IMMUNOGENOMIC ANALYSIS OF THE TUMOUR MICROENVIRONMENT: FOCUSING ON MACROPHAGES



Yuechuan Chi

St. Edmund Hall

Department of Oncology

University of Oxford

This dissertation is submitted for the degree of Doctor of Philosophy

Trinity 2022

To dreamers who dare to think differently and change the world

"To live on a day-to-day basis is insufficient for human beings; we need to transcend, transport, escape; we need meaning, understanding, and explanation; we need to see over-all patterns in our lives. We need hope, the sense of a future. And we need freedom (or, at least, the illusion of freedom) to get beyond ourselves, whether with telescopes and microscopes and our ever-burgeoning technology, or in states of mind that allow us to travel to other worlds, to rise above our immediate surroundings."

- *Oliver Sacks*

DECLARATION

I declare that no part of this thesis has been accepted or submitted for any degree, diploma, certificate or other qualification at this University or elsewhere. This thesis is entirely my own work, and except where otherwise stated, describes my own research.

Bill Chi

Oxford

ABSTRACT

Immune signatures are important tools to investigate the prognostic significance and therapeutic possibilities of different immune populations in the context of the tumour microenvironment (TME). Accompanied by the wealth of cancer transcriptomic data, many immune signatures have been developed using different samples, technologies, and methods. But there has been limited effort to comprehensively assess these tools in cancer, especially in the context of macrophage signatures.

In this study, I first exanimated the performance of six bioinformatic immune dissection tools (Chapter 3). I conducted in-depth comparisons on two of the most comprehensive and methodologically distinctive tools (CIBERSORT and xCell) in 33 different cancer types across the Cancer Genome Atlas (TCGA). The effort identified high agreements on five immune subsets and an unexpected positive relationship between the theoretically polarized M1 and M2 signatures. Subsequently, I assessed the impact of tumour hypoxia on infiltrating immune cells (Chapter 4), since hypoxia is a hallmark of the immunosuppressive microenvironment. I uncovered a reciprocal relationship between tumour hypoxia and the stromal contents within the breast cancer TME.

In the work concerned with tumour associated macrophage (TAM) (Chapter 5), I systematically compared 16 different *ex vivo* and *in vivo* derived macrophage signatures. This comparison indicated these macrophage signatures have high divergence and the *ex vivo* derived signatures were not representative of TAMs in breast cancer. Following this, I derived a new independent TAM signature from single-cell data, which featured lipid metabolism and lysosomal activities. I further refined and improved this signature using the least absolute shrinkage and selection operator (LASSO). Using the refined TAM signature, I demonstrated the lipid-associated macrophages were immunosuppressive, enriched in aggressive HER2-enriched and Basal-like subtypes, and associated with poor prognoses across breast cancer subtypes. Therefore, this lipid-associated TAM could be a promising target in breast cancer.

AIMS

1. Examine the performance of different digital dissection tools (Chapter 3)
2. Assess the impact of tumour hypoxia on infiltrating immune cells (Chapter 4)
3. Compare the existing *ex vivo* and *in vivo* derived macrophage signatures (Chapter 5)
4. Derive improved tumour associated macrophage transcriptomics signatures (Chapter 5)
5. Evaluate the spatial patterns and functions of macrophages within tumour samples (Chapter 5)
6. Estimate the clinical implications of macrophage in breast cancer: survival and target (Chapter 5)

ACKNOWLEDGEMENTS

I would like to thank everyone that I have encountered during my time in Oxford. You have taught me important lessons and have profoundly shaped how I think about life and science. Oxford will always be a special place in my heart.

I would like to thank my supervisors Professor Adrian Harris and Professor Francesca Buffa for their guidance throughout the DPhil journey and for helping me to become a better scientist. Words cannot express my gratitude for your efforts and mentorship. Some of my fondest memories took place during our office meetings. Before COVID when we could still meet face-to-face, I would come to Adrian's office at the WIMM and go over my new data together. The meetings usually lasted around two hours and we spent a lot of time going through the details and the individual genes together. His ability to see patterns from noise and link new findings with the literature has never failed to astonish me. I also remember queuing up with colleagues outside of Francesca's office to compete for a spot to meet with her, since she's always in high demand. The learning environment was truly intellectually stimulating and enjoyable.

I would like to thank Professor Tim Maughan for inviting me to shadow him in the oncology clinic. The clinical experience added a new practical dimensionality to my research and created a sense of urgency to continuously improve the standard of care with new tools and therapeutics. I would like to express my gratitude to my collaborators: Professor Kim Midwood and Anis Gammage for their advice on macrophage biology.

To my friends: Tiffany, Hannah, Alessandro, Philip, Wei-Chen, Matteo, Joachim, Ben, Fiona, May, Clara, Crystal, Ava, Pedro, Karen, Susan, Andrew, Elliot, James, Cristiano, Selena, Samuel, Linda, Shudong, Pauline, Yizhou, Jessica, Katherine, Kay, Syed, Mattia, Dimitris, Benedict, Teofanija, Jade, Debra, Biao, Clive and many others. Thank you for the incredible memories in Oxford!

I would like to thank my family, especially my mother Wenxia, for their unconditional support and belief in me. You have always encouraged me to follow my passions and dare to dream big.

TABLE OF CONTENTS

1 INTRODUCTION **1**

1.1 THE TUMOUR MICROENVIRONMENT AND ITS KEY COMPONENTS 2

1.1.1 Lymphocytes *5*

1.1.2 Myeloid cells *12*

1.1.3 Hypoxic microenvironment *22*

1.2 CANCER IMMUNOGENOMICS 27

1.2.1 Characterization of tumour infiltrating immune cells (TIICs) *31*

1.2.2 Assumptions of estimating immune infiltrates using transcriptomics *36*

1.2.3 Validations of the immune infiltration estimations *37*

2 DATASETS AND METHODS **40**

2.1 DATASETS 40

2.1.1 Bulk transcriptomic data *40*

2.1.2 Single-cell transcriptomic data *41*

2.2 GENE SIGNATURES 45

2.2.1 Macrophage immune signatures *45*

2.2.2 Calculating gene signature scores *46*

2.3 SINGLE-CELL TRANSCRIPTOMIC ANALYSIS 47

2.3.1 Data processing *47*

2.3.2 Clustering *47*

2.3.3 Deriving TAM signatures *49*

2.4 REFINING MACROPHAGE SIGNATURE USING LASSO 50

2.5 SURVIVAL ANALYSIS 51

2.6 STATISTICAL ANALYSIS 52

3 ASSESSMENT OF TIICS ESTIMATING TOOLS IN TCGA **53**

3.1 IMMUNOGENOMIC TIICS ESTIMATING METHODS AVAILABLE 53

3.3 RESULTS AND DISCUSSION 56

3.3.1 Assessment plan of immunogenomic dissection methods *56*

3.3.2 Multiple collinearity check of CIBERSORT and xCell *56*

3.3.3 Assessment of CIBERSORT and xCell in TCGA *61*

3.3.4 Unexpected correlations on macrophage M1 and M2 estimations *64*

3.3.5 ConsensusTME *64*

3.4 CONCLUSION 65

4 ASSOCIATION BETWEEN HYPOXIA AND IMMUNE POPULATIONS IN BREAST CANCER 68

4.1 HYPOXIA AND THE TUMOUR MICROENVIRONMENT 68

4.1.1 Establishing hypoxia score from hypoxia signature in breast cancer *68*

4.1.2 Hypoxia levels in PAM50 and triple-negative breast cancer subtypes *70*

4.2 IMMUNE CELL COMPOSITIONS AND HYPOXIA IN BREAST CANCER 73

4.2.1 Immune compositions in PAM50 subtypes *73*

4.2.2 Immune compositions in TNBC *73*

4.3 STROMAL AND IMMUNE LANDSCAPE IN HYPOXIC ENVIRONMENT 78

4.3.1 Relationship between hypoxia, stroma and immune cells in breast cancer *78*

4.3.2 Relationship between hypoxia, macrophages and mast cells *82*

4.4 PATHWAY ANALYSIS OF HYPOXIA ASSOCIATED GENES IN BREAST CANCER 84

4.4.1 Identify hypoxia associated genes in BRCA *84*

4.4.2 Protein-protein interaction and KEGG analyses of hypoxia signature *85*

4.4.3 Immune GO Pathways in breast cancer *85*

4.5 CONCLUSION 88

5 MACROPHAGE SIGNATURE DERIVED FROM SINGLE-CELL SEQUENCED TUMOUR ASSOCIATED MACROPHAGES 89

5.1 INTRODUCTION 89

5.2 RESULTS 91

5.2.1 TAM and M1/M2 signatures share few genes in common *91*

5.2.2 Published macrophage signatures do not support M1/M2 dichotomy in breast cancer *94*

5.2.3 Tumour associated macrophage and monocyte diversity in breast cancer *101*

5.2.4 Alternatively activated macrophages in breast cancer favoured lipid metabolism *105*

5.2.5 Lipid-associated macrophage spatially concentrated within the invasive front *110*

5.2.6 Lipid-associated macrophages have high expression in TNBC, HER2 & Hypoxic tumours *112*

5.2.7 Lipid-associated TAM signature is associated with poor prognosis in breast cancer *114*

5.2.8 LASSO refined lipid-associated macrophage signature is more prognostic than HER2 *117*

5.2.9 Lipid-associated TAM associated with poor prognoses in hypoxic, ER, PR, and HER2 tumours *123*

5.2.10 Lipid-associated macrophages are associated with low anti-tumour immunity *125*

5.3 DISCUSSION 128

5.3.1 Summary of results *128*

5.3.2 M1/M2 ex vitro models are less relevant in tumour pathology *129*

5.3.3 Source of tumour-associated macrophages *129*

5.3.4 Tumour associated macrophage and estrogen signalling *130*

5.3.5 Lipid associated macrophage – linking obesity and cancer *131*

5.3.6 Prognostic value of LA-TAM signature *132*

5.4 CONCLUSION 132

6 SUMMARY AND CONCLUSIONS **133**

6.1 SUMMARY 133

6.2 TARGETING TAM IN CANCER 135

6.3 CLUSTER STABILITY AND REPRODUCIBILITY IN THE SINGLE-CELL ERA 136

6.4 CONCLUSION 137

7 APPENDICES **138**

APPENDIX 1: CELLS AND MARKERS USED TO DEVELOP IN SILICO METHODS **139**

APPENDIX 2: HYPOXIA LEVEL IN DIFFERENT BREAST CANCER SUBTYPES **140**

APPENDIX 3: RELATIONSHIP BETWEEN PAM50 SUBTYPES AND TRIPLE-NEGATIVE BREAST CANCER **141**

APPENDIX 4: HYPOXIA CORRELATION WITH 64 CELL TYPES IN BREAST CANCER **142**

8 REFERENCES **143**

LIST OF TABLES

TABLE 3.1 CURRENT TIICS ESTIMATING METHODS IN THE LITERATURE 55

TABLE 3.2 SIZE OF CORRELATION AND INTERPRETATION 57

TABLE 4.1: SUMMARY OF SELECTED GENE LISTS FOR PATHWAY ANALYSES 84

TABLE 5.1 POOLED RESULTS OF MULTIVARIATE SURVIVAL ANALYSES USING MACROPHAGE SIGNATURES 120

TABLE 5.2 POOLED RESULTS OF MULTIVARIATE SURVIVAL ANALYSES IN NKI BREAST CANCER DATA 122

List of Figures

Figure 1.1 Overview of the immune landscape within the tumour microenvironment. 4

Figure 1.2 Illustration of T cell activation and antitumour response. 7

Figure 1.3 Immune cell functions and adaptations in the tumour microenvironment. 10

Figure 1.4 Evolving model of macrophage activation, from M1/M2 to TAM. 16

Figure 1.5 Hypoxia inducible factor (HIF) oxygen dependent post-translational regulations 26

Figure 1.6 Relationship between data type and immunogenomic applications. 30

Figure 1.7 An illustration of gene signature method. 34

Figure 1.8 An illustration of deconvolution approach to characterise immune cells in tumour 35

Figure 1.9 An illustration of generating immune profiles from human PBMC 39

Figure 2.1 Overview of the droplet-based single-cell sequencing platforms 44

Figure 3.1 CIBERSORT multicollinearity assessment in 11,000 TCGA samples. 59

Figure 3.2 xCell multicollinearity assessment in 11,000 TCGA samples 60

Figure 3.3 CIBERSORT and xCell comparison in 11,000 TCGA samples. 62

Figure 3.4 CIBERSORT and xCell comparison in TCGA breast cancer samples. 63

Figure 3.5 Positive correlations observed among macrophage M1 and M2 estimations. 66

Figure 3.6 Macrophage canonical markers and their expression in TCGA breast cancer. 67

Figure 4.1 Hypoxia scores and their correlations with hypoxic reference gene CA9. 69

Figure 4.2 Triple-negative samples have the highest hypoxia level in breast cancer samples. 71

Figure 4.3 Relationship between TNBC and PAM50 in TCGA breast cancer samples. 72

Figure 4.4 Variations of hypoxia levels in PAM50 subtypes. 72

Figure 4.5 Relationships between PAM50 subtypes and immune cell compositions. 75

Figure 4.6 Relationships between triple-negative breast cancer and immune compositions. 76

Figure 4.7 Relationships between hypoxia and immune compositions in triple-negative breast cancer. 77

Figure 4.8 Cos^2 variable analysis of the first 16 PCs 80

Figure 4.9 Variable plot of PC1 and PC2 from xCell estimations in BRCA 81

Figure 4.10 Association between mast cell and breast cancer hypoxia 83

Figure 4.11 STRING predictions protein-protein with genes in the hypoxia signature 86

Figure 4.12 Hypoxia associated immune pathways in breast cancer 87

Figure 5.1 Comparison of 16 publicly available macrophage signatures. 93
Figure 5.2 Comparing 16 macrophage signatures in bulk and single-cell RNA-seq data. 96
Figure 5.3 Macrophage signature expression in bulk RNA-seq data. 97
Figure 5.4 Macrophage signature expression in single-cell RNA-seq data. 99
Figure 5.5 Establishing macrophage clusters with consensus clustering method. 103
Figure 5.6 Integrating and mapping the newly derived four monocyte and macrophage clusters in another breast cancer single-cell dataset. 107
Figure 5.7 Functional characterization of the two TAMs. 109
Figure 5.8 Spatial distributions of the two TAMs. 111
Figure 5.9 Expressions and associations of the two TAMs in TCGA breast samples with respect to PAM50, ER, PR, HER2 and hypoxia. 113
Figure 5.10 Pooled univariate proportional hazards regression results from 18 different macrophage signatures and four reference signatures (ER, PR, HER2 and hypoxia) 115
Figure 5.11 Multivariate proportional hazards regression model of lipid-associated TAM signature in METABRIC. 116
Figure 5.12 LASSO refinement of the LA-TAM signature. 119
Figure 5.13 Verifications of LA-TAM signature in independent datasets. 121
Figure 5.14 Impact of LA-TAM on overall survival in four different stratifications 124
Figure 5.15 Functional characterisation of LASSO refined LA-TAM signature. 127

List of Abbreviations and Acronyms

ACP2	Acid Phosphatase 2
ADAMDEC1	ADAM (a Disintegrin and Metalloproteinase) like decysin 1
ADCC	Antibody-Dependent Cellular Cytotoxicity
ADM	Adrenomedullin
AJCC	American Joint Committee on Cancer
ANOVA	Analysis of Variance
APC	Antigen-Presenting Cell
APOC1	Apolipoprotein C1
APOE	Apolipoprotein E
ARGI	Arginase I
AUC	Area Under the Curve
BISCUIT	Bayesian Inference for Single-cell Clustering and Imputing Algorithm
BRCA	Breast Invasive Carcinoma Cohort (TCGA)
CAIX	Carbonic anhydrase IX or CA9 gene
CCL	C-C Motif (Chemokine) Ligand
CCR	C-C Motif (Chemokine) Receptor
CD	Cluster of Differentiation
CDF	Cumulative Distribution Function
CEP17	Chromosome Enumeration Probe 17
CHI3L1	Chitinase 3 Like 1
CIBERSORT	Cell-type Identification by Estimating Relative Subsets of RNA Transcripts
CMP	Common Myeloid Progenitor
CNV	Copy Number Variation
CSF1	Macrophage Colony Stimulating Factor 1
CSF1R	Macrophage Colony Stimulating Factor 1 Receptor
CTL	Cytotoxic T Lymphocyte
CTLA4	Cytotoxic T Lymphocyte Antigen 4
CTS	Cathepsin
CXCL1	C-X-C Motif (Chemokine) Ligand
CXCR	C-X-C Motif (Chemokine) Receptor
DC	Dendritic Cell

DCIS	Ductal Breast Carcinoma in situ
DGE	Differential Gene Expression
DLL4	Delta-like 4
DNA	Deoxyribonucleic Acid
ECM	Extracellular Matrix
EGF	Epidermal Growth Factor
ENCODE	Encyclopedia of DNA Elements Project
ER	Estrogen Receptor (ESR1)
ERBB2	Erb-B2 Receptor Tyrosine Kinase 2 or HER2
ERK2	Extracellular Receptor Kinase 2
ESMO	European Society for Medical Oncology
ESR1	Estrogen Receptor 1 (ER)
F13A1	Coagulation Factor XIII A Chain
FACS	Fluorescence-Activated Cell Sorting
FANTOM5	Functional Annotation Of the Mammalian Genome Project
FCGR3A	Fc Gamma Receptor IIIa or CD16
FDA	Food and Drug Administration
FDX1	Ferredoxin 1
FISH	Fluorescence in situ Hybridization
fMRI	Functional Magnetic Resonance Imaging
FN1	Fibronectin Type I
GEO	Gene Expression Omnibus
GEP	Gene Expression Profile
GLUT	Glucose Transporter Protein
GM-CSF	Granulocyte-Macrophage Colony Stimulating Factor
GO	Gene Ontology
GSEA	Gene Set Enrichment Analysis
GSVA	Gene Set Variation Analysis
HER2	Human Epidermal Growth Factor Receptor 2
HIF	Hypoxia Inducible Factor
HLA	Human Leukocyte Antigen
HPC	High Performance Computing
HR	Hazard Ratio
HSC	Haematopoietic Stem Cells

ICGC	The International Cancer Genome Consortium
ICOS	Inducible T Cell Costimulator
IFNγ	Interferon Gamma
IHC	Immunohistochemistry
IL	Interleukin
IRIS	Immune Response in silico Project
KEGG	Kyoto Encyclopedia of Genes and Genomes
KNN	k-Nearest Neighbour
LASSO	Least Absolute Shrinkage and Selection Operator
LDHA	Lactate Dehydrogenase A
LIPA	Lipase A
LM22	Leukocyte Gene Expression Matrix 22
LPL	lipoprotein lipase
LPS	Lipopolysaccharide
LUSC	lung squamous cell carcinoma
M0	Resting Macrophage
M1	Macrophage M1
M2	Macrophage M2
MAPK	Mitogen-Activated Protein Kinases
MARCO	Macrophage Receptor With Collagenous Structure
MCT4	Monocarboxylate Transporter 4
MDSC	Myeloid-Derived Suppressor Cell
METABIRC	Molecular Taxonomy of Breast Cancer International Consortium
MHC	Major Histocompatibility Complex
MIP-1α	macrophage Inflammatory Protein 1-alpha (CCL3)
MLE	Maximum Likelihood Estimation
MMP	Matrix Metalloproteinase
NGS	Next-Generation Sequencing
NK	Natural Killer Cell
NKG2D	Natural Killer Group 2D
NOS	Nitric Oxide Synthase
OXPHOS	Oxidative Phosphorylation
PAM50	Predictor Analysis of Microarray 50
PBMC	Peripheral Blood Mononuclear Cells

PC	Principal Component
PCR	Polymerase Chain Reaction
PD1	Programmed Cell Death Protein 1
PHD	Prolyl Hydroxylase
PPAR	Peroxisome Proliferator-Activated Receptors
PR	Progesterone Receptor
QC	Quality Control
RNA	Ribonucleic Acid
ROS	Reactive Oxygen Species
RSEM	RNAseq by Expectation-Maximization Algorithm
SCARB3	scavenger receptor class B member 3 (CD36)
SEMA3A	Semaphorin 3A
SIGLEC15	Sialic Acid Binding Ig Like Lectin 15
SIRP	Signal Regulatory Protein α
SLAMF8	SLAM Family Member 8
SNP	Single Nucleotide Polymorphism
SPP1	Secreted Phosphoprotein 1 (Osteopontin)
STAT	Signal Transducer And Activator Of Transcription
STRING	Search Tool for the Retrieval of Interacting Genes
TAA	Tumour-Associated Antigens
TAM	Tumour-Associated Macrophage
TBK1	TANK Binding Kinase 1
TCGA	The Cancer Genome Atlas
TCR	T Cell Receptor
TDLN	Tumor-Draining Lymph Nodes
TERM2	Triggering Receptor Expressed on Myeloid Cells 2
TGFβ	Transforming Growth Factor Beta
TIIC	Tumour Infiltrating Immune Cells
TIL	Tumour-Infiltrating Lymphocyte
TLR	Toll-Like Receptor
TMB	Tumour Mutation Burden
TME	Tumour Microenvironment
TNBC	Triple-Negative Breast Cancer
TNF	Tumour Necrosis Factor

TNFRSF1A	Tumour necrosis factor receptor 1
TREM2	Triggering Receptor Expressed on Myeloid Cells 2
UMAP	Uniform Manifold Approximation and Projection
UMI	Unique Molecular Identifier
UV	Ultraviolet light
VEGF	Vascular Endothelial Growth Factor
VHL	Von Hippel-Lindau
VSIG4	V-Set And Immunoglobulin Domain Containing 4
WES	Whole-Exome Sequencing
WGS	Whole-Genome Sequencing
XBP1	X-box Binding Protein 1
YKL-40	Chitinase 3-like 1

List of Appendices

Appendix 1: Cells and markers used to develop in silico methods 139

Appendix 2: Hypoxia level in different breast cancer subtypes 140

Appendix 3: Relationship between PAM50 subtypes and triple-negative breast cancer 141

Appendix 4: Hypoxia correlation with 64 cell types in breast cancer 142

1 INTRODUCTION

Cancer is a leading cause of death in the world. It has been estimated that in the year 2020, there has been 19.3 million new cancer cases and 10 million cancer-related deaths globally (Sung et al., 2021). The word cancer refers to a heterogenous set of complex diseases that begin with genetic alternations. The majority of cancer cases (90-95%) have resulted from accumulations of deleterious somatic mutations in cells, which are genomic alterations that occurred post conception (Macintyre et al., 2018). In contrast, germline mutations that are inherited from generation to generation only constitute 5-10% of overall cancer cases (Anand et al., 2008). Accumulation of somatic mutations occurs throughout life (Macintyre et al., 2018; Martincorena et al., 2018). Whilst most of the somatic mutations are harmless, occasionally a critical gene and its function are altered. This alteration increases the evolutionary fitness of mutant cells and subsequent deleterious mutations in these cells may shift them toward becoming neoplastic cells (Anglesio et al., 2017; Aubert and Lansdorp, 2008; Martincorena et al., 2017). These somatic mutations are the consequences of environmental interactions over the ageing process. Some of the key risks of cancer have been linked with lifestyle factors (e.g. alcohol, smoking, western diet, obesity, inactivity, UV exposure) (Anand et al., 2008; Gupta et al., 2018; Kennedy et al., 2012). Furthermore, both the incidence rate and mortality of cancer are on the rise due to improvements in socioeconomic developments and the ageing populations (Edwards et al., 2010; Keum and Giovannucci, 2019; UNDP, 2019). It is expected the number of new

cases will increase by nearly 50% in the next 20 years to 28.4 million per year, which would become a substantial disease burden for many emerging economies when their main workforces enter into older age (Sung et al., 2021). Therefore, there is an urgent need to improve the current diagnostic and treatment options to fulfil the unmet needs. Accumulating evidence suggests targeting the tumour microenvironment and its components could provide novel solutions to ease the disease burden of cancer (Ho et al., 2020; Singh and Singh, 2022; Smyth et al., 2016; Ward et al., 2013; Yang et al., 2013).

1.1 The tumour microenvironment and its key components

The concepts of the tumour microenvironment and tumour immunology have been intertwined since its inception, dating back to 1863 when a German polymathic physician Rudolf Virchow reported leukocyte infiltration within solid tumours and hypothesised that chronic inflammatory environment would lead to carcinogenesis (Maman and Witz, 2018; Schmidt and Weber, 2006). Following this initial observation, in 1889, Stephen Paget reported the distributions of 735 metastatic breast cancer and concluded that distal metastases (i.e. “seed”) do not occur randomly, but rather they are sustained by a specific environment (i.e. “soil”) (Paget, 1889). This “seed and soil” principle proposed by Paget has also been corroborated in primary neoplastic lesions and their tumour microenvironment (Witz and Levy-Nissenbaum, 2006). A few years later, in 1891, an American physician William Bradley Coley, inspired by spontaneous regression of Streptococcus infected sarcoma patients, began to treat sarcoma and other solid tumours with a cocktail of attenuated *Streptococcus pyogenes* and *Serratia marcescens*, but the outcomes were mixed (Richardson et al., 1999). Coley hypothesized his bacterial cocktail could stimulate the immune system and eliminate cancer (Coley, 1991; Kucerova and Cervinkova, 2016). His attempt was also recognized as an initial attempt of treating cancer patients using immunotherapy (Carlson et al., 2020).

However, despite the early discoveries by these trailblazers, much of the cancer research and therapeutic developments until the 1980s have been dominated by cancer-centric models (Nguyen et

al., 2012; Vogelstein and Kinzler, 2004). These simplified models facilitated the discovery of oncogenes and tumour suppressors and these models also helped to uncover important cellular and molecular mechanisms that are shared across different cancer types. Seminal work from Hanahan and Weinberg have theorized these convergent features of cancer into six key hallmarks, which included 1) sustaining proliferative signalling, 2) evading growth suppressors, 3) enabling replicative immortality, 4) activating invasion and metastasis, 5) inducing angiogenesis, and 6) resisting cell death (Hanahan and Weinberg, 2000). More recently, the cancer-centric view has gradually been expanded to include other cells and conditions within the tumour microenvironment (TME) due to their roles in regulating tumour progression, especially the immune populations, which can serve both tumour-promoting and tumour-eliminating roles (Hanahan and Weinberg, 2011; Kroemer et al., 2015; Liotta and Kohn, 2001). As a result of an improved understanding of tumour biology and the microenvironment, eight additional hallmarks have been added that include 7) avoiding immune destruction, 8) tumour-promoting inflammation, 9) genome instability, 10) deregulating cellular metabolism, 11) non-mutational epigenetic reprogramming, 12) polymorphic microbiomes, 13) senescent cells, and 14) unlocking phenotypic plasticity (Hanahan, 2022).

TME is highly heterogeneous and it contains a mixture of cellular and non-cellular components (Figure 1.1). Besides the neoplastic cells within tumours, the rest of the cells can be generally classified into three groups: immune cells, endothelial cells and stromal cells (e.g. adipocytes, fibroblasts) (Anderson and Simon, 2020). The non-cellular components include cellular secreted products (e.g. growth factors, cytokines), extracellular matrix (ECM), and other metabolites (e.g. hydrogen peroxide) (Maman and Witz, 2018). Furthermore, environmental conditions such as hypoxia and acidosis are also hallmarks of TME (Corbet and Feron, 2017a; Wilson and Hay, 2011).

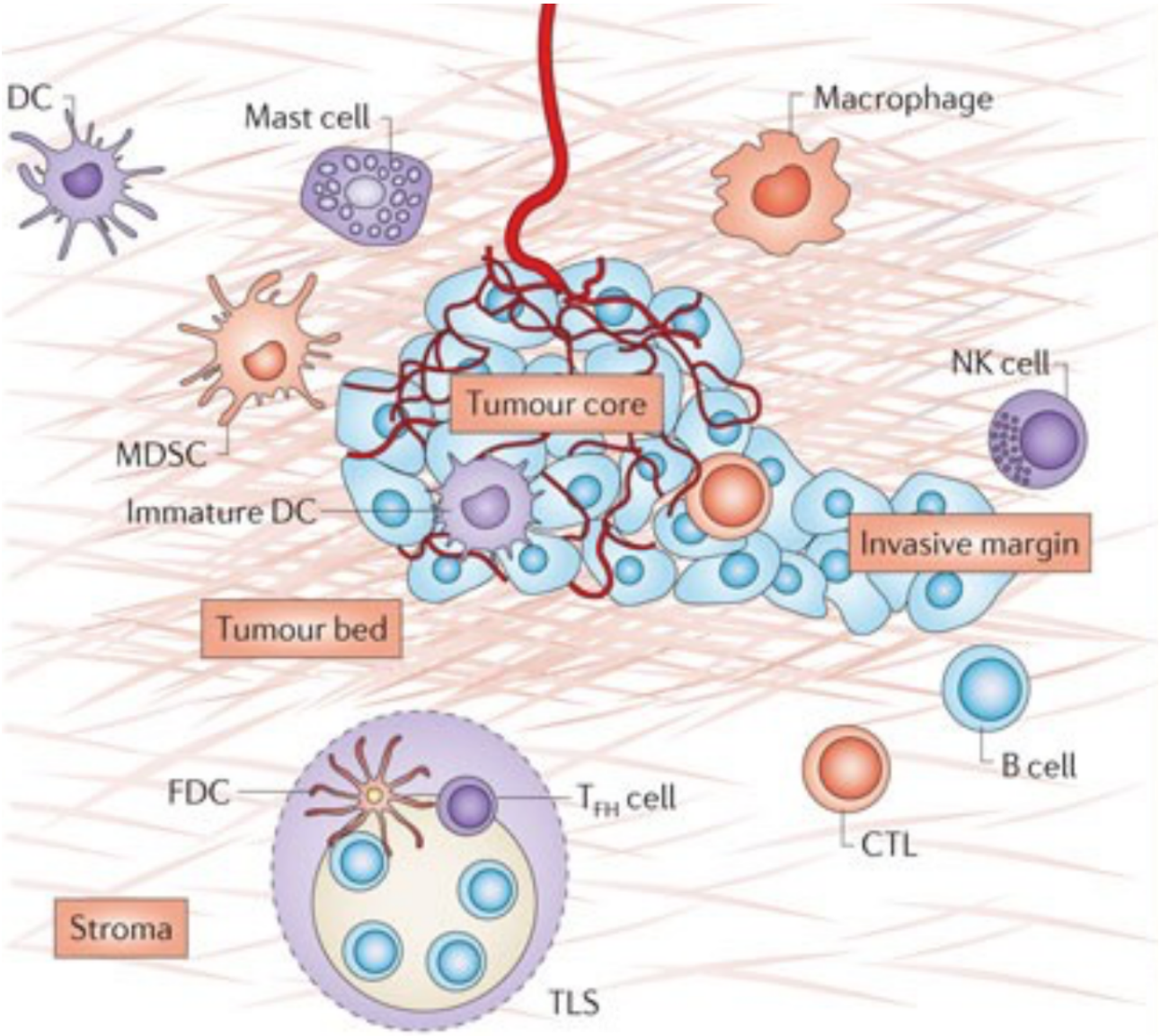


Figure 1.1 Overview of the immune landscape within the tumour microenvironment.

The anatomical structure of a tumour can be divided into several components, which include the tumour core, the invasive margin, and tertiary lymphoid structures. Within the tumour, diverse sets of immune cells exist that have distinctive spatial distributions and serve different functions. Besides the immune compartment, hypoxia, metabolism (e.g. glycolysis, lactate), and acidosis are also hallmarks of the tumour microenvironment that negatively regulates immunity (Jin and Jin, 2020). Abbreviations: CTL = Cytotoxic T lymphocyte; DC = Dendritic Cell; FDC = Follicular Dendritic Cell; MDSC = Myeloid-Derived Suppressor Cell; NK = Natural Killer Cell. Figure adapted from Fridman et al 2012.

1.1.1 Lymphocytes

Lymphocytes are white blood cells in the immune system and they can be categorised into three main groups: T cells, B cells and natural killer cells (NK). Today we already know that infiltrating CD8+ cytotoxic T lymphocyte (CTL) is the main anti-tumour effector cell type within TME, but this concept only emerged in 1975 when the presence of T lymphocytes in breast cancer and adjacent lymph nodes was initially reported (Richters and Kaspersky, 1975). Soon after the discovery of tumour infiltrating T cells, subsequent studies then demonstrated these T cells can have both anti-tumour and immune suppressive properties (Theodor Brunner et al., 1981; Vose and Moore, 1985; Vose et al., 1977). Furthermore, it was also observed that the cytotoxicity of tumour infiltrating T-lymphocytes was reduced than the T cells found from distant sites (Chiou et al., 2005). However, the link between immune suppression, immune evasion and TME was only established a decade later (Drake et al., 2006; Fischer et al., 2000; Gajewski et al., 2006).

1.1.1.1 T cells

Despite the challenges presented by the TME, the immune system can still establish an anti-tumour response that is centred around antigen-presenting cells (APCs) and CTLs (Figure 1.2). To achieve this, APCs such as dendric cells (DCs) infiltrate into the TME to take up tumour-associated antigens (TAAs) from dying cancer cells (Kroemer et al., 2015), followed by maturation and migration of DCs into the tumour-draining lymph nodes (TDLNs) (Wculek et al., 2020). Within the TDLNs, DCs activated antigen-specific CD8 T cells through TAA cross-presentation via the major histocompatibility complex class I (MHC I) and T cells receptor (TCR). These activated T cells then undergo proliferation and differentiation, then migrate to TME to target cancer cells that express the same TAAs by releasing cytotoxic granules enclosing granzymes and perforin (Chen and Mellman, 2013).

The activation of CTLs not only can be stimulated by the cross-presentation of TAAs (signal 1), but also regulated by costimulatory signals (signal 2) from co-receptors such as CD28, inducible costimulatory (ICOS), cytotoxic T-lymphocyte antigen-4 (CTLA-4), programmed death (PD1) (Rudd

et al., 2009). For example, CD28 and CTLA-4 both form homodimers and bind with the ligands from the B7 family proteins (e.g. CD80/B7-1 and CD86/B7-2), but they serve different functions in T cells (Rudd et al., 2009). The costimulatory signal provided by CD28 is often required for T cells to respond to peptide antigens (Bluestone, 1995; Noel et al., 1996). Without the CD28 signal, the TCR signal alone tends to induce anergic or apoptotic T cells, which would not elicit an effective anti-tumour immune response (Linsley and Ledbetter, 1993). CD28 and CTLA-4 also have different expression patterns. CD28 is expressed in resting and activated T cells on the cell surface, but CTLA-4 is an intracellular protein and its expression is induced by TCR/CD28 co-stimulation (Noel et al., 1996; Takahashi et al., 2000). Upon TCR/CD28 co-stimulation, CTLA-4 is translocated onto the cell surface, where it binds with CD80/CD86 with higher affinity than CD28. This binding dampens T cell activation and proliferation (Chambers et al., 2001; Stamper et al., 2001). The discovery of CTLA-4 and other costimulatory receptors have revolutionized the field of cancer immunotherapy. The anti-CTLA-4 antibody was the first drug to significantly increase the survival of melanoma patients compared with other conventional treatment options (Robert et al., 2011). Therefore, the success of blocking co-stimulatory receptors in cancer treatment has helped to establish a new class of therapeutics: checkpoint inhibitors.

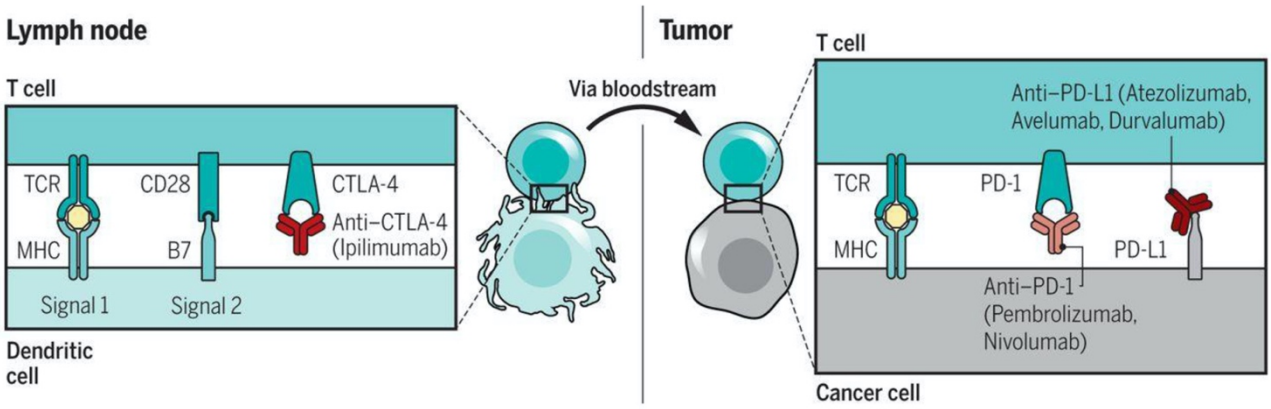


Figure 1.2 Illustration of T cell activation and antitumour response.

Left) An antigen presenting cell (e.g. dendritic cell) presents tumour associated antigens (TAAs) in lymph nodes that it harvested from dying cancer cells from the tumour microenvironment. The activation of both signal 1 (TCR/TAA/MHC) and signal 2 (CD28/B7 family members) leads to the translation of CTLA-4 to the cell surface, which then inhibits the activation and proliferation of T cells by binding to the B7 family members. The anti-CTLA4 antibody can block the inhibition by CTLA4.

Right) Activated T cells are then homed to TME, upon recognition of the TAA via TCR, PD-1 expression is activated. Other factors in the TME such as interferon-gamma or hypoxia can turn on the expression of PD-L1 in cancer cells. The interaction of PD-1/PD-L1 shuts down the anti-tumour response from T cells, but this negative interaction can be reversed by the introduction of anti-PD-1 or anti-PD-L1 antibodies. Figure adapted from Ribas & Wolchok 2018. Abbreviations: TCR = T cell receptor; TAA = tumour associated antigen; MHC = major histocompatibility complex; CTLA-4 = cytotoxic T-lymphocyte-associated protein 4; PD-1 = programme cell death protein 1.

1.1.1.2 Natural Killer (NK) cells

NK cells are a critical component of the tumour immunosurveillance with similar tumour lysing functions as CTLs, but they distinguish cancer cells through different cellular recognition mechanisms (Liu et al., 2021). Instead of recognizing cancer cells through TCR/MHC interactions, NKs detect cancer cells through a repertoire of activating and inhibitory cell surface receptors, which are derived from the germinal sequence (Nausch and Cerwenka, 2008; Spear et al., 2013). One of the main ways that NKs recognize cancer cells is through the interaction of MHC I molecules on the cancer surface and the killer immunoglobulin-like (KIR) family inhibitory receptors on the NKs (Handgretinger et al., 2016). Cancer cells and virally infected cells have downregulation of MHC I and upregulation of ligands for NK cells activating receptors, therefore prompting the activation of NK cells to release perforin and granzymes to lyse the cancer cell (Osińska et al., 2014; Smyth et al., 2005). Additionally, NKs promote anti-tumour activities within the TME through the secretion of pro-inflammatory cytokines and chemokines (e.g. IFN-γ, TNF, IL-6, CCL5) (Guillerey et al., 2016). However, the immunosurveillance capacity and anti-tumour cytotoxicity of NK cells are mostly restricted to circulating cancer cells and distal metastatic sites during initial formations (Brodbeck et al., 2014; Jacobs et al., 2012). The solid tumour microenvironment presents some formidable challenges for NK cells (Figure 1.3). The first layer of challenge is that NK cells have limited ability to penetrate TME (Sconocchia et al., 2012). Even after NK cells successfully infiltrate into the TME, they have limited capacities to overcome the immunosuppressive products within the TME (Guillerey et al., 2016). It has been hypothesized that NK cells may overcome solid TME suppression with the aid of immune checkpoint blockade (Guillerey et al., 2016).

1.1.1.3 B Cells

B cells are known for their ability to produce antibodies, but they can also present antigens to CTLs and secrete cytokines (Anderson and Simon, 2020). Although the presence of tumour infiltrating B cells and their secreted pro-tumour antibodies were reported back in the 1970s (Ran and Witz, 1972; Witz, 1977), the role of B cells in TME still remains ambiguous and controversial (Lei et al., 2020;

Sharonov et al., 2020). The amount of infiltrating B cells tends to be less than T cells in TME (Bindea et al., 2013; Chevrier et al., 2017; Jackson et al., 1996; Schoorl et al., 1976), but the presence of B cells serves as important prognostic factors in cancer (Charoentong et al., 2017; Dieu-Nosjean et al., 2008; Germain et al., 2014). B cells are often found at the margin of tumours and they can be found within the complex tumour-associated immune aggregates such as the tertiary lymphoid structures (TLS), which support the interactions of B cells and T cells to develop anti-tumour T cell responses (Figure 1.3) (Bindea et al., 2013; Shi et al., 2013). Furthermore, B cells or plasma cells in TME can produce a significant amount of cytokines and antibodies, which can then stimulate antitumour immunity via antibody-dependent cellular cytotoxicity (ADCC) (Dang et al., 2014; Gilbert et al., 2011; Kurai et al., 2007). However, B cells can also have pro-tumour effects. In mouse models, depletion of B cells led to increasing T cell infiltration, decreasing chemotherapy resistance, and slower tumour growth (Inoue et al., 2006; Kroemer et al., 2013; Shalapour et al., 2015). In humans, the tumour-promoting effects of B cells tend to associate with the presence of a group of immunosuppressive B cell subsets that express inhibitory ligands and cytokines (Ammirante et al., 2010; Van Den Beucken et al., 2014; Mahoney et al., 2019; Shalapour et al., 2015). As the role of B cells in TME is just emerging, it is expected additional research efforts will help to resolve their functions in TME (Sharonov et al., 2020).

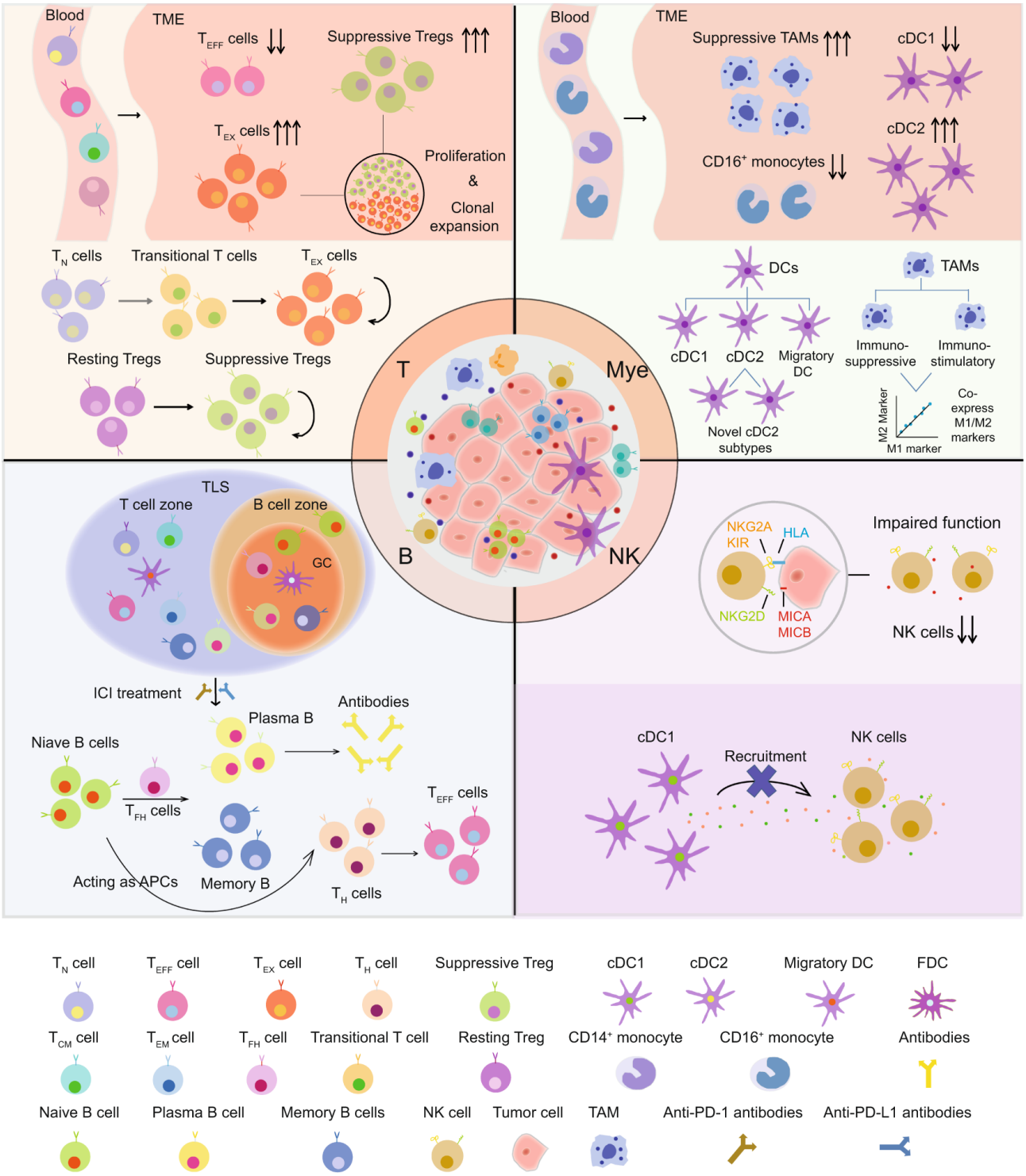


Figure 1.3 Immune cell functions and adaptations in the tumour microenvironment.

1) T cells infiltrate into tumours from circulation and undergo changes driven by the immunosuppressive TME. Within the TME, naïve CD8 T cells and CD4 cells may transition into exhausted states, which feature a reduction in effector functions. Additionally, regulator T cells can transit into suppressive states, which inhibit anti-tumour immunity. 2) B cells have important roles in

anti-tumour immunity and immunotherapy. B cells are often found at the margin of tumours and they can also be found within the complex tumour-associated immune aggregates such as the tertiary lymphoid structures (TLS), which support the interactions of B cells and T cells to develop anti-tumour T cell responses. B cells can differentiate into plasma cells and produce a significant amount of cytokines and antibodies, which can then stimulate antitumour immunity via antibody-dependent cellular cytotoxicity (ADCC). B cells can also present antigens to CD4 cells that promote the activation of CD8 T cells. C) NK cells are a critical component of the tumour immunosurveillance system with similar tumour lysing cytotoxicity as CTLs, but NKs detect cancer cells through a repertoire of activating and inhibitory cell surface receptors, which are derived from the germinal sequence. However, the immunosurveillance capacity and anti-tumour cytotoxicity of NK cells are mostly restricted to circulating cancer cells and distal metastatic sites during initial formations. The solid TME presents some formidable challenges for NK cells. Factors in TME such as VEGF can impact DC maturation, which negatively impacts NK cell recruitment into tumours. D) Myeloid cells often adapt to immunosuppressive phenotypes within TME. Circulating monocytes can enter tumours and differentiate into tumour-associated macrophages and dendritic cells. TAMs serve functions within the TME that include angiogenesis, cancer migration and invasion, metastasis, immunosuppression and more. Single-cell studies have revealed that TAMs can express both M1- and M2-related signatures. Monocyte-derived DCs have limited ability to activate naïve T cells, but they are efficient at suppressing T cell activities and proliferation through the production of TNF-α and inducible nitric oxide synthase (iNOS). Abbreviations: T_N = naïve T cell; T_{CM} = central memory T cell; T_{EM} = effector memory T cell; T_H = T helper cell; T_{FH} = T follicular helper cell; T_{EFF} = effector T-cell; T_{EX} = exhausted T-cell; T_{regs} = regulatory T cells; TAMs tumour-associated macrophages; DCs = dendritic cells; cDC = classical dendritic cell; FDC = follicular dendritic cell; GC = germinal center; TLS = tertiary lymphatic structure; ICIs = immune checkpoint inhibitors; TME = tumour microenvironment; NK = natural killer; APCs = antigen-presenting cells; HLA = human leukocyte antigen; Mye = myeloid cell. Figure adapted from Zhang and Zhang 2020.

1.1.2 Myeloid cells

Myeloid cells are part of the innate immunity and essential components of the TME. The most representative myeloid cells in TME are macrophages, dendritic cells and neutrophils. Each myeloid cell type also contains heterogeneous subpopulations of cells that serve diverse functions (Figure 1.3). Myeloid cells also reshape TME in ways that often promote tumour progression by suppressing the immune system (Zhang and Zhang, 2020).

1.1.2.1 Macrophages

Macrophages were first discovered in 1883 by Élie Metchnikoff for their ability of phagocytosis and the presence of macrophages in tumours has been reported from examining Jensen rat Sarcoma in the 1920s (Carrel and Ebeling, 1926, 1928). With improvements in immune cell isolation methods in the 1970s, the characterisation of infiltrating macrophages in tumour samples began to proliferate (Haskill et al., 1982; Hersh et al., 1976; Russel et al., 1976; Tötterman et al., 1980). More recently, clinicopathological data have indicated that there is a strong relationship between high macrophage infiltration and poor prognosis in different cancer types including lung, liver, kidney and breast cancer (Campbell et al., 2011; Komohara et al., 2011; Mahmoud et al., 2012; Medrek et al., 2012).

1.1.2.1.1 Macrophage origins

The developmental origins of macrophages come from two distinctive sources: embryonic progenitor cells from the yolk sac during embryogenesis and monocytes derived from hematopoietic stem cells (HSCs) after birth (Hashimoto et al., 2013; Hoeffel et al., 2015; Mass et al., 2016; Wynn et al., 2013). Embryogenic precursor cells become much of the tissue-resident macrophages that are capable of self-renewal, especially for privileged tissues such as the brain where macrophage subsets like microglia come exclusively from the embryonic development (Ginhoux et al., 2010; Mass, 2018). In other tissues such as the pancreas, breast and lung, monocytes from HSCs gradually replace a portion or all embryonic-derived macrophages (Calderon et al., 2015; Epelman et al., 2014; Gibbings et al., 2015). The main function of macrophages in healthy tissues is to help maintain homeostasis, but in

tumours, the embryonic- and monocytic-derived macrophages can have different phenotypes and functions (Loyher et al., 2018; Zhu et al., 2017).

The source of tumour-associated macrophages (TAMs) is highly tissue-dependent. For example, it has been demonstrated in mouse mammary cancer models that TAMs are exclusively derived from circulating monocytes (Movahedi et al., 2010; Strachan et al., 2013). But in other tumours (e.g. brain, lung, and pancreas), TAM populations contain mixed cells from both monocytic and embryonic sources (Bowman et al., 2016; Loyher et al., 2018; Zhu et al., 2017). Regarding the different functions, TAMs derived from monocytes have been associated with immunosuppression and antigen presentation, whereas TAMs from embryonic origins are associated with tissue remodelling and wound healing (DeNardo and Ruffell, 2019; Zhu et al., 2017). Therefore, macrophage heterogeneity in different tissues is rooted in embryogenesis and further evaluations of these functional diverse cells in human cancer may help to uncover new therapeutic targets.

1.1.2.1.2 Macrophage recruitment

Tumours produce a variety of cytokines, chemokines and growth factors, which help to recruit circulating monocytes into TME and stimulate their differentiation into macrophages (Van Overmeire et al., 2016). Two of the main factors that promote the recruitment of inflammatory monocytes into tumours are CC-chemokine ligand 2 (CCL2) and colony-stimulating factor 1 (CSF1 or M-CSF). CCL2 has been shown to attract monocytes with CC-chemokine receptor (CCR2) into tumours (Lenzo et al., 2012; Van Overmeire et al., 2016). Furthermore, CCL2 supports tumour progression in a variety of cancer types (e.g. breast, prostate, oral, liver, colon, kidney) (Hourani et al., 2021; Wang et al., 2015).

Besides chemokines, cytokines also aid the recruitment of monocytes into tumours. One of the leading examples is CSF1, which has major functions in regulating monocytic cell biology that includes growth, differentiation and survival (Kawasaki and Ladner, 1990; Murdoch et al., 2004). Overexpression of CSF1 and its receptor CSF1R have been found in diverse sets of cancer types

(Murdoch et al., 2004). For instance, CSF1 expression is strongly correlated with monocyte infiltration in breast cancer and CSF1 knockout reduced the amount of TAM infiltration in tumours and halted tumour progression (Lin et al., 2001; Tang et al., 1990). Moreover, the hypoxic microenvironment also recruits and traps TAMs in tumours. Hypoxia can induce the production of VEGF-A, CXCL12 Angiopoietin-2 (Ang-2), endothelin-2, endothelial cell monocyte-activating polypeptide-II (EMAP-II), which recruit monocytes into tumours (Du et al., 2008; Murdoch et al., 2004; Tripathi et al., 2014). Once entered into the hypoxic regions, hypoxia can lead to the downregulation of chemokine receptors (e.g. CCR2, CCR5), which trap TAMs within the hypoxic environment (Bosco et al., 2004; Henze and Mazzone, 2016a; Sica et al., 2000).

1.1.2.1.3 Macrophage activation model

Macrophages are sensitive to external stimuli and can acquire specialized functional phenotypes in tissues (Wynn et al., 2013). This process is called macrophage activation. Our understanding of macrophage activation processes was mostly derived from *ex vivo* experimental studies of circulating monocytes in conditions that do not capture the complexity of microenvironments in healthy or pathological tissues (Ransohoff, 2016). Accumulating evidence suggests macrophage activation models derived from these oversimplified *ex vivo* settings do not represent the *in vivo* macrophages (Cavallo et al., 2011; Ransohoff, 2016; Xue et al., 2014). The traditional macrophage activation model divided activated macrophages into two polarizing states: M1 and M2, which mirrored the Th1 and Th2 polarization in T cells (Mills et al., 2000). The initial M1/M2 model was based on evidence from the early 1990s that monocytic cells treated with IL-4 and IFN-γ/LPS could lead to both morphological and transcriptomic changes in these cells (Martinez and Gordon, 2014; Stein et al., 1992) (Figure 1.4A&B). Upon stimulation of IFN-γ or LPS, macrophages acquire a pro-inflammatory phenotype (M1 or classically activated) that generates pro-inflammatory cytokines (e.g. IL6, IL12, IL23), TNFα, and reactive oxygen species to promote inflammation and host defence (Aras and Raza Zaidi, 2017). In contrast, the alternative macrophage (M2) phenotype requires the stimulation of IL4, IL13, and IL10 to promote anti-inflammatory, tissue repair and pro-angiogenic functions (Fan et al.,

2016). However, with the advancement of new single-cell sequencing technologies, the distinction of the M1/M2 polarizing model appeals to be less obvious. In 2017, Chung and colleagues reported the expression of 515 cells from 11 breast cancer patients and reported that TAMs expressed both M1 (e.g. CXCL9, CXCL10) and M2 (e.g. CD163, CXCL2, TGFBI, MS4A6A) markers (Chung et al., 2017) (Figure 1.4C). In the subsequent year, Azizi and colleagues reported a similar finding that breast cancer TAMs from 8 patients have a continuous expression of both M1 and M2 signature genes (Figure 1.4D). Of note, the M1 and M2 signatures used by Azizi et al were not derived from specific macrophage transcriptomic profiles, but from a collection of genes in the literature that have previously been used to annotate M1 and M2 populations (Azizi et al., 2018). Therefore, these studies demonstrated the M1 and M2 states generated from *ex vivo* experiments may not capture the *in vivo* TAM phenotypes.

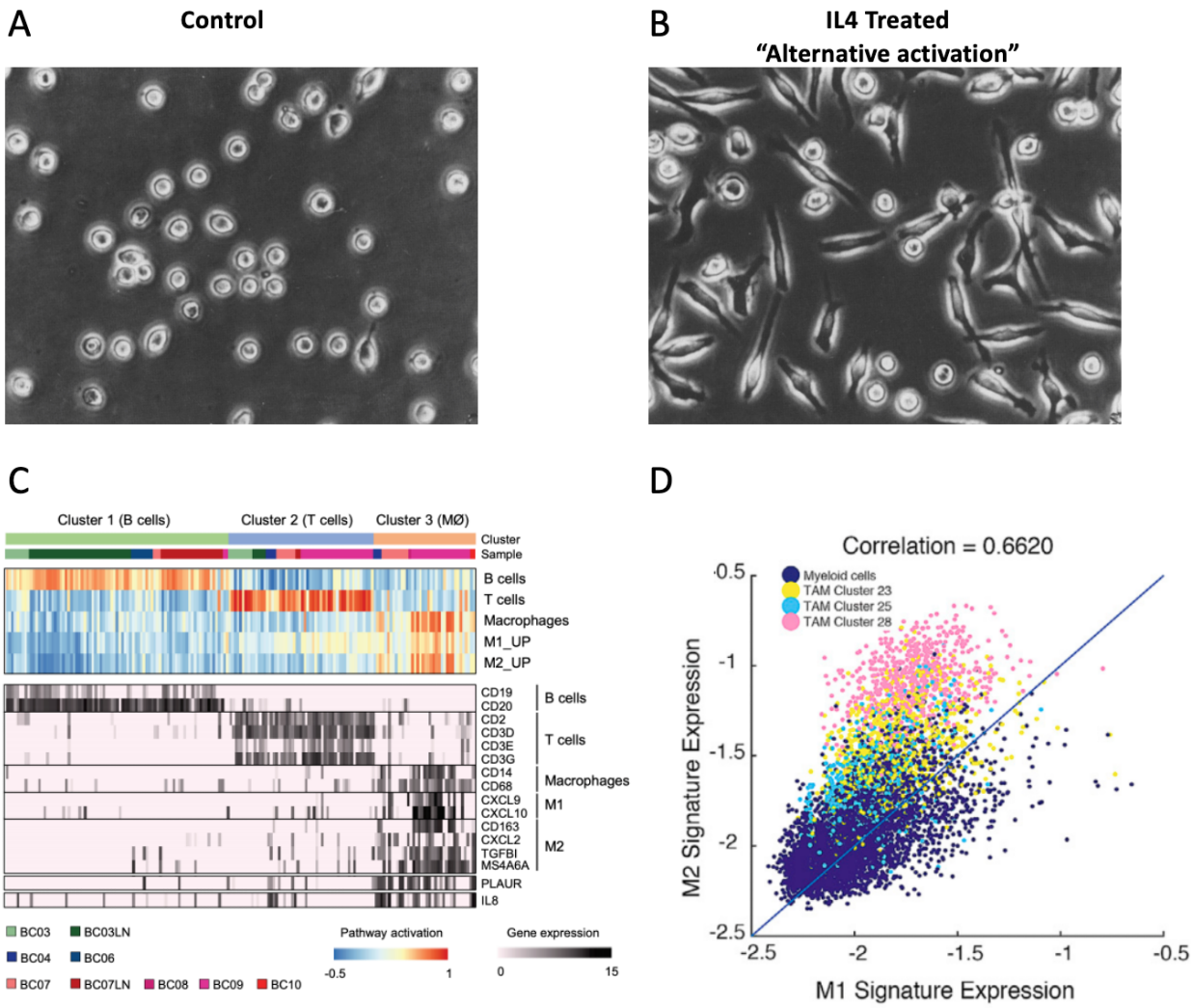


Figure 1.4 Evolving model of macrophage activation, from M1/M2 to TAM.

A & B) The initial experimental evidence demonstrated the morphological changes associated with macrophage alternative activation *ex vivo*. B) Macrophages were isolated from adult BALB/c mice and were treated with IL4 for 8 hours. Both morphological and transcriptomic changes were observed in these alternatively activated macrophages. Subsequently, the M1 and M2 terminology and models were created to describe M1 and M2 as classically and alternatively activated macrophages, respectively. Figures adapted from Stein et al 1992. C) Selected gene expression from immune cells isolated from 11 breast cancer patients. This was one of the first demonstrations at the single-cell resolution that TAMs expressed both M1 and M2 markers, which were previously thought to be exclusive for each of the polarizing populations. Figure adapted from Chung et al 2017. D) Three

TAM populations were isolated from 8 breast cancer samples using single-cell sequencing technology. Analysis of the data using Pearson correlation revealed that these TAMs expressed both M1 and M2 signatures. The signatures were a collection of genes from the literature that were used to annotate M1 and M2 instead of gene signatures directly derived from macrophage populations. Figure adapted from Azizi et al 2018.

1.1.2.1.4 Macrophage functions in TME

TAMs have diverse functions within the TME that include angiogenesis, cancer migration and invasion, metastasis, immunosuppression and more (Aras and Raza Zaidi, 2017; Qian and Pollard, 2010). Tumour growth is intertwined with the growth of new blood vessels from the existing vasculature in a process called angiogenesis. Tumour vasculatures transport oxygen and nutrients to allow cancer cells to proliferate and metastasize (Lin and Pollard, 2007). TAMs significantly contribute to the process of angiogenesis in tumours through the secretions of pro-angiogenic factors (e.g. VEGF, adrenomedullin (ADM), Platelet-derived growth factor (PDGF), TGF-β and matrix metalloproteinases (MMPs)) (Aras and Raza Zaidi, 2017; Jiang et al., 2016). For instance, VEGF-A, a potent angiogenic factor, is produced by TAM and neoplastic cells and its expression is further upregulated under hypoxic conditions through HIF (Henze and Mazzone, 2016a; Murdoch et al., 2008). VEGF-producing TAMs contribute to the neo-angiogenesis and recruitment of additional macrophages to the TME that promote tumour progression (Linde et al., 2012). Additional to VEGF, TAMs also produce MMP9, which has been shown to induce angiogenesis and tumour growth by modifying the extracellular matrix (Giraudo et al., 2004). In contrast, depletion of TAM recruiting chemokine, CSF-1, using genetic deletion in a mouse model demonstrated a significant reduction in both tumour angiogenesis and tumour burden, which indicate TAMs are critical for angiogenesis (Qian and Pollard, 2010).

TAMs and cancer cells can work together to reshape the TME to promote malignancy. For instance, cancer cells release CSF-1 to recruit macrophages and to promote the production of epidermal growth factor (EGF) by macrophages (Chitu and Stanley, 2006). CSF-1 and EGF production from both cells work in a paracrine loop fashion, which promotes the migration of TAMs and cancer cells towards blood vessels (Condeelis and Pollard, 2006; Wyckoff et al., 2007). When TAMs are in proximity to the vasculatures, they create an opening in the blood vessel to allow the cancer cells to escape (Arlauckas et al., 2017; Wyckoff et al., 2007). Furthermore, TAMs can produce additional molecules to support the invasion of cancer cells include cathepsin proteases and osteonectin, which modify the

extracellular matrix and promote tumour cell migration, respectively (Laoui et al., 2011; Sangaletti et al., 2008). Interestingly, genetic depletion of CSF-1 can reduce the number of circulating cancer cells and decrease metastasis (Mitchem et al., 2013). Therefore, TAMs serve critical roles in promoting invasion, migration and tumour metastasis in TME.

TAMs can suppress antitumour immunity through the production of a collection of effector molecules, either through direct cell-cell interaction or indirect signalling by releasing cytokines, chemokines, and enzymes (Noy and Pollard, 2014). For example, TAMs express the ligands of the inhibitory receptors programmed cell death protein 1 (PD-1) and cytotoxic T-lymphocyte antigen 4 (CTLA-4), which are upregulated on activated lymphocytes (T cell, B cell, NK) as a safety mechanism to control overactivated immunity (Belai et al., 2014; Duraiswamy et al., 2013; Greaves and Gribben, 2013). Activation of these receptors inhibits effector cell functions and reduces the potential anti-tumour cytotoxic effects (Kuang et al., 2009).

Hypoxia also plays an important role in TAM-related T cell suppression. It has been shown that hypoxic TAMs have increased PD-L1 expression and induce T cell suppression via HIF-1α activity (Doedens et al., 2010a; Noman et al., 2014). Furthermore, TAMs also express both classical (e.g. HLA-C) and non-classical MHC molecules (e.g. HLA-E, HLA-G) either in attached or soluble forms (Kren et al., 2010; Morandi et al., 2007). These non-classical HLA molecules, can block the activation of NK cells via interaction with NKG2 and inhibit leukocyte activation via interaction with the inhibitory leukocyte immunoglobulin-like receptors LIT-2 (Borrego et al., 1998). Therefore, TAMs can inhibit effector cell functions through direct cell-cell interactions.

TAMs can indirectly inhibit T cell functions by releasing different cytokines, chemokines, and enzymes that help to recruit immunosuppressive regulatory T cells (Treg) into the TME or by educating CD4 T cells to become Treg. To recruit Treg into TME, TAMs may use different chemokines in different cancer types (Adeegbe and Nishikawa, 2013). It has been reported that TAMs in human ovarian cancer secrete CCL22 to recruit CCR4+ Treg (Curiel et al., 2004). Whereas in

colorectal cancer, TAMs recruit CCR6+ Treg by producing CCL20 (Liu et al., 2011). TAM in melanoma can recruit Treg through the production of CCR5 ligands (Schlecker et al., 2012). Furthermore, TAMs can also induce the Treg phenotype from CD4 T cells by producing TGF-β and IL-10 in different human and mouse cancers (Denning et al., 2007; Pollard, 2004). Interestingly, TGF-β and IL-10 can also directly regulate T cell functions, with TGF-β blocking CTL functions and IL-10 inhibiting T helper functions (Ng et al., 2013; Oh and Li, 2013). Finally, TAMs can suppress T cell activity by depleting necessary nutrients in the TME through the production of enzymes (Biswas and Mantovani, 2010). For instance, TAMs release L-arginine processing enzymes: nitric-oxide synthase (NOS) and arginase I (ARGI) (Biswas and Mantovani, 2010). ARGI depletes L-arginine in TME by converting it to urea and L-ornithine (Doedens et al., 2010b; Sharda et al., 2011). This depletion of L-arginine prevents the re-expression of the TCR ζ chain after antigen presentation, which significantly decreases T cell proliferation (Taheri et al., 2001). Together, the data suggest TAMs are critical immunosuppressors within the tumour microenvironment. Developing robust TAM biomarkers could help to predict prognosis and to augment the development of TAM targeting therapeutics.

1.1.2.2 Dendritic cells

Dendritic cells (DCs) form a critical link between innate immunity and adaptive immune through their potent antigen presentation capability (Merad et al., 2013). In tumours, DCs can infiltrate the TME at different stages. The presence of mature DCs in TME is associated with good prognoses in patients, but TME-associated factors can also change the DC phenotypes and cause defects in DCs (Dou and Fang, 2021; Gabrilovich, 2004). DCs can be divided into three major subtypes that operate distinctive anti-tumour and pro-tumour activities (Collin and Bigley, 2018). The first subtype is classical or conventional DCs (cDCs), which activate CTLs and cytotoxic response through antigen cross-representation via TCR/MHC complex (Broz et al., 2014). The high infiltration of cDC in tumours is associated with good prognoses in many different cancer types (Broz et al., 2014; Goc et al., 2014). The second subtype is plasmacytoid DCs (pDCs), which share similar histological

morphology with plasma cells. pDCs are producers of potent anti-tumour type I interferons (Medrano et al., 2017), but cancer cells in TME can halt the interferon production by pDCs through the secretion of TGF-β and TNF-α (Contractor et al., 2007; Palucka et al., 2005; Swiecki and Colonna, 2015). Furthermore, pDCs can also coordinate with regulatory T cells (Tregs) to induce tolerance of cancer cells (Ito et al., 2007; Sisirak et al., 2012). Therefore, infiltration of these tolerogenic pDCs often correlates with poor prognosis in cancer (Munn et al., 2004; Treilleux et al., 2004). The third subtype is a pro-inflammatory group that is known as monocytic DCs (mDCs), which are derived from circulating monocytes (Gopalakrishnan et al., 2018; Inoue et al., 2006). These mDCs have limited ability to activate naïve T cells, but they are efficient at suppressing T cell activities and proliferation through the production of TNF-α and inducible nitric oxide synthase (iNOS), which makes them immunosuppressive in TME (Georgoudaki et al., 2016; Laoui et al., 2016).

1.1.2.3 Neutrophils

Neutrophils make up the majority of leukocytes in circulation and act as the first line of defence against microbial pathogens (Welch et al., 1989). Neutrophil phenotypes can be altered by the secreted factors of TME (Piccard et al., 2012). During the initial tumour development phase, neutrophils can reduce tumour growth through direct cytotoxicity by releasing reactive oxygen species (ROS) (Clark and Klebanoff, 1975; Granot et al., 2011). However, in advanced and metastatic cancer, increasing neutrophil count or elevated neutrophil-to-lymphocyte ratio are associated with poor outcomes and the development of new metastases due to neutrophil reprogramming towards a more immunosuppressive phenotype (Guthrie et al., 2013; Liu et al., 2008; Schmidt et al., 2005; Zucker et al., 2020). Tumours recruit neutrophils into TME through the secretion of C-X-C chemokines (e.g. CXCL8), which interact with the C-X-C motif chemokine receptor 2 (CXCR2) on the neutrophil surface to initiate their migration (Jamieson et al., 2012; Stadtmann and Zarbock, 2012; Wislez et al., 2006). CXCR2 also recruits neutrophils into pre-metastatic sites (Steele et al., 2016; Yu et al., 2017). The contrasting effects of neutrophils during different stages of tumour development result in different neutrophil subsets: the anti-tumour subset N1 and the pro-tumour subset N2

(Brandau et al., 2013; Fridlender et al., 2009). It has been demonstrated that different cytokines within the TME can reprogramme neutrophils toward N1 (e.g. TGF-β, IL-6, G-CSF) or N2 (e.g. IFN-β IL-12) (Andzinski et al., 2016; Medina-Echeverz et al., 2011; Yan et al., 2013; Zou et al., 2017). N1 neutrophils can help to recruit and facilitate the activation of CD8 T cells in TME (Fridlender et al., 2009; Georgoudaki et al., 2016). In contrast, N2 neutrophils can promote tumour growth by negatively regulating cytotoxic lymphocyte functions and recruiting immunosuppressive Tregs into TME (Mishalian et al., 2014). Therefore, TME and its secreted factors could impact tumour progression through the tuning of neutrophil N1 and N2 ratio (Sagiv et al., 2015).

1.1.3 Hypoxic microenvironment

Hypoxia is a common feature in cancer and it creates an immunosuppressive environment within TME through a multitude of mechanisms (Arnaiz and Harris, 2022). High tumour hypoxia is significantly associated with adverse prognosis and clinical outcomes (Evans and Koch, 2003; Hockel and Vaupel, 2001; Vaupel and Mayer, 2007). The main cause of tumour hypoxia is due to an imbalance supply of oxygen in neoplastic tissues where the growing number of cancerous cells consume a significant amount of oxygen, but the leaky and disorganised tumour vasculature cannot cope with oxygen demand from the cancerous cells (De Bock et al., 2011; Nagy et al., 2009). Cancer cells are better adapted to the hypoxic stress than other cells within the TME primarily through the hypoxia inducible factors (HIFs).

1.1.3.1 Hypoxia inducible factors (HIFs)

HIFs are a family of transcriptional factors (e.g. HIF1, HIF2, HIF3) that are critical for cellular response to fluctuations in oxygen levels (LaGory and Giaccia, 2016; Muz et al., 2015). Here I use HIF-1 as a key example to illustrate the HIF activation mechanism. In brief, the HIF response pathways under normal physiological conditions get inactivated by posttranslational hydroxylation of the proline residues on the HIF1α via prolyl hydroxylase domain enzymes (PHDs). This prolyl hydroxylation of HIF1α permits its binding with von Hippel–Lindau protein (VHL) E3 ubiquitin

ligase complex, which subsequently polyubiquitinates HIF1α and targets it for proteasomal degradation (Huang et al., 1998; Ratcliffe, 2007). Under hypoxic conditions (O_2 <2%), the HIF1α and HIF1β transcription factors dimerise to form a heterodimer that binds to hypoxia response elements (HREs) at the promotor regions of hypoxia target genes, which subsequently activate the transcription and translation of these genes (Figure 1.5). These hypoxia-targeting genes have critical roles in almost every aspect of cancer biology: angiogenesis (Lahat et al., 2011), metabolism (Zois and Harris, 2016), invasion and metastasis (Lu and Kang, 2010), acidosis (Swietach et al., 2007), stem cell maintenance (Heddleston et al., 2010), and immune evasion (Abou Khouzam et al., 2021).

1.1.3.2 Angiogenesis

Extensive data has demonstrated hypoxia and HIF play a central role in promoting tumour angiogenesis (Krock et al., 2011; LaGory and Giaccia, 2016; Liao and Johnson, 2007). Angiogenesis is the development of new blood vessels from the existing vasculature (De Palma et al., 2017). Although the concept of angiogenesis and its association with diseases have been described first in 1794 by John Hunter (Turk, 1994), it was only in 1971 that Judah Folkman hypothesised a causal link between angiogenesis and tumorigenesis. In the subsequent years, other research groups then confirmed the importance of angiogenesis in tumour development and demonstrated that cancer cells induce angiogenesis, angiogenesis controls tumour growth, and angiogenesis inhibition leads to tumour growth arrest (Auerbach et al., 1975; Folkman, 1992; Gimbrone et al., 1972; Langer et al., 1980). These observations led to a race to identify angiogenic factors and their associated molecular pathways. In 1989, Ferrara and Henzel identified vascular endothelial growth factor (VEGF) from the conditioned medium of bovine pituitary follicular cells (Ferrara and Henzel, 1989). VEGF was arguably one of the most important factors that direct angiogenesis. Following its discovery, anti-VEGF therapies have been developed and approved by the FDA (Bergers and Hanahan, 2008). The characterization of VEGF and anti-VEGF therapy resistance encouraged the discovery of other angiogenic factors from tumour and non-tumour tissues, which included the Delta-like ligand

(DLL4)-Notch signalling pathway that promotes the formation of vessels, independent of VEGF signalling (Cao et al., 2011; Li et al., 2011a).

1.1.3.3 Hypoxia and immunosuppression in cancer

Hypoxia orchestrates the suppression and evasion of anti-tumour immune responses within the TME through diverse mechanisms that include recruitment of immunosuppressive cells, upregulation of immunosuppressive molecules, and augmentation of acidosis (Labiano et al., 2015; Rapisarda and Melillo, 2012). One of the hallmarks of tumour hypoxia niche is the high infiltration of immunosuppressive cells, which include macrophages, neutrophils, myeloid-derived suppressor cells (MDSCs) and Tregs (Chiu et al., 2016; Facciabene et al., 2011; Ren et al., 2016). Hypoxic cancer cells recruit immunosuppressive cells by releasing HIF1-dependent chemokines such as VEGFA, CXCL12, and SEMA3A (Casazza et al., 2013; Du et al., 2008; Fu et al., 2021). Once recruited into the hypoxic niche, these immunosuppressive cells then further promote tumour progression by stimulating proliferation, angiogenesis and metastasis.

Hypoxia and HIF1α also upregulate cell surface proteins (e.g. PD-L1, CD47, HLA-G) in stromal and cancer cells that directly suppress anti-cancer effector cell functions (Garziera et al., 2017; Noman et al., 2014; Zhang et al., 2015). For example, HIF1α can induce the expression of checkpoints and their ligands in tumour cells and immune cells to suppress anti-cancer immune responses in TME. In cancer cells and macrophages, HIF1α activity upregulated the expression of PD-L1. In CD8 T cells, HIF1α augment the expression of inhibitory receptors, PD-1 and CTLA-4 (Palazon et al., 2017; Topalian et al., 2016). The interaction of immune inhibitory molecules on cancer cells and T cells promotes resistance and T cell exhaustion (Emami Nejad et al., 2021). Furthermore, the HIF-dependent upregulation of CD47 on the cell surface of hypoxic tumour cells can prevent them from being phagocytosed by macrophages (Zhang et al., 2015). CD47 functions as a ligand of signal regulatory protein α (SIRPα), which is expressed on macrophages. The interaction of CD47 and SIRPα prevents the phagocytic or “don’t eat me” signal in macrophages and prevents cancer cells from destruction (Catalán et al., 2020). It has also been shown that this “don’t eat me” signal in TME

can suppress CTL activities and CD47 blockade induced anti-tumour response by CTLs (Liu et al., 2015b; Soto-Pantoja et al., 2014). Therefore, hypoxia suppresses anti-cancer effector cells through the expression of HIF-dependent cell surface inhibitory receptors and ligands.

Acidosis or the reduction of pH is another hallmark of TME and tumour hypoxia can augment the severity of acidosis that further supporting hypoxia-induced immunosuppression (Fu et al., 2021; Singleton et al., 2021). The cause of acidosis is mainly due to the metabolic reprogramming of tumour cells, which features increased glucose metabolism that favours glycolysis instead of oxidative phosphorylation (OXPHOS). This metabolic reprogramming in tumour cells generates and excretes excess amounts of H^+, lactic acid and CO_2 in the extracellular space (Corbet and Feron, 2017a; Persi et al., 2018). Hypoxia and HIF1 response can enhance tumour acidosis by upregulating the expression of key proteins. First, HIF1 increases the expression of glucose transporters (e.g. GLUT1 and GLUT3) and several glycolytic enzymes that increase the intake of glucose (Ebert et al., 1995; Liu et al., 2009). Second, HIF1 upregulates the expression of lactate dehydrogenase A (LDHA) and monocarboxylate transporter 4 (MCT4) (Corbet and Feron, 2017b; Firth et al., 1995). LDHA converts pyruvate into lactate inside the cancer cells. MCT4, a lactate/ H^+ symporter, efflux both lactate and H^+ into the TME (Doherty and Cleveland, 2013; Draoui and Feron, 2011). Third, HIF1 greatly upregulate the expression of CA9, which encodes carbonic anhydrase 9 (CAIX) (Kaluz et al., 2009). CAIX catalyses the hydration of extracellular CO_2 to H^+ and HCO_3^- at a high rate, which generates a concentration gradient that promotes the efflux of CO_2 from the plasma membrane (Li et al., 2011b). Consequently, hypoxia and HIF target genes help to strengthen tumour acidosis that suppresses anti-tumour immunity in TME. For example, extracellular protons and lactic acids suppress the effector functions of CTLs through the inhibition of CTLs proliferation, production of proinflammatory cytokines (e.g., IL-2, TNFs and IFN-γ), upregulation of checkpoint expression (e.g. CTLA-4) and inhibition of anti-tumour cytotoxicity (Fischer et al., 2007; Li et al., 2018; Nakagawa et al., 2015). Therefore, hypoxia and acidosis synergistically suppress immunity and promote tumour progression.

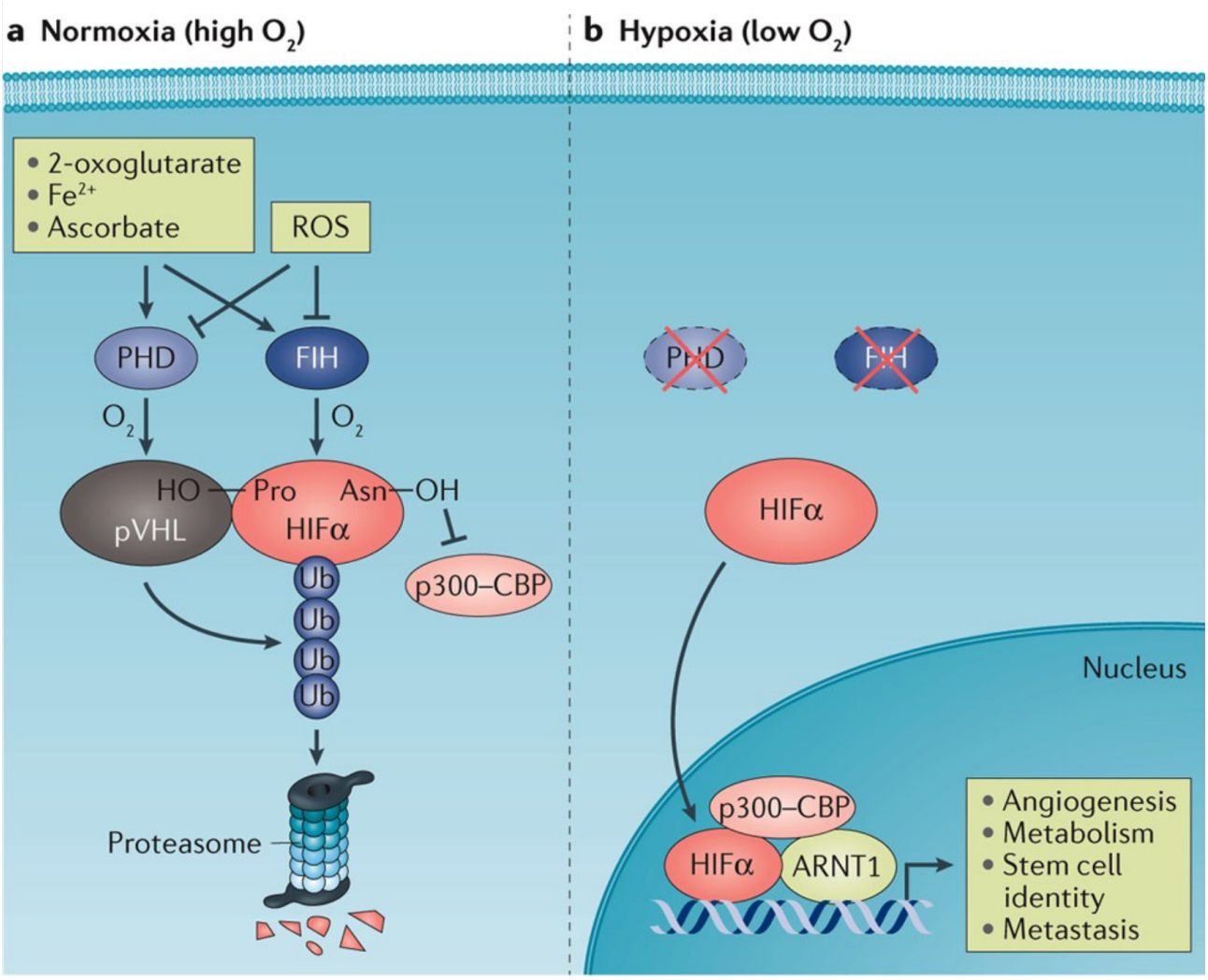


Figure 1.5 Hypoxia inducible factor (HIF) oxygen dependent post-translational regulations

A) Under the normal physiological oxygen concentration, The α-subunits of HIFs are targeted for proteasomal degradation. The chain of reactions initiated by prolyl hydroxylases (PHDs) catalyses the hydroxylation of proline residues on HIFα, which are recognised by the Von Hippel-Lindau (VHL) E3 ubiquitin ligase complex that lead to HIFα ubiquitination and proteasomal degradation. Additional to prolyl hydroxylation, an asparagine residue of HIFα gets hydroxylated by Factor Inhibiting HIF (FIT) and the hydroxylation disrupts the binding of transcriptional co-factor p300 to HIF. B) Under the hypoxia condition, both PHD and FIH become inactivated. HIF-α units translocate to the nucleus and dimerised with HIFβ units to activate the transcription of HIF target genes, which are involved in angiogenesis, metabolism, stem cell maintenance, and metastasis. Figure adapted from Nakazawa et al 2016.

1.2 Cancer immunogenomics

With the advancement of next-generation sequencing technologies, there has never been a time before with more abundance of genomic and transcriptomic datasets. Hence, many computational tools have been developed to mine the biological and disease-relevant insights associated with the cohort datasets. A special set of tools specifically designed to untangle the role of immunity were built for this purpose, and a new subfield has been created: cancer immunogenomics. Immunogenomic methodologies can help to investigate the genetic, cellular, and environmental components within the tumour microenvironment (Hackl et al., 2016; Liu and Mardis, 2017). The developments of cancer immunogenomics are propelled by two milestones in cancer research: breakthroughs in cancer immunotherapy and the expansion of large cancer cohort data using the next-generation sequencing (NGS) technologies.

Cancer immunotherapies have transformed the field of oncology by demonstrating the anti-cancer potential of host immunity in clinics (Galon and Bruni, 2019). The principle of immunotherapy depends on the pre-existing immune components such as T cells that recognize and eliminate cancer cells with specific tumour-associated antigens (TAAs) (Galon et al., 2006; Robert et al., 2011). However, tumour cells can avoid immune destruction by upregulating the ligands of immune checkpoint molecules (e.g. PD1, CTLA4) on the cell surface (Van Allen et al., 2015; Tumeh et al., 2014), which act as breaks of T cell activation and proliferation (reviewed previously). Antibody-based T cell immunomodulators such as anti-PD1 and anti-CTLA4 block interactions of immune checkpoints and their ligands, therefore unleashing the pre-existing anti-cancer immunity. These immune checkpoint inhibitors have been used to treat more than 50 cancer types either as single agents or in combination with chemotherapies as first or second lines of treatments (Van Allen et al., 2015; Huang et al., 2017; Robert, 2020).

However, despite the promise of immunotherapies, not all patients develop durable responses and many patients can acquire undesirable side effects (Wang et al., 2018). Therefore, it's important to

identify patients that are likely to respond to the immune therapies and those who may develop severe side effects prior to treatments using immunogenomic toolsets. For instance, deficiency in the mismatch repair protein (MMR) leads to high mutational loads, microsatellite instability (MMS) and increased TAA levels in multiple cancer types (Li et al., 2020). Patients with MMR deficiency (dMMR) or high tumour mutational burden (TMB) are associated with increased clinical benefit from immune checkpoint blockade therapy (Lauss et al., 2017). Traditionally, it's unreliable and expensive to detect both dMMR and MSI at the same time across multiple samples (Cohen et al., 2019). In order to solve these problems, NGS and immunogenomic technologies can be used to reliable detect dMMR and MSI by targeting the sequence of one hundred genes at the same time (Konishi et al., 2019). In 2017, the FDA approved the use of NGS detection (MSK-IMPACT™) of MSI in multiple cancers and also approved the use of Pembrolizumab (anti-PD1) for the treatment of cancer patients with microsatellite instability-high (MSI-H) or mismatch repair deficient (dMMR) solid tumours (Cheng et al., 2017; Marcus et al., 2019). This was the first approved cancer treatment based on a shared biomarker instead of the primary tumour origin (Colli et al., 2016; Palmeri et al., 2022). Thus, immunogenomic methods have promising potential to augment therapeutic development and biomarker discovery in the era of immunotherapy.

Large cancer genomics projects such as The Cancer Genome Atlas (TCGA) and International Cancer Genome Consortium (ICGC) have a wealth of information that thoroughly covers the mutational landscape across tumour types (Darrell et al., 2008; Hudson et al., 2010). Together these two datasets alone comprise more than 30 million mutations from more than 26,000 cancer samples (Hackl et al., 2016). In addition, there are more than 100,000 cancer genomics and transcriptomics data available on the Genomic Data Commons (GDC) and Gene Expression Omnibus (GEO) repository (Creighton, 2018). Immunogenomic tools can help to mine these achieved genomic data and supplement them with immune related information. For instance, chemotherapy and radiotherapy have been recently shown to augment the immunogenicity of tumour cells through the induction of new mutations and the release of TAAs (Beatty and Gladney, 2015; Demaria et al., 2016; Obeid et al., 2007) and they

have the potential to be used as priming therapy before immunotherapy (Whiteside et al., 2016). Instead of running fresh clinical trials to test new therapeutic regimes, immunogenomics can be used to retrospectively investigate genomic data from previous clinical trials and to gain a deeper understanding of the molecular mechanisms that are triggered by different doses, fractionation, and timing within the tumour microenvironment (Demaria et al., 2016). Therefore, immunogenomics could enrich datasets with immune elements to augment the analyses of prior datasets and help to derive clinical rationale for combining different therapies with immunotherapies.

The applications of immunogenomics in cancer are built mostly upon the advancements of next-generation sequencing technologies, which can be divided into four general categories: 1) characterization of tumour-infiltrating immune cells (TIICs), 2) discovery of tumour-associated antigens (TAA), 3) identification of tumour related immunologic receptors, and 4) human leukocyte antigen (HLA) typing (Figure 1.6). The primary focus of this thesis is to investigate the composition and establish the functions of TIICs, particularly the tumour-associated macrophages, within the tumour microenvironment in large transcriptomic datasets.

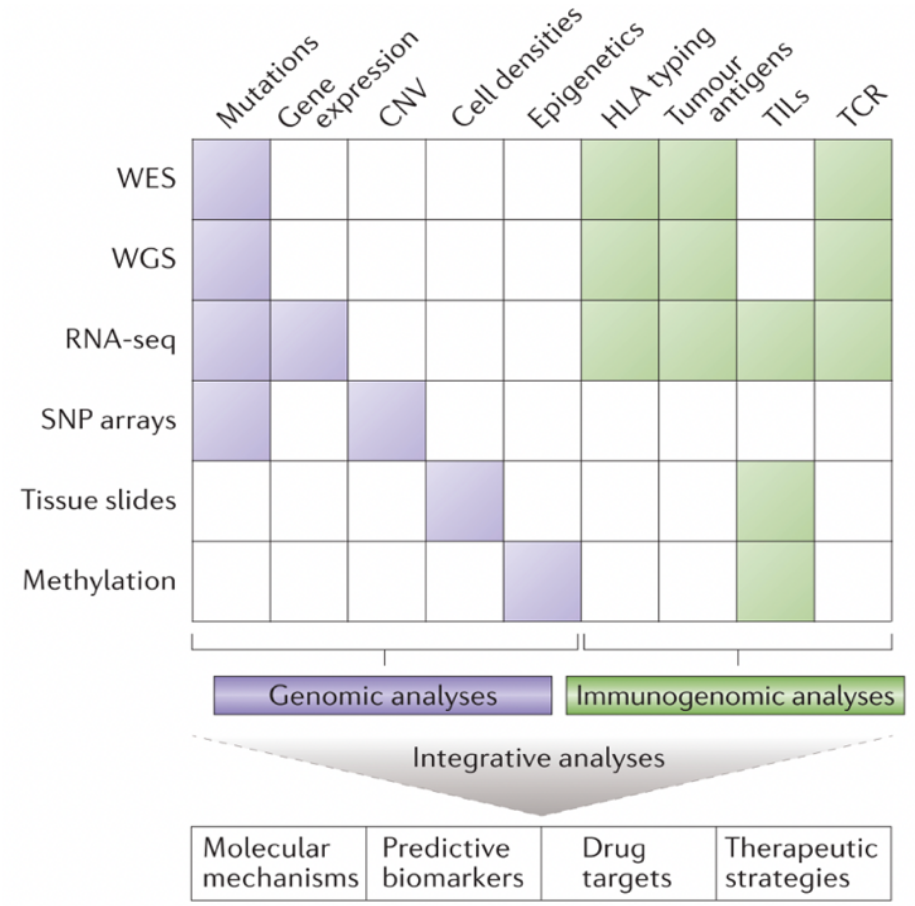


Figure 1.6 Relationship between data type and immunogenomic applications.

Different genomic and immunogenomic analyses can be conducted on various types of NGS data, which are listed on the y-axis. The x-axis listed different types of computational analyses. For example, transcriptomic data generated using RNA-seq can be used for gene expression analysis and diverse sets of immunogenomic analyses. This thesis focused on investigating the tumour infiltrating immune cells (TIICs) in breast cancer and their associated gene signatures using transcriptomic data generated from both bulk and single-cell RNA sequencing technologies. Moreover, amalgamating the results of genomic and immunogenomic analyses could help to reveal new molecular mechanisms, to develop predictive biomarkers and drug targets, and to design therapeutic strategies. Figure adapted from (Hackl et al., 2016). Abbreviations: NGS = Next Generation Sequencing; WES = Whole Exome Sequencing; WGS = Whole Genome Sequencing; RNA-seq = RNA Sequencing; SNP = Single Nucleotide Polymorphism; CNV = Copy Number Variation; HLA = Human Leukocyte Antigens; TILs = Tumour Infiltrating Leukocytes; TCR = T-cell Receptor

1.2.1 Characterization of tumour infiltrating immune cells (TIICs)

The composition and functions of tumour infiltrating immune cells provide important insights to aid the understanding of cancer development and drug response (Hegde et al., 2016). Several methods have been implemented to characterise immune cell contents from tumour samples, which include fluorescence-activated cell sorting (FACS), immunohistochemistry (IHC), and digital pathology tools (Galon et al., 2006; Petitprez et al., 2018). However, despite automation, these tools still have relatively low throughput, high cost, and low-to-medium dimensionality (i.e. markers) (Sturm et al., 2019). In contrast, by combining Next Generation Sequencing (NGS) based transcriptomic profiling and immunogenomic tools, it is possible to estimate a whole range of TIICs and predict their functions in both prospective and retrospective clinical samples (Li et al., 2016a; Liu and Mardis, 2017). The immunogenomic tools for estimating TIICs in bulk transcriptomic data can be divided into *deconvolution* and *gene signature* approaches (Finotello and Trajanoski, 2018). In addition, single-cell technologies can provide insights into cancer immunity at higher cellular resolution and discoveries of novel immune cell types (Villani et al., 2017).

1.2.1.1 Deconvolution approach

Tumour gene expression data can be collected using RNA-sequencing or microarray technologies, which resulted in a gene expression matrix that contains aggregated information from different cell populations (e.g. cancer cells, lymphocytes, myeloid cells) due to the heterogeneity within the tumour microenvironment (Charoentong et al., 2012; Lambrechts et al., 2018). This aggregated data does not directly provide information on tumour immune infiltrates, but the immune infiltration estimates can be reconstructed using deconvolutional-based immunogenomic tools (Sturm et al., 2019). The deconvolution approach assumes that the total gene expression matrix (m) of a tumour sample equals to the linear combination of each cellular component (Shen-Orr and Gaujoux, 2013). To estimate the relative immune cell fractions (f) of immune cell types, these deconvolution tools require inputs from prior knowledge to determine the gene expression profile of individual immune cells, which is called a signature matrix (S) (Finotello and Trajanoski, 2018). The objective of gene expression

deconvolution (Figure 1.8) is to solve the following linear equation to determine f: m = S x f (Chen et al., 2018).

For example, CIBERSORT, a deconvolution tool, estimates the immune cell fractions using support vector regression. It has a built-in immune signature matrix, which was derived from microarray data that composed the expression profiles of 22 immune cells (LM22 signature matrix) (Newman et al., 2015). Users can use CIBERSORT and the LM22 signature matrix to calculate the immune fractions from different bulk tumour transcriptomic samples from microarrays and RNA-seq technologies. Since its introduction, CIBERSORT has become a popular tool to mine the immunogenomic information in transcriptomic datasets and recently it has been used to examine the immune components of 39 different solid and haematological cancers across 18,000 transcriptomic profiles (Gentles et al., 2015).

1.2.1.2 Gene signature approach

Another approach to studying the immune cell infiltration in the bulk transcriptomic dataset is through gene signatures and their summary scores. A “gene signature” is defined as a list of genes that have coherent mRNA expression patterns that reflect a particular biological process, cell population, or pathological outcome (Liberzon et al., 2015). Gene signatures can be derived from the dataset of interests using a variety of methods such as supervised, unsupervised, seed-based approaches or other machine-learning methodologies (Buffa et al., 2010; Kourou et al., 2015; Li et al., 2009; Shipp et al., 2002). More specifically, gene signatures that are specifically derived to characterise immune populations are called “immune signatures”, which can either be derived from well-known immune cell surface markers based on empirical observations, from purified immune cells from tissue samples using cell sorting, or more recently, from single cell sequenced samples (Leahy et al., 1992; Newman et al., 2015, 2019; Zhao et al., 2021). These immune signatures can be used to infer immune

infiltrations in different clinical samples by summarizing the expression of the list of genes using a single “signature score” for each sample (Hänzelmann et al., 2013; Tomfohr et al., 2005).

For example, xCell is an immunogenomic tool using a gene signature-based approach to study the immune composition of tumours (Aran et al., 2017). There are 489 gene signatures in xCell, which are annotated to 64 different cell types including myeloid cells, lymphocytes, epithelial cells and endothelial cells. These 489 signatures were derived from 1725 transcriptomic profiles of purified cells achieved in large expression consortiums such as FANTOM5, ENCODE, BLUEPRINT, IRIS, HPCA, and GEO (Abbas et al., 2005; Dunham et al., 2012; Fernández et al., 2016; Lizio et al., 2015; Mabbott et al., 2013). Multiple gene signatures correspond to each of the 64 cell types to avoid bias associated with a particular dataset during the collection process. To calculate the signature score for each cell type, xCell uses single-sample gene set enrichment analysis (ssGSEA) to score each sample using all 489 gene signatures (Hänzelmann et al., 2013). ssGSEA produces a single aggregated score for each pairing of a sample and a gene signature (Barbie et al., 2009). Then xCell averages the multiple gene signature scores corresponding to each cell type to produce a final signature score. The final xCell signature score for each cell type is arbitrary and cannot directly represent cell fractions, but it has been demonstrated xCell signature final scores correlated with the immune cell fractions from flow cytometry counts (Aran et al., 2017).

There is a clear difference between the previously mentioned deconvolution approach and the gene signature approach. The deconvolution approach requires an accurate signature matrix from purified immune cells, which in practice can be difficult to obtain and potentially biased due to variations in experimental settings, laboratory measurements, and sequencing technologies (e.g. microarray and RNA-seq have different expression ranges) (Barbie et al., 2009). In contrast, the gene signature approach captures a highly representative list of genes for any immune cell type across different laboratories, which may be better at examining immune populations across clinical transcriptomic profiles generated using different sequencing technologies (Le et al., 2021).

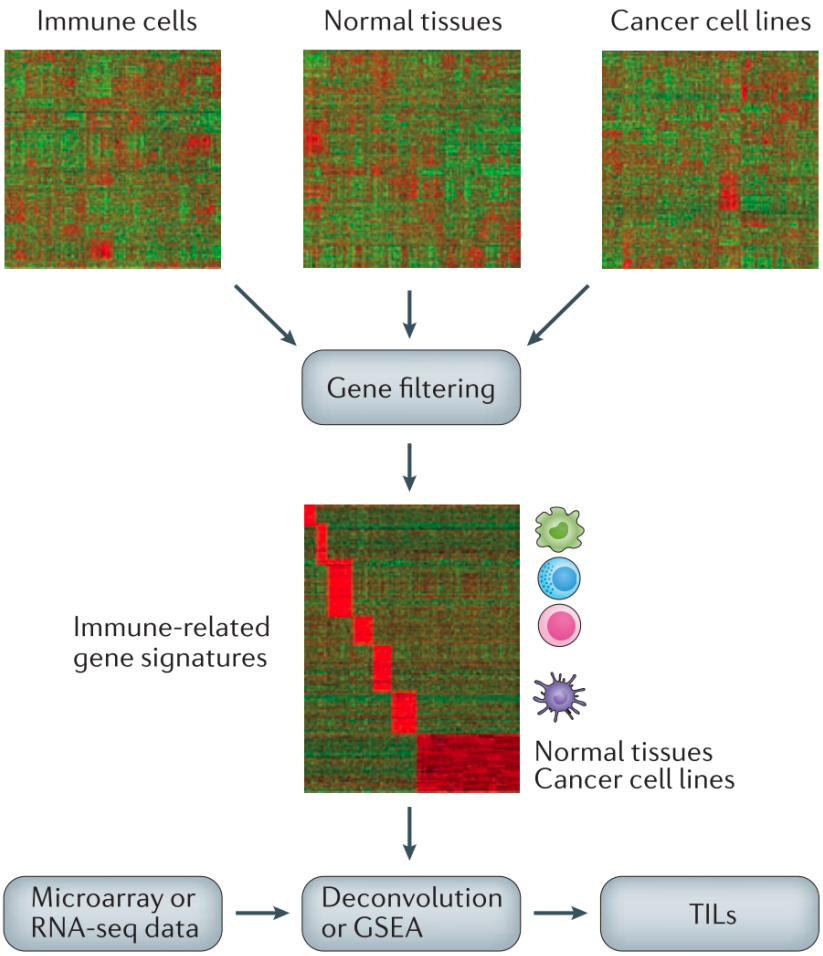


Figure 1.7 An illustration of gene signature method.

To estimate the tumour infiltrating immune cells (TIICs) in bulk transcriptomic samples, several key steps are required. The first step is to derive immune signatures from reference transcriptomic profiles of different cells. As illustrated in the figure: transcriptomic matrixes from purified immune cells, normal tissues, and cancer cells are analysed to identify the most differentially expressed genes among different cell populations, which will result in a unique list of genes for each immune cell type. These lists of genes are then defined as specific immune signatures for each particular immune population. These immune signatures can then be applied to estimate the composition of TIICs in other bulk transcriptomic datasets generated from microarray and RNA-seq experiments. The TIICs estimation is calculated using gene set enrichment methods and summarised into a single sore for each cell type. Figure adapted from (Hackl et al., 2016). Abbreviations: TIL: Tumour Infiltrating Leukocytes; GSEA: Gene Set Enrichment Analysis.

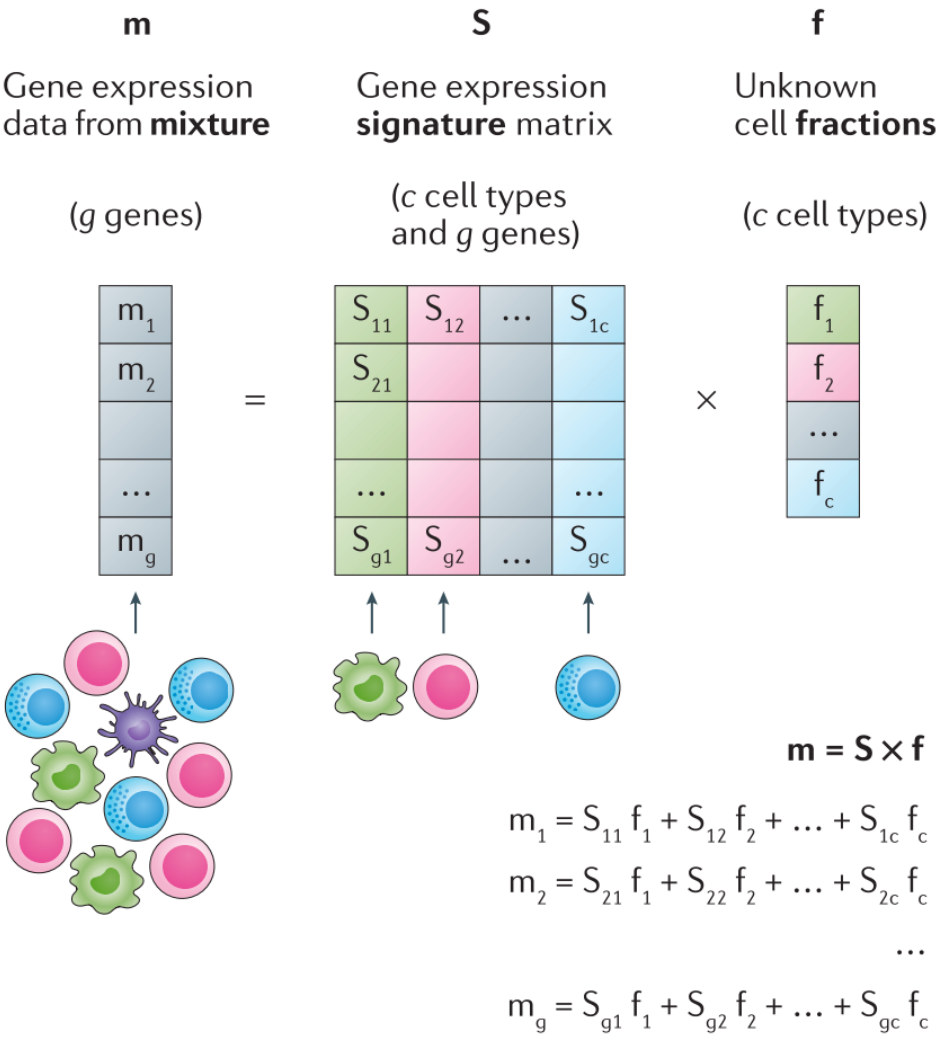


Figure 1.8 An illustration of deconvolution approach to characterise immune cells in tumour

The goal of gene expression deconvolution is to solve the linear equation for f: m = S x f. The deconvolution approach assumes the total gene expression (m) in bulk transcriptomic data of a tumour equals to the linear combination of individual cellular components. To estimate the relative immune cell fractions (f) of each immune cell type, these deconvolution tools require inputs of an immune signature matrix (S), which is derived from gene expression profiles of purified immune cells. Figure adapted from Hackl et al 2016. Abbreviations: m = transcriptomic data from bulk tumour samples; S =immune cell signature matrix derived from purified immune profiles; f = immune cell fractions within the bulk gene expression dataset.

1.2.2 Assumptions of estimating immune infiltrates using transcriptomics

There are three key assumptions behind estimating immune infiltrates using transcriptomics, which are important for understanding and interpreting the results of immunogenomic estimations. The first assumption of methodologies based on gene signatures or deconvolution is that immune cells do not alter their transcriptional profiles from the sites where the signatures were derived and at the sites where their immune signatures were applied, which included the tumour microenvironment and other applicable tissue types (Li et al., 2017a). This is, of course, a convenient hypothesis to increase the scope of the potential applications of these tools. Otherwise, individual signatures will have to be tailored to each tissue environment (Finotello and Trajanoski, 2018). All cells, including senescent cells, can still change transcriptomes to react to their surrounding environments, despite how small those changes may be (Hudgins et al., 2018).

The second assumption is that genes in immune signatures are exclusively expressed in the immune cell compartment, and not the epithelial or cancerous cells. This again is an oversimplification. The immune system is complex. Through evolution, many signalling pathways have been conserved and adapted for various purposes, some of which are often shared by immune and epithelial cells. For example, the interferon response pathways are critical for both the immune responses in epithelial and immune cells, which also feature many shared genes (Saleiro and Platanias, 2019). This created difficulties when using gene signatures to examine bulk transcriptomic datasets which have signals coming from both the immune and cancerous origins. It is difficult to decipher the exact source of interferon signalling within the tumour microenvironment. Some tools may avoid this by filtering out genes specific to cancer cells, but this remains to be a challenge (Li et al., 2017a). Thus, in immunogenomic analyses, it is often assumed that gene signature signals originate from the immune cells, not the epithelial cells, which led to the next assumption.

The third assumption is that there is always enough signal from each of the cellular components to estimate the level of TIICs within the bulk transcriptomic samples. Before the success of immunotherapy in clinical applications, much of the cancer research was focused on characterising

the cancer cells themselves. As a result, many important cohort datasets were carefully examined and collected by pathologists to select tumour areas with the least "impurity" or non-cancerous cells. For instance, TCGA datasets originally aimed to achieve an 80% or higher percentage of cancer cells for each sample, but this was later reduced to 60% (Aran et al., 2015). The ability to study TIICs using transcriptomics on such important datasets may be considered a fortunate coincidence, which allowed us to understand the complexity and biology of the tumour microenvironment.

1.2.3 Validations of the immune infiltration estimations

The performance of immune infiltration estimating tools has been tested against several ground-truth measurements of immune cells in tissues or peripheral blood. The validation approaches can be largely divided into either experimental or computational methods. For experimental validations, flow cytometry was a popular choice. To compare the flow cytometry results, a tissue sample or peripheral blood sample would be divided into two different downstream processes. A portion of it would be processed for transcriptomic processing and the other portion would be processed via flow cytometry for cell counts using the relevant cell surface markers. The results from the two processes were then compared.

For instance, Newman and colleagues compared the results of CIBERSORT against the flow cytometry sorted fractions of immune cells from normal lung tissues, follicular lymphoma tumours, and peripheral blood mononuclear cells (PBMC) (Newman et al., 2017). The results showed a significant association between the *in silico* estimated immune fractions and the flow cytometry results (Newman et al., 2015). However, the comparisons were primarily focused on immune cells that are abundant in PBMC, but the benchmarking on tissue-specific cells such as macrophages was not included.

Other attempts performed computational validation using datasets that were generated from gene expression profiles of purified immune cells (Newman et al., 2019). For example, Aran and colleagues created synthetic datasets from transcriptomic profiles of different immune cells, which

have known immune fractions (Aran et al., 2017). They then compared *in silico* estimated immune scores with the known fractions of the synthetic datasets. The results also showed high agreements on immune cells that are present in PBMC, especially on B cells and CD8 T cells. However, for cells that are present only in tissues, different signatures displayed divergence on the *in silico* immune estimations. One possible explanation for such divergence might be the requirement of *ex vivo* treatments in generating these cells (Appendix 1: Different *ex vivo* protocols used to activate macrophages). For instance, a significant portion of macrophage signatures was derived using monocytes harvested from PBMC with different protocols, whereas, T cells and B cells could be directly isolated from PBMC samples without additional *ex vivo* differentiating steps (Figure 1.9). I investigated the relevance of *ex vivo* derived macrophage signatures in breast cancer samples and compared them with signatures derived from TAMs in chapter 5.

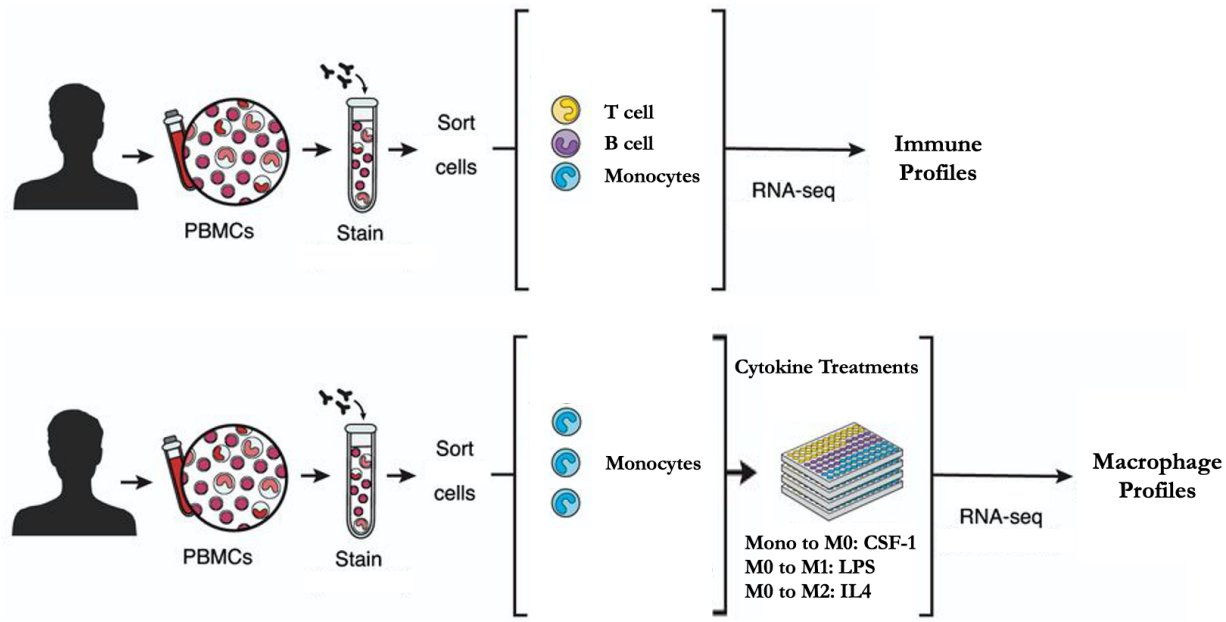


Figure 1.9 An illustration of generating immune profiles from human PBMC

Transcriptomic profiles of purified immune cells are required to derive immune signatures. Top) For cells that already exist in the circulation (e.g. T cell, B cell, monocytes), they can be sorted and harvested via flow cytometry using canonical markers and then sequenced. Bottom) However, for cells that are tissues specific, for instance, macrophages, additional *ex vivo* steps are required to differentiate them into different activation states. These individual transcriptomic profiles can then be used to derive gene signatures or models as described previously. It has been hypothesized that these *ex vivo* derived macrophages may not reflect the phenotypes of TAMs in tumours, but these *ex vivo* signatures have been used to infer TAM biology in many publications. Abbreviations: PBMC = Peripheral Blood Mononuclear Cells; M1 = Macrophage M1; M2 = Macrophage M2; CSF-1 = Colony Stimulating Factor 1; LPS = Lipopolysaccharide.

2 DATASETS AND METHODS

2.1 Datasets

2.1.1 Bulk transcriptomic data

The Cancer Genome Atlas (TCGA) contains transcriptomic data and clinical data from 11,000 cancer patients across 33 different neoplastic types, which makes it one of the largest collections of cancer genomic archives and it serves as a valuable resource for scientific research (Hutter and Zenklusen, 2018). TCGA data were accessed from the Broad Fire browse (www.firebrowse.org). The transcripts were quantified and normalised using RNA-seq by Expectation-Maximization (RSEM) method by the publisher. The RSEM normalised gene expression data for each cancer type was downloaded and then transformed using $\log_2(x+1)$. This was then used for most of the downstream analyses, otherwise specified by different tools. For example, CIBERSORT deconvolution requires untransformed, non-log gene expression data to avoid violating linearity assumptions and to match the expression levels in the signature matrix LM22 (Chen et al., 2018; Steen et al., 2020).

The METABRIC dataset (Curtis et al., 2012) contains transcriptomic data from 1991 primary breast cancer samples. The raw METABRIC files were downloaded from the European genome-phenome archive (EGA) (Study ID: EGAS00000000083) and the dataset was pre-processed, summarized and quantile-normalized from the raw expression files generated by Illumina BeadStudio (R packages: beadarray v2.4.2 and illuminaHuman v3.db_1.12.2). The retrieval and processing of METABRIC expression data were performed by members from the Oncology Department, University of Oxford. Clinical metadata was retrieved from the cBioPortal for Cancer Genomics (www.cbioportal.org).

2.1.2 Single-cell transcriptomic data

Two single-cell transcriptomic datasets of breast cancer immune infiltrating cells were used in the study: Azizi et al 2018 (GSE114725) and Wu et al 2020 (PRJEB23405).

Since its emerging in 2009, single-cell transcriptomic sequencing technologies have revolutionised our understanding of cellular heterogeneity across many different tissue types, especially in investigating the tumour microenvironment (Azizi et al., 2018; Patel et al., 2014; Treutlein et al., 2014; Venteicher et al., 2017). Microfluid-based droplet technologies have enabled scientists to study the whole-transcriptomic profiles of a large number of cells at an acceptable cost. There are three leading droplet-based systems for high-throughput scRNA-seq: inDrop, Drop-seq, and Chromium of 10x Genomics (Briggs et al., 2018; Farrell et al., 2018; Klein et al., 2015; Zheng et al., 2017). In my studies, I used datasets generated using inDrop and 10x Chromium, which will be discussed in more detail. All three technologies similarly generate single-cell droplets through multichannel microfluidic devices that produce water-in-oil droplets in oil emulsion. The goal of generating these droplets is to encapsulate individual cells with on-bead primers with barcodes to identify unique cells (Figure 2.1).

These on-bead primers form the critical component in the droplet-based single-cell sequencing technologies, which contain a 1) PCR handle, 2) cell barcode, 3) unique molecular identifier (UMI) and 4) poly-T (Zhang et al., 2019). Both the 10x and inDrop systems constructed the beads using soft and deformable hydrogel and, in contrast, the Drop-seq used hard resin. The difference in bead materials can impact the quantity and density of DNA primers. For example, the hydrogel-based beads allow more immobilization of primers throughout the beads, but the Drop-seq resin beads can only have limited primers covalently bound to the surface (Abate et al., 2009).

For inDrop and 10x Chromium, after encapsulation of single cells with on-bead primers, cell lysis and reverse transcription (RT) was performed inside of the water-in-oil droplets, which allows uniform results and enhanced yield due to volume restrictions of many local reactions and reduction of reaction competition (Streets et al., 2014). The three platforms use different strategies for cDNA

amplification. Drop-seq and 10x use a template-switching protocol whereas inDrop follows a Cellular Indexing of Transcriptomes and Epitopes by Sequencing (CEL-seq) (Hashimshony et al., 2012). The CEL-seq involves an *in vitro* transcription linear amplification step, which extends the library preparation time of inDrop beyond 24 hrs, which is much longer than both Drop-seq and 10x that can be completed within a day (Zhang et al., 2019). Post sequencing of the cDNA products, the individual cells and the number of transcripts per can be quantified by using cell barcode and UMI sequences, which also help to remove amplification bias (Zilionis et al., 2017). The difference in the on-bead primer design and downstream protocols significantly changes the number of cells that can be processed by these systems and the number of reads per cell obtained post sequencing. Several reviews have compared these technologies in detail (Mereu et al., 2020; Zhang et al., 2019; Zilionis et al., 2017). To what is relevant to my analyses, the 10x Chromium protocol yield a higher number of genes per cell than the inDrop method, which contributed to significant improvements in gene detection and marker expression, but they shared a similar level of clusterability (Mereu et al., 2020).

At the time of analysis, the Azizi dataset was the largest available breast cancer single-cell dataset with annotated TAM populations. which were sequenced from 45,000 $CD45^+$ immune cells from 8 patients. The specimens were collected from primary breast tumours, contralateral prophylactic mastectomies for normal samples, peripheral blood mononuclear cells and adjacent lymph nodes (Azizi et al., 2018). Tissues were enzymatically digested and filtered, then the $CD45^+$ cells were sorted using FACS. Single-cell library was prepared using a droplet-based single-cell sequencing technology, InDrop v2 protocol (Klein et al., 2015; Zilionis et al., 2017). The dataset was retrieved from the Gene Expression Omnibus (GEO), along with sample information and cell labels from the original authors. The cells were classified into 95 clusters of immune cells by the original authors. I reconstructed a monocytic matrix using clusters that were annotated to monocytes (C24, C27, C36,

C40, C42, C49, C54, C64, C67, C68, C84, C86, C91, C94) and macrophages (C23, C25, C28). The macrophage clusters were also used to construct a TAM expression matrix.

The second single-cell dataset from Wu et al 2020 was used to check the presence of TAM populations derived from the Azizi et al dataset using consensus clustering. Fresh tissues were collected from 5 primary triple-negative breast patients and the tissues were dissociated into single-cell suspensions, which contained both cancer cells and other infiltrates in tumours. The cells were processed using a droplet-based sequencing technology from 10x Genomics. The single-cell library was processed using the 10x Chromium 3' v2 library (Zheng et al., 2017). In total, 24,271 cells were sequenced using Illumina NextSeq 500 platform. The sequence data were processed using Cell Ranger Single Cell Software from 10x Genomics. The processed single-cell data by the publishers was retrieved from the European Nucleotide Achieve (ENA). The myeloid cell clusters annotated by the publishers were extracted to generate a separate single-cell matrix with 4,606 cells for subsequent analyses.

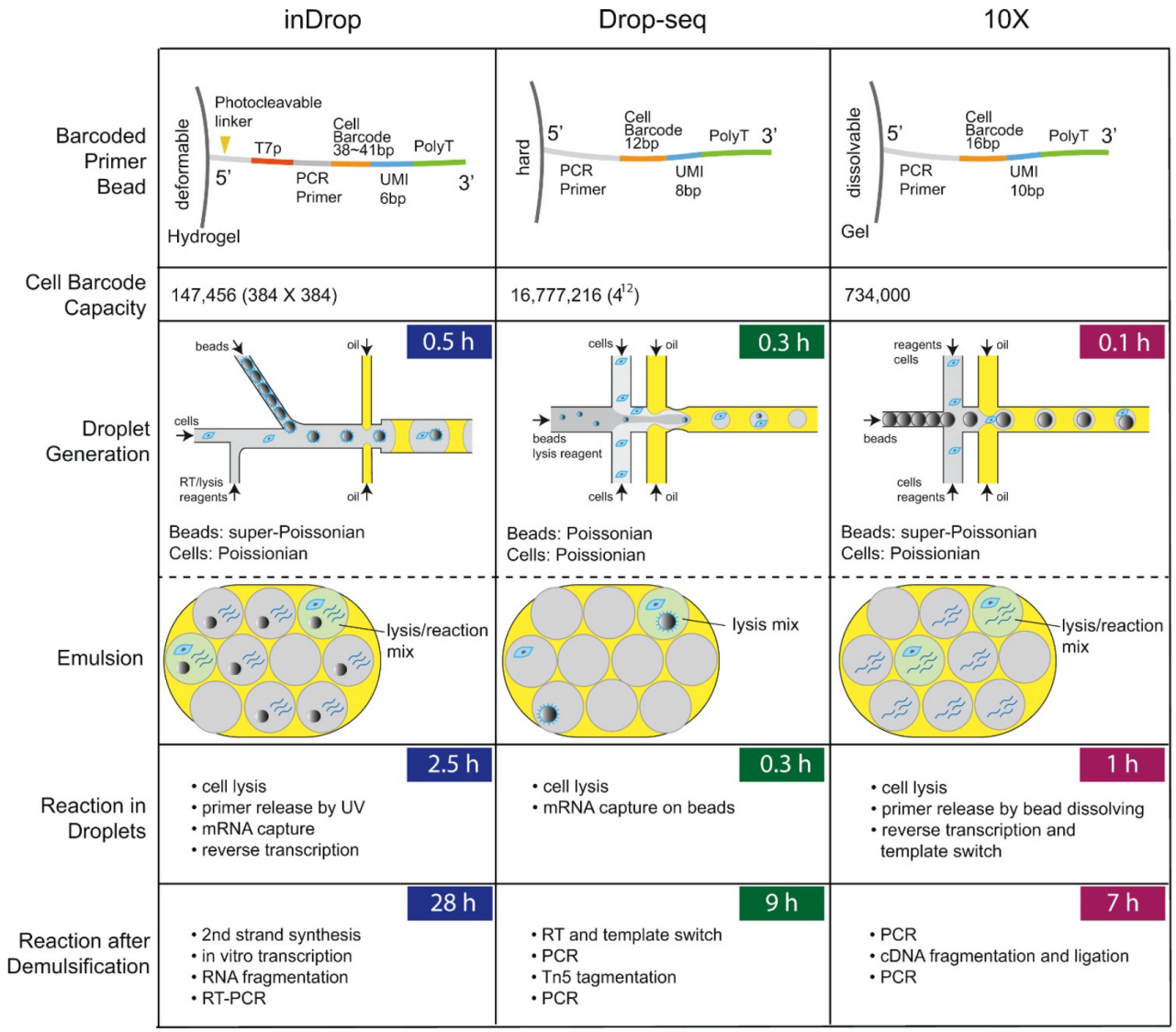


Figure 2.1 Overview of the droplet-based single-cell sequencing platforms

Illustration of the three platforms (inDrop, Drop-seq and 10x Chromium) and their differences. 1) the first difference is due to the on-bead primer design, which includes the bead, PCR primer, cell barcode, UMI and PolyT. Both inDrop and 10x use hydrogel-based beads that allow a higher capacity of primer binding than the hard beads used in Drop-seq. 2) the second difference is cell barcode capacity, which is determined by the barcode library size. In theory, Drop-seq has the largest barcode capacity, however, the effective barcode library size may be smaller (Zhang et al., 2019). 3) the third difference is in the droplet generation. The variabilities come from the design of the microfluidic channels, but the final results are similar with all three technologies producing water-in-oil droplets that deposit in emulsions. 4) the next difference is the type of reactions performed in the droplets. Both inDrop and

10x perform reverse transcriptions in droplets, which can increase the reaction yield. 5) the three platforms use different strategies for cDNA amplification. The *in vitro* transcription step in the inDrop protocol can take up to more than 24 hours, which makes it the slowest droplet-based method. Finally, the single-cell libraries generated using droplet-based platforms can be sequenced using NGS technology and the individual cells and the number of genes per cell is identified using the information from cell barcode and UMI. Figure adapted from Zhang et al 2019.

2.2 Gene signatures

2.2.1 Macrophage immune signatures

A "gene signature" is defined as a list of gene IDs that have coherent mRNA expression patterns that reflect a particular biological process or a cell population. For gene signatures that reflect the mRNA expression patterns of specific immune cells I call them "immune signatures". To comprehensively compare the *ex vivo* and *in vivo* derived human macrophage immune signatures that are commonly used in the literature, a list of 16 macrophage signatures were collected. Macrophage immune signatures were retrieved from previously published literature and their associated supplementary documents. Most of the macrophage signatures could be directly accessed, but additional considerations were needed for Xue et al and xCell. These two studies generated multiple macrophage signatures that correspond to the same macrophage polarization state, and only the most representative signatures were selected in this research. Xue et al treated macrophages with a variety of factors, but I only selected two signatures derived from LPS and IL4 treated cells, which represented the classically described M1 and M2 (Murray et al., 2014). xCell included multiple signatures of M0, M1, and M2 from different transcriptomic datasets, I only selected the most robust

macrophage signatures from the BLUEPRINT dataset after passing them through sigQC with TCGA breast cancer data.

2.2.2 Calculating gene signature scores

To apply immune gene signatures to infer the composition of specific immune cells in bulk samples, I can summarize the expression of the genes in the signatures into a single summary score, which provides information on the signature overall expression and allows inter-sample comparisons. Various approaches exist to calculating the summary score. This can be a simple statistic (e.g. mean, median expression of all genes), or a model (e.g. linear predictor of population fraction using the expression of the genes as covariates). To compute gene signature score calculation and its robustness I used an R package developed by members of the Buffa lab: sigQC (https://cran.r-project.org/web/packages/sigQC/). This package computes several summary scores (e.g. mean, median, PC1) for each gene signature. The mean score is equal to the mean expression of the signature gene in each given sample. It is very simple to use but not all signature genes will have a similar distribution of expression, especially when the distribution is skewed. The median score is the median expression value of all signature genes in each sample. The median expression score is advantageous at reducing the impact of outliners and the score is stable, especially in large datasets. Principal component analysis orthogonally transforms the expression value of a gene signature into multiple linearly uncorrelated variables that are called principal components (PCs). By definition, the first PC (PC1) has the largest possible variance of the gene signature expression. Here, if not otherwise specified, the median expression value of all signature genes for each sample in bulk transcriptomics was used as the signature score. Due to the high dropout nature (more zero values) of single-cell sequencing data, which makes the median expression of the gene signature less reliable and frequently produces zero as result; therefore Gene Set Variation Analysis (GSVA) was used as signature scores in single-cell data. GSVA first calculates an enrichment score as a function of genes inside of the gene signature and outside of the gene signature, then it estimates the variability of the gene signature enrichment across all the samples, and finally summarises the variation into a score for each sample

(Hänzelmann et al., 2013). For head-to-head comparison of gene signatures in bulk and single-cell transcriptomics, GSVA was used in both types of samples.

2.3 Single-cell transcriptomic analysis

2.3.1 Data processing

The initial data processing, normalization, and visualization were performed using Seurat v3.1 package under the R 3.6.0 environment. The single-cell data and additional clinical meta-data were imported into the Seurat v3.1 package using the CreatSeuratObject() function. The data was then normalised using the NormalizeData() function, which by default applied a global-scaling normalization 'LogNormalize' method that normalized the number of genes (i.e. features equal to genes) by the total expression, then multiplied this by 10,000 (a scale factor), and the results were log-transformed. Next, the top 2000 genes that displayed high variability across cells were calculated using the FindVariableFeatures() function, which then created a reduced transcriptomic matrix focusing on these 2000 genes for downstream analyses instead of the full transcriptome. The purpose of using this reduced matrix with the highly variable genes is to account for the highly sparse nature of single-cell data (e.g. high dropouts and zero values in gene expression data) and to improve biological signal in downstream analysis. Afterwards, linear dimensionality reduction using principal component analysis (PCA) was performed to identify the dimensions that represent robust compression of the dataset. The biological meaningful dimension of PCs was estimated using the JackStraw() function. The JackStraw method identifies the meaningful PCs strongly enriched with genes with low p-values (Chung and Storey, 2015). More details on the data pre-processing can be found in the achieved Seurat tutorial (https://satijalab.org.seurat/archive/v3.1).

2.3.2 Clustering

Clustering is a task that divides populations or data points into groups (i.e. clusters) based on their similarities in gene expression. Cluster analysis is an essential task in single-cell data analysis that

can help to identify cellular groups, which can then be used to investigate their biological functions. In my study, I compared results from three different clustering algorithms and tested their performance on the same monocytic dataset. The clustering algorithms were 1) k-nearest neighbour (KNN) graph, 2) Bayesian inference for single-cell clustering and imputing (BISCUIT), and 3) k-means consensus clustering. The clustering outputs were visualized using the ClusterTree package (Zappia and Oshlack, 2018).

Seurat, a popular single-cell analysis package, implements a graph-based method to generate single-cell clusters (Macosko et al., 2015). The first step applied KNN graph-based method that embedded cells in a graph structure, and then edges were drawn between cells based on their Euclidean distance in the PCA space (Altman, 1992). The significant PCs were selected from the previous data processing steps. This first step was performed using the FindNeighbors() function. Next, FindClusters() function was used to identify clusters with a resolution equal to 0.6. This clustering step implemented a modularity optimization algorithm named Louvain that was used to iteratively group cells together to identify highly connected clusters. The Louvain method detects clusters in large networks by maximizing a modularity score for each cluster (Blondel et al., 2008). The resolution parameter defined the granularity of the clustering and the default value of this parameter was 0.6. The increasing value of the resolution parameter leads to a higher number of clusters.

BISCUIT is a single-cell clustering method that normalises single-cell data, imputes dropouts, clusters cells based on similar gene expression, and teases apart technical variation from biological signals (Prabhakaran et al., 2016). Single-cell data are sparse. BISCUIT imputes dropouts in single-cell data using similar co-expression patterns that are shared across comparable cells. Furthermore, it also normalises data by inferring parameters of each cell type that accounts for different cells having a distinctive number of transcripts. As a result, the BISCUIT model concurrently clusters cells and identifies co-expression structures in an iterative process. The BISCUIT clustering results and cell

annotations of the Azizi et al 2018 (GSE114725) dataset were retrieved from GEO and this meta-data were incorporated into a Seurat object for downstream analysis.

Estimating the optimum number of clusters in a dataset is a challenging task and it can also be computationally intensive, but the popular single-cell clustering packages (e.g. Seurat and Scanpy) were developed to focus on speed and scalability to allow fast runs with large single-cell data. Given the relatively small number of cells used in this study and a desire for more in-depth analyses, an R package, Consensus Clustering Plus, was chosen to estimate the number of possible single-cell clusters and to assess the stability of the clusters (Wilkerson and Hayes, 2010). The samples were clustered with a k-means approach, which calculates the distance between each data point and a centroid to select it to a cluster. It aims to partition data into a number of clusters and the number is defined by the parameter k (Hartigan and Wong, 1979). Typically, a predefined k is used to cluster the data, but the consensus clustering method performs a series of clustering that involves subsampling from the data with a range of k values to identify the optimum number of cluster counts (k) in the data. In this case, k was selected between 2 and 20 with 80% cells resampled, repeated 2000 times. Then a consensus matrix was produced for each k to show the proportion of cells that fall into the same cluster over the times they occurred in the same subsample. The optimum number of clusters was determined using a combination of delta area plots, consensus matrix, and biological interpretation of the differentially expressed genes from the clusters. The delta area plot shows the relative change in area under the cumulative distribution function (CDF) with each increase in cluster count (k) (Monti et al., 2003). The recommended k based on the delta area plot is usually around the "knee point". The cluster information of each cell was then integrated into single-cell analysis pipelines in Seurat.

2.3.3 Deriving TAM signatures

Following consensus clustering analysis to establish the four clusters, the TAM signatures were derived using differential genes among the four monocyte and macrophage clusters, which aimed at

developing four orthogonal signatures that have no overlapping genes. The conventional method build in Seurat (i.e. FindAllMarkers() function) tended to produce differentially expressed genes from clusters with overlapping genes, which may affect the performance in the subsequent analysis. The Seurat FindAllMarker() function calculated differentially expressed genes from each cluster against the rest of the clusters combined. A more stringent approach was used in this study. Differentially expressed genes were calculated using the FindMarkers() function, which implements the Wilcoxon rank sum test by default. Between each cluster pair, only the overlapping genes across different cluster pairs were selected to be the most represented genes of the designated cluster. The differential gene expression analyses were limited to genes detected in at least 25% of the cells in each cluster and on average resulted in at least 0.25 log2 fold change. Only positively changed genes were calculated for each cluster. As a result, four orthogonal signatures with nonoverlapping genes were produced.

2.4 Refining macrophage signature using LASSO

To obtain the most predictive gene subsets within the LA-TAM signature without filtering out critical genes (Friedman et al., 2010), a gene selection step using the least absolute shrinkage and selection operator (LASSO) was performed. LASSO is a regression method that performs both variable selection and regularization in order to improve the prediction accuracy and interpretability of the gene signature (Tibshirani, 1996).

LASSO Regression equation $\sum_{i=1}^{n}(y_i - \sum_j x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} |\beta_j|$ is used to minimize the sum of squared errors between the observed outcomes and the predicted outcomes, while also minimizing the absolute value of the coefficients (βj). In this equation, *n* is the number of observations, *p* is the

number of features, y is the vector of dependent variables, x is the matrix of independent variables, and λ is the regularization parameter which controls the amount of shrinkage of the coefficients.

The first part of the equation is the residual sum of squares, which measures the difference between the predicted values and the actual values. The second part of the equation is the regularization term, which penalizes large coefficients and encourages the model to use only a subset of the features. The regularization parameter, λ, controls the strength of the regularisation, which it is usually chosen by cross-validation.

An R package, glmnet, was used to fit the LASSO regression. The dependent variable y was estimated using two methods. The first method was splitting samples into high and low categories using the median value of the gene signature score across METABRIC data. The second method was using maximum likelihood estimation with the gene signature scores and survival data to divide samples into high and low categories (Myung, 2003). LASSO was then performed with 10-fold cross-validation on samples that have agreements between both methods. The penalty parameter λ was selected based on the highest area under the curve (AUC) achieved from the dataset.

2.5 Survival Analysis

To understand the impact of macrophage populations on survival times of cancer patients. I performed Cox proportional hazards regressions, both univariate and multivariate. The Cox model describes the relationship between the event incidence (e.g. survival, disease progression) in a hazard function and a set of covariates (Bradburn et al., 2003). For univariate Cox regression, independent Cox regressions were performed using immune signature scores as continuous variables. Then the outputs (regression beta coefficients and p values) of multiple univariate Cox regressions were pooled together in a meta-table. For multiple multivariate analyses, individual immune signature scores were introduced into Cox multivariate analyses that included other clinical covariates (e.g. age, size, grade). Age and size were introduced as continuous variables and grade and stages were introduced

as categorical variables. To compare the regression beta coefficients of different macrophage signature scores, I extracted the exponentiated regression coefficients and p values from individual multivariate analyses and pooled the data in a second meta-table. An exponentiated regression beta coefficient (also known as hazard ratio) above 1 means the variable is negatively associated with survival. Furthermore, to understand the clinical relevance of the LA-TAM infiltration in breast cancer, I performed Kaplan-Meier analyses using LA-TAM signature and LASSO refined LA-TAM signatures with respect to hypoxia level and key receptor expression (ER, PR, HER2) using METABRIC data (Clark et al., 2003; Curtis et al., 2012; Kishore et al., 2010).

2.6 Statistical analysis

For parametric hypothesis tests between two populations, Student's t-test was performed. For multiple hypothesis testing of parametric data, multiple t-tests were performed and p values were adjusted using Benjamini-Hochberg (BH) procedure, which helps to avoid false positives. For parametric hypothesis tests of three or more populations, a one-way Analysis of Variance (ANOVA) was performed.

For nonparametric hypothesis tests, Wilcoxon's rank sum test was used between two populations. For multiple comparisons, Benjamini-Hochberg (BH) test was performed for corrections and to avoid false positives. For non-parametric hypothesis of three or more populations, the Kruskal-Wallis test was performed. A significant threshold of (adjusted) $p < 0.05$ or was selected.

3 ASSESSMENT OF TIICS ESTIMATING TOOLS IN TCGA

The tumour microenvironment (TME) is highly heterogeneous and contains a wide spectrum of immune cells, cytokines, growth factors, and the extracellular matrix (Cavallo et al., 2011). The ability to detect the interactions and signalling of tumour infiltrating immune cells (TIICs) within the TME could help to advance mechanistic understanding of cancer and potentially help to develop novel therapeutics (Wan et al., 2013). One way to decipher the complex interactions among the cellular components within TME is to apply immunogenomic methods on multi-omics datasets, which can be applied in both retrospective and prospective studies. With the ever-growing amount of new transcriptomics data from both bulk and single-cell levels, immunogenomic tools have become highly promising and rapidly adapted among medical researchers. However, several important questions must be addressed before one can be confident in the results obtained from immunogenomic analyses. This chapter is aimed at comparing the performance of the different TIICs estimating tools and assessing their limitations.

3.1 Immunogenomic TIICs estimating methods available

I comprehensively reviewed a list of TIICs estimating tools that are available in the public domains (Table 3.1). As an area of active development and new emerging tools becoming available, this list is focused on categorising the most representative tools, which are divided into two groups based on their different algorithmic approaches. The rationale behind estimating TIICs using computational tools and transcriptomics data can generally be divided into two categories: gene signature and

deconvolution. A more detailed description of both approaches has been provided in the introduction chapter.

In brief, the gene signature approach utilises a list of genes often bioinformatically derived using differential gene expressions from the gene expression profiles of specific immune cell types (Conesa et al., 2016). Such profiles can be retrieved from multiple gene expression datasets including FANTOM5, CAGE, ENCODE, HPCA, and BLUEPRINT. However, it is important to note that different research groups may have different protocols and prepared the cells differently, which could result in variations in gene expressions simply due to variations of protocols. Then these immune cell-specific signatures can be applied to different RNA-seq datasets to calculate a summary enrichment score, using algorithms such as gene set enrichment analysis (GSEA) (Subramanian et al., 2005). The leading example of the signature-based TIICs estimating tool is xCell, which utilises 489 different gene signatures derived from 1822 human primary cell samples (Aran et al., 2017).

In contrast, the deconvolution methods can be viewed as linear models. The deconvolution approach assumes that the total gene expression matrix (m) of a tumour sample equals the linear combination of each cellular component (Shen-Orr and Gaujoux, 2013). To estimate the relative immune cell fractions (f) of immune cell types, these deconvolution tools require inputs from prior knowledge to determine the gene expression profile of individual immune cells, which is called a signature matrix (S) (Finotello and Trajanoski, 2018). The objective of the gene expression model is to solve the following linear equation to determine f: $m = S \times f$ (Chen et al., 2018). For instance, CIBERSORT, a leading deconvolution tool, estimates the fractions of 22 immune types from tumour bulk transcriptomic datasets, using a signature matrix named LM22 (Newman et al., 2015).

There is no clear consensus on the optimum method for estimating TIICs in bulk RNA-seq; therefore I assessed the performance of the most representative tools (CIBERSORT and xCell) for in-depth comparison in 11,000 TCGA samples across 33 different cancer types.

Table 3.1 Current TIICs estimating methods in the literature

Method	Approach	Score	Cell types
CIBERSORT Relative mode (Newman et al., 2015)	Deconvolution	Immune cell fractions, relative to total immune cell content	22 immune cell types
CIBERSORT Absolute mode (Thorsson et al., 2018) (Deprecated)	Deconvolution	Absolute proportion of each cell type using methylation data	22 immune cell types
EPIC (Racle et al., 2017)	Deconvolution	Cell fractions, relative to all cells in sample	6 immune cell types, fibroblasts, endothelial cells
MCP-counter (Liu et al., 2015a)	Signature	Arbitrary units, comparable between samples	8 immune cell types, fibroblasts, endothelial cells
quanTIseq (Finotello et al., 2017)	Deconvolution	Cell fractions, relative to all cells in sample	10 immune cell types
TIMER (Li et al., 2016b)	Deconvolution	Arbitrary units, comparable between samples (not different cancer types)	6 immune cell types
xCell (Aran et al., 2017)	Signature	Arbitrary units, comparable between samples	64 immune and non-immune cell types
consensusTME (Jimenez-Sanchez et al., 2019)	Signature	Normalised enrichment score (NES)	18 immune and non-immune cell types

3.3 Results and Discussion

3.3.1 Assessment plan of immunogenomic dissection methods

I first sought to examine multicollinearity in each method, which is a challenge in estimating immune cell types using *in silico* methods. Multicollinearity refers to the high correlation of the transcriptomic profiles of closely related cell types, which could include shared genes in gene signature or deconvolutional signature matrix. This issue may cause an overestimation of low abundance cell types when a low abundance cell type expression is masked by a more abundant cell type (Finotello and Trajanoski, 2018; Shen-Orr and Gaujoux, 2013).

The tumour infiltrating immune compositions (TIICs) of 11,000 TCGA RNA-seq samples were estimated using CIBERSORT and xCell methods. I designed a series of bioinformatic studies to check if the two methods could produce comparable TIICs estimations on the 13 different immune cell types that CIBERSORT and xCell have in common. I also compared the performance of CIBERSORT and xCell in TCGA breast cancer separately, since it's a disease of interest in this thesis.

3.3.2 Multiple collinearity check of CIBERSORT and xCell

I assessed for multicollinearity in xCell and CIBERSORT to determine the utility of these algorithms in predicting independent cell types. No collinearity or low collinearity indicates independence in cell type prediction. Multicollinearity is likely due to non-mutually exclusive genes within immune signatures of closely related cell types. Hence, the cell predictions with multiple-collinearity will be noted and removed, if necessary, in the subsequent analyses.

CIBERSORT was used to estimate the immune fractions of 11,000 TCGA cancer samples using the LM22 immune signature matrix. To understand the associations among the 22 immune cell types and their estimations, Pearson's correlations were performed to measure the linear relationship between each pair of immune cell types. The results were visualized in a heatmap (Figure 3.1). Overall, the Pearson's correlation coefficients (i.e. Pearson's r) between different immune estimates were

negligible (|Pearson's r| < 0.3). However, a group of lymphocytes showed moderate or positive correlations (Pearson's r > 0.5), which include CD8 T cells, follicular CD4 T cells, regulatory CD4 T cells, and NK cells. Therefore, the results indicated that CIBERSORT has some level of multicollinearity among a subset of lymphocytes.

Of note, the strength of correlation coefficient was estimated using a guide from the literature that summarised in the following table (Mukaka, 2012).

Table 3.2 Size of correlation and interpretation

Size of correlation	Interpretation
±0.9 to 1.0	Very high positive/negative correlation
±0.7 to 0.9	High positive/negative correlation
±0.5 to 0.7	Moderate positive/negative correlation
±0.3 to 0.5	Low positive/negative correlation
±0.0 to 0.3	Negligible correlation

xCell was used to estimate the immune and non-immune infiltrations of 64 different cell types in 11,000 TCGA samples. Pearson's correlation coefficients were calculated among the 64 different cell estimates and the results were plotted in a heatmap for visualisation (Figure 3.2). Overall, most of the cell types have negligible correlations (|Pearson's r| < 0.3). However, several clusters were showing moderate and high positive correlations (Pearson's r > 0.5). The most noticeable two clusters included a subpopulation of myeloid cells and lymphocytes. The myeloid cell cluster included basophils, erythrocytes, monocytes, neutrophils, mast cells, common myeloid progenitor (CMP), granulocyte-monocyte progenitor (GMP), and multipotent progenitors (MPP). The lymphocyte cluster included two sub-clusters of T lymphocytes (CD4 T cell, CD4 naïve T cell, CD8 T cell, CD4 memory T cell) and B lymphocytes (B cell, pro B cell, memory B cell, naïve B cell class-switched memory B cell,

and plasma cell). The results suggested that closely related cells along the hematopoietic hierarchy tended to have higher associations, which may reduce xCell's performance in separating these cells apart (Seita and Weissman, 2010).

It is important to note that the multiple collinearity assessments were based on Pearson's correlations calculated from the TCGA pan-cancer data, which contains various cancer types. The correlations may vary between cancer types due to differences in sample size and cell type composition. For instance, solid tumours may have different correlations compared to blood cancers due to their distinct cellular makeup. Therefore, in the subsequent analysis, we focused on specifically assessing CIBERSORT and xCell in the context of breast cancer.

In summary, I demonstrated that both CIBERSORT and xCell have some multicollinearity with xCell demonstrating the greatest multicollinearity among groups of myeloid and lymphocyte cells that have a shared hematopoietic lineage. Therefore, xCell may predict more cell types than CIBERSORT, but it may not separate closely related cell types.

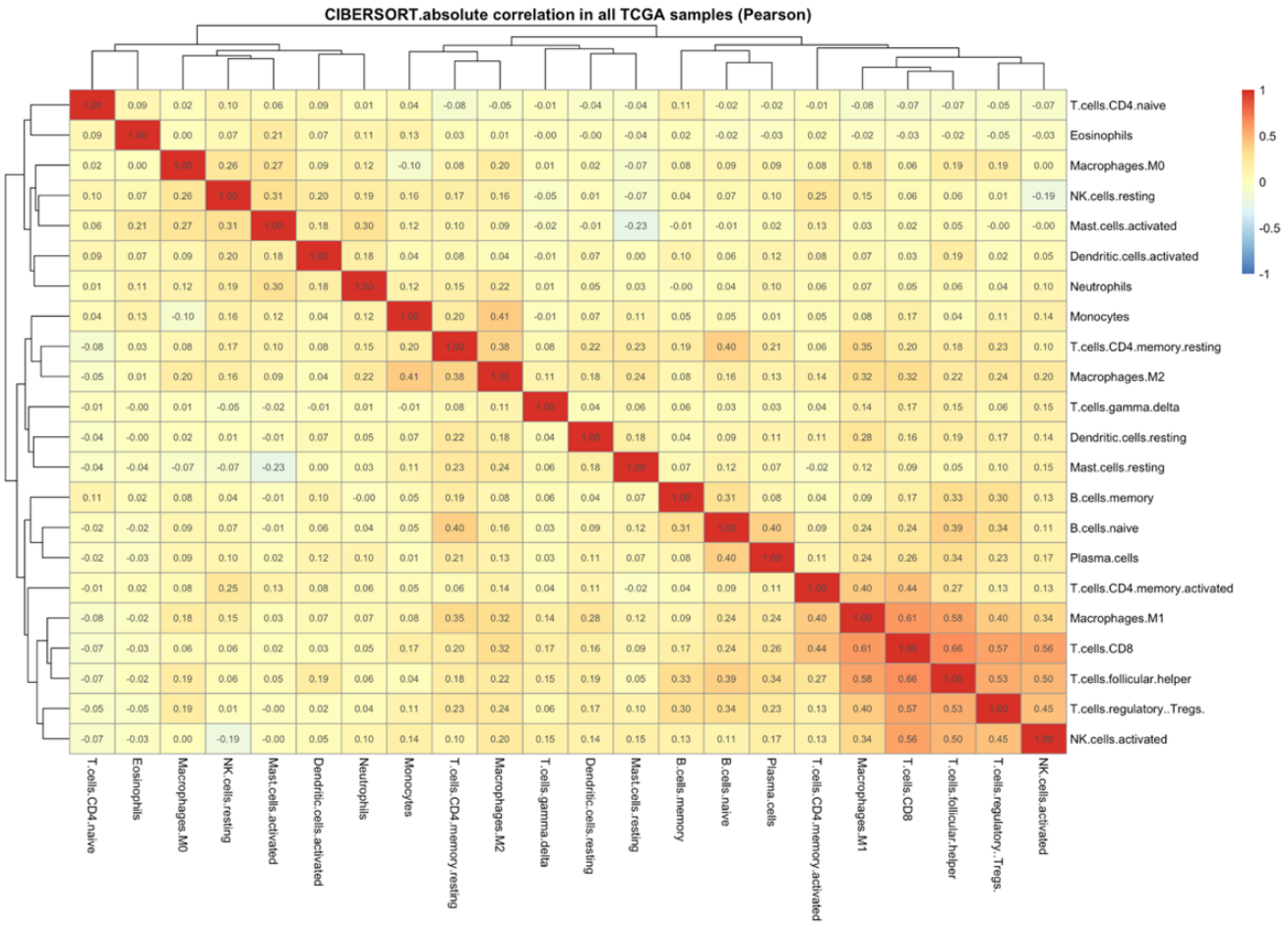


Figure 3.1 CIBERSORT multicollinearity assessment in 11,000 TCGA samples.

CIBERSORT estimations of 22 immune cell types were computed in 11,000 TCGA samples. Pearson's correlation analyses were performed among the 22 immune cell predictions and then the Pearson's correlation coefficients were plotted in a heatmap. Most of the immune cell predictions showed negligible correlations (|Pearson's r| < 0.3), but moderate or high correlations (Pearson's r > 0.6) were observed in closely related lymphocytes that include CD8 T cell, follicular T helper cell, regulatory T cell, and NK cells. Therefore, the result indicated that CIBERSORT has some level of multicollinearity among a subset of T and NK lymphocytes.

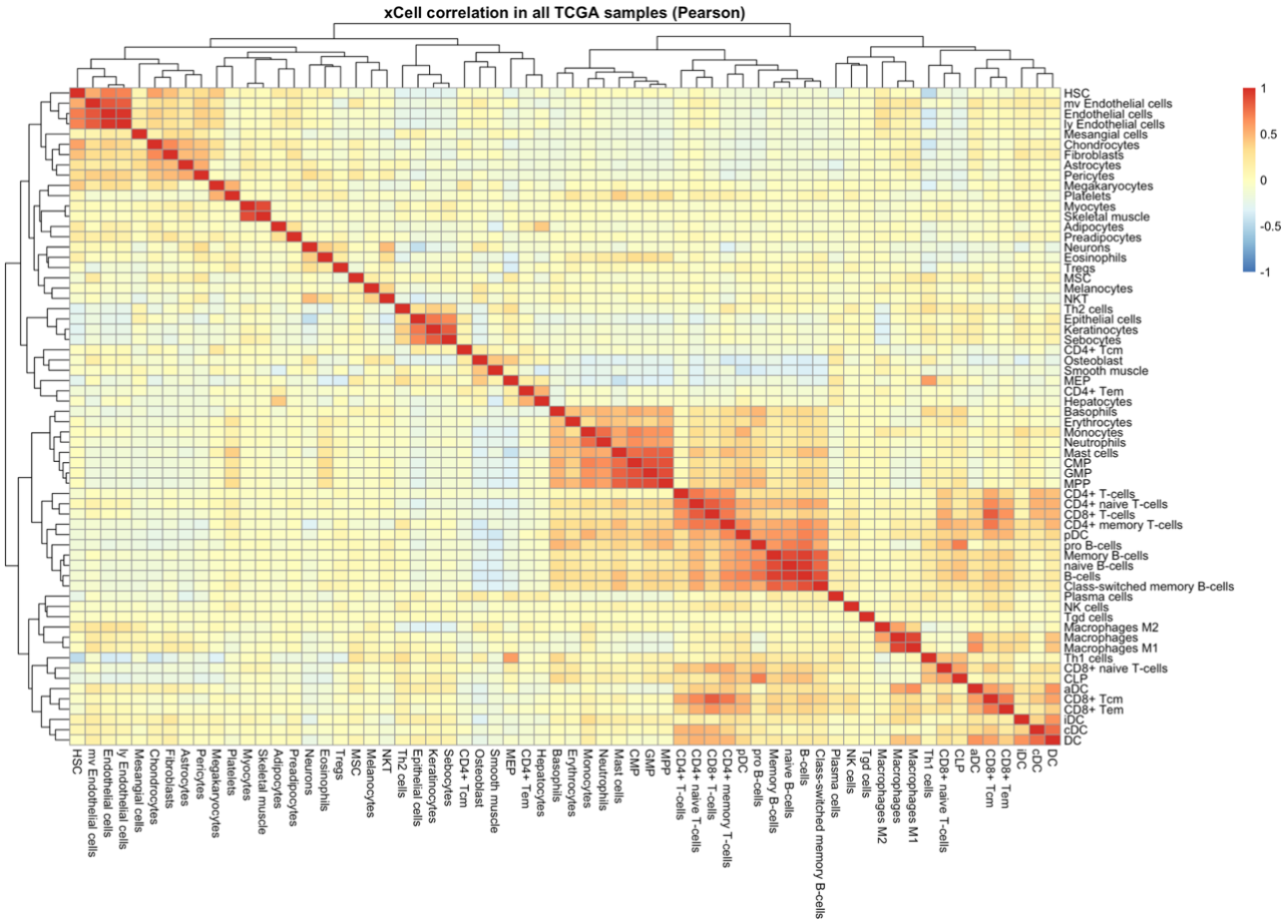


Figure 3.2 xCell multicollinearity assessment in 11,000 TCGA samples.

xCell estimated 64 different cell types from the gene expression data of 11,000 TCGA samples. Pearson's correlation coefficients were computed among the cell estimates and the results were plotted using a heatmap for visualisation. Overall, most of the cell types have negligible correlations (|Pearson's r| < 0.3). However, several cell clusters have moderate and high correlations (Pearson's r > 0.5), which include a myeloid cluster, a lymphocyte cluster and an endothelial cluster. The myeloid and lymphocyte clusters indicated that cells following a close related hematopoietic lineage may be difficult to resolve using xCell method. Hence, xCell may suffer from multicollinearity.

3.3.3 Assessment of CIBERSORT and xCell in TCGA

CIBERSORT and xCell immune estimations were calculated in 11,000 TCGA samples across 33 different cancer types. CIBERSORT can estimate 22 different immune cell types and xCell can assess 64 different cell types, but only 13 cell types were shared between the two tools. For head-to-head comparison purposes, only the overlapping 13 immune cell types and their estimations were considered in the subsequent analyses. Pearson's correlation coefficients were computed among the 13 immune cell estimations from the two tools, and the results were visualised in a heatmap (Figure 3.3). I observed 5 cell types that displayed low or moderate associations (Pearson's $r > 0.4$) between the two tools, which were: B cells, CD8 T cells, macrophage M1, macrophage M2, and neutrophils.

Next, I repeated the Pearson's correlation analysis of CIBERSORT and xCell predictions in 1,093 TCGA breast cancer samples. This breast cancer dataset would be used in further analyses in the subsequent chapters. The results were visualised using a heatmap (Figure 3.4). Similar to the pan-cancer analysis, only a subset of the 13 immune cell types had moderate comparable results between the two tools. I observed moderate associations in B cells, CD8 T cells, Macrophage M1, plasma cells, and neutrophils. Therefore, the comparison of CIBERSORT and xCell in TCGA pan-cancer and breast cancer indicated that the two *in silico* tools have comparable estimations in only selective groups of immune cells, which would be considered as the main cell types of interest in the next chapter.

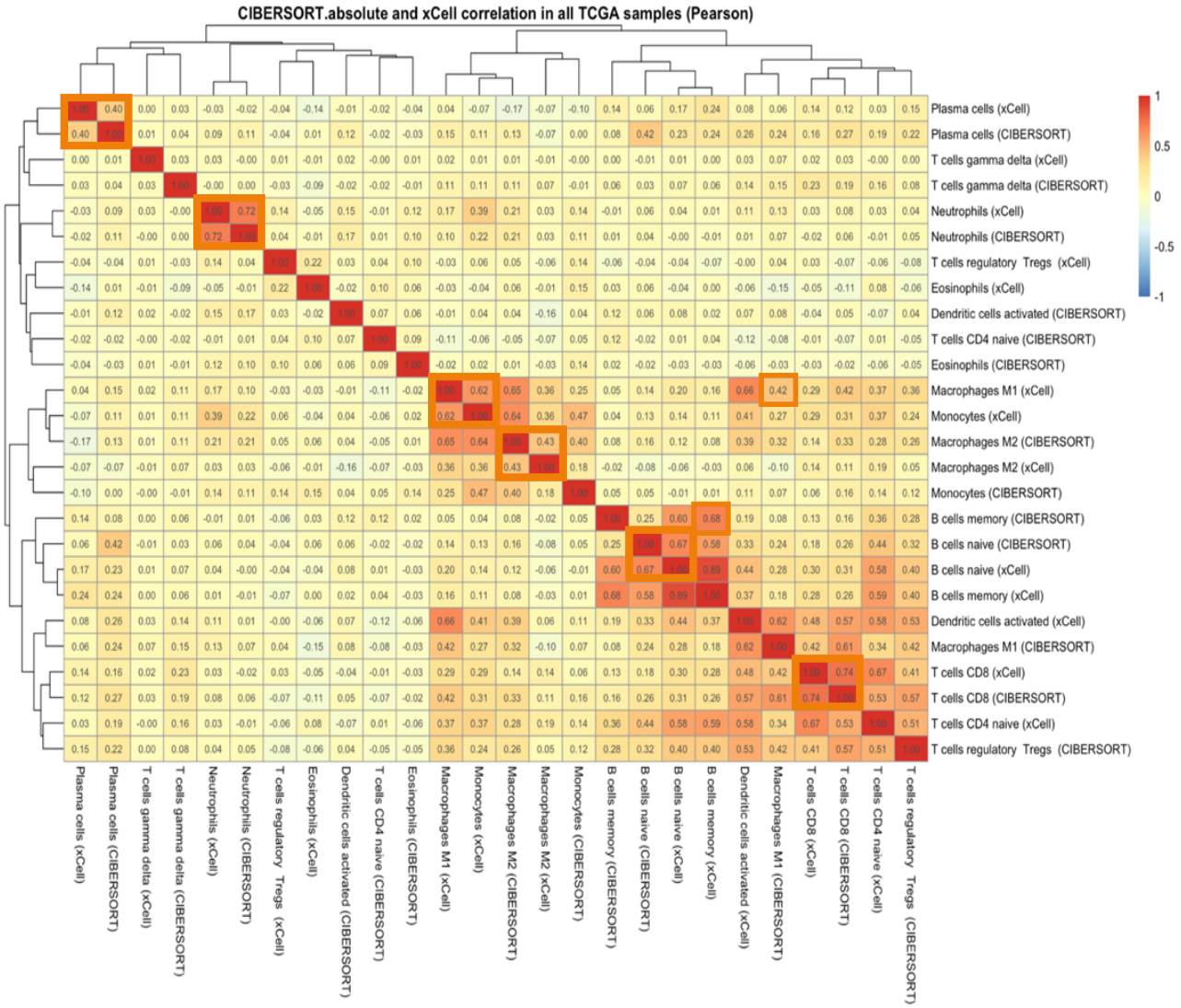


Figure 3.3 CIBERSORT and xCell comparison in 11,000 TCGA samples.

CIBERSORT and xCell immune estimations have 13 cell types in common, which were used in the head-to-head comparison in the pan-cancer TCGA dataset with 11,000 samples. Pearson's correlation coefficients were calculated using the immune estimations from the 13 cell types and the results were visualised in the above heatmap. Only five immune cell types had low or moderate associations (Pearson's r > 0.4). These cell types were B cells, CD8 T cells, macrophage M1, macrophage M2 and neutrophils. These cell types would be focused on in the subsequent analyses.

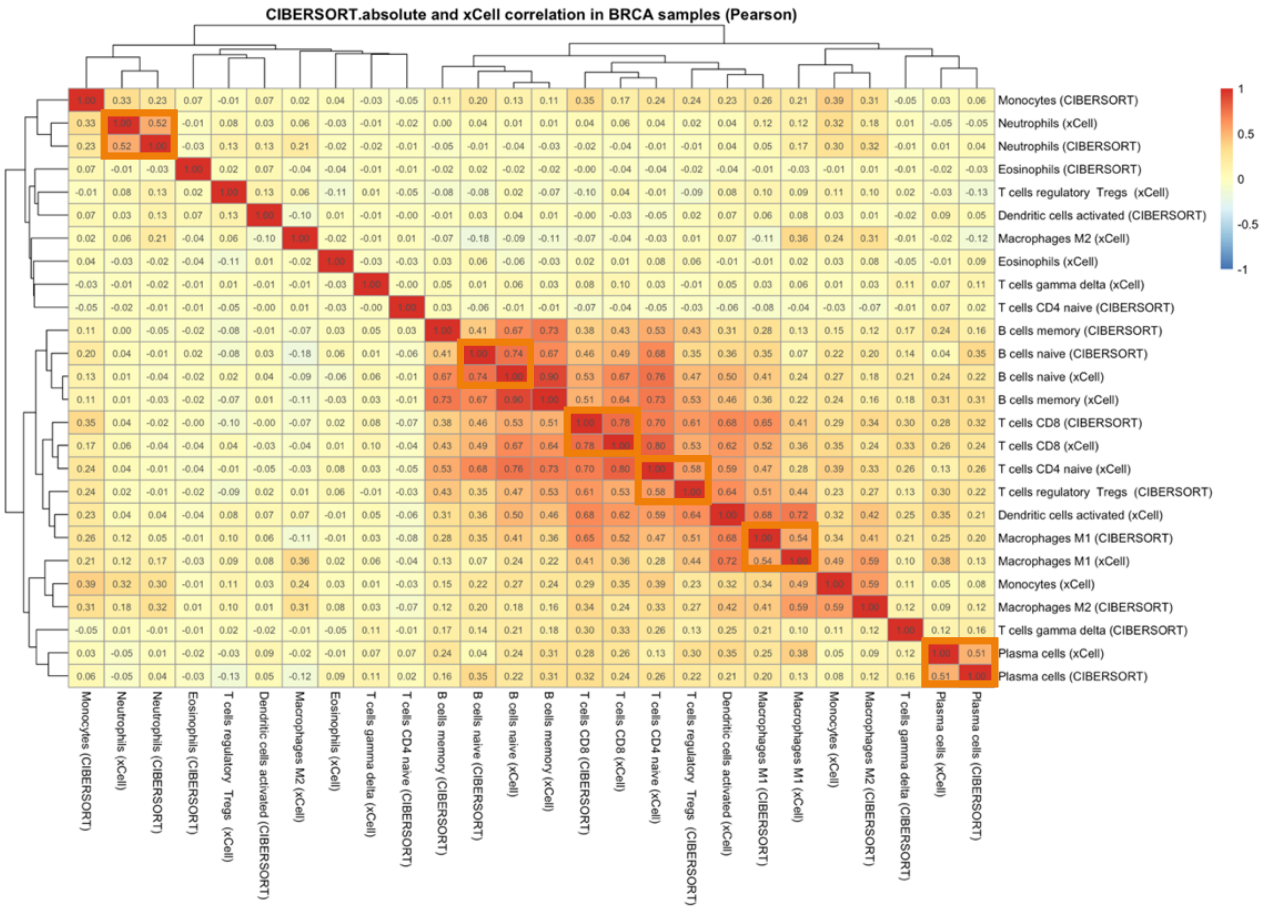


Figure 3.4 CIBERSORT and xCell comparison in TCGA breast cancer samples.

CIBERSORT and xCell immune estimations have 13 cell types in common, which were used in the head-to-head comparison in TCGA breast cancer dataset with 1093 samples. Pearson's correlation coefficients were calculated using the immune estimations from the 13 cell types and the results were visualised in the above heatmap. Only five immune cell types had low or moderate associations (Pearson's r > 0.4). These cell types were B cells, CD8 T cells, macrophage M1, plasma cells, and neutrophils.

3.3.4 Unexpected correlations on macrophage M1 and M2 estimations

To understand how different *ex vivo* macrophage signatures perform in tumour samples, I closely inspected the Pearson's correlation results from TCGA breast cancer analysis (Figure 3.4). Different from the traditional M1/M2 dichotomic model (Martinez and Gordon, 2014), I observed positive correlations among the theoretically polarised macrophage M1 and M2 populations across the CIBERSORT and xCell estimations (Figure 3.5).

Since tumour samples are heterogenous and may contain different immune populations, including the theoretical macrophage M1 and M2, I then assessed the expression of canonical markers of macrophage M1 and M2 in TCGA breast cancer dataset to see if the individual markers can resolve these cells instead of using *ex vivo* immune signatures (Figure 3.6). I noticed the macrophage general marker CD14 and a common TAM marker APOE were positively co-expressed with both M1 and M2 markers. Furthermore, macrophage M1 (TNF, CD80, CD86, TLR2, TLR4) and M2 (TGFB1, CD163, TLR8) markers also showed positive correlations (Murray et al., 2014). The results indicated that in bulk transcriptomic data of breast cancer, macrophage M1/M2 immune signatures and macrophage M1/M2 marker genes all displaced positive correlations with both of the polarised M1 and M2 populations; therefore unable to resolve them. A comprehensive set of analyses were designed to investigate this unexpected observation, the details can be found in Chapter 5.

3.3.5 ConsensusTME

Computational estimation of the TME components is an active field of research. One recent tool that has been developed in this area is called ConsensusTME (Jimenez-Sanchez et al., 2019). In a publication describing ConsensusTME, the authors compared its performance to that of seven other widely used immunogenomic TIIC tools and found that none of the methods consistently outperformed the others in all benchmarks. However, they also reported that ConsensusTME ranked in the top three in cancer-related benchmarks and was the overall best performing tool. It is important to note that ConsensusTME's gene signatures are derived from the union of signature genes from the

seven other tools, which could potentially introduce biases. In my analysis (Chapter 3&5), I have identified a critical issue with the use of *ex vivo* derived macrophage M1 and M2 signatures. These signatures have an unexpected positive correlation and may not accurately represent *in vivo* tumour-associated macrophages. Since ConsensusTME also contains components of macrophage M1 and M2 signatures originally derived from *ex vivo* signatures, it should be used cautiously when estimating macrophage populations. In general, the development of computational tools is a significant contribution to the field of cancer research, but it is important to remember that these tools and their results should be bound by or reflect biological principles.

3.4 Conclusion

In this chapter, a comprehensive review of the available immunogenomic TIIC estimating tools was conducted. I then focused on two of the most representative tools (CIBERSORT and xCell) for head-to-head comparisons. The results indicated high agreements on five cell types including CD8 T cells, B cells, macrophage M1, macrophage M2, and neutrophils, which will be focused on in the next chapter. In addition, I also found unexpected positive correlations among the theoretically distinctive macrophage M1 and M2 populations. This will be further discussed later in the thesis and a detailed analysis of M1 and M2 will be covered in chapter 5.

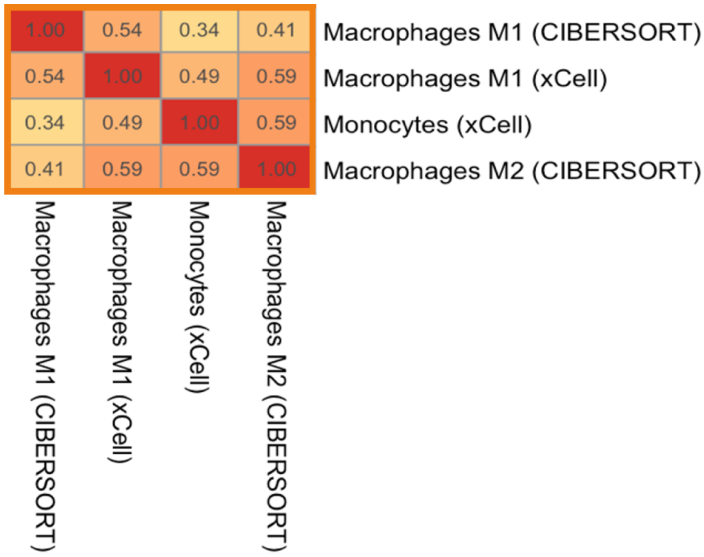


Figure 3.5 Positive correlations observed among macrophage M1 and M2 estimations.

CIBERSORT and xCell estimations of the macrophage M1 and M2 in TCGA breast cancer were compared. The resulted showed positive correlation between CIBERSORT macrophage M1 and M2 (Pearson's r = 0.41) and between CIBERSORT M1 and xCell M1 (Pearson's r = 0.54), and between CIBERSORT M2 and xCell M1 (Pearson's r = 0.59). According to the traditional macrophage M1 and M2 theory, these two populations should be polarised instead of co-expressed or positively correlated as demonstrated here.

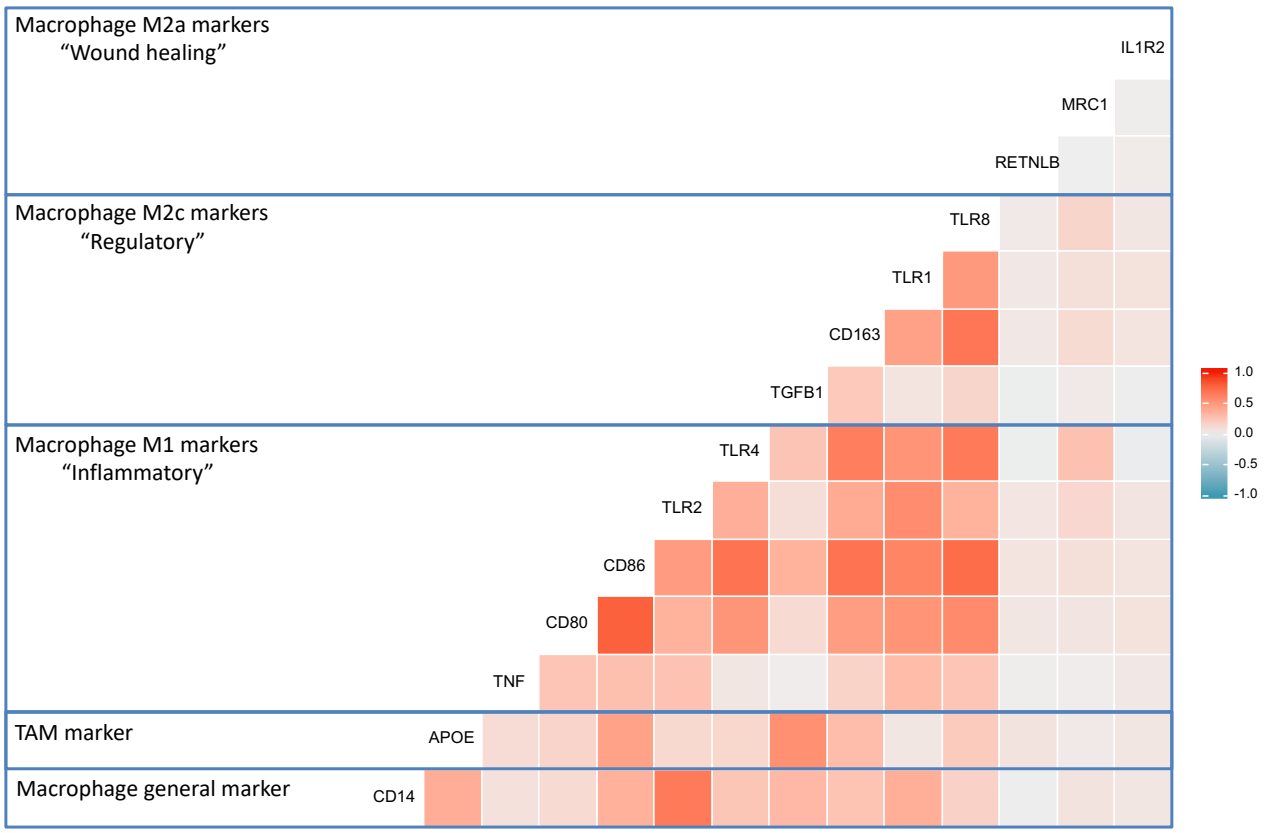


Figure 3.6 Macrophage canonical markers and their expression in TCGA breast cancer.

Macrophage general markers and polarization markers were identified from the literature (Murray et al., 2014). Pearson's correlation coefficients among the marker gene expressions in TCGA breast cancer were calculated and plotted in a heatmap. The result indicated a common TAM marker, APOE, co-expressed with both M1 and M2 markers. Furthermore, M1 and M2 individual markers positively correlated. For simplicity M2c is used interchangeable with M2 in this analysis. The result suggested canonical marker genes cannot distinguish macrophage M1 and M2 population in bulk transcriptomic data.

4 Association between Hypoxia and Immune Populations in Breast Cancer

4.1 Hypoxia and the tumour microenvironment

Tumours can display varying levels of hypoxia due to continual changes in the oxygen consumption requirements and insufficient oxygen delivery. These hypoxic environments subsequently change the gene expression networks in both the cancer cells and surrounding TME compartments (Bristow and Hill, 2008). In particular, hypoxia has been extensively shown to increase HIF pathway signalling (Heddleston et al., 2010). In addition to the hypoxic condition, the TME also contains a wide spectrum of cells, including immune cells, stromal cells, secreted factors and the extracellular matrix (Cavallo et al., 2011). The interactions among these cellular subpopulations and environmental factors may be an important mechanism driving disease progression. Therefore, identifying and characterising these changes using a streamlined quantitative approach may inform future mechanistic insights and guide the development of novel therapeutic strategies (Wan et al., 2013). In this chapter, I investigated the relationship between hypoxia and immune infiltration within the tumour microenvironment in breast cancer.

4.1.1 Establishing hypoxia score from hypoxia signature in breast cancer

To estimate tumour hypoxia levels, I utilised a hypoxia gene signature, which forms a proxy of hypoxia-related transcriptomic changes in cancer (Buffa et al., 2010). I calculated three summary scores using the hypoxia signature: mean, median and Principal Component 1 (PC1) in TCGA breast

cancer samples. The underlying assumptions are different, but by default, I use the median score when it has a high correlation with the mean and PC1. The mean hypoxia score works best when the dataset is normally distributed. The median score is best used when the data has a skewed distribution or the sample size is large. The PC1 hypoxia score captures the most critical variance of the signature in the dataset. Additional details of these summary scores are described in the method section.

In order to determine which hypoxia score was a better proxy of tumour hypoxia, I calculated the correlations among all three different hypoxia scores and a reference gene Carbonic Anhydrase 9 (CA9), which is a prominent marker for tumour hypoxia (Kaluz et al., 2009). The results showed that all three hypoxia scores strongly correlate with CA9 (Figure 4.1). Thus, all three scores were good proxies of hypoxia-induced transcription, and for consistency, I used the median score in all subsequent analyses.

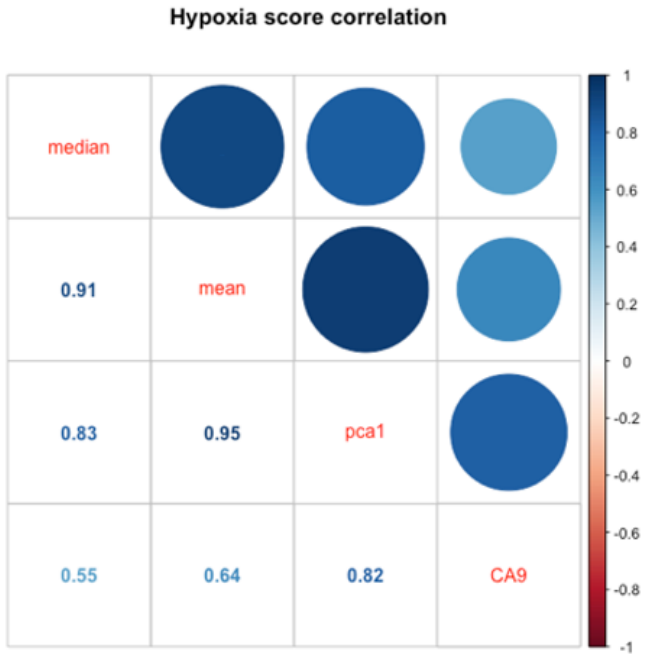


Figure 4.1 Hypoxia scores and their correlations with hypoxic reference gene CA9.

All three hypoxia summary scores (mean, median, PC1) were strongly correlated with a prominent tumour hypoxia marker, CA9. Pearson correlation coefficients are shown in the bottom left boxes.

4.1.2 Hypoxia levels in PAM50 and triple-negative breast cancer subtypes

Breast cancer is a heterogeneous disease, which is comprised of distinctive subtypes that have different molecular, cellular, and prognostic features (Loibl et al., 2021a). Different breast cancer subtypes can be classified using the expression of estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), and proliferation marker Ki-67 (Johnson et al., 2021). In the TCGA dataset, immunohistochemistry (IHC) staining was used to measure the expression of ER and PR. IHC supplemented with fluorescence in situ hybridization (FISH) was used to determine HER2 amplification or overexpression (Koboldt et al., 2012). HER2 positive is defined as HER2 protein overexpression measured by IHC or HER2 gene copy number of six or more or a HER2/CEP17 ratio of 2.0 or greater measured by FISH (Loibl and Gianni, 2016). Using the receptor information, breast cancer can be classified into four major subtypes: Luminal A ($ER^+/PR^+/HER2^-/Ki\text{-}67^-$), Luminal B ($ER^+/PR^+/HER2^-/Ki\text{-}67^+$), HER2-Positive ($ER^-/PR^-/HER2^+$), and triple-negative ($ER^-/PR^-/HER2^-$). Luminal B tumours can be distinguished from luminal A tumours by higher histologic grade, and reduced PR expression (Johnson et al., 2021). Triple-negative breast cancer (TNBC or TNB) accounts for 15-20% of all breast cancers and they are the most aggressive group among all four clinical subtypes. It has been recently recognised that there are additional heterogeneities within the TNBC subtype (Marra et al., 2020). Using hypoxia signature, I have shown that TNBCs have significantly higher hypoxia levels than the rest of the breast cancer subtypes in TCGA (Figure 4.2). Future work could explore the hypoxia levels within the TNBC heterogeneity and their effects on prognosis.

PAM50 is a gene signature specific for breast cancer that helps to stratify tumours into five intrinsic biological subtypes (Liu et al., 2016). The subtypes are luminal A, luminal B, HER2-enriched, Basal-like and normal-like, each of which has different prognostic profiles. It has been reported that patients with the Luminal A subtype have a more favourable outcome and a high sensitivity to endocrine therapy. In contrast, Basal-like subtypes are associated with poor clinical outcomes and high treatment resistance (Liu et al., 2016). The Basal-like and TNBCs have a lot of overlaps. Using the

available TCGA metadata, I demonstrated that almost 80% of TNBCs are Basal-like, and not all Basal-like are TNBCs (Figure 4.3).

To understand the relationship between hypoxia and PAM50 subtypes, I calculated a hypoxia score for each of the TCGA breast cancer samples and then stratified the samples using PAM50 subtypes. Here I report a significant difference in hypoxia levels in the PAM50 subtypes (Figure 4.4). The Basal-like subtype has the highest hypoxia level, followed by the HER2-enriched subtype. Luminal A has the lowest hypoxia level. Both HER2 and Basal-like represent more aggressive breast cancer phenotypes and feature higher hypoxia scores (Martin-Castillo et al., 2013). This pattern is consistent with previous studies, which have also shown that high tumour hypoxia correlates with a poor survival rate (Harris, 2002).

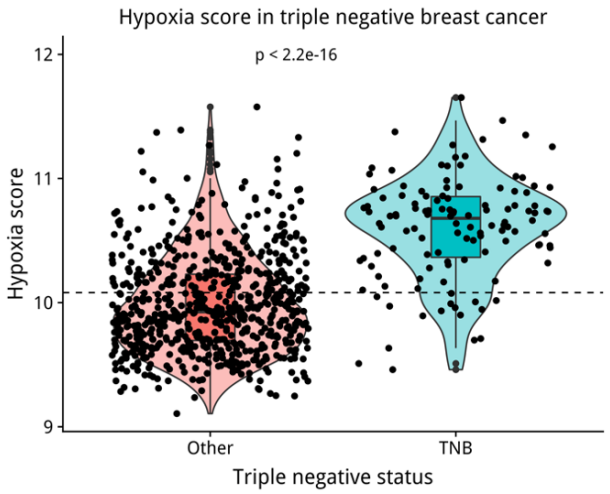


Figure 4.2 Triple-negative samples have the highest hypoxia level in breast cancer samples.

P-values were obtained using the Wilcoxon rank-sum test.

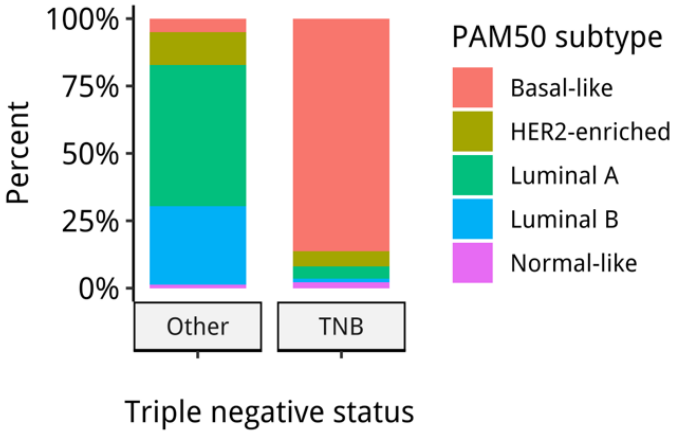


Figure 4.3 Relationship between TNBC and PAM50 in TCGA breast cancer samples.

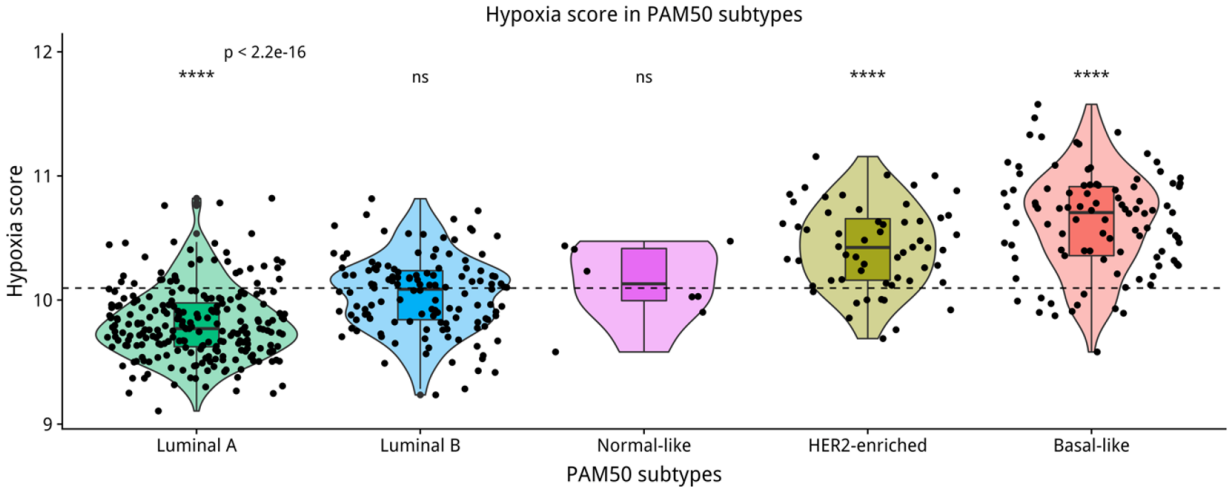


Figure 4.4 Variations of hypoxia levels in PAM50 subtypes.

Basal-like subtype has the highest hypoxia level and is also the more aggressive subtype clinically. The Kruskal-Wallis test was used to obtain p-values for global comparison of multiple groups, and then the Wilcoxon rank-sum test was used to compare each group with the mean of all groups. Dashed lines were used to label the group means, and the following symbols were used to indicate the level of statistical significance: ns $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

4.2 Immune cell compositions and hypoxia in breast cancer

The hypoxic tumour microenvironment has been shown to influence the function of a wide range of immune cells and influence tumour development (Petrova et al., 2018). Here I wanted to combine tumour infiltrating immune cell estimating tools and the 47-gene hypoxia signature to understand which immune cells are enriched under hypoxia in breast cancer. I concentrated the immune cell analysis on the 5 cell types that were verified across CIBERSORT and xCell with an additional CD4 T cell (Chapter 3).

4.2.1 Immune compositions in PAM50 subtypes

I first investigated the relationship between immune composition and PAM50 subtypes in TCGA breast samples (Figure 4.5). I observed a significant difference in B cells, macrophage M1 and M2, but not on CD8 T cells, CD4 T cells and Neutrophils. Interestingly, I noticed a reciprocal relationship between M1 and M2 among the PAM50 subtypes, which featured a high level of M1 and a low level of M2 in the Basal-like subtype. This is surprising as previous literature has suggested that M2 is associated with high hypoxia and aggressiveness (Henze and Mazzone, 2016b). However, from a metabolic perspective, this finding seemed logical, due to the fact that both hypoxic cells and macrophage M1 have heavily relied on glycolysis instead of oxidative phosphorylation (Sadiku and Walmsley, 2019). Additionally, this finding suggested that hypoxia could be a key driver of tumour associated macrophage polarization since the intrinsic PAM50 subtypes have very different hypoxia levels (Figure 4.5). Additional experiments could help to validate this metabolic relationship.

4.2.2 Immune compositions in TNBC

TNBC is the most aggressive breast cancer based on clinical subtypes. We, therefore, wanted to understand the immune composition behind the aggressive biological features. I stratified the TCGA breast cancer samples into TNBC subtypes and other subtypes. In TNBC samples, I observed a

significant increase of B cells and macrophage M1, but a reduction of macrophage M2 and neutrophils (Figure 4.6).

It was unclear from my data if these B cells were antigen-specific since TNBCs also have increased genomic instability which may result in the production of neoantigens (Karn et al., 2017). A recent study has shown that tumour infiltrating B cell level was elevated in TNBC and these cells presented IgG clonal expansion. Additionally, IgG humoral immunity was associated with a good prognosis in TNBCs (Harris et al., 2021). Thus, high infiltrating B cells within the TNBCs may be associated with a favourable prognosis and could be explored in novel immunotherapy-based strategies.

In PAM50 subtypes, I found that significant enrichment of macrophage M1 in more hypoxic tumours such as Basal-like and HER2-like (Figure 4.5). I again found a significant enrichment of macrophage M1 in TNBC (Figure 4.6). Interestingly, the relationship between M1 and M2 was also shown in a reciprocal fashion in TNBCs. I hypothesised that the hypoxic environment within the TNBCs offered a better physiological fit with the macrophage M1, since both hypoxic cancer cells and macrophage M1 are better adapted within the low oxygen level via alternative metabolism featuring glycolysis (Sadiku and Walmsley, 2019). Hence, targeting alternative metabolism pathways could be synergistic to address both cancer and tumour associated macrophage in TNBC.

Next, I further stratified the TNBC samples using the hypoxia signature (Figure 4.7). The results showed a significant difference in B cells, macrophage M1, macrophage M2, and neutrophils. I observed the TNBCs with low hypoxia level has a higher level of B cell infiltration level, which suggested that TNBCs may yield neoantigens that led to B cell infiltration but too much hypoxia may negatively affect the level of B cells. TNBCs with high hypoxia also have the highest level of macrophage M1 infiltration level, which again resonated with altered metabolic profiles of these macrophages. Overall, I can conclude that TNBCs and hypoxia synergistically featured macrophage M1 enrichment in breast cancer, which I will cover more in-depth in the subsequent chapter 5.

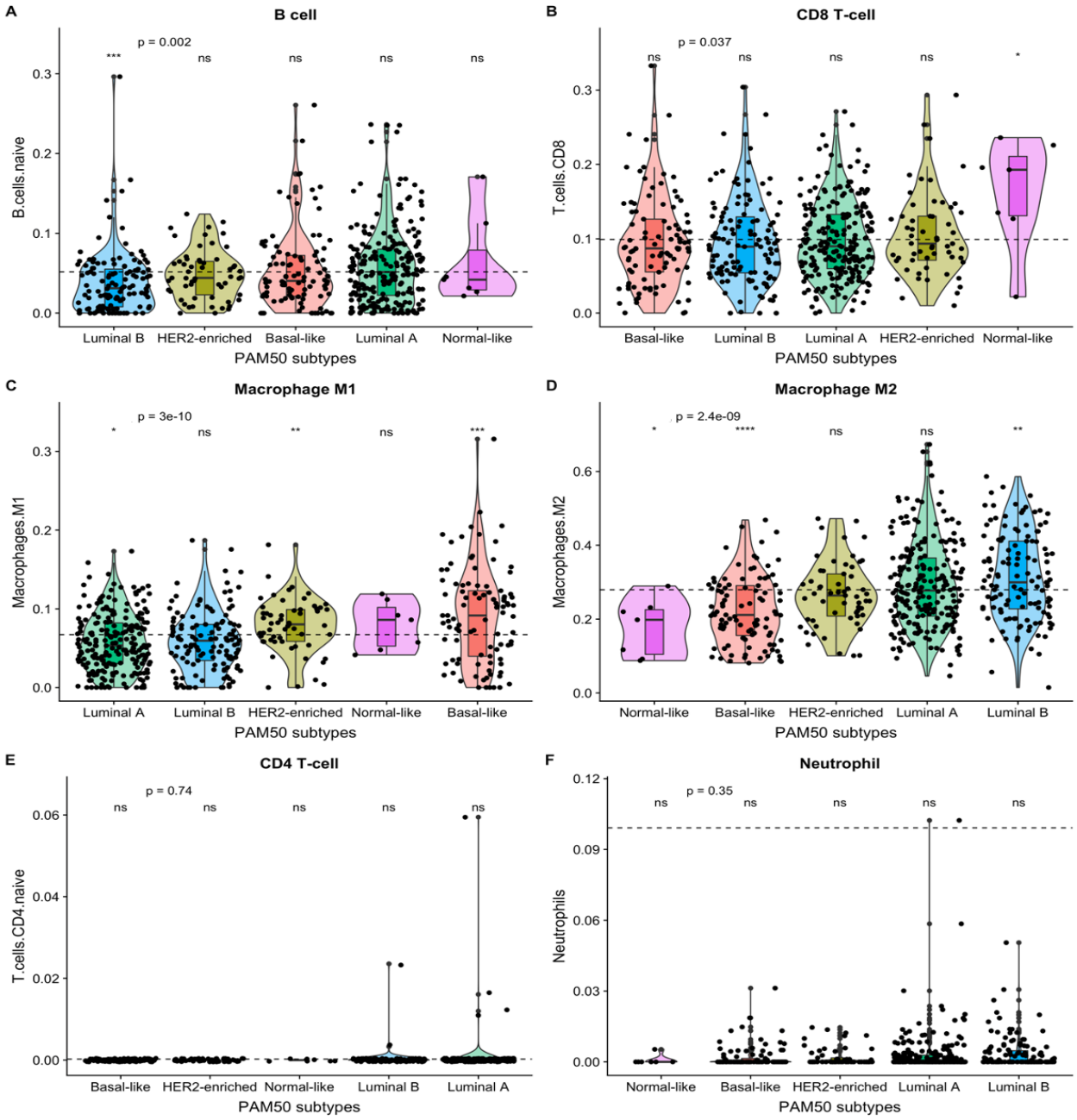


Figure 4.5 Relationships between PAM50 subtypes and immune cell compositions.

More aggressive subtypes of Basal-like and HER2 are enriched with macrophage M1. The x-axis is reorganized using the mean expression level of each group from low to high. P-values for global comparison of multiple groups were obtained using the Kruskal-Wallis test, followed by the Wilcoxon rank-sum test to compare each group with the mean. The group means were labelled with dashed lines, and the following symbols were used to indicate the level of statistical significance: ns $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

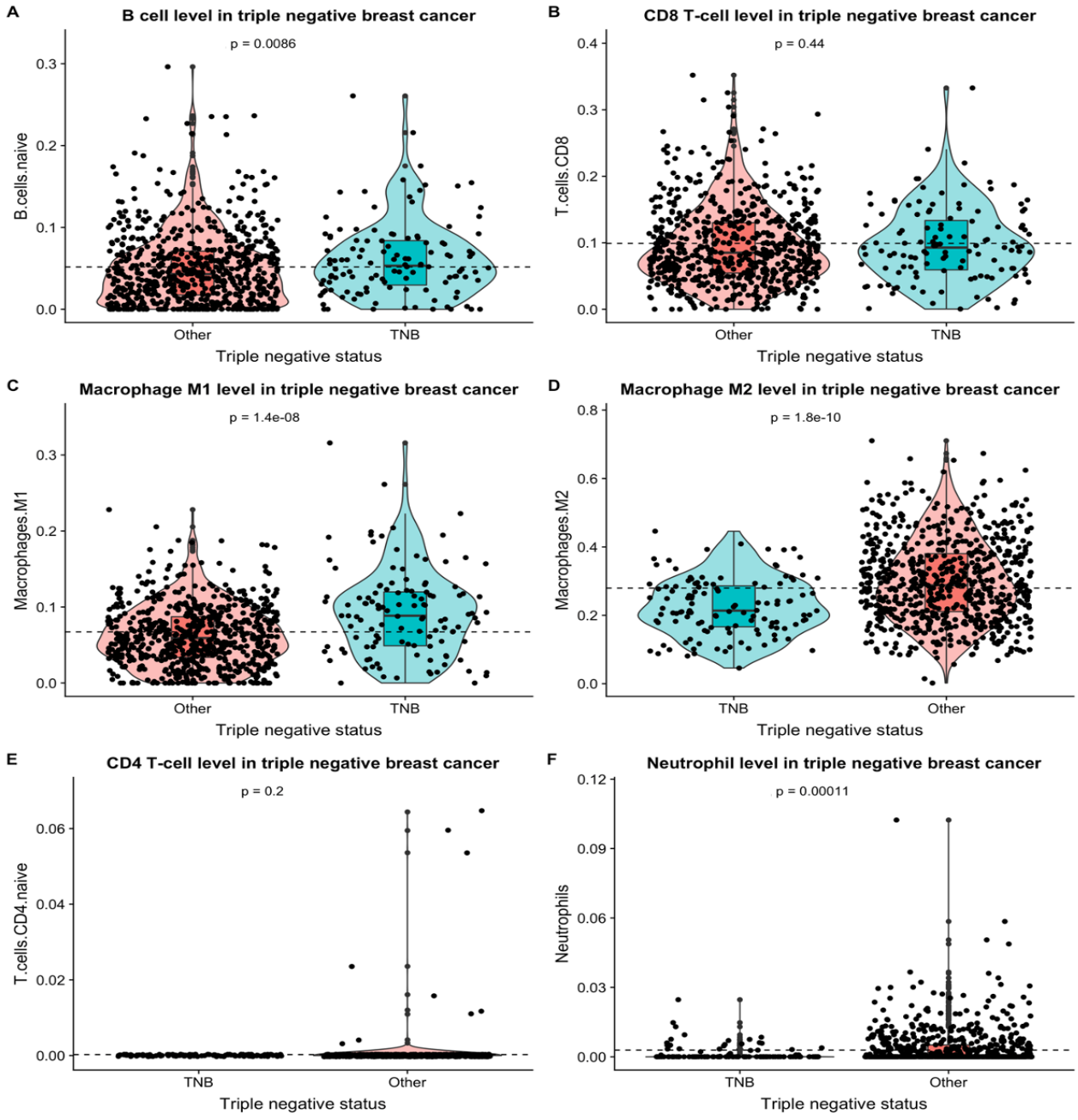


Figure 4.6 Relationships between triple-negative breast cancer and immune compositions.

The aggressive subtype TNBC is enriched with macrophage M1. The x-axis is reorganized using the mean expression level of each group from low to high. P-values were obtained using the Wilcoxon rank-sum test for comparison of two groups. The group means were labelled with dashed lines.

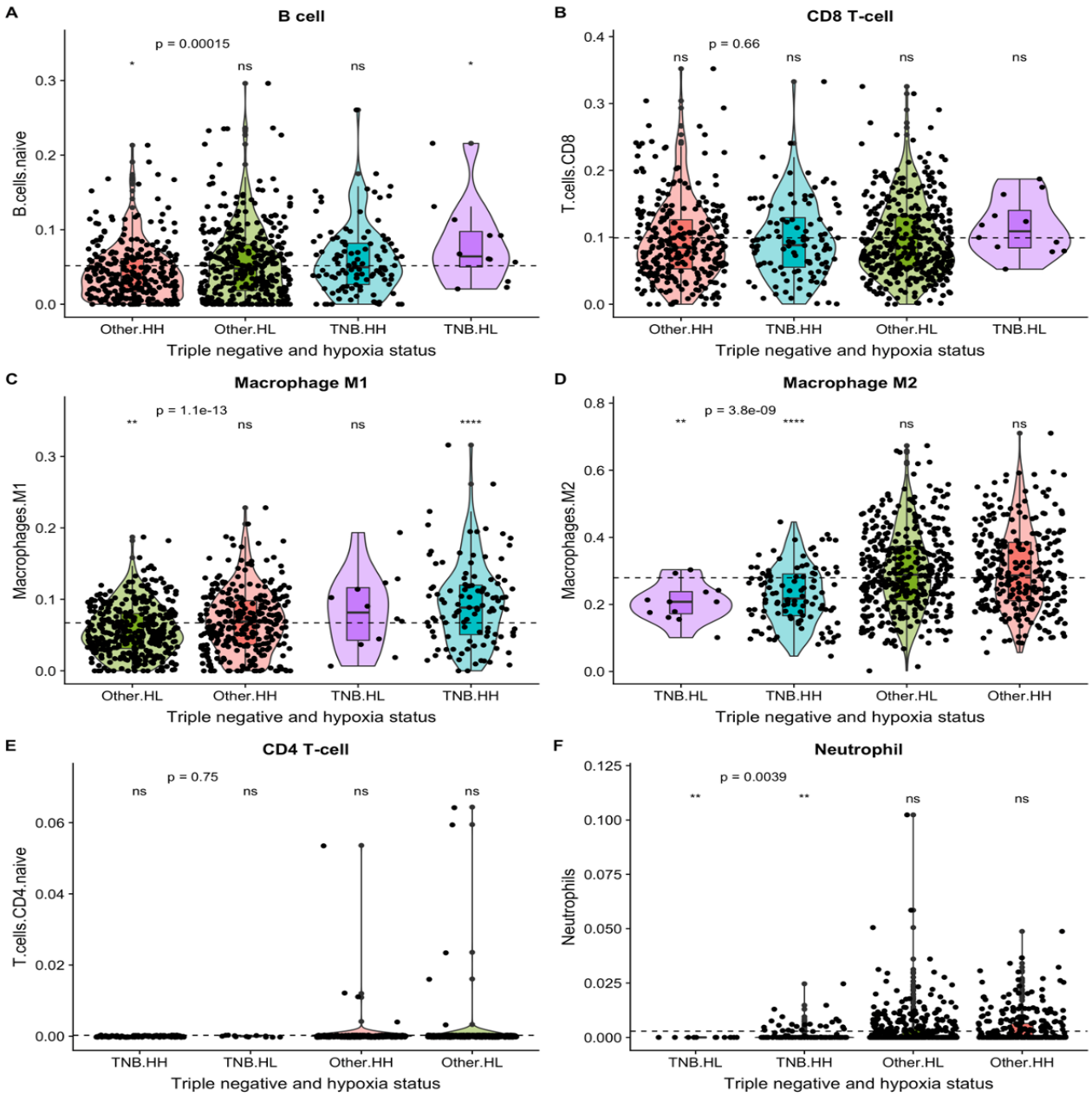


Figure 4.7 Relationships between hypoxia and immune compositions in triple-negative breast cancer.

TNBC with high hypoxia is enriched with macrophage M1. The x-axis is reorganized using the mean expression level of each group from low to high. P-values for global comparison of multiple groups were obtained using the Kruskal-Wallis test, followed by the Wilcoxon rank-sum test to compare each group with the mean. The group means were labelled with dashed lines, and the following symbols were used to indicate the level of statistical significance: ns $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

4.3 Stromal and immune landscape in hypoxic environment

Beyond the immune compartment, the stromal cells also play important functions within the tumour microenvironment (Munn and Bronte, 2016; Wu et al., 2020). I sought to understand the relationship among hypoxia, stroma and immune cells in breast cancer using gene signatures from xCell, which can estimate the infiltrations of a range of cells in the tumour microenvironment. The goal of the following non-biased exploratory analyses was to discover novel biological relationships and to generate new hypotheses.

Principal Component Analysis (PCA) is a multivariate statistical technique which is commonly used to emphasize variation and visualize patterns in a dataset (Ringnér, 2008). I applied PCA to study the relationship between different cell populations and hypoxia in xCell predictions. The conventional application of PCA in biology is to reduce the dimensionality of a dataset (Yao et al., 2012). Here I focused on studying the relationships of the PCA variables to understand the strength of correlations among the individual cell types.

4.3.1 Relationship between hypoxia, stroma and immune cells in breast cancer

xCell estimations of 64 different cell types were calculated in 1,093 TCGA breast cancer samples. In addition, I incorporated the Buffa hypoxia signature and its summary score in the PCA analysis. The xCell summary scores and hypoxia scores of the cancer samples were aggregated together to form a data table for the PCA analysis. I performed PCA analysis on the data table and extracted the Cos^2 score (the squared cosine) for each variable (i.e. immune cell scores or hypoxia score). The Cos^2 score of variables in different PCs was assembled in a summary table (**Figure 4.8**).

Cos^2 is an indicator of the quality of representation for each variable. The Cos^2 also shows the most relevant PCs for each variable (Abdi and Williams, 2010). In practice, I used the Cos^2 table to identify the best PCs to investigate the relationship between variables of interest. For example, if I wanted to understand the relationship between CD8 T cells and mast cells, I can check the Cos^2 table and determine that a significant portion of these two cells' variance is captured by PC1 and PC2. High

Cos^2 from two PCs indicates the variable is better represented in the specific two PCs; therefore, the two PCs should be focused on in the subsequent analyses (Abdi and Williams, 2010).

Since I wanted to understand the relationship of three groups of variables: hypoxia, stromal and immune cells, I searched the PCs where a significant portion of their variants was captured. From the Cos^2 table, I found that the variances of immune and stromal contents are predominately captured by PC1 and PC2 respectively. The Cos^2 result indicated that the hypoxia variances were divided mainly in PC2, PC3, and PC4. Therefore, plotting PC1 vs PC2 variables will provide a good estimate of the relative relationship among hypoxia, stromal and immune cells.

Plotting the variables of PC1 vs PC2 revealed an inverse relationship between hypoxia and stromal content featuring the endothelial cells (Figure 4.9). This finding matched with previous reports that tumour hypoxia is due to inadequate oxygen supply caused by extensive vasculature abnormalities, which include low vessel density, disorganized vessels, aberrant vessel diameters and abnormal vessel walls (Gaustad et al., 2017; Laitala and Erler, 2018). In addition, I also identified a positive relationship between hypoxia and macrophages (Figure 4.9), which matched with previous reports that tumour hypoxic regions enriched with the infiltrations of immunosuppressive cells that included macrophages (Casazza et al., 2013; Chiu et al., 2016). I will investigate further the relationship between hypoxia and different macrophages in more detail in the subsequent chapters.



Figure 4.8 Cos^2 variable analysis of the first 16 PCs

Cos^2 is a measure of the representation quality of each variable in principal component analysis. The sum of Cos^2 for each variable is equal to 1, and the circles in the figure show the Cos^2 scores for each variable as a percentage in each principal component (i.e., a full circle equals 100%). The most relevant principal component (PC) pairs for hypothesis generation analyses were selected using the Cos^2 table from the figure. For example, the relationship between hypoxia and T cells was investigated using PC1 (Dim. 1) and PC2 (Dim. 2). PCs with high Cos^2 for hypoxia, mast cells, and macrophages are labelled with red boxes, and PCs with high Cos^2 for CD8 T-cells, CD4 T-cells, and B cells are labelled with green boxes.

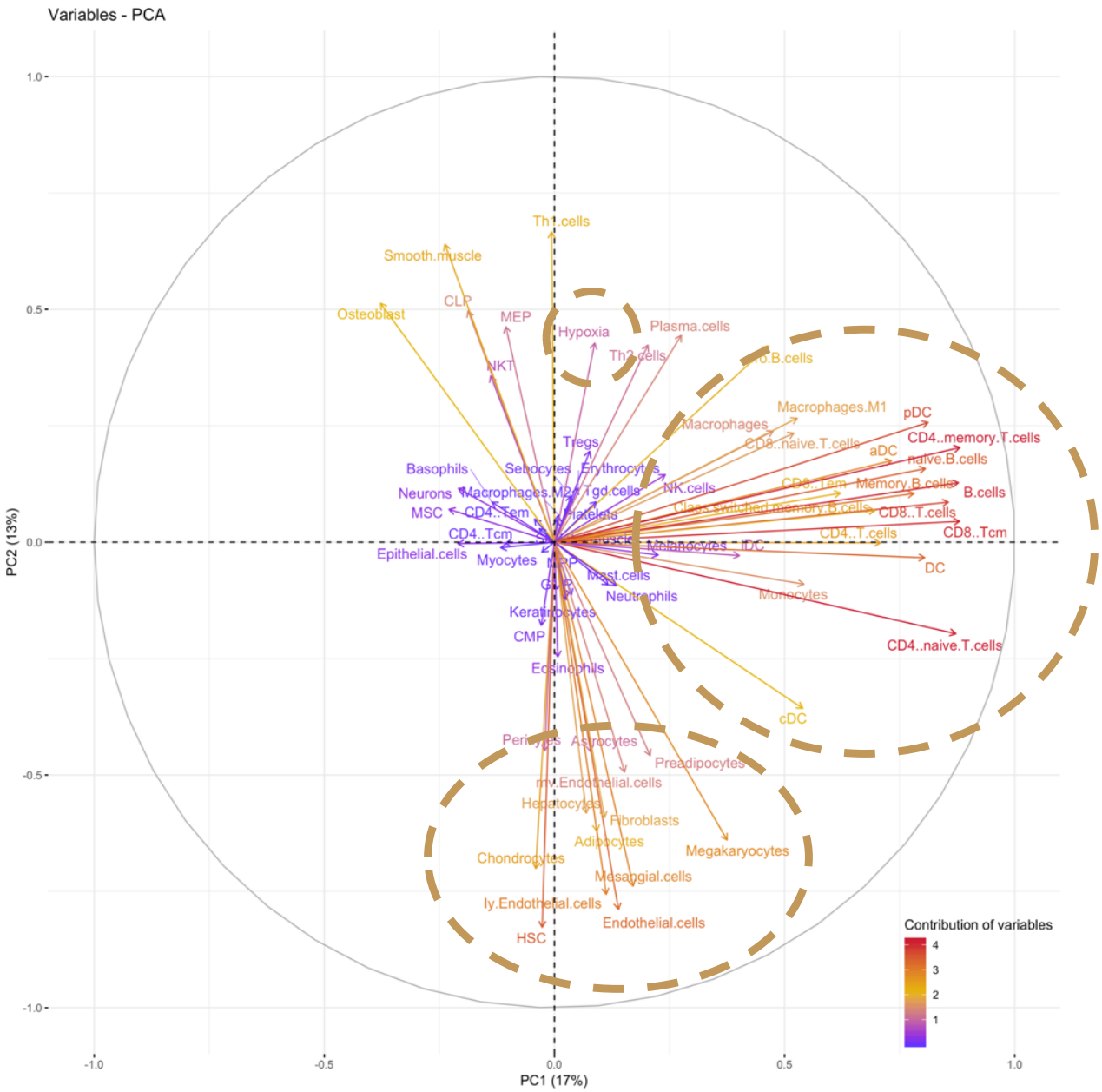


Figure 4.9 Variable plot of PC1 and PC2 from xCell estimations in BRCA

xCell estimations of 64 different cell types and hypoxia scores were calculated in TCGA breast cancer samples. The first two PCs were selected to characterise the relationship among hypoxia, stromal cells and immune cells. An inverse relationship was identified between hypoxia and the endothelial components. Macrophage was found to be positively associated with tumour hypoxia. Each arrow indicated a cell type and its impact on the PCs. For each variable, the sum of the Cos2 of all components is equal to one, which is marked by the outer circle with a radius of one. Dash-cycles highlighted hypoxia, stroma, and immune components.

4.3.2 Relationship between hypoxia, macrophages and mast cells

I conducted an exploratory analysis to understand the relationship between hypoxia and mast cells, because the role of mast cells in breast cancer with respect to hypoxia is less well understood (Aponte-López et al., 2018; Hanes et al., 2021; Varricchi et al., 2017). Searching the Cos^2 table (**Figure 4.8**) indicated that the majority of mask cells and hypoxia's variance are captured by PC3 and PC5 (**Figure 4.8**), which also contain a significant amount of variances from epithelial cells and macrophages. I then plotted PC3 and PC5 (Figure 4.10) in a biplot to investigate the relationship between hypoxia and mast cells. The biplot contained information from the variables with overlaid hypoxia information in three ellipses, which divided the hypoxia scores into three levels: high, medium, and low.

I observed a strong inverse relationship between mast cells and hypoxia (Figure 4.10), which implied human breast tumours with high infiltrating mast cells have reduced hypoxia levels. The role of mast cells in cancer remains elusive, but some reports suggested mast cells may have protective roles in breast cancer. For example, in a cohort of 4444 breast cancer patients, high tumour-associated mast cells were associated with a favourable prognosis, independent of age, tumour grade, tumour size, lymph node, and molecular subtype (Aponte-López et al., 2018). Studies of murine breast cancer models have shown that mast cell-deficient mice develop more abundant and larger tumours than mass cell competent controls (Samoszuk and Corwin, 2003a), which also had more extensive hypoxia with haemorrhagic regions (Samoszuk and Corwin, 2003b). Therefore, my data is in accordance with previously published literature that high mast cells in breast cancer could reduce tumour hypoxia and increase anti-tumour activities, but the detailed mechanisms of hypoxia reduction by mast cells require further investigation.

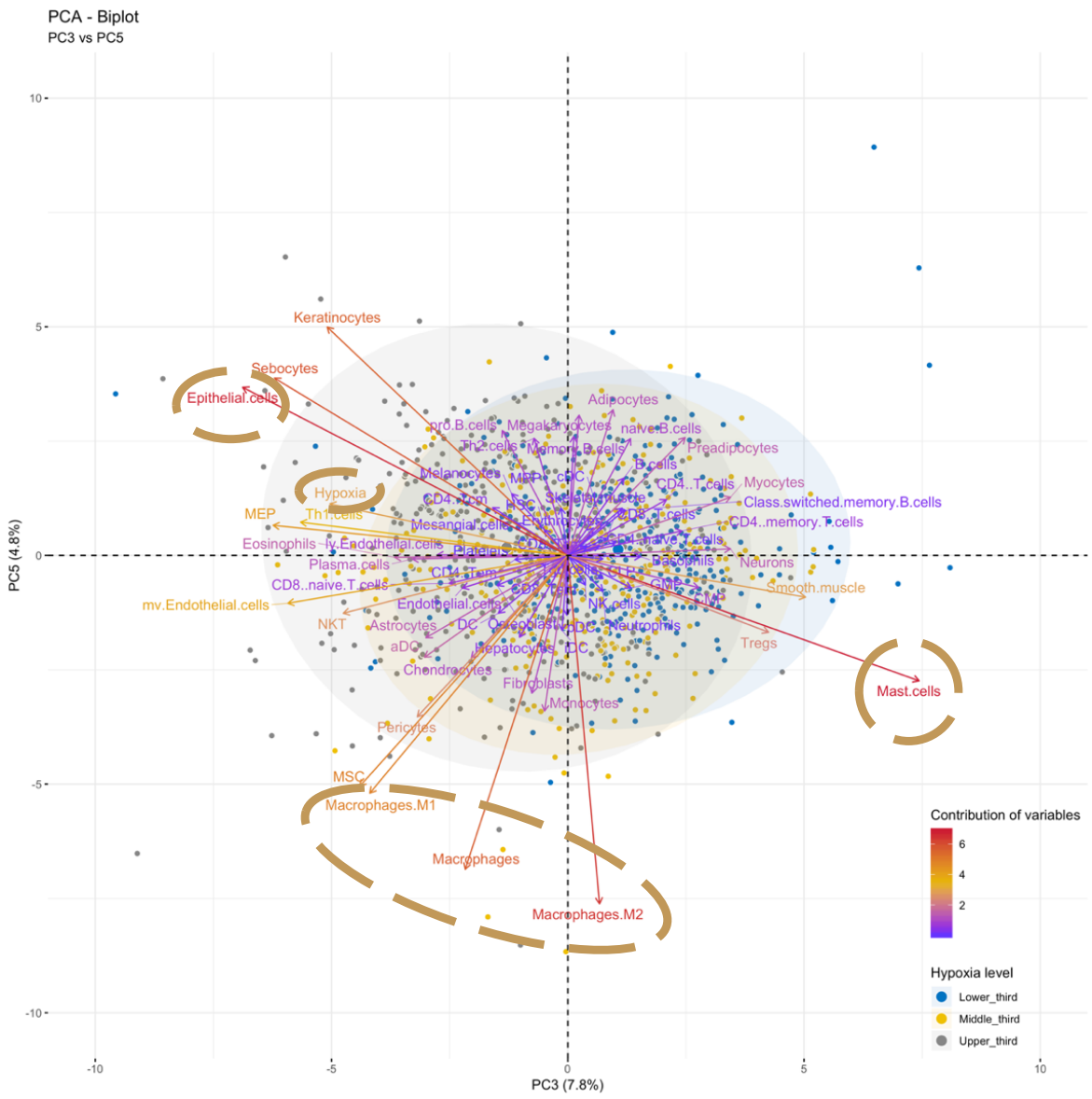


Figure 4.10 Association between mast cell and breast cancer hypoxia

The relationship between mast cells and hypoxia was demonstrated in an exploratory analysis using xCell signatures and the hypoxia signature. Biplot contains variables as cell types with overlaid hypoxia scores from individual samples summarised in three ellipses. The biplot indicated a strong reverse relationship between mast cells and hypoxia, which has been previously observed in mouse breast cancer models. Green dash cycles highlight hypoxia, epithelial cells, mast cells and macrophages. Hypoxia levels of samples are categorised into three groups in ellipses (upper third = high hypoxia, middle third = medium hypoxia, lower third = low hypoxia).

4.4 Pathway analysis of hypoxia associated genes in breast cancer

Having estimated the degree of immune and stomal infiltrations within the breast cancer tumour microenvironment using immunogenomic methods, I then wanted to explore the functions of these immune infiltrates with respect to tumour hypoxia. To achieve this, I applied functional genomic approaches to infer the activities of these cells from a bulk transcriptomic dataset.

4.4.1 Identify hypoxia associated genes in BRCA

To infer the functions of immune infiltrates within the hypoxic tumour microenvironment of breast cancer, I sought to identify a list of genes that are significantly associated with the hypoxia score (calculated using hypoxia signature), then to characterise the immune pathways connected with the list of genes. I used the breast cancer data from TCGA for the following analyses. The Buffa hypoxia signature (Buffa et al., 2010) was used to calculate hypoxia scores for each tumour sample. For consistency, I used the median as the hypoxia score. Spearman's rank correlation was chosen to calculate the correlation between the hypoxia score and gene expression. I did not use Pearson's correlation since it is sensitive to outliers (Mukaka, 2012). Two cut-offs (Spearman coefficient $> \pm 0.4$ and $p < 0.001$) were used to capture genes with moderate to high correlations which yielded 1060 genes positively associated with hypoxia and 1250 genes negatively associated with hypoxia. These gene lists were used for subsequent pathway analyses.

Table 4.1: Summary of selected gene lists for pathway analyses

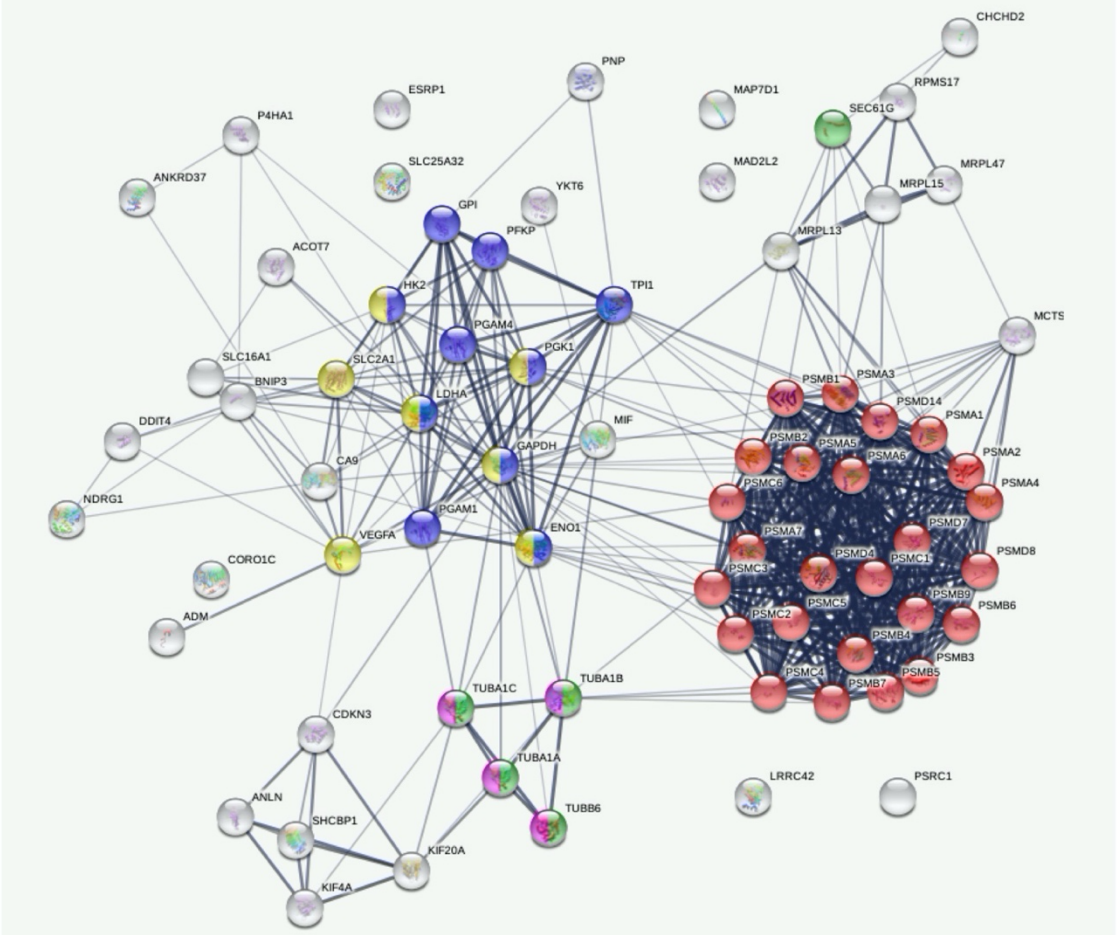
Cancer Type	*Number of Samples*	*Number of Genes selected*	*Correlation Type*
Breast Cancer (BRCA)	1093	1060	Spearman's - Positive
Breast Cancer (BRCA)	1093	1250	Spearman's - Negative

4.4.2 Protein-protein interaction and KEGG analyses of hypoxia signature

As a quality control step, I checked the pathways associated with the hypoxia signature and whether there is an established link with any immune pathways. The 47-gene hypoxia signature was used to calculate their likely protein interactors and related KEGG pathways. The gene list was uploaded to the STRING database, which helps to predict protein-protein interactions and the functions of the interactions (Szklarczyk et al., 2021). The network algorithm was given a limited flexibility to include additional 20 proteins with the highest likelihood for protein-protein interactions with the 47 genes in the hypoxia signature. The result (Figure 4.11) indicated that the hypoxia signature genes would most likely interact with proteins involved in the proteasome, glycolysis, and HIF-1 signalling. No immune-related pathways were identified using the genes in the hypoxia signature.

4.4.3 Immune GO Pathways in breast cancer

To characterise the functions of the 1060 hypoxia positively associated genes, I performed geneset enrichment analysis using immune-specific genesets from the Immune Gene Ontology (ImmGo) database (Bindea et al., 2009). The significant pathways were determined using hypergeometric Fisher's exact test (Benjamini-Hochberg adjusted $p<0.05$). The results indicated hypoxia is positively correlated with antigen processing and presentation via MHC class I; regulation of hematopoietic stem cell differentiation; stimulatory C-type lectin receptor signalling pathway; T cell receptor signalling pathway; positive regulation of innate immune response and susceptibility to natural killer cell-mediated cytotoxicity (Figure 4.12). My data suggest under hypoxia conditions, there is an increase in antigen presentation through the interactions between myeloid cells and T cells in TME. It has been previously demonstrated that T cells under persistent antigen presentation could attenuate their effector functions and proliferation within the TME (Etxeberria et al., 2020), which is similar to the inhibited T cell phenotype identified during chronic viral infection (Zajac et al., 1998). Therefore, my data is in accordance with prior findings that tumour hypoxia acts as a negative regulator of anti-tumour immunity through the suppression of T cell activities (Labiano et al., 2015; Rapisarda and Melillo, 2012).



KEGG Pathways

pathway	*description*	*count in gene set*	*false discovery rate*
hsa03050	Proteasome	25 of 43	2.22e-43
hsa00010	Glycolysis / Gluconeogenesis	10 of 68	9.70e-12
hsa01200	Carbon metabolism	9 of 116	2.13e-08
hsa05169	Epstein-Barr virus infection	10 of 194	7.37e-08
hsa01230	Biosynthesis of amino acids	7 of 72	2.55e-07
hsa04066	HIF-1 signaling pathway	7 of 98	1.54e-06
hsa05230	Central carbon metabolism in cancer	6 of 65	2.73e-06
hsa05130	Pathogenic Escherichia coli infection	4 of 53	0.00058
hsa04145	Phagosome	5 of 145	0.0021
hsa00051	Fructose and mannose metabolism	3 of 33	0.0025
hsa04540	Gap junction	4 of 87	0.0026

Figure 4.11 STRING predictions protein-protein with genes in the hypoxia signature.

To check whether genes in the hypoxia signature serve any immune-related functions, the 47 genes in the signature and an additional 20 genes with the highest likelihood for protein-protein interactions were uploaded to the STRING database, which predicts the functions of protein interactions. The result indicated that these genes most likely interact with proteins involved in the proteasome, glycolysis, and HIF signalling. No immune-related pathways were identified using these genes. In this network, the nodes represent proteins and the edges represent predicted functional associations.

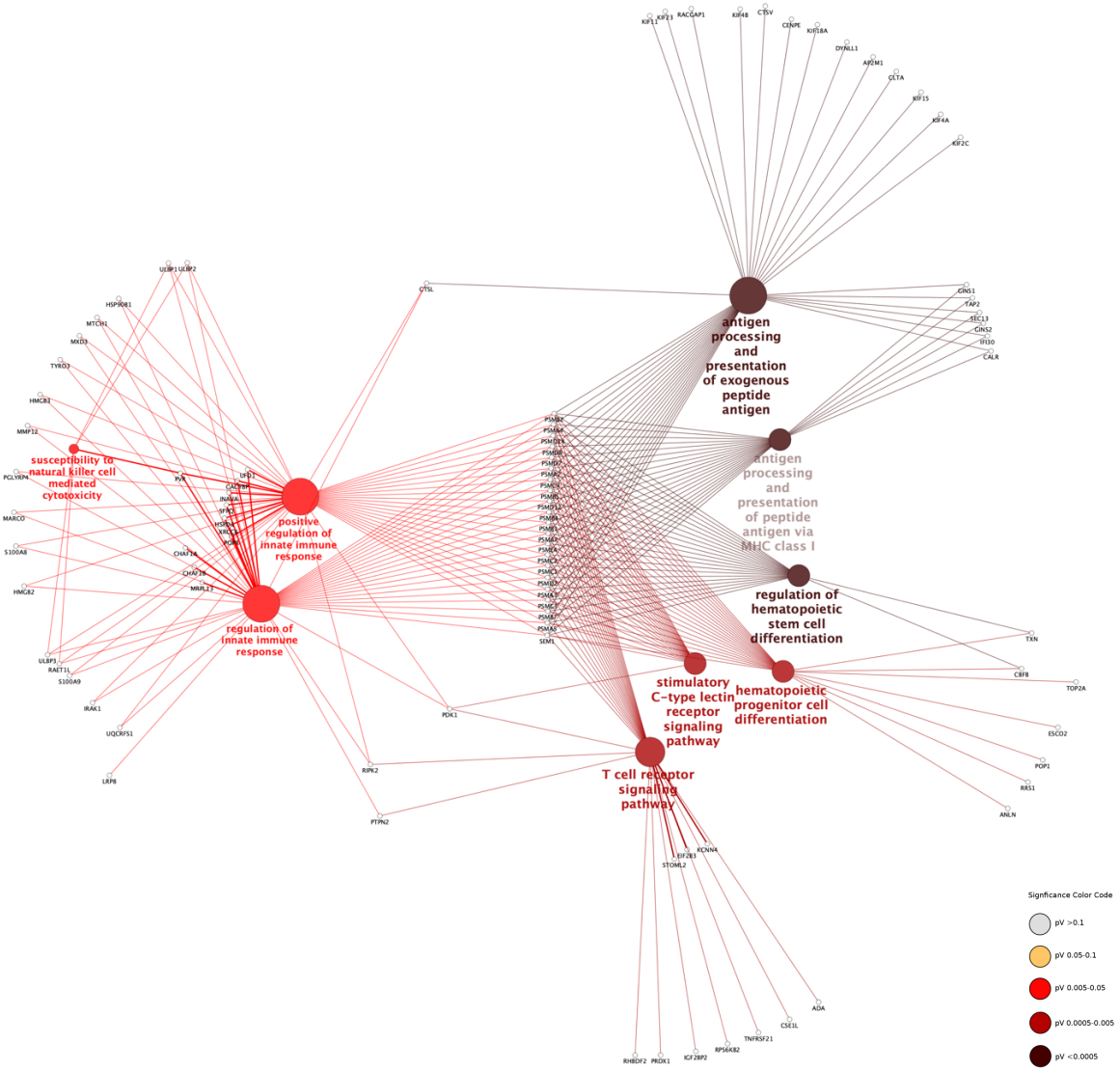


Figure 4.12 Hypoxia associated immune pathways in breast cancer.

The top 1060 genes associated with the hypoxia signature summary score were used to identify immune-specific pathways involved in the breast cancer hypoxic microenvironment. The functions of these genes were characterised using genesets from the Immune Gene Ontology (ImmGo) database. The results indicated that the hypoxic microenvironment featured upregulation of innate immune response, upregulation of antigen processing and presentation, and regulation of T cell receptor signalling. Previous reports have indicated that persistent antigen presentation to T cells in the TME by myeloid cells could lead to T cell exhaustion. In this graph, the nodes represent gene ontology terms and the edges represent genes associated with those terms. The size of the nodes reflects the statistical significance of the terms.

4.5 Conclusion

In this chapter, I characterised the association between tumour hypoxia and different immune and stromal populations in breast cancer by deploying immunogenomic tools to mine large bulk RNA-seq datasets (Hutter and Zenklusen, 2018). Using hypoxia transcriptomic gene signature as a proxy of tumour hypoxia, I identified a strong reverse relationship between hypoxia and endothelial cells, which have been previously reported that low vascular density and vascular abnormality were critical contributors to tumour hypoxia (Gaustad et al., 2017; Laitala and Erler, 2018). By stratifying breast cancer into subtypes, I confirmed an association between high hypoxia and more aggressive subtypes (i.e. TNBC/Basal-like and HER2), which corresponded with previous findings that linked high tumour hypoxia with tumour progression and poor survival (Abou Khouzam et al., 2021; Harris, 2002; Noman et al., 2015). Within these more aggressive subtypes (TNBC/Basal-like and HER2), I found an enrichment of macrophages (M1), which have previously been reported as critical suppressors of anti-tumour immunity (Kiss et al., 2018; Singhal et al., 2019). Finally, using a functional genomics approach, I teased out the key immune pathways involved in the breast cancer hypoxic tumour microenvironment. I found tumour hypoxia was positively associated with regulation of innate immune response, antigen processing and presentation, and regulation of T cell receptor pathway. Upregulation of these pathways corresponded with previous findings that persistent antigen presentations could attenuate effector T cell functions and lead to T cell exhaustion (Etxeberria et al., 2020; Zajac et al., 1998). Therefore, using computational methods I demonstrated that breast cancer hypoxic microenvironment enriched with infiltrating macrophages, which suppressed T cell-mediated anti-tumour immunity (Mantovani et al., 2017; Qiu et al., 2018). I will further investigate the role of tumour-associated macrophages in the subsequent chapter.

5 MACROPHAGE SIGNATURE DERIVED FROM SINGLE-CELL SEQUENCED TUMOUR ASSOCIATED MACROPHAGES

5.1 Introduction

Macrophages are important components of innate immunity and they play critical roles in the pathogenesis of major diseases. Within the context of cancer, it has been demonstrated that macrophages promote tumorigenesis by orchestrating angiogenesis, immunosuppression, invasion, and metastasis (Aras and Raza Zaidi, 2017; Mantovani et al., 2017). Not only these cells are functionally diverse, but macrophages also have a high abundance in solid tumours, in certain cases, they represent nearly 50% of the tumour mass (Qiu et al., 2018). High enrichment of macrophages in tumours correlated with high tumour grade and poor prognosis in cancer (De Palma and Lewis, 2013; Poh and Ernst, 2018; Qian and Pollard, 2010; Sica and Mantovani, 2012).

Much of the earlier studies quantified macrophages using singular immunostaining markers such as CD68, CD163, CD80 or a combination of them (Murray et al., 2014). With the proliferation of transcriptomic data available, more recent works focused on applying gene signatures to infer macrophage infiltration in tumours. These signatures were derived from different experimental models, technologies, tissue samples, and even species. Furthermore, macrophages are exquisitely sensitive to their environment and change phenotype readily to adapt (Xue et al., 2014). Despite this, signatures from one context are often used to infer information about data from another. No one has

systematically examined the relationship between these macrophage signatures, nor applied this comparison at single-cell resolution.

There are two strategies to derive human macrophage signatures: indirect differentiation from monocytes or direct isolation of macrophages from relevant tissues. Most of the published macrophage signatures were derived from *ex vitro* differentiated macrophages from peripheral monocytes using either colony-stimulating factor-1 (CSF-1) or granulocyte macrophage colony-stimulating factor (GM-CSF), they can be future activated by a variety of protocols for activation (Murray et al., 2014). The purified cells from cell culture were then processed for transcriptomics using RNA-Seq or microarray. The gene expression of individual macrophage gene expression profiles (GEP) can then be used to develop gene signatures using differential gene expression (DGE) or other algorithms. For example, two of the most comprehensive digital dissection tools, CIBERSORT and xCell both developed their macrophage signatures from the GEP of purified cells (Aran et al., 2017; Newman et al., 2015). The second strategy isolated macrophages directly from tissue or tumour samples to avoid *ex vivo* differentiation. The isolated cells can then be processed for either bulk or single-cell RNA-seq. The signatures from Azizi et al and Cassetta et al were developed using the second strategy from breast cancer samples (Azizi et al., 2018; Cassetta et al., 2019). There has been no head-to-head comparison of the signatures derived from the two different strategies. We, therefore, assembled 16 frequently cited macrophage signatures that were derived using different technologies, processed from different cell origins, and contained different polarization states.

The traditional theory of macrophage polarization is that mature macrophages can be divided into two different polarization profiles: pro-inflammatory (M1) and anti-inflammatory (M2) (Martinez and Gordon, 2014). However, much of the polarization theory was developed from observations from *in vitro* and *ex vivo* studies. Even before the emergence of single-cell technologies, it has been postulated that such simplified M1/M2 models may not resemble the true phenotype of tumour

associated macrophage (TAM) *in vivo* (Franklin et al., 2014; Martinez and Gordon, 2014; Perdiguero and Geissmann, 2014).

In this chapter, I demonstrated that the *ex vivo* derived macrophage signatures did not reflect the biology of tumour-associated macrophages and these signatures were unable to stratify the TAM populations in breast cancer. Moreover, I developed a new TAM signature from an existing breast cancer single-cell dataset using a more stringent clustering process and a rigorous way to derive the signature, which resulted in a TAM signature with an improved prognostic profile and helped to uncover the connection between TAM and other important factors within the tumour microenvironment.

5.2 Results

5.2.1 TAM and M1/M2 signatures share few genes in common

To understand the relationship between *ex vivo* derived macrophage signatures and tumour associated macrophage (TAM) signatures, I compared 16 published macrophage signatures from popular *in silico* dissection tools (Aran et al., 2017; Newman et al., 2015) and highly cited studies (Azizi et al., 2018; Bindea et al., 2013; Cassetta et al., 2019; Charoentong et al., 2017; Martinez et al., 2006; Xue et al., 2014). The signatures were generated using three different technologies to measure gene expression: microarray, RNA-Seq, and scRNA-Seq. Together the 16 signatures included 807 unique genes, with the largest signature containing 200 genes, and the shortest signature having 16 genes (Figure 5.1A).

I found no universally shared gene that can be used as a pan-macrophage marker, and I found little overlap among the signatures (Figure 5.1A). The most frequent genes across all macrophage signatures were: CCL8, TREM2, VSIG4, CHI3L1, and MARCO, each of them was shared by six different signatures out of the 16 signatures. Another four genes (ACP2, MS4A4A, SLAMF8, FDX1) were shared in at least five different signatures. These genes can be largely categorised into factors

that are secreted by macrophages (CCL8, CHI3L1) or receptors expressed by macrophages (TREM2, VISG4, MARCO, MS4A4A, SLAMF8). I were also surprised to find that the commonly used immunostaining macrophage markers (CD68, CD163, CD80) were not part of the partially shared genes across signatures. Therefore, the published macrophage signatures were divergent at the individual gene level and no shared pan-macrophage marker was identified.

I then compared the 16 human macrophage signatures based on their polarization states in four groups of four signatures: M0, M1, M2 and TAM (Figure 5.1B) In each of the M0, M1, and M2 groups, which were all *ex* vivo derived signatures, I could not find a common gene that encapsulated the dedicated polarization states. Besides common genes, each signature group also showed a high degree of divergence. For example, the M1 signature of Xue et al has 96 genes and only 8/96 (8%) genes were partially in common with the genes from two other M1 signatures of CIBERSORT and Martinez et al. There were no common genes between M1 signatures from Xue et al and xCell. The high divergence among signatures could reflect the different experimental conditions and variations involved in deriving macrophages during the *ex vivo* procedures.

Among the four breast cancer TAM signatures from Azizi et al and Cassetta et al, there was one overlapping gene, V-Set And Immunoglobulin Domain Containing 4 (VSIG4), a B7 family-related protein. It has been previously reported that VSIG4 suppressed T cell proliferation and cytokine production (Li et al., 2017b; Vogt et al., 2006). I observed a higher proportion of genes shared among the TAM signatures than the *ex vivo* derived signatures (Figure 5.1B). For example, Cassetta's 37-gene TAM signature shared 13/37 genes (35%) with two other TAM signatures. Azizi's TAM23 and TAM28 signatures (each contained 200 genes) have 67/200 genes in common (35%). The higher similarity among TAM signatures could be a result of commonality shared within the pathological breast cancer tumour microenvironment. Overall the TAM signatures appeared to be more similar to each other, but very different from *ex vivo* derived macrophage signatures, which echoed previous concerns that the theoretical M1/M2 model did not reflect the macrophage biology in diseases.

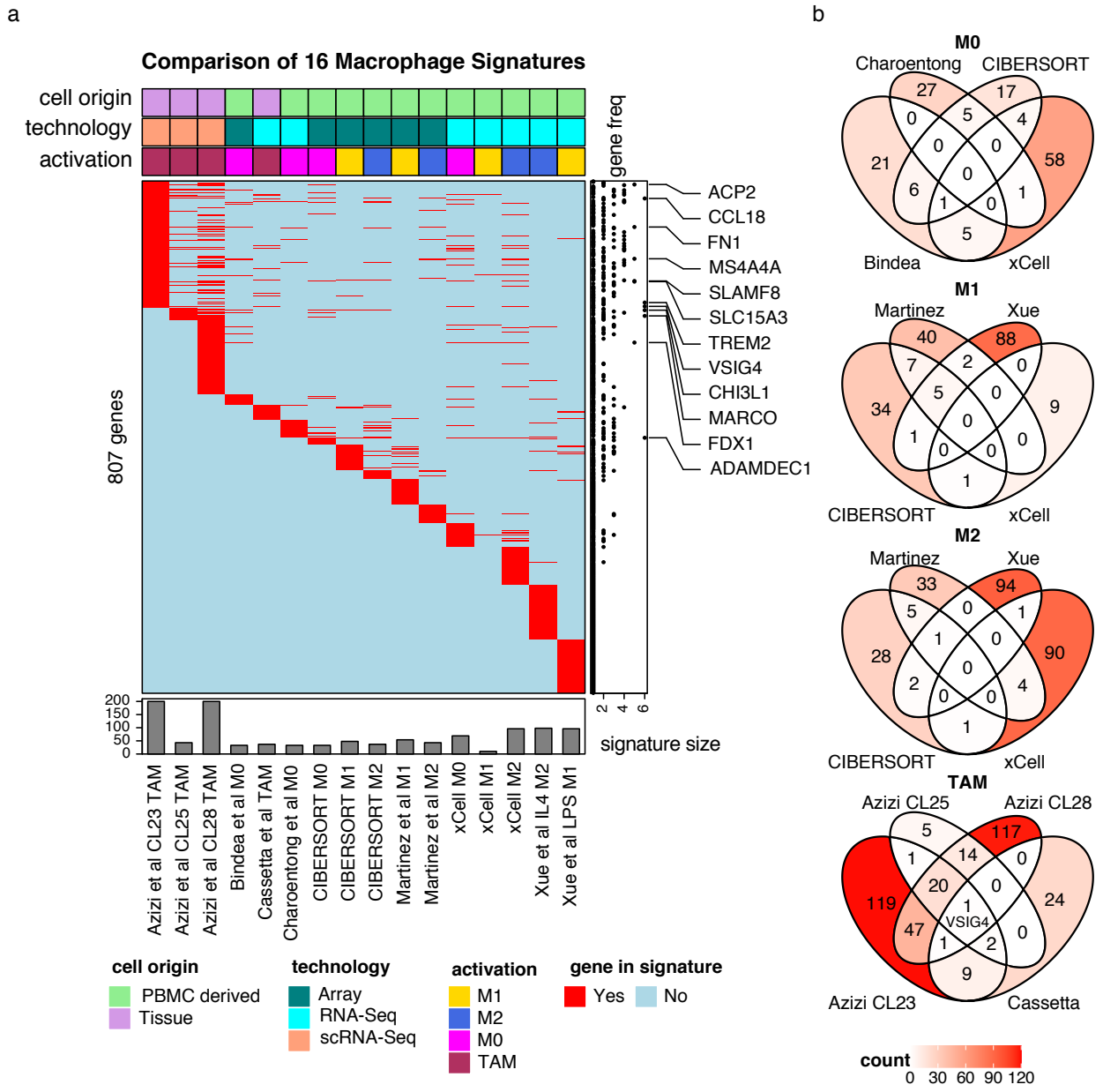


Figure 5.1 Comparison of 16 publicly available macrophage signatures.

The signatures were derived in two ways: *ex vivo* PBMC derived and tumour associate macrophage directly isolation from *in vivo* tissues. A) A gene-by-gene comparison of the 16 published signatures showed a high degree of divergence at the individual gene level. No universal macrophage markers were found. The genes frequently appeared across different signatures are CCL18, TERM2, VSIG4, CHI3L1 and ADAMDEC1. The size of the signatures has a big range from 13 to 200. B) The signatures were grouped using their polarization states into four groups. Even within each group, the

signatures had very little agreement, especially for the *ex vivo* derived M0, M1, and M2 groups. In the TAM group, only one gene, VSIG4, was shared across the signatures.

5.2.2 Published macrophage signatures do not support M1/M2 dichotomy in breast cancer

After discovering the divergence in TAM and *ex vivo* derived macrophage signatures at the individual gene level, I wanted to assess the performance of the 16 signatures in transcriptomic datasets as these signatures were frequently used at both bulk and single-cell levels, I also asked if the gene signatures reflected different macrophage populations within the tumour microenvironment.

Using an established protocol to access gene signatures (Dhawan et al., 2019), I used the 16 curated signatures to calculate macrophage scores in 1093 breast cancer bulk RNA sequencing samples from The Cancer Genome Atlas (TCGA) and 2104 TAM single-cell RNA sequencings samples from GSE114725 (Azizi et al., 2018). Thereafter, I performed Pearson's correlation analyses on the macrophage signature scores to understand the relationships among the signatures, and I also clustered signatures scores to understand the difference between *ex vivo* derived macrophage signatures and *in vivo* derived TAM signatures.

In TCGA bulk breast cancer samples, Pearson's correlation analyses revealed several insights into these signatures (Figure 5.2A). First, I observed positive correlations among the majority of macrophage signatures (14/16 signatures). Noticeably, I found the three TAM signatures (Azizi CL23, Azizi CL25, Cassetta et al) correlated with five *ex vivo* derived macrophage signatures (CIBERSORT M0, CIBERSORT M2, Bindea M0, Charoentong M0, Martinez M2). Second, all four of the *ex vivo* M1 signatures showed positive correlations among themselves. In contrast, the group of four M2 signatures had negative correlations. Lastly, I noticed the M1 and M2 signatures from the two digital dissection tools, CIBERSORT and xCell, all showed positive correlations. Since the bulk RNA sequence contained a mix of signal from a variety of cells, I cannot conclusively establish whether the correlations were have resulted from different underlying macrophage biology or

different performance from the signatures. Therefore, I repeated the signature analyses in a single-cell TAM dataset.

At the single-cell level, nearly all macrophage signatures were positively correlated (15/16 signatures) except the Azizi TAM25 signature (Figure 5.2B). Interestingly, I did not observe a polarized relationship between the *ex vivo* derived M1 and M2 signatures. Therefore, this finding suggested the *ex vivo* M1/M2 signatures would be unlikely to distinguish theoretically distinct cell populations in TAM samples.

To understand the performance of the macrophage signatures at the individual sample level, I performed a hierarchical clustering analysis of the macrophage scores. In TCGA bulk RNA-seq breast cancer data, the 16 signatures appeared to only stratify the samples into general high or low macrophage expression levels (Figure 5.2C), instead of the different macrophage polarization groups derived from *ex vivo*. Based on the dendrogram, the M1 signatures of Xue, CIBERSORT, Martinez and xCell showed a close relationship, but that was not enough to distinguish them from the rest of the signatures in the TCGA breast data. Similar observations were made in the single-cell data: samples were separated into high or low macrophage gene expression levels (Figure 5.2D). Indeed, contrary to the M1/M2 dichotomy hypothesis, no M1 or M2 specific clustering was detected in either of the bulk or single-cell samples.

Furthermore, I checked the expression levels of individual genes in the macrophage signatures in bulk and single-cell datasets. I found all 16 macrophage signatures have a consistent expression at individual gene levels (Figure 5.3). However, I found genes in the 16 macrophage signatures have low or no expression in the single-cell data (Figure 5.4). Whilst this reflected the well-known sparse nature of scRNA-seq data (Hou et al., 2020; Matuła et al., 2020), such observations raised questions on the reliability of applying these signatures to single-cell data as the summary signature score could be highly biased towards a small number of expressed genes. Even stable scores such as the median value of the signature genes, would not be applicable in single-cell samples with high dropouts.

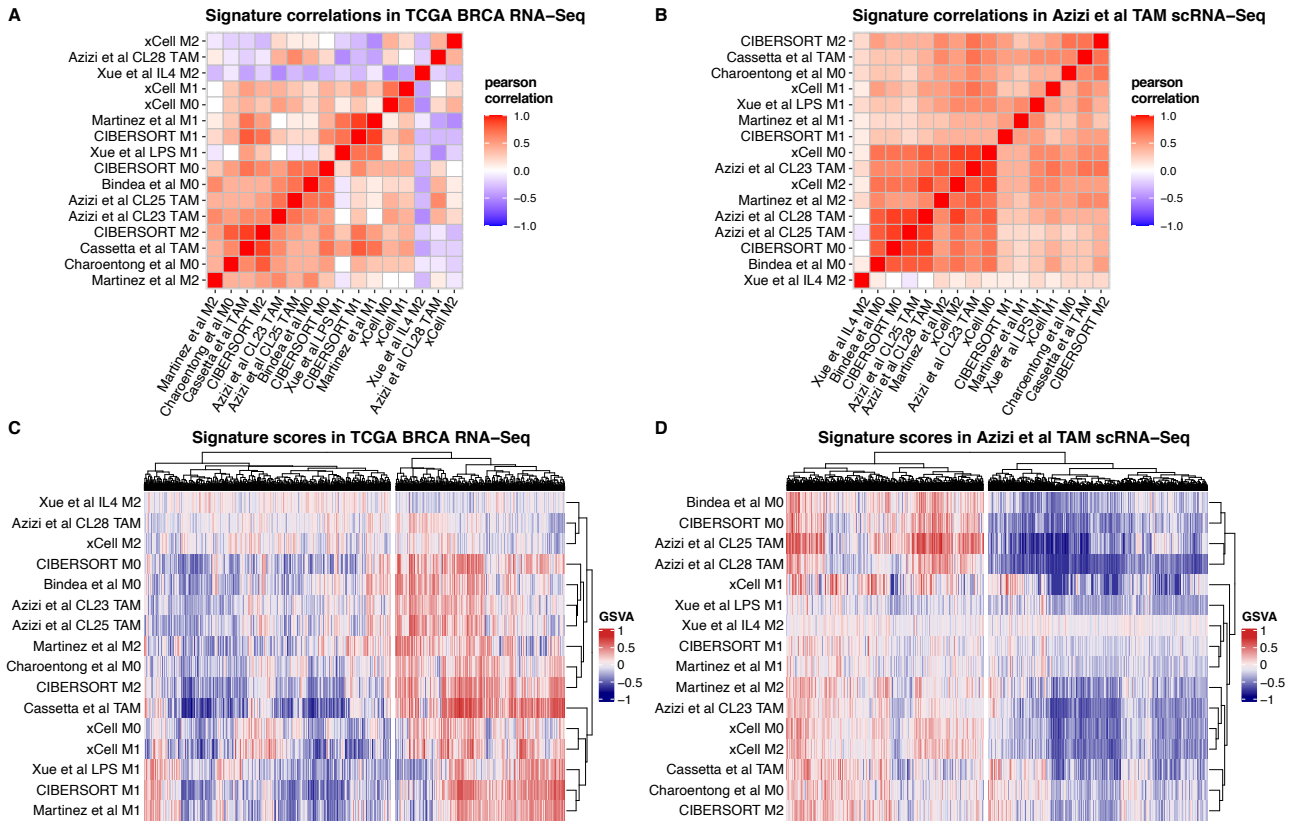


Figure 5.2 Comparing 16 macrophage signatures in bulk and single-cell RNA-seq data.

A) In TCGA breast cancer bulk RNA-Seq samples, high positive correlations were observed among the majority of the signatures including three of the *in vivo* derived TAM and several *ex vivo* derived M0, M1, and M2 signatures. B) In single-cell TAM RNA-Seq samples, nearly all the macrophage signatures showed significant positive correlations, except for *ex vivo* derived Xue M2 signature. C) Clustering of macrophage signature scores in TCGA breast cancer samples, which indicated the *ex vivo* derived signatures were not able to stratify samples into different polarization states. It appeared that these collective 16 signatures were only able to stratify bulk samples into either high or low macrophage expressing samples. D) Clustering of macrophage signatures scores in single-cell TAM RNA-Seq samples followed a similar observation with the bulk samples, that the curated 16 signatures were not able to identify unique clusters of macrophages.

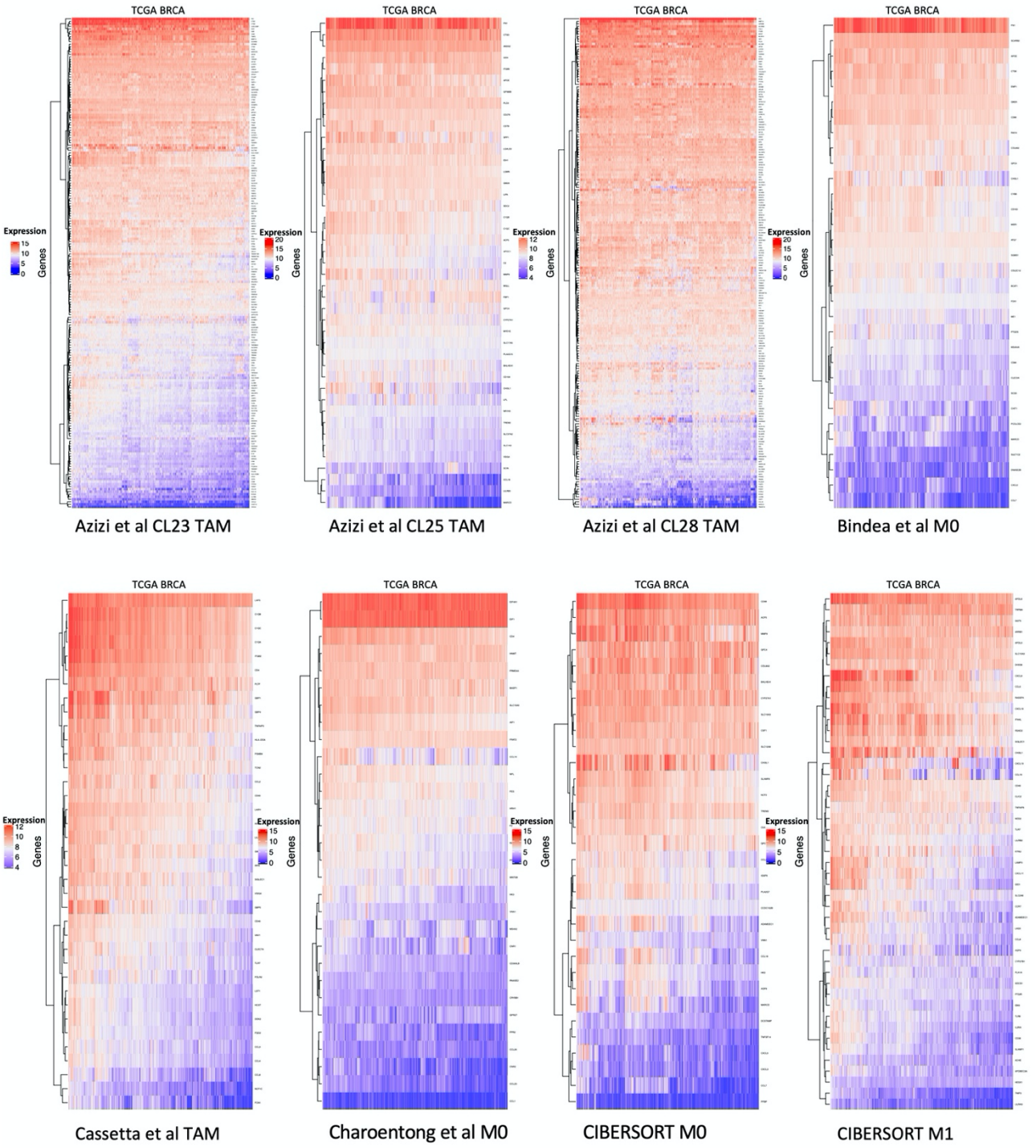


Figure 5.3 Macrophage signature expression in bulk RNA-seq data.

The expression of 16 macrophage signatures was examined in TCGA breast cancer data at the individual gene level. The results indicated that macrophage signatures in bulk transcriptomic data have a range of expression, which reflected the variability in gene signature expression across samples. y-axis: individual genes of each signature. x-axis: individual samples.

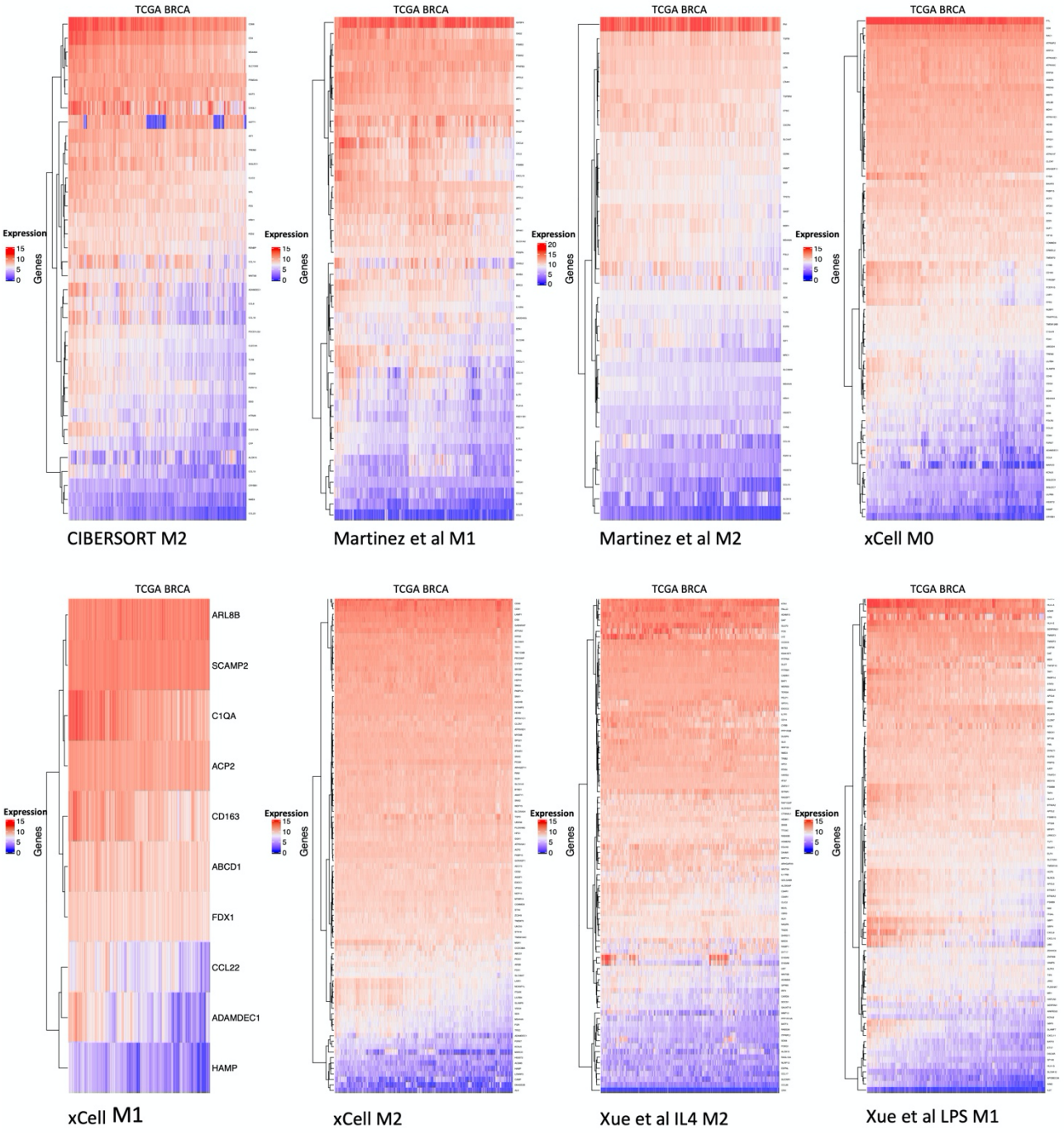


Figure 5.3 (Continue) Macrophage signature expression in bulk RNA-seq data.

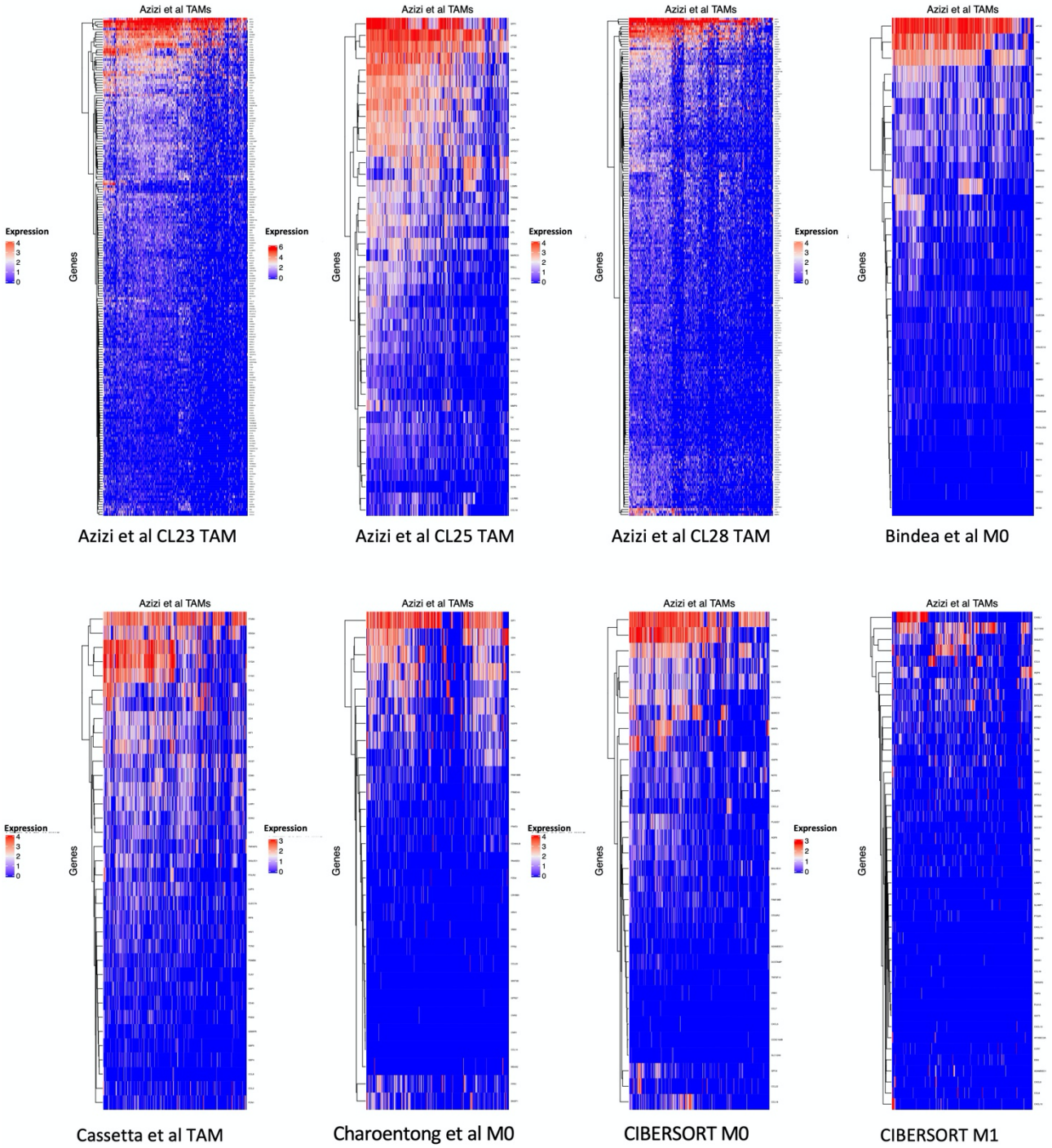


Figure 5.4 Macrophage signature expression in single-cell RNA-seq data.

The expression levels of macrophage signatures were examined at the individual gene level in a single-cell TAM dataset (GSE114725). The expression patterns displaced high sparsity, which reflected the high drop-out nature of single-cell transcriptomic data. Such high sparsity could affect the reliability of applying macrophage signatures to single-cell data. y-axis: individual genes of each signature. x-axis: individual samples.

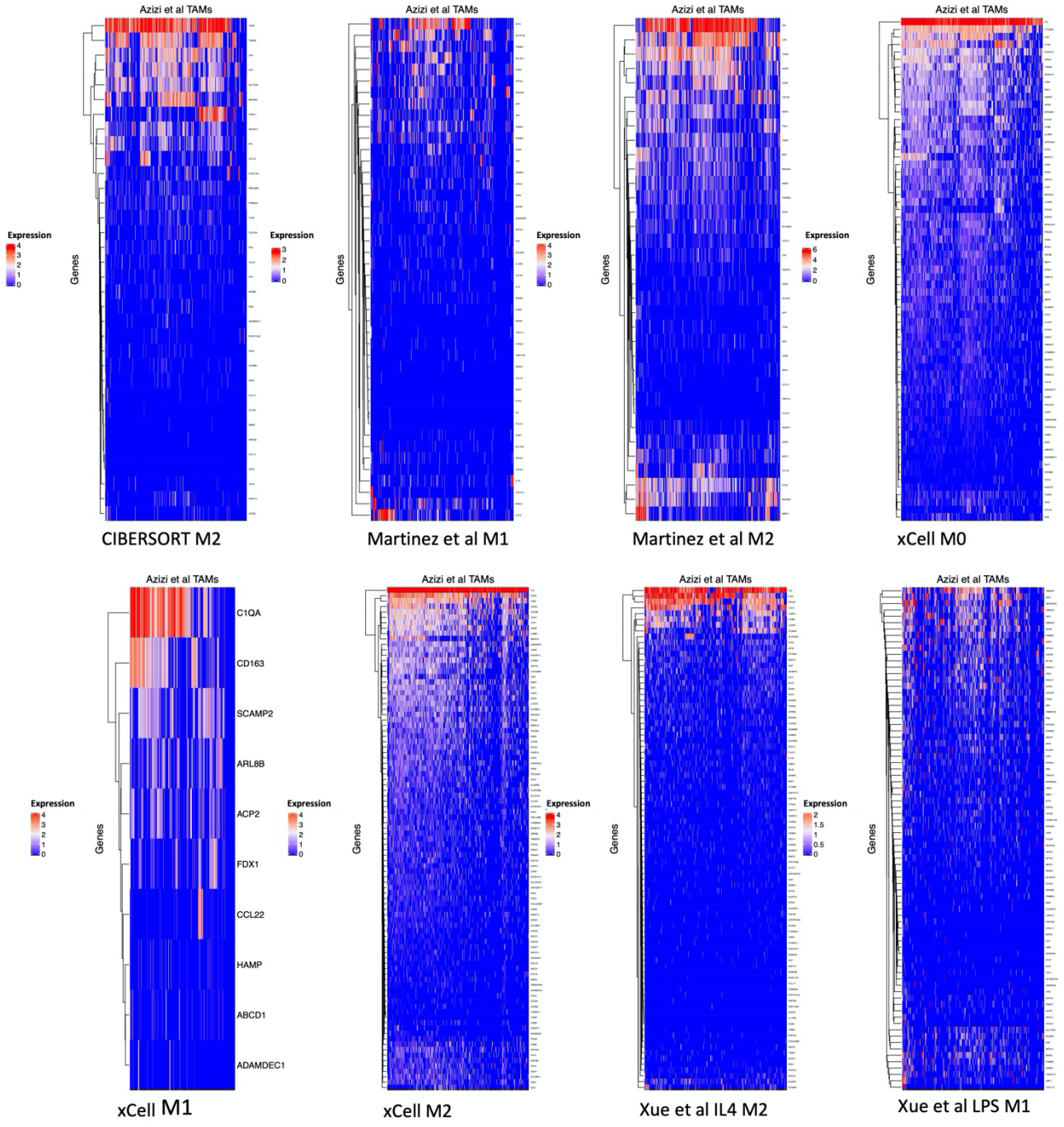


Figure 5.4 (Continue) Macrophage signature expression in single-cell RNA-seq data.

5.2.3 Tumour associated macrophage and monocyte diversity in breast cancer

Recent single-cell studies reported a high degree of diversity of monocytes and macrophage within the tumour microenvironment, but my signature analysis on both TCGA and single-cell macrophage levels were not to able reflect such diversity. Given the challenges often associated with single-cell unsupervised clustering analysis (Kiselev et al., 2019), I decided to re-analyse a subset of 4597 cells from GSE114725, which were previously classified as 17 unique clusters of monocytes and macrophages, using a more rigorous clustering algorithm (More details are available in the method chapter). I performed k-means consensus clustering with 2000-times resampling on these cells, which also estimated cluster stability (Wilkerson and Hayes, 2010). Based on the delta area plot, which estimates the relative change with each increase in cluster count, I determined that the optimal number of clusters of the data was 5 (k=5, Figure 5.5A). When attempting to identify mutually exclusive markers from the k=5 cluster, one of the clusters did not yield any gene, but I were able to determine specific markers when the cluster number was set to 4. Thus, to allow meaningful biological interpretation, I chose k=4 as the optimal cluster number for the monocyte and macrophage subsets (Figure 5.5B), which was considerably less than the previously reported 17 different monocytes and macrophages within the breast cancer landscape (Azizi et al., 2018).

The clusters were annotated using six previously known markers. Clusters 1 and 2 were characterized as monocytes using CD14, FCGR3A/CD16A, and CSF1R. Clusters 3 and 4 were annotated as macrophages using the monocyte markers and their higher expression of CD163, CD68, and CCR5 (Figure 5.5C). Further, I confirmed the classification using the tissue origin of these cells (Figure 5.5D). The tissue data indicated that clusters 1 and 2 were found in blood, lymph nodes, paired normal, and tumours. Having found both clusters 1 and 2 in blood eliminated the possibility that one of them could be macrophages. In contrast, clusters 3 and 4 were exclusively identified in tumours, which matched with the biology of macrophages originating from blood monocytes and differentiated in tissues. I also modelled the developmental trajectory of the four clusters (Figure 5.5E), which indicated that macrophage clusters 3 and 4 were differentiated from monocyte clusters 1 and 2 in a

continuous linear fashion (Street et al., 2018). The distance between clusters 2 and 3 was the widest in the graph, which likely reflected the significant transcriptional changes involved in switching from monocytes to macrophages.

Next, I asked how the clusters generated using consensus clustering compare with two other methodologically distinctive clustering approaches, namely Biscuit and Seurat. Biscuit produced single-cell clusters using a hierarchical Bayesian mixture model and it was used by the original publishers' to generate the 17 different monocyte and macrophage clusters (Azizi et al., 2018; Prabhakaran et al., 2016). Seurat is based on a graph-based clustering approach (Butler et al., 2018). Both methods did not apply resampling during the clustering step. I summarised the cluster relationships in a cluster tree model (Figure 5.5F). The default parameters of Seurat yielded 13 clusters, which would also suggest a diverse monocytic sub-population within the breast cancer microenvironment. However, upon inspection of the result from the consensus clustering delta area plot (Figure 5.5A), even with the most lenient interpretation, the relative increase in area under the curve approached 0 when k equalled 6, which suggested the possible cluster number within the dataset was equal or less than 6. This suggested that Seurat could generate clusters beyond the signal within the single-cell data structure, which has been argued as a limitation of Seurat in other studies (Andrews and Hemberg, 2018; Grabski et al., 2022; Kiselev et al., 2019). Whilst the slight difference in methodology and criteria used could show different numbers of optimal clusters of monocytes and macrophages, I found 2 monocyte and 2 macrophage clusters, together 4 clusters. this number was significantly fewer than the previously reported 17 in breast cancer (Azizi et al., 2018).

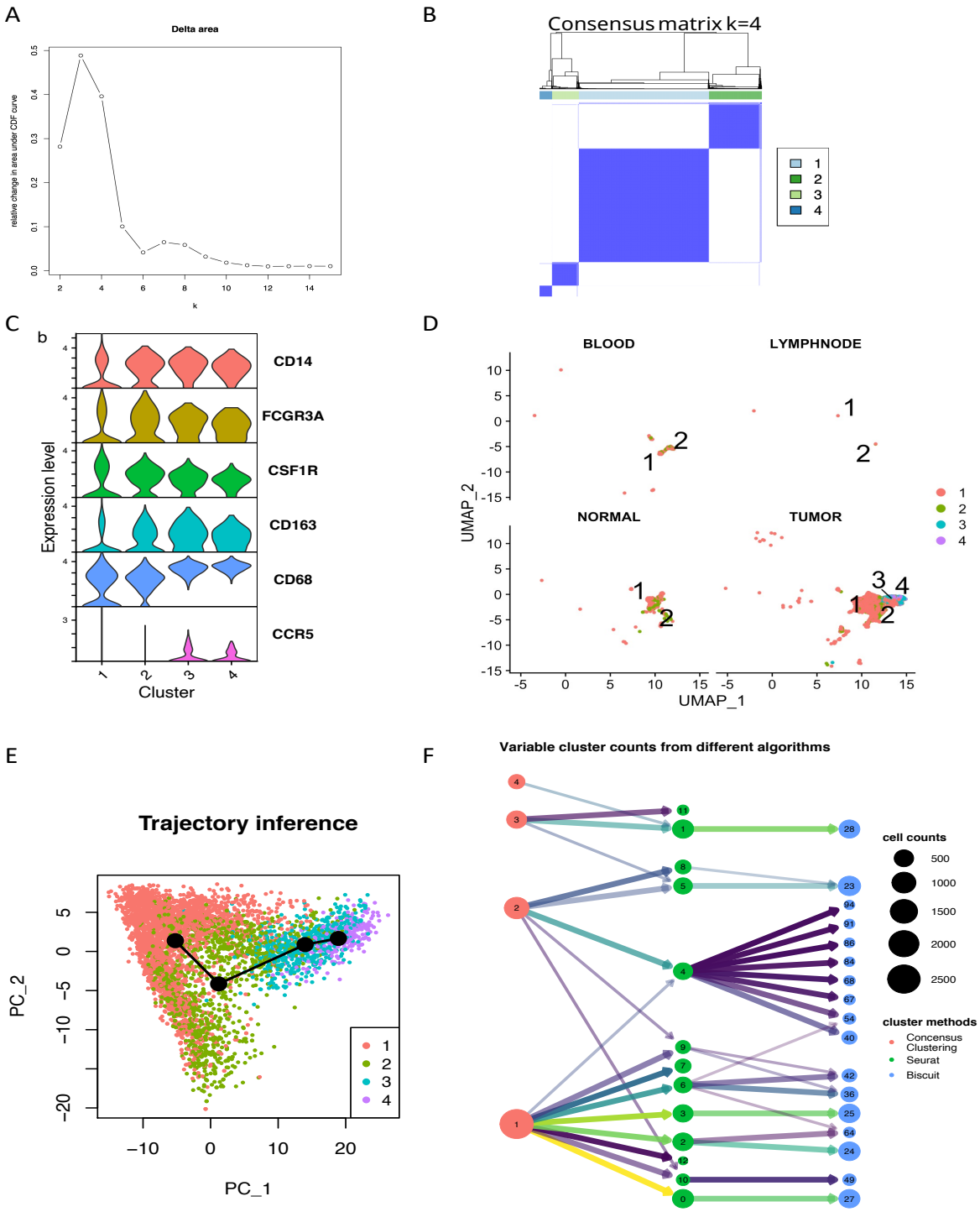


Figure 5.5 Establishing macrophage clusters with consensus clustering method.

I reanalysed 4597 previously classified monocyte and macrophage cells from GSE114725, which were previously annotated to 17 different unique clusters. A) Delta area plot of consensus clustering

indicating the optimum number of clusters in the dataset was around 5, but at k=5 no unique markers could be identified for a cluster. For meaningful biological interpretation, k=4 was chosen. B) Consensus matrix when k equalled to 4, which showed very few cluster reshuffling after 2000-times of resampling. The cluster reshuffling was marked as the lines outside of the main four cluster squares in the figure. C) Top markers used to differentiate monocyte and macrophage clusters. D) Tissue distributions of the four clusters showing clusters 1 and 2 (monocytes) can be found in all tissues including blood, lymph node, normal-pair samples, and tumours. Clusters 3 and 4 (macrophages) were only found in tumours. E) Developmental trajectory of the four clusters, which suggested the relationship of the four clusters following a linear relationship and there is a significant amount of transcriptomic remodelling required to jump from monocyte (clusters 1 & 2) to macrophage (clusters 3 & 4). F) Comparison of the outputs from different single-cell clustering algorithms on the same dataset, which demonstrated flexibility in the clustering algorithms when no resampling (Seurat and Biscuit) was taken into consideration.

5.2.4 Alternatively activated macrophages in breast cancer favoured lipid metabolism

Given the rapid evolvements in single-cell technologies and the observed differences in macrophage signatures (Figure 5.1), I then investigated whether the macrophage clusters identified in an older dataset can be found in newer breast cancer datasets, which would help to understand the landscape of macrophage diversity across different studies. I stratified 4606 myeloid cells from a published triple-negative breast cancer scRNA-Seq of five tumours (Wu et al., 2020). I first performed canonical correlation analysis (CCA) to integrate Azizi and Wu's datasets, which also corrected the batch effect (Stuart et al., 2019). Examination of the t-SNE plots before and after integration (Figure 5.6A&B) and the first two PCs revealed the datasets were successfully integrated, regardless of the sequencing platform (Figure 5.6C&D). Cluster labels were transferred by projecting the PCA structure from Aziz to Wu's dataset. Based on similar expressions of top canonical markers (CD14, FCGR3A, CSF1R, CD163, CD68, CCR5), I confirmed that all four clusters were present in the TNBC dataset (Figure 5.6E). Whilst the clusters were identified across datasets, the proportions of clusters were different (Figure 5.6F). This difference in cell proportions likely reflected the heterogeneity of the tumour microenvironment across different breast cancer subtypes and samples.

After confirming the same macrophage clusters can be found in different breast cancer datasets, I investigated the cell functions of the two macrophage clusters through differential gene expression analysis (DGE). The two datasets were produced using different scRNA-Seq technologies: inDrop and 10x Chromium. In a recent review, 13 common scRNA-Seq technologies were compared. The result indicated both inDrop and 10x Chromium have high clusterability (Mereu et al., 2020). However, inDrop scored poorly in both gene detection and marker expression categories, which are important for DGE analysis. I decided to perform DGE analysis on the 10x Chromium data with the cluster labels transferred from the inDrop dataset.

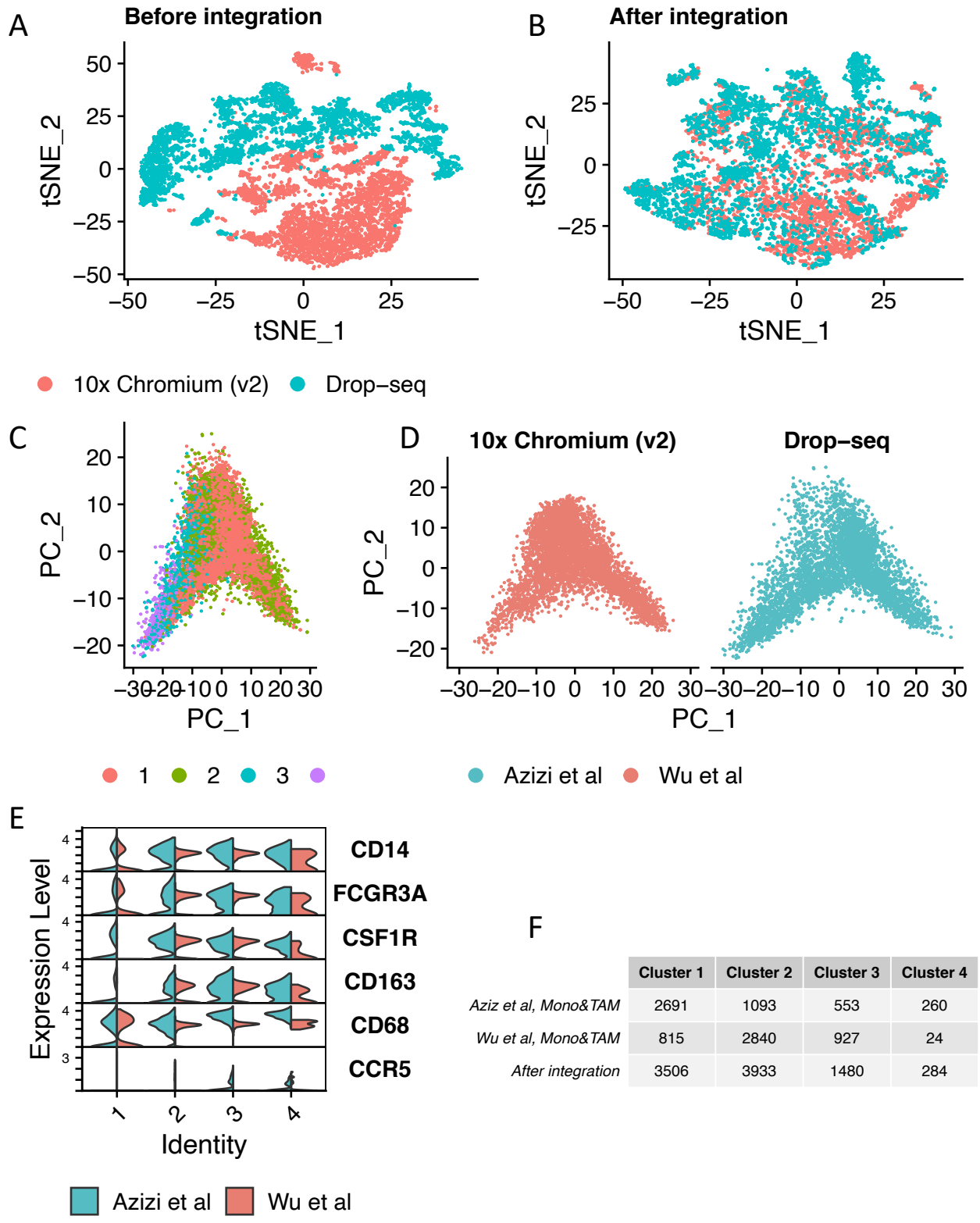
The top two differentially expressed genes from both TAM clusters have been previously reported to serve important roles in cancer progression. TAM1 expressed high F13A1 and TAM2 expressed high CHI3L1 (Figure 5.7A). For example, Factor XIIIA (F13A1) crosslinks fibrin and has roles in clot

stabilization and wound healing (Porrello et al., 2018). A high expression level of F13A1 was correlated with worse outcomes in lung squamous cell carcinoma (LUSC). In addition, F13A1 was also demonstrated as a marker for infiltrating pro-inflammatory monocytes within LUSC (Porrello et al., 2018). Chitinase-3-like protein 1 (CHI3L1) or better known as YKL-40, a growth factor has previously been shown as a serum marker for reduced survival rate in breast cancer (Johansen et al., 2003). CHI3L1 was important for tumour remodelling through VEGF-dependent and independent pathways. Thus, based on the top differentially expressed genes, the two TAMs appeared to have different phenotypes: TAM1 shifting towards pro-inflammatory and TAM2 skewing towards microenvironmental remodelling.

The top differentially expressed genes from each cluster were then used in KEGG enrichment analysis to determine their associated pathways (Yu et al., 2012). The KEGG analysis highlighted that TAM1 was involved in viral protein interaction, IL-17, NFkB, TLR, TNF and MAPK signal pathways and TAM2 was involved in PPAR, lysosome, cholesterol metabolism and ECM-receptor interactions (Figure 5.7B). The results supported the hypothesis that different TAMs serve distinct functions within the tumour microenvironment (Mantovani et al., 2017).

I also identified the core pathways associated with TAMs using the differential gene expression and gene ontology analysis (Figure 5.7C&D). TAM1 has high chemokine activities (CCL2, CCL4, CCL8, CXCL2, CCL13) and genes represented the classically activated pro-inflammatory macrophages (IL1B, CSF1R, TLR7, TNFRSF1A). In comparison, TAM2 was enriched with cathepsins (CTSS, CTSK), which are proteases enclosed in the lysosomes. TAM2 also exhibited genes in extracellular matrix (ECM) interaction (SPP1, FN1, ITGA3, COL6A1), cholesterol metabolism (APOC1, LPL, LIPA), and PPAR signalling pathways (DBI, FABP3, FABP4, FABP5). The enrichment analysis indicated the two TAMs have different functions, with TAM1 matched with the description of a more pro-inflammatory macrophage (Murray et al., 2014) and TAM2 resembling lipid-associated macrophages found within plaques (Eijk and Aerts, 2021). Therefore, I called TAM clusters

according to their key functions: TAM1 to Pro-Inflammatory TAM (PI-TAM) and TAM2 to Lipid-Associated TAM (LA-TAM), respectively.



F

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Aziz et al, Mono&TAM	2691	1093	553	260
Wu et al, Mono&TAM	815	2840	927	24
After integration	3506	3933	1480	284

Figure 5.6 Integrating and mapping the newly derived four monocyte and macrophage clusters in another breast cancer single-cell dataset.

Monocytic cells from two single-cell datasets (Azizi et al., 2018; Wu et al., 2020) were integrated using Seurat . The performance of the integration was evaluated using using two different dimensional reduction methods: t-Distributed Stochastic Neighbour Embedding (t-SNE) and Principal Component Analysis (PCA). A) The distribution of the cells in the two datasets before integration was visualized using t-SNE, with Azizi et al. in green and Wu et al. in red. B) The distribution of the cells after integrating the datasets was visualized using t-SNE. C) A PCA plot of the monocytic cells from two single-cell datasets before integration was created. Mapping clusters across datasets required integration of the different datasets using the principal components as anchors (Butler et al., 2018). D) The integrated dataset was compared using the two different single-cell technologies, 10x Chromium and inDrop, in PC1 and PC2. E) The expression of top macrophage canonical markers in newly identified clusters were compared across the two datasets. F) The counts of each cluster in the integrated dataset and in the individual datasets were recorded.

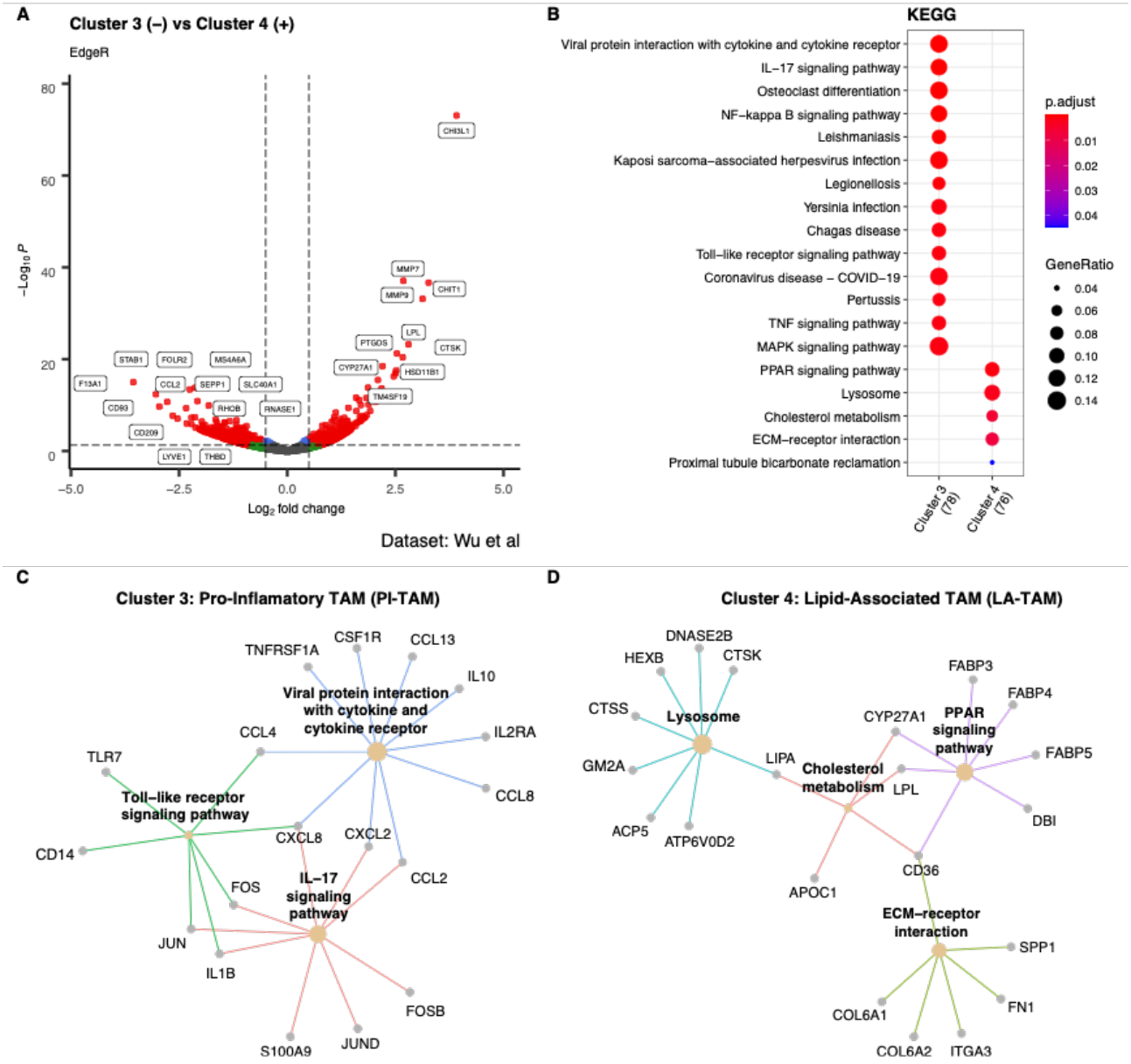


Figure 5.7 Functional characterization of the two TAMs.

A) Differential gene expression from the two TAMs. TAM1 featured a pro-inflammatory marker F13A1 and TAM2 featured a microenvironmental remodelling marker CHI3L1. B) Top KEGG functional pathways associated with TAM1 and TAM2. C & D) Core KEGG pathways shared genes from TAM1 and TAM2. TAM1 resembled classical pro-inflammatory macrophages (IL1B, CSF1R, TLR7, TNFRSF1A). TAM2 appeared to be a lipid-associated macrophage featuring upregulated lysosome and extracellular matrix interactions. Note: (Cluster 3 = TAM1. Cluster 4 = TAM2).

5.2.5 Lipid-associated macrophage spatially concentrated within the invasive front

To better understand the role of the pro-inflammatory and lipid-associated macrophages within the tumour microenvironment (TME), I examined the spatial distribution PI-TAM and LA-TAM clusters using spatial transcriptomics in a sample with ductal breast carcinoma in situ (DCIS) and invasive carcinoma (Data was released by 10x Genomics as a public dataset). Locations of the TAMs were predicted using a deconvolution-based method (Stuart et al., 2019), which used single-cell profiles as classifiers and predict the probabilistic presence of a given cell profile in each of the 50um dot areas on the histology slide (10x Visium).

The spatial transcriptomic slide was clustered into nine areas based on the spatial variations of RNA expression (Figure 5.8A). The relationships of the distinctive areas based on gene expressions were visualized in UMAP space (Figure 5.8B). I then selected area one for more detailed analysis due to its clearly defined immune infiltrations at the histology level. In area one, PI-TAM were distributed in the periphery of the DCIS and along the border between the tumour and stromal areas (Figure 5.8C). In contrast, LA-TAM were mostly concentrated in the micro-invasive front of the tumour and intermingled with cancer cells without a clear boundary or structure (Figure 5.8D). The results indicated the functional distinct PI-TAM and LA-TAM also occupied unique spatial locations within tumours: the outer boundary and within the micro-invasive fronts, respectively.

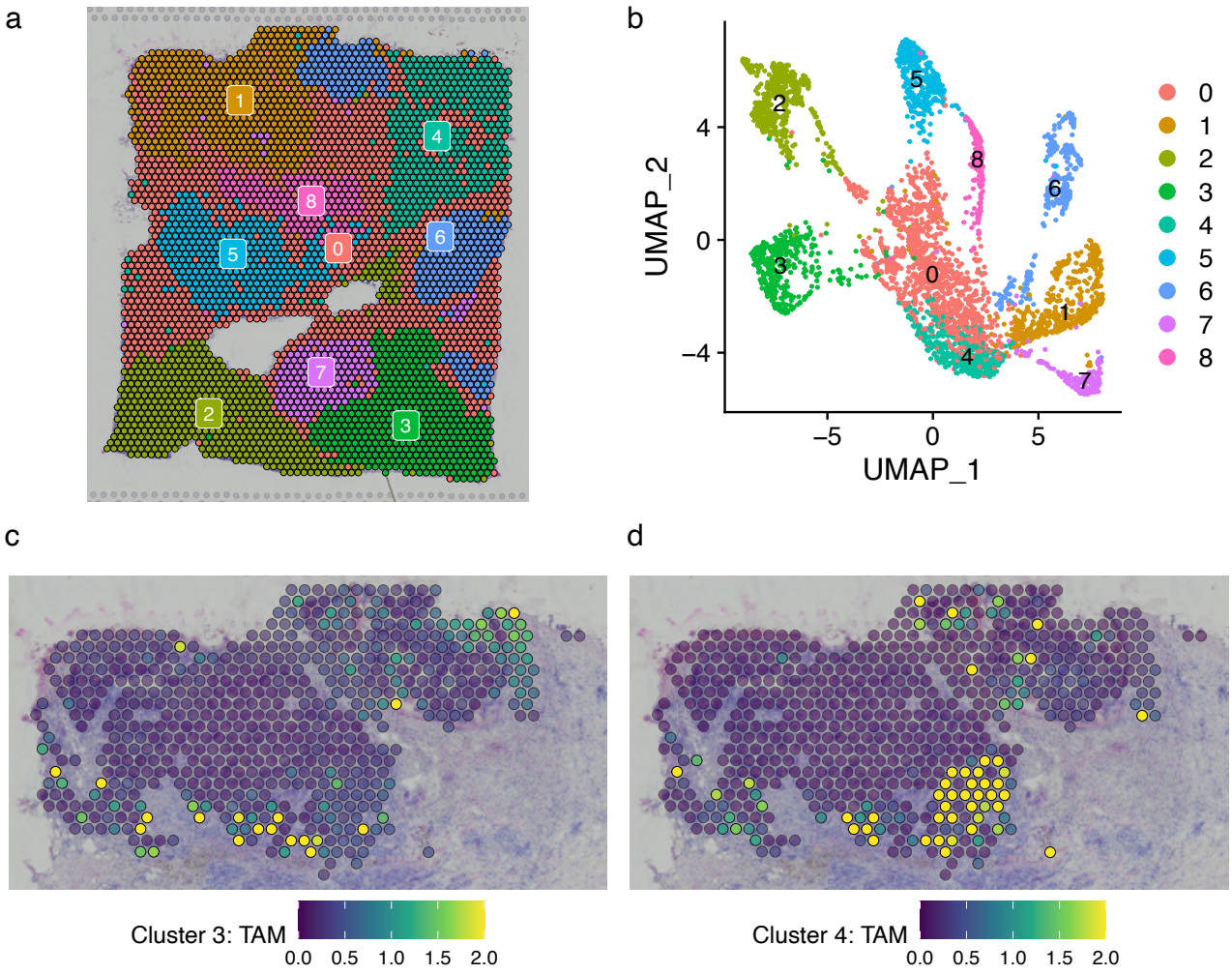


Figure 5.8 Spatial distributions of the two TAMs.

A) A DCIS breast cancer sample was divided into 9 areas based on the shared similarity of gene expression patterns from each of the 50nm areas. B) The gene expression of each area was visualised using UMAP, which showcased the relationships among the groups. C) TAM1/PI-TAM distribution within area 1, which showed a pattern of spatial distribution along the outer boundary of the DCIS. D) TAM2/LA-TAM was located mostly in the micro-invasive fronts of the tumour and intermingled with cancer cells and other immune cells. Abbreviations: TAM = Tumour associated macrophage; UMAP = Uniform manifold approximation and projection; PI-TAM = Pro-inflammatory TAM; LA-TAM = Lipid-associated TAM. Note: Cluster 3 = TAM1/PI-TAM. Cluster 4 = TAM2/LA-TAM. The histology interpterion was supported by staff pathologist Dr Alistair Easton of Oncology Department, Oxford

5.2.6 Lipid-associated macrophages have high expression in TNBC, HER2 & Hypoxic tumours

Next, I investigated the clinical relevance of the TAMs in relationship to PAM50 molecular subtypes (Luminal A, Luminal B, HER2, Basal), hormone receptor status, and hypoxia in 1093 TCGA breast samples. I derived two mutually exclusive signatures from the consensus clusters. Unlike the TAM signatures from Azizi et al, the newly derived signatures have no overlapping genes between each other (See method chapter for more details). The gene signature scores were calculated using the median expression values of all signature genes in each sample.

Both PI-TAM and LA-TAM have higher expressions in more aggressive diseases, Basal-like and HER2, than in Luminal A and Luminal B groups (Figure 5.9A&B). Furthermore, the overall expression of PI-TAM appeared to be higher than LA-TAM across all four PAM50 subtypes (Figure 5.9C). There was a significant positive correlation between LA-TAM and hypoxia (Figure 5.9D), which was stronger than the correlation between PI-TAM and hypoxia (Figure 5.9D). Estrogen receptor (ER) and progesterone receptor (PR) expression levels negatively correlated with both TAMs. HER2 expression did not have a significant correlation with any of the TAMs (Figure 5.9D). Therefore, TAMs have high expression in more aggressive breast cancer subsets, Basal-like and HER2, and the hypoxic microenvironment preferentially enriched with LA-TAMs.

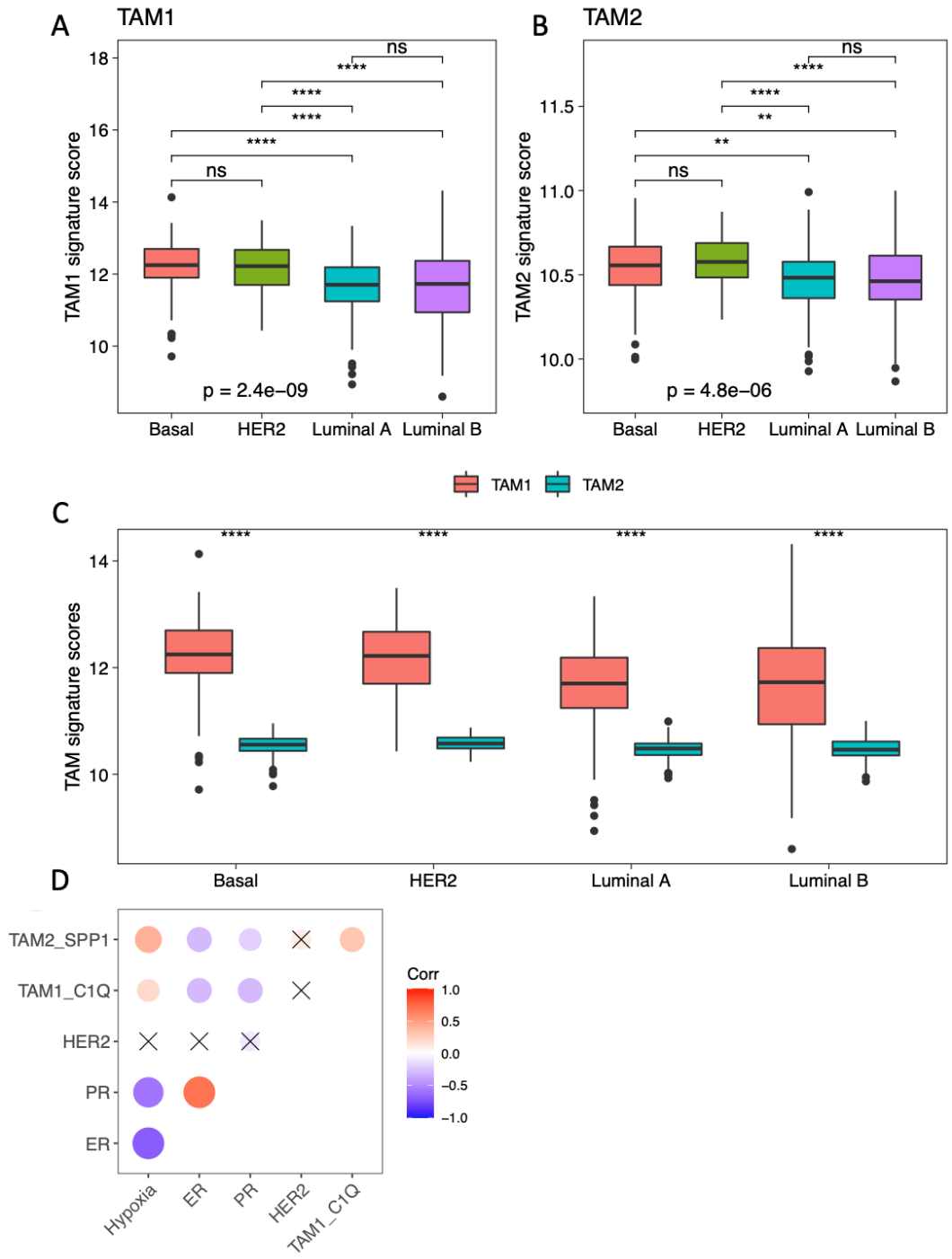


Figure 5.9 Expressions and associations of the two TAMs in TCGA breast samples with respect to PAM50, ER, PR, HER2 and hypoxia.

A & B) TAM1/PI-TAM and TAM2/LA-TAM have high expressions high in more aggressive breast cancer subtypes: Basal-like and HER2-ENRICHED. C) Across subtypes, TAM1 has a higher expression level than TAM2. D) TAM2/LA-TAM strongly correlated with tumour hypoxia, but not the TAM1/PI-TAM. Symbols: ns $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

5.2.7 Lipid-associated TAM signature is associated with poor prognosis in breast cancer

I investigated the prognostic relevance of the two TAM signatures in METABRIC, which were examined together with the curated 16 macrophage signatures and four other signatures as a reference, hypoxia, ESR1 (ER), PGR (PR), and ERBB2 (HER2).

First, I performed univariate proportional hazards regression analyses using signature scores and summarised the results in a forest plot (Figure 5.10). Among the curated macrophage signatures, strikingly only 3 of them were statistically significant. LA-TAM signature was significantly associated with survival (hazard ratio: 1.51, $p = 0.035$), but not the PI-TAM signature (hazard ratio: 1.05, $p = 0.312$). The univariable regression result indicated that LA-TAM was associated with poor prognosis in breast cancer.

Next, I performed a multivariate proportional hazards regression analysis to determine the prognostic value of the LA-TAM in relationship to four key clinical parameters: patient age, tumour size, grade, and stage (Figure 5.11). The result indicated that the LA-TAM signature was an independent prognostic marker (HR 1.7, $p = 0.007$).

The LA-TAM signature was found to be significantly associated with more aggressive Basal-like and HER-enriched phenotypes (Figure 5.9A&B), which have been shown to have poorer survival rates compared to Luminal A and Luminal B breast cancers. In order to evaluate the prognostic value of the LA-TAM signature whilst accounting for the influence of breast cancer subtypes, different breast cancer subtypes were included as covariates in subsequent Kaplan-Meier analyses. This analysis aimed to determine the prognostic power of the LA-TAM signature independent of subtype.

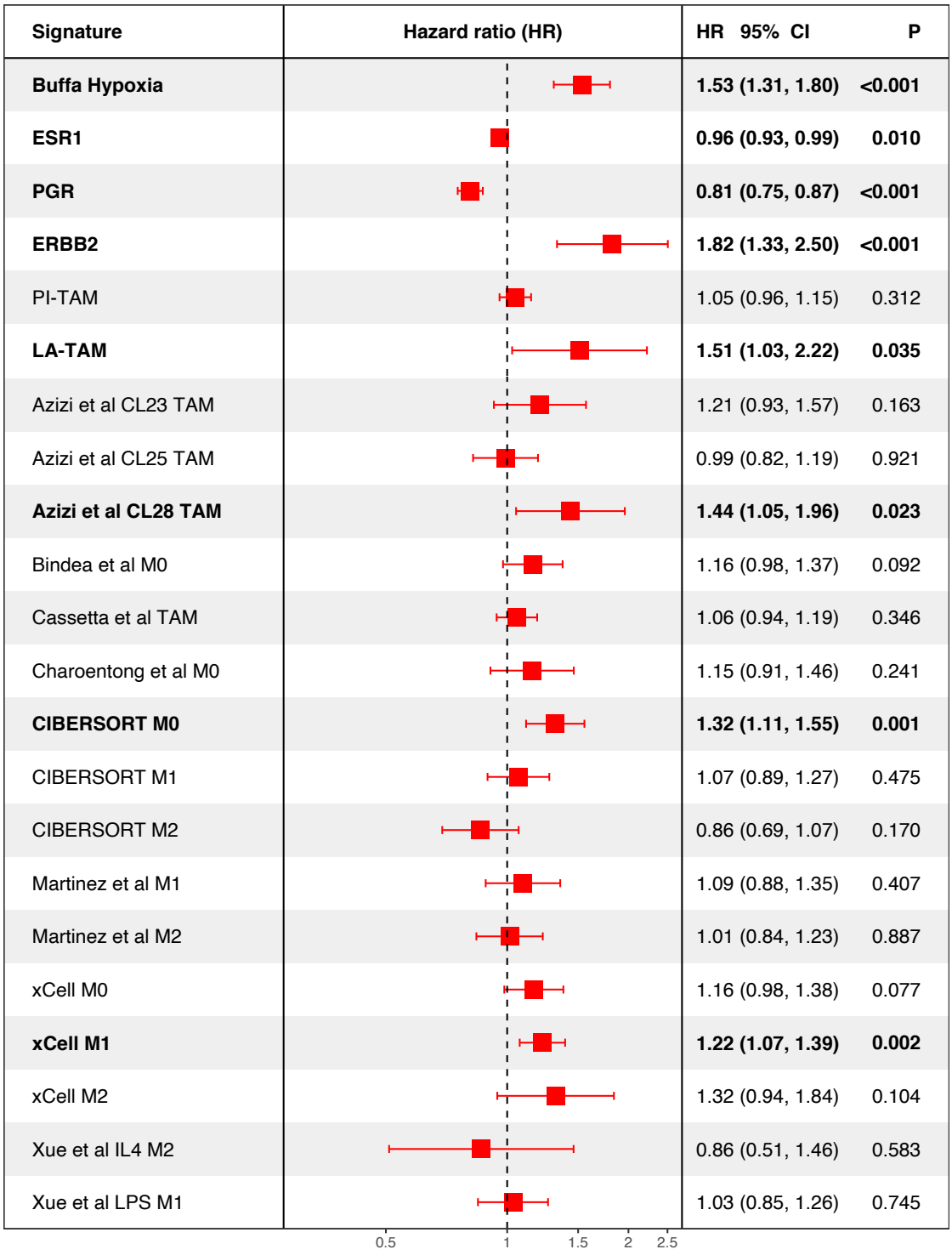


Figure 5.10 Pooled univariate proportional hazards regression results from 18 different macrophage signatures and four reference signatures (ER, PR, HER2 and hypoxia).

Besides the four references, only four other signatures reached statistical significance, including the newly derived lipid-associated TAM signature (LA-TAM). Signatures that reached statistical significance were labelled in bold.

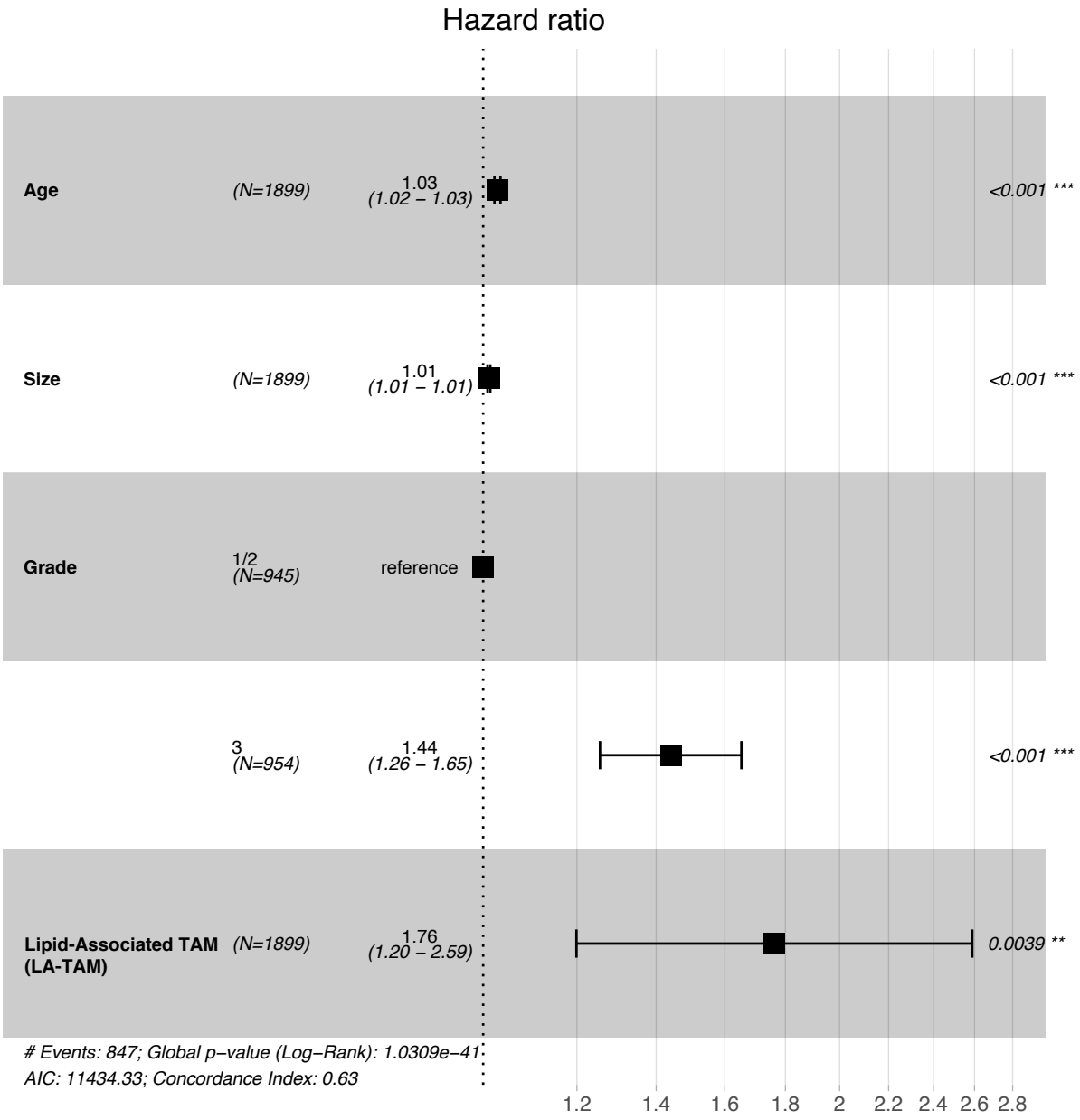


Figure 5.11 Multivariate proportional hazards regression model of lipid-associated TAM signature in METABRIC.

The LA-TAM signature score and four clinical parameters were used as variables (age, size, grade) in the multivariate analysis. This demonstrated the LA-TAM signature is an independent prognostic marker in breast cancer (HR: 1.76, p = 0.0039)

5.2.8 LASSO refined lipid-associated macrophage signature is more prognostic than HER2

To isolate the key variables within the LA-TAM signature and to improve the prognostic performance of the signature, I performed the least absolute shrinkage and selection operator method (LASSO). This is a penalized regression approach that helps to reduce a set of known covariates with the goal to improve the prediction accuracy and interpretability of the model (Tibshirani, 1996). In this case, the covariates were the 346 genes in the LA-TAM signature. Similar to any other supervised machine learning algorithms, LASSO required labelled data; hence, before performing LASSO, I first had to label the data or establish clusters within the METABRIC dataset.

In brief, I calculated the LA-TAM signature scores in 1991 METABRIC samples. I then used two different techniques to establish LA-TAM high and low clusters within the dataset: median expression value and maximally selected rank statistics (Hothorn and Lausen, 2003). The median method divided the samples into two equal size clusters. The maximally selected rank statistics optimized the best cut-off point of a continuous variable with respect to the retrospective survival outcomes (Figure 5.12A). I then compared the results of the two methods, which shared high agreements on classifying the samples into high and low LA-TAM groups, only 139 samples or 7% of the total samples did not show agreement (Figure 5.12B). Therefore, I discarded the samples without agreements on classifications and then performed LASSO regression on the remaining 2008 samples. LASSO shrank the parameters of LA-TAM signature from 346 genes to 160, which was a 54% reduction in the overall parameters (Figure 5.12C).

I then tested the performance of the LASSO refined LA-TAM signature along with the curated 16 macrophage and four reference signatures (ESR1, PGR, ERBB2, hypoxia). The multivariate proportional hazards regression analyses used the same four clinical parameters (age, size, grade), identical to the previous analysis in Figure 5.11. The results of the multivariate analyses were

summarised in a meta-table (Table 5.1). Two important observations were made from the meta-analysis of multivariate regression analyses.

First, only 3/16 of curated macrophage signatures were prognostic, featuring Bindea M0, CIBERSORT M0 and xCell M1, all three were derived from *ex vivo* experiments. And none of the previous reported *in vivo* derived TAM signatures was prognostic, which was unexpected. For instance, Cassetta et al have reported their TAM signature as a prognostic signature in METABRIC, but I could not reproduce the result.

Second, the LASSO refined LA-TAM signature (HR: 1.876, BH adjusted p = 0.001442) has achieved the highest performance among all signatures including ERBB2/HER2 (HR: 1.690, BH adjusted p = 0.009263). LASSO was able to refine and improve the original LA-TAM signature (HR: 1.691, BH adjusted p = 0.01812). Additionally, I verified the prognostic performance of LASSO refined LA-TAM signature in three independent breast cancer datasets (Figure 5.13 & Table 5.2). Therefore, the results indicated the two LA-TAM signatures were the only *in vivo* derived TAM signatures that have reached statistical significance after multivariate proportional hazards regression analyses followed by multiple hypothesis testing. Given the high performance of the LASSO refined LA-TAM signature, I believe it would be a good biomarker candidate to aid the discovery and understanding of this particular group of lipid-associated TAM in breast cancer.

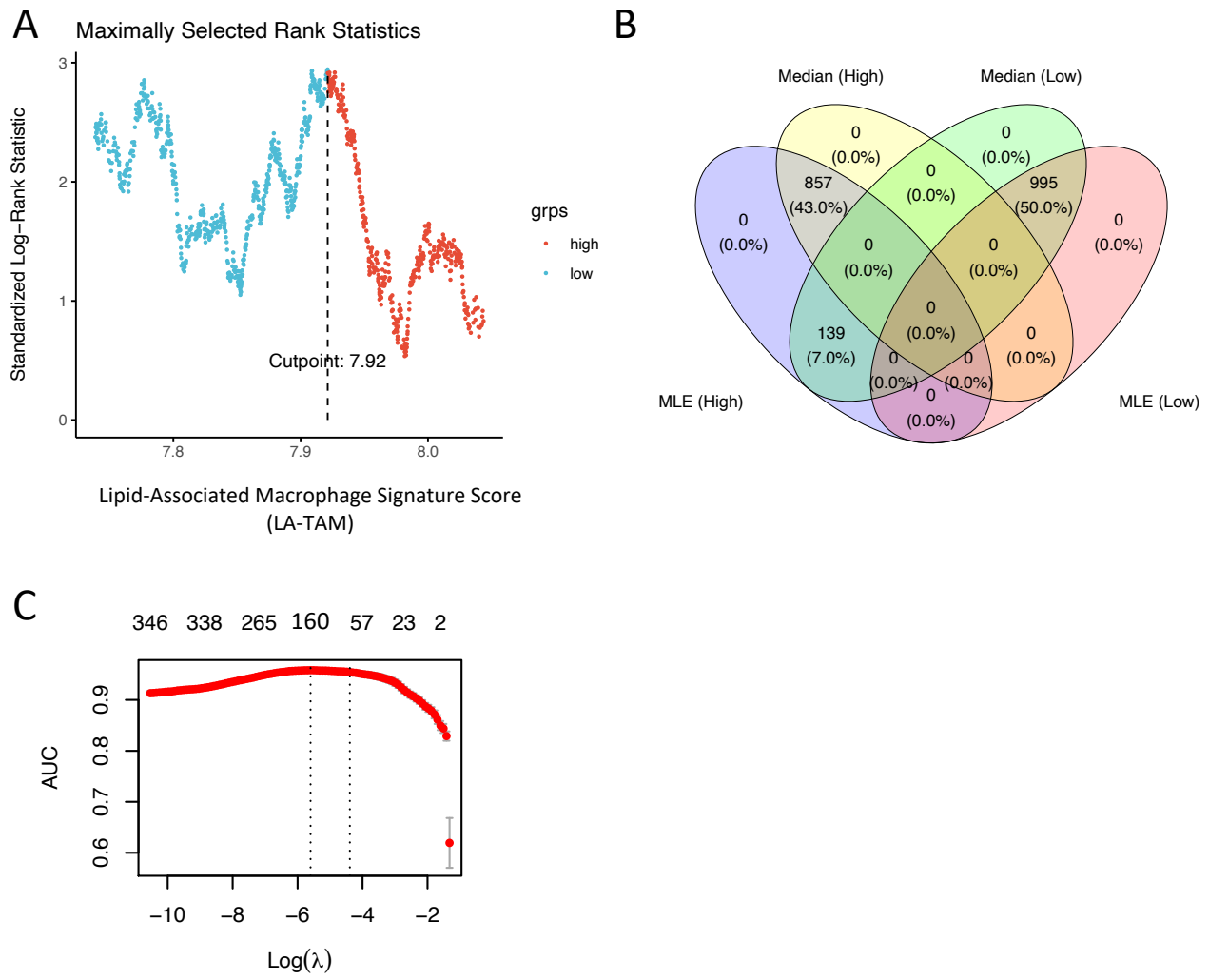


Figure 5.12 LASSO refinement of the LA-TAM signature.

A) METABRIC samples were divided into the high or low expression groups of LA-TAM signatures using two ways. The first method was using the median expression of LA-TAM, which divided the samples into two equal-size groups (not shown). The second method used Maximum Likelihood Estimation (MLE) to identify a cut-off point. B) The samples with agreements from both the median split and MLE method were moved forward into the LASSO refinement step. C) LASSO output indicated the signature has been refined to 160 genes, a significant 54% reduction from the original 346 genes.

Table 5.1 Pooled results of multivariate survival analyses using macrophage signatures

Signatures	HR	95% CI (Lower)	95% CI (Upper)	P	P.adj (BH)
LA-TAM (LASSO Refined)	**1.876**	**1.327**	**2.652**	**3.707E-04**	**1.442E-03**
LA-TAM	**1.691**	**1.152**	**2.481**	**7.249E-03**	**1.812E-02**
ERBB2	**1.690**	**1.189**	**2.402**	**3.441E-03**	**9.263E-03**
Buffa Hypoxia	**1.606**	**1.349**	**1.911**	**1.000E-07**	**3.500E-06**
Xue et al IL4 M2	1.477	0.853	2.556	1.639E-01	2.390E-01
CIBERSORT M0	**1.328**	**1.105**	**1.597**	**2.556E-03**	**7.454E-03**
Azizi et al CL28 TAM	1.273	0.919	1.763	1.466E-01	2.298E-01
Bindea et al M0	**1.258**	**1.053**	**1.502**	**1.131E-02**	**2.639E-02**
Charoentong et al M0	1.232	0.967	1.57	9.178E-02	1.691E-01
xCell M1	**1.195**	**1.04**	**1.374**	**1.214E-02**	**2.656E-02**
Martinez et al M1	1.191	0.952	1.489	1.265E-01	2.213E-01
Martinez et al M2	1.184	0.975	1.437	8.843E-02	1.691E-01
Azizi et al CL23 TAM	1.152	0.883	1.504	2.971E-01	3.714E-01
xCell M0	1.145	0.958	1.367	1.358E-01	2.263E-01
Xue et al LPS M1	1.100	0.897	1.35	3.595E-01	4.317E-01
CIBERSORT M1	1.094	0.899	1.33	3.700E-01	4.317E-01
Azizi et al CL25 TAM	1.074	0.89	1.296	4.548E-01	5.135E-01
Cassetta et al TAM	1.068	0.946	1.205	2.894E-01	3.714E-01
PI-TAM	1.062	0.965	1.168	2.158E-01	2.905E-01
xCell M2	0.976	0.678	1.405	8.971E-01	8.971E-01
CIBERSORT M2	0.959	0.764	1.205	7.219E-01	7.543E-01
ESR1	**0.924**	**0.892**	**0.957**	**1.190E-05**	**1.983E-04**
PGR	**0.845**	**0.782**	**0.912**	**1.700E-05**	**1.983E-04**

This table summarised the HR and statistical significance of a series of multivariate survival analyses. For each signature in the table, a multivariate analysis identical to Figure 5.11 was performed using METABRIC dataset. The newly derived LA-TAM and LASSO refined LA-TAM outperformed all macrophage signatures and reference signatures (ERBB2/HER2 and hypoxia). These two TAM signatures were the only *in vivo* derived macrophage signatures that reached statistical significance. Only 3/16 curated macrophage signatures reached statistical significance. All 3 signatures were derived from *ex vivo* experiments.

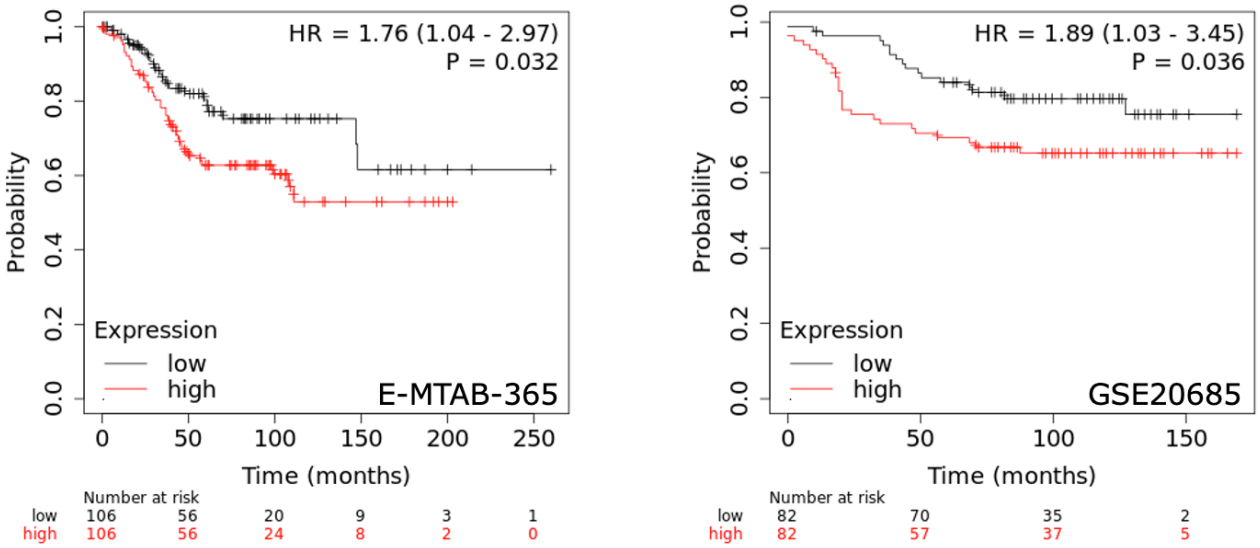


Figure 5.13 Verifications of LA-TAM signature in independent datasets.

The LASSO refined LA-TAM signature was verified in two independent breast cancer datasets. The results indicated the signature was prognostic in multiple datasets.

Table 5.2 Pooled results of multivariate survival analyses in NKI breast cancer data

Signatures	HR	95% CI (Lower)	95% CI (Upper)	P	P.adj (BH)
Buffa Hypoxia	**1.810**	**1.458**	**2.248**	**1.000E-07**	**2.300E-06**
LA-TAM	**1.478**	**1.153**	**1.897**	**2.087E-03**	**1.200E-02**
LA-TAM (LASSO Refined)	**1.397**	**1.113**	**1.753**	**3.952E-03**	**1.818E-02**
Azizi et al CL28 TAM	1.282	1.018	1.614	3.437E-02	1.317E-01
CIBERSORT M0	1.203	0.979	1.478	7.886E-02	2.591E-01
Azizi et al CL25 TAM	1.177	0.948	1.46	1.402E-01	4.030E-01
ERBB2	1.163	0.928	1.459	1.897E-01	4.849E-01
Bindea et al M0	1.125	0.9	1.405	2.999E-01	6.325E-01
xCell M2	1.117	0.882	1.415	3.567E-01	6.700E-01
Cassetta et al TAM	1.114	0.907	1.367	3.025E-01	6.325E-01
Xue et al IL4 M2	1.100	0.867	1.396	4.307E-01	6.700E-01
xCell M1	1.095	0.884	1.357	4.048E-01	6.700E-01
xCell M0	1.078	0.865	1.345	5.020E-01	6.700E-01
Chi et al C1Q TAM	1.078	0.878	1.323	4.742E-01	6.700E-01
Martinez et al M1	1.073	0.877	1.312	4.941E-01	6.700E-01
Xue et al LPS M1	1.069	0.87	1.315	5.244E-01	6.700E-01
Azizi et al CL23 TAM	1.068	0.857	1.33	5.577E-01	6.751E-01
CIBERSORT M2	1.031	0.833	1.277	7.762E-01	8.150E-01
CIBERSORT M1	1.029	0.847	1.25	7.714E-01	8.150E-01
Charoentong et al M0	1.020	0.825	1.261	8.530E-01	8.530E-01
Martinez et al M2	0.968	0.771	1.216	7.795E-01	8.150E-01
PGR	**0.615**	**0.506**	**0.748**	**1.000E-06**	**7.700E-06**
ESR1	**0.587**	**0.477**	**0.724**	**6.000E-07**	**6.900E-06**

This table summarised the HR and statistical significance of a series of multivariate survival analyses. For each signature in the table, a multivariate analysis identical to Figure 5.11 was performed using the Netherlands Cancer Institute (NKI) breast cancer dataset, which has 295 breast cancer samples. The newly derived LA-TAM and LASSO refined LA-TAM outperformed all macrophage signatures. These two TAM signatures were the only *in vivo* derived macrophage signatures that reached statistical significance.

5.2.9 Lipid-associated TAM associated with poor prognoses in hypoxic, ER, PR, and HER2 tumours

Next, I assessed the clinical relevance of the LA-TAM signature in METABRIC samples. I stratified the samples in four different ways using hypoxia level and receptor status (ER, PR, HER2). LA-TAM score was calculated using the LASSO refined 160-gene signature. I used hypoxia as a proxy for aggressive and high-grade tumours, which tended to be HER2+ or triple-negative breast cancer.

For each analysis, I stratified samples into high and low groups based on the median values of variables across samples such as ER, PR, HER2, hypoxia gene signature expression, and the LA-TAM gene signature expression. It is also important to note in the METABRIC trial, patients were treated with chemotherapy and hormonal therapies, but no anti-HER2 therapies since they were not available at the time (Curtis et al., 2012).

For both hormonal receptor stratifications, high expression of the hormonal receptor was associated with good outcomes. For ER and LA-TAM stratifications (Figure 5.14A), high ER expression was associated with good prognoses irrespective of the LA-TAM expression level. Low ER expression has poorer outcomes than high ER. Within low ER groups, high LA-TAM has the worst outcomes, which suggested a high level of LA-TAM aggravated already poor outcomes of the low ER groups.

For the high PR group, high LA-TAM worsened the good outcome (Figure 5.14B). No difference was observed in the low PR group with respect to LA-TAM level. High HER2 with high LA-TAM was associated with the worst outcome, and low HER2 with high LA-TAM was associated with the most favourable outcome (Figure 5.14C). However, the relationship between HER2, LA-TAM and survival would be significantly changed with the introduction of anti-HER2 therapies. I observed that low hypoxia was associated with favourable outcomes irrespective of LA-TAM level (Figure 5.14D). In high hypoxic samples, having high LA-TAM further exacerbate the poor outcomes, which suggested an intrinsic link between LA-TAM and the hypoxic tumour microenvironment. Therefore, having high LA-TAM expression in breast cancer would make the already poor outcome groups worse, and make the good outcome groups less favourable.

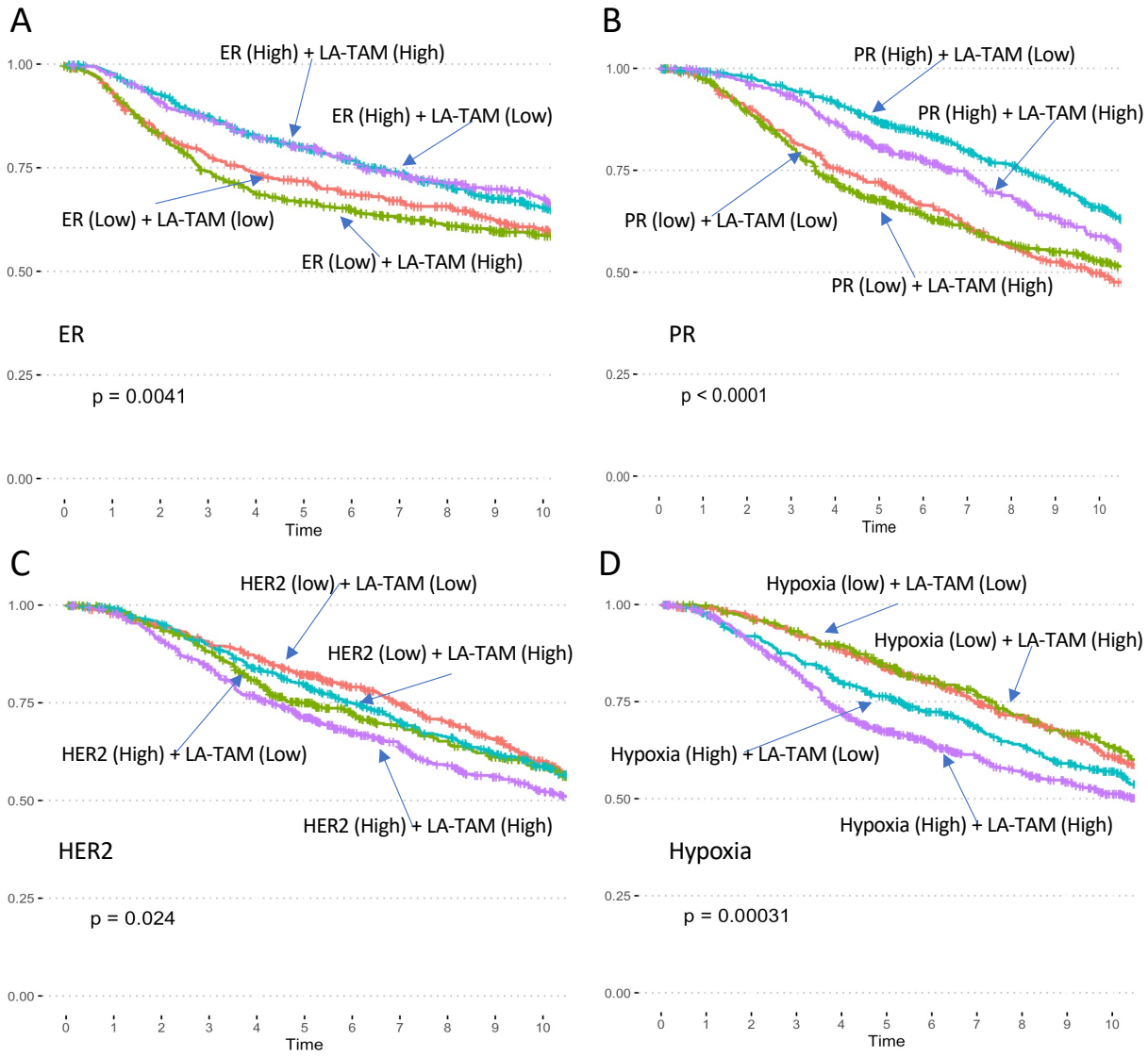


Figure 5.14 Impact of LA-TAM on overall survival in four different stratifications

The METABRIC samples were stratified using ER, PR, HER2 and hypoxia to assess the clinical relevance of the LA-TAM signature. A) ER high was associated with good outcome irrespective of the LA-TAM level. ER low group was associated with poor survival and having high LA-TAM made this group even worse. B) High PR status was associated with good outcomes, but high LA-TAM made high PR less favourable. C) High HER2 in this dataset was associated with poor outcomes, since anti-HER2 therapy was not introduced when the dataset was collected. In the high HER2 group, having a high amount of LA-TAM aggravated the poor outcomes. D) hypoxia high was associated with poor outcomes and a high LA-TAM level further intensified the unfavourable prognosis.

5.2.10 Lipid-associated macrophages are associated with low anti-tumour immunity

Lastly, I analysed the core functions of LA-TAM within the tumour microenvironment by conducting a geneset overrepresentation analysis on the 160 genes in the LASSO refined LA-TAM signature, enriched with GO Molecular Functions and GO Immune Process gene lists (Figure 5.15). The pathway analysis indicated the genes in the LA-TAM signature were associated with myeloid cell activation, regulation of macrophage migration, cytokine receptor binding, integrin binding, regulation of leukocyte chemotaxis, and regulation of T cell activation.

I also inspected the individual genes involved in multiple pathways. For instance, CCL3, also better known as the macrophage inflammatory protein 1-alpha protein (MIP-1α), exhibited a central role in linking three key macrophage-specific pathways, namely: myeloid cell activation involved in immune response, regulation of myeloid cell differentiation and positive regulation of leukocyte chemotaxis. Previous studies have shown that macrophage secreted CCL3 was important in facilitating breast cancer metastasis and CCL3 expression in the macrophages can be upregulated via the CCL2-CCR2 chemokine cascade (Kitamura et al., 2015). Recent studies have shown that CCL3 was important in recruiting T cells into inflammatory sites (Gibaldi et al., 2020), contributing to secondary damage post injury (Pelisch et al., 2020), and relating to mature myeloid lineage differentiation (Staversky et al., 2018).

LA-TAM signature also exhibited an association with CCL22, which promotes the recruitment of and cell-cell interactions with Tregs (Rapp et al., 2019). Tregs in the tumour microenvironment inhibited antitumour immunity. Furthermore, LA-TAM expressed multiple immune checkpoints including CD276 and SIGLEC15. CD276 serves as a T cell inhibitor to promote tumour proliferation and invasion. SIGLEC15 have recently been shown as a critical immune suppressor of antigen-specific T cells in cancer. In addition, MMP14 could promote tumour neo-angiogenesis and immune-modulatory actions in triple-negative breast cancer, which is a clinically relevant target in metastatic cancers (Ling et al., 2017). MMP14 blockade in breast cancer murine model demonstrated decreasing immunosuppressive TGFβ and changes in macrophage phenotype (Ager et al., 2015). Finally, since

the phenotype was defined as lipid-associated TAM, it was not surprising to find CD36 in the LA-TAM signature. CD36, also known as scavenger receptor class B member 3 (SCARB3), imports fatty acids into cells and the efflux of cholesterol could negatively regulate the immunity within TME (Ma et al., 2021).

Together the pathway analysis of the LASSO refined LA-TAM signature depicted a terminally activated macrophage phenotype found within chronic and hypoxic wounds that features: lipid metabolism, angiogenesis, and matrix remodelling. Most importantly, they negatively regulate the immunity within TME, possibly via cholesterol efflux, Treg recruitment, and checkpoint inhibition on antigen-specific T cells.

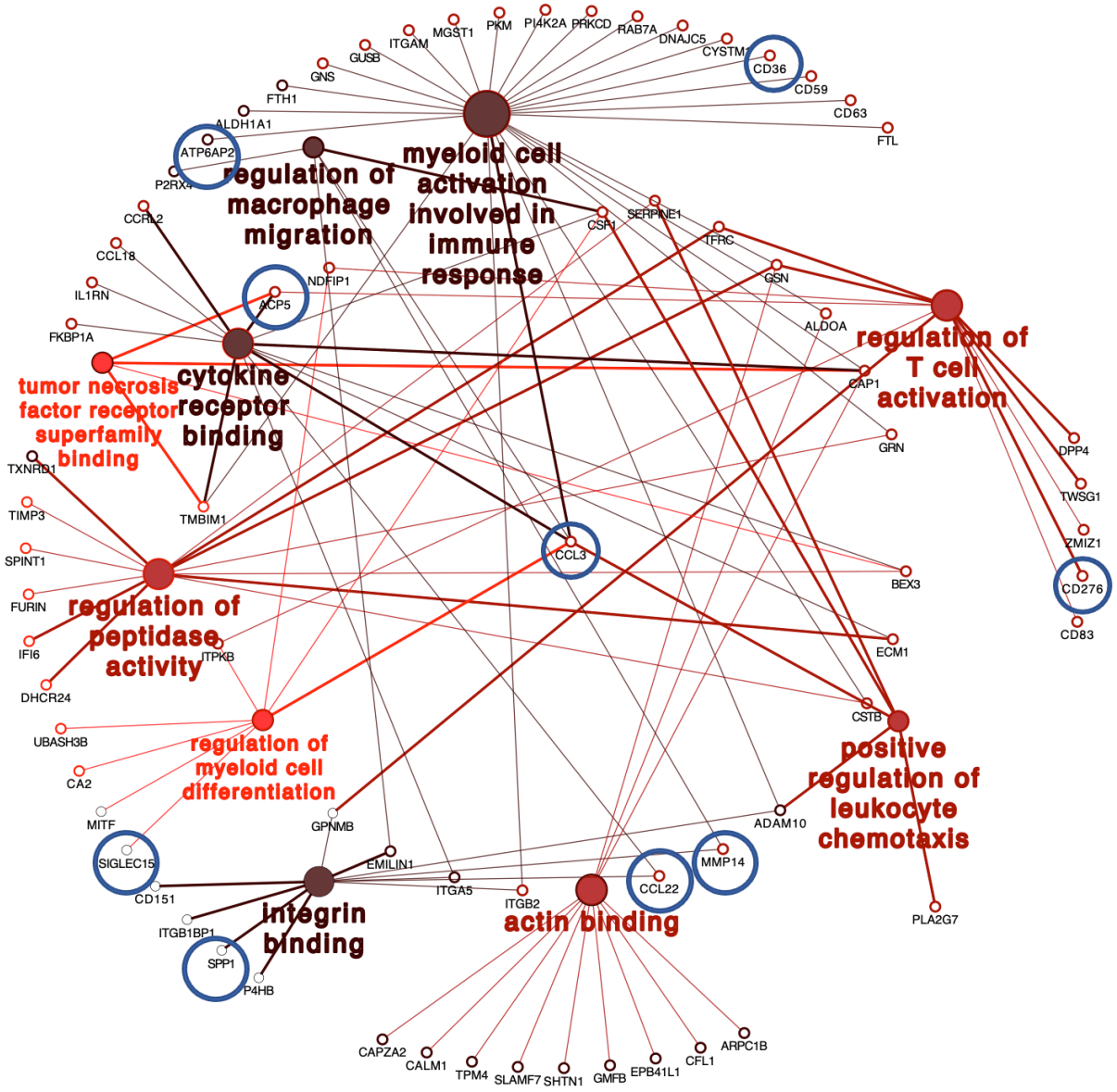


Figure 5.15 Functional characterisation of LASSO refined LA-TAM signature.

The core functions of the LA-TAM centred around the regulation of macrophage migration, cytokine receptor binding, and regulation of T cell activation. Some of the genes involved in negative regulations of anti-tumour immunity are highlighted in blue circles. For example, CD276 is a newly identified member of the B7 family proteins, which inhibits T-cell and promotes tumour proliferation. In addition, LA-TAM features another immune checkpoint SIGLEC15, which has been shown to inhibit antigen-specific T cell functions. In this graph, the nodes represent gene ontology terms and the edges represent genes associated with those terms.

5.3 Discussion

5.3.1 Summary of results

To my knowledge, this is the first study to systematically demonstrate macrophage signatures derived from TAM and *ex vivo* monocytes have a high degree of divergence and there was little agreement among the 16 published macrophage signatures. It has long been hypothesised that the *ex vivo* derived macrophage signatures do not resemble the features of TAM (Martinez and Gordon, 2014). Recent literature from single-cell sequenced immune samples also help to reinforce that M1/M2 dichotomy may not exist in TAMs. Specifically to breast cancer, one of the first studies showed that some of the canonical markers have shared expression in different TAM groups in breast cancer (Chung et al., 2017), then another study demonstrated different groups of TAM all have a comparable level of either M1 or M2 signature expression (Azizi et al., 2018). Furthermore, I also revealed the *ex vivo* derived M1/M2 signatures could not stratify samples into different groups in bulk tumour samples and single-cell TAM RNA-seq datasets using the *ex vivo* derived macrophage signatures. This is concerning as these *ex vivo* signatures were all collected from highly cited publications or popular digital dissection tools and these signatures were often used to infer TAM subpopulations in tumour transcriptomic datasets, when in fact, there was little resembling between these *ex vivo* signatures and TAMs.

I also revealed different clustering algorithms could lead to diverse monocyte and macrophage numbers in the same single-cell dataset. By applying a more rigorous approach using consensus clustering with resampling I found the TAM diversity was less than previously described. Consensus clustering with resampling is more computationally intensive than the clustering algorithms build in Seurat, due to this consensus clustering approach may not be feasible for large single-cell datasets (Kiselev et al., 2019). But for my purpose: an in-depth analysis of a relatively small number of TAMs, it has helped to reveal novel biological insights and to develop a better TAM signature. To confirm the presence of the two TAMs I have identified, I mapped them in a triple-negative breast cancer single-cell dataset and then characterised their functions using the differentially expressed genes. I

found the two TAMs seemed to serve distinctive functions within the TME: pro-inflammatory and lipid-metabolic TAMs. Although the LA-TAM subgroup has a relatively low expression level compared with the pro-inflammatory TAM, my spatial transcriptomics and hypoxia signature correlation analyses have shown these LA-TAM cells occupied a special hypoxic niche within TME with close interactions with the cancer cells and other immune cells. TAM variability appeared to be a result of adaptation to the diverse cellular and conditional heterogeneity within the TME. Furthermore, the survival analyses also indicated the presence of this relatively small group of lipid-metabolic TAM correlated with poor survival in different breast cancer subtypes.

5.3.2 M1/M2 *ex vitro* models are less relevant in tumour pathology

The two TAM clusters I identified from single-cell breast cancer datasets do not resemble any of the previously described *ex vivo* macrophage profiles, which is likely a result of the heterogeneous nature and complexity signalling of TME. For example, the pro-inflammatory TAM I found featured key pathways in viral protein interaction, TLR signalling, IL17 signalling, and NF-kB (Figure 5.7). From a functional view, this TAM profile resembled the classical definition of macrophage or M1, which elicits pathogen recognition via pattern recognition receptors and activates the pro-inflammatory responses. However, key genes of the viral protein interaction pathway also involved IL10 and IL2RA, both are anti-inflammatory and can be used to describe a glucocorticoid and IL10 induced M2 profile (van de Garde et al., 2014; Park-Min et al., 2005). In addition, there is no lipid metabolic macrophage from the *ex vivo* M1/M2 models. Thus, in my opinion, the *ex vivo* derived M1/M2 models may still be useful to facilitate discussions and theoretically conceptualise macrophage functions, but the M1/M2 models are less relevant in cancer or other pathological conditions.

5.3.3 Source of tumour-associated macrophages

The source of macrophages in different tissues can originate from two sources: the yolk sac during development and blood-borne monocytes (Aras and Raza Zaidi, 2017; Gomez Perdiguero et al., 2015; Gordon and Martinez-Pomares, 2017). My clustering analysis (Figure 5.5) revealed TAMs in human

breast cancer are only found in tumours, but not in blood, lymph node, and adjacent normal samples. The trajectory inference model suggested that these TAMs fall within a spectrum of continuous activated infiltrating monocyte-macrophages. Previous developmental studies have demonstrated using mouse models that TAMs in breast cancer and the tissue-resident macrophages all originated from $Ly6C^{hi}$ $CCR2^{+}$ monocytes (Franklin et al., 2014; Tymoszuk et al., 2014), which is unique compared with the origins of macrophages in other tissues (Franklin and Li, 2016). Furthermore, it has also been shown TAMs have a higher capacity to expand their population through proliferation than healthy mammary tissue-resident macrophages (Franklin et al., 2014). My data were in agreement with these previous findings that in human breast cancer, the source of TAMs seemed to be exclusively derived from infiltrating monocytes. Future work could explore this further by employing higher single-cell sequencing depth and including tissue-resident macrophages from healthy controls.

5.3.4 Tumour associated macrophage and estrogen signalling

I found LA-TAM negatively associated with estrogen receptor (ESR1 encoding ERα) mRNA levels in breast cancer. The estrogen receptor is a critical transcriptional regulator in breast cancer biology and the loss of ERα is associated with hormonal therapy resistance and adversary outcome (Stossi et al., 2012). It has been previously demonstrated that macrophage canonical marker CD68 or CD163 expression was associated with ERα- and PR-negative breast tumours, which suggested another possible process that TAMs can negatively impact the breast cancer outcome (Gwak et al., 2015; Lindsten et al., 2017; Mahmoud et al., 2012; Medrek et al., 2012; Mohammed et al., 2012; Tiainen et al., 2015). The transcriptional repression of ERα is likely through the activation of the MAPK/ERK2 kinase cascade that initiated is by macrophage-derived cytokines, especially IL6, IL8, and TNFα (Chavey et al., 2007; Lee and Nam, 2008; Stossi et al., 2012). Therefore, early intervention of TAMs

may offer new opportunities to improve the efficacy of hormonal therapy and overcome hormonal therapy resistance.

5.3.5 Lipid associated macrophage – linking obesity and cancer

In my study, I have found a type of lipid-associated TAM (LA-TAM) that featured cholesterol metabolism paired with increased lysosome activities. This lipid metabolic TAM profile was not part of the *ex vivo* derived M1/M2 model, but they share a resemblance with certain macrophage profiles that have been previously reported in obesity, atherosclerotic plaques, and certain cancer.

In obesity, it has been reported that adipose tissue macrophages displayed a foamy appearance and increased expression of CD36 and lipoprotein lipase (LPL) (Prieur et al., 2011), both genes were also upregulated in my LA-TAM profile. Furthermore, obesity-associated lipid macrophages also displayed increased lysosomal processing and storage activities complemented by lipid metabolism (Xu et al., 2013), which also matched with my LA-TAM upregulated pathways. Similar lipid metabolic macrophages have also been found in the atherosclerotic plaques using single-cell sequencing. A recent study reported the identification of a TREM2 high macrophage in mouse atherosclerotic plaques with high expression of SPP1, CTSD, and CD9. These TREM2 high macrophages also seemed to specialise in lipid metabolism and with similar gene expression patterns to osteoclasts (Cochain et al., 2018). In colorectal cancer, it has been recently demonstrated that lipid accumulations in TAMs can lead to increase fatty acid oxidation and immunosuppressive functions. By targeting lipid accumulations in these macrophages, both immunosuppression and tumour growth has been prevented (Wu et al., 2019). Additionally, cancer cells can also educate TAMs to efflux cholesterol into the TME (Goossens et al., 2019). Cholesterol efflux can drive tumour progression by modulating the immune system (Ma et al., 2021). Together, these foamy macrophages in plaque, obesity and cancer share many similarities with the LA-TAM in breast cancer that specialised to

scavenge and catabolizing lipids, and then efflux the excess cholesterol to promote tumour growth and suppress effector cell functions.

5.3.6 Prognostic value of LA-TAM signature

I have shown most of the curated macrophage signatures were not prognostic in univariate and multivariate proportional hazard analyses, but my LA-TAM signature was significant in both cases. In METABRIC, Kaplan-Meier survival analyses have indicated the presence of LA-TAM aggravated the poor prognosis of low ER, high HER2, or high hypoxia. LA-TAM also reversed good prognoses of high PR or low HER2 to bad. Of note, METABIRC patients did not receive anti-HER2 therapies.

Our multivariate proportional hazards regression analyses have shown high LA-TAM is more hazardous than high HER2. Anti-HER2 therapies have transformed the standard of care for HER2-positive breast cancer and potentially other cancer types with HER2 amplifications (Loibl et al., 2021b). Moreover, I believe my LASSO refined LA-TAM signature could be a valuable tool for breast cancer research and TAM-specific therapy developments.

5.4 Conclusion

Translating the *ex vivo* derived M1/M2 model to *in vivo* TAMs has been difficult, partly due to the complexity and pleiotropic nature of the macrophage biology *in situ*; more importantly, the M1/M2 model is over-simplified to meaningfully capture the macrophage phenotypes in tissues or in tumours. I have found two TAMs from single-cell data that do not resemble subtypes of the M1/M2 model. More specifically I characterised lipid metabolic TAMs, established an independent gene signature using LASSO, and studied its clinical association with different subtypes of breast cancer and their prognostic significance. It is my hope that these carefully derived TAM signatures will facilitate new therapeutic designs and potential change the prospective standard of care in breast cancer and beyond.

6 SUMMARY AND CONCLUSIONS

6.1 Summary

In this dissertation, I have assessed a diverse set of immunogenomic methods, evaluated the performance of different immune signatures, and derived a prognostic tumour associated macrophage signature for breast cancer. The main goal of the studies was to identify clinical relevant insights and to derive novel immune signatures from the transcriptomic data generated from the tumour microenvironment, which could then be used for the development of next-generation cancer diagnostics and immunotherapies.

In Chapter 3, I studied six digital dissection tools that could be used to estimate tumour infiltrating lymphocyte levels in bulk transcriptomic datasets. I test the performance of CIBERSORT and xCell in a series of head-to-head comparisons in 33 different cancer types using the TCGA dataset, because these two tools were the most comprehensive (22 immune types and 68 cell types, respectively) and methodologically distinctive (deconvolution vs gene signature). The results showed high agreements on five cell types that were CD8 T cells, B cells, macrophage M1, macrophage M2, and neutrophils. More importantly, I also observed unexpected positive correlations among the theoretically distinctive macrophage M1 and M2 subtypes, at both gene signature level and individual gene level. This finding led us to investigate the relationship between *ex vivo* M1/M2 signatures and *in vivo* TAM signatures in Chapter 5.

In Chapter 4, hypoxia is a major hallmark of the immunosuppressive tumour microenvironment and I wanted to understand the impact of hypoxia on different immune populations in breast cancer (Arnaiz and Harris, 2022; Fu et al., 2021; Singleton et al., 2021). I established the relationship between tumour hypoxia and different immune infiltrations using an established hypoxia signature

(Buffa et al., 2010). I demonstrated the variation of hypoxia-associated gene expression in PAM50 subtypes and triple-negative breast cancer samples, which agreed with previous reports that tumour hypoxia was associated with more aggressive diseases in breast cancer (Chen et al., 2014; Petrova et al., 2018). In both Basal-like and TNBC, I found an enrichment of macrophage M1, which connected these macrophages with more aggressive diseases and poorer outcomes. Moreover, using hypoxia signature and xCell estimations, I examined the cellular landscape of breast cancer. I identified a strong reciprocal relationship between hypoxia and the stromal content, especially the endothelial compartments. Finally, I applied functional genomics to study the key immune pathways and genes that were associated with hypoxia in breast cancer, which echoed the previous finding that hypoxia negatively regulated the anti-tumour immune response (Fu et al., 2021; Li et al., 2018; Nakagawa et al., 2015).

In Chapter 5, I systematically examined 16 different *ex vivo* and *in vivo* derived macrophage signatures. There have been concerns raised in the medical research community that the traditional M1/M2-based classification was too simplified and the theory needs an urgent update (Cheng et al., 2021; Martinez and Gordon, 2014; Murray et al., 2014; Ren et al., 2021). Nevertheless, these *ex vivo* derived macrophage signatures are still frequently used in analysing transcriptomics of clinical samples at both bulk and single-cell levels. I demonstrated that these macrophage signatures were divergent at the gene level and the *ex vivo* derived signatures were not representative of TAMs in breast cancer. Furthermore, I derived a novel TAM signature using a stringent clustering protocol from consensus clustering. This TAM profile featured lipid metabolism and lysosomal activity and correlated positively with hypoxia. Survival analysis indicated this TAM profile was strongly associated with poor survival in METABRIC data, it even outperformed HER2, which was traditionally linked with bad prognoses until the emergence of anti-HER2 therapies. Therefore, using a combination of different immunogenomic and computational methods, I have identified a novel

macrophage profile featuring lipid metabolism and I derived a novel TAM prognostic signature, specifically for breast cancer.

6.2 Targeting TAM in cancer

The application of a novel TAM prognostic signature offers much potential in the development of TAM targeting agents in Oncology. The conventional TAM targeting strategies can be divided into three groups: macrophage elimination, recruitment inhibition, and reprogramming (Xiang et al., 2021). Macrophages are required to maintain tissue homeostasis and to serve other important effector functions. Hence, an effective therapeutic approach targeting TAM must have high specificity and distinguish TAM from tissue-resident macrophages. The lipid-associated TAM identified in my work offers such an opportunity due to its orthogonal markers and a relatively small population within the tumours.

Both spatial and temporal distributions of these cells should be considered in future work. As I have shown that the two TAM populations occupied spatial distinctive locations within the TME, but the exact developmental relationship between the two TAMs remained unknown. My lineage analysis seemed to suggest that these cells were developed in a linear continuum fashion. I hypothesized that the pro-inflammatory macrophages served anti-tumour functions, but over time these macrophages altered the microenvironment, which helped cancer cells to break free and metastases. I have found that lipid-associated TAMs are associated with more aggressive tumours that exhibit high proliferation due to HER2 amplification and high levels of hypoxia. Targeting these LA-TAMs through removal or conversion to PI-TAMs alone may not be enough to restore balance in such hostile tumour microenvironment. However, combining the inhibition of LA-TAMs' immune-modulatory

function with the introduction of anti-tumour agents like CAR-T cells or antigen-specific T-cells may be a more promising approach for treating high-grade tumours.

6.3 Cluster stability and reproducibility in the single-cell era

I compared the cluster numbers generated from Biscuit, Seurat and Consensus Clustering methods, and found the optimum number of monocyte and macrophage in the tumour was significantly less than previously reported. Clustering at root is an unsupervised classification, and its reproducibility issues have been highlighted in the past with microarray and fMRI data generated from medical sciences (Aarts et al., 2015; Carp, 2012; McShane et al., 2002). In my opinion, identifying stable and reproducible clusters are even more difficult with single-cell data than with other data types used in cancer research, largely due to the high sparsity of the data and ever-increasing sample size. And a recent review has shown that imputation methods have not improved clustering outcomes (Hou et al., 2020). Clustering the high quantity and high dimensional data is computational intensive; thus, many cost-efficient algorithms have been developed to overcome the challenge, which can allow a sizable portion of the single-cell analysis performed on regular laptops or desktops instead of high-performance computing (HPC) platforms. However, these cost-effective algorithms without resampling may trade computational efficiency with the overestimation of single-cell cluster numbers (Grabski et al., 2022).

Many of the clustering methods are described as “unbiased”; however, in practice, such algorithms have multiple tuneable parameters which affect the final results. There is no gold standard for establishing the “right” number of cellular clusters at the moment. Previous efforts have benchmarked the clustering algorithms using normal tissues pronominally isolated from the peripheral blood or the brain (Duò et al., 2018; Kiselev et al., 2019; Mereu et al., 2020). Cells in healthy PBMC samples tended to have more predictable gene expression profiles and there is an abundance of data available for cross-referencing. However, a similar effort is still needed to robustly characterise cells derived from pathological or cancerous tissues, which is challenging as the cell states within the specific

pathological environment are just emerging. I believe with increasing research efforts in the single-cell space and international initiatives like the Human Cell Atlas and continuous technological advancements, I will eventually be able to capture the true cellular diversity in health and diseases.

6.4 Conclusion

Overall, this work has systematically demonstrated that the *ex vivo* and *in vivo* derived macrophage signatures are divergent. These *ex vivo* derived macrophage signatures and M1/M2 models did not reflect the biological functions of TAM in tissues. The new lipid-associated TAM and its prognostic signature could offer new opportunities in immunogenomic-related applications.

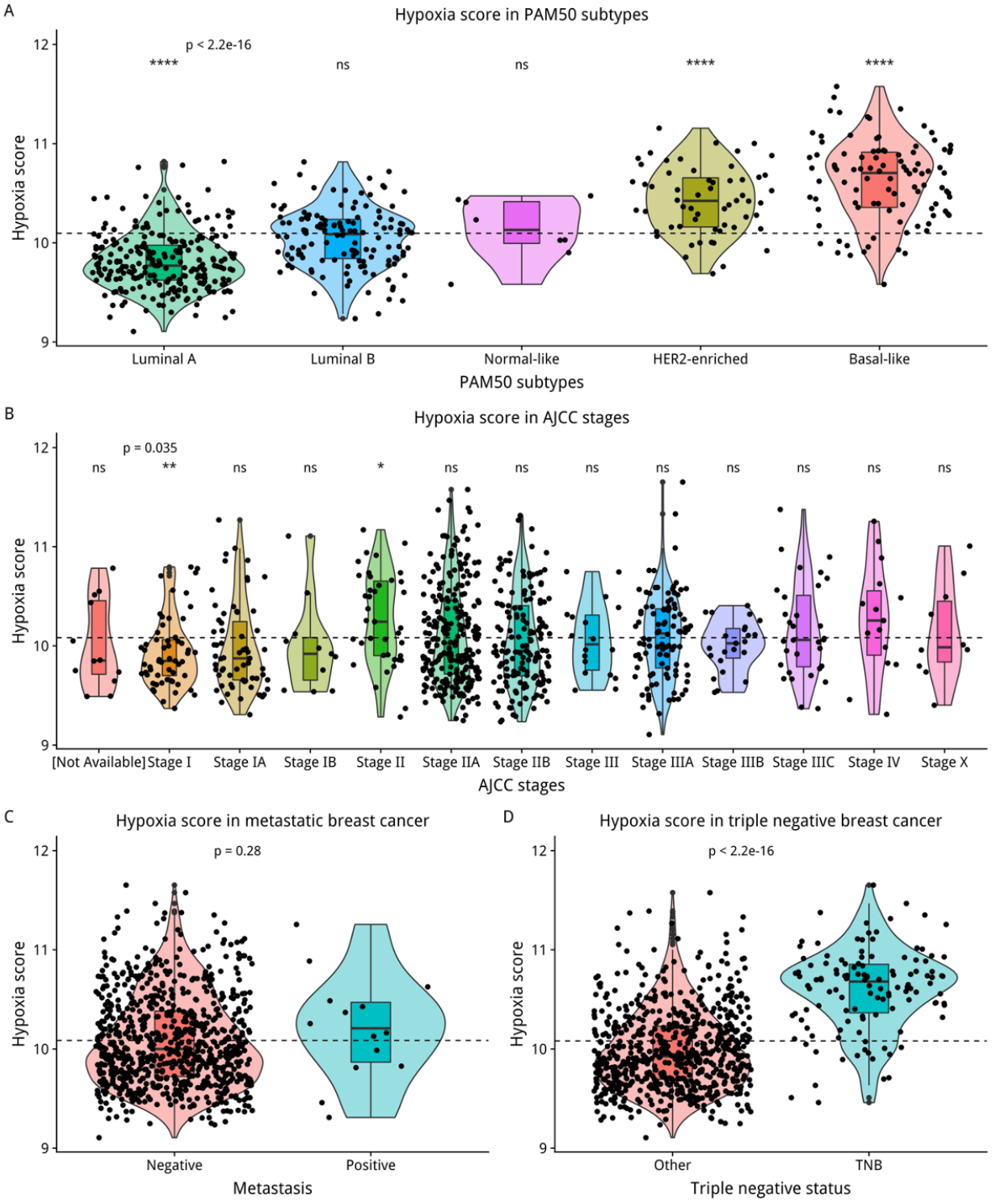
7 APPENDICES

APPENDIX 1: CELLS AND MARKERS USED TO DEVELOP IN SILICO METHODS

	Cell Type Description	Cell Separation Method (CIBERSORT)	Markers (CIBERSORT)	Cell Separation Method (xCell)	Markers (xCell)
B cells	B cells naive	MACS® CD138 microbeads and CD19 microbeads	CD19+CD27- IgG/A-		CD19+,CD20+,CD38low,CD23+
	B cells memory	MACS® CD138 microbeads and CD19 microbeads, then FACS	CD19+ CD27+		CD19+ CD27+ IgD+
CD8 T	T cells CD8	RosetteSep™ CD8+ T-cell enrichment cocktail, CD-8 subset	CD3, CD8, CD45RA		CD3, CD8, CD45RA
CD4 T cells	T cells CD4 naive	Ficoll, then MACS CD4+ T cell isolation kit	CD4+		CD3+ CD4+ CD45RA+
	T cells regulatory (Tregs)	Ficoll-Hypaque, then MACS CD4+ T cell isolation kit, then FACS	CD4+ $CD25^{hi}$		CD3+ CD4+ CD25+ CD127low
Monocytes and Macrophages	Monocytes	MACS® CD14 Microbeads, monocyte subset from blood	N/A	Monocytes isolated from blood	CD14, CD42b, CD64, CD45
	Macrophages M0	Differentiated from monocytes	None known; identified by morphology and phagocytic capability	Differentiated from monocytes using M-CSF (50ng/ml) for 8 days	None known; identified by morphology
	Macrophages M1	Histopaque 1.077, then Miltenyi negative selection monocyte isolation kit and LS columns, then differentiated with 1% medium supplement nutridoma-HU + 100 nM M-CSF, then activated with 20 ng/ml IFN-g+ 100 ng/ml LPS	None known; identified by morphology and phagocytic capability	Stimulated with M-CSF at 50ng/ml for 7 days, then LPS alone at 100ng/ml for the last 18 hours of the culture, cells harvested at day 8	None known; identified by morphology
	Macrophages M2	Histopaque 1.077, then Miltenyi negative selection monocyte isolation kit and LS columns, then differentiated with 1% medium supplement nutridoma-HU + 100 nM M-CSF, then activated with 20 ng/ml IFN-g+ 100 ng/ml LPS and 20 ng/ml IL-4	None known; identified by morphology and phagocytic capability	cells cultured in the presence of 15ng/ml IL-13 AND 0.1uM Rosiglitazone. Rosiglitazone must be added at the start of the culture as it is only effective on monocytes. The media should be changed at day 3 or day 4	None known; identified by morphology
PMNs	Neutrophils	0.6% Dextran T500, then Percoll gradient (70%, 80%), then negative selection with MACS anti-CCR3 + anti-mouse IgG Microbeads	CD62L	CD16 microbeads	CD66b, CD16

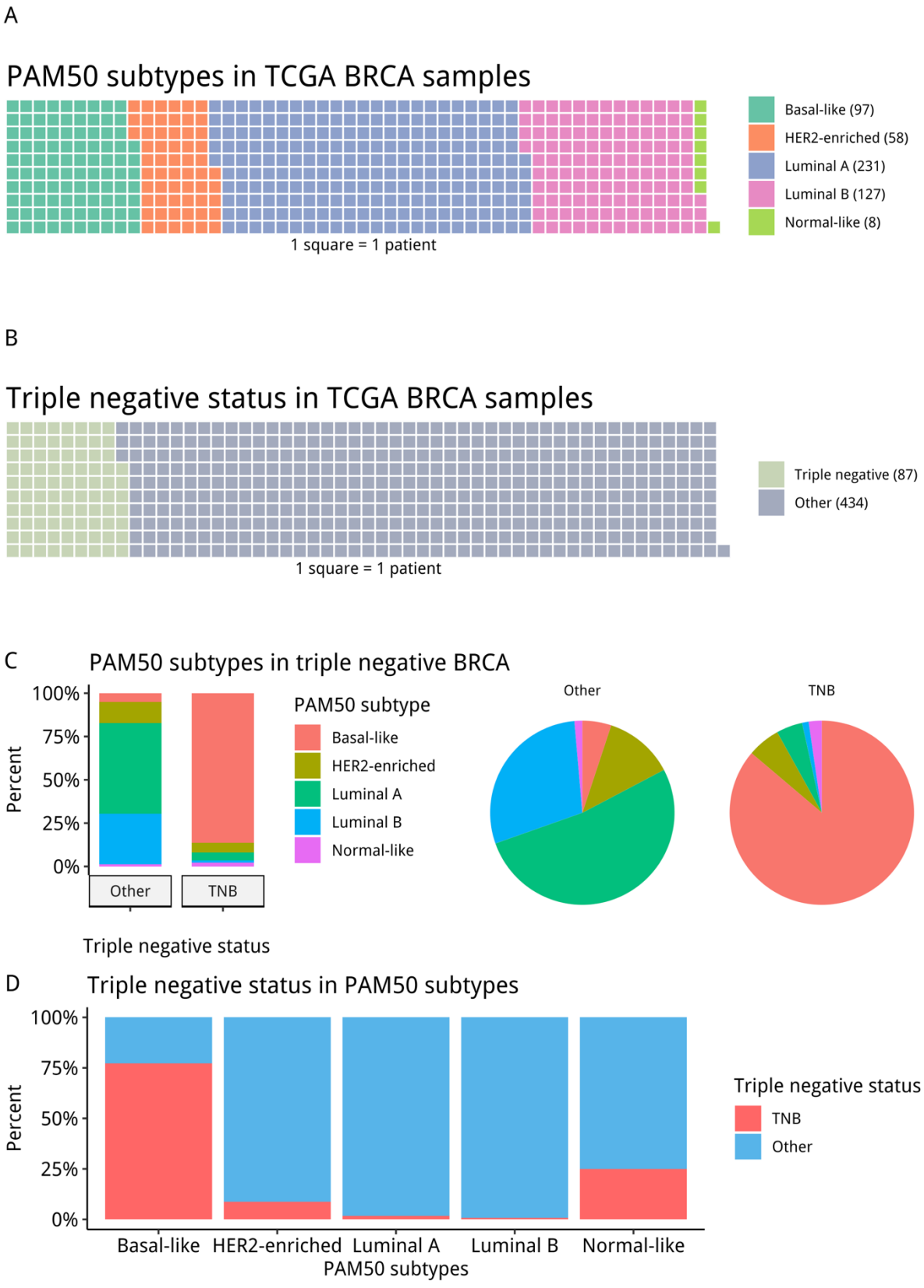
Comparison of cell separation method and markers CIBERSORT and xCell used to derived purified immune cell profiles. For cells that can be directly harvest from blood, they protocols were similar. However, for cells required *ex vivo* treatments, the protocols were divergent.

APPENDIX 2: HYPOXIA LEVEL IN DIFFERENT BREAST CANCER SUBTYPES



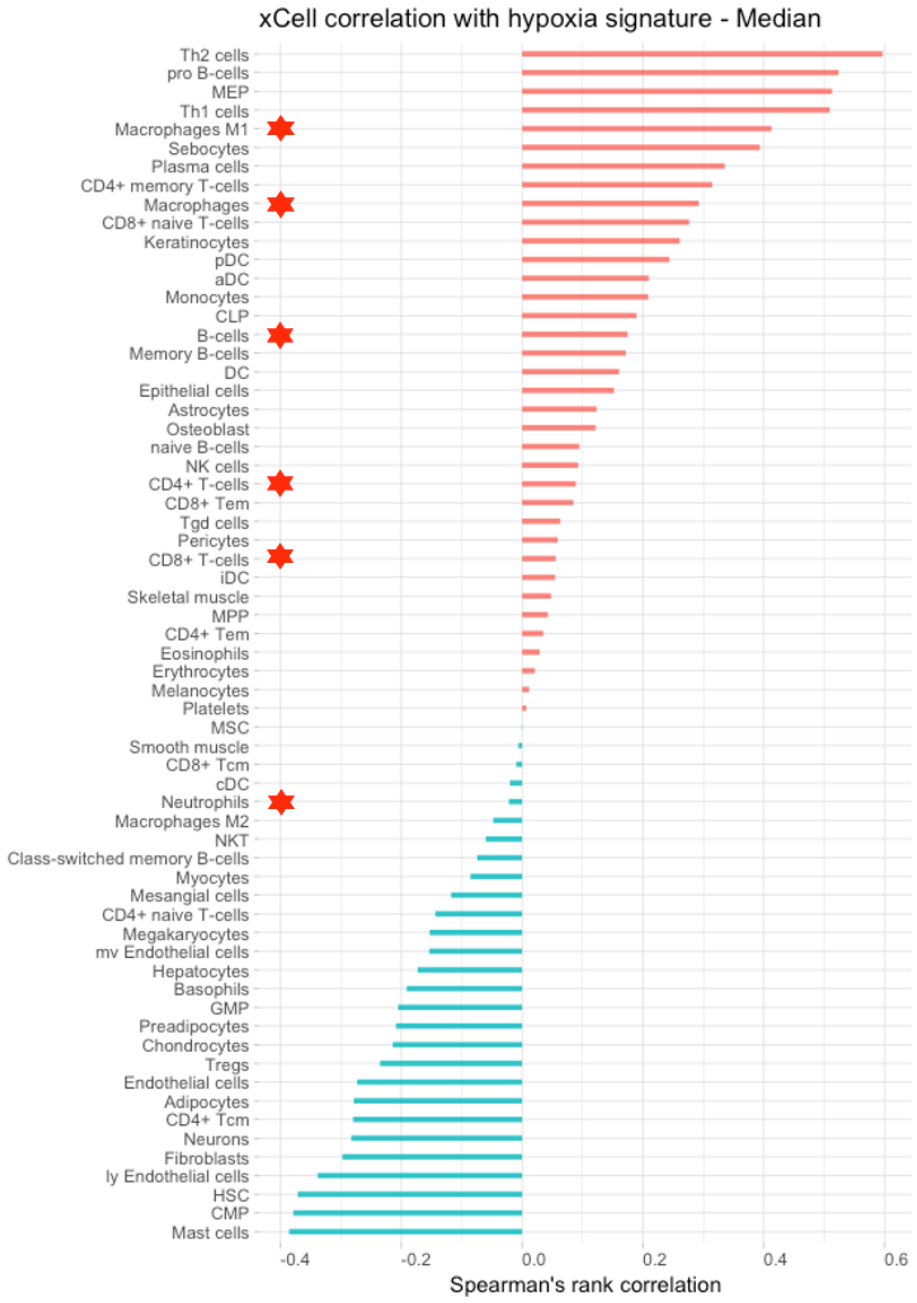
Relationships between hypoxia and different breast cancer classifications. A) PAM50 B) AJCC C) Metastasis D) Triple-negative. Symbol meaning: ns $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$

APPENDIX 3: RELATIONSHIP BETWEEN PAM50 SUBTYPES AND TRIPLE-NEGATIVE BREAST CANCER



Relationships between PAM50 and triple-negative subtypes in TCGA breast cancer

APPENDIX 4: HYPOXIA CORRELATION WITH 64 CELL TYPES IN BREAST CANCER



Relationship between hypoxia and 64 different cells within the breast cancer TME. Star signs labelled cells that have been previously validated in chapter 3.

8 REFERENCES

Aarts, A.A., Anderson, J.E., Anderson, C.J., Attridge, P.R., Attwood, A., Axt, J., Babel, M., Bahník, Š., Baranski, E., Barnett-Cowan, M., et al. (2015). Estimating the reproducibility of psychological science. Science (80-.). *349*. .

Abate, A.R., Chen, C.H., Agresti, J.J., and Weitz, D.A. (2009). Beating Poisson encapsulation statistics using close-packed ordering. Lab Chip *9*, 2628–2631. .

Abbas, A.R., Baldwin, D., Ma, Y., Ouyang, W., Gurney, A., Martin, F., Fong, S., van Lookeren Campagne, M., Godowski, P., Williams, P.M., et al. (2005). Immune response in silico (IRIS): Immune-specific genes identified from a compendium of microarray expression data. Genes Immun. *6*, 319–331. .

Abdi, H., and Williams, L.J. (2010). Principal component analysis. Wiley Interdiscip. Rev. Comput. Stat. *2*, 433–459. .

Abou Khouzam, R., Brodaczewska, K., Filipiak, A., Zeinelabdin, N.A., Buart, S., Szczylik, C., Kieda, C., and Chouaib, S. (2021). Tumor Hypoxia Regulates Immune Escape/Invasion: Influence on Angiogenesis and Potential Impact of Hypoxic Biomarkers on Cancer Therapies. Front. Immunol. *11*, 1–16. .

Adeegbe, D.O., and Nishikawa, H. (2013). Natural and induced T regulatory cells in cancer. Front. Immunol. *4*, 190. .

Ager, E.I., Kozin, S. V., Kirkpatrick, N.D., Seano, G., Kodack, D.P., Askoxylakis, V., Huang, Y., Goel, S., Snuderl, M., Muzikansky, A., et al. (2015). Blockade of MMP14 activity in murine breast carcinomas: Implications for macrophages, vessels, and radiotherapy. J. Natl. Cancer Inst. *107*, 1–12.

.

Van Allen, E.M., Miao, D., Schilling, B., Shukla, S.A., Blank, C., Zimmer, L., Sucker, A., Hillen, U., Foppen, M.H.G., Goldinger, S.M., et al. (2015). Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. Science *350*, 207–211. .

Altman, N.S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. Am. Stat. *46*, 175–185. .

Ammirante, M., Luo, J.-L., Grivennikov, S., Nedospasov, S., and Karin, M. (2010). B-cell-derived lymphotoxin promotes castration-resistant prostate cancer. Nature *464*, 302–305. .

Anand, P., Kunnumakara, A.B., Sundaram, C., Harikumar, K.B., Tharakan, S.T., Lai, O.S., Sung, B., and Aggarwal, B.B. (2008). Cancer is a preventable disease that requires major lifestyle changes. Pharm. Res. *25*, 2097–2116. .

Anderson, N.M., and Simon, M.C. (2020). The tumor microenvironment. Curr. Biol. *30*, R921–R925. .

Andrews, T.S., and Hemberg, M. (2018). Identifying cell populations with scRNASeq. Mol. Aspects Med. *59*, 114–122. .

Andzinski, L., Kasnitz, N., Stahnke, S., Wu, C.-F., Gereke, M., von Köckritz-Blickwede, M., Schilling, B., Brandau, S., Weiss, S., and Jablonska, J. (2016). Type I IFNs induce anti-tumor polarization of tumor associated neutrophils in mice and human. Int. J. Cancer *138*, 1982–1993. .

Anglesio, M.S., Papadopoulos, N., Ayhan, A., Nazeran, T.M., Noë, M., Horlings, H.M., Lum, A., Jones, S., Senz, J., Seckin, T., et al. (2017). Cancer-Associated Mutations in Endometriosis without Cancer. N. Engl. J. Med. *376*, 1835–1848. .

Aponte-López, A., Fuentes-Pananá, E.M., Cortes-Muñoz, D., and Muñoz-Cruz, S. (2018). Mast Cell, the Neglected Member of the Tumor Microenvironment: Role in Breast Cancer. J. Immunol. Res.

2018. .

Aran, D., Sirota, M., and Butte, A.J. (2015). Systematic pan-cancer analysis of tumour purity. Nat. Commun. *6*, 1–12. .

Aran, D., Hu, Z., and Butte, A.J. (2017). xCell: Digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. *18*, 1–14. .

Aras, S., and Raza Zaidi, M. (2017). TAMeless traitors: Macrophages in cancer progression and metastasis. Br. J. Cancer *117*, 1583–1591. .

Arlauckas, S.P., Garris, C.S., Kohler, R.H., Kitaoka, M., Cuccarese, M.F., Yang, K.S., Miller, M.A., Carlson, J.C., Freeman, G.J., Anthony, R.M., et al. (2017). In vivo imaging reveals a tumor-associated macrophage-mediated resistance pathway in anti-PD-1 therapy. Sci. Transl. Med. *9*. .

Arnaiz, E., and Harris, A.L. (2022). Role of Hypoxia in the Interferon Response. Front. Immunol. *13*, 1–14. .

Aubert, G., and Lansdorp, P.M. (2008). Telomeres and Aging. Physiol. Rev. *88*, 557–579. .

Auerbach, R., Arensman, R., Kubai, L., and Folkman, J. (1975). Tumor-induced angiogenesis: Lack of inhibition by irradiation. Int. J. Cancer *15*, 241–245. .

Azizi, E., Carr, A.J., Plitas, G., Cornish, A.E., Konopacki, C., Prabhakaran, S., Nainys, J., Wu, K., Kiseliovas, V., Setty, M., et al. (2018). Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. Cell *174*, 1293-1308.e36. .

Barbie, D.A., Tamayo, P., Boehm, J.S., Kim, S.Y., Moody, S.E., Dunn, I.F., Schinzel, A.C., Sandy, P., Meylan, E., Scholl, C., et al. (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature *462*, 108–112. .

Beatty, G.L., and Gladney, W.L. (2015). Immune escape mechanisms as a guide for cancer immunotherapy. Clin. Cancer Res. an Off. J. Am. Assoc. Cancer Res. *21*, 687–692. .

Belai, E.B., de Oliveira, C.E., Gasparoto, T.H., Ramos, R.N., Torres, S.A., Garlet, G.P., Cavassani,

K.A., Silva, J.S., and Campanelli, A.P. (2014). PD-1 blockage delays murine squamous cell carcinoma development. Carcinogenesis *35*, 424–431. .

Bergers, G., and Hanahan, D. (2008). Modes of resistance to anti-angiogenic therapy. Nat. Rev. Cancer *8*, 592–603. .

Van Den Beucken, T., Koch, E., Chu, K., Rupaimoole, R., Prickaerts, P., Adriaens, M., Voncken, J.W., Harris, A.L., Buffa, F.M., Haider, S., et al. (2014). Hypoxia promotes stem cell phenotypes and poor prognosis through epigenetic regulation of DICER. Nat. Commun. *5*. .

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.H., Pagès, F., Trajanoski, Z., and Galon, J. (2009). ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics *25*, 1091–1093. .

Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A.C., Angell, H., Fredriksen, T., Lafontaine, L., Berger, A., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity *39*, 782–795. .

Biswas, S.K., and Mantovani, A. (2010). Macrophage plasticity and interaction with lymphocyte subsets: Cancer as a paradigm. Nat. Immunol. *11*, 889–896. .

Blondel, V.D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. J. Stat. Mech. Theory Exp. *2008*, P10008. .

Bluestone, J.A. (1995). New perspectives of C1328-137-mediated T cell costimulation. Immunity *2*, 555–559. .

De Bock, K., Mazzone, M., and Carmeliet, P. (2011). Antiangiogenic therapy, hypoxia, and metastasis: risky liaisons, or not? Nat. Rev. Clin. Oncol. *8*, 393–404. .

Borrego, F., Ulbrecht, M., Weiss, E.H., Coligan, J.E., and Brooks, A.G. (1998). Recognition of human histocompatibility leukocyte antigen (HLA)-E complexed with HLA class I signal sequence–derived peptides by CD94/NKG2 confers protection from natural killer cell–mediated lysis. J. Exp.

Med. *187*, 813–818. .

Bosco, M.C., Reffo, G., Puppo, M., and Varesio, L. (2004). Hypoxia inhibits the expression of the CCR5 chemokine receptor in macrophages. Cell. Immunol. *228*, 1–7. .

Bowman, R.L., Klemm, F., Akkari, L., Pyonteck, S.M., Sevenich, L., Quail, D.F., Dhara, S., Simpson, K., Gardner, E.E., Iacobuzio-Donahue, C.A., et al. (2016). Macrophage Ontogeny Underlies Differences in Tumor-Specific Education in Brain Malignancies. Cell Rep. *17*, 2445–2459. .

Bradburn, M.J., Clark, T.G., Love, S.B., and Altman, D.G. (2003). Survival Analysis Part II: Multivariate data analysis- An introduction to concepts and methods. Br. J. Cancer *89*, 431–436. .

Brandau, S., Dumitru, C.A., and Lang, S. (2013). Protumor and antitumor functions of neutrophil granulocytes. Semin. Immunopathol. *35*, 163–176. .

Briggs, J.A., Weinreb, C., Wagner, D.E., Megason, S., Peshkin, L., Kirschner, M.W., and Klein, A.M. (2018). The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. Science (80-.). *360*. .

Bristow, R.G., and Hill, R.P. (2008). Hypoxia and metabolism. Hypoxia, DNA repair and genetic instability. Nat. Rev. Cancer *8*, 180–192. .

Brodbeck, T., Nehmann, N., Bethge, A., Wedemann, G., and Schumacher, U. (2014). Perforin-dependent direct cytotoxicity in natural killer cells induces considerable knockdown of spontaneous lung metastases and computer modelling-proven tumor cell dormancy in a HT29 human colon cancer xenograft mouse model. Mol. Cancer *13*, 244. .

Broz, M.L., Binnewies, M., Boldajipour, B., Nelson, A.E., Pollack, J.L., Erle, D.J., Barczak, A., Rosenblum, M.D., Daud, A., Barber, D.L., et al. (2014). Dissecting the tumor myeloid compartment reveals rare activating antigen-presenting cells critical for T cell immunity. Cancer Cell *26*, 638–652. .

Buffa, F.M., Harris, A.L., West, C.M., and Miller, C.J. (2010). Large meta-analysis of multiple

cancers reveals a common, compact and highly prognostic hypoxia metagene. Br. J. Cancer *102*, 428–435. .

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420. .

Calderon, B., Carrero, J.A., Ferris, S.T., Sojka, D.K., Moore, L., Epelman, S., Murphy, K.M., Yokoyama, W.M., Randolph, G.J., and Unanue, E.R. (2015). The pancreas anatomy conditions the origin and properties of resident macrophages. J. Exp. Med. *212*, 1497–1512. .

Campbell, M.J., Tonlaar, N.Y., Garwood, E.R., Huo, D., Moore, D.H., Khramtsov, A.I., Au, A., Baehner, F., Chen, Y., Malaka, D.O., et al. (2011). Proliferating macrophages associated with high grade, hormone receptor negative breast cancer and poor clinical outcome. Breast Cancer Res. Treat. *128*, 703–711. .

Cao, Y., Arbiser, J., D'Amato, R.J., D'Amore, P.A., Ingber, D.E., Kerbel, R., Klagsbrun, M., Lim, S., Moses, M.A., Zetter, B., et al. (2011). Forty-year journey of angiogenesis translational research. Sci. Transl. Med. *3*. .

Carlson, R.D., Flickinger, J.C., and Snook, A.E. (2020). Talkin' toxins: From Coley's to modern cancer immunotherapy. Toxins (Basel). *12*, 1–23. .

Carp, J. (2012). On the plurality of (methodological) worlds: estimating the analytic flexibility of FMRI experiments. Front. Neurosci. *6*, 149. .

Carrel, A., and Ebeling, A.H. (1926). THE FUNDAMENTAL PROPERTIES OF THE FIBROBLAST AND THE MACROPHAGE: I. THE FIBROBLAST. J. Exp. Med. *44*, 261. .

Carrel, A., and Ebeling, A.H. (1928). THE FUNDAMENTAL PROPERTIES OF THE FIBROBLAST AND THE MACROPHAGE : IV. THE MALIGNANT FIBROBLAST OF JENSEN

SARCOMA. J. Exp. Med. *48*, 285–298. .

Casazza, A., Laoui, D., Wenes, M., Rizzolio, S., Bassani, N., Mambretti, M., Deschoemaeker, S., Van Ginderachter, J.A., Tamagnone, L., and Mazzone, M. (2013). Impeding macrophage entry into hypoxic tumor areas by Sema3A/Nrp1 signaling blockade inhibits angiogenesis and restores antitumor immunity. Cancer Cell *24*, 695–709. .

Cassetta, L., Fragkogianni, S., Sims, A.H., Swierczak, A., Forrester, L.M., Zhang, H., Soong, D.Y.H., Cotechini, T., Anur, P., Lin, E.Y., et al. (2019). Human Tumor-Associated Macrophage and Monocyte Transcriptional Landscapes Reveal Cancer-Specific Reprogramming, Biomarkers, and Therapeutic Targets. Cancer Cell *35*, 588-602.e10. .

Catalán, R., Orozco-Morales, M., Hernández-Pedro, N.Y., Guijosa, A., Colín-González, A.L., Ávila-Moreno, F., and Arrieta, O. (2020). CD47-SIRPα Axis as a Biomarker and Therapeutic Target in Cancer: Current Perspectives and Future Challenges in Nonsmall Cell Lung Cancer. J. Immunol. Res. *2020*, 9435030. .

Cavallo, F., De Giovanni, C., Nanni, P., Forni, G., and Lollini, P.-L. (2011). 2011: the immune hallmarks of cancer. Cancer Immunol. Immunother. *60*, 319–326. .

Chambers, C.A., Kuhns, M.S., Egen, J.G., and Allison, J.P. (2001). CTLA-4-mediated inhibition in regulation of T cell responses: Mechanisms and manipulation in tumor immunotherapy. Annu. Rev. Immunol. *19*, 565–594. .

Charoentong, P., Angelova, M., Efremova, M., Gallasch, R., Hackl, H., Galon, J., and Trajanoski, Z. (2012). Bioinformatics for cancer immunology and immunotherapy. Cancer Immunol. Immunother. *61*, 1885–1903. .

Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., Hackl, H., and Trajanoski, Z. (2017). Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype

Relationships and Predictors of Response to Checkpoint Blockade. Cell Rep. *18*, 248–262. .

Chavey, C., Bibeau, F., Gourgou-Bourgade, S., Burlinchon, S., Boissière, F., Laune, D., Roques, S., and Lazennec, G. (2007). Oestrogen receptor negative breast cancers exhibit high cytokine content. Breast Cancer Res. *9*, R15. .

Chen, D.S., and Mellman, I. (2013). Oncology meets immunology: The cancer-immunity cycle. Immunity *39*, 1–10. .

Chen, B., Khodadoust, M.S., Liu, C.L., Newman, A.M., and Alizadeh, A.A. (2018). Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol. Biol. *1711*, 243–259. .

Chen, X., Iliopoulos, D., Zhang, Q., Tang, Q., Greenblatt, M.B., Hatziapostolou, M., Lim, E., Tam, W.L., Ni, M., Chen, Y., et al. (2014). XBP1 promotes triple-negative breast cancer by controlling the HIF1alpha pathway. Nature *508*, 103–107. .

Cheng, D.T., Prasad, M., Chekaluk, Y., Benayed, R., Sadowska, J., Zehir, A., Syed, A., Wang, Y.E., Somar, J., Li, Y., et al. (2017). Comprehensive detection of germline variants by MSK-IMPACT, a clinical diagnostic platform for solid tumor molecular oncology and concurrent cancer predisposition testing. BMC Med. Genomics *10*, 33. .

Cheng, S., Li, Z., Gao, R., Xing, B., Gao, Y., Yang, Y., Qin, S., Zhang, L., Ouyang, H., Du, P., et al. (2021). A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. Cell *184*, 792-809.e23. .

Chevrier, S., Levine, J.H., Zanotelli, V.R.T., Silina, K., Schulz, D., Bacac, M., Ries, C.H., Ailles, L., Jewett, M.A.S., Moch, H., et al. (2017). An Immune Atlas of Clear Cell Renal Cell Carcinoma. Cell *169*, 736-749.e18. .

Chiou, S.H., Sheu, B.C., Chang, W.C., Huang, S.C., and Hong-Nerng, H. (2005). Current concepts of tumor-infiltrating lymphocytes in human malignancies. J. Reprod. Immunol. *67*, 35–50. .

Chitu, V., and Stanley, E.R. (2006). Colony-stimulating factor-1 in immunity and inflammation. Curr.

Opin. Immunol. *18*, 39–48. .

Chiu, D.K.-C., Xu, I.M.-J., Lai, R.K.-H., Tse, A.P.-W., Wei, L.L., Koh, H.-Y., Li, L.L., Lee, D., Lo, R.C.-L., Wong, C.-M., et al. (2016). Hypoxia induces myeloid-derived suppressor cell recruitment to hepatocellular carcinoma through chemokine (C-C motif) ligand 26. Hepatology *64*, 797–813. .

Chung, N.C., and Storey, J.D. (2015). Statistical significance of variables driving systematic variation in high-dimensional data. Bioinformatics *31*, 545–554. .

Chung, W., Eum, H.H., Lee, H.O., Lee, K.M., Lee, H.B., Kim, K.T., Ryu, H.S., Kim, S., Lee, J.E., Park, Y.H., et al. (2017). Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. Nat. Commun. *8*, 1–12. .

Clark, R.A., and Klebanoff, S.J. (1975). Neutrophil-mediated tumor cell cytotoxicity: role of the peroxidase system. J. Exp. Med. *141*, 1442–1447. .

Clark, T.G., Bradburn, M.J., Love, S.B., and Altman, D.G. (2003). Survival Analysis Part I: Basic concepts and first analyses. Br. J. Cancer *89*, 232–238. .

Cochain, C., Vafadarnejad, E., Arampatzi, P., Pelisek, J., Winkels, H., Ley, K., Wolf, D., Saliba, A.E., and Zernecke, A. (2018). Single-cell RNA-seq reveals the transcriptional landscape and heterogeneity of aortic macrophages in murine atherosclerosis. Circ. Res. *122*, 1661–1674. .

Cohen, R., Hain, E., Buhard, O., Guilloux, A., Bardier, A., Kaci, R., Bertheau, P., Renaud, F., Bibeau, F., Fléjou, J.-F., et al. (2019). Association of Primary Resistance to Immune Checkpoint Inhibitors in Metastatic Colorectal Cancer With Misdiagnosis of Microsatellite Instability or Mismatch Repair Deficiency Status. JAMA Oncol. *5*, 551–555. .

Coley, W.B. (1991). The treatment of malignant tumors by repeated inoculations of erysipelas: With a report of ten original cases. Clin. Orthop. Relat. Res. *105*, 3–11. .

Colli, L.M., Machiela, M.J., Myers, T.A., Jessop, L., Yu, K., and Chanock, S.J. (2016). Burden of nonsynonymous mutations among TCGA cancers and candidate immune checkpoint inhibitor

responses. Cancer Res. *76*, 3767–3772. .

Collin, M., and Bigley, V. (2018). Human dendritic cell subsets: an update. Immunology *154*, 3–20. .

Condeelis, J., and Pollard, J.W. (2006). Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. Cell *124*, 263–266. .

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X., et al. (2016). A survey of best practices for RNA-seq data analysis. Genome Biol. *17*, 1–19. .

Contractor, N., Louten, J., Kim, L., Biron, C.A., and Kelsall, B.L. (2007). Cutting edge: Peyer's patch plasmacytoid dendritic cells (pDCs) produce low levels of type I interferons: possible role for IL-10, TGFbeta, and prostaglandin E2 in conditioning a unique mucosal pDC phenotype. J. Immunol. *179*, 2690–2694. .

Corbet, C., and Feron, O. (2017a). Tumour acidosis: From the passenger to the driver's seat. Nat. Rev. Cancer *17*, 577–593. .

Corbet, C., and Feron, O. (2017b). Cancer cell metabolism and mitochondria: Nutrient plasticity for TCA cycle fueling. Biochim. Biophys. Acta (BBA)-Reviews Cancer *1868*, 7–15. .

Creighton, C.J. (2018). Making use of cancer genomic databases. Curr. Protoc. Mol. Biol. *2018*, 19.14.1-19.14.13. .

Curiel, T.J., Coukos, G., Zou, L., Alvarez, X., Cheng, P., Mottram, P., Evdemon-Hogan, M., Conejo-Garcia, J.R., Zhang, L., and Burow, M. (2004). Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. Nat. Med. *10*, 942–949. .

Curtis, C., Shah, S.P., Chin, S.F., Turashvili, G., Rueda, O.M., Dunning, M.J., Speed, D., Lynch, A.G., Samarajiwa, S., Yuan, Y., et al. (2012). The genomic and transcriptomic architecture of 2,000

breast tumours reveals novel subgroups. Nature *486*, 346–352. .

Dang, V.D., Hilgenberg, E., Ries, S., Shen, P., and Fillatreau, S. (2014). From the regulatory functions of B cells to the identification of cytokine-producing plasma cell subsets. Curr. Opin. Immunol. *28*, 77–83. .

Darrell, C.G.A.R.N.T. source sites: D.U.M.S.M.R. 1 F.A. 2 B., 5, E.U.V.M.E.G. 3 4 5 B.D.J. 5 6 M.M.G. 3 O.J.J. 3 4, 8, H.F.H.M.T. 7 L.N., 11, M.D.A.C.C.A.K. 9 A.Y.W.K. 10 B.O., and 13, U. of C.S.F.V.S. 12 B.M. 13 P.M. (2008). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature *455*, 1061–1068. .

Demaria, S., Coleman, C.N., and Formenti, S.C. (2016). Radiotherapy: Changing the Game in Immunotherapy. Trends in Cancer *2*, 286–294. .

DeNardo, D.G., and Ruffell, B. (2019). Macrophages as regulators of tumour immunity and immunotherapy. Nat. Rev. Immunol. *19*, 369–382. .

Denning, T.L., Wang, Y., Patel, S.R., Williams, I.R., and Pulendran, B. (2007). Lamina propria macrophages and dendritic cells differentially induce regulatory and interleukin 17–producing T cell responses. Nat. Immunol. *8*, 1086–1094. .

Dhawan, A., Barberis, A., Cheng, W.C., Domingo, E., West, C., Maughan, T., Scott, J.G., Harris, A.L., and Buffa, F.M. (2019). Guidelines for using sigQC for systematic evaluation of gene signatures. Nat. Protoc. *14*, 1377–1400. .

Dieu-Nosjean, M.-C., Antoine, M., Danel, C., Heudes, D., Wislez, M., Poulot, V., Rabbe, N., Laurans, L., Tartour, E., de Chaisemartin, L., et al. (2008). Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures. J. Clin. Oncol. *26*, 4410–4417. .

Doedens, A.L., Stockmann, C., Rubinstein, M.P., Liao, D., Zhang, N., DeNardo, D.G., Coussens, L.M., Karin, M., Goldrath, A.W., and Johnson, R.S. (2010a). Macrophage expression of hypoxia-inducible factor-1 alpha suppresses T-cell function and promotes tumor progression. Cancer Res. *70*,

7465–7475. .

Doedens, A.L., Stockmann, C., Rubinstein, M.P., Liao, D., Zhang, N., DeNardo, D.G., Coussens, L.M., Karin, M., Goldrath, A.W., and Johnson, R.S. (2010b). Macrophage expression of hypoxia-inducible factor-1α suppresses T-cell function and promotes tumor progression. Cancer Res. *70*, 7465–7475. .

Doherty, J.R., and Cleveland, J.L. (2013). Targeting lactate metabolism for cancer therapeutics. J. Clin. Invest. *123*, 3685–3692. .

Dou, A., and Fang, J. (2021). Heterogeneous Myeloid Cells in Tumors. Cancers (Basel). *13*. .

Drake, C.G., Jaffee, E., and Pardoll, D.M. (2006). Mechanisms of immune evasion by tumors. Adv. Immunol. *90*, 51–81. .

Draoui, N., and Feron, O. (2011). Lactate shuttles at a glance: from physiological paradigms to anti-cancer treatments. Dis. Model. Mech. *4*, 727–732. .

Du, R., Lu, K. V, Petritsch, C., Liu, P., Ganss, R., Passegué, E., Song, H., Vandenberg, S., Johnson, R.S., Werb, Z., et al. (2008). HIF1alpha induces the recruitment of bone marrow-derived vascular modulatory cells to regulate tumor angiogenesis and invasion. Cancer Cell *13*, 206–220. .

Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J., Kaul, R., et al. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74. .

Duò, A., Robinson, M., and Soneson, C. (2018). A systematic performance evaluation of clustering methods for single-cell RNA-seq data. F1000Research *7*, 1141. .

Duraiswamy, J., Freeman, G.J., and Coukos, G. (2013). Therapeutic PD-1 Pathway Blockade Augments with Other Modalities of Immunotherapy T-Cell Function to Prevent Immune Decline in Ovarian CancerPD-1 Blockade Restores Immune Function in Ovarian Cancer. Cancer Res. *73*, 6900–

6912. .

Ebert, B.L., Firth, J.D., and Ratcliffe, P.J. (1995). Hypoxia and mitochondrial inhibitors regulate expression of glucose transporter-1 via distinct Cis-acting sequences. J. Biol. Chem. *270*, 29083–29089. .

Edwards, B.K., Ward, E., Kohler, B.A., Eheman, C., Zauber, A.G., Anderson, R.N., Jemal, A., Schymura, M.J., Lansdorp-Vogelaar, I., Seeff, L.C., et al. (2010). Annual report to the nation on the status of cancer, 1975-2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. Cancer *116*, 544–573. .

Eijk, M. van, and Aerts, J.M.F.G. (2021). The unique phenotype of lipid-laden macrophages. Int. J. Mol. Sci. *22*. .

Emami Nejad, A., Najafgholian, S., Rostami, A., Sistani, A., Shojaeifar, S., Esparvarinha, M., Nedaeinia, R., Haghjooy Javanmard, S., Taherian, M., Ahmadlou, M., et al. (2021). The role of hypoxia in the tumor microenvironment and development of cancer stem cell: a novel approach to developing treatment. Cancer Cell Int. *21*, 62. .

Epelman, S., Lavine, K.J., Beaudin, A.E., Sojka, D.K., Carrero, J.A., Calderon, B., Brija, T., Gautier, E.L., Ivanov, S., Satpathy, A.T., et al. (2014). Embryonic and adult-derived resident cardiac macrophages are maintained through distinct mechanisms at steady state and during inflammation. Immunity *40*, 91–104. .

Etxeberria, I., Olivera, I., Bolaños, E., Cirella, A., Teijeira, Á., Berraondo, P., and Melero, I. (2020). Engineering bionic T cells: signal 1, signal 2, signal 3, reprogramming and the removal of inhibitory mechanisms. Cell. Mol. Immunol. *17*, 576–586. .

Evans, S.M., and Koch, C.J. (2003). Prognostic significance of tumor oxygenation in humans. Cancer Lett. *195*, 1–16. .

Facciabene, A., Peng, X., Hagemann, I.S., Balint, K., Barchetti, A., Wang, L.-P., Gimotty, P.A.,

Gilks, C.B., Lal, P., Zhang, L., et al. (2011). Tumour hypoxia promotes tolerance and angiogenesis via CCL28 and T(reg) cells. Nature *475*, 226–230. .

Fan, X., Zhang, H., Cheng, Y., Jiang, X., Zhu, J., and Jin, T. (2016). Double Roles of Macrophages in Human Neuroimmune Diseases and Their Animal Models. Mediators Inflamm. *2016*, 8489251. .

Farrell, J.A., Wang, Y., Riesenfeld, S.J., Shekhar, K., Regev, A., and Schier, A.F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. Science (80-.). *360*. .

Fernández, J.M., de la Torre, V., Richardson, D., Royo, R., Puiggròs, M., Moncunill, V., Fragkogianni, S., Clarke, L., Flicek, P., Rico, D., et al. (2016). The BLUEPRINT Data Analysis Portal. Cell Syst. *3*, 491-495.e5. .

Ferrara, N., and Henzel, W.J. (1989). Pituitary follicular cells secrete a novel heparin-binding growth factor specific for vascular endothelial cells. Biochem. Biophys. Res. Commun. *161*, 851–858. .

Finotello, F., and Trajanoski, Z. (2018). Quantifying tumor-infiltrating immune cells from transcriptomics data. Cancer Immunol. Immunother. *67*, 1031–1040. .

Finotello, F., Mayer, C., Plattner, C., Laschober, G., Rieder, D., Hackl, H., Krogsdam, A., Posch, W., Wilflingseder, D., Sopper, S., et al. (2017). We introduce quanTIseq, a method to quantify the tumor immune contexture, determined by the type and density of tumor-infiltrating immune cells. quanTIseq is based on a novel deconvolution algorithm for RNA sequencing data that was validated with independ.

Firth, J.D., Ebert, B.L., and Ratcliffe, P.J. (1995). Hypoxic regulation of lactate dehydrogenase A: Interaction between hypoxia-inducible factor 1 and cAMP response elements. J. Biol. Chem. *270*, 21021–21027. .

Fischer, B., Müller, B., Fischer, K.G., Baur, N., and Kreutz, W. (2000). Acidic pH inhibits non-MHC-restricted killer cell functions. Clin. Immunol. *96*, 252–263. .

Fischer, K., Hoffmann, P., Voelkl, S., Meidenbauer, N., Ammer, J., Edinger, M., Gottfried, E.,

Schwarz, S., Rothe, G., and Hoves, S. (2007). Inhibitory effect of tumor cell–derived lactic acid on human T cells. Blood *109*, 3812–3819. .

Folkman, J. (1992). The role of angiogenesis in tumor growth. Semin. Cancer Biol. *3*, 65–71. .

Franklin, R.A., and Li, M.O. (2016). Ontogeny of Tumor-Associated Macrophages and Its Implication in Cancer Regulation. Trends in Cancer *2*, 20–34. .

Franklin, R.A., Liao, W., Sarkar, A., Kim, M. V., Bivona, M.R., Liu, K., Pamer, E.G., and Li, M.O. (2014). The cellular and molecular origin of tumor-associated macrophages. Science (80-.). *344*, 921–925. .

Fridlender, Z.G., Sun, J., Kim, S., Kapoor, V., Cheng, G., Ling, L., Worthen, G.S., and Albelda, S.M. (2009). Polarization of tumor-associated neutrophil phenotype by TGF-beta: "N1" versus "N2" TAN. Cancer Cell *16*, 183–194. .

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. J. Stat. Softw. *33*, 1–22. .

Fu, Z., Mowday, A.M., Smaill, J.B., Hermans, I.F., and Patterson, A. V (2021). Tumour Hypoxia-Mediated Immunosuppression: Mechanisms and Therapeutic Approaches to Improve Cancer Immunotherapy. Cells *10*. .

Gabrilovich, D. (2004). Mechanisms and functional significance of tumour-induced dendritic-cell defects. Nat. Rev. Immunol. *4*, 941–952. .

Gajewski, T.F., Meng, Y., and Harlin, H. (2006). Immune suppression in the tumor microenvironment. J. Immunother. *29*, 233–240. .

Galon, J., and Bruni, D. (2019). Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. Nat. Rev. Drug Discov. *18*, 197–218. .

Galon, J., Costes, A., Sanchez-Cabo, F., Kirilovsky, A., Mlecnik, B., Lagorce-Pagès, C., Tosolini, M., Camus, M., Berger, A., Wind, P., et al. (2006). Type, density, and location of immune cells within

human colorectal tumors predict clinical outcome. Science (80-.). *313*, 1960–1964. .

van de Garde, M.D.B., Martinez, F.O., Melgert, B.N., Hylkema, M.N., Jonkers, R.E., and Hamann, J. (2014). Chronic Exposure to Glucocorticoids Shapes Gene Expression and Modulates Innate and Adaptive Activation Pathways in Macrophages with Distinct Changes in Leukocyte Attraction. J. Immunol. *192*, 1196–1208. .

Garziera, M., Scarabel, L., and Toffoli, G. (2017). Hypoxic Modulation of HLA-G Expression through the Metabolic Sensor HIF-1 in Human Cancer Cells. J. Immunol. Res. *2017*, 4587520. .

Gaustad, J.V., Simonsen, T.G., Andersen, L.M.K., and Rofstad, E.K. (2017). Vascular abnormalities and development of hypoxia in microscopic melanoma xenografts. J. Transl. Med. *15*, 1–10. .

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., et al. (2015). The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat. Med. *21*, 938–945. .

Georgoudaki, A.M., Prokopec, K.E., Boura, V.F., Hellqvist, E., Sohn, S., Östling, J., Dahan, R., Harris, R.A., Rantalainen, M., Klevebring, D., et al. (2016). Reprogramming Tumor-Associated Macrophages by Antibody Targeting Inhibits Cancer Progression and Metastasis. Cell Rep. *15*, 2000–2011. .

Germain, C., Gnjatic, S., Tamzalit, F., Knockaert, S., Remark, R., Goc, J., Lepelley, A., Becht, E., Katsahian, S., Bizouard, G., et al. (2014). Presence of B cells in tertiary lymphoid structures is associated with a protective immunity in patients with lung cancer. Am. J. Respir. Crit. Care Med. *189*, 832–844. .

Gibaldi, D., Vilar-Pereira, G., Pereira, I.R., Silva, A.A., Barrios, L.C., Ramos, I.P., Mata dos Santos, H.A., Gazzinelli, R., and Lannes-Vieira, J. (2020). CCL3/Macrophage Inflammatory Protein-1α Is Dually Involved in Parasite Persistence and Induction of a TNF- and IFNγ-Enriched Inflammatory

Milieu in Trypanosoma cruzi-Induced Chronic Cardiomyopathy. Front. Immunol. *11*, 1–21. .

Gibbings, S.L., Goyal, R., Desch, A.N., Leach, S.M., Prabagar, M., Atif, S.M., Bratton, D.L., Janssen, W., and Jakubzick, C. V (2015). Transcriptome analysis highlights the conserved difference between embryonic and postnatal-derived alveolar macrophages. Blood *126*, 1357–1366. .

Gilbert, A.E., Karagiannis, P., Dodev, T., Koers, A., Lacy, K., Josephs, D.H., Takhar, P., Geh, J.L.C., Healy, C., Harries, M., et al. (2011). Monitoring the systemic human memory B cell compartment of melanoma patients for anti-tumor IgG antibodies. PLoS One *6*, e19330. .

Gimbrone, M.A., Leapman, S.B., Cotran, R.S., and Folkman, J. (1972). Tumor dormancy in vivo by prevention of neovascularization. J. Exp. Med. *136*, 261–276. .

Ginhoux, F., Greter, M., Leboeuf, M., Nandi, S., See, P., Gokhan, S., Mehler, M.F., Conway, S.J., Ng, L.G., Stanley, E.R., et al. (2010). Fate Mapping Analysis Reveals That Adult Microglia Derive from Primitive Macrophages. Science (80-.). *330*, 841–845. .

Giraudo, E., Inoue, M., and Hanahan, D. (2004). An amino-bisphosphonate targets MMP-9-expressing macrophages and angiogenesis to impair cervical carcinogenesis. J. Clin. Invest. *114*, 623–633. .

Goc, J., Germain, C., Vo-Bourgais, T.K.D., Lupo, A., Klein, C., Knockaert, S., de Chaisemartin, L., Ouakrim, H., Becht, E., Alifano, M., et al. (2014). Dendritic cells in tumor-associated tertiary lymphoid structures signal a Th1 cytotoxic immune contexture and license the positive prognostic value of infiltrating CD8+ T cells. Cancer Res. *74*, 705–715. .

Gomez Perdiguero, E., Klapproth, K., Schulz, C., Busch, K., Azzoni, E., Crozet, L., Garner, H., Trouillet, C., De Bruijn, M.F., Geissmann, F., et al. (2015). Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. Nature *518*, 547–551. .

Goossens, P., Rodriguez-Vita, J., Etzerodt, A., Masse, M., Rastoin, O., Gouirand, V., Ulas, T., Papantonopoulou, O., Van Eck, M., Auphan-Anezin, N., et al. (2019). Membrane Cholesterol Efflux

Drives Tumor-Associated Macrophage Reprogramming and Tumor Progression. Cell Metab. *29*, 1376-1389.e4. .

Gopalakrishnan, V., Spencer, C.N., Nezi, L., Reuben, A., Andrews, M.C., Karpinets, T. V., Prieto, P.A., Vicente, D., Hoffman, K., Wei, S.C., et al. (2018). Gut microbiome modulates response to anti – PD-1 immunotherapy in melanoma patients. Science (80-.). *359*, 97–103. .

Gordon, S., and Martinez-Pomares, L. (2017). Physiological roles of macrophages. Pflugers Arch. Eur. J. Physiol. *469*, 365–374. .

Grabski, I.N., Street, K., Irizarry, R.A., and Isabella N. Grabski, Kelly Street, R.A.I. (2022). Significance Analysis for Clustering with Single-Cell RNA-Sequencing Data. BioRxiv *37*, 2022.08.01.502383. .

Granot, Z., Henke, E., Comen, E.A., King, T.A., Norton, L., and Benezra, R. (2011). Tumor entrained neutrophils inhibit seeding in the premetastatic lung. Cancer Cell *20*, 300–314. .

Greaves, P., and Gribben, J.G. (2013). The role of B7 family molecules in hematologic malignancy. Blood, J. Am. Soc. Hematol. *121*, 734–744. .

Guillerey, C., Huntington, N.D., and Smyth, M.J. (2016). Targeting natural killer cells in cancer immunotherapy. Nat. Immunol. *17*, 1025–1036. .

Gupta, S., Gupta, R., Sinha, D.N., and Mehrotra, R. (2018). Relationship between type of smokeless tobacco & risk of cancer: A systematic review. Indian J. Med. Res. Suppl. *148*, 56–76. .

Guthrie, G.J.K., Charles, K.A., Roxburgh, C.S.D., Horgan, P.G., McMillan, D.C., and Clarke, S.J. (2013). The systemic inflammation-based neutrophil-lymphocyte ratio: experience in patients with cancer. Crit. Rev. Oncol. Hematol. *88*, 218–230. .

Gwak, J.M., Jang, M.H., Kim, D. Il, Seo, A.N., and Park, S.Y. (2015). Prognostic value of tumor-associated macrophages according to histologic locations and hormone receptor status in breast

cancer. PLoS One *10*, e0125728. .

Hackl, H., Charoentong, P., Finotello, F., and Trajanoski, Z. (2016). Computational genomics tools for dissecting tumour-immune cell interactions. Nat. Rev. Genet. *17*, 441–458. .

Hanahan, D. (2022). Hallmarks of Cancer: New Dimensions. Cancer Discov. *12*, 31–46. .

Hanahan, D., and Weinberg, R. (2000). The Hallmarks of Cancer. Cell *100*, 57–70. .

Hanahan, D., and Weinberg, R. a (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674. .

Handgretinger, R., Lang, P., and André, M.C. (2016). Exploitation of natural killer cells for the treatment of acute leukemia. Blood *127*, 3341–3349. .

Hanes, M.R., Giacomantonio, C.A., and Marshall, J.S. (2021). Mast cells and skin and breast cancers: A complicated and microenvironment-dependent role. Cells *10*. .

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics *14*, 1–15. .

Harris, A.L. (2002). Hypoxia--a key regulatory factor in tumour growth. Nat. Rev. Cancer *2*, 38–47. .

Harris, R.J., Cheung, A., Ng, J.C.F., Laddach, R., Chenoweth, A.M., Crescioli, S., Fittall, M., Dominguez-Rodriguez, D., Roberts, J., Levi, D., et al. (2021). Tumor-infiltrating B lymphocyte profiling identifies IgG-biased, clonally expanded prognostic phenotypes in triple-negative breast cancer. Cancer Res. *81*, 4290–4304. .

Hartigan, J.A., and Wong, M.A. (1979). Algorithm AS 136: A k-means clustering algorithm. J. R. Stat. Soc. Ser. c (Applied Stat. *28*, 100–108. .

Hashimoto, D., Chow, A., Noizat, C., Teo, P., Beasley, M.B., Leboeuf, M., Becker, C.D., See, P., Price, J., Lucas, D., et al. (2013). Tissue-Resident Macrophages Self-Maintain Locally throughout

Adult Life with Minimal Contribution from Circulating Monocytes. Immunity *38*, 792–804. .

Hashimshony, T., Wagner, F., Sher, N., and Yanai, I. (2012). CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. Cell Rep. *2*, 666–673. .

Haskill, S., Becker, S., Fowler, W., and Walton, L. (1982). Mononuclear-cell infiltration in ovarian cancer. I. Inflammatory-cell infiltrates from tumour and ascites material. Br. J. Cancer *45*, 728–736. .

Heddleston, J.M., Li, Z., Lathia, J.D., Bao, S., Hjelmeland, a B., and Rich, J.N. (2010). Hypoxia inducible factors in cancer stem cells. Br. J. Cancer *102*, 789–795. .

Hegde, P.S., Karanikas, V., and Evers, S. (2016). The Where, the When, and the How of Immune Monitoring for Cancer Immunotherapies in the Era of Checkpoint Inhibition. Clin. Cancer Res. an Off. J. Am. Assoc. Cancer Res. *22*, 1865–1874. .

Henze, A.-T., and Mazzone, M. (2016a). The impact of hypoxia on tumor-associated macrophages. J. Clin. Invest. *126*, 3672–3679. .

Henze, A., and Mazzone, M. (2016b). The Impact of Hypoxia on TAMS. *126*, 3672–3679. .

Hersh, E.M., Mavligit, G.M., Gutterman, J.U., and Barsales, P.B. (1976). Mononuclear cell content of human solid tumors. Med. Pediatr. Oncol. *2*, 1–9. .

Ho, W.J., Jaffee, E.M., and Zheng, L. (2020). The tumour microenvironment in pancreatic cancer — clinical challenges and opportunities. Nat. Rev. Clin. Oncol. *17*, 527–540. .

Hockel, M., and Vaupel, P. (2001). Tumor hypoxia: definitions and current clinical, biologic, and molecular aspects. J. Natl. Cancer Inst. *93*, 266–276. .

Hoeffel, G., Chen, J., Lavin, Y., Low, D., Almeida, F.F., See, P., Beaudin, A.E., Lum, J., Low, I., Forsberg, E.C., et al. (2015). C-Myb(+) erythro-myeloid progenitor-derived fetal monocytes give rise to adult tissue-resident macrophages. Immunity *42*, 665–678. .

Hothorn, T., and Lausen, B. (2003). On the exact distribution of maximally selected rank statistics.

Comput. Stat. Data Anal. *43*, 121–137. .

Hou, W., Ji, Z., Ji, H., and Hicks, S. (2020). A Systematic Evaluation of Single-cell RNA-sequencing Imputation Methods. 1–30. .

Hourani, T., Holden, J.A., Li, W., Lenzo, J.C., Hadjigol, S., and O'Brien-Simpson, N.M. (2021). Tumor Associated Macrophages: Origin, Recruitment, Phenotypic Diversity, and Targeting. Front. Oncol. *11*, 1–20. .

Huang, A.C., Postow, M.A., Orlowski, R.J., Mick, R., Bengsch, B., Manne, S., Xu, W., Harmon, S., Giles, J.R., Wenz, B., et al. (2017). T-cell invigoration to tumour burden ratio associated with anti-PD-1 response. Nature *545*, 60–65. .

Huang, L.E., Gu, J., Schau, M., and Bunn, H.F. (1998). Regulation of hypoxia-inducible factor 1alpha is mediated by an O2-dependent degradation domain via the ubiquitin-proteasome pathway. Proc. Natl. Acad. Sci. U. S. A. *95*, 7987–7992. .

Hudgins, A.D., Tazearslan, C., Tare, A., Zhu, Y., Huffman, D., and Suh, Y. (2018). Age- and tissue-specific expression of senescence biomarkers in mice. Front. Genet. *9*, 1–6. .

Hudson, T.J., Anderson, W., Artez, A., Barker, A.D., Bell, C., Bernabé, R.R., Bhan, M.K., Calvo, F., Eerola, I., Gerhard, D.S., et al. (2010). International network of cancer genome projects. Nature *464*, 993–998. .

Hutter, C., and Zenklusen, J.C. (2018). The Cancer Genome Atlas: Creating Lasting Value beyond Its Data. Cell *173*, 283–285. .

Inoue, S., Leitner, W.W., Golding, B., and Scott, D. (2006). Inhibitory effects of B cells on antitumor immunity. Cancer Res. *66*, 7741–7747. .

Ito, T., Yang, M., Wang, Y.-H., Lande, R., Gregorio, J., Perng, O.A., Qin, X.-F., Liu, Y.-J., and Gilliet, M. (2007). Plasmacytoid dendritic cells prime IL-10-producing T regulatory cells by

inducible costimulator ligand. J. Exp. Med. *204*, 105–115. .

Jackson, P.A., Green, M.A., Marks, C.G., King, R.J.B., Hubbard, R., and Cook, M.G. (1996). Lymphocyte subset infiltration patterns and HLA antigen status in colorectal carcinomas and adenomas. Gut *38*, 85–89. .

Jacobs, N., Langers, Renoux, Thiry, and Delvenne (2012). Natural killer cells: role in local tumor growth and metastasis. Biol. Targets Ther. *6*, 73. .

Jamieson, T., Clarke, M., Steele, C.W., Samuel, M.S., Neumann, J., Jung, A., Huels, D., Olson, M.F., Das, S., Nibbs, R.J.B., et al. (2012). Inhibition of CXCR2 profoundly suppresses inflammation-driven and spontaneous tumorigenesis. J. Clin. Invest. *122*, 3127–3144. .

Jiang, S., Yang, Y., Fang, M., Li, X., Yuan, X., and Yuan, J. (2016). Co-evolution of tumor-associated macrophages and tumor neo-vessels during cervical cancer invasion. Oncol. Lett. *12*, 2625–2631. .

Jimenez-Sanchez, A., Cast, O., and Miller, M.L. (2019). Comprehensive benchmarking and integration of tumor microenvironment cell estimation methods. Cancer Res. *79*, 6238–6246. .

Jin, M.Z., and Jin, W.L. (2020). The updated landscape of tumor microenvironment and drug repurposing. Signal Transduct. Target. Ther. *5*. .

Johansen, J.S., Christensen, I.J., Riisbro, R., Greenall, M., Han, C., Price, P.A., Smith, K., Brünner, N., and Harris, A.L. (2003). High serum YKL-40 levels in patients with primary breast cancer is related to short recurrence free survival. Breast Cancer Res. Treat. *80*, 15–21. .

Johnson, K.S., Conant, E.F., and Soo, M.S. (2021). Molecular Subtypes of Breast Cancer: A Review for Breast Radiologists. J. Breast Imaging *3*, 12–24. .

Kaluz, S., Kaluzová, M., Liao, S.Y., Lerman, M., and Stanbridge, E.J. (2009). Transcriptional control of the tumor- and hypoxia-marker carbonic anhydrase 9: A one transcription factor (HIF-1) show? Biochim. Biophys. Acta - Rev. Cancer *1795*, 162–172. .

Karn, T., Jiang, T., Hatzis, C., Sänger, N., El-Balat, A., Rody, A., Holtrich, U., Becker, S., Bianchini,

G., and Pusztai, L. (2017). Association between genomic metrics and immune infiltration in triple-negative breast cancer. JAMA Oncol. *3*, E1–E5. .

Kawasaki, E.S., and Ladner, M.B. (1990). Molecular biology of macrophage colony-stimulating factor. Immunol. Ser. *49*, 155–176. .

Kennedy, S.R., Loeb, L.A., and Herr, A.J. (2012). Somatic mutations in aging, cancer and neurodegeneration. Mech. Ageing Dev. *133*, 118–126. .

Keum, N.N., and Giovannucci, E. (2019). Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. Nat. Rev. Gastroenterol. Hepatol. *16*, 713–732. .

Kiselev, V.Y., Andrews, T.S., and Hemberg, M. (2019). Challenges in unsupervised clustering of single-cell RNA-seq data. Nat. Rev. Genet. *20*, 273–282. .

Kishore, J., Goel, M., and Khanna, P. (2010). Understanding survival analysis: Kaplan-Meier estimate. Int. J. Ayurveda Res. *1*, 274. .

Kiss, M., Van Gassen, S., Movahedi, K., Saeys, Y., and Laoui, D. (2018). Myeloid cell heterogeneity in cancer: not a single cell alike. Cell. Immunol. *330*, 188–201. .

Kitamura, T., Qian, B.Z., Soong, D., Cassetta, L., Noy, R., Sugano, G., Kato, Y., Li, J., and Pollard, J.W. (2015). CCL2-induced chemokine cascade promotes breast cancer metastasis by enhancing retention of metastasis-associated macrophages. J. Exp. Med. *212*, 1043–1059. .

Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell *161*, 1187–1201. .

Koboldt, D.C., Fulton, R.S., McLellan, M.D., Schmidt, H., Kalicki-Veizer, J., McMichael, J.F., Fulton, L.L., Dooling, D.J., Ding, L., Mardis, E.R., et al. (2012). Comprehensive molecular portraits of human breast tumours. Nature *490*, 61–70. .

Komohara, Y., Hasita, H., Ohnishi, K., Fujiwara, Y., Suzu, S., Eto, M., and Takeya, M. (2011).

Macrophage infiltration and its prognostic relevance in clear cell renal cell carcinoma. Cancer Sci. *102*, 1424–1431. .

Konishi, H., Komura, D., Katoh, H., Atsumi, S., Koda, H., Yamamoto, A., Seto, Y., Fukayama, M., Yamaguchi, R., Imoto, S., et al. (2019). Capturing the differences between humoral immunity in the normal and tumor environments from repertoire-seq of B-cell receptors using supervised machine learning. BMC Bioinformatics *20*, 267. .

Kourou, K., Exarchos, T.P., Exarchos, K.P., Karamouzis, M. V, and Fotiadis, D.I. (2015). Machine learning applications in cancer prognosis and prediction. Comput. Struct. Biotechnol. J. *13*, 8–17. .

Kren, L., Muckova, K., Lzicarova, E., Sova, M., Vybihal, V., Svoboda, T., Fadrus, P., Smrcka, M., Slaby, O., and Lakomy, R. (2010). Production of immune-modulatory nonclassical molecules HLA-G and HLA-E by tumor infiltrating ameboid microglia/macrophages in glioblastomas: a role in innate immunity? J. Neuroimmunol. *220*, 131–135. .

Krock, B.L., Skuli, N., and Simon, M.C. (2011). Hypoxia-induced angiogenesis: good and evil. Genes Cancer *2*, 1117–1133. .

Kroemer, G., Galluzzi, L., Kepp, O., and Zitvogel, L. (2013). Immunogenic cell death in cancer therapy. Annu. Rev. Immunol. *31*, 51–72. .

Kroemer, G., Senovilla, L., Galluzzi, L., André, F., and Zitvogel, L. (2015). Natural and therapy-induced immunosurveillance in breast cancer. Nat. Med. *21*, 1128–1138. .

Kuang, D.-M., Zhao, Q., Peng, C., Xu, J., Zhang, J.-P., Wu, C., and Zheng, L. (2009). Activated monocytes in peritumoral stroma of hepatocellular carcinoma foster immune privilege and disease progression through PD-L1. J. Exp. Med. *206*, 1327–1337. .

Kucerova, P., and Cervinkova, M. (2016). Spontaneous regression of tumour and the role of microbial infection - possibilities for cancer treatment. Anticancer. Drugs *27*, 269–277. .

Kurai, J., Chikumi, H., Hashimoto, K., Yamaguchi, K., Yamasaki, A., Sako, T., Touge, H., Makino,

H., Takata, M., Miyata, M., et al. (2007). Antibody-dependent cellular cytotoxicity mediated by cetuximab against lung cancer cell lines. Clin. Cancer Res. an Off. J. Am. Assoc. Cancer Res. *13*, 1552–1561. .

Labiano, S., Palazon, A., and Melero, I. (2015). Immune response regulation in the tumor microenvironment by hypoxia. Semin. Oncol. *42*, 378–386. .

LaGory, E.L., and Giaccia, A.J. (2016). The ever-expanding role of HIF in tumour and stromal biology. Nat. Cell Biol. *18*, 356–365. .

Lahat, N., Bitterman, H., Engelmayer-Goren, M., Rosenzweig, D., Weiss-Cerem, L., and Rahat, M. a (2011). Reduced TIMP-2 in hypoxia enhances angiogenesis. Am. J. Physiol. Cell Physiol. *300*, 557–566. .

Laitala, A., and Erler, J.T. (2018). Hypoxic Signalling in Tumour Stroma. *8*, 1–13. .

Lambrechts, D., Wauters, E., Boeckx, B., Aibar, S., Nittner, D., Burton, O., Bassez, A., Decaluwé, H., Pircher, A., Van den Eynde, K., et al. (2018). Phenotype molding of stromal cells in the lung tumor microenvironment. Nat. Med. *24*, 1277–1289. .

Langer, R., Conn, H., Vacanti, J., Haudenschild, C., and Folkman, J. (1980). Control of tumor growth in animals by infusion of an angiogenesis inhibitor. Proc. Natl. Acad. Sci. U. S. A. *77*, 4331–4335. .

Laoui, D., Movahedi, K., Van Overmeire, E., Van den Bossche, J., Schouppe, E., Mommer, C., Nikolaou, A., Morias, Y., De Baetselier, P., and Van Ginderachter, J.A. (2011). Tumor-associated macrophages in breast cancer: distinct subsets, distinct functions. Int. J. Dev. Biol. *55*, 861–867. .

Laoui, D., Keirsse, J., Morias, Y., Van Overmeire, E., Geeraerts, X., Elkrim, Y., Kiss, M., Bolli, E., Lahmar, Q., Sichien, D., et al. (2016). The tumour microenvironment harbours ontogenically distinct dendritic cell populations with opposing effects on tumour immunity. Nat. Commun. *7*, 13720. .

Lauss, M., Donia, M., Harbst, K., Andersen, R., Mitra, S., Rosengren, F., Salim, M., Vallon-Christersson, J., Törngren, T., Kvist, A., et al. (2017). Mutational and putative neoantigen load predict

clinical benefit of adoptive T cell therapy in melanoma. Nat. Commun. *8*, 1738. .

Le, T., Aronow, R.A., Kirshtein, A., and Shahriyari, L. (2021). A review of digital cytometry methods: estimating the relative abundance of cell types in a bulk of cells. Brief. Bioinform. *22*. .

Leahy, D.J., Axel, R., and Hendrickson, W.A. (1992). Crystal structure of a soluble form of the human T cell coreceptor CD8 at 2.6 A resolution. Cell *68*, 1145–1162. .

Lee, S.H., and Nam, H.S. (2008). TNFα-induced down-regulation of estrogen receptor α in MCF-7 breast cancer cells. Mol. Cells *26*, 285–290. .

Lei, X., Lei, Y., Li, J.-K., Du, W.-X., Li, R.-G., Yang, J., Li, J., Li, F., and Tan, H.-B. (2020). Immune cells within the tumor microenvironment: Biological functions and roles in cancer immunotherapy. Cancer Lett. *470*, 126–133. .

Lenzo, J.C., Turner, A.L., Cook, A.D., Vlahos, R., Anderson, G.P., Reynolds, E.C., and Hamilton, J.A. (2012). Control of macrophage lineage populations by CSF-1 receptor and GM-CSF in homeostasis and inflammation. Immunol. Cell Biol. *90*, 429–440. .

Li, A., Walling, J., Ahn, S., Kotliarov, Y., Su, Q., Quezado, M., Oberholtzer, J.C., Park, J., Zenklusen, J.C., and Fine, H.A. (2009). Unsupervised analysis of transcriptomic profiles reveals six glioma subtypes. Cancer Res. *69*, 2091–2099. .

Li, B., Li, T., Pignon, J.-C., Wang, B., Wang, J., Shukla, S.A., Dou, R., Chen, Q., Hodi, F.S., Choueiri, T.K., et al. (2016a). Landscape of tumor-infiltrating T cell repertoire of human cancers. Nat. Genet. *48*, 725–732. .

Li, B., Severson, E., Pignon, J.C., Zhao, H., Li, T., Novak, J., Jiang, P., Shen, H., Aster, J.C., Rodig, S., et al. (2016b). Comprehensive analyses of tumor immunity: Implications for cancer immunotherapy. Genome Biol. *17*, 1–16. .

Li, B., Liu, J.S., and Liu, X.S. (2017a). Revisit linear regression-based deconvolution methods for

tumor gene expression data. Genome Biol. *18*, 1–5. .

Li, J., Diao, B., Guo, S., Huang, X., Yang, C., Feng, Z., Yan, W., Ning, Q., Zheng, L., Chen, Y., et al. (2017b). VSIG4 inhibits proinflammatory macrophage activation by reprogramming mitochondrial pyruvate metabolism. Nat. Commun. *8*, 1322. .

Li, J.L., Sainson, R.C.A., Oon, C.E., Turley, H., Leek, R., Sheldon, H., Bridges, E., Shi, W., Snell, C., Bowden, E.T., et al. (2011a). DLL4-Notch signaling mediates tumor resistance to anti-VEGF therapy in vivo. Cancer Res. *71*, 6073–6083. .

Li, K., Luo, H., Huang, L., Luo, H., and Zhu, X. (2020). Microsatellite instability: a review of what the oncologist should know. Cancer Cell Int. *20*, 16. .

Li, Y., Tu, C., Wang, H., Silverman, D.N., and Frost, S.C. (2011b). Catalysis and pH control by membrane-associated carbonic anhydrase IX in MDA-MB-231 breast cancer cells. J. Biol. Chem. *286*, 15789–15796. .

Li, Y., Patel, S.P., Roszik, J., and Qin, Y. (2018). Hypoxia-driven immunosuppressive metabolites in the tumor microenvironment: new approaches for combinational immunotherapy. Front. Immunol. *9*, 1591. .

Liao, D., and Johnson, R.S. (2007). Hypoxia: a key regulator of angiogenesis in cancer. Cancer Metastasis Rev. *26*, 281–290. .

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database Hallmark Gene Set Collection. Cell Syst. *1*, 417–425. .

Lin, E.Y., and Pollard, J.W. (2007). Tumor-associated macrophages press the angiogenic switch in breast cancer. Cancer Res. *67*, 5064–5066. .

Lin, E.Y., Nguyen, A. V, Russell, R.G., and Pollard, J.W. (2001). Colony-stimulating factor 1 promotes progression of mammary tumors to malignancy. J. Exp. Med. *193*, 727–740. .

Linde, N., Lederle, W., Depner, S., van Rooijen, N., Gutschalk, C.M., and Mueller, M.M. (2012).

Vascular endothelial growth factor-induced skin carcinogenesis depends on recruitment and alternative activation of macrophages. J. Pathol. *227*, 17–28. .

Lindsten, T., Hedbrant, A., Ramberg, A., Wijkander, J., Solterbeck, A., Eriksson, M., Delbro, D., and Erlandsson, A. (2017). Effect of macrophages on breast cancer cell proliferation, and on expression of hormone receptors, uPAR and HER-2. Int. J. Oncol. *51*, 104–114. .

Ling, B., Watt, K., Banerjee, S., Newsted, D., Truesdell, P., Adams, J., Sidhu, S.S., and Craig, A.W.B. (2017). A novel immunotherapy targeting MMP-14 limits hypoxia, immune suppression and metastasis in triple-negative breast cancer models. Oncotarget *8*, 58372–58385. .

Linsley, P.S., and Ledbetter, J.A. (1993). The role of the CD28 receptor during T cell responses to antigen. Annu. Rev. Immunol. *11*, 191–212. .

Liotta, L.A., and Kohn, E.C. (2001). The microenvironment of the tumour - Host interface. Nature *411*, 375–379. .

Liu, X.S., and Mardis, E.R. (2017). Applications of Immunogenomics to Cancer. Cell *168*, 600–612. .

Liu, H., Tabuchi, T., Takemura, A., Kasuga, T., Motohashi, G., Hiraishi, K., Katano, M., Nakada, I., Ubukata, H., and Tabuchi, T. (2008). The granulocyte/lymphocyte ratio as an independent predictor of tumour growth, metastasis and progression: Its clinical applications. Mol. Med. Rep. *1*, 699–704. .

Liu, J., Zhang, N., Li, Q., Zhang, W., Ke, F., Leng, Q., Wang, H., Chen, J., and Wang, H. (2011). Tumor-associated macrophages recruit CCR6+ regulatory T cells and promote the development of colorectal cancer via enhancing CCL20 production in mice. PLoS One *6*, e19495. .

Liu, M.C., Pitcher, B.N., Mardis, E.R., Davies, S.R., Friedman, P.N., Snider, J.E., Vickery, T.L., Reed, J.P., Deschryver, K., Singh, B., et al. (2016). PAM50 gene signatures and breast cancer prognosis with adjuvant anthracycline- and taxane-based chemotherapy : correlative analysis of

C9741 (Alliance). Nat. Publ. Gr.

Liu, P., Boyle, A.J., Lu, Y., Adams, J., Chi, Y., Reilly, R.M., and Winnik, M.A. (2015a). Metal-Chelating Polymers (MCPs) with Zwitterionic Pendant Groups Complexed to Trastuzumab Exhibit Decreased Liver Accumulation Compared to Polyanionic MCP Immunoconjugates. Biomacromolecules *16*, 3613–3623. .

Liu, S., Galat, V., Galat4, Y., Lee, Y.K.A., Wainwright, D., and Wu, J. (2021). NK cell-based cancer immunotherapy: from basic biology to clinical development. J. Hematol. Oncol. *14*. .

Liu, X., Pu, Y., Cron, K., Deng, L., Kline, J., Frazier, W.A., Xu, H., Peng, H., Fu, Y.-X., and Xu, M.M. (2015b). CD47 blockade triggers T cell-mediated destruction of immunogenic tumors. Nat. Med. *21*, 1209–1215. .

Liu, Y., Li, Y., Tian, R., Liu, W., Fei, Z., Long, Q., Wang, X., and Zhang, X. (2009). The expression and significance of HIF-1α and GLUT-3 in glioma. Brain Res. *1304*, 149–154. .

Lizio, M., Harshbarger, J., Shimoji, H., Severin, J., Kasukawa, T., Sahin, S., Abugessaisa, I., Fukuda, S., Hori, F., Ishikawa-Kato, S., et al. (2015). Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome Biol. *16*, 22. .

Loibl, S., and Gianni, L. (2016). Breast cancer 2 HER2-positive breast cancer. Lancet *6736*, 1–15. .

Loibl, S., Poortmans, P., Morrow, M., Denkert, C., and Curigliano, G. (2021a). Breast cancer. Lancet *397*, 1750–1769. .

Loibl, S., Poortmans, P., Morrow, M., Denkert, C., and Curigliano, G. (2021b). Breast cancer. Lancet (London, England) *397*, 1750–1769. .

Loyher, P.-L., Hamon, P., Laviron, M., Meghraoui-Kheddar, A., Goncalves, E., Deng, Z., Torstensson, S., Bercovici, N., Baudesson de Chanville, C., Combadière, B., et al. (2018). Macrophages of distinct origins contribute to tumor development in the lung. J. Exp. Med. *215*, 2536–

2553. .

Lu, X., and Kang, Y. (2010). Hypoxia and hypoxia-inducible factors: Master regulators of metastasis. Clin. Cancer Res. *16*, 5928–5935. .

Ma, L., Cho, W., and Nelson, E.R. (2021). Our evolving understanding of how 27-hydroxycholesterol influences cancer. Biochem. Pharmacol. 114621. .

Mabbott, N.A., Baillie, J.K., Brown, H., Freeman, T.C., and Hume, D.A. (2013). An expression atlas of human primary cells: inference of gene function from coexpression networks. BMC Genomics *14*, 632. .

Macintyre, G., Goranova, T.E., De Silva, D., Ennis, D., Piskorz, A.M., Eldridge, M., Sie, D., Lewsley, L.A., Hanif, A., Wilson, C., et al. (2018). Copy number signatures and mutational processes in ovarian carcinoma. Nat. Genet. *50*, 1262–1270. .

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., and Martersteck, E.M. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell *161*, 1202–1214. .

Mahmoud, S.M.A., Lee, A.H.S., Paish, E.C., Macmillan, R.D., Ellis, I.O., and Green, A.R. (2012). Tumour-infiltrating macrophages and clinical outcome in breast cancer. J. Clin. Pathol. *65*, 159–163. .

Mahoney, K.M., Shukla, S.A., Patsoukis, N., Chaudhri, A., Browne, E.P., Arazi, A., Eisenhaure, T.M., Pendergraft, W.F. 3rd, Hua, P., Pham, H.C., et al. (2019). A secreted PD-L1 splice variant that covalently dimerizes and mediates immunosuppression. Cancer Immunol. Immunother. *68*, 421–432. .

Maman, S., and Witz, I.P. (2018). A history of exploring cancer in context. Nat. Rev. Cancer *18*, 359–376. .

Mantovani, A., Marchesi, F., Malesci, A., Laghi, L., and Allavena, P. (2017). Tumour-associated

macrophages as treatment targets in oncology. Nat. Rev. Clin. Oncol. *14*, 399–416. .

Marcus, L., Lemery, S.J., Keegan, P., and Pazdur, R. (2019). FDA approval summary: Pembrolizumab for the treatment of microsatellite instability-high solid tumors. Clin. Cancer Res. *25*, 3753–3758. .

Marra, A., Trapani, D., Viale, G., Criscitiello, C., and Curigliano, G. (2020). Practical classification of triple-negative breast cancer: intratumoral heterogeneity, mechanisms of drug resistance, and novel therapies. Npj Breast Cancer *6*, 1–16. .

Martin-Castillo, B., Oliveras-Ferraros, C., Vazquez-Martin, A., Cufí, S., Moreno, J.M., Corominas-Faja, B., Urruticoechea, A., Martín, Á.G., López-Bonet, E., and Menendez, J.A. (2013). Basal/HER2 breast carcinomas Integrating molecular taxonomy with cancer stem cell dynamics to predict primary resistance to trastuzumab (Herceptin). Cell Cycle *12*, 225–245. .

Martincorena, I., Raine, K.M., Gerstung, M., Dawson, K.J., Haase, K., Van Loo, P., Davies, H., Stratton, M.R., and Campbell, P.J. (2017). Universal Patterns of Selection in Cancer and Somatic Tissues. Cell *171*, 1029-1041.e21. .

Martincorena, I., Fowler, J.C., Wabik, A., Lawson, A.R.J., Abascal, F., Hall, M.W.J., Cagan, A., Murai, K., Mahbubani, K., Stratton, M.R., et al. (2018). Somatic mutant clones colonize the human esophagus with age. Science (80-.). *362*, 911–917. .

Martinez, F.O., and Gordon, S. (2014). The M1 and M2 paradigm of macrophage activation: Time for reassessment. F1000Prime Rep. *6*, 1–13. .

Martinez, F.O., Gordon, S., Locati, M., and Mantovani, A. (2006). Transcriptional Profiling of the Human Monocyte-to-Macrophage Differentiation and Polarization: New Molecules and Patterns of Gene Expression. J. Immunol. *177*, 7303–7311. .

Mass, E. (2018). Delineating the origins, developmental programs and homeostatic functions of

tissue-resident macrophages. Int. Immunol. *30*, 493–501. .

Mass, E., Ballesteros, I., Farlik, M., Halbritter, F., Günther, P., Crozet, L., Jacome-Galarza, C.E., Händler, K., Klughammer, J., Kobayashi, Y., et al. (2016). Specification of tissue-resident macrophages during organogenesis. Science (80-.). *353*. .

Matuła, K., Rivello, F., and Huck, W.T.S. (2020). Single-Cell Analysis Using Droplet Microfluidics. Adv. Biosyst. *4*. .

McShane, L.M., Radmacher, M.D., Freidlin, B., Yu, R., Li, M.C., and Simon, R. (2002). Methods for assessing reproducibility of clustering pattens observed in analyses of microarray data. Bioinformatics *18*, 1462–1469. .

Medina-Echeverz, J., Fioravanti, J., Zabala, M., Ardaiz, N., Prieto, J., and Berraondo, P. (2011). Successful colon cancer eradication after chemoimmunotherapy is associated with profound phenotypic change of intratumoral myeloid cells. J. Immunol. *186*, 807–815. .

Medrano, R.F. V, Hunger, A., Mendonça, S.A., Barbuto, J.A.M., and Strauss, B.E. (2017). Immunomodulatory and antitumor effects of type I interferons and their application in cancer therapy. Oncotarget *8*, 71249–71284. .

Medrek, C., Pontén, F., Jirström, K., and Leandersson, K. (2012). The presence of tumor associated macrophages in tumor stroma as a prognostic marker for breast cancer patients. BMC Cancer *12*, 306. .

Merad, M., Sathe, P., Helft, J., Miller, J., and Mortha, A. (2013). The dendritic cell lineage: ontogeny and function of dendritic cells and their subsets in the steady state and the inflamed setting. Annu. Rev. Immunol. *31*, 563–604. .

Mereu, E., Lafzi, A., Moutinho, C., Ziegenhain, C., McCarthy, D.J., Álvarez-Varela, A., Batlle, E., Sagar, Grün, D., Lau, J.K., et al. (2020). Benchmarking single-cell RNA-sequencing protocols for

cell atlas projects. Nat. Biotechnol.

Mills, C.D., Kincaid, K., Alt, J.M., Heilman, M.J., and Hill, A.M. (2000). M-1/M-2 Macrophages and the Th1/Th2 Paradigm. J. Immunol. *164*, 6166–6173. .

Mishalian, I., Bayuh, R., Eruslanov, E., Michaeli, J., Levy, L., Zolotarov, L., Singhal, S., Albelda, S.M., Granot, Z., and Fridlender, Z.G. (2014). Neutrophils recruit regulatory T-cells into tumors via secretion of CCL17--a new mechanism of impaired antitumor immunity. Int. J. Cancer *135*, 1178–1186. .

Mitchem, J.B., Brennan, D.J., Knolhoff, B.L., Belt, B.A., Zhu, Y., Sanford, D.E., Belaygorod, L., Carpenter, D., Collins, L., Piwnica-Worms, D., et al. (2013). Targeting tumor-infiltrating macrophages decreases tumor-initiating cells, relieves immunosuppression, and improves chemotherapeutic responses. Cancer Res. *73*, 1128–1141. .

Mohammed, Z.M.A., Going, J.J., Edwards, J., Elsberger, B., Doughty, J.C., and McMillan, D.C. (2012). The relationship between components of tumour inflammatory cell infiltrate and clinicopathological factors and survival in patients with primary operable invasive ductal breast cancer. Br. J. Cancer *107*, 864–873. .

Monti, S., Tamayo, P., Mesirov, J., and Golub, T. (2003). Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. Mach. Learn. *52*, 91–118. .

Morandi, F., Levreri, I., Bocca, P., Galleni, B., Raffaghello, L., Ferrone, S., Prigione, I., and Pistoia, V. (2007). Human neuroblastoma cells trigger an immunosuppressive program in monocytes by stimulating soluble HLA-G release. Cancer Res. *67*, 6433–6441. .

Movahedi, K., Laoui, D., Gysemans, C., Baeten, M., Stangé, G., Van den Bossche, J., Mack, M., Pipeleers, D., In't Veld, P., De Baetselier, P., et al. (2010). Different tumor microenvironments contain functionally distinct subsets of macrophages derived from Ly6C(high) monocytes. Cancer

Res. *70*, 5728–5739. .

Mukaka, M. (2012). Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research. Malawi Med. J. *24*, 69–71. .

Munn, D.H., and Bronte, V. (2016). Immune suppressive mechanisms in the tumor microenvironment. Curr. Opin. Immunol. *39*, 1–6. .

Munn, D.H., Sharma, M.D., Hou, D., Baban, B., Lee, J.R., Antonia, S.J., Messina, J.L., Chandler, P., Koni, P.A., and Mellor, A.L. (2004). Expression of indoleamine 2,3-dioxygenase by plasmacytoid dendritic cells in tumor-draining lymph nodes. J. Clin. Invest. *114*, 280–290. .

Murdoch, C., Giannoudis, A., and Lewis, C.E. (2004). Mechanisms regulating the recruitment of macrophages into hypoxic areas of tumors and other ischemic tissues. Blood *104*, 2224–2234. .

Murdoch, C., Muthana, M., Coffelt, S.B., and Lewis, C.E. (2008). The role of myeloid cells in the promotion of tumour angiogenesis. Nat. Rev. Cancer *8*, 618–631. .

Murray, P.J., Allen, J.E., Biswas, S.K., Fisher, E.A., Gilroy, D.W., Goerdt, S., Gordon, S., Hamilton, J.A., Ivashkiv, L.B., Lawrence, T., et al. (2014). Macrophage Activation and Polarization: Nomenclature and Experimental Guidelines. Immunity *41*, 14–20. .

Muz, B., de la Puente, P., Azab, F., and Azab, A.K. (2015). The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy. Hypoxia (Auckland, N.Z.) *3*, 83–92. .

Myung, I.J. (2003). Tutorial on maximum likelihood estimation. J. Math. Psychol. *47*, 90–100. .

Nagy, J.A., Chang, S.-H., Dvorak, A.M., and Dvorak, H.F. (2009). Why are tumour blood vessels abnormal and why is it important to know? Br. J. Cancer *100*, 865–869. .

Nakagawa, Y., Negishi, Y., Shimizu, M., Takahashi, M., Ichikawa, M., and Takahashi, H. (2015). Effects of extracellular pH and hypoxia on the function and development of antigen-specific cytotoxic T lymphocytes. Immunol. Lett. *167*, 72–86. .

Nausch, N., and Cerwenka, A. (2008). NKG2D ligands in tumor immunity. Oncogene *27*, 5944–

5958. .

Newman, A.M., Liu, C.L., Green, M.R., Gentles, A.J., Feng, W., Xu, Y., Hoang, C.D., Diehn, M., and Alizadeh, A.A. (2015). Robust enumeration of cell subsets from tissue expression profiles. Nat. Methods *12*, 453–457. .

Newman, A.M., Gentles, A.J., Liu, C.L., Diehn, M., and Alizadeh, A.A. (2017). Data normalization considerations for digital tumor dissection. Genome Biol. *18*, 1–6. .

Newman, A.M., Steen, C.B., Liu, C.L., Gentles, A.J., Chaudhuri, A.A., Scherer, F., Khodadoust, M.S., Esfahani, M.S., Luca, B.A., Steiner, D., et al. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat. Biotechnol.

Ng, T.H.S., Britton, G.J., Hill, E. V., Verhagen, J., Burton, B.R., and Wraith, D.C. (2013). Regulation of adaptive immunity; the role of interleukin-10. Front. Immunol. *4*, 129. .

Nguyen, L. V., Vanner, R., Dirks, P., and Eaves, C.J. (2012). Cancer stem cells: an evolving concept. Nat. Rev. Cancer *12*, 133–143. .

Noel, P.J., Boise, L.H., and Thompson, C.B. (1996). Regulation of T cell activation by CD28 and CTLA4. Adv. Exp. Med. Biol. *406*, 209–217. .

Noman, M.Z., Desantis, G., Janji, B., Hasmim, M., Karray, S., Dessen, P., Bronte, V., and Chouaib, S. (2014). PD-L1 is a novel direct target of HIF-1α, and its blockade under hypoxia enhanced MDSC-mediated T cell activation. J. Exp. Med. *211*, 781–790. .

Noman, M.Z., Hasmim, M., Messai, Y., Terry, S., Kieda, C., Janji, B., and Chouaib, S. (2015). Hypoxia: a key player in antitumor immune response. A Review in the Theme: Cellular Responses to Hypoxia. Am. J. Physiol. Physiol. *309*, C569–C579. .

Noy, R., and Pollard, J.W. (2014). Tumor-Associated Macrophages: From Mechanisms to Therapy. Immunity *41*, 49–61. .

Obeid, M., Tesniere, A., Ghiringhelli, F., Fimia, G.M., Apetoh, L., Perfettini, J.-L., Castedo, M.,

Mignot, G., Panaretakis, T., Casares, N., et al. (2007). Calreticulin exposure dictates the immunogenicity of cancer cell death. Nat. Med. *13*, 54–61. .

Oh, S.A., and Li, M.O. (2013). TGF-β: Guardian of T Cell Function. J. Immunol. *191*, 3973–3979. .

Osińska, I., Popko, K., and Demkow, U. (2014). Perforin: An important player in immune response. Cent. Eur. J. Immunol. *39*, 109–115. .

Van Overmeire, E., Stijlemans, B., Heymann, F., Keirsse, J., Morias, Y., Elkrim, Y., Brys, L., Abels, C., Lahmar, Q., Ergen, C., et al. (2016). M-CSF and GM-CSF Receptor Signaling Differentially Regulate Monocyte Maturation and Macrophage Polarization in the Tumor Microenvironment. Cancer Res. *76*, 35–42. .

Paget, S. (1889). the Distribution of Secondary Growths in Cancer of the Breast. Lancet *133*, 571–573. .

Palazon, A., Tyrakis, P.A., Macias, D., Veliça, P., Rundqvist, H., Fitzpatrick, S., Vojnovic, N., Phan, A.T., Loman, N., Hedenfalk, I., et al. (2017). An HIF-1α/VEGF-A Axis in Cytotoxic T Cells Regulates Tumor Progression. Cancer Cell *32*, 669-683.e5. .

De Palma, M., and Lewis, C.E. (2013). Macrophage regulation of tumor responses to anticancer therapies. Cancer Cell *23*, 277–286. .

De Palma, M., Biziato, D., and Petrova, T. V. (2017). Microenvironmental regulation of tumour angiogenesis. Nat. Rev. Cancer *17*, 457–474. .

Palmeri, M., Mehnert, J., Silk, A.W., Jabbour, S.K., Ganesan, S., Popli, P., Riedlinger, G., Stephenson, R., de Meritens, A.B., Leiser, A., et al. (2022). Real-world application of tumor mutational burden-high (TMB-high) and microsatellite instability (MSI) confirms their utility as immunotherapy biomarkers. ESMO Open *7*, 100336. .

Palucka, A.K., Blanck, J.-P., Bennett, L., Pascual, V., and Banchereau, J. (2005). Cross-regulation of

TNF and IFN-alpha in autoimmune diseases. Proc. Natl. Acad. Sci. U. S. A. *102*, 3372–3377. .

Park-Min, K.H., Antoniv, T.T., and Ivashkiv, L.B. (2005). Regulation of macrophage phenotype by long-term exposure to IL-10. Immunobiology *210*, 77–86. .

Patel, A.P., Tirosh, I., Trombetta, J.J., Shalek, A.K., Gillespie, S.M., Wakimoto, H., Cahill, D.P., Nahed, B. V., Curry, W.T., Martuza, R.L., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science (80-.). *344*, 1396–1401. .

Pelisch, N., Rosas Almanza, J., Stehlik, K.E., Aperi, B. V., and Kroner, A. (2020). CCL3 contributes to secondary damage after spinal cord injury. J. Neuroinflammation *17*, 1–16. .

Perdiguero, E.G., and Geissmann, F. (2014). Identifying the infiltrators. Science (80-.). *344*, 801–802. .

Persi, E., Duran-Frigola, M., Damaghi, M., Roush, W.R., Aloy, P., Cleveland, J.L., Gillies, R.J., and Ruppin, E. (2018). Systems analysis of intracellular pH vulnerabilities for cancer therapy. Nat. Commun. *9*, 2997. .

Petitprez, F., Sun, C.M., Lacroix, L., Sautès-Fridman, C., de Reyniès, A., and Fridman, W.H. (2018). Quantitative Analyses of the Tumor Microenvironment Composition and Orientation in the Era of Precision Medicine. Front. Oncol. *8*, 1–6. .

Petrova, V., Annicchiarico-Petruzzelli, M., Melino, G., and Amelio, I. (2018). The hypoxic tumour microenvironment. Oncogenesis *7*. .

Piccard, H., Muschel, R.J., and Opdenakker, G. (2012). On the dual roles and polarized phenotypes of neutrophils in tumor development and progression. Crit. Rev. Oncol. Hematol. *82*, 296–309. .

Poh, A.R., and Ernst, M. (2018). Targeting macrophages in cancer: From bench to bedside. Front. Oncol. *8*, 1–16. .

Pollard, J.W. (2004). Tumour-educated macrophages promote tumour progression and metastasis.

Nat. Rev. Cancer *4*, 71–78. .

Porrello, A., Leslie, P.L., Harrison, E.B., Gorentla, B.K., Kattula, S., Ghosh, S.K., Azam, S.H., Holtzhausen, A., Chao, Y.L., Hayward, M.C., et al. (2018). Factor XIIIA-expressing inflammatory monocytes promote lung squamous cancer through fibrin cross-linking. Nat. Commun. *9*. .

Prabhakaran, S., Azizi, E., Carr, A., and Pe'er, D. (2016). Dirichlet process mixture model for correcting technical variation in single-cell gene expression data. 33rd Int. Conf. Mach. Learn. ICML 2016 *3*, 1691–1715. .

Prieur, X., Mok, C.Y.L., Velagapudi, V.R., Núñez, V., Fuentes, L., Montaner, D., Ishikawa, K., Camacho, A., Barbarroja, N., O'Rahilly, S., et al. (2011). Differential lipid partitioning between adipocytes and tissue macrophages modulates macrophage lipotoxicity and M2/M1 polarization in obese mice. Diabetes *60*, 797–809. .

Qian, B.Z., and Pollard, J.W. (2010). Macrophage Diversity Enhances Tumor Progression and Metastasis. Cell *141*, 39–51. .

Qiu, S.Q., Waaijer, S.J.H., Zwager, M.C., de Vries, E.G.E., van der Vegt, B., and Schröder, C.P. (2018). Tumor-associated macrophages in breast cancer: Innocent bystander or important player? Cancer Treat. Rev. *70*, 178–189. .

Racle, J., de Jonge, K., Baumgaertner, P., Speiser, D.E., and Gfeller, D. (2017). Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. Elife *6*, 1–25. .

Ran, M., and Witz, I.P. (1972). Tumor-associated immunoglobulins. enhancement of syngeneic tumors by igg2-containing tumor eluates. Int. J. Cancer *9*, 242–247. .

Ransohoff, R.M. (2016). A polarizing question: Do M1 and M2 microglia exist. Nat. Neurosci. *19*, 987–991. .

Rapisarda, A., and Melillo, G. (2012). Overcoming disappointing results with antiangiogenic therapy

by targeting hypoxia. Nat. Rev. Clin. Oncol. *9*, 378–390. .

Rapp, M., Wintergerst, M.W.M., Kunz, W.G., Vetter, V.K., Knott, M.M.L., Lisowski, D., Haubner, S., Moder, S., Thaler, R., Eiber, S., et al. (2019). CCL22 controls immunity by promoting regulatory T cell communication with dendritic cells in lymph nodes. J. Exp. Med. *216*, 1170–1181. .

Ratcliffe, P.J. (2007). HIF-1 and HIF-2: working alone or together in hypoxia? J. Clin. Invest. *117*, 862–865. .

Ren, L., Yu, Y., Wang, L., Zhu, Z., Lu, R., and Yao, Z. (2016). Hypoxia-induced CCL28 promotes recruitment of regulatory T cells and tumor growth in liver cancer. Oncotarget *7*, 75763–75773. .

Ren, X., Zhang, L., Zhang, Y., Li, Z., Siemers, N., and Zhang, Z. (2021). Insights Gained from Single-Cell Analysis of Immune Cells in the Tumor Microenvironment. Annu. Rev. Immunol. *39*, 583–609. .

Richardson, M.A., Ramirez, T., Russell, N.C., and Moye, L.A. (1999). Coley toxins immunotherapy: A retrospective review. Altern. Ther. Health Med. *5*, 42–47. .

Richters, A., and Kaspersky, C.L. (1975). Surface immunoglobulin positive lymphocytes in human breast cancer tissue and homolateral axillary lymph nodes. Cancer *35*, 129–133. .

Ringnér, M. (2008). What is principal component analysis? Nat. Biotechnol. *26*, 303–304. .

Robert, C. (2020). A decade of immune-checkpoint inhibitors in cancer therapy. Nat. Commun. *11*, 10–12. .

Robert, C., Thomas, L., Bondarenko, I., O'Day, S., Weber, J., Garbe, C., Lebbe, C., Baurain, J.-F., Testori, A., Grob, J.-J., et al. (2011). Ipilimumab plus Dacarbazine for Previously Untreated Metastatic Melanoma. N. Engl. J. Med. *364*, 2517–2526. .

Rudd, C.E., Taylor, A., and Schneider, H. (2009). CD28 and CTLA-4 coreceptor expression and signal transduction. Immunol. Rev. *229*, 12–26. .

Russel, S.W., Doe, W.F., and Cochrane, C.G. (1976). Number of macrophages and distribution of

mitotic activity in regressing and progressing Moloney sarcomas. J. Immunol. *116*, 164–166. .

Sadiku, P., and Walmsley, S.R. (2019). Hypoxia and the regulation of myeloid cell metabolic imprinting: consequences for the inflammatory response. EMBO Rep. *20*, 1–11. .

Sagiv, J.Y., Michaeli, J., Assi, S., Mishalian, I., Kisos, H., Levy, L., Damti, P., Lumbroso, D., Polyansky, L., Sionov, R. V, et al. (2015). Phenotypic diversity and plasticity in circulating neutrophil subpopulations in cancer. Cell Rep. *10*, 562–573. .

Saleiro, D., and Platanias, L.C. (2019). Interferon signaling in cancer. Non-canonical pathways and control of intracellular immune checkpoints. Semin. Immunol. *43*, 101299. .

Samoszuk, M., and Corwin, M.A. (2003a). Acceleration of tumor growth and peri-tumoral blood clotting by imatinib mesylate (Gleevec^{TM}). Int. J. Cancer *106*, 647–652. .

Samoszuk, M., and Corwin, M.A. (2003b). Mast cell inhibitor cromolyn increases blood clotting and hypoxia in murine breast cancer. Int. J. Cancer *107*, 159–163. .

Sangaletti, S., Di Carlo, E., Gariboldi, S., Miotti, S., Cappetti, B., Parenza, M., Rumio, C., Brekken, R.A., Chiodoni, C., and Colombo, M.P. (2008). Macrophage-derived SPARC bridges tumor cell-extracellular matrix interactions toward metastasis. Cancer Res. *68*, 9050–9059. .

Schlecker, E., Stojanovic, A., Eisen, C., Quack, C., Falk, C.S., Umansky, V., and Cerwenka, A. (2012). Tumor-infiltrating monocytic myeloid-derived suppressor cells mediate CCR5-dependent recruitment of regulatory T cells favoring tumor growth. J. Immunol. *189*, 5602–5611. .

Schmidt, A., and Weber, O.F. (2006). In memoriam of Rudolf virchow: a historical retrospective including aspects of inflammation, infection and neoplasia. In Contributions to Microbiology, (Karger Publishers), pp. 1–15.

Schmidt, H., Bastholt, L., Geertsen, P., Christensen, I.J., Larsen, S., Gehl, J., and von der Maase, H. (2005). Elevated neutrophil and monocyte counts in peripheral blood are associated with poor

survival in patients with metastatic melanoma: a prognostic model. Br. J. Cancer *93*, 273–278. .

Schoorl, R., Brutel De La Riviere, A., Von Dem Borne, E.G., and Feltkamp Vroom, T.M. (1976). Identification of T and B lymphocytes in human breast cancer with immunohistochemical techniques. Am. J. Pathol. *84*, 529–544. .

Sconocchia, G., Arriga, R., Tornillo, L., Terracciano, L., Ferrone, S., and Spagnoli, G.C. (2012). Melanoma cells inhibit NK cell functions. Cancer Res. *72*, 5428–5429. .

Seita, J., and Weissman, I.L. (2010). Hematopoietic stem cell: Self-renewal versus differentiation. Wiley Interdiscip. Rev. Syst. Biol. Med. *2*, 640–653. .

Shalapour, S., Font-Burgada, J., Di Caro, G., Zhong, Z., Sanchez-Lopez, E., Dhar, D., Willimsky, G., Ammirante, M., Strasner, A., Hansel, D.E., et al. (2015). Immunosuppressive plasma cells impede T-cell-dependent immunogenic chemotherapy. Nature *521*, 94–98. .

Sharda, D.R., Yu, S., Ray, M., Squadrito, M.L., De Palma, M., Wynn, T.A., Morris, S.M., and Hankey, P.A. (2011). Regulation of macrophage arginase expression and tumor growth by the Ron receptor tyrosine kinase. J. Immunol. *187*, 2181–2192. .

Sharonov, G. V., Serebrovskaya, E.O., Yuzhakova, D. V., Britanova, O. V., and Chudakov, D.M. (2020). B cells, plasma cells and antibody repertoires in the tumour microenvironment. Nat. Rev. Immunol. *20*, 294–307. .

Shen-Orr, S.S., and Gaujoux, R. (2013). Computational deconvolution: extracting cell type-specific information from heterogeneous samples. Curr. Opin. Immunol. *25*, 571–578. .

Shi, J.-Y., Gao, Q., Wang, Z.-C., Zhou, J., Wang, X.-Y., Min, Z.-H., Shi, Y.-H., Shi, G.-M., Ding, Z.-B., Ke, A.-W., et al. (2013). Margin-Infiltrating CD20+ B Cells Display an Atypical Memory Phenotype and Correlate with Favorable Prognosis in Hepatocellular Carcinoma. Clin. Cancer Res. *19*, 5994–6005. .

Shipp, M.A., Ross, K.N., Tamayo, P., Weng, A.P., Kutok, J.L., Aguiar, R.C.T., Gaasenbeek, M.,

Angelo, M., Reich, M., Pinkus, G.S., et al. (2002). Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. Nat. Med. *8*, 68–74. .

Sica, A., and Mantovani, A. (2012). Macrophage plasticity and polarization: In vivo veritas. J. Clin. Invest. *122*, 787–795. .

Sica, A., Saccani, A., Bottazzi, B., Bernasconi, S., Allavena, P., Gaetano, B., Fei, F., LaRosa, G., Scotton, C., Balkwill, F., et al. (2000). Defective expression of the monocyte chemotactic protein-1 receptor CCR2 in macrophages associated with human ovarian carcinoma. J. Immunol. *164*, 733–738. .

Singh, S.K., and Singh, R. (2022). Nanotherapy: targeting the tumour microenvironment. Nat. Rev. Cancer *22*, 258. .

Singhal, S., Stadanlick, J., Annunziata, M.J., Rao, A.S., Bhojnagarwala, P.S., O'Brien, S., Moon, E.K., Cantu, E., Danet-Desnoyers, G., Ra, H.-J., et al. (2019). Human tumor-associated monocytes/macrophages and their regulation of T cell responses in early-stage lung cancer. Sci. Transl. Med. *11*, eaat1500. .

Singleton, D.C., Macann, A., and Wilson, W.R. (2021). Therapeutic targeting of the hypoxic tumour microenvironment. Nat. Rev. Clin. Oncol. *18*, 751–772. .

Sisirak, V., Faget, J., Gobert, M., Goutagny, N., Vey, N., Treilleux, I., Renaudineau, S., Poyet, G., Labidi-Galy, S.I., Goddard-Leon, S., et al. (2012). Impaired IFN-α production by plasmacytoid dendritic cells favors regulatory T-cell expansion that may contribute to breast cancer progression. Cancer Res. *72*, 5188–5197. .

Smyth, M.J., Cretney, E., Kelly, J.M., Westwood, J.A., Street, S.E.A., Yagita, H., Takeda, K., Dommelen, S.L.H.V., Degli-Esposti, M.A., and Hayakawa, Y. (2005). Activation of NK cell cytotoxicity. Mol. Immunol. *42*, 501–510. .

Smyth, M.J., Ngiow, S.F., Ribas, A., and Teng, M.W.L. (2016). Combination cancer

immunotherapies tailored to the tumour microenvironment. Nat. Rev. Clin. Oncol. *13*, 143–158. .

Soto-Pantoja, D.R., Terabe, M., Ghosh, A., Ridnour, L.A., DeGraff, W.G., Wink, D.A., Berzofsky, J.A., and Roberts, D.D. (2014). CD47 in the tumor microenvironment limits cooperation between antitumor T-cell immunity and radiotherapy. Cancer Res. *74*, 6771–6783. .

Spear, P., Wu, M.R., Sentman, M.L., and Sentman, C.L. (2013). Nkg2d ligands as therapeutic targets. Cancer Immun. *13*, 8. .

Stadtmann, A., and Zarbock, A. (2012). CXCR2: From Bench to Bedside. Front. Immunol. *3*, 263. .

Stamper, C.C., Zhang, Y., Tobin, J.F., Erbe, D. V., Ikemizu, S., Davis, S.J., Stahl, M.L., Seehra, J., Somers, W.S., and Mosyak, L. (2001). Crystal structure of the B7-1/CTLA-4 complex that inhibits human immune responses. Nature *410*, 608–611. .

Staversky, R.J., Byun, D.K., Georger, M.A., Zaffuto, B.J., Goodman, A., Becker, M.W., Calvi, L.M., and Frisch, B.J. (2018). The Chemokine CCL3 Regulates Myeloid Differentiation and Hematopoietic Stem Cell Numbers. Sci. Rep. *8*, 1–9. .

Steele, C.W., Karim, S.A., Leach, J.D.G., Bailey, P., Upstill-Goddard, R., Rishi, L., Foth, M., Bryson, S., McDaid, K., Wilson, Z., et al. (2016). CXCR2 Inhibition Profoundly Suppresses Metastases and Augments Immunotherapy in Pancreatic Ductal Adenocarcinoma. Cancer Cell *29*, 832–845. .

Steen, C.B., Liu, C.L., Alizadeh, A.A., and Newman, A.M. (2020). Profiling Cell Type Abundance and Expression in Bulk Tissues with CIBERSORTx. Methods Mol. Biol. *2117*, 135–157. .

Stein, M., Keshav, S., Harris, N., and Gordon, S. (1992). Interleukin 4 potently enhances murine macrophage mannose receptor activity: A marker of alternative immunologic macrophage activation. J. Exp. Med. *176*, 287–292. .

Stossi, F., Madak-Erdoğan, Z., and Katzenellenbogen, B.S. (2012). Macrophage-elicited loss of estrogen receptor-α in breast cancer cells via involvement of MAPK and c-Jun at the ESR1 genomic

locus. Oncogene *31*, 1825–1834. .

Strachan, D.C., Ruffell, B., Oei, Y., Bissell, M.J., Coussens, L.M., Pryer, N., and Daniel, D. (2013). CSF1R inhibition delays cervical and mammary tumor growth in murine models by attenuating the turnover of tumor-associated macrophages and enhancing infiltration by CD8(+) T cells. Oncoimmunology *2*, e26968. .

Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics *19*, 477. .

Streets, A.M., Zhang, X., Cao, C., Pang, Y., Wu, X., Xiong, L., Yang, L., Fu, Y., Zhao, L., and Tang, F. (2014). Microfluidic single-cell whole-transcriptome sequencing. Proc. Natl. Acad. Sci. *111*, 7048–7053. .

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888-1902.e21. .

Sturm, G., Finotello, F., Petitprez, F., Zhang, J.D., Baumbach, J., Fridman, W.H., List, M., and Aneichyk, T. (2019). Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. Bioinformatics *35*, i436–i445. .

Subramanian, A., Subramanian, A., Tamayo, P., Tamayo, P., Mootha, V.K., Mootha, V.K., Mukherjee, S., Mukherjee, S., Ebert, B.L., Ebert, B.L., et al. (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. .

Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36

Cancers in 185 Countries. CA. Cancer J. Clin. *71*, 209–249. .

Swiecki, M., and Colonna, M. (2015). The multifaceted biology of plasmacytoid dendritic cells. Nat. Rev. Immunol. *15*, 471–485. .

Swietach, P., Vaughan-Jones, R.D., and Harris, A.L. (2007). Regulation of tumor pH and the role of carbonic anhydrase 9. Cancer Metastasis Rev. *26*, 299–310. .

Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N.T., Legeay, M., Fang, T., Bork, P., et al. (2021). The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets'. Nucleic Acids Res. *49*, 10800–10800. .

Taheri, F., Ochoa, J.B., Faghiri, Z., Culotta, K., Park, H.J., Lan, M.S., Zea, A.H., and Ochoa, A.C. (2001). L-Arginine regulates the expression of the T-cell receptor zeta chain (CD3zeta) in Jurkat cells. Clin. Cancer Res. an Off. J. Am. Assoc. Cancer Res. *7*, 958s-965s. .

Takahashi, T., Tagami, T., Yamazaki, S., Uede, T., Shimizu, J., Sakaguchi, N., Mak, T.W., and Sakaguchi, S. (2000). Immunologic self-tolerance maintained by CD25+CD4+ regulatory T cells constitutively expressing cytotoxic T lymphocyte-associated antigen 4. J. Exp. Med. *192*, 303–309. .

Tang, R.P., Kacinski, B., Validire, P., Beuvon, F., Sastre, X., Benoit, P., dela Rochefordière, A., Mosseri, V., Pouillart, P., and Scholl, S. (1990). Oncogene amplification correlates with dense lymphocyte infiltration in human breast cancers: a role for hematopoietic growth factor release by tumor cells? J. Cell. Biochem. *44*, 189–198. .

Theodor Brunner, K., Robson Macdonald, H., and Cerottini, J.C. (1981). Quantitation and clonal isolation of cytolytic T lymphocyte precursors selectively infiltrating murine sarcoma virus-induced tumors. J. Exp. Med. *154*, 362–373. .

Thorsson, V., Gibbs, D.L., Brown, S.D., Wolf, D., Bortone, D.S., Ou Yang, T.H., Porta-Pardo, E., Gao, G.F., Plaisier, C.L., Eddy, J.A., et al. (2018). The Immune Landscape of Cancer. Immunity *48*,

812-830.e14. .

Tiainen, S., Tumelius, R., Rilla, K., Hämäläinen, K., Tammi, M., Tammi, R., Kosma, V.M., Oikari, S., and Auvinen, P. (2015). High numbers of macrophages, especially M2-like (CD163-positive), correlate with hyaluronan accumulation and poor outcome in breast cancer. Histopathology *66*, 873–883. .

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B *58*, 267–288. .

Tomfohr, J., Lu, J., and Kepler, T.B. (2005). Pathway level analysis of gene expression using singular value decomposition. BMC Bioinformatics *6*, 1–11. .

Topalian, S.L., Taube, J.M., Anders, R.A., and Pardoll, D.M. (2016). Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. Nat. Rev. Cancer *16*, 275–287. .

Tötterman, T.H., Parthenais, E., Häyry, P., Timonen, T., and Saksela, E. (1980). Cytological and functional analysis of inflammatory infiltrates in human malignant tumors. III. Further functional investigations using cultured autochthonous tumor cell lines and freeze-thawed infiltrating inflammatory cells. Cell. Immunol. *55*, 219–226. .

Treilleux, I., Blay, J.-Y., Bendriss-Vermare, N., Ray-Coquard, I., Bachelot, T., Guastalla, J.-P., Bremond, A., Goddard, S., Pin, J.-J., Barthelemy-Dubois, C., et al. (2004). Dendritic cell infiltration and prognosis of early stage breast cancer. Clin. Cancer Res. an Off. J. Am. Assoc. Cancer Res. *10*, 7466–7474. .

Treutlein, B., Brownfield, D.G., Wu, A.R., Neff, N.F., Mantalas, G.L., Espinoza, F.H., Desai, T.J., Krasnow, M.A., and Quake, S.R. (2014). Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature *509*, 371–375. .

Tripathi, C., Tewari, B.N., Kanchan, R.K., Baghel, K.S., Nautiyal, N., Shrivastava, R., Kaur, H., Bhatt, M.L.B., and Bhadauria, S. (2014). Macrophages are recruited to hypoxic tumor areas and

acquire a pro-angiogenic M2-polarized phenotype via hypoxic cancer cell derived cytokines Oncostatin M and Eotaxin. Oncotarget *5*, 5350–5368. .

Tumeh, P.C., Harview, C.L., Yearley, J.H., Shintaku, I.P., Taylor, E.J.M., Robert, L., Chmielowski, B., Spasic, M., Henry, G., Ciobanu, V., et al. (2014). PD-1 blockade induces responses by inhibiting adaptive immune resistance. Nature *515*, 568–571. .

Turk, J.L. (1994). Inflammation: John Hunter's "A treatise on the blood, inflammation and gun-shot wounds." Int. J. Exp. Pathol. *75*, 385–395. .

Tymoszuk, P., Evens, H., Marzola, V., Wachowicz, K., Wasmer, M.H., Datta, S., Müller-Holzner, E., Fiegl, H., Böck, G., van Rooijen, N., et al. (2014). In situ proliferation contributes to accumulation of tumor-associated macrophages in spontaneous mammary tumors. Eur. J. Immunol. *44*, 2247–2262. .

UNDP (2019). Human Development Report 2019: Inequalities in Human Development in the 21st Century Jordan. United Nations Dev. Progr. 366. .

Varricchi, G., Galdiero, M.R., Loffredo, S., Marone, G., Iannone, R., Marone, G., and Granata, F. (2017). Are mast cells MASTers in cancer? Front. Immunol. *8*, 1–13. .

Vaupel, P., and Mayer, A. (2007). Hypoxia in cancer: significance and impact on clinical outcome. Cancer Metastasis Rev. *26*, 225–239. .

Venteicher, A.S., Tirosh, I., Hebert, C., Yizhak, K., Neftel, C., Filbin, M.G., Hovestadt, V., Escalante, L.E., Shaw, M.L., Rodman, C., et al. (2017). Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq. Science (80-.). *355*. .

Villani, A., Satija, R., Reynolds, G., Sarkizova, S., Shekhar, K., Fletcher, J., Griesbeck, M., Butler, A., Zheng, S., Lazo, S., et al. (2017). Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *4573*. .

Vogelstein, B., and Kinzler, K.W. (2004). Cancer genes and the pathways they control. Nat. Med. *10*,

789–799. .

Vogt, L., Schmitz, N., Kurrer, M.O., Bauer, M., Hinton, H.I., Behnke, S., Gatto, D., Sebbel, P., Beerli, R.R., Sonderegger, I., et al. (2006). VSIG4, a B7 family-related protein, is a negative regulator of T cell activation. J. Clin. Invest. *116*, 2817–2826. .

Vose, B.M., and Moore, M. (1985). Human tumor-infiltrating lymphocytes: a marker of host response. Semin. Hematol. *22*, 27–40. .

Vose, B.M., Vánky, F., Argov, S., and Klein, E. (1977). Natural cytotoxicity in man: activity of lymph node and tumor-infiltrating lymphocytes. Eur. J. Immunol. *7*, 753–757. .

Wan, L., Pantel, K., and Kang, Y. (2013). Tumor metastasis: moving new biological insights into the clinic. Nat. Med. *19*, 1450–1464. .

Wang, D.Y., Salem, J.-E., Cohen, J. V, Chandra, S., Menzer, C., Ye, F., Zhao, S., Das, S., Beckermann, K.E., Ha, L., et al. (2018). Fatal Toxic Effects Associated With Immune Checkpoint Inhibitors: A Systematic Review and Meta-analysis. JAMA Oncol. *4*, 1721–1728. .

Wang, J., Zhuang, Z.-G., Xu, S.-F., He, Q., Shao, Y.-G., Ji, M., Yang, L., and Bao, W. (2015). Expression of CCL2 is significantly different in five breast cancer genotypes and predicts patient outcome. Int. J. Clin. Exp. Med. *8*, 15684–15691. .

Ward, C., Langdon, S.P., Mullen, P., Harris, A.L., Harrison, D.J., Supuran, C.T., and Kunkler, I.H. (2013). New strategies for targeting the hypoxic tumour microenvironment in breast cancer. Cancer Treat. Rev. *39*, 171–179. .

Wculek, S.K., Cueto, F.J., Mujal, A.M., Melero, I., Krummel, M.F., and Sancho, D. (2020). Dendritic cells in cancer immunology and immunotherapy. Nat. Rev. Immunol. *20*, 7–24. .

Welch, D.R., Schissel, D.J., Howrey, R.P., and Aeed, P.A. (1989). Tumor-elicited polymorphonuclear cells, in contrast to" normal" circulating polymorphonuclear cells, stimulate invasive and metastatic potentials of rat mammary adenocarcinoma cells. Proc. Natl. Acad. Sci. *86*,

5859–5863. .

Whiteside, T.L., Demaria, S., Rodriguez-Ruiz, M.E., Zarour, H.M., and Melero, I. (2016). Emerging Opportunities and Challenges in Cancer Immunotherapy. Clin. Cancer Res. an Off. J. Am. Assoc. Cancer Res. *22*, 1845–1855. .

Wilkerson, M.D., and Hayes, D.N. (2010). ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. Bioinformatics *26*, 1572–1573. .

Wilson, W.R., and Hay, M.P. (2011). Targeting hypoxia in cancer therapy. Nat. Rev. Cancer *11*, 393–410. .

Wislez, M., Fujimoto, N., Izzo, J.G., Hanna, A.E., Cody, D.D., Langley, R.R., Tang, H., Burdick, M.D., Sato, M., Minna, J.D., et al. (2006). High expression of ligands for chemokine receptor CXCR2 in alveolar epithelial neoplasia induced by oncogenic kras. Cancer Res. *66*, 4198–4207. .

Witz, I.P. (1977). Tumor-bound immunoglobulins: In situ expressions of humoral immunity. Adv. Cancer Res. *25*, 95–148. .

Witz, I.P., and Levy-Nissenbaum, O. (2006). The tumor microenvironment in the post-PAGET era. Cancer Lett. *242*, 1–10. .

Wu, H., Han, Y., Rodriguez Sillke, Y., Deng, H., Siddiqui, S., Treese, C., Schmidt, F., Friedrich, M., Keye, J., Wan, J., et al. (2019). Lipid droplet-dependent fatty acid metabolism controls the immune suppressive phenotype of tumor-associated macrophages. EMBO Mol. Med. *11*, 8–10. .

Wu, S.Z., Roden, D.L., Wang, C., Holliday, H., Harvey, K., Cazet, A.S., Murphy, K.J., Pereira, B., Al-Eryani, G., Bartonicek, N., et al. (2020). Stromal cell diversity associated with immune evasion in human triple-negative breast cancer. EMBO J. *39*. .

Wyckoff, J.B., Wang, Y., Lin, E.Y., Li, J., Goswami, S., Stanley, E.R., Segall, J.E., Pollard, J.W., and Condeelis, J. (2007). Direct visualization of macrophage-assisted tumor cell intravasation in

mammary tumors. Cancer Res. *67*, 2649–2656. .

Wynn, T.A., Chawla, A., and Pollard, J.W. (2013). Macrophage biology in development, homeostasis and disease. Nature *496*, 445–455. .

Xiang, X., Wang, J., Lu, D., and Xu, X. (2021). Targeting tumor-associated macrophages to synergize tumor immunotherapy. Signal Transduct. Target. Ther. *6*. .

Xu, X., Grijalva, A., Skowronski, A., Van Eijk, M., Serlie, M.J., and Ferrante, A.W. (2013). Obesity activates a program of lysosomal-dependent lipid metabolism in adipose tissue macrophages independently of classic activation. Cell Metab. *18*, 816–830. .

Xue, J., Schmidt, S. V., Sander, J., Draffehn, A., Krebs, W., Quester, I., DeNardo, D., Gohel, T.D., Emde, M., Schmidleithner, L., et al. (2014). Transcriptome-Based Network Analysis Reveals a Spectrum Model of Human Macrophage Activation. Immunity *40*, 274–288. .

Yan, B., Wei, J.-J., Yuan, Y., Sun, R., Li, D., Luo, J., Liao, S.-J., Zhou, Y.-H., Shu, Y., Wang, Q., et al. (2013). IL-6 cooperates with G-CSF to induce protumor function of neutrophils in bone marrow by enhancing STAT3 activation. J. Immunol. *190*, 5882–5893. .

Yang, J., Liao, D., Chen, C., Liu, Y., Chuang, T.H., Xiang, R., Markowitz, D., Reisfeld, R.A., and Luo, Y. (2013). Tumor-associated macrophages regulate murine breast cancer stem cells through a novel paracrine egfr/stat3/sox-2 signaling pathway. Stem Cells *31*, 248–258. .

Yao, F., Coquery, J., and Lê Cao, K.A. (2012). Independent Principal Component Analysis for biologically meaningful dimension reduction of large biological data sets. BMC Bioinformatics *13*, 1–15. .

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). ClusterProfiler: An R package for comparing biological themes among gene clusters. Omi. A J. Integr. Biol. *16*, 284–287. .

Yu, P.F., Huang, Y., Han, Y.Y., Lin, L.Y., Sun, W.H., Rabson, A.B., Wang, Y., and Shi, Y.F. (2017). TNFα-activated mesenchymal stromal cells promote breast cancer metastasis by recruiting

CXCR2(+) neutrophils. Oncogene *36*, 482–490. .

Zajac, A.J., Blattman, J.N., Murali-Krishna, K., Sourdive, D.J., Suresh, M., Altman, J.D., and Ahmed, R. (1998). Viral immune evasion due to persistence of activated T cells without effector function. J. Exp. Med. *188*, 2205–2213. .

Zappia, L., and Oshlack, A. (2018). Clustering trees: a visualization for evaluating clusterings at multiple resolutions. Gigascience *7*, giy083. .

Zhang, Y., and Zhang, Z. (2020). The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. Cell. Mol. Immunol. *17*, 807–821. .

Zhang, H., Lu, H., Xiang, L., Bullen, J.W., Zhang, C., Samanta, D., Gilkes, D.M., He, J., and Semenza, G.L. (2015). HIF-1 regulates CD47 expression in breast cancer cells to promote evasion of phagocytosis and maintenance of cancer stem cells. Proc. Natl. Acad. Sci. U. S. A. *112*, E6215-23. .

Zhang, X., Li, T., Liu, F., Chen, Y., Yao, J., Li, Z., Huang, Y., and Wang, J. (2019). Comparative Analysis of Droplet-Based Ultra-High-Throughput Single-Cell RNA-Seq Systems. Mol. Cell *73*, 130-142.e5. .

Zhao, Y., Cai, H., Zhang, Z., Tang, J., and Li, Y. (2021). Learning interpretable cellular and gene signature embeddings from single-cell transcriptomic data. Nat. Commun. *12*, 1–15. .

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., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat. Commun. *8*, 14049. .

Zhu, Y., Herndon, J.M., Sojka, D.K., Kim, K.-W., Knolhoff, B.L., Zuo, C., Cullinan, D.R., Luo, J., Bearden, A.R., Lavine, K.J., et al. (2017). Tissue-Resident Macrophages in Pancreatic Ductal Adenocarcinoma Originate from Embryonic Hematopoiesis and Promote Tumor Progression.

Immunity *47*, 323-338.e6. .

Zilionis, R., Nainys, J., Veres, A., Savova, V., Zemmour, D., Klein, A.M., and Mazutis, L. (2017). Single-cell barcoding and sequencing using droplet microfluidics. Nat. Protoc. *12*, 44–73. .

Zois, C.E., and Harris, A.L. (2016). Glycogen metabolism has a key role in the cancer microenvironment and provides new targets for cancer therapy. J. Mol. Med. *94*, 137–154. .

Zou, J.-M., Qin, J., Li, Y.-C., Wang, Y., Li, D., Shu, Y., Luo, C., Wang, S.-S., Chi, G., Guo, F., et al. (2017). IL-35 induces N2 phenotype of neutrophils to promote tumor growth. Oncotarget *8*, 33501–33514. .

Zucker, A., Winter, A., Lumley, D., Karwowski, P., Jung, M.-K., and Kao, J. (2020). Prognostic role of baseline neutrophil-to-lymphocyte ratio in metastatic solid tumors. Mol. Clin. Oncol. *13*, 25. .