanensis

S. sudanensis was cultured and high molecular weight DNA extracted as described in Chapter 2. Illumina sequencing of *S. sudanensis* was carried out by the NU-OMICS facility at Northumbria University, while ONT sequencing was run by Demuris Ltd. Assembly and annotation of the genome was then performed by Dr Nick Allenby of Demuris Ltd. The program Canu was used for assembly, while Prokka was used for annotation (Seemann, 2014, Koren et al., 2017). This resulted in a single contig assembly of 5,259,635 bp in length. As Table 4.1 shows, this a relatively small size compared to other *Streptomyces* spp. The genome also had a GC content of 73.9%, 4622 protein coding sequences (CDSs), 18 rRNAs in 6 operons and 74 tRNAs (summarised in Figure 4.1 along with other salient details).

Organism	Genome size (Mb)
S. coelicolor	8.6
S. albus	7.8
S. venezuelae	8.0
S. avermitilis	9.9
S. griseus	8.1
S. sudanensis	5.3

Table 4.1 Comparison of genome sizes of other, well studied *Streptomyces spp.* with that of *S. sudanensis*. Median genome sizes obtained by manually searching database curated by NCBI (n.d.-a).

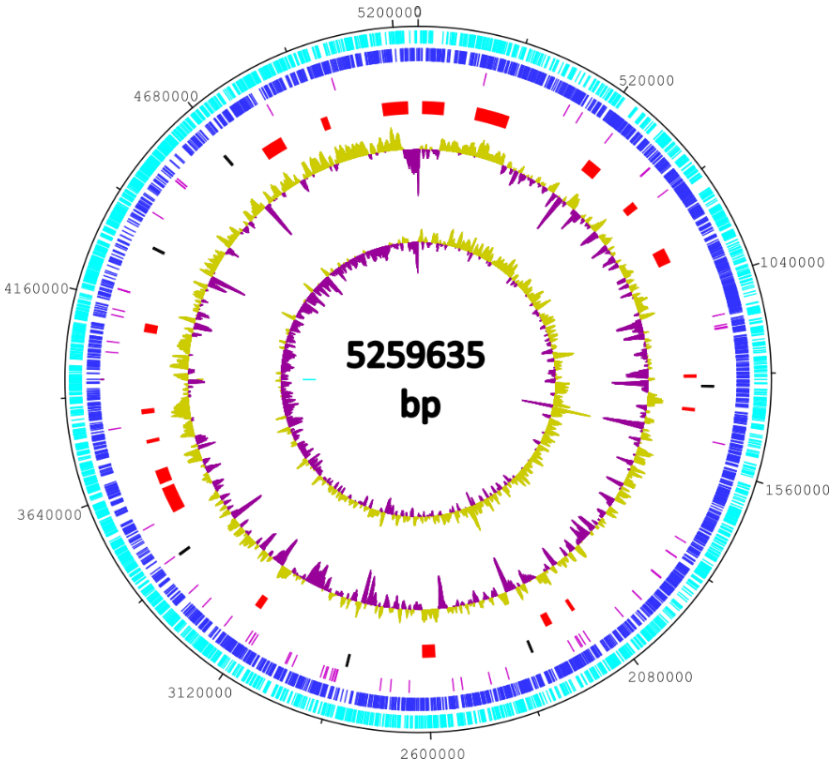


Figure 4.1 Representation of the genome of *S. sudanensis*. From the outer edge: ring 1 (cyan) – protein coding sequences on forward strand, ring 2 (blue) – protein coding sequences on reverse strand, ring 3 (magenta) – tRNAs, ring 4 (black) – rRNAs, ring 5 (red) – biosynthetic gene clusters predicted by antiSMASH, ring 6 – GC content (calculated variation from the average GC content of the whole sequence, for a window of 10,000 bp, above average in yellow, below average in purple), ring 7 – GC skew ((G-C)/(G+C) shown as variation from the average GC skew of the whole sequence, for a window of 10,000 bp, above average in yellow, below average in purple).

4.2.2 Prediction of BGCs within the S. sudanensis genome using antiSMASH

The genome was submitted to antiSMASH (version 4.2.0) for analysis (Blin et al., 2017). This resulted in identification of a total of nineteen BGCs (Table 4.2). Their locations within the genome are marked in red within Figure 4.1. Of these, four showed high homology (≥75%) with known BGCs.

Cluster	BGC type	Start in genome	End in genome	Highest homology BGC (% similarity)	MIBIG BGC-ID
1	NRPS	12195	80339	Griseoviridin/viridogrisein (11%)	BGC0000459_c1
2	T3PKS-Terpene	178244	280710	Isorenieratene (100%)	BGC0000664_c1
3	NRPS	553145	599683	Arginomycin (20%)	BGC0000883_c1
4	Lantipeptide	736313	760835	-	-
5	NRPS	903935	951164	Sch47554/Sch47555 (3%)	BGC0000268_c1
6	Bacteriocin	1299434	1309892	-	-
7	Lantipeptide	1401098	1411496	Ectoine (100%)	BGC0000853_c1
8	Siderophore	2128154	2139947	Desferrioxamine B (100%)	BGC0000940_c1
9	Lantipeptide	2208430	2231033	Ochronotic pigment (75%)	BGC0000918_c1
10	Other	2576919	2617593	-	-
11	Terpene	3132141	3154162	Hopene (15%)	BGC0000663_c1
12	NRPS	3527243	3607797	Desotamide (22%)	BGC0001196_c1
13	Other	3620860	3664732	Auricin deoxysugar moieties (24%)	BGC0000727_c1
14	Bacteriocin	3749500	3760960	-	-
15	Siderophore	3839181	3854675	Macrotetrolide (33%)	BGC0000244_c1
16	Terpene	4088361	4114890	Hopene (69%)	BGC0000663_c1
17	Lantipeptide	4760641	4828236	Daptomycin (10%)	BGC0000336_c1
18	Terpene	4959416	4980525	Hygrocin (6%)	BGC0000075_c1
19	Thiopeptide-T3PKS-Lantipeptide	5149681	5229010	Lactazole (66%)	BGC0000606_c1

Table 4.2 BGCs predicted to occur within the *S. sudanensis* genome by antiSMASH. BGCs are coloured based on type.

Cluster 2 has a 100% match with isorenieratene. This a carotenoid pigment produced by photosynthetic bacteria to protect against photooxidative stress induced by exposure to excess light (Krugel et al., 1999). Closer inspection showed cluster 2 to contain 103 genes. Only seven genes are reported to make up the isorenieratene BGC and these could be found in the centre of cluster 2 (Krugel et al., 1999). The inclusion of the other 95 genes appears to be an artefact of the antiSMASH algorithm. Closer analysis showed the methylesterase of the isorenieratene BGC (*crtV*) to be truncated, giving a predicted protein of 77 amino acids (aa), compared with the expected 338 aa. However, analysis with the alignment tool BLASTP showed that the 77 aa protein aligned with the N-terminus of CrtV, indicating a level functional similarity. The *crtE* gene encoding geranylgeranyl pyrophosphate synthase is absent, replaced with a predicted farnesyl diphosphate synthase. These differences imply that a different carotenoid could be synthesised by this cluster.

Cluster 7 was predicted to produce ectoine, with 100% similarity to the known BGC. Ectoine and its derivatives are osmolytes shown to protect cells under high osmotic pressure (Prabhu et al., 2004). All three genes necessary for ectoine biosynthesis were found in cluster 7. It also features a fourth gene, annotated as an ectoine dioxygenase. However, a BLASTP search showed its product to be homologous to ThpD, an ectoine hydroxylase that is common amongst *Streptomyces* spp. (Prabhu et al., 2004). The primary product of cluster 7 is therefore likely to be hydroxyectoine.

Cluster 8 was predicted to encode for desferrioxamine B, with 100% similarity. This compound is a siderophore produced by a wide variety of *Streptomyces* spp. and all four biosynthetic genes were confirmed to be present within cluster 8 (Barona-Gómez et al., 2004).

Cluster 9 was shown to have 75% similarity to ochronotic pigment. This pigment is also known as pyomelanin and is yellow/brown in colour (Moran, 2005, Singh et al., 2018). It is produced from the oxidation and polymerisation of excess homogentisic acid (HGA), an intermediate formed during the catabolism of tyrosine (Yabuuchi and Ohyama, 1972, Arias-Barrau et al., 2004, Bolognese et al., 2019). Only two genes are described as necessary for biosynthesis of this pigment, an AsnC-family transcriptional regulator and *hpd*, a 4-hydroxyphenylpyruvate dioxygenase (Ōmura et al., 2001). Hpd catalyses the second reaction step in the breakdown of tyrosine and is directly responsible for HGA production (Arias-Barrau et al., 2004). Both genes were found in cluster 9, so it seems highly likely that ochronotic pigment is its product. As with cluster 2, the antiSMASH algorithm grouped other genes into this cluster which do not contribute to biosynthesis of the predicted product. This accounts for the lower similarity score, despite the necessary biosynthetic genes being present.

Of the remaining fifteen predicted BGCs in the genome, three had low homology (≥25%, <75%) to known BGCs, eight had very low homology (<25%) and four had no homology. As these comprise the majority of predicted BGCs, it seems likely that *S. sudanensis* could be a source of novel specialised metabolites.

4.3 Overview of *S. sudanensis* RNAseq dataset

RNA was harvested from *S. sudanensis* incubated within the assay system after 48 hrs. Two conditions were used for the bacteria in the assay: *S. sudanensis* in isolation (S) and *S. sudanensis* + macrophages (S+M). This was done for three biological repeats, with samples numbered accordingly. RNAseq libraries were prepared as described in Chapter 2. Illumina sequencing of the libraries was performed by Edinburgh Genomics. The quality of the resulting sequencing data was then assessed using FastQC, which showed no significant problems.

Reads were then mapped to *S. sudanensis* CDS regions using Salmon (Patro et al., 2017). The number of reads mapped to the transcriptome for each sample are shown in Table 4.3. The mapping rates for each were extremely low, ranging from 0.6% to 4.77%. It was expected that at least a majority of reads should have mapped.

Sample	Reads	Mapped reads	Mapped (%)
S_1	35,372,840	212,338	0.6
S_2	21,011,915	276,573	1.32
S_3	51,665,232	2,462,069	4.77
S+M_1	29,735,568	429,937	1.45
S+M_2	20,338,432	606,465	2.98
S+M_3	28,493,614	915,977	3.21

Table 4.3 Number of *S. sudanensis* RNAseq reads obtained for each sample taken from the indirect interaction assay and their mapping rates to *S. sudanensis* CDS regions.

The nature of the unmapped reads was investigated by three routes. Mapping using Salmon was repeated, but with rRNA sequences included in the mapping index, in addition to CDSs. Mapping of the reads to the whole genome of *S. sudanensis* was then also performed for each sample, using the alignment software Bowtie2 (Langmead and Salzberg, 2012). The mapping rates for these two different references are plotted in Figure 4.2, alongside the original mapping rates to CDSs.

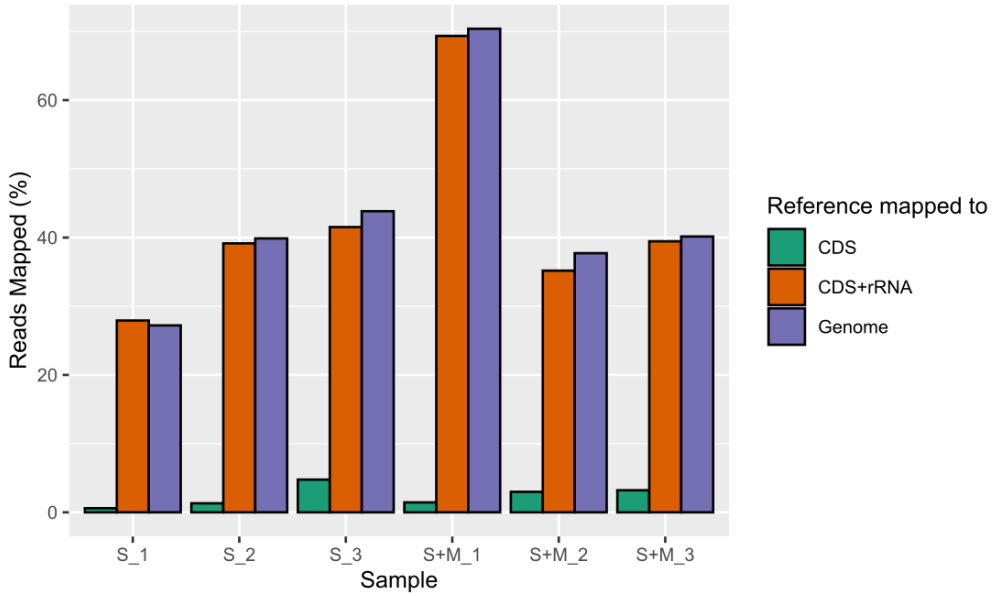


Figure 4.2 Mapping rates of *S. sudanensis* RNAseq reads to *S. sudanensis*: CDSs, CDSs + rRNA, whole genome. Sample key: S: *S. sudanensis* in isolation. S+M: *S. sudanensis* exposed to macrophages.

The addition of rRNA sequences led to a dramatic increase in mapping rate for all samples. This indicates that attempts to remove rRNA prior to cDNA library preparation were not successful. High proportions of the sequenced samples were therefore rRNA. Even with mapping to rRNA however, the highest mapping rate achieved was only 67.89%, for S+M_1. For the other five samples, the mapping rate was still below 50%, the lowest being 27.32% for S_1. Mapping reads to the whole genome yielded only marginal increases for five samples and a marginal decrease for S_1, compared to that achieved with CDSs and rRNA. The decrease seen for S_1 is likely due to differences in the mapping algorithms of Salmon and Bowtie2. Small changes in mapping rate when aligning to the genome indicates that only negligible amounts of unwanted genomic DNA were present in the samples.

Finally, the reads were also mapped to the mouse transcriptome, to check if macrophage RNA had contaminated the bacterial samples. The mapping percentages for all six samples were below 0.2% and so were lower than those obtained for the *S. sudanensis* CDSs. Therefore, contamination with macrophage mRNA was not an issue. Following these repeated mappings, the origins of large proportions of reads within each sample therefore remain unaccounted for.

During set up of the interaction assay trials, the *S. sudanensis* cultures used were always examined by microscopy first. The cultures appeared as expected and showed no sign of contamination with other organisms. Additionally, for each trial 100 µL of the living and autoclaved bacterial cultures were streaked out on nutrient agar plates and incubated for one week to check for contaminants. For each trial no contaminants were observed growing on any of the plates. It therefore seems unlikely that a contaminant was the source of the unmapped reads. A possibility is that the *S. sudanensis* genome assembly itself may not be fully accurate. Reads would not be able to map to inaccurately assembled regions of the genome, as their sequences would be absent.

Despite the low mapping rates, it was decided to continue with differential expression analysis of bacterial genes. Initial processing of the data was performed as previously described in Chapter 3 for the macrophage data. The Salmon output was transferred to DESeq2 using Tximport and the data were normalised (Love et al., 2014, Soneson et al., 2015). As the normalisation corrects the read counts for each gene based on the size of the read library, the unusually low numbers of mapped reads in all six samples were accounted for by DESeq2. Following normalisation, the dispersion estimates for each gene were checked and the resulting plot is shown in Figure 4.3 (Love et al., 2014). The trend of dispersion decreasing as the mean counts increase is shown by the plot.

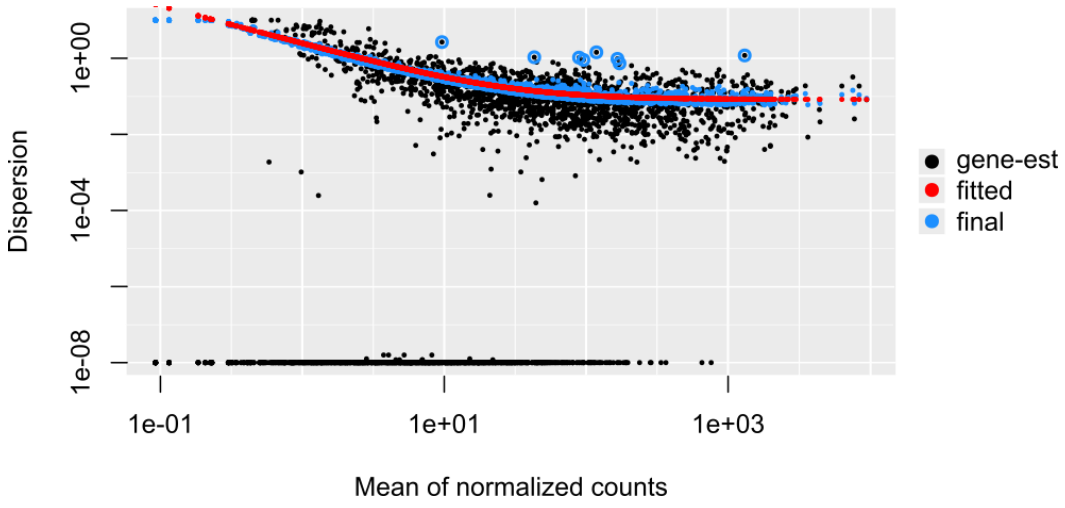


Figure 4.3 Plot produced by DESeq2 showing the relationship between dispersion and mean counts for each gene in the *S. sudanensis* RNAseq dataset.

As in Chapter 3, principal component analysis (PCA) was performed to check for outliers among the biological repeats. The resulting plot shown in Figure 4.4.

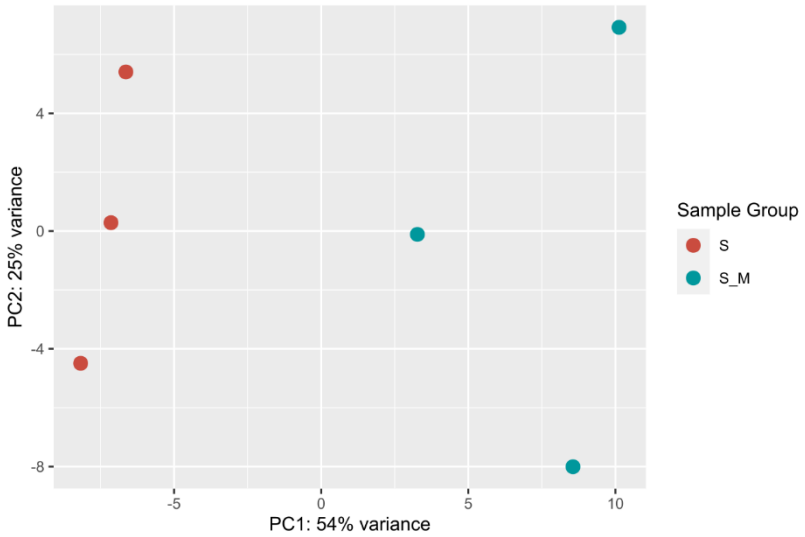


Figure 4.4 Plot of PCA of the *S. sudanensis* RNAseq dataset, featuring all biological repeats.

The PCA plot shows that 54% of the variance between samples is along principal component (PC) 1. The two sample groups are clustered at either end of PC1. Therefore, the majority of variation between samples was attributable to the experimental conditions used during the experimental assay i.e., the presence and absence of macrophages. The PCA plot does show intra-condition variation, with samples in both conditions distributed along PC2. However, this only accounts for 25% of the variance between samples and could be a sign of natural biological heterogeneity. As the samples cluster reasonably well along PC1, no clear outliers were identified for removal.

As no outliers were seen in the PCA analysis, it was decided to carry out an additional check using hierarchical clustering. This process groups the samples together based on the similarities of their read counts for all genes, with the output shown as a dendrogram. The hclust function in R was used for this analysis and the resulting dendrogram is shown in Figure 4.5.

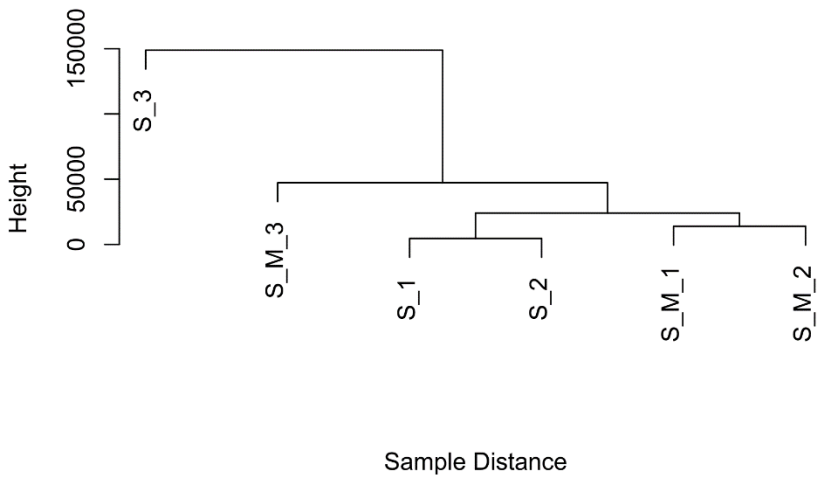


Figure 4.5 Dendrogram showing hierarchical clustering of all samples in the *S. sudanensis* RNAseq dataset.

The dendrogram shows that neither S_3 nor S_M_3 have clustered with the rest of their condition groups. S_3 is the furthest from the other samples, with a height of 150,000 and is therefore the most dissimilar. S_M_3 is the next furthest, but its height is far lower, being roughly 50,000. S_3 is a clear outlier and was removed from the RNAseq dataset. S_M_3 was kept in, being significantly less dissimilar than S_3 and allowing three biological repeats to be kept for at least one of the conditions. Following the removal of S_3, the DESeq2 normalisation and statistical analyses described above were repeated with the remaining five samples.

With the final dataset processed, differentially expressed genes were identified by selecting for those with a fold change in expression greater than +/-2 and an adjusted p-value of less than 0.05. These were the same thresholds used for the macrophage differential expression analysis in Chapter 3. The logic of choosing these thresholds was the same: to look for the largest induced transcriptional changes. The unfiltered dataset is visualised as a volcano plot in Figure 4.6.

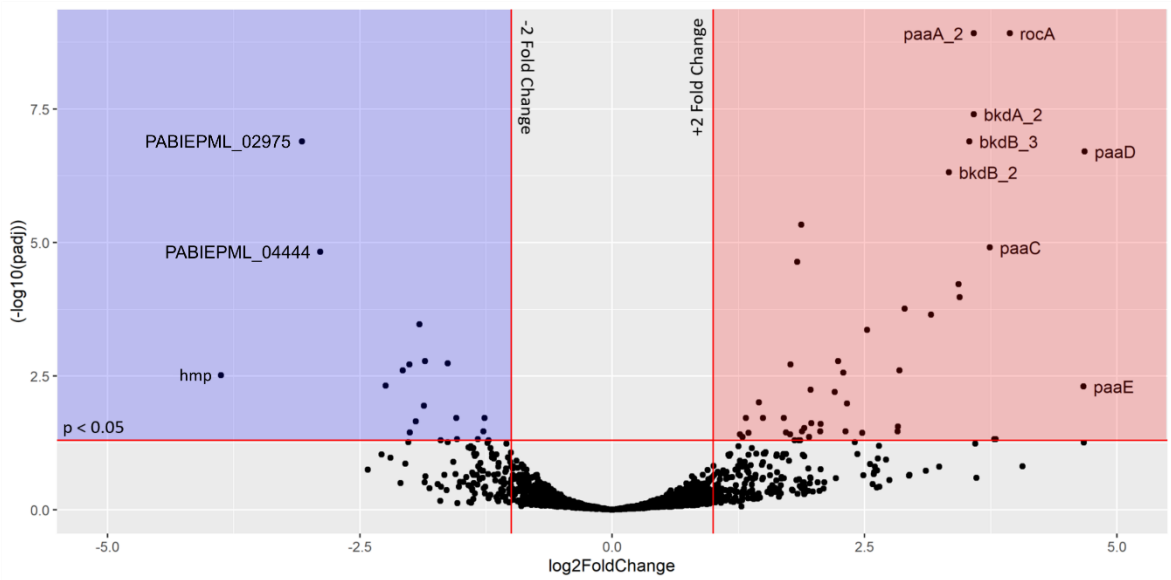


Figure 4.6 Volcano plot of transformed fold changes versus transformed adjusted p-values for each gene in the *S. sudanensis* RNAseq data set. Red lines mark the thresholds applied for a gene to be considered differentially expressed. A fold change of +/- 2 is equivalent to a log2(FoldChange) of 1. A p-value of 0.05 is equivalent to -log10(padj) of 1.3. Genes within the red box were upregulated in the presence of macrophages, while those in the blue box were downregulated.

The volcano plot indicates that a large number of *S. sudanensis* genes underwent fold changes more extreme than +/-2 during exposure to the macrophages. However, only a small proportion had an associated adjusted p-value that takes them above the threshold set for statistical significance. This gave a total of 60 differentially expressed CDSs, from a possible 4622. Of these 60 genes, 43 were upregulated and 17 downregulated. The upregulated genes are shown in Figure 4.7, as a heatmap of read counts.

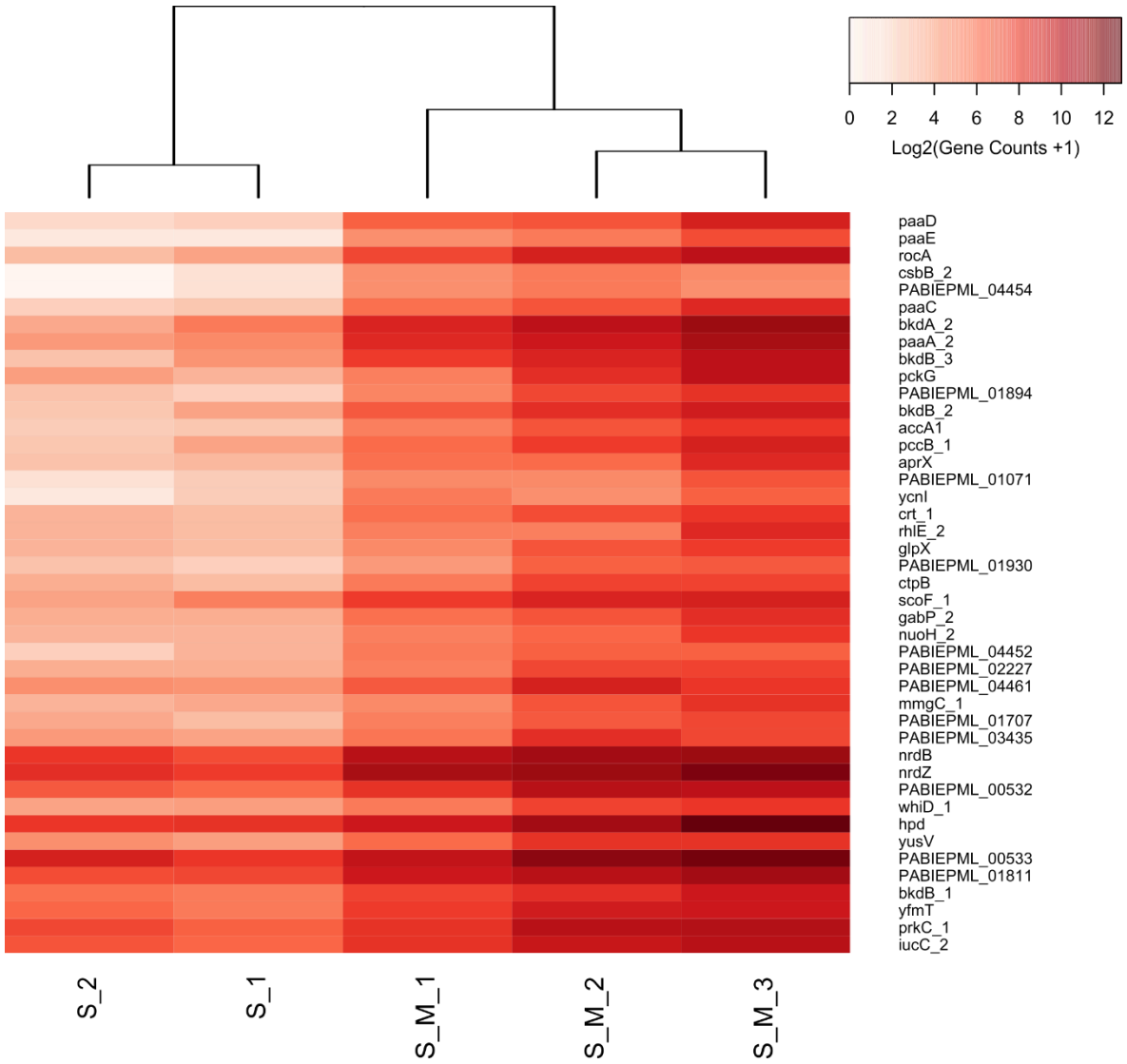


Figure 4.7 Heat map of 43 *S. sudanensis* genes upregulated in the presence of macrophages. Genes are in descending order of magnitude of upregulation. Normalised read counts for each gene were used to make the plot. The darker the shading, the more counts a gene had. Samples are also clustered by similarity, shown by the dendrogram above the heat map.

The dendrogram shows the S and S+M samples have clustered together into their respective groups. This demonstrates that, for this set of genes, individual biological repeats were most similar to others of the same condition. This also vindicates the decision made during hclust analysis to not eliminate sample S_M_3 from the dataset.

The downregulated genes are displayed in Figure 4.8, as a heat map. The dendrogram illustrates that the samples have not clustered together into their respective condition groups. For many of the genes, the heat map shows less of a contrast in colour shade between the S and S+M samples than is seen in the heat map of upregulated genes in Figure 4.7. This indicates that the magnitudes of fold changes in expression for many downregulated genes

are not as great as those for upregulated genes. This is visualised in Figure 4.6, where there are eighteen genes with a log2FoldChange greater than +2.5, while three have a log2FoldChange greater than -2.5. The mean fold changes for the down- and upregulated genes were also calculated, as absolute values. For the downregulated genes the mean was 4.63, while for the upregulated genes it was 7.25, which is 63.8% greater. So, downregulated genes on average underwent a lower magnitude of change in expression than the upregulated genes. This accounts for the S and S+M conditions appearing more similar to one another when looking only at downregulated genes and may explain the disordered clustering in Figure 4.8.

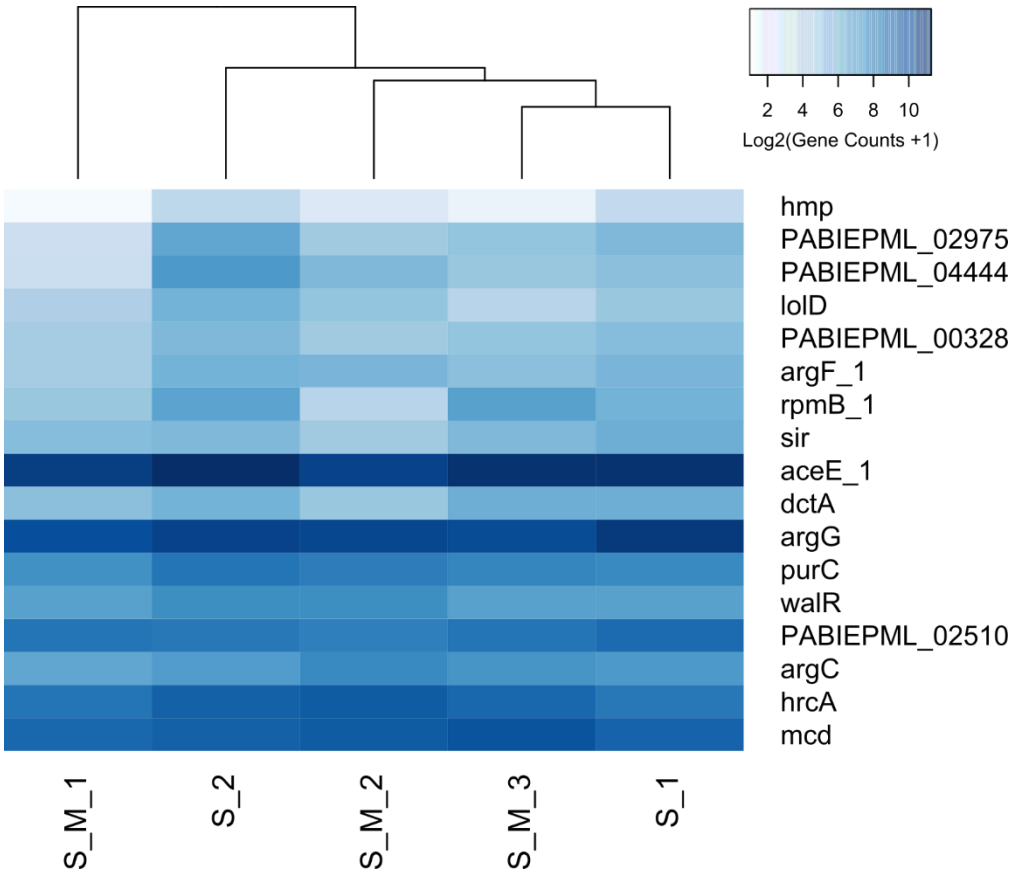


Figure 4.8 Heat map of the 17 *S. sudanensis* genes downregulated in the presence of macrophages. Genes are in descending order of magnitude of downregulation. Normalised read counts for each gene were used to make the plot. The darker the shading, the more counts a gene had. Samples are also clustered by similarity, shown by the dendrogram above the heat map.

4.4 Expression of *S. sudanensis* BGCs during the indirect interaction assay

4.4.1 BGCs expressed in the presence and absence of macrophages

A BGC that is transcriptionally silent in the control condition but expressed during exposure to macrophages would be a clear candidate for producing a virulence factor. It was therefore considered useful to directly visualise which BGCs were transcriptionally active or silent during the assay, under either condition. Means of the normalised read counts for each gene were calculated for both conditions. The mean counts for genes identified as belonging to any of the nineteen BGCs predicted by antiSMASH were extracted and plotted. The plot for the S control condition is shown in Figure 4.9(a).

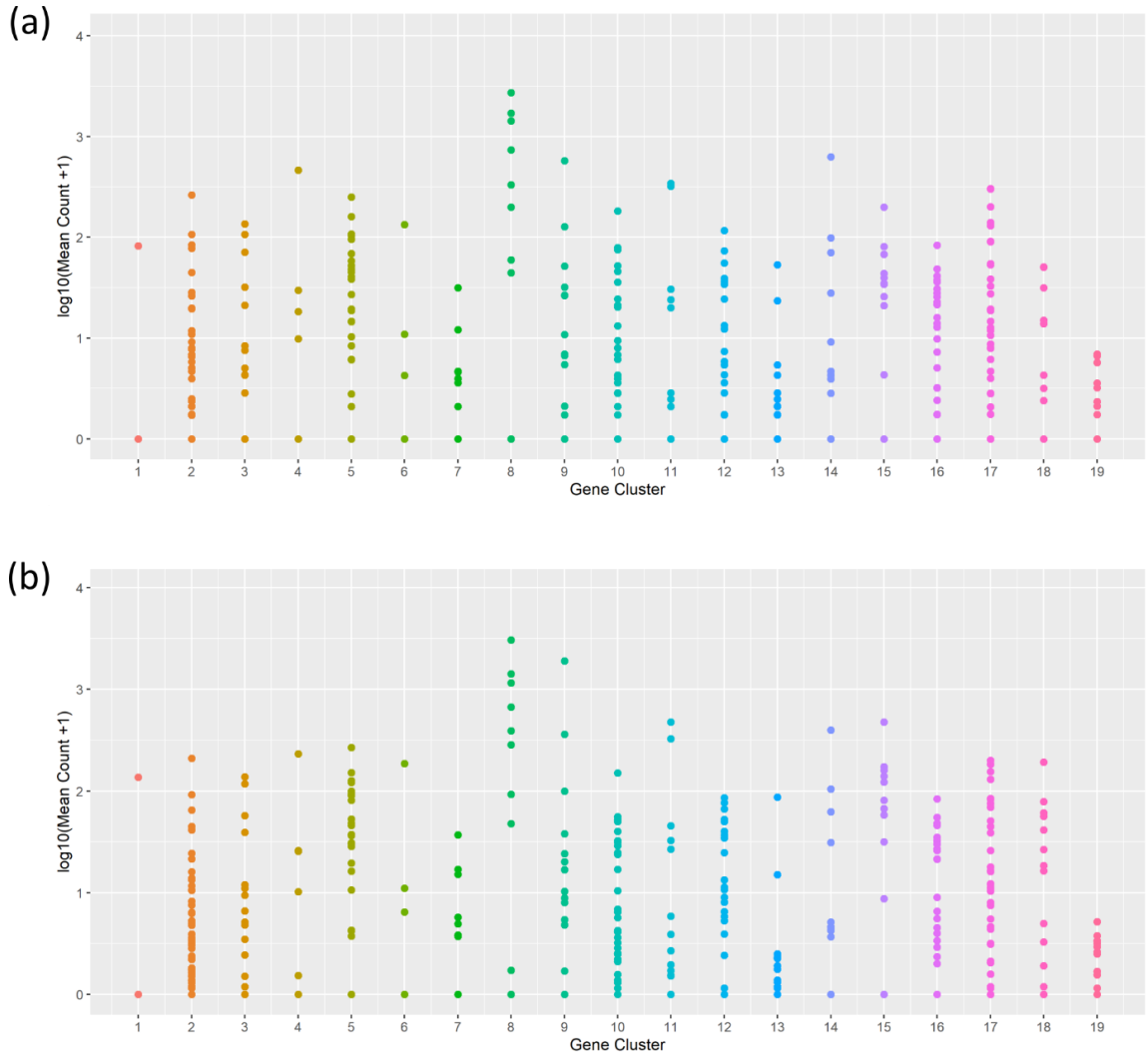


Figure 4.9 Plots of the transformed means of normalised read counts for every gene present in the nineteen *S. sudanensis* BGCs predicted by antiSMASH, grouped by gene cluster. Data from RNAseq performed on samples harvested at 48 hrs. Each point represents 1 gene. **(a)** Data for the S condition – *S. sudanensis* in isolation. **(b)** Data for the S+M condition – *S. sudanensis* + macrophages.

The plot shows that in the absence of macrophages, all nineteen BGCs had at least one gene expressed. While the degree of transcriptional activity varies between BGCs, none were transcriptionally silent. The equivalent plot for the S+M condition (Figure 4.9(b)) is highly similar to that for the S condition. All nineteen BGCs contained at least one expressed gene. Count levels for genes in sixteen of the BGCs appear to be the same or only marginally different. Clusters 9, 15 and 18 were the only ones to show strong changes: all three had at least one gene that was upregulated in the presence of macrophages. With all BGCs showing some degree of expression under both conditions, none were eliminated from contention at this stage.

4.4.2 Differentially expressed BGCs

The list of 60 differentially expressed genes was inspected for members of any of the nineteen BGCs. In total, seven genes from three BGCs (listed in Table 4.4) were found to be differentially expressed. Expression of all seven genes was upregulated in the presence of macrophages. This supported the observations made above from the mean counts plots in Figure 4.9.

Cluster	Gene	Fold Change	Adjusted p-value	Function
9	*hpd*	3.3	0.036	4-hydroxyphenylpyruvate dioxygenase
15	*iucC_2*	2.4	0.039	Aerobactin synthase
	PABIEPML_03435	3.7	0.034	Hypothetical protein
	yusV	3.2	0.019	Putative siderophore transport system ATP-binding protein
18	*PABIEPML_04452*	4.2	0.034	Hypothetical protein
	PABIEPML_04454	13.8	0.048	Hypothetical protein
	PABIEPML_04461	3.9	0.006	Universal stress protein

Table 4.4 Differentially expressed genes from predicted *S. sudanensis* BGCs.

In cluster 9, only the gene *hpd* was upregulated. As described in section 4.1.2, the protein product of this gene is responsible for the single step biosynthesis of HGA from 4-hydroxyphenylpyruvate (Arias-Barrau et al., 2004). When present in excess, HGA will oxidise and polymerise to form ochronotic pigment, which is the predicted product of cluster 9 (Bolognese et al., 2019). The core biosynthetic gene of this cluster had therefore been upregulated, so it is likely that HGA production was increased in the presence of macrophages.

Cluster 15 had three upregulated genes and is predicted to produce a siderophore. The first of the three is *iucC_2*, homologous to aerobactin synthase. Aerobactin is itself a siderophore

and known virulence factor and the synthase enzyme is part of its core biosynthetic machinery (de Lorenzo et al., 1986, de Lorenzo and Neilands, 1986). The next gene is *PABIEPML_03435*, annotated as a hypothetical protein. A BLAST search yielded only matches to hypothetical proteins from other *Streptomyces* spp. and so the potential function of this gene is unknown. Finally, *yusV* was also upregulated and is homologous to genes encoding a siderophore transport system (Ollinger et al., 2006). The expression of *iucC_2* and *yusV* provides evidence that cluster 15 does indeed produce a siderophore and their upregulation suggests biosynthesis of its siderophore product was increased in the presence of macrophages.

Cluster 18 is predicted to produce a terpene and had three genes upregulated. *PABIEPML_04461* is annotated as encoding a universal stress protein. These genes are expressed in response to environmental stressors, helping the bacteria to adapt and survive (Siegele, 2005, Tkaczuk et al., 2013). Their protein products have numerous effects on gene regulation and cell metabolism. *PABIEPML_04452* and *PABIEPML_04454* are annotated as hypothetical proteins. It was noted that with a fold change in expression of +13.8, *PABIEPML_04454* is the fifth most highly upregulated gene in the dataset. A BLASTP search showed their proteins to be homologous to a sensor histidine kinase and a helix-turn-helix (HTH) protein of unknown function respectively. HTH domains are DNA binding motifs and so it is probable that *PABIEPML_04454* encodes a transcriptional regulator (Aravind et al., 2005).

These two genes are therefore likely to collectively encode a two component signalling system, wherein the histidine kinase detects an environmental stimulus and then phosphorylates the transcriptional regulator (Groisman, 2016, Zschiedrich et al., 2016). This then becomes active and alters bacterial gene expression. Such systems are the dominant form of altering transcription in response to changing environmental conditions within bacteria. Evidence has shown that these systems do regulate specialised metabolite production (Martin, 2004). However, they themselves do not perform steps of biosynthetic processes. Thus, the protein products of *PABIEPML_04452* and *PABIEPML_04454* would not participate in the production of the cluster 18 product.

The presence of this potential two component system within the non-pathogen *S. coelicolor* was checked using BLASTP. For PABIEPML_04452, the highest percentage identity match within *S. coelicolor* was 42%, with homology confined to the histidine kinase ATPase domain. For PABIEPML_04454, the highest identity match was 49%, with the homology focused in the

HTH domain. Both hypothetical proteins therefore showed low homology to proteins within *S. coelicolor*, indicating them to be absent within the non-pathogen.

None of the three genes upregulated in cluster 18 appear to play a direct role in terpene biosynthesis. As the core terpene biosynthetic genes within the cluster were not upregulated, it seems unlikely that production of the product of this cluster was increased in response to the presence of macrophages.

4.5 Identification of other genes of interest in the *S. sudanensis* differential expression data

4.5.1 A potential BGC composed of bkd and paa genes

A manual search of the list of upregulated genes for others of potential interest revealed that eight of the twelve most highly upregulated were clustered together in the *S. sudanensis* genome. These eight genes and their levels of upregulation are shown in a heat map in Figure 4.10.

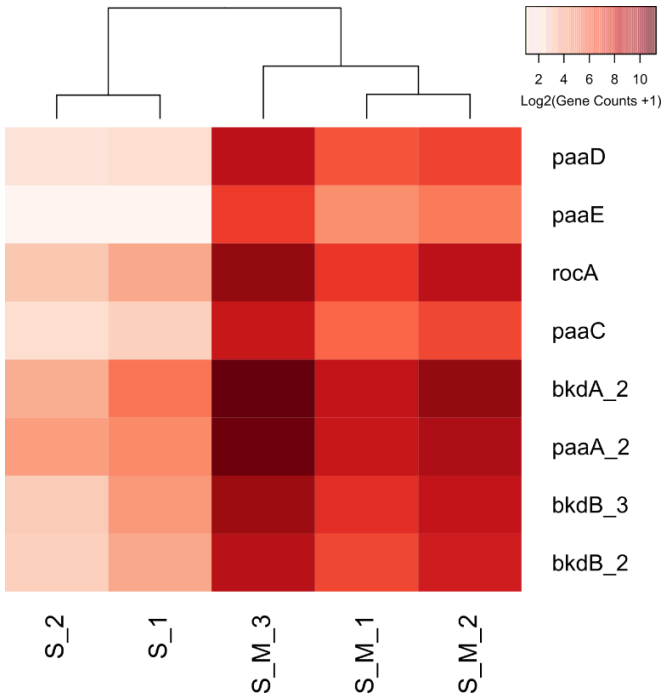


Figure 4.10 Heat map of the 8 *S. sudanensis* genes upregulated in the presence of macrophages and comprising part of a potential BGC. Genes are in descending order of magnitude of upregulation. Normalised read counts for each gene were used to make the plot. The darker the shading, the more counts a gene had. Samples are also clustered by similarity, shown by the dendrogram above the heat map.

The positions of these genes were also labelled in the volcano plot of the entire RNAseq dataset, shown in Figure 4.6. This shows that seven of the eight lowest adjusted p-values for the upregulated genes were assigned to members of this cluster. Therefore, the very strong upregulation of these genes was also associated with some of the greatest statistical significances, indicating high consistency across the biological repeats.

The cluster itself is shown in Figure 4.11. It is 14.65 Kb in length and consists of fifteen genes, distributed across both DNA strands. Of these genes, ten were confirmed to be directly involved in the biosynthesis of molecules. It therefore seems likely that this cluster is a BGC that was not recognised by antiSMASH.

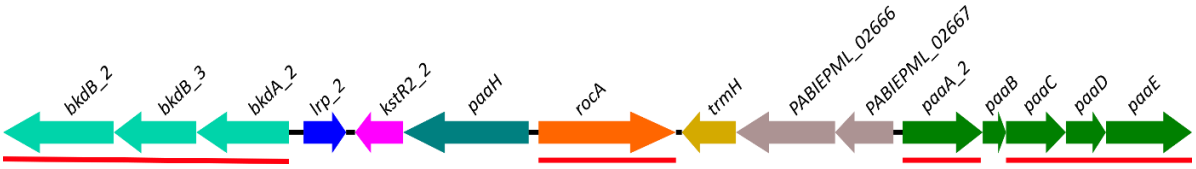


Figure 4.11 Diagram of the strongest upregulated BGC of the *S. sudanensis* genome. The cluster runs from position 2,956,832 in the genome, to 2,971,482. Red lines indicate genes that were upregulated in the presence of macrophages.

The genes *bkdA* and *bkdB* encode the E1α dehydrogenase and E1β decarboxylase subunits of the branched-chain α-keto acid dehydrogenase (BCDH) complex (Denoya et al., 1995). These genes are usually found as part of the *bkd* operon, along with two further genes, which encode the remaining two subunits of the BCDH complex. A transcriptional regulator, *bkdR*, is also commonly found in the operon (Madhusudhan et al., 1993). The BCDH complex forms part of the pathway that catabolises the branched chain amino acids leucine, isoleucine and valine into acyl-CoA derivatives (Mercier et al., 2012). Specifically, the BCDH complex generates isobutyryl-CoA, α-methylbutyryl-CoA and isovaleryl-CoA. The *S. sudanensis* genome contains two copies of *bkdA* and three copies of *bkdB*. Of these, *bkdA_1* and *bkdB_1* are clustered together elsewhere in the genome, while the other three are present in this cluster. *S. sudanensis* lacks genes encoding the remaining subunits of the BCDH complex.

Lrp_2 encodes a leucine responsive regulatory protein. This class of proteins regulate transcription of genes in response to the presence of leucine and are abundant in prokaryotes (Brinkman et al., 2003). As this annotation for *lrp_2* is non-specific, it was decided to conduct BLASTP analysis, to determine whether *lrp_2* could be a *bkdR* homologue. The BkdR protein is

a member of the Lrp family, meaning there was a likelihood of homology (Madhusudhan et al., 1993). The Lrp_2 amino acid sequence was aligned with that of SCO3832, the BkdR protein of *S. coelicolor* A3(2). This gave a percentage identity match of 87%. It therefore seems likely that *lrp_2* is homologous to *bkdR*. The protein product of *bkdR* is a repressor of the *bkd* operon in *S. coelicolor* A3(2), but an activator of its transcription in other species, such as *Pseudomonas putida* (Madhusudhan et al., 1993, Sprusansky et al., 2005).

KstR2_2 encodes a HTH-type transcriptional repressor, with annotation indicating it to be homologous to *kstR2* from *M. tuberculosis* (Kendall et al., 2010). However, BLASTP analysis gave identity matches of over 90% to an unspecified TetR family repressor found in multiple *Streptomyces*. In contrast, the identity match to the *M. tuberculosis* KstR2 was found to be only 25%.

The genes *paaABCDE* and *paaH* are found in other bacteria as part of the larger *paa* operon, responsible for the degradation of phenylacetate into succinyl-CoA (Ferrandez et al., 1998, Teufel et al., 2010). The composition of the full operon varies between species, but it can be comprised of up to sixteen genes, namely *paaABCDEFGHIJKLNWXYZ* (Ferrandez et al., 1998, Navarro-Llorens et al., 2005, Di Gennaro et al., 2007). No *paa* genes other than the eight in this cluster are present within *S. sudanensis*. *paaABCDE* collectively encode subunits of an oxygenase complex, which usually acts in the second step of the degradation pathway, converting phenylacetyl-CoA into 1,2-epoxyphenylacetyl-CoA (Teufel et al., 2010). *paaH* encodes 3-hydroxyadipyl-CoA dehydrogenase and acts in the final step of the pathway, completing the synthesis of succinyl-CoA. The *S. sudanensis* genome contains only one copy of each of these genes, except for *paaA*, of which there are two.

The gene *rocA* encodes a dehydrogenase enzyme. A BLASTP search was conducted to check this annotation. This gave identity matches of 95% and higher to the phenylacetate degradation protein PaaN. The gene *paaN* is yet another member of the paa operon. The product of this gene is an aldehyde dehydrogenase, which has been found to specifically act as a ring-opening enzyme within the phenylacetate degradation pathway (Di Gennaro et al., 2007, Wang et al., 2016).

The eighth gene in the cluster, *trmH*, is predicted to encode a tRNA methyltransferase. BLASTP analysis instead found greater than 90% identity matches with rRNA methyltransferases from multiple *Streptomyces*.

The remaining two genes in the cluster both encode hypothetical proteins. Analysis of their potential functions was undertaken using BLASTP. *PABIEPML_02666* had over 90% identity matches with horizontally transferred transmembrane domain containing hypothetical proteins from other *Streptomyces*. *PABIEPML_02667* had high matches only with hypothetical *Streptomyces* proteins and was described to contain domain of unknown function 5819.

A search of the literature revealed that Zhao et al. (2015) reported a similar cluster in the genome of *Streptomyces pristinaespiralis*. The *S. pristinaespiralis* and *S. sudanensis* clusters are compared in Figure 4.12. They have a highly similar gene arrangement, including annotation of the Lrp encoding gene *bkdR*, in accordance with the above analysis. Additionally, *SSDG_03036 S. pristinaespiralis* is annotated as an rRNA methyltransferase, matching the BLASTP analysis of *trmH* in the *S. sudanensis* cluster. Further BLASTP analysis was performed to determine if the remaining two genes encoding hypothetical proteins were homologous to *PABIEPML_02666* and *PABIEPML_02667* of the *S.* sudanensis cluster. PABIEPML_02666 was found to have a 68% identity match to SSDG_03037 and PABIEPML_02667 had a 66.5% identity match to SSDG_03038. The two pairs of genes therefore share homology, adding to the similarity between clusters. There are two notable differences however: the *S. pristinaespiralis* cluster has a *bkdC* gene, rather than having two copies of *bkdB*; and it has a *paaZ* gene, in place of *rocA*/*paaN*.

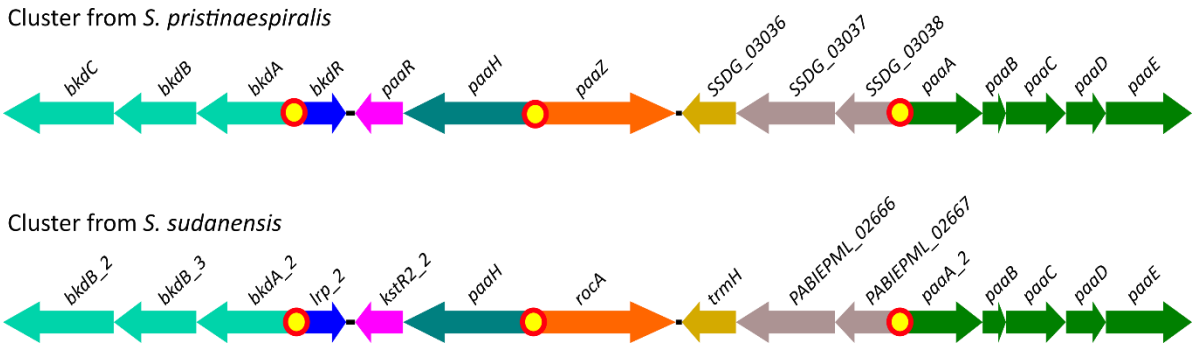


Figure 4.12 Comparison of the bkd-paa BGCs of *S. pristinaespiralis* and *S. sudanensis*. Diagram of the *S. pristinaespiralis* was adapted from Figure 2 of Zhao et al. (2015). The yellow dots indicate binding sites for PaaR, a TetR family transcriptional repressor.

The TetR family repressor was also specifically identified to be PaaR, based on homology to the PaaR of *Corynebacterium glutamicum*. The authors found that the highly conserved PaaR recognition sequence 5'-ACCGA-n4-TCGGT-3' was present at three points within this cluster, shown by yellow dots in Figure 4.12. They also confirmed experimentally that PaaR represses transcription of *paaABCDE* and *bkdABC* within the cluster. Addition of phenylacetyl-CoA is known to inhibit PaaR binding in *C. glutamicum* and the authors showed it to abolish PaaR-mediated repression in *S. pristinaespiralis*.

The *S. sudanensis* genome was manually searched for occurrences of the PaaR recognition sequence. It was found in the equivalent locations within the *S. sudanensis bkd-paa* cluster, as shown by the yellow dots in Figure 4.12. These locations are all upstream of the genes that were upregulated within the cluster. The product of the gene *kstR2_2* was next aligned to PaaR of *S. pristinaespiralis* using BLASTP. This gave a high identity match of 82% and a strong alignment between the proteins. It was accordingly decided that *kstR2_2* is highly likely to also be a *paaR* gene and the regulator of the cluster in *S. sudanensis*.

Zhao et al. (2015) further noted that a similar cluster is present in four other *Streptomyces* spp.: *S. venezuelae, S. avermitilis, S. griseus* and *S. clavuligerus*. *S. coelicolor, S. lividans* and *S. scabies* possess similar gene arrangements but split across two separate clusters: one containing *bkdC* through to *paaZ*, with the remaining genes in the other cluster.

4.5.2 Upregulation of whiD, a transcriptional regulator

The gene annotated as *whiD* was found to be upregulated, with a fold change of +3.39. The product of this gene is a member of the WhiB-like family of transcriptional regulators, which are found only in the Actinobacteria (Gao et al., 2006). This makes *whiD* one of only two transcription regulatory genes to be upregulated in response to macrophages, along with *PABIEPML_04454*. The *whiD* gene of *S. coelicolor* A3(2) plays a key role in the maturation of prespores and the formation of spore septa (Molle et al., 2000). However, *S. sudanensis* has not been observed to sporulate under lab conditions. A BLASTP alignment of the *S. sudanensis* and *S. coelicolor* A3(2) proteins revealed only a 49% identity match between the two. Additionally, multiple sequence alignment of the proteins was carried out using the Clustal Omega tool (Figure 4.13). The *S. sudanensis* WhiD is 12 amino acids longer than that of *S. coelicolor* and differences in sequence are concentrated at the N terminus of the protein. The

whiD gene and protein of *S. sudanensis* could therefore have different functions to those of *S. coelicolor*.

```
S. sudanensis WhiD   MLQLPHQPLQVAAVPSQRVPARQDEAGPWHTEAVCRRDEAGLFFAPSKEPTAARLAREEA       60
S. coelicolor WhiD   ------------MADFSRLPGPNADLWDWQLLAACRGVDSSLFFHPEGERGAARSARENS       48
                                 .  .*:*. : :   *:  *.**  ::.*** *. *  *** ***::

S. sudanensis WhiD   AKQVCARCPVMVDCREHALLQPEPYGVWGGLTAAERRVVLARRRRREVELKKAASADRRI       120
S. coelicolor WhiD   AKEVCMRCPVRAECAAHALAVREPYGVWGGLTEDEREELMGRARNRLVAATASASAGGEA       108
                     **:** **** .:*  ***   **********  **. ::.* *.* *  . :***. .

S. sudanensis WhiD   AAAG       124
S. coelicolor WhiD   AGPH       112
                     *.
```

Figure 4.13 Clustal Omega multiple sequence alignment of WhiD proteins of *S. sudanensis* and *S. coelicolor* A3(2). '*' = a fully conserved residue. ':' = conservation of residues with strongly similar chemical properties. '.' = conservation of residues with weakly similar chemical properties.

4.5.3 Analysis of the most highly downregulated genes

A group of three genes are labelled in the volcano plot in Figure 4.6. Within the downregulated genes, three stood out for the scale of their downregulation. *PABIEPML_04444*, *PABIEPML_02975* and *hmp* underwent fold changes of -7.43, -8.42 and -14.7 respectively. This is well above the mean for all seventeen downregulated genes of -4.63. High statistical significances also sets two of these genes apart from the rest of the group.

Hmp is annotated as encoding a flavohemoprotein. This carries out detoxification of nitric oxide (NO), converting it to nitrate and protecting the bacterium from nitrosative stress (Poole, 2020). Null mutants for this gene are hyper-sensitive to NO. It is therefore an important gene for promoting pathogen survival within a host. BLASTP analysis was carried out to check the accuracy of the annotated function. *S. sudanensis* Hmp was aligned against the *S. coelicolor* A3(2) proteins HmpA1 (SCO7428) and HmpA2 (SCO7094), both of which are identified by homology as flavohemoproteins on the UniProt database (UniProt, 2021). These alignments gave identity matches of 74% and 65% respectively. Additionally, it is known that *hmp* is usually clustered with *rrf2* in Actinobacteria (Rodionov et al., 2005). This gene was found to be immediately upstream of *hmp* in the *S. sudanensis* genome. Taken together, it was concluded from these findings that the gene is indeed *hmp*.

PABIEPML_04444 and *PABIEPML_02975* encode hypothetical proteins. BLASTP searches were run for both. For PABIEPML_04444, this gave matches only to unnamed crystallin family

proteins. Crystallins are also called small heat shock proteins and are expressed to help cells survive heat stress (de Jong et al., 1993, Singh et al., 2007). PABIEPML_02975 matched to unnamed ATP-binding proteins. No further insights into specific functions could easily be inferred.

4.6 Summary and conclusions

A genome mining approach was employed to discover whether *S. sudanensis* produces novel specialised metabolites. *De novo* sequencing of the *S. sudanensis* genome was performed using ONT and Illumina sequencing technologies, resulting in a single contig assembly of 5.3 Mb. AntiSMASH analysis of the genome predicted a total of nineteen BGCs to be present, of which fifteen had the potential to encode novel metabolites.

With a fully assembled genome, RNAseq could then be performed on *S. sudanensis*, using the indirect interaction assay system, in the presence and absence of macrophages. During analysis of the RNAseq reads, it was found that an unusually low number of reads mapped to *S. sudanensis* CDSs, the highest mapping percentage being 4.77% and the lowest 0.6%. Mapping to rRNA sequences substantially increased mapping percentages, but still left large portions of reads unmapped. The source of these reads remains unknown.

Statistical analysis of the CDS mapped reads showed that sample S_3 was an outlier and thus was removed from the dataset. Differential expression analysis was conducted on the remaining samples, which resulted in the identification of 60 differentially expressed genes. Of these, 43 were upregulated and 17 downregulated.

All predicted BGCs were found to contain at least one gene that was expressed under both the S and S+M conditions. However, only cluster 9, 15 and 18 contained genes that were differentially expressed. Through studying annotations and BLASTP analyses, upregulated genes within clusters 9 and 15 were surmised to boost biosynthesis of their cluster products. Their products are ochronotic pigment and an unknown siderophore respectively. Through the same analytical methods, the upregulated genes of cluster 18 were thought not to have a role in the biosynthesis of their cluster product. A potential two component system was identified within the upregulated genes of cluster 18, with *PABIEPML_04452* and

PABIEPML_04454 predicted to encode a sensor histidine kinase and HTH transcriptional regulator respectively by homology.

A highly upregulated BGC not predicted by antiSMASH, comprising mainly of *bkd* and *paa* biosynthetic genes, was identified within the *S. sudanensis* genome: eight of its fifteen genes were among the most highly upregulated in the RNAseq data. This cluster had previously been identified in several other *Streptomyces* spp. and its transcriptional regulation has been shown to be controlled by the TetR protein PaaR. A homologue of the *paaR* gene was found in the *S. sudanensis* cluster and the protein's conserved recognition sequence was found to be upstream of the upregulated cluster genes. The product of this cluster remains unknown.

Upregulation of a homologue of the Actinobacteria transcriptional regulator *whiD* was demonstrated by the dataset. This was one of only two regulators to be upregulated. WhiD has been shown to have a key role in sporulation of *S. coelicolor*. *S. sudanensis* has not been observed to sporulate, however, so this gene's function remains unclear. Among the downregulated genes, three had fold changes significantly greater than the rest. The precise cellular function of two of these genes remains unclear. The third gene however was identified as encoding flavohemoprotein, a vital bacterial defence against NO toxicity.

In conclusion, RNAseq yielded numerous genes of interest that gave an insight into the actinomycetoma pathogen response to macrophages. While only a small proportion of *S. sudanensis* CDSs were differentially expressed, these included members of three predicted BGCs, a previously unpredicted BGC, a transcriptional regulator which usually plays a role in key cellular processes and a gene vital to bacterial defence against one key weapon of the immune response.

Chapter 5. Structural and functional elucidation of compounds of interest from *S. sudanensis*

5.1 Introduction

As described in Chapter 3, *S. sudanensis* culture supernatant (SN) was found to have pyroptotic activity against THP-1 monocytes. Chemical purification methods were deployed to isolate and identify the compound or compounds responsible for this activity. Specifically, organic solvent extractions were used to isolate active compounds from culture SN, followed by preparative high performance liquid chromatography (HPLC) fractionation of the active organic phase. Liquid chromatography-mass spectrometry (LC-MS) and proton nuclear magnetic resonance (^{1}H-NMR) analyses were then employed for structural elucidation and identification of active compounds. Functional characterisation of the identified compounds was then attempted using a bioassay with THP-1 monocytes.

HPLC systems are commonly used for the purification of specialised metabolites from extracts (Latif and Sarker, 2012). They are highly versatile systems, wherein normal phase, reversed phase, ion exchange and size exclusion chromatographic columns can all be used, thus allowing compounds with a wide range of chemical properties to be isolated. HPLC systems generally feature the set up shown in Figure 5.1.

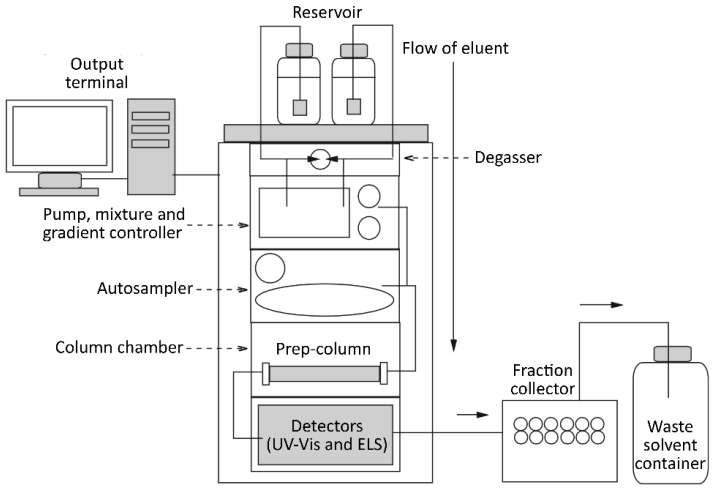


Figure 5.1 General set up of a HPLC purification system. Image taken from Figure 1 of Latif and Sarker (2012)**.**

The core innovation of HPLC systems lies in the structure of their columns, which are packed with a stationary phase comprised of particles that are 10 µm or smaller in diameter (Latif and Sarker, 2012). These are far smaller than particles found within lower pressure chromatography systems. Their small sizes mean the particles pack together tightly and in a regular manner, ensuring uniform flow of the mobile phase through the column. This aids chromatographic separation. Additionally, the particles maximise the stationary phase surface area that compounds in the mobile phase can interact with, enhancing the resolution of the chromatography. Thus, separation of compounds within a crude extract is greatly improved. The drawback of the tightly packed stationary phase is that very high pressures of up to 4000 psi are needed to push the mobile phase through the column.

HPLC is considered to be either preparative or analytical depending on the scale of the purification process (Latif and Sarker, 2012). Analytical HPLC is smaller scale, with column diameters of up to 5 mm and smaller loading volumes, while preparative HPLC is larger scale, with column diameters of up to 100 mm and larger loading volumes. The choice of method is dependent on the quantity of purified material needed. As ^{1}H-NMR analysis was planned following purification, a larger quantity of material was required. Thus, preparative HPLC was chosen for use.

LC-MS couples HPLC to a mass spectrometer, allowing for highly precise measurement of the masses of molecules present within a sample (Zhou et al., 2012). Mass spectrometers function by ionising molecules within a sample and then calculating their mass to charge ratio, from which their actual mass can then be derived (Pitt, 2009). The instruments are usually comprised of an ion source, a mass analyser and an ion detector, as shown in Figure 5.2. There are several types of ion sources that function via different methods (Zhou et al., 2012). They include electrospray ionisation (ESI), matrix-assisted laser desorption/ionisation, atmospheric pressure chemical ionisation, atmospheric pressure photoionisation and fast atom bombardment (Zhou et al., 2012, Singhal et al., 2015).

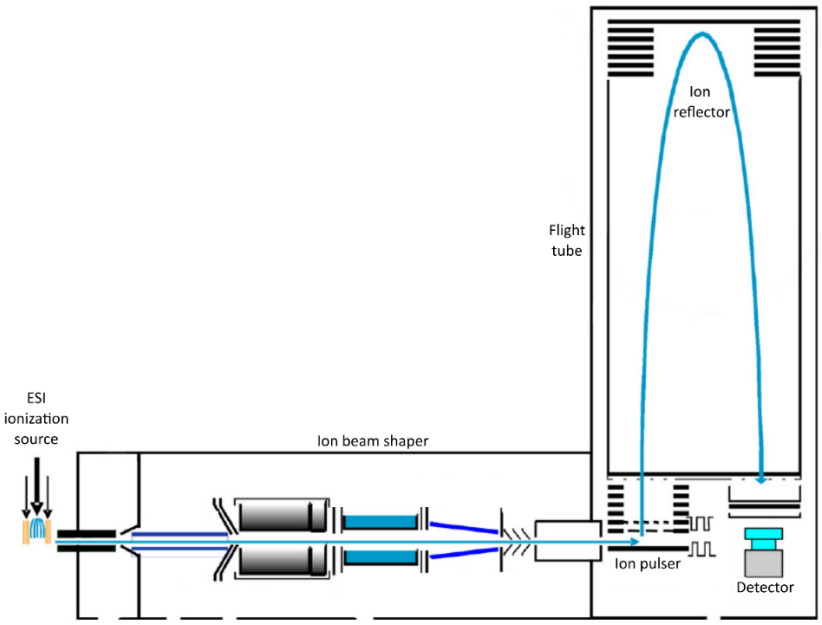


Figure 5.2 Schematic of a mass spectrometer featuring an ESI ion source, TOF mass analyser and a detector. Image from Creative Proteomics (n.d.).

The LC-MS system used in this study featured ESI, whereby a high voltage is applied to the liquid sample, causing it to disperse as an ionised aerosol (Pitt, 2009). The solvent gradually evaporates, leaving only the ionised compounds, which then enter the mass analyser. As with ion sources, there are multiple types of mass analyser, including time-of-flight (TOF), quadrupole, ion trap, Orbitrap and Fourier transform ion cyclotron (Zhou et al., 2012). The LC-MS system used here contained a TOF analyser. TOF systems accelerate ions down a flight tube, through an electric field, as shown in Figure 5.2 (Pitt, 2009). The time taken for the ions to reach the detector at the end of tube is dependent on their mass to charge ratio. The TOF of the ion can therefore be used to calculate this ratio and thus the mass of the ion.

The overall advantages of using an LC-MS system are that a sample can be kept in a liquid phase and any compounds present are separated as the sample flows through the HPLC column, meaning complex mixtures can be analysed (Pitt, 2009, Zhou et al., 2012).

NMR is a spectroscopic method that measures the number of atoms that exist within distinct chemical environments within a molecule (Pavia et al., 2015). Specific atomic nuclei are studied using NMR, with hydrogen and carbon being the most commonly targeted, due to their abundance within organic compounds. NMR can determine how many chemical environments a specific nucleus exists in, how many nuclei are within each of those environments and which atom types make up those environments. Using a combination of

NMR data and a compound's known mass from LC-MS analysis, it is possible to determine the complete chemical structure of a previously unidentified compound. NMR spectroscopy is therefore very useful in specialised metabolite research for the accurate and rapid identification of active compounds.

5.2 Purification of compounds of interest from *S. sudanensis* liquid culture and structural study

5.2.1 Compound isolation from liquid culture

An extraction with the solvent ethyl acetate was performed on *S. sudanensis* culture SN as described in Chapter 2, giving aqueous and organic extracts. The activity of these respective phases was assessed using a bioassay against THP-1 monocytes. Results from the assay are shown in Figure 5.3.

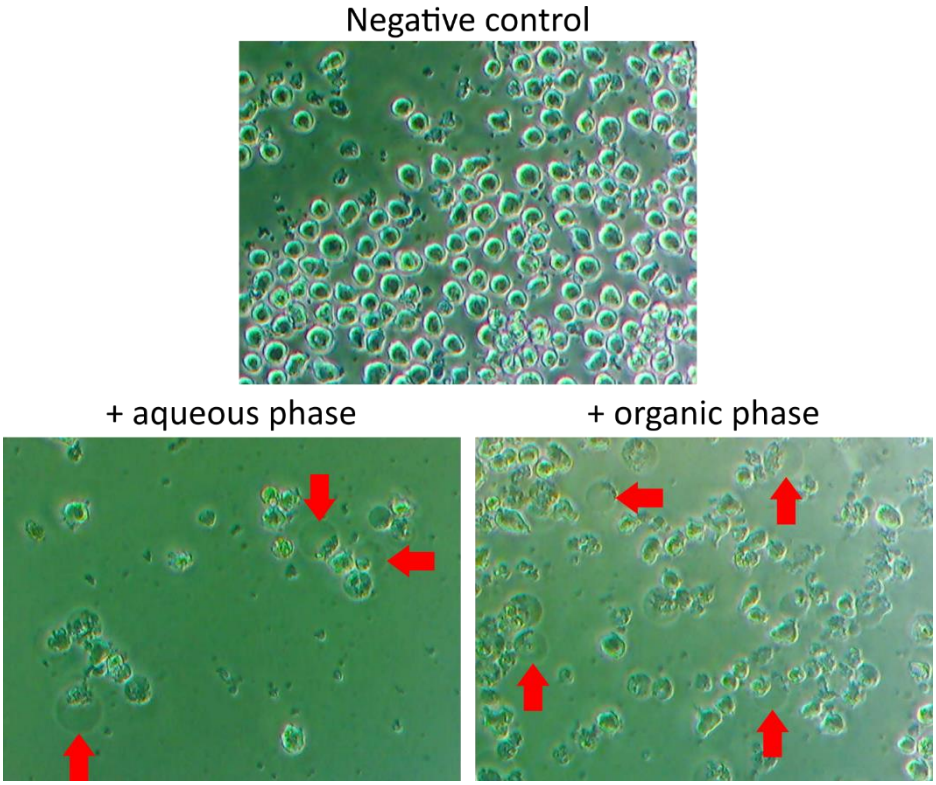


Figure 5.3 Images of THP-1 monocytes in isolation and challenged with the aqueous and organic phases of *S. sudanensis* culture SN. A small sample of the organic phase was dried and resuspended in water for use in the assay. 100 µL of each phase was used, comprising 50% of the assay well volumes. Images were taken at 48hrs and 100X magnification. Red arrows indicate examples of pyroptotic cells.

The organic phase induced pyroptosis, as shown by the presence of swollen, faded cells in the corresponding image of Figure 5.3. The compound(s) responsible for this activity therefore did move into the solvent during extraction. Pyroptotic activity was also seen in the aqueous phase, demonstrating that either the extraction was not 100% efficient or that non-extractable compounds were present.

In collaboration with Dr Yousef Dashti, the organic phase was evaporated to dryness, resuspended in a small volume of methanol and loaded onto a preparative HPLC column. HPLC fractionation of the concentrated extract was performed using a solvent gradient starting from 5% acetonitrile (ACN) and 95% water, moving up to 100% ACN over the course of 60 minutes. Fractions were collected every minute, giving a total of 60. A small sample of each fraction was dried and resuspended in water for use in a bioassay to detect activity against THP-1 cells. During this assay, fractions 14 and 18 were found to have strong apoptotic activity, as shown in Figure 5.4. These active fractions were taken forward for LC-MS and ^{1}H-NMR analysis to identify the compounds present within them.

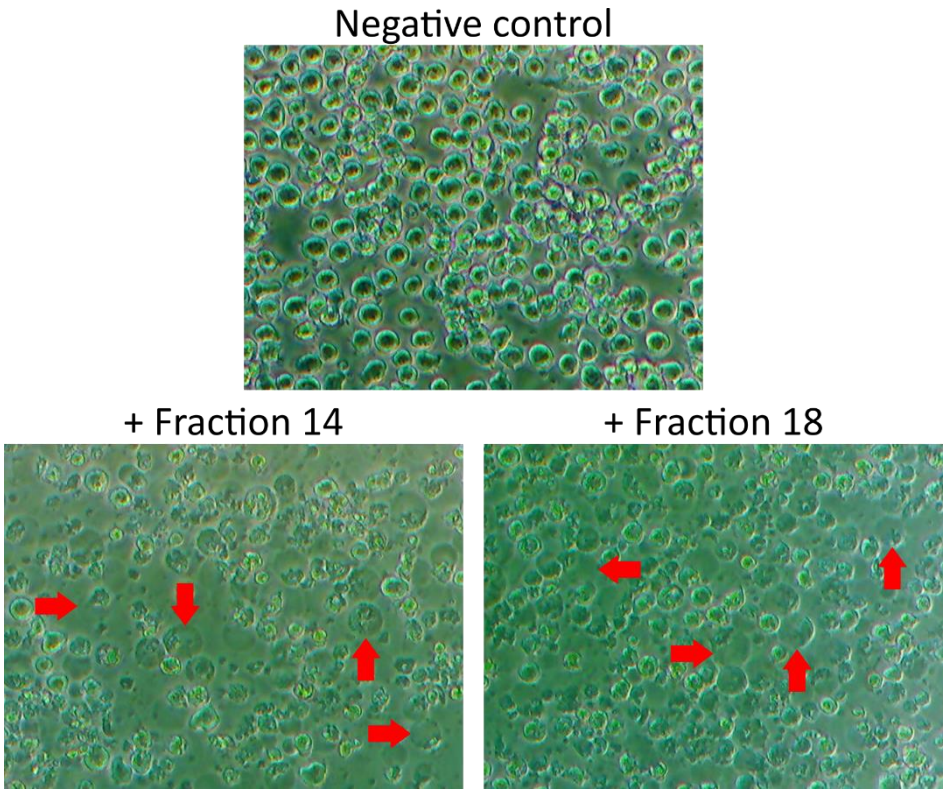


Figure 5.4 Images of THP-1 monocytes in isolation or challenged with HPLC fractions derived from *S. sudanensis* SN ethyl acetate phase. Fractions were added to a concentration of 20%. Images were taken at 24 hrs and 100X magnification. Red arrows indicate examples of pyroptotic cells.

5.2.2 Determination of compound structures by LC-MS and NMR

LC-MS analysis was performed on both fractions 14 and 18, in collaboration with Dr Yousef Dashti. The resulting chromatograms are shown in Figure 5.5.

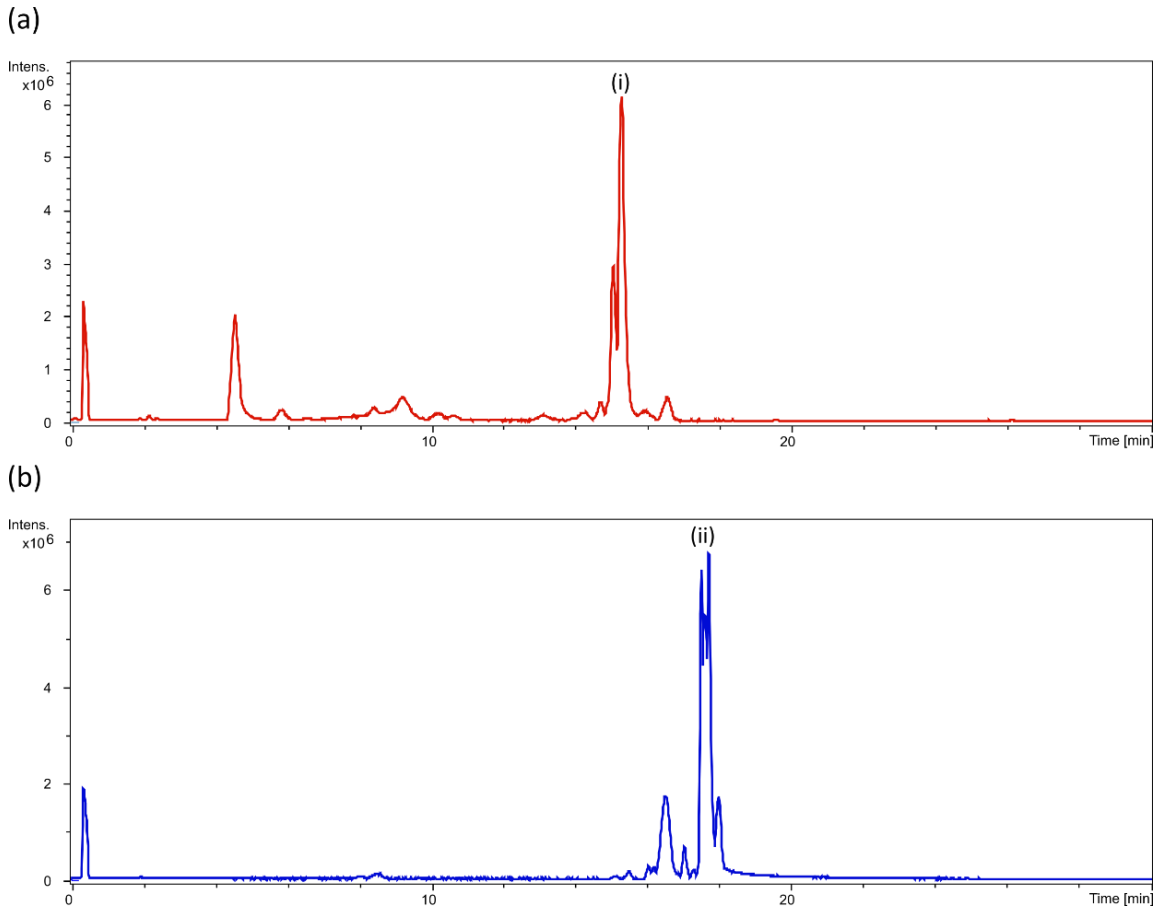


Figure 5.5 (a) LC-MS base peak chromatogram for HPLC fraction 14. Major peak is labelled as (i). **(b)** LC-MS base peak chromatogram for HPLC fraction 18. The major peak is labelled as (ii). For both fractions, the LC-MS method was run for 60 minutes but, as the spectra showed no peaks beyond 20 mins, they have been cropped to reduce their size. The Y-axes of the chromatograms show the relative intensity of the signal measured by the LC-MS detector, which is indicative of the abundance of a compound within the samples.

The chromatograms for both samples had relatively low numbers of peaks, indicating that the purification process was successful in removing large numbers of unwanted compounds. Both chromatograms showed one major peak, (i) and (ii) in Figure 5.5, with an intensity many times greater than that of other peaks. Each fraction was therefore mostly comprised of one compound. The top of peak (ii) in Figure 5.5(b) was split into smaller peaks. This indicated that the compound was present at such high levels in the fraction that the upper detection limit of the MS detector was surpassed. The true relative abundance of the compound in fraction 18 therefore remains unknown. The compounds responsible for peaks (i) and (ii) in Figure 5.5 are

hereafter referred to by their corresponding peak numbers. The mass spectra for these two compounds are shown in Figure 5.6.

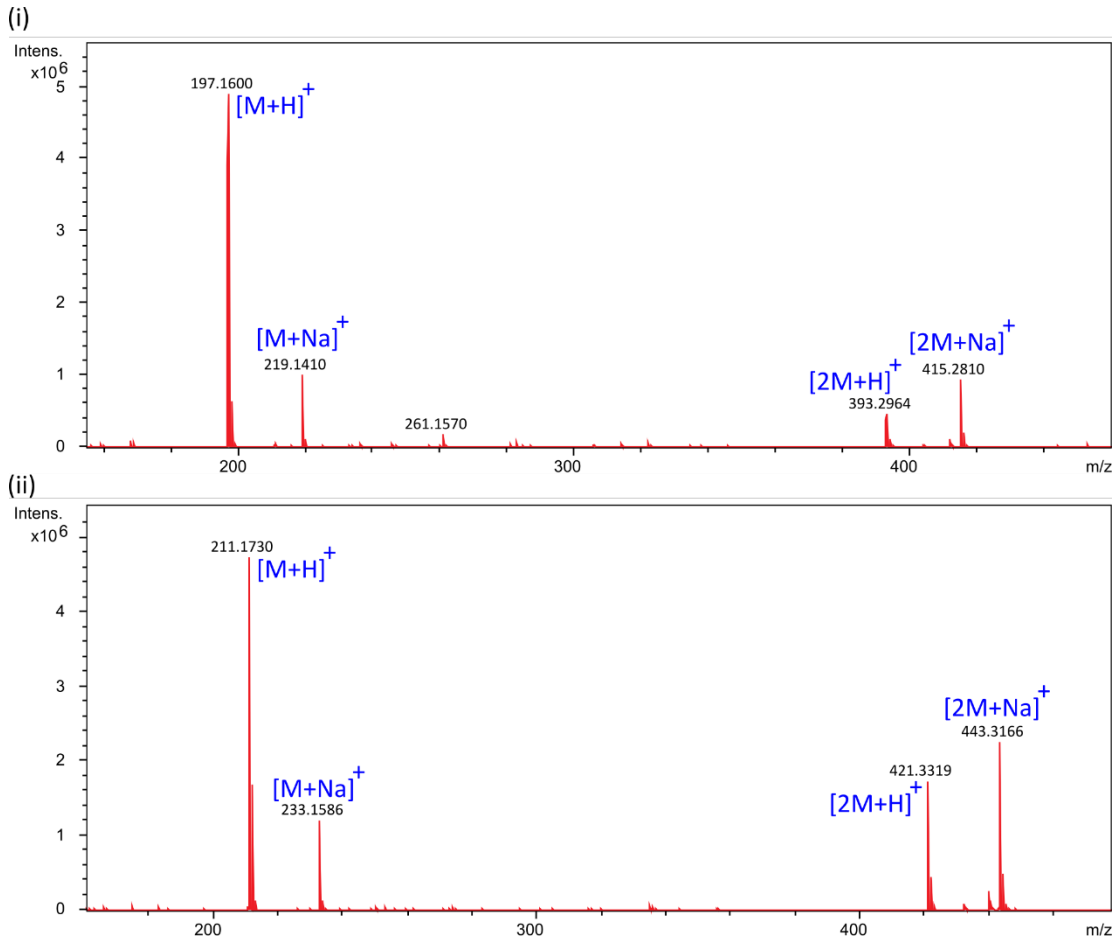


Figure 5.6 High resolution mass spectra of compounds (i) and (ii), identified in LC-MS chromatograms of HPLC fractions 14 and 18.

The mass spectrum for compound (i) showed four peaks at 197.1600, 219.1410, 393.2964 and 415.2810, which were assigned as the $[M+H]^+$, $[M+Na]^+$, $[2M+H]^+$ and $[2M+Na]^+$ ions respectively because of their relative mass differences. The exact molar mass of compound (i) was therefore assigned as 196.1522. The mass spectrum for compound (ii) showed four peaks at 211.1730, 233.1586, 421.3319 and 443.3166, which were assigned as the $[M+H]^+$, $[M+Na]^+$, $[2M+H]^+$ and $[2M+Na]^+$ ions respectively because of their relative mass differences. The exact molar mass of compound (ii) was assigned as 210.1652. With such high abundances within each fraction, it was deemed extremely likely that these compounds were responsible for the pyroptotic activity of the two fractions.

^{1}H-NMR structural analysis was then performed on fractions 14 and 18 by Dr Dashti. The ^{1}H-NMR and high-resolution MS data were compared with values published in the literature and compounds (i) and (ii) were identified as cyclo(-Val-Pro) and cyclo(-Leu-Pro) respectively (Schmidtz et al., 1983, Dashti et al., 2014). Their structures consist of two amino acids cyclised together and are shown in Figure 5.7.

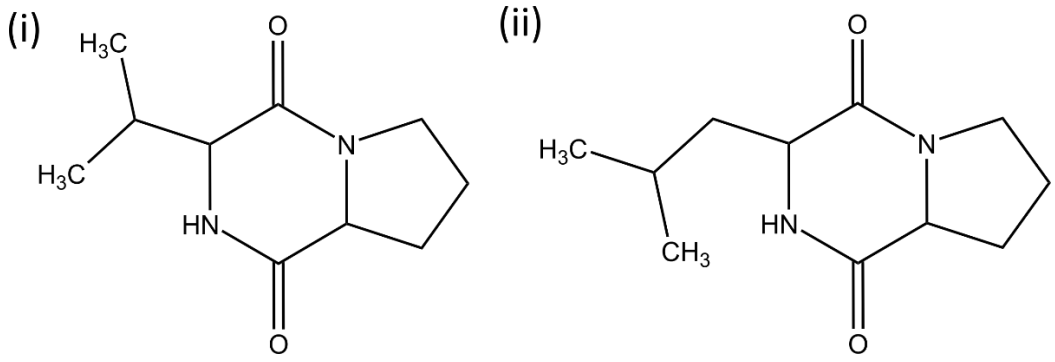


Figure 5.7 Structures of the two compounds identified in LC-MS of pyroptosis-inducing HPLC fractions of *S. sudanensis* SN, as determined by ^{1}H-NMR. (i) Cyclo(-Val-Pro) – 196.25 g/mol (ii) Cyclo(-Leu-Pro) – 210.28 g/mol.

As an additional confirmation of the identity of the compounds, commercial synthetic forms of each were obtained and ^{1}H-NMR performed. As shown in Figure 5.8, these spectra matched those obtained from the HPLC fractions, verifying the identities of the two compounds.

Cyclo(-Leu-Pro) and cyclo(-Val-Pro) are 2,5-diketopiperazines, otherwise known as cyclic dipeptides. These compounds are found across all domains of life and generally function as signalling molecules, being involved in quorum sensing in bacteria and acting as hormone-like molecules in mammals (Bellezza et al., 2014, Chmielewski, 2021). Beyond their signalling roles however, they have been found to have a wide range of activities, including: antibiotic, antiviral, antioxidant, antitumour, anti-inflammatory and neuroprotective. This diversity of activity has been attributed to their strong hydrogen bonding activity, as they contain two hydrogen bond donors and two acceptors within their 2,5-diketopiperazine rings (Chmielewski, 2021). It is postulated that this allows the compounds to bind to a range of different receptors.

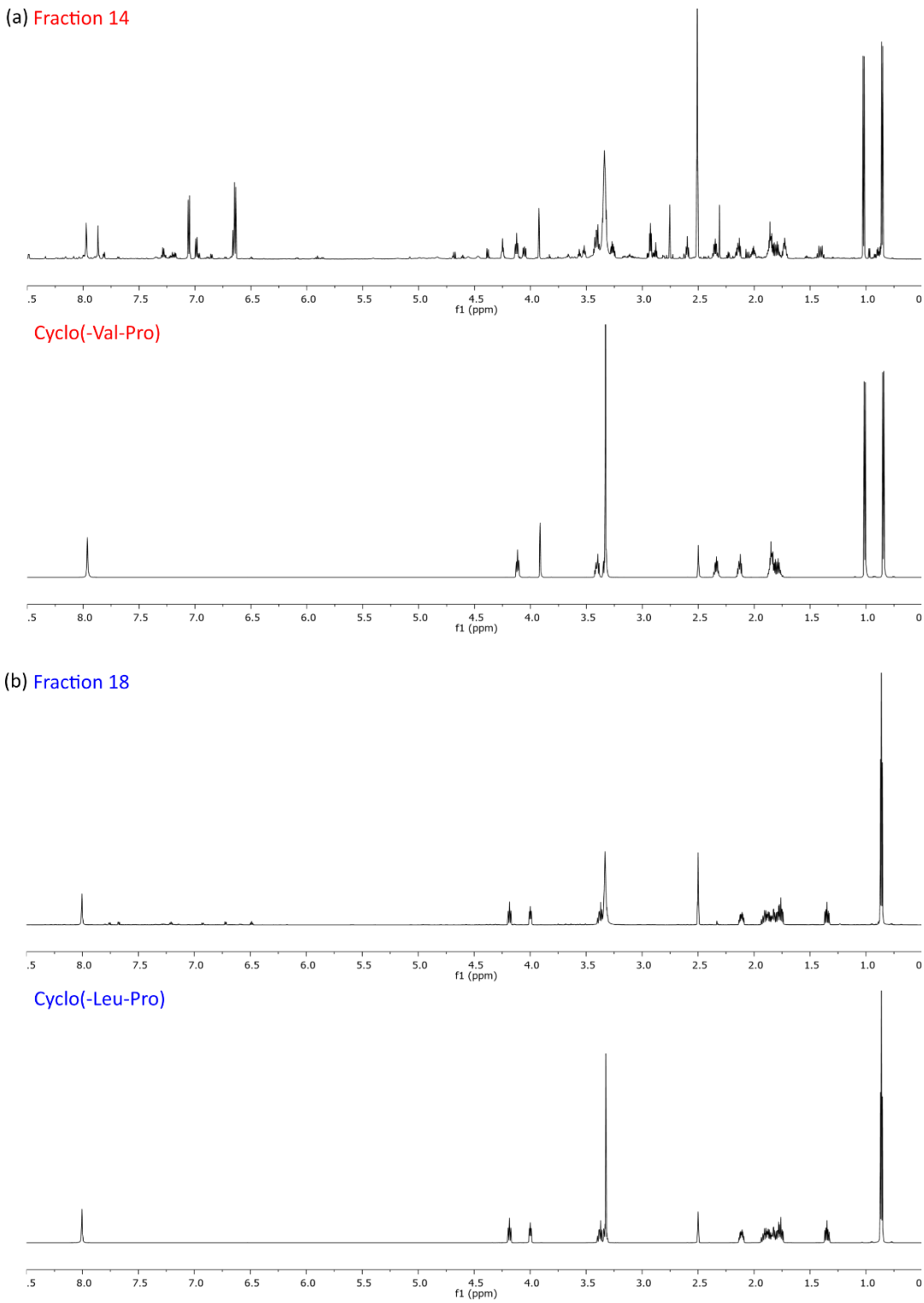


Figure 5.8 Comparison of ^{1}H-NMR spectra of **(a)** HPLC fraction 14 and synthetic cyclo(-Val-Pro) **(b)** HPLC fraction 18 and cyclo(-Leu-Pro).

Within bacteria and fungi, these compounds are synthesised by either non-ribosomal peptide synthetases (NRPSs) or cyclodipeptide synthases (CDPSs) (Belin et al., 2012, Gondry et al., 2018). In the case of NRPSs, the dipeptides are either synthesised by a dedicated NRPS system,

or they are produced as by-products from spontaneous cyclisation of substrates during NRPS synthesis of another product (Belin et al., 2012). CDPSs are single enzymes, encoded by one gene and are dedicated only to the production of cyclic dipeptides (Gondry et al., 2018). They take amino acids bound for the ribosome in the form of aminoacyl-tRNAs and cyclise them together. CDPSs are promiscuous, synthesising a primary cyclic dipeptide alongside several secondary ones.

Four NRPSs were predicted within the *S. sudanensis* genome by antiSMASH, but none were annotated as being dedicated to cyclic dipeptide synthesis. Therefore, if the two compounds are produced by any of the four NRPSs, they would likely be spontaneous by-products as described above. A single CDPS was identified in the *S. sudanensis* genome, namely *PABIEPML_02330*, which is annotated as encoding a cyclo(-Tyr-Tyr) synthase. However, within the RNAseq data, no sequencing reads mapped to this gene, giving a read count of zero in all three repeats, both in the presence and absence of macrophages. It therefore appears that this gene was not expressed. However, the low mapping rates of the *S. sudanensis* RNAseq reads means it is likely only the most highly expressed genes with the greatest quantity of mRNA were sequenced. Therefore, false negatives (genes showing no expression) are likely to be present in the dataset, of which this gene could be one.

5.3 Activity of synthetic 2,5-diketopiperazines against THP-1 cells

To test if the two identified compounds do have pyroptotic activity, synthetic forms were tested against THP-1 monocytes. The synthetic compounds had poor to no solubility in water, so stock solutions were prepared using 100% DMSO. Only 1 µL of each stock was added to wells for the activity assay, meaning the final DMSO concentration in each well was no greater than 0.5%. Additionally, preliminary experiments surprisingly didn't show activity, so it was speculated that perhaps the highly pure synthetic dipeptides are unable to prime the inflammasome, but instead only activate it following priming by another stimulus. Therefore, the activities of the compounds were tested in the presence and absence of lipopolysaccharide (LPS) priming. Selected results from the activity assays are shown in Figure 5.9.

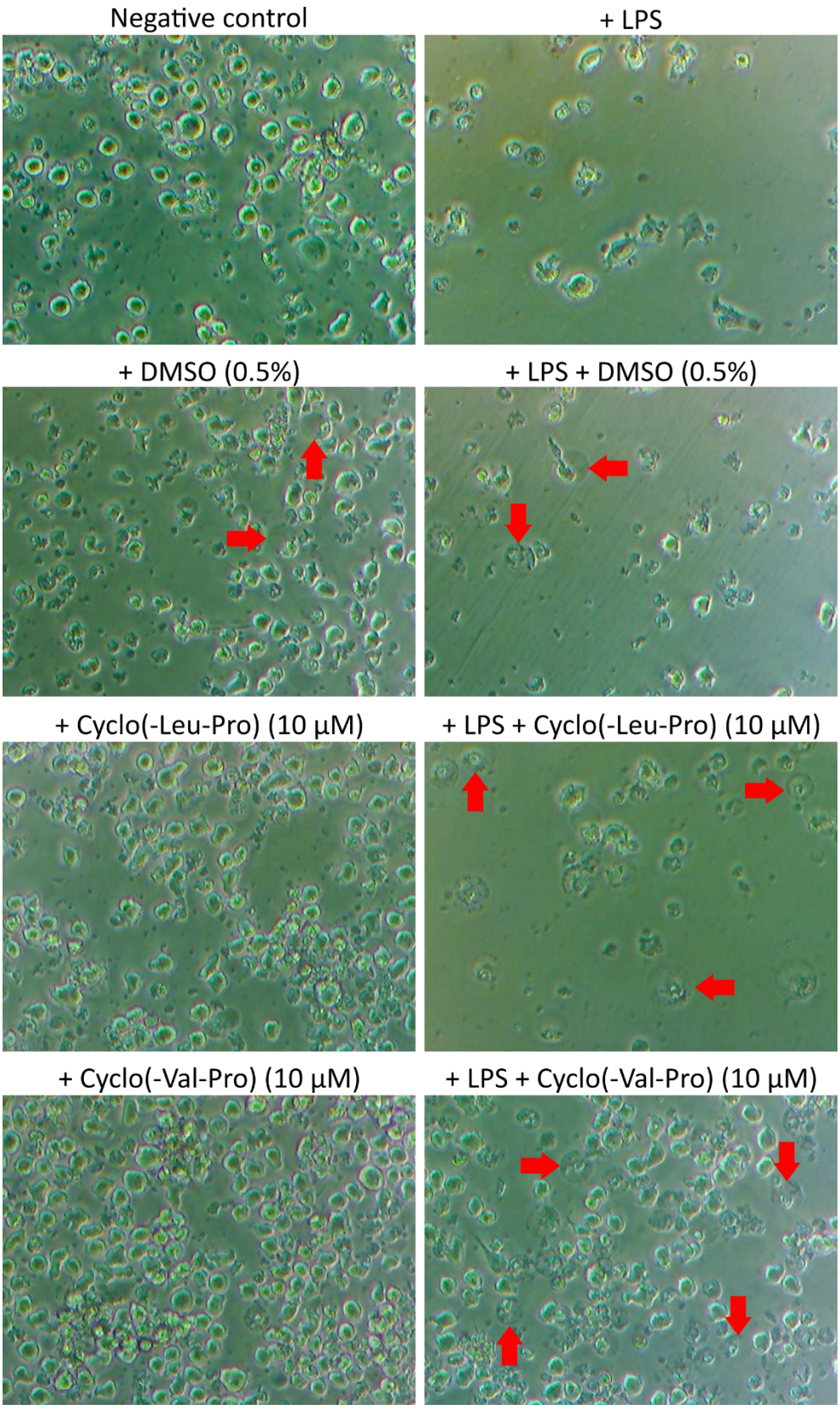


Figure 5.9 Images of THP-1 monocytes during assay testing the activity of two synthetic cyclic dipeptides. Images were taken at 24 hrs, at 100X magnification and represent only one biological repeat. Red arrows indicate examples of pyroptotic cells.

The ‘+ LPS’ control demonstrated that exposure to LPS triggered immune activation in the monocytes. This can be seen in the corresponding image in Figure 5.9, where the cells are enlarged and have irregular, starry morphologies. The low cell number in most of the ‘+ LPS’ images compared to the negative control was not due to cytotoxicity. The LPS priming method involved removal of the LPS after 3 hrs and during this pipetting process many cells were lost as they detached from the cell culture well. The final two control conditions involved adding DMSO to a concentration of 0.5%. Both in the presence and absence of LPS priming, the DMSO had a cytotoxic effect on the cells, with multiple shrivelled, dead cells visible in both control wells. Additionally, a very small number of what appeared to be pyroptotic cells were visible in both wells. This weak effect was attributed to environmental stress caused by the presence of the DMSO. Even so, the ability of DMSO to induce this form of death, even at such low levels, was considered when examining the effects of the cyclic dipeptides.

In the absence of LPS priming, neither cyclo(-Leu-Pro) or cyclo(-Val-Pro) showed pyroptotic activity. Some cytotoxicity was observed but was likely due to the DMSO, as was seen in the ‘+ DMSO’ controls. When added following LPS priming however, both compounds appeared to induce pyroptosis. High quantities of pyroptotic cells could be seen in both wells, as exemplified in the images for ‘+ LPS + Cyclo(-Leu-Pro)’ and ‘+ LPS + Cyclo(-Val-Pro)’ in Figure 5.9. It was also noted that the induction of pyroptosis in these wells was greatly increased in comparison to the low levels seen in the ‘+ DMSO’ controls. This effect was observed for each concentration of the compounds tested, which ranged from 5 mM down to 10 µM. The synthetic compounds triggering pyroptosis only following LPS priming contrasts with the active HPLC fractions from the *S. sudanensis* culture SN, which appeared to induce pyroptosis without LPS priming. A likely explanation for this is that the HPLC fractions still contained traces of bacterial pathogen associated molecular patterns, e.g. cell wall components, that primed the inflammasome.

5.4 Summary and conclusions

Having identified NLRP3-mediated pyroptotic activity in the SN of *S. sudanensis* cultures, a purification pipeline was developed to isolate and identify the compound or compounds responsible for this activity. Following organic extraction and HPLC fractionation, LC-MS and NMR analyses identified two compounds to be present within these fractions: cyclo(-Val-Pro) and cyclo(-Leu-Pro). Both are examples of 2,5-diketopiperazines, or cyclic dipeptides. Synthetic forms of these compounds were obtained and tested for activity against THP-1 monocytes. Both appeared to induce pyroptosis in cells primed with LPS, down to the lowest concentration tested of 10 μM.

In conclusion, two cyclic dipeptides identified in *S. sudanensis* SN appear to possess the ability to induce inflammatory pyroptosis in human monocytes. Their presence in *S. sudanensis* culture suggests that this bacterium produces these compounds. Further biological repeats testing the activities of these compounds are needed to confirm this conclusion.

Chapter 6. Discussion

6.1 Introduction

This study was aimed to investigate the nature of actinomycetoma pathogens' interactions with the immune system and to identify virulence factors within these pathogens, in particular any novel specialised metabolites that might play a role in pathogenesis or have therapeutic potential for treatment of other conditions. A custom host-pathogen interaction assay was designed to interrogate the early stages of actinomycetoma pathology and to assess the involvement of secreted specialised metabolites in such interactions. RNA sequencing (RNAseq), cytokine enzyme-linked immunosorbent assays (ELISAs), an NF-κB activity assay and microscopy were deployed to observe how murine macrophages and the pathogen *S. sudanensis* adapt and respond to each other's presence within this interaction assay. *De novo* genome sequencing was also used to provide insight into the specialised metabolite production potential of *S. sudanensis*. Bioassays and chemical purification processes were additionally utilized to isolate and identify compounds responsible for effects observed during the interaction assay. The significance and implications of results gained from these methods are discussed herein.

6.2 The macrophage response to actinomycetoma pathogen

The macrophage response to the actinomycetoma pathogen *S. sudanensis* was investigated using the indirect interaction assay and RNAseq (Chapter 3). Upregulation of multiple inflammatory marker genes was observed (Figure 3.15) for all four challenge conditions. A majority of these genes showed greatest upregulation in the M+S condition. The NF-κB activity assay (Figure 3.23) also gave evidence of stronger inflammatory activation compared to the non-pathogen, with M+S having induced greater NF-κB activity overall compared to M+C.

Additionally, the specific observations from the M+S condition fit with previous *in vivo* and *in vitro* findings that actinomycetoma pathogens induce an initial pro-inflammatory response in the host upon infection (Solis-Soto et al., 2008, Santiago-Tellez et al., 2019). Furthermore, the upregulation of the key inflammatory genes *Il1β*, *Il12β* and *Tnfα*, as well as the raised TNFα levels shown by the ELISA in Figure 3.20, correlates with data from eumycetoma patients

where the levels of the proteins encoded by these genes were raised (Nasr et al., 2016, Abushouk et al., 2019).

These data provide strong evidence that macrophages under all four challenge conditions were polarised to an M1, pro-inflammatory phenotype. These results also point to living *S. sudanensis* having induced an enhanced pro-inflammatory response compared to the other conditions. The strength of this response appeared even further pronounced when compared to that induced by living *S. coelicolor*. *S. sudanensis* therefore appears to be more immunogenic than *S. coelicolor*. This could be indicative of the production of extracellular antigenic molecules by *S. sudanensis*. The greater upregulation of pro-inflammatory genes induced by living compared to dead *S. sudanensis* also supports the notion of production of antigenic molecules with immunostimulating properties by the pathogen.

Of particular note was the sizeable upregulation of the macrophage matrix metalloproteinase (MMP) encoding genes *Mmp9* and *Mmp13* in M+S. This is significant as both of the proteins encoded by these genes are involved in cutaneous wound repair, as well as regulation of inflammation (Nissinen and Kahari, 2014). Their excessive upregulation has been associated with the pathologies of other chronic inflammatory diseases, such as cartilage destruction in murine arthritis (Joronen et al., 2005). Additionally, they are highly expressed during tuberculosis infections, where they are thought to play a role in granuloma formation and tissue damage, aiding pathogen persistence (Quiding-Jarbrink et al., 2001). Also, as described in Chapter 1, their collagenase and gelatinase activities implicate them in the development of fibrosis, including around granulomas (Geneugelijk et al., 2014, Nissinen and Kahari, 2014). Two studies found protein levels of MMP9 to be raised in mycetoma patients, with one also finding a positive correlation between MMP9 expression and disease duration (Geneugelijk et al., 2014, Siddig et al., 2019b). MMPs could therefore be integral to the development of mycetoma pathology.

The high upregulation of *Mmp9* in the RNAseq data aligns with the observations from the mycetoma patient studies. It was also notable that living *S. sudanensis* induced a seven-fold greater upregulation of *Mmp9* compared to dead *S. sudanensis* and induced upregulation of *Mmp13* while the dead bacteria failed to do so (Figure 3.15). This suggests that the living bacteria actively secrete a molecule that promotes MMP expression, which is not produced in the metabolically inert dead cells. This molecule is likely secreted, as if it were a cell wall

component the high density of dead bacteria (containing a high density of cell wall material) used in the assay would probably have induced similar expression of the two genes. The greater expression of these genes in M+S compared to M+C further suggests that this is a pathogenicity associated mechanism and thus the molecule responsible may constitute a virulence factor.

Expression of the macrophage gene *Nos2* was only slightly upregulated in response to the living pathogen compared to the other three challenge conditions. This gene encodes nitric oxide synthase, which produces nitric oxide (NO), a key bactericidal weapon deployed by immune cells (MacMicking et al., 1997). This lower upregulation therefore represents one of the only major components of the pro-inflammatory immune response not to be enhanced in response to *S. sudanensis* compared to the control conditions. The significance of a weakened NO response to these bacteria is highlighted by the fact that an allele associated with reduced NO production makes humans more vulnerable to mycetoma (van de Sande et al., 2007). Lower NO production equates to reduced bacterial killing and thus would contribute to mycetoma pathogen persistence. In addition to a pro-inflammatory immune response, elements of an anti-inflammatory phenotype were also observed in macrophages under the M+S condition.

Multiple anti-inflammatory, wound healing genes were shown to be uniquely upregulated in M+S (Figure 3.16). Observations at the protein level by ELISA also measured a significant release of anti-inflammatory IL10 from macrophages challenged with living *S. sudanensis* (Figure 3.22). This is consistent with studies that found IL10 to be dominant within mycetoma lesions (El Hassan et al., 2001, Salinas-Carmona et al., 2012). In addition to this, the unique elongated morphology induced in macrophages by living *S. sudanensis* (Figure 3.10) may point towards an anti-inflammatory profile. A study by McWhorter et al. (2013) found that such an elongated cell shape in macrophages signifies polarisation to the anti-inflammatory M2 phenotype. A transition to an anti-inflammatory immune environment has previously been recorded in mycetoma infections (El Hassan et al., 2001, Mendez-Tovar et al., 2004, Salinas-Carmona et al., 2012, Nasr et al., 2016, Abushouk et al., 2019). As the indirect interaction assay was only run for a short time period, it did not model the chronic nature of actinomycetoma. The observations above are perhaps indicative of the beginnings of the transition to a pathogen-permissive anti-inflammatory environment. It therefore seems likely that if given a

longer incubation time, the immune response would eventually polarise further to an anti-inflammatory state. Furthermore, these effects not being seen in the presence of the non-pathogenic control or dead bacteria once again leads to a conclusion that they were mediated by the pathogen secreting one or more molecules that diffused into the macrophage culture.

Il4rα, encoding the receptor for IL4, was upregulated in M+S, meaning the macrophages may have been primed to respond to this cytokine. IL4 is implicated in granuloma formation and addition of IL4 to macrophage cultures has been shown to induce differentiation to epithelioid cells (Cronan et al., 2016, Pagan and Ramakrishnan, 2018). Additionally, animals deficient in the IL4 receptor were unable to form organised, epithelioid granulomas (Cronan et al., 2021). Macrophages do not produce IL4, which is instead secreted by mast cells, basophils, eosinophils and Th2 cells (Gadani et al., 2012). This means there was no source of this cytokine within the indirect interaction assay and therefore no ligand for the IL4 receptors to detect. In any case, through increasing expression of its receptor, the living mycetoma pathogen likely made macrophages more sensitive to IL4 and thus granuloma formation, while the non-pathogen did not. As the interaction assay was run over a short time period (48 hrs) it may have given insight into the earliest stages of granuloma development.

Dead *S. coelicolor* also induced this effect, albeit not to the same extent. A potential explanation for this difference between living and dead *S. coelicolor* may be the fact that the dead cells were added to the assay at a much higher density than the living. The macrophages could therefore have been exposed to a larger quantity of pathogen associated molecular pattens (PAMPs), such as cell wall components, during incubation, altering the extent and nature of immune activation.

The unique upregulation of *Vegfa* in M+S should also be noted, as this cytokine promotes the fibrotic development of granulomas, commonly seen in mycetoma (Pagan and Ramakrishnan, 2018). The already mentioned MMP9 is also known to proteolytically activate vascular endothelial growth factor (VEGF) A and so it seems probable that VEGFA would indeed be active within the context of the interaction assay.

The mevalonate pathway, which is responsible for cholesterol and isoprenoid biosynthesis, was observed to be transcriptionally repressed in M+S (Figure 3.17) (Gruenbacher and Thurnher, 2017). It could be that the living pathogen disrupted the metabolism of the

macrophages in some manner, meaning resources that would normally go to this pathway were diverted elsewhere. The literature is lacking in studies on disruption of the pathway within the context of bacterial infection, but within the cancer field it has been found that a reduction of metabolic flux through the mevalonate pathway triggers a pro-inflammatory response in immune cells (Gruenbacher and Thurnher, 2017). This could therefore be the mechanism underlying the enhanced pro-inflammatory response observed in macrophages challenged with *S. sudanensis*. However, the reason for the repression of the pathway remains unclear.

Genes encoding the transcription factors (TFs) AP-1 and CEBPB were upregulated by living and dead *S. sudanensis* (Section 3.6.1). Both are known to enhance the expression of pro-inflammatory genes (Uematsu et al., 2007, Fitzgerald and Kagan, 2020). Whatever the trigger of the enhanced pro-inflammatory response to *S. sudanensis*, the transcriptional changes may therefore have been mediated partly by the activities of these TFs.

In summary, a unique immune response to an actinomycetoma pathogen, comprising of an enhanced pro-inflammatory profile coupled to the initiation of a transition to an anti-inflammatory phenotype, has been observed and characterised in macrophages.

6.3 Identification of potential virulence factors in an actinomycetoma pathogen

How the pathogen *S. sudanensis* responds to macrophages was also investigated during the indirect interaction assay using RNAseq (Chapter 4). Upregulation of *hpd* within biosynthetic gene cluster (BGC) 9 of *S. sudanensis* was observed in the RNAseq data (Table 4.4). As described in Chapter 4, this gene encodes 4-hydroxyphenylpyruvate dioxygenase, which produces homogentisic acid that in turn polymerises to form pyomelanin, the predicted product of cluster 9. In eumycetoma, production of fungal melanins has been shown to constitute part of grain formation following initial infection (Sheehan et al., 2020). Biosynthesis of pyomelanin specifically has also been demonstrated in eumycetoma fungal pathogens (Lim et al., 2021). Upregulation of *hpd* in *S. sudanensis* therefore suggests that pyomelanin biosynthesis could also be part of actinomycetoma grain development.

Pyomelanin production has been observed within *S. avermitilis*, which is not known to be a mycetoma pathogen (Denoya et al., 1994, Ōmura et al., 2001). However, this should not

eliminate pyomelanin from contention as having a key role in mycetoma development. It is possible that pathogens such as *S. sudanensis* have evolved a transcriptional regulator which is able to respond to the presence of host immune factors and strongly upregulate pyomelanin biosynthesis. Evolution of such a regulator could distinguish pathogen from non-pathogen.

Production of pyomelanin has also been identified across a wide variety of bacterial genera, including *Klebsiella*, *Burkholderia*, *Pseudomonas*, *Vibrio* and *Legionella* (Yabuuchi and Ohyama, 1972, Rdest et al., 1991, Plonka and Grabacka, 2006, Singh et al., 2018). Within these bacterial species, pyomelanin serves a variety of virulence-related functions. It can reduce iron to trigger its release from the mammalian siderophores ferritin and transferrin, as well as promoting its uptake by bacteria (Chatfield and Cianciotto, 2007, Zheng et al., 2013). It protects bacteria from oxidative stress by removing immune cell produced reactive oxygen species (ROS) from the environment (Zughaier et al., 1999). It offers growth and energetic advantages to bacteria by acting as an alternative terminal electron acceptor in respiratory chains (Turick et al., 2002). Finally, it has also been shown to have haemolytic activity, promoting the release of sequestered nutrients into the surrounding environment for bacterial uptake (Liang et al., 2016, Liang et al., 2018). With these known functions, a potential role in grain formation and upregulation of its biosynthesis in the indirect interaction assay, there is strong evidence that pyomelanin constitutes a mycetoma virulence factor. Therefore, *hpd* upregulation is likely to have implications for actinomycetoma pathology.

Competition between host and pathogen for micronutrients such as metal ions is a core process in the development of disease (Palmer and Skaar, 2016). Host cells produce various factors that sequester such resources away from pathogens, thus limiting their growth and spread, while pathogens take in micronutrients from the host environment. Siderophores are common molecules produced by both host cells and pathogens during such interactions, in order to control iron resources (Kramer et al., 2020). Iron is utilised as a cofactor in a variety of enzymes that catalyse key metabolic reactions within both eukaryotic and prokaryotic cells, hence its importance as a resource. The ability of a bacterium to harvest iron within the host environment therefore partially determines its virulence. The upregulation in cluster 15 of core siderophore biosynthetic and transport machinery (Table 4.4) in the presence of macrophages accordingly points to this cluster producing another virulence factor.

While the siderophore predicted to be produced by cluster 15 remains unknown, due to the cluster's low homology to known BGCs, the context of the upregulation of its biosynthesis means it likely contributes to *S. sudanensis* virulence. There is further evidence of competition for iron during the indirect interaction assay, as the siderophore encoding genes *Lcn2* and *S100a8*, as well as the transferrin receptor gene *Tfrc*, were upregulated in macrophages under the M+S condition (Figure 3.15).

A BGC not predicted by antiSMASH and comprised of *bkd* and *paa* genes was identified in the *S. sudanensis* genome (Figure 4.11) from its strong upregulation. Despite previously identifying this BGC and characterising its transcriptional regulation, Zhao et al. (2015) did not discover or predict a potential product from it. They only concluded that the clustering of *paa* and *bkd* genes must mean the process of branched chain fatty acid catabolism is regulated by the same system as phenylacetate catabolism.

It has been highlighted in the literature that *Streptomyces* spp. often contain multiple *bkd* gene clusters, at least one copy of which is usually dedicated to the production of specialised metabolites (Stirrett et al., 2009). Clusters of *bkdABC* have been shown to be essential for the synthesis of avermectin, actinorhodin and virginamycin in *S. avermitilis*, *S. coelicolor* and *S. virginiae* respectively (Denoya et al., 1995, Pulsawat et al., 2007, Stirrett et al., 2009). The genes *paaABCDE* have also been demonstrated to be active beyond their normal pathway, having roles in the synthesis of tropolenes in *Streptomyces* spp., as well as the antibiotic tropodithietic acid in *Phaeobacter inhibens* (Brock et al., 2014, Wang et al., 2016, Chen et al., 2018). In the above processes, the *bkd* and *paa* genes do not operate as part of the core biosynthetic machinery for the natural products, but instead synthesise starter units/precursors for the biosynthetic pathways. The existing literature does not address whether the enzymes encoded by these two groups of genes can function coherently together to produce a single chemical product.

Interestingly, *paaABCDE* have been found to be associated with virulence in the Gram-negative bacteria *Burkholderia cenocepacia*, *Acinetobacter baumanii* and *Pseudomonas aeruginosa* (Law et al., 2008, Musthafa et al., 2012, Wang et al., 2013a, Pribytkova et al., 2014, Bhuiyan et al., 2016). Specifically, deletion of these genes or inhibition of their encoded enzyme complex attenuated virulence in the above strains. This has been attributed to two mechanisms. Firstly, inhibition of phenylacetate degradation led to an accumulation of this

molecule, which then inhibited bacterial quorum sensing and thus triggered a change in gene expression (Musthafa et al., 2012, Wang et al., 2013a, Pribytkova et al., 2014). Secondly, accumulation of phenylacetate was found to enhance neutrophil migration to the site of infection, leading to increased bacterial killing (Bhuiyan et al., 2016).

The upregulation level of the *bkd-paa* cluster in *S. sudanensis* in the presence of macrophages does suggest these genes are likely to play a role in the bacterium's response to immune cells. The cluster is therefore likely to be involved in a pathogenic process. It is interesting to note that *S. sudanensis* lacks the complete *bkd* or *paa* operons within its genome, giving evidence in addition to its small genome size (Table 4.1) that it has an evolutionarily reduced genome, which has previously been observed among bacterial pathogens (Cole et al., 2001, Moran and Plague, 2004, Song et al., 2010). However, exactly how the BGC relates to virulence is difficult to interpret.

The transcriptional regulator *whiD* was also found to be upregulated in *S. sudanensis* (section 4.5.2), as well as a potential two component system (TCS) in the form of the histidine kinase *PABIEPML_04452* and transcriptional regulator *PABIEPML_04454* (section 4.4.2). TCSs and transcriptional regulators can contribute to a pathogen's virulence, as their differential expression allows an organism to sense and adapt to the environmental stresses induced by a host immune response (Geiman et al., 2006, Groisman, 2016). The upregulation of these genes therefore warranted further investigation.

As described in Chapter 4, the homologue of *whiD* in *S. coelicolor* controls sporulation pathways, but *S. sudanensis* has not been observed to sporulate and so the gene's function in this bacterium remains unclear (Molle et al., 2000). A homologue of this gene, *whiB3*, has been found in non-sporulating *Mycobacterium* spp., namely *M. tuberculosis*, *M. marinum* and *M. bovis* (Ramakrishnan et al., 2000, Steyn et al., 2002). Within *M. tuberculosis* specifically, environmental signals, such as acid stress, altered the expression of this regulator *in vivo* (Banaiee et al., 2006, Geiman et al., 2006). More interestingly from the perspective of mycetoma is that the regulator was observed to be preferentially expressed within tuberculosis (TB) granulomas and essential for their formation (Ramakrishnan et al., 2000, Mehta and Singh, 2019). Additionally, *M. tuberculosis* mutants lacking *whiB3* had reduced pathogenicity (Steyn et al., 2002). This is attributed to WhiB3 being a redox sensor and promoting expression of genes to help the bacteria survive oxidative stress upon detection of

ROS and NO produced by immune cells (Saini et al., 2012, Mehta and Singh, 2019). As TB and mycetoma share many features, it can be speculated that WhiD in *S. sudanensis* may function in similar processes to WhiB3 and therefore could be vital for *S. sudanensis* virulence.

BLASTP searches showed the protein products of *PABIEPML_04452* and *PABIEPML_04454* had homology only to other hypothetical histidine kinases and helix-turn-helix proteins in *Streptomyces* spp. There were no characterised homologues to compare them to. There are numerous examples of TCSs being important to the regulation of pathogenicity within bacteria (Ryndak et al., 2008, Jenul and Horswill, 2019, Schaefers, 2020). This potential TCS could be fulfilling a similar role within *S. sudanensis*, allowing it to adapt to and survive the host immune response. Furthermore, the absence of the products of these genes within the non-pathogen *S. coelicolor* suggests this TCS may be associated with pathogenicity. It could thus be involved in the establishment of actinomycetoma infections.

In summary, multiple potential virulence factors were identified within the actinomycetoma pathogen *S. sudanensis*, including a combination of secreted molecules, signalling systems and transcriptional regulators.

6.4 A potential role for inflammasome activation in actinomycetoma pathology

As described in Chapter 3, pyroptosis is an inflammatory form of cell death, that releases pro-inflammatory cytokines and host-derived danger associated molecular patterns (DAMPs) into the surrounding environment (Kelley et al., 2019). The NLRP3 inflammasome is a multimeric protein assembly responsible for detecting the presence of pathogens and initiating pyroptosis to combat the infection. This inflammasome is controlled by a two-signal system, whereby signal one primes the inflammasome, inducing its assembly, and signal two activates it. The NLRP3 inflammasome was observed to be primed by both living and dead *S. sudanensis* (Table 3.7). Further investigation using THP-1 monocytes revealed that sterile supernatant (SN) from *S. sudanensis* culture appeared to induce NLRP3-mediated pyroptosis (section 3.9). Repeated induction of this form of cell death by pathogens could therefore explain the chronic inflammation seen in mycetoma infections. Deliberate activation of a bactericidal pro-inflammatory response by a pathogen does seem counterintuitive at first. However, the literature does provide examples where other pathogens have evolved to do this.

Mycolactone, the toxin produced by *M. ulcerans* and responsible for much of the pathology of Buruli ulcer, has been shown to induce NLRP3 activation *in vitro* (Foulon et al., 2020). This mechanism accounts for its known cytotoxic activity and could also provide an immunosuppressive effect via killing activated immune cells recruited to the infection site. *Staphylococcus aureus* activates NLRP3 via the activity of α-haemolysin, causing necrotic tissue damage in the lung (Kebaier et al., 2012). *Nlrp3*$^{-/-}$ mice on the other hand developed a less severe Staphylococcal pneumonia. *Bacillus cereus* produces another haemolysin, HBL, that also activates NLRP3, leading to rapid death of the host (Mathur et al., 2019). Inhibition of NLRP3 with MCC950 blocked the lethality of the infection. These examples serve to highlight that inappropriate activation of the inflammasome can exacerbate disease phenotypes, via causing tissue damage, and constitute part of disease pathology. Such a mechanism may therefore be being exploited by *S. sudanensis* during actinomycetoma infections. Killing immune cells would favour persistence and thus explain the chronic nature of the infection.

There is another aspect to how inflammasome activation could contribute to chronic disease and benefit the pathogen, which is illustrated by *B. cenocepacia*. This bacterium activates the Pyrin inflammasome via the toxin TecA (Aubert et al., 2016). In the inverse of the above examples however, knocking out the toxin results in increased lethality in mice, as the bacteria are not detected by the inflammasome and so the infection grows unchecked. The authors of this study speculated that toxin production served to allow the pathogen to restrain its own replication, thus keeping the host alive and maintaining the bacteria's replicative niche. Via such a mechanism, the bacteria could then transition from pathogen to commensal. To speculate further beyond this study, controlling growth in this manner would also limit the production of bacterial PAMPs, thus making the pathogen less visible to the immune system. This may then prevent the triggering of a stronger immune response that would fully clear the bacteria. As the disease is chronic, mycetoma pathogens are clearly adept at not killing their hosts while also persisting within them. The dynamics described above could therefore underpin mycetoma pathology.

6.5 Isolation of specialised metabolites from *S. sudanensis* involved in actinomycetoma pathology

With *S. sudanensis* SN demonstrated as having pyroptotic activity, the compounds responsible for this effect needed to be isolated and identified. Through a chemical purification pipeline and bioassays to detect activity, the cyclic dipeptides cyclo(-Leu-Pro) and cyclo(-Val-Pro) were identified as likely candidates responsible for triggering pyroptosis (Chapter 5). A single cyclodipeptide synthase, *PABIEPML_02330*, was also identified in the *S. sudanensis* genome, annotated as encoding a cyclo(-Tyr-Tyr) synthase. While the RNAseq data showed this gene to have an expression level of zero in both the presence and absence of macrophages, it was concluded that the low mapping rates of the bacterial RNAseq data meant false negatives were likely present in the data. It therefore remains a possibility that this gene was expressed and is responsible for synthesis of the two compounds. Any of the four predicted non-ribosomal peptide synthetases within the *S. sudanensis* genome could also be producing these compounds as spontaneous by-products (Belin et al., 2012).

As described in Chapter 5, cyclic dipeptides have a wide range of activities, including: quorum sensing, antibiotic, antiviral, antioxidant, antitumour, anti-inflammatory, and neuroprotective (Bellezza et al., 2014, Chmielewski, 2021). Pyroptotic activity does not appear to have been reported for this class of molecules in the literature. This suggests that while the two compounds identified in this study are not themselves novel, their potential activities may be. However, the mechanism by which these molecules could trigger NLRP3 is unknown. It has previously been found that cyclic dipeptides can interfere with the functioning of ion channels in the plasma membrane of mammalian cells (Milne et al., 1998). NLRP3 can be activated by potassium ion efflux from the cell and so, to purely speculate, perhaps these molecules can bind to potassium ion channels and trigger their opening or alternatively prevent them from closing once opened.

It can be further speculated that if they have this ability, these compounds could also interfere with the generation of action potentials in neurons, contributing to the analgesia experienced by mycetoma patients. The opening of potassium ion channels in a neuron makes its membrane potential far more negative, a process termed hyperpolarisation (Berg et al., 2012). When in this state, the neuron is significantly less sensitive to stimulation, as its greater negative potential means it takes a larger stimulus to reach the threshold for the generation

of an action potential. If the potassium channels were prevented from closing, this desensitisation of the neurons would be prolonged, resulting in an analgesic effect in the host. Such a mechanism could explain the painless lesion development observed in patients with mycetoma. There is at least one precedent in the literature for this mechanism playing a role in Actinobacterial disease. Mycolactone was found to induce hyperpolarisation in neurons, accounting for the analgesia experienced during Buruli ulcer infections (Marion et al., 2014, Song et al., 2017). To reiterate though, in the context of mycetoma this is only speculation and has to be tested by further experiments outside of this study.

6.6 Summary

The work above has successfully provided insight into features of the host response to an actinomycetoma pathogen and potentially uncovered key mechanisms of the disease pathology. Within the indirect interaction assay, a living mycetoma pathogen induced an enhanced pro-inflammatory phenotype in murine macrophages, in comparison to control conditions. This demonstrates that the pathogen produces highly immunogenic antigenic molecules, at least in the initial phase of a mycetoma infection. This inflammatory response included upregulation of MMP expression, which correlated with published patient data. Induction of MMP activity therefore appears to be important to mycetoma development.

The macrophages also showed signs of transitioning to an anti-inflammatory phenotype in the presence of the pathogen, implying that this change may occur in the early stages of a mycetoma infection. This included increased upregulation of receptors for IL4, a cytokine responsible for the initiation of granuloma formation, a key feature of mycetoma pathology. The living mycetoma pathogen also triggered repression of the macrophage mevalonate pathway, indicating a disruption to macrophage cell homeostasis. A reduction in flux through this pathway may have contributed to the enhanced pro-inflammatory phenotype observed. Significantly, the pathogen was found to have the ability to prime and activate the NLRP3 inflammasome, triggering pyroptosis. Dysregulation of inflammasome activation could damage host tissues and kill immune cells, therefore exacerbating mycetoma pathology and playing a key role in its development. It may also promote persistence of the pathogen.

Additionally, multiple potential virulence factors have been identified in the pathogen *S. sudanensis*, via upregulation of genes responsible for their production during the indirect interaction assay. These factors are: pyomelanin, an unknown siderophore produced by BGC 15, a previously unpredicted BGC comprised of *bkd* and *paa* genes, the transcriptional regulator *whiD*, and a potential TCS encoded by *PABIEPML_04452* and *PABIEPML_04454*. Furthermore, two cyclic dipeptides have been isolated from *S. sudanensis*, which initial experiments showed are likely to be responsible for the observed pyroptotic activity. Activation of the NLRP3 inflammasome would represent a novel activity for this class of compounds. Production of these compounds and activation of this inflammasome may therefore represent targets for therapeutic inhibition for the treatment of mycetoma in patients.

In conclusion, a specific macrophage immune response to an actinomycetoma pathogen has been identified and characterised as consisting of an enhanced pro-inflammatory response, mixed with anti-inflammatory phenotypic features. NLRP3-mediated pyroptosis has also been identified as a potential mechanism that promotes the development of actinomycetoma pathology. Potential actinomycetoma virulence factors have also been identified in *S. sudanensis*, including two cyclic dipeptides that may have pyroptotic activity.

6.7 Future work

A key limitation of this study is that the majority of the data for both the macrophages and bacteria was obtained only at the transcriptional level. As post translational regulation of proteins can occur, e.g. proteolytic cleavage of a propeptide to an active form, it remains unknown as to whether many of the transcriptional changes observed were reflected in the proteome. Additional experiments should therefore be performed at the protein level to examine the relevance of key results found in this study.

A further limitation was that the indirect interaction assay was only run for a short period of 48 hrs, meaning that chronic characteristics of the disease were likely unable to develop. An assay method with a much longer incubation time could be optimised in an attempt to model the chronic nature of mycetoma. It is likely that further polarisation towards an anti-inflammatory phenotype could then be observed. Additionally, the assay featured only one

type of immune cell. The majority of complex immune interactions that occur during immune responses to infections were therefore absent from this assay, which likely impacted the phenotypes of the macrophages within the assay. Co-culture of multiple immune cell types within the assay would address this limitation. Alternatively, moving to a whole mouse infection model would provide the full complexity of a complete immune system.

More specific future experiments would include investigating IL4 receptor upregulation in macrophages by adding exogenous IL4 to the interaction assay and comparing any induced effects between pathogen and non-pathogen. Metabolic experiments that specifically disrupt the mevalonate pathway in macrophages could also be performed to define the impact of disrupting this pathway. Measurement of NO levels upon exposure to mycetoma pathogen would also be interesting to perform, in light of reduced *Nos2* upregulation in the presence of *S. sudanensis*.

BGCs of interest should also be knocked out of *S. sudanensis*, to confirm their roles in the response to macrophages. Specifically, *hpd*, cluster 15 and the *bkd-paa* cluster should be targeted. The CRISPR-Cas9 gene editing technique could be used to create such knockouts. Additionally, the siderophore product of cluster 15 needs to be identified, as does a potential product from the *bkd-paa* cluster.

Furthermore, attempts could be made to purify the proteins encoded by *whiD*, *PABIEPML_04452* and *PABIEPML_04454* (the potential TCS). CHIPseq could then be used to locate the DNA binding sites of WhiD and PABIEPML_04454 within the *S. sudanensis* chromosome and thus identify the genes they exert transcriptional control over. This would give insight into the nature of the role these regulators play in the *S. sudanensis* response to immune cells. Additionally, structural studies could be performed on purified PABIEPML_04452 (the sensor histidine kinase) to determine what the ligand of this sensor protein is. This would uncover what environmental stimulus it allows the pathogen to respond to. Alternatively, the potential TCS and *whiD* should be knocked out of *S. sudanensis*. The indirect interaction assay and RNAseq could then be repeated with the mutants generated. This could then allow for the roles of these regulators to be determined, through monitoring how bacterial gene expression changes in their absence.

The CDPS identified in *S. sudanensis* should be heterologously expressed in *S. coelicolor*, to determine whether it is responsible for production of the identified cyclic dipeptides. Further biological repeats of the activity assay with these dipeptides also need to be completed to confirm their pyroptotic activity. This would also include generating additional evidence of NLRP3 activation at the protein level using Western blots and immunofluorescent staining.

If the cyclic dipeptides are confirmed to have pyroptotic activity, their mode of action would then need to be determined. The hypothesis that they interfere with mammalian ion channel opening and closing could be investigated. The patch-clamp technique could be used to measure mammalian cell membrane potentials during exposure to the cyclic dipeptides. If the compounds trigger potassium ion efflux, this could be measured as a change in membrane potential.

This study represents one of the first attempts to simultaneously elucidate the molecular mechanisms underpinning actinomycetoma pathological processes in both the host and a bacterial pathogen. This has been done in a neglected field, where much of the fundamental scientific research to establish and characterise disease models is still ongoing. It has provided novel insights into how macrophages, the core cell type involved in mediating mycetoma pathology through formation of granulomas, respond to the presence of an actinomycetoma pathogen. It has developed an innovative new interaction assay method to allow the disease process to be examined from a unique perspective. Furthermore, it has helped to characterise a bacterial pathogen, *S. sudanensis*, that previously did not even have a complete genome assembly published. Despite its limitations, this study has revealed a number of previously uncharacterised actinomycetoma pathology factors and therefore prospective new targets for drug therapies and preventative treatments to help patients suffering from this neglected disease.

References

ABUSHOUK, A., NASR, A., MASUADI, E., ALLAM, G., SIDDIG, E. E. & FAHAL, A. H. 2019. The Role of Interleukin-1 cytokine family (IL-1beta, IL-37) and interleukin-12 cytokine family (IL-12, IL-35) in eumycetoma infection pathogenesis. *PLoS Negl Trop Dis,* 13**,** e0007098.

ADAMS, D. O. 1974. The structure of mononuclear phagocytes differentiating in vivo. I. Sequential fine and histologic studies of the effect of Bacillus Calmette-Guerin (BCG). *American Journal of Pathology,* 76**,** 17-48.

ADAMS, D. O. 1976. The granulomatous inflammatory response: a review. *American Journal of Pathology,* 84**,** 164-191.

ADUSUMILLI, S., MVE-OBIANG, A., SPARER, T., MEYERS, W., HAYMAN, J. & SMALL, P. L. 2005. Mycobacterium ulcerans toxic macrolide, mycolactone modulates the host immune response and cellular location of M. ulcerans in vitro and in vivo. *Cell Microbiol,* 7**,** 1295-304.

AKIRA, S., ISSHIKI, H., SUGITA, T., TANABE, O., KINOSHITA, S., NISHIO, Y., NAKAJIMA, T., HIRANO, T. & KISHIMOTO, T. 1990. A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. *EMBO Journal,* 9**,** 1897-1906.

AKIRA, S., UEMATSU, S. & TAKEUCHI, O. 2006. Pathogen recognition and innate immunity. *Cell,* 124**,** 783-801.

ALCOVER, A., ALARCON, B. & DI BARTOLO, V. 2018. Cell Biology of T Cell Receptor Expression and Regulation. *Annu Rev Immunol,* 36**,** 103-125.

ALEX, H., SCHERER, A., KREUZBURG, S., ABERS, M. S., ZERBE, C. S., DINAUER, M. C., MANSOUR, M. K. & IRIMIA, D. 2020. Neutrophil swarming delays the growth of clusters of pathogenic fungi. *Nat Commun,* 11**,** 2031.

ALI, R. S., NEWPORT, M. J., BAKHIET, S. M., IBRAHIM, M. E. & FAHAL, A. H. 2020. Host genetic susceptibility to mycetoma. *PLoS Negl Trop Dis,* 14**,** e0008053.

ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. 1990. Basic local alignment search tool. *J Mol Biol,* 215**,** 403-410.

AMBROSIONI, J., LEW, D. & GARBINO, J. 2010. Nocardiosis: updated clinical review and experience at a tertiary center. *Infection,* 38**,** 89-97.

ARAKAWA, K. 2010. Characterization of the Biosynthetic Gene Clusters Located on the Streptomyces Linear Plasmid. *Actinomycetologica,* 24**,** 24-30.

ARASTEHFAR, A., LIM, W., DANESHNIA, F., VAN DE SANDE, W. W. J., FAHAL, A. H., DESNOS-OLLIVIER, M., DE HOOG, G. S., BOEKHOUT, T. & AHMED, S. A. 2020. Madurella real-time PCR, a novel approach for eumycetoma diagnosis. *PLoS Negl Trop Dis,* 14**,** e0007845.

ARAVIND, L., ANANTHARAMAN, V., BALAJI, S., BABU, M. M. & IYER, L. M. 2005. The many faces of the helix-turn-helix domain: transcription regulation and beyond. *FEMS Microbiol Rev,* 29**,** 231-62.

ARENAS, R., FERNANDEZ MARTINEZ, R. F., TORRES-GUERRERO, E. & GARCIA, C. 2017. Actinomycetoma: An Update on Diagnosis and Treatment. *Cutis,* 99**,** e11.

ARIAS-BARRAU, E., OLIVERA, E. R., LUENGO, J. M., FERNANDEZ, C., GALAN, B., GARCIA, J. L., DIAZ, E. & MINAMBRES, B. 2004. The homogentisate pathway: a central catabolic pathway involved in the degradation of L-phenylalanine, L-tyrosine, and 3-hydroxyphenylacetate in Pseudomonas putida. *J Bacteriol,* 186**,** 5062-77.

ARNISON, P. G., BIBB, M. J., BIERBAUM, G., BOWERS, A. A., BUGNI, T. S., BULAJ, G., CAMARERO, J. A., CAMPOPIANO, D. J., CHALLIS, G. L., CLARDY, J., COTTER, P. D., CRAIK, D. J., DAWSON, M., DITTMANN, E., DONADIO, S., DORRESTEIN, P. C., ENTIAN, K. D., FISCHBACH, M. A., GARAVELLI, J. S., GORANSSON, U., GRUBER, C. W., HAFT, D. H., HEMSCHEIDT, T. K., HERTWECK, C., HILL, C., HORSWILL, A. R., JASPARS, M., KELLY, W. L., KLINMAN, J. P., KUIPERS, O. P., LINK, A. J., LIU, W., MARAHIEL, M. A., MITCHELL, D. A., MOLL, G. N., MOORE, B. S., MULLER, R., NAIR, S. K., NES, I. F., NORRIS, G. E., OLIVERA, B. M., ONAKA, H., PATCHETT, M. L., PIEL, J., REANEY, M. J., REBUFFAT, S., ROSS, R. P., SAHL, H. G., SCHMIDT, E. W., SELSTED, M. E., SEVERINOV, K., SHEN,

B., SIVONEN, K., SMITH, L., STEIN, T., SUSSMUTH, R. D., TAGG, J. R., TANG, G. L., TRUMAN, A. W., VEDERAS, J. C., WALSH, C. T., WALTON, J. D., WENZEL, S. C., WILLEY, J. M. & VAN DER DONK, W. A. 2013. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat Prod Rep,* 30**,** 108-60.

ASRAT, S., DAVIS, K. M. & ISBERG, R. R. 2015. Modulation of the host innate immune and inflammatory response by translocated bacterial proteins. *Cell Microbiol,* 17**,** 785-795.

AUBERT, D. F., XU, H., YANG, J., SHI, X., GAO, W., LI, L., BISARO, F., CHEN, S., VALVANO, M. A. & SHAO, F. 2016. A Burkholderia Type VI Effector Deamidates Rho GTPases to Activate the Pyrin Inflammasome and Trigger Inflammation. *Cell Host Microbe,* 19**,** 664-74.

BALTZ, R. H. 2010. Streptomyces and Saccharopolyspora hosts for heterologous expression of secondary metabolite gene clusters. *J Ind Microbiol Biotechnol,* 37**,** 759-72.

BANAIEE, N., JACOBS, W. R., JR. & ERNST, J. D. 2006. Regulation of Mycobacterium tuberculosis whiB3 in the mouse lung and macrophages. *Infect Immun,* 74**,** 6449-57.

BANCHEREAU, J. & STEINMAN, R. M. 1998. Dendritic cells and the control of immunity. *Nature,* 392**,** 245-252.

BARKA, E. A., VATSA, P., SANCHEZ, L., GAVEAU-VAILLANT, N., JACQUARD, C., KLENK, H. P., CLEMENT, C., OUHDOUCH, Y. & VAN WEZEL, G. P. 2016. Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiol Mol Biol Rev,* 80**,** 1-43.

BARONA-GÓMEZ, F., WONG, U., GIANNAKOPULOS, A. E., DERRICK, P. J. & CHALLIS, G. L. 2004. Identification of a Cluster of Genes that Directs Desferrioxamine Biosynthesis in Streptomyces coelicolor M145. *J Am Chem Soc,* 126**,** 16282–16283.

BAUERNFEIND, F. G., HORVATH, G., STUTZ, A., ALNEMRI, E. S., MACDONALD, K., SPEERT, D., FERNANDES-ALNEMRI, T., WU, J., MONKS, B. G., FITZGERALD, K. A., HORNUNG, V. & LATZ, E. 2009. Cutting edge: NF-kappaB activating pattern recognition and cytokine receptors license NLRP3 inflammasome activation by regulating NLRP3 expression. *J Immunol,* 183**,** 787-91.

BEAMER, G. L., FLAHERTY, D. K., ASSOGBA, B. D., STROMBERG, P., GONZALEZ-JUARRERO, M., DE WAAL MALEFYT, R., VESOSKY, B. & TURNER, J. 2008. Interleukin-10 Promotes Mycobacterium tuberculosis Disease Progression in CBA/J Mice. *The Journal of Immunology,* 181**,** 5545-5550.

BECKER, B. & COOPER, M. A. 2013. Aminoglycoside antibiotics in the 21st century. *ACS Chem Biol,* 8**,** 105-15.

BECKWITH, K. S., BECKWITH, M. S., ULLMANN, S., SAETRA, R. S., KIM, H., MARSTAD, A., ASBERG, S. E., STRAND, T. A., HAUG, M., NIEDERWEIS, M., STENMARK, H. A. & FLO, T. H. 2020. Plasma membrane damage causes NLRP3 activation and pyroptosis during Mycobacterium tuberculosis infection. *Nat Commun,* 11**,** 2270.

BELIN, P., MOUTIEZ, M., LAUTRU, S., SEGUIN, J., PERNODET, J. L. & GONDRY, M. 2012. The nonribosomal synthesis of diketopiperazines in tRNA-dependent cyclodipeptide synthase pathways. *Nat Prod Rep,* 29**,** 961-79.

BELLEZZA, I., PEIRCE, M. J. & MINELLI, A. 2014. Cyclic dipeptides: from bugs to brain. *Trends Mol Med,* 20**,** 551-8.

BENTLEY, S. D., CHATER, K. F., CERDENO-TARRAGA, A. M., CHALLIS, G. L., THOMSON, N. R., JAMES, K. D., HARRIS, D. E., QUAIL, M. A., KIESER, H., HARPER, D., BATEMAN, A., BROWN, S., CHANDRA, G., CHEN, C. W., COLLINS, M., CRONIN, A., FRASER, A., GOBLE, A., HIDALGO, J., HORNSBY, T., HOWARTH, S., HUANG, C. H., KIESER, T., LARKE, L., MURPHY, L., OLIVER, K., O'NEIL, S., RABBINOWITSCH, E., RAJANDREAM, M. A., RUTHERFORD, K., RUTTER, S., SEEGER, K., SAUNDERS, D., SHARP, S., SQUARES, R., SQUARES, S., TAYLOR, K., WARREN, T., WIETZORREK, A., WOODWARD, J., BARRELL, B. G., PARKHILL, J. & HOPWOOD, D. A. 2002. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2).

BERG, J. M., TYMOCZKO, J. L. & STRYER, L. 2012. *Biochemistry,* New York, W. H. Freeman Publishing.

BHUIYAN, M. S., ELLETT, F., MURRAY, G. L., KOSTOULIAS, X., CERQUEIRA, G. M., SCHULZE, K. E., MAHAMAD MAIFIAH, M. H., LI, J., CREEK, D. J., LIESCHKE, G. J. & PELEG, A. Y. 2016.

Acinetobacter baumannii phenylacetic acid metabolism influences infection outcome through a direct effect on neutrophil chemotaxis. *Proc Natl Acad Sci U S A,* 113**,** 9599-604.

BIEBER, K. & AUTENRIETH, S. E. 2020. Dendritic cell development in infection. *Mol Immunol,* 121**,** 111-117.

BIORENDER. 2020. *Sanger Sequencing* [Online]. Available: https://app.biorender.com/biorender-templates/t-5ef132f6c7bcd500b388a9c3-sanger-sequencing [Accessed 14/09/2021].

BLIN, K., SHAW, S., KLOOSTERMAN, A. M., CHARLOP-POWERS, Z., VAN WEZEL, G. P., MEDEMA, M. H. & WEBER, T. 2021. antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res,* 49**,** W29-W35.

BLIN, K., WOLF, T., CHEVRETTE, M. G., LU, X., SCHWALEN, C. J., KAUTSAR, S. A., SUAREZ DURAN, H. G., DE LOS SANTOS, E. L. C., KIM, H. U., NAVE, M., DICKSCHAT, J. S., MITCHELL, D. A., SHELEST, E., BREITLING, R., TAKANO, E., LEE, S. Y., WEBER, T. & MEDEMA, M. H. 2017. antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res,* 45**,** W36-W41.

BOGGARAM, V., GOTTIPATI, K. R., WANG, X. & SAMTEN, B. 2013. Early secreted antigenic target of 6 kDa (ESAT-6) protein of Mycobacterium tuberculosis induces interleukin-8 (IL-8) expression in lung epithelial cells via protein kinase signaling and reactive oxygen species. *J Biol Chem,* 288**,** 25500-11.

BOLOGNESE, F., SCANFERLA, C., CARUSO, E. & ORLANDI, V. T. 2019. Bacterial melanin production by heterologous expression of 4 hydroxyphenylpyruvate dioxygenase from Pseudomonas aeruginosa. *Int J Biol Macromol,* 133**,** 1072-1080.

BRINKMAN, A. B., ETTEMA, T. J. G., DE VOS, W. M. & VAN DER OOST, J. 2003. The Lrp family of transcriptional regulators. *Mol Microbiol,* 48**,** 287-294.

BROCK, N. L., NIKOLAY, A. & DICKSCHAT, J. S. 2014. Biosynthesis of the antibiotic tropodithietic acid by the marine bacterium Phaeobacter inhibens. *Chem Commun (Camb),* 50**,** 5487-9.

BUSSI, C. & GUTIERREZ, M. G. 2019. Mycobacterium tuberculosis infection of host cells in space and time. *FEMS Microbiol Rev,* 43**,** 341-361.

CARDENAS-DE LA GARZA, J. A., WELSH, O., CUELLAR-BARBOZA, A., SUAREZ-SANCHEZ, K. P., DE LA CRUZ-VALADEZ, E., CRUZ-GOMEZ, L. G., GALLARDO-ROCHA, A., OCAMPO-CANDIANI, J. & VERA-CABRERA, L. 2020. Clinical characteristics and treatment of actinomycetoma in northeast Mexico: A case series. *PLoS Negl Trop Dis,* 14**,** e0008123.

CHALLIS, G. L. 2014. Exploitation of the Streptomyces coelicolor A3(2) genome sequence for discovery of new natural products and biosynthetic pathways. *J Ind Microbiol Biotechnol,* 41**,** 219-32.

CHANPUT, W., MES, J. J. & WICHERS, H. J. 2014. THP-1 cell line: an in vitro cell model for immune modulation approach. *Int Immunopharmacol,* 23**,** 37-45.

CHATFIELD, C. H. & CIANCIOTTO, N. P. 2007. The secreted pyomelanin pigment of Legionella pneumophila confers ferric reductase activity. *Infect Immun,* 75**,** 4062-70.

CHEN, X., XU, M., LU, J., XU, J., WANG, Y., LIN, S., DENG, Z. & TAO, M. 2018. Biosynthesis of Tropolones in Streptomyces spp.: Interweaving Biosynthesis and Degradation of Phenylacetic Acid and Hydroxylations on the Tropone Ring. *Appl Environ Microbiol,* 84.

CHMIELEWSKI, J. 2021. Emerging designs and applications of peptide-based materials. *Peptide Science,* 113.

COLE, S. T., EIGLMEIER, K., PARKHILL, J., JAMES, K. D., THOMSON, N. R., WHEELER, P. R., HONORÉ, N., GARNIER, T., CHURCHER, C., HARRIS, D., MUNGALL, K., BASHAM, D., BROWN, D., CHILLINGWORTH, T., CONNOR, R., DAVIES, R. M., DEVLIN, K., DUTHOY, S., FELTWELL, T., FRASER, A., HAMLIN, N., HOLROYD, S., HORNSBY, T., JAGELS, K., LACROIX, C., MACLEAN, J., MOULE, S., MURPHY, L., OLIVER, K., QUAIL, M. A., RAJANDREAM, M.-A., RUTHERFORD, K. M., RUTTER, S., SEEGER, K., SIMON, S., SIMMONDS, M., SKELTON, J., SQUARES, R., SQUARES, S., STEVENS, K., TAYLOR, K., WHITEHEAD, S., R., W. J. & BARRELL, B. G. 2001. Massive gene decay in the leprosy bacillus. *Nature,* 409**,** 1007-1011.

COLL, R. C., HILL, J. R., DAY, C. J., ZAMOSHNIKOVA, A., BOUCHER, D., MASSEY, N. L., CHITTY, J. L., FRASER, J. A., JENNINGS, M. P., ROBERTSON, A. A. B. & SCHRODER, K. 2019. MCC950 directly

targets the NLRP3 ATP-hydrolysis motif for inflammasome inhibition. *Nat Chem Biol,* 15**,** 556-559.

COLL, R. C., ROBERTSON, A. A., CHAE, J. J., HIGGINS, S. C., MUNOZ-PLANILLO, R., INSERRA, M. C., VETTER, I., DUNGAN, L. S., MONKS, B. G., STUTZ, A., CROKER, D. E., BUTLER, M. S., HANEKLAUS, M., SUTTON, C. E., NUNEZ, G., LATZ, E., KASTNER, D. L., MILLS, K. H., MASTERS, S. L., SCHRODER, K., COOPER, M. A. & O'NEILL, L. A. 2015. A small-molecule inhibitor of the NLRP3 inflammasome for the treatment of inflammatory diseases. *Nat Med,* 21**,** 248-55.

CORTES, J., HAYDOCK, S. F., ROBERTS, G. A., BEVITT, D. J. & LEADLAY, P. F. 1990. An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of Saccharopolyspora erythraea. *Nature,* 348**,** 176-178.

CREATIVE PROTEOMICS. n.d. *Agilent 6540 UHD Quadrupole Time-of-Flight Accurate-Mass Mass Spectrometer* [Online]. Available: https://www.creative-proteomics.com/support/agilent-6540-uhd-quadrupole-time-of-flight-accurate-mass-mass-spectrometer.htm [Accessed 19/09/2021].

CRONAN, M. R., BEERMAN, R. W., ROSENBERG, A. F., SAELENS, J. W., JOHNSON, M. G., OEHLERS, S. H., SISK, D. M., JURCIC SMITH, K. L., MEDVITZ, N. A., MILLER, S. E., TRINH, L. A., FRASER, S. E., MADDEN, J. F., TURNER, J., STOUT, J. E., LEE, S. & TOBIN, D. M. 2016. Macrophage Epithelial Reprogramming Underlies Mycobacterial Granuloma Formation and Promotes Infection. *Immunity,* 45**,** 861-876.

CRONAN, M. R., HUGHES, E. J., BREWER, W. J., VISWANATHAN, G., HUNT, E. G., SINGH, B., MEHRA, S., OEHLERS, S. H., GREGORY, S. G., KAUSHAL, D. & TOBIN, D. M. 2021. A non-canonical type 2 immune response coordinates tuberculous granuloma formation and epithelialization. *Cell,* 184**,** 1757-1774 e14.

CYKTOR, J. C. & TURNER, J. 2011. Interleukin-10 and immunity against prokaryotic and eukaryotic intracellular pathogens. *Infect Immun,* 79**,** 2964-73.

D'SOUZA, L. & BHATTACHARYA, D. 2019. Plasma cells: You are what you eat. *Immunol Rev,* 288**,** 161-177.

DAIRI, T. 2005. Studies on Biosynthetic Genes and Enzymes of Isoprenoids Produced by Actinomycetes. *J. Antibiot.,* 58**,** 227-243.

DANILCHANKA, O., SUN, J., PAVLENOK, M., MAUERODER, C., SPEER, A., SIROY, A., MARRERO, J., TRUJILLO, C., MAYHEW, D. L., DOORNBOS, K. S., MUNOZ, L. E., HERRMANN, M., EHRT, S., BERENS, C. & NIEDERWEIS, M. 2014. An outer membrane channel protein of Mycobacterium tuberculosis with exotoxin activity. *Proc Natl Acad Sci U S A,* 111**,** 6750-5.

DASHTI, Y., GRKOVIC, T., ABDELMOHSEN, U. R., HENTSCHEL, U. & QUINN, R. J. 2014. Production of induced secondary metabolites by a co-culture of sponge-associated actinomycetes, Actinokineospora sp. EG49 and Nocardiopsis sp. RV163. *Mar Drugs,* 12**,** 3046-59.

DAUM, M., HERRMANN, S., WILKINSON, B. & BECHTHOLD, A. 2009. Genes and enzymes involved in bacterial isoprenoid biosynthesis. *Curr Opin Chem Biol,* 13**,** 180-8.

DAVIS, M. A., FAIRGRIEVE, M. R., DEN HARTIGH, A., YAKOVENKO, O., DUVVURI, B., LOOD, C., THOMAS, W. E., FINK, S. L. & GALE, M., JR. 2019. Calpain drives pyroptotic vimentin cleavage, intermediate filament loss, and cell rupture that mediates immunostimulation. *Proc Natl Acad Sci U S A,* 116**,** 5061-5070.

DE JONG, W. W., LEUNISSEN, J. A. & VOORTER, C. E. 1993. Evolution of the alpha-crystallin/small heat-shock protein family. *Mol Biol Evol,* 10**,** 103-126.

DE LORENZO, V., BINDEREIF, A., PAW, B. H. & NEILANDS, J. B. 1986. Aerobactin biosynthesis and transport genes of plasmid ColV-K30 in Escherichia coli K-12. *J Bacteriol,* 165**,** 570-578.

DE LORENZO, V. & NEILANDS, J. B. 1986. Characterization of iucA and iucC genes of the aerobactin system of plasmid ColV-K30 in Escherichia coli. *J Bacteriol,* 167**,** 350-355.

DE SIMEIS, D. & SERRA, S. 2021. Actinomycetes: A Never-Ending Source of Bioactive Compounds-An Overview on Antibiotics Production. *Antibiotics (Basel),* 10.

DEEM, R. L., DOUGHTY, F. A. & BEAMAN, B. L. 1983. Immunologically specific direct T lymphocyte-mediated killing of Nocardia asteroides. *The Journal of Immunology,* 130**,** 2401-2406.

DELL, M., DUNBAR, K. L. & HERTWECK, C. 2021. Ribosome-independent peptide biosynthesis: the challenge of a unifying nomenclature. *Nat Prod Rep*.

DEMANGEL, C. & HIGH, S. 2018. Sec61 blockade by mycolactone: A central mechanism in Buruli ulcer disease. *Biol Cell,* 110**,** 237-248.

DENOYA, C. D., FEDECHKO, R. W., HAFNER, E. W., MCARTHUR, H. A., MORGENSTERN, M. R., SKINNER, D. D., STUTZMAN-ENGWALL, K., WAX, R. G. & WERNAU, W. C. 1995. A second branched-chain alpha-keto acid dehydrogenase gene cluster (bkdFGH) from Streptomyces avermitilis: its relationship to avermectin biosynthesis and the construction of a bkdF mutant suitable for the production of novel antiparasitic avermectins. . *Journal of Bacteriology,* 177**,** 3505-3511.

DENOYA, C. D., SKINNER, D. D. & MORGERNSTERN, M. R. 1994. A Streptomyces avermitilis Gene Encoding a 4-Hydroxyphenylpyruvic Acid Dioxygenase-Like Protein That Directs the Production of Homogentisic Acid and an Ochronotic Pigment in Escherichia coli. *Journal of Bacteriology,* 176**,** 5312-5319.

DI GENNARO, P., FERRARA, S., RONCO, I., GALLI, E., SELLO, G., PAPACCHINI, M. & BESTETTI, G. 2007. Styrene lower catabolic pathway in Pseudomonas fluorescens ST: identification and characterization of genes for phenylacetic acid degradation. *Arch Microbiol,* 188**,** 117-25.

DI NOIA, J. M. & NEUBERGER, M. S. 2007. Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem,* 76**,** 1-22.

DURR, C., SCHNELL, H. J., LUZHETSKYY, A., MURILLO, R., WEBER, M., WELZEL, K., VENTE, A. & BECHTHOLD, A. 2006. Biosynthesis of the terpene phenalinolactone in Streptomyces sp. Tu6071: analysis of the gene cluster and generation of derivatives. *Chem Biol,* 13**,** 365-77.

EHRENSTEIN, M. R. & NOTLEY, C. A. 2010. The importance of natural IgM: scavenger, protector and regulator. *Nat Rev Immunol,* 10**,** 778-86.

EL HASSAN, A. M., FAHAL, A. H., AHMED, A. O., ISMAIL, A. & VERESS, B. 2001. The immunopathology of actinomycetoma lesions caused by Streptomyces somaliensis. *Transactions of the Royal Society of Tropical Medicines and Hygiene,* 95**,** 88-92.

EMERY, D. & DENNING, D. W. 2020. The global distribution of actinomycetoma and eumycetoma. *PLoS Negl Trop Dis,* 14**,** e0008397.

EMMANUEL, P., DUMRE, S. P., JOHN, S., KARBWANG, J. & HIRAYAMA, K. 2018. Mycetoma: a clinical dilemma in resource limited settings. *Ann Clin Microbiol Antimicrob,* 17**,** 35.

FAHAL, A., EL TOUM, E. A., EL HASSAN, A. M., MAHGOUB, E. S. & GUMAA, S. A. 1995. The host tissue reaction to Madurella mycetomatis: new classification. *Journal of Medical & Veterinary Mycology,* 33**,** 15-17.

FAHAL, A. H. 2017. Mycetoma: A global medical and socio-economic dilemma. *PLoS Negl Trop Dis,* 11**,** e0005509.

FAHAL, A. H., MAHGOUB EL, S., EL HASSAN, A. M. & ABDEL-RAHMAN, M. E. 2015. Mycetoma in the Sudan: an update from the Mycetoma Research Centre, University of Khartoum, Sudan. *PLoS Negl Trop Dis,* 9**,** e0003679.

FELNAGLE, E. A., JACKSON, E. E., CHAN, Y. A., PODEVELS, A. M., BERTI, A. D., MCMAHON, M. D. & THOMAS, M. G. 2008. Nonribosomal peptide synthetases involved in the production of medically relevant natural products. *Mol Pharm,* 5**,** 191-211.

FERRANDEZ, A., MINABRES, B., GARCIA, B., OLIVERA, E. R., LUENGO, J. M., GARCIA, J. L. & DIAZ, E. 1998. Catabolism of Phenylacetic Acid in Escherichia coli: CHARACTERIZATION OF A NEW AEROBIC HYBRID PATHWAY. *The Journal of Biological Chemistry,* 273**,** 25974-25986.

FINKING, R. & MARAHIEL, M. A. 2004. Biosynthesis of nonribosomal peptides. *Annu Rev Microbiol,* 58**,** 453-88.

FISCHBACH, M. A. & WALSH, C. T. 2006. Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms. *Chemical Reviews,* 106**,** 3468-3496.

FITZGERALD, K. A. & KAGAN, J. C. 2020. Toll-like Receptors and the Control of Immunity. *Cell*.

FONSECA, A. B., SIMON, M. D., CAZZANIGA, R. A., DE MOURA, T. R., DE ALMEIDA, R. P., DUTHIE, M. S., REED, S. G. & DE JESUS, A. R. 2017. The influence of innate and adaptative immune responses on the differential clinical outcomes of leprosy. *Infect Dis Poverty,* 6**,** 5.

FOULON, M., ROBBE-SAULE, M., MANRY, J., ESNAULT, L., BOUCAUD, Y., ALCAIS, A., MALLOCI, M., FANTON D'ANDON, M., BEAUVAIS, T., LABARRIERE, N., JEANNIN, P., ABEL, L., SAINT-ANDRE, J. P., CROUE, A., DELNESTE, Y., BONECA, I. G., MARSOLLIER, L. & MARION, E. 2020. Mycolactone toxin induces an inflammatory response by targeting the IL-1beta pathway: Mechanistic insight into Buruli ulcer pathophysiology. *PLoS Pathog,* 16**,** e1009107.

FRIEDRICH, A., PECHSTEIN, J., BERENS, C. & LUHRMANN, A. 2017. Modulation of host cell apoptotic pathways by intracellular pathogens. *Curr Opin Microbiol,* 35**,** 88-99.

FUJITA, T., IKARI, J., WATANABE, A. & TATSUMI, K. 2016. Clinical characteristics of pulmonary nocardiosis in immunocompetent patients. *J Infect Chemother,* 22**,** 738-743.

GADANI, S. P., CRONK, J. C., NORRIS, G. T. & KIPNIS, J. 2012. IL-4 in the brain: a cytokine to remember. *J Immunol,* 189**,** 4213-9.

GAO, B., PARAMANATHAN, R. & GUPTA, R. S. 2006. Signature proteins that are distinctive characteristics of Actinobacteria and their subgroups. *Antonie Van Leeuwenhoek,* 90**,** 69-91.

GEIMAN, D. E., RAGHUNAND, T. R., AGARWAL, N. & BISHAI, W. R. 2006. Differential gene expression in response to exposure to antimycobacterial agents and other stress conditions among seven Mycobacterium tuberculosis whiB-like genes. *Antimicrob Agents Chemother,* 50**,** 2836-41.

GENEUGELIJK, K., KLOEZEN, W., FAHAL, A. H. & VAN DE SANDE, W. W. 2014. Active matrix metalloprotease-9 is associated with the collagen capsule surrounding the Madurella mycetomatis grain in mycetoma. *PLoS Negl Trop Dis,* 8**,** e2754.

GEORGE, K. M., CHATTERJEE, D., GUNAWARDANA, G., WELTY, D., HAYMAN, J., LEE, R. & SMALL, P. L. C. 1999. Mycolactone: A Polyketide Toxin from Mycobacterium ulcerans Required for Virulence. *Science,* 283**,** 854-856.

GILQUIN, J. M., RIVIERE, B., JURADO, V., AUDOUY, B., KOUATCHE, J. B., BERGERON, E., MOUNIEE, D., MOLINA, T., FAURE, P., SAIZ-JIMENEZ, C. & RODRIGUEZ-NAVA, V. 2016. First Case of Actinomycetoma in France Due to a Novel Nocardia Species, Nocardia boironii sp. nov. *mSphere,* 1.

GIORGIO, S., GALLO-FRANCISCO, P. H., ROQUE, G. A. S. & FLORO, E. S. M. 2020. Granulomas in parasitic diseases: the good and the bad. *Parasitol Res,* 119**,** 3165-3180.

GOMEZ-ESCRIBANO, J. P. & BIBB, M. J. 2011. Engineering Streptomyces coelicolor for heterologous expression of secondary metabolite gene clusters. *Microb Biotechnol,* 4**,** 207-15.

GOMEZ-ESCRIBANO, J. P. & BIBB, M. J. 2014. Heterologous expression of natural product biosynthetic gene clusters in Streptomyces coelicolor: from genome mining to manipulation of biosynthetic pathways. *J Ind Microbiol Biotechnol,* 41**,** 425-31.

GOMEZ-ESCRIBANO, J. P., CASTRO, J. F., RAZMILIC, V., CHANDRA, G., ANDREWS, B., ASENJO, J. A. & BIBB, M. J. 2015. The Streptomyces leeuwenhoekii genome: de novo sequencing and assembly in single contigs of the chromosome, circular plasmid pSLE1 and linear plasmid pSLE2. *BMC Genomics,* 16**,** 485.

GONDRY, M., JACQUES, I. B., THAI, R., BABIN, M., CANU, N., SEGUIN, J., BELIN, P., PERNODET, J. L. & MOUTIEZ, M. 2018. A Comprehensive Overview of the Cyclodipeptide Synthase Family Enriched with the Characterization of 32 New Enzymes. *Front Microbiol,* 9**,** 46.

GONZALEZ-SUAREZ, M. L., SALINAS-CARMONA, M. C. & PEREZ-RIVERA, I. 2009. IgM but not IgG monoclonal anti-Nocardia brasiliensis antibodies confer protection against experimental actinomycetoma in BALB/c mice. *FEMS Immunol Med Microbiol,* 57**,** 17-24.

GOPFRICH, K. & JUDGE, K. 2018. Decoding DNA with a pocket-sized sequencer. *Science in School***,** 17-20.

GROISMAN, E. A. 2016. Feedback Control of Two-Component Regulatory Systems. *Annu Rev Microbiol,* 70**,** 103-24.

GRUENBACHER, G. & THURNHER, M. 2017. Mevalonate metabolism governs cancer immune surveillance. *Oncoimmunology,* 6**,** e1342917.

GUENEAU, R., BLANCHET, D., RODRIGUEZ-NAVA, V., BERGERON, E., SOULIER, M., BESTANDJI, N., DEMAR, M., COUPPIE, P. & BLAIZOT, R. 2020. Actinomycetoma caused by Gordonia westfalica: first reported case of human infection. *New Microbes New Infect,* 34**,** 100658.

HAAS, K. M., POE, J. C., STEEBER, D. A. & TEDDER, T. F. 2005. B-1a and B-1b cells exhibit distinct developmental requirements and have unique functional roles in innate and adaptive immunity to S. pneumoniae. *Immunity,* 23**,** 7-18.

HAMS, E., AVIELLO, G. & FALLON, P. G. 2013. The schistosoma granuloma: friend or foe? *Front Immunol,* 4**,** 89.

HAY, R., DENNING, D. W., BONIFAZ, A., QUEIROZ-TELLES, F., BEER, K., BUSTAMANTE, B., CHAKRABARTI, A., CHAVEZ-LOPEZ, M. G., CHILLER, T., CORNET, M., ESTRADA, R., ESTRADA-CHAVEZ, G., FAHAL, A., GOMEZ, B. L., LI, R., MAHABEER, Y., MOSAM, A., SOAVINA RAMAROZATOVO, L., RAKOTO ANDRIANARIVELO, M., RAPELANORO RABENJA, F., VAN DE SANDE, W. & ZIJLSTRA, E. E. 2019. The Diagnosis of Fungal Neglected Tropical Diseases (Fungal NTDs) and the Role of Investigation and Laboratory Tests: An Expert Consensus Report. *Trop Med Infect Dis,* 4.

HAYASHI, Y., MATSUURA, N., TOSHIMA, H., ITOH, N., ISHIKAWA, J., MIKAMI, Y. & DAIRI, T. 2008. Cloning of the Gene Cluster Responsible for the Biosynthesis of Brasilicardin A, a Unique Diterpenoid. *Journal of Antibiotics,* 61**,** 164-174.

HERTWECK, C. 2009. The biosynthetic logic of polyketide diversity. *Angew Chem Int Ed Engl,* 48**,** 4688-716.

HERTWECK, C. 2015. Decoding and reprogramming complex polyketide assembly lines: prospects for synthetic biology. *Trends Biochem Sci,* 40**,** 189-99.

HETRICK, K. J. & VAN DER DONK, W. A. 2017. Ribosomally synthesized and post-translationally modified peptide natural product discovery in the genomic era. *Curr Opin Chem Biol,* 38**,** 36-44.

HOSKISSON, P. A. & SEIPKE, R. F. 2020. Cryptic or Silent? The Known Unknowns, Unknown Knowns, and Unknown Unknowns of Secondary Metabolism. *mBio,* 11.

HOSKISSON, P. A. & VAN WEZEL, G. P. 2019. Streptomyces coelicolor. *Trends Microbiol,* 27**,** 468-469.

HOZUMI, N. & TONEGAWA, S. 1976. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc Natl Acad Sci U S A,* 73**,** 3628-3632.

HUBRICH, F., LOTTI, A., SCOTT, T. A. & PIEL, J. 2021. Uncovering Novel Peptide Chemistry from Bacterial Natural Products. *Chimia (Aarau),* 75**,** 543-547.

IBRAHIM, A. I., EL HASSAN, A. M., FAHAL, A. & VAN DE SANDE, W. W. 2013. A histopathological exploration of the Madurella mycetomatis grain. *PLoS One,* 8**,** e57774.

IKONEN, E. 2008. Cellular cholesterol trafficking and compartmentalization. *Nat Rev Mol Cell Biol,* 9**,** 125-38.

ILLUMINA. 2017. *NovaSeq™ 6000 System Quality Scores and RTA3 Software* [Online]. Available: https://emea.support.illumina.com/sequencing/sequencing_instruments/novaseq-6000/documentation.html [Accessed 15/09/2021].

INVIVOGEN. n.d. *RAW-Blue Cells* [Online]. Available: https://www.invivogen.com/raw-blue [Accessed 16/09/2021].

IWASAKI, A. & MEDZHITOV, R. 2015. Control of adaptive immunity by the innate immune system. *Nat Immunol,* 16**,** 343-53.

JAREMKO, M. J., DAVIS, T. D., CORPUZ, J. C. & BURKART, M. D. 2020. Type II non-ribosomal peptide synthetase proteins: structure, mechanism, and protein-protein interactions. *Nat Prod Rep,* 37**,** 355-379.

JENUL, C. & HORSWILL, A. R. 2019. Regulation of Staphylococcus aureus Virulence. *Microbiol Spectr,* 7.

JORONEN, K., KAHARI, V. M. & VUORIO, E. 2005. Temporospatial expression of matrix metalloproteinases and tissue inhibitors of matrix metalloproteinases in mouse antigen-induced arthritis. *Histochem Cell Biol,* 124**,** 535-45.

KALLIFIDAS, D., JIANG, G., DING, Y. & LUESCH, H. 2018. Rational engineering of Streptomyces albus J1074 for the overexpression of secondary metabolite gene clusters. *Microb Cell Fact,* 17**,** 25.

KARIN, M., LIU, Z. & ZANDI, E. 1997. AP-1 function and regulation. *Current Opinion in Cell Biology,* 9**,** 240-246.

KATZ, L. & BALTZ, R. H. 2016. Natural product discovery: past, present, and future. *J Ind Microbiol Biotechnol,* 43**,** 155-76.

KAUFFMAN, K. D., SALLIN, M. A., SAKAI, S., KAMENYEVA, O., KABAT, J., WEINER, D., SUTPHIN, M., SCHIMEL, D., VIA, L., BARRY, C. E., 3RD, WILDER-KOFIE, T., MOORE, I., MOORE, R. & BARBER, D. L. 2018. Defective positioning in granulomas but not lung-homing limits CD4 T-cell interactions with Mycobacterium tuberculosis-infected macrophages in rhesus macaques. *Mucosal Immunol,* 11**,** 462-473.

KAYE, P. M. & BEATTIE, L. 2016. Lessons from other diseases: granulomatous inflammation in leishmaniasis. *Semin Immunopathol,* 38**,** 249-60.

KEBAIER, C., CHAMBERLAND, R. R., ALLEN, I. C., GAO, X., BROGLIE, P. M., HALL, J. D., JANIA, C., DOERSCHUK, C. M., TILLEY, S. L. & DUNCAN, J. A. 2012. Staphylococcus aureus alpha-hemolysin mediates virulence in a murine model of severe pneumonia through activation of the NLRP3 inflammasome. *J Infect Dis,* 205**,** 807-17.

KELLEY, N., JELTEMA, D., DUAN, Y. & HE, Y. 2019. The NLRP3 Inflammasome: An Overview of Mechanisms of Activation and Regulation. *Int J Mol Sci,* 20.

KENDALL, S. L., BURGESS, P., BALHANA, R., WITHERS, M., TEN BOKUM, A., LOTT, J. S., GAO, C., UHIA-CASTRO, I. & STOKER, N. G. 2010. Cholesterol utilization in mycobacteria is controlled by two TetR-type transcriptional regulators: kstR and kstR2. *Microbiology (Reading),* 156**,** 1362-1371.

KIENLE, K. & LAMMERMANN, T. 2016. Neutrophil swarming: an essential process of the neutrophil tissue response. *Immunol Rev,* 273**,** 76-93.

KIM, E. S. 2021. Recent Advances of Actinomycetes. *Biomolecules,* 11.

KITAGAWA, Y. & SAKAGUCHI, S. 2017. Molecular control of regulatory T cell development and function. *Curr Opin Immunol,* 49**,** 64-70.

KLOSE, C. S. & ARTIS, D. 2016. Innate lymphoid cells as regulators of immunity, inflammation and tissue homeostasis. *Nat Immunol,* 17**,** 765-74.

KNERR, P. J. & VAN DER DONK, W. A. 2012. Discovery, biosynthesis, and engineering of lantipeptides. *Annu Rev Biochem,* 81**,** 479-505.

KOMAKI, H., NEMOTO, A., TANAKA, Y., TAKAGI, H., YAZAWA, K., MIKAMI, Y., SHIGEMORI, H., KOBAYASHI, J., ANDO, A. & NAGATA, Y. 1999. Brasilicardin A, a New Terpenoid Antibiotic from Pathogenic Nocardia brasiliensis: Fermentation, Isolation and Biological Activity. *The Journal of Antibiotics,* 52**,** 13-19.

KOMATSU, K., TSUDA, M., TANAKA, Y., MIKAMI, Y. & KOBAYASHI, J. 2005. SAR studies of brasilicardin A for immunosuppressive and cytotoxic activities. *Bioorg Med Chem,* 13**,** 1507-13.

KONIG, A., MULLER, R., MOGAVERO, S. & HUBE, B. 2021. Fungal factors involved in host immune evasion, modulation and exploitation during infection. *Cell Microbiol,* 23**,** e13272.

KONONEN, E. & WADE, W. G. 2015. Actinomyces and related organisms in human infections. *Clin Microbiol Rev,* 28**,** 419-42.

KOREN, S., WALENZ, B. P., BERLIN, K., MILLER, J. R., BERGMAN, N. H. & PHILLIPPY, A. M. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res,* 27**,** 722-736.

KRAMER, J., OZKAYA, O. & KUMMERLI, R. 2020. Bacterial siderophores in community and host interactions. *Nat Rev Microbiol,* 18**,** 152-163.

KRUGEL, H., KRUBASIK, P., WEBER, K., SALUZ, H. P. & SANDMANN, G. 1999. Functional analysis of genes from Streptomyces griseus involved in the synthesis of isorenieratene, a carotenoid with aromatic end groups, revealed a novel type of carotenoid desaturase. *Biochimica et Biophysica Acta,* 1439**,** 57-64.

KUDO, F. & EGUCHI, T. 2009. Biosynthetic genes for aminoglycoside antibiotics. *J Antibiot (Tokyo),* 62**,** 471-81.

KUDO, F. & EGUCHI, T. 2016. Aminoglycoside Antibiotics: New Insights into the Biosynthetic Machinery of Old Drugs. *Chem Rec,* 16**,** 4-18.

KUMAR, B. V., CONNORS, T. J. & FARBER, D. L. 2018. Human T Cell Development, Localization, and Function throughout Life. *Immunity,* 48**,** 202-213.

LAND, M., HAUSER, L., JUN, S. R., NOOKAEW, I., LEUZE, M. R., AHN, T. H., KARPINETS, T., LUND, O., KORA, G., WASSENAAR, T., POUDEL, S. & USSERY, D. W. 2015. Insights from 20 years of bacterial genome sequencing. *Funct Integr Genomics,* 15**,** 141-61.

LANGMEAD, B. & SALZBERG, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods,* 9**,** 357-9.

LATIF, Z. & SARKER, S. D. 2012. Isolation of natural products by preparative high performance liquid chromatography (prep-HPLC). *Methods Mol Biol,* 864**,** 255-74.

LAW, R. J., HAMLIN, J. N., SIVRO, A., MCCORRISTER, S. J., CARDAMA, G. A. & CARDONA, S. T. 2008. A functional phenylacetic acid catabolic pathway is required for full pathogenicity of Burkholderia cenocepacia in the Caenorhabditis elegans host model. *J Bacteriol,* 190**,** 7209-18.

LEE, N., KIM, W., HWANG, S., LEE, Y., CHO, S., PALSSON, B. & CHO, B. K. 2020. Thirty complete Streptomyces genome sequences for mining novel secondary metabolite biosynthetic gene clusters. *Sci Data,* 7**,** 55.

LIANG, W., ZHANG, C., LIU, N., ZHANG, W., HAN, Q. & LI, C. 2016. Cloning and characterization of Vshppd, a gene inducing haemolysis and immune response of Apostichopus japonicus. *Aquaculture,* 464**,** 246-252.

LIANG, W., ZHANG, W., SHAO, Y., ZHAO, X. & LI, C. 2018. Dual functions of a 4-hydroxyphenylpyruvate dioxygenase for Vibrio splendidus survival and infection. *Microb Pathog,* 120**,** 47-54.

LIM, W., PAREL, F., DE HOOG, S., VERBON, A. & VAN DE SANDE, W. W. J. 2021. Melanin production in coelomycetous agents of black grain eumycetoma. *Trans R Soc Trop Med Hyg*.

LIU, Y., ZHOU, J. & WHITE, K. P. 2014. RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics,* 30**,** 301-4.

LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol,* 15**,** 550.

LU, L. L., SUSCOVICH, T. J., FORTUNE, S. M. & ALTER, G. 2018. Beyond binding: antibody effector functions in infectious diseases. *Nat Rev Immunol,* 18**,** 46-61.

MACH, B., REICH, E. & TATUM, E. L. 1963. Separation of the biosynthesis of the antibiotic polypeptide tyrocidine from protein biosynthesis. *Proc Natl Acad Sci U S A,* 50**,** 175-181.

MACMICKING, J., XIE, Q.-W. & NATHAN, C. 1997. Nitric oxide and macrophage function. *Annual Review of Immunology,* 15**,** 323-350.

MADEIRA, F., PARK, Y. M., LEE, J., BUSO, N., GUR, T., MADHUSOODANAN, N., BASUTKAR, P., TIVEY, A. R. N., POTTER, S. C., FINN, R. D. & LOPEZ, R. 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res,* 47**,** W636-W641.

MADHUSUDHAN, K. T., LORENZ, D. & SOKATCH, J. R. 1993. The bkdR gene of Pseudomonas putida is required for expression of the bkd operon and encodes a protein related to Lrp of Escherichia coli. *J Bacteriol,* 175.

MADIGAN, C. A., CAMBIER, C. J., KELLY-SCUMPIA, K. M., SCUMPIA, P. O., CHENG, T. Y., ZAILAA, J., BLOOM, B. R., MOODY, D. B., SMALE, S. T., SAGASTI, A., MODLIN, R. L. & RAMAKRISHNAN, L. 2017. A Macrophage Response to Mycobacterium leprae Phenolic Glycolipid Initiates Nerve Damage in Leprosy. *Cell,* 170**,** 973-985 e10.

MALICO, A. A., CALZINI, M. A., GAYEN, A. K. & WILLIAMS, G. J. 2020. Synthetic biology, combinatorial biosynthesis, and chemoenzymatic synthesis of isoprenoids. *J Ind Microbiol Biotechnol,* 47**,** 675-702.

MARAKALALA, M. J., RAJU, R. M., SHARMA, K., ZHANG, Y. J., EUGENIN, E. A., PRIDEAUX, B., DAUDELIN, I. B., CHEN, P. Y., BOOTY, M. G., KIM, J. H., EUM, S. Y., VIA, L. E., BEHAR, S. M., BARRY, C. E., 3RD, MANN, M., DARTOIS, V. & RUBIN, E. J. 2016. Inflammatory signaling in human tuberculosis granulomas is spatially organized. *Nat Med,* 22**,** 531-8.

MARIATHASAN, S., WEISS, D. S., NEWTON, K., MCBRIDE, J., O'ROURKE, K., ROOSE-GIRMA, M., LEE, W. P., WEINRAUCH, Y., MONACK, D. M. & DIXIT, V. M. 2006. Cryopyrin activates the inflammasome in response to toxins and ATP. *Nature,* 440**,** 228-32.

MARION, E., SONG, O. R., CHRISTOPHE, T., BABONNEAU, J., FENISTEIN, D., EYER, J., LETOURNEL, F., HENRION, D., CLERE, N., PAILLE, V., GUERINEAU, N. C., SAINT ANDRE, J. P., GERSBACH, P.,

ALTMANN, K. H., STINEAR, T. P., COMOGLIO, Y., SANDOZ, G., PREISSER, L., DELNESTE, Y., YERAMIAN, E., MARSOLLIER, L. & BRODIN, P. 2014. Mycobacterial toxin induces analgesia in buruli ulcer by targeting the angiotensin pathways. *Cell,* 157**,** 1565-76.

MARTIN, J. F. 2004. Phosphate control of the biosynthesis of antibiotics and other secondary metabolites is mediated by the PhoR-PhoP system: an unfinished story. *J Bacteriol,* 186**,** 5197-201.

MATHUR, A., FENG, S., HAYWARD, J. A., NGO, C., FOX, D., ATMOSUKARTO, II, PRICE, J. D., SCHAUER, K., MARTLBAUER, E., ROBERTSON, A. A. B., BURGIO, G., FOX, E. M., LEPPLA, S. H., KAAKOUSH, N. O. & MAN, S. M. 2019. A multicomponent toxin from Bacillus cereus incites inflammation and shapes host outcome via the NLRP3 inflammasome. *Nat Microbiol,* 4**,** 362-374.

MCCOMBIE, W. R., MCPHERSON, J. D. & MARDIS, E. R. 2019. Next-Generation Sequencing Technologies. *Cold Spring Harb Perspect Med,* 9.

MCWHORTER, F. Y., WANG, T., NGUYEN, P., CHUNG, T. & LIU, W. F. 2013. Modulation of macrophage phenotype by cell shape. *Proc Natl Acad Sci U S A,* 110**,** 17253-8.

MEHTA, M. & SINGH, A. 2019. Mycobacterium tuberculosis WhiB3 maintains redox homeostasis and survival in response to reactive oxygen and nitrogen species. *Free Radic Biol Med,* 131**,** 50-58.

MENDEZ-TOVAR, L. J., MONDRAGON-GONZALEZ, R., VEGA-LOPEZ, F., DOCKRELL, H. M., HAY, R., LOPEZ-MARTINEZ, R., MANZANO-GAYOSSO, P., HERNANDEZ-HERNANDEZ, F., PADILLA-DESGARENNES, C. & BONIFAZ, A. 2004. Cytokine production and lymphocyte proliferation in patients with Nocardia brasiliensis actinomycetoma. *Mycopathologia,* 158**,** 407-414.

MERCIER, R., DOMINGUEZ-CUEVAS, P. & ERRINGTON, J. 2012. Crucial role for membrane fluidity in proliferation of primitive cells. *Cell Rep,* 1**,** 417-23.

METZKER, M. L. 2010. Sequencing technologies - the next generation. *Nat Rev Genet,* 11**,** 31-46.

MHMOUD, N. A., FAHAL, A. H. & VAN DE SANDE, W. W. 2013. The association between the interleukin-10 cytokine and CC chemokine ligand 5 polymorphisms and mycetoma granuloma formation. *Med Mycol,* 51**,** 527-33.

MI, Z., LIU, H. & ZHANG, F. 2020. Advances in the Immunology and Genetics of Leprosy. *Frontiers in Immunology,* 11.

MIETTO, B. S., DE SOUZA, B. J., ROSA, P. S., PESSOLANI, M. C. V., LARA, F. A. & SARNO, E. N. 2020. Myelin breakdown favours Mycobacterium leprae survival in Schwann cells. *Cell Microbiol,* 22**,** e13128.

MILLAN-CHIU, B. E., HERNANDEZ-HERNANDEZ, F., PEREZ-TORRES, A., MENDEZ-TOVAR, L. J. & LOPEZ-MARTINEZ, R. 2011. In situ TLR2 and TLR4 expression in a murine model of mycetoma caused by Nocardia brasiliensis. *FEMS Immunol Med Microbiol,* 61**,** 278-87.

MILNE, P. J., HUNT, A. L., ROSTOLL, K., VAN DER WALT, J. J. & GRAZ, C. J. M. 1998. The Biological Activity of Selected Cyclic Dipeptides. *J Pharm Pharmacol,* 50**,** 1331-1337.

MOLLE, V., PALFRAMAN, W. J., FINDLAY, K. C. & BUTTNER, M. J. 2000. WhiD and WhiB, homologous proteins required for different stages of sporulation in Streptomyces coelicolor A3(2). *Journal of Bacteriology,* 182**,** 1286-1295.

MORAN, G. R. 2005. 4-Hydroxyphenylpyruvate dioxygenase. *Arch Biochem Biophys,* 433**,** 117-28.

MORAN, N. A. & PLAGUE, G. R. 2004. Genomic changes following host restriction in bacteria. *Curr Opin Genet Dev,* 14**,** 627-33.

MORTAZAVI, A., WILLIAMS, B. A., MCCUE, K., SCHAEFFER, L. & WOLD, B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods,* 5**,** 621-8.

MOSMANN, T. R. & SAD, S. 1996. The expanding universe of T-cell subsets: Thl, Th2 and more. *Immunology Today,* 17**,** 138-146.

MOSSER, D. M. & EDWARDS, J. P. 2008. Exploring the full spectrum of macrophage activation. *Nat Rev Immunol,* 8**,** 958-69.

MUSTHAFA, K. S., SIVAMARUTHI, B. S., PANDIAN, S. K. & RAVI, A. V. 2012. Quorum sensing inhibition in Pseudomonas aeruginosa PAO1 by antagonistic compound phenylacetic acid. *Curr Microbiol,* 65**,** 475-80.

NAGALAKSHMI, U., WANG, Z., WAERN, K., SHOU, C., RAHA, D., GERSTEIN, M. & SNYDER, M. 2008. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science,* 320**,** 1344-9.

NAGWA. n.d. *Lesson Explainer: Components of the Human Immune System* [Online]. Available: https://www.nagwa.com/en/explainers/649121945685/ [Accessed 14/09/2021 2021].

NASR, A., ABUSHOUK, A., HAMZA, A., SIDDIG, E. & FAHAL, A. H. 2016. Th-1, Th-2 Cytokines Profile among Madurella mycetomatis Eumycetoma Patients. *PLoS Negl Trop Dis,* 10**,** e0004862.

NAVARRO-LLORENS, J. M., PATRAUCHAN, M. A., STEWART, G. R., DAVIES, J. E., ELTIS, L. D. & MOHN, W. W. 2005. Phenylacetate catabolism in Rhodococcus sp. strain RHA1: a central pathway for degradation of aromatic compounds. *J Bacteriol,* 187**,** 4497-504.

NCBI. n.d.-a. *Genome* [Online]. Available: https://www.ncbi.nlm.nih.gov/genome [Accessed 15/09/2021].

NCBI. n.d.-b. *National Center for Biotechnology Information Genome Browser* [Online]. Available: https://www.ncbi.nlm.nih.gov/genome/browse#!/prokaryotes/ [Accessed 15/09/2021].

NEEFJES, J., JONGSMA, M. L., PAUL, P. & BAKKE, O. 2011. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol,* 11**,** 823-36.

NISSINEN, L. & KAHARI, V. M. 2014. Matrix metalloproteinases in inflammation. *Biochim Biophys Acta,* 1840**,** 2571-80.

O'NEILL, L. A. & PEARCE, E. J. 2016. Immunometabolism governs dendritic cell and macrophage function. *J Exp Med,* 213**,** 15-23.

O'NEILL, L. A. J. & ARTYOMOV, M. N. 2019. Itaconate: the poster child of metabolic reprogramming in macrophage function. *Nat Rev Immunol,* 19**,** 273-281.

OETTINGER, M. A., SCHATZ, D. G., GORKA, C. & BALTIMORE, D. 1990. RAG-1 and RAG-2, Adjacent Genes That Synergistically Activate V(D)J Recombination. *Science,* 248**,** 1517-153.

OLANO, C., GARCIA, I., GONZALEZ, A., RODRIGUEZ, M., ROZAS, D., RUBIO, J., SANCHEZ-HIDALGO, M., BRANA, A. F., MENDEZ, C. & SALAS, J. A. 2014. Activation and identification of five clusters for secondary metabolites in Streptomyces albus J1074. *Microb Biotechnol,* 7**,** 242-56.

OLLINGER, J., SONG, K. B., ANTELMANN, H., HECKER, M. & HELMANN, J. D. 2006. Role of the Fur regulon in iron transport in Bacillus subtilis. *J Bacteriol,* 188**,** 3664-73.

ŌMURA, S., IKEDA, H., ISHIKAWA, J., HANAMOTO, A., TAKAHASHI, C., SHINOSE, M., TAKAHASHI, Y., HORIKAWA, H., NAKAZAWA, H., OSONOE, T., KIKUCHI, H., SHIBA, T., SAKAKI, Y. & HATTORI, M. 2001. Genome sequence of an industrial microorganism Streptomyces avermitilis: Deducing the ability of producing secondary metabolites. *PNAS,* 98.

ORTEGA, M. A. & VAN DER DONK, W. A. 2016. New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products. *Cell Chem Biol,* 23**,** 31-44.

PAGAN, A. J. & RAMAKRISHNAN, L. 2018. The Formation and Function of Granulomas. *Annu Rev Immunol,* 36**,** 639-665.

PALMER, L. D. & SKAAR, E. P. 2016. Transition Metals and Virulence in Bacteria. *Annu Rev Genet,* 50**,** 67-91.

PAPAYANNOPOULOS, V. 2018. Neutrophil extracellular traps in immunity and disease. *Nat Rev Immunol,* 18**,** 134-147.

PARISI, L., GINI, E., BACI, D., TREMOLATI, M., FANULI, M., BASSANI, B., FARRONATO, G., BRUNO, A. & MORTARA, L. 2018. Macrophage Polarization in Chronic Inflammatory Diseases: Killers or Builders? *J Immunol Res,* 2018**,** 8917804.

PARK, J. W., BAN, Y. H., NAM, S. J., CHA, S. S. & YOON, Y. J. 2017. Biosynthetic pathways of aminoglycosides and their engineering. *Curr Opin Biotechnol,* 48**,** 33-41.

PATEL, D. D. & KUCHROO, V. K. 2015. Th17 Cell Pathway in Human Immunity: Lessons from Genetics and Therapeutic Interventions. *Immunity,* 43**,** 1040-51.

PATRO, R., DUGGAL, G., LOVE, M. I., IRIZARRY, R. A. & KINGSFORD, C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods,* 14**,** 417-419.

PAVIA, D. L., LAMPMAN, G. M., KRIZ, G. S. & VYVYAN, J. R. 2015. *Introduction to Spectroscopy,* United States of America, Cengage Learning.

PAYNE, J. A., SCHOPPET, M., HANSEN, M. H. & CRYLE, M. J. 2016. Diversity of nature's assembly lines - recent discoveries in non-ribosomal peptide synthesis. *Mol Biosyst,* 13**,** 9-22.

PELEGRIN, P., BARROSO-GUTIERREZ, C. & SURPRENANT, A. 2008. P2X7 Receptor Differentially Couples to Distinct Release Pathways for IL-1β in Mouse Macrophage. *The Journal of Immunology,* 180**,** 7147-7157.

PEYRON, P., VAUBOURGEIX, J., POQUET, Y., LEVILLAIN, F., BOTANCH, C., BARDOU, F., DAFFE, M., EMILE, J. F., MARCHOU, B., CARDONA, P. J., DE CHASTELLIER, C. & ALTARE, F. 2008. Foamy macrophages from tuberculous patients' granulomas constitute a nutrient-rich reservoir for M. tuberculosis persistence. *PLoS Pathog,* 4**,** e1000204.

PFEIFFER, F., GROBER, C., BLANK, M., HANDLER, K., BEYER, M., SCHULTZE, J. L. & MAYER, G. 2018. Systematic evaluation of error rates and causes in short samples in next-generation sequencing. *Sci Rep,* 8**,** 10950.

PITT, J. J. 2009. Principles and Applications of Liquid Chromatography-Mass Spectrometry in Clinical Biochemistry. *Clinical Biochemistry Reviews,* 30**,** 19-34.

PLONKA, P. M. & GRABACKA, M. 2006. Melanin synthesis in microorganisms - biotechnological and medical aspects. *Acta Biochimica Polonica,* 53**,** 429-443.

POOLE, R. K. 2020. Flavohaemoglobin: the pre-eminent nitric oxide-detoxifying machine of microorganisms. *F1000Res,* 9.

PRABHU, J., SCHAUWECKER, F., GRAMMEL, N., KELLER, U. & BERNHARD, M. 2004. Functional expression of the ectoine hydroxylase gene (thpD) from Streptomyces chrysomallus in Halomonas elongata. *Appl Environ Microbiol,* 70**,** 3130-2.

PRIBYTKOVA, T., LIGHTLY, T. J., KUMAR, B., BERNIER, S. P., SORENSEN, J. L., SURETTE, M. G. & CARDONA, S. T. 2014. The attenuated virulence of a Burkholderia cenocepacia paaABCDE mutant is due to inhibition of quorum sensing by release of phenylacetic acid. *Mol Microbiol,* 94**,** 522-36.

PRIETO, J. M. B. & FELIPPE, M. J. B. 2017. Development, phenotype, and function of non-conventional B cells. *Comp Immunol Microbiol Infect Dis,* 54**,** 38-44.

PULSAWAT, N., KITANI, S., KINOSHITA, H., LEE, C. K. & NIHIRA, T. 2007. Identification of the bkdAB gene cluster, a plausible source of the starter-unit for virginiamycin M production in Streptomyces virginiae. *Arch Microbiol,* 187**,** 459-66.

QUIDING-JARBRINK, M., SMITH, D. A. & BANCROFT, G. J. 2001. Production of matrix metalloproteinases in response to mycobacterial infection. *Infect Immun,* 69**,** 5661-70.

RACINE, R., MCLAUGHLIN, M., JONES, D. D., WITTMER, S. T., MACNAMARA, K. C., WOODLAND, D. L. & WINSLOW, G. M. 2011. IgM production by bone marrow plasmablasts contributes to long-term protection against intracellular bacterial infection. *J Immunol,* 186**,** 1011-21.

RAHMAN, M. M. & MCFADDEN, G. 2011. Modulation of NF-kappaB signalling by microbial pathogens. *Nat Rev Microbiol,* 9**,** 291-306.

RAMAKRISHNAN, L. 2012. Revisiting the role of the granuloma in tuberculosis. *Nat Rev Immunol,* 12**,** 352-66.

RAMAKRISHNAN, L., FEDERSPIEL, N. A. & FALKOW, S. 2000. Granuloma-Specific Expression of Mycobacterium Virulence Proteins from the Glycine-Rich PE-PGRS Family. *Science,* 288**,** 1436-1439.

RDEST, U., WINTERMEYER, E., LUDWIG, B. & HACKER, J. 1991. Legiolysin, a New Hemolysin from L. pneumophila. *Zbl. Bakt.,* 274**,** 471-474.

REFAI, A., GRITLI, S., BARBOUCHE, M. R. & ESSAFI, M. 2018. Mycobacterium tuberculosis Virulent Factor ESAT-6 Drives Macrophage Differentiation Toward the Pro-inflammatory M1 Phenotype and Subsequently Switches It to the Anti-inflammatory M2 Phenotype. *Front Cell Infect Microbiol,* 8**,** 327.

REIS, E. S., MASTELLOS, D. C., HAJISHENGALLIS, G. & LAMBRIS, J. D. 2019. New insights into the immune functions of complement. *Nat Rev Immunol,* 19**,** 503-516.

RHOADS, A. & AU, K. F. 2015. PacBio Sequencing and Its Applications. *Genomics Proteomics Bioinformatics,* 13**,** 278-89.

RICO, G., OCHOA, R., OLIVA, A., GONZALEZ-MENDOZA, A., WALKER, S. M. & ORTIZ-ORTIZ, L. 1982. Enhanced resistance to Nocardia brasiliensis infection in mice depleted of antigen-specific B cells. *The Journal of Immunology,* 129**,** 1688-1693.

ROCA, F. J. & RAMAKRISHNAN, L. 2013. TNF dually mediates resistance and susceptibility to mycobacteria via mitochondrial reactive oxygen species. *Cell,* 153**,** 521-34.

RODIONOV, D. A., DUBCHAK, I. L., ARKIN, A. P., ALM, E. J. & GELFAND, M. S. 2005. Dissimilatory metabolism of nitrogen oxides in bacteria: comparative reconstruction of transcriptional networks. *PLoS Comput Biol,* 1**,** e55.

ROSALES, C. 2020. Neutrophils at the crossroads of innate and adaptive immunity. *J Leukoc Biol*.

ROSSI, M. & FASEL, N. 2018. How to master the host immune system? Leishmania parasites have the solutions! *Int Immunol,* 30**,** 103-111.

RUDD, B. A. M. & HOPWOOD, D. A. 1979. Genetics of Actinorhodin Biosynthesis by Streptomyces coelicolor A3 (2). *Journal of General Microbiology,* 114**,** 35-43.

RUDD, B. A. M. & HOPWOOD, D. A. 1980. A Pigmented Mycelial Antibiotic in Streptomyces coelicolor : Control by a Chromosomal Gene Cluster. *Journal of General Microbiology,* 119**,** 333-340.

RUF, M. T., STEFFEN, C., BOLZ, M., SCHMID, P. & PLUSCHKE, G. 2017. Infiltrating leukocytes surround early Buruli ulcer lesions, but are unable to reach the mycolactone producing mycobacteria. *Virulence,* 8**,** 1918-1926.

RYNDAK, M., WANG, S. & SMITH, I. 2008. PhoP, a key player in Mycobacterium tuberculosis virulence. *Trends Microbiol,* 16**,** 528-34.

SAINI, V., FARHANA, A. & STEYN, A. J. 2012. Mycobacterium tuberculosis WhiB3: a novel iron-sulfur cluster protein that regulates redox homeostasis and virulence. *Antioxid Redox Signal,* 16**,** 687-97.

SALINAS-CARMONA, M. C. & PEREZ-RIVERA, I. 2004. Humoral immunity through immunoglobulin M protects mice from an experimental actinomycetoma infection by Nocardia brasiliensis. *Infect Immun,* 72**,** 5597-604.

SALINAS-CARMONA, M. C., ROSAS-TARACO, A. G. & WELSH, O. 2012. Systemic increased immune response to Nocardia brasiliensis co-exists with local immunosuppressive microenvironment. *Antonie Van Leeuwenhoek,* 102**,** 473-80.

SALINAS CARMONA, M. C., WELSH, O. & CASILLAS, S. M. 1993. Enzyme-Linked Immunosorbent Assay for Serological Diagnosis of Nocardia brasiliensis and Clinical Correlation with Mycetoma Infections. *Journal of Clinical Microbiology,* 31**,** 2901-2906.

SANTIAGO-TELLEZ, A., CASTRILLON-RIVERA, L. E., PALMA-RAMOS, A., BELLO-LOPEZ, J. M., SAINZ-ESPUNES, T., CONTRERAS-PAREDES, A., LUNA-HERRERA, J. & CASTANEDA-SANCHEZ, J. I. 2019. Keratinocyte infection by Actinomadura madurae triggers an inflammatory response. *Trans R Soc Trop Med Hyg,* 113**,** 392-398.

SCHAEFERS, M. M. 2020. Regulation of Virulence by Two-Component Systems in Pathogenic Burkholderia. *Infection and Immunity,* 88**,** e00927-19.

SCHATZ, A., BUGIE, E. & WAKSMAN, S. 1944. Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria. *Proceedings of the Society for Experimental Biology and Medicine,* 55**,** 66-69.

SCHMIDTZ, F. J., VANDERAH, D. J., HOLLENBEAK, K. H., ENWALL, C. E. L., GOPICHAND, Y., SENGUPTA, P. K., HOSSAIN, M. B. & VAN DER HELM, D. 1983. Metabolites from the marine sponge Tedania ignis. A new atisanediol and several known diketopiperazines. *J. Org. Chem.,* 48**,** 3941-3945.

SCHWARZ, P. N., BUCHMANN, A., ROLLER, L., KULIK, A., GROSS, H., WOHLLEBEN, W. & STEGMANN, E. 2018. The Immunosuppressant Brasilicardin: Determination of the Biosynthetic Gene Cluster in the Heterologous Host Amycolatopsis japonicum. *Biotechnol J,* 13.

SCOLLARD, D. M., DACSO, M. M. & ABAD-VENIDA, M. L. 2015. Tuberculosis and Leprosy: Classical Granulomatous Diseases in the Twenty-First Century. *Dermatol Clin,* 33**,** 541-62.

SEEMANN, T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics,* 30**,** 2068-2069.

SERRANO-COLL, H., SALAZAR-PELAEZ, L., ACEVEDO-SAENZ, L. & CARDONA-CASTRO, N. 2018. Mycobacterium leprae-induced nerve damage: direct and indirect mechanisms. *Pathog Dis,* 76.

SHEEHAN, G., KONINGS, M., LIM, W., FAHAL, A., KAVANAGH, K. & VAN DE SANDE, W. W. J. 2020. Proteomic analysis of the processes leading to Madurella mycetomatis grain formation in Galleria mellonella larvae. *PLoS Negl Trop Dis,* 14**,** e0008190.

SHI, L., JIANG, Q., BUSHKIN, Y., SUBBIAN, S. & TYAGI, S. 2019. Biphasic Dynamics of Macrophage Immunometabolism during Mycobacterium tuberculosis Infection. *mBio,* 10**,** e02550-18.

SHIGEMORI, H., KOMAKI, H., YAZAWA, K., MIKAMI, Y., NEMOTO, A., TANAKA, Y., SASAKI, T., IN, Y., ISHIDA, T. & KOBAYASHI, J. 1998. Brasilicardin A. A Novel Tricyclic Metabolite with Potent Immunosuppressive Activity from Actinomycete Nocardia brasiliensis. *Journal of Organic Chemistry,* 63**,** 6900-6904.

SHIMIZU, Y., OGATA, H. & GOTO, S. 2017. Type III Polyketide Synthases: Functional Classification and Phylogenomics. *Chembiochem,* 18**,** 50-65.

SIDDIG, E. E., MHMOUD, N. A., BAKHIET, S. M., ABDALLAH, O. B., MEKKI, S. O., EL DAWI, N. I., VAN DE SANDE, W. & FAHAL, A. H. 2019a. The Accuracy of Histopathological and Cytopathological Techniques in the Identification of the Mycetoma Causative Agents. *PLoS Negl Trop Dis,* 13**,** e0007056.

SIDDIG, E. E., MOHAMMED EDRIS, A. M., BAKHIET, S. M., VAN DE SANDE, W. W. J. & FAHAL, A. H. 2019b. Interleukin-17 and matrix metalloprotease-9 expression in the mycetoma granuloma. *PLoS Negl Trop Dis,* 13**,** e0007351.

SIEGELE, D. A. 2005. Universal stress proteins in Escherichia coli. *J Bacteriol,* 187**,** 6253-4.

SILVA, M. T., PORTAELS, F. & PEDROSA, J. 2009. Pathogenetic mechanisms of the intracellular parasite Mycobacterium ulcerans leading to Buruli ulcer. *The Lancet Infectious Diseases,* 9**,** 699-710.

SILVIS, M. J. M., DEMKES, E. J., FIOLET, A. T. L., DEKKER, M., BOSCH, L., VAN HOUT, G. P. J., TIMMERS, L. & DE KLEIJN, D. P. V. 2021. Immunomodulation of the NLRP3 Inflammasome in Atherosclerosis, Coronary Artery Disease, and Acute Myocardial Infarction. *J Cardiovasc Transl Res,* 14**,** 23-34.

SINGH, B. N., RAO, K. S., RAMAKRISHNA, T., RANGARAJ, N. & RAO CH, M. 2007. Association of alphaB-crystallin, a small heat shock protein, with actin: role in modulating actin filament dynamics in vivo. *J Mol Biol,* 366**,** 756-67.

SINGH, D., KUMAR, J. & KUMAR, A. 2018. Isolation of pyomelanin from bacteria and evidences showing its synthesis by 4-hydroxyphenylpyruvate dioxygenase enzyme encoded by hppD gene. *Int J Biol Macromol,* 119**,** 864-873.

SINGHAL, N., KUMAR, M., KANAUJIA, P. K. & VIRDI, J. S. 2015. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Front Microbiol,* 6**,** 791.

SKELLAM, E. 2021. Biosynthesis of fungal polyketides by collaborating and trans-acting enzymes. *Nat Prod Rep*.

SMANSKI, M. J., PETERSON, R. M., HUANG, S. X. & SHEN, B. 2012. Bacterial diterpene synthases: new opportunities for mechanistic enzymology and engineered biosynthesis. *Curr Opin Chem Biol,* 16**,** 132-41.

SMITH, H. G., BEECH, M. J., LEWANDOWSKI, J. R., CHALLIS, G. L. & JENNER, M. 2021. Docking domain-mediated subunit interactions in natural product megasynth(et)ases. *J Ind Microbiol Biotechnol,* 48.

SOLIS-SOTO, J. M., QUINTANILLA-RODRIGUEZ, L. E., MEESTER, I., SEGOVIANO-RAMIREZ, J. C., VAZQUEZ-JUAREZ, J. L. & SALINAS CARMONA, M. C. 2008. In situ detection and distribution of inflammatory cytokines during the course of infection with Nocardia brasiliensis. *Histology and Histopathology,* 23**,** 573-581.

SONESON, C., LOVE, M. I. & ROBINSON, M. D. 2015. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res,* 4**,** 1521.

SONG, H., HWANG, J., YI, H., ULRICH, R. L., YU, Y., NIERMAN, W. C. & KIM, H. S. 2010. The early stage of bacterial genome-reductive evolution in the host. *PLoS Pathog,* 6**,** e1000922.

SONG, O. R., KIM, H. B., JOUNY, S., RICARD, I., VANDEPUTTE, A., DEBOOSERE, N., MARION, E., QUEVAL, C. J., LESPORT, P., BOURINET, E., HENRION, D., OH, S. B., LEBON, G., SANDOZ, G., YERAMIAN, E., MARSOLLIER, L. & BRODIN, P. 2017. A Bacterial Toxin with Analgesic Properties: Hyperpolarization of DRG Neurons by Mycolactone. *Toxins (Basel),* 9.

SPITS, H., ARTIS, D., COLONNA, M., DIEFENBACH, A., DI SANTO, J. P., EBERL, G., KOYASU, S., LOCKSLEY, R. M., MCKENZIE, A. N., MEBIUS, R. E., POWRIE, F. & VIVIER, E. 2013. Innate lymphoid cells--a proposal for uniform nomenclature. *Nat Rev Immunol,* 13**,** 145-9.

SPITS, H. & CUPEDO, T. 2012. Innate lymphoid cells: emerging insights in development, lineage relationships, and function. *Annu Rev Immunol,* 30**,** 647-75.

SPRUSANSKY, O., STIRRETT, K., SKINNER, D., DENOYA, C. & WESTPHELING, J. 2005. The bkdR gene of Streptomyces coelicolor is required for morphogenesis and antibiotic production and encodes a transcriptional regulator of a branched-chain amino acid dehydrogenase complex. *J Bacteriol,* 187**,** 664-71.

STAVNEZER, J. 1996. Immunoglobulin class switching. *Curr Opin Immunol,* 8**,** 199-205.

STERN, J. J., OCA, M. J., RUBIN, B. Y., ANDERSON, S. L. & MURRAY, H. W. 1988. Role of L3T4+ and LyT-2+ cells in experimental visceral leishmaniasis. *J Immunol,* 140**,** 3971-3977.

STEYN, A. J. C., COLLINS, D. M., HONDALUS, M. K., JACOBS, W. R., KAWAKAMI, R. P. & BLOOM, B. R. 2002. Mycobacterium tuberculosis WhiB3 interacts with RpoV to affect host survival but is dispensable for in vivo growth. *PNAS,* 99**,** 3147-3152.

STINEAR, T. P., MVE-OBIANG, A., SMALL, P. L. C., FRIGUI, W., PRYOR, M. J., BROSCH, R., JENKIN, G. A., JOHNSON, P. D. R., DAVIES, J. K., LEE, R. E., ADUSUMILLI, S., GARNIER, T., HAYDOCK, S. F., LEADLAY, P. F. & COLE, S. T. 2004. Giant plasmid-encoded polyketide synthases produce the macrolide toxin of Mycobacterium ulcerans. *Proc Natl Acad Sci U S A,* 101**,** 1345-1349.

STIRRETT, K., DENOYA, C. & WESTPHELING, J. 2009. Branched-chain amino acid catabolism provides precursors for the Type II polyketide antibiotic, actinorhodin, via pathways that are nutrient dependent. *J Ind Microbiol Biotechnol,* 36**,** 129-37.

STODDART, D., HERON, A. J., MIKHAILOVA, E., MAGLIA, G. & BAYLEY, H. 2009. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. *Proc Natl Acad Sci U S A,* 106**,** 7702-7.

TAKAKI, K. K., RINALDI, G., BERRIMAN, M., PAGAN, A. J. & RAMAKRISHNAN, L. 2021. Schistosoma mansoni Eggs Modulate the Timing of Granuloma Formation to Promote Transmission. *Cell Host Microbe,* 29**,** 58-67 e5.

TEUFEL, R., MASCARAQUE, V., ISMAIL, W., VOSS, M., PERERA, J., EISENREICH, W., HAEHNEL, W. & FUCHS, G. 2010. Bacterial phenylalanine and phenylacetate catabolic pathway revealed. *PNAS,* 107**,** 14390-14395.

TKACZUK, K. L., I, A. S., CHRUSZCZ, M., EVDOKIMOVA, E., SAVCHENKO, A. & MINOR, W. 2013. Structural and functional insight into the universal stress protein family. *Evol Appl,* 6**,** 434-49.

TOMLIN, H. & PICCININI, A. M. 2018. A complex interplay between the extracellular matrix and the innate immune response to microbial pathogens. *Immunology,* 155**,** 186-201.

TREANOR, B. 2012. B-cell receptor: from resting state to activate. *Immunology,* 136**,** 21-7.

TREVINO-VILLARREAL, J. H., VERA-CABRERA, L., VALERO-GUILLEN, P. L. & SALINAS-CARMONA, M. C. 2012. Nocardia brasiliensis cell wall lipids modulate macrophage and dendritic responses that favor development of experimental actinomycetoma in BALB/c mice. *Infect Immun,* 80**,** 3587-601.

TRUMAN, A. W. 2016. Cyclisation mechanisms in the biosynthesis of ribosomally synthesised and post-translationally modified peptides. *Beilstein J Org Chem,* 12**,** 1250-68.

TSUCHIYA, S., YAMABE, M., YAMAGUCHI, Y., KOBAYASHI, Y., KONNO, T. & TADA, K. 1980. Establishment and characterization of a human acute monocytic leukemia cell line (THP-1). *International Journal of Cancer,* 26**,** 171-176.

TURICK, C. E., TISA, L. S. & CACCAVO, F., JR. 2002. Melanin production and use as a soluble electron shuttle for Fe(III) oxide reduction and as a terminal electron acceptor by Shewanella algae BrY. *Appl Environ Microbiol,* 68**,** 2436-44.

UEMATSU, S., KAISHO, T., TANAKA, T., MATSUMOTO, M., YAMAKAMI, M., OMORI, H., YAMAMOTO, M., YOSHIMORI, T. & AKIRA, S. 2007. The C/EBP beta isoform 34-kDa LAP is responsible for NF-IL-6-mediated gene induction in activated macrophages, but is not essential for intracellular bacteria killing. *J Immunol,* 179**,** 5378-86.

UNIPROT, C. 2021. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res,* 49**,** D480-D489.

USUI, T., NAGUMO, Y., WATANABE, A., KUBOTA, T., KOMATSU, K., KOBAYASHI, J. & OSADA, H. 2006. Brasilicardin A, a natural immunosuppressant, targets amino Acid transport system L. *Chem Biol,* 13**,** 1153-60.

VAN DE SANDE, W., FAHAL, A., AHMED, S. A., SERRANO, J. A., BONIFAZ, A., ZIJLSTRA, E. & EUMYCETOMA WORKING, G. 2018. Closing the mycetoma knowledge gap. *Med Mycol,* 56**,** 153-164.

VAN DE SANDE, W. W. 2013. Global Burden of Human Mycetoma: A Systematic Review and Meta-analysis. *PLoS Negl Trop Dis,* 7.

VAN DE SANDE, W. W., FAHAL, A. H., GOODFELLOW, M., MAHGOUB EL, S., WELSH, O. & ZIJLSTRA, E. E. 2014. Merits and pitfalls of currently used diagnostic tools in mycetoma. *PLoS Negl Trop Dis,* 8**,** e2918.

VAN DE SANDE, W. W. J., FAHAL, A., VERBRUGH, H. & VAN BELKUM, A. 2007. Polymorphisms in Genes Involved in Innate Immunity Predispose Toward Mycetoma Susceptibility. *The Journal of Immunology,* 179**,** 3065-3074.

VAN DER HEUL, H. U., BILYK, B. L., MCDOWALL, K. J., SEIPKE, R. F. & VAN WEZEL, G. P. 2018. Regulation of antibiotic production in Actinobacteria: new perspectives from the post-genomic era. *Nat Prod Rep,* 35**,** 575-604.

VAN DER WERF, T. S., VAN DER GRAAF, W. T. A., TAPPERO, J. W. & ASIEDU, K. 1999. Mycobacterium ulcerans infection. *The Lancet,* 354**,** 1013-1018.

VAN DIJK, E. L., JASZCZYSZYN, Y., NAQUIN, D. & THERMES, C. 2018. The Third Revolution in Sequencing Technology. *Trends Genet,* 34**,** 666-681.

VERA-CABRERA, L., SALINAS-CARMONA, M. C., WAKSMAN, N., MESSEGUER-PEREZ, J., OCAMPO-CANDIANI, J. & WELSH, O. 2012. Host defenses in subcutaneous mycoses. *Clin Dermatol,* 30**,** 382-8.

VOLKMAN, H. E., POZOS, T. C., ZHENG, J., DAVIS, J. M., RAWLS, J. F. & RAMAKRISHNAN, L. 2010. Tuberculous Granuloma Induction via Interaction of a Bacterial Secreted Protein with Host Epithelium. *Science,* 327**,** 466-469.

WANG, J., DONG, Y., ZHOU, T., LIU, X., DENG, Y., WANG, C., LEE, J. & ZHANG, L. H. 2013a. Pseudomonas aeruginosa cytotoxicity is attenuated at high cell density and associated with the accumulation of phenylacetic acid. *PLoS One,* 8**,** e60187.

WANG, M. & CASEY, P. J. 2016. Protein prenylation: unique fats make their mark on biology. *Nat Rev Mol Cell Biol,* 17**,** 110-22.

WANG, R., GALLANT, E. & SEYEDSAYAMDOST, M. R. 2016. Investigation of the Genetics and Biochemistry of Roseobacticide Production in the Roseobacter Clade Bacterium Phaeobacter inhibens. *MBio,* 7**,** e02118.

WANG, Y., YANG, C., MAO, K., CHEN, S., MENG, G. & SUN, B. 2013b. Cellular localization of NLRP3 inflammasome. *Protein Cell,* 4**,** 425-31.

WARD, A. C. & ALLENBY, N. E. 2018. Genome mining for the search and discovery of bioactive compounds: the Streptomyces paradigm. *FEMS Microbiol Lett,* 365.

WEISSMAN, K. J. 2015. The structural biology of biosynthetic megaenzymes. *Nat Chem Biol,* 11**,** 660-70.

WICK, R. R., JUDD, L. M., GORRIE, C. L. & HOLT, K. E. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb Genom,* 3**,** e000132.

WILLIAMS, K. L., NANDA, I., LYONS, G. E., KUO, C. T., SCHMID, M., LEIDEN, J. M., KAPLAN, M. H. & TAPAROWSKY, E. J. 2001. Characterization of murine BATF: a negative regulator of activator protein-1 activity in the thymus. *Eur J Immunol,* 31**,** 1620-1627.

WILSON, J. L., MAYR, H. K. & WEICHHART, T. 2019. Metabolic Programming of Macrophages: Implications in the Pathogenesis of Granulomatous Disease. *Front Immunol,* 10**,** 2265.

WILSON, J. W. 2012. Nocardiosis: updates and clinical overview. *Mayo Clin Proc,* 87**,** 403-7.

YABUUCHI, E. & OHYAMA, A. 1972. Characterization of "Pyomelanin"-Producing Strains of Pseudomonas aeruginosa. *International Journal of Systemic Bacteriology,* 22**,** 53-64.

YU, Y., ZHANG, Q. & DENG, Z. 2017. Parallel pathways in the biosynthesis of aminoglycoside antibiotics. *F1000Res,* 6.

ZABURANNYI, N., RABYK, M., OSTASH, B., FEDORENKO, V. & LUZHETSKYY, A. 2014. Insights into naturally minimised Streptomyces albus J1074 genome. *BMC Genomics,* 15**,** 97-107.

ZHAO, Y., FENG, R., ZHENG, G., TIAN, J., RUAN, L., GE, M., JIANG, W. & LU, Y. 2015. Involvement of the TetR-Type Regulator PaaR in the Regulation of Pristinamycin I Biosynthesis through an Effect on Precursor Supply in Streptomyces pristinaespiralis. *J Bacteriol,* 197**,** 2062-71.

ZHENG, H., CHATFIELD, C. H., LILES, M. R. & CIANCIOTTO, N. P. 2013. Secreted Pyomelanin of Legionella pneumophila Promotes Bacterial Iron Uptake and Growth under Iron-Limiting Conditions. *Infection and Immunity,* 81**,** 4182-4191.

ZHOU, B., XIAO, J. F., TULI, L. & RESSOM, H. W. 2012. LC-MS-based metabolomics. *Mol Biosyst,* 8**,** 470-81.

ZHU, J., YAMANE, H. & PAUL, W. E. 2010. Differentiation of effector CD4 T cell populations. *Annu Rev Immunol,* 28**,** 445-89.

ZIEGLER, S. F. 2016. Division of labour by CD4(+) T helper cells. *Nat Rev Immunol,* 16**,** 403.

ZIJLSTRA, E. E., VAN DE SANDE, W. W. J., WELSH, O., MAHGOUB, E. S., GOODFELLOW, M. & FAHAL, A. H. 2016. Mycetoma: a unique neglected tropical disease. *The Lancet Infectious Diseases,* 16**,** 100-112.

ZSCHIEDRICH, C. P., KEIDEL, V. & SZURMANT, H. 2016. Molecular Mechanisms of Two-Component Signal Transduction. *J Mol Biol,* 428**,** 3752-75.

ZUGHAIER, S. M., RYLEY, H. C. & JACKSON, S. K. 1999. A Melanin Pigment Purified from an Epidemic Strain of Burkholderia cepacia Attenuates Monocyte Respiratory Burst Activity by Scavenging Superoxide Anion. *Infection and Immunity,* 67**,** 908-913.

ZUMLA, A., RAVIGLIONE, M., HAFNER, R. & VON REYN, C. F. 2013. Tuberculosis. *N Engl J Med,* 368**,** 745-55.