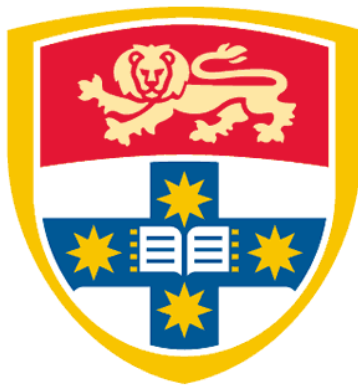


Defining immune system function in humans

A thesis submitted by

Andrew Sawyer



In fulfilment of the requirements for the degree of

Doctor of Philosophy

Faculty of Medicine & Health

The University of Sydney

2023

Table of Contents

Preface.....	7
Publications	8
Acknowledgements	9
Abbreviations	10
List of Figures.....	16
List of Tables	19
Abstract.....	20
Chapter 1 Introduction.....	22
1.1 General Introduction	23
1.2 Tuberculosis (TB) as an important human disease	27
1.2.1 The epidemiology of TB	27
1.2.1.1 Diagnosis of TB	28
1.2.1.2 Treatment of TB.....	29
1.2.1.3 Development of TB vaccines.....	29
1.2.2 Major immune determinants in protection against M.tb	31
1.2.3 Induction of the immune response to M.tb infection	34
1.2.4 The tissue response to persistent M.tb: the formation of granulomas.....	35
1.2.4.1 Evidence for a protective role of the TB granuloma.....	36
1.2.4.2 Evidence for a detrimental role of the TB granuloma	37
1.2.4.3 Heterogeneity of granulomas.....	38

1.2.5 The Human TB granuloma.....	39
1.3 Applications of transcriptomics to understand human disease	42
1.3.1 Bulk transcriptomics.....	42
1.3.1.1 Bulk transcriptomics analysis	43
1.3.1.2 Applications of bulk transcriptomics in disease diagnostics	44
1.3.1.3 Applications of bulk transcriptomics in predicting patient outcomes	45
1.3.1.4 Applications of bulk transcriptomics in longitudinal studies	46
1.3.2 Single cell transcriptomics	48
1.3.2.1 Single cell transcriptomics workflow	48
1.3.2.2 Applications of single cell transcriptomics.....	49
1.3.3 Mining public transcriptomic data.....	50
1.3.3.1 Meta-analyses of public transcriptomic data	52
1.3.3.2 Applications of meta-analyses of public transcriptomic data	53
1.3.3.3 Accessibility of public transcriptomic data to immunologists.....	55
1.3.4 Using public transcriptomic data to characterise novel gene expression	56
1.4 Utilising tissue imaging to advance human immunology	58
1.4.1 Conventional immunofluorescence (IF).....	59
1.4.2 OPAL multiplex imaging system	61
1.4.2.1 Applications of OPAL multiplex imaging.....	62
1.4.3 Co-detection by indexing (CODEX).....	63
1.4.4 Mass Based multiplex imaging	64

1.4.5 Spatially resolved transcriptomics.....	66
1.5 Spatial Biology.....	67
1.5.1 Cell segmentation	67
1.5.2 Cell density	68
1.5.3 Infiltration Analysis.....	68
1.5.4 Neighbourhood analysis	69
1.5.5 Spatial analysis software	69
1.6 The Present Study.....	71
Chapter 2 General Methods.....	72
2.1 OPAL staining.....	73
2.2 Image capture	73
2.3 Cell segmentation in HALO.....	74
2.4 Cell phenotyping in HALO	74
2.5 Statistical Analyses	74
Chapter 3 Transcriptomic profiling identifies neutrophil-specific upregulation of cystatin F as a marker of acute inflammation in humans.....	75
3.1 Introduction	76
3.2 Methods	78
3.2.1 Literature search	78
3.2.2 Single-cohort transcriptomic analysis	78
3.2.3 Multicohort transcriptomic analysis	78
3.2.4 scRNAseq analysis	79

3.2.5 Software.....	80
3.3 Results	83
3.3.1 Cystatin F is upregulated in neutrophils following stimulation with septic plasma and in circulating leukocytes of sepsis patients.....	83
3.3.2 The significance and regulation of cystatin F induction are poorly defined	88
3.3.3 Cystatin F is consistently upregulated across bacterial, viral, and sterile systemic inflammation.....	91
3.3.4 Cystatin F upregulation is neutrophil-specific	94
3.3.5 Cystatin F expression during inflammation is regulated by host factor(s).....	98
3.3.6 Cystatin F expression in neutrophils is associated with type I IFN signalling.....	102
3.4 Discussion	108
Chapter 4 Development of spatial tools to define intra-lesion cell distribution and inter- lesion relationships.....	112
4.1 Introduction	113
4.2 Results and discussion.....	115
4.2.1 TB lesions have diverse intra-lesion cell density	115
4.2.2 Lesion centroid is not a reliable reference point for quantifying cellular centrality of TB granulomas	119
4.2.3 Central and peripheral regions can be accurately defined by the distance of individual cells from the lesion border	122
4.2.4 Total Central Preference Index (tCPI) is a novel algorithm that defines intra-lesion cell density.....	124

4.2.5 Band-based infiltration analysis is not practical for quantifying intra-lesion cell-type distribution.....	129
4.2.6 Immune Cell Preference Index (immCPI) quantifies and enables visualisation of the intra-lesion cell-type distribution across any number of lesions and lesion types ...	132
4.2.7 Lesion neighbourhood Analysis quantifies the inter-lesion tissue superstructure	135
4.2.8 Generation of datasets using the LANDSCAPE method	138
4.2.9 Limitations of LANDSCAPE analysis.....	139
4.3 Concluding remarks	141
Chapter 5 Spatial mapping of human lung tissue reveals granuloma diversity and histopathological superstructure in TB	142
5.1 Introduction	143
5.2 Materials and methods	145
5.2.1 Patient Cohort.....	145
5.2.2 OPAL staining.....	145
5.2.3 Density based tissue region identification.....	148
5.2.4 Density based lesion identification.....	148
5.2.5 Lesion clustering analysis based on intra-lesion immune cell density.....	148
5.2.6 Spatial analysis of necrotising granulomas	149
5.2.7 Macrophage / monocyte distribution withing necrotising lesions.....	149
5.2.8 Lesion trajectory analysis	150
5.3 Results	151
5.3.1 Active pulmonary TB is associated with extensive tissue consolidation and	

development of immune cell aggregates in the lungs	151
5.3.2 Cellular aggregates are highly diverse and fall across a continuous spectrum	
underpinned by the density of CD20+ cells	157
5.3.3 Lesions can be stratified based on immune cell composition and spatial	
organisation	161
5.3.4 Divergent intra-lesion spatial distribution of CD20+ cells and CD68+ cells	
determines the types of non-necrotising granulomas	165
5.3.5 CD14+CD68+ subset of monocyte / macrophage populations preferentially	
localises at the necrotising border	169
5.3.6 The distance to necrotising lesions correlates with the relative abundance of	
specific immune cell populations in the surrounding non-necrotising lesions.....	171
5.3.7 Formation of non-necrotising lesions is shaped by the proximity to necrotising	
granulomas and intra-lesion organisation of CD20+ and CD68+ cells.....	173
5.4 Discussion	175
Chapter 6 General discussion	180
6.1. The current study.....	181
6.2. The future of public transcriptomic data	181
6.3 Single cell spatial analysis at the lesion level.....	182
6.4 Validation of TB animal models	185
6.5 The spectrum of human TB granulomas	187
6.6 Concluding remarks	189
Chapter 7 References.....	191

Preface

The work described in this thesis was performed in the Infection, Immunity and Inflammation theme, School of Medical Sciences, Faculty of Medicine and Health, The University of Sydney between June 2018 and September 2022. The author performed all experiments unless otherwise stated below. This work is entirely original and has not been presented previously for the purpose of obtaining another degree.

Archival clinical specimens were sectioned and stained with H&E by the Pathology laboratory of Shenzhen Third People's Hospital, Shenzhen, Guangdong, China. The characterisation of major histopathologic structures in the H&E-stained tuberculosis lung tissue was performed by Dr. Timothy Fielder, Department of Clinical Pathology, Royal Prince Alfred Hospital, Camperdown, NSW, Australia.

Andrew Sawyer

September 2022, Sydney

Publications

Parts of Chapter 3 have been published as:

SAWYER, A. J., GARAND, M., CHAUSSABEL, D. & FENG, C. G. 2021. Transcriptomic Profiling Identifies Neutrophil-Specific Upregulation of Cystatin F as a Marker of Acute Inflammation in Humans. *Frontiers in Immunology*, 12.

I designed the study, analysed the data, and wrote the drafts of the manuscript.

Parts of Chapter 4 and Chapter 5 have been published as:

SAWYER, A., PATRICK, E., EDWARDS, J., WILMOTT, J. S., FIELDER, T., YANG, Q., BARBER, D. L., ERNST, J. D., BRITTON, W. J., PALENDIRA, U., CHEN, X. & FENG, C. G. 2022. Spatial mapping of human lung tissue reveals granuloma diversity and histopathological superstructure in tuberculosis. *bioRxiv*, 2022.07.31.502240. **Under revision: Journal of Experimental Medicine.**

I designed the study, performed all experiments, analysed the data, and wrote the drafts of the manuscript.

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Andrew Sawyer, , 29/09/2022

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Carl Feng, , 29/09/2022

Acknowledgements

Many people are owed great thanks for all the help that they gave me throughout my PhD.

Firstly, to my supervisor Carl. Thank you for constant support and guidance over these four years. I feel blessed to have a supervisor so dedicated and knowledgeable in science and also so willing to walk alongside your students as we take our first baby steps as scientists. You were always there when I needed you, especially through the more difficult and uncertain parts of this journey and I couldn't have asked for a better supervisor. I'm looking forward to continue working with you in the future!

To my co-supervisor Damien, thank you for teaching me the basics of bioinformatics and opening my eyes its vast potential in complementing bench research. Your kind words and support in my project were much appreciated. To my other co-supervisor Warwick, thank you for your guidance and sharing your immense TB knowledge with me.

To Jarem, thank you for teaching me the basics of OPAL and HALO, without you this project may have never got off the ground! To Ellis, thank you for walking alongside me for much of this project and your patience with me when I struggled to get my head around things. To Tim, I wish we'd met earlier! Your kindness to us and your expertise have made a huge impact on my project. To Mainthan, Angela and James, thank you for the helpful discussions and advice throughout the years. To our lab, Lina, Taylor, Mikaela, Nayan, Seb, and Helen, thanks for your help and encouragement.

To my parents, thank you for constantly supporting me in so many ways while I wrestled with this project. Thank you too for proofreading the thesis for me!

Finally, to God, thank you for sustaining me and revealing the wonders of your creation in the extraordinary intricacy of the immune system.

Abbreviations

AIDS	Acquired immune deficiency syndrome
AFB	Acid fast bacilli
ANXA3	Annexin A3
AOSD	Adult-onset Still's disease
ARDS	Acute respiratory distress syndrome
BCG	Bacillus Calmette–Guérin
CA	Cellular aggregate
cDNA	Complementary deoxyribose nucleic acid
CF	Cystic fibrosis
CLI	Command line interface
CODEX	Co-detection by indexing
COVID-19	Coronavirus disease 2019
<i>CST7</i>	Cystatin F
CT	Computerised tomography
CT	Consolidated tissue
CTSC	Cathepsin C
CTSG	Cathepsin G

DC	Dendritic cell
DEG	Differentially expressed gene
ELANE	Elastase, neutrophil expressed
FFPE	Formalin-fixed paraffin-embedded
FISH	Fluorescence in-situ hybridisation
FITC	Fluorescein isothiocyanate
FRESH	Females Rising through Education Support and Health
FOV	Field of view
$\gamma\delta$ T cells	Gamma delta T cells
GEO	Gene Expression Omnibus
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GSE	Gene Series Expression
GSEA	Gene-set enrichment analysis
GO	Gene Ontology
GUI	Graphical user interface
GXB	Gene Expression Browser
H&E	Haematoxylin and eosin
HIER	Heat-induced epitope retrieval
HIV	Human immunodeficiency virus
HRP	Horseradish peroxidase

HT	Healthy Tissue
IAV	Influenza A virus
IDO1	Indoleamine 2,3-dioxygenase
IF	Immunofluorescence
IFN	Interferon
IFN γ	Interferon gamma
IGRA	Interferon gamma release assay
IHC	Immunohistochemistry
IL-1 β	Interleukin 1 beta
IL-6	Interleukin 6
IL-10	Interleukin 10
IL-12	Interleukin 12
IL-17	Interleukin 17
ILC	Innate lymphoid cell
IMC	Image mass cytometry
ISG	Interferon stimulated gene
iNK T cells	Innate natural killer T cell
LANDSCAPE	Lesion And Neighbourhood Description Stratified by cellular Composition And Position Estimates
LPS	Lipopolysaccharide

LTBI	Latent tuberculosis infection
MDR-TB	Multi-drug resistant tuberculosis
MHC	Major histocompatibility complex
MIBI	Multiplex ion beam imaging
MIBI-TOF	Multiplex ion beam imaging by time of flight
μCT	Micro-computerised tomography
<i>M.marinum</i>	<i>Mycobacterium marinum</i>
mRNA	Messenger ribose nucleic acid
MSI	Multispectral images
MSMD	Mendelian susceptibility to mycobacterial disease
M.tb	<i>Mycobacterium tuberculosis</i>
NAA	Nucleic acid amplification
NHP	Non-human primate
NK	Natural killer
NUDT16	Nudix hydrolase 16
PBMC	Peripheral blood mononuclear cells
PCA	Principal component analysis
PC1	Principal component 1
PD-1	Programmed cell death protein 1
PD-L1	Programmed death-ligand 1

PET	Positron emission tomography
PGN	Peptidoglycan
PRTN3	Proteinase 3
RNA	Ribose nucleic acid
RNAseq	Ribose nucleic acid sequencing
ROI	Region of interest
RSV	Respiratory syntactical virus
SCC	Single colour control
scRNAseq	Single Cell Ribose nucleic acid sequencing
SIRS	Systemic inflammatory response syndrome
SOJIA	Systemic onset juvenile idiopathic arthritis
SLE	Systemic Lupus Erythematosus
TB	Tuberculosis
tCPI	Total central preference index
TGF- β	Transforming growth factor β
TIM-3	T-cell immunoglobulin and mucin domain-3
Th1	T helper 1
Th2	T helper 2
Th17	T helper 17
TLR	Toll-like receptor

TLS	Tertiary lymphoid structure
TMA	Tissue microarray
TNF	Tumour necrosis factor
TOF	Time of flight
TSA	Tyramide signal amplification
TST	Tuberculin skin test
UMAP	Uniform manifold approximation and projection
UV	Ultraviolet
WHO	World Health Organisation
XDR-TB	Extensively drug-resistant tuberculosis

List of figures

Chapter 3

Figure 3.1. Flow chart of the process employed in this study to identify a gene of interest and assess its regulation.....	84
Figure 3.2. Meta analysis of sepsis associated neutrophil genes.	85
Figure 3.3. Identification of <i>CST7</i> as a previously unknown neutrophil-specific, sepsis-associated gene.....	87
Figure 3.4. Cystatin F is identified as a cathepsin C-directed serine protease inhibitor.....	90
Figure 3.5. <i>CST7</i> is upregulated across a spectrum of infectious diseases.	92
Figure 3.6. <i>CST7</i> is upregulated in non-infectious inflammatory diseases.....	93
Figure 3.7. <i>CST7</i> upregulation during inflammation is neutrophil-specific.	95
Figure 3.8. <i>CST7</i> is upregulated in neutrophils under multiple inflammatory conditions.....	97
Figure 3.9. <i>CST7</i> upregulation could not be induced by pro-inflammatory molecules in neutrophils ex vivo.....	99
Figure 3.10. Neutrophil-specific <i>CST7</i> upregulation during inflammation is regulated by host factors.....	101
Figure 3.11. <i>CST7</i> is upregulated in neutrophils during severe COVID-19.	103
Figure 3.12. <i>CST7</i> is expressed in mature and immature neutrophils.....	105
Figure 3.13. <i>CST7</i> is upregulated in neutrophils during severe COVID-19.	107

Chapter 4

Figure 4.1. Diverse pathological presentations in M.tb-infected human lung.....	117
--	-----

Figure 4.2. Cells in non-necrotizing lesions show diverse spatial distribution preference...	118
Figure 4.3. Lesion centroid is not an accurate reference point for quantifying cell centrality within mycobacterial lesions.....	121
Figure 4.4. Lesion border is a useful reference point for defining central and peripheral intra-lesion cell distribution.....	123
Figure 4.5. Intra-lesion cell density can be quantified by comparing the areas covered by the inner and outer 50% of cells.	126
Figure 4.6. tCPI is a single number measurement for intra-lesion cell localisation preference.	128
Figure 4.7. Band-based infiltration analysis is not practical for quantifying intra-lesion cell-type distribution in high numbers of lesions and across lesion types.	131
Figure 4.8. immCPI quantifies intra-lesion cell-type distribution in any number of lesions and across lesion types.....	134
Figure 4.9. Lesion neighbourhood analysis quantifies inter-lesion distances.....	137
Figure 4.10. Example dataset structures generated by LANDSCAPE analysis.	140

Chapter 5

Figure 5.1. Active pulmonary TB is characterised by development of histopathological superstructures in the lung.	152
Figure 5.2. Tissue regions can be stratified by cell density.	154
Figure 5.3. Lesion delineation by immune cell density.	156
Figure 5.4. Necrotising and non-necrotising lesions have diverse immune cell compositions.	158

Figure 5.5. Lesions do not form distinct clusters but fall across a continuous spectrum underpinned by CD20 ⁺ cell density.	160
Figure 5.6. Combined analysis of compositional and spatial features identifies four types of TB lesions.	162
Figure 5.7. Distribution, composition, and activity of immune cells in each lesion type.....	164
Figure 5.8. Divergent intra-lesion spatial distribution of CD20 ⁺ cells and CD68 ⁺ cells determine the types of non-necrotising granulomas.	166
Figure 5.9. Spatial organisation of immune cells around the necrotic interface in necrotising lesions.	168
Figure 5.10. CD14 ⁺ CD68 ⁺ subset of monocyte / macrophage populations preferentially localise at the necrotising border.	170
Figure 5.11. The distance of non-necrotising lesions from necrotising lesions is associated with lesion type.	172
Figure 5.12. Trajectory analysis reveals lesions organise according to type.	174
Figure 5.13. Schematic diagram summarising key findings and illustrating hypothetical lesion progression pathways.	177

List of tables

Chapter 1

Table 1.1. Strengths and weaknesses of animal TB models.	32
--	----

Chapter 3

Table 3.1. Transcriptomic datasets used in this study.....	81
---	----

Chapter 5

Table 5.1. Clinical information of TB patients.	146
---	-----

Abstract

In this thesis I design two novel methods for the analysis of human immunological data from blood and solid tissue respectively.

I have developed a method capable of identifying poorly characterised disease-associated genes and defining mechanisms controlling their expression. By exploring multiple cohorts of public transcriptome data, I showed that *CST7* is upregulated in the blood during a diverse set of inflammatory conditions. Interestingly, this upregulation was neutrophil-specific and is not induced by microbial products or cytokines commonly associated with inflammation but is associated with type I interferon signalling. In this chapter, I demonstrate the value of publicly available transcriptome data in knowledge generation and potential biomarker discovery.

I have also developed a protocol for the single-cell analysis of lesions in multiplex images. I developed a metric called tCPI that can quantify how tight or loose the distribution of cells is in the lesion centre. An additional metric was developed called immCPI that measures the overall distribution of each immune cell population relative to the lesion centre. I developed a final method called lesion neighbourhood analysis that can quantify the relative location of hundreds of individual lesions within a tissue section and can identify patterns in their distributions. Together this workflow provides a robust tool for spatial quantification at the lesion level.

I applied this lesion analysis protocol to lung tissue resection samples from patients with Tuberculosis. Interestingly, lesions did not fall into distinct clusters based on composition. Using both principal component data and tCPI, I grouped lesions into 4 types based on composition and cell distribution. Using immCPI revealed that the intra-lesion location macrophages / monocytes and B cells is dependent on lesion type. Lesion

neighbourhood analysis revealed that B cell enriched lesions tend to locate twice as close to a necrotising lesion compared to other lesions.

The robust analytic power of these new approaches offers important tools in the study of human immunology and will contribute to furthering our understanding of inflammatory diseases.

Chapter 1

Introduction

1.1 General Introduction

The field of immunology has grown exponentially over the past decades. As research has revealed the extraordinary intricacy and efficiency of the immune system, much credit is due to the development of animal models. Many of the foundational discoveries in immunology were made in animals, including clonal selection (Talmage, 1957, Burnet, 1976, Nossal and Lederberg, 1958), dendritic cells (DC) (Steinman and Cohn, 1973), the major histocompatibility complex (MHC) (Hood et al., 1983), thymic T cell development (Miller, 1961, Miller, 2011), T and B cell functions in adaptive immunity (Cooper et al., 1965) and the role of toll-like receptors (TLR) in innate immunity (Poltorak et al., 1998, Lemaitre et al., 2012). The inbred mouse has become the quintessential model for in-vivo immunological research and the ability to knock out (Watts, 2007) or overexpress (Brinster et al., 1981) genes of interest in animal models has revealed the role of countless genes and proteins underpinning the immune response (Mak et al., 2001, Masopust et al., 2017). However, as the field increasingly expands beyond discovery studies to applications in preventing and treating human disease, the limitations of the inbred mouse model are becoming clear (Davis and Brodin, 2018, Jameson and Masopust, 2018, Su et al., 2013). While some effective immune-based therapies in mice have been successfully translated to real world human medicine (Notable recent examples being checkpoint inhibitor therapy for melanoma (Sharma and Allison, 2015, Ai and Curran, 2015) and tumour necrosis factor (TNF) inhibitor therapy for rheumatoid arthritis (Feldmann and Maini, 2008)), these are proving to be the exception rather than the rule (Rice, 2012, von Herrath and Nepom, 2005, Steinman and Mellman, 2004, van der Worp et al., 2010, Mak et al., 2014).

Differences between the immune systems of mice and humans have long been known. (Mestas and Hughes, 2004, Payne and Crooks, 2007) These comprise variations in both

innate (Iwasaki and Medzhitov, 2004, Zschaler et al., 2014) and adaptive immunity (Dekkers et al., 2017). The genetic consistency of the inbred mouse model, one of its greatest strengths in experimental rigour, can be seen as a weakness when translating results to human disease given the high levels of person-to-person immune diversity (Brodin and Davis, 2017, Kaczorowski et al., 2017, Maecker et al., 2012, Orrù et al., 2013), much of which is believed to be driven by environmental factors (Brodin et al., 2015, Carr et al., 2016) which cannot be recapitulated in typical specific-pathogen free housing of experimental mice (Beura et al., 2016, Tao and Reese, 2017). Taken together with the fundamental biological differences in size, lifespan (Furman and Davis, 2015, Wagar et al., 2018, den Braber et al., 2012), disease presentation (Constantinescu et al., 2011, Webster et al., 2014, Ostrand-Rosenberg, 2004), as well as recently identified transcriptional differences (Seok et al., 2013, Godec et al., 2016, Takao and Miyakawa, 2015, Piasecka et al., 2018, Shay et al., 2013) it is becoming clear that the inbred mouse model alone will not be sufficient to develop treatments for all of humanity's most pressing diseases (Pulendran and Davis, 2020, Davis, 2008). An alternative model, the non-human primate (NHP) has been growing in popularity in the last two decades and has significant advantages over the mouse in terms of biological similarity to humans (Flynn et al., 2015, Brok et al., 2001), but its usage is limited by higher cost, husbandry requirements and ethical restraints compared to the mouse model (Padrell et al., 2021). Additionally results in the NHP are still not guaranteed to translate effectively to humans (Bjornson-Hooper et al., 2022) and in some cases successful interventions in the NHP model have even proven to be deleterious when applied in human trials (Suntharalingam et al., 2006, Buchbinder et al., 2008, Sekaly, 2008, Attarwala, 2010).

While studying immunology directly in humans promises to overcome the translational shortfalls of animal models, human immunology had long been relegated to the sidelines of immunological research for several reasons (Davis and Brodin, 2018, Hayday

and Peakman, 2008). Experimental scope in human studies is far less than in animal models for ethical reasons (Wagar et al., 2018). Tissue availability is also a major limiting factor as while blood is readily accessible, solid tissue biopsies may not be practically accessible in many conditions, are often in short supply and in the case of infection and cancer, are usually collected at less-than-ideal times for research objectives (Carpenter et al., 2018, Pulendran and Davis, 2020). Despite these limitations, human immunology has seen a resurgence over the last two decades (Varadé et al., 2021), largely driven by high throughput technologies and systems immunology that have exponentially increased the scope of data that can be acquired from samples (FitzGerald et al., 2018, Martino and Allen, 2015). Transcriptomic and multiplex imaging techniques are two such examples of major developments in high throughput technologies that are revolutionising human immunology. Both single cell and bulk transcriptomic technologies are facilitating breakthroughs including discovery of novel cell types (Björklund Å et al., 2016, Villani et al., 2017), cell development pathways (See et al., 2017, Stubbington et al., 2017) as well as biomarkers for disease diagnostics (Zaas et al., 2009), vaccine efficacy (Lu et al., 2016, Jennewein et al., 2019) and clinical outcomes (Banchereau et al., 2016, Abboud et al., 2016). As multiplex imaging technologies have also advanced this is facilitating development of rapid immunophenotyping of tissue biopsies for diagnostic purposes in cancer (Van Herck et al., 2021, Galon et al., 2006, Allam et al., 2020) and when combined with advances in spatial biology tools, are redefining the types of cellular analyses that can be performed in solid tissue (van Dam et al., 2022, Giladi and Amit, 2017). High-throughput technologies typically produce vastly more data than the original researchers require to address their hypotheses and entire datasets are usually stored in publicly available online repositories, providing invaluable resources to addressing future research questions (Regev et al., 2017, Sparks et al., 2016, Burel et al., 2016). However, as

the amount and complexity of data generated by these technologies continues to grow, they will require continuously evolving analysis protocols to effectively utilise data.

All these data provide invaluable guidance and validation to animal models, which will continue to have an important role in answering mechanistic questions that remain beyond the scope of human studies (Wagar et al., 2018, Pulendran and Davis, 2020).

1.2 Tuberculosis (TB) as an important human disease

TB, caused by infection with the bacterium *Mycobacterium tuberculosis* (M.tb) is one of the oldest diseases that continues to afflict humanity and remains a major cause of morbidity and mortality, second only to Coronavirus disease 2019 (COVID-19) in deaths caused by a single infectious agent (WHO, 2021). While infection with M.tb can result in development of symptomatic and transmissible disease, termed active TB, in an estimated 90% of M.tb-infected individuals the bacteria establish a non-symptomatic and non-transmissible sub-clinical condition termed latent TB infection (LTBI) (Getahun et al., 2015). LTBI can progress to active TB in the months or years following initial infection and it is estimated that one in four people on earth have LTBI (Cohen et al., 2019). There is also growing evidence of a spectrum of M.tb infection states with the term subclinical TB used to covers the transitionary period between LTBI and active TB in which there is evidence of M.tb replication but no clinical symptoms (Lin and Flynn, 2018, Barry et al., 2009, Coleman et al., 2014). In approximately 80% of active TB cases the infection is contained to the respiratory system and termed pulmonary TB (Lyon and Rossman, 2017), however M.tb can also cause disease in almost any other body tissue as extra-pulmonary TB (Baykan et al., 2022).

1.2.1 The epidemiology of TB

TB is a preventable and largely curable disease that has been mostly eradicated in the developed world, however it still remains a major global health problem, predominantly affecting the developing world and in particular, the immunocompromised (WHO, 2021). Human Immunodeficiency Virus (HIV) infection is a major risk factor for developing active TB disease and TB is the leading cause of hospitalisation in HIV infected individuals (Ford et al., 2015). Poor nutrition and indoor air pollution are also significant risk factors in

developing active TB disease (Lönnroth et al., 2010). There were an estimated ten million cases of TB overall in 2020 resulting in 1.5 million deaths, highlighting the high mortality rate of the disease (WHO, 2021). Indeed, untreated TB disease will prove fatal in about 50% of cases (Tiemersma et al., 2011). While the World Health Organisation (WHO) has set targets to reduce TB deaths by 90% and TB incidence by 80% by 2030 compared to 2015 numbers, these efforts are hampered by significant shortfalls in the diagnosis, treatment, and prevention of TB disease (Chakaya et al., 2020).

1.2.1.1 Diagnosis of TB

A diagnosis of LTBI is usually established by testing whether the asymptomatic individual's immune system has encountered M.tb before, either by examining hypersensitivity to M.tb antigens in the tuberculin skin test (TST) (Gualano et al., 2019) or by measuring the recall Interferon Gamma (IFN γ) response to defined M.tb antigens in whole blood using the IFN γ Release Assay (IGRA) (Zhou et al., 2020). While the sensitivity of these tests is high in immunocompetent individuals, it is reduced in immunocompromised patients such as those with HIV (Yang et al., 2022, Primaturia et al., 2020). Additionally, neither of these tests has strong predictive value as to whether individuals with LTBI will progress to clinically significant active TB (Rangaka et al., 2012). Making a diagnosis of active TB disease is more complicated as TB can present similarly to other respiratory illnesses and case finding may utilise a combination of non-specific symptomatic findings (fever, cough, weight loss etc), radiological findings (chest X-ray, CT scan) and immunological findings (TST, IGRA) to screen patients. Traditionally, the definitive diagnosis is made based on identification of M.tb bacilli microscopically either on a sputum smear or in sputum culture with the mycobacteria-specific acid-fast stain (Pai et al., 2016). The recent introduction of nucleic acid amplification

(NAA) tests, most prominently the Gene Xpert MTB/RIF has been an important advance in TB diagnosis in poorer areas, allowing rapid analysis of sputum to detect M.tb with a higher sensitivity compared to microscopic approaches and the ability to detect resistance to rifampicin, one of the front-line drugs to treat TB (Steingart et al., 2014).

1.2.1.2 Treatment of TB

Standard treatment for active TB requires a minimum 6-month course of daily multidrug therapy comprising the drugs rifampicin and isoniazid with the addition of pyrazinamide, and ethambutol for the first two months (Nahid et al., 2016). This regime is highly effective in curing drug-sensitive TB and has remained largely unchanged for the last four decades (Dheda et al., 2016). However, the frontline antibiotics used have a number of side effects and additionally, up to 50% of patients do not complete the full 6-month (sometimes longer) antibiotic course prescribed for them, facilitating the development of drug resistant M.tb strains (Pai et al., 2016, Adisa et al., 2021). Multidrug-resistant TB (MDR-TB) is caused by M.tb that is resistant to rifampicin and isoniazid while extensively drug-resistant TB (XDR-TB) is caused by M.tb strains that are resistant to additional second line drugs as well as rifampicin and isoniazid. Treatment of most drug resistant TB strains is possible, but requires the use of more toxic antimicrobial agents (Calligaro et al., 2014). TB treatment would be greatly improved with the development of shorter and less toxic treatment protocols as well as enhanced understanding of pathogenesis of TB disease.

1.2.1.3 Development of TB vaccines

While many public health measures are prescribed in TB endemic areas to limit the spread of M.tb, disease prevention is hampered by the lack of an effective vaccine. The only current

vaccine in use, *Bacillus Calmette–Guérin* (BCG), an attenuated *M. bovis* strain offers protection against disseminated extrapulmonary active TB in children under five and is currently prescribed to over 90% of children born around the world (Zwerling et al., 2011, Trunz et al., 2006). However, its effectiveness in adolescents and adults is unclear, with highly variable results from studies ranging from 0-80% protection (Mangtani et al., 2014) with a greater change of protective efficacy at latitudes of 40 degrees and above (Fine, 1995). This suggested that factors that vary with latitude, such as sunlight or socio-economic status may be associated with vaccine efficacy (Wilson et al., 1995). It is clear though that despite widespread vaccination, BCG has not been sufficient to prevent the spread of TB given the high numbers of cases that still occur each year.

The goal of any vaccine is to elicit a protective adaptive immune response that persists through immunological memory and is able to prevent subsequent disease and mortality when the patient encounters the actual pathogen. While the effectiveness of BCG in infants has demonstrated that it is possible to generate a successful anti-*M. tb* immune response through vaccination, it remains unclear exactly how BCG does this and what the immune correlates of protection are (Pai et al., 2016). Furthermore, the development of new vaccines is hindered by the lack of knowledge of what constitutes a successful immune response against *M. tb* in both humans and animal models. It is also unclear why most people who are infected with *M. tb* are able to control the infection while a subset go on to develop active disease, both initially following infection and then others in the months or years that follow (Cadena et al., 2017). Curiously, some individuals, though highly exposed to *M. tb* never show a positive TST or IGRA suggesting that *M. tb* may even be cleared in these individuals without the initiation of an adaptive immune response (Morrison et al., 2008). There is a dire need therefore, for an improved understanding of what comprises a successful

anti-M.tb immune response and advances in human immunology approaches are facilitating new and promising discoveries.

1.2.2 Major immune determinants in protection against M.tb

Much of what we know about the progression of the immune response to M.tb comes from animal and in-vitro models of M.tb infection. The strengths and weaknesses of common TB animal models are elaborated in Table 1.1. It has been known for half a century that T cells play an important role in immunity to M.tb based on mouse studies that showed extreme susceptibility of athymic mice to M.tb (North, 1973), and later observations that humans with Acquired immune deficiency syndrome (AIDS), a disease caused by HIV infection that results in CD4⁺ T cell deficiency, are highly susceptible to active TB disease (Geldmacher et al., 2012). Further studies in mice revealed that a subset of T cells that produce IFN γ (later known as T helper 1 (Th1) cells) in particular mediate this protection (Orme et al., 1993, Shimokata et al., 1986, Mogues et al., 2001). A role for IFN γ in human TB was also suggested by the fact that individuals with one of a number of possible genetic defects that result a deficiency in IFN γ production show increased susceptibility to a variety of mycobacterial species in a condition known as mendelian susceptibility to mycobacterial disease (MSMD) (Casanova and Abel, 2002). An important role for TNF in anti-TB was also established based on the fact that mice deficient in TNF succumbed to M.tb infection (Flynn et al., 1995). This was later reaffirmed in humans from observations that patients undergoing anti-TNF therapy were at high risk of active TB disease (Keane et al., 2001). The development of the knockout mouse also established important roles for CD4⁺ T cells, CD8⁺ T cells, Gamma Delta T cells ($\gamma\delta$ T cells) and invariant Natural killer T cells (iNK T cells) in the immune response to M.tb (Mogues et al., 2001, Behar et al., 1999).

Table 1.1. Strengths and weaknesses of animal TB models.

Species	Strengths	Weaknesses	References
Mouse	• Cheap and low husbandry requirements • Plentiful reagents	• Cannot form compact granulomas or necrotising lesions with the exception of C3H/FeJ strain • Cannot model LTBI	(Dharmadhikari and Nardell, 2008, Singh and Gupta, 2018)
Guinea pig	• Spreads M.tb via cough • Forms pathologically similar lesions to human TB	• Relatively few reagents • Relatively expensive and difficult husbandry requirements	(Singh and Gupta, 2018, Gupta and Katoch, 2005)
Zebrafish	• Earliest events of granuloma formation can be modelled • Live granulomas can be visualised in vivo • Cheap and low husbandry requirements • Easily bred	• Embryos used for live visualisation lack adaptive immunity • Uses <i>M.marinum</i> rather than M.tb • Evolutionarily distant from humans	(Singh and Gupta, 2018, Young, 2009, Williams and Orme, 2016)
NHP	• Develop a clinical spectrum of disease similar to human TB • Works with most human reagents • Use with PET/CT scanning can visualise granuloma dynamics in vivo	• Very expensive • Highest animal husbandry requirements • Requires large housing • Low availability	(Scanga and Flynn, 2014, Flynn et al., 2015, Singh and Gupta, 2018)

It was long considered that the central tenet of anti-M.tb immunity was IFN γ produced by CD4⁺ T cells that then activates macrophages and facilitates bacterial killing, however later discoveries have brought this central dogma into question (Nunes-Alves et al., 2014). For example, while those infected with HIV are at heightened risk of developing active TB, this increased susceptibility occurs from the early stages of HIV infection, before depletion of the CD4⁺ T cell population (Sonnenberg et al., 2005) which suggests a more nuanced risk profile than simply a lack of CD4⁺ T cells. Furthermore, the depletion of CD4⁺ T cells and development of AIDS does not correlate with a further increased risk of TB disease compared to those who are only infected with HIV and have normal CD4⁺ T cell counts (Murray et al., 2007). The essential role of IFN γ has also been called into question by the observation that while patients with MSDM are highly susceptible to disseminated infection with a number of mycobacterial species, M.tb does not seem to be among them (Casanova and Abel, 2002), suggesting a level of redundancy in the role of IFN γ in the human anti-M.tb immune response. Most TB patients also have normal CD4⁺ T cell numbers and normal IFN γ levels (Wallis et al., 2013), indicating that even if these components of the immune response are necessary for a successful anti-M.tb immune response, their presence alone is not sufficient to prevent active TB disease. Further, a clinical trial of BCG revealed that protection generated in infants does not correlate with IFN γ production (Kagina et al., 2010). Taken together, there is strong evidence that IFN γ does not play as critical a role in humans to a successful anti-M.tb immune response as it does in the mouse model.

Besides IFN γ and TNF, it has been shown that mice deficient in any of the cytokines, granulocyte-macrophage colony-stimulating factor (GM-CSF) (Gonzalez-Juarrero et al., 2005), Interleukin 1 beta (IL-1 β) (Mayer-Barber et al., 2010) IL-6 (Ladel et al., 1997) or IL-12 (Cooper et al., 1997, Feng et al., 2005) are also highly susceptible to M.tb infection again suggesting a nuanced and intricate balance of cytokines involved in host protection. Other

research into the M.tb immune response takes a more wholistic approach to characterise the full series of immunological events that follow M.tb infection.

1.2.3 Induction of the immune response to M.tb infection

M.tb bacilli spread from person-to-person within aerosolised droplets from an infected individual that are inhaled by those around them. The earliest timepoint of M.tb infection that can be studied in humans is when a TST or IGRA turns positive, which occurs up to 8 weeks following initial infection (Lin and Flynn, 2018), and almost all knowledge of the earliest events of M.tb infection is inferred from other models. Importantly, it has been shown across many animal models of TB that M.tb infection outcomes can vary depending on the circumstances of initial infection, with a higher dose of M.tb bacilli leading to worse disease outcomes compared to infection with a lower dose (Kaushal et al., 2012).

Once M.tb enters the respiratory tract it can infect respiratory epithelial cells (Gold et al., 2010), however its primary target is the alveolar macrophages deep in the lower respiratory tract (Orme et al., 2015). Alveolar macrophages internalise the bacteria via phagocytosis and at this point the bacteria may either be cleared by the alveolar macrophages or may persist by preventing fusion of the phagosome with the lysosome and escape the phagosome to establish infection in the cytosol of these cells (Russell, 2011, Houben et al., 2012, van der Wel et al., 2007). Alveolar macrophages comprise a heterogeneous population of cells and it has been shown in the NHP model that the presence of different subsets of these cells prior to infection is associated with infection outcome (Silver et al., 2016). While the first encounters with M.tb occur in the airway, M.tb often manages to cross the interstitium into the lung parenchyma to continue the infection. It has been suggested based on research in the zebrafish model, using infection with *Mycobacterium Marinum*

(*M.marinum*) that creates a TB-like disease, that infected macrophages can directly translocate across the interstitium to enter the lung parenchyma. It is not clear in other species however whether infection initially spreads this way, and it may also occur via subsequent infection of epithelial cells lining the alveoli and then deeper into the lung tissue (Pai et al., 2016, Pagán and Ramakrishnan, 2014).

While initial encounters with *M.tb* occur in lung-resident cells, there is also rapid recruitment of leukocytes to the infection site and these cells may both contribute to bacterial containment as well as serve as vessels to facilitate bacterial spread (Ramakrishnan, 2012). In the initial days and weeks following the onset of infection, myeloid cells, in particular macrophages, tend to aggregate around areas of infection. After the initiation of the adaptive immune response, after approximately 6 weeks, lymphocytes are additionally recruited to infection sites and also contribute to these granulomas (Cadena Anthony et al., 2016, Capuano et al., 2003). It has been shown in the zebrafish model, using embryos that lack an adaptive immune system, that granulomas can still be formed in response to *M.marinum* with only innate immunity (Davis et al., 2002, Davis and Ramakrishnan, 2009). Following the initial events of infection, the granuloma becomes the interface at which interactions between *M.tb* and the infected host occur.

1.2.4 The tissue response to persistent *M.tb*: the formation of granulomas

The granuloma has been identified for centuries as the histopathological hallmark of TB infection and radiological identification of granulomatous lesions through X-ray and Computerised Tomography (CT) scans remains an important step in diagnosing TB disease. The precise definition of a granuloma varies; however, they are typically considered as an organised aggregate of immune cells, including macrophages and lymphocytes, that forms in

response to persistent stimuli. Granuloma formation is not exclusive to TB and these structures can form in response to other mycobacteria infections (e.g. leprosy) (Modlin and Rea, 1988), other bacterial infections (Lamps and Scott, 2004, Cho et al., 1984) and in response to non-infectious stimuli such as foreign bodies (Molina-Ruiz and Requena, 2015). Another common cause of granulomas is sarcoidosis, although in this case the cause is unknown (Broos et al., 2013). TB granulomas can be broadly divided into non-necrotising and necrotising. Non-necrotising granulomas have a solid core of live cells while necrotising granulomas contain a central area of necrosis consisting of the necrotic remains of dead cells. Necrotising granulomas are associated with poorer patient outcomes and the necrotising core has been identified as a site of *M.tb* persistence and replication (Grosset, 2003, Pagán and Ramakrishnan, 2014, Rosenthal, 1955, Medlar, 1948). Large necrotising granulomas can collapse internally leading to the formation of cavities and the erosion of these cavities into bronchi is believed to play an important role in *M.tb* accessing the airways to continue its spread into additional hosts (Hunter et al., 2014, Hunter, 2011a). While the granuloma was long considered a host-protective structure that walls off *M.tb* to prevent further dissemination, recent studies have called this into question and the role of the granuloma in *M.tb* infection remains poorly understood.

1.2.4.1 Evidence for a protective role of the TB granuloma

Several factors contribute to the view that granulomas are intended as a host-protective structure organised in such a way to contain *M.tb*. The simplest granuloma type is the foreign body granuloma that appears to surround foreign objects too large for phagocytosis in order to enable their gradual digestion (Sheffield, 1990). It follows that the immune response to *M.tb* is utilising the same principal in an attempt to wall off the bacteria and infected cells

prior to their elimination. Additionally, sterile, fibrotic granulomas can be identified in the lungs of healthy patients suggesting that these represent successful immunity to M.tb and the fulfilment of granuloma function (Ramakrishnan, 2012). It also follows that necrotising granulomas containing replicating bacilli represent a breakdown in the host defence against M.tb. Furthermore, it has been shown that TNF is a key cytokine that directs macrophages in the formation of granulomas, and in its absence can result in the formation of loose and disorganised granulomas (Holly et al., 2005, Bean et al., 1999). Interrupting TNF production/signalling in both animal models and humans is associated with increased susceptibility to M.tb and it follows that the inability to form tight granulomas may be contributing to the poor immune response in these patients. Poor granuloma formation is also observed in response to a number of other immunocompromising deficiencies that are associated with heightened susceptibility to M.tb (Lawn et al., 2002, North and Izzo, 1993, Cooper et al., 1993, Sugawara et al., 2003, Scanga et al., 2004, Juffermans et al., 2000, Fremont et al., 2004, Fremont et al., 2007). However, in these cases, a clear causal link between poor granuloma formation and increased susceptibility has not been established and other evidence suggests that the bacteria themselves may directly benefit from granuloma formation.

1.2.4.2 Evidence for a detrimental role of the TB granuloma

Following the paradigm of granulomas as primarily host protective structures, there emerged evidence that granuloma formation may in fact be directly beneficial to bacterial persistence and replication and perhaps even orchestrated by M.tb itself. Rather than simply walling off M.tb, the tight barriers formed by macrophages within the layers of granulomas might serve to prevent the access of adaptive immune cells from reaching the infected cells located

centrally in the lesions and facilitating the killing of infected cells. Interestingly, infection with attenuated mycobacteria that lack a key virulence factor named the ESX-1 secretion system in the zebrafish model does not cause well-formed granulomas, nor does it cause as serious disease (Volkman et al., 2004). Macrophages that arrive at the site of infection in this model are also not as motile as macrophages arriving at the site of infection with unattenuated bacilli and are less susceptible to infection (Davis and Ramakrishnan, 2009). This suggests that unattenuated mycobacteria may have a role in directing additional macrophages to granulomas and orchestrating their infection. Similarly, the ESX-1 secretion system can also activate type 1 IFNs in mice (Stanley et al., 2007) and human cells (Novikov et al., 2011), that in turn recruit a population of macrophages to the granuloma that are highly permissive to M.tb infection in mice (Antonelli et al., 2010) and are associated with enhanced immunopathology in TB patients (Zhang et al., 2018).

1.2.4.3 Heterogeneity of granulomas

More recent studies in the role of the granuloma, particularly in the NHP model have revealed significant heterogeneity between individual granulomas, in their composition, structure and outcomes, even within the same animal in response to the same initial infection (Lin and Flynn, 2018). Rather than considering a host immunity to M.tb as a systemic response, it is becoming clear that there are diverse local immune responses and that each granuloma represents an individual battlefield between host and bacteria.

A combination of CT scans and positron emission tomography (PET) scans have been used in NHP TB to demonstrate that granulomas in the same animal are highly dynamic in both size and metabolic activity (Coleman et al., 2014, Lin et al., 2013). It has also been shown in the same model that sterilising immunity can be achieved in a subset of granulomas

while other granulomas in the same macaque show poorly contained bacterial replication (Lin et al., 2014). Some NHPs infected with M.tb go on to develop LTBI and granulomas and sterilisation was more common when compared to NHPs that developed active TB disease. Further evidence for the heterogeneous nature of responses to M.tb was achieved by infecting NHPs with a number of barcoded strains of M.tb, such that the bacterial source of M.tb in each granuloma could be identified and determined whether this was a primary granuloma or the result of dissemination from another site (Martin et al., 2017). It was shown using this model that only a subset of granulomas go on to cause dissemination and spread M.tb. Taken together, these findings suggest that each granuloma represents an individual battleground against the bacteria.

At the immunological level, sterilising NHP granulomas were not associated with the classic IFN γ producing T cells, rather with T cells that produce TNF, IL-10, or IL-17 (Gideon et al., 2015). This contrasting cytokine profile suggests a balance of pro- and anti-inflammatory cytokines are required for an effective anti-M.tb immune response. A more recent single cell sequencing study compared high and low-burden NHP granulomas and further revealed divergent cellular and molecular profiles associated with each type (Gideon et al., 2022). Low-burden granulomas were associated with Th1, T helper 17 (Th17) and CD8 T cells producing pro-inflammatory cytokines while high-burden granulomas were associated with mast and plasma cells as well as stromal cells and signalling for wound-healing and type 2 immune responses. The discovery of these divergent immune profiles will help guide future research into the factors underpinning granuloma trajectory and outcomes.

1.2.5 The Human TB granuloma

While our understanding of the TB granuloma has greatly expanded in the past few decades, this work has almost entirely been performed in animal models, of which relatively little has been verified in humans. These animal models have the advantages of precise control of the time of initial infection conditions and unhindered access to tissues for analysis. In comparison, given the chronic nature of M.tb infection, human studies into the TB granuloma rarely have any knowledge of for how long the infection has progressed since the patient was initially exposed to M.tb. Non-invasive approaches using PET/CT scans have been applied to map the architecture of human TB granulomas *in vivo* and have demonstrated changes in granuloma architecture through treatment including revealing granuloma resolution upon completion of treatment (Coleman et al., 2014). Much of our understanding of the human TB granuloma comes from *ex vivo* examination of lung tissue, usually removed as part of a lung resection as an end stage treatment for patients with drug resistant disease.

Micro computerised tomography (μ CT) is a new technology that enables CT scanning with resolution down to the level less than $1\mu\text{M}$. μ CT was applied in a recent study to generate images of the TB lesions in large patient biopsy samples and revealed that necrotising lesions adopt branching morphologies as opposed to the typically spherical morphology revealed by conventional 2-dimensional tissue analysis (Wells et al., 2021). The study further showed that these large branching lesions tend to follow the bronchioles and their morphology is likely shaped by airways. This study offers important guidance to other research using imaging of sections from human TB lung tissue as multiple branches of a single necrotising granuloma may appear as separate lesions when looking only at a single plane.

Another study (Mohlopheni et al., 2016) examined lung sections from human TB patients and performed laser capture microdissection, followed by mass spectrometry to characterise the molecular signatures of different spatial regions of solid, necrotising and cavitary granulomas. They revealed that a pro-inflammatory signalling environment characterises the central necrotic region of necrotising granulomas, while non-necrotising granulomas and the solid cuff of necrotising granulomas were characterised by anti-inflammatory signalling.

Two recent multiplex studies have performed single cell analysis on human TB granuloma sections. In one study using OPAL IF, it was shown that granulomas were enriched with myeloid cells and T cells, with low presence of B cells. They further showed that the density of CD68⁺ macrophages and CD8⁺ T cells was increased in granulomas that contained acid fast bacilli (AFB) compared to AFB negative granulomas (Abengozar-Muela et al., 2020). While this study was limited to examining six protein markers concurrently, another multiplex study used 37 parameter multiplex ion Beam imaging by time of flight (MIBI-TOF) to identify 19 unique cell-types in the TB infected lung (McCaffrey et al., 2022). They further identified 8 distinct microenvironments that comprise TB granulomas and surrounding lung tissue based on density and spatial localisation of immune cell subsets. They showed by in-situ hybridisation that the TB granuloma environment was depleted of IFNG transcripts and was rather enriched for transcription of immunoregulatory cytokine transforming growth factor β (TGF- β). They also revealed the presence of immunomodulatory cells within the granuloma including FOXP3⁺ regulatory T cells and PD-L1⁺ and IDO1⁺ myeloid cell subsets.

While the precise roles and effectiveness of the human TB granuloma have not yet been fully characterised, imaging approaches are providing deeper phenotyping than has been possible before and are revealing novel insights into human granuloma form and function.

1.3 Applications of transcriptomics to understand human disease

Transcriptomic technologies are used to identify and quantify genes that are expressed in a sample for a snapshot in time (Lowe et al., 2017). From recent initial applications in human studies (Adams et al., 1991, Pan et al., 2008, Sultan et al., 2008), transcriptomic analyses can now be routinely and rapidly performed at the tissue and single-cell level with ever-evolving technologies. While variations in gene expression at the messenger RNA (mRNA) level do not necessarily correlate with changes at the level of translated functional proteins (Vogel and Marcotte, 2012), changes in gene expression can still provide a plethora of scientific and clinically relevant findings in both coding and non-coding genes (Byron et al., 2016).

1.3.1 Bulk transcriptomics

Bulk transcriptomics refers to the measurement of gene expression at the level of the entire tissue sample and can be performed on both blood and homogenised solid tissues. Two major techniques for the study of bulk transcriptomics are the microarray, which matches RNA to an array of known nucleotide probes (Romanov et al., 2014, Barbulovic-Nad et al., 2006), and RNA sequencing (RNAseq) which identifies the sequence of each RNA molecule (Ozsolak and Milos, 2011, Wang et al., 2009). Both approaches can provide quantification of expression levels for each gene (Shalon et al., 1996, Lowe et al., 2017) however RNAseq can do so with a much higher dynamic range than microarrays (Rao et al., 2019). RNAseq can also be performed with far less genetic material compared to a microarray (Hashimshony et al., 2012, Stears et al., 2000) and provides a more unbiased insight at the transcriptome as it does not require the prior identification of sequences and probe generation that a microarray does (Mantione et al., 2014). While these two approaches have comprised the majority of bulk transcriptomic work over the past two decades, the latter has now emerged as the

dominant method for the measuring of bulk gene expression (Barton et al., 2017, Su et al., 2014).

1.3.1.1 Bulk transcriptomics analysis

The standard analysis protocol following bulk transcriptomics and data processing is to identify Differentially Expressed Genes (DEG) through differential gene expression analysis (Van Verk et al., 2013). DEGs are genes that are found to be either upregulated or downregulated between different sample conditions. While identifying DEGs alone can be informative and especially useful when investigating a particular gene of interest, transcriptomic analyses become more powerful when the transcriptome is analysed in terms of groups of genes known to be associated with the same biological functions (Subramanian et al., 2005). This is termed gene-set enrichment analysis (GSEA) and relies on annotated gene sets that identify genes known to be involved with particular biological processes and pathways (Gene Ontology, 2004). Performing GSEA can identify whether any gene-sets are overrepresented among DEGs between samples and can thus inform about biological changes occurring between study conditions as well as potential regulatory processes underpinning these changes (The Gene Ontology Consortium, 2019, Reimand et al., 2019). Another important downstream analysis protocol is module analysis. While GSEA relies on prior annotation of each gene with its known biological roles, module analysis is based on a network approach that identifies correlations in gene expression patterns across different conditions (Chaussabel et al., 2008). These modules can include large numbers of genes and can generate broad “immune fingerprints” of disease conditions and patient outcomes that can be readily and simply visualised for ease of comparison (Chaussabel and Baldwin, 2014, Altman et al., 2021). The statistical programming language, R (R Core Team, 2021) is the

software typically used for downstream transcriptomic analysis and there are many add on packages that can perform DEG (Trapnell et al., 2013, Robinson et al., 2010, Love et al., 2014, Ritchie et al., 2015), GSEA (Wu et al., 2021, Rue-Albrecht et al., 2016, Yu et al., 2015, Federico and Monti, 2020, Korotkevich et al., 2021, Manjang et al., 2020, Falcon and Gentleman, 2007) and module analysis (Rinchai et al., 2021, Langfelder and Horvath, 2008) available. However, numerous alternative software such as Enrichment Map and Cytoscape (Merico et al., 2010, Shannon et al., 2003) and online applications such as GEO2R and EnrichR (Kuleshov et al., 2016, Ge et al., 2018, Reyes et al., 2019, Li and Andrade, 2017, Raudvere et al., 2019) are also available and represent more accessible tools for performing these analyses.

1.3.1.2 Applications of bulk transcriptomics in disease diagnostics

Bulk transcriptomics approaches have been used in human studies for the last two decades and have proved enormously useful in providing deep phenotyping of diseased and healthy tissues. After initially being applied in cancer research to explore the differences between tumour and healthy tissue (Alon et al., 1999) and identify new cancer sub-types (Alizadeh et al., 2000, Golub et al., 1999, Bittner et al., 2000), transcriptomic approaches were soon adopted into inflammatory diseases with a significant early finding being that the disease systemic lupus erythematosus (SLE) is characterised by an upregulation of IFN stimulated genes (ISG) compared to healthy patients (Baechler et al., 2003, Bennett et al., 2003).

Since then, bulk transcriptomics approaches have driven the discovery of numerous blood transcriptional biomarkers in the quest to improve diagnostics of challenging inflammatory diseases such as systemic onset juvenile idiopathic arthritis (SOJIA) (Allantaz et al., 2007), Crohn's disease (Burczynski et al., 2006), diabetes (Kaizer et al., 2007), rheumatoid arthritis

(Olsen et al., 2004) and systemic inflammatory response syndrome (SIRS) (Sweeney et al., 2015). A major goal of such studies being the eventual development of rapid and non-invasive peripheral blood-based diagnostics. A similar approach has been taken in infectious diseases, with TB as a prominent example due to its current complex diagnostic process (Singhania et al., 2018b). Transcriptomic studies have been undertaken to distinguish TB from other respiratory illnesses (Berry et al., 2010, Bloom et al., 2013, Singhania et al., 2018a, Clayton et al., 2017, Maertzdorf et al., 2012, Anderson et al., 2014), to distinguish between active TB and LTBI (Kaforou et al., 2013, Jacobsen et al., 2007, Leong et al., 2018) and to distinguish different types of active TB (Blankley et al., 2016b). Similar approaches have also been utilised in infectious diseases such as sepsis (Sweeney et al., 2015, Tang et al., 2009a, Balamuth et al., 2020) and influenza (Voigt et al., 2018, Parnell et al., 2012, Woods et al., 2013, Tang et al., 2017).

1.3.1.3 Applications of bulk transcriptomics in predicting patient outcomes

In addition to diagnostics, transcriptional biomarkers have also been sought to differentiate between patient outcomes, in particular to stratify patients according to expected disease severity in advance and to predict responses to treatment. Stratifying patients according to predicted severity has been a particularly pressing need in sepsis due to the rapid progression and high mortality of the condition, and several predictive gene expression signatures have been identified (Baghela et al., 2022, AbdAllah et al., 2021, Wong et al., 2009, Banerjee et al., 2021, Wong et al., 2017, Mohammed et al., 2019) although none have yet made it to clinical application.

Using gene expression to measure and predict successful treatment has been an important goal in the fight against TB given the long treatment program necessary to cure the

disease and the prevalence of patients with poor responses (Walzl et al., 2018). Gene signatures have been identified that both correlate with successful TB treatment (Bloom et al., 2012, Cliff et al., 2013, Berry et al., 2010, van Doorn et al., 2022, Warsinske et al., 2018, Odia et al., 2020) and provide deep phenotyping to characterise an effective treatment response. Similarly, transcriptomics have been utilised to measure response to vaccines, such as those produced against the influenza A virus (IAV) to characterise both short term (Tan et al., 2014, Nakaya et al., 2011, Ovsyannikova et al., 2016) and long term (Henn et al., 2013, Nakaya et al., 2015) patient responses as well as to identify patient characteristics that correlate with strong vaccine responses (Thakar et al., 2015, Cortese et al., 2021, Tsang et al., 2014). These types of longitudinal studies are almost exclusively performed in peripheral blood or its constituents due to ease of accessibility and usually rely on measuring the same individuals across multiple time points.

1.3.1.4 Applications of bulk transcriptomics in longitudinal studies

Some of most ambitious bulk transcriptomics research in humans have involved using longitudinal studies to attempt to elucidate the transcriptional changes that occur at the point of infection or disease initiation (Pulendran and Davis, 2020). While it is unethical to deliberately infect subjects with deadly pathogens for the sake of research, these studies have identified healthy individuals at high risk of disease and formed cohorts of such subjects to take blood samples from regularly with the expectation that some individuals will progress to disease during the study (Barry and Mayer-Barber, 2021). The Females Rising through Education, Support, and Health (FRESH) study gathered a group of 300 women from a community with high HIV infection rates and provided them with tools for poverty alleviation while also acquiring blood samples twice weekly testing for HIV RNA (Ndung'u

et al., 2018). They were able to identify 12 individuals who acquired HIV during the study and for the first time characterised the cellular (Ndhlovu et al., 2015) and transcriptional (Anahtar et al., 2015) landscape immediately following initial infection with HIV, identifying genital inflammation as a potential risk factor in acquiring HIV. A similar approach was taken for studying the earliest transcriptional changes following initial infection with M.tb going on to active TB disease (Tabone et al., 2021). A cohort of several hundred household contacts of active TB patients who were considered at high risk of TB were recruited. They identified minimal transcriptional changes upon initial infection that gradually became more distinct as patients approached clinically diagnosable TB disease.

Another significant longitudinal study in TB research is the Adolescent Cohort Study, which aimed to elucidate the transcriptional changes that occur when a latently infected individual progresses to active TB, rather than looking at the initial events after infection (Zak et al., 2016). Almost 6000 adolescents were recruited, who were identified as latently infected with M.tb and blood samples were taken twice yearly over a two-year period. From these samples a gene signature was identified that could predict active TB disease for the 12 months before progression occurred. An additional approach that was applied to this cohort was cell-type deconvolution, which can be used to estimate the proportions of different cell types in a tissue based on the relative expression of different transcript sets (Vallania et al., 2018, Shen-Orr et al., 2010). This further revealed a decrease in the proportion of NK cells in the blood of individuals with LTBI in the year preceding active TB diagnosis (Roy Chowdhury et al., 2018).

While analysis protocols like cell type deconvolution can deepen the analysis that can be performed on gene expression data, bulk transcriptomic methods are still limited in their power to differentiate which cells in the tissue are contributing which transcripts to the total RNA milieu. To address this and to enable deep transcriptional phenotyping at the individual

cell level another methodology utilising RNAseq has been developed that can probe the transcriptome at the single cell level.

1.3.2 Single cell transcriptomics

The development of RNAseq enabled transcriptomic studies to be performed with far less quantity of RNA compared to microarrays and opened up the possibility of sequencing the transcriptome at the level of individual cells (Hashimshony et al., 2012). After the first demonstration that it was possible to sequence at this level in 2009 (Tang et al., 2009b), rapid advances in sequencing technologies (Aldridge and Teichmann, 2020, Zheng et al., 2017), microfluidics (Klein et al., 2015, Macosko et al., 2015) and robotics (Jaitin et al., 2014) have facilitated the single cell RNAseq (scRNAseq) technologies that are available today. The number of cells that can be sequenced per experiment has increased over time from dozens to hundreds of thousands (Svensson et al., 2018) while the cost has fallen to only cents per cell (Chambers et al., 2019).

1.3.2.1 Single cell transcriptomics workflow

While there are multiple approaches, in brief the latest single cell RNAseq workflows function by separating and lysing individual cells, enriching for mRNA out of other RNA types in the tissue, such as ribosomal RNA, reverse transcribing RNA to complementary DNA (cDNA), uniquely barcoding the cDNA belonging to each cell and each sample and then pooling transcripts to run RNAseq (Ding et al., 2020, See et al., 2018). New multi-modal approaches are also facilitating concurrent measurement of DNA or protein along with RNA to overcome implicit limitations in only studying genetic transcripts (Lee et al., 2020). While

still a relatively new technique, scRNAseq has already had a profound impact on human immunology (Ginhoux et al., 2022).

1.3.2.2 Applications of single cell transcriptomics

Probing gene expression at the level of thousands of individual cells at tissue level has provided an unparalleled opportunity for unbiased cell classification. Indeed, scRNAseq has enabled the discovery of new subtypes of DCs (Villani et al., 2017, Breton et al., 2016) and innate lymphoid cells (ILC) (Björklund Å et al., 2016) by deep phenotyping far beyond what is practical at the protein level alone. While bulk transcriptomics studies rely on longitudinal studies to probe the immune landscape over time, a distinct advantage of scRNAseq studies is that cells of each type, including those at various stages along their developmental pathways, may all be present in a single sample. This has given rise to the field of cellular trajectory analyses termed “pseudotime” that attempts to map cells from scRNAseq studies across a developmental continuum using both longitudinal and single timepoint studies (Stubbington et al., 2017). Pseudotime approaches have been used to generate development / differential profiles of both myeloid (Drissen et al., 2016, Schlitzer et al., 2015, See et al., 2017, Olsson et al., 2016) and lymphoid cells (Stubbington et al., 2016, Yu et al., 2016, Setty et al., 2016) under homeostatic conditions. The approach can also be utilised to reveal the developmental changes that occur during infection and disease (Stephenson et al., 2021, Schulte-Schrepping et al., 2020, Pisu et al., 2021).

Single cell transcriptomic approaches also offer more comprehensive exploration of clinical data compared to bulk transcriptomics. For example, It has long been shown that there is variation in the transcriptional environment between tumour and healthy tissue (Alon et al., 1999). However, an early scRNAseq study looking at melanoma skin biopsies revealed

inter- and intra-patient heterogeneity in tumour cell phenotypes (Tirosh et al., 2016) and identified sub-populations of tumour cells associated with the increased resistance to therapy within the total tumour microenvironment. Many subsequent studies continued to use scRNAseq to probe the tumour microenvironment and have identified diverse immune functions and regulation as well as cellular biomarkers associated with successful and poor patient outcomes (Azizi et al., 2018, Li et al., 2019, Ren et al., 2021).

It is important to note that while single cell sequencing technologies do offer considerably deeper sample phenotyping compared to bulk sequencing, they are unlikely to replace bulk transcriptomic approaches, rather bulk sequencing retains several key strengths going into the future (Li and Wang, 2021). Single cell sequencing remains much more expensive than bulk sequencing per sample and is still prohibitively expensive for most scRNAseq studies to utilise more than a handful of samples and bulk transcriptomics will likely remain the preferred option for studying large sample sizes (Sun et al., 2021). For studies seeking tissue-level biomarkers, bulk cell approaches will likely remain more than sufficient. Furthermore, the relative ease of bulk transcriptomic approaches may make them better suited for potential clinical applications (Kuksin et al., 2021). An additional drawback of scRNAseq is that it requires far more computing power to process the data compared to bulk sequencing, further prohibiting its use from potential researchers. Both single cell and bulk transcriptomics still produce massive datasets following processing and these have presented an unparalleled opportunity for scientists to collate and share data in future studies.

1.3.3 Mining public transcriptomic data

Each transcriptomic dataset generated is probed by the original researchers, however the sheer scale of the data generated means there will remain vast swathes of it that were

unexplored because they are unnecessary for directly answering the original research hypotheses. To prevent all this data from remaining underutilised, several online repositories have been created where entire datasets can be uploaded once the original researchers have finished their investigations (Kolesnikov et al., 2015, Hruz et al., 2008). These repositories allow the datasets to be freely downloaded and used by other researchers to explore their own hypotheses and are particularly useful to bioinformaticians who may not perform bench science themselves. The largest of these repositories, the Gene Expression Omnibus (GEO) contains over 180,000 unique datasets comprising data from over 5 million individual samples (Edgar et al., 2002), while many of the other smaller repositories are targeted toward researchers in particular fields (Heng et al., 2008, Tomczak et al., 2015, Bhattacharya et al., 2014, Hudson et al., 2010).

In addition, many online tools have been built to allow scientists to run re-analyses of public data without needing to use advanced software like R (Liberzon et al., 2011, Coletta et al., 2012, Zhu et al., 2015, Petryszak et al., 2016, Mahi et al., 2019). GEO itself has a built-in tool called GEO2R that facilitates rapid DEG analysis between any groupings of samples from any microarray datasets held within the repository while additionally providing the R code used to generate the analysis to the user (Barrett et al., 2013). Another dataset called Gene Expression Browser (GXB) focusses on human immunology datasets and provides analysis tools to rank genes according to multiple parameters, such as selecting only genes relating to certain biological functions, while also comparing the transcriptome between samples grouped by multiple covariate features (age, gender etc) (Speake et al., 2015, Bougarn et al., 2019).

1.3.3.1 Meta-analyses of public transcriptomic data

The first examples of reuse of transcriptomic datasets were in cancer studies (Segal et al., 2004, Segal et al., 2005). The practice soon found a place in human immunology research, as for example, the prior mentioned IFN signature in SLE was also identified in Sjogren's syndrome and rheumatoid arthritis through reuse of pre-existing datasets (Higgs et al., 2012, Toro-Domínguez et al., 2014). While reuse of transcriptomic datasets in studying human disease can be performed by reinterrogating single datasets, (Park and Lee, 2020, Kim et al., 2021), either as the basis of a study or as an aid to investigating an overarching hypothesis, the most powerful application of transcriptomic re-analysis come in the form of meta-analyses.

Meta-analyses involve pooling transcriptomic data across multiple cohorts and attempting to find consistent signals across studies (Sweeney et al., 2017). Meta-analyses of transcriptomics datasets can increase the rigor of gene signatures and other findings in single studies by testing them across multiple cohorts of samples from the same conditions, which can dramatically increase the sample size from any individual study and overcome study-specific variations, such as the technological platform used to perform each individual assay (Ramilo et al., 2007, Ramasamy et al., 2008, Lofgren et al., 2016, Robinson et al., 2019, Sweeney et al., 2016). Performing meta-analyses across datasets covering multiple diseases, tissue types, patient backgrounds etc, can further increase the robustness and sensitivity of individual biomarkers and gene signatures, while also highlighting disease-shared and disease-specific transcriptional features (Andres-Terre et al., 2015).

1.3.3.2 Applications of meta-analyses of public transcriptomic data

The earliest transcriptional meta-analyses were performed in cancer research and revealed the transcriptional profile of tumours could be compared between different cancers and revealed common transcriptional changes that occur in malignancy even across different types of cancer (Rhodes et al., 2002, Rhodes et al., 2004, Wirapati et al., 2008). More recently cell-type deconvolution has been performed across cancer types to explore the role of infiltrating leukocytes into the tumour microenvironment (Bindea et al., 2013, Li et al., 2016). A 2015 meta-analysis across patient samples from 39 cancer subtypes revealed through cell type deconvolution that an increased frequency of T cells is associated with better patient outcomes across cancers (Gentles et al., 2015). Another study developed a transcriptional profile that could estimate the cytolytic effect of CD8 T cells and NK cells based on mRNA of cytolytic granule proteins and showed that an increase in this signature was associated with better outcomes and other positive prognostic markers across multiple cancer types. Together, these studies demonstrate the power of public data meta-analyses to draw conclusions that are not feasible from single studies.

While meta-analyses are highly suited for comparing transcriptional responses between cancer types, they are also well suited for comparisons between different infectious diseases and responses to vaccination (Tsalik et al., 2016, Blankley et al., 2016a). A 2014 study measured transcriptional responses in patients following two different meningococcal vaccines and then compared those with publicly available datasets following vaccination against influenza and yellow fever. They uncovered shared transcriptional modules across vaccines from different diseases. Interestingly, however transcriptional signatures diverged depending on the type of vaccine (protein-based, polysaccharide-based, or live organism) (Li et al., 2014).

Meta-analyses have also been performed to strengthen signatures from disease classification by combining transcriptional datasets across patients from multiple diseases. A 2015 study compared the transcriptional profiles across a variety of respiratory pathogens and identified a common host signature of viral infection that could differentiate viral infection from bacterial infection (Andres-Terre et al., 2015). They also identified a signature that could classify pre-symptomatic patients from healthy controls as well as patients infected with IAV from other respiratory pathogens, highlighting the multi-faceted analytic power that multi-cohort meta-analyses can provide. Another study combined 27 datasets to develop a robust gene expression signature that can distinguish between infectious and non-infectious sources of inflammation in critically ill patients (Sweeney et al., 2015).

Meta-analyses can also be performed to test non-transcriptional findings generated in a research project. A recent imaging study identified myeloid cells expressing immunoregulatory molecules programmed death ligand 1 (PD-L1) and Indoleamine 2,3-dioxygenase (IDO1) in the lungs of TB patients (McCaffrey et al., 2022). Using several cohorts of public transcriptomic datasets, they showed that expression of the genes underpinning protein expression of IDO1 and PD-L1 were also upregulated in the blood of active TB patients compared to both healthy controls and latently infected individuals. They further showed that expression of these genes returns to baseline following end of treatment. They additionally reanalysed the previously mentioned adolescent cohort study of individuals latently infected with M.tb and showed that expression of PD-L1 increases in the blood over the 8.5 months before latently infected individuals progress to active TB and that this single gene alone provided a predictive power for progression as strong as the previously generated gene signature by the original researchers (Zak et al., 2016). Taken together it is evident that meta-analyses of public transcriptomic datasets can greatly increase the power of gene

expression studies while also supporting and validating immunological findings using other methodological approaches.

1.3.3.3 Accessibility of public transcriptomic data to immunologists

One of the challenges preventing a broader implementation of public data to the toolbox of immunologists used to working solely in the laboratory is that most of the analysis pipelines used to perform transcriptomic meta-analyses still require knowledge of advanced software and thus the field is still largely the domain of bioinformaticians. The R package MetaIntegrator, for example, is a comprehensive platform for performing transcriptomic meta-analyses. It facilitates direct download of any number of the tens of thousands of datasets from GEO and can sort samples and studies according to any chosen arrangement. It can then identify and validate gene signatures across multiple studies, all in a matter of minutes (Haynes et al., 2017). While operating the program only requires a basic knowledge of R, the need to work in the R programming environment is still highly restrictive and has served to limit the role of public data in human immunology. However online platforms are emerging that can facilitate meta-analyses, with one example by the same team that produced the MetaIntegrator package called MetaSignature (Haynes et al., 2017). While MetaSignature currently only has access to about 700 datasets covering 106 human diseases, it has the power to perform meta-analyses of a gene signature of interest between diseases or across datasets of the same disease. It can perform these operations in seconds with a user-friendly interface and produces an assortment of plots to visually display the data. It can also identify genes with the strongest effect size across datasets and can perform meta-analysis of individual genes across diseases in the case of investigations targeted to a single gene of interest. MetaSignature and similar programs represent a promising new type of analysis tools that

may facilitate a widespread uptake of data reuse and meta-analysis in the future of human immunology.

1.3.4 Using public transcriptomic data to characterise novel gene expression

The free accessibility of vast numbers of public transcriptomic datasets covering a plethora of tissue types, diseases and experimental conditions have also given rise to a genre of studies that seek to characterise poorly understood genes using these datasets. In such studies, the expression of a gene of interest can be tested in datasets covering a variety of diseases to identify conditions upon which its expression is differentially regulated (Huang et al., 2021). Single cell datasets and bulk datasets based on separated cell populations can also be used to identify the major cell populations that express the gene of interest under both healthy and diseased conditions. Furthermore, the regulatory mechanisms that directly underpin expression of the gene of interest can be probed using human challenge datasets or ex-vivo cell stimulation datasets. Such studies highlight the strengths of using public data as multi-faceted hypotheses can be tested entirely in silico (Rinchai et al., 2015).

For example, one study used a meta-analysis of bulk transcriptomic datasets to examine the gene Annexin A3 (ANXA3) and showed that the gene is robustly upregulated in the blood during sepsis and that it could also act as a potential biomarker to differentiate between sepsis survivors and non-survivors (Toufiq et al., 2020). They further demonstrated using another bulk transcriptomic dataset of purified cells from sepsis patients that the upregulation of ANXA3 in the whole blood actually corresponds to gene upregulation in neutrophils. Using another dataset that stimulated whole blood with a series of host- and pathogen-specific immunostimulatory molecules they finally showed that ANXA3 upregulation was driven by IFN γ . The fact that such a study can be performed entirely in

silico is a testimony to the comprehensive range of public datasets available. Additionally, all the software and data utilised in that study are freely available online and are accessible to any laboratory groups to access.

Another study looking at the gene nudix hydrolase 16 (NUDT16), similarly identified an upregulation at the whole blood level in sepsis and identified that whole blood expression of the gene was primarily driven by expression in neutrophils and monocytes using the same dataset generated using purified cell populations (Huang et al., 2022). In this case, to assess the regulatory mechanisms underpinning changes in expression they moved to the bench and employed blood samples from human donors. At the whole blood level, they found that stimulation with lipopolysaccharide (LPS) and peptidoglycan (PGN) led to a six-fold change increase in NUTD16 expression. However, upon performing flow cytometry and separating NUTD16⁺ cells from whole blood both before and after stimulation they found that while at the total leukocyte level NUTD16 expression increased upon stimulation, NUTD16 expression decreased in T cells, highlighting a complicated regulation of the gene across leukocyte subsets. Overall, this study highlights the role that public data can play in guiding experimental design and hypothesis generation before commencing laboratory research.

Public data represents an invaluable resource for all biomedical researchers, providing the full datasets from some of the most expensive and comprehensive experiments of the past decade right into the hands of every scientist for reanalysis. Many journals now require that transcriptomic datasets must be uploaded to online repositories, ensuring that the data does not sit idly in the hard drives of the original researchers, but is rather in the public domain to the benefit of any researchers who wish to use it. As the use of public data becomes more widespread and user-friendly programs become more commonplace and comprehensive in their functionality, public datasets may reshape the way hypotheses are generated and experiments are designed into the future.

1.4 Utilising tissue imaging to advance human immunology

While the aforementioned transcriptomic approaches to human immunology can provide deep phenotyping of solid tissues biopsies, down to the level of gene expression in single cells, they rely on prior tissue dissociation to access RNA and cells. This means that any spatial information about the tissue structure and cellular organisation is lost. However, the cellular spatial context of tissue can reveal important prognostic and diagnostic information, particularly in cancer (Maniatis et al., 2021). For example, the extent of leukocyte infiltration into a tumour correlates positively with survival rate of melanoma patients (Attrill et al., 2022a). However, this information cannot be ascertained by conventional transcriptomic approaches with homogenised tissue. Furthermore, there is variation in the rate that different cell populations are recovered following tissue dissociation, and cell-type specific losses can decrease the reliability of downstream transcriptomic analyses (Hickey et al., 2022). Single cell transcriptomic approaches are also unable to directly identify cell-cell interactions, and while these can be inferred based on gene expression of receptors and co-receptors (Armingol et al., 2021), it has also been identified that tissue dissociation can cause rapid transcriptional changes in some cell populations (van den Brink et al., 2017), which may in turn limit the power of transcriptomic interaction inferences. Therefore, approaches that can phenotype cells in situ will remain necessary to complement bulk and single cell transcriptomic methods.

Image based spatial proteomics and transcriptomics approaches are the most widely used methods to study undissociated solid tissue biopsies and have developed rapidly over the past two decades. While transcriptomics detect nucleic acids, the majority of imaging approaches detect proteins using antibody-antigen hybridisation. Single-plex staining, that is limited to detecting a single protein of interest per stain, has long been the norm in both

research and clinical applications (Tan et al., 2020). However, multiplex approaches that can detect multiple proteins in the same tissue using multispectral fluorophores and time of flight (TOF) based separation are now routinely detecting up to nine protein markers at the whole tissue level and fifty or more in selected regions of interests (ROI). Spatial transcriptomic approaches are also allowing gene expression across the entire transcriptome to be probed within spatially resolved tissue areas. New imaging methods continue to drive the field toward highly multiplexed, whole slide imaging with single cell resolution of both protein and nucleic acids. As technology develops, increasingly powerful software solutions are required to process and analyse the data, and the field of spatial biology is emerging to meet the challenge of producing meaningful findings out of the vast array of spatial cellular information that can be produced in a single study.

1.4.1 Conventional immunofluorescence (IF)

It is approaching one hundred years since the first application of immunohistochemistry (IHC). The combination of antibodies specific to pneumococcal antigens with a tagged fluorescent molecule called Fluorescein isothiocyanate (FITC) allowed for visualisation of the bacterial antigens within infected tissue for the first time (Coons et al., 1941). IHC includes staining with both chromogenic markers and fluorescent markers, which is also called IF. Since its inception, IHC has become a standard tool for scientists and clinical pathologists alike. The ability to visualise proteins of interest in situ has been particularly valuable for studying the pathology of malignancies as cancer sub-types and expected sensitivity to treatment are often defined by their expression or upregulation of particular proteins (Duraiyan et al., 2012). IHC has proved particularly useful in identifying non-primary tumours of unknown origin through their antigen expression (Lin and Liu, 2014).

IHC has also proved itself in the field of neuropathology, replacing many other more complicated and less sensitive diagnostic methodologies in identifying infectious, neurodegenerative, and cancerous causes of disease (Kleinman et al., 1994). Overall traditional IHC is a rapid, cheap, and reliable approach to phenotyping cells in solid tissue in both the laboratory and the clinic.

Multiplex IHC is typically performed as IF, tagging each marker to a different fluorophore with non-overlapping fluorescent profiles. However, the number of markers that can be visualised simultaneously with traditional multiplex IHC is limited by several factors. To prevent spectral overlap in IF, a maximum of three fluorescent channels can typically be used in addition to an ultraviolet (UV) channel (that is usually reserved for a nuclear stain to identify individual cells). This has long limited the number of parameters in IF, resulting in only a small number of markers that can be imaged per slide (Joshi and Yu, 2017). The number of parameters in IF is also limited by potential antibody cross-reactivity. All primary antibodies must be raised in different species to prevent the species-specific secondary antibodies from binding to more than one primary antibody and causing multiple epitopes to appear in the same fluorescent channel. Given the relatively low number of species in which antibodies are generated, this also places an inherent limit on multiplex IF. IF staining also suffers from photobleaching, which refers to the gradual degradation of fluorescent molecules following excitation. Photobleaching of fluorophores also occurs rapidly after staining from exposure to light and limits the timescale that IF samples can be kept before imaging (Brodin et al., 2013).

However, several new IF methodologies are emerging that are revolutionising whole tissue IF imaging by using novel approaches to circumvent these limitations while maintaining image quality and are enabling multiplexing with ever increasing numbers of parameters.

1.4.2 OPAL multiplex imaging system

The OPAL multiplex system was introduced a decade ago and allows up to eight unique protein epitopes as well as cell nuclei to be simultaneously visualised (Gorris et al., 2018). To achieve this, the OPAL system relies on the principle of tyramide signal amplification (TSA) at the end of each staining round. Each OPAL fluorophore is bound to an inactive tyramide molecule that becomes active after encountering horseradish peroxidase (HRP)-bound secondary antibodies. The reactive tyramide-fluorophore complex then covalently binds to tyrosine residues in the vicinity of the antibody complex, many of which are within the tissue itself. This allows multiple fluorophores to bind for each antibody, greatly amplifying the fluorescence signal (Viratham Pulsawatdi et al., 2020). The tissue is then subjected to heat-induced epitope retrieval (HIER) that deactivates and strips each antibody complex while leaving the covalently bound fluorophores still attached to the tissue (Stack et al., 2014). Each fluorophore added remains attached to the tissue through additional staining rounds and HIER (Willemsen et al., 2021). As all previously bound primary antibodies are removed during HIER at the end of each staining round, subsequent rounds are free to include primary antibodies raised in the same species, removing a key limitation of traditional IF.

The OPAL system also relies on the principle of linear unmixing to overcome the spectral overlap that limits the number of fluorophores in traditional IF. Linear unmixing is based on the principle that the abundance of signal from each fluorophore will produce consistent wavelengths when mixed depending on the proportional contributions of each colour (Stack et al., 2014). Computer software is therefore able to unmix the fluorescence profile of each channel back into the contributing profiles of each fluorophore. The software that performs linear unmixing for OPAL is called inForm (Perkin Elmer) and relies on previously being taught the profile of each fluorophore via a series of slides stained for only one fluorophore each called single colour controls (SCC). Using linear unmixing now allows

the OPAL system to routinely image eight different proteins in a single tissue. The Vectra imaging system (Akoya) that accompanies the OPAL protocol has a built-in slide scanner that sequentially captures 20x high resolution images across entire tissue sections. Individual image tiles can then be stitched in downstream analysis using HALO software (Indica Labs) to produce single whole slide images, which can comprise an area covering over one million individual cells (Feng et al., 2015). A distinct advantage of OPAL analysis, therefore, is that it can facilitate both tissue-level and single cell analyses in the same image. OPAL can also work with both frozen and formalin-fixed paraffin-embedded (FFPE) sections.

1.4.2.1 Applications of OPAL multiplex imaging

The primary application of OPAL staining has been in cancer studies and has enabled researchers to study both tumour tissue as well as the accompanying peri-tumour area with single cell resolution. Cancer researchers have also taken advantage of the large tissue area offered by OPAL and the Vectra imaging system to perform multiplex staining with tissue microarrays (TMA). With the TMA approach, small regions from dozens of different tumours are arrayed on a single slide, stained simultaneously with OPAL fluorophores, and imaged together in a single slide image with the Vectra system (Rimm et al., 2001).

Individual samples can be later separated during image analysis. This approach allows rapid phenotyping of dozens of tumour samples and offers a promising application of the OPAL multiplex system in clinical pathology (Tan-Garcia et al., 2020).

The multiplex capability of OPAL staining has been used in cancer studies to identify biomarkers of patient prognosis (Sousa et al., 2022, Ansong et al., 2015, Damodaran et al., 2017, Lee et al., 2021) such as that the ratio of CD4⁺ T cells to PDL-1 expressing cells was correlated with disease regression in rectal cancer (Park et al., 2017) and that expression of

PDL-1 and T-cell immunoglobulin and mucin domain-3 (TIM-3) on tumour infiltrating CD8⁺ T cells was associated with larger tumour size and poorer patient outcomes in renal cell carcinoma (Granier et al., 2017). Single cell data acquired through OPAL staining can also use spatial analytics, such as cell interactions as potential biomarkers (Vanguri et al., 2022). For example, it has been shown that expression of programmed cell death protein 1 (PD-1) or PD-L1 alone are not sufficient biomarkers of responsiveness to checkpoint inhibitor therapy in metastatic melanoma. However, a strong interaction score between cells expressing each of the two molecules was a good prognostic biomarker of response to treatment (Johnson et al., 2018).

While OPAL multiplex imaging offers clear advantages over single-plex IHC, the total number of multiplexing parameters is limited by spectral overlap, which prevents channels from being effectively unmixed when too many fluorophores are used. New fluorescence-based imaging technologies are instead relying on multiple imaging rounds that strip off fluorophores between each set of image captures.

1.4.3 Co-detection by indexing (CODEX)

CODEX is a newly developed fluorescence-based imaging technology that can perform imaging with more than 50 different epitopes in a single tissue (Black et al., 2021). Codex is able to achieve such highly multiplexed fluorescence imaging by conjugating each primary antibody to a unique polynucleotide barcode sequence, allowing a theoretically infinite number of barcoded primary antibodies to be simultaneously bound to the tissue (Wang et al., 2021). Fluorophores are attached to probe polynucleotide sequences, each complimentary to one primary barcode sequence. Fluorophore hybridisation and imaging are then performed in cycles of three spectrally separate fluorophores at a time (Tan et al., 2020). At the end of each

imaging round the fluorophores are removed in preparation for the subsequent round, allowing the same three fluorophores to be used in each cycle, only varying in their probe sequence in each iteration. The use of an automated microfluidics system allows the cycles of fluorophore hybridisation, imaging, and fluorescence removal to be performed without user input and greatly increases the efficiency of the process. This novel approach entirely removes the issue of spectral overlap that typically comes with increasing multiplexity in fluorescence staining.

Besides the ability to phenotype far more cell types and functional markers than other IF protocols, the high multiplexed nature of CODEX allows for the identification of distinct microenvironments within the broader tissue landscape. Microenvironments are calculated based on patterns in the composition and spatial arrangement of cells and such calculations are made possible in CODEX and other highly multiplexed imaging approaches by the plethora of cell types and functional distinctions that can be identified. For example, a study looking at colorectal cancer using a 56 parameter CODEX panel, identified 9 distinct microenvironments (termed cellular neighbourhoods in the study) in the tumour and peritumoral area and identified that an enrichment of PD1⁺ CD4⁺ T cells in only one tissue microenvironment was associated with improved patient outcomes (Schürch et al., 2020), highlighting the power of highly multiplexed imaging.

1.4.4 Mass Based multiplex imaging

Besides fluorescence-based imaging methods like CODEX, mass-based, time-of-flight imaging offer an alternative methodology for highly multiplexed imaging. The two commonly implemented mass-based multiplex imaging approaches, Imaging mass cytometry (IMC) (Giesen et al., 2014) and multiplex ion beam imaging (MIBI) (Angelo et al., 2014)

both facilitate highly multiplexed imaging of up to 40 parameters. They differ from fluorescence imaging in that primary antibodies are tagged with metal isotopes of different molecular weights and the varying weights of these ionised metals following tissue ablation facilitate epitope differentiation during imaging (Baharlou et al., 2019). Both approaches ablate the tissue sections in small spots at a time, via a laser in IMC and an ion beam in MIBI, and the ions are fed into a mass spectrometer that determines the makeup of primary antibodies bound to the tissue in that region based on the time of flight of the ions liberated during ablation (Keren et al., 2018). While IMC can provide highly multiplexed images using a relatively simple staining and imaging process, it is limited by the inability to image large areas of tissue and is typically only able to phenotype small ROIs. Furthermore, unlike fluorescence-based approaches, IMC imaging is destructive to the tissue areas visualised, preventing any future imaging of the same regions. Both IMC and MIBI were introduced earlier than CODEX and so far, have found broad usage in characterising microenvironments in cancer (Gerdtsen et al., 2018, Schulz et al., 2018, Malihi et al., 2018, Scher et al., 2016), as well as in autoimmune conditions (Damond et al., 2019, Wang et al., 2019, Ramaglia et al., 2019, Park et al., 2019) and more recently, TB (McCaffrey et al., 2022).

The emergence of multiplexed fluorescence and TOF based imaging methods are greatly expanding the number of protein epitopes than can be simultaneously studied in situ, however they are unlikely to approach the highly multiplexed nature of transcriptomics in the near future. However, recent advances have opened up the possibility of studying the entire transcriptome in situ at the single cell level.

1.4.5 Spatially resolved transcriptomics

Measuring gene expression in tissue has long been possible through technologies like fluorescence in-situ hybridisation (FISH) that identify individual RNA molecules using fluorescent bound probe sequences, however these approaches have been limited by the same constraints on multiplexing as traditional IF (Lewis et al., 2021). The introduction of spatially resolved transcriptomic approaches now enable expression of thousands of different genes to be probed at single cell resolution using cyclic staining with fluorescent barcoded probes. Spatial information can also be applied to scRNAseq studies by either barcoding spatially resolved tissue regions in situ before cell dissociation, or through dissecting tissue regions and performing independent sequencing runs for each tissue region (Asp et al., 2020). These approaches now enable the entire transcriptome to be probed with near single cell spatial resolution (Liao et al., 2021). One notable recent example of this approach performed in human leprosy samples shed light on the layered organisation of inflammatory signalling in the tissue granulomas (Ma et al., 2021). While these approaches can provide the most comprehensive phenotyping of spatially localised tissue regions, they lack the ability to perform single-cell analysis which limits their capacity to perform cellular spatial analyses.

1.5 Spatial Biology

Multiplex imaging approaches to human immunology now enable the capture of over a million cells from a single piece of tissue with single cell resolution (Tan et al., 2020). The boundaries between cells can be delineated, and individual cells quantified through the process of cell segmentation, which produces enormous datasets comprising both the spatial locations of each cell along with its phenotypic information (van Dam et al., 2022). While such datasets offer an invaluable tool in phenotyping diseased tissue, they require powerful software to process and the appropriate algorithms to quantify tissue microenvironments and relationships between cells. In recent years the field of spatial biology has emerged to grapple with these needs and a number of effective software and algorithmic approaches have been developed to effectively process single cell spatial data (Hofman et al., 2019).

1.5.1 Cell segmentation

Cell segmentation is the process used identify the boundaries of single cells in images, delineate cells from each other, and is an essential step for many downstream cellular analyses. Cell segmentation can be performed based on a membrane stain to divide cells using their outer surfaces or alternatively using a nuclear stain to segment nuclei and use individual nuclei as a basis to infer cell boundaries (Hollandi et al., 2022, Schüffler et al., 2015). Membrane-based segmentation approaches are often applied to study in vitro cell culture, while segmentation based on nuclei is the more practical method in solid tissue imaging, given nuclei tend to be spatially separated from one another while membranes are not (Han et al., 2022). The quantitative output of cell segmentation algorithms enable rapid cell counts across large tissue areas and can broadly phenotype the composition of different tissue areas. Cell segmentation also quantifies the XY coordinates of each segmented cell in

the tissue which enables a number of spatial tools to measure distances of cells from major tissue structures (such as a tumour) or to quantify cell-cell relationships (van Dam et al., 2022).

1.5.2 Cell density

A major advantage of image-based approaches to human immunology is that spatial and phenotypic information can be combined together to increase the power of both types of data. One of the simplest methods to apply this is to measure the density of particular cell-types of interest and make comparisons between different tissue regions, for example within a tumour compared to the surrounding tissue (Meylan et al., 2022, Galon et al., 2006, Ozbek et al., 2022). Mapping cell density across large areas of tissue can also be used to visually identify expression patterns of cellular markers (Wada et al., 2006), such as when exploring the depth of penetration of infiltrating immune cells into a tumour (Masugi et al., 2019). While density-based analyses can allow for broad comparisons of cell arrangement within tissue, there are other methods more suited to exact quantification of cell locations within tissue structures.

1.5.3 Infiltration Analysis

Infiltration analysis is the most common name given to analysis methods that measure the distance of cells from defined tissue regions. It is most often used to quantify the distance of penetrating leukocytes from the tumour margin either into the intra-tumoral space or outward in the peri-tumour area, which can be a prognostic biomarker in some cancers (Pagès et al., 2010, Grabovska et al., 2020). Infiltration analysis requires the tumour margin to first be annotated and is then typically performed using a series of software generated bands at fixed distances outward or inward from the tumour margin. The density of the cell of interest is

calculated in each of these bands to produce a density histogram for comparative analysis with other samples/tumours/cells (Attrill et al., 2022a). While infiltration analysis is typically applied to study the relationship between leukocyte populations to a tumour, it can also be used to measure the distance from any cell population to any annotated tissue structure.

1.5.4 Neighbourhood analysis

Neighbourhood analysis refers to the quantification of spatial relationships between different cell populations. A cell-type of interest is typically chosen and the distance of each individual cell to the nearest cell of another cell population is quantified. Taking the average of all these distances can give a single readout to quantify the relationship between the two cell populations and from this can be inferred how closely the two populations are interacting with each other. While neighbourhood analysis does not quantify direct cell-cell interactions, the results can be prognostically significant biomarkers, and in cancer studies it has been frequently shown that differences in the relationship between cell populations across patients correlate with patient outcomes (Attrill et al., 2022b, Attrill et al., 2022a, Thomas et al., 2019, Lee et al., 2019, van Dam et al., 2022).

1.5.5 Spatial analysis software

There are a number of software solutions available for processing and analysing highly multiplexed whole slide images. These range from open-source software requiring source code entry to commercially available software with a more user-friendly graphical user interface (GUI) (Tan et al., 2020). HALO (Indica Labs) is one such commercially available software package that is often used in conjunction with images taken through the Vectra and CODEX imaging platforms. HALO can perform image stitching of tiled images taken with

the Vectra or other platforms to produce any sized whole slide images. HALO can perform cell segmentation based on nuclear, cytoplasm or membrane staining, and tissue can be annotated manually into regions, or alternatively this can be achieved based on cell density or using a tissue classifier that identifies regions with similar cell compositions. HALO has a number of spatial plug-ins that can be used to perform cellular spatial analyses including density calculations, infiltration analysis and neighbourhood analysis. Halo has been applied in many studies looking at the tumour environment both in whole slide images and TMAs (Attrill et al., 2022a, Vanguri et al., 2022, Hofman et al., 2019, Attrill et al., 2022b).

1.6 The Present Study

The present study aims to expand the analytic approaches available to human immunologists in the areas of transcriptomics and multiplex image analysis, with a focus on defining the immune function in infectious diseases.

With regard to transcriptomics this study aims to

- Develop a pipeline for public transcriptomic data meta-analyses that can identify disease association, cell-type specificity, and regulatory mechanisms of genes of interest using bulk and single-cell datasets
- Apply this pipeline to characterise an unknown gene associated with the inflammatory immune response to infection

With regard to imaging this study aims to

- Develop a pipeline for delineating individual granulomas in whole slide multiplex imaging and performing intra and inter-granuloma spatial and compositional single-cell analysis
- Apply this pipeline to assess the diversity and spatial arrangement of human TB granulomas

Chapter 2

General methods

2.1 OPAL staining

FFPE TB lung sections were baked at 65°C for a minimum of 1 hour and dewaxed and rehydrated through gradients of xylene and ethanol concentrations (xylene x 10 mins, 100% ethanol x 10 mins, 90% ethanol x 10 mins, 70% ethanol x 10 mins, 50% ethanol wash). Antigen retrieval was performed by boiling samples to 100°C for 15 minutes in pH 9 Antigen retrieval buffer (Akoya). Tissues were incubated for 45 minutes at room temperature with primary antibodies specific for CD20 (L26, Cell Marque, 1:1000), CD4 (SP35, Cell Marque, 1:100), CD68 (Kp-1, Cell Marque, 1:500), CD8 (ab4055, Abcam, 1:500), CD14 (HPA001887, Cell Marque, 1:100), Ki-67 (D2H10, Cell Signalling technologies, 1:1000) or M.tb (Rabbit polyclonal Ab to M.tb, Cat# ab905, Abcam, 1:500). Primary antibodies were detected with OPAL Polymer HRP Mouse + Rabbit (Akoya Biosciences) for 15 minutes and visualised following a 10-minute incubation with the OPAL fluorophores (Akoya) OPAL 520, OPAL 540, OPAL 570, OPAL 620, OPAL 650, and OPAL 690 at a 1:100 dilution in TSA. Tissue was then boiled at 100°C to strip antibody-HRP complexes from tissue. The process was repeated until all markers had been stained. Following the final marker staining, tissue section was incubated with DAPI for 10 minutes and mounted using ProLong Gold (Thermofisher).

2.2 Image capture

All multiplex images used in this thesis were captured using the Vectra 3.0.5 multispectral imaging system (Akoya) using DAPI, FITC, Cy3, Texas Red and Cy5 channels. A series of high-power 20X multispectral images (MSIs) covering the entirety of tissue sections were captured.

2.3 Cell segmentation in HALO

MSIs were unmixed using Inform 2.3 (Akoya) and exported as .tiff images. Tiff images were then imported in HALO3.3.2541.345 (Indica Labs) and digitally stitched to construct single images of entire tissue sections. HALO was trained to segment individual cells based on characteristic nuclear staining in the DAPI channel using the nuclear detection analysis tool.

2.4 Cell phenotyping in HALO

To distinguish cell identities the positivity threshold for each marker was determined using the analysis panel real time tuning feature in HALO. Any regions containing folded tissue or artifacts were excluded using the exclusion annotation tool. Cell segmentation and cell phenotyping were run together using the “Analyze annotation layer” function. Cell Object Data (including identity and location information for each cell) for of each patient sample was exported and further analysis was performed in R (R Core Team, 2021).

2.5 Statistical Analyses

All statistical analyses were performed in Prism 9.1.2 (GraphPad). Single comparisons were performed using a Mann-Witney test. Multiple comparisons were performed using a Kruskal Wallis test with Dunns multiple comparisons test. Single comparisons using paired samples were performed using a Wilcoxon matched pairs signed rank test. $P < 0.05$ was considered statistically significant for all tests. All plots were generated using Prism or R.

Chapter 3

**Transcriptomic profiling identifies
neutrophil-specific upregulation of cystatin
F as a marker of acute inflammation in
humans**

3.1 Introduction

Transcriptomic profiling has been widely used over the past two decades to uncover the expression and function of genes across the genome. Each transcriptomic profiling study generates enormous amounts of data, the bulk of which has been published in online repositories and is available for download free of charge. These include datasets from across the spectrum of human diseases, tissue types and cell types. Transcriptomic data generated from cells following stimulation or from gene-targeted cells provide additional insights into gene regulation. While transcriptomic datasets have typically been interrogated by the original researchers to identify pathways/genes of interest, most of the data remains incompletely explored and utilised. This wealth of data remains available online and represents a valuable resource for life scientists exploring a gene or pathway of interest. For example, investigations can be used to generate informed hypotheses that guide future studies.

Of particular interest are genes that are found to be differentially expressed despite no prior association with the conditions under investigation. In this case, public transcriptomic datasets can be used to validate findings, facilitate further clarification of the specificity of gene expression across cell types and diseases, and even provide insights into gene regulation. Some studies have already successfully identified a gene of interest across a small cohort of GEO datasets and have been used to formulate a hypothesis for gene function (Rinchai et al., 2015, Toufiq et al., 2020, Huang et al., 2021).

Here, I demonstrate an analytic pipeline that allows identification of differentially expressed genes that are yet to be characterised, then defines mechanisms regulating their expression across a cohort of 43 publicly available transcriptomic datasets. I discovered that the expression of *CST7* in whole blood was markedly upregulated in sepsis as well as across a

range of bacterial, viral, and sterile inflammatory conditions. Cystatin F, encoded by *CST7* is a cathepsin C-directed cysteine protease known to regulate cytotoxicity in natural killer (NK) cells (Nanut et al., 2017, Kos et al., 2018). While cathepsin C is involved in serine peptidase activation in NK cells and neutrophils, the significance and regulation of cystatin F expression in neutrophils remains unexplored. My analysis revealed that this *CST7* upregulation is neutrophil-specific and regulated by host factors. In this study, I demonstrate that transcriptomic profiling using publicly available datasets is a robust approach to identify associations of unknown genes with human diseases.

3.2 Methods

3.2.1 Literature search

PubMed was searched to retrieve the literature associated with each gene using a query that consisted of a concatenation of its official name, symbol and aliases retrieved from NCBI gene. Using *CST7* as an example, the following query was applied: *CST7* [tw] OR 'cystatin F' [tw] OR Leukocystatin [tw]. The text word [tw] function was used to restrict the search to the title and abstract.

3.2.2 Single-cohort transcriptomic analysis

All bulk transcriptomic datasets used in this study were collected from the NCBI GEO as gene series expression (GSE) files (Tab. 3.1). All single-cohort analysis was performed in R and datasets were downloaded using the GEOquery (Davis and Meltzer, 2007) package. Normalised expression counts of genes of interest were exported for plotting. Where necessary, datasets were log-transformed and DEGs were identified using the Limma package (Ritchie et al., 2015) and P values were adjusted with a Benjamini Hochberg false discovery rate of 0.01.

3.2.3 Multicohort transcriptomic analysis

All Multicohort analyses were performed in R using the Metaintegrator package (Haynes et al., 2017). Datasets were accessed from GEO using the `getGEOData` function and samples were separated into control and non-control groups using the `classFunction` function. Meta-analysis was performed using the `RunMetaAnalysis` function and forest plots were

previewed with the `forestPlot` function. The output of the `RunMetaAnalysis` function included summary effect sizes for each genes. These were determined based on the standardised mean difference in each cohort. The summary effect size served as a quantitative measure of the magnitude of difference between the control and non-control groups across cohorts. Effect sizes DEGs between septic patients and controls across the nine whole blood sepsis cohorts were identified using the `filterGenes` function with an effect size threshold of 1 and false discovery rate threshold of 0.05. Heatmaps of sepsis DEG expression across other disease states were generated using the `heatmapPlot` function with “`diplayPooled = FALSE`” to maintain the original gene order.

3.2.4 scRNAseq analysis

The single-cell transcriptomic dataset (EGAS00001004571) used in this study was downloaded from FASTGenomics (Scholz et al., 2018) in a processed form with cluster annotation already applied. All single-cell analysis was performed in R using the Seurat package (Stuart et al., 2019). The ‘DimPlot’ function was used to visualise cell clusters in the Uniform Manifold Approximation and Projection (UMAP) applied dataset. The `FeaturePlot` function was used to visualise gene expression of individual cells within the UMAP space and coloured using the `viridis` package (Garnier. et al., 2021). The `VlnPlot` function was used to visualise gene expression between cell clusters.

Neutrophils that express *CST7* were separated from other neutrophils using the `WhichCells` function and DEG analysis was performed using the `FindMarkers` function using a Wilcoxon Rank Sum Test restricted to genes expressed in at least 25% of cells in each group. Genes with an adjusted p value < 0.05 were considered DEGs. The top 50 DEGs were selected, and Gene Ontology (GO) was performed using the online tool Enrichr

(Kuleshov et al., 2016) (<https://maayanlab.cloud/Enrichr/>) with the “GO Biological Processes 2021” ontology.

3.2.5 Software

Transcriptomic analysis was performed in R_{x64} 4.0.2 (R Core Team, 2021).

Table 3.1. Transcriptomic datasets used in this study.

GSE #	Reference	Age*	Sample type	Conditions compared
GSE2322	Coldren et al., 2006	Adult	Airspace neutrophils Circulating neutrophils	Airspace neutrophils <i>In vitro</i> control <i>In vitro</i> endotoxin <i>In vivo</i> control <i>In vivo</i> endotoxin
GSE8668	Radom-Aizik et al., 2008	12–29	Purified neutrophils	Pre-exercise Post-exercise
GSE13015	Pankla et al., 2009	37–74	Whole blood	Healthy control Sepsis
GSE17755	Lee et al., 2011	2–19	Whole blood	Healthy child SOJIA
GSE19491	Berry et al., 2010	21–72	Whole blood	Healthy control TB
GSE25504	Dickinson et al., 2015	Neonate	Whole blood	Healthy child Sepsis
GSE27131	Berdal et al., 2011	28–59	Whole blood	Healthy control H1N1
GSE28623	Elliott et al., 2015	Adult	Whole blood	Healthy control Active TB
GSE28750	Sutherland et al., 2011	Adult	Whole blood	Healthy control Sepsis
GSE29366	-	0–3	Whole blood	Healthy child Influenza
GSE29536	Banchereau et al., 2012	Adult & children	Whole blood	TB control, TB IAV control, IAV SOJIA control, SOJIA Sepsis control, Sepsis
GSE30101	Obermoser et al., 2013	22–41	Whole blood	Unstimulated blood Stimulated blood
GSE30119	Banchereau et al., 2012	0–18	Whole blood	Healthy child <i>S. Aureus</i> infection
GSE34608	Maertzdorf et al., 2012	Adult	Whole blood	Healthy control TB
GSE38900	Mejias et al., 2013	0–2	Whole blood	Healthy Control RSV, IAV
GSE40012	Parnell et al., 2012	Adult	Whole blood	Healthy Control IAV pneumonia
GSE41055	Verhagen et al., 2013	1–14	Whole blood	Healthy control TB
GSE42026	Herberg et al., 2013	Adult	Whole blood	Healthy control H7N9, RSV
GSE42834	Bloom et al., 2013	Adult	Whole blood	Healthy control TB
GSE49757	Khaenam et al., 2014	Adult	Purified neutrophils	Stimulation with healthy plasma Stimulation with septic plasma
GSE54514	Parnell et al., 2013	24–86	Whole blood	Healthy Sepsis survivor Sepsis non-survivor

GSE57183	Shenoi et al., 2015	1–11	Whole blood	Healthy control SOJIA Kawasaki disease
GSE60424	Linsley et al., 2014	Adult	Purified neutrophils, Whole blood	Healthy control Sepsis
GSE64457	Demaret et al., 2015	Adult	Purified Neutrophils	Healthy control Septic shock patients
GSE66099	Alder et al., 2017	Adult	Whole blood	Healthy control Sepsis
GSE69528	Pankla et al., 2009	Adult	Whole blood	Healthy control Sepsis
GSE71010	Hu et al., 2016	Children	Purified neutrophils	Healthy Cystic fibrosis, SOJIA
GSE76293	Juss et al., 2016	Adult	Purified neutrophils	Healthy control ARDS
GSE80060	Brachat et al., 2017	Children	Whole blood	Healthy control SOJIA prior to treatment
GSE81746	- **	22–65	Whole blood	Healthy control TB
GSE83456	Blankley et al., 2016b	19–82	Whole blood	Healthy control Pulmonary TB
GSE100151	-	Adult	Whole blood	Healthy control HIV
GSE100154	-	Children	Whole blood	Healthy control Kawasaki disease
GSE100159	-	Adult	Whole blood	Healthy control Sepsis
GSE100160	-	Adult	Whole blood	Healthy control IAV
GSE100161	-	Adult	Whole blood	Healthy control RSV
GSE100164	-	Children	Whole blood	Healthy Control SOJIA
GSE100166	-	Adult	Whole blood	Healthy control TB
GSE101584	Shaul et al., 2016	-	Purified neutrophils	Healthy neutrophils Tumour associated neutrophils TAN + SM16
GSE103170	Ter Haar et al., 2018	Children	Purified neutrophils	Healthy control SOJIA
GSE107731	-	Adult	Whole blood	Healthy control TB
GSE111368	Dunning et al., 2018	18–71	Whole blood	Healthy control H1N1

* Numbers indicate age range in years.

** Hyphen indicates no publicised study associated with the dataset.

3.3 Results

3.3.1 Cystatin F is upregulated in neutrophils following stimulation with septic plasma and in circulating leukocytes of sepsis patients

It has been shown previously that transcriptomic changes in public datasets can be explored for biomarker discovery by identifying previously uncharacterised genes (Rinchai et al., 2015, Toufiq et al., 2020, Huang et al., 2021). Here I demonstrate a two-staged approach of literature-searches and public-data mining, which enables not only the identification of a novel gene of interest but also to define regulatory mechanisms controlling its expression (Fig. 3.1). To identify an initial cohort of genes upregulated during inflammation, a dataset deposited in the NCBI GEO database (GSE49757) was chosen. In this dataset, neutrophils from healthy donors had been stimulated with plasma from either healthy controls or bacterial culture-confirmed sepsis patients and mRNA expression was measured using microarray (Fig. 3.2A). DEG analysis was performed, significant genes were sorted according to the fold change in expression and the top 250 genes were selected for further analysis (Fig. 3.2B).

To extend the in vitro observations and identify potential candidate genes for further study, validation was performed in publicly available datasets generated using whole blood samples from healthy controls and sepsis patients. Nine datasets were selected from the NCBI GEO and multicohort analysis was performed to generate a mean fold change in expression across all cohorts. The top 250 significant genes were ranked according to fold change in expression in this cohort (Fig. 3.2C). The highest ranked genes in terms of the effect size were considered as the strongest candidates for further investigation (Fig. 3.2D).

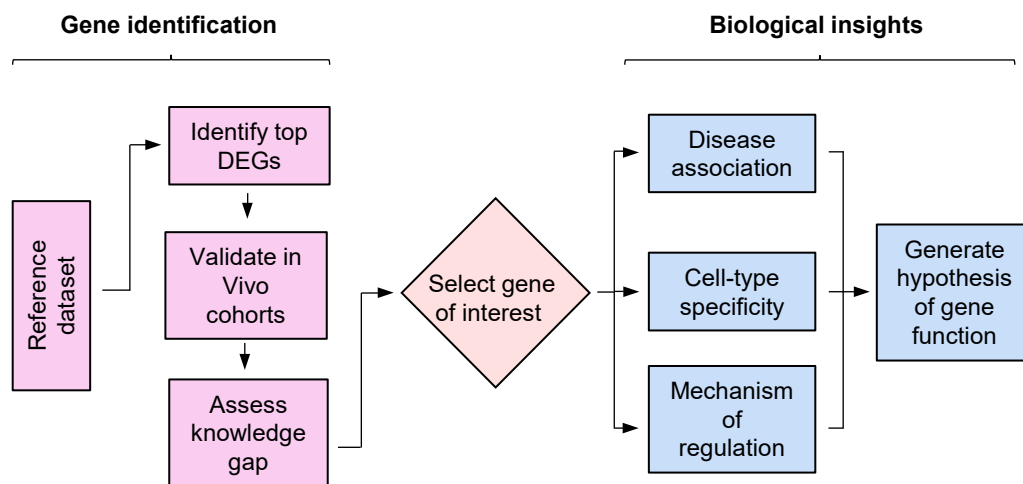


Figure 3.1. Flow chart of the process employed in this study to identify a gene of interest and assess its regulation.

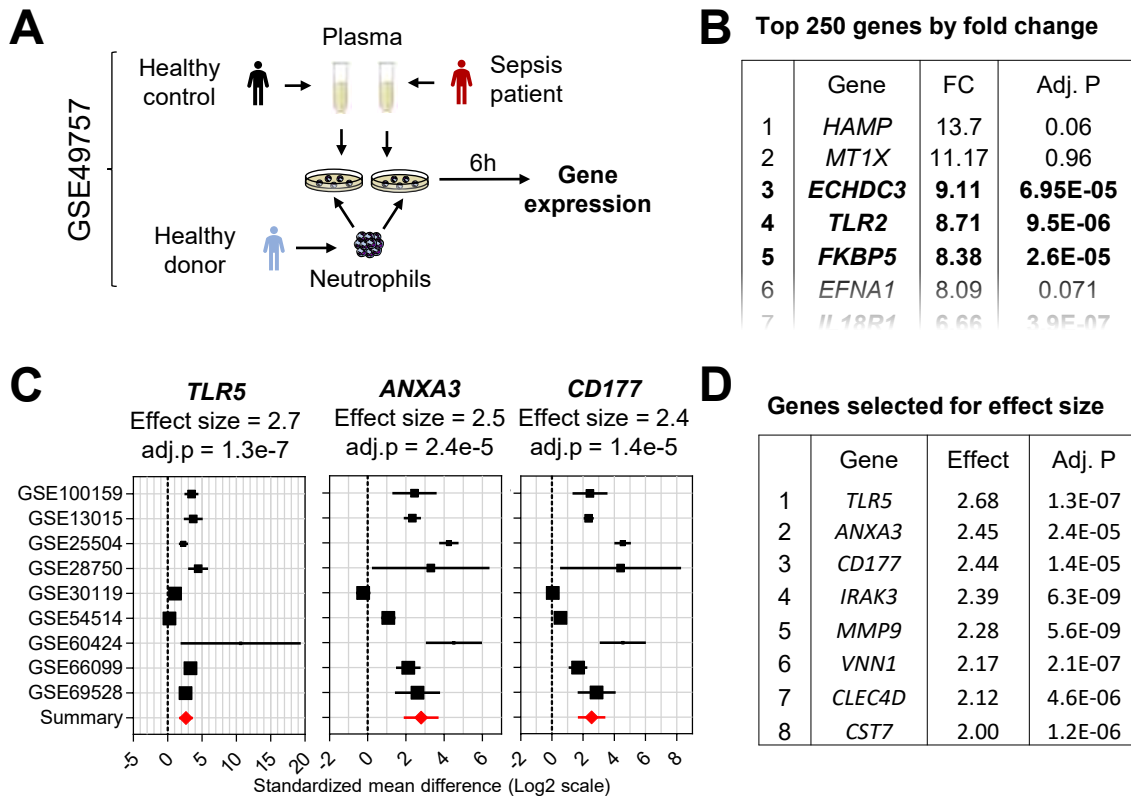


Figure 3.2. Meta analysis of sepsis associated neutrophil genes. (A) Schematic overview of the experimental design for the reference dataset (GSE49757) used in this study. Purified neutrophils from healthy adult donors were stimulated with plasma from either septic or healthy individuals for 6 h and gene expression was measured by microarray. (B) Genes were sorted according to fold change in expression using Limma with Benjamini-Hochberg false discovery rate applied and the top 250 genes with a significant change in expression were selected for further assessment. An adj. P < 0.05 was considered to indicate statistical significance. (C) Gene expression profiles across nine transcriptomic datasets comparing whole blood samples from patients with sepsis to healthy controls are represented in forest plots. Cohort identifiers are shown on the Y-axis and closed black squares represent the standardized mean difference. Square size is proportional to the sample size and whiskers represent a 95% confidence interval. The summary (closed red diamond) represents the overall change in gene expression across cohorts and is also shown as the effect size above each plot along with its corresponding adj. P-value. (D) The top 250 genes identified in panel B were sorted according to overall effect size across sepsis cohorts.

The final step in the process of selecting a gene of interest was to assess the current knowledge on each identified gene. PubMed was searched to identify the total number of articles relating to each gene and alternate names for each gene were identified through NCBI Gene and added to the search. Genes for which a very small number of articles were identified (<10) were excluded as it was considered unlikely that sufficient information about these gene would be available to inform future studies. Additional searches were carried out for each gene paired with the search terms “sepsis”, “neutrophil” and “inflammation” to identify previous gene associations. The number of articles was counted, and a score applied to each gene based on the number of related articles. Genes were examined in order starting from the gene with the highest fold change in expression and those with known associations to sepsis or neutrophils were excluded (Fig. 3.3A). *CST7* (Cystatin F) was ranked 8th in terms of fold change in expression and 72 articles were identified that referred to the gene. Among these, six publications were related to inflammation; however, none of these articles contained the term “sepsis” or linked cystatin F directly to the term neutrophil. *CST7* was therefore deemed the most eligible for further investigation. I further showed that *CST7* expression was significantly increased in neutrophils following stimulation of plasma from patients with severe sepsis compared with those with non-severe sepsis (Fig. 3.3B).

A**Gene selected based on novelty of finding**

Gene	Studies published	"gene" AND sepsis	"gene" AND neutrophil	"gene" AND inflammation
<i>TLR5</i>	2076	65	124	535
<i>ANXA3</i>	454	2	16	17
<i>CD177</i>	258	12	155	36
<i>IRAK3</i>	239	25	16	96
<i>MMP9</i>	28343	162	1503	4904
<i>VNN1</i>	146	1	19	30
<i>CLEC4D</i>	201	10	22	53
<i>CST7</i>	72	0	4	6

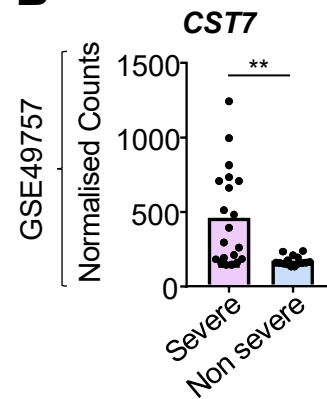
B

Figure 3.3. Identification of *CST7* as a previously unknown neutrophil-specific, sepsis-associated gene. (A) For each gene identified in Fig. 3.2D, PubMed was searched to identify both the total number of articles and the number of articles linking the gene with each of the terms “sepsis”, “neutrophil” and “inflammation” in the title and abstract. Genes were searched in order from the highest effect size until a gene with adequate novelty (determined by the number of publications) was identified. (B) *CST7* expression was compared between patients with severe sepsis vs. non-severe sepsis in the reference dataset (GSE49757). Each symbol represents an individual donor, and the height of each column corresponds to the group mean. The level of statistical significance was determined using a Mann-Whitney test and $P < 0.05$ was considered to indicate statistical significance. ** $P < 0.01$.

3.3.2 The significance and regulation of cystatin F induction are poorly defined

To inform further exploration of public data, a profile of the current knowledge of *CST7* was constructed. The 72 articles that referred to *CST7* were examined for major research themes and tallied (Fig. 3.4A). The cystatins consist of three families of proteins that bind reversibly with high-affinity to inhibit multiple cathepsin family members (Magister and Kos, 2013).

Cystatin F is unique among the cystatins in that its expression is limited to leukocytes (Halfon et al., 1998) and because it is secreted in an inactive form that can be internalised and subsequently activated in target cells (Colbert et al., 2009, Colbert et al., 2011). In the context of malignancy, increased levels of *CST7* have been correlated with a poorer prognosis in liver metastasis following colorectal cancer (Utsunomiya et al., 2002) and with improved survival in the context of pancreatic ductal adenocarcinoma (Yang et al., 2019) and hepatocellular carcinoma (Zhou et al., 2019). Additionally, in the context of neuroinflammation, *CST7* is upregulated during CNS demyelination and is associated with areas of remyelination (Ma et al., 2011, Liang et al., 2016). While *CST7* has not been studied in the context of other inflammatory conditions, it has been identified as playing a role in granulocyte inflammation, specifically as an inhibitor of cathepsin C (Hamilton et al., 2008).

Cathepsins consist of a large family of cysteine proteases and play a diverse set of roles in both physiological and pathophysiological conditions. In the immune system, they are involved in antigen processing and presentation, activation of pattern recognition receptors and as cytotoxic devices utilised by both the host and pathogens (Bird et al., 2009). Cathepsin C is a component of the granules in cytotoxic leukocytes and functions as a key activator of other inflammatory granules in these cells (Hamilton et al., 2008). Cystatin F can inhibit cathepsin C in NK cells, thereby inducing an impaired cytotoxic state known as split anergy (Nanut et al., 2017, Magister et al., 2015). High levels of cystatin F in CD8⁺ T cells are also correlated with decreased cytotoxicity (Prunk et al., 2019).

In neutrophils, cathepsin C is required for activation of neutrophil serine proteases including Elastase, neutrophil expressed (*ELANE*), Cathepsin G (*CTSG*) and Proteinase 3 (*PRTN3*) in the inflammatory granules (Hamilton et al., 2008, McGuire et al., 1993, Adkison et al., 2002). However, whether cystatin F inhibits cathepsin C in neutrophils or exerts anti-inflammatory effects in these cells remains unknown (Fig. 3.4B). Due to their potential connection with *CST7*, I also examined *CTSC*, *ELANE*, *CTSG* and *PRTN3* in the public data cohorts in the remainder of this study. Based on the previous cohort of sepsis datasets, *CST7* was significantly upregulated (effect size = 1.99) in patients with sepsis compared with healthy controls, while there was no significant difference in *CTSC* expression between the two groups. *ELANE*, *CTSG* and *PRTN3* were significantly upregulated (effect size = 1.28, 0.74 and 1.08, respectively) (Fig. 3.4C).

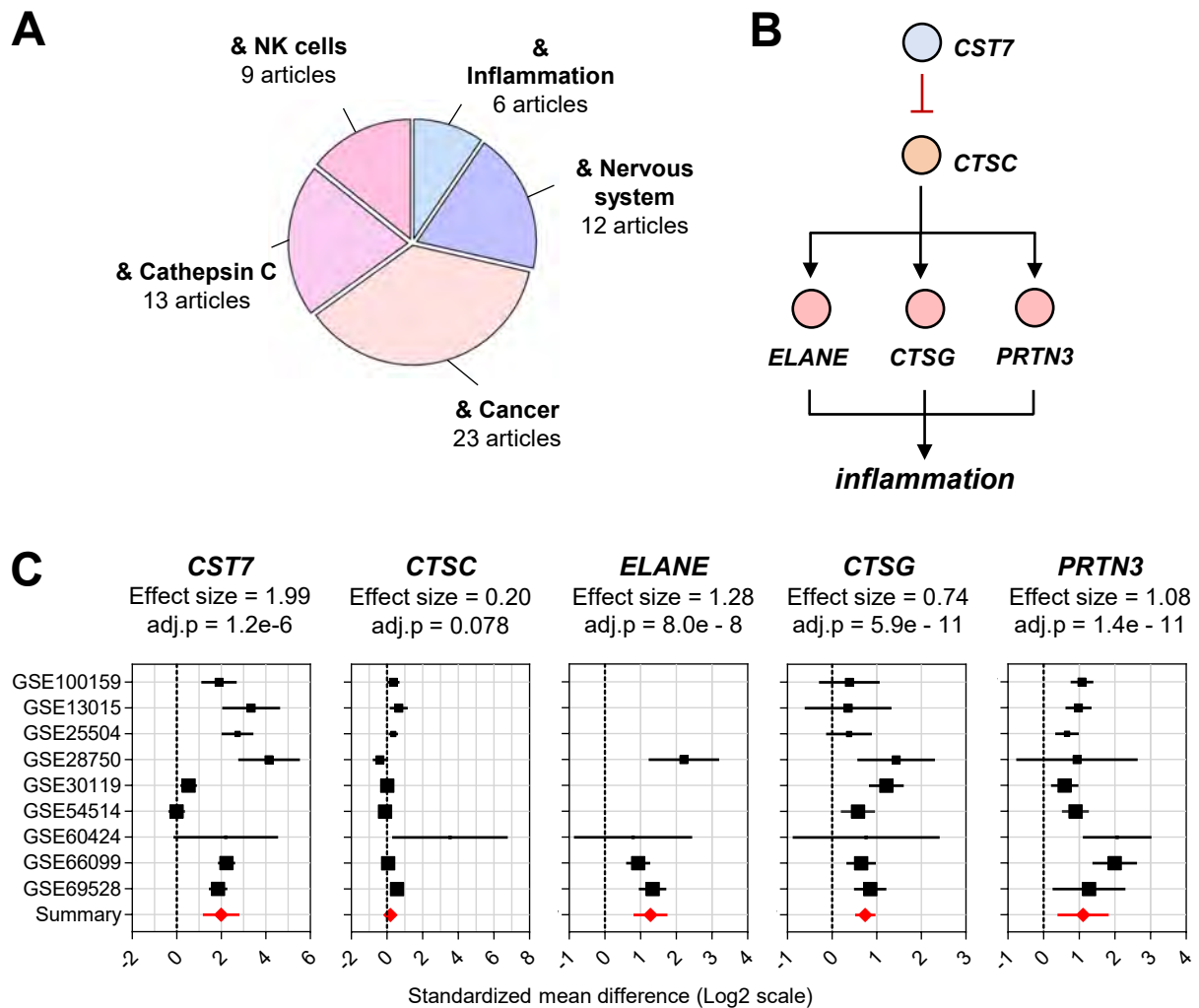


Figure 3.4. Cystatin F is identified as a cathepsin C-directed serine protease inhibitor.

(A) The total number of articles on *CST7* (72 articles) was curated and major reported biological functions were summarized. The size of each segment in the pie chart corresponds to the number of articles matching *CST7* in relation to the indicated area. (B) The proposed function of the *CST7* protein indicating its ability to inhibit *CTSC* protein. *CTSC*, which is known to activate the neutrophil serine proteases *ELANE*, *PRTN3* and *CTSG*, can increase inflammation. (C) Forest plots of *CST7*, *CTSC*, *ELANE*, *PRTN3* and *CTSG* gene expression across sepsis cohorts. The plots were constructed as described in Figure 3.2C.

3.3.3 Cystatin F is consistently upregulated across bacterial, viral, and sterile systemic inflammation

Having identified *CST7* as an upregulated gene in whole blood during sepsis, I sought to determine whether the change was restricted to sepsis or common across other inflammatory diseases. I selected six inflammatory diseases for further examination, each of which was represented by at least two cohorts on the GEO. The diseases were categorised as bacterial, viral, and sterile inflammation. All the selected cohorts compared whole blood samples from patients and healthy controls. The 2,140 genes that were found to be differentially expressed in the sepsis cohort were measured across each of the other disease cohorts.

CST7 was significantly upregulated during TB (effect size = 0.84) (Fig. 3.5A). In contrast, there were no changes in *CTSC* or *CTSG* expression, although expression of *PRTN3* and *ELANE* was increased. *CST7* was also significantly upregulated and with the highest effect sizes during viral-induced inflammation (Fig. 3.5B). *CST7* expression was increased during infection with IAV, respiratory syncytial virus (RSV) and HIV infection (effect size = 1.47, 1.18, 1.54 respectively). Interestingly, *CTSC* expression did not change during IAV infection, but was significantly increased during RSV and HIV infection. *PRTN3*, *ELANE* and *CSTG* were significantly increased in all three viral infection cohorts. Finally, in the case of sterile inflammation, *CST7* expression was increased in both SOJIA (Fig. 3.6A) and Kawasaki disease (Fig. 3.6B) (effect size = 0.6 and 1.02 respectively). *CTSC* was not changed in either cohort, while *CTSG* and *PRTN3* were significantly upregulated in both cohorts and *ELANE* was significantly upregulated during SOJIA. Heat maps of the similarity between the gene expression profile of each disease state and sepsis are presented in Figure 3.5 and Figure 3.6.

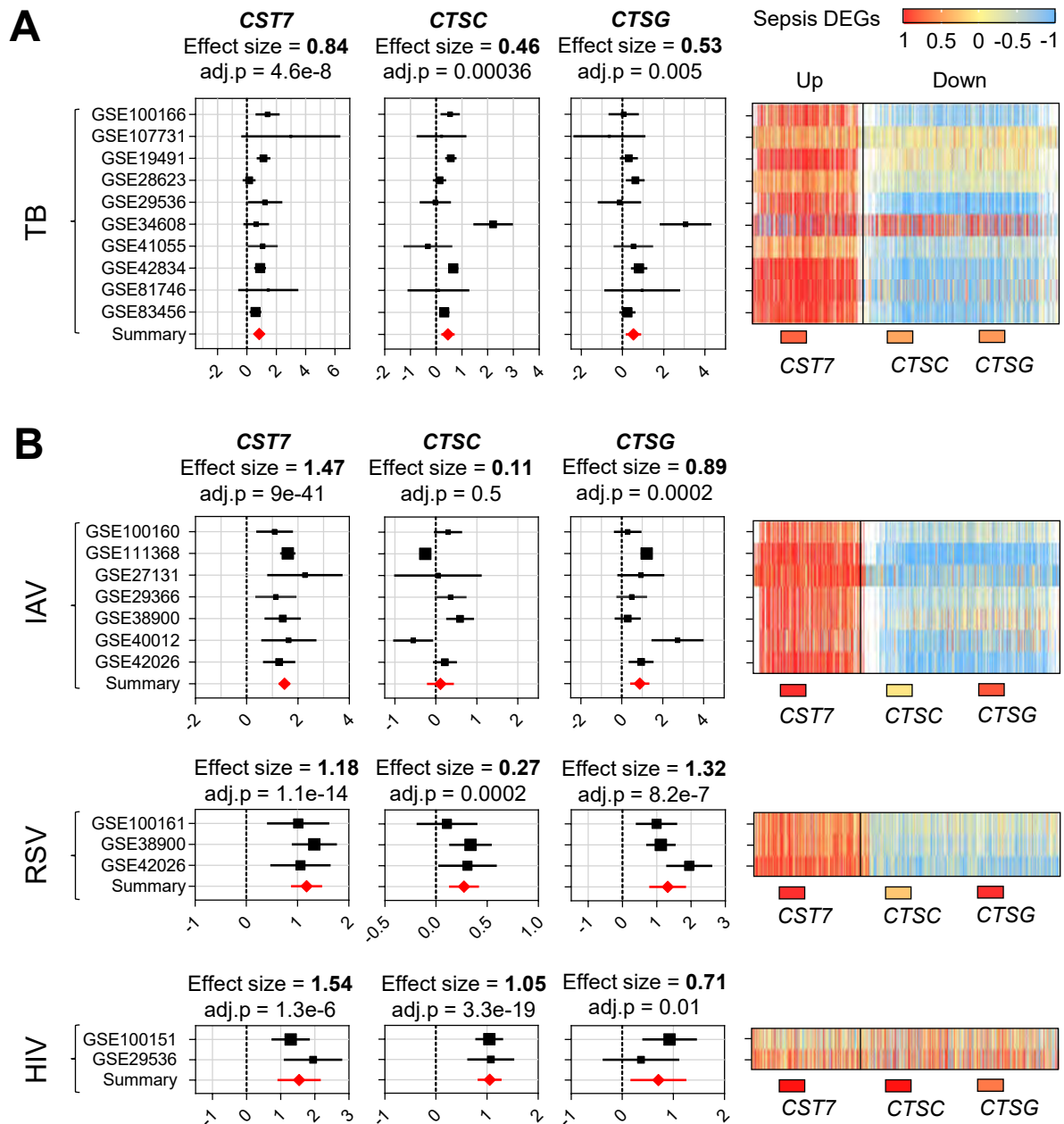


Figure 3.5. *CST7* is upregulated across a spectrum of infectious diseases. (A) Gene expression profiles for *CST7*, *CTSC*, and *CTSG* across 10 transcriptomic datasets comparing whole blood samples from patients with TB to healthy controls are represented in forest plots constructed as described in Figure 3.2C. An accompanying heatmap represents the expression of the 2,140 DEGs across the sepsis cohorts (X-axis, separated into up/downregulated expression during sepsis) in each TB cohort (Y-axis). This was repeated in (B) with transcriptomic datasets comparing whole blood samples from IAV patients to healthy controls, RSV patients to healthy controls and HIV patients to healthy controls.

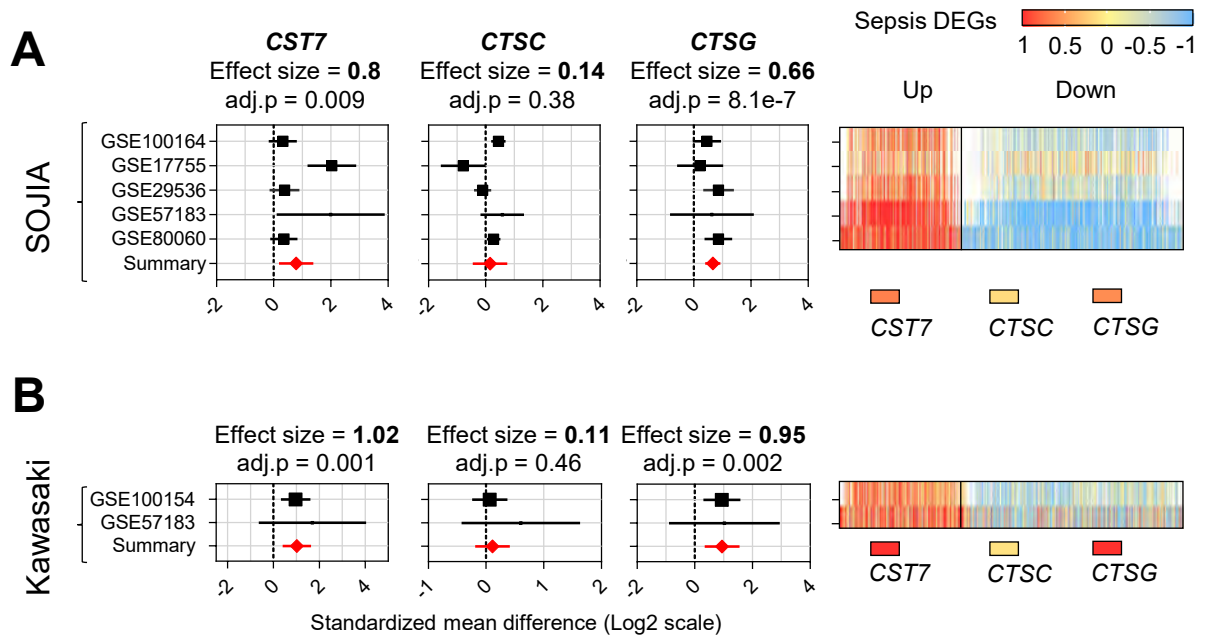


Figure 3.6. *CST7* is upregulated in non-infectious inflammatory diseases. (A) Gene expression profiles for *CST7*, *CTSC*, and *CTSG* comparing whole blood of patients with SOJIA to healthy controls across five transcriptomic datasets represented with forest plots and heatmap, constructed as described in Figure 3.5A. This was repeated in **(B)** with transcriptomic datasets comparing whole blood from patients with Kawasaki disease to healthy controls.

3.3.4 Cystatin F upregulation is neutrophil-specific

Having established that *CST7* is upregulated in bulk leukocyte preparations during inflammation, I next determined whether this change is restricted to certain leukocyte populations in the blood. Given the initial finding that septic plasma induced upregulation of *CST7* in neutrophils, I identified datasets comparing purified neutrophils from healthy controls and sepsis patients. Two datasets were identified and *CST7* was found to be significantly upregulated in neutrophils from sepsis patients compared to those from healthy controls (effect size = 2.3). However, there were no changes in the expression levels of *CTSC*, *CTSG*, *PRTN3* and *ELANE* in neutrophils during sepsis compared to those from healthy controls (Fig. 3.7A).

I next examined whether expression of *CST7* is modulated in other leukocyte populations during sepsis using one cohort from the previous analysis that also included purified CD8⁺ T cells NK cells, CD4⁺ T cell and monocytes as well as neutrophils. While NK cells expressed the highest *CST7* levels under healthy conditions, there was a significant increase in *CST7* expression in neutrophils during sepsis and these cells became the predominant contributor of *CST7* expression (Fig. 3.7B). There was no change in *CST7* expression in NK cells or CD8⁺ T cells during sepsis. *CST7* was expressed lowly in CD4⁺ T cells and monocytes and its expression in these cells did not change during sepsis (Data not shown).

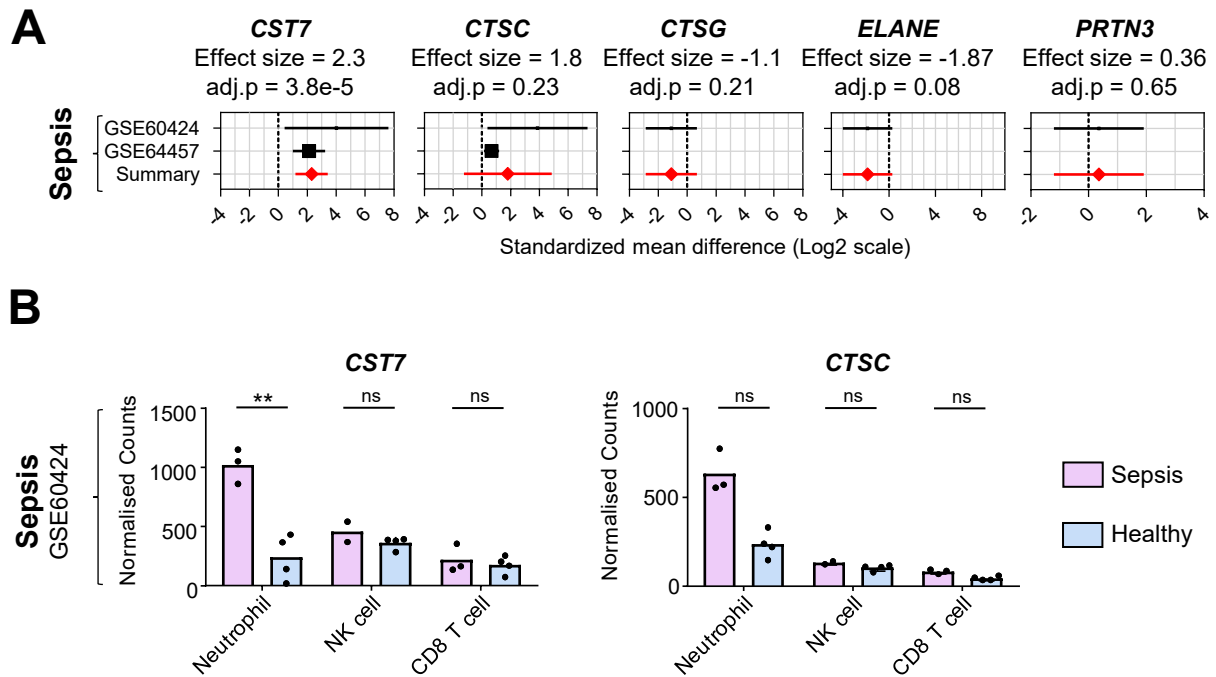


Figure 3.7. *CST7* upregulation during inflammation is neutrophil-specific.

(A) Gene expression profiles for *CST7*, *CTSC*, *CTSG*, *ELANE*, and *PRTN3* across two datasets comparing purified neutrophil populations from patients with sepsis and healthy controls are represented in forest plots constructed as described in Figure 3.2C. **(B)** Normalized expression counts of *CST7* and *CTSC* in purified neutrophils, NK cells and CD8⁺ T from healthy controls compared to those from patients with sepsis. Each symbol represents an individual donor, and the height of each column corresponds to the group mean. The level of statistical significance in each comparison was determined using a Mann-Whitney test and $P < 0.05$ was considered to indicate statistical significance. ** $P < 0.01$. ns, Not Significant.

Given this neutrophil-specific increase during sepsis, I next sought to determine whether *CST7* levels were increased in neutrophils under other inflammatory conditions. I examined four cohorts comparing gene expression from purified neutrophils during acute respiratory distress syndrome (ARDS), adult-onset Still's disease (AOSD), cystic fibrosis (CF), in each case compared to purified neutrophils from healthy controls. An additional cohort compared purified neutrophils from individuals both pre- and post-exercise. In all these conditions, *CST7* expression was significantly increased, suggesting neutrophil-specific upregulation of *CST7* is positively associated with inflammatory conditions (Fig. 3.8A-D).

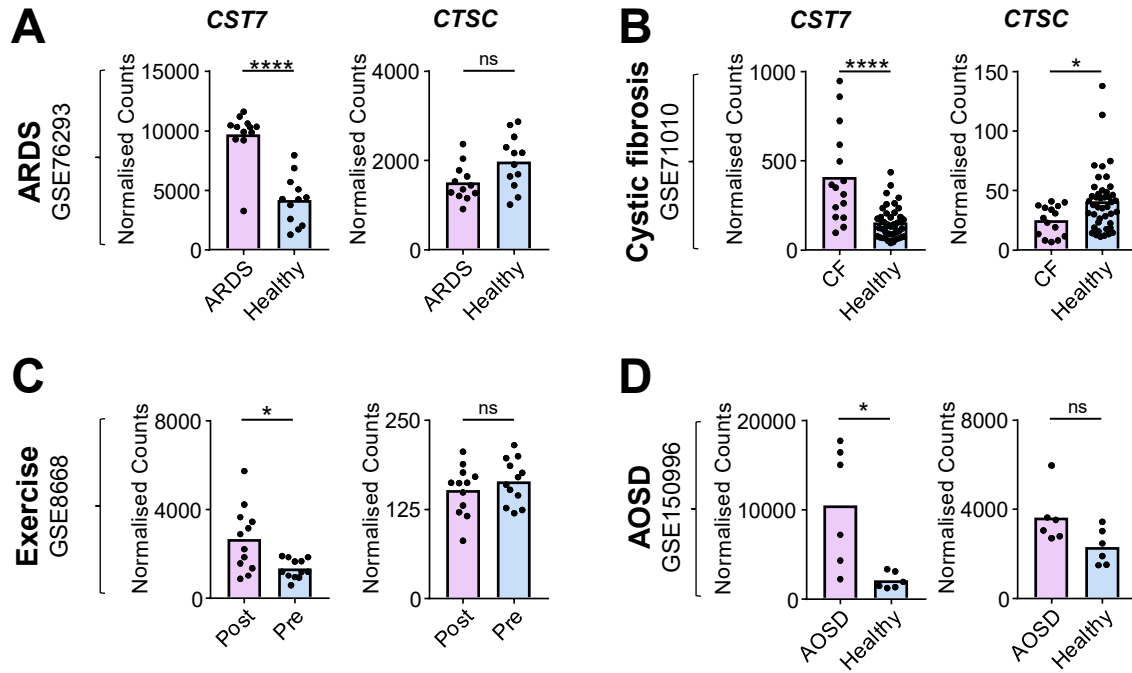


Figure 3.8. *CST7* is upregulated in neutrophils under multiple inflammatory conditions. Expression of *CST7* and *CTSC* in purified neutrophils comparing patients with (A) ARDS (GSE76293), (B) CF (GSE71010) and (C) SOJIA (GSE103170) to healthy controls. (D) Comparison of *CST7* expression in purified neutrophils from healthy individuals pre- and post-exercise (GSE8668). In (A) – (D) each symbol represents an individual donor, and the height of each column corresponds to the group mean. The level of statistical significance in each comparison was determined using a Mann-Whitney test and $P < 0.05$ was considered to indicate statistical significance. * $P < 0.05$. **** $P < 0.0001$. ns, Not Significant.

3.3.5 Cystatin F expression during inflammation is regulated by host factor(s)

Having identified upregulation of *CST7* during inflammation, I subsequently examined whether public transcriptomic data can be used to shed light on the regulatory mechanisms underpinning this increase in expression. Given that *CST7* is increased under a diverse range of inflammatory conditions, I hypothesised that a broadly acting inflammatory mediator may regulate *CST7* expression changes. However, when I analysed another dataset in which whole blood cells were stimulated with a range of TLR ligands, heat killed pathogens and major pro-inflammatory cytokines, none of these factors stimulated any significant change in *CST7* expression (Fig. 3.9).

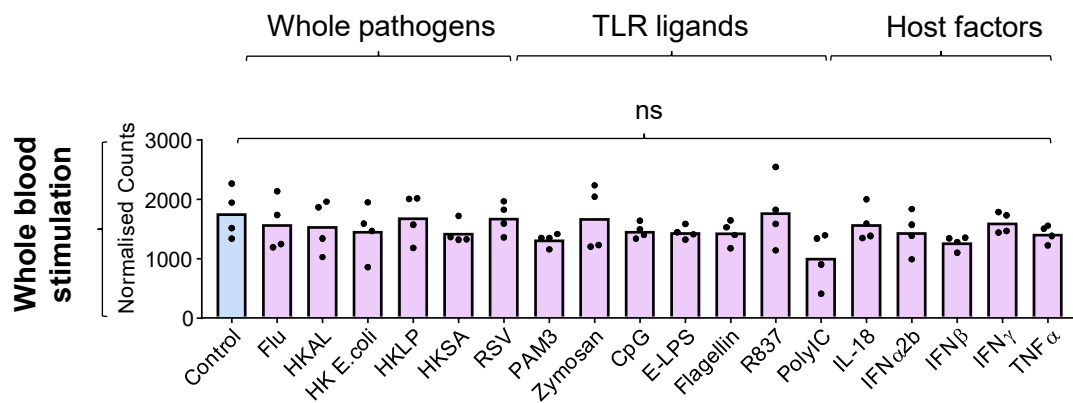


Figure 3.9. *CST7* upregulation could not be induced by pro-inflammatory molecules in neutrophils ex vivo. *CST7* expression in whole blood cells following stimulation in vitro with heat-killed (HK) whole pathogens, purified TLR ligands and host cytokines (GSE30101). Whole blood from four healthy donors was incubated for 6 h with one of the listed stimuli before gene expression was measured by microarray. Each symbol represents an individual donor, and the height of each column corresponds to the group mean. The level of statistical significance is determined using a Kruskal-Wallis Test with Dunns Multiple Comparisons and adj. $P < 0.05$ was considered to indicate statistical significance. ns, Not Significant.

Finally, I analysed a dataset in which neutrophil gene expression was compared following endotoxin exposure in vivo and in vitro. I found that when neutrophils were purified from the blood of healthy donors and exposed to endotoxin in vitro, no change in *CST7* expression was observed (Fig. 3.10A), confirming that microbial products do not contribute to *CST7* induction directly. However, when the donors were exposed to endotoxin by bronchoscopic instillation in vivo and circulating neutrophils were purified from the blood and examined 16 h later, *CST7* was significantly upregulated (Fig. 3.10B). Together, these findings suggest that an endogenous rather than exogenous factor contributes to the *CST7* upregulation during inflammation.

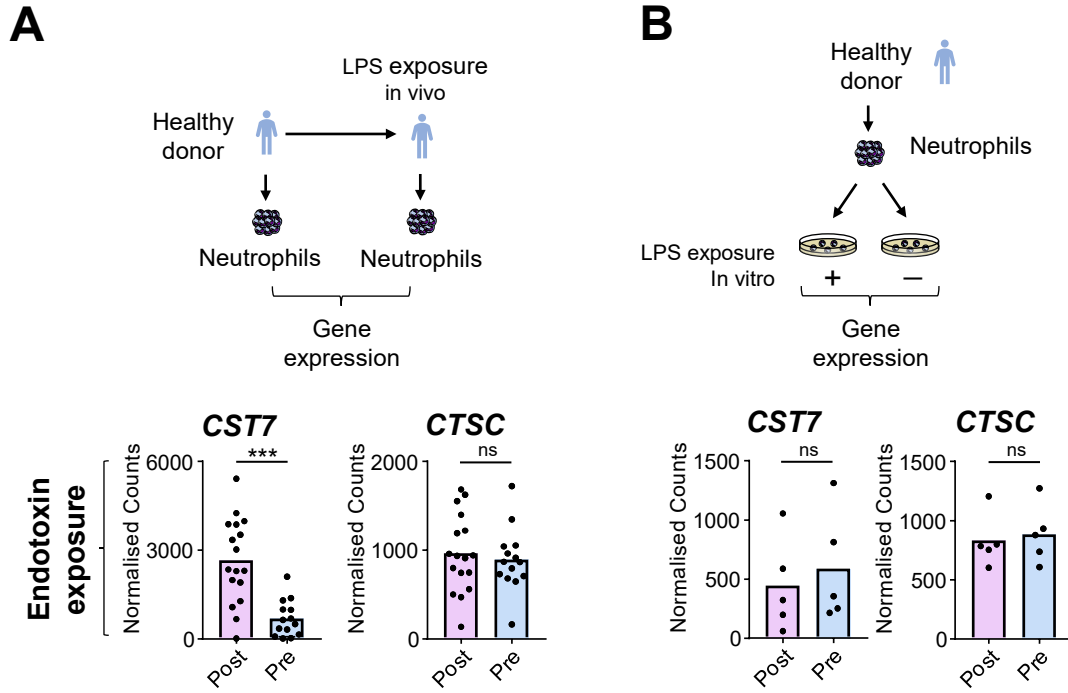


Figure 3.10. Neutrophil-specific CST7 upregulation during inflammation is regulated by host factors. (A) Expression of *CST7* and *CTSC* in circulating neutrophils from healthy individuals taken prior to or 16 h after bronchoscopic instillation of endotoxin (LPS). (B) Expression of *CST7* and *CTSC* in purified neutrophils isolated from healthy donors following stimulation with endotoxin in vitro. In (A) and (B) each symbol represents an individual donor and the height of each column corresponds to the group mean. The level of statistical significance in each comparison was determined using a Mann-Whitney test and $P < 0.05$ was considered to indicate statistical significance. *** $P < 0.001$. ns, Not Significant.

3.3.6 Cystatin F expression in neutrophils is associated with type I IFN signalling

To further explore the regulatory mechanisms underpinning expression of *CST7* in neutrophils, a public scRNAseq dataset was examined (EGAS00001004571) (Schulte-Schrepping et al., 2020). This dataset compared both moderate and severe COVID-19 patients to healthy controls. scRNAseq had been performed on both isolated peripheral blood mononuclear cells (PBMCs) and whole blood, which were later integrated together for analysis to allow for suitable abundance of major immune populations (Fig. 3.11A). The proportion of neutrophils that express *CST7* during severe COVID-19 was double that of the control group, however curiously the proportion of neutrophils that express *CST7* during mild COVID-19 was lower than in the control group. *CST7* was also expressed in CD8⁺ T cells and NK cells under all three conditions (Fig. 3.11B). When *CST7* expression was viewed across all cells in the UMAP space, different populations of neutrophils appeared to express *CST7* under each condition (Fig. 3.11C).

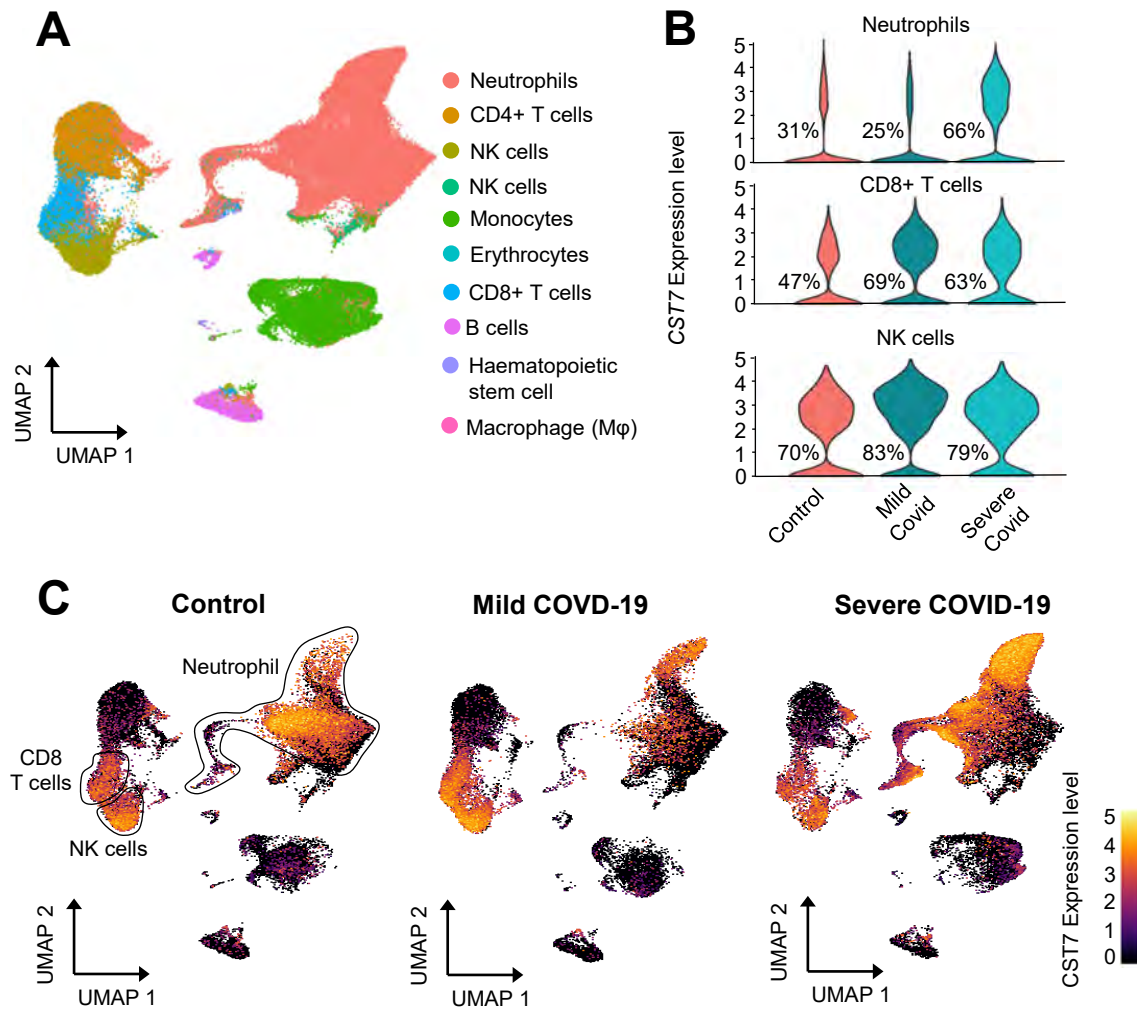


Figure 3.11. *CST7* is upregulated in neutrophils during severe COVID-19. (A) UMAP presentation of major immune populations scRNAseq whole blood and PBMC combined data. **(B)** Expression of *CST7* between three patient conditions in neutrophils, NK cells and CD8⁺ T cells. Percentage accompanying each violin plot indicates the percentage of cells positive for *CST7*. **(C)** UMAP from (A) split into cells from each patient condition with cells coloured according to expression of *CST7*.

The original authors of this dataset identified 9 subsets of neutrophils in the blood, including both mature and immature populations as well as populations specific to control or diseased conditions (Fig. 3.12A). *CST7* was expressed in both mature and immature neutrophil clusters and was most highly expressed in pro-neutrophils as well as a mature cluster specific to severe COVID-19 (Cluster 1, marker genes: PD-L1+, ISG15+, ZC3H12A+, MME-), a COVID-19 specific cluster (Cluster 2, marker genes: CD46+) and a control specific cluster (Cluster 4, marker genes MME+, S100A4+) (Fig. 3.12B and 3.12C).

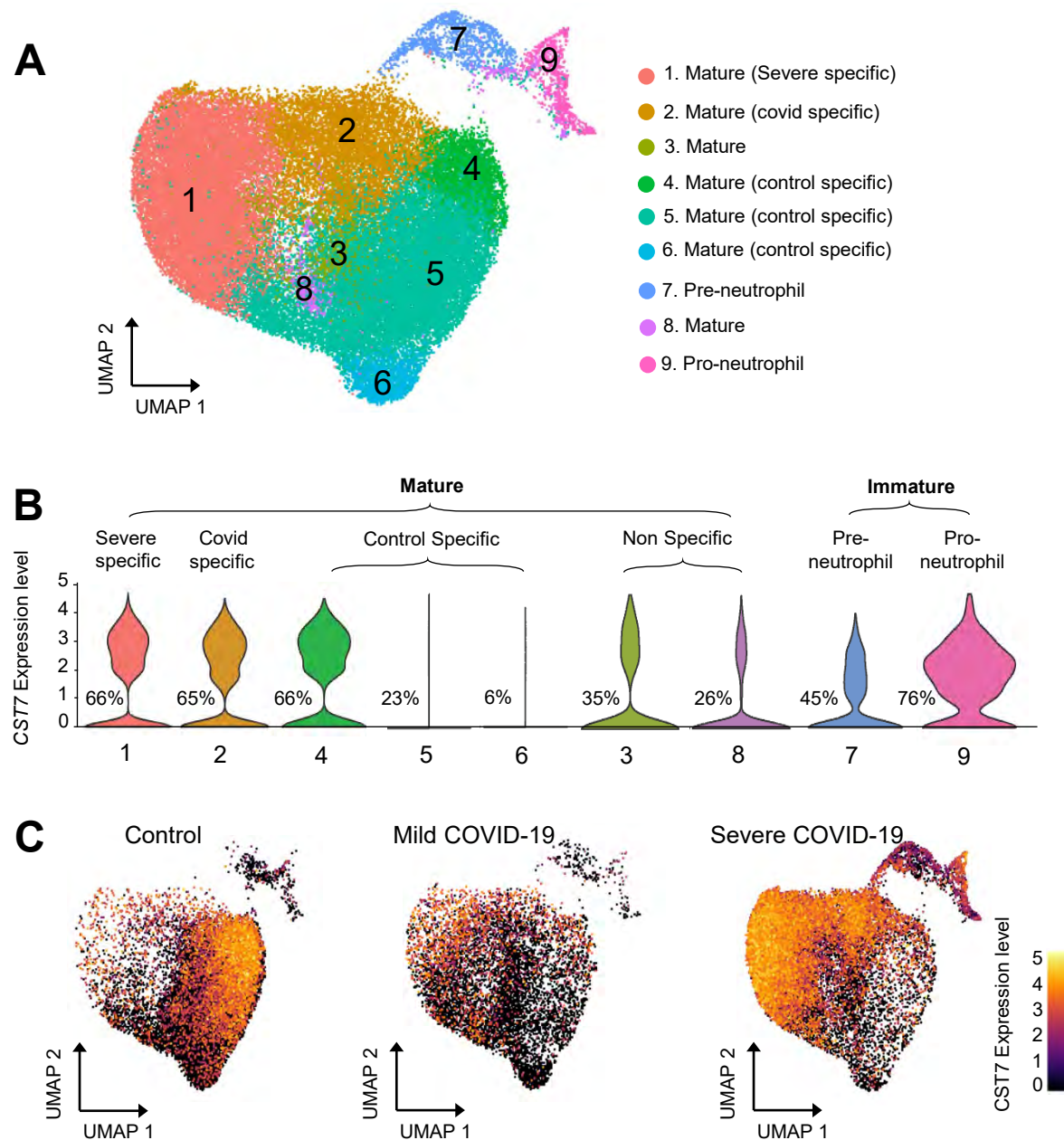


Figure 3.12. *CST7* is expressed in mature and immature neutrophils. (A) UMAP presentation of neutrophils from whole blood scRNAseq coloured according to cluster. **(B)** Violin plots of *CST7* expression across each neutrophil cluster. Percentages indicate the percentage of neutrophils positive for *CST7* in each accompanying cluster. **(C)** UMAP from (A) separated into patient conditions and coloured according to expression of *CST7*.

To explore whether any common regulatory mechanisms could be associated with expression of *CST7* across neutrophil clusters, neutrophils were divided into a *CST7*⁺ population and a *CST7*⁻ population (Fig. 3.13A). DEG analysis was performed and the top fifty significantly upregulated genes, sorted by fold change were used for GO analysis (Fig. 3.13B). Using the Biological Processes 2021 Ontology the most significantly upregulated biological process was cellular response to type I IFN (Fig. 3.13C). The top fifty DEGs used for GO contained several genes known to be induced by type I IFN signalling including *IFITM3*, *IFI6*, *OAS1* and *MX1*. Other top ten most significantly upregulated biological processes included neutrophil activation in immune response, neutrophil degranulation, and neutrophil mediated immunity. Taken together, these findings reveal an association between type 1 IFN signalling and the expression of *CST7* in neutrophils.

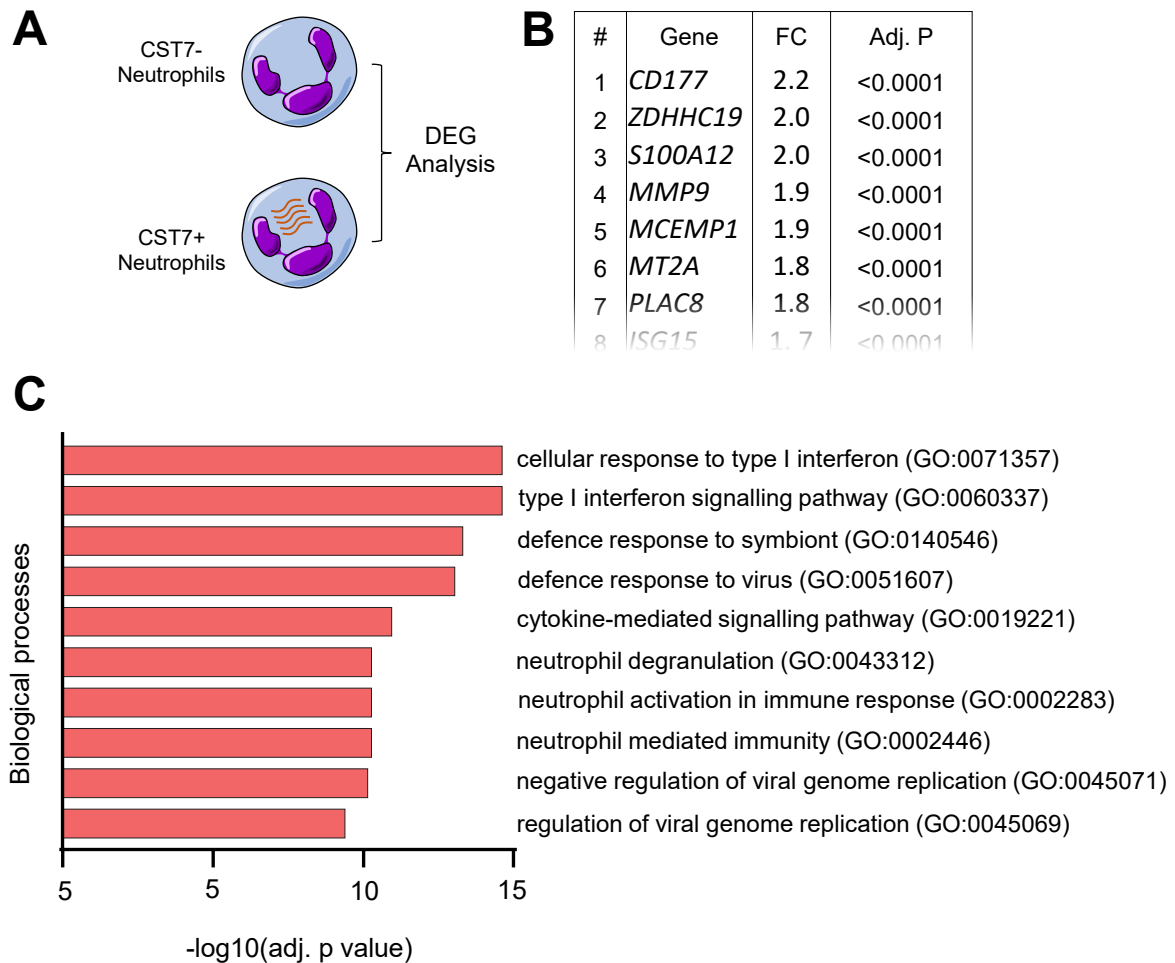


Figure 3.13. CST7 is upregulated in neutrophils during severe COVID-19. (A)

Schematic overview of the comparison performed between *CST7*⁺ neutrophils with *CST7*⁻ neutrophils. **(B)** The top 50 DEGs between *CST7*⁺ neutrophils to *CST7*⁻ neutrophils ranked according to fold change. **(C)** Upregulated biological processes of *CST7*⁺ neutrophils compared with *CST7*⁻ neutrophils ranked according to adj. P-value.

3.4 Discussion

Cystatin F has been poorly characterised as a protease inhibitor and has not been linked to neutrophils until now. By using public transcriptome data, I discovered that *CST7* expression is robustly increased in neutrophils cultured in vitro in the presence of plasma from sepsis patients and also in the circulating neutrophils of sepsis patients in vivo. Further analysis using additional datasets established that this neutrophil-specific *CST7* upregulation occurs across a diverse range of sterile and pathogen-driven inflammatory conditions. Publicly available transcriptomic data also showed that this induction is independent of a broad range of microbial products, suggesting that *CST7* upregulation during inflammation is regulated by a circulating endogenous soluble factor. The heightened expression of interferon stimulated genes (ISGs) in *CST7*⁺ neutrophils compared to *CST7*⁻ suggests type I IFN signalling may be involved in neutrophilic regulation of *CST7*. These novel findings suggest that *CST7* is a previously unrecognised major inflammation-responsive gene in humans. Given the known anti-cytotoxic actions of cystatin F in NK cells (Nanut et al., 2017), further in-depth investigations focusing on its expression patterns and functions may reveal fundamental mechanisms regulating inflammation in human disease.

My literature search of the PubMed database revealed that no link has been established between *CST7* expression in the context of sepsis or in neutrophils. This approach, which has been used in previous studies (Rinchai et al., 2015, Toufiq et al., 2020, Huang et al., 2021) is a simple and rapid method for establishing novel gene expression in the context of a condition or cell of interest. Subsequent full text searches using Google Scholar revealed that *CST7* was identified as a DEG during sepsis in three studies (Chen et al., 2019, Ahmad et al., 2019, Xie et al., 2020). However, in each of these studies *CST7* appeared among dozens of other DEGs and none of these studies commented on the biological

relevance of *CST7* upregulation during sepsis. Thus, I consider that *CST7* upregulation in neutrophils is a novel biological finding in the context of sepsis.

Throughout this chapter I performed multicohort analyses of genes of interest to determine an overall fold change in the expression of a gene across any number of compatible datasets. This type of multicohort gene expression analysis has been used to validate biomarker gene signatures across inflammatory diseases (Andres-Terre et al., 2015, Sweeney et al., 2016), including sepsis (Sweeney et al., 2015). Here I apply this approach to a single gene across a variety of different diseases and conditions. Comprehensive exploration of the NCBI GEO was performed to identify biologically relevant datasets for each disease and cell-type explored in this study, which led to the discovery of the novel association between neutrophil-expressed *CST7* and acute inflammation in humans. Using my analysis pathway, I additionally identified novel neutrophil-specific changes in the expression of two other genes, *SAMSN1* and *BATF* during inflammation, and each of these genes could form the basis of an additional study. Overall, the multicohort analysis process is straightforward and provides a wealth of information on gene regulation to inform hypotheses for future experiments. Moreover, the entire process is free and uses publicly available data and software. The primary limitation of this study (and similar public-data gene discovery research) is that the effects of changes in gene expression cannot be readily determined or explored at the protein level. Changes in gene expression do not always correlate linearly with changes in the level of functional protein or their biological effects and indeed my approach of identifying genes based on highest changes in fold change may have missed genes with a small fold change but high degree of biological significance.

With the current knowledge of *CST7* in the literature, the new findings generated from this study can be used to inform hypotheses for future research. Of particular interest, was the ability of cystatin F to inhibit cathepsin C and thus inhibit the activation of serine proteases

(Hamilton et al., 2008). While cystatin F can function in an autocrine manner, it can also act in trans, being secreted and taken up by other cells. If the upregulation of *CST7* in neutrophils corresponds to an increase in cystatin F protein, it can be speculated that it acts on the neutrophils themselves, other cells, or both. If neutrophilic cystatin F acts in an autocrine manner, this would be of great interest given its known function to suppress cathepsin C.

I hypothesise that cystatin F suppression of cathepsin C in neutrophils may be an as-yet unknown anti-inflammatory function in neutrophils. If neutrophilic cystatin F is secreted, then it may act to suppress cathepsin C in other cells. Furthermore, if cystatin F is secreted by neutrophils in a tumour microenvironment (Zhou et al., 2019), then it may contribute to pro-tumour immunosuppression. I identified a dataset (GSE101584) for tumour-associated neutrophils in a mouse cancer model (Shaul et al., 2016) and found a significant increase in *CST7* expression compared to that in naïve neutrophils. Given that cystatin F also inhibits cathepsin C in mice (Shimizu et al., 2017), this supports the hypothesis that cystatin F plays a role in the tumour microenvironment. The ability of cystatin F to inhibit cathepsin C in neutrophils is a key area for future investigation. To test these hypotheses, future bench-based research is required to determine first whether the increase in *CST7* in neutrophils corresponds to an increase in cystatin F protein and whether the protein is secreted.

Intriguingly, while *CST7* was not induced in whole blood following type I IFN stimulation ex vivo, scRNAseq analysis revealed neutrophils that expressed *CST7* also had heightened expression of genes associated with a type I IFN response. This suggests that type I IFN signalling may be involved in the regulation of *CST7* in neutrophils but does not regulate it directly. There may be an additional host signal/signals required alongside type I IFN signalling to induce *CST7* expression. Future research is required to identify whether stimulation of neutrophils in addition to other inflammatory mediators is able to induce *CST7* expression.

In this study I have both identified *CST7* as a gene of interest and assessed its regulation across a variety of public transcriptomic datasets. To date, no role for *CST7* has been established in neutrophils or in any inflammatory disease. Having identified *CST7* upregulation across inflammatory diseases and in a neutrophil-specific manner, I was able to formulate hypotheses about the potential inhibitory activity of cystatin F. My findings suggest that cystatin F or *CST7* may potentially be used, either independently or in conjunction with C-reactive protein, as an indicator of acute inflammation in the clinic.

Chapter 4

Development of spatial tools to define intra- lesion cell distribution and inter-lesion relationships

4.1 Introduction

Spatial biology is a rapidly developing field that studies cells and molecules within their spatial context using tissue imaging. Many tools have been developed to study the spatial relationships between cells, such as nearest neighbour analysis (Attrill et al., 2022a, Attrill et al., 2022b). While cells can be studied in the context of other cells in tissue, they can also be studied in the context of broader tissue architecture, with the most prominent examples of this being in cancer research with the study of leukocytes in relation to the tumour border.

Infiltration analysis can measure the proximity of surrounding leukocytes to the tumour as well as their infiltration into the tumour mass itself and leukocyte infiltration has been associated with better patient prognosis in many studies (Pagès et al., 2010, Grabovska et al., 2020). While infiltration analysis has been effective to measure the distribution of cells in relation to the border of solid tumour, there remains an absence of spatial tools to characterise cell distribution in other types of lesions, such as granulomas.

Granulomas are organised lesions that occur in a number of diseases (Segal et al., 2000, Modlin and Rea, 1988, Ramakrishnan, 2012). Granulomas are comprised of a diverse assortment of leukocytes and have varying spatial cellular distributions including lesions with a solid cellular core and lesions with a necrotic acellular core (Hunter et al., 2011). TB is a major disease associated with the formation of granulomas, and TB granulomas serve as sites of infection and cause of immunopathology (Hunter et al., 2014, Medlar, 1948). An understanding of the spatial cellular structure of the TB granuloma will be important in understanding why some granulomas resolve while others facilitate bacterial persistence and form necrosis. The development of future drugs targeting *M.tb* will also benefit from an understanding of granuloma structure as drugs must penetrate through granulomas to reach the bacteria within. Previous studies that have imaged the TB granuloma with single cell

resolution have examined cell relationships within selected 20x fields of view (FOV) from the TB infected lung, rather than attempting to characterise the cellular distribution of entire granulomas (McCaffrey et al., 2022, Abengozar-Muela et al., 2020). Therefore, the cellular structure of the TB granuloma remains poorly defined.

In this chapter I have developed a series of spatial algorithms to characterise cell distribution of entire granulomas built on whole slide imaging with single cell resolution. I first develop a tool to quantify the intra-lesion cell density and reveal whether the cells comprising the granuloma tend to distribute more centrally or peripherally in the lesion. I next develop a tool to quantify whether cells major immune cell lineages localise more centrally or peripherally in the lesion relative to the other cells in the lesion. I finally develop a tool to quantify the distance between each lesion in the tissue to address research questions regarding the organisation of lesions across the entire tissue section. For each tool, I also develop an appropriate method to visualise and statistically analyse the data produced. In the following section I describe the methods, results, and discussion specific to developing these spatial analytic tools.

4.2 Results and discussion

4.2.1 TB lesions have diverse intra-lesion cell density

Necrotising TB granulomas are distinguished from non-necrotising granulomas by the presence of a central necrotic area. Necrotising granulomas can be clearly delineated from non-necrotising granulomas in haematoxylin and eosin (H&E) staining by the presence of necrotic tissue in the lesion core (Fig. 4.1A), as well in DAPI staining by an absence of nuclei in the central region of a lesion (Fig. 4.1B). Upon staining TB lung with DAPI, I observed that the cells within individual non-necrotising lesions were unevenly distributed (Fig. 4.2A). This ranged from lesions with a dense, cellular central area to others with a sparse, cellular central area. This diversity in lesion structure was especially evident when visualised with a density heatmap (Fig. 4.2A). Since there is no available analytic tool capable of documenting the spatial organisation unique to mycobacterial granulomas, I thus sought to develop an algorithm to quantify intra-lesion cell distribution. Ideally such an equation should give a single number readout that can quantify intra-lesion cellular structure, ranging from tight cellular distribution through to a loose and finally necrotic structure (Fig. 4.2B).

The software HALO (Indica Labs) was chosen as the program in which to develop this algorithm for several key reasons. Firstly, the annotation tool in HALO allows for the borders of lesions to be defined precisely according to lesion shape. Additionally, the infiltration analysis tool, within the spatial analysis module of HALO, enables cell quantification across spatially defined regions. It can measure the density of cells within set ranges of annotated areas, as well as assign a distance to each individual cell from annotated areas such as lesion borders. Finally, HALO is powerful enough to process large tissues in excess of 1 million cells and is able to run batch analyses of the infiltration tool which can

process hundreds of lesions within a section from a single command without the need for additional user input.

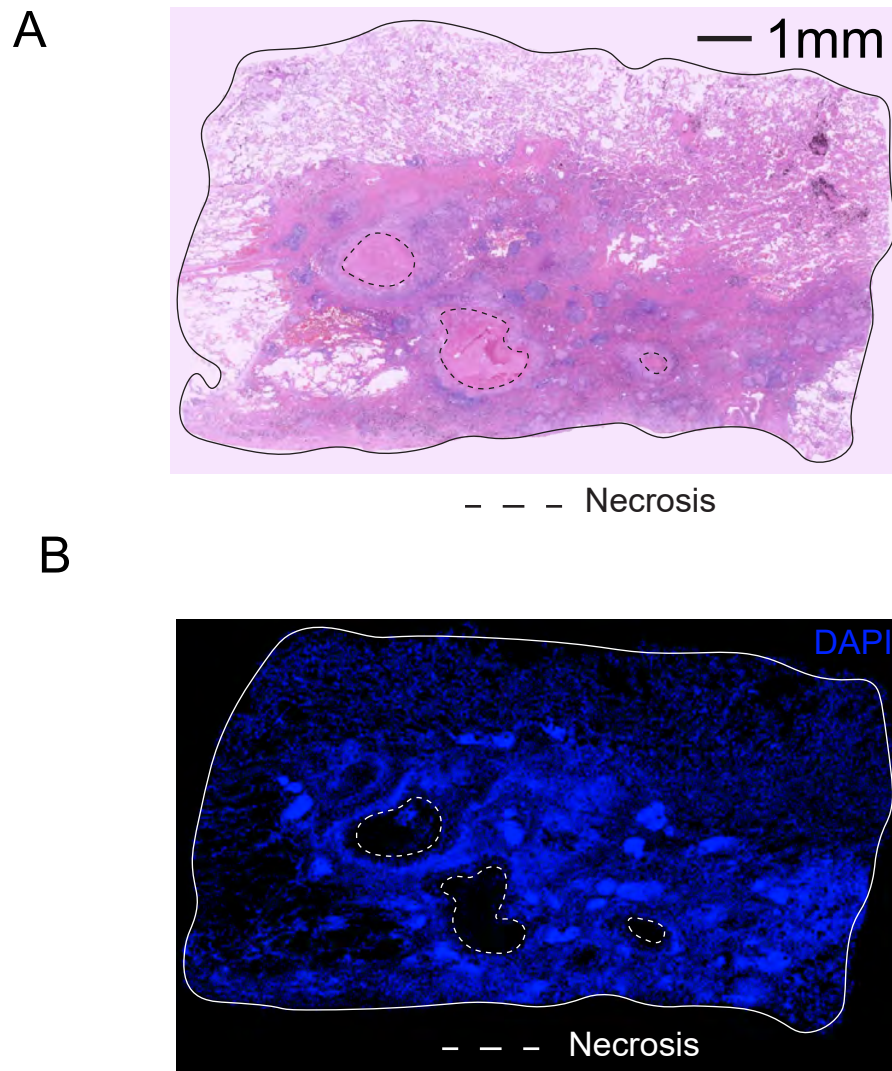


Figure 4.1. Diverse pathological presentations in *M.tb*-infected human lung. Serial lung tissue sections from a TB patient stained with **(A)** H&E or **(B)** DAPI with necrotising areas illustrated by dotted lines.

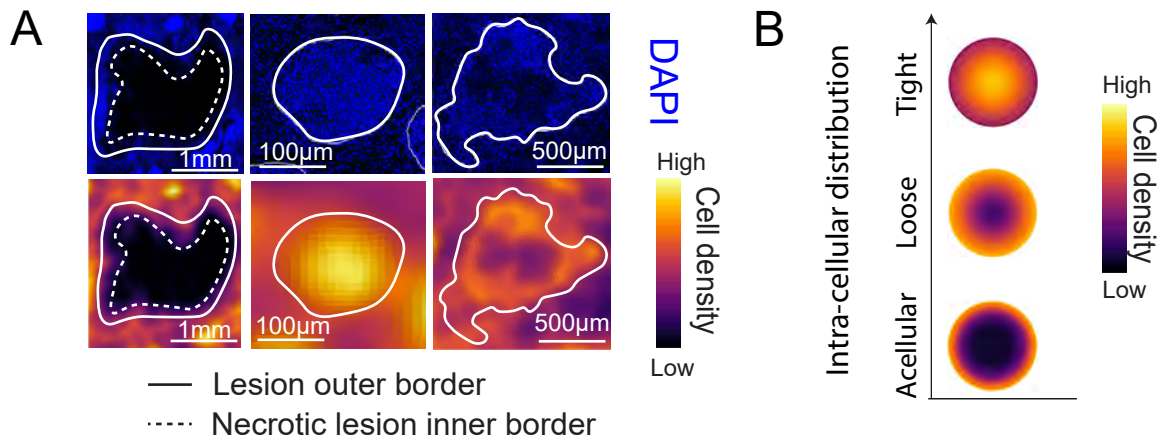


Figure 4.2. Cells in non-necrotizing lesions show diverse spatial distribution preference.

(A) Images of DAPI-stained tissue showing variations in intra-lesion cell distribution (top panel). Heatmap images showing the same three lesions with relative cell density in the central and peripheral regions indicated (bottom panel). **(B)** Schematic diagram illustrating the changing intra-lesion cell distribution ranging from a lesion with a dense centre to a lesion with an acellular centre.

4.2.2 Lesion centroid is not a reliable reference point for quantifying cellular centrality of TB granulomas

To develop an algorithm to quantify intra-lesion cell distribution (central verses peripheral localisation), it was first necessary to define the central and peripheral regions of each lesion. The most intuitive way to approach this was to identify the centroid of each lesion and use this as a reference point around which to measure cell density (Fig. 4.3A). The centroid of each lesion was calculated by first drawing a rectangular bounding box around the lesion using the maximum and minimum X and Y values of the lesion. The X and Y locations of the lesion centroid were then calculated as

$$X_{Centroid} = \frac{X_{Maximum} + X_{Minimum}}{2}$$

$$Y_{Centroid} = \frac{Y_{Maximum} + Y_{Minimum}}{2}$$

The distance of each cell in the lesion from the centroid can then be calculated. To do this, the centroid of each lesion must be indicated in HALO using the annotation tool and all centroid annotations can be stored together in the same annotation layer (Fig. 4.3B). The infiltration analysis tool can then be used to measure the distance of all cells in the lesion from the centroid. The centroid must be selected as the interface layer and the layer containing the lesion border must be selected as the tissue annotation layer. The outside range must be set to a value high enough to cover the entire lesion and “Store object (cell) data” must be set to true in order to acquire an individual readout for each cell rather than simply an average distance readout for all cells in the lesion. Following infiltration analysis, the object

data can be exported into R and the location distance of all cells in a lesion from the centroid can then be visualised with the ggplot2 package.

To determine whether distance from the lesion centroid provides an accurate estimation of the central and peripheral areas of a lesion (Fig. 4.3C). I also coloured all cells in each lesion according to whether they fell above or below the median distance from the lesion centroid (Fig. 4.3D). Ideally the lesion should appear visually separated into the inner 50% of cells and outer 50% of cells. I observed that in lesions with a circular shape (high circularity score), using the lesion centroid was able to accurately divide lesions into central and peripheral regions. However, in lesions with an irregular shape (lower circularity score) the 50% of cells closest to the centroid did not correspond with the innermost 50% of cells in the lesion and even included some cells from the lesion border. This issue became more pronounced the lower the lesion circularity score (Fig. 4.3D). To calculate the lesion circularity score required knowledge of the lesion area and border length, which were both obtained in the annotations panel in HALO. The circularity score was calculated as

$$Circularity = \frac{Perimeter^2}{4\pi \times Area}$$

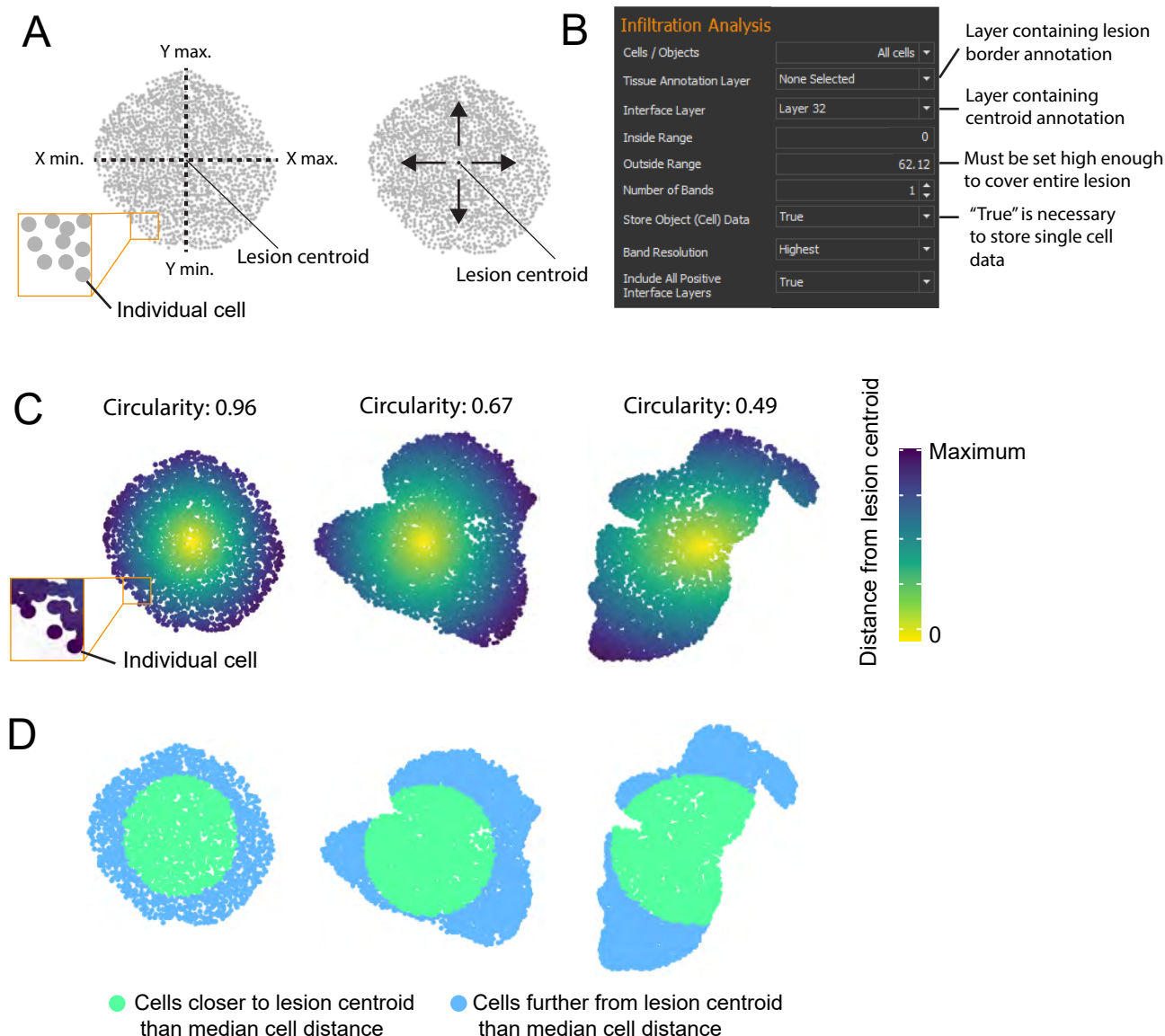


Figure 4.3. Lesion centroid is not an accurate reference point for quantifying cell

centrality within mycobacterial lesions. (A) Schematic lesion images with individual cells, lesion centroid, as well as maximum and minimum X and Y cell values indicated (left panel).

The same lesion image with the direction of measurement outward from the centroid indicated (right panel).

(B) Settings required to measure the distance of cells away from the lesion centroid using the infiltration analysis tool in HALO software.

(C) Three lesion images with different circularity showing the distance of each cell from the lesion centroid indicated by colour. The circularity of each lesion is shown above each lesion images.

(D) The same three lesion images as (C) with cells coloured based on their distance from the centroid. The blue and green colours indicate their distance from the lesion centroid, whether greater or lesser than the median distance of all cells in the lesion respectively.

4.2.3 Central and peripheral regions can be accurately defined by the distance of individual cells from the lesion border

Given the shortcomings of attempting to define the central region of a lesion relative to the lesion centroid, I developed an alternative approach which measures the shortest possible distance of each cell to the lesion border (Fig. 4.4A). With this approach the cells with furthest distance to the lesion border are considered the most central cells in the lesion, while those with the shortest distance to the lesion border were considered the most peripheral.

To calculate the minimum distance of each cell within a lesion to the lesion border the infiltration analysis tool in HALO was employed. The lesion border was selected as both the interface layer and the annotation layer (Fig. 4.4B). The outside range was set to zero and the inside range was set large enough such that the entire lesion area would be analysed. Once again, the “Store object (cell) data” was set to true to obtain single cell data.

Cells were again coloured according to distance, but this time by distance from lesion border (Fig. 4.4C) and I again separated cells according to whether they fell above or below the median distance from the lesion border (Fig. 4.4D). I found that using distance from the lesion border provided excellent separation of the central and peripheral regions of lesions and did so regardless of the lesion circularity score. Therefore, I considered the distance from the lesion border of each cell as a way to classify how centrally each cell in the lesion was located in the intra-lesion space.

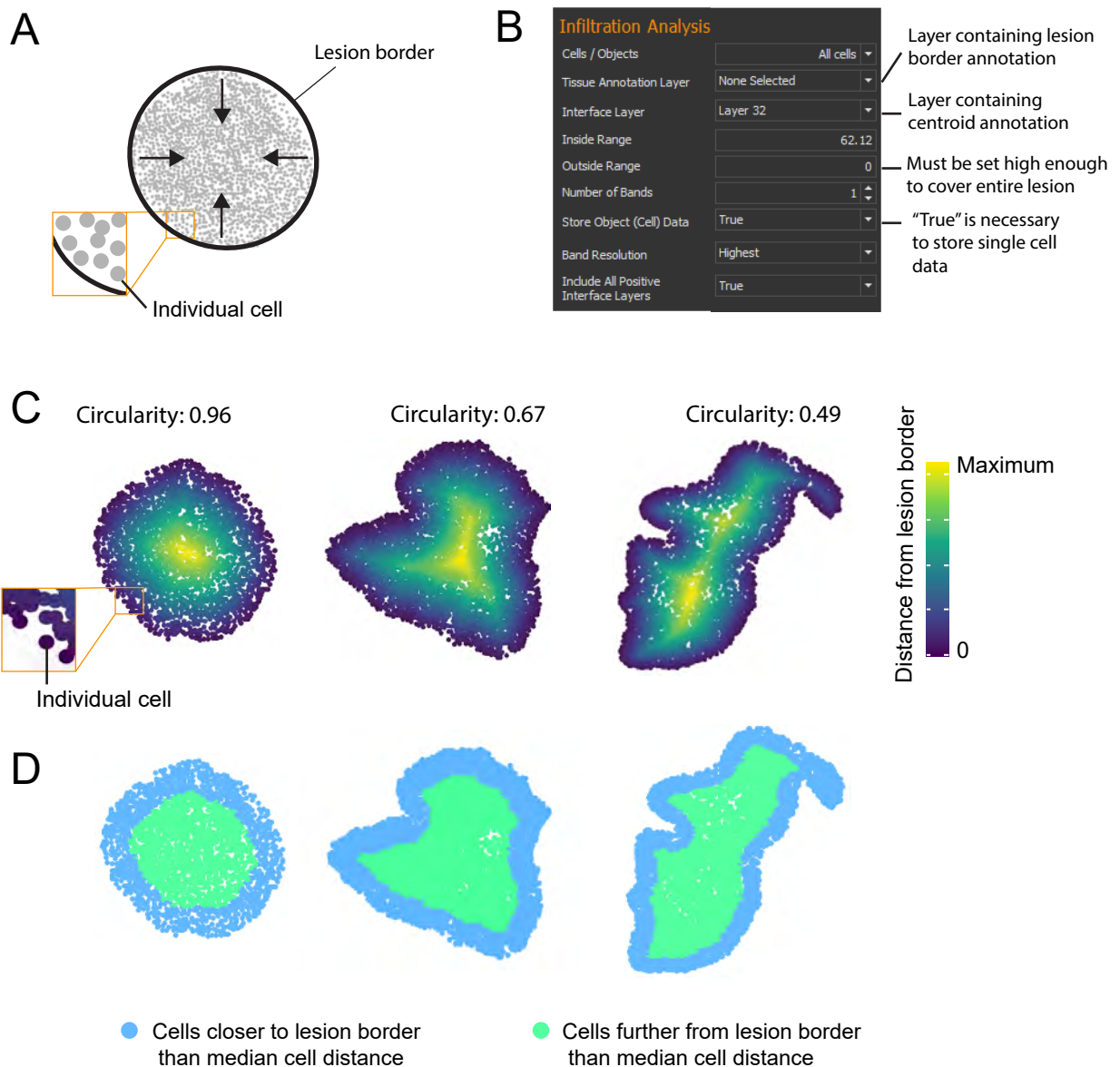


Figure 4.4. Lesion border is a useful reference point for defining central and peripheral intra-lesion cell distribution. (A) A schematic lesion image with the individual cells, lesion border and the direction of measurement inward from the border indicated. (B) Settings required to measure the distance of cells away from the lesion border using the infiltration analysis tool in HALO software. (C) Three lesion images with the distance of each cell from the lesion border indicated by colour. The circularity of each lesion is also indicated above each lesion. (D) The same three lesion images as (C) with cells coloured according to whether their distance from the lesion border is greater or less than the median cell distance in the lesion.

4.2.4 Total Central Preference Index (tCPI) is a novel algorithm that defines intra-lesion cell density

Having a method to quantify central and peripheral lesion regions, I aimed to develop a single digit readout of whether the intra-lesion cell density was highest in the lesion centre (dense core) or in the periphery (sparse or necrotic lesion core). I observed that the intra-lesion cell density changed depending on the median distance of cells from the border. When the median distance of cells from the lesion border was relatively low, this corresponded with a lower density of cells in the lesion centre. On the contrary, when the median distance of cells from the lesion border was relatively high, this corresponded with a higher density of cells in the lesion centre (Fig. 4.5A). However, the median distance of cells from the lesion border could not be used as a tool to quantify intra-lesion cell density across lesions as it also changed depending on lesion size (Fig. 4.5B). Two lesions of different sizes will have different median distances from the lesion border despite the fact they have the same intra-lesion cell density.

I also observed that the relative areas covered by the inner and outer 50% of cells appeared to change depending on the intra-lesion cell density. The inner 50% of cells covered a relatively small area in lesions with a tight structure due to the high cellular density in the lesion centre. Comparatively the inner 50% of cells covered a relatively large area in lesions with a loose structure due to the low density of cells in the lesion centre (Fig. 4.5C). The ratio between size of areas covered by the inner and outer 50% of cells could therefore be used as a tool to quantify the intra-lesion cell density. Additionally, because this measurement is a ratio, it is not affected by lesion size and can be used comparatively between lesions.

The area covered by the outer 50% of cells was calculated in HALO using the infiltration analysis tool and setting the lesion boundary as both the interface and tissue

annotation layer. The median distance of cells from the lesion border was input as the inside range and “Store object (cell) data” set to false as single cell data was not required. Following infiltration analysis, the area covered by the outer 50% of cells was obtained as the “Total interface area (mm²)” from the summary data (Fig. 4.5D). The area covered by the inner 50% of cells was calculated by subtracting the area of the outer 50% of cells from the total lesion area.

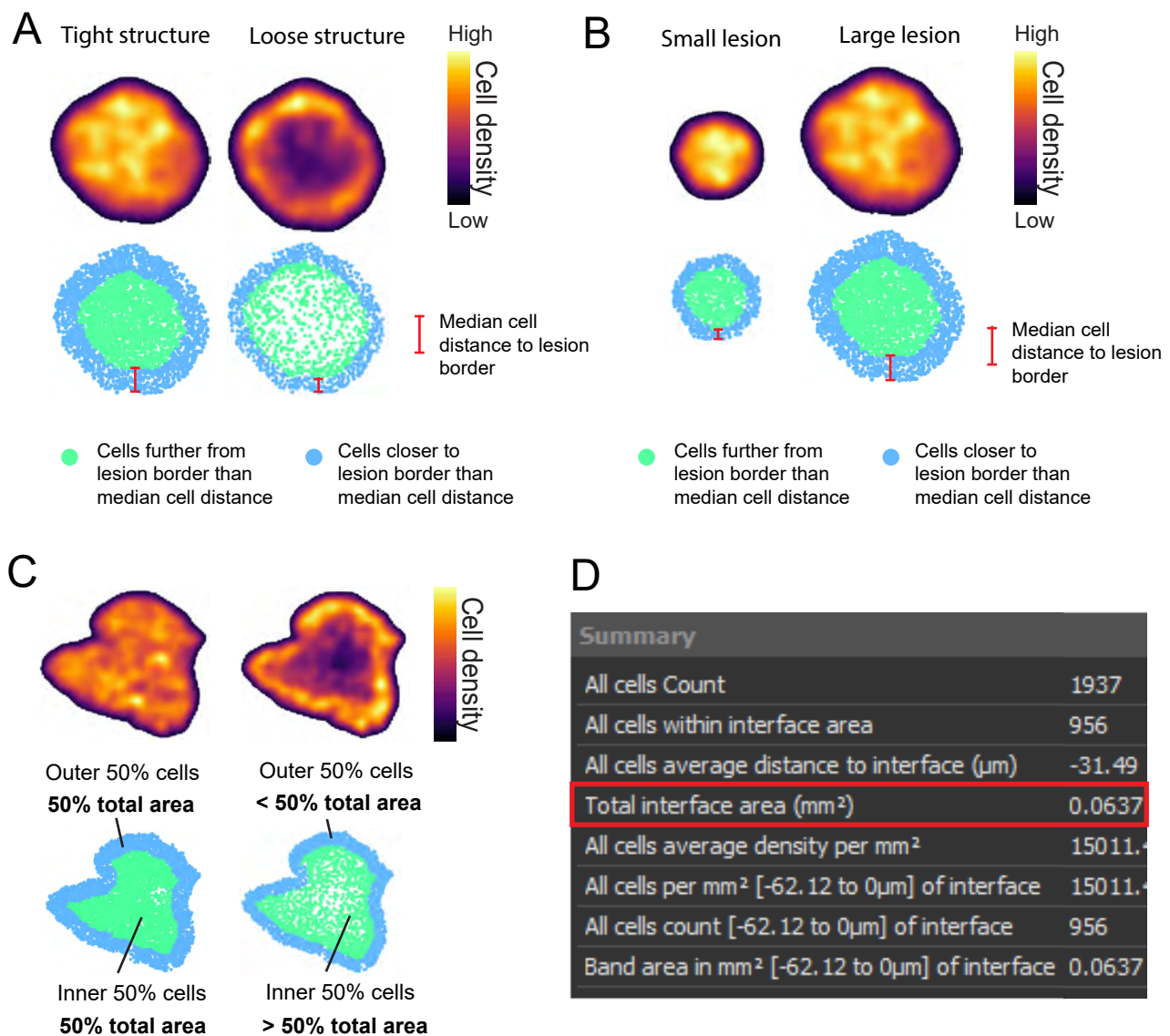


Figure 4.5. Intra-lesion cell density can be quantified by comparing the areas covered by the inner and outer 50% of cells. (A) Density heatmaps of two lesions differing in intra-lesion cell distribution (top panel). Lesion images of the same two lesions coloured according to whether the distance of the individual cells from the lesion border is greater or lesser than the median cell distance in the lesion (bottom panel). The median distance to the lesion border is also indicated. **(B)** Density heatmaps of two lesions with different sizes (top panel). Lesion images of the same two lesions coloured according to whether their distance from the lesion border is greater or lesser than the median cell distance in the lesion (bottom panel). The median distance to the lesion border is also indicated. **(C)** Density heatmaps of two lesions with different intra-lesion cell distributions (top panel). Lesion images of the same two lesions coloured according to whether their distance from the lesion border is greater or lesser than the median cell distance in the lesion (bottom panel). The areas occupied by the inner and outer 50% of cells are indicated. **(D)** Settings required to calculate the area of the outer 50% of cells using the infiltration tool in HALO software.

The final algorithm used to quantify intra-lesion cell density was termed the total Central Preference Index (tCPI) (Fig. 4.6A) and was calculated as

$$tCPI = \frac{\textit{Peripheral area}_{(\textit{Area of outer 50\% of cells})}}{\textit{Total area}_{(\textit{Area of all cells})}}$$

tCPI will return a value greater than 0.5 if the lesion has a tight central structure and a value less than 0.5 if the lesion has a loose or necrotic central structure (Fig. 4.6B). A value of 0.5 implies that the density of the central and peripheral regions of the lesion are the same.

Therefore, tCPI allows for a simple readout that quantifies the overall cellular structure of lesions and is comparable between lesions of any size and any shape.

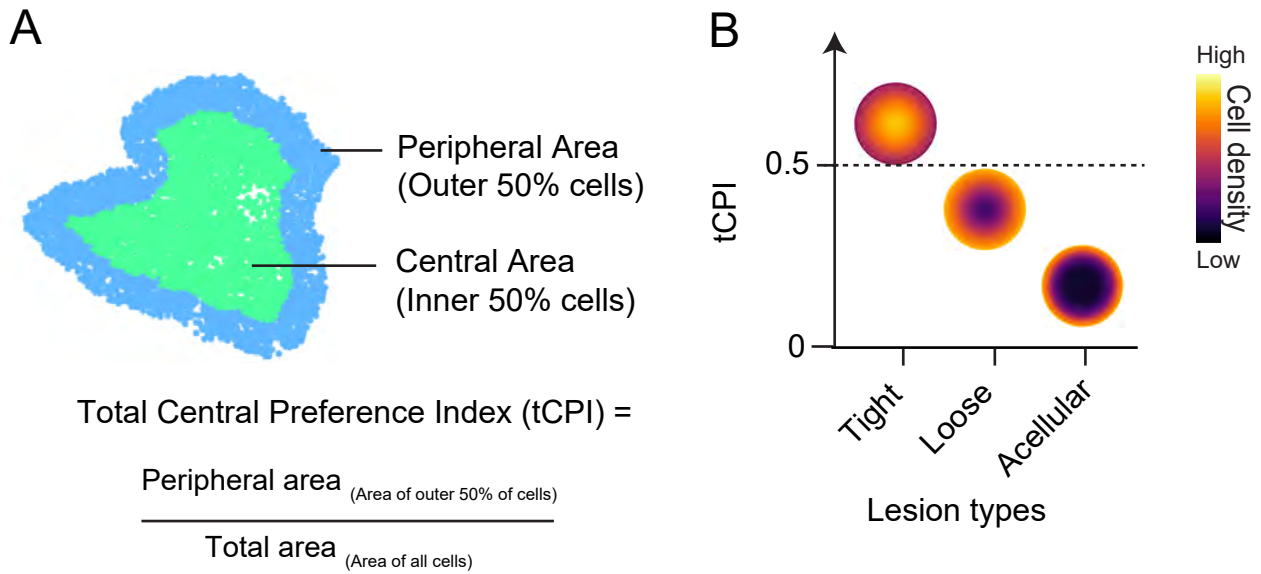


Figure 4.6. tCPI is a single number measurement for intra-lesion cell localisation

preference. (A) A lesion image with areas coloured according to which cells are closer or further from the lesion border relative to the median distance (top). Equation used to calculate tCPI (bottom). **(B)** Schematic diagram illustrating the intra-lesion cell distribution patterns and their corresponding tCPI.

4.2.5 Band-based infiltration analysis is not practical for quantifying intra-lesion cell-type distribution

While tCPI is designed to quantify the spatial distribution of total cells within a lesion, it is also of interest to calculate the relative locations of specific cell types. This can inform whether a given cell type is present in a central or peripheral location within the lesion and may also inform the level of accessibility that cell type has to pathogens within the lesion. In cancer studies, this is usually quantified by measuring the density of each cell type within a series of bands moving inward from the outer border of the tumour and is used to determine the depth of penetration of leukocyte populations into the tumour tissue (Fig. 4.7A). This approach can be performed in HALO using the infiltration tool with as many bands as desired (Fig. 4.7B).

The typical output of band-based infiltration analysis is the density histogram, in which each bar corresponds to the density of a cell of interest within bands, progressively moving inward from the lesion border (Fig. 4.7C). This is an effective way to visualise the data if the goal is to present the overall distribution of cells in a representative lesion or to compare the density of cells in specific bands between multiple lesions. However, to visualise the distribution of cells across multiple lesions requires multiple density histograms, and while this may be practical in a small number of lesions, when the number increases to dozens of lesions this approach becomes impractical as a tool to visualise the intra-lesion distribution of cell-types (Fig. 4.7D).

I designed an alternative approach to visualise cell distribution across multiple lesions, also based on band-based infiltration analysis, but using heatmaps instead of density histograms to visualise the data (Fig. 4.7E). In this approach, one axis of the histogram corresponds to bands from the lesion border, the other axis corresponds to each lesion and the

colour corresponds to cell density. A single histogram is required for each cell type. This approach enables the visualisation of the overall distribution of each cell type in a large number of lesions. However, this approach also has several drawbacks. As the data for each cell-type is stored as the density of multiple bands across a large number of lesions, there is no simple way to quantify and compare the intra-lesion distribution between different cell-types statistically. Further, if a comparison is warranted between both cell-types as well as different types of lesions (i.e., necrotising, non-necrotising), then an array of heatmaps is required (Fig. 4.7F). Additionally, heatmaps are impractical to visualise intra-lesion cell distribution when the number of lesions analysed is in the hundreds, as the heatmap size must become impractically large.

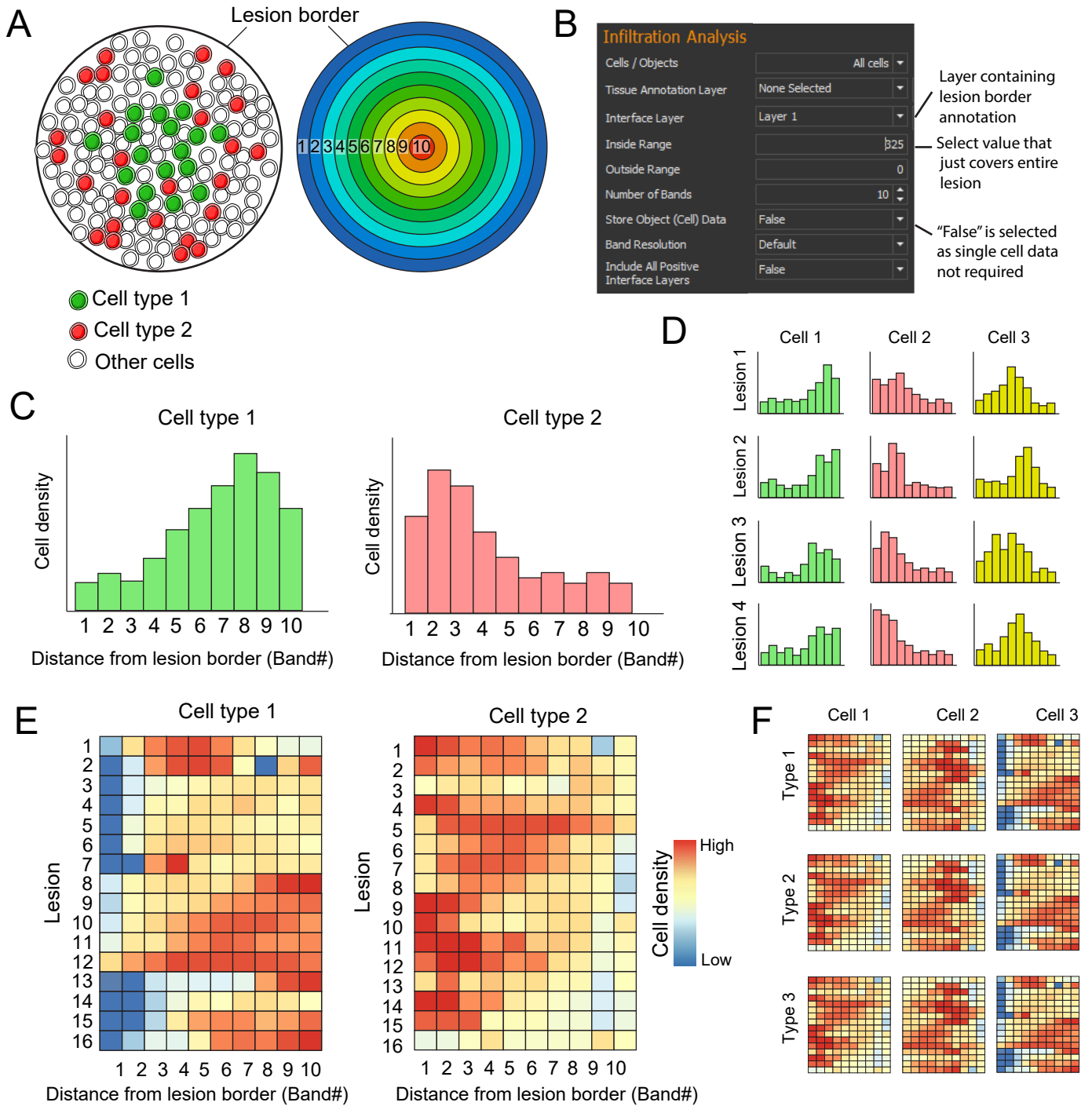


Figure 4.7. Band-based infiltration analysis is not practical for quantifying intra-lesion cell-type distribution in high numbers of lesions and across lesion types. (A) Schematic lesion image with cells coloured according to cell type (left). Lesion image with bands indicating regions in ten increments from lesion border (right). In both lesion images the lesion border is indicated. (B) Settings required to measure the density of a cell population in incremental bands from the lesion border using the infiltration analysis tool in HALO software. (C) Example density histograms indicating the cell density of two cell types in ten bands from the lesion border of a single lesion. (D) Example density histograms to visualise the cell density in ten bands from three cell types in four lesions. (E) Example heatmaps indicating the cell density of two cell types in ten bands from the lesion border from sixteen lesions. (F) Example heatmaps required to indicate the cell density of three cell types and lesions from three lesion types.

4.2.6 Immune Cell Preference Index (immCPI) quantifies and enables visualisation of the intra-lesion cell-type distribution across any number of lesions and lesion types

It was evident that any analysis tool to compare the intra-lesion distribution of each cell-type between lesions should rely on a single readout for each cell type per lesion, rather than a readout across multiple bands. A single readout will enable ease of visualisation and statistical comparison between high numbers of lesions and lesion types. To develop such an analytical tool, I started with the mean distance of each cell-type from the lesion border as well as the mean distance of all cells in the lesion from the lesion border (Fig. 4.8A). This was calculated in HALO using the infiltration analysis tool, with the lesion border set as both the tissue annotation layer and the interface layer (Fig. 4.8B). The cells/objects was set to only the cell of interest and inside range was set high enough to cover the entire lesion. “Store object (cell) data” was set to true to store single cell data.

I observed that the mean distance of each cell type to the lesion border changed depending on whether the cell type was more peripherally or centrally distributed within the lesion (Fig. 4.8C). To avoid this measurement being confounded by lesion size or lesion tCPI, I normalised the mean distance of each cell of interest to the mean distance of all cells in the lesion. This measurement was termed immune cell preference index (immCPI) (Fig. 4.8D) and was calculated as

$$immCPI = \frac{\text{Mean distance from lesion border}_{Cell\ of\ interest}}{\text{Mean distance from lesion border}_{All\ cells}}$$

When the immCPI calculation returns a value of 1, this means that the cell type of interest is evenly distributed in the lesion relative to all cells in the lesion (Fig. 4.8E). If the immCPI is

higher than 1, this implies that the cell type of interest is distributed closer to the lesion centre on average relative to all cells in the lesion. For example, an immCPI measurement of 2 indicates that the cell type being measured is on average twice as close to the lesion centre compared to all cells in the lesion. On the contrary, if the immCPI measurement is below 1 this indicates that the cell type of interest is located more peripherally in the lesion compared to the average cell in the lesion and a value of 0.5 would imply that the cell type is located twice as close to the lesion periphery compared to the average of all cells.

As the immCPI measurement returns a single value for each cell type in a lesion, statistical comparisons can readily be made between different cell types across an indefinite number of lesions or of the same cell type different lesion types (Fig. 4.8F). The immCPI of each cell type in a lesion can be visualised as a single point in a plot and large numbers of lesions can be readily visualised with violin plots or box plots. Thus, the immCPI measurement offers a practical approach to both visualise and statistically analyse the intra-lesion distribution of different cell types across large numbers of lesions and lesion types.

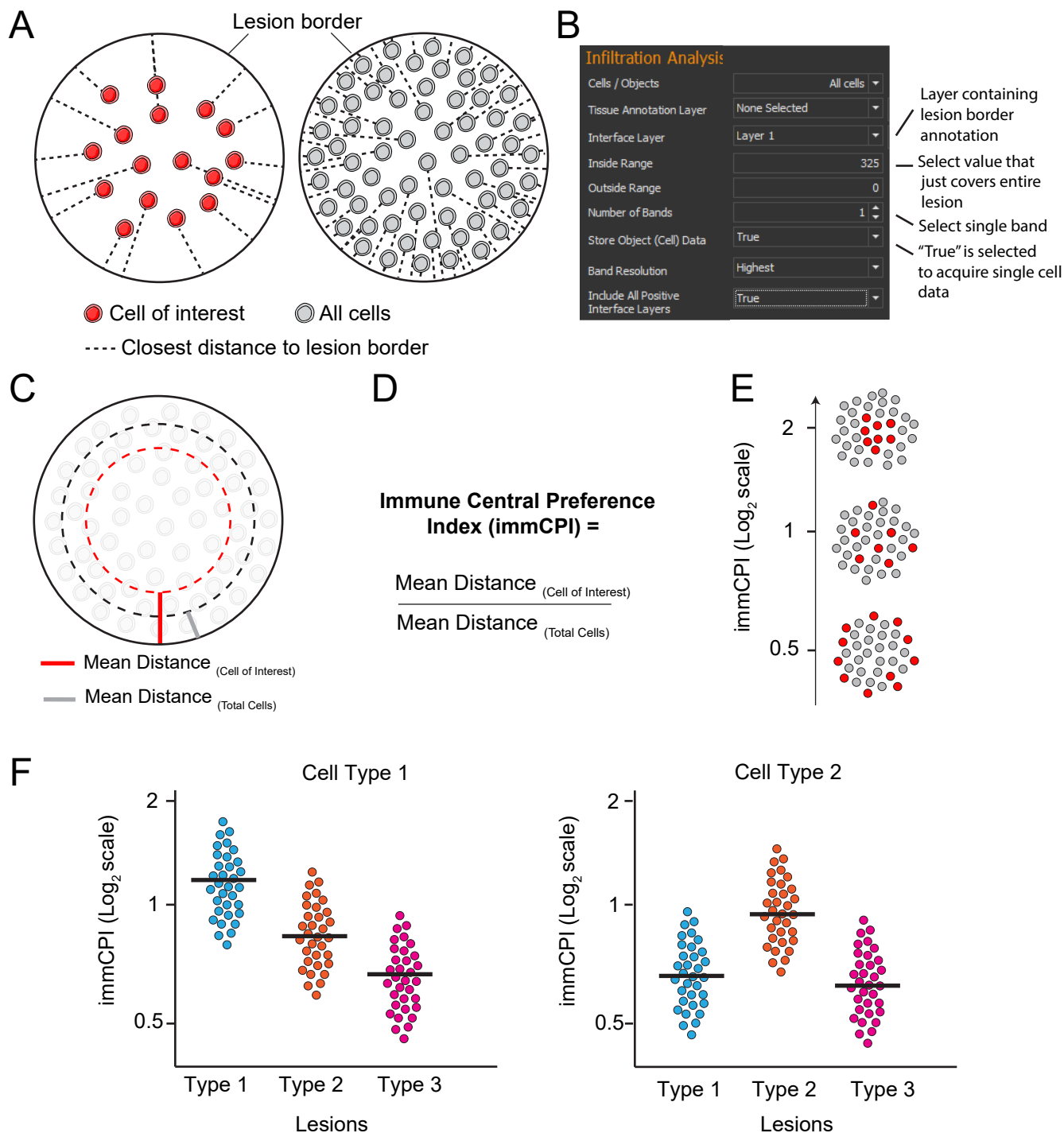


Figure 4.8. immCPI quantifies intra-lesion cell-type distribution in any number of lesions and across lesion types. (A) A lesion image with only the cell of interest visualised and the closest distance of each cell to the lesion border indicated (left). A lesion image with all cells (grey) illustrated and the closest distance of each cell to the lesion border indicated (right). The lesion border is also indicated in both lesion images. (B) Settings required to measure the distance of each cell from the lesion border using the infiltration analysis tool in HALO software. (C) A schematic lesion image with the mean distance of both the cell of interest and all cells indicated by bands within the lesion. (D) Equation used to calculate immCPI. (E) Schematic diagram of lesions with representative immCPI measurements. Red, cells of interest; grey, the rest of the cells. (F) Example plots comparing the immCPI of two cell types between different lesion types.

4.2.7 Lesion neighbourhood analysis quantifies the inter-lesion tissue superstructure

Having established two spatial analysis tools for intra-lesion cellular analysis, I also aimed to develop a tool to quantify the spatial relationships between lesions across the entire tissue section (Fig. 4.9A). This first required the measurement of the shortest possible distance between each pair of lesions in the tissue (Fig. 4.9B). HALO does not have the capacity to perform this measurement, so the object data file for each patient sample was exported into R. R offers greater flexibility in the spatial calculations that can be performed due to its command line interface (CLI) in which user generated functions can be developed. The object data file for each patient was imported into R, which is a data frame consisting of each cell and information about its features, including the cell's spatial location in the tissue. Any additional data such as whether the lesion is a necrotising or non-necrotising lesion can also be included in the object data.

The `pointDistance` function from the Raster package (Hijmans. and Etten., 2012) was employed to calculate the minimum distance between each pair of lesions. This function operated by measuring the distances between every cell in the first lesion to every cell in the second lesion, which created large data tables for each lesion pair, with the cells of each lesion populating the two axes and the distances between every cell pair filling the table cells (Fig. 4.9C). The `minimum` function was used to identify the minimum cell-cell distance between the two lesions, and this was considered the distance between the two lesions. The minimum distances between each lesion pair were stored in a single data table for each patient (Fig. 4.9D).

This measurement was termed Lesion Neighbourhood Analysis and can be used to determine whether lesions of different types tend to locate closer or further to any other particular lesion or lesion type (Fig. 4.9E). For example, it can be tested whether two

different types of non-necrotising lesions locate differently in the tissue relative to necrotising lesions. More broadly, this analysis tool can be used to determine whether the spatial arrangement in a patient tissue is random or follows a certain pattern to form a superstructure.

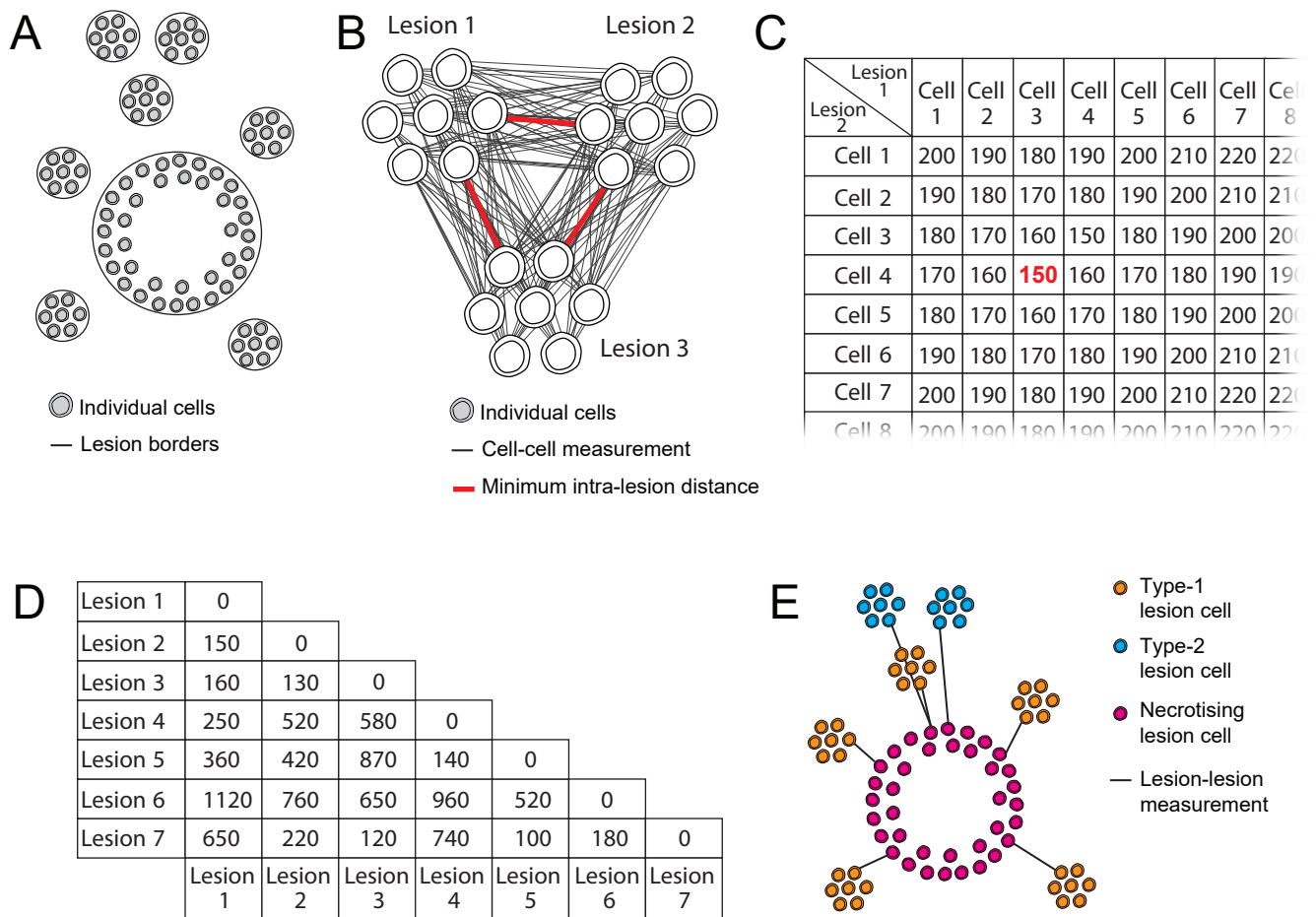


Figure 4.9. Lesion neighbourhood analysis quantifies inter-lesion distances. (A) Diagram of multiple lesions within the tissue space with individual cells and lesion borders indicated. **(B)** Diagram of three lesions with the distance between each cell pair of the three lesions indicated by a line. A red line indicates the distance between the closest cells in each lesion pair **(C)** Table generated to calculate lesion neighbourhood analysis that stores the distance between each pair of cells in two lesions. **(D)** Table generated by lesion neighbourhood analysis that stores the minimum distance between each lesion pair. **(E)** Diagram of multiple lesions of different lesion types within the tissue space. The minimum distance of each non-necrotising lesion to a necrotising lesion is indicated.

4.2.8 Generation of datasets using the LANDSCAPE method

The three tools generated in this chapter for lesion-level spatial analysis were together termed LANDSCAPE (Lesion And Neighbourhood Description Stratified by cellular Composition And Position Estimates). The data generated from performing LANDSCAPE analyses can be pooled together in R to create two types of data table from which many further analyses can be performed.

The cell data table, which is based on the HALO object data, each row represents a single cell, while each column represents a feature of that cell (Fig. 4.10A). Cells from all patients can be pooled together and information about which patient and specific lesion each cell belongs to are stored as features. Any phenotypic information for the cell can be stored here, such as what type of cell it is and whether it is positive for any functional markers. The X and Y location of each cell are also preserved in the table as well the distance of the cell from the lesion border (for cells that are part of a lesion) used in the tCPI calculation. Additionally, the nearest neighbour function in HALO (from the spatial analysis module) can be used to calculate the distance of each cell to another particular cell type. For example, the distance of every cell to the nearest CD4⁺ T cell can be calculated and stored as a feature in this table. Altogether, the data stored in this table can be used to address specific research questions, for example, do replicating B cells in the centre of type 1 lesions tend to be closer to CD4⁺ T cells compared to those in type 2 lesions? Cells with a particular phenotype and particular intra-lesion location can be readily isolated from the cell data table and a plethora of comparisons can be made between lesions.

The lesion data table generated in R is based on the HALO summary data and in this case each row corresponds to a particular lesion while each column corresponds to a feature of that lesion (Fig. 4.10B). Lesion features include basic phenotypic characteristics such as

patient and lesion size, as well as compositional information such as the number of each cell-type present in the lesion. Spatial information about the lesion can also be stored in this table, including the density and immCPI of each cell type as well as the tCPI for the lesion. Lesion neighbourhood information such as the lesion distance to the closest necrotising lesion can also be stored as a feature. The lesion data table can be used to address specific research questions, such as does the immCPI of CD4⁺ T cells in non-necrotising lesions with a low tCPI change depending on the distance of the lesion to a necrotising lesion? The multi-dimensional nature of the lesion data table allow it to be used to address questions involving both compositional and spatial features of the lesion.

Overall, the LANDSCAPE method enables simple readouts of major intra- and inter-lesion spatial characteristics that can be effectively visualised and statistically analysed.

4.2.9 Limitations of LANDSCAPE analysis

There are two primary limitations on the use of the LANDSCAPE tool. HALO software is required to perform the spatial algorithms comprising LANDSCAPE, however, HALO is prohibitively expensive to many researchers. This limits the accessibility of LANDSCAPE and future iterations of the tool could be built as a plugin in a free-to-use image processing system such as ImageJ, which would enable broader access. A further limitation of the LANDSCAPE tool is that it requires basic proficiency in R in order to perform lesion neighbourhood analysis. This could be alleviated in the future through the development of an interactive web app using the Shiny R package that enables lesion neighbourhood to be calculated with a more accessible GUI.

A

Patient #	Cell #	X	Y	Lesion	Identity	Distance from lesion border (µm)	Distance from necrosis (µm)	Distance to nearest CD4+ T cell (µm)
1	1	7	5	NA	NA	NA	146	14
1	2	9	8	NA	CD4 ⁺	NA	157	47
1	3	11	7	1	CD8 ⁺	2	175	11
1	4	15	9	NA	NA	NA	158	7
1	5	18	6	1	CD68 ⁺	5	160	12
1	6	21	8	1	CD20 ⁺	6	147	18

B

Patient #	Lesion #	Total cells	Size (mm ²)	Necrotic Y / N	tCPI	CD4 ⁺ T cell density	CD4 ⁺ T cell immCPI	Lesion group	Distance to necrotic lesion
1	1	257	0.02	N	0.41	1250	0.93	1	10285
1	2	1253	0.1	N	0.55	231	1.28	2	2368
1	3	645	0.07	N	0.52	6531	0.72	1	531
1	4	12953	33.4	Y	0.23	800	0.87	3	NA
1	5	2351	0.4	N	0.45	1200	1.12	2	1245
1	6	806	0.21	N	0.32	3214	0.86	1	780

Figure 4.10. Example dataset structures generated by LANDSCAPE analysis. (A)

Example of cell data table generated by LANDSCAPE analysis that stores each cell as a row with cell feature data in the columns. **(B)** Example of lesion data table generated by LANDSCAPE analysis that stores each lesion as a row with lesion feature information in the columns.

4.3 Concluding remarks

The three analysis protocols developed in this chapter represent the first tools to study both intra-lesion and inter-lesion spatial distribution of granulomas with single cell resolution.

Data generated using these approaches can be readily visualised and statistically analysed.

These approaches can be applied to any granulomatous disease and potentially, more broadly to other types of tissue lesions. The following chapter will apply these tools to study granulomas in TB lung samples.

Chapter 5

**Spatial mapping of human lung tissue reveals
granuloma diversity and histopathological
superstructure in TB**

5.1 Introduction

TB remains a major cause of mortality with 1.5 million deaths in 2020 (WHO, 2021).

Development of new TB vaccines and therapeutics remains a global priority and will depend on a clearer understanding of mechanism mediating host resistance to *M.tb*, the causative agent of TB (O'Garra et al., 2013, Orme et al., 2015, J Libardo et al., 2018). The battle between the host and the pathogen occurs at the level of immune cell-enriched lesions called granulomas (Ramakrishnan, 2012). While being recognized as a protective structure for walling-off infection, granulomas also contribute to the survival of mycobacteria and importantly are the major manifestation of organ damage (Hunter, 2015, Hunter, 2011b, Cohen et al., 2022, Lin et al., 2014).

Despite the realisation of the critical role of granulomas in host defence, the cellular basis of mycobacterial granuloma formation is incompletely defined. Investigations of granulomas in human tissues are hampered by the lack of access to clinical samples and limited imaging technologies. In addition, commonly used laboratory mouse strains do not form lesions with the key features of human granulomas (Basaraba and Hunter, 2017, Williams and Orme, 2016, Barry et al., 2009). These hurdles are further compounded by the challenges in granuloma classification due to the high variation in lesion size and histopathological features in diseased human tissues (Cadena et al., 2017, Hunter et al., 2011). Consequently, most studies choose to characterise one or two types of lesions (e.g., non-necrotising and necrotising) in isolation. The extent of variation across non-necrotising TB lesions is unknown.

Recent advances in spatial transcriptomics and proteomics technologies enable researchers to image mycobacterium-infected tissues at various levels. In this regard, spatial transcriptomics have provided comprehensive gene expression information across tissue sections of human leprosy (Ma et al., 2021) and murine TB (Carow et al., 2019), although the

current technologies do not have the resolution required for single-cell imaging. Antibody-based multiplexed tissue imaging approaches (Abengozar-Muela et al., 2020, McCaffrey et al., 2022), including one with MIBI-TOF, have been employed to examine human TB granulomas with single-cell resolution. The MIBI-TOF study has successfully revealed diverse micro-environments within human TB granulomas (McCaffrey et al., 2022). However, this approach interrogates mainly the selected regions of individual lesions; its scope in tissue-wide cross-lesion comparative analysis is limited (Tan et al., 2020, Hartmann and Bendall, 2020). Because only a subset of lesions were analysed, it is unclear whether the observed difference in microenvironments reflects the cellular structural variations within a lesion or across multiple lesions. Therefore, revealing the level of granuloma diversity by analysing many lesions across large areas of diseased tissue will deepen the understanding of the intra-lesion microenvironment characterised by high-dimension phenotyping.

To capture all cellular structures present in a TB lung tissue sample, I have employed multiplex Opal IF to map the tissue-wide immune landscape (size ranging from 100 to 300 mm², ten TB patients in total) with single cell resolution. Moreover, I developed a novel granuloma identification strategy to minimise the subjectivity and bias commonly associated with lesion analysis. By utilising the novel single-cell spatial analysis developed in chapter 4 on 657 lesions (defined as distinct immune cell aggregates) with approximately 7 million cells, I established that human TB lung tissues contain, in addition to well-recognised necrotising lesions, at least four types of compositionally and spatially distinct non-necrotising lesions. Further spatial analysis demonstrated that the latter lesion structures are shaped by the neighbouring necrotising granuloma as well as by intra-lesion organisation of immune cells, suggesting that granuloma architecture is governed by lesion-intrinsic and extrinsic signals in human TB.

5.2 Materials and methods

5.2.1 Patient Cohort

I utilised a retrospective study cohort of patients from the Shenzhen Third People's Hospital (Shenzhen, China). The study was approved by the Research Ethics Committee of the Shenzhen Third People's Hospital (protocol ID #: 2016-081). All archival clinical specimens were analysed, with no active participation of human subjects. Active TB diagnosis was based on clinical symptoms, chest radiography, pathological result, microscopy for acid-fast bacilli, M.tb culture, and GeneXpert analysis of sputum (Tab. 5.1). For patients without positive TB culture or GeneXpert results, justification for resection was based on both radiological findings indicative of active TB and failure to respond to TB drug treatment. All patients had undergone treatment for between 3-96 months preceding resection and rifampicin resistance was identified in five out of ten patients. Serial sections (5 µm) of each specimen were stained with H&E and inspected by two clinical pathologists to evaluate the presence of granulomatous inflammation. Patients with comorbid HIV were excluded from the study. All patients with TB were undergoing anti-TB therapy.

5.2.2 OPAL staining

Patient tissues were mounted on slides and OPAL multiplex staining, image acquisition, image stitching, cell segmentation and cell phenotyping were performed as described in chapter 2.

Table 5.1. Clinical information of TB patients.

#	Age *	Sputum smear	TB culture	Gen- Xpert	Pathological diagnosis	X-ray or CT data	Surgical resection
1	38	- **	-	ND ***	There are scattered granulomas, some necrotising, consistent with pulmonary TB	Consider the possibility of pulmonary TB in the left lung	Left upper lobe lung
2	46	-	-	+ ****	There are scattered granulomatous nodules, some with caseous necrosis, consistent with pulmonary TB	The lesions in the posterior segment of the upper lobe apex of the left lung were considered to be pulmonary TB	Left upper lobe lung
3	30	-	+	+	It is consistent with pulmonary TB, with caseous necrosis and granuloma formation	Consider the possibility of pulmonary TB (TB bulb) in the left lung	Left upper lobe lung
4	28	-	-	ND	It is consistent with pulmonary TB, with caseous necrosis and granuloma formation	A few inflammatory lesions in both lungs, and the possibility of lymph node TB was considered	Right upper lobe lung
5	22	+	+	+	It is consistent with pulmonary TB, with caseous necrosis and granuloma formation	Double pulmonary infective lesions, raising the possibility of pulmonary TB, with double pulmonary cavity formation	Right upper lobe lung
6	23	ND	-	+	It is consistent with pulmonary TB, with caseous necrosis and granuloma formation	Multiple nodules in the lower lobe of the left lung, TB is not excluded	Left lower lobe lung

7	25	ND	+	+	It is consistent with pulmonary TB, with caseous necrosis and granuloma formation	Nodules in Posterior segment of right upper lobe lung	Posterior segment of right upper lobe lung
8	42	-	ND	ND	It is consistent with pulmonary TB, with caseous necrosis and granuloma formation	Nodules in the upper lobe of the left lung. A few chronic pneumonia changes in both lungs	Left upper lobe lung
9	36	-	-	-	There are scattered granulomatous nodules, some with caseous necrosis, consistent with pulmonary TB.	TB, granuloma of superior lobe of left lung	Left upper lobe lung
10	49	-	-	ND	There are scattered granulomatous nodules, some with caseous necrosis, consistent with pulmonary TB.	Inflammatory lesions, TB was considered in the left upper lobe of lung	Left upper lobe, lung apex

* Age in years

** “-“ indicates a negative result

*** ND, not done

**** “+” indicates a positive result

5.2.3 Density based tissue region identification

Kernel density analysis on the XY location of all cells within each patient sample was performed using the “kde2d” function in the MASS (Venables and Ripley., 2002) package in R using bandwidth of 500 within a 500 by 500 grid. 3D projections were created with the “persp” function and coloured with the “viridis” function from the Viridis package (Garnier et al., 2021). Relative to the maximum cell density in each patient the healthy region was defined by densities between 0-25%, consolidated region between 25-50% and aggregate region between 50-100%. To visualise the regions defined by cellular density on tissue images the Density Heatmap tool in HALO was used to map density across the entire tissue section and the Annotation tool was used to define the regions for export.

5.2.4 Density based lesion identification

To define individual lesions, Kernel Density Analysis was further performed on lesions that were stained positively for at least one of four immune cell markers (CD20, CD68, CD4 or CD8) in R. 3D density projection was again performed with “persp” function and areas between 25-100% of the maximum cell density were considered as diseased tissue areas. Lesions were identified using HALO with the density heatmap function and lesion boundaries were drawn using the annotation tool. Cell segmentation was performed in batch on all lesions and object data was exported to R.

5.2.5 Lesion clustering analysis based on intra-lesion immune cell density

The area of individual lesions was determined using HALO summary data in R. The relative abundance of individual immune cell populations in each lesion was expressed as immune cells per mm² lesion area. Lesion clustering analysis was performed using principal component analysis (PCA) with density of CD8⁺, CD20⁺, CD4⁺ and CD68⁺ cells used as

parameters. PCA was performed in R with the `prcomp` function with “scale = True”. PCA results were visualised using the `factoextra` package (Kassambara and Mundt., 2021) with functions “`fviz_pca_ind`” and “`fviz_pca_var`”.

5.2.6 Spatial analysis of necrotising granulomas

The tCPI, immCPI and lesion neighbourhood measurements were calculated as described in chapter 4. The necrotic distance index was determined by calculating the distance from all cells in each necrotising lesion from the necrotic interface. The borders of necrosis within necrotising lesions were drawn with the annotation tool based on an absence of staining in the DAPI channel. The distance of all cells from the necrotic interface within each necrotising lesion was then measured using the infiltration analysis tool. The necrotic distance index was then calculated as follows

$$\text{Necrotic Distance Index} = \frac{\text{Mean Distance}^{\text{Cell of interest}}}{\text{Mean Distance}^{\text{All cells}}}$$

5.2.7 Macrophage / monocyte distribution withing necrotising lesions

The measurements of macrophage distances from the necrotic interface were calculated using the infiltration analysis tool in HALO. All necrotic border annotations from each patient were stored in the same annotation layer. This layer was then chosen as the interface layer in the infiltration analysis tool and all macrophages within the patient were selected for analysis. The “store object data” option was set to “false” as only average data for each patient was required.

5.2.8 Lesion trajectory analysis

Data from immune cell densities, tCPI and immCPI were collated into a matrix and functions from the SCORPIUS package (Cannoodt et al., 2016) were used to produce the trajectory plot and trajectory heatmap.

5.3 Results

5.3.1 Active pulmonary TB is associated with extensive tissue consolidation and development of immune cell aggregates in the lungs

To gain information on tissue-wide inflammatory response to M.tb infection in human lungs, I first examined H&E-stained tissue sections obtained from ten patients who underwent partial surgical lung resection due to unsuccessful TB drug therapy. Major histopathological changes were evaluated by two independent clinical pathologists. The analysis revealed that, in contrast to healthy lungs with clear alveolar space, the landscape of diseased tissues was highly complex with regions of consolidation and foci of leukocytes (Fig. 5.1A). The well-recognised necrotising granulomas were frequently accompanied by many non-necrotising granulomas and lymphoid aggregates (Fig. 5.1B). The superstructure was consistently observed across all but one patient sample (n=10) (Fig. 5.1A and 5.1C), suggesting that the multifaceted histopathological landscape is common in the lungs of active TB patients.

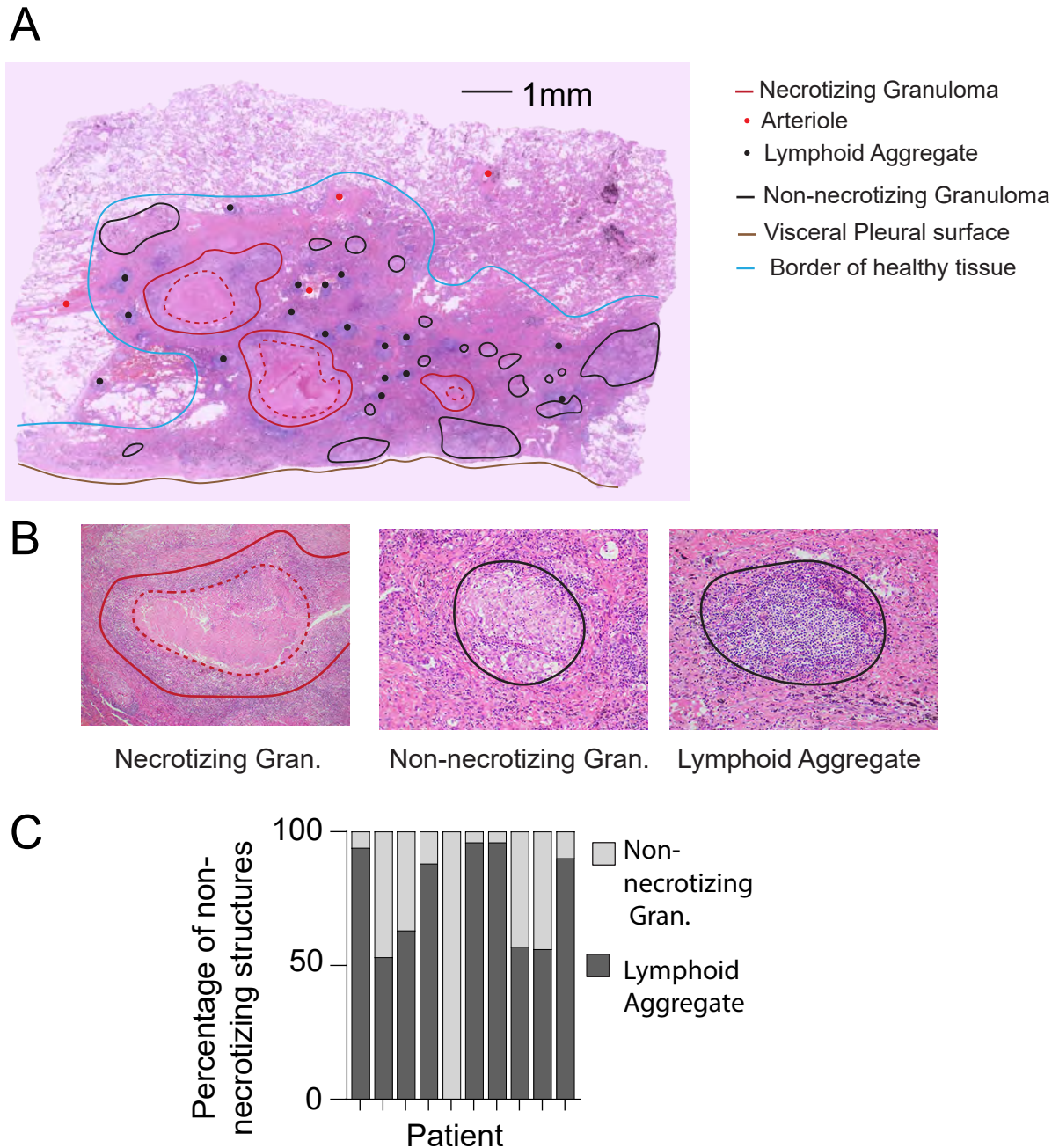


Figure 5.1. Active pulmonary TB is characterised by development of histopathological superstructures in the lung. (A) A TB lung tissue section stained with H&E with various histological and histopathological features indicated. **(B)** Enlarged example images of distinct lesion types identified in (A). The outer and inner borders of necrotising lesions were marked with solid and dotted red lines, respectively. The borders of non-necrotising lesions were marked with solid black lines. **(C)** Proportion of non-necrotising granulomas and lymphoid aggregates across 10 patients. Each colour represents one type of non-necrotising structure.

To capture the immune landscape across entire lung sections, tissues were stained with the major immune lineage markers CD68 (Macrophages), CD20 (B cells), CD8 (CD8⁺ T cells) and CD4 (CD4⁺ T cells), together with nuclear dye DAPI (Fig. 5.2A). High-resolution 20X regional images acquired using the Akoya Vectra system were stitched digitally to reconstruct the images of the entire plane of the tissue section, allowing for visualisation of each tissue section in its entirety and at a single cell resolution. I observed extensive infiltration of immune cells across all tissue sections, which was mirrored in the HE-stained serial sections from the same patient.

To assess whether the diverse structures defined pathologically could be captured and analysed at single-cell level, I first examined nuclear dye DAPI expression on TB lung tissue sections. To objectively distinguish healthy lung tissue from consolidated areas and cellular aggregates in the IF images, healthy and inflamed regions were first stratified based on the relative density of nucleated cells on the tissue-wide Opal IHC image following cell segmentation with DAPI using HALO software (Indica Labs) (Fig. 5.2B). Kernel density estimation was employed to quantify the cell density. The variations in the density across TB lung tissue were visualised with 3D projection (Fig. 5.2C), where height of the peaks (y-axis) represents cell density. Each distinct peak represents a cell aggregate, and a trough (coloured in white) reports an acellular necrotising lesion. Tissue was subsequently divided into regions based on cell density relative to the densest area of each patient sample. Cellular aggregates were identified as areas with density between 50-100% of the maximum density, consolidated tissues between 25-50% and healthy lung tissue as below 25%. When mapped back onto the tissue, the localisation of the cell density-defined, healthy tissue (HT) consolidated tissue (CT) and cellular aggregates (CA) were matched closely to the regions identified on HE-stained sections (Fig. 5.2D).

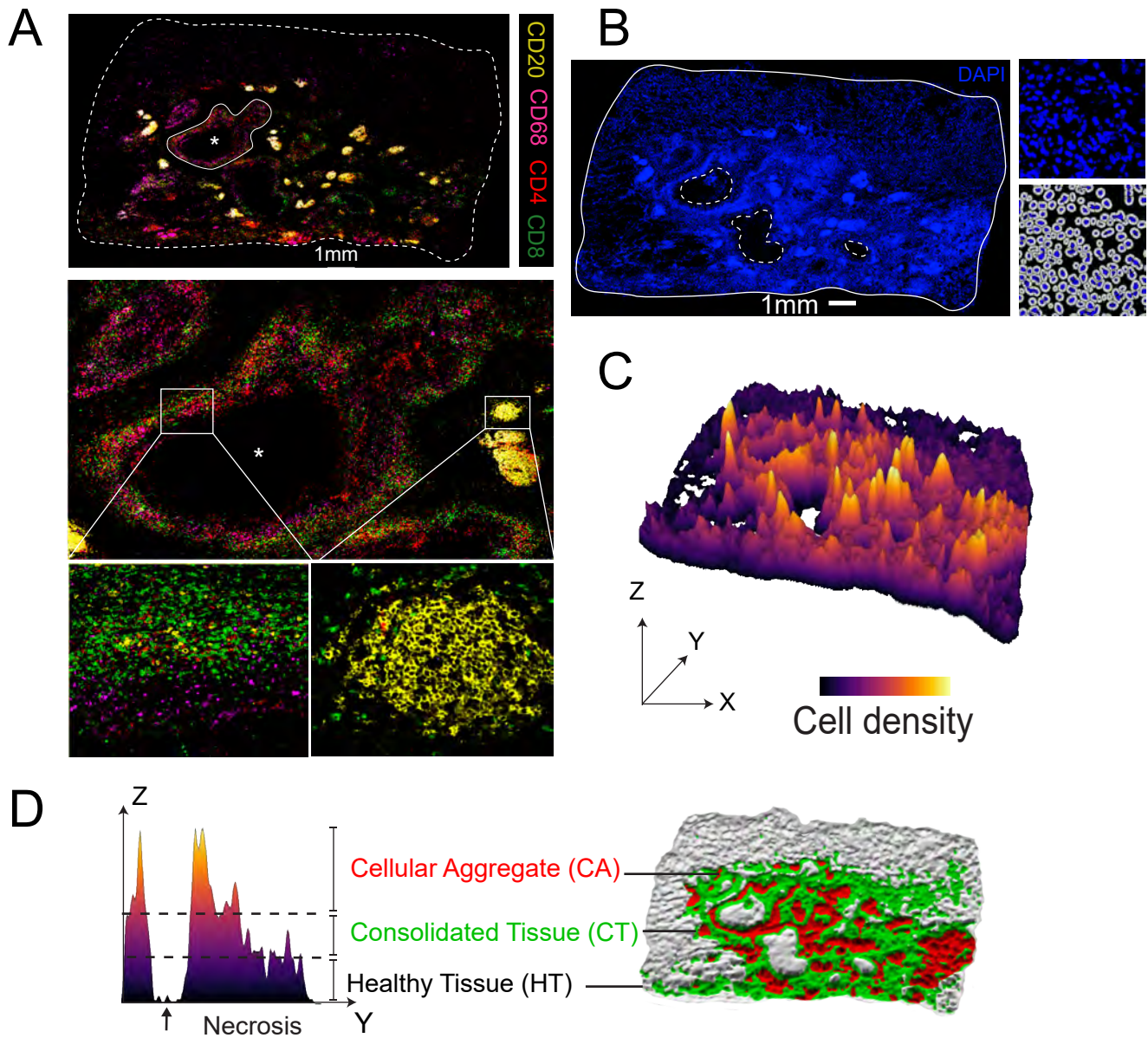


Figure 5.2. Tissue regions can be stratified by cell density. (A) Composite Opal immunofluorescence image of the same tissue section as in Fig. 5.1A stained for major immune cell markers inlaid with magnified views of a necrotising lesion (middle panel, circled by a white line, * necrotising core) and two non-necrotising regions (lower panel, rectangle boxes). (B) The same TB lung section as (A) stained with DAPI with a representative zoomed image and view of segmented cells shown. (C) 3D diagram showing the variations in the density of DAPI-positive segmented cells across entire TB lung section from (A). The height of the peak (Z-axis) indicates cell density. (D) A representative 2D histogram showing a cross-section along Y- and Z-axis of (C). The three Z-axis segments (divided by dotted lines, left panel) are defined based on cell density and correspond to the regions of healthy, consolidated tissue and cellular aggregates shown in the pseudo colour topography image (right panel) of the same TB lung tissue as in (A).

Next, I compared the composition of major immune cell populations in the consolidated region and cellular aggregates and found that, as a proportion of total DAPI⁺ cells, CD68⁺, CD8⁺, CD20⁺ and CD4⁺ cells were all significantly more abundant in the cellular aggregates compared to the consolidated region (Fig. 5.3A). These observations indicate that these immune cell-enriched cellular aggregates, defined by my unbiased cell density-based method, are distinct lesions within the consolidated region and represent the “granulomas” and lymphoid aggregates defined conventionally with histopathological features. To capture the total immune cell aggregates more precisely, I refined the lesion identification strategy by performing kernel density analysis and lesion definition only on cells stained positively for immune markers (Fig. 5.3B). In the remainder of this investigation, I have employed this unbiased analytic method to define individual immune cell aggregates in human tissues. I have focused my analysis on cellular aggregates containing 20-100% of the maximum immune cell density per patient, which included 672 lesions from ten patients including 22 structures containing a central necrotising cavity. Together, this unbiased lesion identification approach has revealed previously unappreciated complexity in the cellular landscape of human TB granulomas.

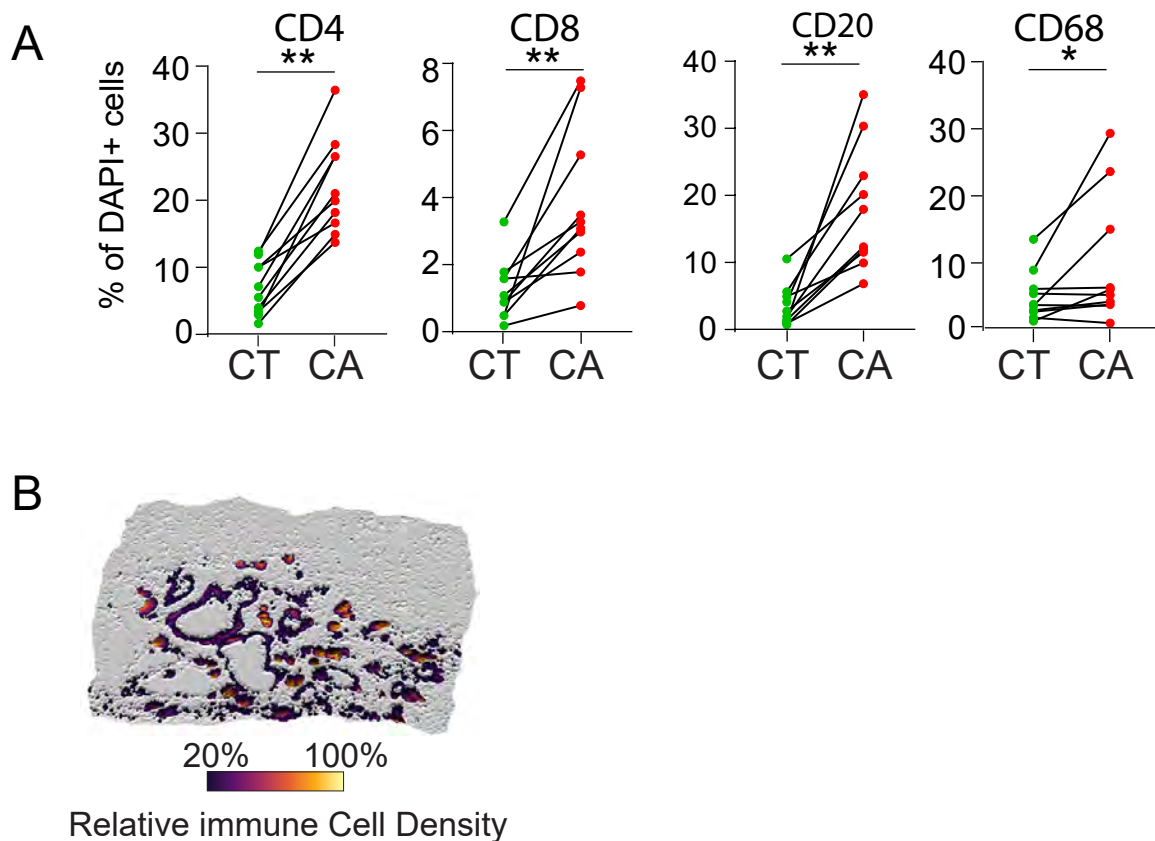


Figure 5.3. Lesion delineation by immune cell density. (A) Percentage of major immune cell populations among total DAPI positive cells in consolidated tissue (CT, green) versus cellular aggregates region (CA, red). Each paired point represents a patient. The level of statistical significance is determined using a Wilcoxon signed-rank test and $P < 0.05$ was considered to indicate statistical significance. * $P < 0.05$. ** $P < 0.01$. **(B)** Distribution of immune cell-containing cell aggregates (coloured) across the same tissue section as in Fig. 5.1A. The lesions selected are high in cellularity and positive for at least one of four major immune cell lineage markers (CD4, CD8, CD20 or CD68). Colour intensity of non-greyed areas correlates inversely with immune cell density of the region relative to the maximum density in the tissue.

5.3.2 Cellular aggregates are highly diverse and fall across a continuous spectrum underpinned by the density of CD20⁺ cells

Opal IHC images have revealed that the immune cell compositions in TB lesions are highly diverse (Fig. 5.4A). To understand the lesion heterogeneity, I analysed the density of each immune cell type across 672 lesions. For necrotising lesions, the central area of necrosis was excluded from the density calculation to allow meaningful comparison with non-necrotising lesions. I observed that the density of individual immune cell population varied significantly across lesions (up to 3 logs) and interestingly the variation was more apparent in non-necrotising lesions than their necrotising counterparts (Fig. 5.4B). CD20⁺ B cells and CD4⁺ T cells were more abundant in non-necrotising lesions while CD68⁺ macrophages were more abundant in necrotising lesions. CD8⁺ T cells distributed indistinguishably between necrotising and non-necrotising lesions.

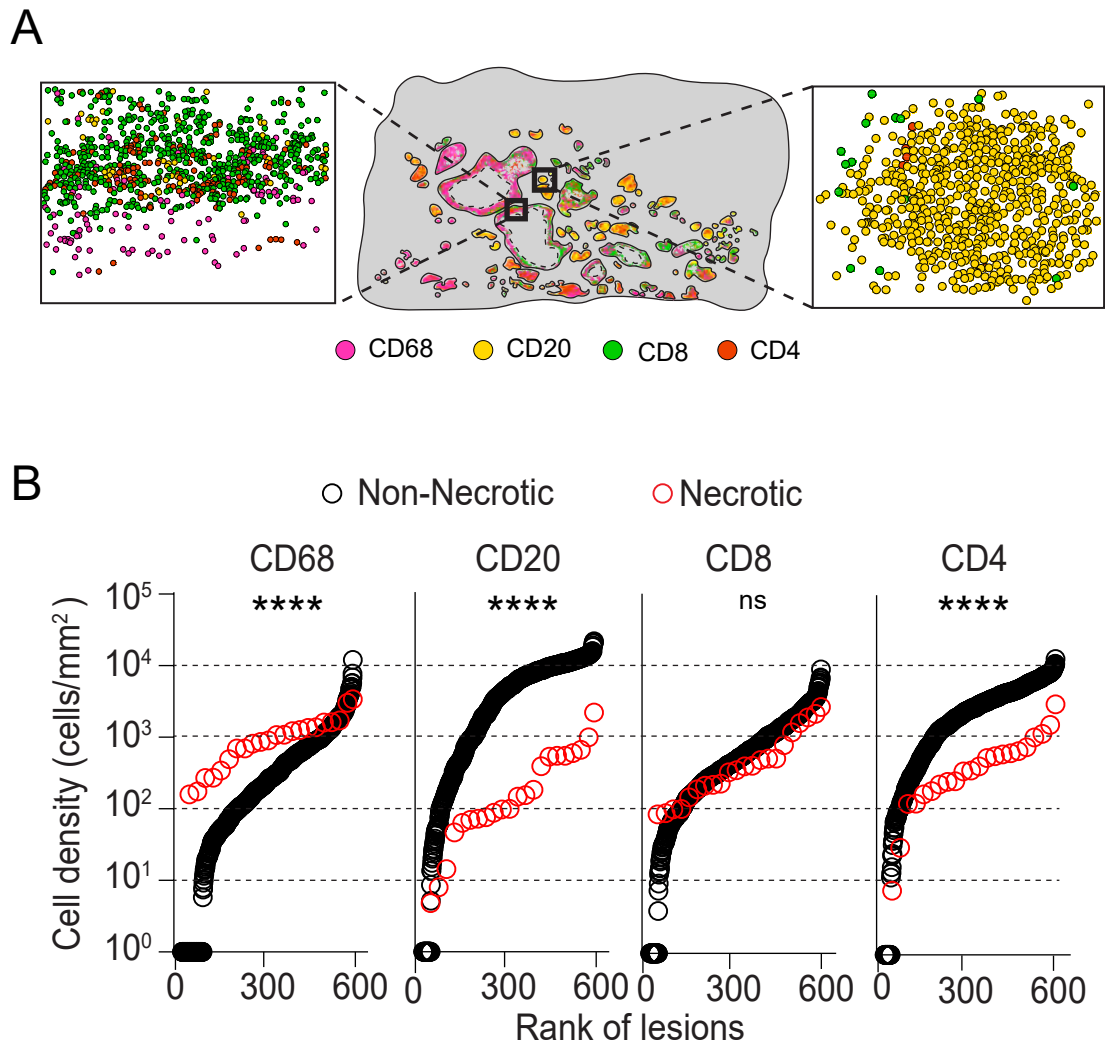


Figure 5.4. Necrotising and non-necrotising lesions have diverse immune cell compositions. (A) Image of segmented immune cells in positively identified lesions in TB lung from Fig. 5.1A with two representative areas (boxed) showing distinct cellular composition. (B) Necrotising and non-necrotising lesions from 10 patients ranked according to the density of each major immune cell population. Each symbol represents an individual lesion. The level of statistical significance between necrotising and non-necrotising lesions is determined using a Mann-Whitney test and $P < 0.05$ was considered to indicate statistical significance. **** $P < 0.0001$. ns, Not Significant.

To assess whether the lesions cluster into distinct groups based on their immune cell composition, I performed PCA on 672 lesions from ten patients based on the density of four major immune cell populations. Interestingly, I did not observe clustering of lesions into distinct groups, although necrotising lesions (coloured in red) tend to cluster closely, indicating a continuous spectrum of lesion compositions (Fig. 5.5A). Principal component 1 (PC1) represented the largest share of variation (37.2%) and was primarily driven by the variation in the density of B cells, indicating the continuous spectrum of the lesions is underlined mainly by the variations in CD20⁺ cell density. Moreover, these lesions did not cluster according to patient with the exception of one individual patient with increased density of CD8⁺ T cell (Fig. 5.5B), suggesting a minimal patient variation in lesion types defined using 4 immune cell markers.

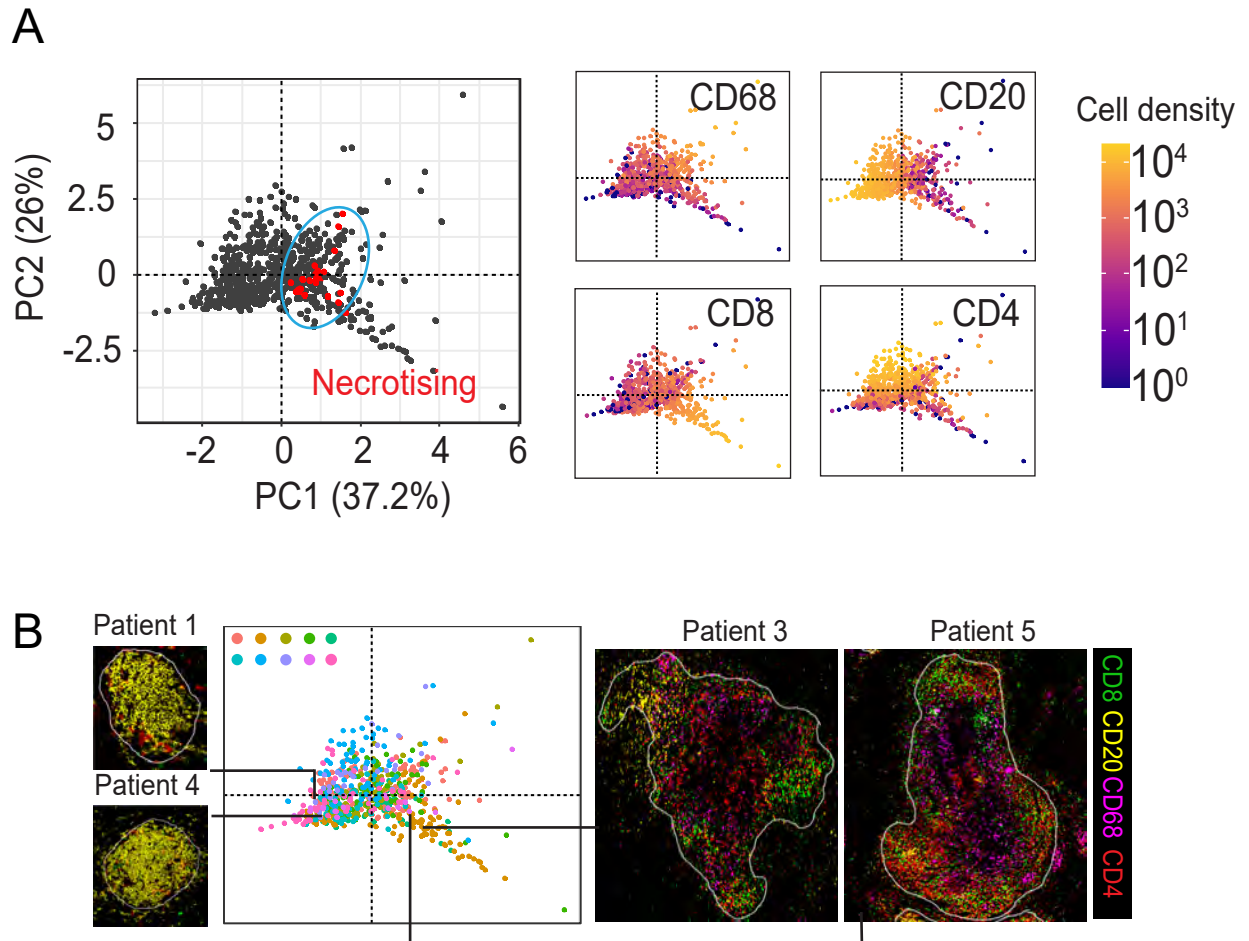


Figure 5.5. Lesions do not form distinct clusters but fall across a continuous spectrum underpinned by CD20⁺ cell density. (A) PCA plot of all lesions (n=657) with necrotising lesions indicated in red and circled by an ellipse. Matching marker specific PCA plots are coloured according to density of marker positive cells. **(B)** PCA plot visualisation of all lesions across 10 patients. Each colour indicates one patient, and each symbol represents an individual lesion. Images of example lesions of four individual patients are shown.

5.3.3 Lesions can be stratified based on immune cell composition and spatial organisation

Having established a spectrum of lesion compositions based on immune cell density, I sought to determine whether lesions have diverse spatial arrangement of cells. As expected, necrotising lesions were characterised by their acellular central core (Fig. 5.6A). However, within non-necrotising solid lesions, cells were also found to distribute unevenly, in some lesions the majority of the cells preferentially localised in the central region whilst in others most cells localised peripherally (Fig. 5.6A) leaving a loose but non-necrotising central area. To analyse this spatial variation, the tCPI for all lesions was measured as described in chapter 4. As expected, when the tCPI was compared between necrotising and non-necrotising lesions, the former lesions had significantly lower tCPI than non-necrotising lesions (Fig. 5.6B), validating the approach in quantifying cell spatial distribution tendency in mycobacterial granulomas.

I next investigated whether lesions can be stratified based on both the density of major immune cell populations as well as the overall distribution of cells within each lesion. I performed pairwise comparison of PC1 (driven mainly by B cells) from the previous PCA (Fig. 5.5A) and the tCPI measurement. In addition to necrotising lesions readily identifiable on H&E-stained sections, I used this analysis to define four types of solid lesions (Fig. 5.6C). Interestingly, all 4 lesion types were identified in each of the 10 patients analysed, although there were variations in the proportion of each lesion type across patients (Fig. 5.6D), suggesting that the lesion types identified here represent common immune features of the local tissue response in human TB lungs.

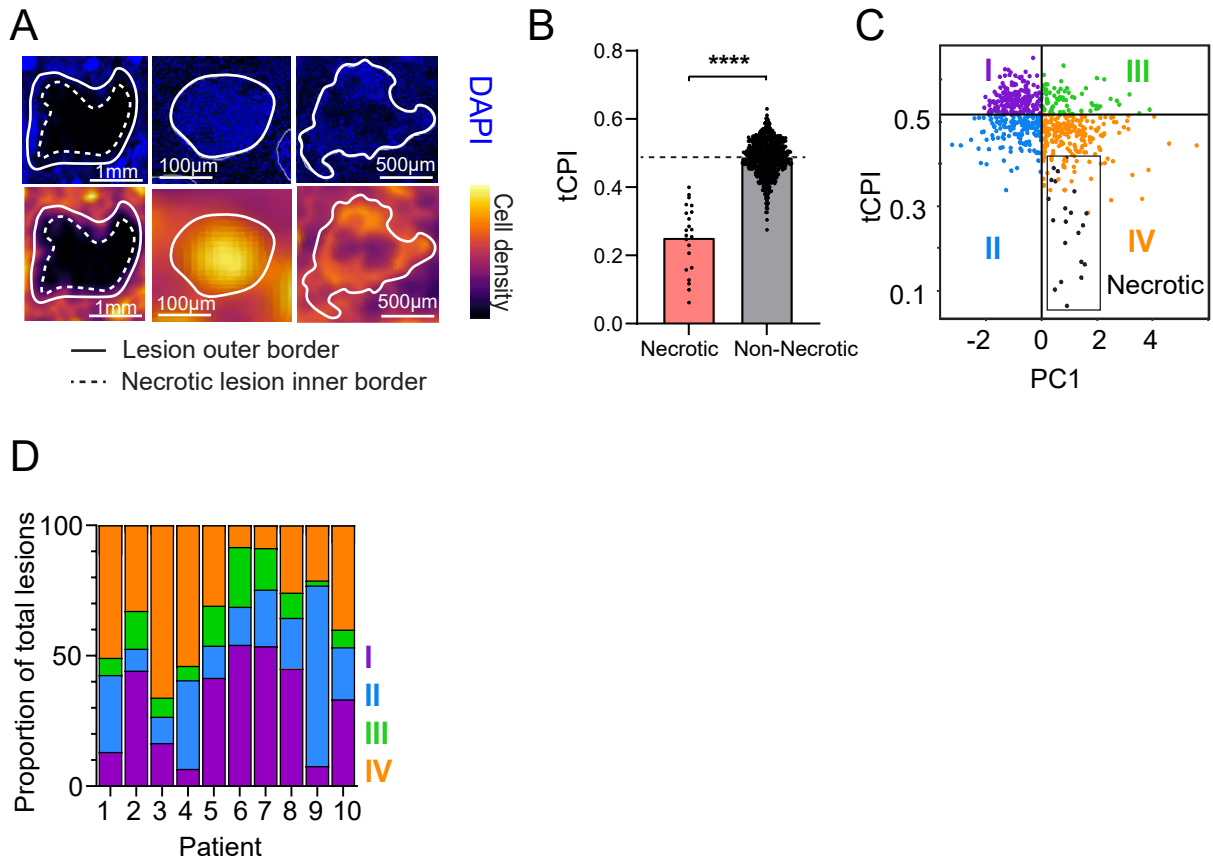


Figure 5.6. Combined analysis of compositional and spatial features identifies four types of TB lesions. (A) Images of DAPI-stained tissue showing variations in intra-lesion cell distribution (top panel). Heatmap images showing the same three lesions with relative cell density indicated (bottom panel). Repeated from Fig. 4.2A. (B) tCPI compared between necrotising and non-necrotising lesions. Each symbol represents an individual lesion. The height of each column corresponds to the group mean. Dashed line corresponds to a tCPI of 0.5. The level of statistical significance is determined using a Mann-Whitney test and $P < 0.05$ was considered to indicate statistical significance. **** $P < 0.0001$. (C) Pairwise comparison of PC1 from Fig. 5.5A against tCPI. Each symbol represents an individual lesion. Lesions ($n=657$) are divided into 4 sub-populations with the quadrants being set at $PC1 = 0$ and $tCPI = 0.5$. The 5th type of lesions, necrotising lesions (black symbols), are additionally indicated in the boxed area. (D) Proportion of each lesion type across 10 patients. Each colour represents one lesion type.

The cellular characteristics of the lesion type defined using LANDSCAPE analysis were highly consistent with those presented on multiplexed IF images of the same lesion type (Fig. 5.7A). Type I lesions frequently have a high density of B lymphocytes and cells tend to localise close to the lesion centre (high tCPI) (Fig. 5.7A and 5.7B). Type II lesions also have a high density of B cells but with reduced cell density in the lesion centre (low tCPI). Both type III and IV lesions had reduced presence of B cells and increased macrophages. However, they exhibited distinct cellular organisation. Type IV lesions displayed lower tCPI than type III lesions, suggesting that in these lesions most cells localised preferentially away from lesion centres. This description is reinforced by the observation that all necrotising / cavitory lesions fell into the type IV lesion category (Fig. 5.6C, lesions shown in the box area). Curiously, necrotising lesions also had a higher proportion of KI-67⁺ replicating CD4⁺ T cells compared to other non-necrotising lesion types (Fig. 5.7C).

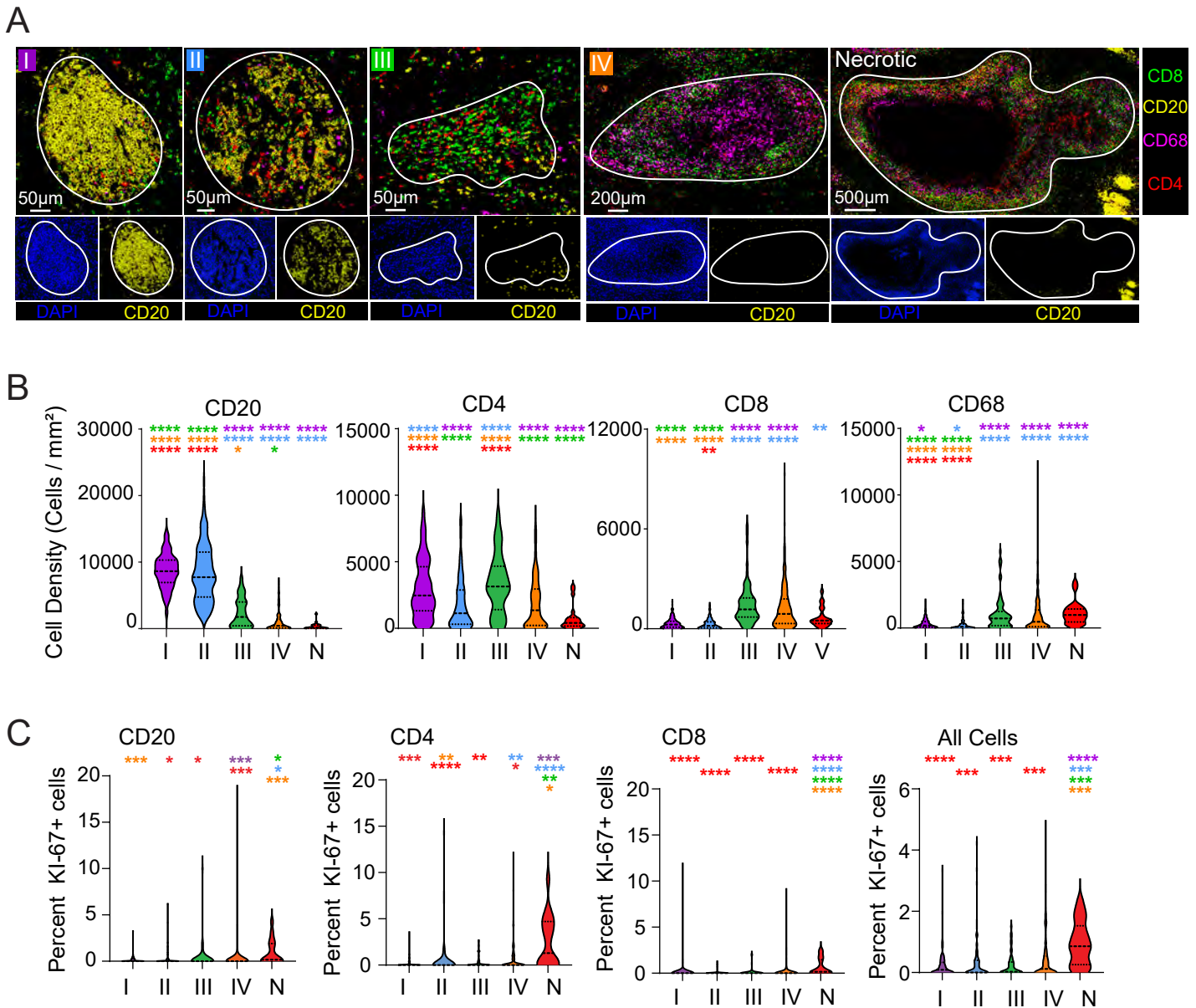


Figure 5.7. Distribution, composition, and activity of immune cells in each lesion type.

(A) Representative multiplex images of each lesion type with correlated images of DAPI and CD20 staining shown below each multiplex image. Solid white line indicates lesion border.

(B) Density of major immune populations according to lesion type. **(C)** Proportion of major cell populations positive for Ki-67 according to lesion type. The level of statistical significance in (B) and (C) is determined using a Kruskal-Wallis Test with Dunns Multiple Comparisons Test and an adj. $P < 0.05$ was considered to indicate statistical significance. * $P < 0.05$. ** $P < 0.01$. *** $P < 0.001$. **** $P < 0.0001$. Asterix colour above each group is used to indicate the matching group to which the statistical significance corresponds.

5.3.4 Divergent intra-lesion spatial distribution of CD20⁺ cells and CD68⁺ cells

determines the types of non-necrotising granulomas

To understand more specifically how the spatial distribution of individual immune cell populations regulates granuloma formation, I quantified the immCPI of each cell type in all lesions as described in chapter 4 (Fig. 5.8A). I found that CD20⁺ cells were located more peripherally in type IV lesions compared to types I, II and III, while CD68⁺ cells were located more centrally in type III and IV compared to type I (Fig. 5.8B). Together, these data highlight further cellular structural differences between lesion types and suggest a progression of CD20⁺ and CD68⁺ cell localisation between lesion types.

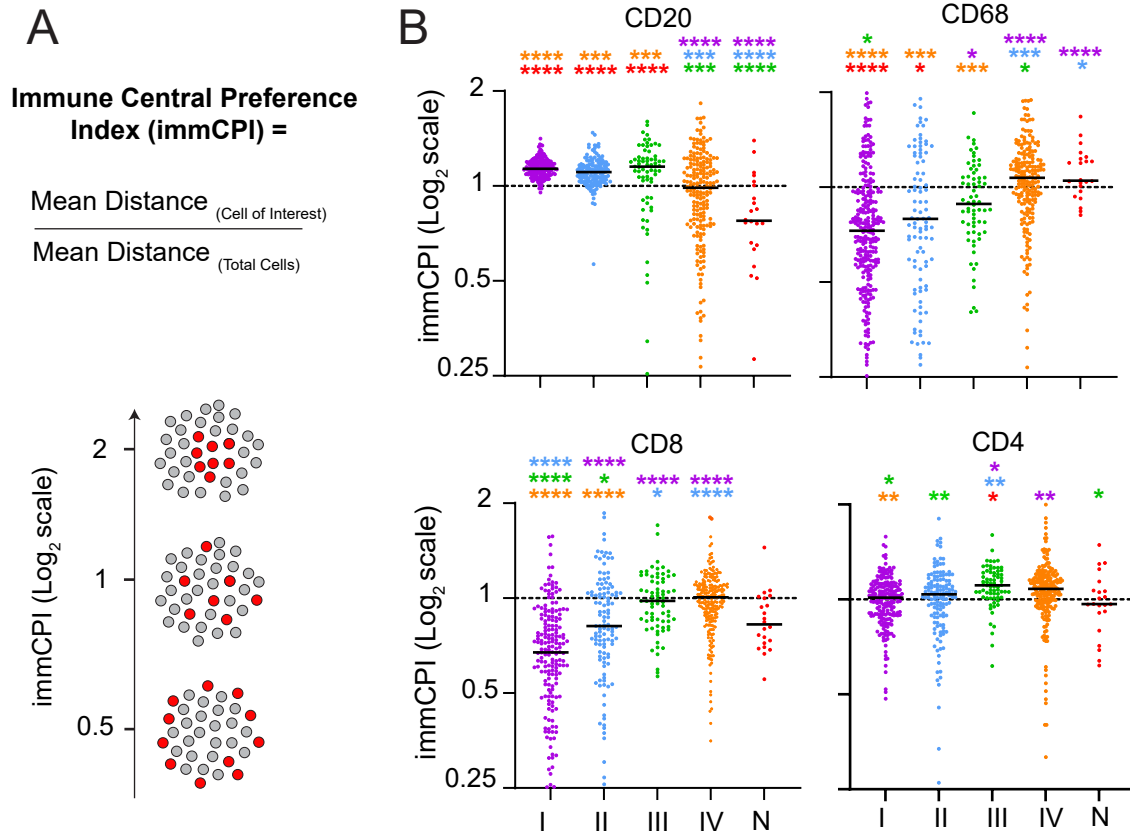


Figure 5.8. Divergent intra-lesion spatial distribution of CD20⁺ cells and CD68⁺ cells

determine the types of non-necrotising granulomas. (A) Formula for calculating the spatial distribution cell-of-interest within a lesion. Schematic diagram of lesions with representative immCPI measurements. Both repeated from Fig. 4.8D and E. **(B)** immCPI of each immune cell type across lesion type. Each symbol represents an individual lesion. Solid black lines indicate group mean. Dashed black line indicates an immCPI of 1. The level of statistical significance is determined using a Kruskal-Wallis Test with Dunns Multiple Comparisons Test and an adj. $P < 0.05$ was considered to indicate statistical significance. * $P < 0.05$. ** $P < 0.01$. *** $P < 0.001$. **** $P < 0.0001$. Asterix colour above each group is used to indicate the matching group to which the statistical significance corresponds.

Necrotising granulomas are formed when a host immune system fails to contain *M.tb* replication or regulate the local inflammatory response (Pagán and Ramakrishnan, 2018). The pathological manifestation is frequently associated with active disease, tissue injury, and transmission of *M.tb*. However, there is little knowledge as to how the infection niche is walled off and how immune cells are spatially organised at the border of necrotising (acellular) regions. I co-stained TB lung sections with antibodies specific to immune cells and mycobacteria to determine the localisation of infection foci in lung sections. I found that *M.tb* was only positively identified in a subset of tissue sections, almost exclusively in the cavity region (Fig. 5.9A and 5.9B), which is consistent with previous histopathological studies (Hunter et al., 2014, Medlar, 1948, Medlar, 1955). To map the cellular microenvironment of the necrotising border, I measured the distance of each immune cell in necrotising lesions from the necrotising interface. To compare the distance of immune cell populations between lesions, I developed a tool called the necrotising cell index. It functions similarly to immCPI but instead uses the mean distance of each cell population from the necrotising border divided by the mean distance of all cells in that lesion. This will return a value less than one if the cell population tends to be closer to the necrotising border and a value greater than one if the cell population tends to be further from the necrotising border. I found that CD68⁺ cells were localised closest to the necrotising core, compared to all other cell types. (Fig. 5.9C and 5.9D). Taken together, these findings reveal a further level of cellular organisation within necrotising lesions.

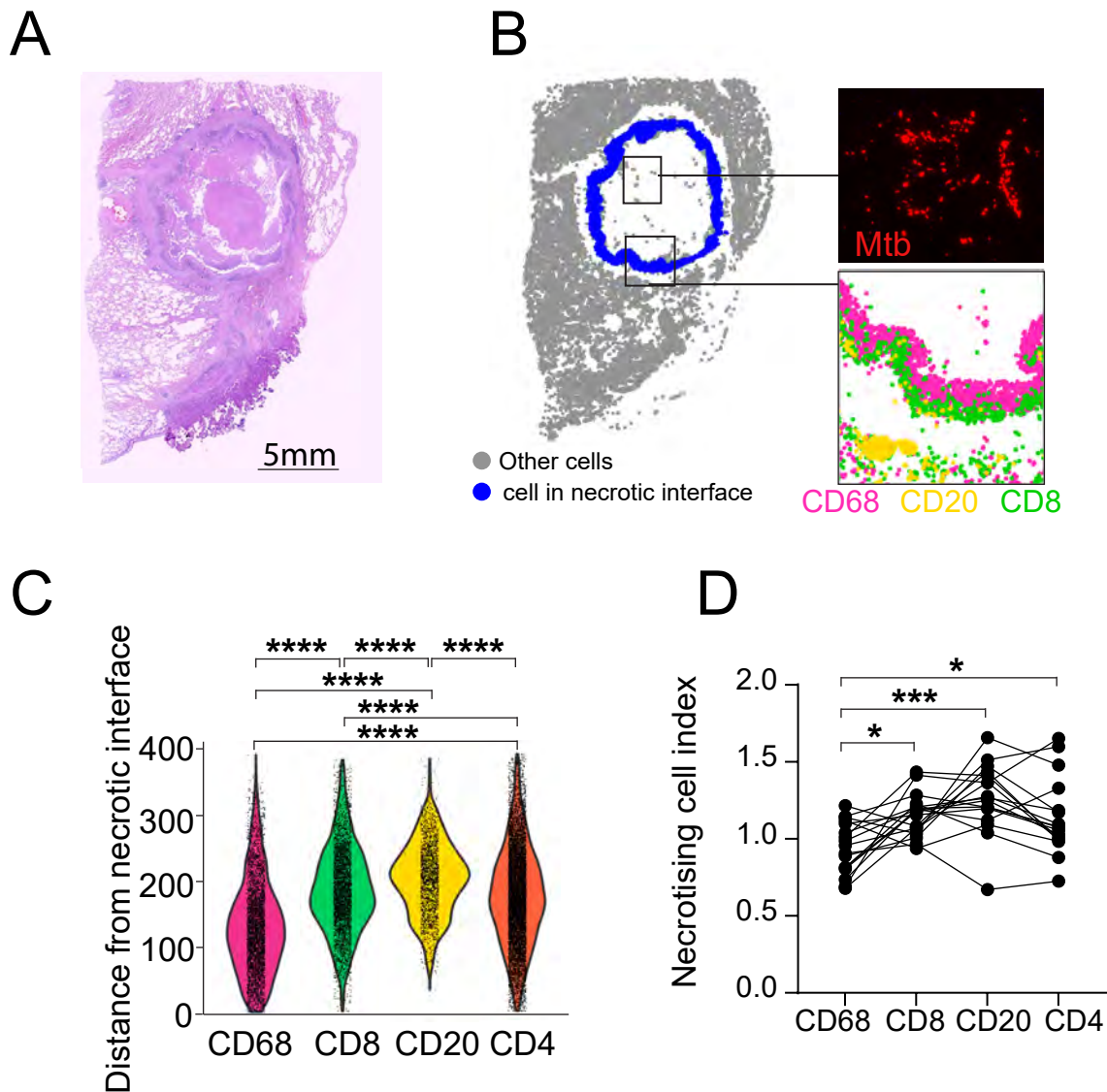


Figure 5.9. Spatial organisation of immune cells around the necrotic interface in necrotising lesions. **(A)** Image of a H&E-stained TB lung sample. **(B)** Pseudo colour image of the section from **(A)** with the necrotising border indicated in blue. The magnified images of the boxed areas show positive M.tb staining and spatial position of CD68⁺, CD20⁺ and CD8⁺ cells. **(C)** Distance of individual immune cell population from necrotising border in a representative lesion. Each symbol represents an individual cell. **(D)** Necrotising Cell Index of major immune populations. Each symbol represents an individual necrotising lesion and lines track data from the same lesion. The level of statistical significance in **(C)** and **(D)** is determined using a Kruskal-Wallis Test with Dunns Multiple Comparisons Test and an adj. $P < 0.05$ was considered to indicate statistical significance. * $P < 0.05$. *** $P < 0.001$. **** $P < 0.0001$.

5.3.5 CD14⁺CD68⁺ subset of monocyte / macrophage populations preferentially localises at the necrotising border

I next stained a subset of patient samples (n = 5) with an additional myeloid marker CD14 to better understand the populations of macrophages/monocytes residing in TB lung tissues (Fig. 5.10A and 5.10B). The relative abundance of the three subtypes, CD14⁺ C68⁻, CD14⁻ CD68⁺ and CD14⁺CD68⁺, was comparable across all four types of granulomas (Fig. 5.10C). However, while at the whole tissue level, the CD68⁺CD14⁺ macrophage/monocyte population was the least abundant subset relative to other two populations, they localised significantly closer to the necrotising core (Fig. 5.10D), revealing previously unrecognised tissue-wide spatial distribution of different myeloid cell populations from the necrotising interface.

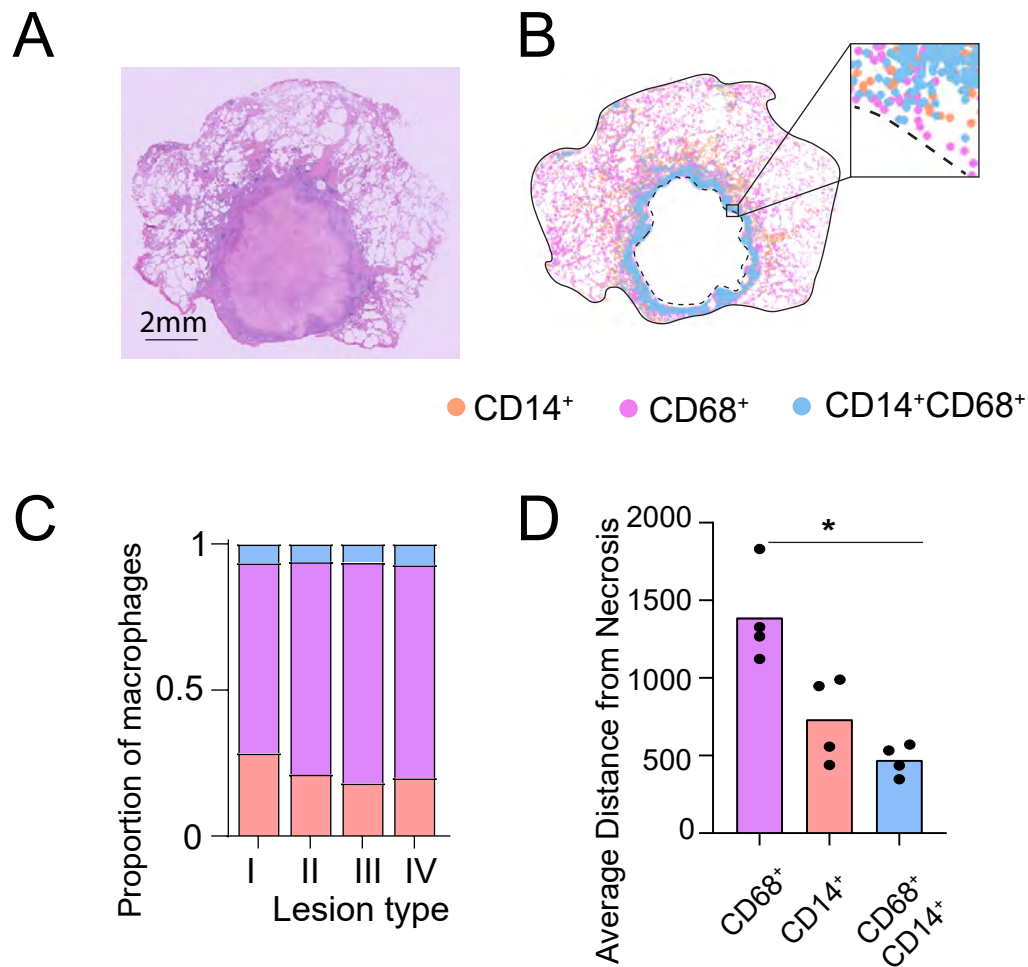


Figure 5.10. CD14⁺CD68⁺ subset of monocyte / macrophage populations preferentially localise at the necrotising border. (A) image of a H&E-stained TB lung sample. (B) Spatial distribution of monocyte / macrophage populations defined by CD14 and CD68 in tissue section from (A) with magnified image showing individual segmented myeloid cells. (C) Proportion of each macrophage subset by lesion type across 4 patients. (D) Distance of each macrophage subset from the closest necrotising area across entire tissue section. Each symbol represents an individual patient, and the height of each column corresponds to the group mean. The level of statistical significance is determined using a Kruskal-Wallis Test with Dunns Multiple Comparisons Test and an adj. $P < 0.05$ was considered to indicate statistical significance. * $P < 0.05$.

5.3.6 The distance to necrotising lesions correlates with the relative abundance of specific immune cell populations in the surrounding non-necrotising lesions

My tissue-wide imaging analysis revealed that necrotising lesions are surrounded by diverse solid lesions (Fig. 5.1A and Fig. 5.11A), which raises the questions of the spatial relationship between necrotising and various non-necrotising lesions. To determine the inter-lesion relationship, I performed lesion neighbourhood analysis as described in chapter 4 (Fig. 5.11B). Given the cellular organisation I observed around necrosis, I compared the distance of lesions of each type to the closest necrotising lesion. I found that lesions of type I and II tend to be twice as close to a necrotising lesion as lesions of type III and IV (Fig. 5.11C).

To further explore the role of non-necrotising lesion distance from necrotising lesions, I divided lesions of each type into those localised greater than 500 μm from necrosis and those less than 500 μm . I found that the lesions of type I, III and IV had more CD4^+ cells when they localised closer to necrotising lesions ($< 500 \mu\text{m}$) compared to those further away from necrosis ($> 500 \mu\text{m}$) (Fig. 5.11D). Type II lesions contain more CD20^+ cells when they were closer than 500 μm compared to when they were further. There was no difference in the density of CD8^+ or CD68^+ cells in any lesion type when comparing the distance of the lesions from the closest necrotising lesion. These results suggest that potential cross-talks between necrotising and certain types of non-necrotising lesions may contribute to the formation of multi-lesion superstructures in the lung.

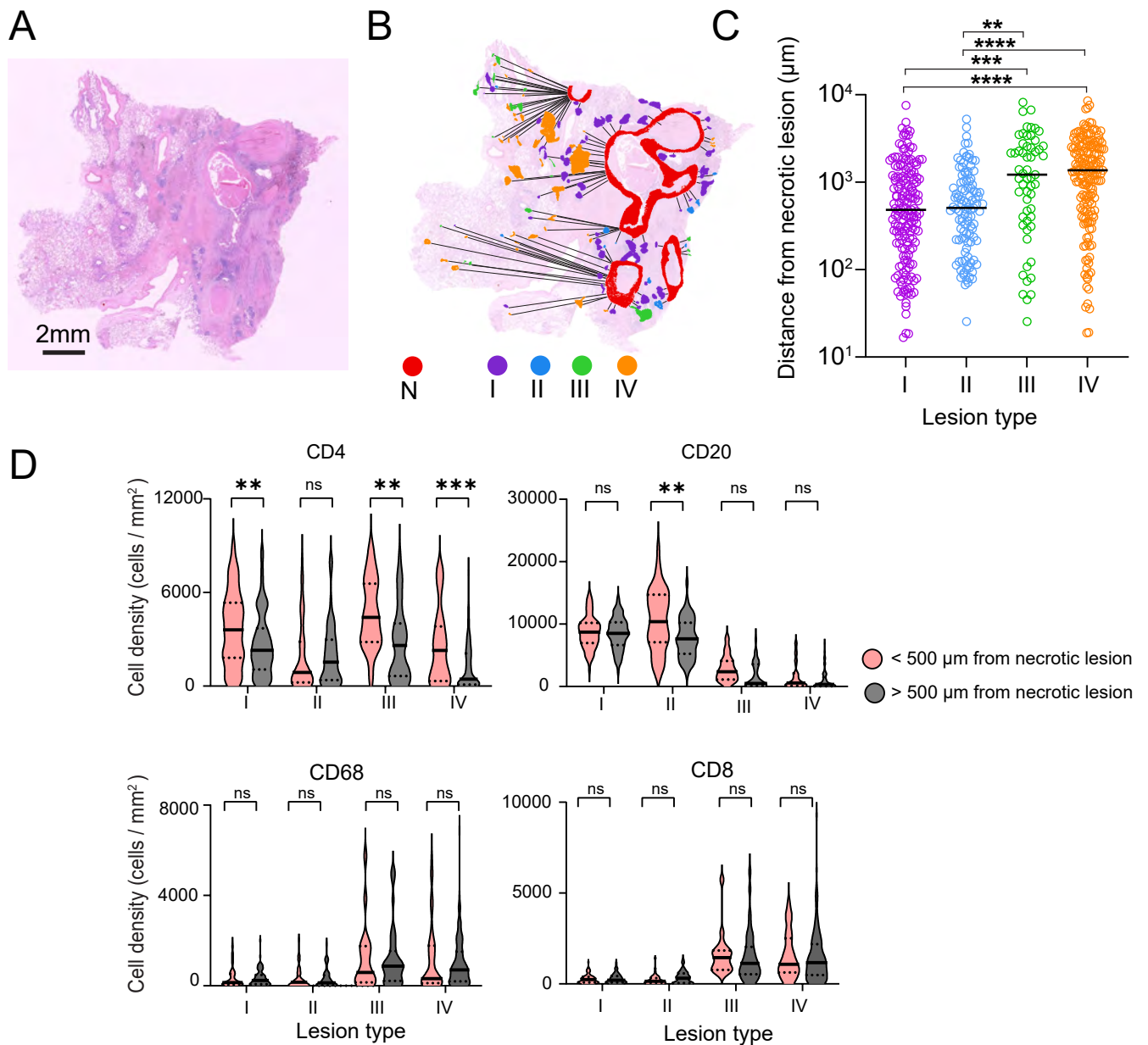


Figure 5.11. The distance of non-necrotising lesions from necrotising lesions is associated with lesion type. (A) Image of a H&E-stained TB lung sample. (B) Same TB lung section as in (A) overlaid with the locations of non-necrotising and necrotising lesions. The relationship of each non-necrotising lesion to closest necrotising lesion is indicated with connecting lines. Each colour represents one lesion type, with necrotising lesions indicated in red. (C) Distance of 4 types of non-necrotising lesions to the closest necrotising granuloma. Each symbol represents an individual non-necrotising lesion. Solid lines indicate group mean. The level of statistical significance is determined using a Kruskal-Wallis Test with Dunns Multiple Comparisons Test and an adj. $P < 0.05$ was considered to indicate statistical significance. ** $P < 0.01$. *** $P < 0.001$. **** $P < 0.0001$. (D) Density of $CD4^+$ and $CD20^+$ cells in lesions located greater or less than 500 μm from necrotising lesions. Solid lines indicate group median. The level of statistical significance is determined using a Mann-Whitney test and $P < 0.05$ was considered to indicate statistical significance. ** $P < 0.01$. *** $P < 0.001$. ns, Not Significant.

5.3.7 Formation of non-necrotising lesions is shaped by the proximity to necrotising granulomas and intra-lesion organisation of CD20⁺ and CD68⁺ cells

Since granulomas are highly dynamic and diverse, lesion-lesion interactions may also play a role in regulating lesion progression and the overall lung pathology. I combined both compositional and spatial measurements taken thus far and performed trajectory analysis on all lesions. I found that lesion progression correlates with the order of the four types of lesions identified and that necrotising lesions consistently fall at one end of spectrum (Fig. 5.12A and 5.12B). CD20⁺ cell density, as well as the central preference of total leukocytes, CD20⁺ and CD68⁺ cells all follow along the progression trajectory closely. Taken together, this study has shed light on both the cellular diversity and spatial organisation of human TB lesions and identified different lesion types that may represent a spectrum across lesion development.

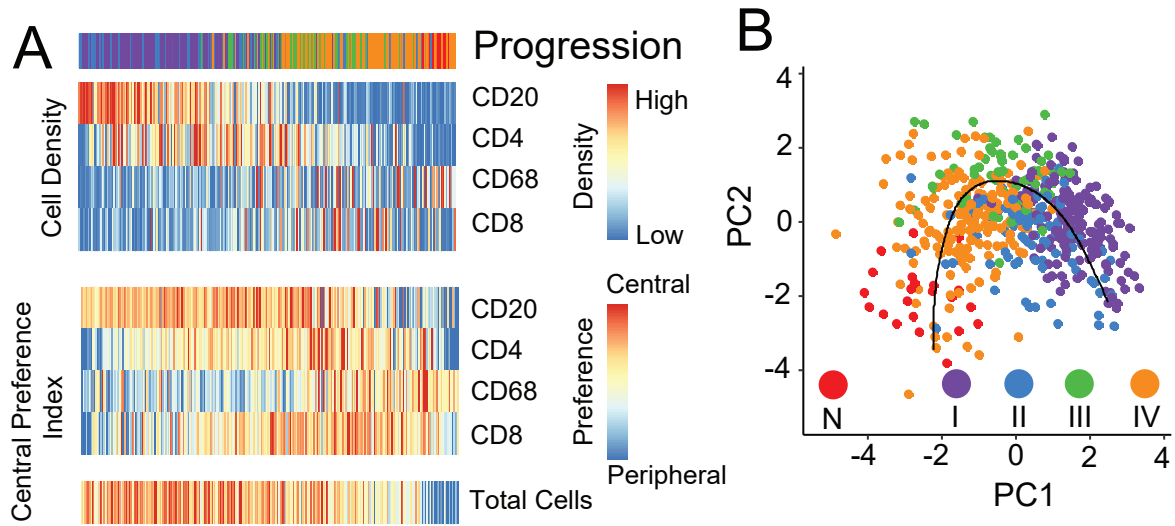


Figure 5.12. Trajectory analysis reveals lesions organise according to type. (A) Heatmap of inter-lesion trajectory analysis based on spatial and compositional parameters of individual lesions defined by LANDSCAPE. Lesions are ranked along the X axis according to trajectory and manually ordered on the Y axis. Top panel indicates lesion type. The analysis was performed using R. **(B)** Trajectory analysis of all lesions based on spatial and compositional lesion parameters as shown in (A). Line indicates inferred trajectory. Each symbol represents an individual lesion (n = 672). Each colour represents a lesion type, with necrotising lesions indicated in red.

5.4 Discussion

Recent work in M.tb-infected NHPs and human patients has revealed that TB lesions are heterogenous in their histopathological features (Subbian et al., 2015, Abengozar-Muela et al., 2020, Wells et al., 2021), metabolic activities (Kim et al., 2010, Mattila et al., 2017), inflammatory signalling (Mohlopheni et al., 2016, Cronan et al., 2021) and pathogen burdens (Gideon et al., 2022, Lin et al., 2014, Martin et al., 2017). Collectively, these previous studies have contributed to the notion that TB represents a spectrum of clinical and sub-clinical manifestations (Lin and Flynn, 2018, Cadena et al., 2017, Barry et al., 2009). By mapping immune landscape at a single cell level across entire lung tissue sections, I demonstrate here that even on the same lung section, human TB lesions, especially non-necrotising cellular aggregates, are highly diverse in their composition and cellular spatial organisation.

Furthermore, these non-necrotising lesions, together with necrotising granulomas, appear to form a histopathological superstructure underlying the immunopathology associated with uncontrolled M.tb infection. Since each lesion type represents a distinct type of immune response and lesions may individually or cooperatively contribute to TB pathogenesis, I suggest that both intra- and inter-lesion analysis should be performed to define host response to M.tb.

By analysing > 650 lesions with a total of 7 million single cells, I defined four types of non-necrotising lesions representing a spectrum of distinct yet overlapping cellular architecture. The frequently documented lesions with a necrotising core represent only one of the several types of lesions in the infected organs, suggesting that examination of only one or two stereotypical lesion types is unlikely to capture all the cellular events contributing to the overall lung pathology and clinical outcome. It is of particular interest that each type of the four different solid lesions was present in all the patients examined. I speculate that the lesion spectrum may reflect the history of ongoing local immune response to continuing seeding of

locally disseminated drug-resistant bacteria. Therefore, this heterogenous and dynamic nature captured in a single tissue section may be explored to identify distinct cellular structures associated with many if not all of the stages of the granuloma maturation process. A similar approach has been used to successfully define T cell population at different maturation stages in the thymus.

Together, I hypothesise that the lesions identified here may represent distinct yet related stages of granuloma maturation, from B cell-enriched lymphocytic foci to macrophage-centred granulomas, and some of the latter lesions may eventually develop into necrotising tissue whereas others become healed or fibrotic lesions. Similarly, type I lesions may contain both distinct tertiary lymphoid structures (TLS) and B cell-enriched infection foci (Fig. 5.13). Future high-dimension phenotyping of the four lesion types on samples with different clinical outcomes will assist in defining the branchpoints along the lesion progression.

I have revealed the extent of granuloma heterogeneity in patients with ongoing M.tb infection. Interestingly, PCA analysis showed that TB granulomas are organised as a spectrum rather than distinct clusters, suggesting overlaps in cellular structure and potentially interactions across lesions. The spatial association between certain types of non-necrotising and necrotising granulomas further suggests that these lesions may function cooperatively as an inflammatory superstructure to determine the outcome of local immune response. These observations raise a question how a TB granuloma should be defined at a single cell level. It is impossible to determine with 2-dimensional IF staining whether these structures were structurally connected as has been shown previously with μ CT imaging (Wells et al., 2021). Nevertheless, my results support the hypothesis that the “double-edged sword” function of granulomas may reflect the heterogeneity among granuloma populations (Lin and Flynn, 2018).

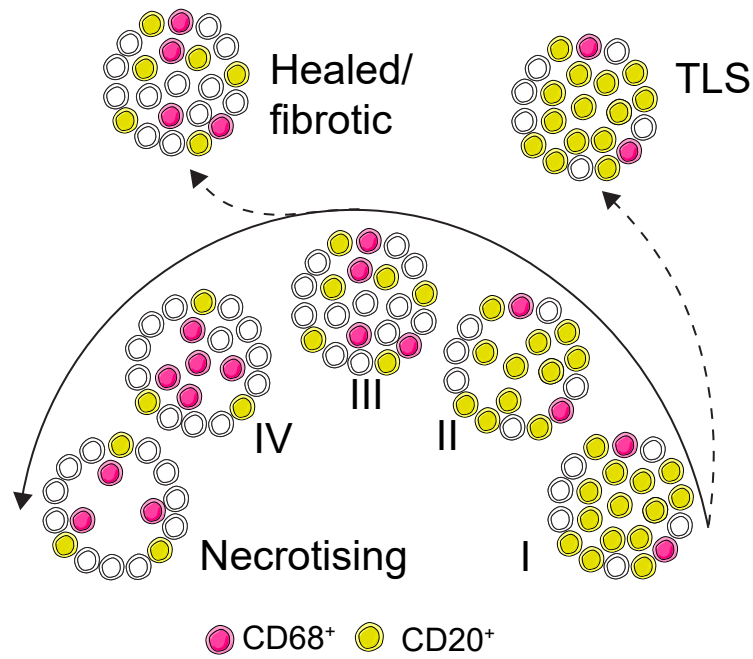


Figure 5.13. Schematic diagram summarising key findings and illustrating hypothetical lesion progression pathways. Dotted lines indicate hypothetical lesion progression branchpoints and pathways. TLS, tertiary lymphoid structure.

Advances in single-cell RNA sequencing have provided unprecedented insights into the composition and function of immune cells in M.tb-infected tissues (Cai et al., 2022, Pisu et al., 2021, Akter et al., 2022, Esaulova et al., 2021). However, this approach does not provide the spatial context of the local immune response (Rao et al., 2021), which is critical in understanding the architecture and function of mycobacterial granulomas. By employing MIBI-TOF, McCaffrey *et al.* have successfully documented high-dimension cellular information of human granulomas (McCaffrey et al., 2022). However, though time-of-flight mass cytometry can be used to analyse the regions of interest within a lesion, it does not reveal comprehensive tissue-level immune landscape or lesion-lesion relationship (Tan et al., 2020). On the other hand, my tissue-wide single-cell analysis is better suited for examining the tissue landscape across lesions at the cost of detailed phenotyping of immune populations. Therefore, the two studies are highly complementary. Comparing and contrasting the data generated in the two investigations may provide additional insights into human TB granuloma structure. For example, some of the B cell-enriched type I lesions I identified, which may resemble the TLS formed in mycobacterial infected tissues (Ulrichs et al., 2004, Cohen et al., 2022), have also been independently identified in the McCaffrey study (McCaffrey et al., 2022). Therefore, it is possible that Type I lesions represent heterogeneous lesions including conventional TLS and B cell-enriched mycobacterial granulomas. Future deep-phenotyping using high-dimensional image technology will be able to address this question.

My study has several limitations. Firstly, a limited number of lung tissue samples were examined. While a minimal patient variation was observed when four immune cell markers were used, the variation between patients may become more apparent when more markers are used to profile a broader variety of immune subsets. Secondly, Opal multiplex technology is only capable of examining small numbers of markers thereby limiting the depth of cellular

analysis. Combining tissue landscape and in-depth region of interest analysis will lead to a better understating of lesion microenvironment in the context of surrounding inflamed tissues. Finally, since my initial samples were collected from TB patients undergoing lobectomy due to a poor treatment response, whether my study has captured a fraction or full spectrum of granulomatous lesions remains to be determined. In future, I will apply my LANDSCAPE strategy to broaden my investigation by including samples from patients with active TB or latent infection in lungs resected for other reasons, which will assist in elucidating the relationship between cellular structure, anti-microbial function, and disease outcomes.

My study has revealed a previously unknown heterogeneity of human TB granulomas. It is unknown at this stage whether any particular type is more permissive or inhibitory of M.tb replication in humans, however the findings of this study will help guide future study into what a successful M.tb immune response looks like at the level of individual granulomas. Additionally, the deeper understanding of the cellular structure of TB granulomas ascertained in this study may provide insights for the development of future anti-M.tb drugs that must penetrate into the granuloma to reach the pathogen.

Chapter 6

General discussion

6.1 The current study

This study details the development and application of two pipelines for high throughput analysis of human immunology data. For transcriptomic data, a pipeline was developed in chapter 3 to characterise gene expression patterns using public data and was applied to identify neutrophil specific upregulation of *CST7* during inflammation in humans. For defining inflammatory responses in human tissue, a pipeline was developed in chapter 4 to perform spatial analyses of inflammatory foci with single cell resolution. This approach was employed in chapter 5 to define distinct lesion types as well as the broader histopathologic superstructure in human TB lung tissue.

6.2 The future of public transcriptomic data

The transcriptomic pipeline I developed is capable of broadly exploring the conditions under which a gene of interest is expressed using only publicly available data and software. While the use of public data in human immunology is expanding, allowing in some cases complete hypothesis testing in silico (Toufiq et al., 2020, Huang et al., 2021, Huang et al., 2022), transcriptomic data reuse remains largely restricted to those with bioinformatics training (Sparks et al., 2016) and as a result, most publicly available transcriptomic datasets continue to be under-utilised. The potential for meta-analyses, which can draw conclusions with greater power than single studies is also lost to many immunologists due to the knowledge barrier required to perform them. While the pipeline I have developed can apply meta-analyses to answer a variety of research questions in a manner both fast and free of cost, it too requires a working knowledge of R in its current form, which limits its accessibility and usefulness to non-bioinformatician scientists.

Any future application in which public data transcriptomic meta-analyses are routinely performed by the broader community of immunologists will require online databases with a simple user interface. Relevant transcriptomic datasets should be able to be quickly identified, accessed, and pooled in preparation for meta-analyses, which would ideally be run using preprogramed functions. Such a database should have patient metadata annotated to each sample to facilitate ease of patient grouping by covariate features such as age, gender etc. The combination of such a tool with a feature to rapidly screen PubMed or other literature databases for previous gene associations would increase its power to characterise previously unrecognised gene associations. The application of the pipeline I developed into such a future online repository would overcome its primary limitation of accessibility and would enable its usage as both a free and simple tool that has the potential to benefit all human immunologists.

6.3 Single cell spatial analysis at the lesion level

In this thesis, I developed a spatial metric termed tCPI that can stratify any cellular lesion with defined borders according to whether its constituent cells distribute more densely in the centre or periphery of the lesion. A peripheral distribution of cells is particularly apparent in lesions that contain a central area of necrosis, which is a type of uncontrolled cell death. Among a cohort of hundreds of TB granulomas, I used the tCPI spatial tool to identify many non-necrotising lesions that had a reduced central density without the histopathologic presence of necrotic tissue. These may represent non-necrotising granulomas enriched with macrophages, which have larger cytoplasm regions relative to lymphocytes. Alternatively, the spatial characteristics may reflect the reduction in cellularity associated with the process of forming a central necrotic region. In contrast, many non-necrotising lesions with dense

cellular core were enriched with lymphocytes. They may represent a local immune response distinct to that associated with macrophage-driven lesions, and one potentially more effective at bacterial containment, thereby mitigating uncontrolled cell death. These questions should be investigated in the future with high-dimension imaging technologies, such as CODEX or IMC, together with tissue samples with a well-characterised clinical outcome.

While tCPI is designed to stratify TB lesions based on their intra-lesion cellular distribution, it can potentially perform the same function to examine inflammatory foci from any disease and would be particularly applicable to the lesions with a necrotic central area. Apart from TB, necrotising lesions can also be observed in both infectious diseases (such as with non-tuberculosis mycobacteria (Dobos et al., 1999, Patel et al., 2014, Walsh et al., 2011), *cryptococcus spp.* (Gave et al., 2004, Ho et al., 2015) and *Coccidioides immitis* (Shah et al., 2017)) as well as non-infectious diseases (such as Granulomatosis with polyangiitis (Comarmond and Cacoub, 2014), granuloma annulare (Shah et al., 2017), and many cancers (Caruso et al., 2012, Richards et al., 2011)). In each case tCPI will be able to reveal the intra-lesion cell densities in both necrotising and non-necrotising lesions present in the tissue and may reveal a spectrum of cell distributions between the two lesion types as was observed in TB lung samples examined in this study.

The immCPI tool developed in this study is able to determine the intra-lesion distribution of each leukocyte population within a lesion and was used to determine central localisation preference of major immune populations according to TB lesion type. As with the tCPI measurement, the immCPI tool can also be directly applied to any tissue lesion with a defined border. The immCPI tool will prove useful in lesions that harbour pathogens by identifying the leukocyte populations in closest proximity to the pathogens. For example, the pathogen localises peripherally within the lesions of Buruli ulcer (James, 2000), while it

localises centrally in the case of tuberculoid leprosy (Adams and Krahenbuhl, 1996), and in each case, immCPI will be able to identify leukocytes in closest proximity.

In cancer research, the penetration of leukocytes into the intra-tumoral space is often associated with favourable patient outcomes (Pagès et al., 2010, Grabovska et al., 2020, Attrill et al., 2022a) and this is typically determined by measuring the density of each cell population within repeated bands inward from the tumour border. The band-based approach to infiltration analysis is useful to visualise the overall distribution of immune cell populations within a tumour with density histograms and for simple comparisons between tumours at distinct distances (for example 0-50µm penetration into the intra-tumoral space). However, the overall penetration of each cell-type cannot be easily quantified or compared between tumours using this approach and any comparisons between tumours may be confounded when the tumours are of different sizes.

When studying tumours histologically, the immCPI tool could supplement infiltration analysis as a measure to determine the depth of penetration of each leukocyte into the intra-tumoral space, providing a single measurement for each leukocyte population that can be used comparatively between tumours in different patient groups.

The third novel approach developed through this thesis work, lesion neighbourhood measurement tool, can quantify the distances between all lesions in a tissue so as to examine the histopathological landscape of the clinical sample more broadly, covering the entire tissue section. This can again be applied in any disease that forms multiple lesions within a reasonably close tissue area and can only be practically applied when using whole slide imaging. It will be of greatest benefit when multiple phenotypically distinct lesion types are present in the same tissue section, for example necrotising and non-necrotising lesions, so as to compare their relative distance to one another between patients.

6.4 Validation of TB animal models

I identified four types of lesions present in the human TB lung, characterised by both cellular composition and intra-lesion cell distribution. While my study was able to explore TB disease directly in human tissue, the majority of TB research, especially studies focusing on defining cellular and molecular mechanisms underlying pathogenesis, are performed in animal models. Animal models allow control of the time of infection, inoculant dosage and bacterial strain, as well as the exact tissue type to be analysed.

However, our understanding of how accurately the granulomas in animal models reflect those of human TB remains poorly understood, largely due to a lack of understanding of the cellular structure of the human TB granuloma and the cellular landscape in the TB lung more broadly. We can be confident the compact organised granulomas seen in the lung of the C3H/FeJ mouse strain and NHPs are a closer recapitulation of human TB pathology than the more diffused pneumonia-like pathology seen in the majority of inbred laboratory mouse strains, including C57BL6 and BALB/c mice (Lanoix et al., 2015). However, no prior study has attempted to extensively characterise the diversity and cellular organisation of human TB granulomas at the cellular level and thus the ability to validate the granulomas generated by animal models, on which so much research relies, remains limited.

The spatial analysis tools developed in chapter 4 were used to identify four subtypes of non-necrotising lesions within the human TB lung, based on their cellular composition and distribution. This expands our understanding of the cellular diversity of the human TB granuloma and offers guidance to ascertain whether animal models of the TB granuloma are able to recapitulate lesions with the same extent of diversity as human TB. The tCPI and immCPI tools developed in this study can each be readily applied to lesions from animal tissue and thus these tools can play a direct role in characterising lesions in each animal

model and determining whether the diversity of lesions in the human TB lung is reflected. Individual lesions in animal models can be delineated based on density of leukocytes in a similar approach that was employed in chapter 5 of this study to allow for valid comparison with human lesions.

One of the interesting observations found in this study is the formation of a multi-lesion superstructure, in which non-necrotising lesions are organised by lesion type relative to large necrotising lesions in the TB lung. This represents a novel histological feature of human TB pathology that can be used to validate animal models. The cynomolgus macaque model of TB can contain multiple granulomas in the lung, including both small solid lesions and large lesions with a central necrotic core (Grant et al., 2022). Some granulomas present in the weeks following initial infection were seeded at the point of infection while others represent dissemination from other granulomas (Martin et al., 2017), and it would be of great interest to identify whether lesion dissemination results in a kind of histological superstructure as was identified in this study in human TB.

Trajectory analysis is a tool that has been used frequently in scRNAseq studies (Van den Berge et al., 2020) with the assumption that the total cell population contains cells across some form of developmental trajectory and attempts to arrange cells according to that progression, for example the development of a particular cell-type from its immature to mature form. To the best of my knowledge my study represents the first time that trajectory analysis has been applied at the level of entire lesions rather than single cells. Trajectory analysis at the lesion level may also prove invaluable in the development of animal models of TB, given that more information about the outcome of each lesion is available. For example, the bacterial burden of each lesion can be reliably measured, whereas in human studies such as the current study, the patients have already undergone multidrug therapy and any attempts to determine bacterial load are likely to be inaccurate.

6.5 The spectrum of human TB granulomas

In this study I developed a method to delineate individual lesions based on density of immune cells. Much of our understanding of the human TB granuloma comes from *ex vivo* examination of lung tissue, usually removed as part of a lung resection as an end stage treatment for patients with drug resistant disease. As these samples are from patients with severe disease, there is usually extensive granuloma formation and tissue remodelling that can complicate analysis and prevent simple delineation of one lesions from another. As a result, the TB lung is typically analysed in imaging studies using only selected 20x FOVs from areas with granulomatous inflammation. This means that any lesions larger than a 20x FOV cannot be analysed in their entirety. The density-based lesion delineation approach I have developed will allow all lesions within a tissue sample to be analysed in their entirety.

I identified a spectrum of cellular compositions and spatial organisations across the landscape of lesions in the human TB lung tissues. While my study cannot ascertain how effectively each lesion is able to control *M.tb*, it establishes that large necrotising lesions represent both a breakdown in host tolerance and, based on mycobacterial staining, in some cases a site of persistent bacterial presence. Evidence has emerged from the NHP model of TB that each granuloma represents an individual battleground and has highlighted that while some granulomas resolve during the course of infection, others remain sites of persistent infection and go on to disseminate *M.tb*, producing new granulomas in the process (Lin et al., 2013, Coleman et al., 2014, Lin et al., 2014, Martin et al., 2017).

If the NHP TB model does indeed reflect the diversity of lesion cellular compositions and cellular distributions identified in human TB, then it can be further explored whether any of these lesion phenotypes correlate with enhanced bacterial burden or lesion sterilisation. If for example NHP lesions resembling the type IV lesions identified in this study correlate with

increased bacterial replication and poorer outcomes, it can be hypothesised that this lesion type may also be associated with increased bacterial replication.

The multiplex imaging employed in chapter 5 used only four primary markers to identify corresponding major immune populations and to perform all the lesion stratification that followed. Four markers is a relatively low number compared to what is possible with the most modern multiplex imaging technology. However, the fact that lesions were able to be stratified in several ways based on such a low number of markers highlights the power of LANDSCAPE analysis tools when working with whole slide images containing hundreds of thousands of cells that in turn comprise a high number of distinct lesions. Further, the tools developed as part of LANDSCAPE analysis are not restricted to the OPAL multiplex system and can be applied to images from any multiplex imaging technology so long as a nuclear stain is incorporated in order to identify the spatial coordinates of each individual cell. Future use of LANDSCAPE in a highly multiplex system that can image dozens of parameters such as CODEX will require some adjustments to data visualisation, given the multi-dimensional nature of the data that can be generated. A newer microscope suite called the PhenoCycler Fusion System (Akoya) has recently been released that updates the CODEX technology to enable rapid visualisation, within minutes, of up to a million cells. This dramatically improves on the runtime of CODEX (which can take many hours to run a single sample), and in the future may pave the way for making highly multiplex imaging at the whole slide level practical for many samples.

A recent study using scRNAseq in the NHP model of TB revealed that lesions demonstrating a low M.tb burden were associated with Th1, T helper 17 (Th17) and cytotoxic T cells producing pro-inflammatory cytokines while lesions containing a high burden of M.tb were associated with T helper 2 (Th2) T cells, plasma cells, mast cells and fibroblasts (Gideon et al., 2022). The pairing of LANDSCAPE tools with highly multiplexed imaging

systems such as CODEX or PhenoCycler, targeting both the cell-types and functional markers identified in that study will be able to extensively phenotype both the composition and spatial architecture of human TB lesions and identify any correlates with the NHP study. Ultimately it is unlikely that M.tb burden will be able to be directly measured in ex-vivo human lung tissue in the near future, given the inaccessibility of TB lung tissue in all but the most severe and treatment resistant cases, therefore identifying correlates of protection in the NHP model represents the most promising approach to identify the spectrum of successful and failed granulomas in the human TB lung.

6.6 Concluding remarks

This thesis represents methodological developments to aid future human immunology research. In chapter 3 I designed a bioinformatic pathway to extensively explore the expression of a gene of interest through meta-analyses and perform hypothesis testing using only publicly available transcriptomic data in a manner both fast and free of cost. The application of this pipeline revealed a novel neutrophil specific marker of acute inflammation across a plethora of inflammatory conditions. In chapter 4 I designed the LANDSCAPE spatial biology analysis pathway that can perform comparative analysis on entire lesions from whole slide images, based on their compositional and spatial architecture, which when applied in chapter 5 revealed novel sub-types of TB granulomas as well as their overall distribution within the M.tb infected lung.

Transcriptomic and multiplex imaging technologies are developing at a rapid pace and are enabling the generation of increasingly large and complicated datasets. This in turn requires the development of continuously evolving analysis protocols to utilise data effectively. The analytical tools developed here are able to generate simple and biologically

relevant findings from enormous datasets and will no doubt prove beneficial to future investigations. Equipped with ever more powerful data generation and analysis tools, future research in human immunology has great potential for the prevention and treatment of human diseases.

Chapter 7

References

- ABBOUD, A., MI, Q., PUCCIO, A., OKONKWO, D., BULIGA, M., CONSTANTINE, G. & VODOVOTZ, Y. 2016. Inflammation following traumatic brain injury in humans: Insights from data-driven and mechanistic models into survival and death. *Front. Pharmacol.*, 7.
- ABDALLAH, N. B., TORAIH, E. A., AL AGEELI, E., ELHAGRASY, H., GOUDA, N. S., FAWZY, M. S. & HELAL, G. M. 2021. MYD88, NFKB1, and IL6 transcripts overexpression are associated with poor outcomes and short survival in neonatal sepsis. *Sci. Rep.*, 11, 13374.
- ABENGOZAR-MUELA, M., ESPARZA, M. V., GARCIA-ROS, D., VÁSQUEZ, C. E., ECHEVESTE, J. I., IDOATE, M. A., LOZANO, M. D., MELERO, I. & DE ANDREA, C. E. 2020. Diverse immune environments in human lung tuberculosis granulomas assessed by quantitative multiplexed immunofluorescence. *Mod. Pathol.*, 33, 2507-2519.
- ADAMS, L. B. & KRAHENBUHL, J. L. 1996. Granulomas Induced by Mycobacterium leprae. *Methods*, 9, 220-32.
- ADAMS, M. D., KELLEY, J. M., GOCAYNE, J. D., DUBNICK, M., POLYMERPOULOS, M. H., XIAO, H., MERRIL, C. R., WU, A., OLDE, B., MORENO, R. F. & ET AL. 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. *Science*, 252, 1651-6.
- ADISA, R., AYANDOKUN, T. T. & IGE, O. M. 2021. Knowledge about tuberculosis, treatment adherence and outcome among ambulatory patients with drug-sensitive tuberculosis in two directly-observed treatment centres in Southwest Nigeria. *BMC Public Health*, 21, 677.
- ADKISON, A. M., RAPTIS, S. Z., KELLEY, D. G. & PHAM, C. T. 2002. Dipeptidyl peptidase I activates neutrophil-derived serine proteases and regulates the development of acute experimental arthritis. *J. Clin. Invest.*, 109, 363-71.
- AHMAD, S., SINGH, P., SHARMA, A., ARORA, S., SHRIWASH, N., RAHMANI, A. H., ALMATROODI, S. A., MANDA, K., DOHARE, R. & SYED, M. A. 2019. Transcriptome Meta-Analysis Deciphers a Dysregulation in Immune Response-Associated Gene Signatures during Sepsis. *Genes*, 10.
- AI, M. & CURRAN, M. A. 2015. Immune checkpoint combinations from mouse to man. *Cancer Immunol. Immunother.*, 64, 885-892.
- AKTER, S., CHAUHAN, K. S., DUNLAP, M. D., CHOREÑO-PARRA, J. A., LU, L., ESAULOVA, E., ZÚÑIGA, J., ARTYOMOV, M. N., KAUSHAL, D. & KHADER, S. A. 2022. Mycobacterium tuberculosis infection drives a type I IFN signature in lung lymphocytes. *Cell Rep.*, 39, 110983.
- ALDER, M. N., OPOKA, A. M., LAHNI, P., HILDEMAN, D. A. & WONG, H. R. 2017. Olfactomedin-4 Is a Candidate Marker for a Pathogenic Neutrophil Subset in Septic Shock. *Crit. Care Med.*, 45, e426-e432.
- ALDRIDGE, S. & TEICHMANN, S. A. 2020. Single cell transcriptomics comes of age. *Nat. Commun.*, 11, 4307.
- ALIZADEH, A. A., EISEN, M. B., DAVIS, R. E., MA, C., LOSSOS, I. S., ROSENWALD, A., BOLDRICK, J. C., SABET, H., TRAN, T., YU, X., POWELL, J. I., YANG, L., MARTI, G. E., MOORE, T., HUDSON, J., JR., LU, L., LEWIS, D. B., TIBSHIRANI, R., SHERLOCK, G., CHAN, W. C., GREINER, T. C., WEISENBURGER, D. D., ARMITAGE, J. O., WARNKE, R., LEVY, R., WILSON, W., GREVER, M. R., BYRD, J. C., BOTSTEIN, D., BROWN, P. O. & STAUDT, L. M. 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*, 403, 503-11.

- ALLAM, M., CAI, S. & COSKUN, A. F. 2020. Multiplex bioimaging of single-cell spatial profiles for precision cancer diagnostics and therapeutics. *NPJ Precis. Oncol.*, 4, 11.
- ALLANTAZ, F., CHAUSSABEL, D., STICHWEH, D., BENNETT, L., ALLMAN, W., MEJIAS, A., ARDURA, M., CHUNG, W., SMITH, E., WISE, C., PALUCKA, K., RAMILO, O., PUNARO, M., BANCHEREAU, J. & PASCUAL, V. 2007. Blood leukocyte microarrays to diagnose systemic onset juvenile idiopathic arthritis and follow the response to IL-1 blockade. *J. Exp. Med.*, 204, 2131-44.
- ALON, U., BARKAI, N., NOTTERMAN, D. A., GISH, K., YBARRA, S., MACK, D. & LEVINE, A. J. 1999. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proc. Natl. Acad. Sci. U. S. A.*, 96, 6745-50.
- ALTMAN, M. C., RINCHAI, D., BALDWIN, N., TOUFIQ, M., WHALEN, E., GARAND, M., SYED AHAMED KABEER, B., ALFAKI, M., PRESNELL, S. R., KHAENAM, P., AYLLÓN-BENÍTEZ, A., MOUGIN, F., THÉBAULT, P., CHICHE, L., JOURDE-CHICHE, N., PHILLIPS, J. T., KLINTMALM, G., O'GARRA, A., BERRY, M., BLOOM, C., WILKINSON, R. J., GRAHAM, C. M., LIPMAN, M., LERTMEMONGKOLCHAI, G., BEDOGNETTI, D., THIEBAUT, R., KHERADMAND, F., MEJIAS, A., RAMILO, O., PALUCKA, K., PASCUAL, V., BANCHEREAU, J. & CHAUSSABEL, D. 2021. Development of a fixed module repertoire for the analysis and interpretation of blood transcriptome data. *Nat. Commun.*, 12, 4385.
- ANAHTAR, M. N., BYRNE, E. H., DOHERTY, K. E., BOWMAN, B. A., YAMAMOTO, H. S., SOUMILLON, M., PADAVATTAN, N., ISMAIL, N., MOODLEY, A., SABATINI, M. E., GHEBREMICHAEL, M. S., NUSBAUM, C., HUTTENHOWER, C., VIRGIN, H. W., NDUNG'U, T., DONG, K. L., WALKER, B. D., FICHOROVA, R. N. & KWON, D. S. 2015. Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. *Immunity*, 42, 965-76.
- ANDERSON, S. T., KAFOROU, M., BRENT, A. J., WRIGHT, V. J., BANWELL, C. M., CHAGALUKA, G., CRAMPIN, A. C., DOCKRELL, H. M., FRENCH, N., HAMILTON, M. S., HIBBERD, M. L., KERN, F., LANGFORD, P. R., LING, L., MLOTHA, R., OTTENHOFF, T. H. M., PIENAAR, S., PILLAY, V., SCOTT, J. A. G., TWAHIR, H., WILKINSON, R. J., COIN, L. J., HEYDERMAN, R. S., LEVIN, M., ELEY, B., GRP, K. T. S. & CONSORTIUM, I. 2014. Diagnosis of childhood tuberculosis and host RNA expression in Africa. *N. Engl. J. Med.*, 370, 1712-1723.
- ANDRES-TERRE, M., MCGUIRE, H. M., POULIOT, Y., BONGEN, E., SWEENEY, T. E., TATO, C. M. & KHATRI, P. 2015. Integrated, Multi-cohort Analysis Identifies Conserved Transcriptional Signatures across Multiple Respiratory Viruses. *Immunity*, 43, 1199-211.
- ANGELO, M., BENDALL, S. C., FINCK, R., HALE, M. B., HITZMAN, C., BOROWSKY, A. D., LEVENSON, R. M., LOWE, J. B., LIU, S. D., ZHAO, S., NATKUNAM, Y. & NOLAN, G. P. 2014. Multiplexed ion beam imaging of human breast tumors. *Nat. Med.*, 20, 436-42.
- ANSONG, E., YING, Q., EKOUE, D. N., DEATON, R., HALL, A. R., KAJDACSY-BALLA, A., YANG, W., GANN, P. H. & DIAMOND, A. M. 2015. Evidence That Selenium Binding Protein 1 Is a Tumor Suppressor in Prostate Cancer. *PLoS One*, 10, e0127295.
- ANTONELLI, L. R., GIGLIOTTI ROTHFUCHS, A., GONÇALVES, R., ROFFÊ, E., CHEEVER, A. W., BAFICA, A., SALAZAR, A. M., FENG, C. G. & SHER, A. 2010. Intranasal Poly-IC treatment exacerbates tuberculosis in mice through the

- pulmonary recruitment of a pathogen-permissive monocyte/macrophage population. *J. Clin. Invest.*, 120, 1674-82.
- ARMINGOL, E., OFFICER, A., HARISMENDY, O. & LEWIS, N. E. 2021. Deciphering cell-cell interactions and communication from gene expression. *Nat. Rev. Genet.*, 22, 71-88.
- ASP, M., BERGENSTRÄHLE, J. & LUNDEBERG, J. 2020. Spatially Resolved Transcriptomes—Next Generation Tools for Tissue Exploration. *Bioessays*, 42, 1900221.
- ATTARWALA, H. 2010. TGN1412: From Discovery to Disaster. *J. Young Pharm.*, 2, 332-6.
- ATTRILL, G. H., LEE, H., TASKER, A. T., ADEGOKE, N. A., FERGUSON, A. L., DA SILVA, I. P., SAW, R. P. M., THOMPSON, J. F., PALENDIRA, U., LONG, G. V., FERGUSON, P. M., SCOLYER, R. A. & WILMOTT, J. S. 2022a. Detailed spatial immunophenotyping of primary melanomas reveals immune cell subpopulations associated with patient outcome. *Front. Immunol.*, 13.
- ATTRILL, G. H., OWEN, C. N., AHMED, T., VERGARA, I. A., COLEBATCH, A. J., CONWAY, J. W., NAHAR, K. J., THOMPSON, J. F., PIRES DA SILVA, I., CARLINO, M. S., MENZIES, A. M., LO, S., PALENDIRA, U., SCOLYER, R. A., LONG, G. V. & WILMOTT, J. S. 2022b. Higher proportions of CD39+ tumor-resident cytotoxic T cells predict recurrence-free survival in patients with stage III melanoma treated with adjuvant immunotherapy. *J. Immunother. Cancer.*, 10, e004771.
- AZIZI, E., CARR, A. J., PLITAS, G., CORNISH, A. E., KONOPACKI, C., PRABHAKARAN, S., NAINYS, J., WU, K., KISELIOVAS, V., SETTY, M., CHOI, K., FROMME, R. M., DAO, P., MCKENNEY, P. T., WASTI, R. C., KADAVERU, K., MAZUTIS, L., RUDENSKY, A. Y. & PE'ER, D. 2018. Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell*, 174, 1293-1308.e36.
- BAECHLER, E. C., BATLIWALLA, F. M., KARYPIS, G., GAFFNEY, P. M., ORTMANN, W. A., ESPE, K. J., SHARK, K. B., GRANDE, W. J., HUGHES, K. M., KAPUR, V., GREGERSEN, P. K. & BEHRENS, T. W. 2003. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. *Proc. Natl. Acad. Sci. U. S. A.*, 100, 2610-5.
- BAGHELA, A., PENA, O. M., LEE, A. H., BAQUIR, B., FALSAFI, R., AN, A., FARMER, S. W., HURLBURT, A., MONDRAGON-CARDONA, A., RIVERA, J. D., BAKER, A., TRAHTEMBERG, U., SHOJAEI, M., JIMENEZ-CANIZALES, C. E., DOS SANTOS, C. C., TANG, B., BOUMA, H. R., COHEN FREUE, G. V. & HANCOCK, R. E. W. 2022. Predicting sepsis severity at first clinical presentation: The role of endotypes and mechanistic signatures. *eBioMedicine*, 75.
- BAHARLOU, H., CANETE, N. P., CUNNINGHAM, A. L., HARMAN, A. N. & PATRICK, E. 2019. Mass Cytometry Imaging for the Study of Human Diseases-Applications and Data Analysis Strategies. *Front. Immunol.*, 10, 2657.
- BALAMUTH, F., ALPERN, E. R., KAN, M., SHUMYATCHER, M., HAYES, K., LAUTENBACH, E. & HIMES, B. E. 2020. Gene Expression Profiles in Children With Suspected Sepsis. *Ann. Emerg. Med.*, 75, 744-754.
- BANCHEREAU, R., HONG, S., CANTAREL, B., BALDWIN, N., BAISCH, J., EDENS, M., CEPIKA, A. M., ACS, P., TURNER, J., ANGUIANO, E., VINOD, P., KAHN, S., OBERMOSER, G., BLANKENSHIP, D., WAKELAND, E., NASSI, L., GOTTE, A., PUNARO, M., LIU, Y. J., BANCHEREAU, J., ROSSELLO-URGELL, J., WRIGHT, T. & PASCUAL, V. 2016. Personalized Immunomonitoring Uncovers Molecular Networks that Stratify Lupus Patients. *Cell*, 165, 551-65.

- BANCHEREAU, R., JORDAN-VILLEGAS, A., ARDURA, M., MEJIAS, A., BALDWIN, N., XU, H., SAYE, E., ROSSELLO-URGELL, J., NGUYEN, P., BLANKENSHIP, D., CREECH, C. B., PASCUAL, V., BANCHEREAU, J., CHAUSSABEL, D. & RAMILO, O. 2012. Host immune transcriptional profiles reflect the variability in clinical disease manifestations in patients with *Staphylococcus aureus* infections. *PLoS One*, 7, e34390.
- BANERJEE, S., MOHAMMED, A., WONG, H. R., PALANIYAR, N. & KAMALESWARAN, R. 2021. Machine Learning Identifies Complicated Sepsis Course and Subsequent Mortality Based on 20 Genes in Peripheral Blood Immune Cells at 24 H Post-ICU Admission. *Front. Immunol.*, 12.
- BARBULOVIC-NAD, I., LUCENTE, M., SUN, Y., ZHANG, M., WHEELER, A. R. & BUSSMANN, M. 2006. Bio-microarray fabrication techniques--a review. *Crit. Rev. Biotechnol.*, 26, 237-59.
- BARRETT, T., WILHITE, S. E., LEDOUX, P., EVANGELISTA, C., KIM, I. F., TOMASHEVSKY, M., MARSHALL, K. A., PHILLIPPY, K. H., SHERMAN, P. M., HOLKO, M., YEFANOV, A., LEE, H., ZHANG, N., ROBERTSON, C. L., SEROVA, N., DAVIS, S. & SOBOLEVA, A. 2013. NCBI GEO: Archive for functional genomics data sets - Update. *Nucleic Acids Res.*, 41, D991-D995.
- BARRY, C. E., 3RD, BOSHOF, H. I., DARTOIS, V., DICK, T., EHRT, S., FLYNN, J., SCHNAPPINGER, D., WILKINSON, R. J. & YOUNG, D. 2009. The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat. Rev. Microbiol.*, 7, 845-55.
- BARRY, C. E., III & MAYER-BARBER, K. D. 2021. Signature required: The transcriptional response to tuberculosis. *J. Exp. Med.*, 218, e20211665.
- BARTON, A. J., HILL, J., POLLARD, A. J. & BLOHMKE, C. J. 2017. Transcriptomics in Human Challenge Models. *Front. Immunol.*, 8.
- BASARABA, R. J. & HUNTER, R. L. 2017. Pathology of Tuberculosis: How the Pathology of Human Tuberculosis Informs and Directs Animal Models. *Microbiol Spectr.*, 5.
- BAYKAN, A. H., SAYINER, H. S., AYDIN, E., KOC, M., INAN, I. & ERTURK, S. M. 2022. Extrapulmonary tuberculosis: an old but resurgent problem. *Insights Imaging*, 13, 39.
- BEAN, A. G., ROACH, D. R., BRISCOE, H., FRANCE, M. P., KORNER, H., SEDGWICK, J. D. & BRITTON, W. J. 1999. Structural deficiencies in granuloma formation in TNF gene-targeted mice underlie the heightened susceptibility to aerosol *Mycobacterium tuberculosis* infection, which is not compensated for by lymphotoxin. *J. Immunol.*, 162, 3504-11.
- BEHAR, S. M., DASCHER, C. C., GRUSBY, M. J., WANG, C. R. & BRENNER, M. B. 1999. Susceptibility of mice deficient in CD1D or TAP1 to infection with *Mycobacterium tuberculosis*. *J. Exp. Med.*, 189, 1973-80.
- BENNETT, L., PALUCKA, A. K., ARCE, E., CANTRELL, V., BORVAK, J., BANCHEREAU, J. & PASCUAL, V. 2003. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J. Exp. Med.*, 197, 711-23.
- BERDAL, J. E., MOLLNES, T. E., WÆHRE, T., OLSTAD, O. K., HALVORSEN, B., UELAND, T., LAAKE, J. H., FURUSETH, M. T., MAAGAARD, A., KJEKSHUS, H., AUKRUST, P. & JONASSEN, C. M. 2011. Excessive innate immune response and mutant D222G/N in severe A (H1N1) pandemic influenza. *J. Infect.*, 63, 308-16.
- BERRY, M. P., GRAHAM, C. M., MCNAB, F. W., XU, Z., BLOCH, S. A., ONI, T., WILKINSON, K. A., BANCHEREAU, R., SKINNER, J., WILKINSON, R. J., QUINN, C., BLANKENSHIP, D., DHAWAN, R., CUSH, J. J., MEJIAS, A., RAMILO, O., KON, O. M., PASCUAL, V., BANCHEREAU, J., CHAUSSABEL, D.

- & O'GARRA, A. 2010. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature*, 466, 973-7.
- BEURA, L. K., HAMILTON, S. E., BI, K., SCHENKEL, J. M., ODUMADE, O. A., CASEY, K. A., THOMPSON, E. A., FRASER, K. A., ROSATO, P. C., FILALI-MOUHIM, A., SEKALY, R. P., JENKINS, M. K., VEZYS, V., HAINING, W. N., JAMESON, S. C. & MASOPIST, D. 2016. Normalizing the environment recapitulates adult human immune traits in laboratory mice. *Nature*, 532, 512-6.
- BHATTACHARYA, S., ANDORF, S., GOMES, L., DUNN, P., SCHAEFER, H., PONTIUS, J., BERGER, P., DESBOROUGH, V., SMITH, T., CAMPBELL, J., THOMSON, E., MONTEIRO, R., GUIMARAES, P., WALTERS, B., WISER, J. & BUTTE, A. J. 2014. ImmPort: disseminating data to the public for the future of immunology. *Immunol. Res.*, 58, 234-239.
- BINDEA, G., MLECNIK, B., TOSOLINI, M., KIRILOVSKY, A., WALDNER, M., OBENAUF, ANNA C., ANGELL, H., FREDRIKSEN, T., LAFONTAINE, L., BERGER, A., BRUNEVAL, P., FRIDMAN, WOLF H., BECKER, C., PAGÈS, F., SPEICHER, MICHAEL R., TRAJANOSKI, Z. & GALON, J. 2013. Spatiotemporal Dynamics of Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer. *Immunity*, 39, 782-795.
- BIRD, P. I., TRAPANI, J. A. & VILLADANGOS, J. A. 2009. Endolysosomal proteases and their inhibitors in immunity. *Nat. Rev. Immunol.*, 9, 871.
- BITTNER, M., MELTZER, P., CHEN, Y., JIANG, Y., SEFTOR, E., HENDRIX, M., RADMACHER, M., SIMON, R., YAKHINI, Z., BEN-DOR, A., SAMPAS, N., DOUGHERTY, E., WANG, E., MARINCOLA, F., GOODEN, C., LUEDERS, J., GLATFELTER, A., POLLOCK, P., CARPTEN, J., GILLANDERS, E., LEJA, D., DIETRICH, K., BEAUDRY, C., BERENS, M., ALBERTS, D. & SONDAK, V. 2000. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature*, 406, 536-40.
- BJÖRKLUND Å, K., FORKEL, M., PICELLI, S., KONYA, V., THEORELL, J., FRIBERG, D., SANDBERG, R. & MJÖSBORG, J. 2016. The heterogeneity of human CD127(+) innate lymphoid cells revealed by single-cell RNA sequencing. *Nat. Immunol.*, 17, 451-60.
- BJORNSON-HOOPER, Z. B., FRAGIADAKIS, G. K., SPITZER, M. H., CHEN, H., MADHIREDDY, D., HU, K., LUNDSTEN, K., MCILWAIN, D. R. & NOLAN, G. P. 2022. A Comprehensive Atlas of Immunological Differences Between Humans, Mice, and Non-Human Primates. *Front. Immunol.*, 13.
- BLACK, S., PHILLIPS, D., HICKEY, J. W., KENNEDY-DARLING, J., VENKATARAAMAN, V. G., SAMUSIK, N., GOLTSEV, Y., SCHÜRCH, C. M. & NOLAN, G. P. 2021. CODEX multiplexed tissue imaging with DNA-conjugated antibodies. *Nat. Protoc.*, 16, 3802-3835.
- BLANKLEY, S., GRAHAM, C. M., LEVIN, J., TURNER, J., BERRY, M. P., BLOOM, C. I., XU, Z., PASCUAL, V., BANCHEREAU, J., CHAUSSABEL, D., BREEN, R., SANTIS, G., BLANKENSHIP, D. M., LIPMAN, M. & O'GARRA, A. 2016a. A 380-gene meta-signature of active tuberculosis compared with healthy controls. *Eur. Respir. J.*, 47, 1873-6.
- BLANKLEY, S., GRAHAM, C. M., TURNER, J., BERRY, M. P., BLOOM, C. I., XU, Z., PASCUAL, V., BANCHEREAU, J., CHAUSSABEL, D., BREEN, R., SANTIS, G., BLANKENSHIP, D. M., LIPMAN, M. & O'GARRA, A. 2016b. The Transcriptional Signature of Active Tuberculosis Reflects Symptom Status in Extra-Pulmonary and Pulmonary Tuberculosis. *PLoS One*, 11, e0162220.

- BLOOM, C. I., GRAHAM, C. M., BERRY, M. P., ROZAKEAS, F., REDFORD, P. S., WANG, Y., XU, Z., WILKINSON, K. A., WILKINSON, R. J., KENDRICK, Y., DEVOUASSOUX, G., FERRY, T., MIYARA, M., BOUVRY, D., VALEYRE, D., GOROCHOV, G., BLANKENSHIP, D., SAADATIAN, M., VANHEMS, P., BEYNON, H., VANCHEESWARAN, R., WICKREMASINGHE, M., CHAUSSABEL, D., BANCHEREAU, J., PASCUAL, V., HO, L. P., LIPMAN, M. & O'GARRA, A. 2013. Transcriptional blood signatures distinguish pulmonary tuberculosis, pulmonary sarcoidosis, pneumonias and lung cancers. *PLoS One*, 8, e70630.
- BLOOM, C. I., GRAHAM, C. M., BERRY, M. P., WILKINSON, K. A., ONI, T., ROZAKEAS, F., XU, Z., ROSSELLO-URGELL, J., CHAUSSABEL, D., BANCHEREAU, J., PASCUAL, V., LIPMAN, M., WILKINSON, R. J. & O'GARRA, A. 2012. Detectable changes in the blood transcriptome are present after two weeks of antituberculosis therapy. *PLoS One*, 7, e46191.
- BOUGARN, S., BOUGHORBEL, S., CHAUSSABEL, D. & MARR, N. 2019. A curated transcriptome dataset collection to investigate the blood transcriptional response to viral respiratory tract infection and vaccination. *F1000Res.*, 8, 284.
- BRACHAT, A. H., GROM, A. A., WULFFRAAT, N., BRUNNER, H. I., QUARTIER, P., BRIK, R., MCCANN, L., OZDOGAN, H., RUTKOWSKA-SAK, L., SCHNEIDER, R., GERLONI, V., HAREL, L., TERRERI, M., HOUGHTON, K., JOOS, R., KINGSBURY, D., LOPEZ-BENITEZ, J. M., BEK, S., SCHUMACHER, M., VALENTIN, M. A., GRAM, H., ABRAMS, K., MARTINI, A., LOVELL, D. J., NIRMALA, N. R. & RUPERTO, N. 2017. Early changes in gene expression and inflammatory proteins in systemic juvenile idiopathic arthritis patients on canakinumab therapy. *Arthritis Res. Ther.*, 19, 13.
- BRETON, G., ZHENG, S., VALIERIS, R., TOJAL DA SILVA, I., SATIJA, R. & NUSSENZWEIG, M. C. 2016. Human dendritic cells (DCs) are derived from distinct circulating precursors that are precommitted to become CD1c⁺ or CD141⁺ DCs. *J. Exp. Med.*, 213, 2861-2870.
- BRINSTER, R. L., CHEN, H. Y., TRUMBAUER, M., SENEAR, A. W., WARREN, R. & PALMITER, R. D. 1981. Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. *Cell*, 27, 223-31.
- BRODIN, P. & DAVIS, M. M. 2017. Human immune system variation. *Nat. Rev. Immunol.*, 17, 21-29.
- BRODIN, P., JOJIC, V., GAO, T., BHATTACHARYA, S., ANGEL, C. J. L., FURMAN, D., SHEN-ORR, S., DEKKER, C. L., SWAN, G. E., BUTTE, A. J., MAECKER, H. T. & DAVIS, M. M. 2015. Variation in the human immune system is largely driven by non-heritable influences. *Cell*, 160, 37-47.
- BRODIN, P., VALENTINI, D., UHLIN, M., MATTSSON, J., ZUMLA, A. & MAEURER, M. J. 2013. Systems level immune response analysis and personalized medicine. *Expert Rev. Clin. Immunol.*, 9, 307-317.
- BROK, H. P. M., BAUER, J., JONKER, M., BLEZER, E., AMOR, S., BONTROP, R. E., LAMAN, J. D. & 'T HART, B. A. 2001. Non-human primate models of multiple sclerosis. *Immunol. Rev.*, 183, 173-185.
- BROOS, C., VAN NIMWEGEN, M., HOOGSTEDEN, H., HENDRIKS, R., KOOL, M. & VAN DEN BLINK, B. 2013. Granuloma Formation in Pulmonary Sarcoidosis. *Front. Immunol.*, 4.
- BUCHBINDER, S. P., MEHROTRA, D. V., DUERR, A., FITZGERALD, D. W., MOGG, R., LI, D., GILBERT, P. B., LAMA, J. R., MARMOR, M., DEL RIO, C., MCEL RATH, M. J., CASIMIRO, D. R., GOTTESDIENER, K. M.,

- CHODAKEWITZ, J. A., COREY, L. & ROBERTSON, M. N. 2008. Efficacy assessment of a cell-mediated immunity HIV-1 vaccine (the Step Study): a double-blind, randomised, placebo-controlled, test-of-concept trial. *Lancet*, 372, 1881-1893.
- BURCZYNSKI, M. E., PETERSON, R. L., TWINE, N. C., ZUBEREK, K. A., BRODEUR, B. J., CASCIOTTI, L., MAGANTI, V., REDDY, P. S., STRAHS, A., IMMERMAN, F., SPINELLI, W., SCHWERTSCHLAG, U., SLAGER, A. M., COTREAU, M. M. & DORNER, A. J. 2006. Molecular classification of Crohn's disease and ulcerative colitis patients using transcriptional profiles in peripheral blood mononuclear cells. *J. Mol. Diagn.*, 8, 51-61.
- BUREL, J. G., APTE, S. H. & DOOLAN, D. L. 2016. Systems Approaches towards Molecular Profiling of Human Immunity. *Trends Immunol.*, 37, 53-67.
- BURNET, F. M. 1976. A modification of Jerne's theory of antibody production using the concept of clonal selection. *CA Cancer J. Clin.*, 26, 119-21.
- BYRON, S. A., VAN KEUREN-JENSEN, K. R., ENGELTHALER, D. M., CARPTEN, J. D. & CRAIG, D. W. 2016. Translating RNA sequencing into clinical diagnostics: opportunities and challenges. *Nat. Rev. Genet.*, 17, 257-71.
- CADENA, A. M., FORTUNE, S. M. & FLYNN, J. L. 2017. Heterogeneity in tuberculosis. *Nat. Rev. Immunol.*, 17, 691-702.
- CADENA ANTHONY, M., FLYNN JOANNE, L. & FORTUNE SARAH, M. 2016. The Importance of First Impressions: Early Events in Mycobacterium tuberculosis Infection Influence Outcome. *mBio*, 7, e00342-16.
- CAI, Y., WANG, Y., SHI, C., DAI, Y., LI, F., XU, Y., ZHANG, P., KONG, F., DENG, G., WEN, Z., ZHOU, Q., KANG, B. C., SINGHAL, A., YANG, Q., FENG, C. G. & CHEN, X. 2022. Single-cell immune profiling reveals functional diversity of T cells in tuberculous pleural effusion. *J. Exp. Med.*, 219.
- CALLIGARO, G. L., MOODLEY, L., SYMONS, G. & DHEDA, K. 2014. The medical and surgical treatment of drug-resistant tuberculosis. *J. Thorac. Dis.*, 6, 186-95.
- CANNOODT, R., SAELENS, W., SICHIE, D., TAVERNIER, S., JANSSENS, S., GUILLIAMS, M., LAMBRECHT, B., PRETER, K. D. & SAEYS, Y. 2016. SCORPIUS improves trajectory inference and identifies novel modules in dendritic cell development. *bioRxiv*, 079509.
- CAPUANO, S. V., 3RD, CROIX, D. A., PAWAR, S., ZINOVIK, A., MYERS, A., LIN, P. L., BISSEL, S., FUHRMAN, C., KLEIN, E. & FLYNN, J. L. 2003. Experimental Mycobacterium tuberculosis infection of cynomolgus macaques closely resembles the various manifestations of human M. tuberculosis infection. *Infect. Immun.*, 71, 5831-44.
- CAROW, B., HAULING, T., QIAN, X., KRAMNIK, I., NILSSON, M. & ROTTENBERG, M. E. 2019. Spatial and temporal localization of immune transcripts defines hallmarks and diversity in the tuberculosis granuloma. *Nat. Commun.*, 10, 1823.
- CARPENTER, D. J., GRANOT, T., MATSUOKA, N., SENDA, T., KUMAR, B. V., THOME, J. J. C., GORDON, C. L., MIRON, M., WEINER, J., CONNORS, T., LERNER, H., FRIEDMAN, A., KATO, T., GRIESEMER, A. D. & FARBER, D. L. 2018. Human immunology studies using organ donors: Impact of clinical variations on immune parameters in tissues and circulation. *Am. J. Transplant.*, 18, 74-88.
- CARR, E. J., DOOLEY, J., GARCIA-PEREZ, J. E., LAGOU, V., LEE, J. C., WOUTERS, C., MEYTS, I., GORIS, A., BOECKXSTAENS, G., LINTERMAN, M. A. & LISTON, A. 2016. The cellular composition of the human immune system is shaped by age and cohabitation. *Nat. Immunol.*, 17, 461-468.
- CARUSO, R., PARISI, A., BONANNO, A., PAPARO, D., QUATTROCCHI, E., BRANCA, G., SCARDIGNO, M. & FEDELE, F. 2012. Histologic coagulative tumour necrosis

- as a prognostic indicator of aggressiveness in renal, lung, thyroid and colorectal carcinomas: A brief review. *Oncol. Lett.*, 3, 16-18.
- CASANOVA, J. L. & ABEL, L. 2002. Genetic dissection of immunity to mycobacteria: the human model. *Annu. Rev. Immunol.*, 20, 581-620.
- CHAKAYA, J. M., HARRIES, A. D. & MARKS, G. B. 2020. Ending tuberculosis by 2030—Pipe dream or reality? *Int. J. Infect. Dis.*, 92, S51-S54.
- CHAMBERS, D. C., CAREW, A. M., LUKOWSKI, S. W. & POWELL, J. E. 2019. Transcriptomics and single-cell RNA-sequencing. *Respirology*, 24, 29-36.
- CHAUSSABEL, D. & BALDWIN, N. 2014. Democratizing systems immunology with modular transcriptional repertoire analyses. *Nat. Rev. Immunol.*, 14, 271-280.
- CHAUSSABEL, D., QUINN, C., SHEN, J., PATEL, P., GLASER, C., BALDWIN, N., STICHWEH, D., BLANKENSHIP, D., LI, L., MUNAGALA, I., BENNETT, L., ALLANTAZ, F., MEJIAS, A., ARDURA, M., KAIZER, E., MONNET, L., ALLMAN, W., RANDALL, H., JOHNSON, D., LANIER, A., PUNARO, M., WITTKOWSKI, K. M., WHITE, P., FAY, J., KLINTMALM, G., RAMILO, O., PALUCKA, A. K., BANCHEREAU, J. & PASCUAL, V. 2008. A Modular Analysis Framework for Blood Genomics Studies: Application to Systemic Lupus Erythematosus. *Immunity*, 29, 150-164.
- CHEN, H., LI, Y., LI, T., SUN, H., TAN, C., GAO, M., XING, W. & XIAO, X. 2019. Identification of Potential Transcriptional Biomarkers Differently Expressed in Both *S. aureus*- and *E. coli*-Induced Sepsis via Integrated Analysis. *Biomed. Res. Int.*, 2019, 2487921.
- CHO, C., LINSCHER, W. G., HIRSCHKORN, M. A. & ASHUTOSH, K. 1984. Sarcoidlike granulomas as an early manifestation of Whipple's disease. *Gastroenterology*, 87, 941-947.
- CLAYTON, K., POLAK, M. E., WOELK, C. H. & ELKINGTON, P. 2017. Gene Expression Signatures in Tuberculosis Have Greater Overlap with Autoimmune Diseases Than with Infectious Diseases. *Am. J. Respir. Crit. Care Med.*, 196, 655-656.
- CLIFF, J. M., LEE, J. S., CONSTANTINOU, N., CHO, J. E., CLARK, T. G., RONACHER, K., KING, E. C., LUKEY, P. T., DUNCAN, K., VAN HELDEN, P. D., WALZL, G. & DOCKRELL, H. M. 2013. Distinct phases of blood gene expression pattern through tuberculosis treatment reflect modulation of the humoral immune response. *J. Infect. Dis.*, 207, 18-29.
- COHEN, A., MATHIASSEN, V. D., SCHÖN, T. & WEJSE, C. 2019. The global prevalence of latent tuberculosis: a systematic review and meta-analysis. *Eur. Respir. J.*, 54.
- COHEN, S. B., GERN, B. H. & URDAHL, K. B. 2022. The Tuberculous Granuloma and Preexisting Immunity. *Annu. Rev. Immunol.*, 40, 589-614.
- COLBERT, J. D., MATTHEWS, S. P., KOS, J. & WATTS, C. 2011. Internalization of exogenous cystatin F suppresses cysteine proteases and induces the accumulation of single-chain cathepsin L by multiple mechanisms. *Biol. Chem.*, 286, 42082-42090.
- COLBERT, J. D., MATTHEWS, S. P., MILLER, G. & WATTS, C. 2009. Diverse regulatory roles for lysosomal proteases in the immune response. *Eur. J. Immunol.*, 39, 2955-65.
- COLDREN, C. D., NICK, J. A., POCH, K. R., WOOLUM, M. D., FOUTY, B. W., O'BRIEN, J. M., GRUBER, M. P., ZAMORA, M. R., SVETKAUSKAITE, D., RICHTER, D. A., HE, Q., PARK, J. S., OVERDIER, K. H., ABRAHAM, E. & GERACI, M. W. 2006. Functional and genomic changes induced by alveolar transmigration in human neutrophils. *Am. J. Physiol. Lung Cell Mol. Physiol.*, 291, L1267-76.
- COLEMAN, M. T., CHEN, R. Y., LEE, M., LIN, P. L., DODD, L. E., MAIELLO, P., VIA, L. E., KIM, Y., MARRINER, G., DARTOIS, V., SCANGA, C., JANSSEN, C.,

- WANG, J., KLEIN, E., CHO, S. N., BARRY, C. E., 3RD & FLYNN, J. L. 2014. PET/CT imaging reveals a therapeutic response to oxazolidinones in macaques and humans with tuberculosis. *Sci. Transl. Med.*, 6, 265ra167.
- COLETTA, A., MOLTER, C., DUQUÉ, R., STEENHOFF, D., TAMINAU, J., DE SCHAETZEN, V., MEGANCK, S., LAZAR, C., VENET, D., DETOURS, V., NOWÉ, A., BERSINI, H. & WEISS SOLÍS, D. Y. 2012. InSilico DB genomic datasets hub: an efficient starting point for analyzing genome-wide studies in GenePattern, Integrative Genomics Viewer, and R/Bioconductor. *Genome Biol.*, 13, R104.
- COMARMOND, C. & CACOUB, P. 2014. Granulomatosis with polyangiitis (Wegener): Clinical aspects and treatment. *Autoimmun. Rev.*, 13, 1121-1125.
- CONSTANTINESCU, C. S., FAROOQI, N., O'BRIEN, K. & GRAN, B. 2011. Experimental autoimmune encephalomyelitis (EAE) as a model for multiple sclerosis (MS). *Br. J. Pharmacol.*, 164, 1079-106.
- COONS, A. H., CREECH, H. J. & JONES, R. N. 1941. Immunological Properties of an Antibody Containing a Fluorescent Group. *Proc. Soc. Exp. Biol. Med.*, 47, 200-202.
- COOPER, A. M., COOPER, A. M., DALTON, D. K., DALTON, D. K., STEWART, T. A., STEWART, T. A., GRIFFIN, J. P., GRIFFIN, J. P., RUSSELL, D. G., RUSSELL, D. G., ORME, I. M. & ORME, I. M. 1993. Disseminated tuberculosis in interferon γ gene-disrupted mice. *J. Exp. Med.*, 178, 2243-2247.
- COOPER, A. M., MAGRAM, J., FERRANTE, J. & ORME, I. M. 1997. Interleukin 12 (IL-12) is crucial to the development of protective immunity in mice intravenously infected with mycobacterium tuberculosis. *J. Exp. Med.*, 186, 39-45.
- COOPER, M. D., PETERSON, R. D. & GOOD, R. A. 1965. DELINEATION OF THE THYMIC AND BURSAL LYMPHOID SYSTEMS IN THE CHICKEN. *Nature*, 205, 143-6.
- CORTESE, M., SHERMAN, A. C., ROUPHAEL, N. G. & PULENDRAN, B. 2021. Systems Biological Analysis of Immune Response to Influenza Vaccination. *Cold Spring Harb. Perspect. Med.*, 11.
- 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.
- DAMODARAN, S., DAMASCHKE, N., GAWDZIK, J., YANG, B., SHI, C., ALLEN, G. O., HUANG, W., DENU, J. & JARRARD, D. 2017. Dysregulation of Sirtuin 2 (SIRT2) and histone H3K18 acetylation pathways associates with adverse prostate cancer outcomes. *BMC Cancer*, 17, 874.
- DAMOND, N., ENGLER, S., ZANOTELLI, V. R. T., SCHAPIRO, D., WASSERFALL, C. H., KUSMARTSEVA, I., NICK, H. S., THOREL, F., HERRERA, P. L., ATKINSON, M. A. & BODENMILLER, B. 2019. A Map of Human Type 1 Diabetes Progression by Imaging Mass Cytometry. *Cell Metab.*, 29, 755-768.e5.
- DAVIS, J. M., CLAY, H., LEWIS, J. L., GHORI, N., HERBOMEL, P. & RAMAKRISHNAN, L. 2002. Real-Time Visualization of Mycobacterium-Macrophage Interactions Leading to Initiation of Granuloma Formation in Zebrafish Embryos. *Immunity*, 17, 693-702.
- DAVIS, J. M. & RAMAKRISHNAN, L. 2009. The Role of the Granuloma in Expansion and Dissemination of Early Tuberculous Infection. *Cell*, 136, 37-49.
- DAVIS, M. M. 2008. A Prescription for Human Immunology. *Immunity*, 29, 835-838.
- DAVIS, M. M. & BRODIN, P. 2018. Rebooting Human Immunology. *Annu. Rev. Immunol.*, 36, 843-864.

- DAVIS, S. & MELTZER, P. S. 2007. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics*, 23, 1846-7.
- DEKKERS, G., BENTLAGE, A. E. H., STEGMANN, T. C., HOWIE, H. L., LISSENBERG-THUNNISSEN, S., ZIMRING, J., RISPENS, T. & VIDARSSON, G. 2017. Affinity of human IgG subclasses to mouse Fc gamma receptors. *MAbs*, 9, 767-773.
- DEMARET, J., VENET, F., FRIGGERI, A., CAZALIS, M. A., PLASSAIS, J., JALLADES, L., MALCUS, C., POITEVIN-LATER, F., TEXTORIS, J., LEPAPE, A. & MONNERET, G. 2015. Marked alterations of neutrophil functions during sepsis-induced immunosuppression. *J. Leukoc. Biol.*, 98, 1081-90.
- DEN BRABER, I., MUGWAGWA, T., VRISEKOOP, N., WESTERA, L., MÖGLING, R., DE BOER, A. B., WILLEMS, N., SCHRIJVER, E. H., SPIERENBURG, G., GAISER, K., MUL, E., OTTO, S. A., RUITER, A. F., ACKERMANS, M. T., MIEDEMA, F., BORGHANS, J. A., DE BOER, R. J. & TESSELAAR, K. 2012. Maintenance of peripheral naive T cells is sustained by thymus output in mice but not humans. *Immunity*, 36, 288-97.
- DHARMADHIKARI, A. S. & NARDELL, E. A. 2008. What Animal Models Teach Humans about Tuberculosis. *Am. J. Respir. Cell Mol. Biol.*, 39, 503-508.
- DHEDA, K., BARRY, C. E. & MAARTENS, G. 2016. Tuberculosis. *Lancet*, 387, 1211-1226.
- DICKINSON, P., SMITH, C. L., FORSTER, T., CRAIGON, M., ROSS, A. J., KHONDOKER, M. R., IVENS, A., LYNN, D. J., ORME, J., JACKSON, A., LACAZE, P., FLANAGAN, K. L., STENSON, B. J. & GHAZAL, P. 2015. Whole blood gene expression profiling of neonates with confirmed bacterial sepsis. *Genom. Data*, 3, 41-8.
- DING, J., ADICONIS, X., SIMMONS, S. K., KOWALCZYK, M. S., HESSION, C. C., MARJANOVIC, N. D., HUGHES, T. K., WADSWORTH, M. H., BURKS, T., NGUYEN, L. T., KWON, J. Y. H., BARAK, B., GE, W., KEDAIGLE, A. J., CARROLL, S., LI, S., HACHOEN, N., ROZENBLATT-ROSEN, O., SHALEK, A. K., VILLANI, A.-C., REGEV, A. & LEVIN, J. Z. 2020. Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. *Nat. Biotechnol.*, 38, 737-746.
- DOBOS, K. M., QUINN, F. D., ASHFORD, D. A., HORSBURGH, C. R. & KING, C. H. 1999. Emergence of a unique group of necrotizing mycobacterial diseases. *Emerg. Infect. Dis.*, 5, 367-78.
- DRISSEN, R., BUZA-VIDAS, N., WOLL, P., THONGJUEA, S., GAMBARDELLA, A., GIUSTACCHINI, A., MANCINI, E., ZRIWIL, A., LUTTEROPP, M., GROVER, A., MEAD, A., SITNICKA, E., JACOBSEN, S. E. W. & NERLOV, C. 2016. Distinct myeloid progenitor-differentiation pathways identified through single-cell RNA sequencing. *Nat. Immunol.*, 17, 666-676.
- DUNNING, J., BLANKLEY, S., HOANG, L. T., COX, M., GRAHAM, C. M., JAMES, P. L., BLOOM, C. I., CHAUSSABEL, D., BANCHEREAU, J., BRETT, S. J., MOFFATT, M. F., O'GARRA, A. & OPENSHAW, P. J. M. 2018. Progression of whole-blood transcriptional signatures from interferon-induced to neutrophil-associated patterns in severe influenza. *Nat. Immunol.*, 19, 625-635.
- DURAIYAN, J., GOVINDARAJAN, R., KALIYAPPAN, K. & PALANISAMY, M. 2012. Applications of immunohistochemistry. *J. Pharm. Bioallied Sci.*, 4, S307-9.
- EDGAR, R., DOMRACHEV, M. & LASH, A. E. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.*, 30, 207-10.

- ELLIOTT, T. O., OWOLABI, O., DONKOR, S., KAMPMANN, B., HILL, P. C., OTTENHOFF, T. H., HAKS, M. C., KAUFMANN, S. H., MAERTZDORF, J. & SUTHERLAND, J. S. 2015. Dysregulation of Apoptosis Is a Risk Factor for Tuberculosis Disease Progression. *J. Infect. Dis.*, 212, 1469-79.
- ESAULOVA, E., DAS, S., SINGH, D. K., CHOREÑO-PARRA, J. A., SWAIN, A., ARTHUR, L., RANGEL-MORENO, J., AHMED, M., SINGH, B., GUPTA, A., FERNÁNDEZ-LÓPEZ, L. A., DE LA LUZ GARCIA-HERNANDEZ, M., BUCSAN, A., MOODLEY, C., MEHRA, S., GARCÍA-LATORRE, E., ZUNIGA, J., ATKINSON, J., KAUSHAL, D., ARTYOMOV, M. N. & KHADER, S. A. 2021. The immune landscape in tuberculosis reveals populations linked to disease and latency. *Cell Host Microbe*, 29, 165-178.e8.
- FALCON, S. & GENTLEMAN, R. 2007. Using GOstats to test gene lists for GO term association. *Bioinformatics*, 23, 257-258.
- FEDERICO, A. & MONTI, S. 2020. hypeR: an R package for geneset enrichment workflows. *Bioinformatics*, 36, 1307-1308.
- FELDMANN, M. & MAINI, S. R. 2008. Role of cytokines in rheumatoid arthritis: an education in pathophysiology and therapeutics. *Immunol. Rev.*, 223, 7-19.
- FENG, C. G., JANKOVIC, D., KULLBERG, M., CHEEVER, A., SCANGA, C. A., HIENY, S., CASPAR, P., YAP, G. S. & SHER, A. 2005. Maintenance of pulmonary Th1 effector function in chronic tuberculosis requires persistent IL-12 production. *J. Immunol.*, 174, 4185-92.
- FENG, Z., PURI, S., MOUDGIL, T., WOOD, W., HOYT, C. C., WANG, C., URBA, W. J., CURTI, B. D., BIFULCO, C. B. & FOX, B. A. 2015. Multispectral imaging of formalin-fixed tissue predicts ability to generate tumor-infiltrating lymphocytes from melanoma. *J. Immunother. Cancer.*, 3, 47.
- FINE, P. E. M. 1995. Variation in protection by BCG: implications of and for heterologous immunity. *Lancet*, 346, 1339-1345.
- FITZGERALD, G., BOTSTEIN, D., CALIFF, R., COLLINS, R., PETERS, K., VAN BRUGGEN, N. & RADER, D. 2018. The future of humans as model organisms. *Science*, 361, 552-553.
- FLYNN, J. L., GIDEON, H. P., MATTILA, J. T. & LIN, PHILANA L. 2015. Immunology studies in non-human primate models of tuberculosis. *Immunol. Rev.*, 264, 60-73.
- FLYNN, J. L., GOLDSTEIN, M. M., CHAN, J., TRIEBOLD, K. J., PFEFFER, K., LOWENSTEIN, C. J., SCHRELBER, R., MAK, T. W. & BLOOM, B. R. 1995. Tumor necrosis factor- α is required in the protective immune response against mycobacterium tuberculosis in mice. *Immunity*, 2, 561-572.
- FORD, N., SHUBBER, Z., MEINTJES, G., GRINSZTEJN, B., EHOLIE, S., MILLS, E. J., DAVIES, M. A., VITORIA, M., PENAZZATO, M., NSANZIMANA, S., FRIGATI, L., O'BRIEN, D., ELLMAN, T., AJOSE, O., CALMY, A. & DOHERTY, M. 2015. Causes of hospital admission among people living with HIV worldwide: a systematic review and meta-analysis. *Lancet HIV*, 2, e438-44.
- FREMOND, C. M., TOGBE, D., DOZ, E., ROSE, S., VASSEUR, V., MAILLET, I., JACOBS, M., RYFFEL, B. & QUESNIAUX, V. F. J. 2007. IL-1 Receptor-Mediated Signal Is an Essential Component of MyD88-Dependent Innate Response to Mycobacterium tuberculosis Infection. *J. Immunol.*, 179, 1178-1189.
- FREMOND, C. M., YEREMEEV, V., NICOLLE, D. M., JACOBS, M., QUESNIAUX, V. F. & RYFFEL, B. 2004. Fatal Mycobacterium tuberculosis infection despite adaptive immune response in the absence of MyD88. *J. Clin. Invest.*, 114, 1790-1799.
- FURMAN, D. & DAVIS, M. M. 2015. New approaches to understanding the immune response to vaccination and infection. *Vaccine*, 33, 5271-5281.

- GALON, J., COSTES, A., SANCHEZ-CABO, F., KIRILOVSKY, A., MLECNIK, B., LAGORCE-PAGÈS, C., TOSOLINI, M., CAMUS, M., BERGER, A., WIND, P., ZINZINDOHOUE, F., BRUNEVAL, P., CUGNENC, P. H., TRAJANOSKI, Z., FRIDMAN, W. H. & PAGÈS, F. 2006. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science*, 313, 1960-4.
- GARNIER, S., ROSS, N., RUDIS, B., SCIAINI, M., CAMARGO, A. P. & SCHERER, C. 2021. *Viridis - Colorblind-Friendly Color Maps for R* [Online]. Available: <https://sjmgarnier.github.io/viridis/> [Accessed].
- GAVE, A. A., TORRES, R. & KAPLAN, L. 2004. Cryptococcal Myositis and Vasculitis: An Unusual Necrotizing Soft Tissue Infection. *Surg. Infect. (Larchmt.)*, 5, 309-313.
- GE, S. X., SON, E. W. & YAO, R. 2018. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics*, 19, 534.
- GELDMACHER, C., ZUMLA, A. & HOELSCHER, M. 2012. Interaction between HIV and Mycobacterium tuberculosis: HIV-1-induced CD4 T-cell depletion and the development of active tuberculosis. *Curr. Opin. HIV AIDS*, 7, 268-75.
- GENE ONTOLOGY, C. 2004. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.*, 32, D258-D261.
- GENTLES, A. J., NEWMAN, A. M., LIU, C. L., BRATMAN, S. V., FENG, W., KIM, D., NAIR, V. S., XU, Y., KHUONG, A., HOANG, C. D., DIEHN, M., WEST, R. B., PLEVITIS, S. K. & ALIZADEH, A. A. 2015. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat. Med.*, 21, 938-945.
- GERDTSSON, E., PORE, M., THIELE, J. A., GERDTSSON, A. S., MALIHI, P. D., NEVAREZ, R., KOLATKAR, A., VELASCO, C. R., WIX, S., SINGH, M., CARLSSON, A., ZURITA, A. J., LOGOTHETIS, C., MERCHANT, A. A., HICKS, J. & KUHN, P. 2018. Multiplex protein detection on circulating tumor cells from liquid biopsies using imaging mass cytometry. *Converg. Sci. Phys. Oncol.*, 4.
- GETAHUN, H., MATTEELLI, A., CHAISSON, R. E. & RAVIGLIONE, M. 2015. Latent Mycobacterium tuberculosis infection. *N. Engl. J. Med.*, 372, 2127-35.
- GIDEON, H. P., HUGHES, T. K., TZOUANAS, C. N., WADSWORTH, M. H., 2ND, TU, A. A., GIERAHN, T. M., PETERS, J. M., HOPKINS, F. F., WEI, J. R., KUMMERLOWE, C., GRANT, N. L., NARGAN, K., PHUAH, J. Y., BORISH, H. J., MAIELLO, P., WHITE, A. G., WINCHELL, C. G., NYQUIST, S. K., GANCHUA, S. K. C., MYERS, A., PATEL, K. V., AMEEL, C. L., COCHRAN, C. T., IBRAHIM, S., TOMKO, J. A., FRYE, L. J., ROSENBERG, J. M., SHIH, A., CHAO, M., KLEIN, E., SCANGA, C. A., ORDOVAS-MONTANES, J., BERGER, B., MATTILA, J. T., MADANSEIN, R., LOVE, J. C., LIN, P. L., LESLIE, A., BEHAR, S. M., BRYSON, B., FLYNN, J. L., FORTUNE, S. M. & SHALEK, A. K. 2022. Multimodal profiling of lung granulomas in macaques reveals cellular correlates of tuberculosis control. *Immunity*, 55, 827-846.e10.
- GIDEON, H. P., PHUAH, J., MYERS, A. J., BRYSON, B. D., RODGERS, M. A., COLEMAN, M. T., MAIELLO, P., RUTLEDGE, T., MARINO, S., FORTUNE, S. M., KIRSCHNER, D. E., LIN, P. L. & FLYNN, J. L. 2015. Variability in tuberculosis granuloma T cell responses exists, but a balance of pro- and anti-inflammatory cytokines is associated with sterilization. *PLoS Pathog.*, 11, e1004603.
- GIESEN, C., WANG, H. A., SCHAPIRO, D., ZIVANOVIC, N., JACOBS, A., HATTENDORF, B., SCHÜFFLER, P. J., GROLIMUND, D., BUHMANN, J. M., BRANDT, S., VARGA, Z., WILD, P. J., GÜNTHER, D. & BODENMILLER, B. 2014. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods*, 11, 417-22.
- GILADI, A. & AMIT, I. 2017. Immunology, one cell at a time. *Nature*, 547, 27-29.

- GINHOUX, F., YALIN, A., DUTERTRE, C. A. & AMIT, I. 2022. Single-cell immunology: Past, present, and future. *Immunity*, 55, 393-404.
- GODEC, J., TAN, Y., LIBERZON, A., TAMAYO, P., BHATTACHARYA, S., BUTTE, A. J., MESIROV, J. P. & HAINING, W. N. 2016. Compendium of Immune Signatures Identifies Conserved and Species-Specific Biology in Response to Inflammation. *Immunity*, 44, 194-206.
- GOLD, M. C., CERRI, S., SMYK-PEARSON, S., CANSLER, M. E., VOGT, T. M., DELEPINE, J., WINATA, E., SWARBRICK, G. M., CHUA, W. J., YU, Y. Y., LANTZ, O., COOK, M. S., NULL, M. D., JACOBY, D. B., HARRIFF, M. J., LEWINSOHN, D. A., HANSEN, T. H. & LEWINSOHN, D. M. 2010. Human mucosal associated invariant T cells detect bacterially infected cells. *PLoS Biol.*, 8, e1000407.
- GOLUB, T. R., SLONIM, D. K., TAMAYO, P., HUARD, C., GAASENBEEK, M., MESIROV, J. P., COLLIER, H., LOH, M. L., DOWNING, J. R., CALIGIURI, M. A., BLOOMFIELD, C. D. & LANDER, E. S. 1999. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, 286, 531-7.
- GONZALEZ-JUARRERO, M., HATTLE, J. M., IZZO, A., JUNQUEIRA-KIPNIS, A. P., SHIM, T. S., TRAPNELL, B. C., COOPER, A. M. & ORME, I. M. 2005. Disruption of granulocyte macrophage-colony stimulating factor production in the lungs severely affects the ability of mice to control Mycobacterium tuberculosis infection. *J. Leukoc. Biol.*, 77, 914-22.
- GORRIS, M. A. J., HALILOVIC, A., RABOLD, K., VAN DUFFELEN, A., WICKRAMASINGHE, I. N., VERWEIJ, D., WORTEL, I. M. N., TEXTOR, J. C., DE VRIES, I. J. M. & FIGDOR, C. G. 2018. Eight-Color Multiplex Immunohistochemistry for Simultaneous Detection of Multiple Immune Checkpoint Molecules within the Tumor Microenvironment. *J. Immunol.*, 200, 347.
- GRABOVSKA, Y., MACKAY, A., O'HARE, P., CROSIER, S., FINETTI, M., SCHWALBE, E. C., PICKLES, J. C., FAIRCHILD, A. R., AVERY, A., COCKLE, J., HILL, R., LINDSEY, J., HICKS, D., KRISTIANSEN, M., CHALKER, J., ANDERSON, J., HARGRAVE, D., JACQUES, T. S., STRAATHOF, K., BAILEY, S., JONES, C., CLIFFORD, S. C. & WILLIAMSON, D. 2020. Pediatric pan-central nervous system tumor analysis of immune-cell infiltration identifies correlates of antitumor immunity. *Nat. Commun.*, 11, 4324.
- GRANIER, C., DARIANE, C., COMBE, P., VERKARRE, V., URIEN, S., BADOUAL, C., ROUSSEL, H., MANDAVIT, M., RAVEL, P., SIBONY, M., BIARD, L., RADULESCU, C., VINATIER, E., BENHAMOUDA, N., PEYROMAURE, M., OUDARD, S., MÉJEAN, A., TIMSIT, M.-O., GEY, A. & TARTOUR, E. 2017. Tim-3 Expression on Tumor-Infiltrating PD-1+CD8+ T Cells Correlates with Poor Clinical Outcome in Renal Cell Carcinoma. *Cancer Res.*, 77, 1075-1082.
- GRANT, N. L., MAIELLO, P., KLEIN, E., LIN, P. L., BORISH, H. J., TOMKO, J., FRYE, L. J., WHITE, A. G., KIRSCHNER, D. E., MATTILA, J. T. & FLYNN, J. L. 2022. T cell transcription factor expression evolves over time in granulomas from Mycobacterium tuberculosis-infected cynomolgus macaques. *Cell Rep.*, 39, 110826.
- GROSSET, J. 2003. Mycobacterium tuberculosis in the extracellular compartment: an underestimated adversary. *Antimicrob. Agents Chemother.*, 47, 833-6.
- GUALANO, G., MENCARINI, P., LAURIA, F. N., PALMIERI, F., MFINANGA, S., MWABA, P., CHAKAYA, J., ZUMLA, A. & IPPOLITO, G. 2019. Tuberculin skin test – Outdated or still useful for Latent TB infection screening? *Int. J. Infect. Dis.*, 80, S20-S22.

- GUPTA, U. D. & KATOCH, V. M. 2005. Animal models of tuberculosis. *Tuberculosis*, 85, 277-293.
- HALFON, S., FORD, J., FOSTER, J., DOWLING, L., LUCIAN, L., STERLING, M., XU, Y., WEISS, M., IKEDA, M., LIGGETT, D., HELMS, A., CAUX, C., LEBECQUE, S., HANNUM, C., MENON, S., MCCLANAHAN, T., GORMAN, D. & ZURAWSKI, G. 1998. Leukocystatin, a new Class II cystatin expressed selectively by hematopoietic cells. *J. Biol. Chem.*, 273, 16400-8.
- HAMILTON, G., COLBERT, J. D., SCHUETTELKOPF, A. W. & WATTS, C. 2008. Cystatin F is a cathepsin C-directed protease inhibitor regulated by proteolysis. *EMBO J.*, 27, 499-508.
- HAN, W., CHEUNG, A. M., YAFFE, M. J. & MARTEL, A. L. 2022. Cell segmentation for immunofluorescence multiplexed images using two-stage domain adaptation and weakly labeled data for pre-training. *Sci. Rep.*, 12, 4399.
- HARTMANN, F. J. & BENDALL, S. C. 2020. Immune monitoring using mass cytometry and related high-dimensional imaging approaches. *Nat. Rev. Rheumatol.*, 16, 87-99.
- HASHIMSHONY, T., WAGNER, F., SHER, N. & YANAI, I. 2012. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep.*, 2, 666-73.
- HAYDAY, A. C. & PEAKMAN, M. 2008. The habitual, diverse and surmountable obstacles to human immunology research. *Nat. Immunol.*, 9, 575-80.
- HAYNES, W. A., VALLANIA, F., LIU, C., BONGEN, E., TOMCZAK, A., ANDRES-TERRÈ, M., LOFGREN, S., TAM, A., DEISSEROTH, C. A., LI, M. D., SWEENEY, T. E. & KHATRI, P. 2017. EMPOWERING MULTI-COHORT GENE EXPRESSION ANALYSIS TO INCREASE REPRODUCIBILITY. *Pac. Symp. Biocomput.*, 22, 144-153.
- HENG, T. S. P., PAINTER, M. W., ELPEK, K., LUKACS-KORNEK, V., MAUERMANN, N., TURLEY, S. J., KOLLER, D., KIM, F. S., WAGERS, A. J., ASINOVSKI, N., DAVIS, S., FASSETT, M., FEUERER, M., GRAY, D. H. D., HAXHINASTO, S., HILL, J. A., HYATT, G., LAPLACE, C., LEATHERBEE, K., MATHIS, D., BENOIST, C., JIANU, R., LAIDLAW, D. H., BEST, J. A., KNELL, J., GOLDRATH, A. W., JARJOURA, J., SUN, J. C., ZHU, Y., LANIER, L. L., ERGUN, A., LI, Z., COLLINS, J. J., SHINTON, S. A., HARDY, R. R., FRIEDLINE, R., SYLVIA, K., KANG, J. & THE IMMUNOLOGICAL GENOME PROJECT, C. 2008. The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.*, 9, 1091-1094.
- HENN, A. D., WU, S., QIU, X., RUDA, M., STOVER, M., YANG, H., LIU, Z., WELLE, S. L., HOLDEN-WILTSE, J., WU, H. & ZAND, M. S. 2013. High-Resolution Temporal Response Patterns to Influenza Vaccine Reveal a Distinct Human Plasma Cell Gene Signature. *Sci. Rep.*, 3, 2327.
- HERBERG, J. A., KAFOROU, M., GORMLEY, S., SUMNER, E. R., PATEL, S., JONES, K. D., PAULUS, S., FINK, C., MARTINON-TORRES, F., MONTANA, G., WRIGHT, V. J. & LEVIN, M. 2013. Transcriptomic profiling in childhood H1N1/09 influenza reveals reduced expression of protein synthesis genes. *J. Infect. Dis.*, 208, 1664-8.
- HICKEY, J. W., NEUMANN, E. K., RADTKE, A. J., CAMARILLO, J. M., BEUSCHEL, R. T., ALBANESE, A., MCDONOUGH, E., HATLER, J., WIBLIN, A. E., FISHER, J., CROTEAU, J., SMALL, E. C., SOOD, A., CAPRIOLI, R. M., ANGELO, R. M., NOLAN, G. P., CHUNG, K., HEWITT, S. M., GERMAIN, R. N., SPRAGGINS, J. M., LUNDBERG, E., SNYDER, M. P., KELLEHER, N. L. & SAKA, S. K. 2022. Spatial mapping of protein composition and tissue organization: a primer for multiplexed antibody-based imaging. *Nat. Methods*, 19, 284-295.

- HIGGS, B. W., ZHU, W., RICHMAN, L., FIORENTINO, D. F., GREENBERG, S. A., JALLAL, B. & YAO, Y. 2012. Identification of activated cytokine pathways in the blood of systemic lupus erythematosus, myositis, rheumatoid arthritis, and scleroderma patients. *Int. J. Rheum. Dis.*, 15, 25-35.
- HIJMANS, R. J. & ETTEEN, J. V. 2012. *raster: Geographic analysis and modeling with raster data* [Online]. Available: <http://CRAN.R-project.org/package=raster> [Accessed].
- HO, S. W. L., ANG, C.-L., DING, C. S. L., BARKHAM, T. & TEOH, L. C. 2015. Necrotizing Fasciitis Caused by *Cryptococcus gattii*. *Am. J. Orthop. (Belle Mead NJ)*, 44, E517-22.
- HOFMAN, P., BADOUAL, C., HENDERSON, F., BERLAND, L., HAMILA, M., LONG-MIRA, E., LASSALLE, S., ROUSSEL, H., HOFMAN, V., TARTOUR, E. & ILIÉ, M. 2019. Multiplexed Immunohistochemistry for Molecular and Immune Profiling in Lung Cancer-Just About Ready for Prime-Time? *Cancers (Basel)*, 11.
- HOLLANDI, R., MOSHKOV, N., PAAVOLAINEN, L., TASNADI, E., PICCININI, F. & HORVATH, P. 2022. Nucleus segmentation: towards automated solutions. *Trends Cell Biol.*, 32, 295-310.
- HOLLY, M. S. A., LIN, P. L. & FLYNN, J. L. 2005. Tumor Necrosis Factor and Chemokine Interactions in the Formation and Maintenance of Granulomas in Tuberculosis. *Clin. Infect. Dis.*, 41, S189-S193.
- HOOD, L., STEINMETZ, M. & MALISSEN, B. 1983. Genes of the major histocompatibility complex of the mouse. *Annu. Rev. Immunol.*, 1, 529-68.
- HOUBEN, D., DEMANGEL, C., VAN INGEN, J., PEREZ, J., BALDEÓN, L., ABDALLAH, A. M., CALEECHURN, L., BOTTAI, D., VAN ZON, M., DE PUNDER, K., VAN DER LAAN, T., KANT, A., BOSSERS-DE VRIES, R., WILLEMSSEN, P., BITTER, W., VAN SOOLINGEN, D., BROSCHE, R., VAN DER WEL, N. & PETERS, P. J. 2012. ESX-1-mediated translocation to the cytosol controls virulence of mycobacteria. *Cell. Microbiol.*, 14, 1287-98.
- HRUZ, T., LAULE, O., SZABO, G., WESSENDORP, F., BLEULER, S., OERTLE, L., WIDMAYER, P., GRUISSEM, W. & ZIMMERMANN, P. 2008. Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv. Bioinformatics*, 2008, 420747.
- HU, Z., JIANG, K., FRANK, M. B., CHEN, Y. & JARVIS, J. N. 2016. Complexity and Specificity of the Neutrophil Transcriptomes in Juvenile Idiopathic Arthritis. *Sci. Rep.*, 6, 27453.
- HUANG, S. S. Y., RINCHAI, D., TOUFIQ, M., KABEER, B. S. A., ROELANDS, J., HENDRICKX, W., BOUGHORBEL, S., BEDOGNETTI, D., VAN PANHUYS, N., CHAUSSABEL, D. & GARAND, M. 2022. Transcriptomic profile investigations highlight a putative role for NUDT16 in sepsis. *J. Cell. Mol. Med.*, 26, 1714-1721.
- HUANG, S. S. Y., TOUFIQ, M., SARAIVA, L. R., VAN PANHUYS, N., CHAUSSABEL, D. & GARAND, M. 2021. Transcriptome and Literature Mining Highlight the Differential Expression of ERLIN1 in Immune Cells during Sepsis. *Biology*, 10.
- HUDSON, T. J., ANDERSON, W., ARETZ, A., BARKER, A. D., BELL, C., BERNABÉ, R. R., BHAN, M. K., CALVO, F., EEROLA, I., GERHARD, D. S., GUTTMACHER, A., GUYER, M., HEMSLEY, F. M., JENNINGS, J. L., KERR, D., KLATT, P., KOLAR, P., KUSUDA, J., LANE, D. P., LAPLACE, F., LU, Y., NETTEKOVEN, G., OZENBERGER, B., PETERSON, J., RAO, T. S., REMACLE, J., SCHAFER, A. J., SHIBATA, T., STRATTON, M. R., VOCKLEY, J. G., WATANABE, K., YANG, H., YUEN, M. M. F., KNOPPERS, B. M., BOBROW, M., CAMBON-THOMSEN, A., DRESSLER, L. G., DYKE, S. O. M., JOLY, Y., KATO, K., KENNEDY, K. L.,

- NICOLÁS, P., PARKER, M. J., RIAL-SEBBAG, E., ROMEO-CASABONA, C. M., SHAW, K. M., WALLACE, S., WIESNER, G. L., ZEPS, N., LICHTER, P., BIANKIN, A. V., CHABANNON, C., CHIN, L., CLÉMENT, B., ALAVA, E. D., DEGOS, F., FERGUSON, M. L., GEARY, P., HAYES, D. N., JOHNS, A. L., KASPRZYK, A., NAKAGAWA, H., PENNY, R., PIRIS, M. A., SARIN, R., SCARPA, A., DE VIJVER, M. V., FUTREAL, P. A., ABURATANI, H., BAYÉS, M., BOWTELL, D. D. L., CAMPBELL, P. J., ESTIVILL, X., GRIMMOND, S. M., GUT, I., HIRST, M., LÓPEZ-OTÍN, C., MAJUMDER, P., MARRA, M., MCPHERSON, J. D., NING, Z., PUENTE, X. S., RUAN, Y., STUNNENBERG, H. G., SWERDLOW, H., VELCULESCU, V. E., WILSON, R. K., XUE, H. H., YANG, L., SPELLMAN, P. T., BADER, G. D., BOUTROS, P. C., FLICEK, P., GETZ, G., GUIGÓ, R., GUO, G., HAUSSLER, D., HEATH, S., HUBBARD, T. J., JIANG, T., et al. 2010. International network of cancer genome projects. *Nature*, 464, 993-998.
- HUNTER, R. L. 2011a. On the pathogenesis of post primary tuberculosis: The role of bronchial obstruction in the pathogenesis of cavities. *Tuberculosis*, 91, S6-S10.
- HUNTER, R. L. 2011b. Pathology of post primary tuberculosis of the lung: An illustrated critical review. *Tuberculosis*, 91, 497-509.
- HUNTER, R. L. 2015. Tuberculosis as a three-act play: A new paradigm for the pathogenesis of pulmonary tuberculosis. *Tuberculosis*, 97, 8-17.
- HUNTER, R. L., ACTOR, J. K., HWANG, S. A., KAREV, V. & JAGANNATH, C. 2014. Pathogenesis of post primary tuberculosis: Immunity and hypersensitivity in the development of cavities. *Ann. Clin. Lab. Sci.*, 44, 365-387.
- HUNTER, R. L., WELSH, K. J., RISIN, S. A. & ACTOR, J. K. 2011. Immunopathology of postprimary tuberculosis: Increased T-regulatory cells and DEC-205-positive foamy macrophages in cavitary lesions. *Clin. Dev. Immunol.*, 2011, 1-9.
- IWASAKI, A. & MEDZHITOV, R. 2004. Toll-like receptor control of the adaptive immune responses. *Nat. Immunol.*, 5, 987-95.
- J LIBARDO, M. D., BOSHOFF, H. I. M. & BARRY, C. E. 2018. The present state of the tuberculosis drug development pipeline. *Curr. Opin. Pharmacol.*, 42, 81-94.
- JACOBSEN, M., REPSILBER, D., GUTSCHMIDT, A., NEHER, A., FELDMANN, K., MOLLENKOPF, H. J., ZIEGLER, A. & KAUFMANN, S. H. 2007. Candidate biomarkers for discrimination between infection and disease caused by *Mycobacterium tuberculosis*. *J. Mol. Med. (Berl.)*, 85, 613-21.
- JAITIN, D. A., KENIGSBERG, E., KEREN-SHAUL, H., ELEFANT, N., PAUL, F., ZARETSKY, I., MILDNER, A., COHEN, N., JUNG, S., TANAY, A. & AMIT, I. 2014. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science*, 343, 776-9.
- JAMES, D. G. 2000. A clinicopathological classification of granulomatous disorders. *Postgrad. Med. J.*, 76, 457-65.
- JAMESON, S. C. & MASOPUST, D. 2018. What is the predictive value of animal models for vaccine efficacy in humans? Reevaluating the potential of mouse models for the human immune system. *Cold Spring Harb. Perspect. Biol.*, 10.
- JENNEWEIN, M. F., GOLDFARB, I., DOLATSHAHI, S., COSGROVE, C., NOELETTE, F. J., KRYKBAEVA, M., DAS, J., SARKAR, A., GORMAN, M. J., FISCHINGER, S., BOUDREAU, C. M., BROWN, J., COOPERRIDER, J. H., ANEJA, J., SUSCOVICH, T. J., GRAHAM, B. S., LAUER, G. M., GOETGHEBUER, T., MARCHANT, A., LAUFFENBURGER, D., KIM, A. Y., RILEY, L. E. & ALTER, G. 2019. Fc Glycan-Mediated Regulation of Placental Antibody Transfer. *Cell*, 178, 202-215.e14.

- JOHNSON, D. B., BORDEAUX, J., KIM, J. Y., VAUPEL, C., RIMM, D. L., HO, T. H., JOSEPH, R. W., DAUD, A. I., CONRY, R. M., GAUGHAN, E. M., HERNANDEZ-AYA, L. F., DIMOU, A., FUNCHAIN, P., SMITHY, J., WITTE, J. S., MCKEE, S. B., KO, J., WRANGLE, J. M., DABBAS, B., TANGRI, S., LAMEH, J., HALL, J., MARKOWITZ, J., BALKO, J. M. & DAKAPPAGARI, N. 2018. Quantitative Spatial Profiling of PD-1/PD-L1 Interaction and HLA-DR/IDO-1 Predicts Improved Outcomes of Anti-PD-1 Therapies in Metastatic Melanoma. *Clin. Cancer Res.*, 24, 5250-5260.
- JOSHI, S. & YU, D. 2017. Chapter 8 - Immunofluorescence. In: JALALI, M., SALDANHA, F. Y. L. & JALALI, M. (eds.) *Basic Science Methods for Clinical Researchers*. Boston: Academic Press.
- JUFFERMANS, N. P., FLORQUIN, S., CAMOGLIO, L., VERBON, A., KOLK, A. H., SPEELMAN, P., VAN DEVENTER, S. J. H. & VAN DER POLL, T. 2000. Interleukin-1 signaling is essential for host defense during murine pulmonary tuberculosis. *J. Infect. Dis.*, 182, 902-908.
- JUSS, J. K., HOUSE, D., AMOUR, A., BEGG, M., HERRE, J., STORISTEANU, D. M., HOENDERDOS, K., BRADLEY, G., LENNON, M., SUMMERS, C., HESSEL, E. M., CONDLIFFE, A. & CHILVERS, E. R. 2016. Acute Respiratory Distress Syndrome Neutrophils Have a Distinct Phenotype and Are Resistant to Phosphoinositide 3-Kinase Inhibition. *Am. J. Respir. Crit. Care Med.*, 194, 961-973.
- KACZOROWSKI, K. J., SHEKHAR, K., NKULIKIYIMFURA, D., DEKKER, C. L., MAECKER, H., DAVIS, M. M., CHAKRABORTY, A. K. & BRODIN, P. 2017. Continuous immunotypes describe human immune variation and predict diverse responses. *Proc. Natl. Acad. Sci. U. S. A.*, 114, E6097-e6106.
- KAFOROU, M., WRIGHT, V. J., ONI, T., FRENCH, N., ANDERSON, S. T., BANGANI, N., BANWELL, C. M., BRENT, A. J., CRAMPIN, A. C., DOCKRELL, H. M., ELEY, B., HEYDERMAN, R. S., HIBBERD, M. L., KERN, F., LANGFORD, P. R., LING, L., MENDELSON, M., OTTENHOFF, T. H., ZGAMBO, F., WILKINSON, R. J., COIN, L. J. & LEVIN, M. 2013. Detection of tuberculosis in HIV-infected and -uninfected African adults using whole blood RNA expression signatures: a case-control study. *PLoS Med.*, 10, e1001538.
- KAGINA, B. M., ABEL, B., SCRIBA, T. J., HUGHES, E. J., KEYSER, A., SOARES, A., GAMIELDIEN, H., SIDIBANA, M., HATHERILL, M., GELDERBLOEM, S., MAHOMED, H., HAWKRIDGE, A., HUSSEY, G., KAPLAN, G. & HANEKOM, W. A. 2010. Specific T cell frequency and cytokine expression profile do not correlate with protection against tuberculosis after bacillus Calmette-Guérin vaccination of newborns. *Am. J. Respir. Crit. Care Med.*, 182, 1073-9.
- KAIZER, E. C., GLASER, C. L., CHAUSSABEL, D., BANCHEREAU, J., PASCUAL, V. & WHITE, P. C. 2007. Gene expression in peripheral blood mononuclear cells from children with diabetes. *J. Clin. Endocrinol. Metab.*, 92, 3705-11.
- KASSAMBARA, A. & MUNDT, F. 2021. *Factoextra: Extract and Visualize the Results of Multivariate Data Analyses* [Online]. Available: <https://CRAN.R-project.org/package=factoextra> [Accessed].
- KAUSHAL, D., MEHRA, S., DIDIER, P. J. & LACKNER, A. A. 2012. The non-human primate model of tuberculosis. *J. Med. Primatol.*, 41, 191-201.
- KEANE, J., GERSHON, S., WISE, R. P., MIRABILE-LEVENS, E., KASZNICA, J., SCHWIETERMAN, W. D., SIEGEL, J. N. & BRAUN, M. M. 2001. Tuberculosis associated with infliximab, a tumor necrosis factor alpha-neutralizing agent. *N. Engl. J. Med.*, 345, 1098-104.

- KEREN, L., BOSSE, M., MARQUEZ, D., ANGOSHTARI, R., JAIN, S., VARMA, S., YANG, S. R., KURIAN, A., VAN VALEN, D., WEST, R., BENDALL, S. C. & ANGELO, M. 2018. A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell*, 174, 1373-1387.e19.
- KHAENAM, P., RINCHAI, D., ALTMAN, M. C., CHICHE, L., BUDDHISA, S., KEWCHAROENWONG, C., SUWANNASAEN, D., MASON, M., WHALEN, E., PRESNELL, S., SUSANENGAT, W., O'BRIEN, K., NGUYEN, Q.-A., GERSUK, V., LINSLEY, P. S., LERTMEMONGKOLCHAI, G. & CHAUSSABEL, D. 2014. A transcriptomic reporter assay employing neutrophils to measure immunogenic activity of septic patients' plasma. *J. Transl. Med.*, 12, 65-65.
- KIM, C. W., OH, J. E. & LEE, H. K. 2021. Single Cell Transcriptomic Re-analysis of Immune Cells in Bronchoalveolar Lavage Fluids Reveals the Correlation of B Cell Characteristics and Disease Severity of Patients with SARS-CoV-2 Infection. *Immune Netw.*, 21, e10.
- KIM, M. J., WAINWRIGHT, H. C., LOCKETZ, M., BEKKER, L. G., WALTHER, G. B., DITTRICH, C., VISSER, A., WANG, W., HSU, F. F., WIEHART, U., TSENOVA, L., KAPLAN, G. & RUSSELL, D. G. 2010. Caseation of human tuberculosis granulomas correlates with elevated host lipid metabolism. *EMBO Mol. Med.*, 2, 258-74.
- KLEIN, A. M., MAZUTIS, L., AKARTUNA, I., TALLAPRAGADA, N., VERES, A., LI, V., PESHKIN, L., WEITZ, D. A. & KIRSCHNER, M. W. 2015. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*, 161, 1187-1201.
- KLEINMAN, G. M., ZAGZAG, D. & MILLER, D. C. 1994. Diagnostic use of immunohistochemistry in neuropathology. *Neurosurg. Clin. N. Am.*, 5, 97-126.
- KOLESNIKOV, N., HASTINGS, E., KEAYS, M., MELNICHUK, O., TANG, Y. A., WILLIAMS, E., DYLAG, M., KURBATOVA, N., BRANDIZI, M., BURDETT, T., MEGY, K., PILICHEVA, E., RUSTICI, G., TIKHONOV, A., PARKINSON, H., PETRYSZAK, R., SARKANS, U. & BRAZMA, A. 2015. ArrayExpress update--simplifying data submissions. *Nucleic Acids Res.*, 43, D1113-6.
- KOROTKEVICH, G., SUKHOV, V., BUDIN, N., SHPAK, B., ARTYOMOV, M. N. & SERGUSHICHEV, A. 2021. Fast gene set enrichment analysis. *bioRxiv*, 060012.
- KOS, J., NANUT, M., PRUNK, M., SABOTIČ, J., DAUTOVIĆ, E. & JEWETT, A. 2018. Cystatin F as a regulator of immune cell cytotoxicity. *Cancer Immunol. Immunother.*, 67, 1931-1938.
- KUKSIN, M., MOREL, D., AGLAVE, M., DANLOS, F.-X., MARABELLE, A., ZINOVYEV, A., GAUTHERET, D. & VERLINGUE, L. 2021. Applications of single-cell and bulk RNA sequencing in onco-immunology. *Eur. J. Cancer*, 149, 193-210.
- KULESHOV, M. V., JONES, M. R., ROUILLARD, A. D., FERNANDEZ, N. F., DUAN, Q., WANG, Z., KOPLEV, S., JENKINS, S. L., JAGODNIK, K. M., LACHMANN, A., MCDERMOTT, M. G., MONTEIRO, C. D., GUNDERSEN, G. W. & MA'AYAN, A. 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.*, 44, W90-7.
- LADEL, C. H., BLUM, C., DREHER, A., REIFENBERG, K., KOPF, M. & KAUFMANN, S. H. 1997. Lethal tuberculosis in interleukin-6-deficient mutant mice. *Infect. Immun.*, 65, 4843-9.
- LAMPS, L. W. & SCOTT, M. A. 2004. Cat-scratch disease: historic, clinical, and pathologic perspectives. *Pathol. Patterns Rev.*, 121, S71-S80.

- LANGFELDER, P. & HORVATH, S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, 9, 559.
- LANOIX, J.-P., LENAERTS, A. J. & NUERMBERGER, E. L. 2015. Heterogeneous disease progression and treatment response in a C3HeB/FeJ mouse model of tuberculosis. *Dis. Model. Mech.*, 8, 603-610.
- LAWN, S. D., BUTERA, S. T. & SHINNICK, T. M. 2002. Tuberculosis unleashed: the impact of human immunodeficiency virus infection on the host granulomatous response to *Mycobacterium tuberculosis*. *Microbes Infect.*, 4, 635-46.
- LEE, H., QUEK, C., SILVA, I., TASKER, A., BATTEN, M., RIZOS, H., LIM, S. Y., NUR GIDE, T., SHANG, P., ATTRILL, G. H., MADORE, J., EDWARDS, J., CARLINO, M. S., GUMINSKI, A., SAW, R. P. M., THOMPSON, J. F., FERGUSON, P. M., PALENDIRA, U., MENZIES, A. M., LONG, G. V., SCOLYER, R. A. & WILMOTT, J. S. 2019. Integrated molecular and immunophenotypic analysis of NK cells in anti-PD-1 treated metastatic melanoma patients. *Oncoimmunology*, 8, e1537581.
- LEE, H. M., SUGINO, H., AOKI, C. & NISHIMOTO, N. 2011. Underexpression of mitochondrial-DNA encoded ATP synthesis-related genes and DNA repair genes in systemic lupus erythematosus. *Arthritis Res. Ther.*, 13, R63.
- LEE, J., HYEON, D. Y. & HWANG, D. 2020. Single-cell multiomics: technologies and data analysis methods. *Exp. Mol. Med.*, 52, 1428-1442.
- LEE, S. W., LEE, H. Y., KANG, S. W., KIM, M. J., LEE, Y. J., SUNG, C. O. & KIM, Y. M. 2021. Application of Immunoprofiling Using Multiplexed Immunofluorescence Staining Identifies the Prognosis of Patients with High-Grade Serous Ovarian Cancer. *Int. J. Mol. Sci.*, 22.
- LEMAITRE, B., NICOLAS, E., MICHAUT, L., REICHHART, J. M. & HOFFMANN, J. A. 2012. Pillars article: the dorsoventral regulatory gene cassette *spätzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. *Cell*. 1996. 86: 973-983. *J. Immunol.*, 188, 5210-20.
- LEONG, S., ZHAO, Y., JOSEPH, N. M., HOCHBERG, N. S., SARKAR, S., PLESKUNAS, J., HOM, D., LAKSHMINARAYANAN, S., HORSBURGH, C. R., JR., ROY, G., ELLNER, J. J., JOHNSON, W. E. & SALGAME, P. 2018. Existing blood transcriptional classifiers accurately discriminate active tuberculosis from latent infection in individuals from south India. *Tuberculosis*, 109, 41-51.
- LEWIS, S. M., ASSELIN-LABAT, M.-L., NGUYEN, Q., BERTHELET, J., TAN, X., WIMMER, V. C., MERINO, D., ROGERS, K. L. & NAIK, S. H. 2021. Spatial omics and multiplexed imaging to explore cancer biology. *Nat. Methods*, 18, 997-1012.
- LI, B., LI, T., PIGNON, J. C., WANG, B., WANG, J., SHUKLA, S. A., DOU, R., CHEN, Q., HODI, F. S., CHOUEIRI, T. K., WU, C., HACHEN, N., SIGNORETTI, S., LIU, J. S. & LIU, X. S. 2016. Landscape of tumor-infiltrating T cell repertoire of human cancers. *Nat. Genet.*, 48, 725-32.
- LI, H., VAN DER LEUN, A. M., YOFI, I., LUBLING, Y., GELBARD-SOLODKIN, D., VAN AKKOOI, A. C. J., VAN DEN BRABER, M., ROZEMAN, E. A., HAANEN, J., BLANK, C. U., HORLINGS, H. M., DAVID, E., BARAN, Y., BERCOVICH, A., LIFSHITZ, A., SCHUMACHER, T. N., TANAY, A. & AMIT, I. 2019. Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. *Cell*, 176, 775-789.e18.
- LI, S., ROUPHAEL, N., DURAISINGHAM, S., ROMERO-STEINER, S., PRESNELL, S., DAVIS, C., SCHMIDT, D. S., JOHNSON, S. E., MILTON, A., RAJAM, G., KASTURI, S., CARLONE, G. M., QUINN, C., CHAUSSABEL, D., PALUCKA, A. K., MULLIGAN, M. J., AHMED, R., STEPHENS, D. S., NAKAYA, H. I. &

- PULENDRAN, B. 2014. Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. *Nat. Immunol.*, 15, 195-204.
- LI, X. & WANG, C.-Y. 2021. From bulk, single-cell to spatial RNA sequencing. *Int. J. Oral Sci.*, 13, 36.
- LI, Y. & ANDRADE, J. 2017. DEApp: an interactive web interface for differential expression analysis of next generation sequence data. *Source Code Biol. Med.*, 12, 2.
- LIANG, J., LI, N., ZHANG, Y., HOU, C., YANG, X., SHIMIZU, T., WANG, X., IKENAKA, K., FAN, K. & MA, J. 2016. Disinhibition of Cathepsin C Caused by Cystatin F Deficiency Aggravates the Demyelination in a Cuprizone Model. *Front. Mol. Neurosci.*, 9, 152.
- LIAO, J., LU, X., SHAO, X., ZHU, L. & FAN, X. 2021. Uncovering an Organ's Molecular Architecture at Single-Cell Resolution by Spatially Resolved Transcriptomics. *Trends Biotechnol.*, 39, 43-58.
- LIBERZON, A., SUBRAMANIAN, A., PINCHBACK, R., THORVALDSDÓTTIR, H., TAMAYO, P. & MESIROV, J. P. 2011. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*, 27, 1739-1740.
- LIN, F. & LIU, H. 2014. Immunohistochemistry in Undifferentiated Neoplasm/Tumor of Uncertain Origin. *Arch. Pathol. Lab. Med.*, 138, 1583-1610.
- LIN, P. L., COLEMAN, T., CARNEY, J. P., LOPRESTI, B. J., TOMKO, J., FILLMORE, D., DARTOIS, V., SCANGA, C., FRYE, L. J., JANSSEN, C., KLEIN, E., BARRY, C. E., 3RD & FLYNN, J. L. 2013. Radiologic Responses in Cynomolgus Macaques for Assessing Tuberculosis Chemotherapy Regimens. *Antimicrob. Agents Chemother.*, 57, 4237-4244.
- LIN, P. L. & FLYNN, J. L. 2018. The End of the Binary Era: Revisiting the Spectrum of Tuberculosis. *J. Immunol.*, 201, 2541-2548.
- LIN, P. L., FORD, C. B., COLEMAN, M. T., MYERS, A. J., GAWANDE, R., IOERGER, T., SACCHETTINI, J., FORTUNE, S. M. & FLYNN, J. L. 2014. Sterilization of granulomas is common in active and latent tuberculosis despite within-host variability in bacterial killing. *Nat. Med.*, 20, 75-79.
- LINSLEY, P. S., SPEAKE, C., WHALEN, E. & CHAUSSABEL, D. 2014. Copy number loss of the interferon gene cluster in melanomas is linked to reduced T cell infiltrate and poor patient prognosis. *PLoS One*, 9, e109760.
- LOFGREN, S., HINCHCLIFF, M., CARNS, M., WOOD, T., AREN, K., ARROYO, E., CHEUNG, P., KUO, A., VALENZUELA, A., HAEMEL, A., WOLTERS, P. J., GORDON, J., SPIERA, R., ASSASSI, S., BOIN, F., CHUNG, L., FIORENTINO, D., UTZ, P. J., WHITFIELD, M. L. & KHATRI, P. 2016. Integrated, multicohort analysis of systemic sclerosis identifies robust transcriptional signature of disease severity. *JCI Insight*, 1, e89073.
- LÖNNROTH, K., CASTRO, K. G., CHAKAYA, J. M., CHAUHAN, L. S., FLOYD, K., GLAZIOU, P. & RAVIGLIONE, M. C. 2010. Tuberculosis control and elimination 2010-50: cure, care, and social development. *Lancet*, 375, 1814-29.
- 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.
- LOWE, R., SHIRLEY, N., BLEACKLEY, M., DOLAN, S. & SHAFEE, T. 2017. Transcriptomics technologies. *PLoS Comput. Biol.*, 13, e1005457.
- LU, L. L., CHUNG, A. W., ROSEBROCK, T. R., GHEBREMICHAEL, M., YU, W. H., GRACE, P. S., SCHOEN, M. K., TAFESSE, F., MARTIN, C., LEUNG, V., MAHAN, A. E., SIPS, M., KUMAR, M. P., TEDESCO, J., ROBINSON, H., TKACHENKO, E., DRAGHI, M., FREEDBERG, K. J., STREECK, H., SUSCOVICH, T. J., LAUFFENBURGER, D. A., RESTREPO, B. I., DAY, C.,

- FORTUNE, S. M. & ALTER, G. 2016. A Functional Role for Antibodies in Tuberculosis. *Cell*, 167, 433-443.e14.
- LYON, S. M. & ROSSMAN, M. D. 2017. Pulmonary Tuberculosis. *Microbiol Spectr*, 5.
- MA, F., HUGHES, T. K., TELES, R. M. B., ANDRADE, P. R., DE ANDRADE SILVA, B. J., PLAZYO, O., TSOI, L. C., DO, T., WADSWORTH, M. H., 2ND, OULEE, A., OCHOA, M. T., SARNO, E. N., IRUELA-ARISPE, M. L., KLECHEVSKY, E., BRYSON, B., SHALEK, A. K., BLOOM, B. R., GUDJONSSON, J. E., PELLEGRINI, M. & MODLIN, R. L. 2021. The cellular architecture of the antimicrobial response network in human leprosy granulomas. *Nat. Immunol.*, 22, 839-850.
- MA, J., TANAKA, K. F., SHIMIZU, T., BERNARD, C. C., KAKITA, A., TAKAHASHI, H., PFEIFFER, S. E. & IKENAKA, K. 2011. Microglial cystatin F expression is a sensitive indicator for ongoing demyelination with concurrent remyelination. *J. Neurosci. Res.*, 89, 639-49.
- MACOSKO, E. Z., BASU, A., SATIJA, R., NEMESH, J., SHEKHAR, K., GOLDMAN, M., TIROSH, I., BIALAS, A. R., KAMITAKI, N., MARTERSTECK, E. M., TROMBETTA, J. J., WEITZ, D. A., SANES, J. R., SHALEK, A. K., REGEV, A. & MCCARROLL, S. A. 2015. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*, 161, 1202-1214.
- MAECKER, H. T., MCCOY, J. P. & NUSSENBLATT, R. 2012. Standardizing immunophenotyping for the Human Immunology Project. *Nat. Rev. Immunol.*, 12, 191-200.
- MAERTZDORF, J., WEINER, J., 3RD, MOLLENKOPF, H. J., BAUER, T., PRASSE, A., MULLER-QUERNHEIM, J. & KAUFMANN, S. H. 2012. Common patterns and disease-related signatures in tuberculosis and sarcoidosis. *Proc. Natl. Acad. Sci. U. S. A.*, 109, 7853-8.
- MAGISTER, S. & KOS, J. 2013. Cystatins in immune system. *J. Cancer*, 4, 45-56.
- MAGISTER, Š., TSENG, H.-C., BUI, V. T., KOS, J. & JEWETT, A. 2015. Regulation of split anergy in natural killer cells by inhibition of cathepsins C and H and cystatin F. *Oncotarget*, 6, 22310-22327.
- MAHI, N. A., NAJAFABADI, M. F., PILARCZYK, M., KOURIL, M. & MEDVEDOVIC, M. 2019. GREIN: An Interactive Web Platform for Re-analyzing GEO RNA-seq Data. *Sci. Rep.*, 9, 7580.
- MAK, I. W., EVANIEW, N. & GHERT, M. 2014. Lost in translation: animal models and clinical trials in cancer treatment. *Am. J. Transl. Res.*, 6, 114-8.
- MAK, T. W., PENNINGER, J. M. & OHASHI, P. S. 2001. Knockout mice: a paradigm shift in modern immunology. *Nat. Rev. Immunol.*, 1, 11-19.
- MALIHI, P. D., MORIKADO, M., WELTER, L., LIU, S. T., MILLER, E. T., CADANEANU, R. M., KNUDSEN, B. S., LEWIS, M. S., CARLSSON, A., VELASCO, C. R., KOLATKAR, A., RODRIGUEZ-LEE, M., GARRAWAY, I. P., HICKS, J. & KUHN, P. 2018. Clonal diversity revealed by morphoproteomic and copy number profiles of single prostate cancer cells at diagnosis. *Converg. Sci. Phys. Oncol.*, 4.
- MANGTANI, P., ABUBAKAR, I., ARITI, C., BEYNON, R., PIMPIN, L., FINE, P. E., RODRIGUES, L. C., SMITH, P. G., LIPMAN, M., WHITING, P. F. & STERNE, J. A. 2014. Protection by BCG vaccine against tuberculosis: a systematic review of randomized controlled trials. *Clin. Infect. Dis.*, 58, 470-80.
- MANIATIS, S., PETRESCU, J. & PHATNANI, H. 2021. Spatially resolved transcriptomics and its applications in cancer. *Curr. Opin. Genet. Dev.*, 66, 70-77.

- MANJANG, K., TRIPATHI, S., YLI-HARJA, O., DEHMER, M. & EMMERT-STREIB, F. 2020. Graph-based exploitation of gene ontology using GOxploreR for scrutinizing biological significance. *Sci. Rep.*, 10, 16672.
- MANTIONE, K. J., KREAM, R. M., KUZELOVA, H., PTACEK, R., RABOCH, J., SAMUEL, J. M. & STEFANO, G. B. 2014. Comparing bioinformatic gene expression profiling methods: microarray and RNA-Seq. *Med. Sci. Monit. Basic. Res.*, 20, 138-42.
- MARTIN, C. J., CADENA, A. M., LEUNG, V. W., LIN, P. L., MAIELLO, P., HICKS, N., CHASE, M. R., FLYNN, J. L. & FORTUNE, S. M. 2017. Digitally Barcoding Mycobacterium tuberculosis Reveals In Vivo Infection Dynamics in the Macaque Model of Tuberculosis. *mBio*, 8, e00312-17.
- MARTINO, D. & ALLEN, K. 2015. Meeting the challenges of measuring human immune regulation. *J. Immunol. Methods*, 424, 1-6.
- MASOPUST, D., SIVULA, C. P. & JAMESON, S. C. 2017. Of mice, dirty mice, and men: Using mice to understand human immunology. *J. Immunol.*, 199, 383-388.
- MASUGI, Y., ABE, T., UENO, A., FUJII-NISHIMURA, Y., OJIMA, H., ENDO, Y., FUJITA, Y., KITAGO, M., SHINODA, M., KITAGAWA, Y. & SAKAMOTO, M. 2019. Characterization of spatial distribution of tumor-infiltrating CD8⁺ T cells refines their prognostic utility for pancreatic cancer survival. *Mod. Pathol.*, 32, 1495-1507.
- MATTILA, J. T., BEAINO, W., MAIELLO, P., COLEMAN, M. T., WHITE, A. G., SCANGA, C. A., FLYNN, J. L. & ANDERSON, C. J. 2017. Positron Emission Tomography Imaging of Macaques with Tuberculosis Identifies Temporal Changes in Granuloma Glucose Metabolism and Integrin $\alpha 4\beta 1$ -Expressing Immune Cells. *J. Immunol.*, 199, 806.
- MAYER-BARBER, K. D., BARBER, D. L., SHENDEROV, K., WHITE, S. D., WILSON, M. S., CHEEVER, A., KUGLER, D., HIENY, S., CASPAR, P., NÚÑEZ, G., SCHLUETER, D., FLAVELL, R. A., SUTTERWALA, F. S. & SHER, A. 2010. Caspase-1 independent IL-1 β production is critical for host resistance to mycobacterium tuberculosis and does not require TLR signaling in vivo. *J. Immunol.*, 184, 3326-30.
- MCCAFFREY, E. F., DONATO, M., KEREN, L., CHEN, Z., DELMASTRO, A., FITZPATRICK, M. B., GUPTA, S., GREENWALD, N. F., BARANSKI, A., GRAF, W., KUMAR, R., BOSSE, M., FULLAWAY, C. C., RAMDIAL, P. K., FORGÓ, E., JOJIC, V., VAN VALEN, D., MEHRA, S., KHADER, S. A., BENDALL, S. C., VAN DE RIJN, M., KALMAN, D., KAUSHAL, D., HUNTER, R. L., BANAEI, N., STEYN, A. J. C., KHATRI, P. & ANGELO, M. 2022. The immunoregulatory landscape of human tuberculosis granulomas. *Nat. Immunol.*, 23, 318-329.
- MCGUIRE, M. J., LIPSKY, P. E. & THIELE, D. L. 1993. Generation of active myeloid and lymphoid granule serine proteases requires processing by the granule thiol protease dipeptidyl peptidase I. *J. Biol. Chem.*, 268, 2458-67.
- MEDLAR, E. M. 1948. The pathogenesis of minimal pulmonary tuberculosis; a study of 1,225 necropsies in cases of sudden and unexpected death. *Am. Rev. Tuberc.*, 58, 583-611.
- MEDLAR, E. M. 1955. The behavior of pulmonary tuberculous lesions; a pathological study. *Am. Rev. Tuberc.*, 71, 1-244.
- MEJIAS, A., DIMO, B., SUAREZ, N. M., GARCIA, C., SUAREZ-ARRABAL, M. C., JARTTI, T., BLANKENSHIP, D., JORDAN-VILLEGAS, A., ARDURA, M. I., XU, Z., BANCHEREAU, J., CHAUSSABEL, D. & RAMILO, O. 2013. Whole blood

- gene expression profiles to assess pathogenesis and disease severity in infants with respiratory syncytial virus infection. *PLoS Med.*, 10, e1001549.
- MERICO, D., ISSERLIN, R., STUEKER, O., EMILI, A. & BADER, G. D. 2010. Enrichment Map: A Network-Based Method for Gene-Set Enrichment Visualization and Interpretation. *PLoS One*, 5, e13984.
- MESTAS, J. & HUGHES, C. C. 2004. Of mice and not men: differences between mouse and human immunology. *J. Immunol.*, 172, 2731-8.
- MEYLAN, M., PETITPREZ, F., BECHT, E., BOUGOÛIN, A., PUIPIER, G., CALVEZ, A., GIGLIOLI, I., VERKARRE, V., LACROIX, G., VERNEAU, J., SUN, C.-M., LAURENT-PUIG, P., VANO, Y.-A., ELAÏDI, R., MÉJEAN, A., SANCHEZ-SALAS, R., BARRET, E., CATHELINEAU, X., OUDARD, S., REYNAUD, C.-A., DE REYNIÈS, A., SAUTÈS-FRIDMAN, C. & FRIDMAN, W. H. 2022. Tertiary lymphoid structures generate and propagate anti-tumor antibody-producing plasma cells in renal cell cancer. *Immunity*, 55, 527-541.e5.
- MILLER, J. F. 1961. Immunological function of the thymus. *Lancet*, 2, 748-9.
- MILLER, J. F. 2011. The golden anniversary of the thymus. *Nat. Rev. Immunol.*, 11, 489-95.
- MODLIN, R. L. & REA, T. H. 1988. Immunopathology of leprosy granulomas. *Springer Semin. Immunopathol.*, 10, 359-374.
- MOGUES, T., GOODRICH, M. E., RYAN, L., LACOURSE, R. & NORTH, R. J. 2001. The relative importance of T cell subsets in immunity and immunopathology of airborne Mycobacterium tuberculosis infection in mice. *J. Exp. Med.*, 193, 271-80.
- MOHAMMED, A., CUI, Y., MAS, V. R. & KAMALESWARAN, R. 2019. Differential gene expression analysis reveals novel genes and pathways in pediatric septic shock patients. *Sci. Rep.*, 9, 11270.
- MOHLOPHENI, J. M., RAVIKIRAN, M. R., KIRTI, S., YANJIA, J. Z., ELISEO, A. E., BRENDAN, P., ISAAC, B. D., PEI-YU, C., MATTHEW, G. B., JIN HEE, K., SEOK YONG, E., LAURA, E. V., SAMUEL, M. B., CLIFTON, E. B. I., MATTHIAS, M., VÉRONIQUE, D. & ERIC, J. R. 2016. Inflammatory signaling in human tuberculosis granulomas is spatially organized. *Nat. Med.*, 22.
- MOLINA-RUIZ, A. M. & REQUENA, L. 2015. Foreign body granulomas. *Dermatol. Clin.*, 33, 497-523.
- MORRISON, J. M. D., PAI, M. M. D. & HOPEWELL, P. C. M. D. 2008. Tuberculosis and latent tuberculosis infection in close contacts of people with pulmonary tuberculosis in low-income and middle-income countries: a systematic review and meta-analysis. *Lancet Infect. Dis.*, 8, 359-368.
- MURRAY, J., SONNENBERG, P., NELSON, G., BESTER, A., SHEARER, S. & GLYNN, J. R. 2007. Cause of death and presence of respiratory disease at autopsy in an HIV-1 seroconversion cohort of southern African gold miners. *AIDS*, 21 Suppl 6, S97-s104.
- NAHID, P., DORMAN, S. E., ALIPANAH, N., BARRY, P. M., BROZEK, J. L., CATTAMANCHI, A., CHAISSON, L. H., CHAISSON, R. E., DALEY, C. L., GRZEMSKA, M., HIGASHI, J. M., HO, C. S., HOPEWELL, P. C., KESHAVJEE, S. A., LIENHARDT, C., MENZIES, R., MERRIFIELD, C., NARITA, M., O'BRIEN, R., PELOQUIN, C. A., RAFTERY, A., SAUKKONEN, J., SCHAAF, H. S., SOTGIU, G., STARKE, J. R., MIGLIORI, G. B. & VERNON, A. 2016. Official American Thoracic Society/Centers for Disease Control and Prevention/Infectious Diseases Society of America Clinical Practice Guidelines: Treatment of Drug-Susceptible Tuberculosis. *Clin. Infect. Dis.*, 63, e147-e195.
- NAKAYA, HELDER I., HAGAN, T., DURASINGHAM, SAI S., LEE, EVA K., KWISSA, M., ROUPHAEL, N., FRASCA, D., GERSTEN, M., MEHTA, ANEESH K., GAUJOUX, R., LI, G.-M., GUPTA, S., AHMED, R., MULLIGAN, MARK J.,

- SHEN-ORR, S., BLOMBERG, BONNIE B., SUBRAMANIAM, S. & PULENDRAN, B. 2015. Systems Analysis of Immunity to Influenza Vaccination across Multiple Years and in Diverse Populations Reveals Shared Molecular Signatures. *Immunity*, 43, 1186-1198.
- NAKAYA, H. I., WRAMMERT, J., LEE, E. K., RACIOPPI, L., MARIE-KUNZE, S., HAINING, W. N., MEANS, A. R., KASTURI, S. P., KHAN, N., LI, G. M., MCCAUSLAND, M., KANCHAN, V., KOKKO, K. E., LI, S., ELBEIN, R., MEHTA, A. K., ADEREM, A., SUBBARAO, K., AHMED, R. & PULENDRAN, B. 2011. Systems biology of vaccination for seasonal influenza in humans. *Nat. Immunol.*, 12, 786-95.
- NANUT, M. P. I. A., SABOTIA, J., AVAJGER, U., JEWETT, A. & KOS, J. 2017. Cystatin F Affects Natural Killer Cell Cytotoxicity.(Brief article). *Front. Immunol.*, 8, 1459.
- NDHLOVU, Z. M., KAMYA, P., MEWALAL, N., KLØVERPRIS, H. N., NKOSI, T., PRETORIUS, K., LAHER, F., OGUNSHOLA, F., CHOPERA, D., SHEKHAR, K., GHEBREMICHAEL, M., ISMAIL, N., MOODLEY, A., MALIK, A., LESLIE, A., GOULDER, P. J., BUUS, S., CHAKRABORTY, A., DONG, K., NDUNG'U, T. & WALKER, B. D. 2015. Magnitude and Kinetics of CD8+ T Cell Activation during Hyperacute HIV Infection Impact Viral Set Point. *Immunity*, 43, 591-604.
- NDUNG'U, T., DONG, K. L., KWON, D. S. & WALKER, B. D. 2018. A FRESH approach: Combining basic science and social good. *Sci. Immunol.*, 3.
- NORTH, R. J. 1973. Importance of thymus-derived lymphocytes in cell-mediated immunity to infection. *Cell. Immunol.*, 7, 166-76.
- NORTH, R. J. & IZZO, A. A. 1993. Granuloma formation in severe combined immunodeficient (SCID) mice in response to progressive BCG infection. Tendency not to form granulomas in the lung is associated with faster bacterial growth in this organ. *Am. J. Pathol.*, 142, 1959-66.
- NOSSAL, G. J. & LEDERBERG, J. 1958. Antibody production by single cells. *Nature*, 181, 1419-20.
- NOVIKOV, A., CARDONE, M., THOMPSON, R., SHENDEROV, K., KIRSCHMAN, K. D., MAYER-BARBER, K. D., MYERS, T. G., RABIN, R. L., TRINCHIERI, G., SHER, A. & FENG, C. G. 2011. Mycobacterium tuberculosis triggers host type I IFN signaling to regulate IL-1 β production in human macrophages. *J. Immunol.*, 187, 2540-7.
- NUNES-ALVES, C., BOOTY, M. G., CARPENTER, S. M., JAYARAMAN, P., ROTHCHILD, A. C. & BEHAR, S. M. 2014. In search of a new paradigm for protective immunity to TB. *Nat. Rev. Microbiol.*, 12, 289.
- O'GARRA, A., REDFORD, P. S., MCNAB, F. W., BLOOM, C. I., WILKINSON, R. J. & BERRY, M. P. R. 2013. The Immune Response in Tuberculosis. *Annu. Rev. Immunol.*, 31, 475-527.
- OBERMOSER, G., PRESNELL, S., DOMICO, K., XU, H., WANG, Y., ANGUIANO, E., THOMPSON-SNIPES, L., RANGANATHAN, R., ZEITNER, B., BJORK, A., ANDERSON, D., SPEAKE, C., RUCHAUD, E., SKINNER, J., ALSINA, L., SHARMA, M., DUTARTRE, H., CEPIKA, A., ISRAELSSON, E., NGUYEN, P., NGUYEN, Q. A., HARROD, A. C., ZURAWSKI, S. M., PASCUAL, V., UENO, H., NEPOM, G. T., QUINN, C., BLANKENSHIP, D., PALUCKA, K., BANCHEREAU, J. & CHAUSSABEL, D. 2013. Systems scale interactive exploration reveals quantitative and qualitative differences in response to influenza and pneumococcal vaccines. *Immunity*, 38, 831-44.
- ODIA, T., MALHERBE, S. T., MEIER, S., MAASDORP, E., KLEYNHANS, L., DU PLESSIS, N., LOXTON, A. G., ZAK, D. E., THOMPSON, E., DUFFY, F. J.,

- KUIVANIEMI, H., RONACHER, K., WINTER, J., WALZL, G. & TROMP, G. 2020. The Peripheral Blood Transcriptome Is Correlated With PET Measures of Lung Inflammation During Successful Tuberculosis Treatment. *Front. Immunol.*, 11, 596173.
- OLSEN, N., SOKKA, T., SEEHORN, C. L., KRAFT, B., MAAS, K., MOORE, J. & AUNE, T. M. 2004. A gene expression signature for recent onset rheumatoid arthritis in peripheral blood mononuclear cells. *Ann. Rheum. Dis.*, 63, 1387.
- OLSSON, A., VENKATASUBRAMANIAN, M., CHAUDHRI, V. K., ARONOW, B. J., SALOMONIS, N., SINGH, H. & GRIMES, H. L. 2016. Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature*, 537, 698-702.
- ORME, I. M., ROBERTS, A. D., GRIFFIN, J. P. & ABRAMS, J. S. 1993. Cytokine secretion by CD4 T lymphocytes acquired in response to Mycobacterium tuberculosis infection. *J. Immunol.*, 151, 518-25.
- ORME, I. M., ROBINSON, R. T. & COOPER, A. M. 2015. The balance between protective and pathogenic immune responses in the TB-infected lung. *Nat. Immunol.*, 16, 57-63.
- ORRÙ, V., STERI, M., SOLE, G., SIDORE, C., VIRDIS, F., DEI, M., LAI, S., ZOLEDZIEWSKA, M., BUSONERO, F., MULAS, A., FLORIS, M., MENTZEN, W. I., URRU, S. A., OLLA, S., MARONGIU, M., PIRAS, M. G., LOBINA, M., MASCHIO, A., PITZALIS, M., URRU, M. F., MARCELLI, M., CUSANO, R., DEIDDA, F., SERRA, V., OPPO, M., PILU, R., REINIER, F., BERUTTI, R., PIREDDU, L., ZARA, I., PORCU, E., KWONG, A., BRENNAN, C., TARRIER, B., LYONS, R., KANG, H. M., UZZAU, S., ATZENI, R., VALENTINI, M., FIRINU, D., LEONI, L., ROTTA, G., NAITZA, S., ANGIUS, A., CONGIA, M., WHALEN, M. B., JONES, C. M., SCHLESSINGER, D., ABECASIS, G. R., FIORILLO, E., SANNA, S. & CUCCA, F. 2013. Genetic variants regulating immune cell levels in health and disease. *Cell*, 155, 242-56.
- OSTRAND-ROSENBERG, S. 2004. Animal models of tumor immunity, immunotherapy and cancer vaccines. *Curr. Opin. Immunol.*, 16, 143-50.
- OVSYANNIKOVA, I. G., SALK, H. M., KENNEDY, R. B., HARALAMBIEVA, I. H., ZIMMERMANN, M. T., GRILL, D. E., OBERG, A. L. & POLAND, G. A. 2016. Gene signatures associated with adaptive humoral immunity following seasonal influenza A/H1N1 vaccination. *Genes Immun.*, 17, 371-379.
- OZBEK, B., ERTUNC, O., ERICKSON, A., VIDAL, I. D., GOMES-ALEXANDRE, C., GUNER, G., HICKS, J. L., JONES, T., TAUBE, J. M., SFANOS, K. S., YEGNASUBRAMANIAN, S. & DE MARZO, A. M. 2022. Multiplex immunohistochemical phenotyping of T cells in primary prostate cancer. *Prostate*, 82, 706-722.
- OZSOLAK, F. & MILOS, P. M. 2011. RNA sequencing: advances, challenges and opportunities. *Nat. Rev. Genet.*, 12, 87-98.
- PADRELL, M., LLORENTE, M. & AMICI, F. 2021. Invasive Research on Non-Human Primates—Time to Turn the Page. *Animals*, 11.
- PAGÁN, A. J. & RAMAKRISHNAN, L. 2014. Immunity and Immunopathology in the Tuberculous Granuloma. *Cold Spring Harb. Perspect. Med.*, 5, a018499.
- PAGÁN, A. J. & RAMAKRISHNAN, L. 2018. The Formation and Function of Granulomas. *Annu. Rev. Immunol.*, 36, 639-665.
- PAGÈS, F., GALON, J., DIEU-NOSJEAN, M. C., TARTOUR, E., SAUTÈS-FRIDMAN, C. & FRIDMAN, W. H. 2010. Immune infiltration in human tumors: a prognostic factor that should not be ignored. *Oncogene*, 29, 1093-1102.

- PAI, M., BEHR, M. A., DOWDY, D., DHEDA, K., DIVANGAHI, M., BOEHME, C. C., GINSBERG, A., SWAMINATHAN, S., SPIGELMAN, M., GETAHUN, H., MENZIES, D. & RAVIGLIONE, M. 2016. Tuberculosis. *Nat. Rev. Dis. Primers*, 2.
- PAN, Q., SHAI, O., LEE, L. J., FREY, B. J. & BLENCOWE, B. J. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.*, 40, 1413-5.
- PANKLA, R., BUDDHISA, S., BERRY, M., BLANKENSHIP, D. M., BANCROFT, G. J., BANCHEREAU, J., LERTMEMONGKOLCHAI, G. & CHAUSSABEL, D. 2009. Genomic transcriptional profiling identifies a candidate blood biomarker signature for the diagnosis of septicemic melioidosis. *Genome Biol.*, 10, R127.
- PARK, C., PONATH, G., LEVINE-RITTERMAN, M., BULL, E., SWANSON, E. C., DE JAGER, P. L., SEGAL, B. M. & PITT, D. 2019. The landscape of myeloid and astrocyte phenotypes in acute multiple sclerosis lesions. *Acta Neuropathol. Commun.*, 7, 130.
- PARK, I. J., AN, S., KIM, S. Y., LIM, H. M., HONG, S. M., KIM, M. J., KIM, Y. J. & YU, C. S. 2017. Prediction of radio-responsiveness with immune-profiling in patients with rectal cancer. *Oncotarget*, 8, 79793-79802.
- PARK, J. H. & LEE, H. K. 2020. Re-analysis of Single Cell Transcriptome Reveals That the NR3C1-CXCL8-Neutrophil Axis Determines the Severity of COVID-19. *Front. Immunol.*, 11.
- PARNELL, G. P., MCLEAN, A. S., BOOTH, D. R., ARMSTRONG, N. J., NALOS, M., HUANG, S. J., MANAK, J., TANG, W., TAM, O.-Y., CHAN, S. & TANG, B. M. 2012. A distinct influenza infection signature in the blood transcriptome of patients with severe community-acquired pneumonia. *J. Crit. Care*, 16, R157.
- PARNELL, G. P., TANG, B. M., NALOS, M., ARMSTRONG, N. J., HUANG, S. J., BOOTH, D. R. & MCLEAN, A. S. 2013. Identifying key regulatory genes in the whole blood of septic patients to monitor underlying immune dysfunctions. *Shock*, 40, 166-74.
- PATEL, S. S., TAVANA, M. L., BOGER, M. S., WIN, S. S. & RIMAWI, B. H. 2014. Necrotizing Soft Tissue Infection Occurring after Exposure to *Mycobacterium marinum*. *Case Rep. Infect. Dis.*, 2014, 702613.
- PAYNE, K. J. & CROOKS, G. M. 2007. Immune-cell lineage commitment: translation from mice to humans. *Immunity*, 26, 674-7.
- PETRYSZAK, R., KEAYS, M., TANG, Y. A., FONSECA, N. A., BARRERA, E., BURDETT, T., FÜLLGRABE, A., FUENTES, A. M., JUPP, S., KOSKINEN, S., MANNION, O., HUERTA, L., MEGY, K., SNOW, C., WILLIAMS, E., BARZINE, M., HASTINGS, E., WEISSER, H., WRIGHT, J., JAISWAL, P., HUBER, W., CHOUDHARY, J., PARKINSON, H. E. & BRAZMA, A. 2016. Expression Atlas update--an integrated database of gene and protein expression in humans, animals and plants. *Nucleic Acids Res.*, 44, D746-52.
- PIASECKA, B., DUFFY, D., URRUTIA, A., QUACH, H., PATIN, E., POSSEME, C., BERGSTEDT, J., CHARBIT, B., ROUILLY, V., MACPHERSON, C. R., HASAN, M., ALBAUD, B., GENTIEN, D., FELLAY, J., ALBERT, M. L. & QUINTANA-MURCI, L. 2018. Distinctive roles of age, sex, and genetics in shaping transcriptional variation of human immune responses to microbial challenges. *Proc. Natl. Acad. Sci. U. S. A.*, 115, E488-e497.
- PISU, D., HUANG, L., NARANG, V., THERIAULT, M., LÊ-BURY, G., LEE, B., LAKUDZALA, A. E., MZINZA, D. T., MHANGO, D. V., MITINI-NKHOMA, S. C., JAMBO, K. C., SINGHAL, A., MWANDUMBA, H. C. & RUSSELL, D. G.

2021. Single cell analysis of *M. tuberculosis* phenotype and macrophage lineages in the infected lung. *J. Exp. Med.*, 218, e20210615.
- POLTORAK, A., HE, X., SMIRNOVA, I., LIU, M. Y., VAN HUFFEL, C., DU, X., BIRDWELL, D., ALEJOS, E., SILVA, M., GALANOS, C., FREUDENBERG, M., RICCIARDI-CASTAGNOLI, P., LAYTON, B. & BEUTLER, B. 1998. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in *Tlr4* gene. *Science*, 282, 2085-8.
- PRIMATURIA, C., RENIARTI, L. & NATAPRAWIRA, H. M. N. 2020. Comparison between the Interferon γ Release Assay—QuantiFERON Gold Plus (QFT-Plus)—and Tuberculin Skin Test (TST) in the Detection of Tuberculosis Infection in Immunocompromised Children. *Pulm. Med.*, 2020, 7159485.
- PRUNK, M., NANUT, M. P., SABOTIC, J., SVAJGER, U. & KOS, J. 2019. Increased cystatin F levels correlate with decreased cytotoxicity of cytotoxic T cells. *Radiol. Oncol.*, 53, 57-68.
- PULENDRAN, B. & DAVIS, M. 2020. The science and medicine of human immunology. *Science*, 369, eaay4014.
- R CORE TEAM. 2021. *R: A language and environment for statistical computing* [Online]. R Foundation for Statistical Computing, Vienna, Austria. Available: <https://www.R-project.org/> [Accessed].
- RADOM-AIZIK, S., ZALDIVAR, F., JR., LEU, S. Y., GALASSETTI, P. & COOPER, D. M. 2008. Effects of 30 min of aerobic exercise on gene expression in human neutrophils. *J. Appl. Physiol.*, 104, 236-43.
- RAMAGLIA, V., SHEIKH-MOHAMED, S., LEGG, K., PARK, C., ROJAS, O. L., ZANDEE, S., FU, F., ORNATSKY, O., SWANSON, E. C., PITT, D., PRAT, A., MCKEE, T. D. & GOMMERMAN, J. L. 2019. Multiplexed imaging of immune cells in staged multiple sclerosis lesions by mass cytometry. *Elife*, 8.
- RAMAKRISHNAN, L. 2012. Revisiting the role of the granuloma in tuberculosis. *Nat. Rev. Immunol.*, 12, 352-366.
- RAMASAMY, A., MONDRY, A., HOLMES, C. C. & ALTMAN, D. G. 2008. Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Med.*, 5, e184.
- RAMILO, O., ALLMAN, W., CHUNG, W., MEJIAS, A., ARDURA, M., GLASER, C., WITTKOWSKI, K. M., PIQUERAS, B., BANCHEREAU, J., PALUCKA, A. K. & CHAUSSABEL, D. 2007. Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood*, 109, 2066-77.
- RANGAKA, M. X., WILKINSON, K. A., GLYNN, J. R., LING, D., MENZIES, D., MWANSA-KAMBAFWILE, J., FIELDING, K., WILKINSON, R. J. & PAI, M. 2012. Predictive value of interferon- γ release assays for incident active tuberculosis: a systematic review and meta-analysis. *Lancet Infect. Dis.*, 12, 45-55.
- RAO, A., BARKLEY, D., FRANÇA, G. S. & YANAI, I. 2021. Exploring tissue architecture using spatial transcriptomics. *Nature*, 596, 211-220.
- RAO, M. S., VAN VLEET, T. R., CIURLIONIS, R., BUCK, W. R., MITTELSTADT, S. W., BLOMME, E. A. G. & LIGUORI, M. J. 2019. Comparison of RNA-Seq and Microarray Gene Expression Platforms for the Toxicogenomic Evaluation of Liver From Short-Term Rat Toxicity Studies. *Front. Genet.*, 9.
- RAUDVERE, U., KOLBERG, L., KUZMIN, I., ARAK, T., ADLER, P., PETERSON, H. & VILO, J. 2019. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.*, 47, W191-W198.
- REGEV, A., TEICHMANN, S. A., LANDER, E. S., AMIT, I., BENOIST, C., BIRNEY, E., BODENMILLER, B., CAMPBELL, P., CARNINCI, P., CLATWORTHY, M.,

- CLEVERS, H., DEPLANCKE, B., DUNHAM, I., EBERWINE, J., EILS, R., ENARD, W., FARMER, A., FUGGER, L., GÖTTGENS, B., HACHOEN, N., HANIFFA, M., HEMBERG, M., KIM, S., KLENERMAN, P., KRIEGSTEIN, A., LEIN, E., LINNARSSON, S., LUNDBERG, E., LUNDEBERG, J., MAJUMDER, P., MARIONI, J. C., MERAD, M., MHLANGA, M., NAWIJN, M., NETEA, M., NOLAN, G., PE'ER, D., PHILLIPAKIS, A., PONTING, C. P., QUAKE, S., REIK, W., ROZENBLATT-ROSEN, O., SANES, J., SATIJA, R., SCHUMACHER, T. N., SHALEK, A., SHAPIRO, E., SHARMA, P., SHIN, J. W., STEGLE, O., STRATTON, M., STUBBINGTON, M. J. T., THEIS, F. J., UHLEN, M., VAN OUDENAARDEN, A., WAGNER, A., WATT, F., WEISSMAN, J., WOLD, B., XAVIER, R. & YOSEF, N. 2017. The Human Cell Atlas. *Elife*, 6.
- REIMAND, J., ISSERLIN, R., VOISIN, V., KUCERA, M., TANNUS-LOPES, C., ROSTAMIANFAR, A., WADI, L., MEYER, M., WONG, J., XU, C., MERICO, D. & BADER, G. D. 2019. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat. Protoc.*, 14, 482-517.
- REN, X., ZHANG, L., ZHANG, Y., LI, Z., SIEMERS, N. & ZHANG, Z. 2021. Insights Gained from Single-Cell Analysis of Immune Cells in the Tumor Microenvironment. *Annu. Rev. Immunol.*, 39, 583-609.
- REYES, A. L. P., SILVA, T. C., COETZEE, S. G., PLUMMER, J. T., DAVIS, B. D., CHEN, S., HAZELETT, D. J., LAWRENSON, K., BERMAN, B. P., GAYTHER, S. A. & JONES, M. R. 2019. GENAVi: a shiny web application for gene expression normalization, analysis and visualization. *BMC Genomics*, 20, 745.
- RHODES, D. R., BARRETTE, T. R., RUBIN, M. A., GHOSH, D. & CHINNAIYAN, A. M. 2002. Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. *Cancer Res.*, 62, 4427-33.
- RHODES, D. R., YU, J., SHANKER, K., DESHPANDE, N., VARAMBALLY, R., GHOSH, D., BARRETTE, T., PANDEY, A. & CHINNAIYAN, A. M. 2004. Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc. Natl. Acad. Sci. U. S. A.*, 101, 9309-14.
- RICE, J. 2012. Animal models: Not close enough. *Nature*, 484, S9-S9.
- RICHARDS, C. H., MOHAMMED, Z., QAYYUM, T., HORGAN, P. G. & MCMILLAN, D. C. 2011. The prognostic value of histological tumor necrosis in solid organ malignant disease: a systematic review. *Future Oncol.*, 7, 1223-1235.
- RIMM, D. L., CAMP, R. L., CHARETTE, L. A., COSTA, J., OLSEN, D. A. & REISS, M. 2001. Tissue microarray: a new technology for amplification of tissue resources. *J. Cancer*, 7, 24-31.
- RINCHAI, D., KEWCHAROENWONG, C., KESSLER, B., LERTMEMONGKOLCHAI, G. & CHAUSSABEL, D. 2015. Increased abundance of ADAM9 transcripts in the blood is associated with tissue damage. *F1000Res.*, 4, 89.
- RINCHAI, D., ROELANDS, J., TOUFIQ, M., HENDRICKX, W., ALTMAN, M. C., BEDOGNETTI, D. & CHAUSSABEL, D. 2021. BloodGen3Module: blood transcriptional module repertoire analysis and visualization using R. *Bioinformatics*, 37, 2382-2389.
- RITCHIE, M. E., PHIPSON, B., WU, D., HU, Y., LAW, C. W., SHI, W. & SMYTH, G. K. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.*, 43, e47.
- ROBINSON, M., SWEENEY, T. E., BAROUCH-BENTOV, R., SAHOO, M. K., KALESINSKAS, L., VALLANIA, F., SANZ, A. M., ORTIZ-LASSO, E., ALBORNOZ, L. L., ROSSO, F., MONTOYA, J. G., PINSKY, B. A., KHATRI, P. &

- EINAV, S. 2019. A 20-Gene Set Predictive of Progression to Severe Dengue. *Cell Rep.*, 26, 1104-1111.e4.
- ROBINSON, M. D., MCCARTHY, D. J. & SMYTH, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26, 139-40.
- ROMANOV, V., DAVIDOFF, S. N., MILES, A. R., GRAINGER, D. W., GALE, B. K. & BROOKS, B. D. 2014. A critical comparison of protein microarray fabrication technologies. *Analyst*, 139, 1303-26.
- ROSENTHAL, S. R. 1955. The Tubercle Bacillus in the Pulmonary Lesion of Man: Histobacteriology and Its Bearing on the Therapy of Pulmonary Tuberculosis. *Am. J. Clin. Pathol.*, 25, 1064-1064.
- ROY CHOWDHURY, R., VALLANIA, F., YANG, Q., LOPEZ ANGEL, C. J., DARBOE, F., PENN-NICHOLSON, A., ROZOT, V., NEMES, E., MALHERBE, S. T., RONACHER, K., WALZL, G., HANEKOM, W., DAVIS, M. M., WINTER, J., CHEN, X., SCRIBA, T. J., KHATRI, P. & CHIEN, Y. H. 2018. A multi-cohort study of the immune factors associated with M. tuberculosis infection outcomes. *Nature*, 560, 644-648.
- RUE-ALBRECHT, K., MCGETTIGAN, P. A., HERNÁNDEZ, B., NALPAS, N. C., MAGEE, D. A., PARNELL, A. C., GORDON, S. V. & MACHUGH, D. E. 2016. GOexpress: an R/Bioconductor package for the identification and visualisation of robust gene ontology signatures through supervised learning of gene expression data. *BMC Bioinformatics*, 17, 126.
- RUSSELL, D. G. 2011. Mycobacterium tuberculosis and the intimate discourse of a chronic infection. *Immunol. Rev.*, 240, 252-68.
- SCANGA, C. A., BAFICA, A., FENG, C. G., CHEEVER, A. W., HIENY, S. & SHER, A. 2004. MyD88-Deficient Mice Display a Profound Loss in Resistance to Mycobacterium tuberculosis Associated with Partially Impaired Th1 Cytokine and Nitric Oxide Synthase 2 Expression. *Infect. Immun.*, 72, 2400-2404.
- SCANGA, C. A. & FLYNN, J. L. 2014. Modeling tuberculosis in nonhuman primates. *Cold Spring Harb. Perspect. Med.*, 4, a018564.
- SCHER, H. I., LU, D., SCHREIBER, N. A., LOUW, J., GRAF, R. P., VARGAS, H. A., JOHNSON, A., JENDRISAK, A., BAMBURY, R., DANILA, D., MCLAUGHLIN, B., WAHL, J., GREENE, S. B., HELLER, G., MARRINUCCI, D., FLEISHER, M. & DITTAMORE, R. 2016. Association of AR-V7 on Circulating Tumor Cells as a Treatment-Specific Biomarker With Outcomes and Survival in Castration-Resistant Prostate Cancer. *JAMA Oncol.*, 2, 1441-1449.
- SCHLITZER, A., SIVAKAMASUNDARI, V., CHEN, J., SUMATOH, H. R., SCHREUDER, J., LUM, J., MALLERET, B., ZHANG, S., LARBI, A., ZOLEZZI, F., RENIA, L., POIDINGER, M., NAIK, S., NEWELL, E. W., ROBSON, P. & GINHOUX, F. 2015. Identification of cDC1- and cDC2-committed DC progenitors reveals early lineage priming at the common DC progenitor stage in the bone marrow. *Nat. Immunol.*, 16, 718-28.
- SCHOLZ, C. J., BIERNAT, P., BECKER, M., BAßLER, K., GÜNTHER, P., BALFER, J., DICKTEN, H., FLÖER, L., HEIKAMP, K., ANGERER, P., HEILIG, M., KARLE, R., KÖHLER, M., MAZURKIEWICZ, T., MÖNNIGHOFF, M., SAUER, C., SCHICK, A., SCHLEMM, G., WEIGELT, R., WINKLER, M., ULAS, T., THEIS, F., HUTHMACHER, S., KRATSCH, C. & SCHULTZE, J. L. 2018. FASTGenomics: An analytical ecosystem for single-cell RNA sequencing data. *bioRxiv*, 272476.

- SCHÜFFLER, P. J., SCHAPIRO, D., GIESEN, C., WANG, H. A. O., BODENMILLER, B. & BUHMANN, J. M. 2015. Automatic single cell segmentation on highly multiplexed tissue images. *Cytometry B*, 87, 936-942.
- SCHULTE-SCHREPPING, J., REUSCH, N., PACLIK, D., BAßLER, K., SCHLICKEISER, S., ZHANG, B., KRÄMER, B., KRAMMER, T., BRUMHARD, S., BONAGURO, L., DE DOMENICO, E., WENDISCH, D., GRASSHOFF, M., KAPPELLOS, T. S., BECKSTETTE, M., PECHT, T., SAGLAM, A., DIETRICH, O., MEI, H. E., SCHULZ, A. R., CONRAD, C., KUNKEL, D., VAFADARNEJAD, E., XU, C. J., HORNE, A., HERBERT, M., DREWS, A., THIBEAULT, C., PFEIFFER, M., HIPPENSTIEL, S., HOCKE, A., MÜLLER-REDETZKY, H., HEIM, K. M., MACHLEIDT, F., UHRIG, A., BOSQUILLON DE JARCY, L., JÜRGENS, L., STEGEMANN, M., GLÖSENKAMP, C. R., VOLK, H. D., GOFFINET, C., LANDTHALER, M., WYLER, E., GEORG, P., SCHNEIDER, M., DANG-HEINE, C., NEUWINGER, N., KAPPERT, K., TAUBER, R., CORMAN, V., RAABE, J., KAISER, K. M., VINH, M. T., RIEKE, G., MEISEL, C., ULAS, T., BECKER, M., GEFFERS, R., WITZENRATH, M., DROSTEN, C., SUTTORP, N., VON KALLE, C., KURTH, F., HÄNDLER, K., SCHULTZE, J. L., ASCHENBRENNER, A. C., LI, Y., NATTERMANN, J., SAWITZKI, B., SALIBA, A. E. & SANDER, L. E. 2020. Severe COVID-19 Is Marked by a Dysregulated Myeloid Cell Compartment. *Cell*, 182, 1419-1440.e23.
- SCHULZ, D., ZANOTELLI, V. R. T., FISCHER, J. R., SCHAPIRO, D., ENGLER, S., LUN, X. K., JACKSON, H. W. & BODENMILLER, B. 2018. Simultaneous Multiplexed Imaging of mRNA and Proteins with Subcellular Resolution in Breast Cancer Tissue Samples by Mass Cytometry. *Cell Syst.*, 6, 25-36.e5.
- SCHÜRCH, C. M., BHATE, S. S., BARLOW, G. L., PHILLIPS, D. J., NOTI, L., ZLOBEC, I., CHU, P., BLACK, S., DEMETER, J., MCILWAIN, D. R., KINOSHITA, S., SAMUSIK, N., GOLTSEV, Y. & NOLAN, G. P. 2020. Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. *Cell*, 182, 1341-1359.e19.
- SEE, P., DUTERTRE, C. A., CHEN, J., GÜNTHER, P., MCGOVERN, N., IRAC, S. E., GUNAWAN, M., BEYER, M., HÄNDLER, K., DUAN, K., SUMATOH, H. R. B., RUFFIN, N., JOUVE, M., GEA-MALLORQUÍ, E., HENNEKAM, R. C. M., LIM, T., YIP, C. C., WEN, M., MALLERET, B., LOW, I., SHADAN, N. B., FEN, C. F. S., TAY, A., LUM, J., ZOLEZZI, F., LARBI, A., POIDINGER, M., CHAN, J. K. Y., CHEN, Q., RÉNIA, L., HANIFFA, M., BENAROCH, P., SCHLITZER, A., SCHULTZE, J. L., NEWELL, E. W. & GINHOUX, F. 2017. Mapping the human DC lineage through the integration of high-dimensional techniques. *Science*, 356.
- SEE, P., LUM, J., CHEN, J. & GINHOUX, F. 2018. A Single-Cell Sequencing Guide for Immunologists. *Front. Immunol.*, 9, 2425.
- SEGAL, B. H., LETO, T. L., GALLIN, J. I., MALECH, H. L. & HOLLAND, S. M. 2000. Genetic, biochemical, and clinical features of chronic granulomatous disease. *Medicine*, 79, 170-200.
- SEGAL, E., FRIEDMAN, N., KAMINSKI, N., REGEV, A. & KOLLER, D. 2005. From signatures to models: understanding cancer using microarrays. *Nat. Genet.*, 37, S38-S45.
- SEGAL, E., FRIEDMAN, N., KOLLER, D. & REGEV, A. 2004. A module map showing conditional activity of expression modules in cancer. *Nat. Genet.*, 36, 1090-1098.
- SEKALY, R. P. 2008. The failed HIV Merck vaccine study: a step back or a launching point for future vaccine development? *J. Exp. Med.*, 205, 7-12.

- SEOK, J., WARREN, H. S., CUENCA, A. G., MINDRINOS, M. N., BAKER, H. V., XU, W., RICHARDS, D. R., MCDONALD-SMITH, G. P., GAO, H., HENNESSY, L., FINNERTY, C. C., LÓPEZ, C. M., HONARI, S., MOORE, E. E., MINEL, J. P., CUSCHIERI, J., BANKEY, P. E., JOHNSON, J. L., SPERRY, J., NATHENS, A. B., BILLIAR, T. R., WEST, M. A., JESCHKE, M. G., KLEIN, M. B., GAMELLI, R. L., GIBRAN, N. S., BROWNSTEIN, B. H., MILLER-GRAZIANO, C., CALVANO, S. E., MASON, P. H., COBB, J. P., RAHME, L. G., LOWRY, S. F., MAIER, R. V., MOLDAWER, L. L., HERNDON, D. N., DAVIS, R. W., XIAO, W. & TOMPKINS, R. G. 2013. Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc. Natl. Acad. Sci. U. S. A.*, 110, 3507-12.
- SETTY, M., TADMOR, M. D., REICH-ZELIGER, S., ANGEL, O., SALAME, T. M., KATHAIL, P., CHOI, K., BENDALL, S., FRIEDMAN, N. & PE'ER, D. 2016. Wishbone identifies bifurcating developmental trajectories from single-cell data. *Nat. Biotechnol.*, 34, 637-45.
- SHAH, K. K., PRITT, B. S. & ALEXANDER, M. P. 2017. Histopathologic review of granulomatous inflammation. *J. Clin. Tuberc. Other Mycobact. Dis.*, 7, 1-12.
- SHALON, D., SMITH, S. J. & BROWN, P. O. 1996. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res.*, 6, 639-45.
- SHANNON, P., MARKIEL, A., OZIER, O., BALIGA, N. S., WANG, J. T., RAMAGE, D., AMIN, N., SCHWIKOWSKI, B. & IDEKER, T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, 13, 2498-504.
- SHARMA, P. & ALLISON, J. P. 2015. The future of immune checkpoint therapy. *Science*, 348, 56-61.
- SHAUL, M. E., LEVY, L., SUN, J., MISHALIAN, I., SINGHAL, S., KAPOOR, V., HORNG, W., FRIDLENDER, G., ALBELDA, S. M. & FRIDLENDER, Z. G. 2016. Tumor-associated neutrophils display a distinct N1 profile following TGF β modulation: A transcriptomics analysis of pro- vs. antitumor TANs. *Oncoimmunology*, 5, e1232221.
- SHAY, T., JOJIC, V., ZUK, O., ROTHAMEL, K., PUYRAIMOND-ZEMMOUR, D., FENG, T., WAKAMATSU, E., BENOIST, C., KOLLER, D. & REGEV, A. 2013. Conservation and divergence in the transcriptional programs of the human and mouse immune systems. *Proc. Natl. Acad. Sci. U. S. A.*, 110, 2946-2951.
- SHEFFIELD, E. A. 1990. The granulomatous inflammatory response. *J. Pathol.*, 160, 1-2.
- SHEN-ORR, S. S., TIBSHIRANI, R., KHATRI, P., BODIAN, D. L., STAEDTLER, F., PERRY, N. M., HASTIE, T., SARWAL, M. M., DAVIS, M. M. & BUTTE, A. J. 2010. Cell type-specific gene expression differences in complex tissues. *Nat. Methods*, 7, 287-9.
- SHENOI, S., OU, J. N., NI, C., MACAUBAS, C., GERSUK, V. H., WALLACE, C. A., MELLINS, E. D. & STEVENS, A. M. 2015. Comparison of biomarkers for systemic juvenile idiopathic arthritis. *Pediatr. Res.*, 78, 554-9.
- SHIMIZU, T., WISESMITH, W., LI, J., ABE, M., SAKIMURA, K., CHETSAWANG, B., SAHARA, Y., TOHYAMA, K., TANAKA, K. F. & IKENAKA, K. 2017. The balance between cathepsin C and cystatin F controls remyelination in the brain of Plp1-overexpressing mouse, a chronic demyelinating disease model. *Glia*, 65, 917-930.
- SHIMOKATA, K., KISHIMOTO, H., TAKAGI, E. & TSUNEKAWA, H. 1986. Determination of the T-cell subset producing gamma-interferon in tuberculous pleural effusion. *Microbiol. Immunol.*, 30, 353-61.

- SILVER, R. F., MYERS, A. J., JARVELA, J., FLYNN, J., RUTLEDGE, T., BONFIELD, T. & LIN, P. L. 2016. Diversity of Human and Macaque Airway Immune Cells at Baseline and during Tuberculosis Infection. *Am. J. Respir. Cell Mol. Biol.*, 55, 899-908.
- SINGH, A. K. & GUPTA, U. D. 2018. Animal models of tuberculosis: Lesson learnt. *Indian J. Med. Res.*, 147, 456-463.
- SINGHANIA, A., VERMA, R., GRAHAM, C. M., LEE, J., TRAN, T., RICHARDSON, M., LECINE, P., LEISSNER, P., BERRY, M. P. R., WILKINSON, R. J., KAISER, K., RODRIGUE, M., WOLTMANN, G., HALDAR, P. & O'GARRA, A. 2018a. A modular transcriptional signature identifies phenotypic heterogeneity of human tuberculosis infection. *Nat. Commun.*, 9, 2308.
- SINGHANIA, A., WILKINSON, R. J., RODRIGUE, M., HALDAR, P. & O'GARRA, A. 2018b. The value of transcriptomics in advancing knowledge of the immune response and diagnosis in tuberculosis. *Nat. Immunol.*, 19, 1159-1168.
- SONNENBERG, P., GLYNN, J. R., FIELDING, K., MURRAY, J., GODFREY-FAUSSETT, P. & SHEARER, S. 2005. How soon after infection with HIV does the risk of tuberculosis start to increase? A retrospective cohort study in South African gold miners. *J. Infect. Dis.*, 191, 150-8.
- SOUSA, L. G. D., RAJAPAKSHE, K., RODRIGUEZ CANALES, J., CHIN, R. L., FENG, L., WANG, Q., BARRESE, T. Z., MASSARELLI, E., WILLIAM, W., JOHNSON, F. M., FERRAROTTO, R., WISTUBA, I., COARFA, C., LEE, J., WANG, J., MELIEF, C. J. M., CURRAN, M. A. & GLISSON, B. S. 2022. ISA101 and nivolumab for HPV-16 cancer: updated clinical efficacy and immune correlates of response. *J. Immunother. Cancer.*, 10, e004232.
- SPARKS, R., LAU, W. W. & TSANG, J. S. 2016. Expanding the Immunology Toolbox: Embracing Public-Data Reuse and Crowdsourcing. *Immunity*, 45, 1191-1204.
- SPEAKE, C., PRESNELL, S., DOMICO, K., ZEITNER, B., BJORK, A., ANDERSON, D., MASON, M. J., WHALEN, E., VARGAS, O., POPOV, D., RINCHAI, D., JOURDE-CHICHE, N., CHICHE, L., QUINN, C. & CHAUSSABEL, D. 2015. An interactive web application for the dissemination of human systems immunology data. *J. Transl. Med.*, 13.
- STACK, E. C., WANG, C., ROMAN, K. A. & HOYT, C. C. 2014. Multiplexed immunohistochemistry, imaging, and quantitation: a review, with an assessment of Tyramide signal amplification, multispectral imaging and multiplex analysis. *Methods*, 70, 46-58.
- STANLEY, S. A., JOHNDROW, J. E., MANZANILLO, P. & COX, J. S. 2007. The Type I IFN Response to Infection with *Mycobacterium tuberculosis* Requires ESX-1-Mediated Secretion and Contributes to Pathogenesis. *J. Immunol.*, 178, 3143.
- STEARNS, R. L., GETTS, R. C. & GULLANS, S. R. 2000. A novel, sensitive detection system for high-density microarrays using dendrimer technology. *Physiol. Genomics*, 3, 93-9.
- STEINGART, K. R., SCHILLER, I., HORNE, D. J., PAI, M., BOEHME, C. C. & DENDUKURI, N. 2014. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database Syst. Rev.*, 2014, Cd009593.
- STEINMAN, R. M. & COHN, Z. A. 1973. Identification of a novel cell type in peripheral lymphoid organs of mice. I. Morphology, quantitation, tissue distribution. *J. Exp. Med.*, 137, 1142-62.
- STEINMAN, R. M. & MELLMAN, I. 2004. Immunotherapy: bewitched, bothered, and bewildered no more. *Science*, 305, 197-200.

- STEPHENSON, E., REYNOLDS, G., BOTTING, R. A., CALERO-NIETO, F. J., MORGAN, M. D., TUONG, Z. K., BACH, K., SUNGNAK, W., WORLOCK, K. B., YOSHIDA, M., KUMASAKA, N., KANIA, K., ENGELBERT, J., OLABI, B., SPEGAROVA, J. S., WILSON, N. K., MENDE, N., JARDINE, L., GARDNER, L. C. S., GOH, I., HORSFALL, D., MCGRATH, J., WEBB, S., MATHER, M. W., LINDEBOOM, R. G. H., DANN, E., HUANG, N., POLANSKI, K., PRIGMORE, E., GOTHE, F., SCOTT, J., PAYNE, R. P., BAKER, K. F., HANRATH, A. T., SCHIM VAN DER LOEFF, I. C. D., BARR, A. S., SANCHEZ-GONZALEZ, A., BERGAMASCHI, L., MESCIA, F., BARNES, J. L., KILICH, E., DE WILTON, A., SAIGAL, A., SALEH, A., JANES, S. M., SMITH, C. M., GOPEE, N., WILSON, C., COUPLAND, P., COXHEAD, J. M., KISELEV, V. Y., VAN DONGEN, S., BACARDIT, J., KING, H. W., BAKER, S., BRADLEY, J. R., DOUGAN, G., GOODFELLOW, I. G., GUPTA, R. K., HESS, C., KINGSTON, N., LEHNER, P. J., MATHESON, N. J., OWEHAND, W. H., SAUNDERS, C., SMITH, K. G. C., SUMMERS, C., THAVENTHIRAN, J. E. D., TOSHNER, M., WEEKES, M. P., BUCKE, A., CALDER, J., CANNA, L., DOMINGO, J., ELMER, A., FULLER, S., HARRIS, J., HEWITT, S., KENNET, J., JOSE, S., KOURAMPA, J., MEADOWS, A., O'BRIEN, C., PRICE, J., PUBLICO, C., RASTALL, R., RIBEIRO, C., ROWLANDS, J., RUFFOLO, V., TORDESILLAS, H., BULLMAN, B., DUNMORE, B. J., FAWKE, S., GRÄF, S., HODGSON, J., HUANG, C., HUNTER, K., JONES, E., LEGCHENKO, E., MATARA, C., et al. 2021. Single-cell multi-omics analysis of the immune response in COVID-19. *Nat. Med.*, 27, 904-916.
- STUART, T., BUTLER, A., HOFFMAN, P., HAFEMEISTER, C., PAPALEXI, E., MAUCK, W. M., HAO, Y., STOECKIUS, M., SMIBERT, P. & SATIJA, R. 2019. Comprehensive Integration of Single-Cell Data. *Cell*, 177, 1888-1902.e21.
- STUBBINGTON, M. J. T., LÖNNBERG, T., PROSERPIO, V., CLARE, S., SPEAK, A. O., DOUGAN, G. & TEICHMANN, S. A. 2016. T cell fate and clonality inference from single-cell transcriptomes. *Nat. Methods*, 13, 329-332.
- STUBBINGTON, M. J. T., ROZENBLATT-ROSEN, O., REGEV, A. & TEICHMANN, S. A. 2017. Single-cell transcriptomics to explore the immune system in health and disease. *Science*, 358, 58-63.
- SU, L. F., HAN, A., MCGUIRE, H. M., FURMAN, D., NEWEL, E. W. & DAVIS, M. M. 2013. The promised land of human immunology. *Cold Spring Harb. Symp. Quant. Biol.*, 78, 203-213.
- SU, Z., FANG, H., HONG, H., SHI, L., ZHANG, W., ZHANG, W., ZHANG, Y., DONG, Z., LANCASHIRE, L. J., BESSARABOVA, M., YANG, X., NING, B., GONG, B., MEEHAN, J., XU, J., GE, W., PERKINS, R., FISCHER, M. & TONG, W. 2014. An investigation of biomarkers derived from legacy microarray data for their utility in the RNA-seq era. *Genome Biol.*, 15, 523.
- SUBBIAN, S., TSENOVA, L., KIM, M. J., WAINWRIGHT, H. C., VISSER, A., BANDYOPADHYAY, N., BADER, J. S., KARAKOUSIS, P. C., MURRMANN, G. B., BEKKER, L. G., RUSSELL, D. G. & KAPLAN, G. 2015. Lesion-Specific Immune Response in Granulomas of Patients with Pulmonary Tuberculosis: A Pilot Study. *PLoS One*, 10, e0132249.
- SUBRAMANIAN, A., TAMAYO, P., MOOTHA, V. K., MUKHERJEE, S., EBERT, B. L., GILLETTE, M. A., PAULOVICH, A., POMEROY, S. L., GOLUB, T. R., LANDER, E. S. & MESIROV, J. P. 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.*, 102, 15545-15550.

- SUGAWARA, I., YAMADA, H. & MIZUNO, S. 2003. Relative importance of STAT4 in murine tuberculosis. *J. Med. Microbiol.*, 52, 29-34.
- SULTAN, M., SCHULZ, M. H., RICHARD, H., MAGEN, A., KLINGENHOFF, A., SCHERF, M., SEIFERT, M., BORODINA, T., SOLDATOV, A., PARKHOMCHUK, D., SCHMIDT, D., O'KEEFFE, S., HAAS, S., VINGRON, M., LEHRACH, H. & YASPO, M. L. 2008. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science*, 321, 956-60.
- SUN, G., LI, Z., RONG, D., ZHANG, H., SHI, X., YANG, W., ZHENG, W., SUN, G., WU, F., CAO, H., TANG, W. & SUN, Y. 2021. Single-cell RNA sequencing in cancer: Applications, advances, and emerging challenges. *Mol. Ther. Oncolytics*, 21, 183-206.
- SUNTHARALINGAM, G., PERRY, M. R., WARD, S., BRETT, S. J., CASTELLO-CORTES, A., BRUNNER, M. D. & PANOSKALTSIS, N. 2006. Cytokine storm in a phase 1 trial of the anti-CD28 monoclonal antibody TGN1412. *N. Engl. J. Med.*, 355, 1018-28.
- SUTHERLAND, A., THOMAS, M., BRANDON, R. A., BRANDON, R. B., LIPMAN, J., TANG, B., MCLEAN, A., PASCOE, R., PRICE, G., NGUYEN, T., STONE, G. & VENTER, D. 2011. Development and validation of a novel molecular biomarker diagnostic test for the early detection of sepsis. *Crit. Care*, 15, R149.
- SVENSSON, V., VENTO-TORMO, R. & TEICHMANN, S. A. 2018. Exponential scaling of single-cell RNA-seq in the past decade. *Nat. Protoc.*, 13, 599-604.
- SWEENEY, T. E., BRAVIK, L., TATO, C. M. & KHATRI, P. 2016. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *The Lancet. Respiratory medicine*, 4, 213-24.
- SWEENEY, T. E., HAYNES, W. A., VALLANIA, F., IOANNIDIS, J. P. & KHATRI, P. 2017. Methods to increase reproducibility in differential gene expression via meta-analysis. *Nucleic Acids Res.*, 45, e1.
- SWEENEY, T. E., SHIDHAM, A., WONG, H. R. & KHATRI, P. 2015. A comprehensive time-course-based multicohort analysis of sepsis and sterile inflammation reveals a robust diagnostic gene set. *Sci. Transl. Med.*, 7, 287ra71.
- TABONE, O., VERMA, R., SINGHANIA, A., CHAKRAVARTY, P., BRANCHETT, W. J., GRAHAM, C. M., LEE, J., TRANG, T., REYNIER, F., LEISSNER, P., KAISER, K., RODRIGUE, M., WOLTMANN, G., HALDAR, P. & O'GARRA, A. 2021. Blood transcriptomics reveal the evolution and resolution of the immune response in tuberculosis. *J. Exp. Med.*, 218, e20210915.
- TAKAO, K. & MIYAKAWA, T. 2015. Genomic responses in mouse models greatly mimic human inflammatory diseases. *Proc. Natl. Acad. Sci. U. S. A.*, 112, 1167-72.
- TALMAGE, D. W. 1957. Allergy and immunology. *Annu. Rev. Med.*, 8, 239-56.
- TAN-GARCIA, A., LAI, F., SHENG YEONG, J. P., IRAC, S. E., NG, P. Y., MSALLAM, R., TATT LIM, J. C., WAI, L. E., THAM, C. Y. L., CHOO, S. P., LIM, T., YOUNG, D. Y., D'AMBROSIO, R., DEGASPERI, E., PERBELLINI, R., NEWELL, E., LE BERT, N., GINHOUX, F., BERTOLETTI, A., CHEN, Q. & DUTERTRE, C. A. 2020. Liver fibrosis and CD206(+) macrophage accumulation are suppressed by anti-GM-CSF therapy. *JHEP Rep*, 2, 100062.
- TAN, W. C. C., NERURKAR, S. N., CAI, H. Y., NG, H. H. M., WU, D., WEE, Y. T. F., LIM, J. C. T., YEONG, J. & LIM, T. K. H. 2020. Overview of multiplex immunohistochemistry/immunofluorescence techniques in the era of cancer immunotherapy. *Cancer Commun.*, 40, 135-153.
- TAN, Y., TAMAYO, P., NAKAYA, H., PULENDRAN, B., MESIROV, J. P. & HAINING, W. N. 2014. Gene signatures related to B-cell proliferation predict influenza vaccine-induced antibody response. *Eur. J. Immunol.*, 44, 285-295.

- TANG, B. M., SHOJAEI, M., PARNELL, G. P., HUANG, S., NALOS, M., TEOH, S., CONNOR, K., SCHIBECI, S., PHU, A. L., KUMAR, A., HO, J., MEYERS, A. F. A., KEYNAN, Y., BALL, T., PISIPATI, A., KUMAR, A., MOORE, E., EISEN, D., LAI, K., GILLET, M., GEFFERS, R., LUO, H., GUL, F., SCHREIBER, J., RIEDEL, S., BOOTH, D., MCLEAN, A. & SCHUGHART, K. 2017. A novel immune biomarker IFI27 discriminates between influenza and bacteria in patients with suspected respiratory infection. *Eur. Respir. J.*, 49, 1602098.
- TANG, B. M. P., MCLEAN, A. S., DAWES, I. W., HUANG, S. J. & LIN, R. C. Y. 2009a. Gene-expression profiling of peripheral blood mononuclear cells in sepsis*. *Crit. Care Med.*, 37.
- TANG, F., BARBACIORU, C., WANG, Y., NORDMAN, E., LEE, C., XU, N., WANG, X., BODEAU, J., TUCH, B. B., SIDDIQUI, A., LAO, K. & SURANI, M. A. 2009b. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods*, 6, 377-82.
- TAO, L. & REESE, T. A. 2017. Making Mouse Models That Reflect Human Immune Responses. *Trends Immunol.*, 38, 181-193.
- TER HAAR, N. M., TAK, T., MOKRY, M., SCHOLMAN, R. C., MEERDING, J. M., DE JAGER, W., VERWOERD, A., FOELL, D., VOGL, T., ROTH, J., LELIEFELD, P. H. C., VAN LOOSDREGT, J., KOENDERMAN, L., VASTERT, S. J. & DE ROOCK, S. 2018. Reversal of Sepsis-Like Features of Neutrophils by Interleukin-1 Blockade in Patients With Systemic-Onset Juvenile Idiopathic Arthritis. *Arthritis Rheumatol.*, 70, 943-956.
- THAKAR, J., MOHANTY, S., WEST, A. P., JOSHI, S. R., UEDA, I., WILSON, J., MENG, H., BLEVINS, T. P., TSANG, S., TRENTALANGE, M., SICONOLFI, B., PARK, K., GILL, T. M., BELSHE, R. B., KAECH, S. M., SHADEL, G. S., KLEINSTEIN, S. H. & SHAW, A. C. 2015. Aging-dependent alterations in gene expression and a mitochondrial signature of responsiveness to human influenza vaccination. *Aging (Albany N. Y.)*, 7, 38-52.
- THE GENE ONTOLOGY CONSORTIUM 2019. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.*, 47, D330-D338.
- THOMAS, D. L., II, MURPHY, A. G., WEISS, M. J., HE, J., MAKARY, M. A., BURKHART, R. A., WOLFGANG, C. L., JAFFEE, E. M., ZHENG, L. & THOMPSON, E. D. 2019. Analysis of spatial relationships between infiltrating immune cells within the tumor microenvironment following combinatorial immunotherapy. *Cancer Res.*, 79, 3102A-3102A.
- TIEMERSMA, E. W., MARIEKE, J. V. D. W., BORGDORFF, M. W., WILLIAMS, B. G. & NAGELKERKE, N. J. D. 2011. Natural History of Tuberculosis: Duration and Fatality of Untreated Pulmonary Tuberculosis in HIV Negative Patients: A Systematic Review. *PLoS One*, 6.
- TIROSH, I., IZAR, B., PRAKADAN, S. M., WADSWORTH, M. H., 2ND, TREACY, D., TROMBETTA, J. J., ROTEM, A., RODMAN, C., LIAN, C., MURPHY, G., FALLAHI-SICHANI, M., DUTTON-REGESTER, K., LIN, J. R., COHEN, O., SHAH, P., LU, D., GENSHAFT, A. S., HUGHES, T. K., ZIEGLER, C. G., KAZER, S. W., GAILLARD, A., KOLB, K. E., VILLANI, A. C., JOHANNESSEN, C. M., ANDREEV, A. Y., VAN ALLEN, E. M., BERTAGNOLLI, M., SORGER, P. K., SULLIVAN, R. J., FLAHERTY, K. T., FREDERICK, D. T., JANÉ-VALBUENA, J., YOON, C. H., ROZENBLATT-ROSEN, O., SHALEK, A. K., REGEV, A. & GARRAWAY, L. A. 2016. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*, 352, 189-96.

- TOMCZAK, K., CZERWIŃSKA, P. & WIZNEROWICZ, M. 2015. Review The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp. Oncol.*, 68-77.
- TORO-DOMÍNGUEZ, D., CARMONA-SÁEZ, P. & ALARCÓN-RIQUELME, M. E. 2014. Shared signatures between rheumatoid arthritis, systemic lupus erythematosus and Sjögren's syndrome uncovered through gene expression meta-analysis. *Arthritis Res. Ther.*, 16, 489.
- TOUFIQ, M., ROELANDS, J., ALFAKI, M., SYED AHAMED KABEER, B., SAADAOU, M., LAKSHMANAN, A. P., BANGARUSAMY, D. K., MURUGESAN, S., BEDOGNETTI, D., HENDRICKX, W., AL KHODOR, S., TERRANEGRA, A., RINCHAI, D., CHAUSSABEL, D. & GARAND, M. 2020. Annexin A3 in sepsis: novel perspectives from an exploration of public transcriptome data. *Immunology*, 161, 291-302.
- TRAPNELL, C., HENDRICKSON, D. G., SAUVAGEAU, M., GOFF, L., RINN, J. L. & PACHTER, L. 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.*, 31, 46-53.
- TRUNZ, B. B., FINE, P. & DYE, C. 2006. Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwide: a meta-analysis and assessment of cost-effectiveness. *Lancet*, 367, 1173-80.
- TSALIK, E. L., HENAO, R., NICHOLS, M., BURKE, T., KO, E. R., MCCLAIN, M. T., HUDSON, L. L., MAZUR, A., FREEMAN, D. H., VELDMAN, T., LANGLEY, R. J., QUACKENBUSH, E. B., GLICKMAN, S. W., CAIRNS, C. B., JAEHNE, A. K., RIVERS, E. P., OTERO, R. M., ZAAS, A. K., KINGSMORE, S. F., LUCAS, J., FOWLER, V. G., JR., CARIN, L., GINSBURG, G. S. & WOODS, C. W. 2016. Host gene expression classifiers diagnose acute respiratory illness etiology. *Sci. Transl. Med.*, 8, 322ra11.
- TSANG, J. S., SCHWARTZBERG, P. L., KOTLIAROV, Y., BIANCOTTO, A., XIE, Z., GERMAIN, R. N., WANG, E., OLNES, M. J., NARAYANAN, M., GOLDING, H., MOIR, S., DICKLER, H. B., PERL, S. & CHEUNG, F. 2014. Global analyses of human immune variation reveal baseline predictors of postvaccination responses. *Cell*, 157, 499-513.
- ULRICH, T., KOSMIADI, G. A., TRUSOV, V., JÖRG, S., PRADL, L., TITUKHINA, M., MISHENKO, V., GUSHINA, N. & KAUFMANN, S. H. E. 2004. Human tuberculous granulomas induce peripheral lymphoid follicle-like structures to orchestrate local host defence in the lung. *J. Pathol.*, 204, 217-228.
- UTSUNOMIYA, T., HARA, Y., KATAOKA, A., MORITA, M., ARAKAWA, H., MORI, M. & NISHIMURA, S. 2002. Cystatin-like metastasis-associated protein mRNA expression in human colorectal cancer is associated with both liver metastasis and patient survival. *Clin. Cancer Res.*, 8, 2591-4.
- VALLANIA, F., TAM, A., LOFGREN, S., SCHAFFERT, S., AZAD, T. D., BONGEN, E., HAYNES, W., ALSUP, M., ALONSO, M., DAVIS, M., ENGLEMAN, E. & KHATRI, P. 2018. Leveraging heterogeneity across multiple datasets increases cell-mixture deconvolution accuracy and reduces biological and technical biases. *Nat. Commun.*, 9, 4735.
- VAN DAM, S., BAARS, M. J. D. & VERCOULEN, Y. 2022. Multiplex Tissue Imaging: Spatial Revelations in the Tumor Microenvironment. *Cancers (Basel)*, 14.
- VAN DEN BERGE, K., ROUX DE BÉZIEUX, H., STREET, K., SAELENS, W., CANNODT, R., SAEYS, Y., DUDOIT, S. & CLEMENT, L. 2020. Trajectory-based differential expression analysis for single-cell sequencing data. *Nat. Commun.*, 11, 1201.

- VAN DEN BRINK, S. C., SAGE, F., VÉRTESY, Á., SPANJAARD, B., PETERSON-MADURO, J., BARON, C. S., ROBIN, C. & VAN OUDENAARDEN, A. 2017. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat. Methods*, 14, 935-936.
- VAN DER WEL, N., HAVA, D., HOUBEN, D., FLUITSMA, D., VAN ZON, M., PIERSON, J., BRENNER, M. & PETERS, P. J. 2007. M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells. *Cell*, 129, 1287-98.
- VAN DER WORP, H. B., HOWELLS, D. W., SENA, E. S., PORRITT, M. J., REWELL, S., O'COLLINS, V. & MACLEOD, M. R. 2010. Can animal models of disease reliably inform human studies? *PLoS Med.*, 7, e1000245.
- VAN DOORN, C. L. R., ECKOLD, C., RONACHER, K., RUSLAMI, R., VAN VEEN, S., LEE, J. S., KUMAR, V., KERRY-BARNARD, S., MALHERBE, S. T., KLEYNHANS, L., STANLEY, K., HILL, P. C., JOOSTEN, S. A., VAN CREVEL, R., WIJMENGA, C., CRITCHLEY, J. A., WALZL, G., ALISJAHBANA, B., HAKS, M. C., DOCKRELL, H. M., OTTENHOFF, T. H. M., VIANELLO, E. & CLIFF, J. M. 2022. Transcriptional profiles predict treatment outcome in patients with tuberculosis and diabetes at diagnosis and at two weeks after initiation of anti-tuberculosis treatment. *EBioMedicine*, 82, 104173.
- VAN HERCK, Y., ANTORANZ, A., ANDHARI, M. D., MILLI, G., BECHTER, O., DE SMET, F. & BOSISIO, F. M. 2021. Multiplexed Immunohistochemistry and Digital Pathology as the Foundation for Next-Generation Pathology in Melanoma: Methodological Comparison and Future Clinical Applications. *Front. Oncol.*, 11.
- VAN VERK, M. C., HICKMAN, R., PIETERSE, C. M. & VAN WEES, S. C. 2013. RNA-Seq: revelation of the messengers. *Trends Plant Sci.*, 18, 175-9.
- VANGURI, R., BENHAMIDA, J., YOUNG, J. H., LI, Y., ZIVANOVIC, O., CHI, D., SNYDER, A., HOLLMANN, T. J. & MAGER, K. L. 2022. Understanding the impact of chemotherapy on the immune landscape of high-grade serous ovarian cancer. *Gynecol. Oncol. Rep.*, 39, 100926.
- VARADÉ, J., MAGADÁN, S. & GONZÁLEZ-FERNÁNDEZ, Á. 2021. Human immunology and immunotherapy: main achievements and challenges. *Cell. Mol. Immunol.*, 18, 805-828.
- VENABLES, W. N. & RIPLEY, B. D. 2002. *Modern Applied Statistics with S*, Springer.
- VERHAGEN, L. M., ZOMER, A., MAES, M., VILLALBA, J. A., DEL NOGAL, B., ELEVELD, M., VAN HIJUM, S. A., DE WAARD, J. H. & HERMANS, P. W. 2013. A predictive signature gene set for discriminating active from latent tuberculosis in Warao Amerindian children. *BMC Genomics*, 14, 74.
- VILLANI, A. C., SATIJA, R., REYNOLDS, G., SARKIZOVA, S., SHEKHAR, K., FLETCHER, J., GRIESBECK, M., BUTLER, A., ZHENG, S., LAZO, S., JARDINE, L., DIXON, D., STEPHENSON, E., NILSSON, E., GRUNDBERG, I., MCDONALD, D., FILBY, A., LI, W., DE JAGER, P. L., ROZENBLATT-ROSEN, O., LANE, A. A., HANIFFA, M., REGEV, A. & HACOEN, N. 2017. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science*, 356.
- VIRATHAM PULSAWATDI, A., CRAIG, S. G., BINGHAM, V., MCCOMBE, K., HUMPHRIES, M. P., SENEVIRATHNE, S., RICHMAN, S. D., QUIRKE, P., CAMPO, L., DOMINGO, E., MAUGHAN, T. S., JAMES, J. A. & SALTO-TELLEZ, M. 2020. A robust multiplex immunofluorescence and digital pathology workflow for the characterisation of the tumour immune microenvironment. *Mol. Oncol.*, 14, 2384-2402.

- VOGEL, C. & MARCOTTE, E. M. 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.*, 13, 227-232.
- VOIGT, E., GRILL, D., ZIMMERMANN, M., SIMON, W., OVSYANNIKOVA, I., KENNEDY, R. & POLAND, G. 2018. Transcriptomic signatures of cellular and humoral immune responses in older adults after seasonal influenza vaccination identified by data-driven clustering. *Sci. Rep.*, 8, 739-739.
- VOLKMAN, H. E., CLAY, H., BEERY, D., CHANG, J. C. W., SHERMAN, D. R. & RAMAKRISHNAN, L. 2004. Tuberculous granuloma formation is enhanced by a Mycobacterium virulence determinant. *PLoS Biol.*, 2, 1946-1956.
- VON HERRATH, M. G. & NEPOM, G. T. 2005. Lost in translation: barriers to implementing clinical immunotherapeutics for autoimmunity. *J. Exp. Med.*, 202, 1159-62.
- WADA, M., YOSHIMI, K., HIGO, N., REN, Y.-R., MOCHIZUKI, H., MIZUNO, Y. & KITAZAWA, S. 2006. Statistical parametric mapping of immunopositive cell density. *Neurosci. Res.*, 56, 96-102.
- WAGAR, L. E., DIFAZIO, R. M. & DAVIS, M. M. 2018. Advanced model systems and tools for basic and translational human immunology. *Genome Med.*, 10, 73.
- WALLIS, R. S., KIM, P., COLE, S., HANNA, D., ANDRADE, B. B., MAEURER, M., SCHITO, M. & ZUMLA, A. 2013. Tuberculosis biomarkers discovery: developments, needs, and challenges. *Lancet Infect. Dis.*, 13, 362-72.
- WALSH, D. S., PORTAELS, F. & MEYERS, W. M. 2011. Buruli Ulcer: Advances in Understanding Mycobacterium ulcerans Infection. *Dermatol. Clin.*, 29, 1-8.
- WALZL, G., MCNERNEY, R., DU PLESSIS, N., BATES, M., MCHUGH, T. D., CHEGOU, N. N. & ZUMLA, A. 2018. Tuberculosis: advances and challenges in development of new diagnostics and biomarkers. *Lancet Infect. Dis.*, 18, e199-e210.
- WANG, N., LI, X., WANG, R. & DING, Z. 2021. Spatial transcriptomics and proteomics technologies for deconvoluting the tumor microenvironment. *Biotechnol. J.*, 16, e2100041.
- WANG, Y. J., TRAUM, D., SCHUG, J., GAO, L., LIU, C., ATKINSON, M. A., POWERS, A. C., FELDMAN, M. D., NAJI, A., CHANG, K. M. & KAESTNER, K. H. 2019. Multiplexed In Situ Imaging Mass Cytometry Analysis of the Human Endocrine Pancreas and Immune System in Type 1 Diabetes. *Cell Metab.*, 29, 769-783.e4.
- WANG, Z., GERSTEIN, M. & SNYDER, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.*, 10, 57-63.
- WARSINSKE, H. C., RAO, A. M., MOREIRA, F. M. F., SANTOS, P. C. P., LIU, A. B., SCOTT, M., MALHERBE, S. T., RONACHER, K., WALZL, G., WINTER, J., SWEENEY, T. E., CRODA, J., ANDREWS, J. R. & KHATRI, P. 2018. Assessment of Validity of a Blood-Based 3-Gene Signature Score for Progression and Diagnosis of Tuberculosis, Disease Severity, and Treatment Response. *JAMA Netw Open*, 1, e183779.
- WATTS, G. 2007. Nobel prize is awarded for work leading to "knockout mouse". *BMJ*, 335, 740.
- WEBSTER, S. J., BACHSTETTER, A. D., NELSON, P. T., SCHMITT, F. A. & VAN ELDIK, L. J. 2014. Using mice to model Alzheimer's dementia: an overview of the clinical disease and the preclinical behavioral changes in 10 mouse models. *Front. Genet.*, 5, 88.
- WELLS, G., GLASGOW, J. N., NARGAN, K., LUMAMBA, K., MADANSEIN, R., MAHARAJ, K., HUNTER, R. L., NAIDOO, T., COETZER, L., LE ROUX, S., DU PLESSIS, A. & STEYN, A. J. C. 2021. Micro-Computed Tomography Analysis of

- the Human Tuberculous Lung Reveals Remarkable Heterogeneity in Three-dimensional Granuloma Morphology. *Am. J. Respir. Crit. Care Med.*, 204, 583-595.
- WHO 2021. *Global Tuberculosis Report 2021*.
- WILLEMSSEN, M., KREBBERS, G., BEKKENK, M. W., TEUNISSEN, M. B. M. & LUITEN, R. M. 2021. Improvement of Opal Multiplex Immunofluorescence Workflow for Human Tissue Sections. *J. Histochem. Cytochem.*, 69, 339-346.
- WILLIAMS, A. & ORME, I. M. 2016. Animal Models of Tuberculosis: An Overview. *Microbiol Spectr.*, 4.
- WILSON, M. E., FINEBERG, H. V. & COLDITZ, G. A. 1995. Geographic latitude and the efficacy of bacillus Calmette-Guérin vaccine. *Clin. Infect. Dis.*, 20, 982-91.
- WIRAPATI, P., SOTIRIOU, C., KUNKEL, S., FARMER, P., PRADERVAND, S., HAIBEKAINS, B., DESMEDT, C., IGNATIADIS, M., SENGSTAG, T., SCHÜTZ, F., GOLDSTEIN, D. R., PICCART, M. & DELORENZI, M. 2008. Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. *Breast Cancer Res.*, 10, R65.
- WONG, H. R., CVIJANOVICH, N., ALLEN, G. L., LIN, R., ANAS, N., MEYER, K., FREISHTAT, R. J., MONACO, M., ODOMS, K., SAKTHIVEL, B. & SHANLEY, T. P. 2009. Genomic expression profiling across the pediatric systemic inflammatory response syndrome, sepsis, and septic shock spectrum. *Crit. Care Med.*, 37, 1558-66.
- WONG, H. R., CVIJANOVICH, N. Z., ANAS, N., ALLEN, G. L., THOMAS, N. J., BIGHAM, M. T., WEISS, S. L., FITZGERALD, J. C., CHECCHIA, P. A., MEYER, K., QUASNEY, M., HALL, M., GEDEIT, R., FREISHTAT, R. J., NOWAK, J., RAJ, S. S., GERTZ, S., GRUNWELL, J. R. & LINDSELL, C. J. 2017. Improved Risk Stratification in Pediatric Septic Shock Using Both Protein and mRNA Biomarkers. PERSEVERE-XP. *Am. J. Respir. Crit. Care Med.*, 196, 494-501.
- WOODS, C. W., MCCLAIN, M. T., CHEN, M., ZAAS, A. K., NICHOLSON, B. P., VARKEY, J., VELDMAN, T., KINGSMORE, S. F., HUANG, Y., LAMBKIN-WILLIAMS, R., GILBERT, A. G., HERO, A. O., III, RAMSBURG, E., GLICKMAN, S., LUCAS, J. E., CARIN, L. & GINSBURG, G. S. 2013. A Host Transcriptional Signature for Presymptomatic Detection of Infection in Humans Exposed to Influenza H1N1 or H3N2. *PLoS One*, 8, e52198.
- WU, T., HU, E., XU, S., CHEN, M., GUO, P., DAI, Z., FENG, T., ZHOU, L., TANG, W., ZHAN, L., FU, X., LIU, S., BO, X. & YU, G. 2021. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation*, 2, 100141.
- XIE, K., KONG, S., LI, F., ZHANG, Y., WANG, J. & ZHAO, W. 2020. Bioinformatics-Based Study to Investigate Potential Differentially Expressed Genes and miRNAs in Pediatric Sepsis. *Med. Sci. Monit.*, 26, e923881.
- YANG, C., YU, T., LIU, Z., YE, X., LIAO, X., WANG, X., HAN, C., ZHU, G., QIN, W. & PENG, T. 2019. Cystatin F as a key family 2 cystatin subunit and prognostic biomarker for earlystage pancreatic ductal adenocarcinoma. *Oncol. Rep.*, 42, 79-90.
- YANG, Y., WANG, H.-J., HU, W.-L., BAI, G.-N. & HUA, C.-Z. 2022. Diagnostic Value of Interferon-Gamma Release Assays for Tuberculosis in the Immunocompromised Population. *Diagnostics*, 12.
- YOUNG, D. 2009. Animal models of tuberculosis. *Eur. J. Immunol.*, 39, 2011-2014.
- YU, G., WANG, L.-G., YAN, G.-R. & HE, Q.-Y. 2015. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics*, 31, 608-609.
- YU, Y., TSANG, J. C., WANG, C., CLARE, S., WANG, J., CHEN, X., BRANDT, C., KANE, L., CAMPOS, L. S., LU, L., BELZ, G. T., MCKENZIE, A. N., TEICHMANN, S. A., DOUGAN, G. & LIU, P. 2016. Single-cell RNA-seq identifies

- a PD-1(hi) ILC progenitor and defines its development pathway. *Nature*, 539, 102-106.
- ZAAS, A. K., CHEN, M., VARKEY, J., VELDMAN, T., HERO, A. O., 3RD, LUCAS, J., HUANG, Y., TURNER, R., GILBERT, A., LAMBKIN-WILLIAMS, R., ØIEN, N. C., NICHOLSON, B., KINGSMORE, S., CARIN, L., WOODS, C. W. & GINSBURG, G. S. 2009. Gene expression signatures diagnose influenza and other symptomatic respiratory viral infections in humans. *Cell Host Microbe*, 6, 207-17.
- ZAK, D. E., PENN-NICHOLSON, A., SCRIBA, T. J., THOMPSON, E., SULIMAN, S., AMON, L. M., MAHOMED, H., ERASMUS, M., WHATNEY, W., HUSSEY, G. D., ABRAHAMS, D., KAFAAR, F., HAWKRIDGE, T., VERVER, S., HUGHES, E. J., OTA, M., SUTHERLAND, J., HOWE, R., DOCKRELL, H. M., BOOM, W. H., THIEL, B., OTTENHOFF, T. H. M., MAYANJA-KIZZA, H., CRAMPIN, A. C., DOWNING, K., HATHERILL, M., VALVO, J., SHANKAR, S., PARIDA, S. K., KAUFMANN, S. H. E., WALZL, G., ADEREM, A. & HANEKOM, W. A. 2016. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *Lancet*, 387, 2312-2322.
- ZHANG, G., DEWEERD, N. A., STIFTER, S. A., LIU, L., ZHOU, B., WANG, W., ZHOU, Y., YING, B., HU, X., MATTHEWS, A. Y., ELLIS, M., TRICCAS, J. A., HERTZOG, P. J., BRITTON, W. J., CHEN, X. & FENG, C. G. 2018. A proline deletion in IFNAR1 impairs IFN-signaling and underlies increased resistance to tuberculosis in humans. *Nat. Commun.*, 9, 85.
- ZHENG, G. X. Y., TERRY, J. M., BELGRADER, P., RYVKIN, P., BENT, Z. W., WILSON, R., ZIRALDO, S. B., WHEELER, T. D., MCDERMOTT, G. P., ZHU, J., GREGORY, M. T., SHUGA, J., MONTESCLAROS, L., UNDERWOOD, J. G., MASQUELIER, D. A., NISHIMURA, S. Y., SCHNALL-LEVIN, M., WYATT, P. W., HINDSON, C. M., BHARADWAJ, R., WONG, A., NESS, K. D., BEPPU, L. W., DEEG, H. J., MCFARLAND, C., LOEB, K. R., VALENTE, W. J., ERICSON, N. G., STEVENS, E. A., RADICH, J. P., MIKKELSEN, T. S., HINDSON, B. J. & BIELAS, J. H. 2017. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.*, 8, 14049.
- ZHOU, G., LUO, Q., LUO, S., TENG, Z., JI, Z., YANG, J., WANG, F., WEN, S., DING, Z., LI, L., CHEN, T., ABI, M.-E., JIAN, M., LUO, L., LIU, A. & BAO, F. 2020. Interferon- γ release assays or tuberculin skin test for detection and management of latent tuberculosis infection: a systematic review and meta-analysis. *Lancet Infect. Dis.*, 20, 1457-1469.
- ZHOU, X., WANG, X., HUANG, K., LIAO, X., YANG, C., YU, T., LIU, J., HAN, C., ZHU, G., SU, H., QIN, W., HAN, Q., LIU, Z., HUANG, J., GONG, Y., YE, X. & PENG, T. 2019. Investigation of the clinical significance and prospective molecular mechanisms of cystatin genes in patients with hepatitis B virus-related hepatocellular carcinoma. *Oncol. Rep.*, 42, 189-201.
- ZHU, Q., WONG, A. K., KRISHNAN, A., AURE, M. R., TADYCH, A., ZHANG, R., CORNEY, D. C., GREENE, C. S., BONGO, L. A., KRISTENSEN, V. N., CHARIKAR, M., LI, K. & TROYANSKAYA, O. G. 2015. Targeted exploration and analysis of large cross-platform human transcriptomic compendia. *Nat. Methods*, 12, 211-214.
- ZSCHALER, J., SCHLORKE, D. & ARNHOLD, J. 2014. Differences in innate immune response between man and mouse. *Crit. Rev. Immunol.*, 34, 433-54.
- ZWERLING, A., BEHR, M. A., VERMA, A., BREWER, T. F., MENZIES, D. & PAI, M. 2011. The BCG World Atlas: a database of global BCG vaccination policies and practices. *PLoS Med.*, 8, e1001012.