Spatial Transcriptomic Profiling of Bone Marrow in Myeloproliferative Neoplasms



Emily Thomas

Green Templeton College

University of Oxford

Department of Engineering Science

supervised by Jens Rittscher, Bethan Psaila and Daniel Royston

Trinity 2025

Acknowledgements

First, I would like to thank my supervisors, Professor Daniel Royston, Professor Bethan Psaila, and Professor Jens Rittscher, for their support throughout this PhD, for the opportunity to work on such novel datasets and freedom to explore analysis solutions.

Thanks also to Dr Rosalin Cooper and Dr Hosuk Ryou for their mentorship, particularly on the CALR/CycleGAN project, which, though not included here, was an invaluable learning experience.

I'm also grateful to the Spatial and Computational Pathology Group, for their insightful discussions, which helped shape many of the ideas and analyses in this thesis.

Finally thank you very much to my family, friends and loved ones for their constant love and support throughout this PhD, especially during the tougher moments.

Generative AI was used to improve grammar and clarity; all text was written by me. I was supported by EPSRC and CRUK funding.

Abstract

This thesis presents a multi-modal, spatially informed profile of bone marrow in myeloproliferative neoplasms (MPNs), integrating spatial transcriptomics and machine learning based image analysis.

In Chapter 3, I develop a pipeline for processing spatial transcriptomic data from bone marrow biopsies. This includes segmentation of non-uniform cells and construction of graphs that preserve neighbourhood relationships across heterogeneous morphology. Cell annotation is improved through multi-layer segmentation and integration with scRNA-seq references, including one enriched for stromal cells. Macrostructures such as bone, adipocytes, vasculature and fibrosis maps are incorporated to provide spatial context.

Chapter 4 presents a spatial transcriptomic analysis of bone marrow from MPN patients and healthy donors. I identify ten cellular neighbourhoods (CNs), several aligning with known regions. CN2, associated with higher fibrosis, may represent early fibrosis or a cellular correlate. Spatial expansion is a MF feature and, in ET and MF, expanded CNs become more distant from CN5 (a lymphoid aggregate), indicating impaired immunosurveillance. At higher fibrosis grades, I observe novel associations between Osteogenic Mesenchymal Stem Cells (Osteo-MSCs) and megakaryocytes. Gene expression suggests these Osteo-MSCs may correlate with a known profibrotic subset. CN1, CN3, and CN4 resist expansion and maintain spatial organisation, while CN3, CN4, and CN9, enriched for HSPCs, are flanked by distinct lineages, supporting spatially encoded niches. Finally, megakaryocyte morphology reflects transcriptional state. This work supports integrating spatial transcriptomics into haematopathology workflows and clinical decision-making.

Contents

List of Figures **viii**

1 Introduction **1**

1.1 Background and Context 1
1.2 Unmet Need 1
1.3 Methodology 2
1.4 Aims 2
1.5 Chapter Overview 3
1.6 Summary of Contributions 4
1.7 List of outputs 5
1.7.1 Publications 5
1.7.2 Invited Talks and Awards 5

2 Literature Review **6**

2.1 Myeloproliferative Neoplasms (MPNs) 6
2.1.1 An Introduction to MPNs 6
2.1.2 Diagnosis of MPNs 8
2.1.3 Treatment of MPNs 10
2.2 Initiation of MPNs and the importance of the microenvironment 11
2.2.1 Driver Mutations 11
2.2.2 Beyond the Driver Mutation 12
2.2.3 The Bone Marrow Microenvironment and MPN 13
2.3 The Bone Marrow Microenvironment 14
2.3.1 Importance of the Bone Marrow Microenvironment 14
2.3.2 Components of the bone marrow microenvironment 14
2.3.3 The HSPC niche 18
2.4 The Bone Marrow Microenvironment in MPN 22
2.4.1 Heterogeneity in Myelofibrosis 22
2.4.2 MPN and the HSPC niche 23
2.4.3 The Inflammatory Cytokine landscape of MF 24

2.4.4 The Spleen in MPN . . . 25
2.5 Fibrosis in MPN . . . 26
2.5.1 Fibrosis across MPN . . . 26
2.5.2 Fibrosis Categorisation . . . 26
2.5.3 Mechanisms of Fibrosis Development . . . 27
2.5.4 Automated Fibrosis Detection . . . 30
2.6 Spatial Transcriptomics . . . 31
2.6.1 The Limitations of Clinically Available H&E and IHC . . . 31
2.6.2 The Advantage of Spatial Gene Expression . . . 32
2.6.3 The Spatial Transcriptomic Landscape . . . 33
2.7 Spatial Analysis Studies of the Bone Marrow . . . 33
2.7.1 Spatial Analysis of Non-MPN Bone Marrow: A Timeline . . . 34
2.7.2 Spatial Analysis of MPN Bone Marrow . . . 37
2.8 Characteristics of Spatial Data . . . 38
2.8.1 Modifiable Areal Unit Problem (MAUP) . . . 38
2.8.2 Spatial Autocorrelation . . . 39
2.8.3 Spatial Heterogeneity . . . 39
2.8.4 Additional Considerations for Biological Tissue . . . 40
2.9 Analysing Spatial Data . . . 41
2.9.1 Unsupervised Tissue Domains . . . 42
2.9.2 Spatial Statistics . . . 47

3 Preparing the Xenium Dataset **49**
3.1 Author Contributions . . . 49
3.2 Introduction . . . 50
3.2.1 Aims . . . 50
3.2.2 Xenium . . . 50
3.3 Data . . . 51
3.3.1 Sample Selection . . . 51
3.3.2 Sample Preparation . . . 51
3.3.3 Cohort Metadata . . . 52
3.4 Results . . . 53

3.4.1 Sample Quality Control . . . 53
3.4.2 Quality from a Haemapathologist perspective . . . 54
3.4.3 H&E to improve segmentation . . . 58
3.4.4 Filtering of low quality cells . . . 60
3.4.5 Batch Correction . . . 62
3.4.6 Cell Annotation . . . 65
3.4.7 Segmentation and Cell Annotation . . . 75
3.4.8 Final Annotation . . . 89
3.4.9 Annotation Resolution . . . 89
3.4.10 Incorporating Macrostructures from H&E . . . 94
3.4.11 Graph Creation . . . 97
3.5 Discussion . . . 100
3.6 Future Work . . . 101

4 Spatial Transcriptomics Analysis of the MPN Bone Marrow 102
4.1 Author Contributions . . . 103
4.2 Introduction . . . 103
4.3 Global cellular abundance in the Bone Marrow in MPNs . . . 104
4.3.1 Introduction . . . 104
4.3.2 Methodology . . . 104
4.3.3 Results . . . 106
4.3.4 Discussion . . . 114
4.3.5 Conclusion . . . 115
4.4 Global cellular expression in the Bone Marrow in MPNs . . . 115
4.4.1 Introduction . . . 115
4.4.2 Methodology . . . 116
4.4.3 Results . . . 116
4.4.4 Discussion . . . 120
4.4.5 Conclusions . . . 121
4.5 The Unsupervised Bone Marrow Architecture . . . 122
4.5.1 Introduction . . . 122
4.5.2 Methodology . . . 122

4.5.3 Results . . . 126
4.5.4 Discussion . . . 150
4.6 Fibrosis and Bone Marrow Changes in MF . . . 154
4.6.1 Introduction . . . 154
4.6.2 Methodology . . . 155
4.6.3 Results . . . 158
4.6.4 Discussion . . . 174
4.7 Gene Expression and Cytomorphology . . . 177
4.7.1 Introduction . . . 177
4.7.2 Methodology . . . 178
4.7.3 Results . . . 179
4.7.4 Discussion . . . 182
4.8 Limitations . . . 183
4.9 Future Work . . . 184

5 Future Work **186**
5.1 Summary of Contributions . . . 186
5.2 Future Work . . . 187

References **189**

List of Figures

2.1 Key cells of the myeloid and lymphoid branches of the immune system. 7

3.1 Xenium Bone marrow trephines 53
3.2 Exemplar of Spatially unpreserved regions 54
3.3 Transcripts per sample by diagnosis 56
3.4 Transcripts per sample by batch 56
3.5 Transcripts and Spatially Unpreserved Regions 56
3.6 Megakaryocytes detected by H&E by Diagnosis 59
3.7 Integration of Megakaryocytes into Segmentation 61
3.8 Venn Diagrams of Filtering Steps 62
3.9 Histogram of transcripts per cell 63
3.10 PCA by batch and diagnosis before and after harmony correction 65
3.11 PCA by batch before and after harmony correction 66
3.12 UMAP of key lineage markers 67
3.13 Hierarchy of Louvain Clusters Resolutions 69
3.14 HSPC determinaion and coexpression of other lineage markers (cropped) 70
3.15 Unsupervised clustering summary 72
3.16 DotPlot of Unsupervised Clustering 73
3.17 SingleR median predictions score for each cell type using Tan Lab Paper 75
3.18 SingleR assignment confidence using Tan Lab reference 76
3.19 SingleR assigned cell types for 15um voronoi segmentation 77
3.20 Clustree of SingleR assigned annotations from Tan Lab and Azimuth Dataset, broader populations 78
3.21 Comparison of the assignation per Tan Lab and Azimuth Dataset by SingleR for the cells in each of our unsupervised clusters 79
3.22 Comparison of the assignation per Tan Lab and Azimuth Dataset by DotR for the cells in each of our unsupervised clusters 80
3.23 Location within cell of Lineage Markers 81
3.24 Examplar of the resegmentation of the same cells at different levels of segmentation 84
3.25 Matched indexes across re-segmentations 84

3.26 Summary of re-annotation using the concordance between unsupervised annotations of the 15um segmentations and singleR annotations of the 5um and nucleus segmenations . . . 85
3.27 SingleR annotations which are the same for resegmented cells (5um and nuclear) as in 15um final annotation for each unsupervised cluster . . . 86
3.28 SingleR annotations which are the same in 5um and nuclear segmentations for each unsupervised cluster . . . 87
3.29 Cells still to be determined within each unsupervised cluster . . . 88
3.30 Final annotations of 15 um Voronoi clusters before and after applying overrides 90
3.31 Number of cells assigned to each cell type . . . 91
3.32 Dot plot of marker gene expression in Final Annotation . . . 91
3.33 UMAP of Final Annotations . . . 92
3.34 Difference in percentage of cells expressing lineage markers between unsupervised and final annotations in each cell type . . . 93
3.35 Dot plot of stromal marker genes . . . 94
3.36 Final Bone Segmentation visualised on H&E for two samples . . . 95
3.37 Fat segmentation and challenges around bone . . . 95
3.38 CIF score of each tile per sample . . . 96
3.39 Distribution of fibrosis scores aross samples by fibrosis grade . . . 97
3.40 Comparison of Delaunay graph constructions with and without distance filtering. 98
3.41 Distance from a Megakaryocyte to other cells . . . 99
3.42 Comparison of Delaunay Graphs at Two Radius Thresholds . . . 99
3.43 Comparison of boundary-based graphs with and without fat segmentation . . . 100

4.1 BCV and MDS plots for cell abundance analysis . . . 106
4.2 Differential Abundance at level 3 annotation across whole trephines . . . 107
4.3 Differential Abundance of Megakaryocyte Phenotypes and Non-phenotyped across conditions . . . 108
4.4 Differential Abundance at level 3 with refined stromal populations . . . 109
4.5 Cell Type Percentage of Cellular Marrow at level 3 . . . 110
4.6 Cell Type Percentage of Cellular and Fatty Marrow at level 3 . . . 111
4.7 Number of each cell type per mm2 of cellular marrow at level 3 . . . 113
4.8 Number of each cell type per mm2 of Cellular and Fatty Marrow at level 3 . . . 114
4.9 Cumulative total of the number times a gene is differentially expressed . . . 117
4.10 Number of times gene is differentially expressed . . . 118

4.11 Count of differentially expressed genes across all cells for all comparisons 118
4.12 Volcano plots for Endothelial and Megakaryocyte populations 119
4.13 Stability of cellcharter cluster numbers 124
4.14 CellCharter clustering and annotations for MF and Normal 124
4.15 CellCharter clustering at 3-hop with varying cluster numbers 125
4.16 Schematic of Bone Regions in a Trephine Section 126
4.17 CellCharter clustering at 3-hop, with 10 clusters, for MF and Normal 127
4.18 CellCharter clustering at 3-hop, with 10 clusters in MF samples 127
4.19 CellCharter clustering at 3-hop, with 10 clusters in ET samples 128
4.20 CellCharter clustering at 3-hop, with 10 clusters in PV samples 129
4.21 CellCharter clustering at 3-hop, with 10 clusters in Normal samples 130
4.22 Cell type enrichment across layers for Cellcharter 10 clusters 130
4.23 Cell type enrichment across layers for Cellcharter 10 clusters, at level 1 annotation with refined stromal groups 131
4.24 Percentage of HSPCs in each CN per patient 132
4.25 DotPlot showing differential gene expression of HSPCs per CN vs all other CNs 133
4.26 CellCharter region variation across MPN subtypes 134
4.27 Differential Abundance of CellCharter Clusters across conditions 135
4.28 CellCharter largest clusters across diagnoses 136
4.29 Distribution of CellCharter Cluster Sizes (0–5) 137
4.30 Distribution of CellCharter Cluster Sizes (6–9) 138
4.31 Median cluster size of the 10 largest clusters per CN per sample across conditions 140
4.32 CellCharter neighbour enrichment across conditions 141
4.33 Enrichment of cell types across CellCharter clusters 142
4.34 CellCharter neighborhood enrichment contrasts vs Normal 143
4.35 Median Proportion of CellCharter Regions Across Distances to Bone 144
4.36 Differential Abundance of CellCharter Regions in each Bone region 145
4.37 Differential Abundance of CellCharter Regions Across Bone Regions 147
4.38 CIF score by CellCharter CN within conditions 149
4.39 Boxplot of CIF scores per sample 150
4.40 CIF score maps for four MF samples with the highest median fibrosis 150
4.41 Correlation between the proportion of CN region 8 and the proportion of each CIF grade per trephine in MF samples 151

4.42 Spatial distribution of fibrosis scores and categories in a trephine section 156

4.43 Comparison of cell-type spatial neighborhoods in Normal and MF samples. Radii are shown at 50-300 pixels in 50-pixel intervals. Cell center points are indicated by dots; cell colours indicate cell types. 157

4.44 Median of the proportion of each cell type per MF sample as CIF cif score increases158

4.45 Cells with a correlation with CIF score 159

4.46 Differential abundance of cell types across MF fibrosis grades 161

4.47 Cell Type Abundance across fibrotic high and low spots in MF 162

4.48 Neighbourhood enrichment analysis across MF grades (MF-0 to MF-2) 163

4.49 Neighbourhood enrichment differences relative to MF-0 across fibrosis grades . . 164

4.50 Co-occurrence of cell types across MF fibrosis grades (MF-0 to MF-2) at radii 150–500 pixels 165

4.51 Co-occurrance of Osteo-MSCs across MF grades at 250 pixel radii 167

4.52 Differential Abundance of CNs across fibrosis grades in MF 168

4.53 Number of differentially expressed genes per comparison per cell type in MF samples 170

4.54 Volcano plots for Osteo-MSC differential comparisons across MF fibrosis grades (excluding MF-3) 171

4.55 Representative image of the megakaryocyte phenotypes from 0-8 178

4.56 Comparison of phenotyped and unphenotyped megakaryocytes 180

4.57 Megakaryocyte Differential Gene Expression by Phenotype 181

1

Introduction

1.1 Background and Context

Myeloproliferative neoplasms (MPNs) arise in the bone marrow following acquisition of a driver mutation in haematopoietic stem cells (HSCs), which give rise to all blood and immune cells. Mutated HSCs gain a clonal advantage, leading to over-proliferation of blood lineages. MPNs include three main subtypes: Essential Thrombocythemia (ET), Polycythaemia Vera (PV), and Myelofibrosis (MF), characterised by excess platelets, red blood cells, or marrow fibrosis, respectively. MF has the poorest prognosis, and both ET and PV can progress to MF, often via a pre-fibrotic stage, and ultimately to secondary acute myeloid leukaemia (AML).

This thesis investigates immune population differences between healthy donors and patients with fibrotic and non-fibrotic MPNs using a spatially resolved view of patient bone marrow.

1.2 Unmet Need

Understanding the single-cell landscape of MPNs is crucial for two main reasons:

Action Point 1: Gaining insight into the MPN landscape for early detection

MPNs often have a long latency, up to 30 years, between driver mutation acquisition and symptom onset. Although ET, PV, and MF share the same mutations, they manifest as distinct disease phenotypes. This suggests that microenvironmental factors may trigger clonal expansion and drive progression, including transformation to fibrosis or AML. Mapping the immune landscape may help define the microenvironments underpinning these phenotypes.

Action Point 2: Enhancing diagnostic utility of bone marrow biopsies

Bone marrow biopsies, routinely analysed with H&E and IHC, form a large archival dataset. While pathologists can distinguish MPN subtypes morphologically, their assessments are qualitative, and machine learning tools often lack interpretability. High-resolution immune profiling may uncover quantitative features that complement current diagnostics and better leverage archival tissue.

Addressing these two needs may lead to improved biomarkers for disease progression and treatment response, informing both clinical care and trial design where curative options remain limited.

1.3 Methodology

I will investigate the immune landscape of MPNs using a Xenium spatial transcriptomics dataset of 32 bone marrow trephine samples from 30 healthy donors and MPN patients (ET, PV, Pre-MF, MF), paired with H&E images. It uses a pan-cancer panel targeting 377 genes.

1.4 Aims

In the Xenium spatial transcriptomics dataset, I address Action Point 1 by profiling bone marrow from healthy donors and MPN subtypes. This dataset captures cells in situ, including stromal populations, enabling assessment of global and local spatial changes. Due to its cost, spatial profiling is limited to a few samples. I therefore begin to extend the analysis to features quantifiable in routine stains, such as fibrosis grade and megakaryocyte morphology, to characterise spatial immune variation in a morphological context.

Linking spatial or transcriptional features to morphology lays the foundation for Action Point 2: enhancing clinical interpretation of diagnostic bone marrow trephines.

1.5 Chapter Overview

Chapter 1: Introduction

This chapter outlines the thesis motivation, key unmet needs, and the methodologies used. It provides an overview of the thesis structure, highlights contributions to the field, and lists relevant outputs.

Chapter 2: Literature Review

This chapter reviews current knowledge of myeloproliferative neoplasms (MPNs), including disease initiation, diagnosis, and treatment. It covers bone marrow organisation, fibrosis development, and spatial transcriptomic studies of the marrow. It then focuses on methodologies and analytical approaches such as unsupervised tissue domain identification and spatial statistics for spatial transcriptomics.

Chapter 3: Preparing the Xenium Dataset

This chapter presents a pipeline for processing Xenium spatial transcriptomics data. It introduces a novel method for integrating irregularly shaped cells into segmentation, a novel graph construction approach to preserve cell–cell neighbourhoods, and improved annotation via multiple segmentation layers and scRNA-seq references. It also incorporates macrostructures (e.g., adipocytes, vasculature, bone) and machine learning–derived features like a spatial fibrosis map.

Chapter 4: Spatial Transcriptomics Analysis of the MPN Bone Marrow

This chapter presents a spatial transcriptomic analysis of the bone marrow. It integrates global differential analyses with unsupervised identification of spatial neighbourhoods, characterising cellular composition, interaction, and gene expression across fibrosis grades. Together, these approaches build a comprehensive view of the bone marrow microenvironment in both healthy individuals and patients with myeloproliferative neoplasms (MPNs). Finally it links gene expression to megakaryocyte morphology, bridging spatial transcriptomics with histopathology.

1.6 Summary of Contributions

Chapter 3: Preparing the Xenium Dataset

This chapter details a novel pipeline for processing Xenium spatial transcriptomic data of the bone marrow across all MPN subtypes and healthy controls. Machine learning outputs and fibrosis maps are integrated into the dataset to enhance spatial context.

Innovations include custom methods for incorporating irregular cell morphologies (e.g., megakaryocytes) into segmentation, and a Voronoi-based graph construction that preserves neighbourhood structure even when dealing with non-uniform cell morphologies. Cell type annotation is improved using multi-layered segmentations and scRNA-seq references, including a stromal-enriched dataset. Macrostructures like adipocytes, bone, and vasculature are also incorporated.

Chapter 4: Spatial Transcriptomics Analysis of the MPN Bone Marrow

This chapter presents the first spatial transcriptomic analysis of the bone marrow across all MPN subtypes. Ten distinct cellular neighbourhoods (CNs) were identified, many aligning with known histological regions. CN9, enriched in megakaryocytes, endothelial cells, and HSPCs, expanded in ET and MF and may represent a precursor fibrotic niche. CN2, linked to high fibrosis scores and spatial shifts from the bone wth disease progression, also emerged as a potential early fibrotic microenvironment.

Stable CNs (CN1, CN3, CN4) maintained spatial architecture across disease states and resisted spatial expansion, indicating a conserved marrow organisation previously undescribed. CN8 expansion in MF correlated with a fibrotic pattern enriched for MF-2, potentially marking a late-stage fibrotic pathway. High fibrosis grades showed novel spatial associations between osteo-MSCs and megakaryocytes, and DC and T cells, along with downregulation of Wnt inhibitors in osteo-MSCs at higher fibrosis grades, suggesting recapitulation of a NCAM1+ profibrotic subtype.

Megakaryocyte morphology was linked to transcriptional state, with smaller, immature cells being more proliferative. These findings support the integration of spatial transcriptomics into clinical workflows and fibrosis monitoring.

1.7 List of outputs

1.7.1 Publications

- Cooper, R.*, **Thomas, E.***, Dawood, M.*, Ryou, H., Sozanska, A., Pescia, C., McCallion, O., Teague, R., Gupta, M., Hester, J. and Issa, F., 2025. Redefining the topology of the human bone marrow using augmented spatial transcriptomic analysis. *bioRxiv, pp.2025-12*.
- Sozanska, A.M., Pescia, C., **Thomas, E.**, Royston, D.J. and Cooper, R.A., 2025. Entering the era of spatial transcriptomics: opportunities and challenges for pathology. *Diagnostic Histopathology*.
- Dawood, M., **Thomas, E.**, et al. Unsupervised Discovery of Spatiotypes and Context-Aware Graph Neural Networks for Modeling Clinical Endpoints, *MICCAI, 2025*
- Ryou, H., **Thomas, E.**, Wojciechowska, M., Harding, L., Tam, K.H., Wang, R., Hu, X., Rittscher, J., Cooper, R. and Royston, D., 2025. Reticulin-Free Quantitation of Bone Marrow Fibrosis in MPNs: Utility and Applications. *EJHaem*, *6*(2), p.e70005.
- Pescia, C., Sozanska, A.M., **Thomas, E.** and Cooper, R.A., 2025. Artificial intelligence in haematopathology: current perspective and future directions. *Diagnostic Histopathology*.
- Cooper, R.A.*, **Thomas, E.***, Sozanska, A.M., Pescia, C. and Royston, D.J., 2025. Spatial transcriptomic approaches for characterising the bone marrow landscape: pitfalls and potential. *Leukemia*, *39*(2), pp.291-295.
- Ryou, H., Lomas, O., Theissen, H., **Thomas, E.**, Rittscher, J. and Royston, D., 2023. Quantitative interpretation of bone marrow biopsies in MPN—What's the point in a molecular age?. *British Journal of Haematology*, *203*(4), pp.523-535.

1.7.2 Invited Talks and Awards

- **Invited Talk**, EMBO workshop "The Spatial Biology of Cancer", Francis Crick Institute, UK (September 2024)
- **Invited Talk**, Spatial Biology Summit, Stanford University, USA (September 2024)
- **$500 NIH travel award**, Spatial Biology Summit, Stanford University, USA (September 2024)
- **Invited Talk given by co-author**, EHA 2024
- **Invited Talk given by co-author**, ASH 2024
- **Invited Talk given by co-author**, EHA 2025

2

Literature Review

2.1 Myeloproliferative Neoplasms (MPNs)

2.1.1 An Introduction to MPNs

The bone marrow is the birthplace of all blood cells, which include those which make up the immune system. Although later maturation of immune cells may occur outside the bone marrow, for example in the thymus and secondary lymphoid organs, all blood cells are originally derived from haematopoietic stem cells (HSCs) which reside in the bone marrow. HSCs are multipotent, meaning they can give rise to many cell types. However, differentiation biases (e.g. towards platelet and myeloid or lymphoid) and various states (dormant, quiescent, activated) of HSCs are documented. The cause of these biases and states is still under investigation (Figure 2.1).

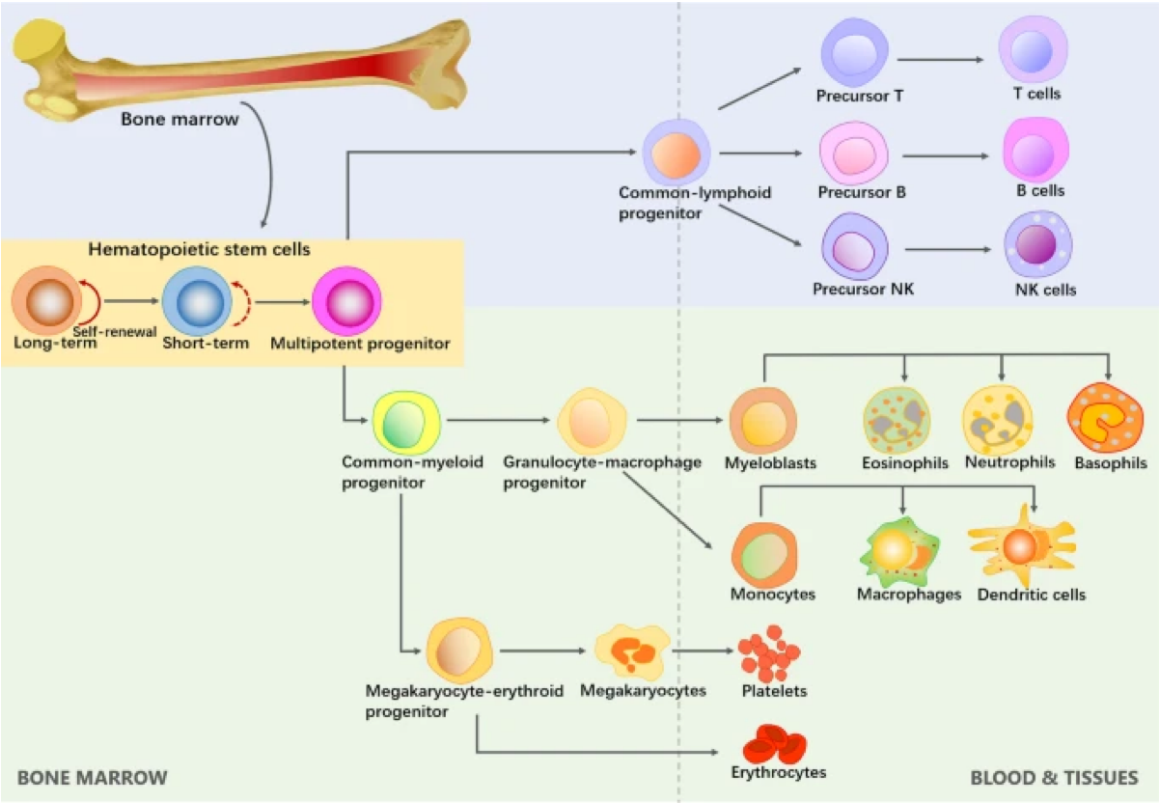


Figure 2.1: Key cells of the myeloid and lymphoid branches of the immune system. Taken from: Zhang et al. [1]

Many different types of blood cancer exist, including leukaemia, lymphoma, myeloma, myelodysplastic syndromes and myeloproliferative neoplasms (MPNs). Within this thesis I will focus on exploring the immune landscape of MPNs.

MPNs are a group of blood cancers that originate from haematopoietic stem cells (HSCs) in the bone marrow. They are a heterogeneous disease, characterized by proliferation of one or more blood lineages. Their overall prevalence is approximately 1 in 1,000–2,000 individuals, rising sharply with age [2].

MPNs include three principal subtypes, Essential Thrombocythaemia (ET), Polycythaemia Vera (PV), and Myelofibrosis (MF), with Pre-PMF often described as an intermediate stage between ET and MF. ET and PV present primarily with increased platelet production or red blood cells respectively. MF is hallmarked by increased fibrosis (excessive thickening/scarring of connective tissue) and cytopenias (decreased blood cell count). Thrombosis and haemorrhage risk is increased in all subtypes [3]. Although many PV and ET patients achieve near-normal life expectancy with appropriate treatment, all MPN subtypes can transform into fibrotic disease (MF five-year survival $\bar{3}5$ %) or acute leukaemia. ET and PV progresses to MF in 10-30%

of patients [4].Specifically, AML develops in 10–20 % of MF patients and in 1–4 % of those with ET or PV. Treatment decisions and risk of progression differ between and within each subgroup, making accurate distinctions vital. However, overlapping clinical and laboratory features between subgroups can make this challenging.

Immune dysregulation is known to be a key pathological feature of MPNs, facilitating clonal evolution, mediating symptoms, and leading to immunocompromise. Immunocompromise leaves patients vulnerable both to infection and failure of tumour immunosurveillance. To date, no study has fully delineated the immune landscape of each MPN subtype. Understanding the immune landscape of MPN may offer insight into the mechanism of initiation of MPNs, including how different subtypes arise from the same driver mutations, and provide insight into vulnerabilities for treatment.

In this thesis, I will therefore document the immune landscape of MPNs, aiming to identify subtype-specific patterns and vulnerabilities [5].

2.1.2 Diagnosis of MPNs

Treatment decisions and risk of progression differ between and within each subgroup, making accurate distinctions vital. However, overlapping clinical and laboratory features between subgroups can make this challenging. The 2016 WHO diagnostic criteria emphasizes the importance of incorporating clinical, morphological and genetic information for making accurate distinctions between MPNs [6].

Genetically, more than 95% of MPNs have a mutually exclusive driver mutation in either the JAK2, CALR or MPL gene. These mutations activate MPL-JAK-STAT signalling in the MPN stem cell [7]. The prevalence of each of these mutations differs between subtypes.

Morphologically, samples from the bone marrow are distinguished by clinicians through H&E and immunohistochemical staining. Diagnoses of MPN are generally made by experienced haemato-pathologists who assess histology slides of the bone marrow (in the context of clinical information). These slides are thin sections of bone marrow (4-5um) adhered to a glass slide [8].

To create the slides, the sample is placed in neutral buffered formalin for fixation and then decalcified in EDTA [9]. Counterstains can be added to provide contrast, for example haematoxylin and eosin. Haematoxylin stains the nucleus blue, whilst eosin stains the cationic group of proteins, causing the majority of structures in a tissue to stain red [8]. Every tissue being sent for pathology analysis will have a tissue slice stained with H&E for analysis. Under this stain, a cell's size, shape and location can be identified, including in relation to macrostructures such as the bone. Whilst some cell types are identifiable in this H&E stain, such as (a)typical megakaryocytes, and indeed some cellular structures can be observed e.g. erythrocyte islands, this is not the case for every cell or structure. Whilst expert, the assessment is largely qualitative. For MPN samples, other stains on sequential tissue slices of the tissue block will be requested to highlight the position of specific proteins in the tissue e.g. CD3. This will give more certainty to the location and number of cell types that can't be reliably distinguished from cell shape alone. This is known as immunohistochemistry (IHC).

Detection antibodies are added specific to the antigens of interest. The antibodies for IHC are bound to a chromogenic or fluorogenic conjugate, which can be detected either directly or on addition of an enzyme. The concentration of most antigens means that signal amplification is required. This is achieved by using a secondary multivalent antibody bound to a reporter as the conjugate. One of the most common reporters is horseradish peroxidase, which oxidizes a molecule to produce a brown precipitate. However, the non-linear nature can result with poor correlation with the original antigen of interest [10]. In addition, once a tissue is removed from their source, the tissue's oxygen and nutrient supply is restricted. This can cause changes in protein phosphorylation within minutes and alter protein levels over hours. Some proteins, such as Ki-67, are more fragile if the sample is not fixed rapidly. Thus sample preparation, in particular if analysing samples from a variety of sites, may introduce batch effects [8], [11].

Whilst the classifications from these MPN tissue slides are subjective, studies have found the use of bone marrow features reproducible. A range of 76-88% diagnostic consensus with the formal assessment was found in 2033 patient samples taken from several independent studies, including one which included a high numbers of "reactive" samples, for example due to chronic inflammation [6]. This suggests true morphological distinctions do exist between MPN subtypes and that there should be qantifiable changes in the bone marrow.

The potential mimicry of MPNs makes morphological classification harder. Beyond the three subtypes, pre-PMF (prefibrotic/early PMF) and masked PV (mPV) can mimic an ET-like phenotype. Characterisation studies are improving this distinction, especially between PV, pre-PMF and ET [6], [12], [13], [14]. Indeed, there is suggestion that perceived changes between MPN subtypes in a patient are actually highly dependent on initial diagnosis. A study of six haematopathologists found inter-rater agreement of 93% of 98 blinded samples, including 48 mPV, 31 PV and 19 control cases of JAK2 mutated ET, PM, MDS and various reactive conditions [15]. This suggests that identifying quantitative morphological distinctions, even for cases with similar presentation, is both possible and useful.

Whilst consensus on the final histological subtype is reproducible, experience of the pathologist is a key factor. Consensus in one study rose from 49% to 72% post a workshop chaired by the author of the 2008 WHO diagnostic criteria ([16]). The reproducibility of specific identification of individual morphological characteristics is also still controversial, with a range of 49% to 100% prior to the same workshop, with no significant effect after [16], [6].

Advancements have already been made in developing quantitative assessment of MPN histology. One recent tool focussed on changes between megakaryocyte cytological and topological features between reactive and MPN (PMF, ET and PV subtypes) (n=131) with an AUC= 0.98. The authors also identified 9 megakaryocyte subtypes, 8 of which were significantly associated with key MPN driver mutations. This suggests that new biological insights may be gained from more quantitative analysis of histology [17].

2.1.3 Treatment of MPNs

Currently, therapy for MPNs aims to reduce symptoms, such as abnormal spleen enlargement and improve abnormal blood counts. Unfortunately, treatment is only very rarely curative. The only potentially curative therapy is allogeneic stem cell transplantation from a healthy donor. However, this proceduce involves very intensive treatment and carries a high risk of treatment-related death and morbidity e.g. due to graft versus host disease and other complications. Treatments used in MPNs are focused on normalising blood counts and improving symptoms and spleen size. This might include cytoreductive agents such as hydroxycarbamide or interferon. Patients with myelofibrosis are often treated with JAK inhibitors. The two most commonly used JAK inhibitors (ruxolitinib and momelotinib) reduce signalling via both JAK1 and JAK2.

The tumour microenvironment refers to the bone marrow environment surrounding the diseased cells, for example non-clonal immune cells, niche cells, signalling molecules and structural tissue. Recent studies have shown a high latency between the acquisition of known MPN driver mutations and the development of disease, with one theory suggesting that micro-environmental change may be needed to provide a hospitable niche for disease initiation [2]. Likewise, inflammation has been associated with development of fibrosis [18]. Understanding how cells interact in the microenvironment gives increased resolution into immune cell (and stromal cell) interactions and potential ways that these could be disrupted.

Current therapies for MPN have documented immunomodulatory effects. Ruxolitinib decreases circulating inflammatory cytokines, a signalling molecule produced by cells. Other treatments, such as IFN-alpha, are artificial mimics of cytokines produced by the immune system [5]. Immunotherapy is another branch of treatment which manipulates the body's immune system to attack cancerous cells. Successes include (CAR)-T cells and immune checkpoint blockades. Whilst more modern immunotherapy strategies have been pioneered in other types of blood cancers, a recent review has argued the challenge is harder in MPNs due to the immunologically hostile microenvironment. While patients with mutant CALR-driven MPNs have expression of a cancer cell-specific neoantigen on their MPN cells (mutant CALR), there is no MPN-specific target known for patients with JAK2V617F or MPL driven MPN. Also, the cellular and molecular heterogeneity between patients and MPN subtypes and the immune dysregulation caused by the disease make it harder to produce reliable anti-tumour immunity [5]. Characterization of the immune interactions may help with development of effective immunotherapeutic strategies for MPNs.

2.2 Initiation of MPNs and the importance of the microenvironment

2.2.1 Driver Mutations

HSCs are the cell of origin for MPNs. MPNs have been used as a model of cancer evolution as most patients present in early disease phase with only a single cancer driver mutation, but the cancer then progresses slowly with time. In MPNs, HSCs acquire a driver mutations that leads

to upregulation of JAKSTAT signalling pathway. The genes affected are JAK2, calreticulin (CALR) and the thrombopoietin (TPO) receptor MPL in 95% of patients [4]. The mutations in these genes occur mutually exclusively. The mutant HSCs gain an advantage over wildtype (WT) HSCs and this results in clonal expansion and the resulting MPN phenotype.

The JAK2V617F mutation is the most common driver mutation in MPN patients. It is detected in 95% of all PV patients and more than 50% of ET/MF [19]. This mutation leads to constitutive activation of JAK2. Mutations in the CALR gene are classed as type 1 or type 2, depending on the mechanism of frameshift, resulting in a changed C-terminus. Type 1 resulting from CALRdel52 compared to type 2 CALRins5. CALR mutations occur in approximately 30-50% of ET or MF patients and can lead to a targetable mutant CALR (mutCALR) protein on the cell surface. mutCALR binds and activates MPL in the endoplasmic reticulum. MPL mutations are the least common and occur in approximately 5-8% of ET and MF patients. MPLW515L leads to constitutive MPL activation leading to aberrent activation of the JAK-STAT pathway. Additional mutations can also affect prognosis and response to therapy, for example in ASXL1, SRSF2 and U2AF1 which result in higher risk of progression [20]. Additional mutations to driver mutations are associated with worse survival. Williams et al. [2] suggest that there is parallel evolution in clonal expansions within the same patient i.e. cells evolve similarly even though not from the same common ancestor.

Previously, the prevailing hypothesis was that acquisition of driver mutations within HSCs initiated MPNs. However, increasing avenues of research are looking into whether a permissive bone marrow environment is necessary for the initiation of MPNs and their progression.

2.2.2 Beyond the Driver Mutation

The presence of an MPN driver mutation alone isn't enough to implement the phenotype. Both JAK2 V617F and CALR type 1 and 2 mutations have been detected in healthy individuals who do not have an MPN. Within a Danish cohort of approximately 20 000, 3.1% (613 participants) had a JAK2 V617F mutation, of which only 16 had MPN. The CALR mutation was found in 0.16% of participants, with 75% being type 1 mutation. The Variant Allele Frequency (VAF) measures the fraction of sequencing reads with a variant allele. The number ranges between 0-1 with 1 representing clonal homogeneity of the variant. Within these participants the majority

had a VAF of less than 1. Patients with MPN had a higher allele burden of the mutation than patients without MPN. [21] .

In addition, two recent pieces of work have shown that these driver mutations are acquired early in life and MPN disease has a long latency until symptoms present [2, 22]. The first study used whole geneome sequencing from 12 patients to reconstruct haematopoietic phylogenies and evolution histories by benchmarking against the collection of mutations acquired over time within a patient. They found that the average timing of acquisition of a JAK2 mutation was from 3 weeks of gestation to 10.8 years. The mean latency of acquisition to diagnosis was 30 years [2]. A smaller study [23], found the same conclusions [2]. The second study showed that monozygotic twins who both presented with myelofibrosis in their fourth decade had a common clonal origin of their CALR-mutant clone, which therefore must have ocurred in utero by transplacental transmission [22].

The genetic background has an influence on MPN initiation. First degree relatives have a 5-7 fold increase likelihood of MPN and genome wide association studies have also identified a 15 gene signature that accounts for 20% of the familial risk [24]. However, this latency between acquisition of the mutation and presentation of MPN phenotype, suggests that other factors, such as the bone marrow microenvironent, cytokine homeostasis, inflammation, or individual genetic variation, also contribute to the likelihood of developing an MPN.

2.2.3 The Bone Marrow Microenvironment and MPN

The bone marrow microenvironment includes structural tissue such as blood vessels and stromal tissue, alongside infiltrating immune cells. HSCs reside in a bone marrow microenvironment known as a niche, with significant interaction between stromal (structural cells e.g. MSCs, osteoblasts, endothelial cells, adipocytes and peripheral neurons and their associated Schwann cells) and other haematopoietic cells. Bone marrow niches have been associated with immune evasion and chemoresistance, raising the possibility of the niche remodelling as a therapeutic target [25]. Two potentially complementary theories associate the niche with malignancy; one, the niche predisposing HSCs towards disease initiation or two, malignant cells transforming the niche for their own survival

MPN are known to have increased levels of circulating inflammatory cytokines [26]. This results from both increased production of cytokines per cell and an increase in the percentage of cytokine secreting cells [5], [26]. However, the non-clonal haematopoietic and stromal cells also contribute to immune dysregulation.

Development of fibrosis is possible in all MPN subtypes and characterises MF. Fibrosis causes scarring of the bone marrow, disrupting normal blood cell production. Increased fibrosis is associated with poorer prognosis. Fibrosis has been linked to inflammation, although whether inflammation occurs prior to or as a consequence of fibrosis is still an open question. Thus, understanding the immune landscape will increase our ability to manipulate prognostic outcome and emphasizes the importance of cell interactions involved in immune dysregulation and delineating the interactions of these cells in the bone marrow through spatial analysis.

2.3 The Bone Marrow Microenvironment

2.3.1 Importance of the Bone Marrow Microenvironment

As mentioned above, whilst driver mutations are required for MPN initiation, they aren't sufficient alone. Two seminal studies, [2, 22], found latency between the date of acquisition of the driver mutations and development of disease, with a median latency of 30 years. Therefore, it is suggested that other factors cause initiation of clonal expansion of the mutated HSCs leading to the disease phenotype. Additionally, it is still unknown how the same driver mutations give rise to different MPN subtypes. Hypotheses for initiation include changes in the bone marrow microenvironent, cytokine homeostatis, inflammation, epigenetic changes or individual genetic variation [5]. In Chapter 5 we will investigate changes in the bone marrow microenvironment.

2.3.2 Components of the bone marrow microenvironment

The bone marrow is the site of haematopoeisis. The bone marrow micoenvironment is perfused with blood vessels and made up of haematopoetic cells, non-haematopoeic cells (such as stromal and neural cells) and extracellular matrix. When a sample is taken during a biopsy, the sample will also contain bone, from trabeculae which make up a lattice within the larger bone, as well

as red marrow where the haematopoetic cells are situated. Where the marrow meets the bone is the endosteal surface [27, 28]. Below I have detailed the main components of the bone marrow, their spatial arrangement and function.

Stromal Cells

Outside the haematopoietic lineages, mesenchymal stromal cells (MSCs) play a crucial role in bone marrow function. These multipotent cells can differentiate into osteoblasts, adipocytes, or chondrocytes, the latter contributing to the maintenance of the extracellular matrix (ECM). MSCs are a heterogeneous population, and many studies have attempted to classify them into subtypes based on biomarker expression and to associate each subtype with specific functions.

Among these, leptin receptor (LepR)-expressing MSCs are the most abundant in the bone marrow. A key subset of LepR+ MSCs express CXCL12, the ligand for CXCR4 which is expressed on haematopoietic stem cells (HSCs). [28]. LepR+ MSCs that produce high levels of CXCL12 are referred to as CARc cells. These are the most numerous LepR+ MSCs and are distributed throughout the bone marrow [27]. CXCL12 forms a chemoattractant gradient that helps retain HSCs within the bone marrow. Deletion of CXCL12 causes HSCs to relocate to extramedullary sites, such as the spleen and liver [28]. CARc cells are also the primary source of adipocytes and osteoblasts [27]. CARc cells can be further subdivided based on their lineage bias. Adipo-biased CARc cells are located near sinusoids and are thought to support HSC maintenance. In contrast, osteo-biased CARc cells express higher levels of alkaline phosphatase and tend to localize closer to arterioles or away from the vasculature[27, 28]. Another MSC subset is defined by expression of Nestin (Nestin+ MSCs). These cells express low levels of CXCL12 and are typically found near arterioles. Depletion of Nestin+ MSCs has minimal impact on HSC abundance but does lead to a reduction in lymphoid progenitors [28].

Osteo-lineage cells

Of the three differentiated MSCs (osteolineage, adipo-lineage and chondrocytes) only osteoblasts and adipocytes are formed continuously throughout life [27]. Osteo-lineage cells are found around bone, whilst adipocytes increase with age and are interspersed in the marrow [28]. Osteo-lineage cells, found in the endosteal reagion, were the first population to be identified

with regulation of HSCs. Osteo-lineage cells associated with HSCs were shown to have increased Notch signalling with Jag1 in particular associated with HSC quiescence. In addition TPO and ANGPT1 are suggested to regulate quiescence [28].

Extracellular Matrix

The ECM contains fibrous proteins (collagen and elastin) as well as glycoproteins, produced by stromal cells. Collagens I-XI are the most abundant. The composition of the ECM varies across bone marrow regions. The endosteal niche is enriched in collagen I and fibronectin (FN), while laminin becomes more prevalent in areas with increased vascularization. The mechanical properties of the bone marrow microenvironment also influence stromal cell fate. Stiffer substrates promote osteogenesis, whereas softer substrates bias mesenchymal stromal cells (MSCs) toward adipogenesis. Additionally, tissue geometry plays a role: long, concave curves have been shown to favor adipogenic differentiation [27].

Vasculature

The vasculature of the bone marrow consists of arterioles branching from the main artery towards the endosteal regions. Here, arterioles fuse into a network of capillaries called the transitional vessels which give rise to the sinosoids which drain venus blood towards a large sinus. Arterioles and sinusoids can be identified in H&E by expert pathologists. However, these vessels are also lined by endothelial cells with distinct morphological, phenotypic, and molecular characteristics. These endothelial cells influence not only vascular permeability but also the behaviour and interaction of surrounding cells. Even within morphologically similar sinusoidal structures, local microenvironments can vary depending on vascular position. For example, CSF1 expression has been shown to generate specialised sinusoidal niches for monocyte development [27].

Distinct molecular markers define different vascular components. Nestin-positive, NG2-coexpressing cells are associated with both large arteries and arteriole branches. Pericytic cells, which wrap around endothelial cells, are found only around larger arterioles and typically express PDGFR and Podoplanin. These pericytes are multipotent and capable of differentiating into either adipocytic or osteoblastic lineages [27]. Smooth muscle cells are also present around large arteries.

Endothelial cells (ECs) contribute to the maintenance of haematopoietic stem cells (HSCs), although they produce much lower levels of CXCL12 and stem cell factor (SCF) compared to mesenchymal stromal cells (MSCs). ECs can be broadly classified into arteriolar endothelial cells (AECs) and sinusoidal endothelial cells (SECs). AECs exhibit low permeability and reduced reactive oxygen species (ROS) levels, thereby promoting HSC quiescence. In contrast, SECs are more permeable and associated with higher ROS levels, leading to increased HSC activation, migration, and differentiation. Despite this, megakaryocytes located near sinusoids have been shown to promote HSC quiescence through the expression of TGF-beta, CXCL4, and thrombopoietin [28].

Notably, patients with primary myelofibrosis exhibit elevated levels of circulating endothelial cells compared to healthy controls, suggesting vascular damage is a component of disease pathology. Intriguingly, these circulating endothelial cells often carry myeloid-associated mutations, some of which are also present in haematopoietic stem and progenitor cells (HSPCs); a finding not observed in healthy individuals. This suggests that in myelofibrosis, circulating endothelial cells and HSPCs may originate from a common mutated progenitor [29].

Neural

A specialised group of neurons, known as nociceptive neurons, has been shown to facilitate haematopoietic stem cell (HSC) mobilisation through the production of the neuropeptide CGRP. Pleiotrophin, a neurite growth factor, is primarily expressed by LepR-positive stromal cells during steady-state haematopoiesis, where it supports HSC maintenance. Following injury, pleiotrophin expression also increases in sinusoidal vessels, contributing to HSC regeneration [30].

Haematopoetic Lineage Cells

As mentioned earlier, the bone marrow is the primary site where haematopoietic stem cells (HSCs) differentiate into immune and blood cells. Beyond their specialised immune functions, these differentiated cells can also influence the HSC differentiation process. Quiescence is a state of reversible growth arrest in which HSCs remain in the G0 phase of the cell cycle. Megakaryocytes are thought to promote HSC quiescence through the secretion of CXCL4, TGF-beta, and thrombopoietin (TPO). Additionally, TPO produced by osteoblasts can attract

megakaryocytes to the endosteal region, where they further support osteolineage expansion via PDGFRB signalling and help re-establish HSC quiescence [28].

Macrophages also contribute to the HSC niche. In their absence, osteoblast activity is suppressed, and the expression of HSC-supporting cytokines in the endosteum is reduced. This can result in HSPCs entering the peripheral blood. These findings suggest that macrophages play a role in maintaining the endosteal HSPC niche [28]. Regulatory T cells (Tregs) help suppress HSC proliferation and contribute to the maintenance of quiescence. Interleukin-10 (IL-10), produced by lymphocytes and monocytes, has also been shown to support HSC survival following transplantation[28].

2.3.3 The HSPC niche

The concept of a haematopoietic stem cell (HSC) niche was first proposed by Schofield in 1978, describing it as a specialised tissue environment that supports undifferentiated HSC self-renewal and regulates their differentiation [31]. As Williams et al. [2] showed, driver mutation bearing HSCs exist alongside normal blood production, suggesting that different niches for the different HSCs may occur. There are two main suggested niches in the healthy bone marrow. The more hypoxic endosteal osteoblastic niche and the more oxygenated perivascular niche [28]. The idea of an adipocytic niche is also gaining traction.

Endosteal Niche

The endosteum is a thin layer of cells lining the inner surface of the bone. The endosteal niche has been associated with the maintenance of haematopoietic stem cell (HSC) quiescence [28]. Even before Schofield formally introduced the concept of a stem cell niche, haematopoietic stem and progenitor cells (HSPCs) were observed to be preferentially enriched near the bone surface, suggesting a spatial relationship with the endosteum. The enrichment of osteoblasts in this region led to the hypothesis that they may contribute to the formation and maintenance of the HSC niche [31].

There is evidence to suggest that osteoblasts can regulate HSCs. These cells produce several key cytokines, including G-CSF, TPO, ANGPT1, and CXCL12. These factors support long-term maintenance of HSCs in ex vivo cultures and influence the localisation of transplanted HSCs

to the endosteal surface. Furthermore, a correlation has been reported between the number of osteolineage cells and HSC abundance [30–32].

However, it is important to note that cell ablation experiments may not provide definitive insight into the role of specific cell populations in HSC maintenance due to indirect or secondary effects. For example, ablation of CARc cells, LepR-positive cells, Nestin-positive cells, NG2-positive cells, megakaryocytes, and macrophages has all been shown to reduce HSC frequency and quiescence indiscriminately [30]. Evidence also suggests that the interaction between HSCs and the endosteal niche may be indirect. Advanced imaging studies indicate that HSCs actually tend to localise near microvessels, even when they are positioned close to the bone surface. Furthermore, neither genetic nor pharmacological depletion of osteoblasts results in acute HSC loss, and the removal of SCF or CXCL12 from osteolineage cells does not significantly affect HSC abundance or retention in the bone marrow [30–32]. It has also been proposed that the function of the endosteal may vary depending on the stage of bone remodelling. For instance, one study found that proliferating HSCs are almost exclusively located in regions of active bone remodelling following stimulation with G-CSF [33]. This suggests that the concept of an osteoblastic niche is now outdated and alternative concepts of HSC niches are being explored.

Perivascular Niche

Haematopoietic stem cells (HSCs) localise predominantly to the sinusoidal blood vessels within the bone marrow. At this sinusoidal niche, reticular cells expressing high levels of CXCL12, known as CARc cells, have been observed in direct contact with HSCs. Nestin-positive mesenchymal stromal cells (Nestin+ MSCs) are also found in close proximity and express key HSC maintenance factors, including ANGPT1, SCF, and CXCL12. Using Nes-GFP as a reporter for Nestin expression, two distinct subtypes of Nestin+ cells have been identified. A rare population of cells with high Nestin expression localises exclusively around arterioles and is associated with NG2 expression. In contrast, the more abundant intermediate Nestin+ cells form a fine morphological network and are primarily associated with sinusoids. These sinusoidal Nestin+ cells co-express LepR [30].

There is evidence supporting the existence of a sinusoidal niche that contributes to HSC maintenance and quiescence. CARc cells within this niche produce SCF and pleiotrophin, factors essential for HSC function. Pleiotrophin is also produced by SECs. Depletion of pleiotrophin

from CARc cells results in a reduction in HSC numbers. CARc cells also produce IL-7 and CSF1, which are involved in myeloid and erythroid lineage differentiation [27]. Additionally, HSC numbers are reduced when SCF is specifically ablated from arteriolar endothelial cells (AECs)[27].

Adipocytic Niche

Single-nucleus RNA sequencing (snRNA-seq), which is better suited than single-cell RNA sequencing (scRNA-seq) for capturing larger, more fragile cells such as adipocytes and megakaryocytes, has revealed that adipogenic progenitors can differentiate into both LepR-positive stromal cells and osteolineage cells. In acute myeloid leukaemia (AML), these adipogenic and osteogenic differentiation pathways are disrupted, leading to impairment of the adipocytic niche and resulting in perturbed haematopoietic cell production [30].

A more comprehensive atlas of healthy human bone marrow and AML, was recently developed to include both stromal and non-stromal cell populations. Bandyopadhyay et al. [34] found that cells with the highest expression of CXCL12 also expressed adipocyte-related genes, such as APOE and LPL. They identified a population of THY1-positive mesenchymal stromal cells (MSCs) that does not appear to have a clear murine equivalent. In their study, THY1-positive and adipo-stromal cells were most closely associated with haematopoietic stem and progenitor cells (HSPCs). Spatial profiling further revealed that adipocytes were frequently found in proximity to HSCs [30, 34].

Niche supporting factors

Yip et al. [30] has a good review of the HSC niche supporting factors with their summary table included below.

Soluble Factor	Cellular Sources	Impact on HSCs
SCF	Perivascular stromal cells, endothelial cells	Maintains HSC quiescence and bone marrow retention
CXCL12	Perivascular stromal cells, endothelial cells, osteoblasts	Maintains HSC quiescence and retention; interacts with CXCR4 on HSPCs
CXCL4	Megakaryocytes	Regulates HSC quiescence
Thrombopoietin	Hepatocytes	Regulates HSC quiescence
Angiopoietin-1	Perivascular stromal cells, megakaryocytes	Delays HSC regeneration
Pleiotrophin	Perivascular stromal cells, endothelial cells	Promotes HSC maintenance and regeneration
TGFB-1	Megakaryocytes, Schwann cells, macrophages	Regulates HSC quiescence
CGRP	Sympathetic/nociceptive neurons	Promotes HSC mobilisation

Table 2.1: Soluble factors affecting HSCs and their cellular sources. Adapted from Yip et al. [30]

.

CXCL12 is a chemoattractive cytokine that binds to CXCR4 on HSCs and other haematopoietic cells. While deletion of CXCL12 from HSCs or mature osteoblasts does not impact HSC maintenance, its ablation from Osterix-expressing stromal cells, including CARc cells and osteoprogenitors, leads to HSPC mobilisation. Similarly, deletion of CXCL12 from perivascular stromal cells and osteoblasts, but not from endothelial cells, results in a reduction of HSC numbers and loss of quiescence. These perivascular cells are believed to be mesenchymal progenitors that do not express Nestin or LepR [30].

Nestin-positive cells produce high levels of both CXCL12 and SCF. Deletion of CXCL12 from Nestin-high, NG2-positive cells, typically found around arterioles, results in reduced HSC frequency in the bone marrow and displacement of a subset of HSCs away from the arteriolar region. However, this effect appears to be more relevant in early bone marrow development, as short-term tamoxifen treatment (used to activate Cre recombinase for CXCL12 ablation) in adult mice does not produce the same outcome [30]. In contrast, deletion of CXCL12 from Nestin-intermediate, LepR-positive cells typically associated with sinusoidal vessels, leads to HSC mobilisation into the bloodstream without altering HSC abundance or localisation within the bone marrow. For distinguishing these overlapping stromal populations, Osteolectin has been identified as a specific marker of peri-arteriolar LepR-positive cells [30].

Regarding other signalling factors, deletion of thrombopoietin (TPO) from hepatocytes results in depletion of bone marrow HSCs. However, deletion of TPO from osteoblasts, haematopoietic cells, or LepR-positive HSCs does not appear to affect haematopoiesis [30].

2.4 The Bone Marrow Microenvironment in MPN

Within this section, I focus on what we currently know about the bone marrow microenvironment in MPN, with a particular emphasis on fibrosis.

2.4.1 Heterogeneity in Myelofibrosis

Myeloproliferative neoplasms (MPNs) are a heterogeneous group of diseases that include essential thrombocythaemia (ET), polycythaemia vera (PV), and myelofibrosis (MF). Among these, myelofibrosis itself is notably heterogeneous, encompassing a clinical spectrum that ranges from a myelodepletive phenotype, characterised by cytopenias and low blood counts, to a myeloproliferative phenotype.

This heterogeneity across MPNs, and particularly within MF, has been recognised since the 1950s. Patients with cytopenic MF generally experience worse outcomes. For example, a lower JAK2 variant allele frequency (VAF), particularly below 25%, is associated with reduced leukocyte and haemoglobin levels, and correlates with shortened survival [35]. Cytopenia is also linked to more advanced marrow fibrosis, often grade 2 or higher. One explanation for this variability is the presence of additional, non-driver mutations commonly seen in myelodysplastic syndromes (MDS), which reinforces the difficulty in the clinical distinction between MF and other myeloid malignancies [35].

Histologically, trephine H&E sections from MF patients frequently show megakaryocytic hyperplasia, an increase in megakaryocyte number. Interestingly, this feature is paradoxically associated with thrombocytopenia (reduced platelet counts), especially in cytopenic MF, although it is not exclusive to this subgroup. Thrombocytopenia is a well-established marker of poor prognosis. Patients with severe thrombocytopenia often exhibit a shorter and less durable response to ruxolitinib, though this may reflect drug-specific limitations rather than a class effect of JAK inhibition. In addition to higher-grade fibrosis, these patients also face a greater risk of leukaemic transformation [35]. Further molecular studies support the idea

that co-mutations influence disease phenotype. For example, Collins et al. [36] reported that JAK2 mutations promote increased erythroid colony formation, while mutations in ASXL1—a non-driver mutation—are associated with erythroid colony defects.

2.4.2 MPN and the HSPC niche

Clonal haematopoiesis refers to the expansion of a haematopoietic stem cell (HSC) carrying a somatic mutation, resulting in a population of cells that share the same genetic alteration. In this study, we investigate the bone marrow microenvironment with the hypothesis that changes within this environment may permit the clonal expansion of mutated HSCs, an expansion that may not have occurred under homeostatic conditions.

Two previous studies have explored the spatial distribution of CD34+ cells in MPNs. Grockowiak et al. [37] used intravital imaging and immunohistochemically (IHC) stained bone marrow trephines to compare HSC localisation in essential thrombocythaemia (ET) and polycythaemia vera (PV). They found that HSCs in ET were situated closer to the bone surface than those in PV. Specifically, ET HSCs were frequently located near arterioles, which were expanded and positioned closer to the bone, whereas PV HSCs were more often found adjacent to sinusoids. The authors also examined the effects of ruxolitinib on the HSC niche. In ET, ruxolitinib treatment led to the redistribution of multipotent progenitors (MPPs) toward the central marrow, accompanied by an increase in sinusoidal vessels. In PV, the treatment restored the endosteal niche, increased HSC quiescence, and expanded arteriolar networks. Additionally, the study reported a loss of polarity in PV HSCs, attributed to dysregulated CDC42b. Inhibition of CDC42b restored HSC polarity, relocated PV HSCs toward the endosteal region, increased their quiescence, and reduced peripheral blood counts. They also noted that MPN progression is accelerated in prematurely aged mice following transplantation of cells carrying MPN-associated driver mutations.

Young et al. [38] investigated the spatial distribution of somatically mutated cells in the femoral head of a 50-year-old male with PV. They found that haematopoietic clones were not evenly mixed throughout the bone marrow. Instead, the marrow was composed of spatially distinct regions, some with high clonal diversity (where both normal and neoplastic clones coexisted), and others with low diversity, dominated by a single clone. These clonally distinct areas were

often located just millimetres apart, highlighting the patchy and regionally compartmentalised nature of clonal expansion in PV.

2.4.3 The Inflammatory Cytokine landscape of MF

In myelofibrosis (MF), one common theory is that abnormal signalling from mutated haematopoietic cells drives fibroblast proliferation, resulting in reticulin deposition and disruption of the normal bone marrow microenvironment. An alternative hypothesis posits that inflammation arises as a consequence of fibrosis. Regardless of sequence, MF is characterised by a highly inflammatory microenvironment, with elevated levels of TNF-alpha, IL-6, and IL-1 [20].

These inflammatory cytokines have been shown to be produced by monocytes. TGF-B has been implicated as a critical driver of bone marrow fibrosis. Elevated levels of CXCL8/IL-8, IL-2R, IL-12, and IL-15 in plasma have been associated with poorer prognosis. TNF-alpha, in particular, is believed to promote clonal dominance in MPNs. Hyperactivation of the JAK-STAT, MAPK/PI3K, and NF-kB pathways has been observed in MF, with NF-kB activation reported in haematopoietic stem cells (HSCs), B cells, and T cells [39]. NF-kB is a central transcriptional regulator of many cytokines implicated in MF, including IL-6, IL-1beta, IL-8, TNF-alpha, and TGF-beta. It can be activated via Toll-like receptors (TLRs) or through constitutive activation of the JAK/STAT pathway. JAK/STAT signalling promotes degradation of TP53, relieving inhibition of NF-kB. In addition, epigenetic changes downstream of JAK2 signalling can further enhance NF-kB activation [35]. Monocytes in MPN patients exhibit hyper-responsiveness to TLR ligands. Engagement of TLRs and IL-1 receptors (IL1Rs), both of which are overexpressed in MPNs, leads to the release of inflammatory cytokines such as TNF-alpha, IL-6, IL-1, CCL4, and type I/II interferons. IL-10 acts as an anti-inflammatory cytokine in this context [35].

The temporal relationship between inflammation and fibrosis remains unresolved, whether inflammation precedes fibrosis or arises in response to it. However, excess TNF-alpha production supports the clonal expansion of JAK2-mutated HSCs while suppressing normal progenitors. IL-1R/myddosome signalling is also indispensable for clonal expansion, megakaryocyte proliferation, and MF progression in murine models. Furthermore, JAK2-mutant HSPCs produce serum IL-1beta, and its expression, along with IL-1 receptor levels and cytokine output, correlates with JAK2 variant allele frequency (VAF). Neutralisation of IL-1beta reduces myelofibrosis and osteosclerosis [35]. Interferon gamma has been shown to disrupt the endothelial niche through

release of HMGB1. HMGB1, in turn, worsens anaemia by competing with erythropoietin and reducing the number of erythroid precursors. Gli1 is upregulated in MF fibrocytes via sustained STAT3 activation, leading to activation of pro-fibrotic pathways [35].

The NLRP3 inflammasome, which is abundant in myeloid cells, is also implicated in MF and can be activated by damage-associated molecular patterns (DAMPs). AIM2, another inflammasome, has similarly been implicated. S100A8/A9 (calprotectin) is upregulated in both mouse models of MF and human MSCs from MF patients, with levels increasing alongside disease progression, suggesting its potential as a biomarker. Small-molecule inhibition of calprotectin has been shown to ameliorate fibrosis. Calprotectin also exacerbates cytopenia and is particularly elevated in patients with lower JAK2 VAF. S100A9 alone is sufficient to cause ineffective haematopoiesis, which can be rescued by NLRP3 inhibition. This process is linked to the induction of PD-1 on HSCs and PD-L1 on myeloid-derived suppressor cells [35]. IL-33 and its receptor IL1RL1 are upregulated in HSPCs in MPNs. In mice, IL-33 promotes HSPC colony formation and induces the secretion of DAMPs, including the S100A8/A9 heterodimer [35]. Genetic deletion of IL1R1 in murine models normalises peripheral blood counts, reduces splenomegaly, and improves bone marrow fibrosis. IL-1beta induction enhances myeloid expansion and accelerates fibrosis in JAK2-mutant mice [40]. In addition, Rai et al. [41] demonstrated that IL-1beta levels in peripheral blood correlate with JAK2 VAF and that the primary source of IL-1beta overexpression is JAK2-mutant HSCs. Deletion of Il1beta specifically in HSCs reduces the expression of inflammatory cytokines, protects Nestin-positive niche cells, and suppresses megakaryopoiesis.

Finally, disruption of vascular architecture has also been implicated in MF, with deregulation of angiogenic genes such as angiopoietin-1 and -2 [20].

2.4.4 The Spleen in MPN

It is important to note that organs beyond the bone marrow are also affected in myeloproliferative neoplasms (MPNs). In myelofibrosis (MF), excessive extramedullary haematopoiesis frequently occurs in the spleen as a result of a disrupted bone marrow microenvironment. Pathological changes observed in the spleen often parallel those found in the bone marrow in MF [20]. The spleen can also serve as a reservoir for malignant haematopoietic stem cells (HSCs). Using spatial transcriptomics Peroni et al. [20] identified regional gene expression gradients in the spleen, particularly related to inflammation and iron metabolism.

2.5 Fibrosis in MPN

2.5.1 Fibrosis across MPN

Fibrosis in the bone marrow is characterised by the excessive accumulation of reticulin fibres, and in some cases, collagen. This fibrotic scarring disrupts the marrow architecture, impairs normal blood cell production, and contributes to both symptom burden and poorer prognostic outcomes, including reduced overall survival [42]. In myelofibrosis (MF), fibrosis may present in several clinical forms: prefibrotic primary MF (pre-PMF), overt primary MF, or secondary MF following transformation from essential thrombocythaemia (ET) or polycythaemia vera (PV). Median survival for these subtypes is approximately 14 years, 6–7 years, and 9 years, respectively [43].

Fibrosis in MF is not uniformly distributed across the bone marrow, and disease severity is highly variable. Life expectancy can range from less than two years to more than 20 years in some patients, reflecting the heterogeneity of disease progression and response to treatment [44]. Mascarenhas et al. [35] provide an in-depth discussion of this heterogeneity, particularly in MF, as we outlined above. Thrombocytopenia in MF is linked to worse clinical outcomes, including higher circulating blast counts, advanced fibrosis (MF-3), and an increased risk of leukaemic transformation. Moreover, approximately 20% of ET patients and a variable proportion of those with PV also exhibit increased reticulin deposition in the bone marrow. In both ET and PV, the presence of bone marrow fibrosis is associated with a higher risk of progression to overt MF and acute myeloid leukaemia [42].

2.5.2 Fibrosis Categorisation

Bone marrow fibrosis is classified into four grades based on the European consensus grading system [45]. This system categorises fibrosis as a function of increasing reticulin, fiber intersections, bundling of collagen and/or osterosclerosis. The overall fibrosis grade assigned to a patient corresponds to the highest grade observed in at least 30% of the evaluable marrow area[46].

The IPSS model for prediction of survival was introduced in 2009. This has since been updated with mutation enhanced IPSS (MIPSS-70 or 70+) and the dynamic IPSS (DIPSS) and MYSEC-

MF-0	Scattered linear reticulin with no intersection (cross-overs) corresponding to normal bone marrow
MF-1	Loose network of reticulin with many intersections, especially in perivascular areas
MF-2	Diffuse and dense increase in reticulin with extensive intersections, occasionally with only focal bundles of collagen and/or focal osteosclerosis
MF-3	Diffuse and dense increase in reticulin with extensive intersections with coarse bundles of collagen, often associated with significant osteosclerosis

Table 2.2: Fibrosis Grades in MPN. Table taken from: Ghosh et al. [45]

PM. A good review of the characteristics which predict prognosis in each model is found Passamonti and Mora [43]. Fibrosis at grade 2 or above is identified as a significant risk factor as part of MIPSS70 [46].

2.5.3 Mechanisms of Fibrosis Development

The mechanisms driving bone marrow fibrosis and its relationship with inflammation remain incompletely understood. Nonetheless, numerous studies have significantly advanced our understanding.

Chronic inflammation is a hallmark of myeloproliferative neoplasms (MPNs), particularly myelofibrosis (MF). However, it is still unclear whether inflammation initiates myeloproliferation or arises as a consequence of the disease [35]. Key cytokines implicated in the disruption of bone marrow homeostasis include TGF-beta, fibroblast growth factors (FGFs), and proinflammatory cytokines such as TNF-alpha [20].

Several mutations are associated with more severe fibrosis and poor outcomes. Mutations in ASXL1 and EZH2 accelerate disease progression, leading to higher fibrosis grades and lower survival. ASXL2 mutations reduce responsiveness to JAK2 inhibitors and increase the risk of leukemic transformation. SRSF2 and IDH1/2 impair cell differentiation and promote myeloproliferation, raising the risk of AML. EZH2 mutations contribute to fibrosis via disrupted epigenetic regulation [20].

Gleitz, Benabid, and Schneider [47] proposed a model in which chronic inflammation, originating from mutated HSCs, alters the bone marrow niche and establishes a disease-permissive microenvironment. Stromal cells are active participants in sustaining this inflammation. Age-related inflammation, or “inflammaging,” is an example of low grade inflammation; stromal IL-6 production has been linked to age-associated erythropoietic decline, and stromal cells are known to become increasingly proinflammatory with age.

In mouse models, Gli1+ and LepR+ mesenchymal stromal cells (MSCs) were identified as progenitors of myofibroblasts. Targeted inhibition of these populations reduced fibrosis [48].

CXCL4 (PF4) emerged as a marker of fibrotic progression; its knockout reduced JAK-STAT signalling and inflammation, reflected by decreased IL-6 in megakaryocytes. These Gli1+LepR+ cells also contribute to osteogenesis [49]. Interestingly, CXCL4 expression shifts during disease, from haematopoietic cells in early stages to stromal cells in advanced fibrosis [47]. The same group also noted co-occurrence of S100A8/9 (calprotectin) in fibrotic regions [47]. Capitanio et al. [50] confirmed that CXCL4, produced by megakaryocytes, characterises overt fibrosis and is taken up by stromal cells via glycosaminoglycans (GAGs), promoting myofibroblast transdifferentiation and ECM production. Ruxolitinib treatment was shown to suppress CXCL4 in vivo.

Li et al. [51] described a proinflammatory niche composed of MSCs, eosinophils, basophils, and mast cells that drives MF progression. Basophils and mast cells serve as hubs of TNF signalling. Galectin-1 inhibition reduced both fibrosis and cell proliferation. These authors also identified a population of fibroblasts enriched for TNF, NF-kB, IL-6, and IFN-gamma signalling. In parallel work, Li, Stanojevic, and Garmire [52] reported expansion of proinflammatory and angiogenic megakaryocytes in MF. Together the authors identified basophils and mast cells as a hub of TNF and IL4 signalling with fibroblasts, iFibs, MSCs, neutrophils and monocytes/macrophages as the key interacting partners.

TNF-alpha and JAK-STAT signalling are upregulated in pre-fibrotic MSCs, while TNF-beta becomes predominant in overt fibrosis[53]. However, inflammation alone is not unique to MF, many other myeloid malignancies are inflammatory but do not develop fibrosis. It is hypothesised that the specific cytokine profile in MPNs is critical, particularly elevated IL-8, IL-12, and IL-15. Another theory points to dysregulated reactive oxygen species (ROS) activating TGF-beta, which is both pro-fibrotic and a known regulator of HSC quiescence. TGF-beta is produced by dysplastic megakaryocytes, but JAK2-mutant HSCs may be resistant to its suppressive effects, gaining a clonal advantage [47].

In MF, early multipotent stem cells show a bias toward megakaryocyte differentiation, contributing to aberrant megakaryopoiesis [54]. These megakaryoyte progenitors also expressed inflammatory mediators and earlier progenitors have been shown to be enriched for inflammatory pathways involving TNF alpha and interferons [47].

Of the cellular components of the bone marrow, megakaryocytes are the main cells implicated in fibrosis. They secrete proinflammatory cytokines and activate Gli1+ stromal cells. As mentioned above, CXCL4, a megakaryocyte-specific cytokine, is directly implicated in stromal remodelling [47]. Monocytes have also be found to be pro-inflammatory on disease progresson in PV and a source of myelofibrosis associated cytokines. Later work has suggested that a role for macrophage interaction with osetoCAR stromal cells via upregulation of Spp1 [55].

Recent fate-mapping studies in mice suggest that chronic inflammatory stimuli in the central marrow niche leads to reprogramming of perivascular CXCL12-abundant reticular (CAR) cells. These cells exit the perivascular space for the central marrow, adopt proinflammatory phenotypes, and expose the endothelium to inflammatory signals, promoting endothelial-to-mesenchymal transition (EndMT) [56]. Focusing on the metaphyseal region, the same study identified the peritrabecular space as a reservoir for fibrosis-driving cells. An osteolineage subset was identified as mesenchymal progenitors enriched in Gli1 and marked by NCAM1/CD56 which expand in fibrotic conditions. Gli inhibition reverses fibrotic transformation. Expansion of these NCAM1/CD56 Osteolineage cells correlated with increased fibrosis and the authors therefore suggest the peritrabecular as a hotspot of fibrosis formation. The authors show that peritrabecular skeletal cell's activation and increase in frequency with bias towards osteoblast and fibroblast differentiation is Wnt dependant with upregulation of beta catenin. This fits with the knowledge that Osterosclerosis also occurs in late stage fibrotic MPN, although the pathogenesis is still unknown [56].

Additional work has also focussed on the interplay between the endothelium and mylofibrosis and osteosclerosis. Sivaraj et al. [49] demonstrated that Lats2 inactivation or YAP1 overexpression in endothelial cells promotes fibrosis and osteosclerosis via EndMT. SRF, a transcription factor, appears necessary for YAP1/TAZ-mediated EndMT, as SRF inactivation rescues the phenotype. Endothelial cells in the metaphysis normally support perivascular osteoprogenitors and osteoblasts by secreting growth factors. This behaviour is regulated by external growth factors, but also mechanical cues such as matrix stiffness and shear stress. EndMT alters not only endothelial gene expression but also cell morphology and vascular architecture [49]. The authors suggest that the link between endothelial cell changes and hematopoetic cells may be an increase in TGF beta, which has been shown to be up-regulated in megakaryocytes and monocyte/macrophages in myelofibrosis, but this is so far unknown [49]. TGF beta has however

been linked to the loss of microvessels, endothelial to mesenchymal trasition and can induce MRTF-SRF signalling.

2.5.4 Automated Fibrosis Detection

Recently techniques to automatically grade fibrosis in reticulin stained slides have been developed. Our group has developed an automated method of ranking fibrosis severity in a way that can be overlaid onto a stained slide. Rather than relying on traditional categorical grading (MF-0 to MF-3), the method uses a Learning to Rank framework that compares pairs of image tiles and assigns them scores based on relative fibrosis severity. This ranking is performed using a Ranking Convolutional Neural Network (Ranking-CNN) that combines a convolutional feature extractor with a RankNet loss function, enabling the model to learn from expert-labeled image pairs.

Expert hematopathologists then reviewed pairs of these tiles and identified which tile exhibited more severe fibrosis based on WHO criteria. These judgments were then used to train the Ranking-CNN in a human-in-the-loop setup, where iterative feedback refined the model. The output of the model is a CIF score ranging from 0 (least fibrotic) to 1 (most fibrotic), which can be overlaid on the original tissue image to produce a heatmap of fibrosis severity[46]. Previously this algorithm used reticulin stained slides which, whilst from the same biopsy block, are on a different tissue slice to H&E stain. This means that our automated fibrosis detection is not on the same cells which hold the highest morphological information.

Our group has developed an extension, by training on slides which have been stained and destained so that both reticulin and H&E images can be taken of the slides. To address stain variability histogram equalisation was performed prior to preprocessing. Using affine transformations, corresponding H&E tiles were spatially aligned to their reticulin counterparts. The method involved training a ranking convolutional neural network (ranking-CNN) using matched H&E and reticulin image tiles. The model was trained on 9,252 tile pairs extracted from 30 bone marrow samples, where each reticulin tile had an associated Continuous Indexing of Fibrosis (CIF) score generated by a validated reticulin-based model [57].

2.6 Spatial Transcriptomics

2.6.1 The Limitations of Clinically Available H&E and IHC

Clinically diagnostic slides provide an image of cells in the spatial landscape. H&E provides details of cells morphology through nuclear and cytoplasmc locations, whilst IHC provide the binary location (positive or negative) of an individual protein in the context of the nucleus. Each stain (H&E or IHC) is on a separate slice of tissue from the biopsy block and so it is difficult to relate the two at a single cell level.

Detection antibodies are added specific to antigens of interest. The antibodies for IHC are bound to a chromogenic or fluorogenic conjugate, which can be detected either directly or on addition of an enzyme. The concentration of most antigens means that signal amplification is required. This is achieved by using a secondary multivalent antibody bound to a reporter as the conjugate. One of the most common reporters is horseradish peroxidase, which oxidizes a molecule to produce a brown precipitate. However, the non-linear nature can result with poor correlation with the original antigen of interest ([10]). Control stains and antibody titration should be considered when judging staining intensity ([8]).

Using machine learning (ML) algorithms, we can associate visual patterns in histological image data with diagnostically relevant features. However, a common criticism of ML approaches in pathology is their limited interpretability. While spatial patterns in MPNs have been effectively leveraged for fibrosis scoring and disease classification, the underlying molecular mechanisms contributing to these changes cannot be directly inferred from histology alone.

To investigate these mechanisms, researchers typically rely on in vitro experiments, especially high-resolution methods like single-cell sequencing. These approaches offer insight into cell type–specific gene or protein expression and enable hypotheses about interactions (e.g. ligand–receptor signalling or pathway activity) and changes in cell-type abundance. However, these are global estimates across the whole sample and cannot be directly mapped back to the detailed cellular architecture seen in a trephine biopsy. Moreover, suspension-based single-cell analyses tend to underrepresent stromal cell populations due to dissociation biases during sample preparation.

Some cell types can be inferred from H&E-stained sections based on cytomorphological features. For example, megakaryocytes are easily identifiable by their large size and distinct morphology, allowing their spatial distribution to be mapped at single-cell resolution. Our group has developed an automated pipeline to identify megakaryocytes from H&E slides and has correlated their distribution and morphological phenotypes with mutational status [17]. Other morphologically distinguishable cells include eosinophils, which can be recognised by their polylobated nuclei, and clusters of erythroid precursors, which have small, dark nuclei.

Although not yet used clinically, various groups have attempted to classify cell types directly within bone marrow trephines using computational methods. These studies, however, often suffer from a lack of ground truth for validation. There is also an implicit assumption that cells consistently match canonical morphological definitions which, as we demonstrate in our megakaryocyte analysis, is not always the case. Additionally, the effects of disease on tissue architecture must be considered. In myelofibrosis, for instance, hypercellularity is a hallmark feature. The resulting high cell density can distort cell morphology, further complicating interpretation.

2.6.2 The Advantage of Spatial Gene Expression

While H&E staining enables the identification of cell types based on morphology, it provides no direct insight into gene expression. Unless a gene alters a cell's shape or spatial behaviour, its activity cannot be inferred from an image alone. Immunohistochemistry (IHC) offers spatial context for protein expression, but it is typically limited to detecting a single marker per sample. Moreover, IHC does not yield quantitative protein expression, and cell type identification is more challenging due to the absence of eosin staining.

Single-cell suspension studies can identify differentially expressed genes across cell types, particularly when changes are widespread. However, rare cell populations may be lost during sample preparation. If these rare cells cluster spatially or interact in a targeted way within the tissue, their importance might be visually detectable, even if they are underrepresented in suspension data. Spatial structure may therefore serve as a proxy for biological relevance. Furthermore, single-cell studies often rely on indirect inference—linking gene expression changes to functional effects based on known pathways or results from other systems. In contrast,

combining spatial location with gene expression removes the need for inference and allows for direct observation of spatially resolved transcriptional activity.

Spatial transcriptomics thus offers two major advantages: it provides cell-level spatial resolution in situ, and it serves as a ground truth for interpreting and extracting meaningful insights from archival histological images.

2.6.3 The Spatial Transcriptomic Landscape

Spatial transcriptomics builds upon earlier techniques such as laser capture microdissection and in situ hybridization, but now encompasses advanced sequencing and imaging-based strategies with far superior resolution and throughput. In addition, H&E is preserved and so results can be linked back to imaging data in the clinical workflow.

Current ST platforms fall broadly into two categories: sequencing-based and imaging-based. Sequencing approaches, like Visium by 10x Genomics, utilize spatial barcodes on slides to capture RNA from predefined spots or pixels. These techniques can profile the entire transcriptome and detect isoforms, but their spatial resolution is limited, often encompassing multiple cells per spot. In contrast, imaging-based methods, such as MERFISH and CosMx, use fluorescence-based in situ hybridization to detect RNA transcripts at subcellular resolution. While imaging approaches offer excellent spatial fidelity and detection efficiency, they are limited to predefined gene panels and cannot identify novel transcripts. Despite rapid advances, no single platform meets the ideal trifecta of whole-transcriptome coverage, single-cell or subcellular resolution, and high detection efficiency. Sequencing platforms excel in transcriptomic breadth but lack spatial granularity, whereas imaging techniques offer fine spatial detail at the cost of transcriptomic depth. Ongoing developments, such as Visium HD and upcoming whole-transcriptome imaging panels, aim to bridge this gap [58].

2.7 Spatial Analysis Studies of the Bone Marrow

Two recent reviews have summarised previous applications of spatial omics applied to the bone marrow microenvironment and spatial immuno-onclogy respectively [30, 59]. Here, I summarise spatial analyses conducted in the bone marrow, focusing on the key methods used and, where applicable to MPNs, their findings. I also highlight how the field has evolved over time.

2.7.1 Spatial Analysis of Non-MPN Bone Marrow: A Timeline

2018: Gomariz et al. [60] used 3D microscopy of the bone marrow of young and aged mice to show sinusoidal endothelial cells and meschenchymal reticular cells are more abundant than estimated in flow analyses and that all tissue resident cells are situated close to these two cell types.

In terms of spatial statistics they meaured the distance of every voxel of BM to a CARc or sinusoid to create a 3D distance map and then calculated the cumulative distribution function of this empty space transform. This would give results such as 60% of the tissue is within 10um of a sinusoid. To mitigate the effects of image borders they assessed whether the same results were found in different artficial boundaries. To assess whether CARc were located close to sinusoids they formulated the null hypothesis that the position of CARc was not affected by distance to sinusoids. A two sample Kolmorogorov-Smirnov test was used to analyse the significance. However, it's worth noting that the size of the structures being analysed may strongly influence the results.

2019: Baccin et al. [61] used laser capture micro-dissection of approximately 200-300 cells from 76 microdissected regions from mice bone marrow. The characterised each of these regions into arteriole, sinosoid or endosteum, based on the prescence of these features. Their findings suggested that Adipo-CAR and Osteo-CAR localise to sinosoid or arteriole, by looking at the abundance of these cells in each region category.

2023: Xiao et al. [62] performed Visium on the bone marrow of mice. They used CellChat to understand interactions between receptors and ligands in their data. They then performed SpatialTime Analysis, on the distance from blood vessels, trabecular and cortical surfaces using Monocle. These distances were normalised between 0 and 1. They then used the Monocle DifferentialGeneTest function replacing pseudotime with self calculated spatial time values. They then calculated the module scores across spatial time values, smoothing the curve by averging neighbouring spots.

In the same year, Coulis et al. [63], performed Visium on muscle to find that galectin-3 in macrophages is chronically activated in dystrophic muscle. They performed diffential gene expression on regions of interest and CellChat for receptor ligand interactions.

2024: Sudupe et al. [64] performed visium on mouse and human myeloma bone marrow samples. They performed differential analysis between zones using Seurat's FindAllMarkers function.

Stanley et al. [65] have performed Xenium on human MDS bone marrow samples. Preliminary results suggest that they have performed Nearest neighbour analysis and have investigated the size of immune aggregates.

In 2024, Bandyopadhyay et al. [34] published the results of their 53 plex CODEX data on human bone marrow samples.

In terms of spatial analysis, Bandyopadhyay et al. [34] divided the bone marrow into neighbourhoods using the approach from Schürch et al. [66], one of the first domain tissue finders. This method, took the 10 nearest cells of each cell, calculated the composition of the n cell types. Each n length vector was then clustered with minibatchkmeans where k=j. They then looked at differential enrichment of cell types in these neighbourhoods using hypergeometric tests to annotate the neighbourhoods by the cellular composition. The authors also performed cellchat and sought to validate receptor-ligand interactions in their data, based on whether the cell types co-localise. To discover the distance to nearest cell type of interest, distances of target cells were split into quartiles based on their receptor expression, and a one sided wilcox was performed on the distances to determine which expression binned group was closer.

For distance to structures, Bandyopadhyay et al. [34] structures were represented with 10% of their contour points. Distances to each structure was calculated for each cell and a poisson point process model based on spatial coordinates and distance was used. The proximity metric of interest, beta1, captures how strongly cell presence decays with distance to the structure where lower values indicate closer association. To ensure comparability across marrow cavities, beta1 values were within each sample and used normalized ranks for analysis as the marrow cavity size may differ. Within each sample beta1 was ranked for all cell types from 0-1. Significance was assessed via permutation testing (cell labels were permutated 100 times and the median distance for those 100 permutations was used). The proportion of distances smaller than that of the cell type is the p value and pvalues were aggregated using Stouffer's method. For pairwise cell-cell interactions (< 51um), a significance was calculated using permutation of cell labels 500 times and Stouffer's method to combine the multiple p values into one for significance.

Briefly the results from this analysis showed that there was a hyper oxygenated arterio-endosteal niche for early myelopoesis, and that HSPCs were localised to adipocytes. In AML samples MSCs were expanded and there were spatial neighbourhoods enriched for leukaemic blasts and MSCs. CXCL12 was highest in adipo and Thy1 MSCs, which were a new population not defined in mice. These cells produced IL7, whilst ECs and VSMCs contributed to the majority of Notch signalling. The authors also found different GATA+/- megakaryocyte clusters.

In another paper from 2024, Dasdemir et al. [67] used GeoMx and Visium to interrogate AML. The authors performed deconvolution of visium spots to obtain cell composition. The pairwise correlation between abundances was calculated using the Pearson correlation matrix. Again, CellChat was used for receptor-ligand interactions and SpatialTime with values given as distance to bone. Results were visualised with FeatureStatPlot (normalised distance to bone).

John et al. [68] are using RNA tomography and Visium to analyse extramedullary sites in multiple myeloma, although have only presented a preliminary approach so far.

2025: Gui et al. [69] investigated AML using CosMx. They used a Poisson linear mixed effect model for each cell type. Distance was defined in bins of increasing 5 microns with each leukaemic cell at the centre. The input was the count of each cell type at different distance bins. Cells were compared between responders and non-responders to treatment, with adjustment for the total number of cells. As the authors had sampled patients at different time points, they also wanted to know whether there was a cell type density shift. They calculated the probability density function of the distance between two cell types at point A and B and compared significance by undertaking permutation testing 100 times. For ligand-receptor analysis the authors downloaded data from CellTalkDB and found ligands in the leukaemic cell. Cells were bined into close and far distances from the leukaemic cell and expression of the cognate receptor compared.

Mathavan et al. [70] invetigated the mechanical microenvironment using Visium in mice samples. The authors compared DEG between defined regions.

Cenzano et al. [71] also looked into the aged bone marrow using visium which was used mostly for visualisation within their paper.

Lin et al. [72] used visium to investigate human bone marrow samples. They used PROGENy for pathway signalling activity. They subset their graph of cells into ligand and receptor nodes then use Earth Movers Distance based on degree of distribution of the network to quantify

spatial connectivity. The authors used Louvain clustering for niche finding and performed differential gene expression using find markers (with clusterprofiler and enrichGO for pathway enrichment).

2.7.2 Spatial Analysis of MPN Bone Marrow

Three studies have also explored the spatial landscape of MPNs. We previously mentioned the findings of Young et al. [38] and Grockowiak et al. [37]. In addition, Bräunig et al. [73] have investigated the 3D landscape of MPN bone marrow.

Grockowiak et al. [37] used intravital imaging experiemnts and IHC stained bone marrow trephines, to investigate the differences between ET and PV HSCs. They found ET HSCS were situated closer to the bone than PV HSCs and that HSCs in ET were near arterioles (which were expanded and located closer to bone) whilst in PV HSCs were closer to sinusoids. They also investigated the effect of ruxolitinib on the HSC niche and found that treatment did result in changes to the bone marrow. In ET ruxolitinib returned MPPs to the central bone marrow, with concurrent increase in sinusoids. Whilst in PV ruxolitinib restored the endosteal niche, increasing HSC quiescence and expanding arterioles. This was also observed in CALR ET and WT mice. The authors also commented on the loss of polarity in PV HSCs as a consequence of CDC42b. Inhibition of CDC42b restored polarity, and relocated HSCs in PV back to the endosteal increasing quiescence and reducing blood counts. MPN development also accelerates in the prematurely aged mice, on transplantation of cells with MPN driver mutations [74].

Whilst, Young et al. [38] investigated the spatial distribution of somatically mutated cells in the femur head of a 50 year old male with PV. They found haematopoetic clones were not well mixed and so were not evenly distributed in the bone marrow. Instead, they found distinct regions in the bone marrow, either consisting of high clonal variability e.g. a region where both normal and neoplastic clones existed, or regions of low clonal diversity. These distinct regions were often within milimeters of each other.

In addition, Bräunig et al. [73] have investigated the 3D landscape of MPN bone marrow using multi-color immunoflorescence with sequential staining. They then measured distances, after normalising so that distance between each vertex was 0.5 um, and cellular volume. They analysed the full range of MPN, with aged matched normal as controls. In PV they found no differences

in cellularity or normalised megakaryocyte volumes vs controls. However, megakaryocytes in PV were closer to vasculature. In ET, they found decreased cellularity in vascular region but slightly increased megakaryocyte volme. The distribution of megakaryocytes in ET is more similar to controls but the megakaryocytes are more hyperlobated. In PreMF, there was comparable cellularity in vascular regions but an increased perivascular megakaryocyte volume. Megakaryocytes had decreased volume in the endosteal and there was increased cellularity in the endosteal. Finally in MF, there was increased cellularity especially in the endosteal. There was also increased megakaryocyte volume in endosteal and vascular region and megakaryocytes often had hyperlobated nuclei [29]

2.8 Characteristics of Spatial Data

There are three main considerations when dealing with spatial data as opposed to non-spatial data detailed well in [75]. First, is the modifiable aerial unit problem where zoning and aggregation when deciding regions can lead to different results. Second, spatial autocorrelation where spatially close observations are similar and so usual assumptions of independence in classical statistical tests are not obeyed. Finally, there will be spatial heterogeneity, for example a global relationship may not be the same as a local relationship or local relationships may vary depending on the location in the tissue.

2.8.1 Modifiable Areal Unit Problem (MAUP)

How we divide the tissue matters. To analyse a piece of tissue we commonly divide it into regions in either a supervised or unsupervised manner. The MAUP is affected by both scale, i.e. the size of the regions, and two zoning, i.e. even if we choose regions of the same size, how we choose the boundaries matters. In addition, the spatial support (the area over which an observation is made) and spatial sampling will affect the MAUP. To identify an appropriate division and aggregation of the tissue, it is recommended to identify spatial scales at which the observations are stable.

It is important to be aware that spatial aggregation, for example a grid or biologically defined region, can be unpredicatable. One, because unlike non-spatial aggregation which tends to be summaries of sets, spatial aggregation can be non-nested. For example, the boundaries of one

region may overlap with the boundaries of another distinct unit; ecotones may exist, which are transition zones at the borders of two biological communities [75]. In addition, smaller units of patterns may not always be predictably aggregated into larger units.

Biological knowledge can help us in this regard in defining the spatial scales which are sensible to consider, for example, defined tissue regions or biologically relevant actionable distance of cytokines. For immune interactions of cell-cell receptors we may only want to consider immediate neighbours.

2.8.2 Spatial Autocorrelation

This reflects Tobler's first law of geography "everything is related to everything else but near things are more related than distant things". Tests for spatial autocorrelation generate a measure of similarity by comparing each observation to those in their neighbourhood, where a neighbourhood is defined as a distance or adjacency. This can either be local or global i.e. across the whole tissue. Examples of global measures of spatial autocorrelation include Morans I, Geary's C and Getis-Ord G. Each of these global indicators also has a local equivalent. These local indicators of spatial autocorrelation (LISA) provide three kinds of information. First, the degree of spatial clustering around that observation, second, an indicator of local significance, and finally how much each observation contributes to the global spatial autocorrelation statistic. For hypothesis testing we rely on a pseudo p value, as the classical assumptions do not hold. This is either generated with a Z score calculation or a Monte-Carlo method of permutation simulation.

These methods are univariate and analyse one variable at a time. LIMSA (local indicator of multivariate spatial association) are used to discover co-regulated variables, by measuring the extent to which similar neighbours in a multivariate feature space are also neighbours spatially. However, this metric is affected by the curse of dimensionality [75].

2.8.3 Spatial Heterogeneity

Spatial Heterogeneity involves taking into account that relationships between variables may not be consistent across space. Geographically weighted regression (GWR) is an extension of ordinary linear regression but which allows model parameters to vary across spatial locations to

capture local variations. It undertakes a series of local regressions using a moving window or a kernel. Defining a bandwidth, the kernel size and so distance decay, is a key consideration. There are two options, standard GWR which assumes all relationships between the independent and dependant variables operate on the same scale, therefore giving a best on average scale. Multiscale GWR, instead defines a bandwidth for each covariate, allowing the scale to vary for each relationship [75].

2.8.4 Additional Considerations for Biological Tissue

In addition, to these well-defined features of spatial data from computational geography, Summers, Wills, and Rees [76] has highlighted where these assumptions may often not be met in the context of biological samples when applying traditional spatial statistics.

For example, they highlight the fact that many spatial measures assume spatial homogeneity. For example, Ripleys K compares the number of cells in a radius to a spatial Poisson process. However, as mentioned above we may assume heterogeneity as the norm. In addition certain areas may not be available for spatial analysis such as bone splitting the field. Therefore, we have to be careful that any spatial findings are not just a reflection of overall tissue shape and division. Generally to mitigate this problem, randomisation is used to set the null hypothesis using Monte Carlo methods. Another option is to use graph theory (or other domain detection methods), to identify regions of similar topology and perform analysis within these regions. This has been described for crypts in gut tissue. One key challenge in our analysis is that the bone marrow is known to be a less organised structure. Nevertheless, regions may be identified, for example peritrabecular fat spaces, areas of megakaryocyte clustering (as the cells are typically larger), or areas of cellular proliferation.

In addition, Summers, Wills, and Rees [76] also highlight the importance of edge effects where incomplete coverage of a local neighborhood needs to be corrected for. In Velten and Stegle [77], the authors discuss how to gain mechanistic insight from spatial analysis. They suggest using spatial distance as a proxy for closeness to refine pseudotemporal clusterings. In addition, receptor-ligand interactions could be validated in the same way.

2.9 Analysing Spatial Data

One of the fundamental principles in spatial biology is that structure informs function. This concept has long been recognized in developmental biology, such as during gastrulation, and is routinely applied in clinical settings, particularly in the pathological analysis of biopsy samples[78]. Just as identifying cell types from gene expression enhances interpretability, methods have been developed to cluster transcripts and spatial information into discrete units [78]. These units aim to reveal distinct tissue structures, which can then be further analyzed to understand their roles within the broader tissue system. Just as spatial-level analysis differs from suspension-level analysis, examining tissue at the regional level provides insights that differ from those gained at the whole-tissue level [78].

The decision on how to divide a domain is non-trivial and depends on the specific biological question being investigated. In biology, common spatial scales include: the paracrine scale, which captures close interactions between neighboring cells; the mesoscale, such as brain layers or the peritrabecular region, and the macroscale, which reflects the overall composition of the tissue; keeping in mind that a biopsy represents only a subset of a larger system. As discussed in the context of the Modifiable Areal Unit Problem (MAUP), the way tissue is partitioned can significantly influence the conclusions drawn [79].

In their recent review, Gulati et al. [80] propose that the fundamental unit for interpreting spatial transcriptomic data is the multicellular neighbourhood, a cell together with its surrounding local environment. This local environment can be defined in various ways, such as by a fixed radius, a fixed number of neighbouring cells, or graph-based methods like Delaunay triangulation. The goal is to extract feature sets that characterize each multicellular neighbourhood and then group those with similar features into spatial ecotypes; recurrent multicellular neighbourhoods composed of co-associated cell states. These ecotypes are defined by three key properties: cell type composition, cell state heterogeneity (e.g., gene expression), and spatial organization.

A central challenge in defining multicellular neighbourhoods and, by extension, spatial ecotypes, is determining which features best represent them. This involves not only selecting the types of features but also deciding how they are processed for clustering. For example, features may be based on the abundance of cell types and states, or derived by aggregating gene expression

across cells using approaches such as weighted averaging, gene–gene covariance matrices, or deep learning methods [80].

Analyzing multicellular neighbourhoods using gene expression alone can capture more complex biological variation than relying on cell type composition. However, this approach is often more susceptible to batch effects. Methods such as CellCharter and CCST address this by performing batch correction prior to neighbourhood analysis, while NicheCompass integrates embedding and batch correction into a single step.

As a result, two key parameters must be considered when defining multicellular neighbourhoods: (1) the format of the input data, and (2) the spatial scale at which the neighbourhood is defined.

2.9.1 Unsupervised Tissue Domains

In many tissues, the domain to be analyzed is well defined. For example, tumour regions are often morphologically distinguishable from normal tissue in H&E-stained sections. Analyses in such contexts may focus on tumour infiltration, or on the spatial aggregation and gene expression profiles of cells within the tumour itself. Other canonical examples of well-defined domains include the crypts of the colon and the layered architecture of the brain. Failing to account for these regions can introduce noise and obscure the biological processes specific to each domain. For instance, combining portions of adjacent brain layers may lead to misleading interpretations of cellular interactions.

When such regions are morphologically distinct, the most reliable approach is expert annotation by a pathologist. However, when domains are not clearly visible on H&E, they must instead be inferred using alternative features such as cell type composition, gene expression profiles, spatial organization, or a combination of these. Numerous computational methods have been developed to identify tissue domains from spatial transcriptomic data. Many of these have been benchmarked on tissues with well-characterized morphological structure, such as the brain and colon.

Domain Expectations

Walker et al. [81] provide a useful framework for thinking about how to define domains in spatial transcriptomic data, emphasizing that the choice of method depends heavily on the type of domain one aims to capture. The most straightforward approach, particularly when no morphologically obvious regions are present, is to define domains as spatially contiguous groups of cells with maximally similar gene expression. For instance, monocytes expressing inflammatory markers might cluster into one region, while T cells with granulocytic capacity form another. However, this approach may overlook biologically meaningful tissue architecture, where distinct cell types with different functional gene expression together form a coordinated functional unit.

Walker et al. [81] therefore suggest three other ways to define a domain.

1. **Heterogeneous expression domains**: Domains have different combinations/distributions of gene expression or heterogeneous collections of cells with different gene expression but these are distributed so there is no clear sub-regions.
2. **Spatial pattern domains**: Domains differentiated by the spatial organization of gene expression. For example, one region may exhibit a layered structure, while another may display a dispersed, "salt-and-pepper" distribution of cell types.
3. **Interaction-based domains**: Regions defined by patterns of cell–cell interactions, where the functional relationships among cells, rather than their individual profiles, drive domain boundaries.

This framework highlights that the most appropriate domain definition depends on the underlying biological question and tissue context.

Methods for detecting domains

A wide range of computational methods have been developed to identify tissue domains or spatial ecotypes from spatial transcriptomic data. One class of methods incorporates spatial information during the preprocessing stage, before clustering. These methods are often extensions of traditional scRNA-seq clustering approaches but are adapted to account for spatial context. For example, spatialLIBD uses spatial information to identify spatially variable genes (SVGs),

particularly those showing layer-specific expression, and performs clustering only on these selected genes. In contrast, stLearn integrates spatial information by using computer vision techniques to assess morphological similarity from H&E images and then smooths the gene expression data based on this similarity prior to clustering [81].

Other methods incorporate spatial information directly at the clustering stage. A prominent subset of these approaches employs probabilistic models such as Hidden Markov Random Fields (HMRFs); a framework that assigns weights to possible cell-to-cluster assignments based. Classical examples of HMRF applications include the Ising model (from statistical physics), the Potts model (from ecology), and image denoising. Spatial transcriptomics tools in this category include SmfishHmrf and BayesSpace.

Other approaches have used machine learning techniques. SpaGCN detects both SVGs and spatial regions by learning a representation from gene expression aggregation from the surrounding points through a Graph Convolutional Network. SEDR creates an embedding which captures gene expression and spatial information via an autoencoder. Such deep learning-based methods often suffer from limited interpretability, making it more challenging to understand the biological basis of the resulting clusters.

A benchmark analysis of domain detection methods

Yuan et al. [82] conducted one of the most recent and comprehensive benchmarking studies of spatial domain detection methods in 2024, evaluating performance on both 10x Visium and MERSCOPE datasets. The study compared a wide range of spatial methods, including SpaGCN, BayesSpace, stLearn, SEDR, CCS, SCAN-IT, STAGATE, SpaceFlow, conST, BASS, and GraphST. These methods were grouped based on how they incorporate spatial information: preprocessing-based (e.g., stLearn), probabilistic (e.g., BayesSpace, BASS), and deep learning-based (all others).

A key challenge was selecting appropriate evaluation metrics, as each captures different aspects of performance. The authors recommend prioritising methods that perform well on the metrics most relevant to the specific biological question and tissue context. Normalized Mutual Information (NMI) was used as the primary accuracy metric. It measures how much information is shared between the predicted clusters and the ground truth, making it the most widely used metric for

spatial domain evaluation. Additional metrics included the homogeneity score (HOM), which favors methods that subdivide known domains, and the completeness score (COM), which rewards merging small regions into larger domains. Three other metrics were used to evaluate continuity (spatial coherence and clear domain-domain interfaces), CHAOS, PAS and ASW. The authors also evaluated the ability of each method to define domains with high quality marker gene expression, using spatial autocorrelation with Moran's I or Gearys C.

The authors suggest prioritising a method that performs well for the scale on which you are interested. For example for finer grained features this would be SpaceFlow. BASS has a better continuity score, indicating smoother boundaries. Whilst BASS, GraphST and BayesSpace have high NMI on Visium data. In merFish data the authors instead recommend BASS, SpaceFlow and SCA-IT.

Another 2024 paper has performed a similar benchmarking study [83]. BASS, GraphST, ADEPT, BANKSY and STAGATE were top performers. As with the previous study, performance varied depending on the technology. For example, GraphST performed less well on single-cell–resolved data, while SpaceFlow ranked third on MerFISH data. Performance was also tissue-specific: ConGI performed well in tumours but poorly in brain tissue.

Domain Interpretability

While domain detection methods can identify spatial regions, these regions still require post hoc biological interpretation. Classical dimensionality reduction methods, such as PCA, NMF and latent Drichlet allocation, return interpretable embeddings. However, they often fail to capture spatial context or the full complexity of the data, such as non-linear relationships or intricate spatial patterns. Given this, as the field developed, better domain capture with autoencoders or graph neural networks were balanced with the need for interpretability post hoc. More recently, the field has shifted toward improving interpretability by incorporating spatial structure directly into dimensionality reduction frameworks. Approaches using Gaussian processes or hidden Markov random fields, for example, model spatial correlations explicitly and can yield latent factors that align more intuitively with biological or spatial domains. While these methods are often more interpretable than deep learning black-box models, they tend to be computationally intensive. Other examples, for example STAMP, uses topic modelling to link gene expression to the latent representation of spatially organised topics computed from a simplified graph

convolutional network [84]. This provides a more transparent and spatially aware representation of tissue organization.

Method Scalability

Another challenge is ensuring these domains are comparable across conditions and multiple samples. Yuan et al. [82] also highlight the problem of handling large scale data, i.e. domains across multiple samples.

The first challenge in scaling to multi-tissues is batch correction of gene expression. To solve this Yuan et al. [82] integrated spatial domains from SpaceFlow across multiple tissues using Harmony, in a new method named SpaceFlow-DC. MNN, Harmony and canonical correlation methods are commonly used for scRNA-seq data integration. The problem of integrating large spatial datasets has been subject of a recent review by Guo et al. [85], who discuss using the spatial coordinates to help reduce noise when performing integration in latent space. Hu et al. [83] therefore also benchmarked techniques for alignment and integration of spatial transcriptomic slices. They evaluated PRECAST, DeepST, STAligner and SPIRAL. STAligner had mildly better UMAP results. For domain identification in consecutive slices, BASS, STAligner then SPIRAL scored best. STAligner however performed better when the tissue slices were from distinct anatomical regions in the same tissue (e.g. aligning clusters from the anterior and posterior brain regions).

Two other challenges in scalability are the computational requirements and balancing analysing many cells whilst still detecting domains of low prevalence. scNiche, BANKSY, CellCharter and UTAG are currently some of the only methods which can scale to more than 3 million cells, as in our dataset [86]. A more recent method Quiche, has also been designed to work across multiple conditions, and identify niches at low prevalence. However, Quiche only takes cell labels as input rather than gene expression [**ranek**]. Other techniques such as COVET and GASTON rely on gene expression and gradients or covariance matrices [87]. ENVI an extension of COVET also aims to identify spatial gradients from snRNA-seq data[88].

2.9.2 Spatial Statistics

Spatial transcriptomics data is a collection of points where each point has a collection of associated features, either cell label or gene expression, as well as a spatial coordinate location. In their recent review, [89] suggested categorising analysis of spatial data into two categories, lattice based and point process.

Lattice based assumes locations are fixed and spatially dependant. We are therefore studying observations, whilst taking into account that the observations are not independent due to the spatial structure. Point processes assumes that point locations were generated by a stochastic process and the goal is then to study the properties of the process that created the pattern. Examples of lattice based analysis include local Morans I and bivariate Lees L, whilst Ripley's K and the empty space function are examples of point processes. Either approach can be used for Xenium data, the lattices constructed will be irregular lattices.

Point Pattern Analysis

By considering a tissue as a point pattern analysis, we aim to find the mechanism responsible for a pattern. This can be done in two ways. One, by considering that all the cells in the tissue depend on each other and the distribution is from one overarching process. This is known as multi-type. Two, we consider that N patterns can be created by N point processes which can all be individually analysed. This would be a univariate view point. These processes are not necessarily independent, but this is a common assumption.

There are multiple assumptions built into point pattern analysis. First, related to sampling, is that we do not observe the entire pattern only a small sample of it; this is known as window sampling where we are only seeing part of a bigger pattern. Another related concept is the small world model, where points are only observed in a finite space and not beyond those boundaries, for example Emons et al. [89] suggest the arrangement of the epithelium as an example. Second are the statistical properties of the point pattern. Key amongst these are whether the null state of the points is homogeneous. Homogeneity assumes that the number of points in a given region is proportional to the area and that there is uniform intensity within the area of sampling. However, it is often hard to distinguish between genuinely inhomogeneous intensity vs a genuine interaction (clustering) of points [89]. Inhomogeneity is a still random process but there are

changes in intensity, for example, there are more trees in a wetter area than a dry area but the distribution of the trees is still random.

Key challenges include identifying whether a whole tissue can be represented as a single point process, whereas analysing cell types individually might underestimate the dependancies between cells. It is also possible that the sample is not representative of the larger process.

Lattice based Analysis

Alternatively, we can regard the locations of the cells as fixed in a regular or irregular lattice. Each point is assigned a weight and the stronger the weight the higher the connection. The connections can either be direct contact, graph based, distance based or at a scale of n neighbours. The most common exploratory analysis for this lattice is spatial autocorrelation, in which the degree of association of features are measured, where features are assumed to be dependant in space. This can either be univariate, where one feature is compared over multiple locations or allow for the comparison of two (bivariate) or more (multivariate) features at once. Generally these features are assumed to be continuous, but some metrics such as join count statistics do allow categorical variables. As before spatial autorrelation can be measured on a local or global level [89]. The choice of how to construct the weights matrix greatly affects the results.

3

Preparing the Xenium Dataset

Chapter Contributions to the Literature

- A processed Xenium ST Dataset of Healthy and MPN bone marrow
- Pipeline for processing Xenium Spatial Transcriptomic Data from raw data to analysable output
- Function for incorporating cells with irregular shapes into ST voronoi segmentation
- Improved quality control of cells with spatial annotations on the tissue
- Improved cell annotation by integrating 3 layers of cell segmentation and supervised learning from full scRNA-seq datasets
- Identified significantly more megakaryocytes in the bone marrow trephine from gene expression than we could previously detect morphologically, with consequences for H&E analysis
- Integrated macrostructures and ML outputs from image data to augment the ST analysis
- Improved the graph creation to capture neighbours of irregular shaped cells more accurately

3.1 Author Contributions

This project was conceived by Professor Daniel Royston and Dr Rosalin Cooper. I thank the patients, and those involved in ethics and sample acquisition. Dr Cooper generated the data and advised on clustering, also reviewing the unsupervised clustering output. The pipeline for megakaryocyte detection, segmentation, and phenotyping was developed by Dr Korsuk Sirinukunwattana. I received assistance from Dr Hosuk Ryou for initial implementation of the

megakaryocyte pipeline. He also performed bone segmentation and generated the fibrosis map. Prof. Royston and Dr Carlo Pescia annotated spatial regions, and Dr Pescia annotated arterioles and sinusoids. I performed all data analysis.

3.2 Introduction

3.2.1 Aims

In this section, I present a pipeline for processing Xenium spatial transcriptomic data from raw input to analysis-ready output, integrating image data from stained bone marrow trephines used in the clinical workflow. First, I enhance quality control by incorporating pathologist-informed review, revealing that standard workflows may include low-quality regions as still deemed analysable without expert input. Second, I integrate machine learning algorithms trained on H&E slides to detect, segment, and phenotype megakaryocytes, cells with complex morphologies often missegmented in standard pipelines. I developed a function to integrate these cells back into the original Xenium segmentation without disrupting existing cell boundaries, which can also be applied to other large irregularly shaped cells. Third, I improve upon existing unsupervised cell annotation (15um Voronoi segmentation) by combining multiple segmentation layers and supervised reference-based annotations. This yields better specificity and sensitivity in lineage marker detection. Fourth, I integrate macrostructures, adipocytes, bone, and vessels, from H&E images to contextualise the spatial transcriptomic data. This enables recovery of adipocytes, which are often lost during tissue processing, and supports analysis of suggested HSPC niches. Finally, I implement a graph construction method based on Xenium's Voronoi outputs that preserves neighbour relationships even for irregularly shaped cells. I highlight the risks of relying on default graph-building tools and demonstrate that this method ensures accurate spatial representation of all cells, including large megakaryocytes.

3.2.2 Xenium

Xenium is a spatial transcriptomic technology that allows single cell resolution on fresh frozen or FFPE tissue sections. Selected probes hybridise to RNA targets in the tissue and bound probes are then ligated at RNA target sites to allow rolling circle amplification of the probe and

a unique barcode. A series of cycles allows sequencing and allows the identification and location of each RNA molecule.

3.3 Data

3.3.1 Sample Selection

We analyse samples from bone marrow trephine biopsies, typically taken from the hip using a needle to extract a 1–2 cm core of tissue. These biopsies are routinely used to diagnose MPN. In clinical workflows, the trephine is formalin-fixed, paraffin-embedded (FFPE), decalcified, and sectioned into 4 um slices. One section is stained with H&E for morphological assessment, while others are stained with immunohistochemistry (IHC) to detect specific cell types based on protein expression (e.g. B cells). When MPN is suspected, a standard panel of eight IHC stains (CD3, CD20, CD79a, MPO, glycophorin C, CD61, CD34, and c-kit) and a reticulin stain is typically used. Additional diagnostic data include bone marrow aspirates, peripheral blood counts, and variant allele frequencies of detected driver mutations.

Samples for this study were selected from the archive based on high-quality H&E morphology and architectural preservation. Pathological review excluded samples with artefacts (e.g., crush, haemorrhage, folding) or significant sclerosis, which can impair RNA quality. Prior to Xenium processing, RNA was extracted from adjacent sections to assess integrity via DV200 or RIN (RNA Integrity) scores.

3.3.2 Sample Preparation

After morphological review, a thin section was cut from the FFPE block and transferred to a water bath for rehydration. Blocks were lightly scored before sectioning to aid macrodissection. Tissue handling was minimised to preserve integrity and avoid detachment, critical for spatial analyses where larger continuous areas improve interpretability. Bone detachment, in particular, can confound distance-based analyses. Decalcification was performed using EDTA, selected over stronger acids to preserve nucleic acids while still enabling sectioning by removing mineral content.

Diagnosis	Mutational Status	Count	Total Trephine Count
Normal	NA	5	5
ET	JAK2	5	8
	CALR	2	
	TN	1	
PV	JAK2	3	3
PrePMF	JAK2	1	3
	TN	2	
MF	JAK2	10	13
	CALR	3	

Table 3.1: Distribution of mutational status across diagnoses. The table shows the count of bone marrow samples with each mutational status (TN, JAK2, CALR) or NA (non-applicable) and the total count of samples for each diagnosis.

3.3.3 Cohort Metadata

Our cohort comprised of 32 bone marrow trephines taken from 30 patients. Our cohort was selected to cover the full range of the MPN disease spectrum and importantly included samples from patients who had required a bone marrow biopsy but analysis did not determine them to have a disease of the bone marrow. These are commonly known as reactive samples, which we will call normal and use as a control (Table 3.1).

Of the 13 MF samples, two were post-PV MF (18286_R1 and 18669_R2). In addition, in MF, 2 patients had two bone marrow trephines each (18612_R1 and 18612_R4 belong to one patient, and 18612_R2 and 18612_R3 belong to the another patient). Therefore in total we have 11 unique MF patients (13 trephines), 3 PrePMF patients, 8 ET patients, 3 PV patients and 5 Normal patients to compare.

It is important to note the VAF (Variant allele frequency) for mutational status when running comparisons, as not all cells may contain the driver mutation. The bone marrow microenvironment is known to change with age [90]. Therefore, patients were age matched.

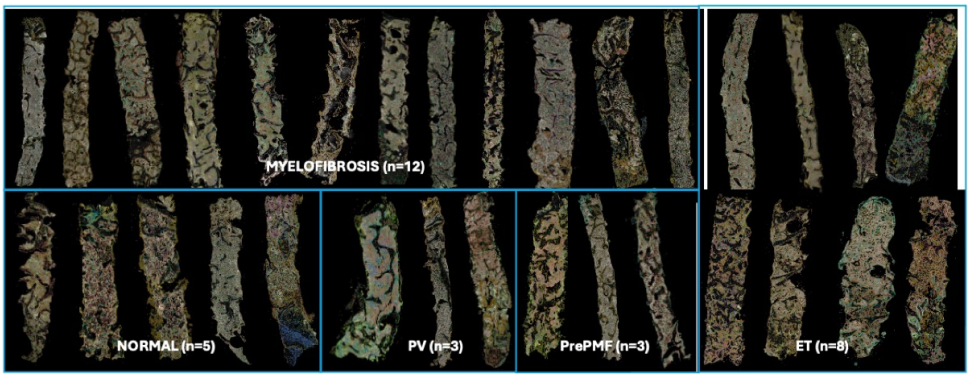


Figure 3.1: Output from Xenium for each bone marrow in our samples (n=32). Colours are assigned by Xenium platform.

3.4 Results

3.4.1 Sample Quality Control

Xenium Quality Control of the tissue

Xenium incorporates several built-in quality control measures, including the Q-score, which reflects confidence in transcript detection and spatial assignment. A Q-score of 20 corresponds to 99% accuracy (1% error), and a score of 30 indicates 99.9% accuracy. Each 10-point increase reflects a tenfold reduction in error probability.

Q-scores are derived by comparing signal intensities from negative control probes (which do not bind RNA) with target-specific probes. A signal significantly above the negative background is considered a true transcript, contributing to a higher Q-score. Transcripts with Q-scores below 20 are excluded from the dataset, ensuring only high-confidence reads are retained. In our 32 samples, the median number of false positives per cell was just 0.03, and the median percentage of high-quality transcripts was 83.1%, with the lowest sample still scoring 78.5%.

We also observed high transcript counts per cell, supporting confidence in cell annotation. Before any processing, the median transcript count per cell was 33; after processing, this remained high at 35. For comparison, a 10x Genomics study using Xenium in breast cancer tissue reported a median of 62 transcripts per cell. Given the additional challenges of bone marrow tissue, including decalcification, our data are within a comparable range [91].

3.4.2 Quality from a Haemapathologist perspective

Spatially Unpreserved Regions

The bone marrow trephines processed with Xenium are the same FFPE samples used for clinical diagnosis by haematopathologists. Since Xenium preserves the tissue post-imaging, we were able to apply standard H&E staining and assess morphology directly. To ensure only spatially preserved regions were analysed, two expert pathologists reviewed each section, annotating high-quality areas and flagging regions with poor spatial preservation. This quality control step is critical for accurate spatial analysis, as artefacts like bone detachment, crush, haemorrhage, folding, or complete architectural disruption can still produce nuclei and segmentation, potentially misleading downstream analysis. As shown below, without this pathology-led review, low-quality regions could have been incorrectly included (Fig. 3.2).

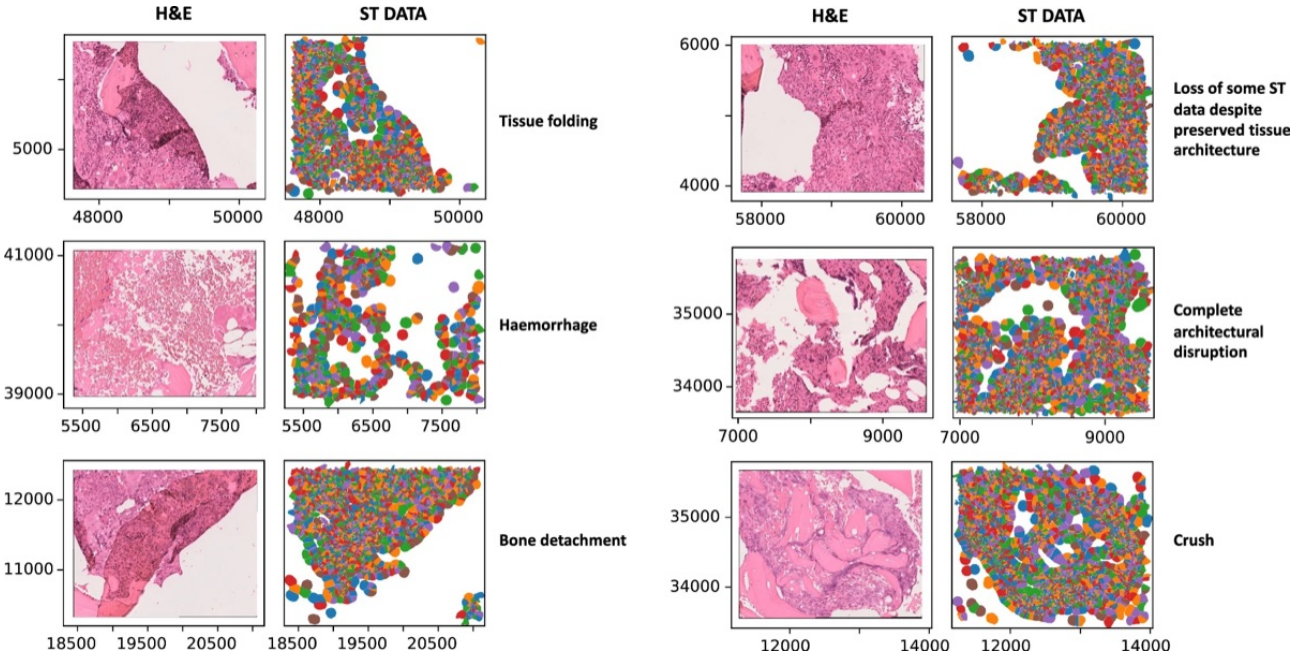


Figure 3.2: Areas in H&E that are spatially unpreserved vs the xenium output within each section

Registration

To incorporate the H&E annotations into the Xenium tissue space, we used an the 10x recommended registration method adapted from Horn (1987). This rigid body transformation aligns each Xenium view with its corresponding H&E image, preserving distances and angles, though it does not account for shearing or non-uniform scaling, unlike an affine transform. The approach proved effective due to the close structural similarity between DAPI and H&E images.

Registration relied on shared landmarks and least-squares optimisation, with accuracy improved by increasing the number of well-matched keypoint pairs. Although the rigid transformation was not perfect across entire samples, manual checks showed misalignments of less than half a nucleus; unlikely to affect spatial annotations that span multiple cells.

The registration outputs annotations in three coordinate systems: (1) on the H&E ome.tiff image, (2) in global Xenium reference space, and (3) on the DAPI image in microns. One pixel in global space corresponds to 0.2125 um; therefore, annotations from Xenium Explorer must be divided by 0.2125 to map accurately onto the H&E image in pixels.

Transcript threshold in the context of H&E

Our dataset combines elements of both imaging and single-cell analysis. A key quality control step in single-cell workflows is filtering out cells that cannot be confidently assigned to a lineage, typically due to low transcript counts.

Although gene expression can be analysed directly, biological interpretation is greatly enhanced by assigning cell types. This requires inferring cell identity based on the genes expressed within each cell, while filtering out low-quality cells lacking sufficient transcripts for reliable annotation. We used the in-house Xenium 15 um Voronoi-based segmentation, supplemented with manually integrated megakaryocytes from H&E. Details of segmentation modifications are described in later sections. Transcripts were assigned to cells if they fell within the 2D polygon boundaries in any z-layer. Transcript aggregation per cell was performed using the SOPA package [92].

Establishing a transcript count threshold for excluding low-quality cells is a key step. To determine this, I assessed the distribution of transcripts per cell across all 32 samples, ensuring no systematic variation by condition or experimental batch. Before any filtering, but after megakaryocyte insertion, the overall median transcript count was 33 per cell, with sample-level variation that appeared unrelated to batch (Fig. 3.3, 3.4).

In regions marked as spatially unpreserved, the median transcript count per cell was 13, compared to 35 in regions deemed high quality, supporting the decision to exclude the former from analysis. While a threshold of 13 transcripts might seem appropriate, applying this cut-off to high-quality intertrabecular regions (containing 4,474,667 cells) would remove 685,542 cells (15.3%). Given the good spatial preservation of these regions in the H&E and no morphological reason to

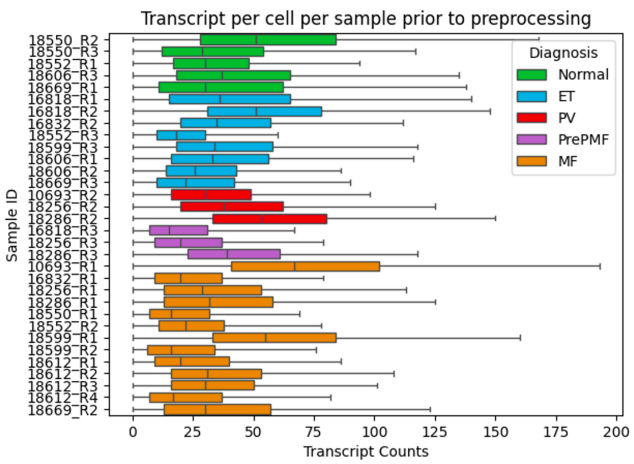


Figure 3.3: Transcripts per sample (n=32) coloured by diagnosis

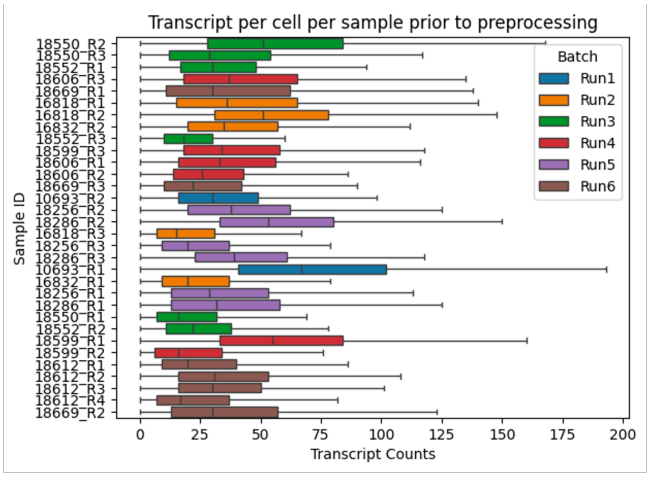


Figure 3.4: Transcripts per sample (n=32) coloured by batch

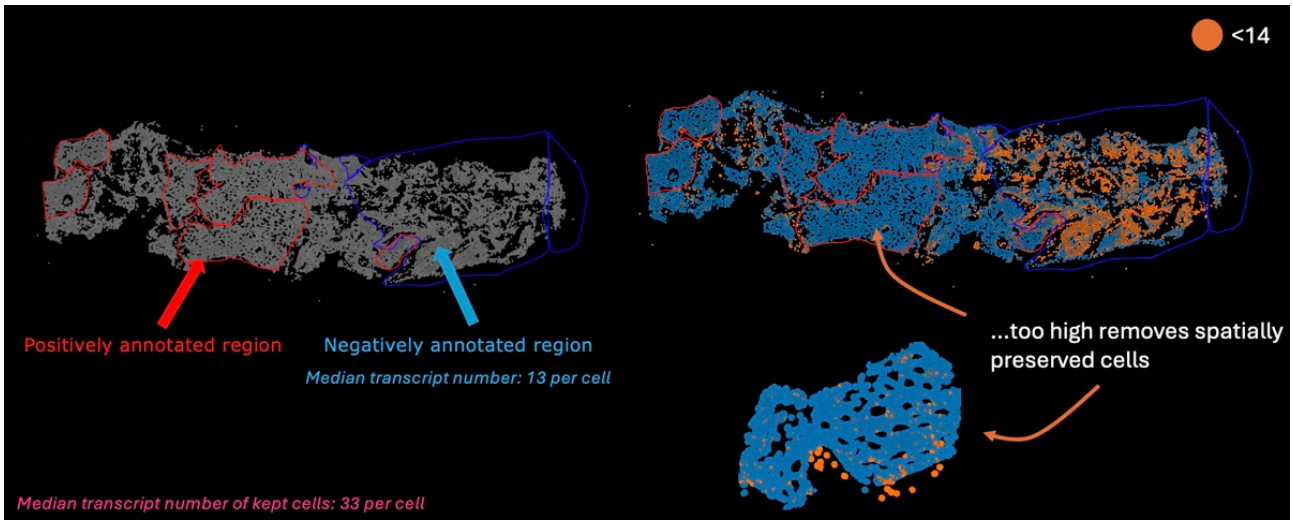


Figure 3.5: Example of a bone marrow trephine with spatially preserved (red) and unpreserved regions (blue) marked on the left. The median number of transcripts in the spatially unpreserved region is 13. However if we set the threshold <14 we can see how we start losing cells (orange) in the spatially preserved region in the same sample on the right. Below the trephine on the right is a blown up spatially preserved region to demonstrate the number of cells (orange) that we lose versus those kept (blue)

consider these cells low quality, such a threshold appears overly conservative. Moreover, cells with 13 transcripts can still be meaningfully annotated if they express canonical lineage markers (Fig. 3.5).

We instead defined the transcript threshold based on both the number of cells removed and the amount removed from preservd vs unpreserved annotated regions. Filtering out cells with more than 3 transcripts results in a noticeable shift; cells begin to be excluded from regions previously annotated as spatially preserved (Table. 3.2). Since our analyses rely on accurate spatial relationships, we aim to retain as many cells from these regions as possible to avoid biases

from missing neighbours. At a threshold of 3 transcripts, only 2.5% of cells in preserved regions are removed, a reasonable trade-off. Moreover, cells with less than or equal to 3 transcripts lack sufficient information for confident cell type annotation, justifying their exclusion(Table. 3.2).

Threshold transcript number	Number of cells to remove (% of all cells)	% of filtered cells in spatially unpreserved region	% of filtered cells in spatially preserved region	% of the spatially preserved region we are removing	% of the spatially unpreserved region we are removing
0	45195 (0.78%)	30.02%	32.67%	0.33%	2.92%
1	108808 (1.88%)	31.34%	35.22%	0.86%	7.34%
2	184545 (3.19%)	31.05%	38.15%	1.57%	12.33%
3	268052 (4.64%)	30.0%	40.83%	2.45%	17.30%
4	355160 (6.15%)	28.9%	43.15%	3.43%	22.07%
5	445035 (7.7%)	27.65%	45.29%	4.5%	26.47%

Table 3.2: Increasing filtering out of cells with more than 3 transcripts starts decreasing more rapidly the cells taken from the spatially unpreserved region. However, I would not be confident making a call on a cell lineage from less than 3 transcripts.

Another consideration when filtering low-transcript cells is whether this introduces bias by disproportionately removing a specific cell population. To assess this, I examined the top genes expressed in the filtered-out cells. Most genes had a median expression of 1 across unfiltered cells (considering only non-zero values), but a subset showed higher expression, potentially due to biological abundance or technical factors, such as stronger probe binding or higher local transcript concentration. These highly expressed genes included AHSP, ALAS2, CTSG, FCN1, GYPA, HEMGN, MKI67, MNDA, PCNA, S100A12, SLC41A1, SNCA, and VCAN.

In cells with less than 3 transcripts, 16.9% express no transcripts. 24.8% express only one transcript, of which 5% are S100A12, 2.9% are SLC4A1, 2.78% are MNDA and 1.72% are CTSG. 29.9% express two transcripts, the most common combinations of genes: MNDA-S100A12, S100A12-SLC4A1, ALAS2-SLC4A1, AHSP-SLC4A1, PTPRC-S100A12.

Of the 28.4% that express 3 transcripts, again the most common are the highly expressed genes, MNDA-PTPRC-S100A12 and MNDA-S100A12-SLC4A1. MNDA is a myeloid cell marker, PTPRC (CD45) haematopoetic cell marker, S100A12 mediates calcium dependant interactions and SLC4A1 codes for anion exchange. This does not indicate a specific cell type and it would be hard to charcterise these cells into a useful lineage. Choosing a threshold of >3 removes 268052 cells from all cells, and 109434 cells from the spatially preserved regions.

3.4.3 H&E to improve segmentation

Segmentation of DAPI

Xenium provides two primary image types: the morphology focus image, a sharply focused capture at a single z-plane, and the morphology MIP (maximum intensity projection), which combines signal across z-planes to highlight the brightest features. DAPI is used to stain nuclei by binding to A-T rich regions in DNA. Xenium segments nuclei using a custom neural network trained on thousands of annotated image patches. Nuclei with over 95% of pixel intensity below a threshold are excluded.

For cell segmentation, Xenium initially recommended 15 um Voronoi expansion from the nuclear boundary, which we followed. This was supported by their 2023 white paper [91]. More recently, they have recommended a 5 um expansion as the default. Both 5 um and nuclear-only partitions were later validated in this thesis, as described in subsequent sections.

More recent options from Xenium include segmentation using Cell boundary stains (e.g., ATP1A1, E-Cadherin, CD45), which provide accurate segmentation near the true membrane and can identify multinucleated or anucleated cells and nuclear and interior stains (e.g., 18S rRNA and DAPI), which infer boundaries without membrane markers, though this may be less accurate in dense tissues. Since our dataset lacked membrane or interior stains, segmentation relied on the 15 um Voronoi expansion from DAPI-stained nuclei.

Two additional approaches were tested; Cellpose and Baysor. Cellpose is a generalist deep learning algorithm that uses an optical flow-inspired method to identify object boundaries. It predicts pixel gradients from the center of each cell toward its edges, enabling generalizable segmentation across diverse image types. Baysor is a Bayesian segmentation tool for spatial transcriptomics. It assigns RNA spots to cells based on spatial proximity, gene expression similarity, and assumptions about cell shape and smoothness. It uses prior segmentation (e.g., DAPI) and optimizes assignments using MCMC and stochastic methods to maximize posterior probability. Other topological approaches exist, such as TopACT, as well as assessment of quality of transcripts within a given segmentation by FastReseg [93].

Although I did attempt Cellpose and Baysor, I did not pursue these segmentations for two reasons. First, both require tuning of parameters that were difficult to optimize across all

tissue regions. Second, the Xenium DAPI segmentation appeared reliable and was supported by benchmarking that showed nuclei alone can robustly capture cell identity [94].

Megakaryocytes: The challenge of multinucleated cells

A limitation of DAPI-based segmentation is its failure to accurately capture cells with polylobated nuclei, an issue particularly relevant for megakaryocytes, which often exhibit this morphology and are central to myelofibrosis. However, to solve this issue we again make use of the accompanying H&E. We already have a pipeline to detect (Single Shot Detection), segment (U-NET) and phenotype (autoencoder) megakaryocytes from H&E images [17]. Phenotypes are created by clustering the latent vector representations of the segmented lab colour space megkaryocytes, so that the phenotypes are defined by megakaryocyte cytomorphology. Therefore, we can apply this algorithm to the accompanying H&E and use our registration to insert the megakaryocytes to replace the voronoi segmentation in that region (Fig. 3.6).

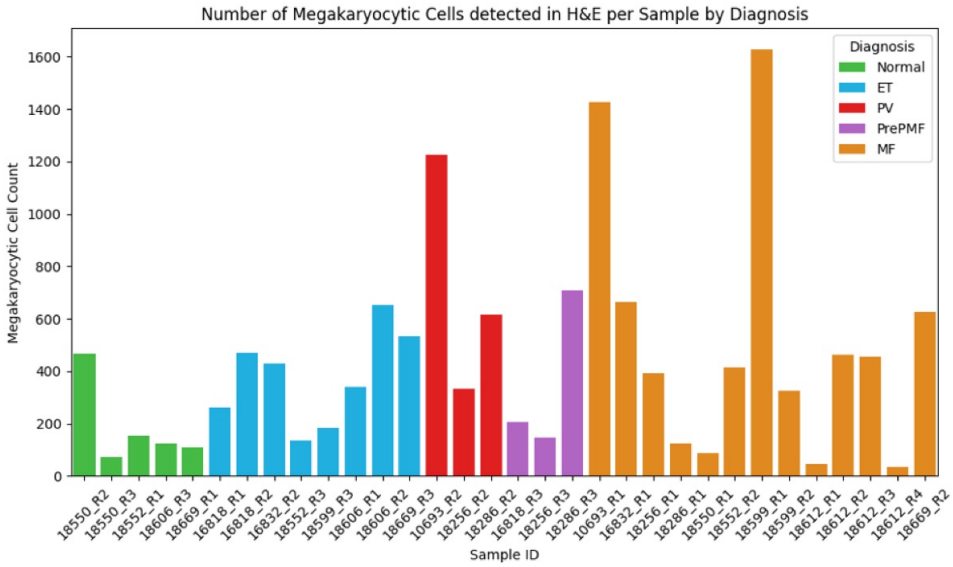


Figure 3.6: Number of megakaryocytes detected and segmented by H&E-based algorithm per bone marrow sample (n=32), coloured by diagnosis.

Inserting megkaryocytes

Using the previous transformation for registration derived from landmarks, I could register megakaryocytes into the same space as the original xenium segmentation.

To integrate megakaryocytes into the original cell segmentation, I developed a custom function which cleans, validates, and merges megakaryocyte geometries with existing cell boundaries.

This process ensures that the final segmentation reflects both the original cell structure and the addition of biologically relevant megakaryocytes, while also accounting for artefacts and cell fragments introduced by overlapping geometries.

The core of the function involves subtracting the megakaryocyte geometries from the existing cell boundaries. This is done using a geometric difference operation, which removes any overlapping areas between original cells and megakaryocytes. The resulting geometries include fully intact cells, fragmented cells (caused by partial overlaps with megs), and leftover polygon fragments. To distinguish meaningful fragments from artefacts, the function checks whether each fragment contains a nucleus centroid. Fragments with a nucleus are considered valid and are labeled as "altered" cells.

Fragments without a nucleus are further classified by size. Those smaller than half the median cell area are labeled as "small cells" and are merged into the nearest valid full cell, based on proximity to the nearest nucleus. Larger fragments without nuclei are retained as "artefacts" but are not merged into existing cells. These classification steps help ensure that biologically relevant segments, that may contain transcripts, are preserved, while artefactual regions introduced by geometric processing are appropriately managed.

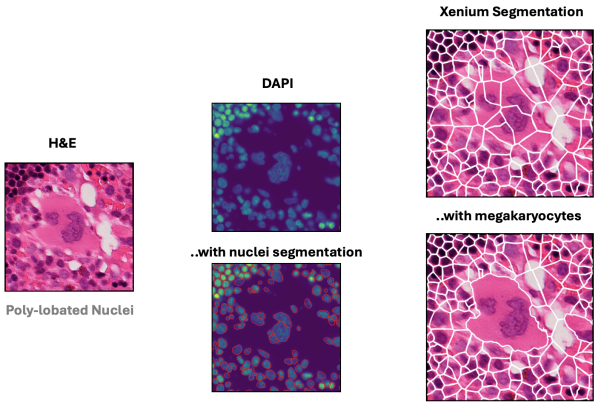
After merging small fragments into their corresponding full cells, all updated geometries, including original, altered, altered with small fragments, artefacts, and megakaryocytes, are combined into a single segmentation. Cells that were modified during processing are given new unique identifiers, while original cells retain their existing IDs. All geometries are re-validated and transformed to the global coordinate space before final export(Fig. 3.7a, 3.7b).

3.4.4 Filtering of low quality cells

We applied three filtering steps to remove low-quality cells from our samples, none of which were used for downstream annotation. First, we excluded cells located in spatially unpreserved regions. Second, we removed cells with fewer than four transcripts. Third, we filtered out artefacts introduced during megakaryocyte insertion.

The number of cells removed at each step was:

1. Spatially unpreserved: 464,793 cells
2. Low transcript count: 268,052 cells



(a) Raw H&E and DAPI images of a polylobated megakaryocyte. The original 15 um Voronoi segmentation splits the cell, but our inserted megakaryocyte segmentation correctly captures it as one cell.



(b) Example of segmentation artefacts following megakaryocyte insertion. Red cells lack nuclei and are excluded from analysis; blue cells are altered with small merged fragments; light blue cells are unchanged from the original segmentation.

Figure 3.7: Integration of megakaryocytes into the segmentation. (A) Correction of polylobated megakaryocyte over-segmentation. (B) Handling of artefacts and alterations introduced during megakaryocyte insertion.

3. Artefacts from megakaryocyte insertion: 4,574 cells

Some cells met multiple exclusion criteria, but the overlap was limited enough to justify applying each filter independently. The Venn diagram below illustrates this overlap (Fig. 3.8a). Given the overlap, this left 5,120,330 cells taken forward to cell annotation.

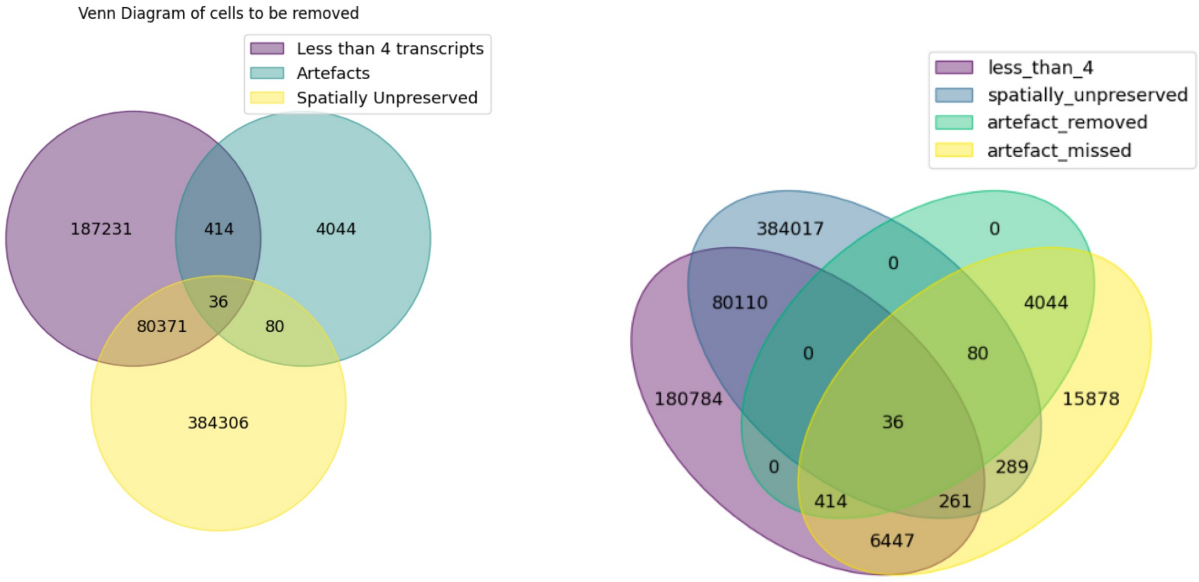
Artefact Correction

After performing cell annotation, I identified an issue with the original method for assigning artefacts. When a megakaryocyte was inserted, the remaining cell was often a multipolygon. My initial approach only assessed the largest fragment of the multipolygon for artefact criteria, inadvertently allowing smaller artefactual fragments to remain classified as original cells and proceed to annotation.

Upon recognizing this error, I corrected the filtering and removed 27,449 artefacts from spatial analysis. Of these, 7,158 (26.1%) had fewer than four transcripts, 666 (2.4%) were located in spatially unpreserved regions, and 4,574 (16.7%) had been correctly identified originally. After accounting for overlaps among these categories, only 11,571 (42.2%) of the artefacts were removed initially, while 15,878 (57.8%) remained and were included in cell annotation.

These missed artefacts represented just 0.31% of the 5,120,330 cells that passed initial transcript and spatial filters. Given this small proportion, it is unlikely they significantly affected cell

annotation, and since they are excluded in downstream spatial analyses, they should not impact spatial statistics (Fig. 3.8b).



(a) Cells removed based on three criteria: spatially unpreserved regions, low transcript counts, and artefacts from megakaryocyte insertion. The overlap between categories is shown.

(b) Overlap of missed artefacts with other filtering criteria. Artefacts that passed initial filters made up only 0.31% of annotated cells and were later removed from spatial analysis.

Figure 3.8: Summary of low-quality cell filtering. (A) Initial filtering by spatial preservation, transcript count, and artefact status. (B) Missed artefacts and their overlap with other filters, showing minimal impact on downstream annotation.

3.4.5 Batch Correction

Raw Expression Matrix

With low-quality cells removed, we proceeded to annotate the remaining cells based on their expression profiles. Unlike typical scRNA-seq datasets, which profile 10,000–20,000 genes with 500–3,000 genes detected per cell, the Xenium panel includes far fewer genes. In our dataset, the median number of genes detected per cell is 37, with a mean of 47.9, an interquartile range (IQR) of 43 (Q1=20, Q3=63), a minimum of 4, and a maximum of 4,436.

This high maximum initially appeared suspicious but was confirmed to correspond to a large megakaryocyte identified from H&E staining (cell ID: 150_18599_R3). Of the 1,277 cells with more than 1,000 transcripts, 1,251 (97.96%) were annotated as megakaryocytes based on H&E segmentation, while 26 belonged to the original segmentation. On visual inspection, these 26

high-transcript cells appear to be additional megakaryocytes lacking polylobated nuclei, which were missed by the H&E-based segmentation.

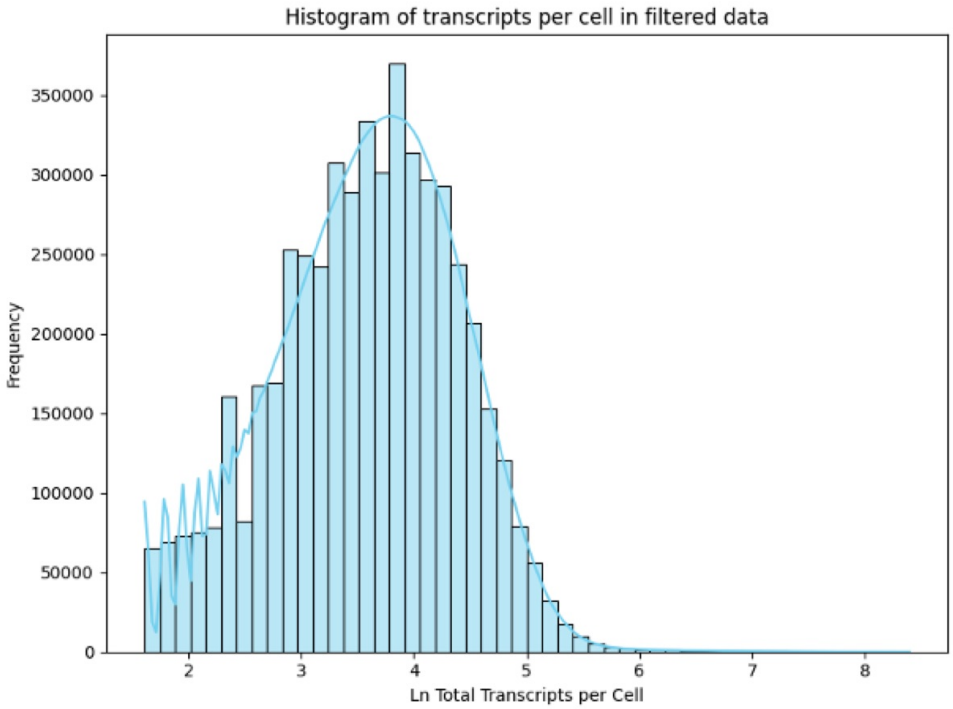


Figure 3.9: Histogram of transcripts per cell in the filtered dataset.

The next step was to examine the range of gene expression levels across cells. Some genes are expressed at much higher levels than others. Additionally, transcript counts per cell vary widely, from as few as four to over 1,000 (Fig.3.9), reflecting differences in sequencing depth. To ensure these technical differences did not dominate cell type identification, normalization was necessary.

We applied SCTransform, a normalization method commonly used in single-cell RNA sequencing (scRNA-seq) analysis. SCTransform addresses challenges such as technical noise, uneven sequencing depth, and high variability in gene expression by fitting a regularized negative binomial model. This approach improves the comparability of expression values across cells, especially when traditional methods like log-normalization fall short.

A key advantage of SCTransform is its use of Pearson residuals to stabilize variance across genes. In single-cell data, lowly expressed genes often show disproportionately high variability, complicating downstream analysis. Pearson residuals measure the difference between observed and expected counts, scaled by the expected variance. This transformation helps control for

Batch	Normal	ET	PV	PrePMF	MF
Run1			1		1
Run2		3		1	1
Run3	3	1			2
Run4	1	3			2
Run5			2	2	2
Run6	1	1			5

Table 3.3: Samples and Conditions per Batch

variability driven by technical factors, making gene expression values more consistent across both genes and cells.

Overall, SCTransform reduces the influence of technical artifacts and enhances the detection of true biological signals, resulting in a dataset better suited for accurate cell type identification.

Harmony for Batch Correction

Once technical noise had been minimized, we assessed the data for batch effects. No substantial batch effect was observed. Conditions were distributed across batches to ensure that any variation between them would be primarily technical and therefore correctable if necessary (Table 3.3).

To correct for any remaining batch effects, I applied Harmony, which operates on a dimensionally reduced representation of the dataset. Prior to this, I used PCA on the SCTransform-normalized data, which had 377 dimensions. I retained the top 30 principal components (PCs), which are linear combinations of genes designed to capture the major sources of variation in the data. SCTransform standardizes gene expression (mean 0, standard deviation 1), ensuring that no single gene disproportionately influences the PCA. An elbow plot showed a clear inflection between PCs 30 and 50, suggesting that additional components beyond 30 primarily capture noise. Since variance explained by each PC is the square of its standard deviation, we can quantify the contribution of each. Assuming the top 50 PCs account for 100% of the retained variation, the first 30 PCs explain 82.5% cumulatively, with PC30 contributing approximately 1%. Harmony adjusts the PCA embeddings rather than the raw data (Fig. 3.10).

It is difficult to quantitatively assess the impact of Harmony, aside from visualizing batch distributions in the reduced dimensional space (Fig.3.10). Distinguishing technical variation

from true biological variation is inherently challenging. In the PCA plot, we observed a distinct right-arm structure dominated by Run 6, which raised concerns about potential biological signal, as this run included a higher proportion of MF samples. However, MF samples were also present in other batches, and after applying Harmony, this arm showed greater mixing across batches. Specifically, when examining Run 6, the separation was no longer pronounced (Fig.3.10). We retained the batch correction, as Harmony operates only in the reduced space and does not affect differential expression statistics. However, it may improve the identification of comparable clusters across batches.

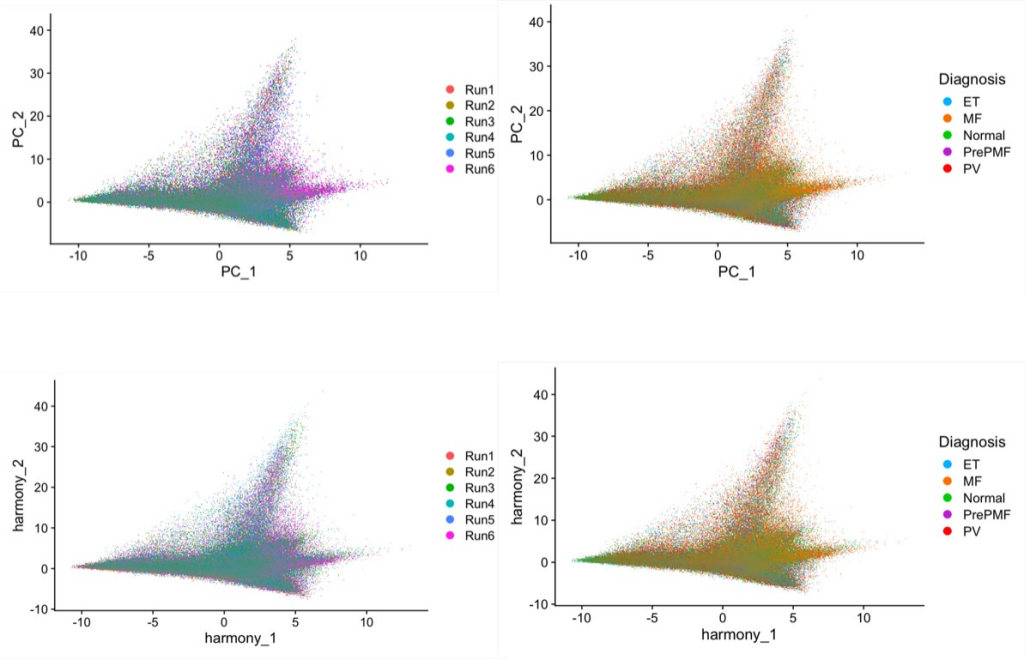


Figure 3.10: PCA of the downsampled object (20000 cells per sample) coloured by batch (left) and diagnosis (right). Top row demonstrates before harmony. Bottom row is after batch correction with harmony.

I also include a plot of batch before and after harmony where points were plotted randomly using shuffle=TRUE within the Seurat DimPlot function (Fig. 3.11)

3.4.6 Cell Annotation

Lineage Markers

Once I had removed technical variation as much as possible, it was time to find cell types within the data. Cells can be annotated by which canonical lineage markers they express. We are using

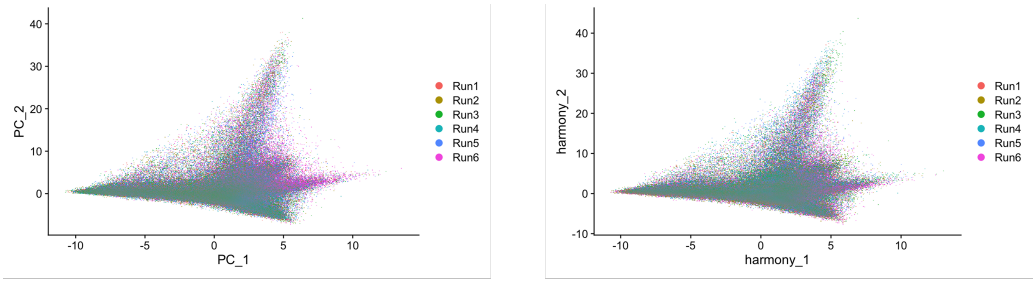


Figure 3.11: PCA of the downsampled object (20000 cells per sample) coloured by batch. Left demonstrates before harmony. Right is after batch correction with harmony. Cells are plotted at random

a pan cancer gene panel of 377 genes, however, there are certain populations we would expect to find within the bone marrow as in table 3.4.

Cell type	**Markers**
HSPC	CD34, KIT, GATA2
Megakaryocyte	MMRN1, PF4, GNG11, VWF
T Cells	CD3D, CD3E, CD4, CD8A, TRAC
NK Cells	KLRD1, KLRB1, NKG7, GNLY
B Cells	MS4A1, CD19, CD79A
Erythroid Cells	ALAS2, GYPA, GYPB, AHSP
Stromal Cells	VCAN, COL5A2, PDGFRA, PDGFRB
Endothelial Cells	PECAM1, VWF, EGFL7, CD34
Monocytes	CD14, FCGR1A, FCGR3A, CD68
Macrophages	CD163, CD68, MRC1
Granulocytes	CTSG, MPO
Dendritic Cells	CD1C, CLEC10A, LILRA4

Table 3.4: Canonical Markers for the major cell types we would expect in the bone marrow

Due to the large size of the dataset, whilst most analyses were performed on the full object, visual exploration was done on a downsampled version containing 20,000 cells per sample. To begin examining cell populations, I visualized the expression of canonical marker genes across the dataset. These patterns help indicate where we might expect specific populations to be located spatially (Fig. 3.12).

A UMAP was constructed using the Harmony batch-corrected, PCA-reduced space. I selected 20 principal components from Harmony as input. Using Harmony reduces batch effects that might otherwise lead to spurious clustering, while PCA helps denoise the data and emphasizes major sources of global variation. However, this approach also introduces limitations: PCA captures only linear relationships, and focusing on global variation may obscure finer differences between closely related cell types. UMAP parameters were chosen to prioritize resolution of

local structure. Specifically, I set the number of neighbors to 5, meaning each cell is connected to its 5 nearest neighbors, and the minimum distance to 0.05, which controls how tightly points are packed in the 2D embedding. These settings enhance separation of closely related populations, such as monocyte subtypes.

It's important to note that while UMAP provides an intuitive visualization of the data, clustering is not performed in UMAP space. UMAP is optimized for visual clarity rather than preserving exact distances or relationships between cells. Unlike PCA, which is deterministic, UMAP is stochastic and can yield slightly different results on each run. There are two methods for annotating cells, either supervised using a reference dataset or unsupervised.

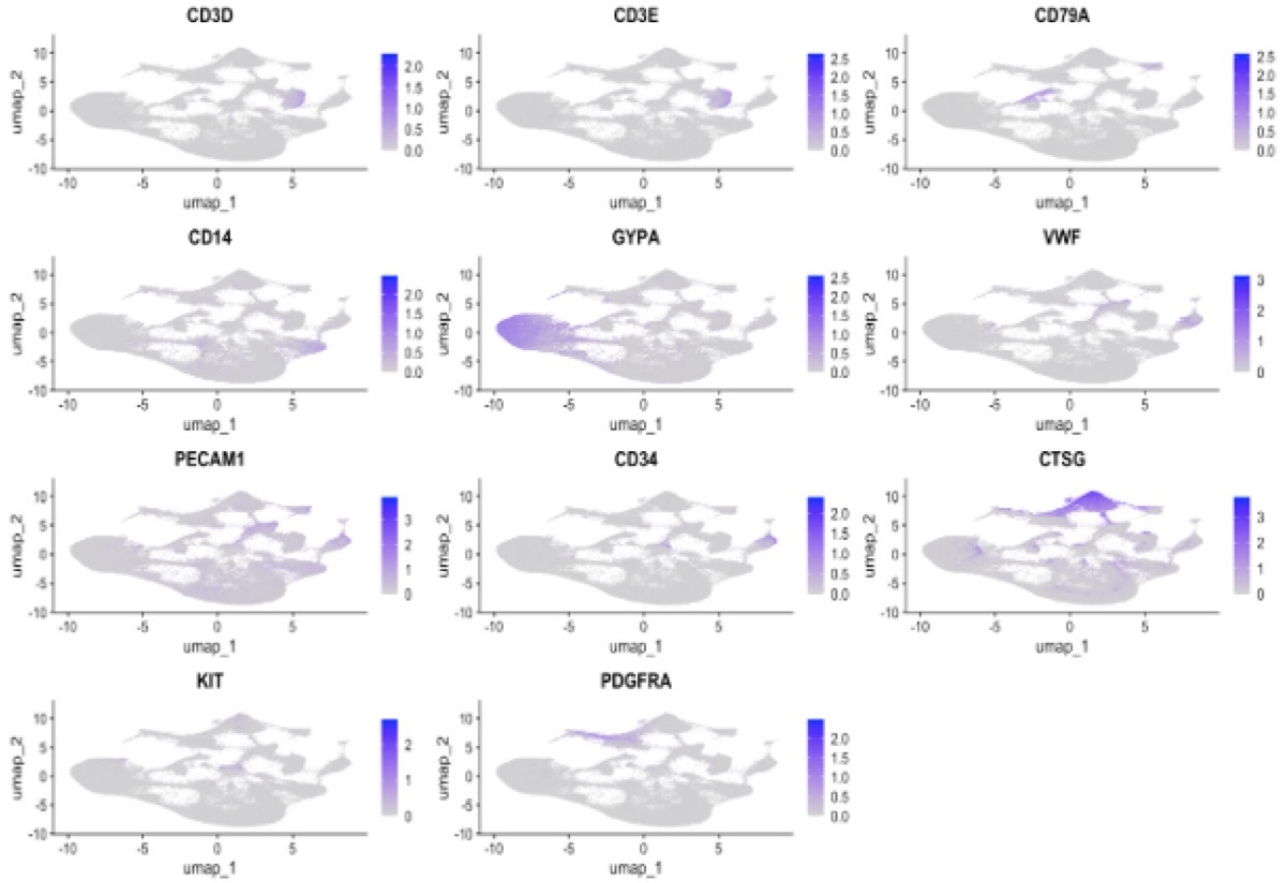


Figure 3.12: UMAP of the downsampled object (20000 cells per sample). I have visualised the expression of key genes used for lineage determination. Each point is a cell and they are coloured by their expression the the gene of interest. In this way we can get a sense of where key lineage populations will be visualised on our UMAP

Unsupervised Cell Annotation: Louvain Clustering

For unsupervised clustering, I used the Louvain algorithm, which requires a graph-based representation of the data. To construct this graph, I used the first 20 Harmony-corrected principal components and applied the annoy method to identify nearest neighbors. Annoy builds a forest of random projection trees to efficiently approximate the k-nearest neighbors for each cell, in this case, using k=20. From the kNN graph, I then constructed a shared nearest neighbor (SNN) graph to reduce noise and minimize spurious edges. The SNN graph strengthens connections between cells that share many neighbors, quantified using the Jaccard index which measures the overlap between two cells' neighbor sets.

Louvain clustering was then applied to this SNN graph. The algorithm optimizes modularity, which quantifies how strongly cells are connected within clusters compared to between clusters. Initially, each cell is assigned to its own cluster, and the algorithm iteratively merges cells into neighboring clusters if it increases overall modularity. This continues until no further improvements are detected. The resolution parameter controls the granularity of clustering. Lower resolutions yield broader, coarser clusters, while higher resolutions produce more refined clusters by encouraging faster modularity optimization.

To determine the optimal clustering resolution, I tested values from 0.2 to 1.0 in increments of 0.2, resulting in cluster counts of 18, 27, 30, 41, and 51. Smaller clusters remained stable across resolutions, while larger clusters at lower resolutions progressively split into subclusters as resolution increased (Fig. 3.13). To identify defining markers for each cluster, we used the Wilcoxon rank-sum test via FindAllMarkers. However, this clustering did not produce a distinct cluster corresponding to HSPCs.

Defining HSPCs

HSPCs were of particular interest given the bone marrow's role in hematopoietic differentiation. Due to the limited gene panel, we manually defined HSPCs as cells co-expressing KIT and CD34, which serve as reasonable proxies. While we initially considered excluding cells that also expressed lineage-specific markers, this criterion proved too restrictive (Fig. 3.14). Ultimately, all KIT+CD34+ cells were retained as HSPCs, regardless of lineage marker expression.

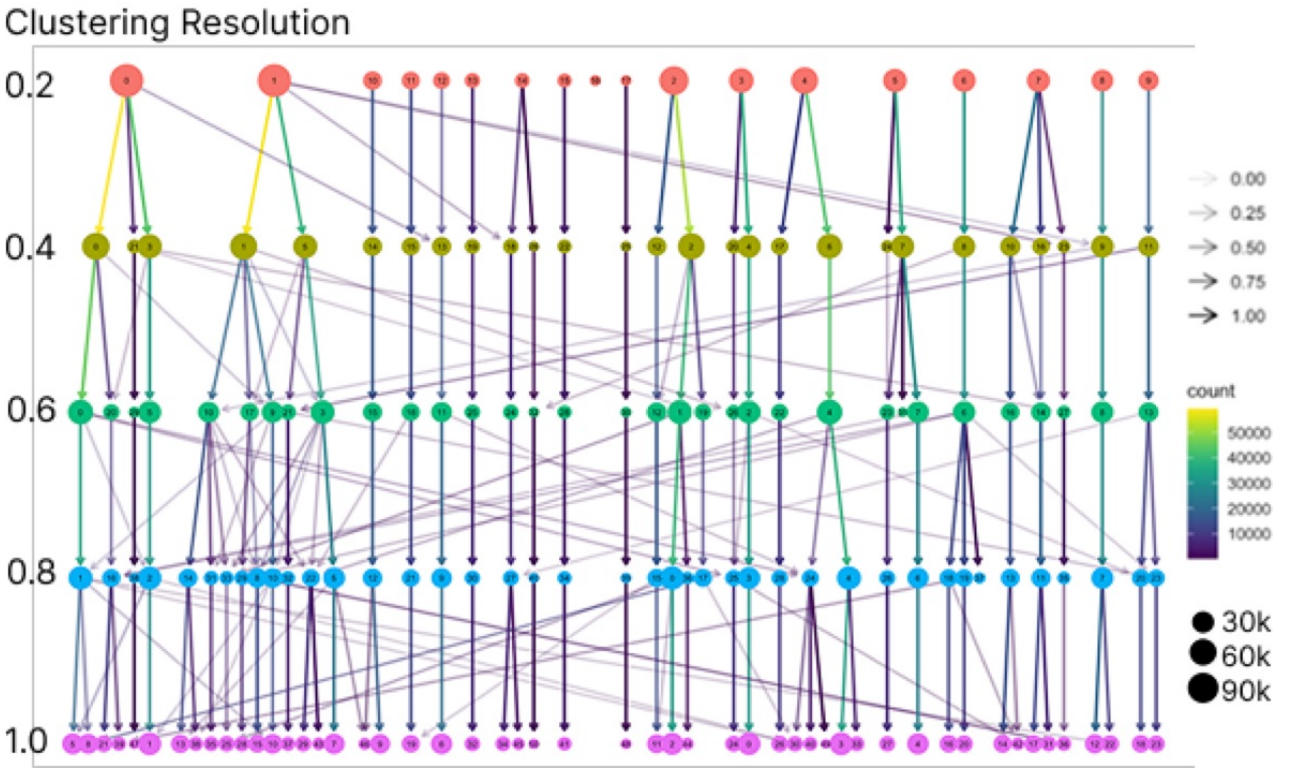


Figure 3.13: Hierarchy of clustering with Louvain at different resolutions. Arrows represent the segregation of cells into the various clusters at different resolutions. This is the result of clustering on the whole object, visualised on the downsampled object (20000 cells per sample)

From a total of 5,104,321 cells, 27,607 (0.54%) co-expressed KIT and CD34. After removing 548 cells identified as megakaryocytes via H&E staining, 27,059 HSPCs remained for analysis. Lineage markers were defined as: PDGFRA, GYPA, GYPB, VWF, MMRN1, PECAM1, CLEC14A, CD3D, CD3E, CD8A, CD4, CD14, CD68, CTSG, and FCGR3A. Of the 27,059 HSPCs, 1,250 (4.62%) expressed none of these lineage markers. Among those that did, the most common co-expression patterns involved combinations of granulocyte, megakaryocyte, and endothelial markers. These patterns may reflect segmentation artifacts (e.g., neighboring cells merged), ongoing differentiation, or misclassification of true HSPCs (Fig. 3.14).

Unsupervised Re-clustering

After defining and removing HSPCs from the dataset, I reran clustering at resolutions of 0.6 and 1.0, resulting in 32 and 51 clusters, respectively. Most clusters remain consistent across resolutions, but the higher resolution reveals additional smaller clusters due to finer modularity optimization. Given this, I chose to annotate the higher-resolution (1.0) clustering and merge clusters if needed. Marker genes were identified using the Wilcoxon rank-sum test, requiring a

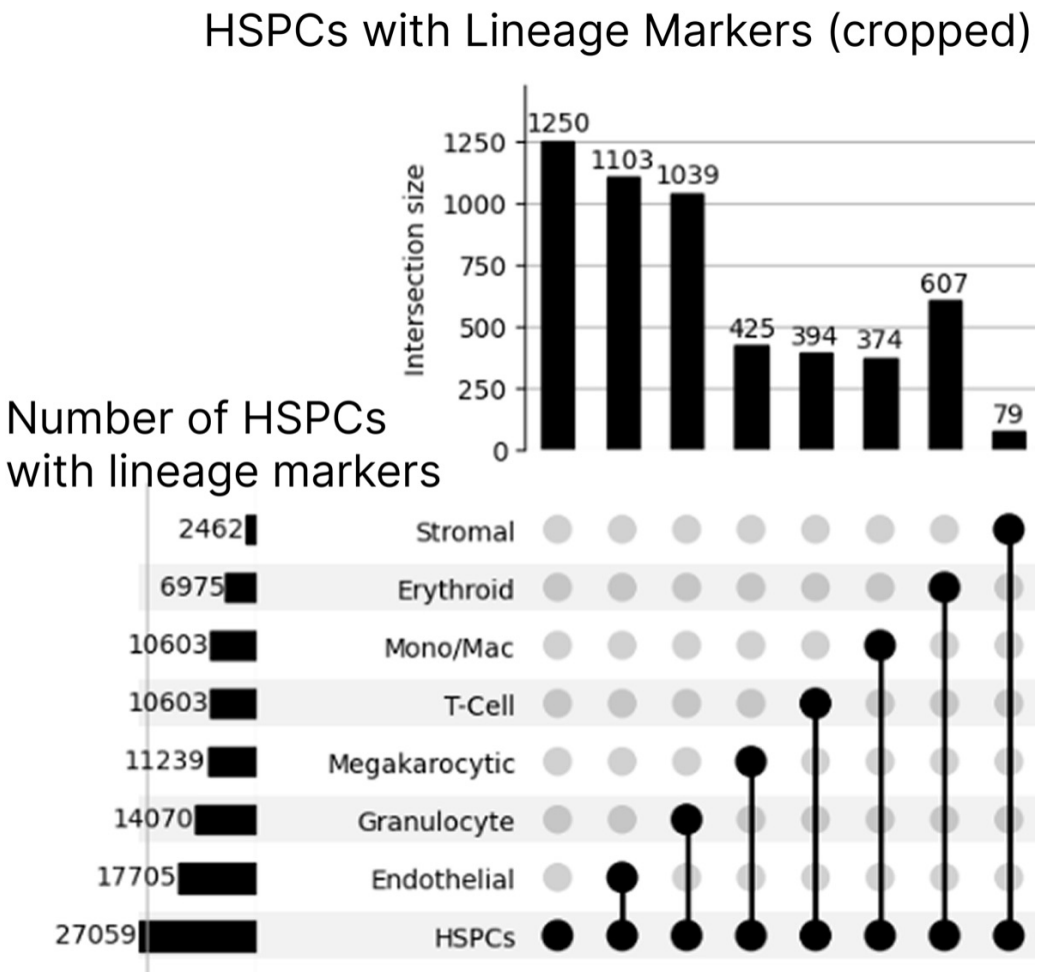


Figure 3.14: I have grouped the above lineage markers into cells types. Then for every HSPC identified I asked whether it also expressed, as well as at least 1 CD34 and KIT, any other lineage markers from each cell lineage group. The above plot are the most likely alternative cell type markers expressed in the HSPCs (dot plots and intersection size only shown for 8 combinations).

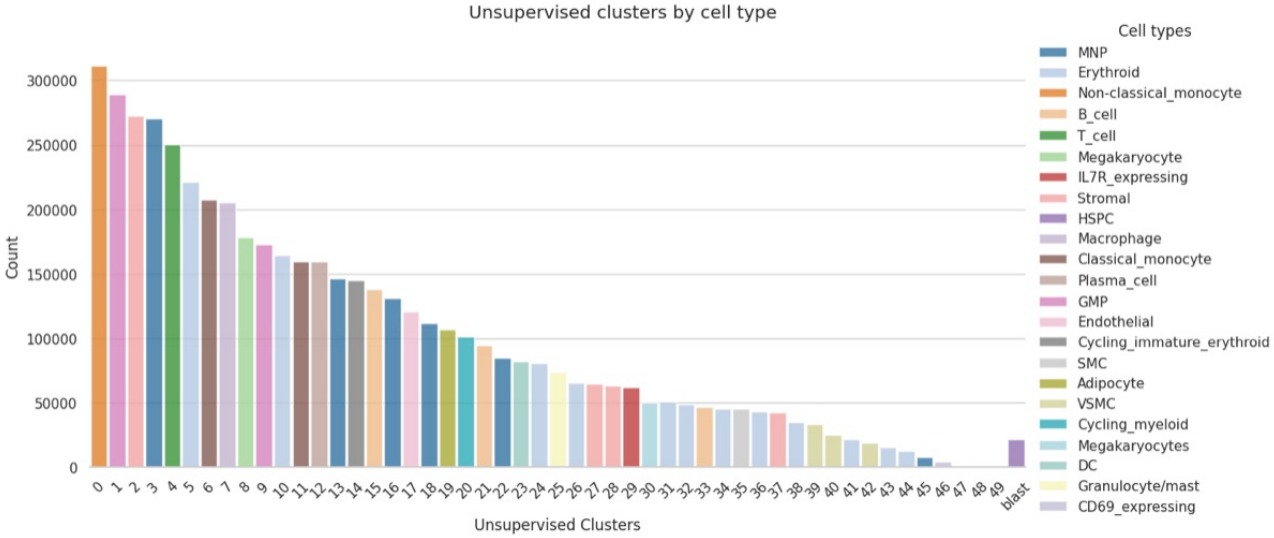
minimum log fold change of 1 and expression in at least 25% of cells within a cluster. Genes were ranked by significance to guide annotation (Table 3.5). This results in the following cell counts and annotations for each cluster across the full dataset (Fig. 3.15a, 3.15b). These populations were justified by the corresponding marker expression for each cell type (Fig. 3.16). Dr R. Cooper also kindly reviewed these annotations.

Supervised Clustering: The reference datasets

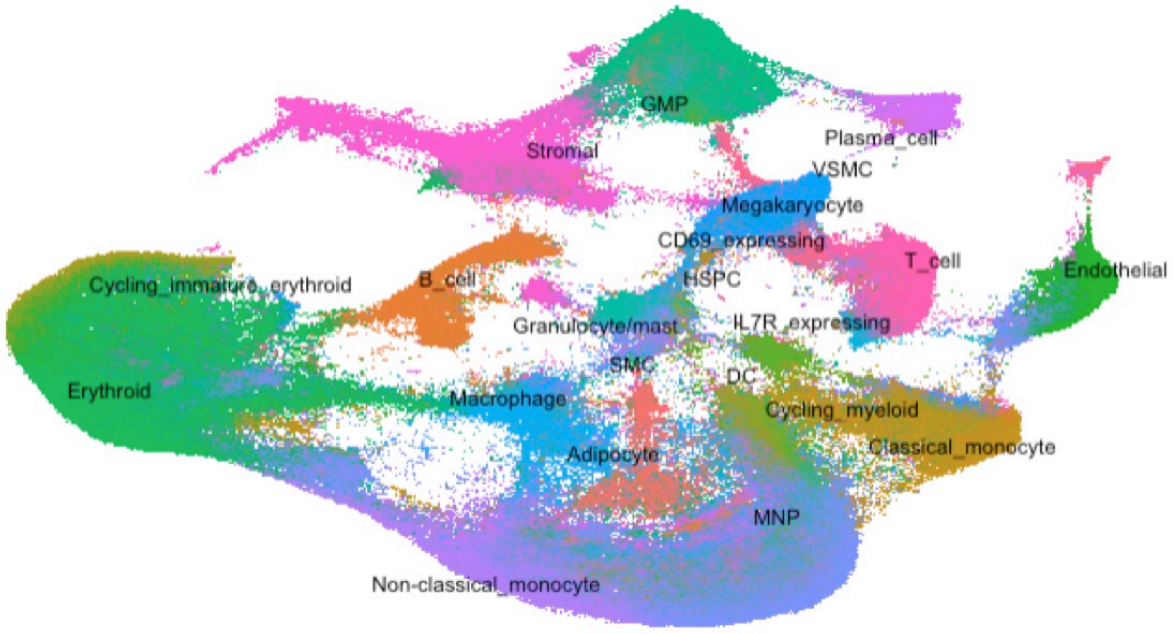
For the supervised annotation, I selected two reference bone marrow datasets. The first was the Azimuth reference, which includes 297,627 bone marrow cells from 39 donors across three studies, all generated using 10x Genomics v2 and integrated into a single dataset. This reference provides strong coverage of the immune mononuclear compartment of the bone marrow [95, 96].

Cluster	New Annotation	Key Markers
0	Non-classical_monocyte	AQP9, STEAP4, CRISPLD2, APOBEC3A, CLEC4E, SELL, FCGR3A
1	GMP	CTSG, NKG7, MNDA, SPI1, BASP1, AIF1
2	Stromal	TNC, PDGFRA, LTBP2, PDGFRB, PCOLCE, FBN1
3	Mono/Mac	S100A12, MCEMP1, FCN1, LILRB2, CSF2RA, APOBEC3A, MNDA
4	T_cell	TRAC, CD247, CD3E, CD8A, CD3D, GZMA, CD2
5	Erythroid	ALAS2, SLC4A1, GYPB, AHSP, GYPA, SNCA
6	Classical_monocyte	CCR2, VCAN, CD14, MS4A6A, FGL2, MPEG1, FCGR1A, GLIPR1, IRF8
7	Macrophage	CD163, PLA2G7, VSIG4, CD5L, KCNMA1, MS4A4A, CD68, CD4
8	Megakaryocyte	MMRN1, CAVIN2, MYLK, BANK1, PECAM1, ACTG2, MDM2
9	GMP	CTSG, SMIM24, NKG7, MYC, ERG, ANPEP, AIF1
10	Erythroid	CD5L, GYPB, GYPA, HEMGN, VSIG4, AHSP, CD163
11	Classical_monocyte	VCAN, FGL2, MPEG1, MS4A6A, FCN1, C1orf162, SELL
12	Plasma_cell	MZB1, TNFRSF17, SLAMF7, PRDM1, FKBP11, TENT5C, CYTIP
13	MNP	S100A12, FCN1, MCEMP1, MNDA, ALAS2, BASP1, TENT5C
14	Cycling_immature_erythroid	GYPB, CENPF, CCNB2, GYPA, CDK1, TOP2A, HEMGN
15	B_cell	TCL1A, CD19, MZB1, CXCR4, PLCG2, MDM2
16	MNP	DNASE1L3, ADGRL4, S100A12, MNDA, BASP1, ALAS2
17	Endothelial	BTNL9, NPDC1, TCIM, CLEC14A, MMRN2, CD34, ECSCR
18	MNP	FCER1A, FCN1, LILRB2, C1orf162, FGL2, SPI1, SELL
19	Adipocyte	PLIN4, APOBEC3A, CSF2RA, LILRB2, SELL, MCEMP1, MNDA, S100A12
20	Cycling_myeloid	TOP2A, UBE2C, MKI67, RETN, CDK1, CCNB2, NKG7, ERG
21	B_cell	MS4A1, CD79A, CD19, SPIB, BANK1, TCL1A, ADAM28
22	MNP	S100A12, APOBEC3A, MCEMP1, FCN1, CCL5, CSF2RA, LILRB2, PTPRC
23	DC	LILRA4, IRF8, MPEG1, CSF2RA, S100A12, SELL
24	Erythroid	GYPB, GYPA, AHSP, TNC, HEMGN, ALAS2, SLC4A1
25	Granulocyte/mast	MS4A2, CPA3, GATA2, FCER1A, CTSG, AIF1
26	Erythroid	MMRN1, CAVIN2, GYPB, SNCA, GYPA, FHL2, HEMGN
27	Stromal	DPT, COL5A2, FCN1, S100A12, MNDA
28	Stromal	SFRP4, PDGFRA, TNC, LTBP2, FBN1, CTSG
29	IL7R_expressing	IL7R, SLC4A1, HEMGN, ALAS2, AHSP, S100A12
30	Megakaryocytes	MMRN1, CAVIN2, MYLK, BANK1, CXCL2, GNG11, VWF
31	Erythroid	ALAS2, SLC4A1, GYPB, AHSP, GYPA, HEMGN, SNCA
32	Erythroid	PLIN4, TIMP4, LPL, GYPB, GYPA, HEMGN, SLC4A1
33	B_cell	MS4A1, CD79A, APOBEC3A, MCEMP1, S100A12, FCN1, LILRB2
34	Erythroid	ALAS2, SLC4A1, GYPB, HEMGN, GYPA, AHSP, SNCA
35	SMC	ACTA2, GYPA, ALAS2, HEMGN, AHSP, SLC4A1, MKI67
36	Erythroid	MS4A1, GYPB, GYPA, SLC4A1, HEMGN, ALAS2, AHSP
37	Stromal	PDGFRB, THY1, STEAP4, FBN1, CAVIN1
38	Erythroid	FCER1A, GYPB, GYPA, HEMGN, SLC4A1, ALAS2, AHSP
39	VSMC	CCL5, CAVIN2, GNG11, ACTG2
40	VSMC	MYH11, ACTA2, CAVIN1, PDGFRB, MYLK, PECAM1
41	Erythroid	TCL1A, GYPB, HEMGN, GYPA, AHSP, ALAS2, SLC4A1, SNCA
42	VSMC	CAVIN2, MMRN1, CTSG, CCL5, MYLK, ACTG2, NKG7
43	Erythroid	DPT, COL5A2, GYPB, GYPA, SLC4A1, HEMGN, ALAS2
44	Erythroid	MZB1, GYPB, GYPA, HEMGN, TENT5C, ALAS2, SLC4A1, AHSP
45	MNP	RTKN2, CD1A, OPRPN, CXCL6, COL5A2, CYP2B6, DPT, SNTN
46	CD69_expressing	CD69
47	SMC	MYH11, ACTA2
48	T_cell	KLRB1, CCL5
49	Granulocyte/mast	CPA3, MS4A2, TNC
blast	HSPC	KIT, CD34

Table 3.5: For each cluster in the Louvain clustering, I performed a Wilcoxon test to identify the most likely cell type using differentially expressed genes. Key genes were selected using thresholds: avglog2FC > 1, expression in more than or equal to 25% of cells in that cluster, and significance $p < 0.05$.



(a) Total number of cells per cluster across all samples, labelled by most likely cell type.



(b) UMAP of downsampled dataset (20,000 cells per sample) showing clusters merged by cell type.

Figure 3.15: Summary of unsupervised Louvain clustering. (A) Cell counts per cluster with inferred cell type. (B) UMAP visualization of cell type annotations across clusters.

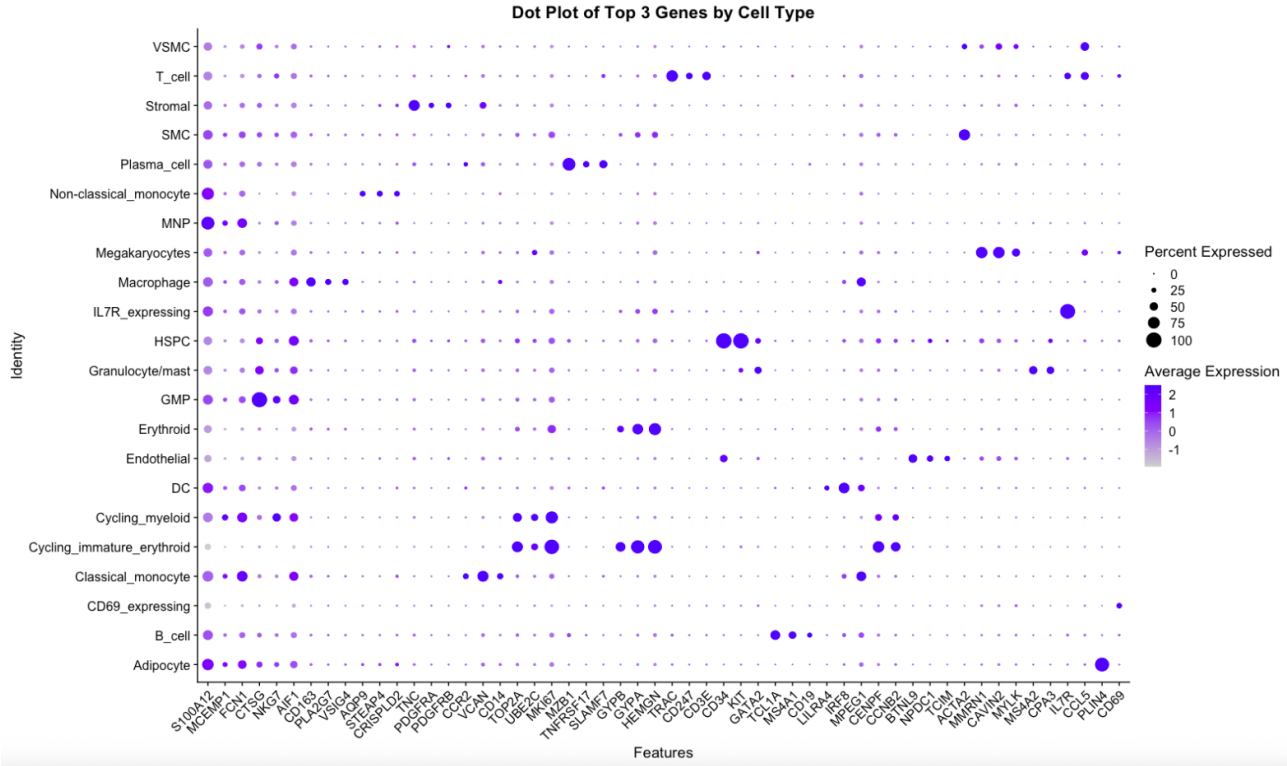


Figure 3.16: Top 3 differentially expressed genes per cluster (avglog2FC > 1, >25% of cells expressing) in the downsampled object (20 000 cells). Dot size indicates % of cells expressing the gene; color indicates average expression.

However, it lacks robust representation of granulocytes and stromal cells, which are more difficult to capture in scRNA-seq due to their fragility, short in vitro lifespan, and susceptibility to RNA degradation. Stromal cells also require extensive tissue dissociation, which can compromise RNA quality. I refer to this dataset as the Azimuth dataset. The second reference is from a recent study by Tan Lab, which specifically enriched for stromal populations. This was achieved through enzymatic digestion followed by sorting of CD34-positive and CD34-negative fractions before pooling. Across 12 samples, they captured 82,742 single-cell transcriptomes representing a wide range of bone marrow populations, including HSCs, progenitors (HSPC, MPP, GMP, MEP), megakaryocyte/erythroid precursors, lymphoid progenitors (CLP), immune cells (e.g., pDC, basophils, eosinophils), stromal cells (MSC), and endothelial subsets (AEC, SEC, VSMC) [34]. I refer to this dataset as the Tan Lab dataset.

Both reference datasets have broader gene panels than our own. The Tan Lab dataset contains 28,659 genes, 365 of which are shared with our Xenium panel (13 are missing). The Azimuth dataset includes 69,349 genes, with 375 shared (2 missing: TCIM and TENT5C). Using scRNA-seq references for annotation provides two main advantages: first, it ensures alignment with

established literature; second, it helps address the limitation that our reduced gene panel may not fully resolve cellular lineages.

Supervised Cell Annotation: Methods

There are several reference mapping methods available, most of which do not incorporate spatial information. Examples include Transfer Anchors, DotR, SingleR, and CellTypist. Of these, I found SingleR to perform best and used it for cell type prediction [97]. As a form of spatial validation, I also applied DotR, which incorporates spatial context [98].

Prior to mapping, reference datasets were subsetted to include only genes shared with the Xenium panel. SingleR then selects marker genes for each reference cell type using the formula 500*(2/3)**log(n). n is the number of reference genes. In our case, this resulted in 176 markers deemed most informative. SingleR uses Spearman rank correlation to compare the expression profile of each query cell against reference profiles, focusing on gene ranking rather than absolute expression. For each label, the 0.8 quantile of these correlations is used as the final score, reducing the influence of outliers and accommodating variation within heterogeneous cell types.

Using the Tan Lab dataset as a reference, SingleR assigned cell types to our data. However, as shown in Figures 3.17, 3.18a, 3.18b, predictions were generally low-confidence, with small differences between the top-scoring and next-best labels (delta.next). This likely reflects the limited number of available marker genes in the Xenium panel. In the end, the data is split up into these clusters below (Fig. 3.19). Using the canonical lineage genes we defined earlier we check the expression of our clusters defined by singleR. I repeated this process with the Azimuth dataset, although I have not included figures for conciseness.

We can now see the overlap between the two datasets. Across the broad cell type categories, several show perfect overlap between the annotations from the Azimuth reference and the Tan Lab dataset when applied to our data using singleR (Fig. 3.20). When compared to our unsupervised annotations, we observe both alignment and divergence. In the Fig. 3.21, each unsupervised cluster is compared to the top five predicted cell types from supervised annotation using SingleR with the Tan Lab (singleR_cell) and Azimuth (singleR_azi) references. Some clusters, such as 6 and 24, show strong agreement across methods, while others display more

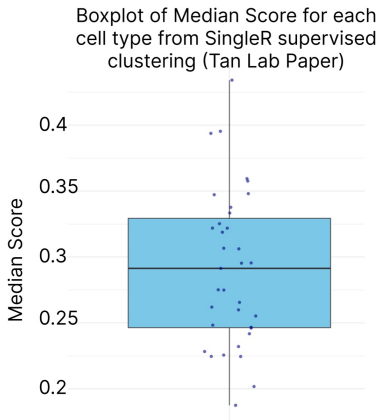


Figure 3.17: Median score of each cell type predicted from SingleR using the Tan Lab paper as reference. The median score of all cell types was less than 0.3

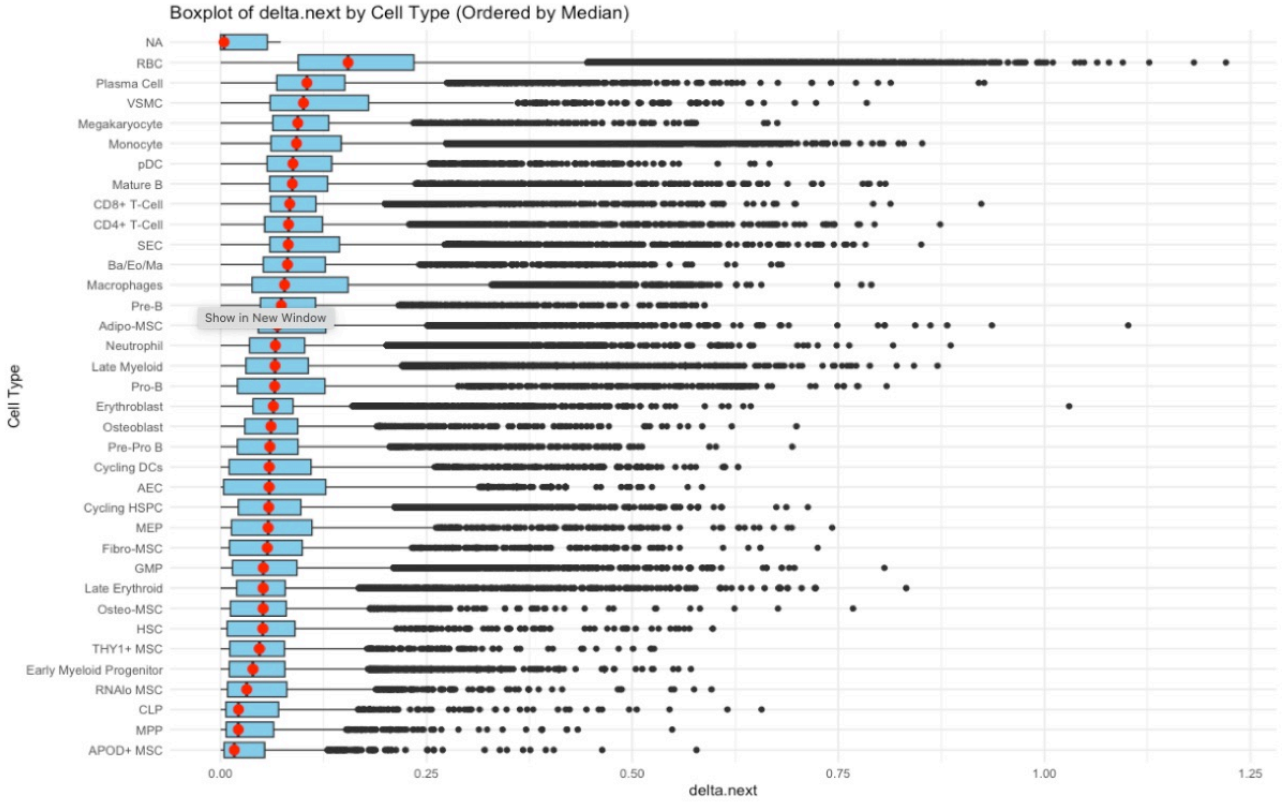
heterogeneous assignments in the supervised predictions e.g. 13 and 16. I have included four examples for visualisation (Fig. 3.21).

To assess whether differences in annotation were method-dependent, I also applied dotR, an algorithm designed for spatial cell type annotation, using the same reference datasets (Fig. 3.22). Given differences in marker sets and the limitations of the reference datasets, some discrepancies between supervised and unsupervised annotations were expected. However, clusters 16 and 13 stood out due to a clear division between erythroid and monocyte or macrophage populations, lineages for which we have adequate marker coverage. This raised concerns about potential misclassification. As a result, I investigated whether segmentation might be influencing cell lineage assignments.

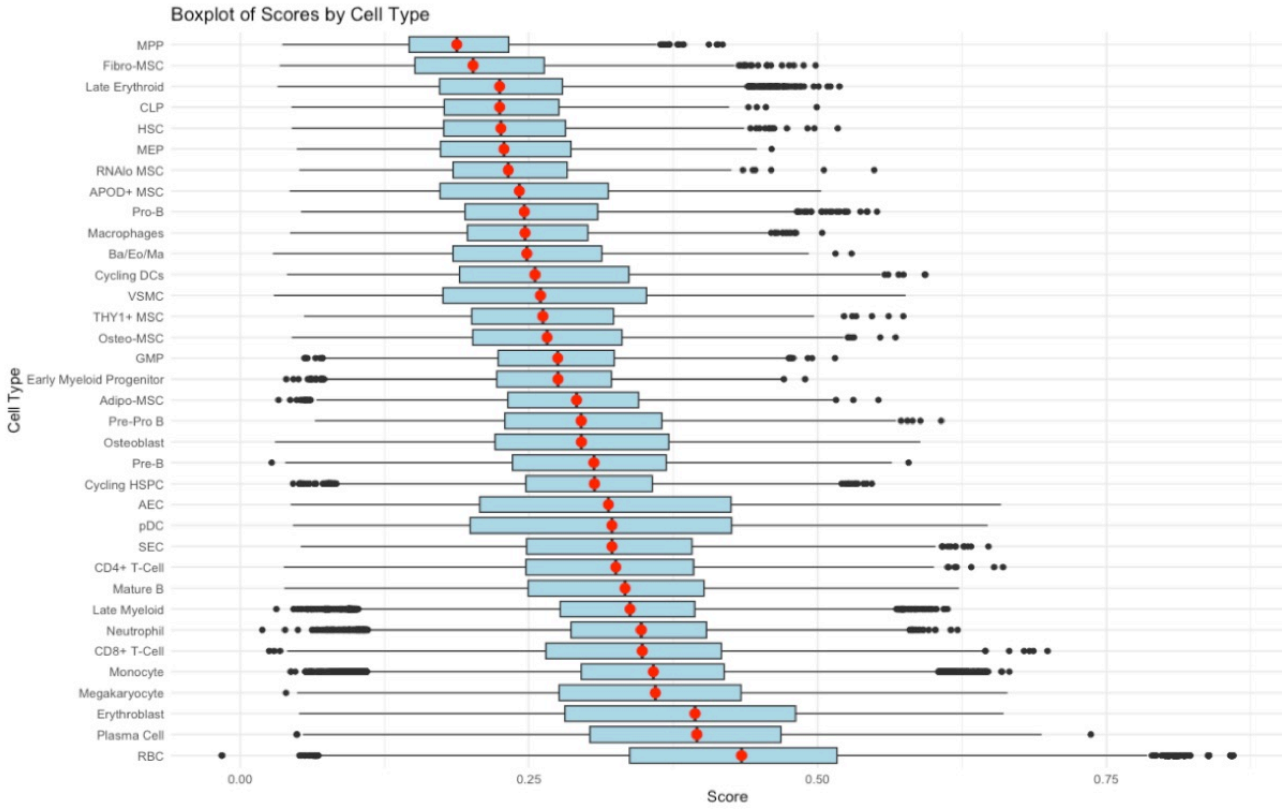
3.4.7 Segmentation and Cell Annotation

Gene Transcripts location within cells

I was concerned that discrepancies between the unsupervised and supervised clustering results could be due to segmentation errors, specifically overlapping cells that might result in marker expression being misattributed. Since our analysis focuses on single-cell interactions, accurate segmentation is critical. To investigate this, I assessed the spatial localization of key lineage markers within each cell, comparing the proportion of marker expression within the nucleus versus the cytoplasm. If a marker is primarily nuclear, segmentation errors are less likely to impact lineage assignment. Conversely, if expression is mostly cytoplasmic, mis-segmentation may lead to incorrect annotations (Fig.3.23).



(a) Delta score (difference to next-best match) per cell type using SingleR and the Tan Lab reference. Lower deltas indicate lower confidence.



(b) Prediction scores for each assigned cell type using SingleR and the Tan Lab reference, ordered by median confidence.

Figure 3.18: Confidence of cell type assignments using SingleR with the Tan Lab dataset. (A) Delta scores reflect confidence in the top prediction. (B) Prediction scores indicate overall match strength per cell type.

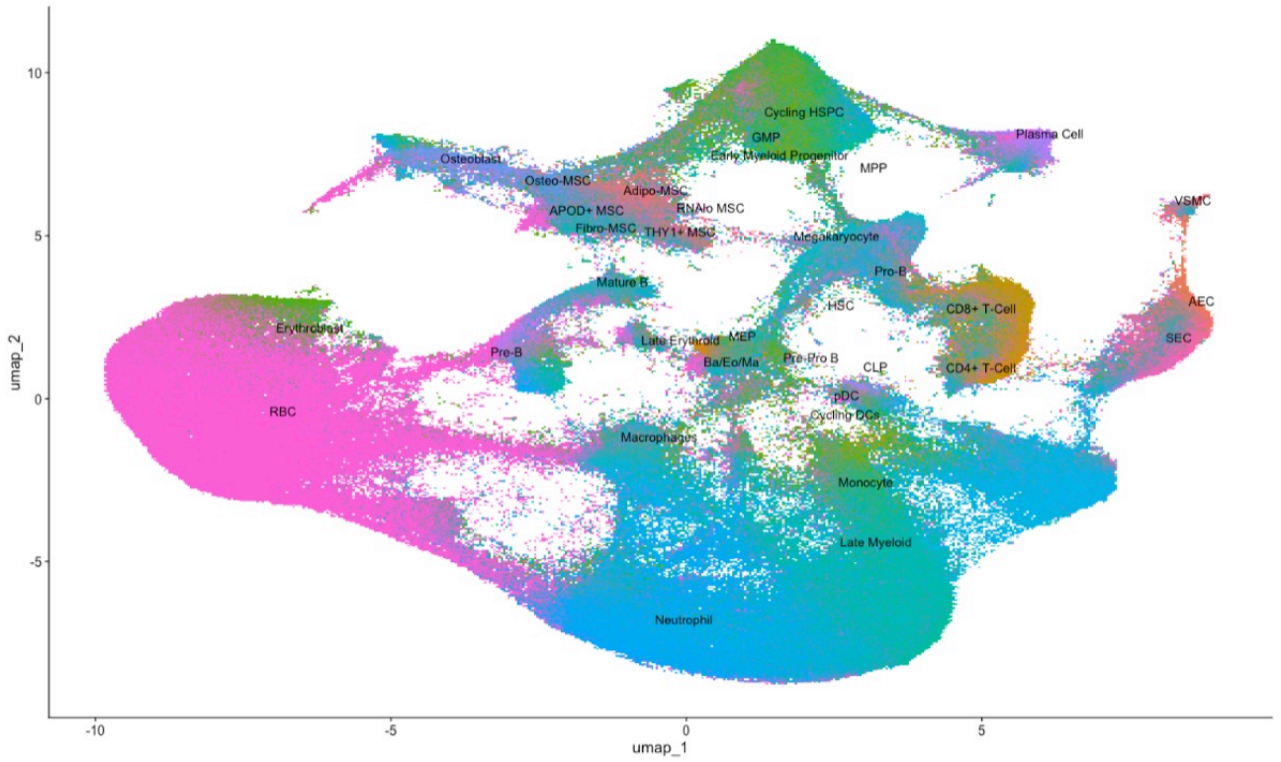


Figure 3.19: Final clusters for 15um voronoi segmentation, visualised in UMAP space, as designated by SingleR using the Tan Lab paper as reference

For example, erythroid markers such as GYPA and GYPB showed predominantly nuclear expression, increasing our confidence in accurate erythroid identification. In contrast, markers like CD4 and CD68 were largely cytoplasmic, suggesting lineage assignment for these populations may be more susceptible to segmentation errors. To evaluate the impact of segmentation on annotation, I re-segmented the data at two additional resolutions: (1) nuclear boundaries only, and (2) Voronoi expansion up to 5 um or until reaching a neighboring cell. I then re-applied SingleR using both the Azimuth and Tan Lab references, comparing the resulting annotations to those from the original 15 um Voronoi segmentation (Figs. 3.24). For polylobated megakaryocytes, the nuclear segmentation was defined as the concave hull encompassing all associated nuclear boundaries.

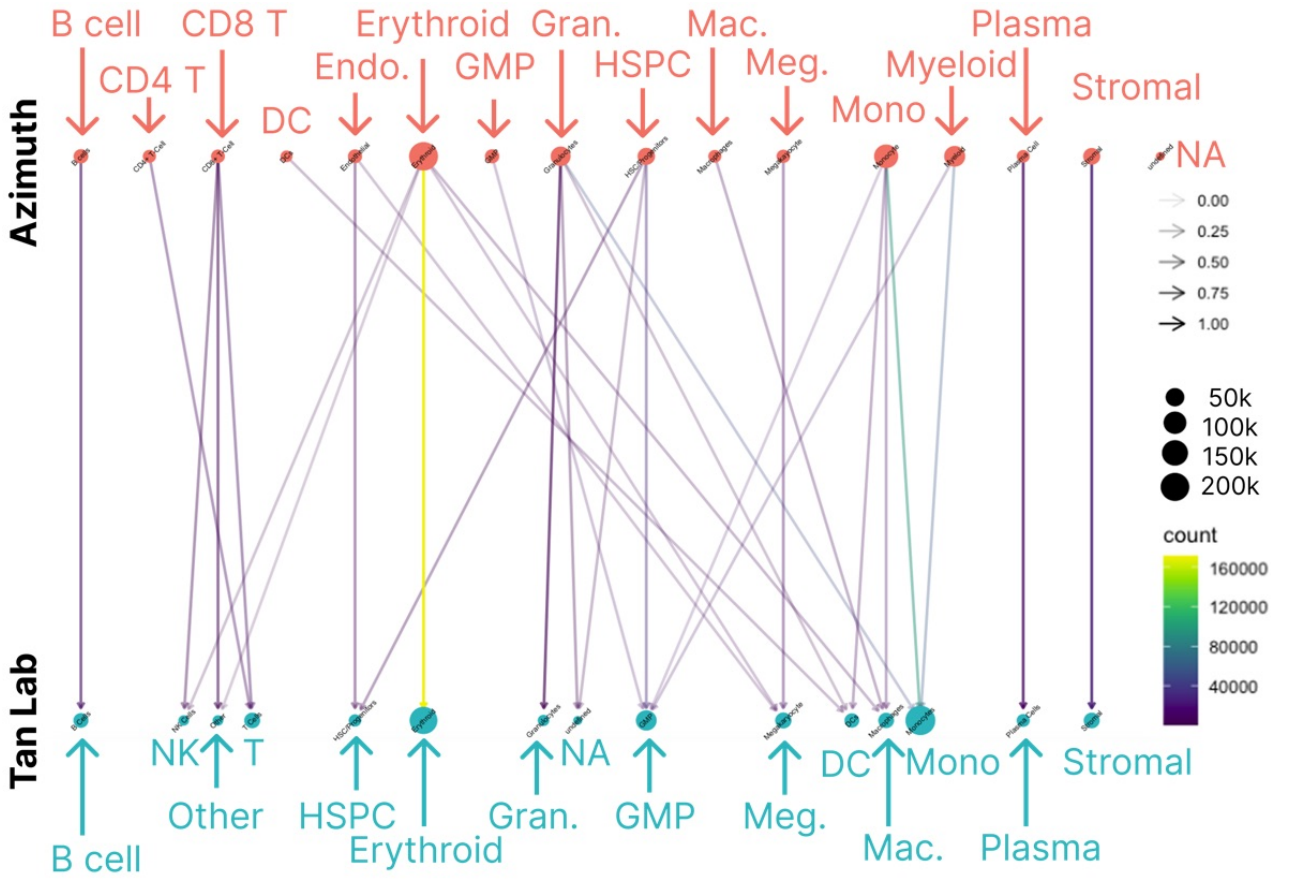


Figure 3.20: Clustree of the overlap between cell groups from SingleR on azimuth (top) and singleR on Tan Lab paper as a reference (bottom). I have merged smaller groups in both sets into larger macro groups for better comparison

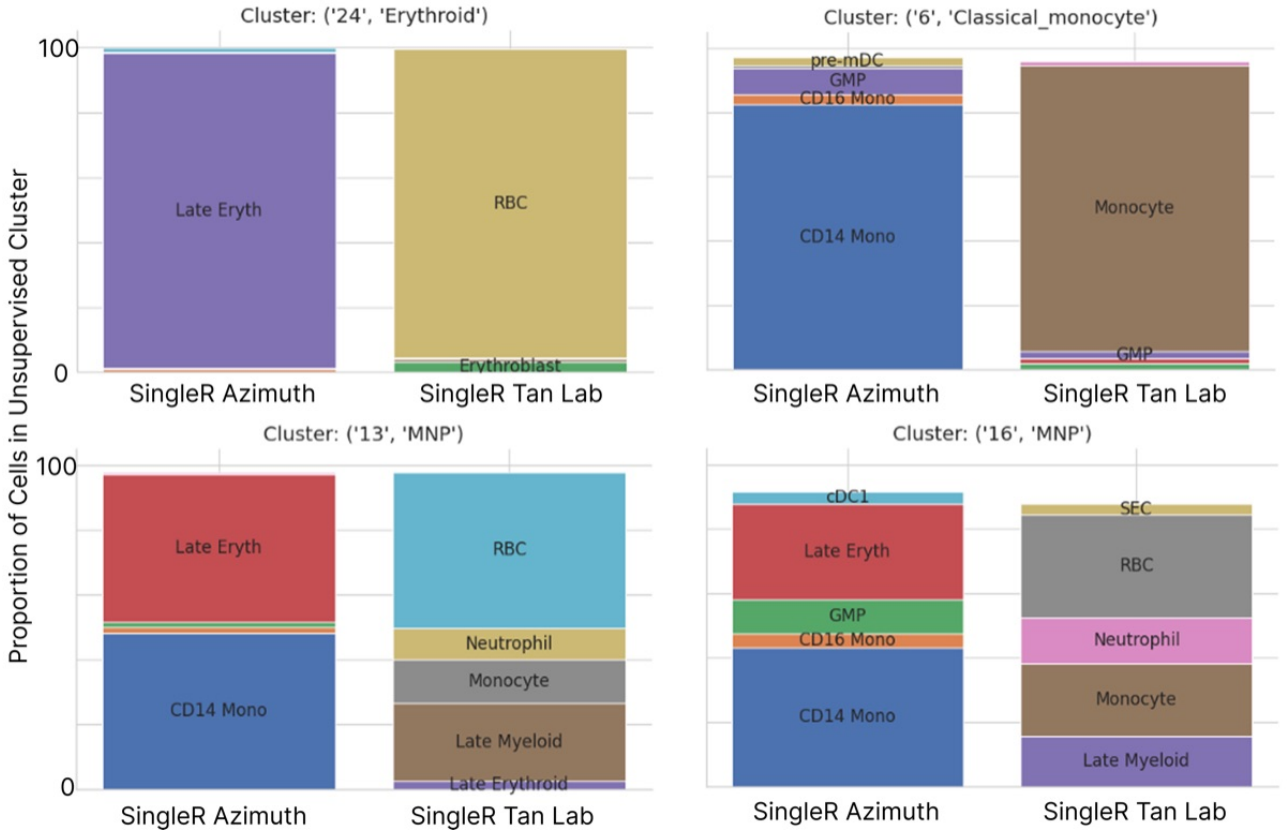


Figure 3.21: Each of the 50 unsupervised clusters has been given a cell annotation for the cells in this group, based on differential gene expression. The cell annotation from unsupervised clustering is the title of each stacked barplot. The stacked barplots represent the proportion of cells within that cluster assigned to a cell anotation by singleR using either the Tan Lab paper reference (right, singleR_cell) or the azimuth reference (left, singleR_azi). I've only shown the top 5 most prevalent clusters from each singleR annotation, which accounts for the proportions not always covering all the cells in the unsupervised cluster.

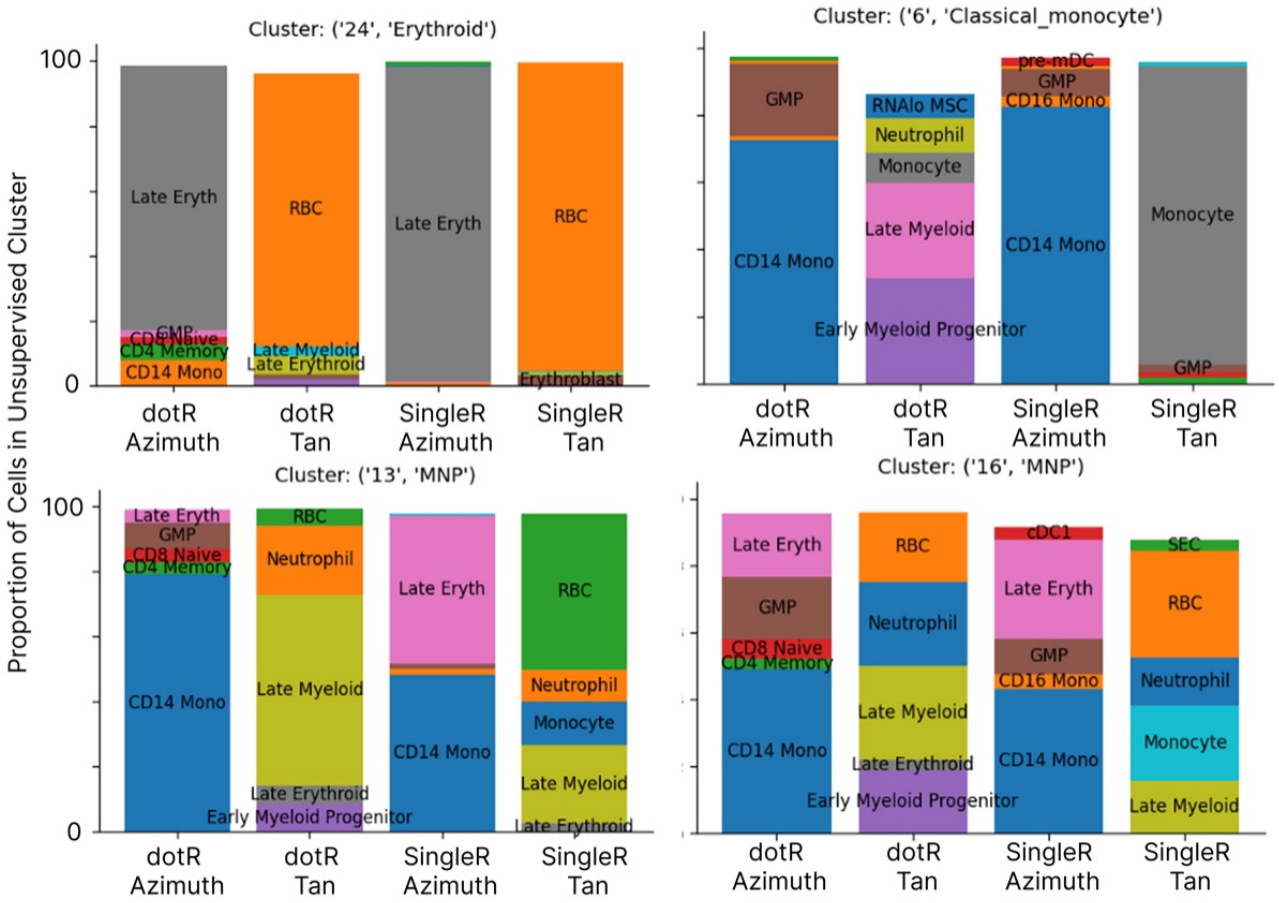


Figure 3.22: Each of the 50 unsupervied clusters has been given a cell annotation for the cells in this group, based on differential gene expression. The cell annotation from unsupervised clustering is the title of each stacked barplot. The stacked barplots represent the proportion of cells within that cluster assigned to a cell anotation by singleR, right, or dotR, left, using either the Tan Lab paper reference, right, or the azimuth reference, left. I've only shown the top 5 most prevalent clusters from each singleR annotation, which accounts for the proportions not always covering all the cells in the unsupervised cluster.

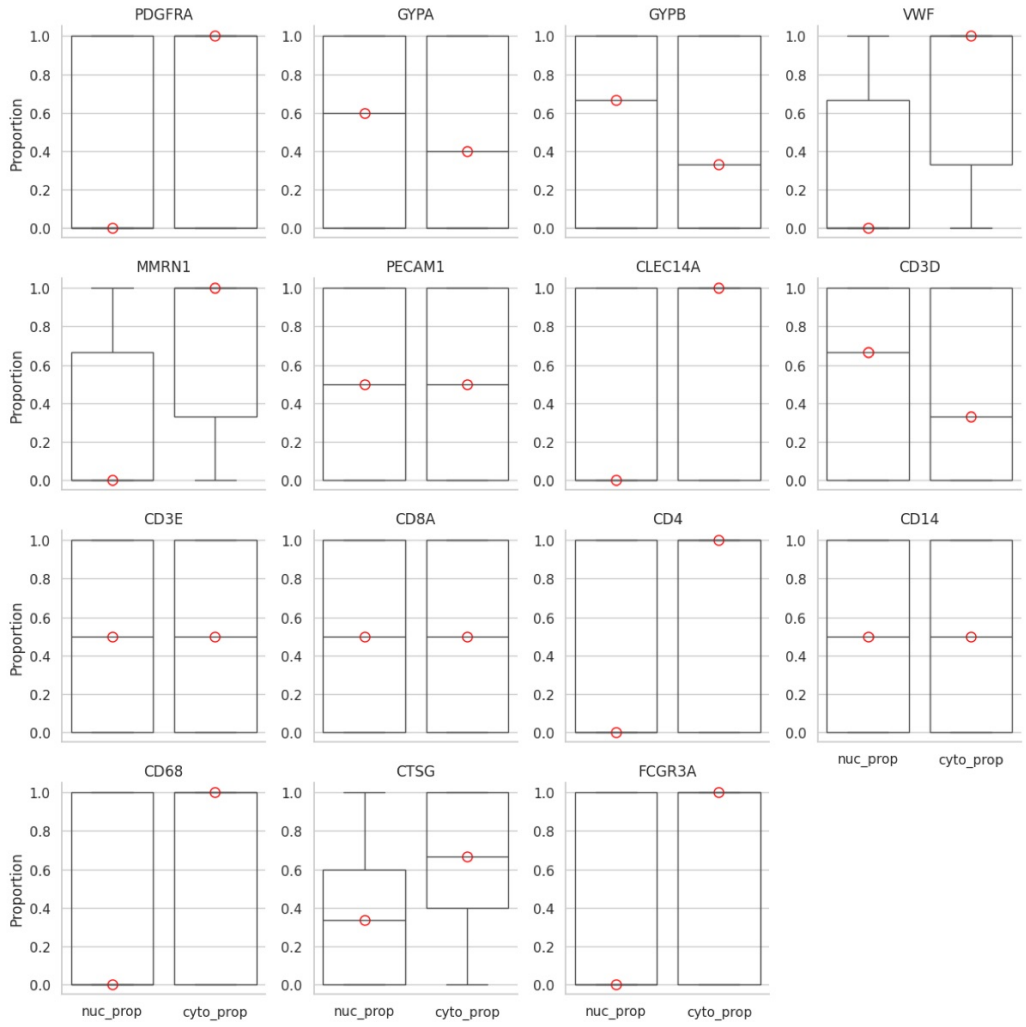


Figure 3.23: For each cell I calculated the number of each lineage marker in the figure. For cells with at least one count of that marker, I then calculated the proportion of that lineage marker that was expressed in the cytoplasm compared to the nucleus. The boxplots are made up of all cells that contained that cell marker within it's boundary at 15um voronoi. The median is represented as a red dot.

Comparing resegmentations: 5um voronoi and nuclear

To enable comparison, I have renamed categories across level 2 cell annotations of the Azimuth and the Tan Lab papers, as seen below to match my unsupervised cluster names (Table 3.6 3.7).

Tan Lab paper	My annotations
AEC	Endothelial
APOD+MSC	Stromal
Adipo-MSC	Adipocyte
Ba+Eo+Ma	Granylocyte/mast
CD4+T-Cell	T_cell
CD8+T-Cell	T_cell
CLP	HSPC
Cycling-DCs	DC
Cycling-HSPC	HSPC
Early-Myeloid-Progenitor	GMP
Erythroblast	Cycling_immature_erythroid
Fibro-MSC	Stromal
GMP	GMP
HSC	HPC
Late-Erythroid	Erythroid
Late-Myeloid	MNP
MEP	Cycling_myeloid
MPP	Cycling_myeloid
Macrophages	Macrophage
Mature-B	B_cell
Megakaryocyte	Megakaryocyte
Monocyte	Classical_monocyte
Neutrophil	Granylocyte/mast
Osteo-MSC	Stromal
Osteoblast	Stromal
Plasma-Cell	Plasma_cell
Pre-B	B_cell
Pre-Pro-B	B_cell
Pro-B	B_cell
RBC	Erythroid
RNAlo-MSC	Stromal
SEC	Endothelial
THY1+MSC	Stromal
VSMC	VSMC
pDC	DC

Table 3.6: Tan Lab paper annotations and corresponding classifications.

Azimuth	My annotations
ASDC	DC
BaEoMa	Granylocyte/mast
CD14-Mono	Classical_monocyte
CD16-Mono	Non-classical_monocyte
CD4-Effector	T_cell
CD4-Memory	T_cell
CD4-Naive	T_cell
CD8-Effector_1	T_cell
CD8-Effector_2	T_cell
CD8-Effector_3	T_cell
CD8-Memory	T_cell
CD8-Naive	T_cell
CLP	HSPC
EMP	Cycling_immature_erythroid
Early-Eryth	Cycling_immature_erythroid
GMP	GMP
HSC	HSPC
ILC	T_cell
LMPP	Cycling_myeloid
Late-Eryth	Erythroid
MAIT	T_cell
Macrophage	Macrophage
Memory-B	B_cell
NK	T_cell
NK-CD56-	T_cell
NK-Proliferating	T_cell
Naive-B	B_cell
Plasma	Plasma_cell
Platelet	Megakaryocyte
Prog-Mk	Megakaryocyte
Stromal	Stromal
T-Proliferating	T_cell
cDC1	DC
cDC2	DC
pDC	DC
pre-B	B_cell
pre-mDC	DC
pre-pDC	DC
pro-B	B_cell
transitional-B	B_cell

Table 3.7: Azimuth annotations and corresponding classifications.

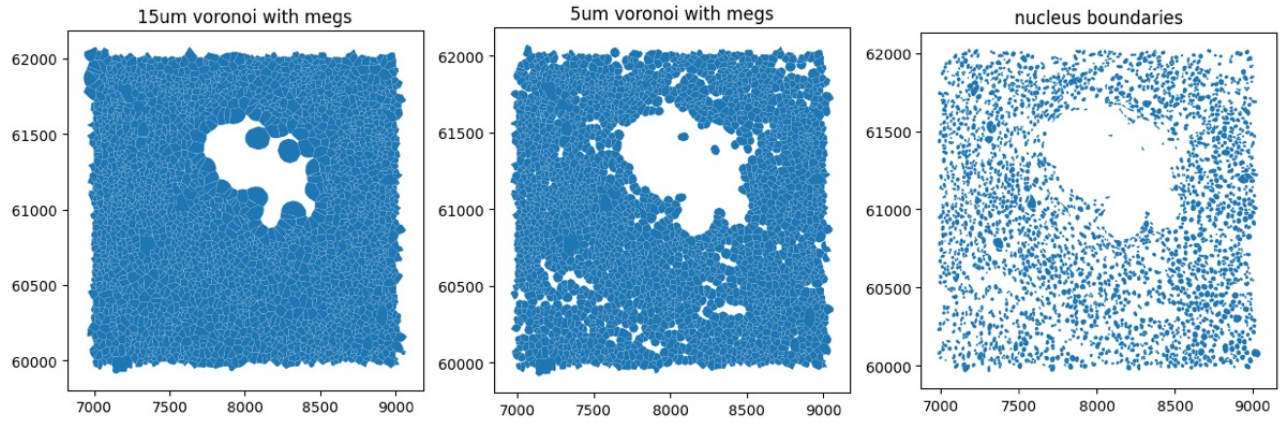


Figure 3.24: Examplar of the resegmentation of the same cells at different levels of segmentation. Left: 15um voronoi expansion, Middle: 5um voronoi expansion, Right: nuclear segmentation

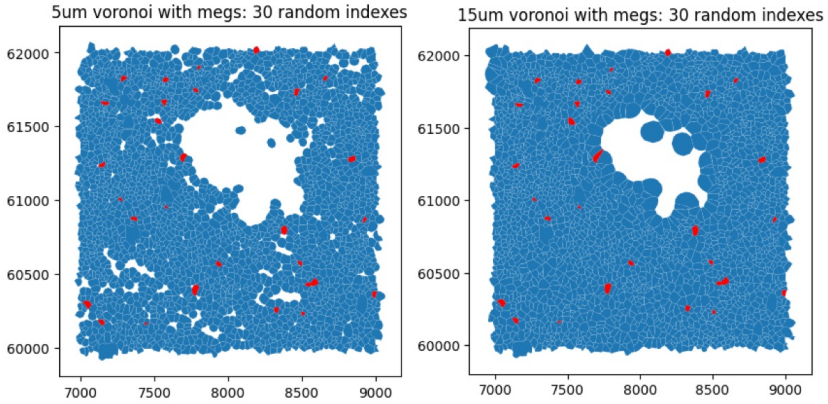


Figure 3.25: As I compared the annotation of each level of segmentation, I wanted to include a schematic showing if I choose 30 random indexes I will get the same cells in each exemplar (red cells). This is to show that the indexes are matched across segmentations to facilitate comparisons. I have not compared segmentations that are not matched.

I next subsetted the dataset to include only cells present in all of the 15um Voronoi, 5um Voronoi, and nuclear segmentations, yielding 5,103,091 cells. Minor differences in cell inclusion arose due to changes in segmentation around megakaryocytes (Fig. 3.25).

I applied SingleR-based annotation using both the Tan Lab reference and Azimuth, separately for the 5 um Voronoi and nuclear segmentations. This produced two annotations per cell for each segmentation, four annotations in total. Using the unsupervised annotations of the 15um segmentations and the SingleR annotations from the Tan lab and Azimuth references for the 5um segmentations and nucleus boundaries, I kept or reassign annotations as detailed below and in Fig. 3.26.

After standardising the SingleR labels, 25.7% of all cells (1,312,860) had consistent annotations across all 5 methods; the 4 annotations for nuclear and 5 um segmentations, from the Tan Lab

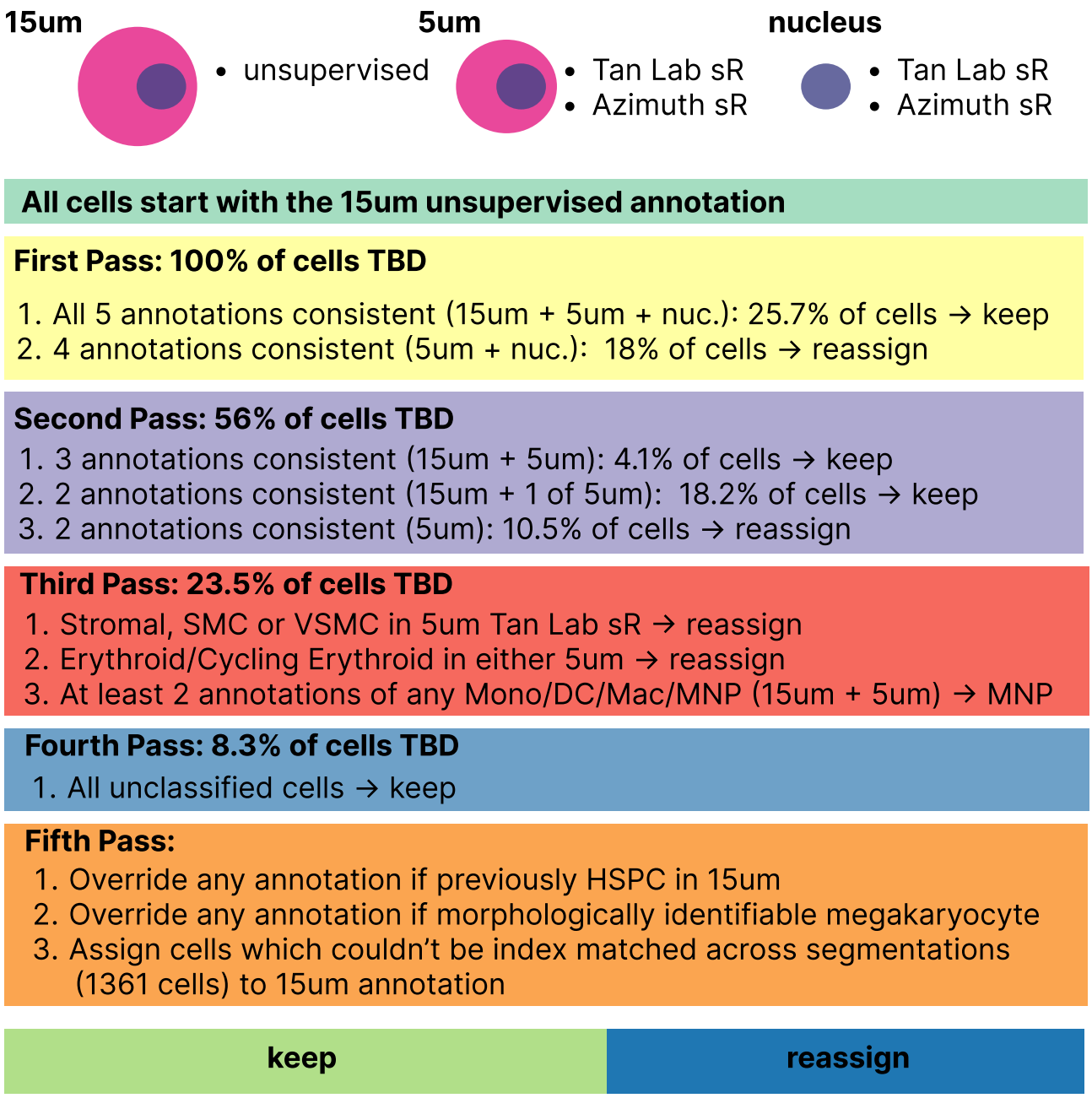


Figure 3.26: Summary of re-annotation using the concordance between unsupervised annotations of the 15um segmentations and singleR annotations of the 5um and nucleus segmenations. Approximately 56% of cells had the unsupervied annotations kept.

and Azimuth references, as well as the unsupervised annotations for the 15 um Voronoi.

The majority of these consistently annotated cells were erythroid, which aligns with expectations, as erythroid lineage markers are more frequently nuclear-localised (Fig. 3.27).

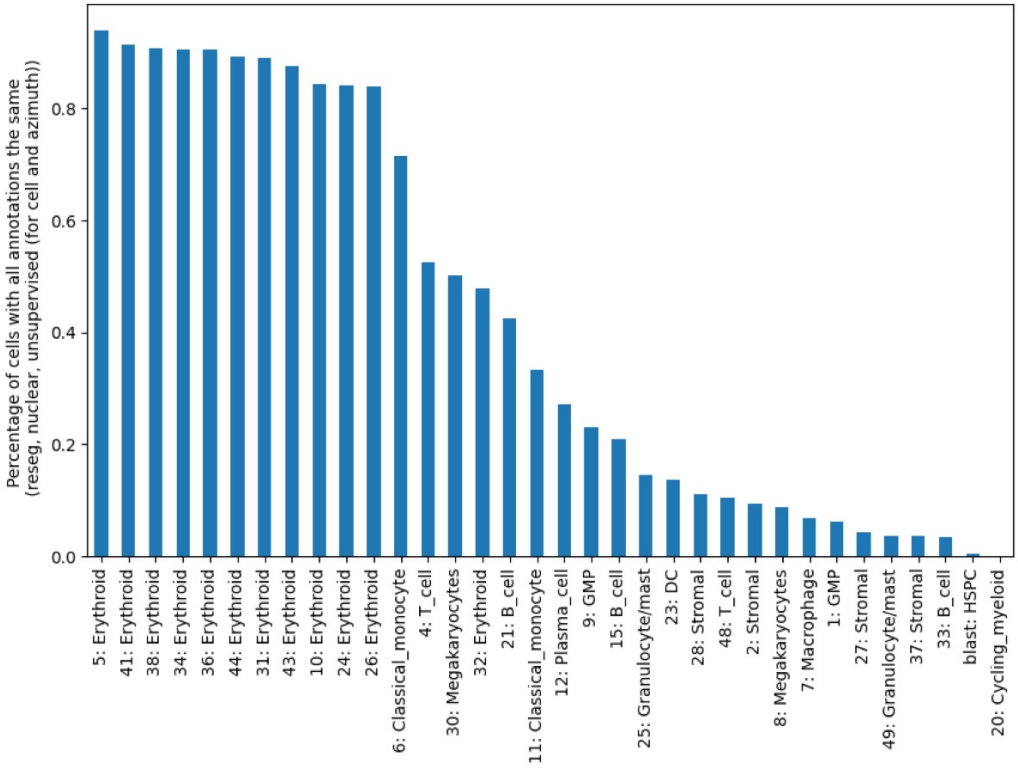


Figure 3.27: For each unsupervised cluster the percentage of that cluster that has the same annotation from singleR (with Tan Lab paper and azimuth as references) in the 5um voronoi, nuclear segmentation and the same annotation in my unsupervised clustering. These are the cells that I am most confident in my lineage assessment.

Next, I examined the remaining cells whose annotations were consistent across all 4 SingleR references (Tan Lab and Azimuth for both 5 um Voronoi and nuclear segmentations) but differed from the unsupervised annotation at 15 um. I hypothesised that the original 15 um segmentation was too large for these cells, leading to inaccurate lineage assignments by capturing extraneous gene expression.

These cells accounted for 18% of the dataset. Given the consistency across all SingleR annotations, for 5um and nuclear boundaries, I reclassified them according to the SingleR-predicted identity. Most of these were reassigned as erythroid cells or classical monocytes (Fig. 3.28).

The proportion of cells with undetermined lineage in each original cluster is shown in blue (TBD). These include cells with inconsistent annotations across segmentations; across the four SingleR annotations for the 5 um and nuclear segmentations. This group accounts for 56.3%

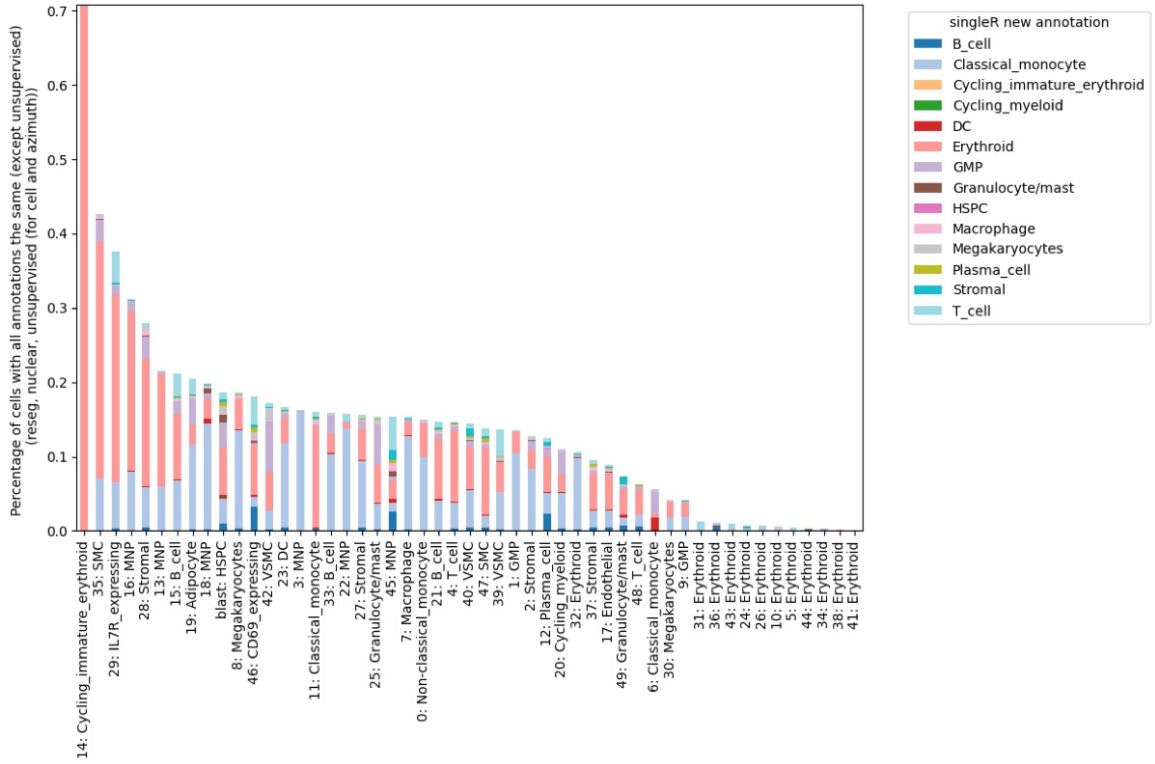


Figure 3.28: For each cluster, the percentage of cells where all annotations from singleR in 5um and nuclear where the same, but they are not the same as the original unsupervised cluster. The colours within each bar represent the cell type that these cells are now reassigned to. The majority are reassigned to erythroid (pink) or classical monocytes (baby blue)

of all cells $100 - (25.7 + 18)$, meaning that just over half of the dataset still requires lineage assignment (Fig. 3.29).

As not all lineage markers are expressed in the nucleus, I used only the 5 um Voronoi SingleR annotations to guide reannotation of the TBD population. In 7.3% of TBD cells, the annotation was consistent across the unsupervised and SingleR methods for both 5 um segmentations, 3 annotations, so I retained the original label. Among the remaining 92.7%, 35% had unsupervised annotations supported by at least one SingleR reference (Tan Lab or Azimuth) in the 5 um segmentation; these annotations were also retained. For the rest, I identified cells where Tan Lab and Azimuth gave the same SingleR annotation in the 5 um segmentation. This applied to 31% of the remaining cells, which I then reclassified accordingly.

After these steps, 1,283,765 cells (23.5% of the total) remain without confident lineage assignment. For these remaining cells, I assign labels according to the following criteria:

1. Stromal, VSMC, or SMC: If a cell was identified as stromal, vascular smooth muscle cell (VSMC), or smooth muscle cell (SMC) in the 5 um resegmentation and Tan Lab supervised annotation, I retained this label. The Tan Lab dataset specifically enriches

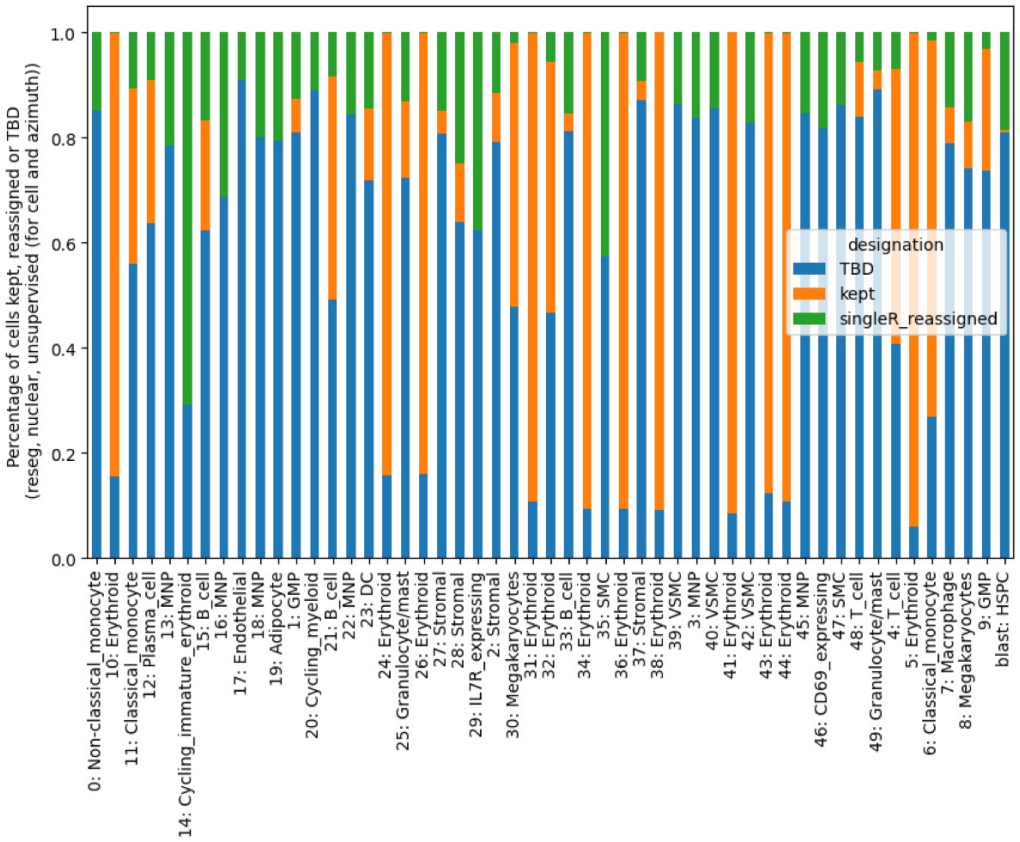


Figure 3.29: For each unsupervised cluster, I am showing the proportion of cells within that cluster where the original unsupervised annotation was 1) kept (orange) because the singleR annotations as 5um and nucear where in concordance with it, 2) reassigned (green) because singleR at 5um and nuclear annotations were in concordance but differed from the unsupervised label; here my hypothesis is that 15um voronoi captured to many markers from neighbouring cells, 3) still to be determined (blue).

for stromal markers, making their annotations more reliable than those derived from our limited panel.

2. Erythroid or Cycling Immature Erythroid: Cells classified into these categories in either unsupervised or 5um singleR were retained, as erythroid markers are more frequently nuclear-localised, reducing the likelihood of segmentation-related misclassification.

3. MNP (Myeloid/Monocyte Lineage): Cells labeled as classical monocytes, non-classical monocytes, macrophages, dendritic cells (DC), or MNPs in any of the resegmented annotations (Azimuth, Tan Lab) or unsupervised clustering were grouped under a broad MNP category; provided the cell was assigned to this lineage more than once. This grouping avoids over-interpreting noisy distinctions among myeloid subtypes while still capturing lineage identity.

4. All other unclassified cells (n = 422,365) were assigned to their original unsupervised cluster.

These final annotations reflect how cells were reassigned using the two additional segmentations (nuclear and 5um voronoi resegmentation), supported by reference-based labeling (Fig. 3.30a).

However, some annotations were manually overridden based on high-confidence criteria. Cells identified as megakaryocytes from H&E morphology and those meeting blast cell criteria (e.g., co-expression of at least one CD34 and one KIT marker, denoting HSPCs) were reassigned accordingly. These adjustments resulted in the final annotation set shown in Fig. 3.30b.

Note that the final annotations only include cells with matched segmentations across the 15 um Voronoi, 5 um Voronoi, and nuclear datasets. In some cases, particularly near megakaryocyte insertions, cells could not be index-matched due to substantial segmentation changes. Across all samples, only 1,361 cells were excluded for this reason. These unmatched cells were assigned their original unsupervised annotations (Fig. 3.31).

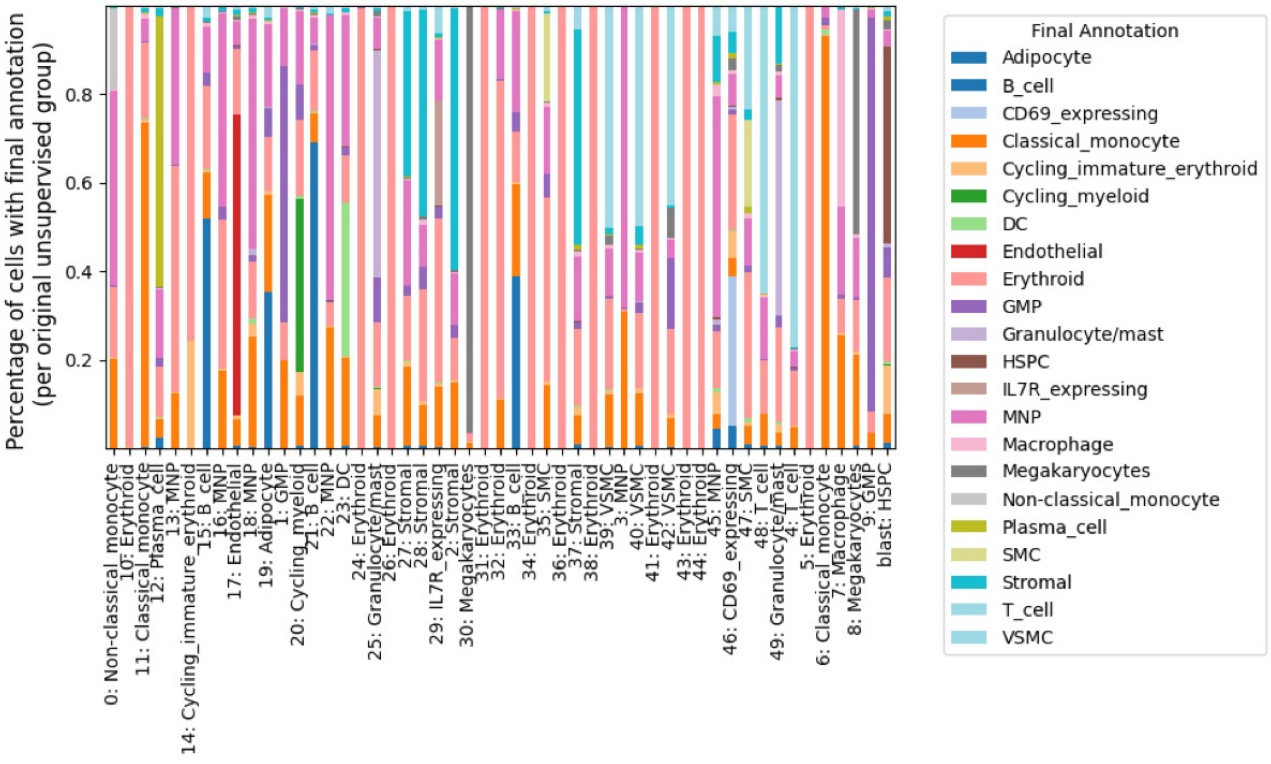
3.4.8 Final Annotation

Next, I performed differential gene expression analysis per cluster using the downsampled dataset (20,000 cells per sample). Gene counts were aggregated based on the 15 um Voronoi segmentations. As expected, expression profiles reflect the updated, expanded cell annotations (Fig. 3.32).

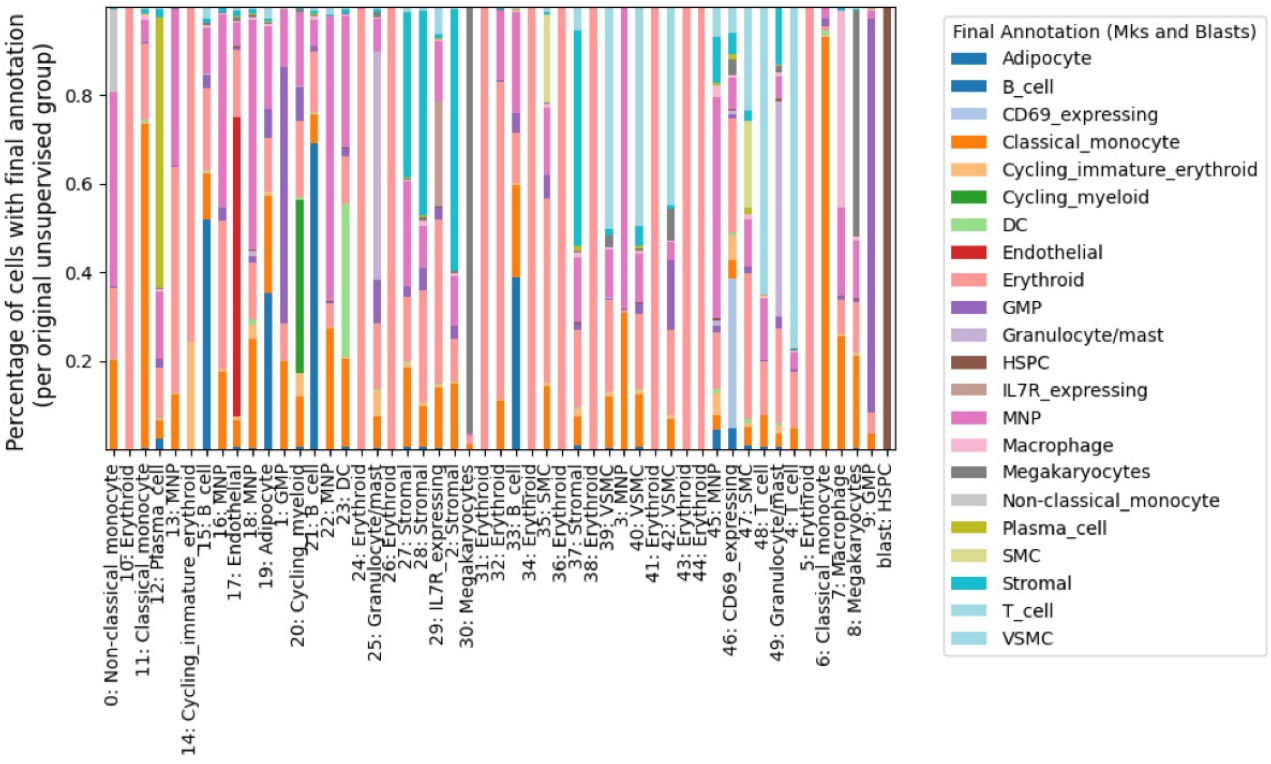
These dot plots demonstrate improved and clearer assignment of cells to their respective lineages. This is particularly evident when comparing the top 3 defining markers in each celltype in the dot plots for the final annotations compared to the original unsupervised annotations (Fig. 3.16 vs (Fig. 3.32). We can visualise these annotations on a UMAP as before (Fig. 3.33), and as a change in the percentage of cells in each cell type which express the lineage markers that we previously suggested define each population in Table 3.4 between the final annotations and unsupervised annotations (Fig. 3.34)

3.4.9 Annotation Resolution

In practice, it is often more informative to examine cells at multiple levels of resolution. For instance, identifying lymphocyte aggregation may be more relevant than distinguishing between T and B cells individually. To accommodate this, I defined five hierarchical annotation levels by progressively merging the original clusters (Table 3.8).



(a) Final annotations per unsupervised cluster in 15 um Voronoi segmentations, based on 5 um and nuclear segmentation results. Colours denote cell labels; bar length indicates cluster proportions.



(b) Same as (a), but with final overrides: megakaryocytes identified from H&E morphology, and blast cells defined by co-expression of CD34 and KIT in the 15 um Voronoi segmentation.

Figure 3.30: Comparison of final annotations for unsupervised clusters from 15 um Voronoi segmentations, before and after applying morphological (H&E) and marker-based overrides.

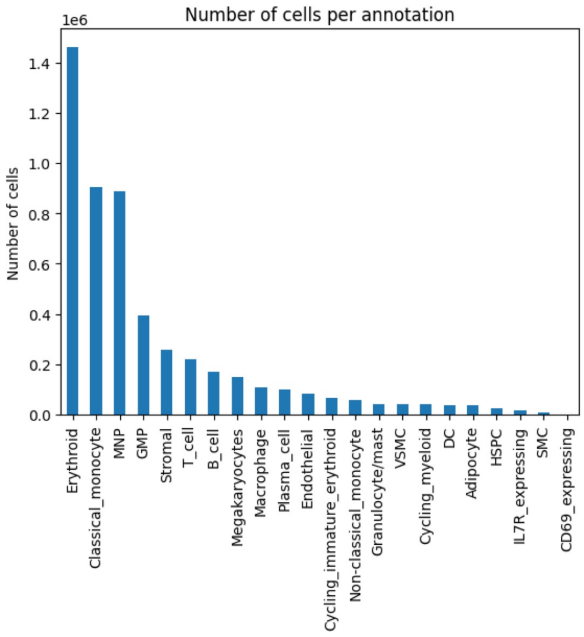


Figure 3.31: The number of cells assigned to each category across all samples (n=32) after processing

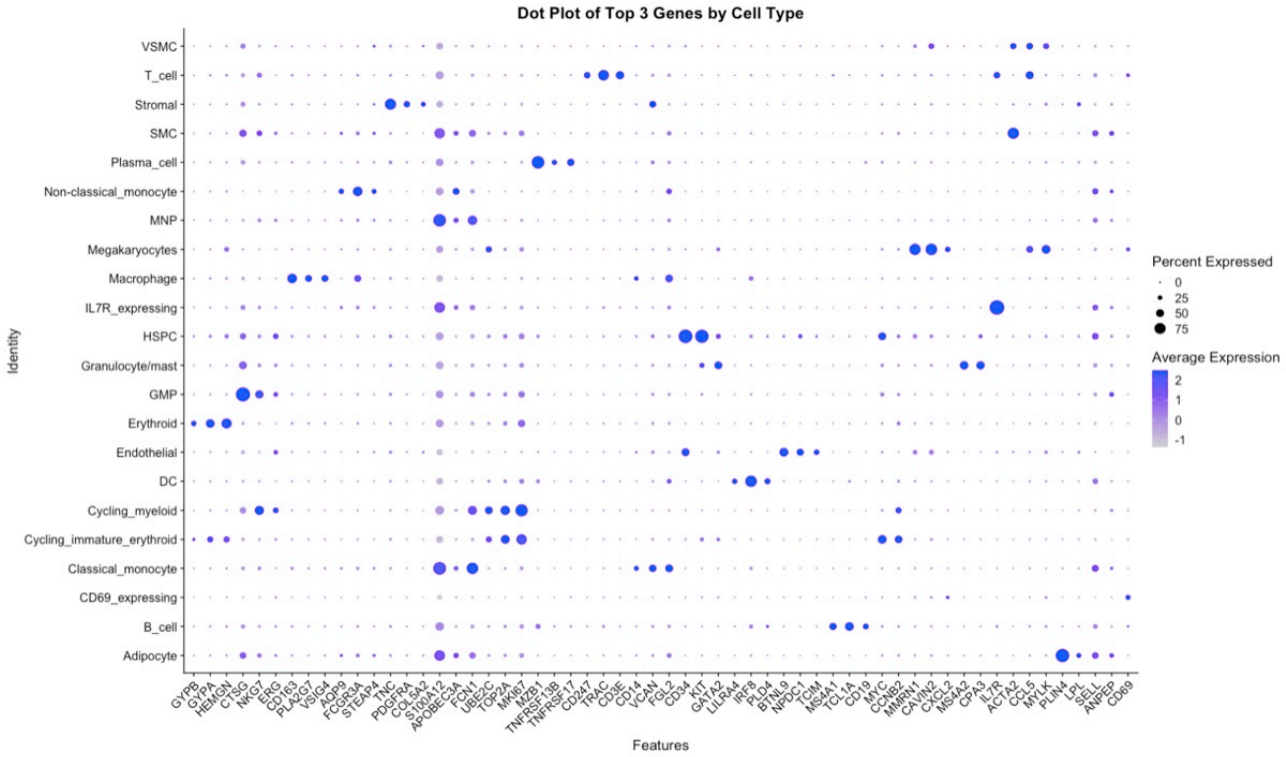


Figure 3.32: Top 3 differentially expressed genes per cluster (avglog2FC > 1, >25% of cells expressing). Dot size indicates % of cells expressing the gene; color indicates average expression.

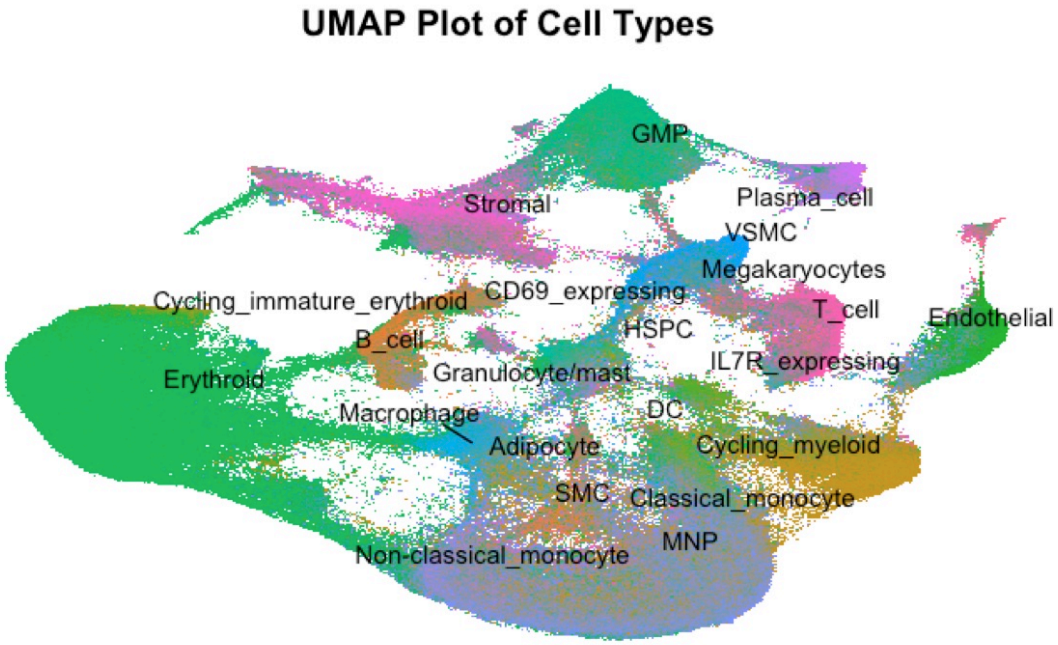


Figure 3.33: UMAP of the final annotations (visualised on a downsampled (20000 cells per sample) object)

First Pass	Level 1	Level 2	Level 3	Level 4	Level 5
Erythroid	Erythroid	Erythroid	Erythroid	Erythroid	Erythroid
MNP	Myeloid	Myeloid	Myeloid	Myeloid	Myeloid
Classical_monocyte	Classical_monocyte	Classical_monocyte	Monocyte	MNP	MNP
GMP	GMP	GMP	GMP	Immature_Myeloid	Immature_Myeloid
Stromal	Stromal	Stromal	Stromal	Stromal	Stromal
T_cell	T_cell	T_cell	T_cell	T_cell	Lymphocyte
B_cell	B_cell	B_cell	B_cell	B_cell	Lymphocyte
Megakaryocytes	Megakaryocytes	Megakaryocytes	Megakaryocytes	Megakaryocytes	Megakaryocytes
Macrophage	Macrophage	Macrophage	Macrophage	MNP	MNP
Plasma_cell	Plasma_cell	Plasma_cell	Plasma_cell	B_cell	Lymphocyte
Endothelial	Endothelial	Endothelial	Endothelial	Endothelial	Endothelial
Cycling_immature_erythroid	Cycling_immature_erythroid	Erythroid	Erythroid	Erythroid	Erythroid
Non-classical_monocyte	Non-classical_monocyte	Non-classical_monocyte	Monocyte	MNP	MNP
VSMC	VSMC	VSMC	SMC	SMC	Stromal
Granulocyte/mast	Granulocyte/mast	Granulocyte/mast	Granulocyte/mast	Granulocyte/mast	Granulocyte/mast
DC	DC	DC	DC	MNP	MNP
Adipocyte	Adipocyte	Adipocyte	Stromal	Stromal	Stromal
Cycling_myeloid	Cycling_myeloid	Myeloid	Myeloid	Myeloid	Myeloid
HSPC	HSPC	HSPC	HSPC	HSPC	HSPC
IL7R_expressing	IL7R_expressing	T_cell	T_cell	T_cell	Lymphocyte
SMC	SMC	SMC	SMC	SMC	Stromal
CD69_expressing	CD69_expressing	CD69_expressing	CD69_expressing	CD69_expressing	CD69_expressing

Table 3.8: Table of hierarchical classifications for cell annotations.

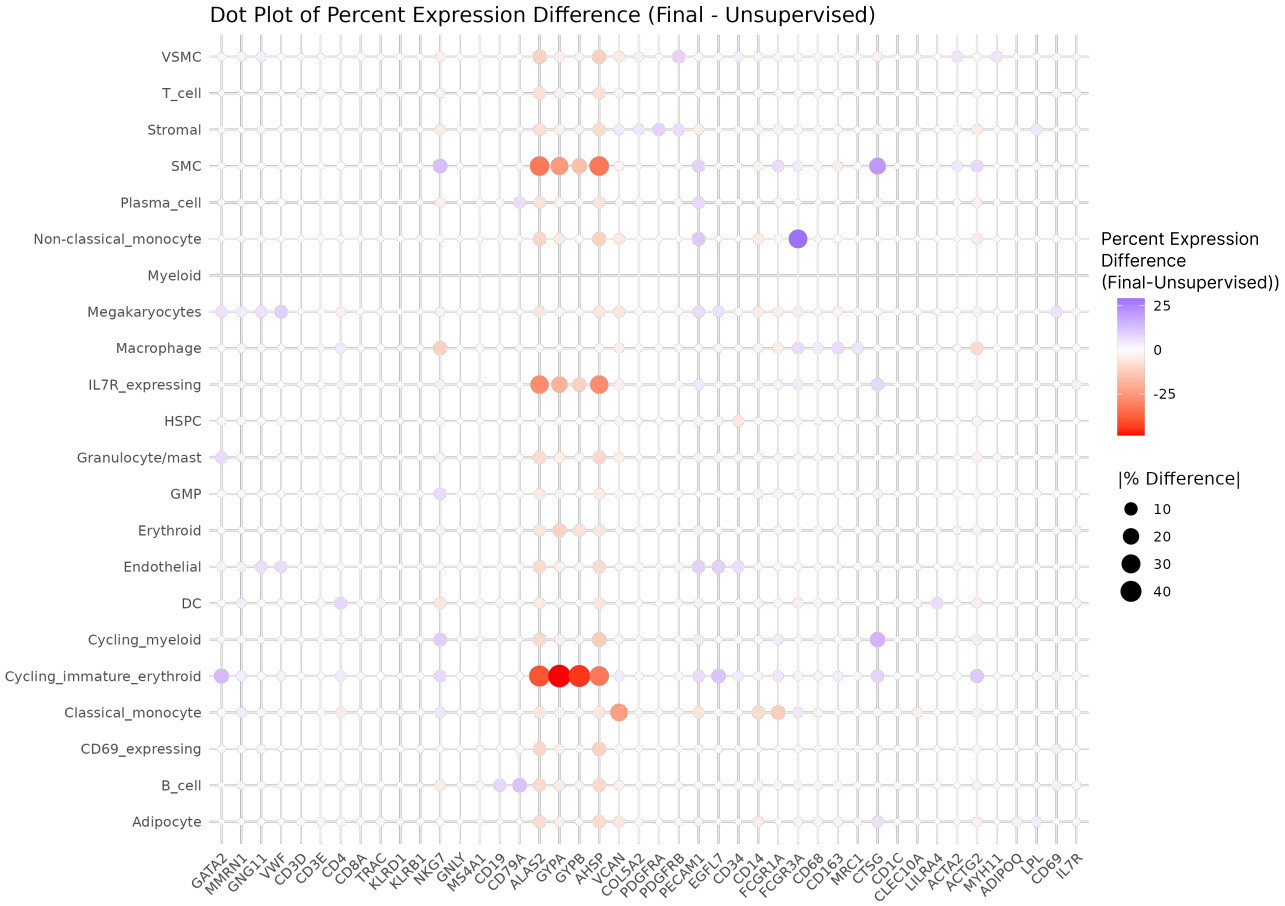


Figure 3.34: Difference in percentage of cells expressing lineage markers between unsupervised and final annotations in each cell type. Dot size corresponds to the absolute difference in the percentage of cells expressing each marker in the cell type. Colour corresponds to direction of difference, red (greater percentage of cells in the population express the marker in the unsupervised annotations), blue (percentage is greater in the final annotations)

I also aimed to investigate the stromal annotations in greater depth, as these cells are rarely captured in single-cell datasets and represent a unique aspect of our data. To refine these annotations, I applied the following procedure:

1. Extract all stromal cells from annotation level 5.
2. Compare these cells to the SingleR annotations from the Tan Lab reference using the 15 um segmentation.
3. Reclassify any cell not annotated as Osteo-MSC, Osteoblast, Adipo-MSC, VSMC, or SMC as Stromal-NOS (Not Otherwise Specified).
4. Merge Osteo-MSC and Osteoblast into a single category: Osteo-MSC.
5. Merge VSMC and SMC into a single category: SMC.
6. If any cell originally annotated as SMC in annotation level 1 is now labelled Stromal-NOS, reassign it back to SMC.

This gives us an additional 5 levels, with stromal subsets incorporated (Fig. 3.35).



Figure 3.35: Top 5 differentially expressed genes per stromal cluster (Wilcoxon test; avglog2FC > 1, pct > 25%), showing expression level and percent of cells expressing each gene. Based on downsampled data (20,000 cells per sample).

3.4.10 Incorporating Macrostructures from H&E

We have already demonstrated how megakaryocyte segmentations, identified from H&E images, were integrated into our single-cell segmentation. However, other macrostructures within the bone marrow also play important roles in spatial context, and we sought to incorporate these to investigate how cells behave relative to them.

Bone

Dr H. Ryou developed and applied an in-house algorithm to detect bone structures in each bone marrow trephine from the H&E images. Although the algorithm is still under active development and generates some noise, true bone segmentations were present within the output. I therefore applied a threshold based on area to the noisy segmentations to isolate accurate segmentations for each trephine (Fig. 3.36).

Fat

To segment adipocytes (fat cells) in the H&E images, we used MarrowQuant2, an adaptation of StarDist, trained specifically on bone marrow trephines [99]. Post-processing was necessary, including area-based thresholding, as bone detachment artifacts were often misclassified as fat (Fig. 3.37a).

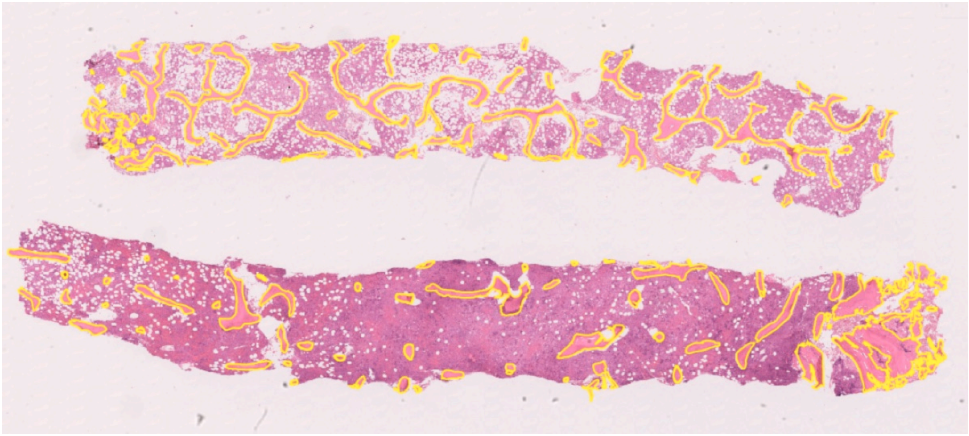
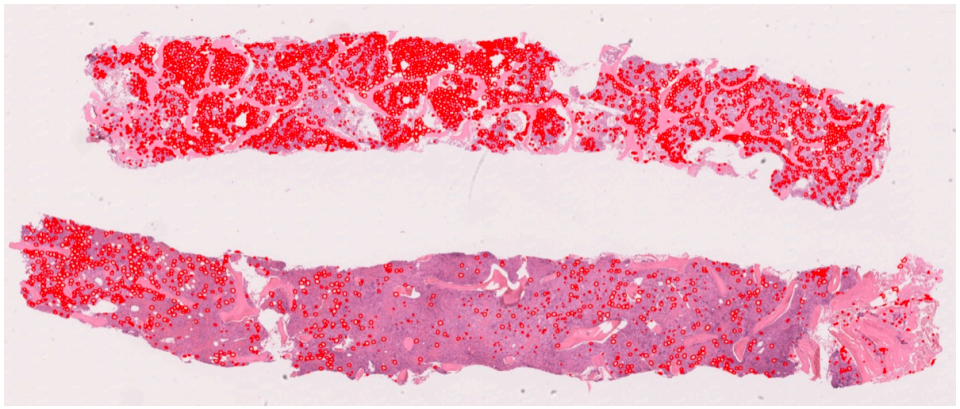
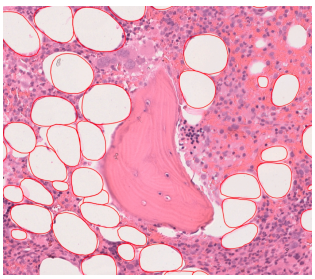


Figure 3.36: Final bone segmentations overlaid on H&E images for two samples.

While the thresholded MarrowQuant2 results were generally reliable, some adipocytes were still missed. To address this, we manually segmented fat cells in each trephine. Despite this, missing adipocytes remain, although quantifying this is difficult due to the lack of a ground truth. A particularly challenging area was around the bone, where distinguishing between tissue detachment and true fat space was often subjective (Fig. 3.37b).



(a) Fat segmentation using MarrowQuant2 after area thresholding and manual refinement. The top trephine is from a PV patient, and the bottom from an MF patient.



(b) Segmentation challenges around bone structures. The bottom tip of the bone shows an example of tissue detachment (not fat). Some adipocytes are also missed.

Figure 3.37: (A) Fat segmentation using MarrowQuant2 in PV and MF trephines. (B) Segmentation challenges near bone, where tissue detachment and some missed adipocytes are evident.

Vessels

Arterioles are small branches of arteries that lead to capillaries and carry oxygenated blood. Sinusoids, by contrast, are wider capillaries with higher permeability, facilitating nutrient exchange. To capture these structures, an expert pathologist, Dr C. Pescia, manually annotated arterioles and sinusoids in each H&E-stained trephine.

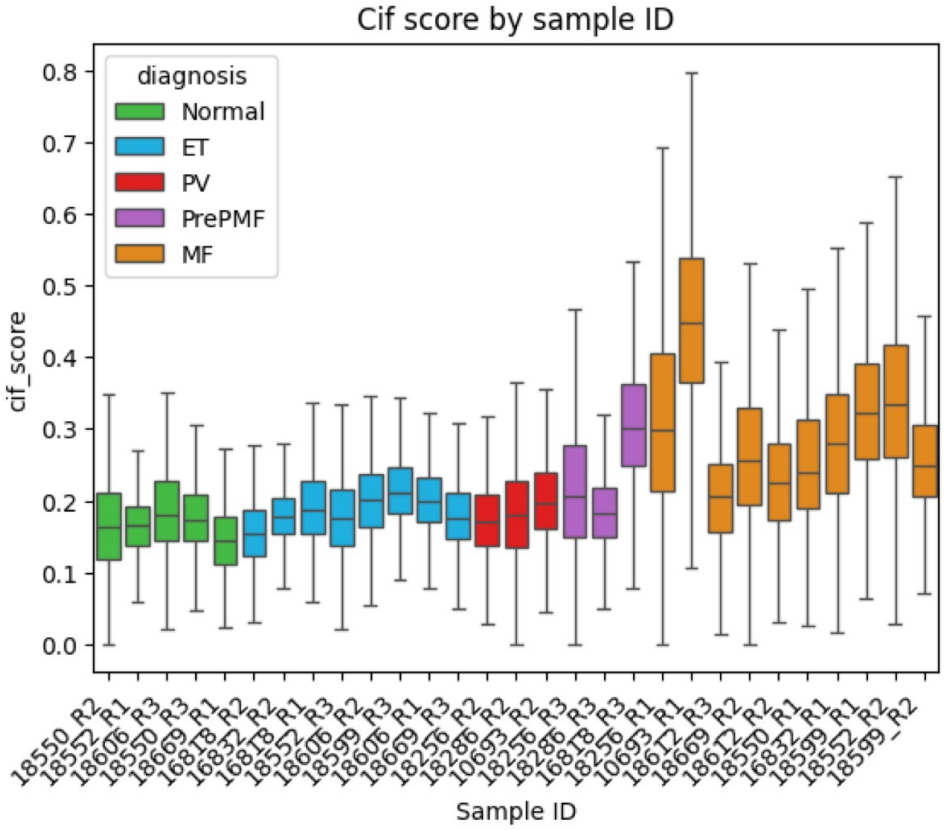


Figure 3.38: CIF score of each tile per sample (n=30), coloured by condition aross the spataly preserved area.

Fibrosis

We also applied a recently developed method for detecting fibrosis directly from H&E images. This approach identifies fibrotic regions at the tile level (512×512 pixels) with a stride of 128 pixels [46, 57]. Dr H. Ryou ran the algorithm on the accompanying H&E slides, and I registered and processed the raw coordinate output, assigning scores only to areas of spatially preserved tissue. Each tile is given a score within a range of 0-1. We can plot the distribution of tiles within each condition. One MF patient (two trephines, 18612 R1 and R4), does not have a fibrosis score as the sample was too bony to be able to calculate. I've calculated the range of fibrosis scores for each sample (Fig. 3.38).

We also binned the scores which broadly correspond to different fibrosis scores as outlined in [46] (Fig. 3.39).

- MF-0: Bin 1 (CIF score 0 - 0.267)
- MF-1: Bin 2 (CIF score 0.2671 - 0.439)
- MF-2: Bin 3 (CIF score 0.4391 - 0.647)

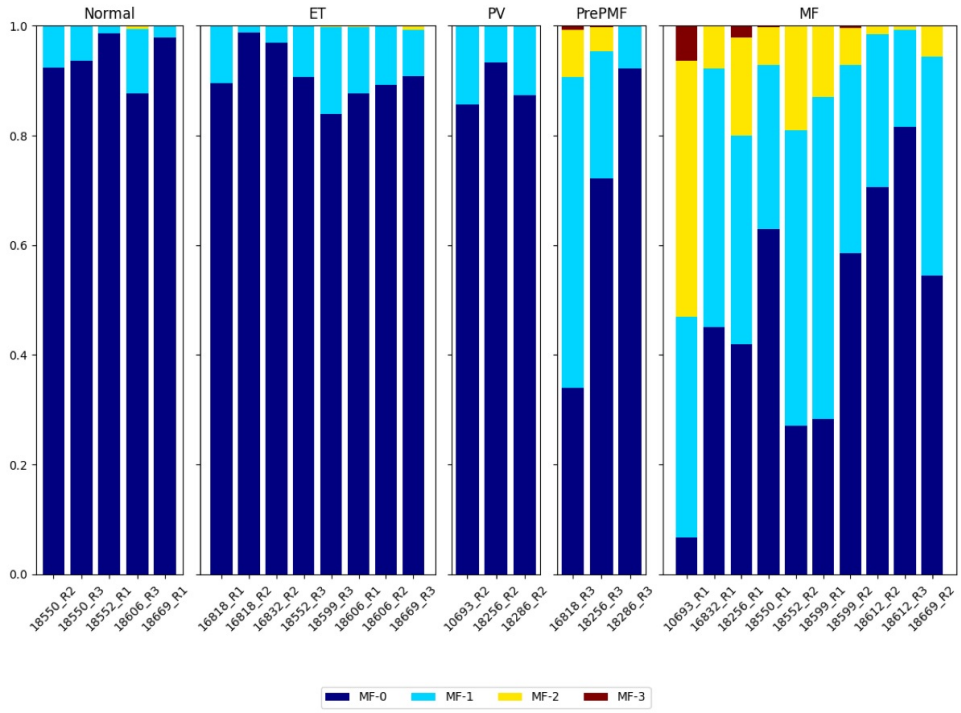


Figure 3.39: Distribution of fibrosis scores across samples in each condition across the spatially preserved area. Scores were binned to follow most closely WHO criterea (MF-0: Bin 1 (CIF score 0 - 0.267), MF-1: Bin 2 (0.2671 - 0.439), MF-2: Bin 3 (0.4391 - 0.647), MF-3: Bin 4 (0.6471 – 1))

- MF-3: Bin 4 (CIF score 0.6471 – 1)

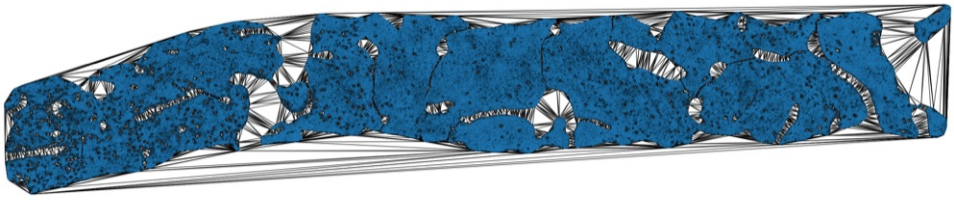
3.4.11 Graph Creation

Graph Without Fat

Constructing a representative graph of cell–cell interactions is essential for downstream spatial analysis. The most widely used Python package for spatial statistics, Squidpy, constructs such graphs using Delaunay triangulation by default. Delaunay triangulation favors fatter triangles whose circumcircles do not contain other points, which can be advantageous in uniform datasets. However, this approach may not be ideal for large irregular cells with many small neighbours (e.g. megakaryocytes). Squidpy creates graphs using the centroid of each segmented cell as a node and defines edges based on the distance between these centroids.

An example of a default Delaunay graph, constructed without filtering, is shown in Fig.3.40a. The bone marrow is a long, narrow tissue with many bone fragments. Connecting cells across bone segments introduces artefacts, as real biological interactions are unlikely in such cases. To explore distance-based filtering, I examined the radius around the centroids of megakaryocytes, the largest cells in the tissue, at 100, 150, and 200 pixels (21.25, 31.88, and 42.5 um) (Fig. 3.41).

Applying a threshold of 100 pixels produced a much more biologically plausible graph (Fig. 3.40b). To assess the performance of Delaunay-based graphs, I constructed graphs with multiple edge-length thresholds (100 and 150 pixels) on spatially preserved regions only. Delaunay graphs with different radius thresholds are shown in Figs. 3.42a and 3.42b.



(a) Delaunay graph constructed on spatially preserved cells from a bone marrow trephine. Blue dots represent cell nodes; black lines represent edges. No distance filtering or spatial exclusions applied.



(b) Delaunay graph using a maximum edge length of 100 pixels. Same trephine as Fig. 3.40a. Shorter edges yield a more realistic representation of potential cell–cell interactions.

Figure 3.40: Comparison of Delaunay graph constructions with and without distance filtering.

While both thresholds generate reasonable graphs, they struggle with irregular cell shapes. At 100 pixels, some neighbours of large megakaryocytes are missed. Moreover, elongated megakaryocytes can cause their immediate neighbours to be incorrectly connected to each other, skipping the megakaryocyte entirely (Fig. 3.42a). At 150 pixels, this issue persists, and additional spurious connections appear as longer triangles are allowed (Fig. 3.42b). Given these limitations, and the fact that we already have Voronoi-based segmentations (extending up to 15 um), I constructed an alternative graph where edges are created between cells whose boundaries touch. To address artefact-induced isolation of megakaryocytes, I merged any artefact polygons that touched megakaryocytes for the purpose of graph construction only. This approach results in a more biologically faithful graph. While it may still allow for some 30 um gaps in edge formation (based on 2 cell radii of the maximum Voronoi allowed size), this is acceptable given the 2D nature of the dataset and unknown 3D extensions of the cells (Fig. 3.43a). This "neighbor graph" was constructed across all spatially preserved regions in the dataset.

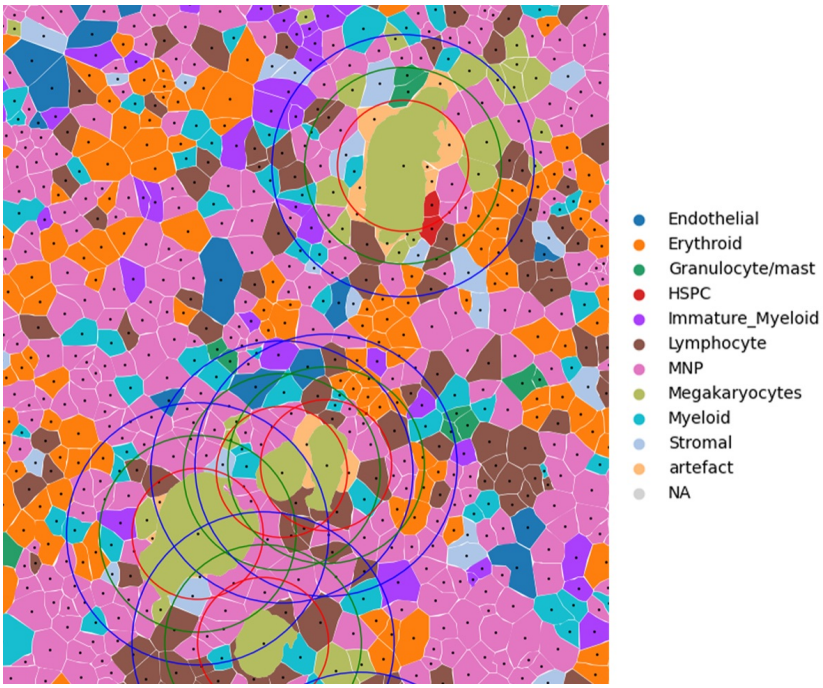
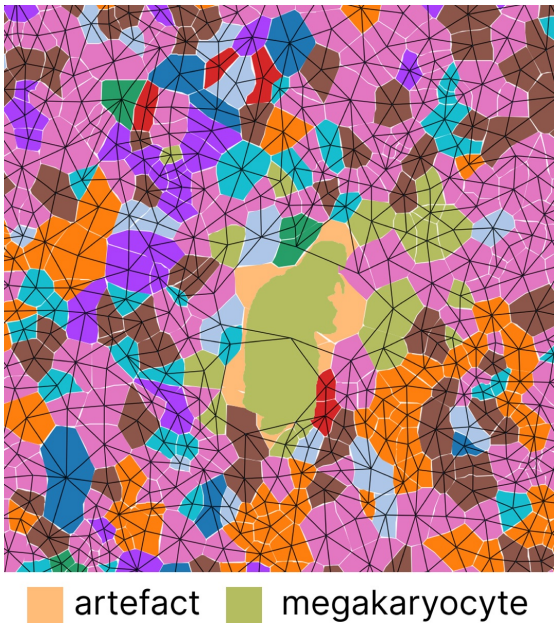
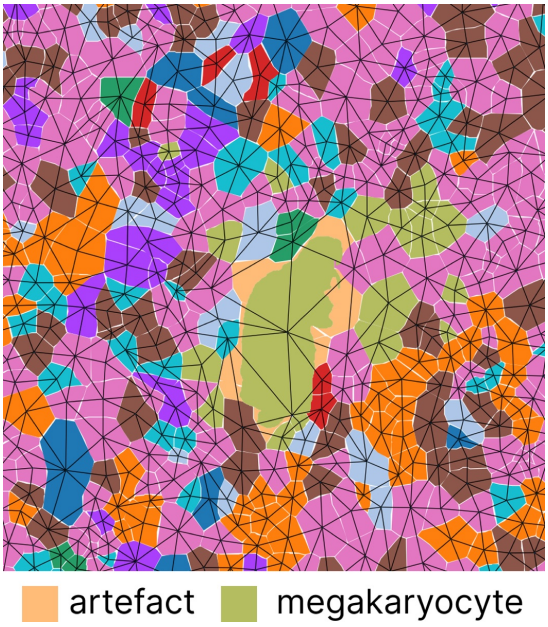


Figure 3.41: Megakaryocytes (green) are the largest cells in the tissue. Circles represent search radii at 100 (red), 150 (green), and 200 (blue) pixels. Dots represent the centroids of each cell (a node). These circles define which nearby nodes would be included in the graph under different edge length thresholds.



(a) Graph constructed with a 100-pixel edge length. Cell types are color-coded.

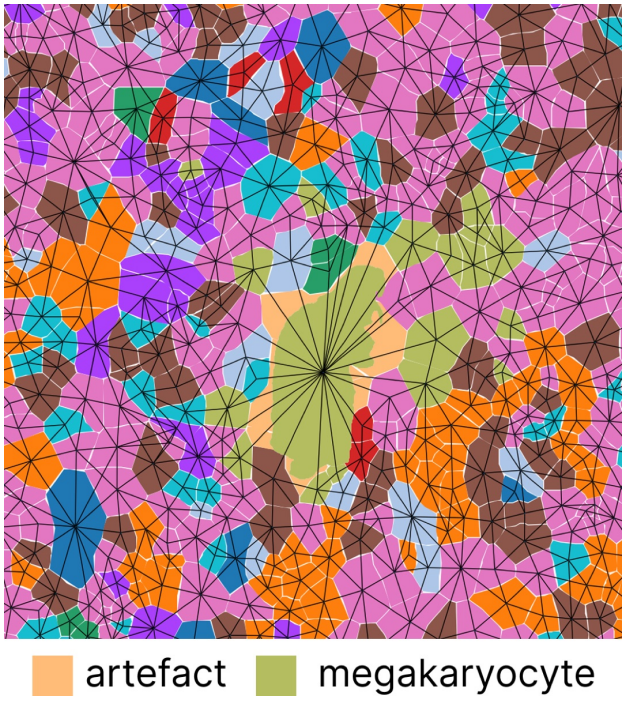


(b) Graph constructed with a 150-pixel edge length. Larger radius captures more cells but introduces artifacts.

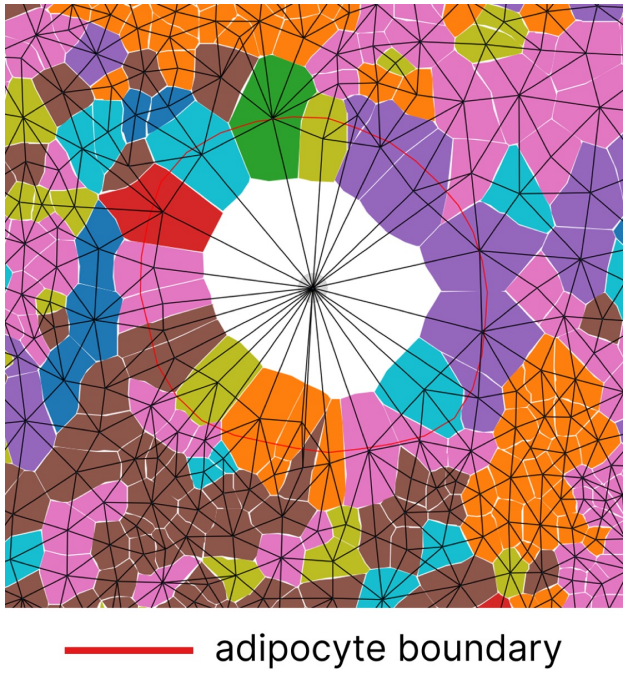
Figure 3.42: Comparison of Delaunay graphs constructed using 100-pixel (a) and 150-pixel (b) edge length thresholds on spatially preserved regions.

Graph Construction With Fat

The same boundary-based graph construction was extended to include fat (adipocyte) segmentations. Edges were created if an adipocyte boundary touched any cell (Fig.3.43b). This addition is crucial for investigating potential adipocyte–cell interactions, particularly within the emerging context of the peri-adipocytic niche.



(a) Graph constructed using boundary-touching relationships in the 15um Voronoi segmentation. Artefacts were merged into megakaryocytes for edge construction only. Colours represent cell types.



(b) Graph with fat cells included (red boundaries). Nested fat cells remain due to preprocessing limitations. Only cells in intertrabecular spaces are included.

Figure 3.43: Comparison of neighbor-based graph construction from 15um Voronoi segmentations, without (a) and with (b) fat segmentation.

3.5 Discussion

In this chapter, I present a novel pipeline for processing Xenium spatial transcriptomic data from bone marrow across all MPN subtypes and healthy controls. To enhance spatial context, we integrate machine learning–based fibrosis maps and incorporate macrostructures such as adipocytes, bone, and vasculature. Key innovations include methods for handling irregular cell morphologies (e.g. megakaryocytes), including addition into an existing segmentation and a boundary-aware Voronoi graph that maintains neighbourhood structure even in the prescence of irregular cells. I also improved cell type annotation using multi-scale segmentations and

scRNA-seq references, including a stromal-enriched dataset. At the time of writing, no standard pipeline existed for Xenium data, and this work lays a foundational pipeline for its meaningful analysis.

While 10x Genomics now recommends the 5um Voronoi expansion as the default segmentation strategy, our multi-resolution approach offers a more robust framework for capturing biologically relevant variation in cell identity and spatial organization. Comparing annotations across nuclear, 5um, and 15um segmentations and to reference datasets allowed us to assess annotation stability and resolve ambiguities, critical in bone marrow, where complex architecture, large multinucleated cells, and rare populations challenge traditional unsupervised annotation.

Our method for megakaryocyte insertion illustrates the benefit of incorporating morphological information. By registering H&E-derived segmentations with Xenium DAPI images, we accurately inserted large megakaryocytes into the segmentation and so expression matrix. While registration shifts can affect smaller or irregular cells like granulocytes or fibroblasts, the size of megakaryocytes makes this less problematic. We note, however, that some polylobated megakaryocytes may still be missed or fragmented, leading to potential overestimation.

Finally, while SingleR-based annotation is limited by the relatively small gene overlap (365 genes) between Xenium and scRNA-seq references, we mitigated this by requiring consistency across segmentations and reference datasets (e.g., Azimuth and Tan Lab). This multi-layered validation ensures that reannotations are biologically meaningful, not artifacts of a single segmentation or method.

3.6 Future Work

The functions for megakaryocyte integration into Xenium segmentations and for graph construction that better captures neighbours of irregularly shaped cells will be released as a standalone package. This tool will be valuable for other groups working with bone marrow spatial transcriptomic data who seek accurate incorporation of megakaryocytes and biologically meaningful graph representations.

4

Spatial Transcriptomics Analysis of the MPN Bone Marrow

Chapter Contributions to the Literature

- Identified increased megakaryocytes in ET and MF, driven by gene-expression-defined cells and phenotype 8.
- Observed global expansion of Osteo-MSCs and widespread gene expression changes in MF, including fibrosis- and EndMT-associated markers.
- Defined 10 unsupervised Cellular Neighbourhoods (CNs), each enriched for distinct cell types and mapped to anatomical compartments (e.g., CN3: endosteal, CN5: lymphoid aggregate).
- MF is marked by spatial expansion of CN0, CN2, CN9; ET by CN6, CN7, CN9. CN5 loses CN neighbours prone to spatial expansion in MF/ET, suggesting impaired local immunosurveillance allows expansion.
- CN1, CN3, and CN4 resist expansion and have conserved relationships across conditions, indicating previously unreported preserved spatial architecture in the healthy bone marrow.
- CN3/4/9 are enriched for HSPCs and flanked by lineage-restricted cells, supporting spatially organised differentiation. CN9 expansion aligns with megakaryocyte-biased differentiation in MPNs.
- CN2 has higher fibrosis scores and relocates away from bone across disease stages, suggesting it marks early fibrosis or a cellular correlate of fibrosis.
- CN9 invades peritrabecular space in MF, displacing GMP-rich CN6, suggesting an altered myeloid niche in MF.
- CN8 expands in MF at the expense of CN2 and correlates with higher fibrosis grades (MF-2), possibly marking late-stage progression.
- High fibrosis shows increased monocytes, DCs, B cells, and decreased Osteo-MSCs and megakaryocytes.

- Osteo-MSCs spatially associate with DCs, T and B cells at high fibrosis, whilst monocytes remain distant from Osteo-MSCs, suggesting indirect or subpopulation-specific monocyte-Oteo-MSC interactions explains CN2 cell enrichment.
- Found previously unreported spatial proximity between Osteo-MSCs and megakaryocytes at high fibrosis grades.
- Osteo-MSCs downregulate Wnt inhibitors (SFRP4, APCCD1, NTN4) in high fibrosis, supporting loss of regulatory function and possibly aligning with a previously identifed NCAM-1+ pro-fibrotic osteolineage cells.
- CN9 is enriched in MF-0 grade, spatially expanded in ET and MF, loses CN5 association, and megakaryocytes and endothelial cells expresses CCL5 (endothelial injury marker) in MF-0 and SNAI1 (EndMT) in MF. CN9 may represent a precursor fibrotic niche.
- Linked megakaryocyte morphology to gene expression: smaller cells are more proliferative and immature, while larger ones (e.g. phenotype 8) express CCL5.

4.1 Author Contributions

I received guidance from Dr Rosalin Cooper, Dr Muhammad Dawood, Dr Hosuk Ryou, Dr Carlo Pescia, Dr Anna Sozanska, Professor Dan Woodcock, Professor Jens Rittscher, and Professor Daniel Royston. I performed all data analysis.

4.2 Introduction

In this section, I use the integrated Xenium dataset to explore the immune landscape of MPNs, focusing on both haematopoietic and stromal populations, and their relationship to bone marrow architecture and cytomorphology. First, I compare global cell abundances in bone marrow from healthy donors and patients with the four MPN subtypes (ET, PV, Pre-MF, MF). Unlike cell suspensions or proxies (e.g. blood or aspirate), in situ spatial data avoids preprocessing losses and captures all marrow cells, offering a comprehensive view of cellular shifts across health and disease. This global perspective is essential for interpreting spatial changes within and across disease states. Second, I examine global gene expression profiles by performing differential expression analysis between MF and healthy donors within each cell type. This highlights broad transcriptional changes across the disease spectrum and reveals what may be missed without spatial context. Third, I use an unsupervised approach to identify distinct bone marrow cellular neighbourhoods and track how they change across disease stages. I examine their spatial relationships with each other, the trabeculae, and fibrosis, and interpret these in the context of HSC niches and fibrotic progression. Next, I investigate how cell abundance, expression, and

interactions vary across increasing fibrosis grades in MF, focusing on changes between lower and higher grades. Finally, I demonstrate the ability of spatial transcriptomics to link gene expression with morphology identifiable on standard H&E slides. By correlating morphology and molecular profiles, we can infer genetic phenotypes from routine histology.

4.3 Global cellular abundance in the Bone Marrow in MPNs

4.3.1 Introduction

This dataset provides a unique in situ view of human MPN bone marrow at single-cell resolution, including stromal populations. Its scale allows for robust exploration of global cellular abundance across healthy donors and MPN subtypes. While our primary interest lies in the spatial relationships between cell populations, we first examine global changes. This contextualises local spatial patterns and helps profile broad shifts in the immune landscape.

4.3.2 Methodology

Choice of Statistical Test

A key challenge in differential abundance analysis is the compositional nature of the data; as cell populations are proportions of a whole, an increase in one necessarily results in a relative decrease in another. This interdependence can obscure the true source of change if not accounted for. A second challenge is that each sample represents a subset of the tissue, so cell counts should be interpreted as proportions rather than absolute values.

An alternative suggestion is to report cells normalised by marrow area. When interpreting this metric we should be mindful of the effect of differing size of cells between both conditions (hyper-cellular vs hypo-cellular tissues) and cell types (larger megakaryocytes vs smaller lymphocytes).

Methods for differential abundance generally fall into three categories: Wilcoxon rank-sum tests, negative binomial models, and compositional models such as scCODA. For this analysis, I used the generalised linear model likelihood ratio test (GLM-LRT) from the edgeR package. While it

does not explicitly model compositional effects, it is well-suited for overdispersed count data and supports complex experimental designs, including sample-level variability. Overdispersion, where variance exceeds the mean, is common in biological data.

edgeR uses a negative binomial model to account for this, making it more robust. The GLM-LRT compares a full model including all conditions with a reduced model excluding the condition of interest, to test whether that condition significantly improves model fit. No additional normalisation is applied, as total counts (i.e., library size) are already incorporated into the model [100]. For differential abundance across the full trephine, I used the model design: design <- model.matrix(~ 0 + diagnosis2). This coding avoids assigning a reference group, allowing for unbiased comparisons between all conditions.

An Example of Checking Model Fit

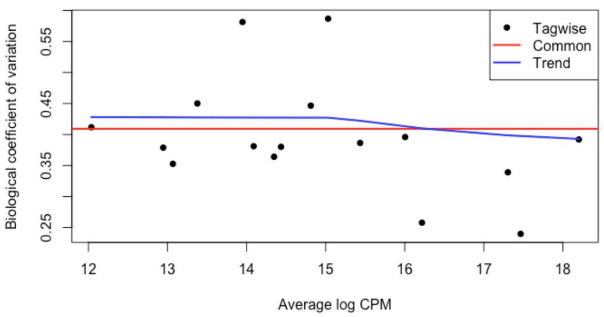
I performed differential abundance analysis at four annotation levels: level 1, level 3, level 5, and level 3 with refined stromal annotations (Table 3.8). Additionally, I included adipocytes, which were segmented separately from the H&E slide, as an extra cell type.

Below, I show an example of model fit checking using the level 3 annotations without stromal refinement. This analysis was applied at the whole-trephine level. Biological measurements are inherently noisy. One sense check when fitting a model with edgeR is the Biological Coefficient of Variation (BCV). Unlike in a Poisson model, where $Var(Y) = \mu$ (where $\mu =$ the mean), a negative binomial model used in edgeR accounts for overdispersion by modelling $Var(Y) = \mu + \phi\mu^2$ (where $\phi =$ the dispersion (extra variability)). In edgeR the BCV is the square root of the dispersion (ϕ), where dispersion controls how much bigger the variation is than poisson (through the $\phi\mu^2$ term). The BCV is modelled on the concept of the coefficient of variation (CV), $\frac{\sigma}{\mu}$, which allows comparisons of variability across scales. For large μ, the μ^2 dominates and so $\mathrm{Var}(Y) \approx \phi\mu^2$, thus that $\sigma \approx \sqrt{\phi}\,\mu$ and so $\mathrm{CV} \approx \sqrt{\phi}$, the BCV.

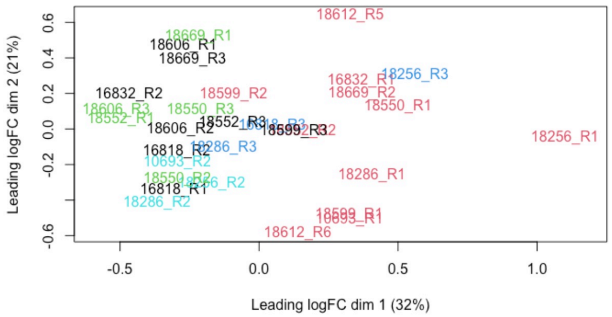
Given this, the BCV can be used as an indicator of inter-sample variability and so how confidently changes can be detected; higher BCV implies greater heterogeneity and reduced power to detect differences unless effects are large and consistent across samples. Within edgeR, the test for significance asks whether including a condition significantly improves the model fit compared with a null model, given inter-sample variability captured by the dispersion. In this way BCV is taken into account in significance testing. The average biological coefficient of variation (BCV)

was 0.35, within expectations for biological data. Two populations, B cells and adipocytes, showed higher BCV. The B cell population is supported by distinct lineage markers, and the adipocytes were directly segmented and visually validated, so I would suggest that their variability more likely reflects true biological heterogeneity and is less likely to be technical variation that is unaccounted for (Fig. 4.1a, 4.1b).

With 30 samples and 5 conditions, we have 25 degrees of freedom. To evaluate model fit, I calculated the residual deviances, which reflect how well the model predicts the observed data. Under a well-fitting model, residual deviances approximately follow a chi-squared distribution. Using the 95th percentile of the chi-squared distribution (37.65) as a threshold (which would allow a deviance above 37.65 5% of the time under a well fitting model), all residual deviances fell below this value. The highest was 30.9 (adipocytes), indicating acceptable model fit. To test for significant differences in abundance, I used a likelihood ratio test (LRT) to compare full and reduced models. False discovery rate (FDR) correction was applied using the Benjamini-Hochberg procedure to account for multiple testing across cell types.



(a) BCV plot showing the relationship between variability and abundance. More abundant cell types tend to be less variable. Each point represents a cell population. The red line shows common dispersion; the blue line shows the trend. The y-axis is the square root of ϕ, the BCV, and the x-axis is average log CPM.



(b) MDS plot of differential abundance between conditions at the whole-trephine level. MF samples (in red) separate from other subtypes. Each sample is labeled by ID.

Figure 4.1: Model quality assessment: (A) Biological coefficient of variation (BCV) by cell population. (B) MDS plot showing sample separation by diagnosis.

4.3.3 Results

Megakaryocytes are elevated in MF, driven by phenotype 8 and morphologically indistinct subtypes.

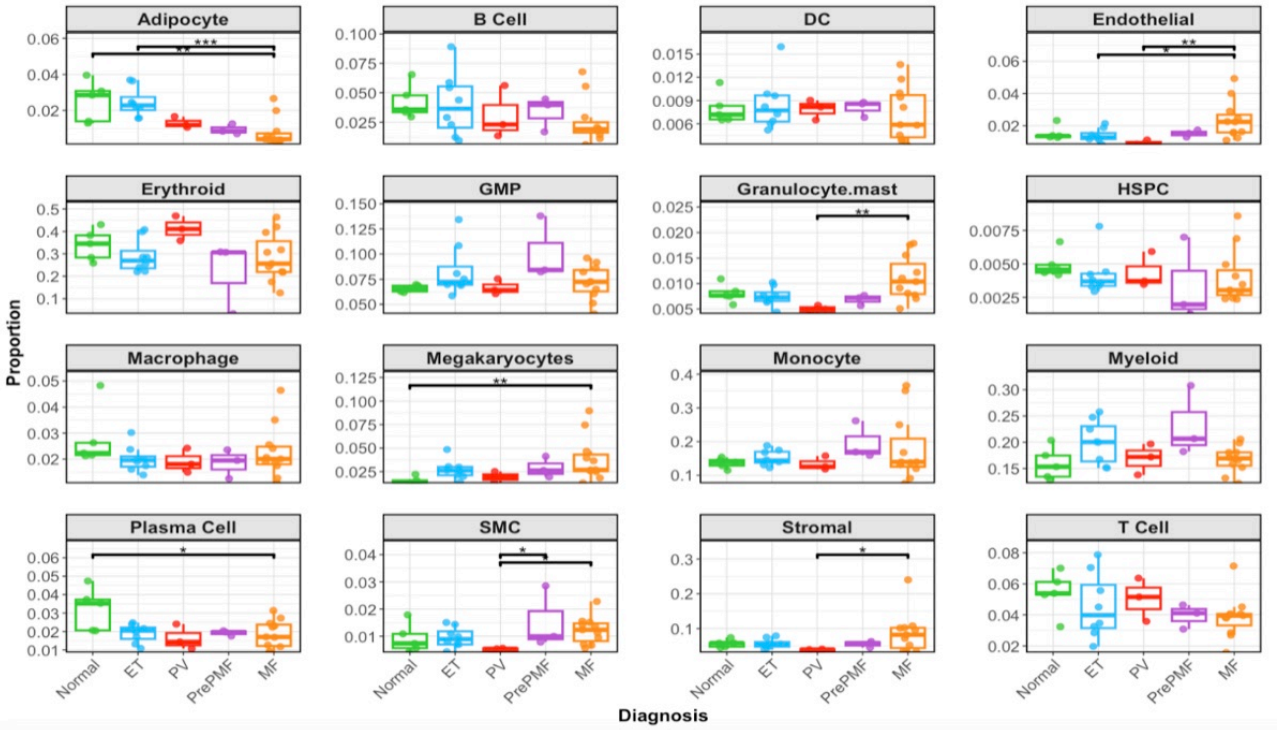


Figure 4.2: Differential Abundance at level 3 annotation across whole trephines. p values *** < 0.001, ** < 0.01, *<0.05.

Across the whole trephine, megakaryocyte abundance is increased in MF compared to healthy controls (Fig. 4.2), consistent with H&E observations of elevated megakaryocyte numbers in MPN samples. Prior work has identified nine morphologically distinct megakaryocyte phenotypes from H&E slides (Fig. 4.55) [17].

Interestingly, gene expression identifies more megakaryocytes than are visible in H&E-stained trephines, which may explain the lack of statistically significant differences between MF and other MPN subtypes despite histological expectations. To investigate this further, I separated megakaryocytes into phenotypic groups (Megakaryocytes_0–8) and a separate group for those not classifiable in H&E (Megakaryocytes). Only phenotype 8 showed a significant increase in MF and ET compared to healthy bone marrow (Fig. 4.3).

Adipocytes are decreased in MF and ET compared to healthy donors

Adipocytes are decreased in MF compared to Normal and ET, consistent with the increased marrow cellularity in MF due to expanded haematopoiesis. Stromal cells are increased in MF compared to PV, and SMCs are elevated in MF and Pre-PMF relative to PV. Endothelial cells are also increased in MF compared to ET and PV. This may reflect a reduction in arterioles in PV marrow, as previously described in comparisons of PV and ET [37]. Plasma cells are reduced

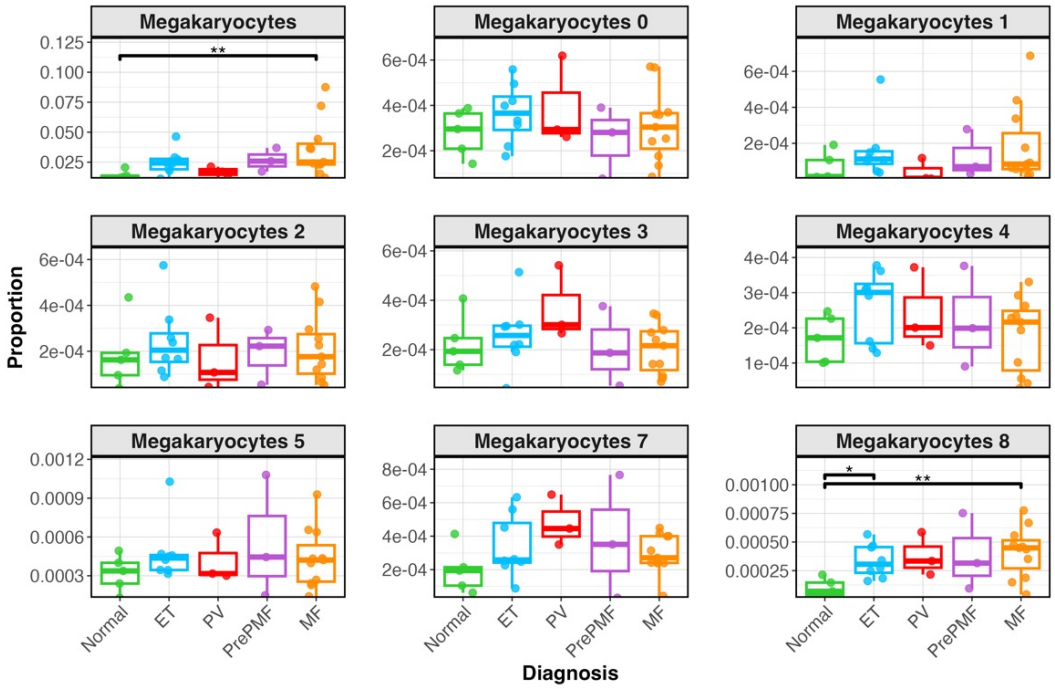


Figure 4.3: Differential Abundance of Megakaryocyte Phenotypes (Megakaryocytes n) and Non-phenotyped Megakaryocytes (Megakaryocytes) across conditions. Megakaryocytes indicate those megakaryocytes which can only be determined from gene expression. Proportions were analysed in the context of all cells at level 3 stromal refined. p values *** < 0.001, ** < 0.01, *<0.05.

in MF compared to Normal, an unexpected finding. However, given MPNs are myeloid-dominant, a relative decrease in lymphoid populations like plasma cells may be anticipated.

While this analysis used level 3 annotations, I repeated the differential abundance analysis using coarser, level 5, and finer, level 1, annotation levels (not shown). The observed trends were consistent across resolutions.

Osteo-MSCs are increased in MF compared to all other conditions

Given the observed differences in Stromal cells, particularly SMCs, between PV and MF, I further examined differential abundance within the refined stromal subsets at level 3.

Osteo.MSCs were significantly increased in MF compared to all other conditions. SMCs remained elevated in both Pre-PMF and MF relative to PV, and the remaining stromal population also showed increased abundance in MF compared to PV.

Cell abundance normalised by trephine area

I have repeated the differential abundance analysis by quantifying cells normalised by trephine area. This is in line with how pathologists would make a rough quantification of cells when

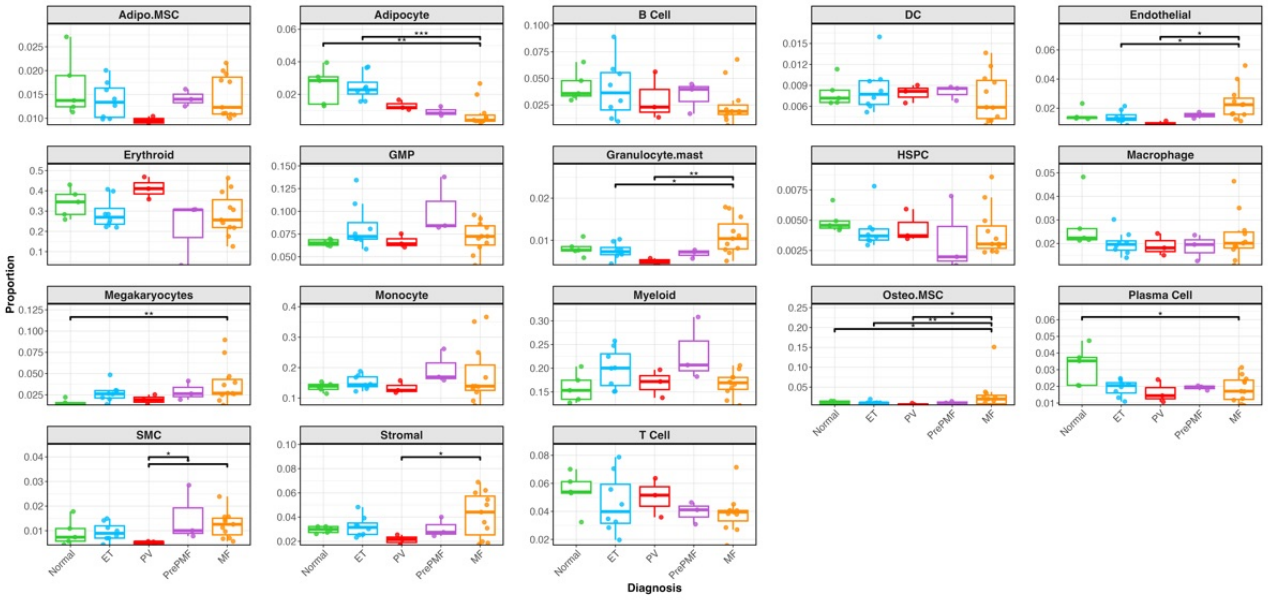


Figure 4.4: Differential Abundance at level 3 with refined stromal populations

evaluating a trephine in the clinical workflow. I repeated each analysis to consider normalisation where trephine area is either the total area of all spatially preserved cells, or normalisation where trephine area is the total area of all spatially preserved cells and spatially preserved adipocytes. I also only included the area of cells which passed other quality controls e.g. adequate transcript number.

Cell Type Area normalised by Trephine Area

For each trephine, I have calculated the total area of cellular marrow, with and without adipocytes. Bone area was not included when calculating total trephine area, and I only included cells and adipocytes which were in spatially preserved regions.

I have then calculated the percentage of the marrow area (cellular marrow only (Fig. 4.5), or cellular marrow with adipocytes (Fig. 4.6)), taken up by each cell type.

To calculate significance I performed MannWhitneyU tests between each combination of conditions for each cell type, and corrected for multiple hypothesis testing with Benjamini-Hochberg. I have considered results significant if the p value was less than 0.05.

Using this method, there were no significant differences in the percentage of trephine area taken up by each cell type between any conditions, where trephine area is cellular marrow only (Fig. 4.5). When normalising by trephine area which also includes fat (cellular marrow and

adipocytes), only the percentage of the trephine area occupied by adipocytes between ET and MF was found to be significantly different (adjusted p value = 0.017) (Fig. 4.6).

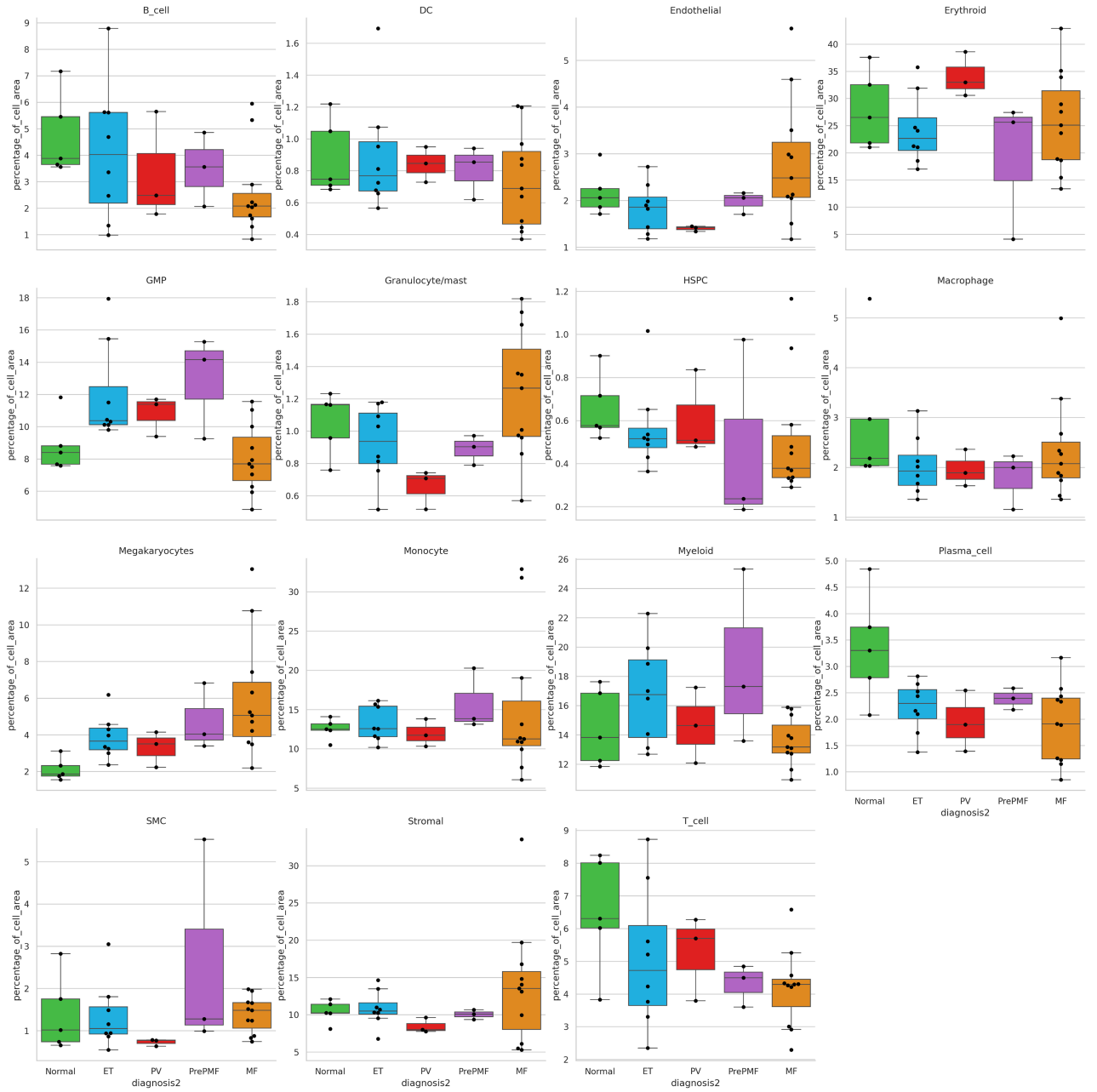


Figure 4.5: Cell Type Percentage of Cellular Marrow at level 3 annotation. Y axis indicates the percentage of the trephine (total cell area only, without fat) taken up by each cell type. Dots indicate patients within each condition, boxplots are coloured by condition.

Cell Type Count normalised by Trephine Area

Next, rather than the area of cells normalised by trephine area, I proceeded to normalise the number of cells by trephine area. As above, I calculated two values, either the number of cells normalised by only cellular marrow (without adipocytes) or the number of cells normalised by

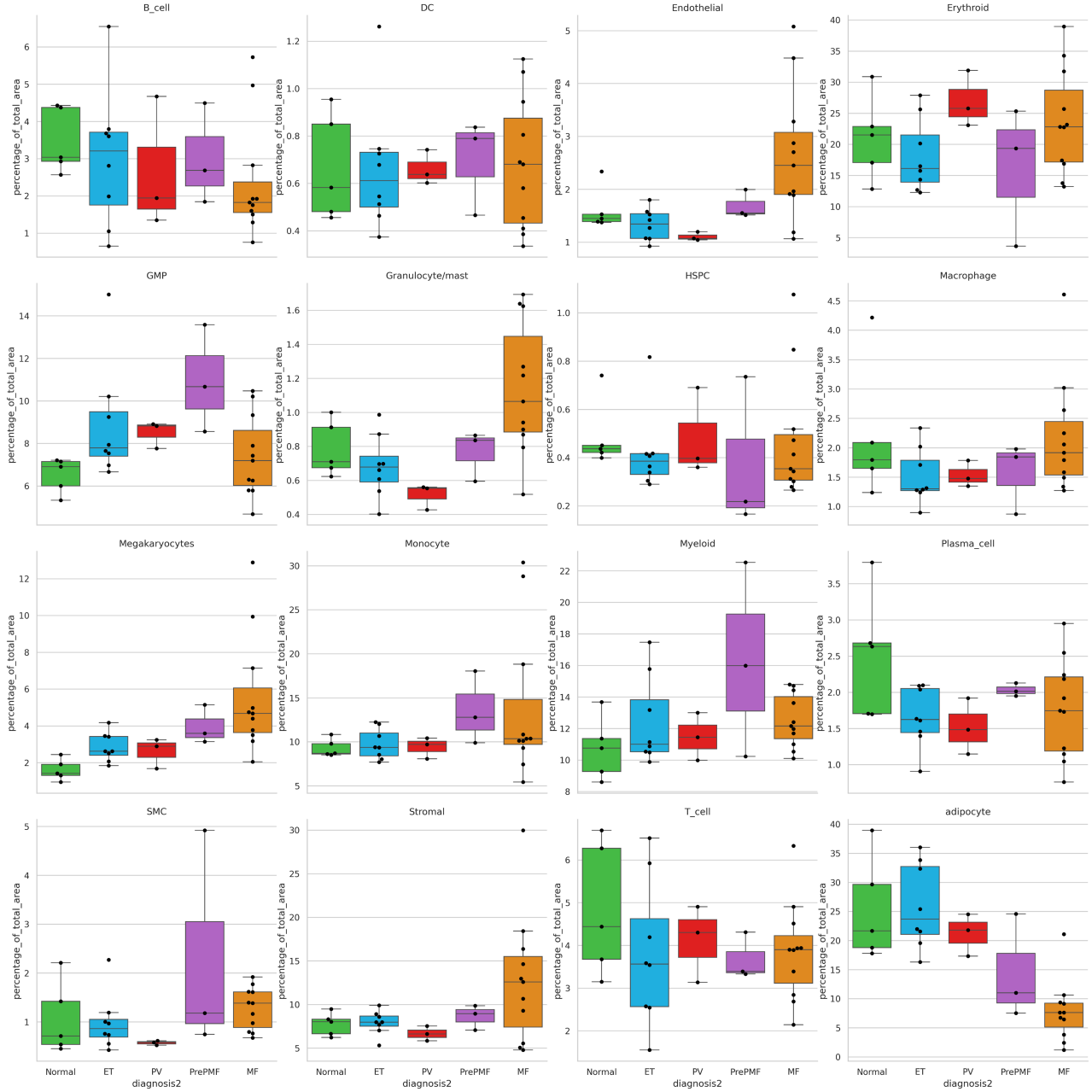


Figure 4.6: Cell Type Percentage of Cellular and Fatty Marrow at level 3 annotation. Y axis indicates the percentage of the trephine (total cell and adipocyte area) taken up by each cell type. Dots indicate patients within each condition, boxplots are coloured by condition.

Cell Type	Condition1	Condition2	Median Difference	p value	adj. p value
Endothelial	ET	MF	154.324863	0.000185	0.012210
Endothelial	Normal	MF	165.970243	0.000458	0.012210
GMP	Normal	MF	358.259394	0.001832	0.036630
Granulocyte/mast	ET	MF	53.620219	0.000185	0.012210
Granulocyte/mast	Normal	MF	54.454408	0.000458	0.012210
Megakaryocytes	Normal	MF	245.642138	0.000458	0.012210
Stromal	Normal	MF	622.888366	0.000458	0.012210
adipocyte	ET	MF	-104.915292	0.001191	0.027217

Table 4.1: Significant changes (adjusted p value < 0.05) between the number of cells per mm2 of trephine area (cellular and fatty marrow) between conditions, after performed a MannWhitneyU test for each cell type between each pair of condition and correcting for multiple hypothesis testing with Benjamini-Hochberg method.

cellular marrow with adipocytes included. This value can be interpreted as the number of cells per unit area. As the unit area is calculated in pixel space, where 1 pixel is 0.2125 microns, I converted the number of cells per pixels squared to the number of cells per mm2 to ensure the subsequent numbers and plots were more interpretable.

Again, to calculate significance I performed MannWhitneyU tests between each combination of conditions for each cell type, and corrected for multiple hypothesis testing with Benjamini-Hochberg. I have considered results significant if the p value was less than 0.05.

Using this method, I found no significant differences between the number of cells per mm2 between conditions, when trephine area was considered to be only cellular marrow (without adipocytes) (Fig. 4.7). However, when I calculated the number of cells per mm2 between conditions, when trephine area was considered to be cellular marrow and adipocytes, I found significant changes between 6 cell types (Fig. 4.8, Table 4.1). These significant changes were between endothelial cells, GMP, Granulocytes/mast, megakaryocytes and stromal cells between Normal and MF, and endothelial cells, granulocytes/mast and adipocytes between ET and MF.

It is reassuring that similar significant changes in cell types between conditions are found between Fig. 4.8, number of cells per mm2 of cellular and fatty marrow (MannWhitneyU), and Fig. 4.2, number of cells normalised by number of all cells (edgeR), when using different methods for quantification. For example, significant changes are observed in adipocytes, endothelial cells, granulocytes/masts, megakaryocytes, and stromal cells in both (Fig. 4.2 and Fig. 4.8, Table 4.1). A possible explanation for differences may be that adipocytes are significantly larger than all other cell types, and we know from Fig. 4.2 and Fig. 4.6 that there are changes in

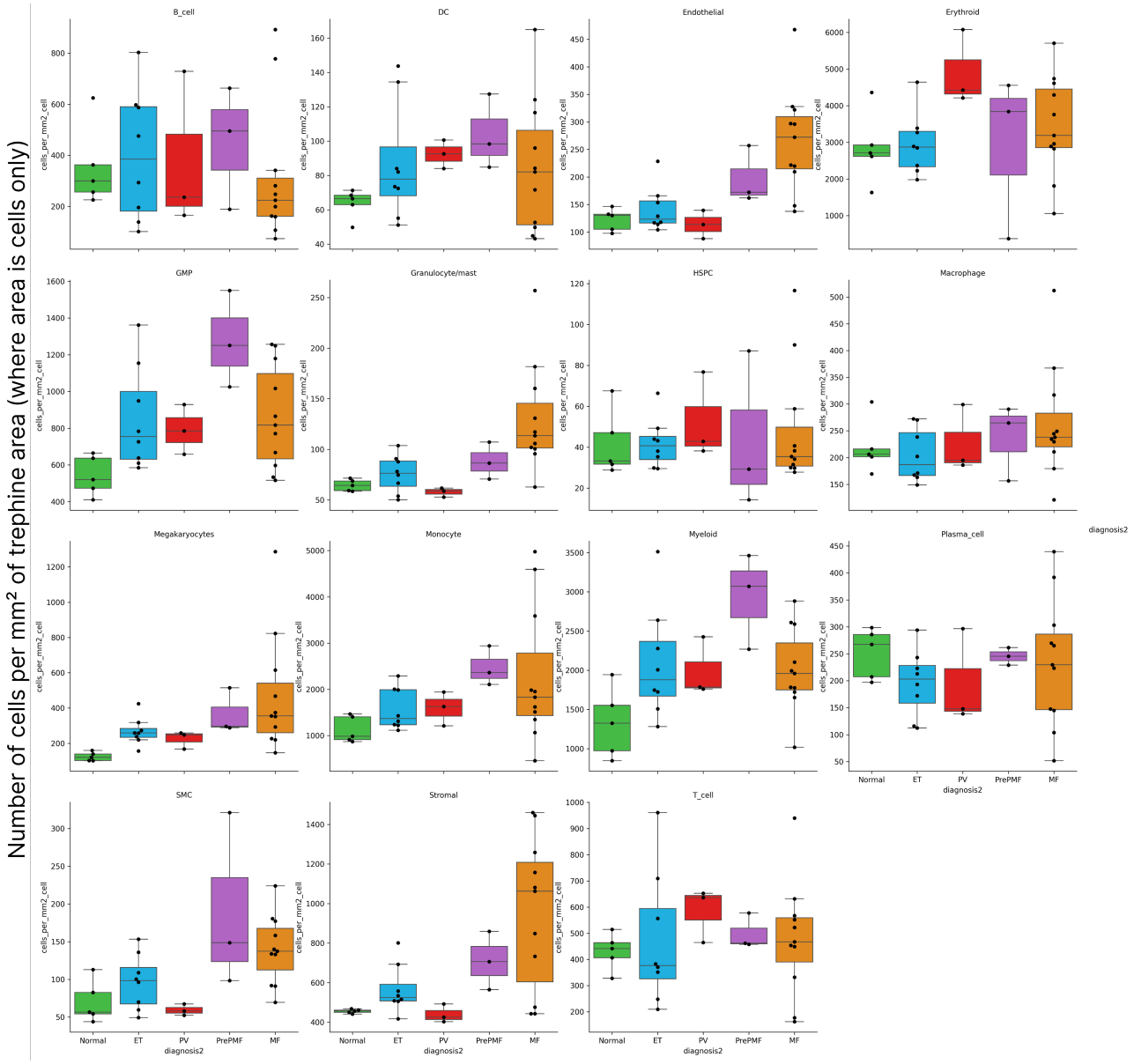


Figure 4.7: Number of each cell type per mm2 of cellular marrow at level 3. Y axis indicates the number of cells per mm2 of the trephine (total cell area) for each cell type. Dots indicate patients within each condition, boxplots are coloured by condition.

adipocyte number and area particularly between Normal/ET and MF. This may influence some of the changes between number of cells per mm2 between Normal/ET and MF (where we are considering trephine area as cellular and fatty marrow). We also know MF is hypercellular (more cells in the same unit of space). Normalising by trephine area (cellular and fatty marrow), when there are known differences in adipocytes between conditions, could exaggerate trends in cell type differences which already exist, but which weren't significant when evaluating number of cells normalised by number of all cells (Fig. 4.2), or number of cells per mm2 of cellular marrow only (Fig. 4.7), or when the differing area of cell types was taken into account (Fig. 4.5,

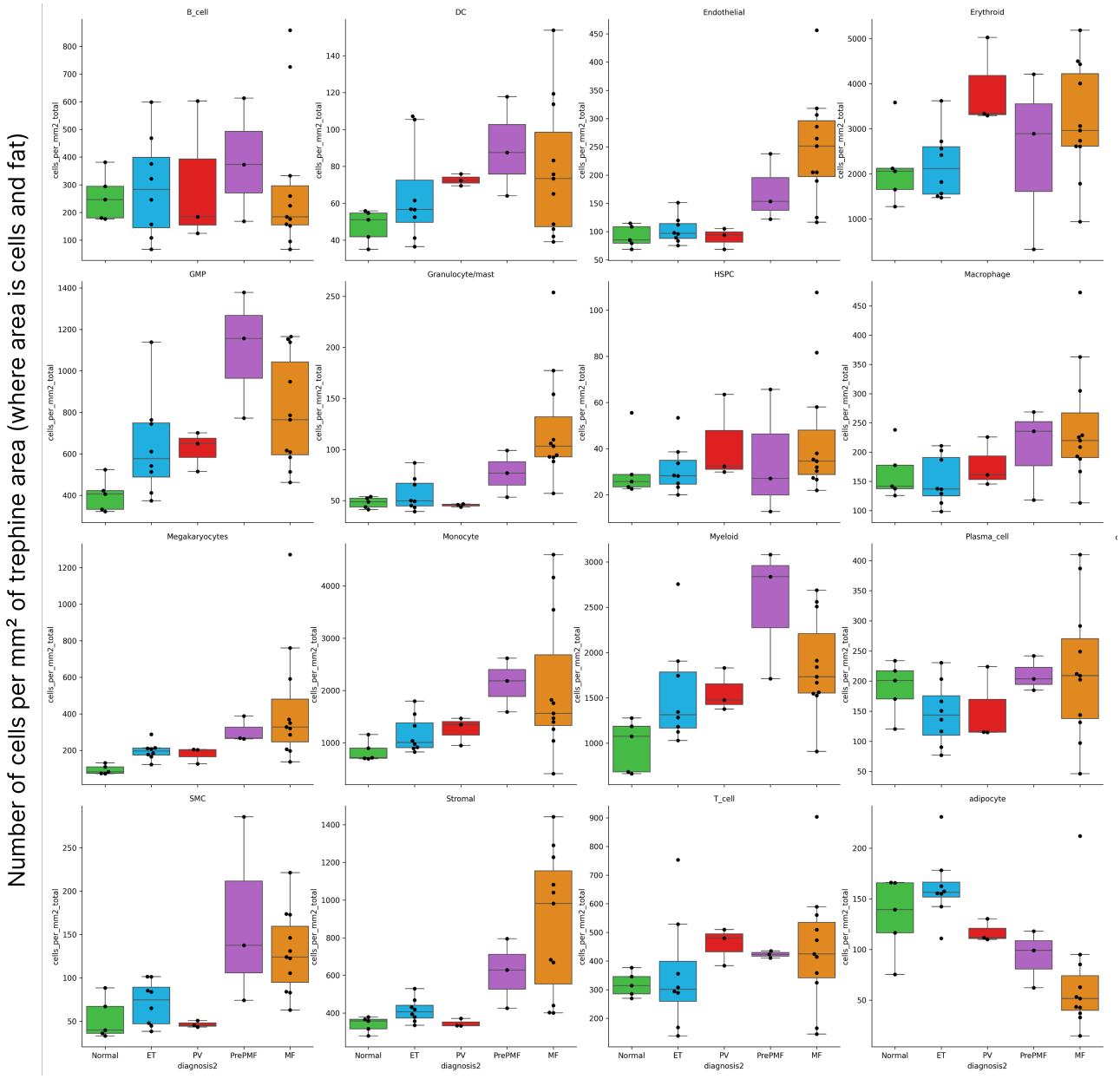


Figure 4.8: Number of each cell type per mm2 of Cellular and Fatty Marrow at level 3 annotation. Y axis indicates the number of cells per mm2 of the trephine (total cell and adipocyte area) taken up by each cell type. Dots indicate patients within each condition, boxplots are coloured by condition.

Fig. 4.6). Additional differences may be due to the different sensitivity of statistical tests used between Fig. 4.2(edgeR) and Fig. 4.8 (MannWhitneyU).

4.3.4 Discussion

Differential abundance analysis revealed two main findings. First, megakaryocytes were significantly increased in MF compared to healthy bone marrow, consistent with previous reports [54]. This increase was largely driven by megakaryocytes not distinguishable on H&E.

While overall megakaryocyte abundance was not significantly different between MF, ET, and PV, phenotype 8, the largest group, marked by atypical nuclei, was significantly increased in both MF and ET relative to healthy donors. In every sample, we identified numerous megakaryocytes that were not morphologically visible (Fig. 4.56a), comprising the majority of the megakaryocyte population. This has important implications for clinical interpretation of H&E slides (Action Point 2). Future work should assess whether these "hidden" megakaryocytes share spatial distributions with morphologically defined ones and explore their contribution to marrow organisation. Phenotype 8, in addition, may play a key role in MPNs and warrants correlation with gene expression to uncover its functional significance.

Second, we found a marked increase in Osteo-MSCs in MF compared to all other conditions. This may reflect their involvement in fibrosis, a defining feature of MF. Although our panel lacks the resolution to fully distinguish osteoblasts and osteoclasts, previous work has implicated a CD56+ osteolineage subset in fibrotic processes [56]. Spatial analysis of Osteo-MSCs, especially across fibrosis grades, will help clarify their role (Fig. 4.4).

4.3.5 Conclusion

These findings have two key implications. First, they highlight that not all megakaryocytes are morphologically visible, which may alter how we interpret changes observed in archival H&E-stained trephines. Second, the increase in Osteo-MSCs in MF suggests a potential role in fibrosis, consistent with existing literature on MF pathogenesis.

4.4 Global cellular expression in the Bone Marrow in MPNs

4.4.1 Introduction

I next investigated global gene expression changes in MPNs. As with abundance, analysing the whole bone marrow highlights genes markedly altered between conditions, which can then be explored spatially. For example, given the chronic inflammatory phenotype of MPNs, identifying spatial regions of inflammation becomes a meaningful follow-up. I focused on genes differentially

expressed in each cell type in MF compared to healthy donors. This global landscape provides a baseline, later identifying cells lacking MF-associated expression patterns may indicate preserved, unaffected regions within the diseased tissue.

4.4.2 Methodology

Choice of Statistical Test

I treated the dataset similarly to a scRNA-seq experiment when analysing the whole trephine. Like scRNA-seq, our data is zero-inflated, as not all genes are expressed in every cell, a common challenge in scRNA-seq experiments. Differential expression (DE) methods for scRNA-seq fall into two main categories: cell-level and pseudobulk analyses. Pseudobulk approaches are generally more robust, as they avoid treating cells as independent replicates and can incorporate sample-level variability. Both approaches outperform naive methods that ignore random effects such as patient ID.

For trephine-level DE analysis, I used DESeq2 on pseudobulked data, summing counts per gene across diagnosis, sample ID, and cell type (level 3 with stromal refinement). Sum aggregation closely mimics a bulk RNA-seq experiment, allowing DESeq2 to apply its internal library size normalization. Pairwise comparisons were performed using the Wald test, and genes with a Benjamini-Hochberg adjusted p-value < 0.05 were considered significant.

4.4.3 Results

Given the large number of genes, conditions, and cell types, summarising the findings is not straightforward. I therefore approached the analysis in three ways: (1) counting the total number of differentially expressed genes (DEGs) between conditions, (2) identifying the top DEGs across all cell types, and (3) determining which cell types show the most DEGs. To illustrate how gene expression changes at the cell-type level, I then focused on DEGs between Normal and MF within each cell type.

MF shows widespread gene expression changes compared to Normal, ET, and PV

MF shows substantial differential expression compared to all other subtypes, particularly Normal, followed by ET and then PV. The smaller difference between MF and PV may reflect the inclusion of two Post-PV MF samples, though this classification depends on full clinical history of the samples. Notably, the other subtypes appear strikingly similar to one another, even before considering specific genes or cell types, especially ET and PrePMF, as well as Normal and PV (Fig. 4.9).

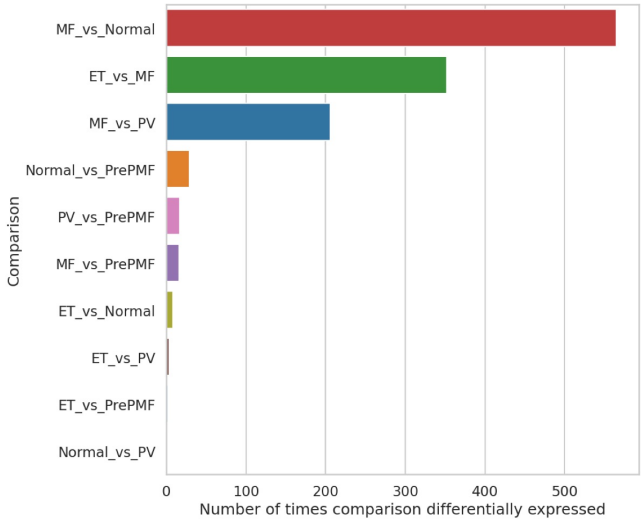


Figure 4.9: Cumulative total of the number times a gene is differentially expressed across all cell types between the conditions. The count is 1 for each occurence and does not relate to the level of differential expression - a small log fold change is represented the same as a large log fold change in expression

Breaking this down by gene, most differential expression occurs between MF, and Normal, ET, and PV. Many genes are differentially expressed across multiple cell types in each comparison, as shown by the horizontal continuity of each colour band within the bars. The top genes consistently altered across cell types in MF compared to ET, Normal, and PV include: ALDH1A3, PLIN4, ADIPOQ, PRG4, THBS2, THY1, VWF, PDGFRB, COL5A2, and TNC (Fig. 4.10)

Stromal cells, endothelial cells, megakaryocytes and HSPCs have the most altered gene expression across comparisons

We can also identify the cell types most affected across comparisons. Stromal cells, endothelial cells, megakaryocytes, and HSPCs show the greatest number of differentially expressed genes. These cell types are all implicated in fibrosis (Fig. 4.11).

Megakaryocytes and Endothelial cells have increased expression of an endothelial-meschenchymal transition gene in MF compared to healthy donors

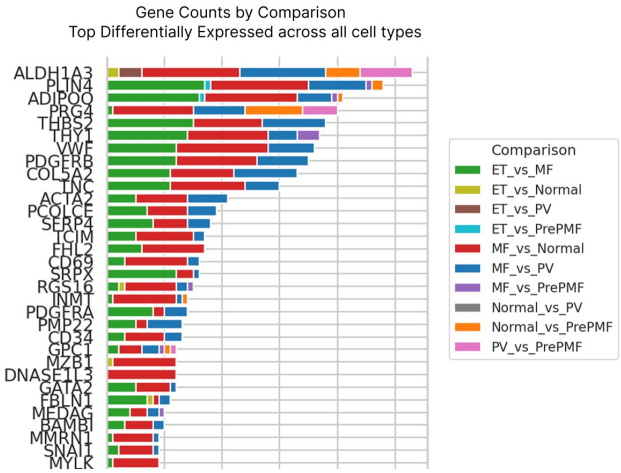


Figure 4.10: Number of times gene is differentially expressed across all cell types and comparisons, coloured by the number of times per comparison. Displaying only the top genes

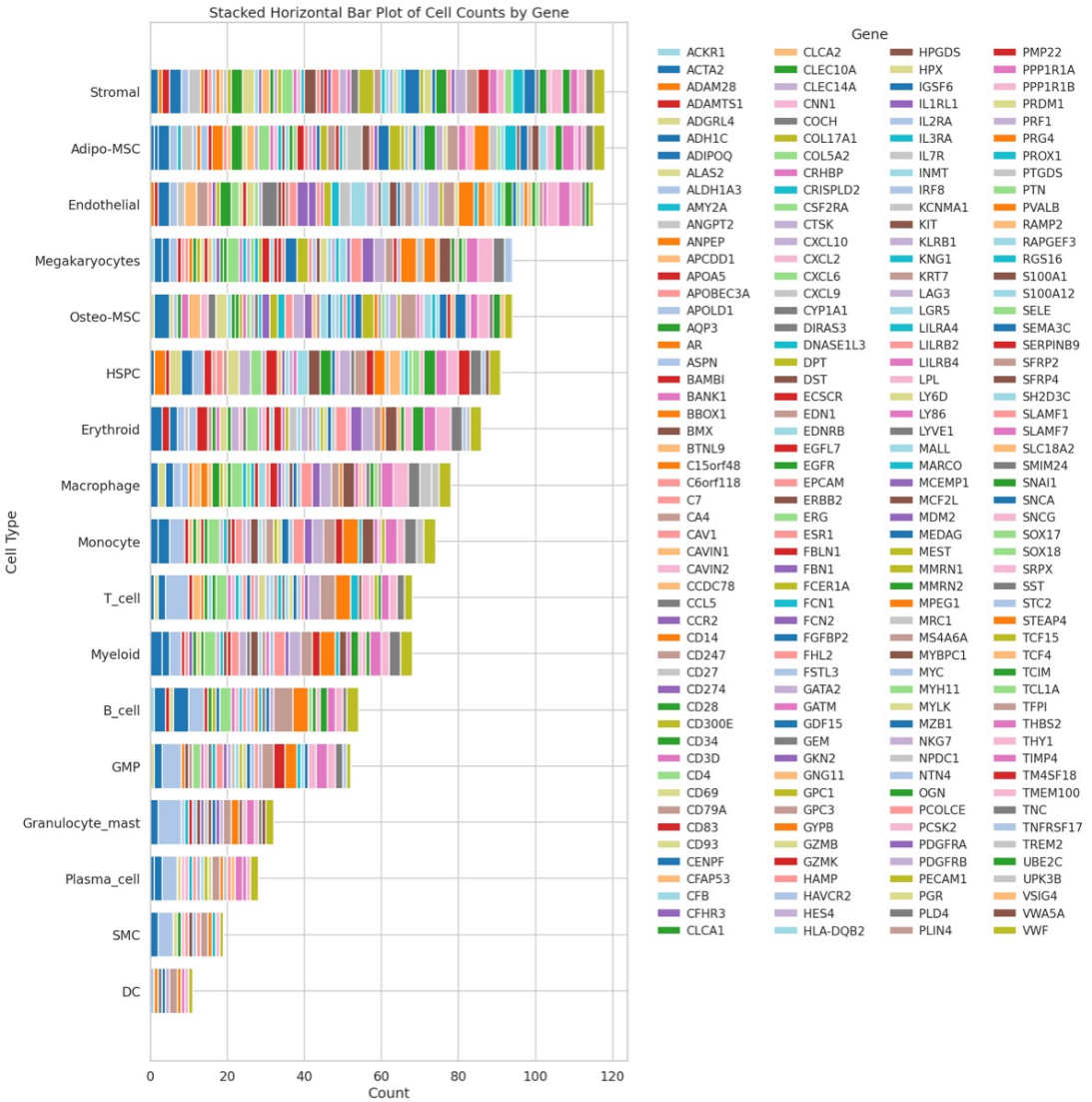
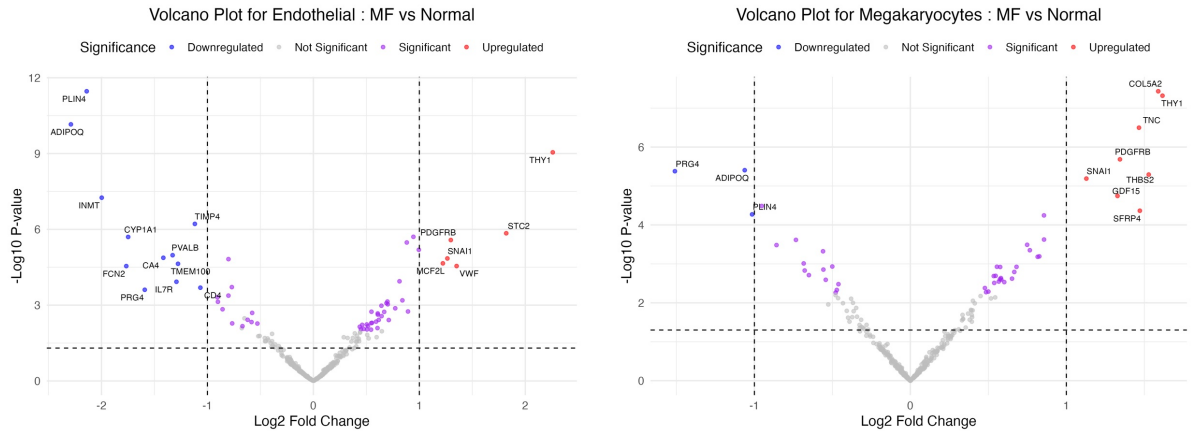


Figure 4.11: Count of differentially expressed genes across all cells for all comparisons within each cell type. Larger colour blocks indicate a gene that appears in multiple comparisons.

I focused on MF because other MPN subtypes can progress to myelofibrosis, which is associated with poorer prognosis. Identifying genes linked to fibrosis may help position other subtypes along the disease continuum. To determine whether these genes are up or downregulated in MF, I performed differential expression analysis for each cell type individually. In Fi. 4.12, I show volcano plots comparing Normal and MF, the most distinct conditions, highlighting genes with significant differential expression and a log fold change greater than 1.



(a) Endothelial **(b)** Megakaryocyte

Figure 4.12: Differential gene expression in Endothelial and Megakaryocyte populations between normal and MF. Red indicates genes upregulated in MF.

Endothelial cells are increased in MF compared to ET and PV and show significant expression changes, suggesting they may contribute to fibrosis. These changes include upregulation of SNAI1, linked to endothelial-mesenchymal transition [101], MCF2L, associated with osteogenic differentiation and fibrosis [49, 102], and THY1, which promotes myeloid cell adhesion and migration during inflammation [103]. VWF, a marker of endothelial aging, is also upregulated, while IL7R is downregulated, consistent with JAK2-mutant endothelial cells [104].

Megakaryocytes, also increased in MF, similarly show elevated SNAI1 expression, suggesting a role in mesenchymal transition. SNAI1 has been linked to MPN development and predisposition to leukemia in aged mice [105], and in AML is associated with poor prognosis, impaired differentiation, increased progenitor proliferation, and HSC expansion with GMP bias. Previous work also shows epithelial-to-mesenchymal gene expression in MF megakaryocytes [106].

GDF15 is also upregulated in MF megakaryocytes. This cytokine promotes MSC growth and osteoblastic differentiation [107], and in AML inhibits adipogenesis, improves chemotherapy

response, and reduces relapse [108]. This aligns with the reduced adipocyte population in MF.

Interestingly, MF megakaryocytes also overexpress SFRP4, a Wnt inhibitor. While Wnt inhibition may reduce fibrosis [56], SFRP4 may also suppress osteoblast proliferation and increase trabecular bone mass [109]. SFRP4 is also overexpressed in monocytes and erythroid cells in MF.

4.4.4 Discussion

MF shows markedly altered gene expression compared to healthy donors, ET, and PV (Fig. 4.9). Many genes are differentially expressed across multiple subtypes and cell types (Fig. 4.10). The top genes altered in several cell types, ALDH1A3, PLIN4, ADIPOQ, PRG4, THBS2, THY1, VWF, PDGFRB, COL5A2, and TNC, are largely associated with fibrosis and megakaryocytic changes, rather than inflammation. This pattern likely reflects our panel's design. The upregulation of fibrotic genes in MF is reassuring and supports the validity of our findings. Notably, THY1/CD90 is upregulated in multiple cell types in MF. THY1 marks fibroblasts, and a recent study identified a human-specific THY1+ MSC population that supports hematopoiesis via IL7, SCF, and CXCL12 [34]. These MSCs are more frequent in AML than in healthy marrow and localize centrally near blast cells. In mice, THY1 promotes MSC osteogenic differentiation, and its absence reduces Wnt ligand expression [103]. Wnt signalling is linked to fibrosis and osteogenesis in MPN [56], and THY1+ profibrotic stromal cells are absent in younger individuals [71]. That known fibrotic markers are upregulated despite our panel's lower resolution lends credibility to our results.

The increased abundance of megakaryocytes and endothelial cells in MF, and their shared upregulation of SNAI1, a gene involved in endothelial-to-mesenchymal transition (EndMT), suggest they may jointly contribute to this process. EndMT has been previously linked to megakaryocytes in MF [106], though further confirmation with a broader panel is needed. Understanding this process is important given emerging insights into EndMT's role in fibrosis. Banjanin et al. [56] used fate-tracing in mice to show that inflammation causes CAR cells to move out from the perivascular niche, reprogramming them into proinflammatory cells. As CAR cells vacate the niche, endothelial cells are exposed to inflammation and undergo EndMT. Sivaraj et al. [49] demonstrated that Lats2 inactivation or YAP1 overexpression induces EndMT, leading

to fibrosis and osteosclerosis. They found that SRF upregulation is essential for this transition, and that its deletion rescues the fibrotic phenotype. In the bone marrow, endothelial cells also regulate osteoprogenitors and osteoblasts via growth factors and mechanical cues. EndMT alters their morphology, gene expression, and vascular function. Sivaraj et al. [49] further proposed that elevated TGF-beta links endothelial changes to hematopoietic cells. TGF-beta is known to be upregulated in MF megakaryocytes and monocytes. It also drives microvessel loss and activates EndMT and SRF signaling, although our panel does not include TGF-beta to investigate this. Both endothelial cells and megakaryocytes show upregulation of genes involved in osteogenic differentiation and SNAI1 in our dataset, suggesting a potential role for megakaryocytes in driving EndMT. EndMT affects around 30% of endothelial cells in bone marrow microvessels in MF [110], and megakaryocyte supernatants have been shown to induce EndMT in vitro. Megakaryocytes are often found adjacent to microvessels undergoing EndMT in MF [111].

Finally, we observed a paradoxical upregulation of SFRP4 in megakaryocytes, monocytes, and erythroid cells. While SFRP4 is a canonical Wnt inhibitor and has been shown to suppress osteoblast proliferation [109], it is also linked to elevated ECM gene expression in fibroblasts in prostate cancer [112].

4.4.5 Conclusions

MF shows a markedly altered gene expression profile compared to Normal, ET, and PV. One globally upregulated gene is Thy1, which is associated with fibrosis, osteogenic differentiation, and marks a human MSC subset not observed in mice.

In MF, both endothelial cells and megakaryocytes upregulate genes linked to endothelial-to-mesenchymal transition (EndMT), a process implicated in fibrosis and osteosclerosis. While our panel lacks the resolution to confirm this conclusively, the findings support a previously proposed role for megakaryocytes in driving EndMT [49, 106].

These results align well with existing literature. A key next step is to investigate the spatial distribution of these genes and the cell types expressing them.

4.5 The Unsupervised Bone Marrow Architecture

4.5.1 Introduction

I then moved from global to spatially resolved analysis. As described in the background chapter, a key challenge is the Modifiable Areal Unit Problem (MAUP), where different ways of dividing tissue can yield conflicting results. In some tissues, such as tumours, regions of interest are morphologically distinct on H&E, allowing clear delineation for spatial analyses. Including surrounding tissue without considering boundaries can obscure interactions and introduce noise. Pathologist-defined annotations help avoid such pitfalls. However, regions like HSPC niches lack visible boundaries. In these cases, spatial domains must be defined using features such as cell type, gene expression, and local organization. To address this, I applied unsupervised cellular neighbourhood detection to identify regions with similar cellular composition.

In bone marrow, certain morphological landmarks have traditionally guided the definition of functional regions, such as proximity to trabeculae (endosteal niche), blood vessels (perivascular niche), or fibrotic areas. Other regions may include erythroblast islands or lymphoid aggregates. This section explores whether functionally distinct bone marrow regions, beyond those visible on H&E, can be revealed by unsupervised analysis of cell composition and expression. If known structures (e.g. endosteal niches) emerge from these data alone, it suggests they represent meaningful biological domains.

4.5.2 Methodology

Defining unsupervised cellular neighbourhoods in the bone marrow using CellCharter

As outlined in the background chapter, there are multiple strategies for identifying cellular neighbourhoods in tissue. For bone marrow, I used CellCharter to define spatial domains.

CellCharter is scalable across multiple samples and uses a graph-based approach. I used the voronoi neighbour graph (described in the previous chapter) to define cell connectivity. Gene expression was log-normalised, and dimensionality reduction with batch correction was performed using the scVI model (scvi.model.SCVI). The model was trained with sample ID as the batch variable and cell annotations as labels, resulting in a 128-dimensional latent representation per

cell. To capture spatial context, CellCharter concatenates each cell's latent vector, from gene expression and cell annotations, with those of its neighbourhood up to k hops. For k = 3, this results in four vectors: the cell itself, and the averaged latent vectors of its 1st-, 2nd-, and 3rd-hop neighbours. This representation is constructed for every cell. These combined vectors are then clustered into a user-defined number of spatial domains using a Gaussian Mixture Model. To select the number of clusters, CellCharter computes a stability score across cluster numbers n by running the pipeline 10 times and comparing clustering consistency using the average Fowlkes-Mallows Index between n–1, n, and n+1 [113].

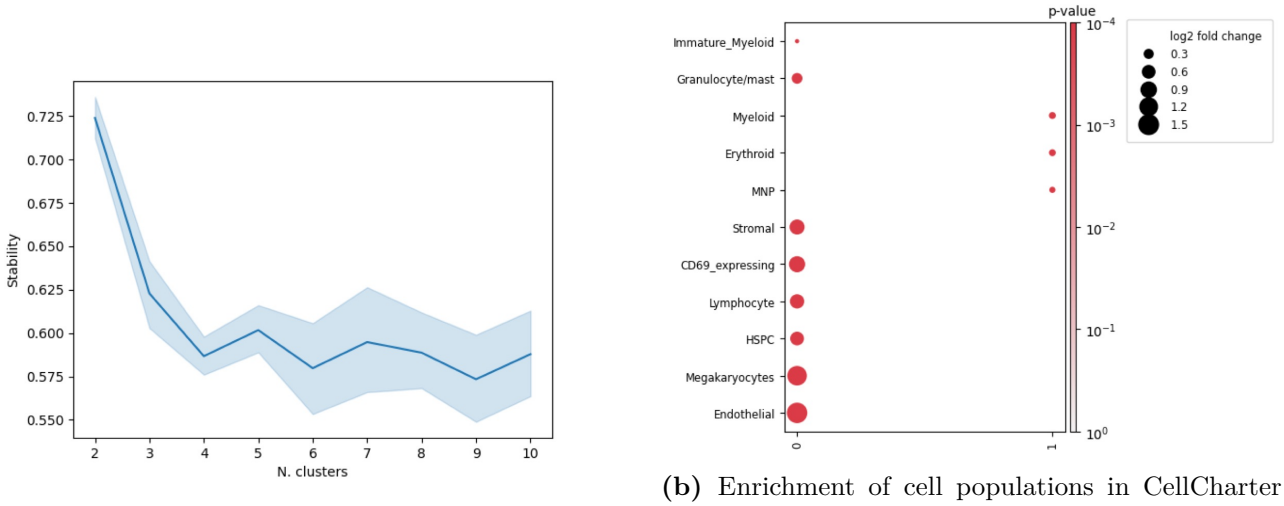
Assessing the Stability of CellCharter cellular neighbourhoods

To select the number of clusters, I ran CellCharter with n=2 to n=10 clusters and evaluated cluster stability (Fig. 4.13a). The most stable solution was at n=2. Interpreting CellCharter's cellular neighbourhoods requires post hoc analysis. A primary step is testing for enrichment of each cell type in each domain. To do this, I permuted cell type labels 100 times across the whole trephine to create a null distribution of expected proportions. Comparing observed vs expected proportions gives a log fold-change and p-value (Fig. 4.13b). I also visualised domain composition across cell types in representative Normal and MF samples (Fig. 4.14).

This suggests one compartment consists primarily of Erythroid, Myeloid, and MNP cells, while the other includes all remaining cell types. Although this clustering is stable, two groups seem too coarse to meaningfully capture tissue complexity, especially given that stromal and lymphoid cells likely play distinct roles.

Choosing the Number of CellCharter Cellular Neighbourhoods

Instead, I chose to proceed with 10 CNs, as stability was similar across most values of n. This decision was also guided by biological expectations, specifically, the presence of an endosteal region, a key feature of bone marrow architecture. Clustering at 7 and 10 corresponded well to endosteal regions in some samples. Although this endosteal CN also appeared in the 7 CN model, the region also included additional cells (e.g., the brown CN in Fig. 4.17) with distinct identities, suggesting that the CNs were hierarchically nested. Choosing 10 CNs allowed better separation of biologically meaningful compartments.



(a) Stability of cell clusters found by CellCharter

(b) Enrichment of cell populations in CellCharter clusters

Figure 4.13: (a) Stability of cluster numbers ($n = 2$ to 10) from CellCharter with $k = 3$ hops, using level 3 annotations. (b) Cell population enrichment across CellCharter domains at annotation level 5, based on 100 permutations.

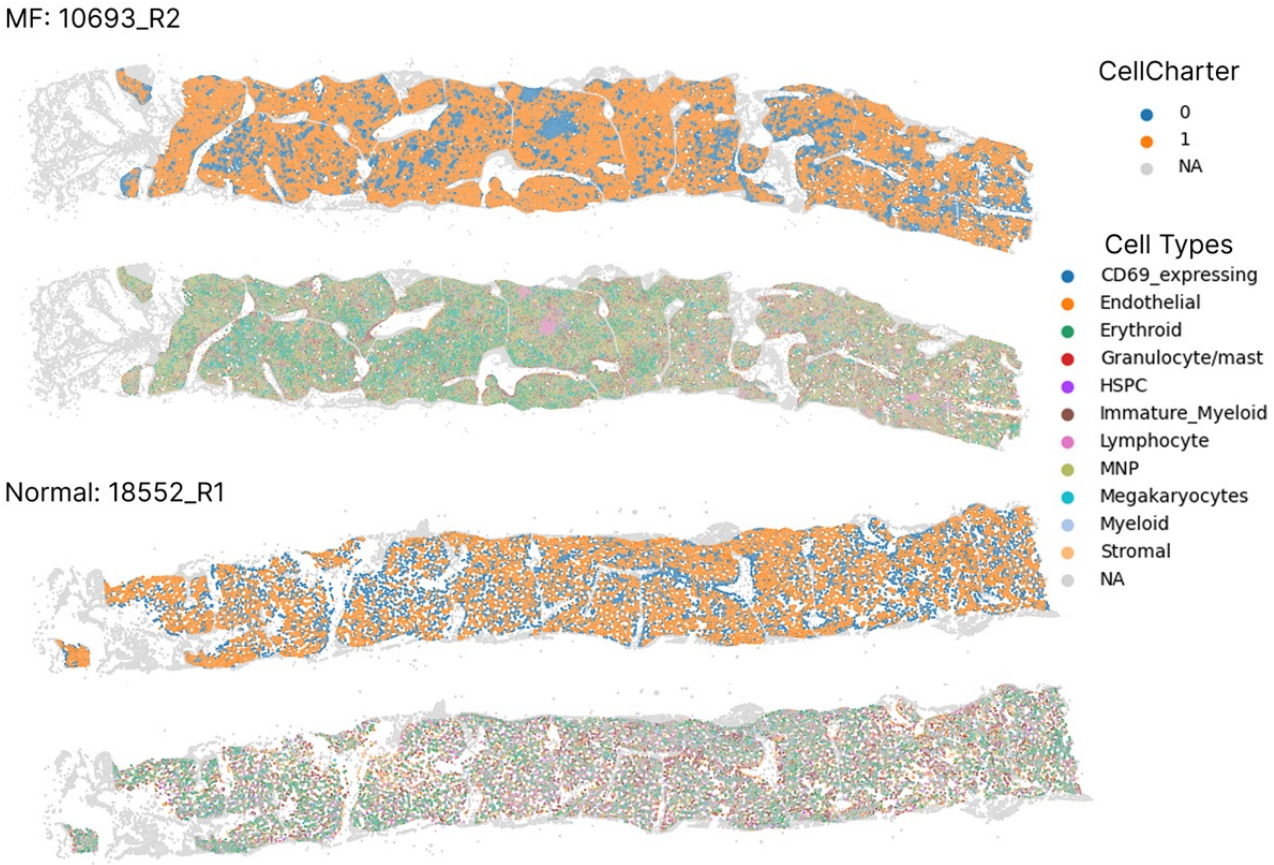


Figure 4.14: CellCharter clustering at 3-hop with log normalization of genes across all samples. Automatic stability selection was used to choose the number of clusters. Shown are exemplar MF and Normal samples with level 5 cell type annotations and the corresponding samples with cellcharter clusters at n=2.

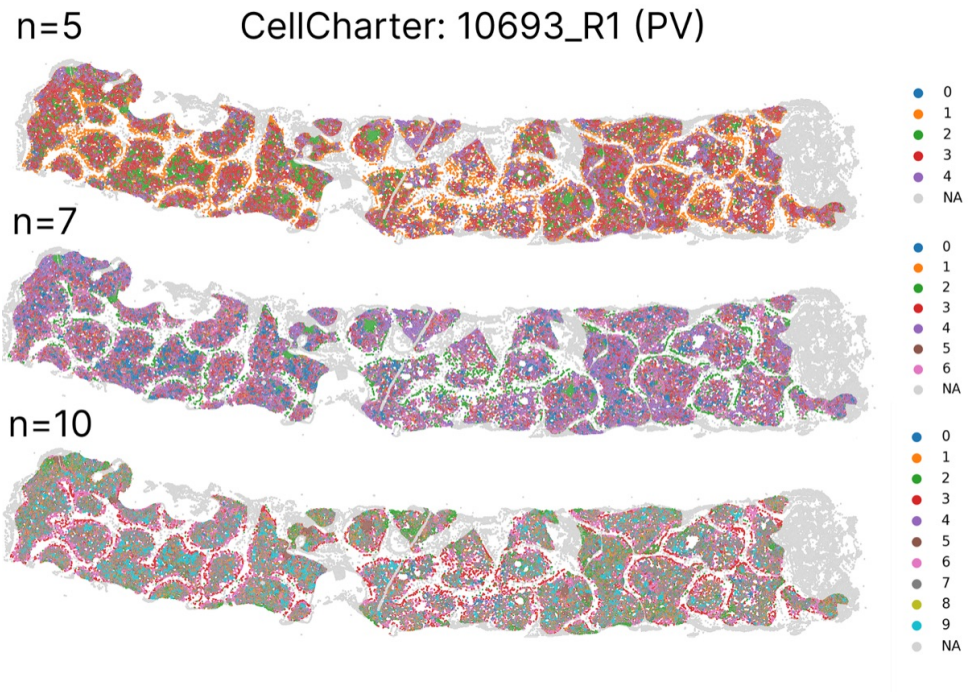


Figure 4.15: CellCharter clustering at 3-hop with 5, 7, and 10 clusters across the same PV sample. A distinct endosteal and peritrabecular niche is only evident in the n = 10 clustering.

Defining the Endosteal and Peritrabcular Regions

Given prior knowledge of bone marrow structure, we defined regions corresponding to the endosteal and peritrabecular niches, then identified the cells within those areas. To calculate each cell's distance from bone, I created two GeoDataFrames: one for bone-segmentation polygons and one for 15 um Voronoi cell polygons. I then used GeoPandas' sjoin_nearest (with a 30,000-unit search limit) to determine the shortest distance from each cell to the nearest bone polygon. Each sample was divided into four spatial compartments: endosteal (0–1 pixels, where cells touch bone), peritrabecular (1–301 pixels, 64 um), intertrabecular (>301 pixels, >64 um), and within_bone (cells entirely enclosed by bone). These definitions were chosen pragmatically with input from expert haematopathologists, for example, the peritrabecular zone roughly corresponds to two fat spaces from the bone surface (Fig. 4.16).

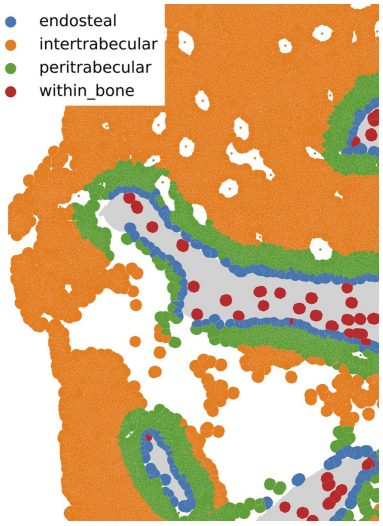


Figure 4.16: Schematic of Bone Regions in a Trephine Section

4.5.3 Results

Cellular neighbourhoods in MF show visually striking differences both within MF and compared to other conditions

After applying cellcharter, we obtained 10 distinct regions across all bone marrow trephines. I have shown an example of one MF sample and one Normal sample below (Fig. 4.17). Even from just two examples, it's clear that the distribution of cellular neighbourhoods differs markedly between MF and Normal samples (Fig. 4.17).

I've included three representative samples from each condition (Fig. 4.18, 4.19, 4.20, 4.21). In healthy marrow, ET, and PV, cellular neighbourhoods appear smaller and more intermixed. In contrast, MF samples show expanded, spatially distinct regions with greater heterogeneity between intertrabecular spaces. Notably, across all samples, the unsupervised analysis consistently identifies two regions that align with the expected locations of the endosteal and peritrabecular niches (Fig. 4.18, 4.19, 4.20, 4.21).

Characterising CellCharter Regions by Cell Enrichment

To characterise each region, I assessed the enrichment of cell annotations using two annotation schemes: level 5 and level 3 with refined stromal annotations (Fig. 4.22). The level 3 scheme with refined stromal annotations proved more informative.

Four clusters show HSPC enrichment: CN3, CN4, CN8, and CN9. CN3 represents an HSPC

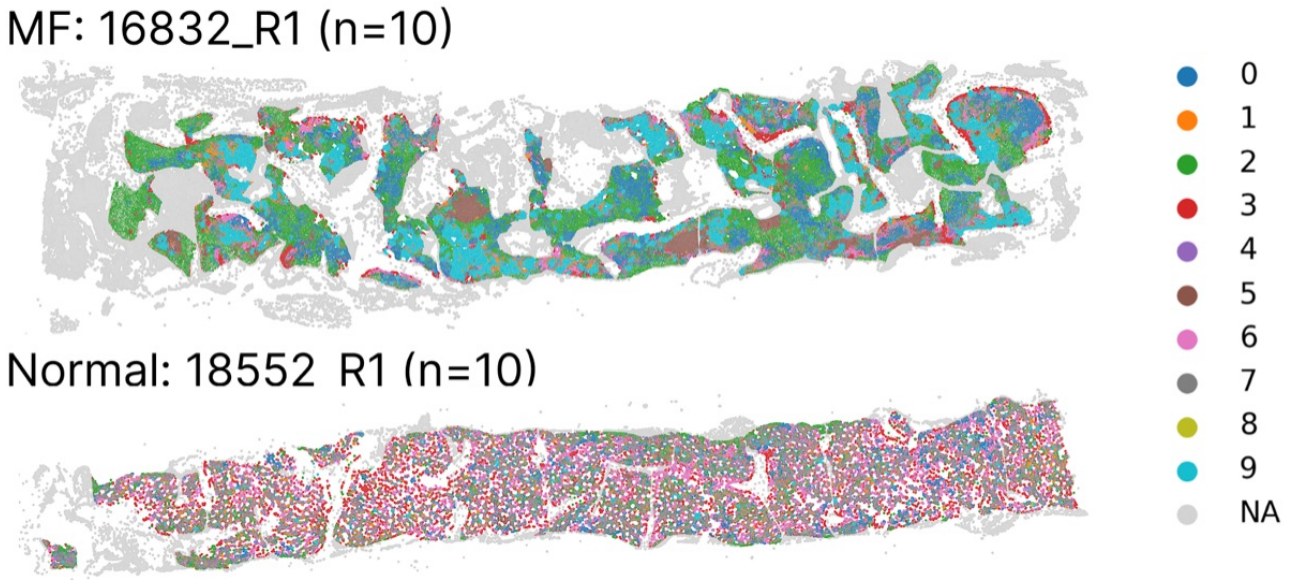


Figure 4.17: CellCharter clustering at 3-hop with log-normalized gene expression across all samples. Clustering was performed with 10 clusters (n = 10). Shown are representative MF and Normal samples for comparison.

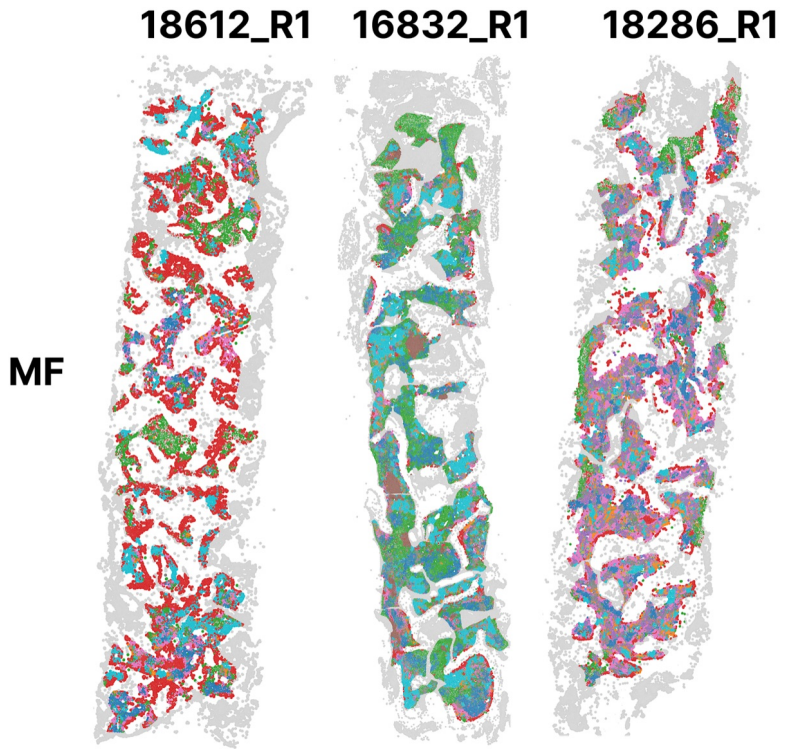


Figure 4.18: CellCharter clustering at 3-hop with 10 clusters (k=3, n=10) across three exemplar MF samples.

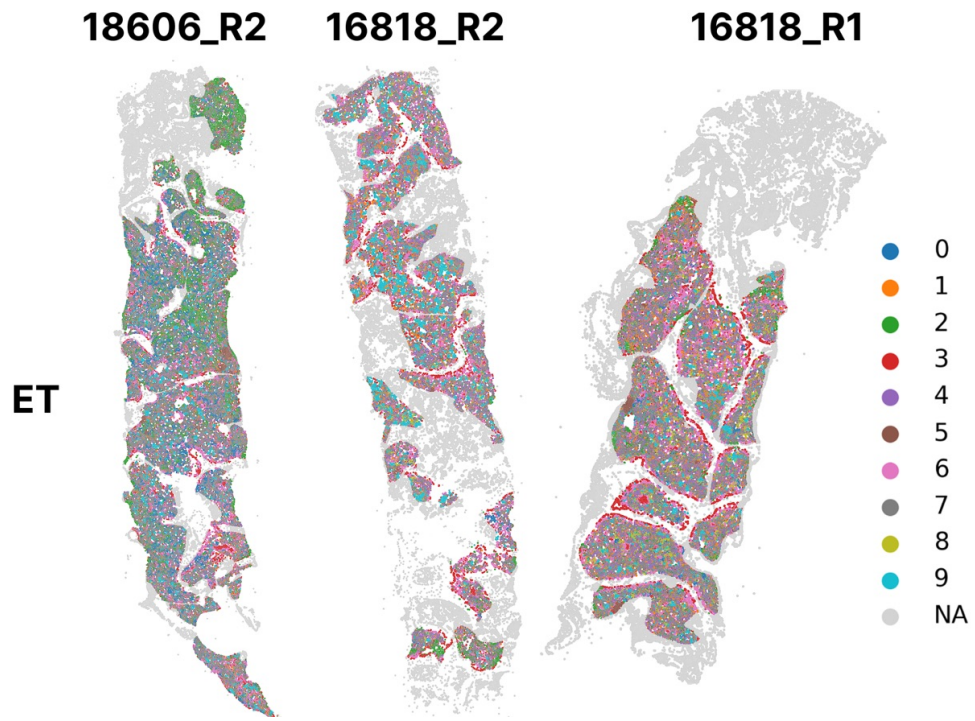


Figure 4.19: CellCharter clustering at 3-hop with 10 clusters (k = 3, n = 10) across three exemplar ET samples (18606_R2, 16818_R2, and 16818_R1).

niche with Endothelial/SMC, Osteo-MSCs, and Plasma/B cells. CN4 is enriched for HSPCs and Erythroid cells. CN8 includes HSPCs and Monocytes, while CN9 contains HSPCs alongside Megakaryocytes and Endothelial/SMC. Among the remaining clusters, CN0 appears to reflect a myeloid trajectory (Stromal, GMP, Myeloid, and Monocytes). CN1 is enriched for Erythroid cells and Macrophages, and CN2 contains Osteo/Stromal cells along with DCs and B cells. CN5 is a Lymphocyte/DC-rich region, CN6 includes GMP and Stromal cells, and CN7 is a non-specific myeloid cluster. Interestingly, DCs are consistently enriched in clusters with lymphocytes, suggesting potential antigen-presenting roles in these regions.

To further investigate regional enrichment of functional states, I used refined annotations, including classical and non-classical monocytes (Fig.4.23).

Most regions remained defined by the same cell groups. However, CN2 showed strong enrichment for non-classical monocytes and was renamed to a "Non-classical Monocyte/Osteo-Stromal/DC-B cell" region. This is notable given the pro-inflammatory role often attributed to non-classical monocytes. CN4 was also redefined as an "HSPC niche with cycling erythroid cells," and CN8 as an "HSPC niche with classical monocytes."

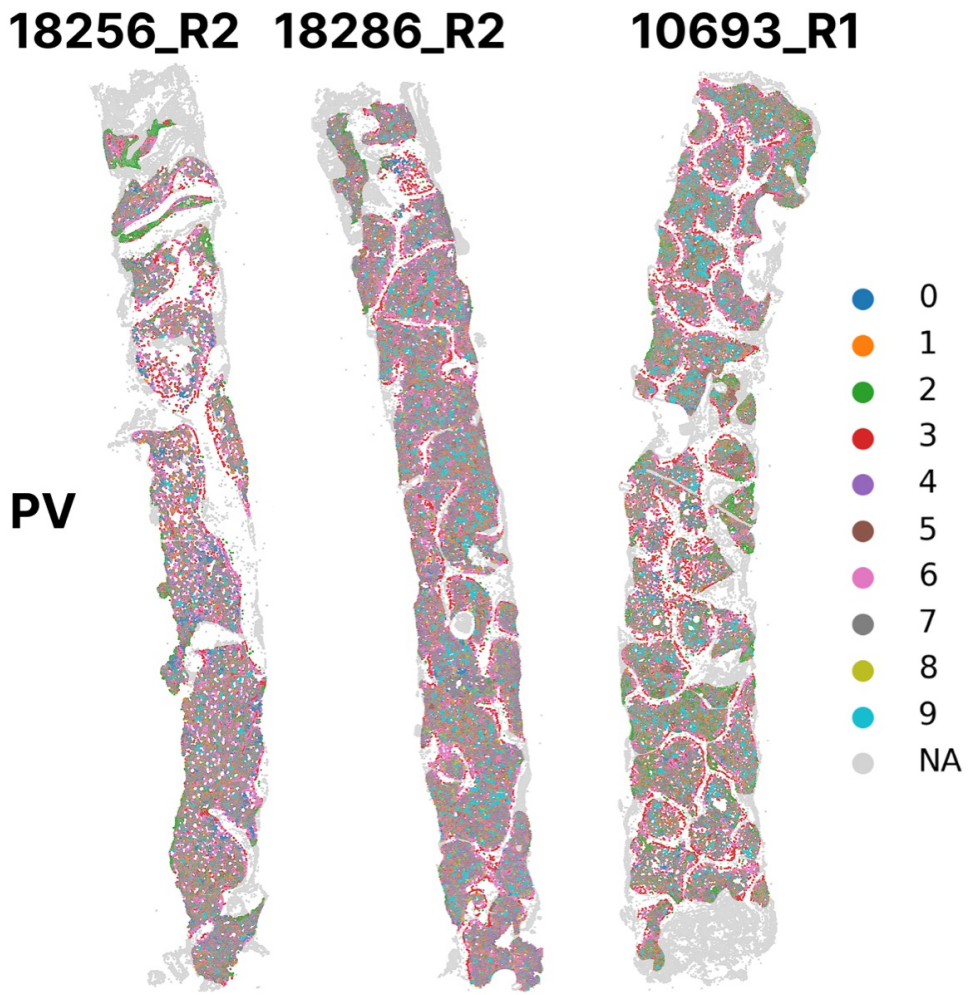


Figure 4.20: CellCharter clustering at 3-hop with 10 clusters (k = 3, n = 10) across three exemplar PV samples (18256_R2, 18286_R2, and 10693_R2).

I chose to proceed with the highest level of annotation resolution, as it allows for later simplification if needed. This yields the following final CellCharter cellular neighbourhoods:

- CN0 is a myeloid trajectory with stromal region.
- CN1 is an Erythroid/Macrophage region.
- CN2 is a Non-classical Monocytes with Osteo/Stromal and DC/B region.
- CN3 is an HSPC niche with Endothelial/SMC, Osteo-MSCs and Plasma/B cells.
- CN4 is an HSPC niche with Cycling Erythroid cells
- CN5 is a Lymphocyte/DC region
- CN6 is a GMP/Stromal region
- CN7 is a non-specific myeloid region
- CN8 is an HSPC niche with Classical Monocytes
- CN9 is an HSPC niche with Megakaryocytes and Endothelial/SMC.

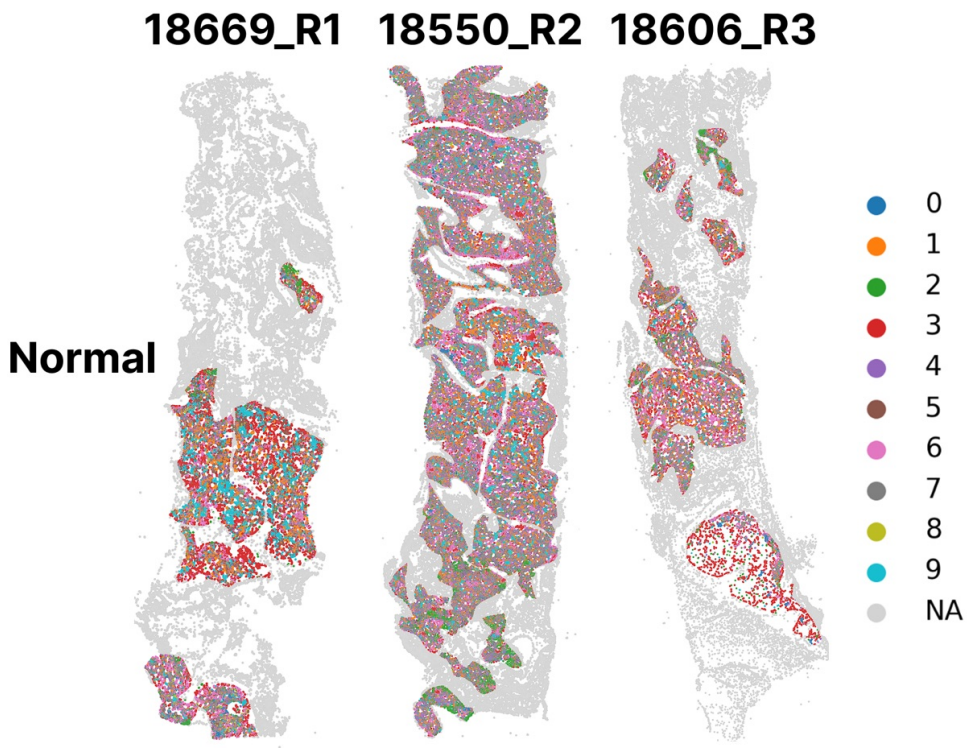
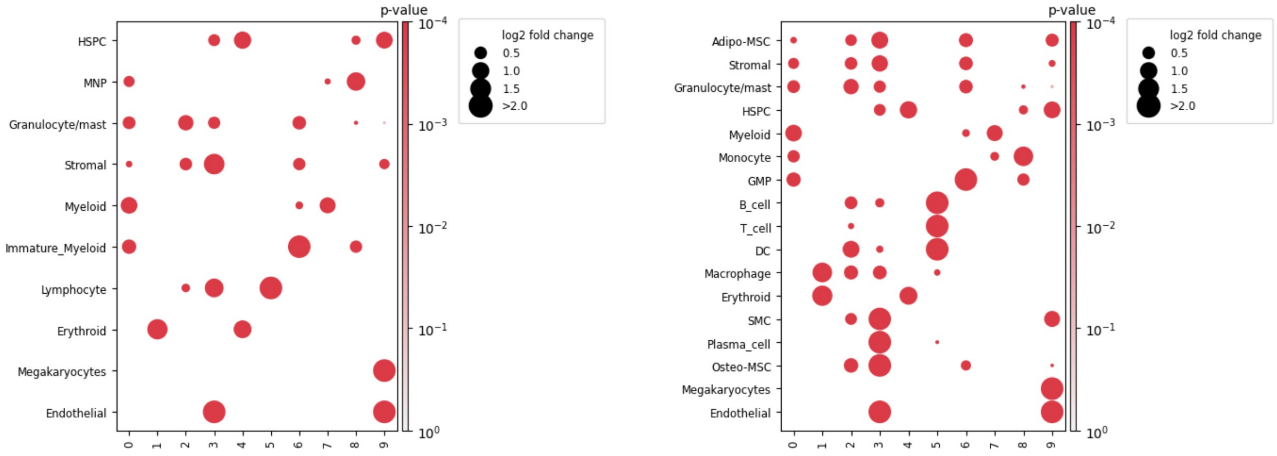


Figure 4.21: CellCharter clustering at 3-hop with 10 clusters (k = 3, n = 10) across three exemplar Normal samples (18669_R1, 18550_R2, and 18606_R3).



(a) Layer 5 enrichment

(b) Layer 3 with stromal enrichment

Figure 4.22: Dot plot showing enrichment of cell types across 10 CellCharter clusters for layers 5 (left) and 3 with stromal refinement (right). Circle size indicates log2 fold change; color reflects statistical significance (p-value).

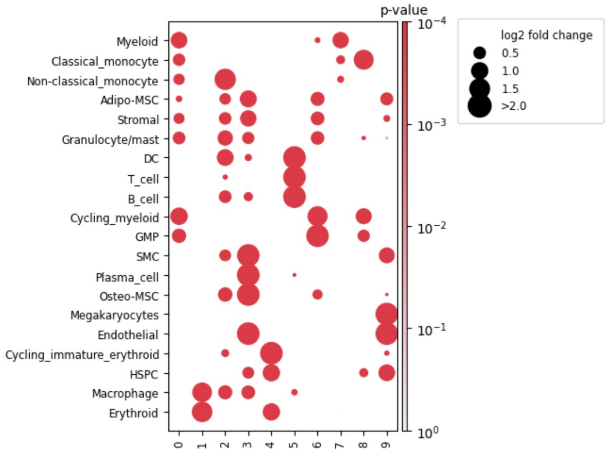


Figure 4.23: Cell type enrichment across layers for Cellcharter 10 clusters, at level 1 annotation with refined stromal groups. Circle size indicates log2 fold change; color reflects statistical significance (p-value).

Characterising CellCharter Cellular Neighbourhoods by Gene Enrichment

I questioned whether cell-type enrichment alone was the best way to define each region, or whether some were better characterised by expression of inflammatory or fibrotic markers. To explore this, I identified region-specific genes using Scanpy's rank_genes_groups with the Wilcoxon test. I selected markers expressed in more than 25% of cells within a region and in fewer than 50% of cells outside it, and filtered for genes with a log fold change morethan or equal to 1.

As expected, most marker genes reflect the dominant cell lineages within each CellCharter region. CN2 (Non-classical Monocytes with Osteo/Stromal and DC/B cells) does not show any differentially expressed genes at the applied thresholds. CXCR4 is upregulated in CN0 (myeloid trajectory) and CN6 (GMP/Stromal), which may suggest sites of HSPC retention and chemoattraction, despite the absence of its ligand CXCL12 in our panel. Proliferation markers such as MKI67 are enriched in CN1 (erythroid/macrophage) and CN4 (HSPC niche with cycling erythroid cells), likely reflecting the higher turnover rate of erythroid cells.

Investigating HSPCs within CNs

In addition to the log fold enrichment shown in Fig. 4.22 to quantify cell enrichment in each CN, I also investigated the percentage of cells which are HSPCs in each CN per patient. This is shown in Fig. 4.24.

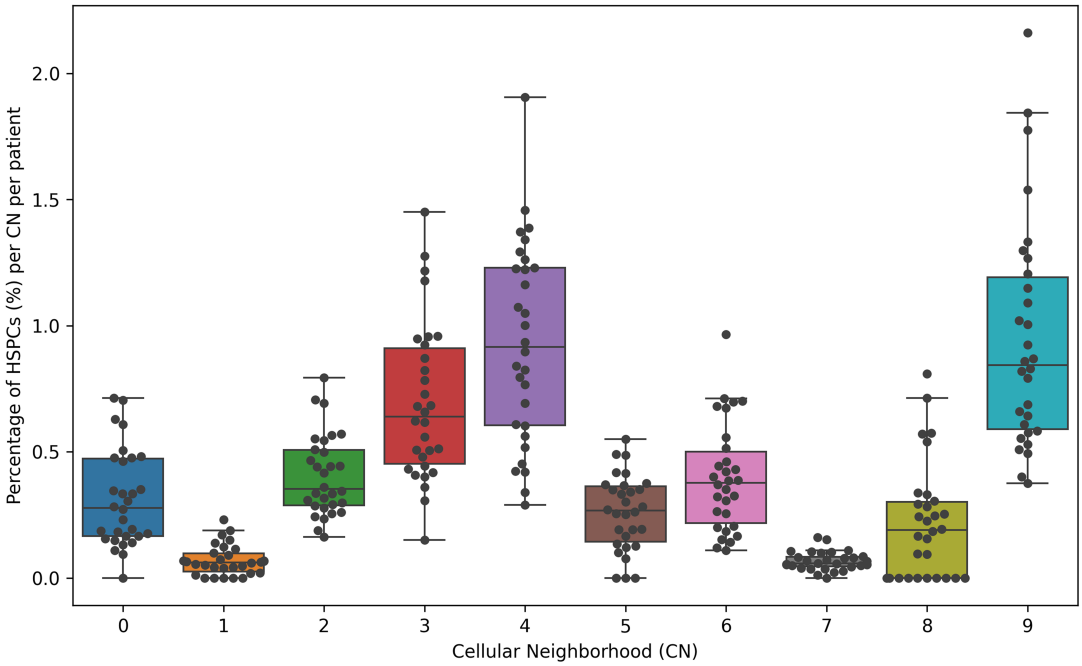


Figure 4.24: Percentage of HSPCs in each CN per patient. Colours represent CNs and dots are the percentage of cells within that CN which are HSPCs in each patient.

As in Fig. 4.22 CN3, CN4 and CN9 appear to be most enriched for HSPCs. To quantify, I performed Wilcoxon signed-rank test between each pair of CNs and used the Benjamini-Hochberg multiple hypothesis test.

CN3 is significantly different (adjusted p value < 0.05) to all other CNs, CN4 is significantly different to all other CNs except CN9, CN9 is significantly different to all other CNs except CN4.

I also investigated whether the HSPCs in each CN expressed a unique set of genes to those HSPCs in other CNs. Finding differential gene expression may give greater insight into the role of these HSPCs. Differential expression analysis was of course limited by the pan-cancer panel, which was not weighted for HSPC lineage markers. Nevertheless, after performing differential expression analysis between each CN and all other cells using the wilcoxon method within scanpy's rank_genes_groups, I returned the top 5 differentially expressed genes which had a minimum log fold change of 1 (Fig. 4.25).

In line with the cell types defining each CNs, HSPC gene expression correlated with the main cell types defining the CNs. HSPCs in CN4 (HSPC niche with Cycling Erythroid cells)

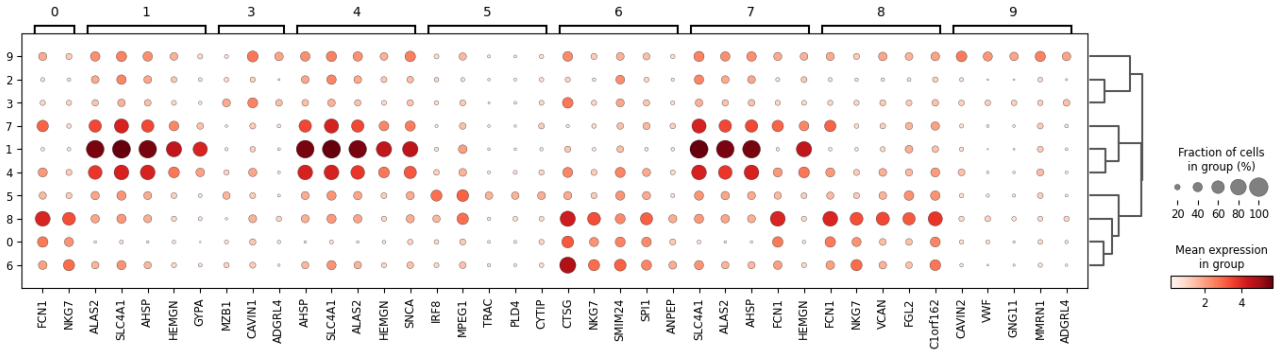


Figure 4.25: DotPlot showing the top 5 differentially expressed genes (only genes which have a minimum log fold change of 1 are shown) of HSPCs per CN vs all other CNs. Analysis was performed with scanpy's rank genes groups using the wilcoxon method on log normalised expression. Multiple hypothesis testing was performed with benjamini-hochberg

expressed erythroid lineage markers (AHSP, SLC4A1, ALAS2, HEMGN, SNCA). HSPCs in CN3 (HSPC niche with Endothelial/SMC, Osteo-MSCs and Plasma/B cells) expressed endothelial (CAVIN1 ADGRL4) and plasma/B cell markers (MZB1). HSPCs in CN9 (HSPC niche with Megakaryocytes and Endothelial/SMC) expressed endotheial and megakaryocyte markers (CAVIN2, VWF, GNG11, MMRN1, ADGRL4).

Whilst this could indicate these HSPCs were primed for differentiation into these lineages, I would be cautious in interpreting these results. Expression of lineage markers in HSPCs may indicate close physical association, due to overlap from segmentation and resolution of transcript detection, rather than HSPC expression.

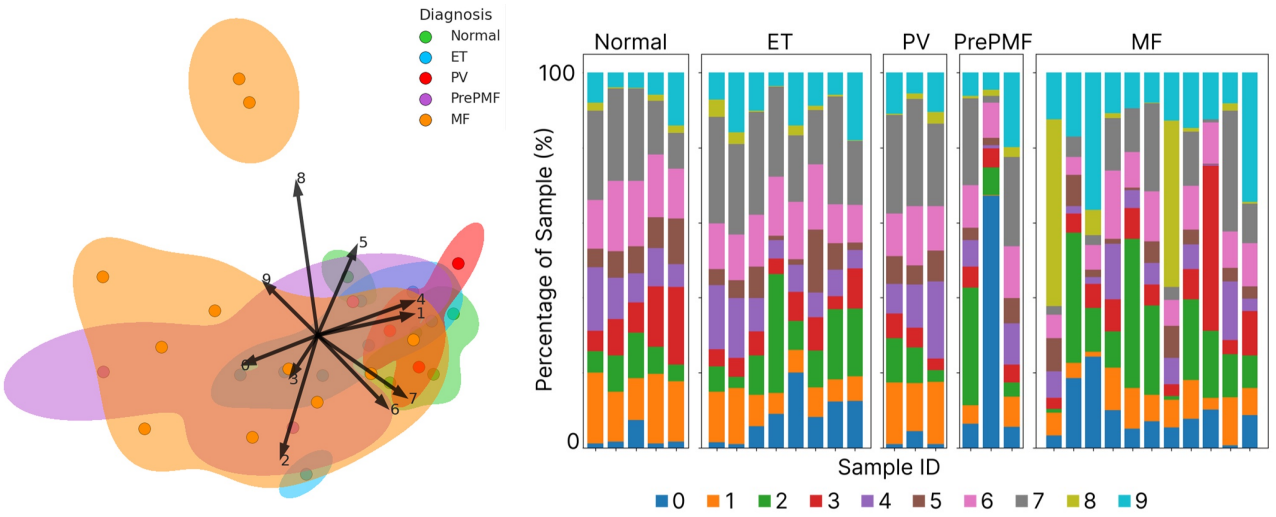
Heterogeneity within MF samples is driven by CN2 and CN8

To determine whether specific cellular neighbourhoods (CNs) were unique to particular MPN subtypes, I examined their distribution across samples. All Normal, ET, and PV samples contained every CN. In contrast, MF samples showed greater heterogeneity in CN composition, as seen in both the stacked bar plot and PCA (though PCA does not capture all variance (PC2:19.08% and PC1:34.95%)); Fig. 4.26).

Two MF samples (10693_R1 and 18599_R1) were dominated by CN8 (HSPC niche with classical monocytes), with a corresponding reduction in CN2 (non-classical monocytes with Osteo/Stromal and DC/B cells) (Fig. 4.26b). Conversely, three MF samples (16832_R1, 18550_R1, and 18612_R5) showed a marked reduction in CN8, with samples 16832_R1 and 18550_R1 instead showing expansion of CN2. This reciprocal pattern suggests a potential

regulatory mechanism driving the expansion of either CN8 or CN2. However, as these patterns appear in only a subset of samples, further validation in larger cohorts is needed. Given the high cost of spatial transcriptomics, identifying a morphological correlate in archival H&E sections would be a valuable next step.

One additional MF patient, 18612_R5, characterised by severe osteosclerosis, was dominated by CN3 (HSPC niche with Endothelial/SMC, Osteo-MSCs, and Plasma/B cells).



(a) PCA of cellcharter regions colored by diagnosis.

(b) Stacked barplot of cellcharter region composition across samples

Figure 4.26: (A) PCA showing region variation across diagnoses. (B) Stacked barplot of region composition by diagnosis, faceted by patient.

I also tested for compositional differences in CNs across MPN subtypes using differential abundance analysis of cell counts per region. Only Pre-PMF showed significant changes: CN0 (myeloid trajectory) was elevated relative to Normal and PV, while CN1 (erythroid/macrophage) was reduced compared to Normal (Fig. 4.27). These findings suggest that most variation between subtypes arises from spatial organisation rather than overall composition, though intra-subtype heterogeneity may contribute. Interestingly, HSPC niche CNs did not differ significantly in abundance across subtypes. One explanation may be that CellCharter was tuned to identify regions common to all samples, potentially biasing detection toward the most ubiquitous CNs.

MF samples have greater CN spatial expansion

One of the most visually striking differences in MF is the spatial expansion of Cellular Neighbourhoods (CNs) compared to other conditions. To quantify this, I used CellCharter's

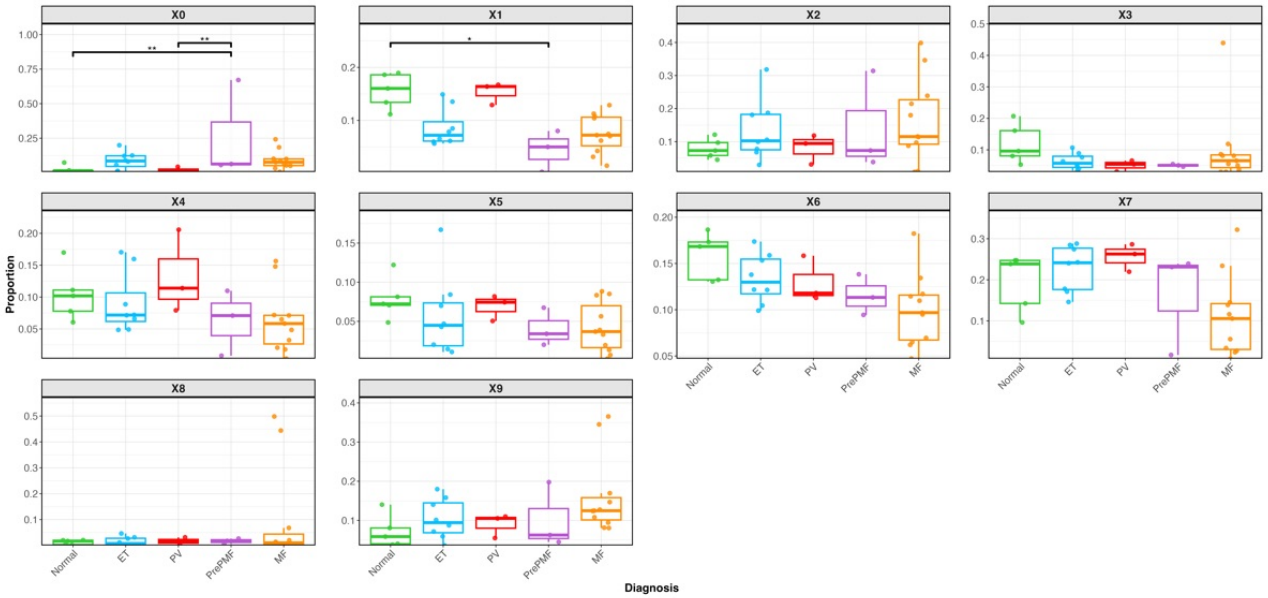
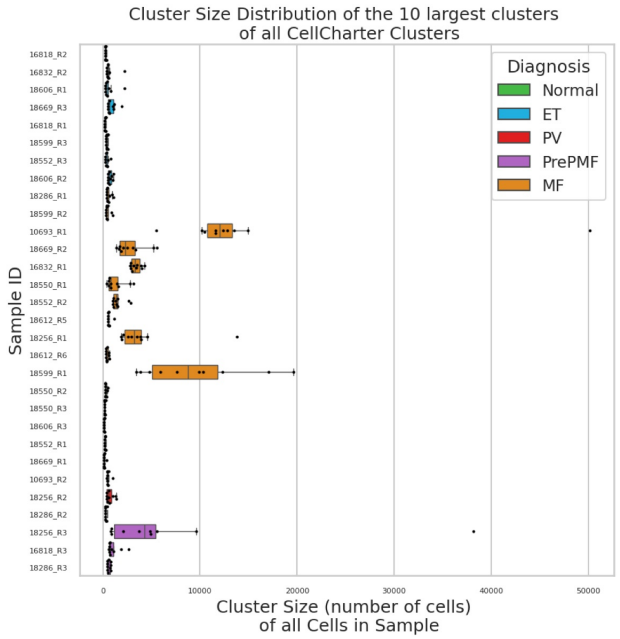


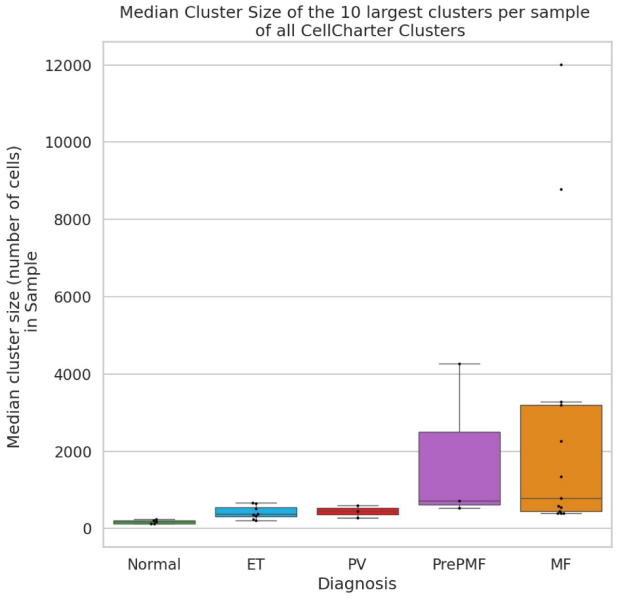
Figure 4.27: Differential Abundance of CellCharter Clusters across conditions, Xn represents CNn where n is 0 to 10

connected-components tool to identify clusters of at least five adjacent cells in each sample, based on the Voronoi-neighbour graph. I then measured the size of each connected component. Since larger clusters occupy more space and thus occur less frequently, I focused on the ten largest components per sample to compare spatial expansion. MF samples consistently had larger top-10 connected components than other subtypes (Fig. 4.28). Because larger samples naturally contain more cells, component cell number may reflect sample area rather than true spatial expansion. To account for this, I normalized each component's size by the total number of cells in the sample and plotted the ten largest component proportions per sample. MF samples showed significantly larger maximum cluster proportions than ET, PV, or Normal, suggesting a more heterogeneous and spatially expanded microenvironment (Fig. 4.28).

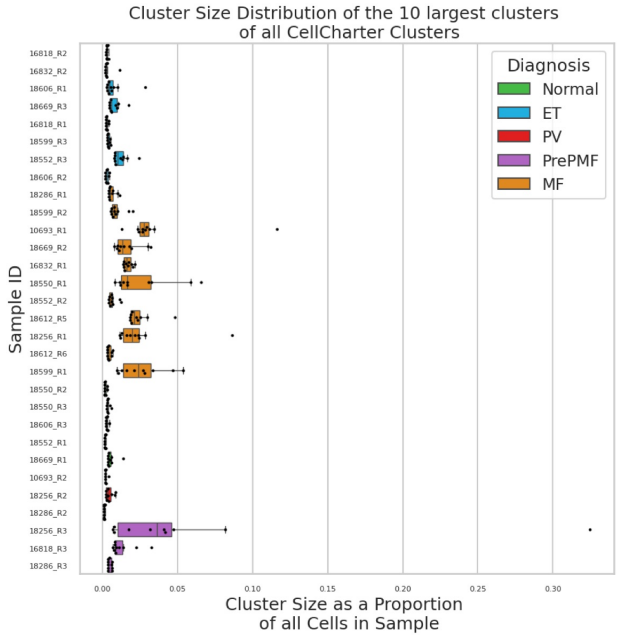
MF samples show larger regional clusters than Normal, PV, or ET, suggesting a less mixed microenvironment dominated by expanded CNs. To identify which regions drive this expansion, I examined cluster-size distributions for each CN. In the resulting dot plots, colour indicates the number of clusters in each size bin, while dot size reflects the proportion of total cells within that bin (Fig. 4.29, 4.30). These results show that even CNs with large connected components also contain many smaller ones. Thus, regions are not defined by a critical minimum size but can exist as both large and small domains.



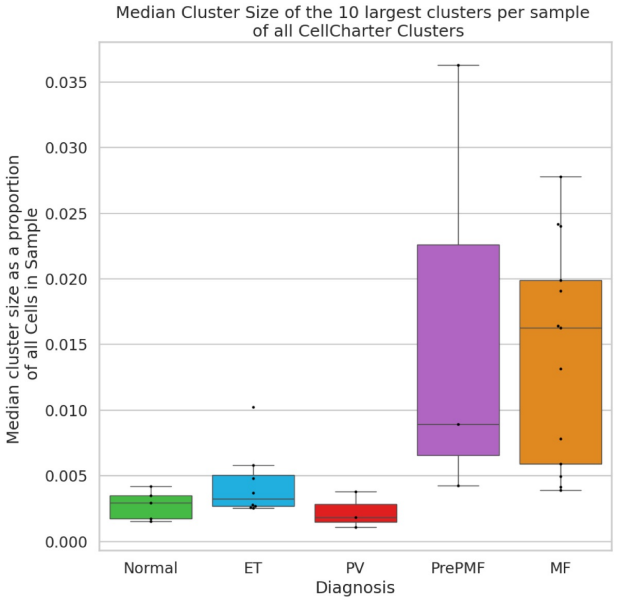
(a) Top 10 largest CellCharter connected components (count)



(b) Median cell count of top 10 components per sample



(c) Top 10 largest CellCharter components (proportion)



(d) Median proportion of top 10 components per sample

Figure 4.28: Quantification of the 10 largest CellCharter connected components per sample. (A) Raw cell counts; (B) Median cell count per sample by diagnosis. We see a clear expansion of region size in MF samples, with significant (show on plot) difference in the maximum cluster size between MF and Normal (p=0.002334) and ET (p=0.025816), and ET and Normal (p=0.025816) after a MannWhitneyU test with BH MHT (C) Proportion of all cells in the top 10 components; (D) Median proportion per sample by diagnosis. We see a clear expansion of region size in MF samples, with significant difference (show on plot) in the maximum cluster size between MF and Normal (p=0.006683), PV (p=0.011905), and ET (p=0.006683) after a MannWhitneyU test with BH MHT

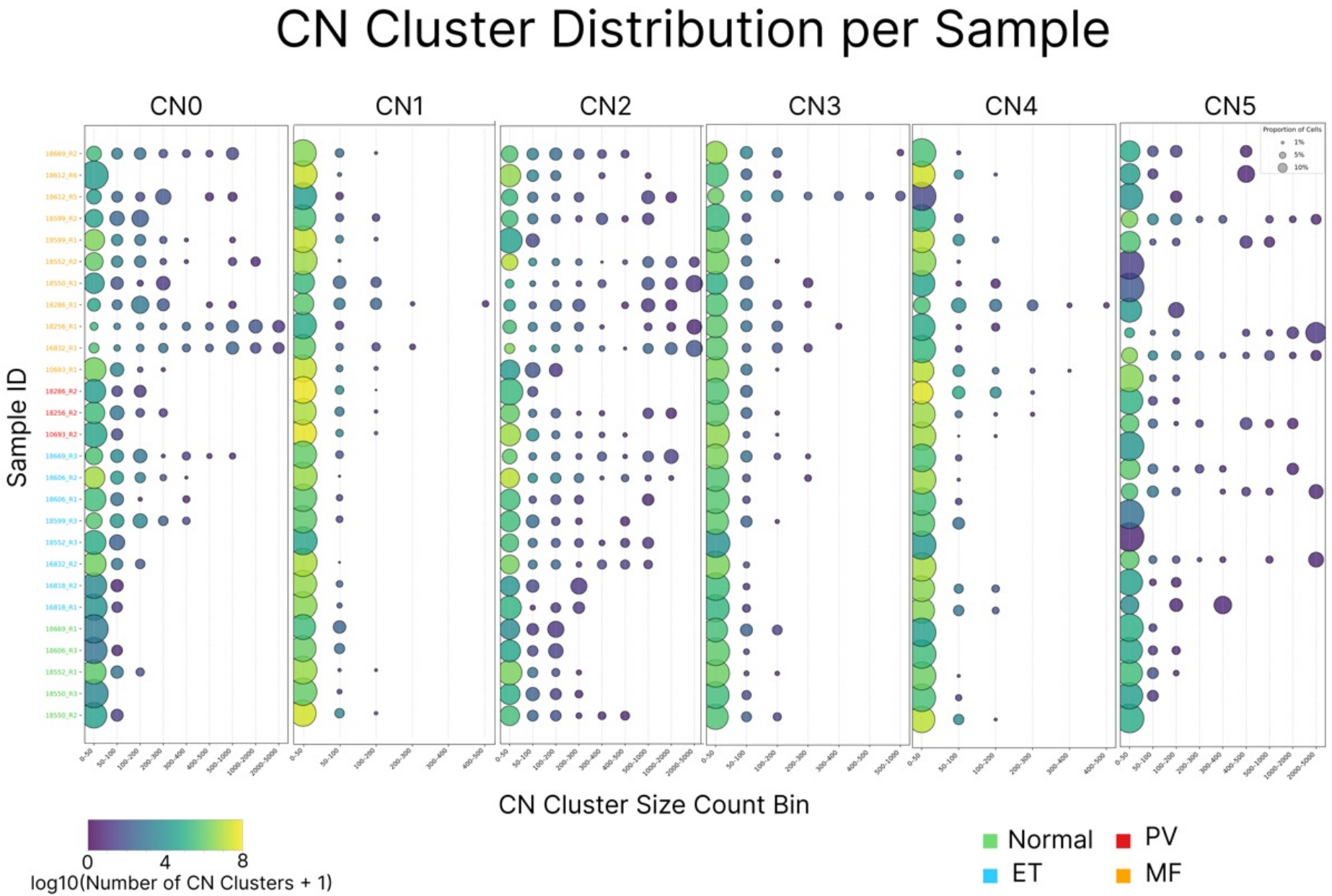


Figure 4.29: Dot plot showing the distribution of CellCharter cluster sizes for each sample (clusters 0–5). The x-axis represents cells per bin. Each subplot shows a CellCharter region. Connected components are grouped into bins by the number of cells per component. For each bin, the dot size reflects the total proportion of all cells in that sample, and colour shows the number of connected components in that bin and region.

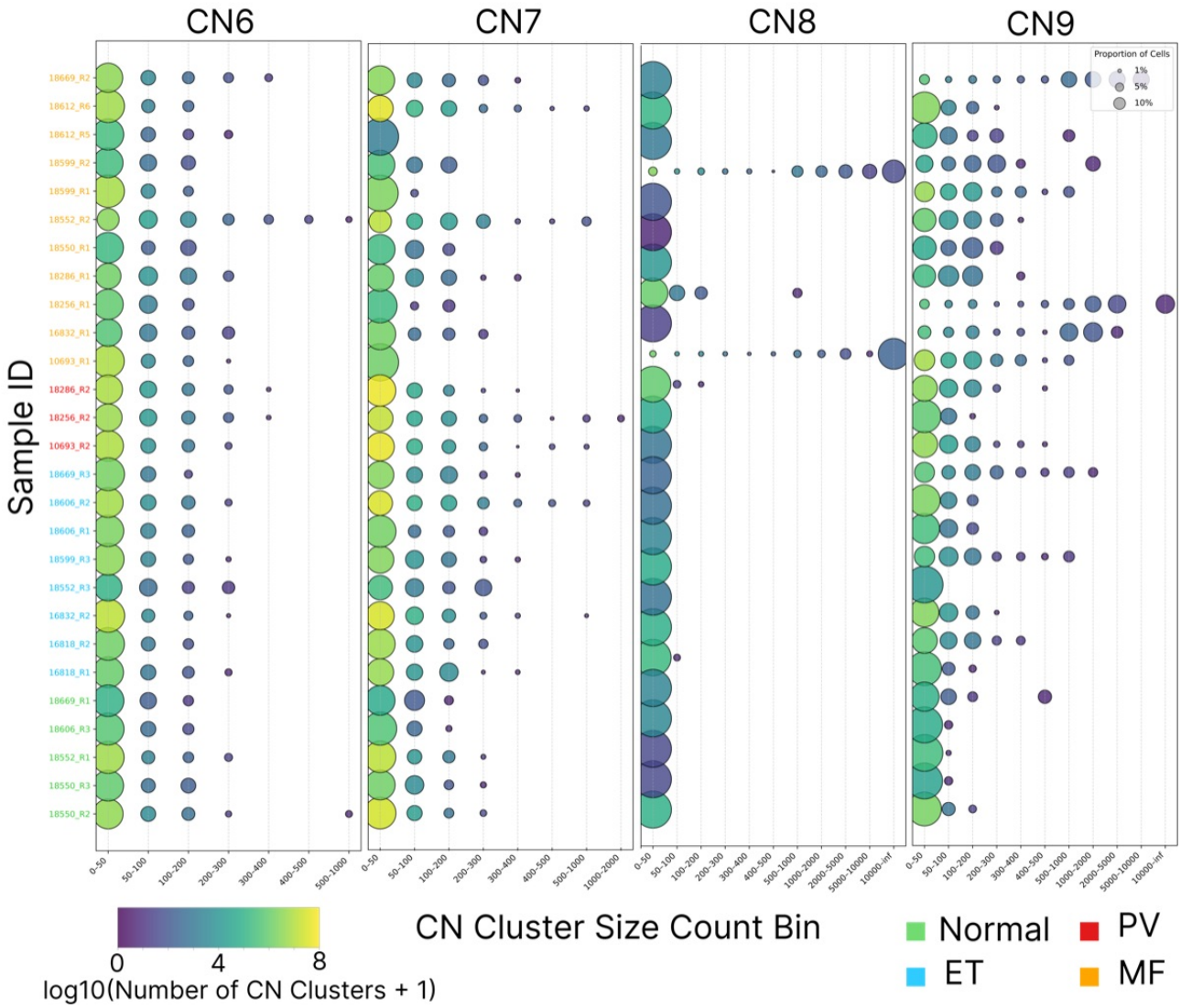


Figure 4.30: Dot plot showing the distribution of CellCharter cluster sizes for each sample (clusters 6–9). Each subplot represents a CellCharter region. Connected components are grouped into bins by the number of cells per component. Dot size corresponds to the summed proportion of the total sample from all connected components in that bin. Colour reflects the number of such connected components.

CN 0, 2 and 9 are spatially expanded in MF

Some Cell Neighbourhoods (CNs), such as CN1 (erythroid/macrophage), CN3 (HSPC niche with Endothelial/SMC, Osteo-MSCs, and Plasma/B cells), and CN4 (HSPC niche with Erythroid cells), rarely form large clusters. The lack of spatial expansion in CN3 and CN4 raises the possibility that they contain non-mutated HSPCs, or alternatively, that cells generated in these niches migrate to form other neighbourhoods.

The dot plots show a clear pattern: MPNs, especially MF, are characterised by expanded CNs, while healthy bone marrow has a higher number of smaller, dispersed connected components (Fig. 4.29, 4.30). The most pronounced expansions in MF are in CN0 (myeloid trajectory with stromal),

CN2 (non-classical monocytes with Osteo/Stromal and DC/B), and CN9 (HSPC niche with megakaryocytes and Endothelial/SMC). To test this formally, I used the Wilcoxon signed-rank test to compare the sizes of the top 10 largest clusters per CN within each sample.

In MF, CN0 was significantly larger than CN4; CN2 was significantly larger than CN1, CN4, CN6, and CN7; and CN9 was significantly larger than CN1, CN4, CN5, and CN6. In ET, CN9 was also expanded relative to CN1, CN3, and CN8, but the most striking expansions were in CN6 (vs. CN1, CN3, CN4, CN8) and CN7 (vs. CN1, CN3, CN4, CN6, CN8). No significant differences in top cluster size were observed in PV or Normal samples. CN6 is a GMP/Stromal region whilst CN7 is a non-specific myeloid region. Although not significant in both, the expansion of CN0, 2 and 9 appears to be a pattern seen in ET and MF, whilst the expansion of CN6 and CN7 is more commonly seen in non-fibrotic and healthy samples (Fig. 4.31).

CN8 (HSPC niche with classical monocytes) expands notably in only two MF samples (Fig. 4.26). Given that large CNs are a hallmark of MF, while smaller, dispersed CNs are more typical of healthy bone marrow, the expansion of CN8 in these cases may indicate more severe disease or poorer prognosis. I therefore examined CN8 expansion in relation to fibrosis grade in subsequent analyses.

Conserved CN-CN Co-localisation across all conditions

Within each condition, I assessed whether specific CNs are more or less likely to neighbour one another than expected by chance (Fig. 4.32). As expected, each CN most frequently neighbours itself due to the k=3-hop construction of CellCharter clusters. However, consistent inter-CN associations were observed across conditions, suggesting the existence of conserved regional building blocks within the bone marrow. Notable enriched proximities include:

- CN1 and CN4 (Erythroid/Macrophage – HSPC niche with Cycling Erythroid cells)
- CN3 and CN6 (HSPC niche with Endothelial/SMC, Osteo-MSCs, and Plasma/B cells – GMP/Stromal)
- CN4 and CN7 (HSPC niche with Cycling Erythroid cells – Non-specific Myeloid region)

These two HSPC niches (CN3 and CN4) are less spatially expanded and exhibit stable CN associations, supporting the hypothesis that they may represent healthy stem cell microenvironments. Interestingly, only healthy donor samples show a low-probability but significant association

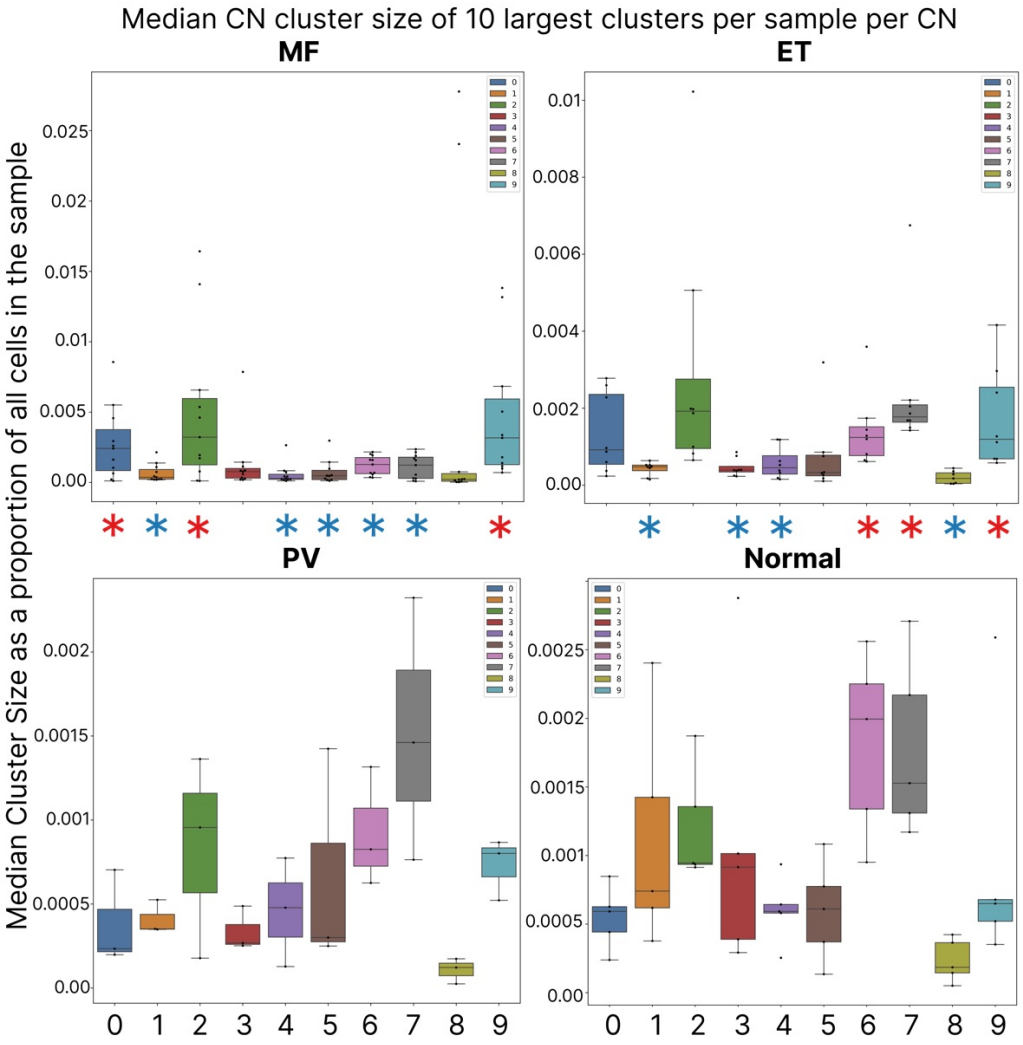


Figure 4.31: Median cluster size of the 10 largest clusters per CN per sample across conditions. Red/Blue stars indicate significantly expanded/restricted compared to other CNs ($p < 0.5$)

between CN8 (HSPC niche with Classical Monocytes) and CN5 (Lymphocyte/DC region) (Fig. 4.32). If CN8 expansion marks poor prognosis, as previously suggested, its association with CN5 in healthy samples may reflect a role in immunosurveillance.

Differential HSPC enrichment in CNs within conditions

Given the differential associations between CN5 and CN8, I next asked whether cell-type enrichment within CNs might vary subtly between conditions. To investigate this, I assessed cell-type enrichment per CN within individual conditions rather than across the full dataset. We observe slight shifts in enrichment patterns, particularly within HSPC niches. Notably, CN8 is enriched for HSPCs only in MF, whereas in ET and PV, CN2 shows HSPC enrichment instead. CN3, CN4, and CN9 consistently remain HSPC-enriched across all conditions (Fig. 4.33).

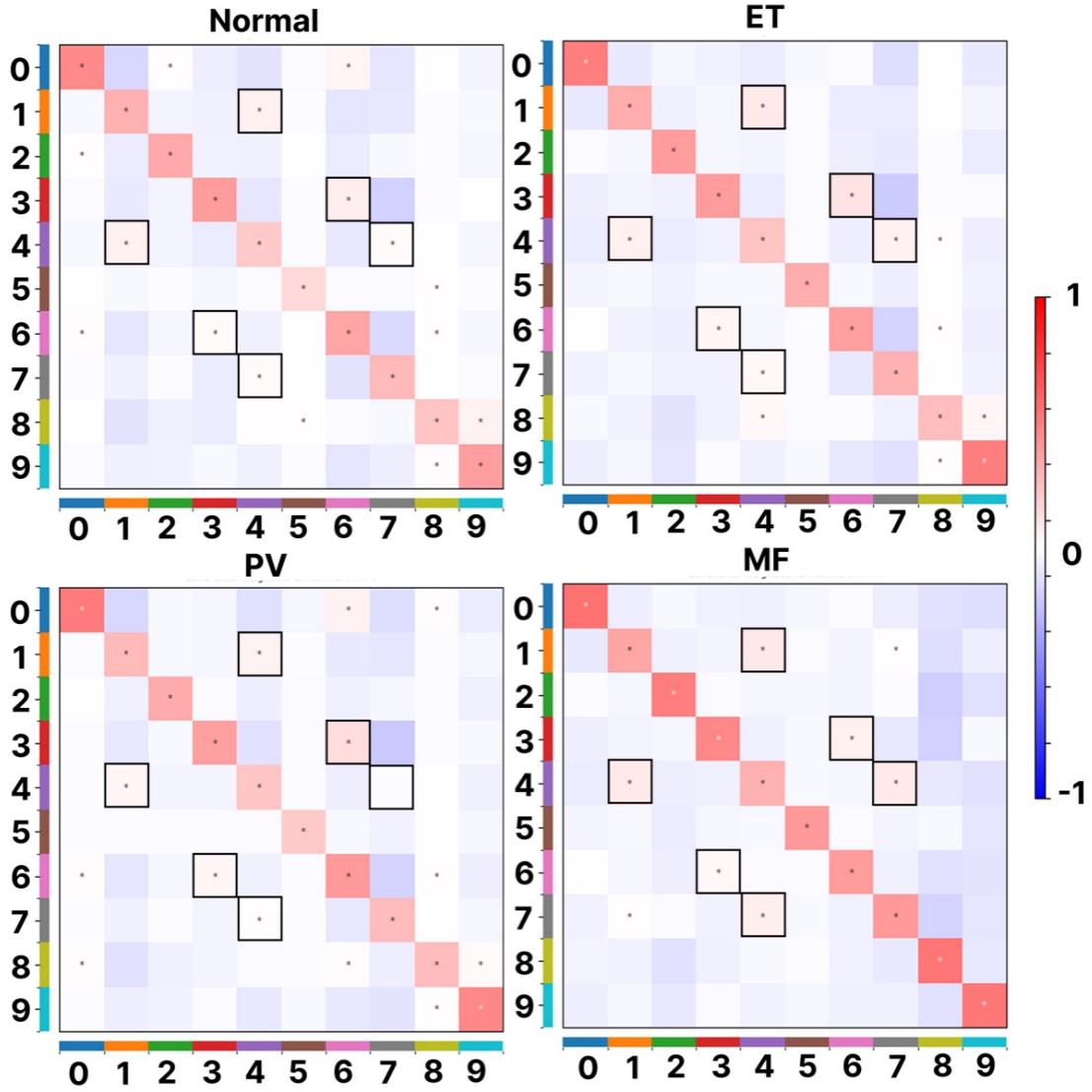


Figure 4.32: CellCharter neighbour enrichment matrices for each condition. Colours represent $\log_2$(observed/expected) after 100 permutations of cell labels. Asterisks indicate significance at $p < 0.05$. Black boxes indicates preserved CN-CN relationships across all conditions.

Differential CN-CN Co-localisation across Conditions

The previous neighbourhood enrichment analysis tested for significant CN–CN associations within each condition. However, to understand how MPN alters the spatial architecture of bone marrow, we next asked which CN–CN relationships differ significantly between MPN subtypes and healthy controls. To do this, we conducted a permutation test on sample labels to assess condition-specific enrichment relative to Normal samples (Fig. 4.34).

Although CN8 (HSPC niche with Monocytes) and CN5 (Lymphocyte/DC region) appeared associated in healthy marrow, this was not statistically more enriched in Normal than in MPN subtypes. Given CN5's high lymphocyte content and possible immunosurveillance role, we asked

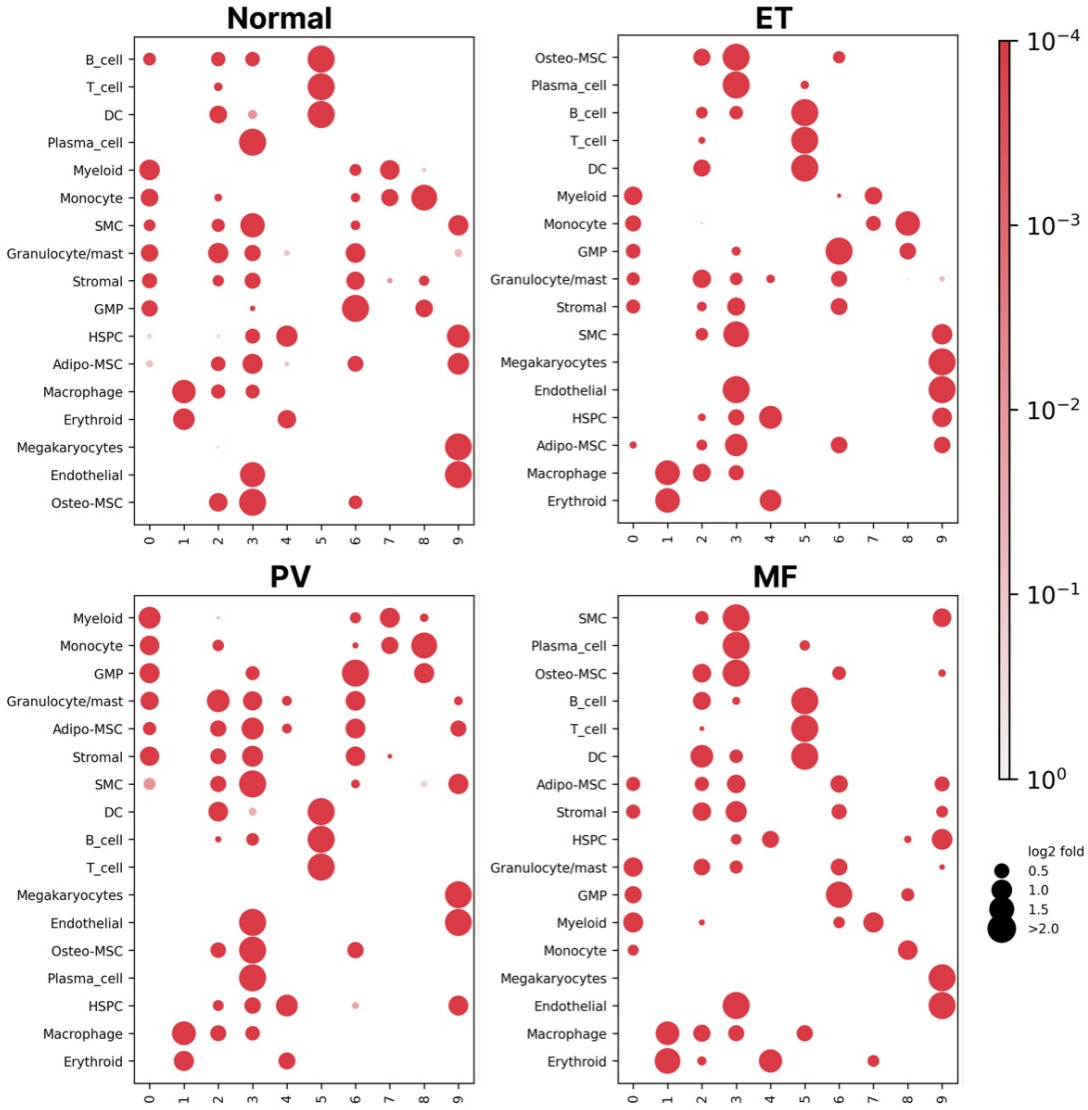


Figure 4.33: Enrichment of cell types across CellCharter clusters in Normal, ET, PV, and MF conditions. Dot size represents $\log_2$ fold change in observed vs. expected probability; color indicates p-value.

whether CN5 had different neighbours in MPN compared to Normal.

Indeed, in Normal samples, CN5 is more likely to neighbour CN0, CN2, CN7, and CN9 than in ET, and CN0 and CN9 than in MF (Figs. 4.34). These CNs were previously associated with expansion in ET and MF, suggesting a potential loss of immunosurveillance may permit their spatial expansion in MPN.

In PV, CN3 (the HSPC niche with Endothelial/SMC, Osteo-MSCs, and Plasma/B cells) is more likely to neighbour CN6 (GMP/Stromal) compared to Normal, and shows broader connectivity to multiple CNs (Fig. 4.34).

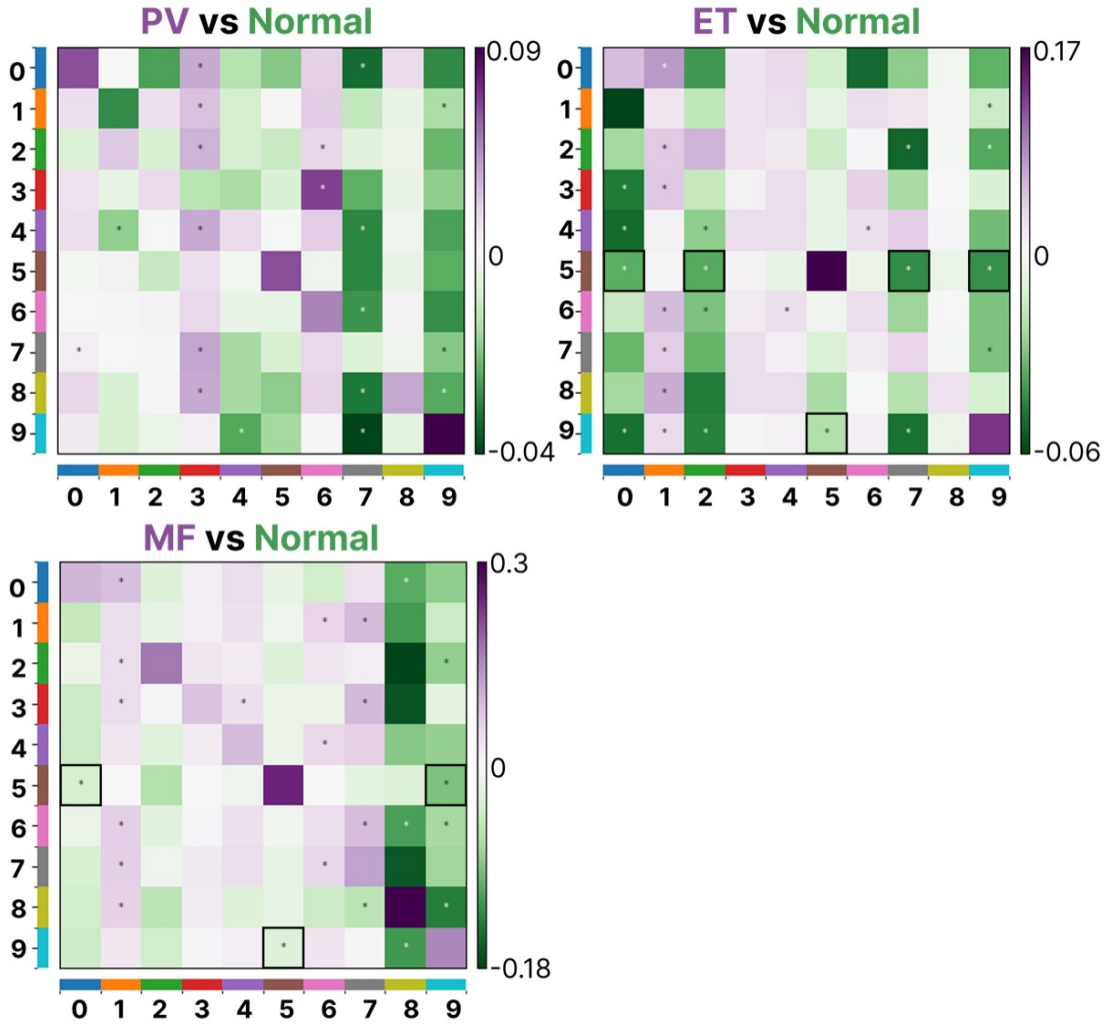


Figure 4.34: Differential CellCharter neighborhood enrichment matrices showing contrast between PV, ET, and MF compared to Normal. Asterisks indicate statistically significant changes ($p < 0.05$) in cluster neighbour relationships. Green indicates enrichment in Normal compared to MPN.

Co-localisation of CellCharter Regions with bone marrow features

Having characterised these CNs of the bone marrow with respect to each other, I wanted to understand how they related to features of the bone marrow. Within MPNs we are particularly interested in fibrosis. Fibrosis is believed to start close to the bone and move into the central regions. Distance to bone is also relevant for HSPC location as the first hypothesised niche was suggested to be endosteal. Now, studies have started to suggest that the endosteal niche may just be perivascular or even raised suggestions of an peri-adipocytic niche. Nevertheless distance to bone has been an important landmark in analysing the spatial organisation of the bone marrow.

Investigating CNs in relation to bone identifies endosteal and peritrabecular

regions

Visual inspection suggests that CN3 (HSPC niche with Endothelial/SMC, Osteo-MSCs, and Plasma/B cells) and CN6 (GMP/Stromal region) correspond to the endosteal and peritrabecular regions, respectively (Fig. 4.21, 4.19, 4.20, 4.18). Supporting this, we previously observed in Fig. 4.32 that CN3 and CN6 frequently co-localize across all conditions. To further investigate this potential spatial relationship, I visualised the distribution of CellCharter CNs as a function of distance from bone (Fig. 4.35).

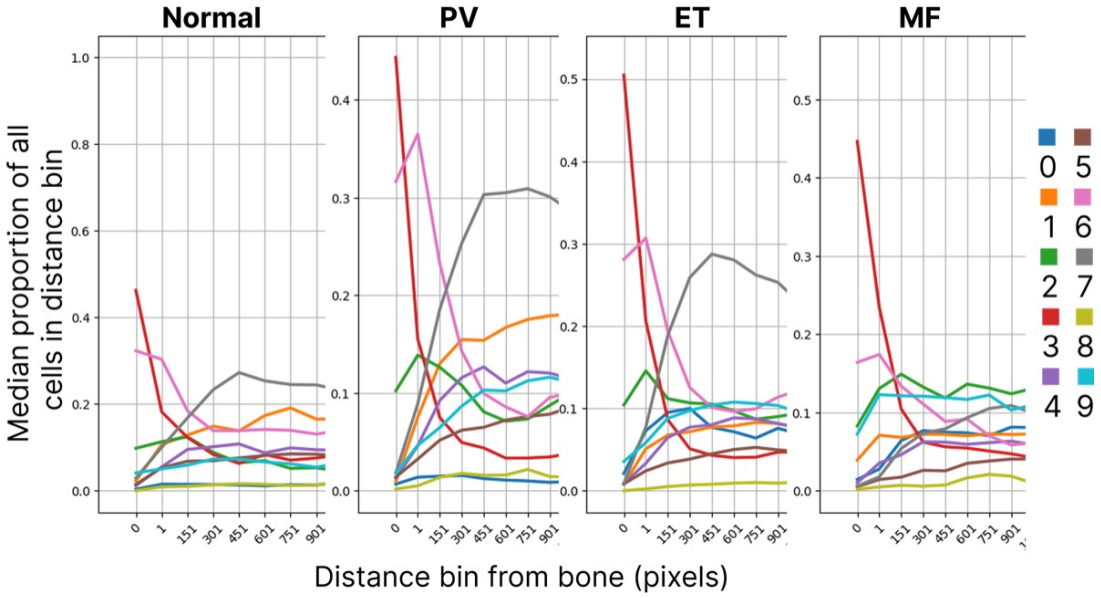


Figure 4.35: Median proportion of CellCharter regions across distances to bone by condition (0–1, then in 150-pixel intervals).

Across all bone regions, CN3 (HSPC niche with Endothelial/SMC, Osteo-MSCs, and Plasma/B cells) is the dominant cluster at the endosteal surface (0–1 pixels from bone), followed by CN6 (GMP/Stromal), consistently across all conditions. Although there is no fixed definition for the peritrabecular zone, we observe that CN3 is gradually replaced by CN6 at increasing distances from bone. Specifically, this transition occurs between 1–151 pixels in most conditions, but is delayed to 151–301 pixels in MF. As this is an aggregated view without accounting for sample variability, the apparent expansion of CN3 in MF may reflect a few highly trabecular samples, a known feature of this condition. At greater distances, CN3 and CN6 decrease in abundance, aligning with other CNs (Fig. 4.35).

Manual definitions of the endosteal and peritrabecular regions confirm CN3 and CN6 respectively

To further explore whether unsupervised CNs correspond to known anatomical regions, I divided each sample into endosteal (0 pixels; directly touching bone), peritrabecular (1–300 pixels, 64 um), and intertrabecular (>300 pixels) zones. These definitions were developed in consultation with pathologists, with the peritrabecular zone approximating two fat-cell widths from bone (Fig. 4.16). Across all conditions, CN3 (HSPC niche with Endothelial/SMC, Osteo-MSCs, and Plasma/B cells) and CN6 (GMP/Stromal) are enriched near bone and decrease in abundance with distance (Fig. 4.36). CN3 shows a stepwise decline across compartments, with highest abundance at the endosteal surface. In contrast, CN6 shows no significant difference between endosteal and peritrabecular zones, but decreases beyond. Based on these trends, CN3 can be defined as the endosteal region and CN6 as peritrabecular. Most other CNs increase in abundance with greater distance from bone.

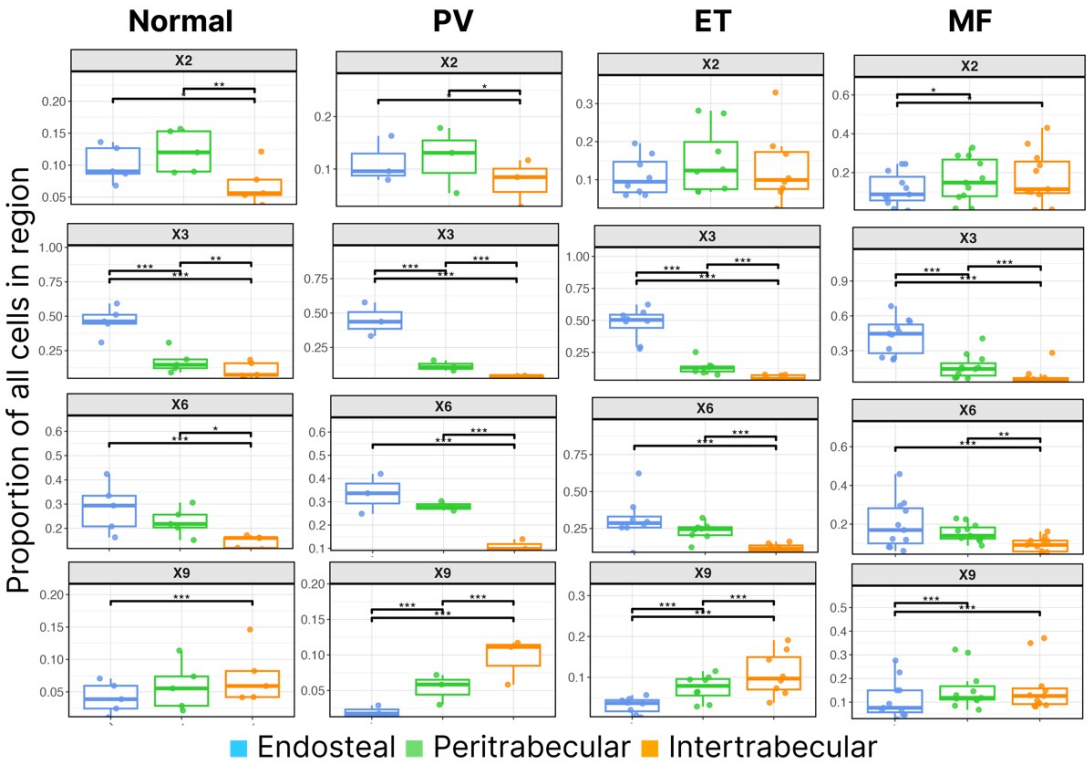


Figure 4.36: Differential abundance of CellCharter regions (CN2/3/6/9) in each bone region within each condition. Xn denotes CNn, X3 is CN3.

CN2 may move out from the endosteal as MPNs progress towards fibrosis

CN2 also shows condition-specific localization relative to bone. This CN, enriched for HSPCs in non-fibrotic MPNs (ET and PV), demonstrates distinct spatial patterns across conditions. In Normal and PV samples, CN2 is less abundant in the intertrabecular region compared to the endosteal and peritrabecular zones. In ET, CN2 shows no significant variation across

compartments. In MF, however, CN2 (Non-classical Monocytes with Osteo/Stromal and DC/B cells) is enriched in the peritrabecular and intertrabecular regions relative to the endosteal (Fig. 4.36). This pattern supports a potential trajectory in which CN2 localizes near bone in healthy marrow and progressively shifts toward central marrow in MPN progression; from Normal and PV, through ET, to MF.

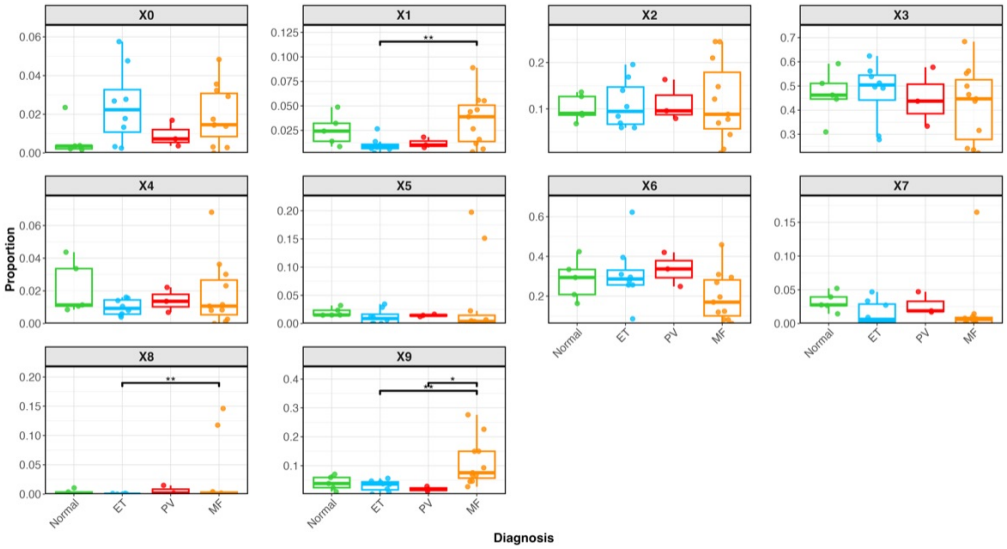
CN9 expands towards the bone in MF and replaces the usual peritrabecular region, CN6

Given these localization patterns, I next investigated whether the composition of CNs within bone-defined regions (endosteal, peritrabecular, intertrabecular) differs across conditions. I performed differential abundance analysis stratified by region and condition (Fig.4.37). Although whole-trephine CN abundance did not differ significantly between conditions, region-specific analysis reveals compositional shifts. This suggests that CN localization differs between disease states.

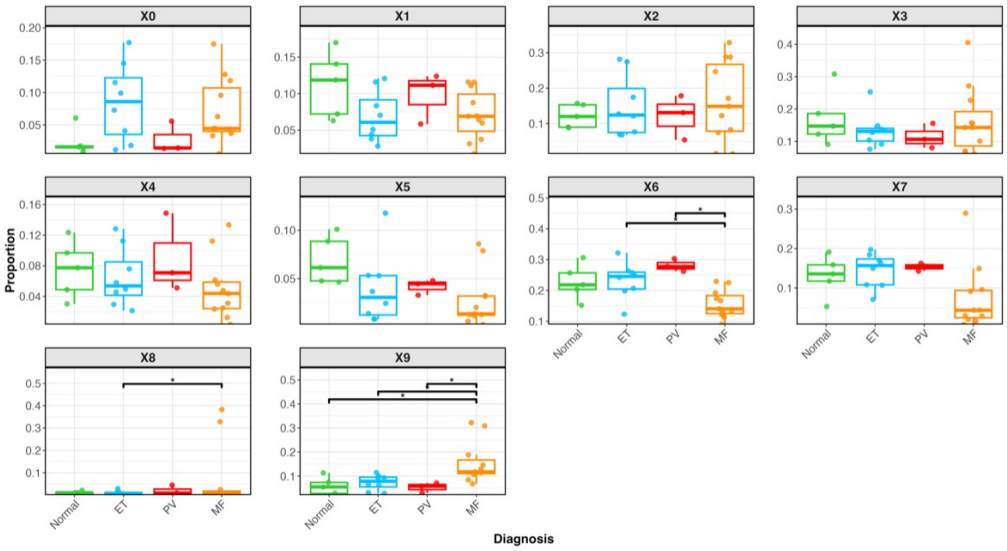
In the endosteal region, CN3, the designated endosteal region, shows no significant difference in abundance across conditions. However, MF samples show increased CN9 (HSPC niche with Megakaryocytes and Endothelial/SMC) compared to non-fibrotic MPNs (ET and PV) (Fig. 4.37). In the peritrabecular region, CN9 is significantly increased in MF compared to all other conditions, while CN6 (the designated peritrabecular region) is decreased in MF relative to ET and PV. This suggests that in MF, CN9 replaces the typical peritrabecular composition (Fig. 4.37). CN2 shows no significant differences between conditions within any bone region. However, Figure 4.37c suggests that intra-condition heterogeneity, particularly within MF, may obscure any group-level differences. In the intertrabecular space, MF samples show increased abundance of CN0 (myeloid trajectory with stromal) and CN9, and decreased CN6 (GMP/Stromal) and CN1 (Erythroid/Macrophage) compared to Normal (Fig. 4.37).

CN2 may be a cellular correlate of early fibrosis

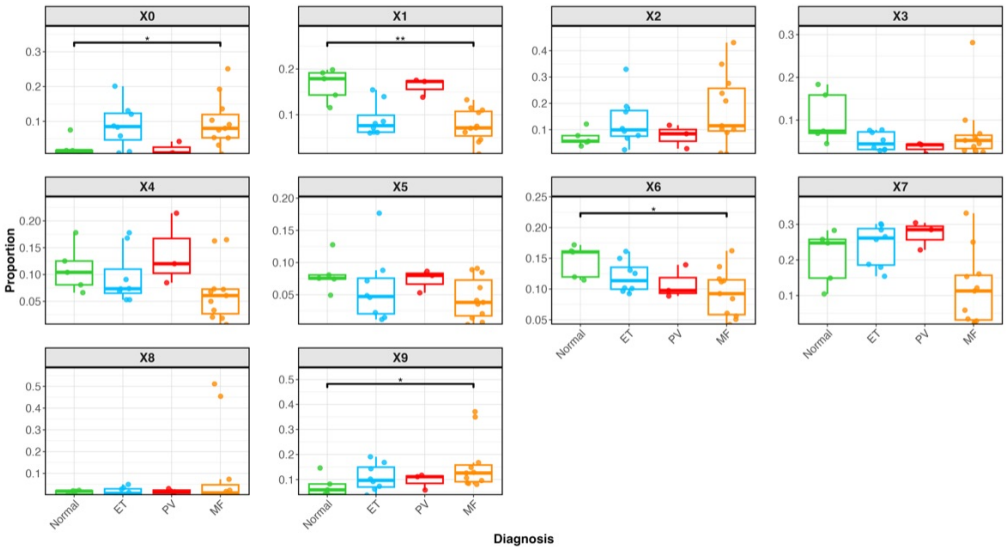
Given that key CN differences are observed near the bone, and fibrosis is believed to initiate in these areas and progress outward, I investigated CNs in relation to fibrosis. This also allowed me to test the earlier hypothesis that CN2 may serve as a cellular correlate of fibrosis progression, given its apparent shift away from the bone as MPNs progress.



(a) CellCharter Regions in Endosteal



(b) CellCharter Regions in Peritrabecular



(c) CellCharter Regions in Intertrabecular

Figure 4.37: Differential Abundance of CellCharter Regions Across Bone Regions

When visualising the continuum of fibrosis scores across individual trephines, I observed that CN2 tends to localize in regions with high fibrosis in MF samples. This observation is particularly intriguing, as we previously noted a negative relationship between CN2 and CN8 (HSPC niche with Classical Monocytes), along with CN2's altered localization in MF compared to ET, Normal, and PV (Fig. 4.26b, 4.36). To evaluate whether CN2 is genuinely associated with fibrosis, I compared fibrosis scores across CNs, within each condition (Fig. 4.38).

To assess the association between CNs and fibrosis, I applied a mixed-effects model, treating sample ID as a random factor to account for the paired design. I tested whether the median CIF (Continuous Index of Fibrosis) score differed significantly between each CN and CN2.

In MF, CN2 exhibited a significantly higher CIF score than all other CNs. Repeating this analysis on the non-MF samples (Normal, ET, and PV), CN2 again showed a higher CIF score than all other CNs in ET. In Normal and PV cells in CN2 had significantly higher CIF scores than CN1/CN3/CN4/CN6/CN9 (the spatially restricted CNs and the three HSPC enriched CNs) and CN3/CN6 (the endosteal and peritrabecular) respectively (Fig. 4.38).

These findings support the hypothesis that CN2 is associated with fibrosis. Although the median CIF score for CN2 is elevated relative to other CNs, it remains below 0.35 in all conditions and is lower in ET than in MF (Fig. 4.38). Scores below 0.439 correspond to fibrosis grades MF-0 and MF-1. This suggests that CN2 is not merely a correlate of pre-existing, advanced fibrosis.

Instead, several features point to CN2 as a potential correlate of early fibrotic change: its shifting localisation away from the bone in parallel with fibrosis progression (Fig. 4.36), its enrichment for Osteo-MSCs (Fig. 4.33), a cell subset implicated in fibrosis initiation, and its association with elevated but not high fibrosis scores (Fig. 4.38). Together, these observations suggest that CN2 may represent a microenvironmental signature of early fibrotic remodelling.

Expansion of CN8 is associated with high MF-2

We previously observed heterogeneity among MF samples, where expansion of CN8, a region enriched for HSPCs and classical monocytes in MF, was associated with a corresponding reduction in CN2 (the Non-classical Monocytes with Osteo/Stromal and DC/B region) (Fig. 4.26).

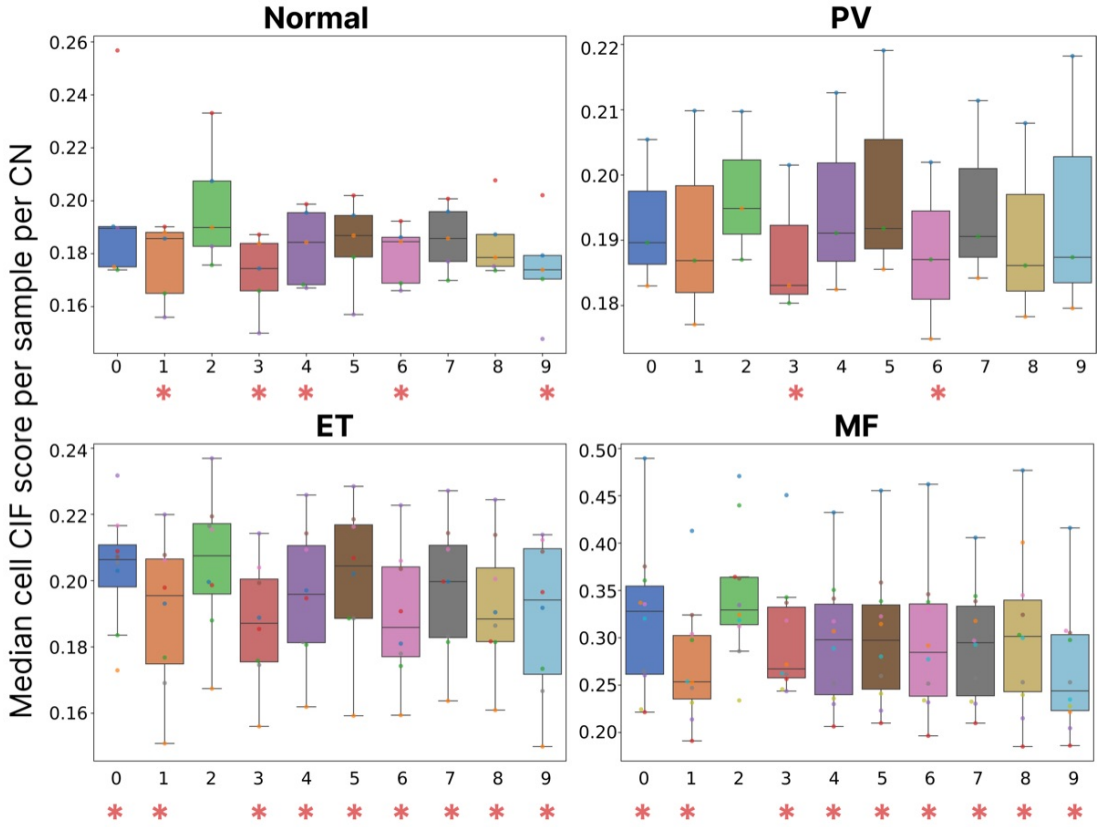


Figure 4.38: CIF score by CellCharter CN within conditions (MF, ET, PV, and Normal). Each dot represents a sample's median CIF score in a given CN, dots are coloured by sample ID. Red stars indicate CNs which have significantly lower CIF scores than CN2.

Initially, I hypothesised that CN8 expansion might reflect overall higher fibrosis. Indeed, CN8 is expanded in two MF samples with widespread fibrosis (10693_R1 and 18599_R1). While 10693_R1 has the highest overall fibrosis, two other samples (18552_R2 and 18256_R1) show higher median fibrosis than 18599_R1. This suggests that total CIF score alone does not explain CN8 expansion (Fig. 4.39).

To investigate further, I examined the fibrosis patterns in all four samples. Prior studies have used fibrosis pattern, rather than total score, to distinguish disease states such as ET and pre-PMF [46]. Visualising fibrosis distribution, a common feature of the two CN8-expanded trephines is the near-complete absence of MF-0 regions (Fig. 4.40). In contrast, the two high-fibrosis samples without CN8 expansion show heterogeneous patterns, containing both fibrotic and non-fibrotic areas.

This led to the hypothesis that CN8 expansion may be linked not to the amount of fibrosis, but to the loss of non-fibrotic regions (MF-0). To test this across all MF samples, I quantified the correlation between CN8 abundance and the proportion of each CIF grade per trephine. This

analysis confirmed a statistically significant negative correlation between CN8 and MF-0, and a positive correlation between CN8 and MF-2 (Fig. 4.41).

Expansion of MF-2 is associated with poorer prognosis. While these results support the idea that CN8 expansion correlates with increased MF-2 and loss of MF-0, additional samples would be needed to confirm whether this represents a consistent subtype within MF.

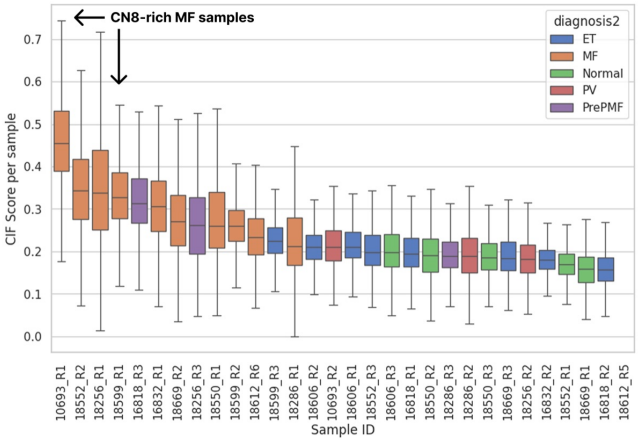


Figure 4.39: Boxplot of CIF scores per sample, calculated at a cell level

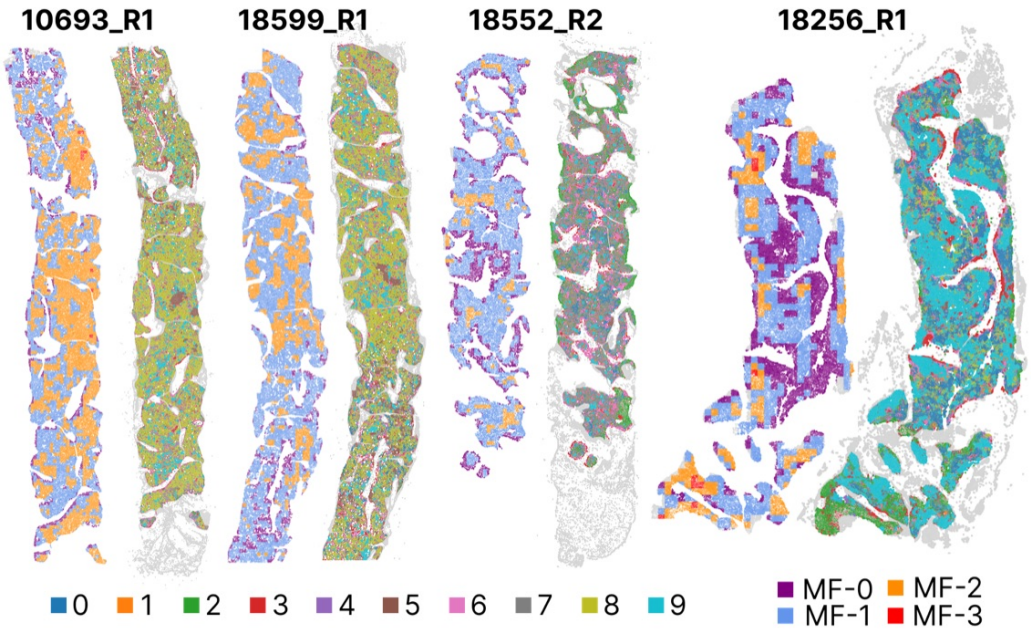


Figure 4.40: CIF score maps for four MF samples with the highest median fibrosis, with accompanying CIF categories and CN maps. First two left samples have expanded CN8.

4.5.4 Discussion

In this section, I used CellCharter to define 10 unsupervised cellular neighbourhoods (CNs) across all bone marrow trephines. These CNs were annotated using cell type enrichment analysis,

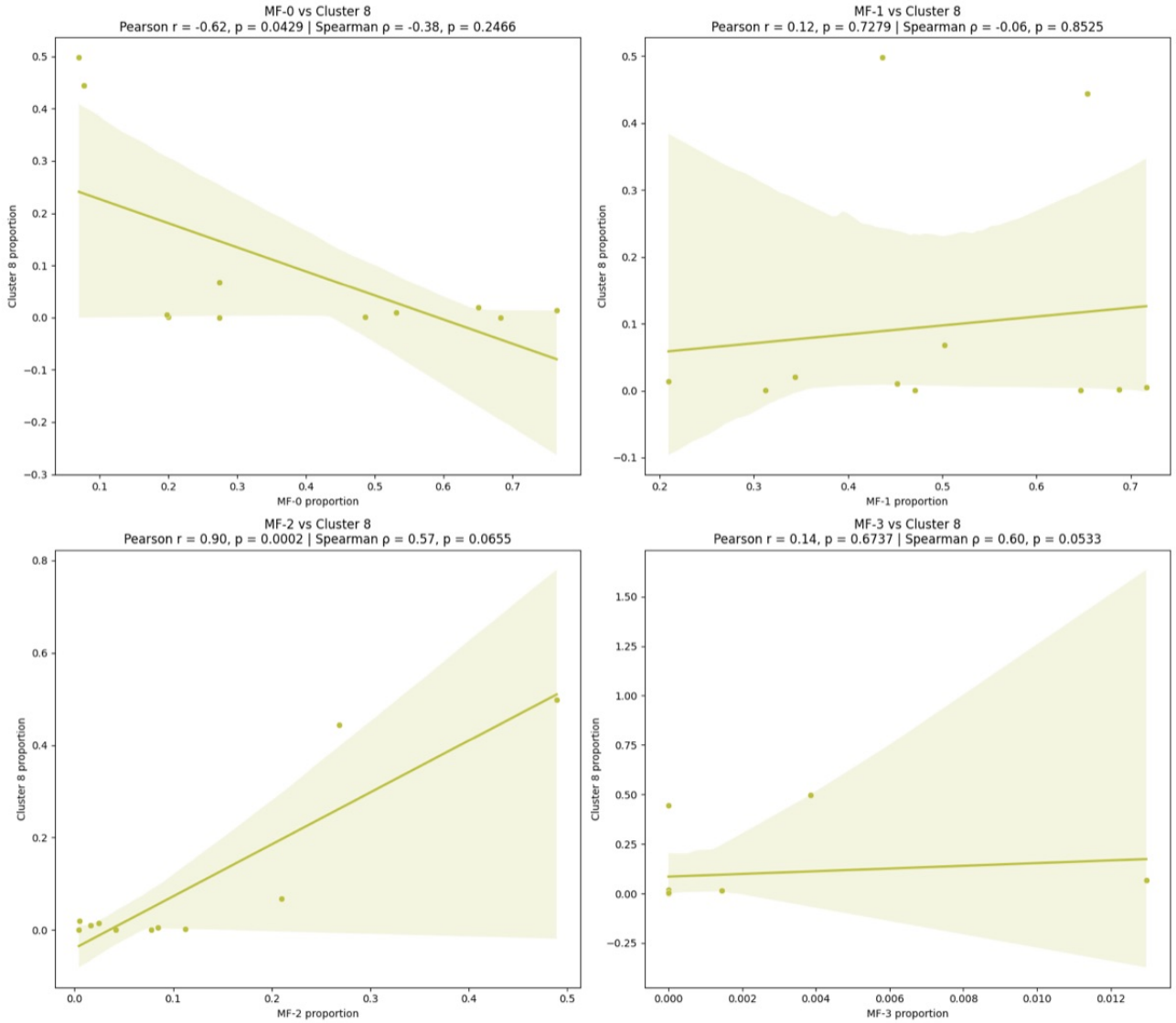


Figure 4.41: Correlation between the proportion of CellCharter region 8 and the proportion of each CIF grade per trephine in MF samples

identifying three stable HSPC-enriched CNs and seven others spanning immune, stromal, and myeloid populations (Fig.4.23). Differential gene expression between CNs revealed expected lineage-associated markers, except for CN2, which lacked any distinguishing gene expression signature compared to other CNs.

At the whole-sample level, no statistically significant differences in CN abundance were observed between conditions (Normal, ET, PV, MF) (Fig.4.27). Despite this, MF samples appeared strikingly distinct upon visual inspection and also showed substantial inter-sample heterogeneity. This prompted the hypothesis that differences between conditions, particularly between fibrotic and non-fibrotic MPNs, might manifest through spatial distribution rather than total abundance. Indeed, two MF samples were outliers with expanded CN8 (an HSPC niche with Classical

Monocytes) and a concurrent loss of CN2 (Non-classical Monocytes with Osteo/Stromal and DC/B) (Fig.4.26b).

To quantify spatial expansion, I assessed the size of the ten largest connected components (regions of spatially contiguous cells from the same CN) in each sample. These were significantly larger in MF than in other conditions (Fig. 4.28). However, MF samples also retained small clusters, suggesting that CN expansion does not replace smaller components but coexists with them (Fig.4.29, 4.30). Thus, spatial expansion of CNs is a distinguishing feature of MF.

When investigating which CNs are prone to expansion, Normal samples rarely showed large connected components. CN1 (Erythroid/Macrophage), CN3 (HSPC niche with Endothelial/SMC and Plasma/B cells), and CN4 (HSPC niche with Erythroid cells) were consistently small, suggesting they represent stable, potentially non-mutated or WT niches. In contrast, CN0 (Myeloid trajectory with stromal), CN2, and CN9 (HSPC niche with Megakaryocytes and Endothelial/SMC) were significantly expanded in MF. CN9 was also expanded in ET, which is notable given ET's hallmark of increased platelet production, implicating CN9 in aberrant megakaryopoiesis. Interestingly, CN6 (GMP/Stromal) and CN7 (non-specific myeloid) were expanded in ET but not in MF. In Normal and PV samples, they were also the most commonly expanded CNs, though not significantly. This raises the possibility that CN6 and CN7 could play a stabilizing or protective role, buffering against expansion of CN0 and CN2, which are prevalent in MF. CN8, while present at low levels in other conditions, is HSPC-enriched only in MF and highly skewed toward large clusters in just two MF samples. Its selective expansion and HSPC content suggest it may mark a unique disease microenvironment, potentially associated with poor prognosis (Fig. 4.33).

Next, I examined inter-CN relationships across spatial neighborhoods. Certain CN-CN associations, such as CN1-CN4, CN3-CN6, and CN4-CN7, were conserved across all conditions (Fig.4.32). These CNs were also less prone to expansion and enriched for HSPCs, supporting their role as components of the healthy marrow microenvironment. CN3 and CN4 may represent WT HSPC niches with myeloid and erythroid bias respectively, particularly given CN3's consistent colocalisation with CN6 (GMP-rich).

Cell enrichment analyses revealed that CN3, CN4, and CN9 are HSPC-enriched across all conditions, whereas CN2 and CN8 are condition-specific. CN2 is enriched for HSPCs in non-fibrotic MPNs (ET, PV), while CN8 is enriched only in MF.

CN5, enriched in B, T, and DC cells (and macrophages in MF), is likely a site of lymphocyte activation and antigen presentation. In Normal samples, CN5 was significantly associated with CN8, but was not in disease states. Given CN8's spatial enlargement and shift to HSPC enrichment in MF, I hypothesised that CN5 might represent a site of immunosurveillance of CNs with poor prognostic outcomes which is lost in MPN. Supporting this, in the healthy bone marrow, CN5 is significantly enriched as a neighbour of CN0, CN2, CN7 and CN9 compared to ET, and with CN0 and CN9 compared to MF. CN0, CN2 and CN9 are CNs which we previously found significantly expanded in MF, whilst CN7 and CN9 were expanded in ET (Fig. 4.31). This suggests that the loss of immunosurveillance, normal in healthy bone marrow, may allow these clusters to expand in ET and MF.

I then investigated CN localisation relative to bone. CN3 and CN6, previously defined as stable and non-expanded in MF, aligned with traditionally defined regions, the endosteal and peritrabecular, respectively (Fig. 4.35, 4.36). This supports the biological plausibility of CellCharter's unsupervised spatial segmentation.

Disruption of the region close to bone appears to be a key feature of MPN. In MF, CN9 (Megakaryocyte/Endothelial niche) expands into the endosteal and peritrabecular regions, displacing CN6. CN9 is associated with loss of immunosurveillance, expansion in fibrotic disease, and contains endothelial and megakaryocytic cells, both of which globally expressed SNAI1 in MF, an endothelial–mesenchymal transition marker implicated in fibrosis [49]. Given the known pro-fibrotic role of TGF-beta and CXCL4, both secreted by megakaryocytes [47, 56], CN9 may represent a fibrotic niche. Future studies could investigate whether megakaryocytes in CN9 express markers of endothelial-mesenchymal transition or possess distinctive morphologies that could be recognized on H&E.

CN2, enriched for Osteo-lineage and stromal cells, was significantly expanded in MF and became increasingly distal from bone as disease progressed (Normal/PV to ET to MF) (Fig. 4.36). This spatial drift mirrors fibrosis progression and suggests CN2 could be an early correlate of fibrotic remodeling. Supporting this, CN2 had higher CIF scores than other CNs in both ET and MF,

but still below MF-2 thresholds, reinforcing the notion that it may mark early-stage fibrosis rather than late fibrotic transformation (Fi. 4.38).

Finally, within MF, CN8 expansion was observed only in two samples, both of which had reduced CN2 and were associated with MF-2 rather than MF-0 scores (Fig. 4.41). This suggests that CN8 may emerge later in fibrosis, potentially following CN2-driven microenvironmental changes. Whether CN8 represents an alternate fibrotic pathway or a terminal stage of CN2 progression is an open question.

More fundamentally, future work will need to address whether CN2 represents the beginnings of well-defined aberrations in the marrow microenvironment that may precede the development of fibrosis and predict disease progression. Importantly, if CN2 is a marker of fibrosis development, it would be interesting if we could link it to morphology available in the clinically available H&E slide. Identifying latent features in non-fibrotic MPN that predispose to the development of fibrotic MF, especially with the explainability that comes from spatial transcriptomics for therapeutic intervention, would be a major goal.

4.6 Fibrosis and Bone Marrow Changes in MF

4.6.1 Introduction

The analysis above relied on 10 CellCharter defined CNs. Whilst CNs have been helpful for understanding changes between conditions and do correspond to known regions of the bone marrow such as the endosteal and peritrabecular space, the division into 10 CNs, rather than any other number, was manually determined. This will of course affect our interpretation of the tissue.

Our focus is on understanding the progression of fibrosis. Myelofibrosis (MF), the most advanced form of MPN, is associated with high-grade fibrosis and poor prognosis. ET and PV can evolve into MF, but the underlying drivers of fibrotic transformation remain unclear. Identifying early fibrotic changes is key to developing interventions that prevent or delay disease progression.

Fibrosis is traditionally assessed using a reticulin stain on bone marrow trephines, which highlights reticulin fibers. In healthy marrow, these fibers form a fine, uniform meshwork. In fibrosis, they become thickened, intersecting, and disorganized. Fibrosis is graded from MF-0

to MF-3, reflecting increasing reticulin deposition, collagen bundling, and osteosclerosis. The reticulin stain is typically performed on a separate section from the H&E.

To quantify fibrosis, our group previously developed an algorithm that assigns each tile of a trephine a continuous fibrosis (CIF) score between 0 and 1 based on the reticulin stain [46]. Recently, Ryou et al. [57] extended this approach to predict CIF scores directly from H&E images; eliminating the need for a reticulin stain. We have applied this updated method to our dataset.

To build on the CellCharter analysis, I now investigate how cellular composition, cellular relationships and expression change across increasing CIF scores and between fibrosis grades.

4.6.2 Methodology

Visualisation of Fibrotic Grades and Domains in MF

As shown in Figure 4.42, fibrosis is spatially heterogeneous within trephines. To capture this, we quantified fibrosis at the 512×512 pixel tile level on H&E-stained images, generating spatial maps of CIF scores and corresponding fibrosis grades across each sample (Fig.4.42). Tiles were extracted with a stride of 128 pixels. Following Ryou et al. [46], CIF scores, with a range 0-1, were mapped to fibrosis grades using these thresholds: MF-0 <= 0.267, MF-1 > 0.267, MF-2 > 0.439, MF-3 > 0.647

To assign each cell a fibrosis score, I calculated the mean CIF score of all overlapping tiles in which the cell was located, using a geopandas.sjoin with the "within" predicate to identify spatial overlap. Although MF grades are clinically useful, they are based on morphological criteria and may not reflect the underlying cellular processes. Therefore, in addition to analysing the tissue by continuous CIF scores and standard MF grade boundaries, I also introduced an intermediate classification: fibrosis hotspots. Given that a whole-trephine CIF score MF-2 is associated with poorer prognosis, I defined hotspots as regions with CIF > 0.543 (midway through MF-2), and lowspots as regions with CIF <0.35 (midway between MF-1 and MF-2) (Fig. 4.42).

Cellular Composition of Fibrosis Grades in MF

To understand which cells are present across different levels of fibrosis in MF, I analysed cellular composition relative to fibrosis severity using three approaches. While CIF is a continuous score,

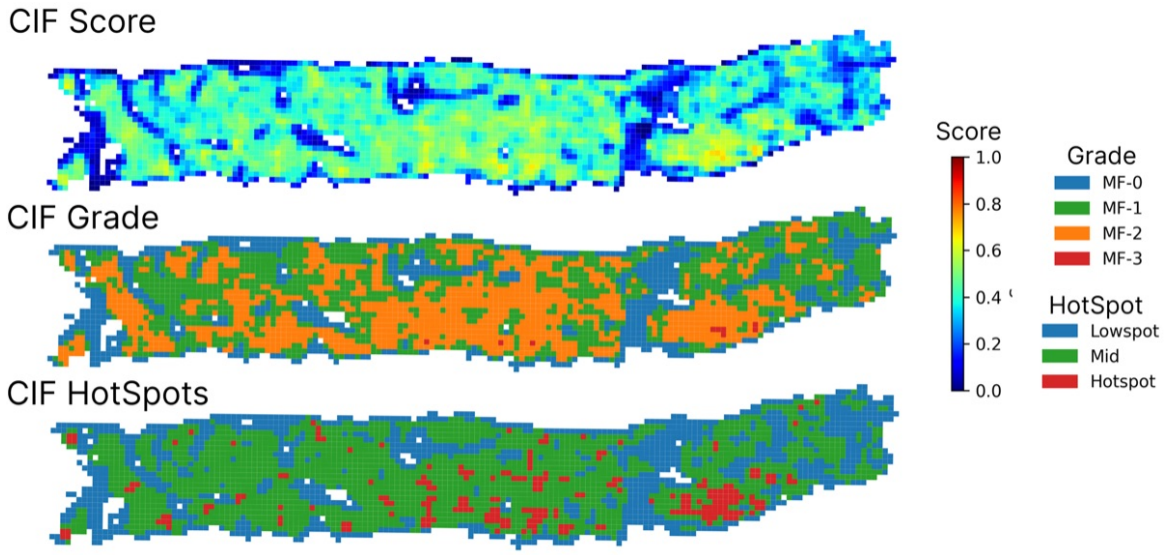


Figure 4.42: 512×512 tiles in an MF trephine, coloured by fibrosis score, fibrosis category, and fibrosis hotspot classification.

conventional fibrosis grading (MF-0 to MF-3) is categorical. Therefore, I first examined cell type composition as a function of continuous CIF score, then by discrete MF grades. Since MF-3 is rare in our dataset, even within MF samples, I also redefined the CIF thresholds to divide the tissue into low, mid, and high fibrosis regions, enabling a more balanced analysis across fibrosis severity.

Continuous Evaluation

To evaluate how cellular composition varies with fibrosis severity in MF, I binned CIF scores into intervals of 0.025. I then used a mixed effects model to assess the relationship between CIF score and the proportion of each cell type, while accounting for patient ID as a random effect. This was implemented using smf.mixedlm, with the model formula: proportion ~score_bin (Table 4.2).

Categorical Evaluation

I grouped cells into categorical fibrosis regions according to their fibrosis score (either MF grades or fibrosis hotspots) and performed differential abundance analysis using edgeR controlling for patient id. CIF boundaries for MF fibrosis grades were: MF-0 <= 0.267,MF-1 > 0.267, MF-2 > 0.439, MF-3 > 0.647. Whilst I defined fibrosis hotspot analysis according to the following boundaries: hotspot CIF > 0.543, lowspot CIF < 0.35, all other regions marked as mid.

Cellular Interactions within Fibrosis Grades in MF

Neighbourhood Enrichment

I used an adaptation of Squidpy's neighbourhood enrichment function, as used in CellCharter, to assess whether pairs of cell types are spatially enriched beyond random expectation [113]. I used the Voronoi-derived connectivity graph, devloped in the previous chapter, and compared observed neighbourhood frequencies against 100 random permutations of cell labels, allowing same-type neighbours. 100 permutations limits the minimum p-value to 0.01. This approach does not account for inter-patient variability, as sample ID is not included in the model.

Differential Neighbourhood Enrichment

I again used a CellCharter function, diff_nhood_enrichment, which tests for differential neighbourhood enrichment by permuting sample labels between conditions. Since our samples are paired, I created a composite column combining sample ID and grade to use as the library key. Observed differences were compared against expectations calculated from 100 permutations.

Determining a suitable radii in MPN for co-occurence analysis

To assess the correct radii range to use, given the differences in cellularity, I visualised potential bands on both MF and Normal. A cell is only included within a band if the center point, represented by a dot, is within the band. Each pixel corresponds to 0.2125um. Therefore, 50 pixels is only around 10 um.

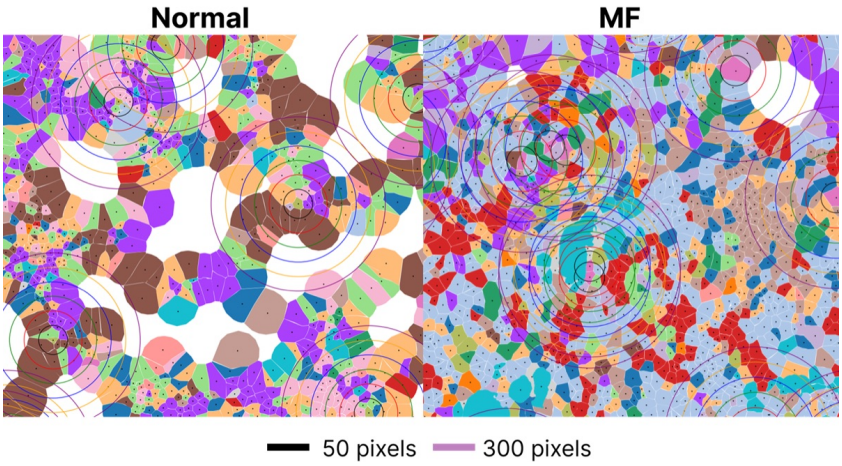


Figure 4.43: Comparison of cell-type spatial neighborhoods in Normal and MF samples. Radii are shown at 50-300 pixels in 50-pixel intervals. Cell center points are indicated by dots; cell colours indicate cell types.

Cellular Expression across Fibrosis Grades in MF

I used pseudo-bulk DESeq2 controlling for sample id in the paired analysis. I performed pairwise comparisons across all 6 combinations of fibrosis grades, and after multiple hypotheisis testing only considered those with an adjusted p value less than 0.05.

4.6.3 Results

Cellular Composition of Fibrosis Grades in MF

First, I aimed to identify which cells and cell–cell relationships are associated with increasing fibrosis grades in MF. Cell types or interactions that become more prevalent with fibrosis severity may represent key drivers of fibrotic progression.

Correlation of cell type abundance with CIF score in MF

To visualise the general trend, I plotted the proportions of each cell type as CIF score increased, binned at intervals of 0.025 (Fig. 4.44). Immediate observations suggest there may be a decrease in Erythroid cells and an increase in Monocytes and Myeloid cells with increasing CIF score.

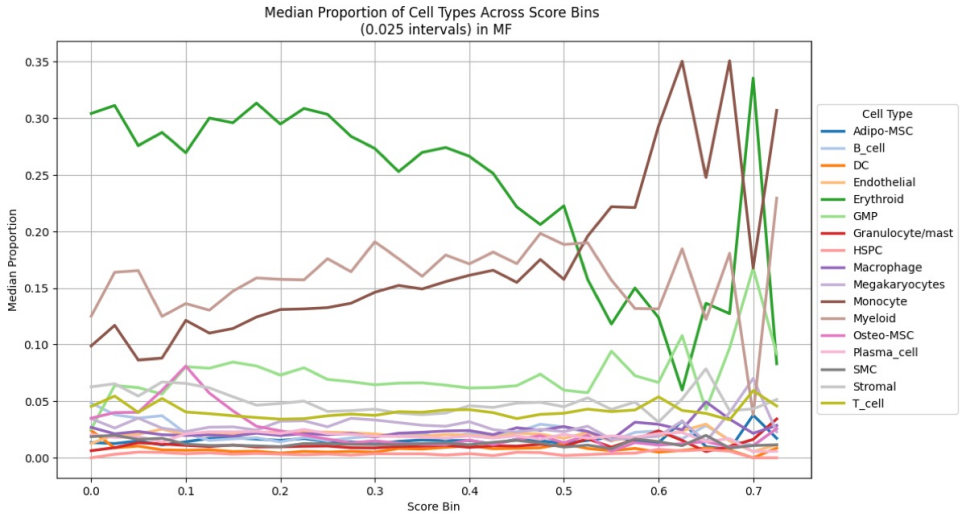


Figure 4.44: Median of the proportion of each cell type per MF sample as CIF score increases at bins of 0.025. Of note, there are fewer cells at high cif tiles, accounting for the loss in trend

I used a mixed effect model to control for patient id as a random effect, whilst modelling the relationship between CIF score (at 0.025 intervals) and the proportion of each cell type. I used smf.mixedlm with the comparison proportion ~score_bin (Table 4.2). Plotting only the cells

which significantly correlate with CIF score we can evaluate correlations at an individual cell type level (Fig. 4.45).

Cell Type	Coef	P-value	CI Lower	CI Upper	P-value Adjusted
Monocyte	0.174622	4.35e-36	0.147336	0.201909	7.40e-35
Myeloid	0.057262	7.80e-06	0.032157	0.082366	3.32e-05
Macrophage	0.011543	3.89e-04	0.005166	0.017920	1.32e-03
Granulocyte/mast	0.005418	1.06e-02	0.001265	0.009572	2.99e-02
Megakaryocytes	-0.013636	2.03e-02	-0.025148	-0.002123	4.92e-02
Osteo-MSC	-0.079328	2.83e-19	-0.096652	-0.062005	2.41e-18
Erythroid	-0.151327	5.08e-15	-0.189230	-0.113423	2.88e-14

Table 4.2: Statistical results for different cell types using a mixed effects model with sample_id as a random effect and modelling the relatonship between CIF score and proportion at 0.025 intervals for CIF score. p values adjusted with FDR

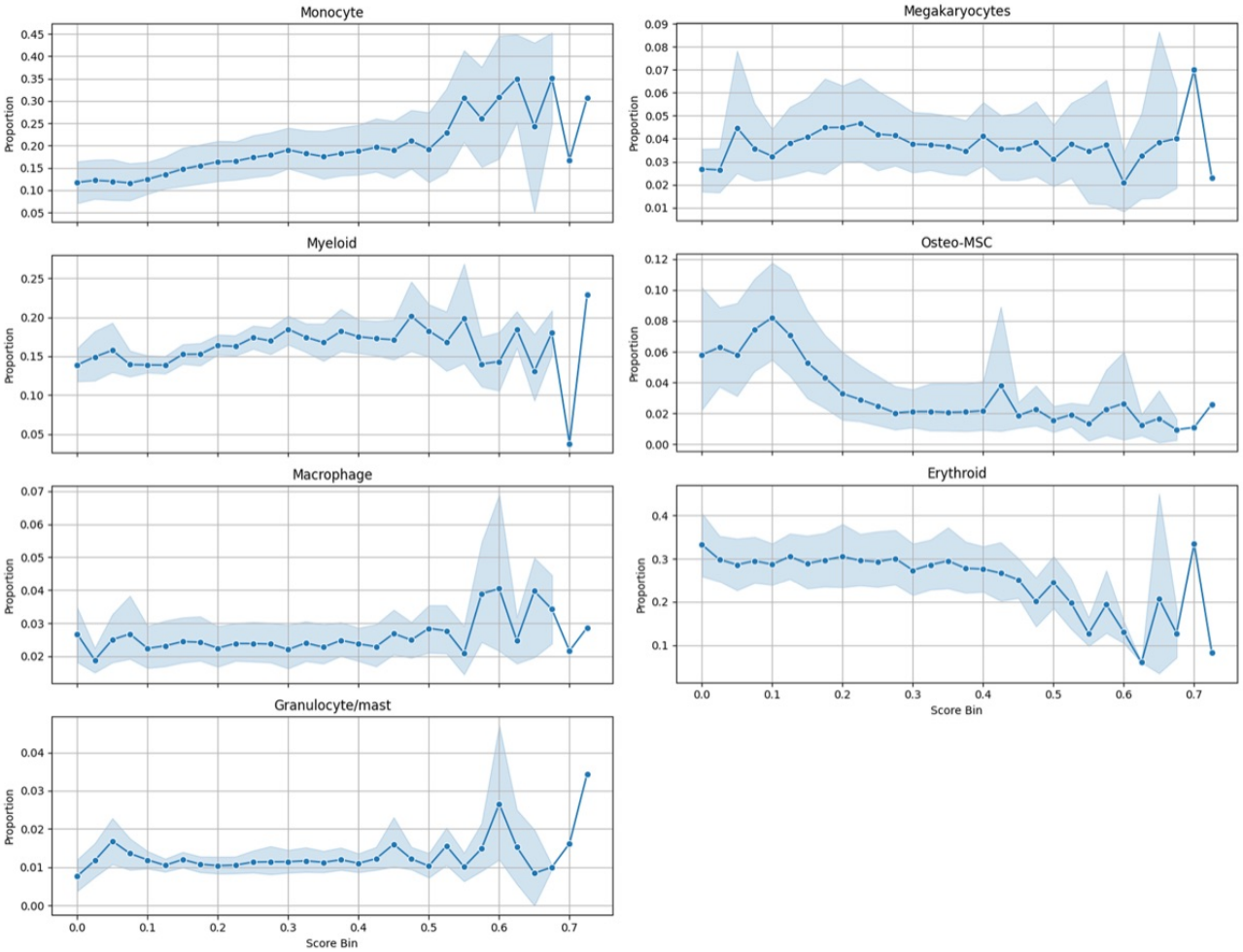


Figure 4.45: Mean proportion (as a percentage of all cells) across all samples per 0.025 CIF score increase. Alpha represents 95% confidence interval.

Although effect sizes are small, monocytes, myeloid cells, macrophages, and granulocytes/mast cells increase with CIF score, while megakaryocytes, Osteo-MSCs, and erythroid cells decrease

as fibrosis progresses (Fig. 4.45). These small coefficients may reflect averaging over the whole CIF range; future work could focus on specific CIF score ranges to better capture localized effects.

Monocytes are known to secrete inflammatory cytokines and may help establish a pro-fibrotic environment. Surprisingly, however, both megakaryocytes and Osteo-MSCs, cell types implicated in fibrosis, decline with increasing fibrosis, despite being globally elevated in MF compared to other conditions (Fig. 4.4, 4.45). This suggests they may play a role early in fibrosis development but are reduced once fibrosis is established.

The global rise in megakaryocytes in MF is largely due to non-morphologically phenotyped cells, though phenotype 8 is also elevated in ET and MF relative to healthy bone marrow (Fig. 4.3). To assess whether megakaryocyte phenotype is linked to fibrosis severity, I binned cells by CIF score (0.1 intervals) and tested for associations (not shown). However, no significant correlations were found after FDR correction.

Differential abundance across MF fibrosis grades

As cell proportions may not change linearly with CIF score, continuous modeling may miss associations, especially if they occur only within narrow fibrosis intervals. To address this, I performed differential abundance analysis within MF using edgeR, controlling for patient ID. Analysis was conducted across fibrosis grades, aligning with clinical relevance and pathology workflows (Fig. 4.46).

As with the continuous CIF score analysis, Osteo-MSCs decreased consistently with increasing fibrosis grade, despite being globally elevated in MF compared to other conditions. This supports the idea that a reservoir of Osteo-lineage cells may exist at earlier fibrosis stages. Prior studies have linked NCAM1-expressing Osteo-lineage cells, expanding from peritrabecular to central marrow, with fibrosis development [56]. Given the limited resolution of our panel, our Osteo-MSC annotation likely includes both osteoblasts and osteoclasts. It is plausible that the Osteo-MSCs persisting at higher fibrosis grades are the NCAM1+ cells driving fibrotic progression.

As before, Erythroid cells decreased with fibrosis grade. GMPs now also show a reduction at MF-2 vs MF-0. In contrast, Monocytes increased from MF-0 to MF-2, echoing earlier results. Stratifying by fibrosis grade also revealed new findings: both B cells and DCs increased in MF-2

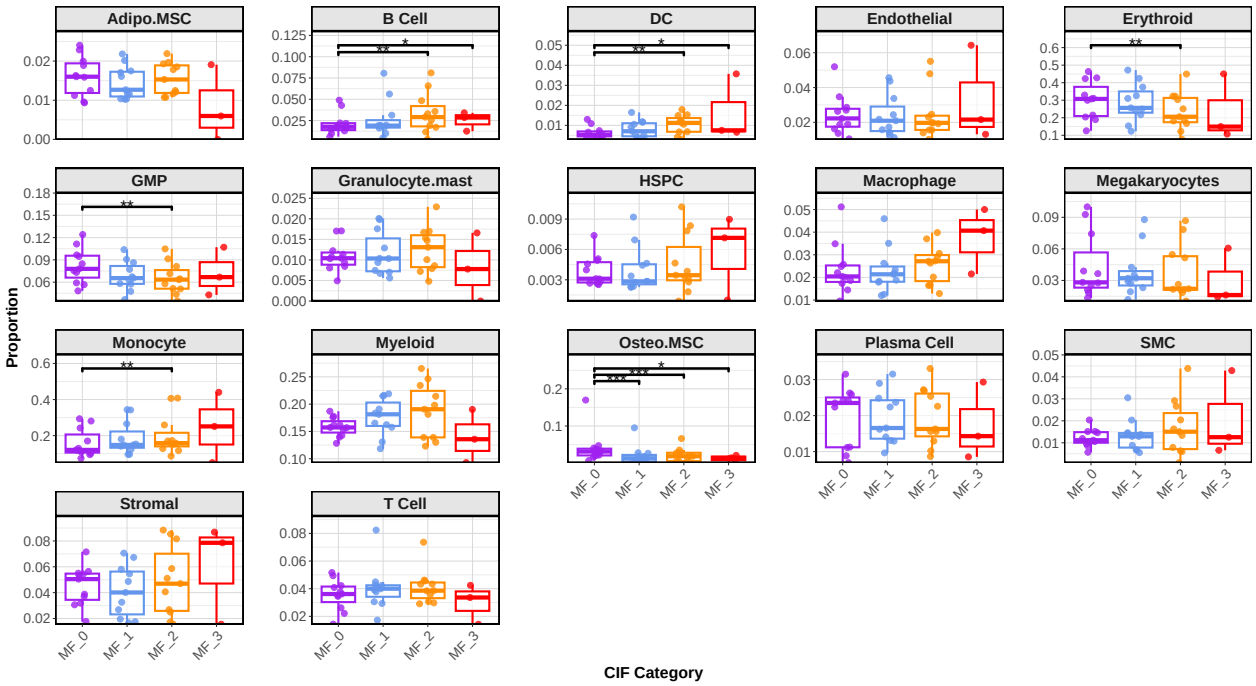


Figure 4.46: Cells were grouped into fibrosis grades across MF samples only. Differential abundance was calculated using edgeR controlling for patient id.

and MF-3 compared to MF-0, suggesting their enrichment in more fibrotic regions. However, very few samples (n=3) had extensive MF-3 fibrosis, which likely limited statistical power.

Differential abundance across MF hotspots

Due to the low number of MF-3 samples (n=3), I redefined the fibrosis categories by lowering the threshold for high fibrosis to the midpoint of MF-2 (CIF more than 0.543). Tiles with CIF less than 0.35 (midpoint of MF-1) were defined as lowspots, with all others as midspots. This expanded the high fibrosis group to n=7 samples.

Focusing on lowspots vs highspots, both osteo-MSCs and erythroid cells were consistently reduced in highspots. They were also lower in highspots compared to midspots, suggesting a gradual decline with increasing fibrosis (Fig. 4.47). This threshold adjustment also revealed a decrease in megakaryocytes in highspots. DCs remained elevated in both highspots and midspots compared to lowspots, while B cells were increased in midspots only. However, we no longer observed the increase in monocytes or myeloid cells seen in previous analyses using continuous CIF or MF grades. These patterns align with findings from CN2, which is enriched for osteo-MSCs, B cells, DCs, and monocytes. We had previously identified CN2 as a potential

early marker of fibrosis. Future work should investigate whether classical and non-classical monocytes contribute differently to fibrosis.

Most changes in cell abundance occur between MF-0 and MF-2, the latter associated with worse prognosis. This suggests that cellular changes may accompany fibrotic progression and represent potential therapeutic targets. Although CN2 was originally proposed as an early fibrosis correlate, the enrichment of its cell types in MF-2, paired with a drop in osteo-MSCs, suggests it may be a cellular correlate of fibrosis. Delineating whether the CN2 osteo-MSCs represent the profibrotic NCAM1+ osteo cells may explain the decline of osteo-mscs with fibrosis but enrichment in CN2.

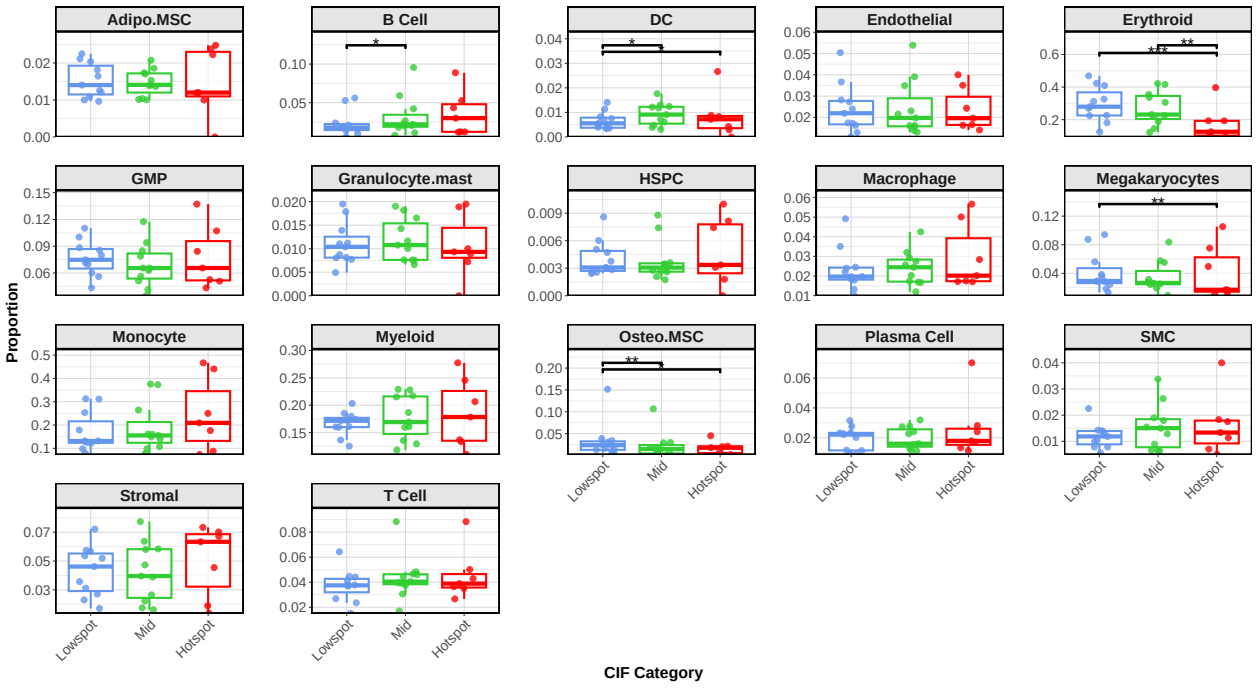


Figure 4.47: MF samples ($n = 11$) comparison of abundance of cell types in fibrotic hotspots using edgeR corrected for sample_id. Hotspot are cells with cif score more than 0.543 and Lowspot is cif score below 0.35

Cellular Interactions across Fibrosis Grades in MF

An increase in cell abundance with rising CIF scores suggests potential involvement in fibrosis. To further explore which cell types might actively contribute to fibrotic progression, I next examined whether these increasing cell types also showed greater spatial association with each other, indicative of potential interaction. For completeness, I also included adipocytes in the analysis and assigned them mean CIF scores using the same method as for other cell types.

Neighbour Enrichment within Fibrosis Grades in MF

I first assessed neighbourhood enrichment within fibrosis grades by comparing observed cell–cell associations against a null distribution generated from 100 permutations of cell labels within each region. (Fig. 4.48).

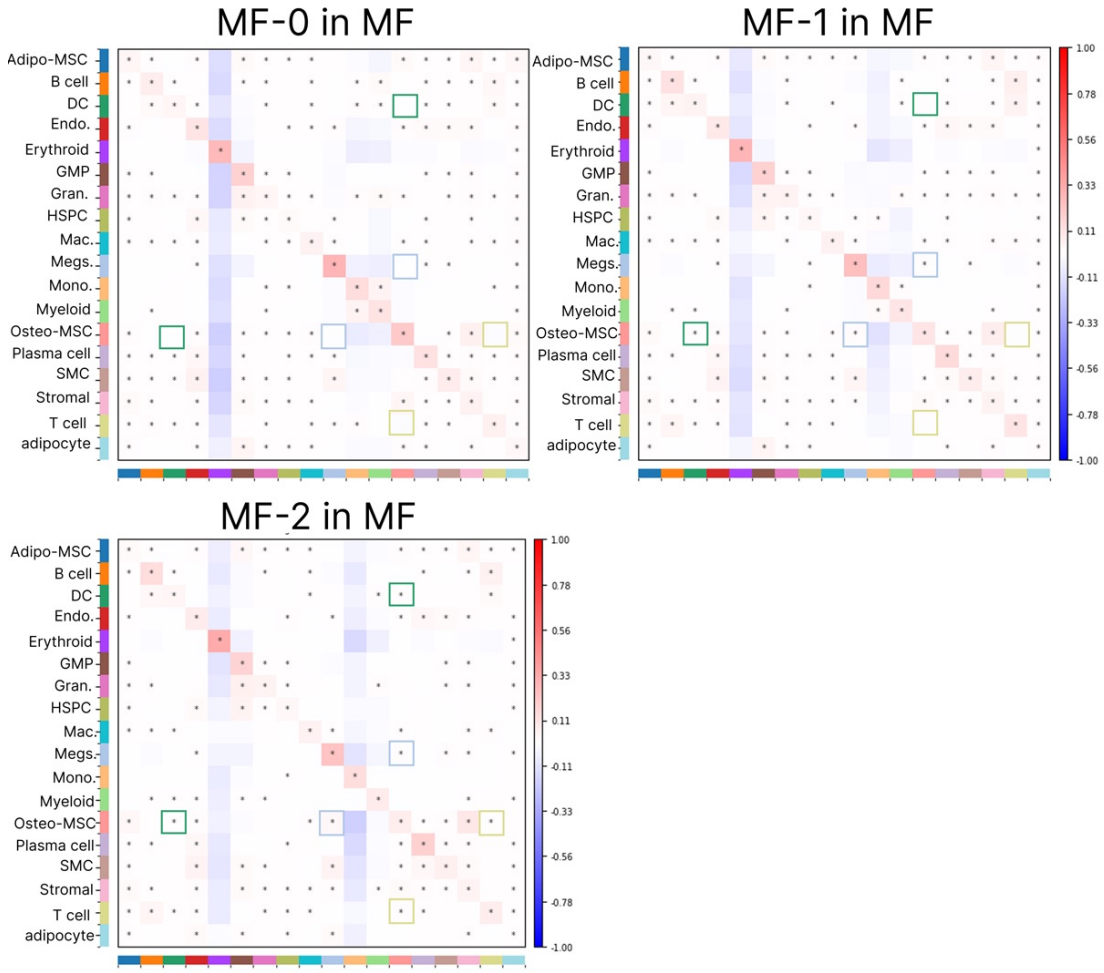


Figure 4.48: Neighbourhood enrichment analysis across MF fibrosis grades (MF-0 to MF-2) using `anno_3` with refined stromal populations and fat cells. Colours represent $\log_2$(observed/expected) after 100 permutations of cell labels. Asterisks indicate significance at $p < 0.05$.

Across MF grades 0 to 2, a conserved T–B–DC module was significantly enriched, defined as a group of cell types with mutually enriched spatial interactions. This likely corresponds to CN5, previously identified as a lymphoid activation region. Similarly, an Endothelial–Osteo–Stromal module, also including Plasma cells and SMCs at MF-1 and MF-2, was consistently enriched, likely reflecting CN3 (the endosteal space) (Fig. 4.48).

Enrichment between HSPCs and endothelial cells, and between HSPCs and GMPs (but not GMPs and endothelial cells), was also preserved across grades. Megakaryocytes were consistently

associated with endothelial cells across all grades, but only showed significant enrichment with Osteo-MSCs at MF-1 and MF-2. Notably, while Osteo-MSCs were not associated with the T–B–DC module in MF-0, they became enriched with DC cells at MF-1, and formed a T–DC–Osteo-MSC module at MF-2 (Fig. 4.48).

Neighbour Enrichment compared to MF-0

To assess whether cellular relationships change significantly with fibrosis progression, I performed differential neighbour enrichment analysis comparing each fibrosis grade to MF-0. MF-0, representing early-stage fibrosis, was used as the reference baseline to identify changes as fibrosis severity increases (Fig. 4.49).

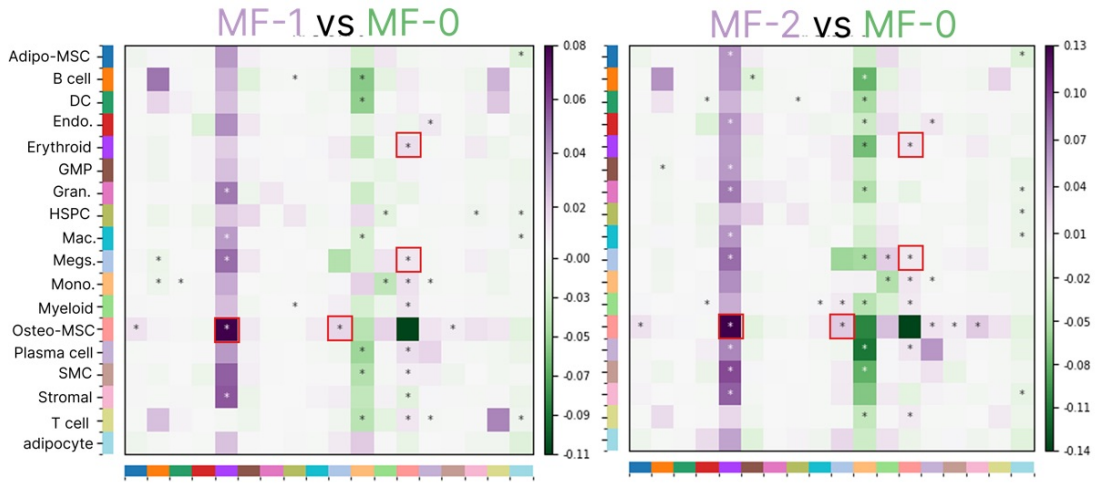


Figure 4.49: Neighbourhood enrichment differences relative to MF-0 across fibrosis grades. Colours represent $\log_2$(observed/expected) of neighbourhood interactions; significance assessed via 100 permutations of condition labels. Purple indicates greater neighbourhood likelihood in MF-x; green indicates greater likelihood in MF-0. Asterisks represent $p < 0.05$.

In both MF-2 vs MF-0 and MF-1 vs MF-0, one of the most striking findings is the significant enrichment of various cell types with erythroid cells at higher fibrosis grades. Among these, only the interaction between erythroid cells and Osteo-MSCs is reciprocally enriched in both comparisons. Conversely, at MF-0, several cell types show significantly stronger associations with monocytes compared to higher fibrosis grades. Notably, only the monocyte–plasma cell interaction is reciprocally enriched at MF-0 vs MF-2, while monocyte interactions with plasma cells, DCs, and B cells are reciprocally enriched at MF-0 vs MF-1. This suggests that monocyte–plasma cell interactions are consistently enriched at lower fibrosis levels. A final shared observation is the increased association of many cell types with Osteo-MSCs at higher fibrosis grades. However, only the interactions between Osteo-MSCs and megakaryocytes, as

well as with erythroid cells, are reciprocally enriched at both MF-1 and MF-2 compared to MF-0 (Fig. 4.49).

Together, these findings suggest that monocyte–plasma cell interactions may play a role early in fibrosis (or may not be involved in fibrosis), while Osteo-MSC interactions with megakaryocytes and erythroid cells become more prominent later. This is especially interesting given that all three cell types decrease in abundance as fibrosis progresses.

Co-occurrence within MF fibrosis grades

The previous analysis focused solely on direct neighbour enrichment. To explore whether different patterns emerge at longer spatial ranges, I calculated co-occurrence probabilities across multiple radii for each fibrosis grade. Using Squidpy's co-occurrence function, I assessed cell–cell interactions at 50-pixel intervals from 150 to 500 pixels, ensuring sensitivity to interactions involving smaller cells. In Fig. 4.50, I show the median co-occurrence probabilities across all samples for each cell–cell pair at each radius.

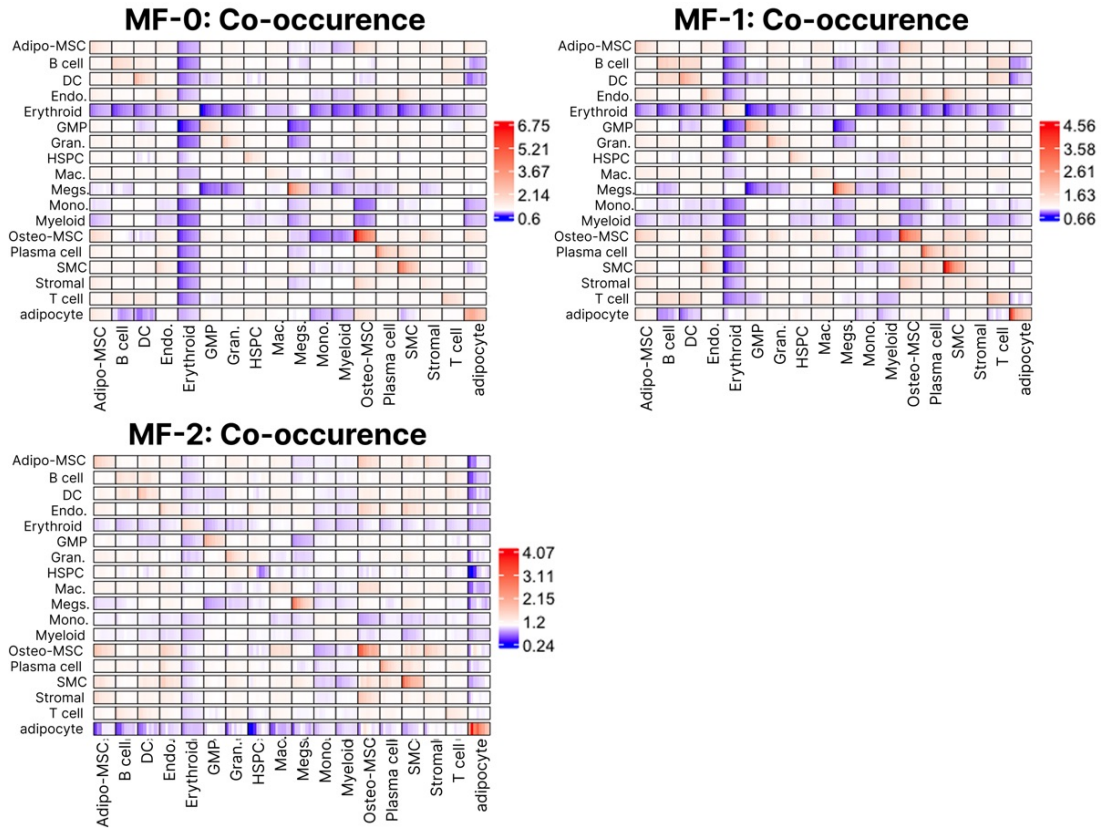


Figure 4.50: Co-occurrence of cell types across MF fibrosis grades (MF-0 to MF-2) at radii 150–500 pixels using `anno_3` with refined stromal and fat populations. Colours show the median $\frac{p(\text{cellA}|\text{cell_of_interest})}{p(\text{cellA})}$ across samples per fibrosis grade. X-axis: cell of interest; Y-axis: co-occurring cell type. White (1) indicates no enrichment.

Although this analysis does not include significance testing, it reveals patterns not captured by direct neighbourhood enrichment. For instance, HSPCs in MF-0 and MF-1 consistently co-localise across all radii, while in MF-2 they are only positively co-localised at short distances and negatively at longer ones, suggesting increased spatial dispersion and possible niche disruption in higher fibrosis (Fig. 4.50).

The association between Osteo-MSCs and Megakaryocytes persists at all distances in MF-2, but is absent in MF-0, consistent with previous enrichment findings. Interestingly, Osteo-MSCs no longer co-localise with Erythroid cells at any grade. We observe Osteo-MSC associations with B and DC cells in higher fibrosis but not in MF-0. T cell co-localisation with Osteo-MSCs appears unchanged across grades. Additionally, while Monocyte–Plasma cell co-localisation is not enriched in MF-0, we observe increasing negative co-localisation in higher fibrosis, suggesting the earlier pattern reflects de-enrichment at later stages rather than enrichment at MF-0 (Fig. 4.50).

To assess significance, I next examined co-localisation at a fixed radius. I selected 250 pixels (approx. 50 microns) and tested whether observed co-occurrence exceeded a null model (probability = 1), correcting for multiple comparisons. Given our focus on Osteo-MSCs, I specifically assessed whether they significantly co-localise with other cell types at this scale across MF grades (Fig. 4.51).

Across all fibrosis grades, Osteo-MSCs significantly co-occur with endothelial cells and SMCs, suggesting a persistent association with arteriolar structures. Surprisingly, despite the overall increase in monocytes with fibrosis, particularly in MF-2, Osteo-MSCs do not significantly co-localise with monocytes at any grade, nor with myeloid cells at MF-0 and MF-1. However, they consistently co-localise with macrophages across all grades (Fig. 4.51).

At 250-pixel radius, Osteo-MSCs still show positive co-localisation with megakaryocytes in MF-1, but not in MF-2. Notably, they co-localise with T cells and DCs in MF-1 and with B cells in MF-2. Analyses across CNs and co-localisation suggest a key interaction network among Osteo-MSCs, B cells, DCs, and T cells, and also Osteo-MSCs and Megakaryocytes in advanced fibrosis (Fig. 4.51).

CN2 increases with higher fibrosis grades, while CN9 is enriched at lower grades

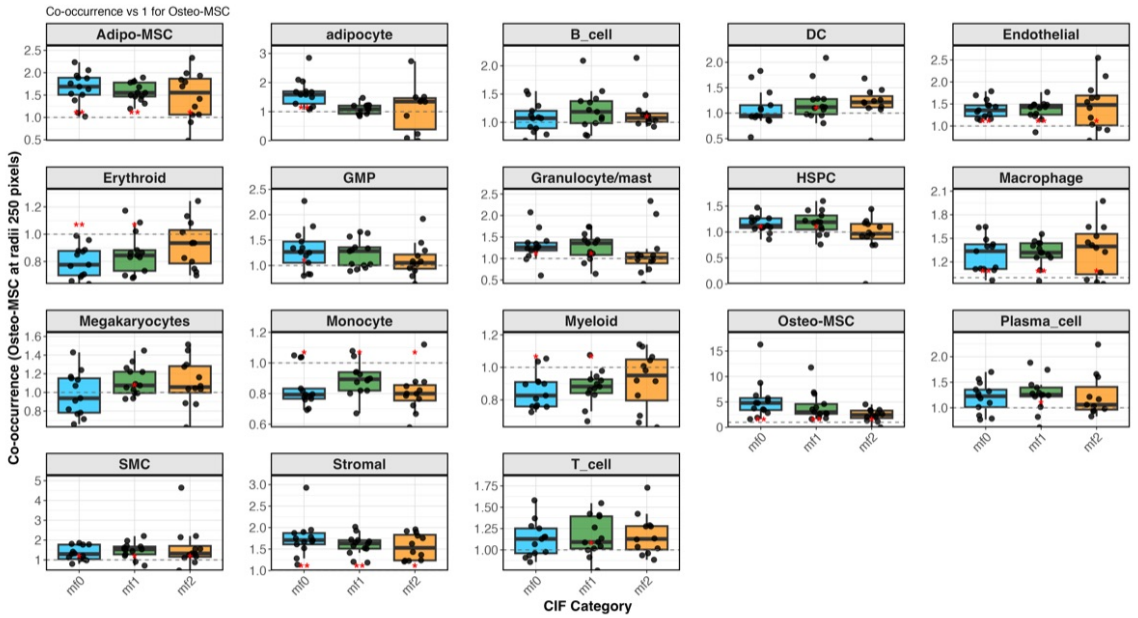


Figure 4.51: Co-occurrance of Osteo-MSCs across MF grades at 250 pixel radii. Wilcoxon test was performed comparing the distribution of each grade to the null hypothesis of 1. Red star indicates significance < 0.05, compared to the null

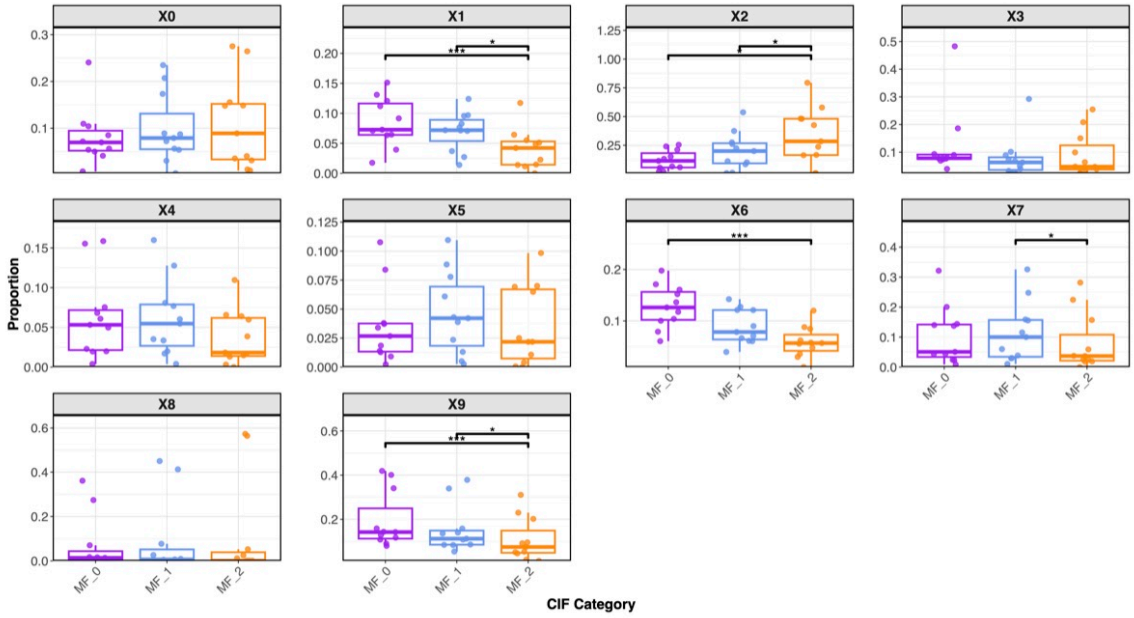
Although defined using genome-wide expression rather than lineage markers alone, the previously identified CNs offer another layer to examine cellular relationships across MF grades. As expected, CN2 (Non-classical Monocytes with Osteo/Stromal and DC/B) increases with fibrosis grade, consistent with the enrichment of the Osteo–DC–T module in advanced fibrosis (Fig. 4.52a). This supports the idea that these interactions may drive fibrosis progression.

Interestingly, despite their co-enrichment in CN2, Osteo-MSCs show no significant short-range association with monocytes at any fibrosis grade. This may suggest that their interaction occurs over longer distances, possibly by production of inflammatory cytokines, or that the monocyte population should be further resolved into classical and non-classical subsets when studying spatial relationships.

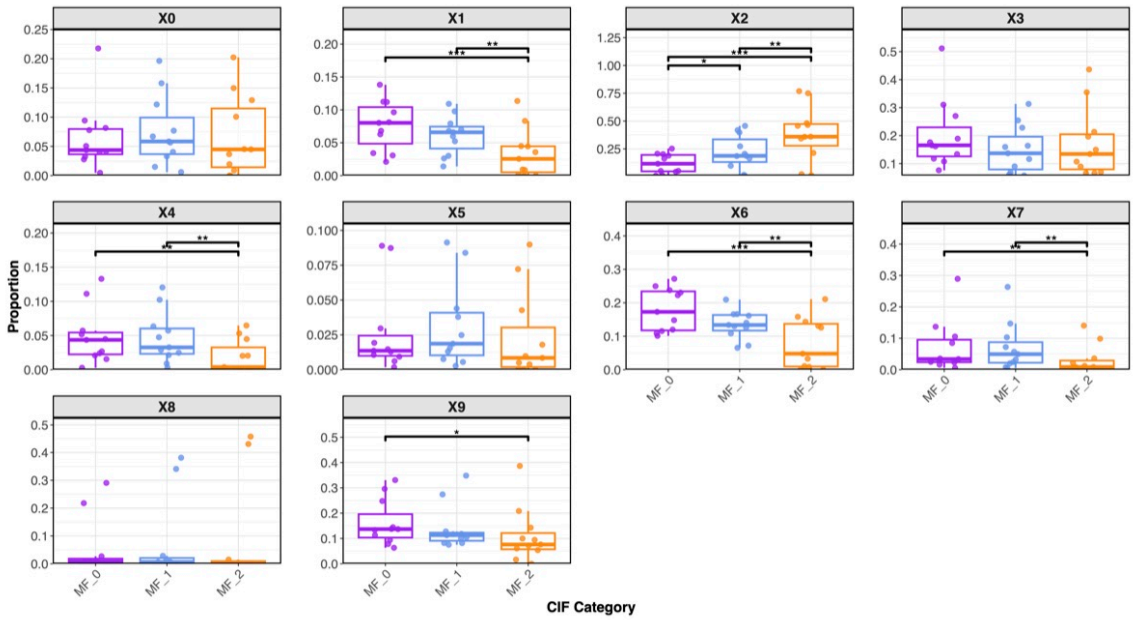
CN6 (the peritrabecular region) and CN1 (Erythroid/Macrophage, associated with normal bone marrow) both decrease from MF-0 to MF-2. In contrast, CN9, a CN expanded in MF and de-enriched for immunosurveillance by CN5, was most abundant in MF-0, suggesting it may represent an earlier disease state (Fig. 4.52a).

To confirm that CN changes were associated with fibrosis grade rather than proximity to bone, I repeated the analysis within only the bone-adjacent regions (combined endosteal and

peritrabecular) (Fig. 4.52b). The same trends were observed: CN2 increased with higher fibrosis grades, while CN6 and CN9 increased as fibrosis declined (Fig. 4.52b).



(a) Whole trephine



(b) Bone-adjacent regions

Figure 4.52: Differential abundance of each CN across fibrosis grades in MF, shown for the whole trephine (a) and the combined endosteal and peritrabecular regions (b).

I previously hypothesised that CN2 marks early fibrosis, based on its expansion from the trabeculae in Normal/PV to ET and MF (Fig. 4.36). However, its strong association with MF-2, even near bone, challenges this view (Fig. 4.52b). Determining whether CN2 reflects

early fibrosis initiation or is simply a correlate of established fibrosis is beyond the scope of this dataset, but represents an important avenue for future work. Either possibility could have predictive or therapeutic implications.

The increased abundance of CN9 (HSPC niche with Megakaryocytes and Endothelial/SMC) at lower fibrosis grades is also notable (Fig. 4.52). CN9 was previously identified as a defining region in MF, given its spatial expansion and its disrupted association with CN5 in ET and MF (Fig. 4.31, 4.34). While CN9 extends into the bone region in MF (Fig. 4.37), it is also enriched in MF-0, even within the endosteal/peritrabecaular. This raises the possibility that CN9 may represent a precursor state to fibrosis. Prior studies suggest that inflammatory or other insults in MF may trigger endothelial-to-mesenchymal transition (EndMT), producing myofibroblasts and contributing to fibrosis and osteosclerosis [49, 56, 110]. Megakaryocytes have been implicated in driving this process [111]. In our global expression data, we observed upregulation of SNAI1 in both endothelial cells and megakaryocytes (Fig. 4.12). Future studies should compare gene expression within CN9 to the rest of the trephine to further test this hypothesis.

Resolving key questions across fibrosis grades requires gene expression analysis

Beyond CN analysis, key questions remain. One is the co-localisation of Osteo-MSCs and megakaryocytes at higher fibrosis grades. Another is the observed loss of CN5–CN2 association in ET and MF compared to Normal, which we previously interpreted as reduced immunosurveillance. Yet, CN2 is also enriched in DCs, B cells, and T cells, and Osteo-MSCs associate with T cells and DCs a higher fibrosis grades. Future studies should assess whether these lymphoid populations in CN2 share similar gene expression profiles with those in CN5. Interestingly, despite the presence of a conserved T–DC–B module across all MF grades, fibrosis still progresses and CN2 remains associated with high CIF scores. This suggests possible dysfunction in the T cell compartment within CN2, either impaired activation or a shift toward pro-inflammatory T cell subsets. Comparing lymphoid gene expression between CN2 and CN5 may help clarify this. Finally, it remains an open question whether the Osteo-MSCs observed at higher fibrosis grades correspond to the NCAM-1+ osteolineage cells previously shown to expand from the peritrabecular region and associate with fibrosis progression.

Cellular Expression across Fibrosis Grades in MF

To investigate whether gene expression changes underlie the altered cell relationships and abundances observed across fibrosis grades, I examined differential expression across cell populations. Specifically, I tested whether cells involved in changes at higher fibrosis grades also showed distinct transcriptional profiles.

Using pseudo-bulk DESeq2, controlling for sample ID in a paired design, I performed all six pairwise comparisons between fibrosis grades. Only genes with an adjusted p-value < 0.05 after multiple hypothesis correction were considered significant. To identify which cell populations exhibited the most transcriptional changes, I visualised the number of unique differentially expressed genes per cell type for each comparison (Fig. 4.53). For most cell types, the largest expression differences occur between MF-0 and MF-1/MF-2, suggesting transcriptional shifts between low and high fibrosis grades (Fig. 4.53).

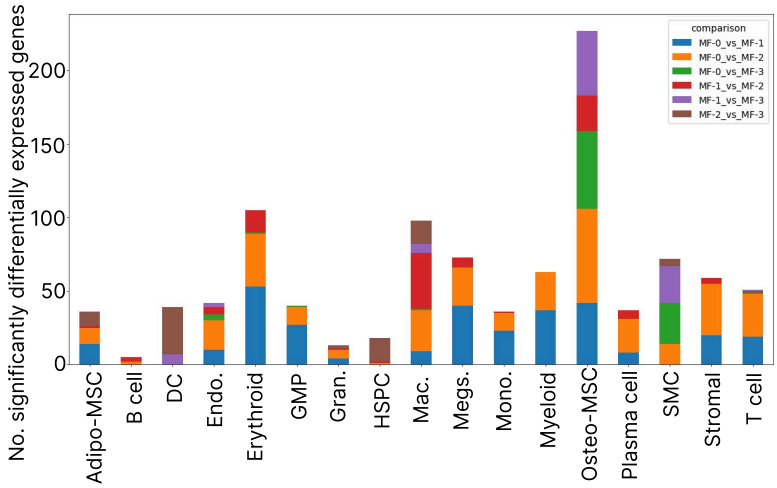


Figure 4.53: Number of differentially expressed genes per comparison per cell type in MF samples only (n=13).

Supporting the idea that Osteo-MSCs in higher fibrosis grades may represent a distinct subset, they show the most extensive transcriptional differences of any cell type, especially between MF-0 and all other grades (Fig. 4.53). Osteo-MSCs are significantly reduced in high fibrosis regions, despite being globally increased in MF compared to other conditions. A Wnt-dependent subset of Osteo-MSCs, defined by NCAM1, has previously been implicated in fibrosis development. To explore this further, I examined Osteo-MSCs expression across fibrosis grades (Fig. 4.54). In MF-1, MF-2, and MF-3 compared to MF-0, Osteo-MSCs show reduced expression of PCOLCE, MYC, CAVIN1, KCNMA1, GPC1, PDGFRA, NPDC1, VWA5A, ACTA2, TNC, PTGDS,

COL5A2, IGF1, DPT, MAMDC2, SFRP4, MEF2C, and APCDD1. Several of these genes—such as MYC, IGF1, and PDGFRA—are associated with fibroblast proliferation (GO:0048146), while MEF2C, BAMBI, and IGF1 are linked to osteoblast differentiation (GO:0045667) [114]. These findings suggest that fibroblast and osteoblast activity is more prominent at lower fibrosis grades.

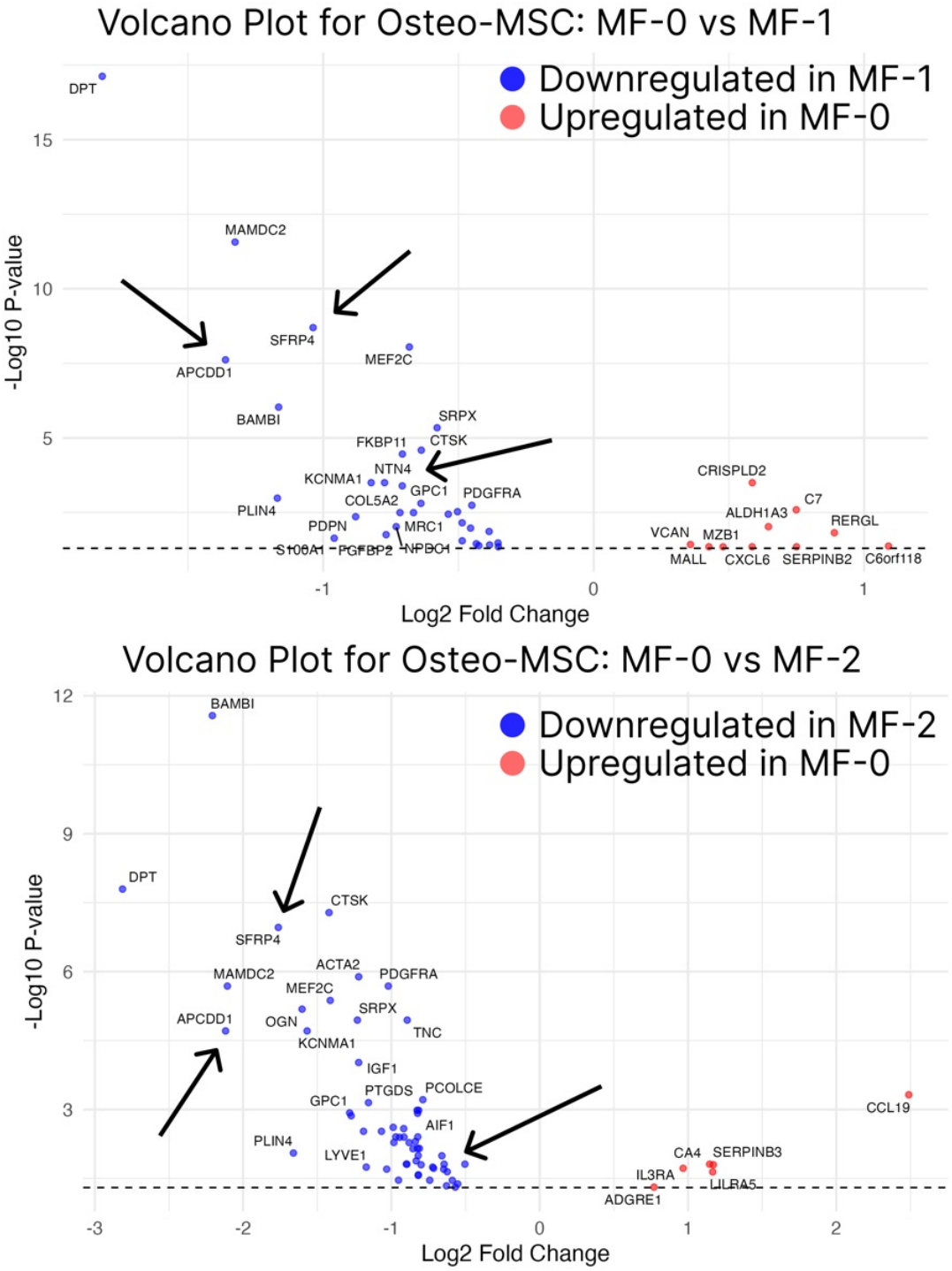


Figure 4.54: Volcano plots showing Osteo-MSC differential expression across MF fibrosis grades, excluding MF-3 due to low sample number. Arrows indicate Wnt inhibitors

Osteo-MSCs at higher fibrosis grades have reduced Wnt inhibition

DPT, BAMBI, APCCD1, MAMDC2, and SFRP4 are upregulated in Osteo-MSCs in MF-0 compared to all other grades, suggesting these genes may define key differences between Osteo-MSCs in low and high fibrosis (Fig. 4.54).

DPT is upregulated during osteogenic differentiation in the bone marrow [115] and may inhibit haematopoietic–endothelial adhesion and transendothelial migration via VLA-4 [116], suggesting a loss of osteogenic potential in higher fibrosis grades. BAMBI suppresses TGFB1 expression and reduces fibrosis in multiple tissues [117]. TGF-beta is a major driver of fibrosis and has been used in bone marrow organoid models to induce fibrosis [118]. Its decreased inhibition by BAMBI in Osteo-MSCs at higher grades suggests elevated TGF-beta signalling.

However, most interestingly, SFRP4, APCCD1, and NTN4 are Wnt pathway inhibitors upregulated in Osteo-MSCs in MF-0 but downregulated at higher grades. Wnt inhibition has been shown to suppress fibrogenesis and osteosclerosis [56]. Their reduced expression supports the idea that Osteo-MSCs lose control over Wnt signalling as fibrosis progresses. NTN4 also inhibits osteoclast differentiation and bone loss [119, 120]. APCCD1 regulates adipogenesis, and its silencing upregulates Wnt signalling while suppressing adipocyte differentiation [121], potentially contributing to the fat loss observed in fibrotic marrow. Loss of adipose tissue is a hallmark of MF. A peri-adipocytic niche has recently been proposed for HSPCs [34]. Future work should determine whether this niche exists in healthy marrow and how fibrosis affects its integrity.

General downregulation of Wnt inhibitor SRFP4 in higher fibrosis grades

There is also a general downregulation of Wnt inhibition across multiple cell types in higher fibrosis grades. SFRP4, a Wnt pathway inhibitor, is reduced in monocytes, macrophages, megakaryocytes, and endothelial cells in both MF-1 and MF-2 compared to MF-0. In B cells, SFRP4 is also downregulated in MF-2 relative to MF-1. These findings suggest widespread activation of Wnt signalling in advanced fibrosis, consistent with observations from Banjanin et al. [56].

We observed enriched spatial relationships between Osteo-MSCs and both megakaryocytes and B cells at higher fibrosis grades, and a consistent enrichment between Osteo-MSCs and endothelial cells across all grades. However, there were no significant close spatial associations

between Osteo-MSCs and monocytes or macrophages in higher grades. Still, CN2, enriched for Osteo-MSCs, stromal cells, DCs, and B cells, also includes non-classical monocytes, raising the possibility that distinct monocyte subsets play different roles in fibrosis. Future work should explore this further.

We do not have inflammatory markers enriched in our gene panel. However, interestingly, VSIG4, a negative regulator of inflammation, is downregulated in monocytes in MF-1 and MF-2 compared to MF-0. Loss of VSIG4 has been linked to increased inflammation via NLRP3 and IL-1beta activation in macrophages [122, 123]. The NLRP3 inflammasome is already implicated in MPN pathogenesis, with increased expression associated with splenomegaly, JAK2 mutations, and elevated leukocyte counts [124]. Although these findings should be interpreted cautiously, they suggest that monocytes may contribute to inflammation in intermediate and advanced fibrosis.

CCL5 is upregulated in MF-0 compared to all higher MF grades in endothelial cells (and in MF-1 compared to MF-2), and in megakaryocytes in MF-0 compared to MF-1. Macrophages show increased expression in MF-0 and MF-1 versus MF-2, and T cells in MF-0 versus MF-2. CCL5 has been associated with megakaryocyte expansion under inflammatory conditions, and its expression in endothelial cells increases after injury as part of the inflammatory response [90, 125–127]. EndMT has also been identified as a feature of myelofibrosis in transcriptomic analyses, and has been hypothesised to occur after an injury stimuli such as chronic inflammation [49, 56]. Megakaryocytes may play a key role, as EndMT tends to occur in microvessels near megakaryocytes, and megakaryocyte-derived supernatant can induce this process in vitro [111].

I previously hypothesised that CN9, enriched for HSPCs, endothelial cells, and megakaryocytes, could represent a region where EndMT occurs. The CCL5 expression in these cell types at lower fibrosis grades supports the idea of endothelial injury, although there is not yet enough evidence to confirm CN9 as an EndMT niche. Endothelial cells and megakaryocytes remain associated across all fibrosis grades (Fig. 4.48). Global expression analysis showed upregulation of SNAI1, a marker of EndMT, in endothelial cells and megakaryocytes in MF compared to normal bone marrow (Fig. 4.12). However, within MF grades, only Osteo-MSCs showed higher SNAI1 expression in MF-0 compared to MF-1 and MF-2 (Fig. 4.54). Future work should

examine whether endothelial cells and megakaryocytes in CN9 express EndMT markers and whether distinct megakaryocyte phenotypes are involved in this process.

4.6.4 Discussion

In this section, I evaluated cellular changes in abundance, relationships, and gene expression, across fibrosis grades in MF to better understand drivers of fibrosis.

Somewhat unexpectedly, two cell types globally increased in MF, Osteo-MSCs and Megakaryocytes, show a linear decrease in abundance with increasing fibrosis. In contrast, monocytes, macrophages, and myeloid cells increase, while erythroid cells decrease (Fig. 4.45).

However, cell proportions may not vary linearly across trephines, therefore I also investigated fibrosis grades changes. This reveals increasing B and DC cells and decreasing GMPs from MF-0 to MF-2. Among monocytes, myeloid cells, and macrophages, only monocytes remain elevated at MF-2 (Fig. 4.46). Osteo-MSCs decline significantly with fibrosis, while megakaryocytes decrease only when fibrosis grades are extended to high-spot tiles (Fig. 4.47).

The increase in B cells, DCs, and monocytes is reminiscent of CN2, which we previously associated with higher CIF scores (Fig. 4.22). CN2 also includes Osteo-MSCs, which paradoxically decline with fibrosis. One explanation is that higher fibrosis grades retain only the NCAM-1+ Osteo-MSC subset implicated in fibrosis [56], while more general Osteo-MSCs are left behind at MF-0. Though my panel lacks NCAM-1/CD56, Osteo-MSCs show the most differential expression across grades, mainly between MF-0 and others, suggesting functional divergence (Fig. 4.53).

Given its clinical relevance, I maintained fibrosis grades as the framework for further analyses. Neighbour enrichment analysis revealed conserved cell modules across all grades, likely corresponding to known CNs e.g., a T-DC-B module (CN5), an Endothelial-Osteo-Stromal module (CN3), and consistent HSPC-GMP and HSPC-Endothelial associations, suggestive of a peri-vascular niche. Megakaryocyte-Endothelial interactions were also conserved across all grades (Fig. 4.48).

In line with CN2, we observed new relationships at higher fibrosis grades: Osteo-MSC–DC cell enrichment at MF-1 and a T-DC–Osteo-MSC module at MF-2. Comparing MF-1 and MF-2 to MF-0, only Osteo–Megakaryocyte and Osteo–Erythroid relationships were reciprocally enriched, while Monocyte–Plasma associations were diminished (Fig. 4.49). Notably, the

Osteo–Megakaryocyte relationship is enriched at higher fibrosis despite both populations decreasing, suggesting a selective interaction (Fig. 4.4).

Expanding to longer-range co-occurrence (150–500 pixels), we observed disrupted HSPC–HSPC associations in MF-2 at increased radii, perhaps hinting at niche fragmentation (Fig. 4.50). The Osteo–Erythroid association disappeared, but Osteo–Megakaryocyte, Osteo–DC, and Osteo–B cell co-localisation persisted at higher grades. Interestingly, Osteo–T cell co-localisation was maintained across all grades. There was still no spatial association between Osteo-MSCs and Monocytes, inconsistent with CN2-level Monocyte relationships. The Monocyte–Plasma trend appeared to reflect increased de-enrichment in fibrosis, not MF-0-specific enrichment (Figs. 4.50, 4.48).

Focusing on a 250-pixel radius, Osteo-MSCs showed no significant co-localisation with Monocytes at any grade, despite Monocyte expansion in abundance. Conversely, Osteo–Macrophage co-localisation was consistently positive. Osteo–Megakaryocyte association remained significant at MF-1 but not MF-2. Associations with T cells and DCs were significant at MF-1, and with B cells at MF-2 (Fig. 4.51).

Together, these findings reinforce CN2's relevance to fibrosis: enriched for DC, B, and Osteo-MSCs. However, I cannot validate a clear Osteo-MSC-Monocyte association. Monocytes may contribute via long-range effects or may differ functionally between classical and non-classical subsets, and increased resolution is required.

CN2 remained fibrosis enriched even when analysis was restricted to bone-adjacent regions (Fig. 4.52a, 4.52b). While I had hypothesised CN2 as a marker of early fibrosis due to its spatial expansion, association with increased CIF, and changed localisation in ET and MF, its persistence at MF-2, even near bone, challenges this. Future work is needed to determine whether CN2 reflects fibrosis onset or is merely a cellular correlate of fibrosis progression.

CN9 (an HSPC niche with Megakaryocytes and Endothelial/SMC), which expands in MF and loses CN5 association, paradoxically shows greatest enrichment at MF-0, even near bone (Figs. 4.52a, 4.52b). This raises the possibility that CN9 precedes fibrosis. Endothelial-to-mesenchymal transition (EndMT), possibly triggered by megakaryocyte-derived signals, has been proposed as a fibrosis mechanism [49, 111]. Our global gene expression analysis already showed increased SNAI1, an EndMT marker, in MF megakaryocytes and endothelial cells (Fig. 4.12). Surprisingly,

within MF grades, SNAI1 expression was only increased in Osteo-MSCs at MF-0 compared to MF-1/2 (Fig. 4.54). Future work should test whether megakaryocytes or endothelial cells in CN9 differentially express EndMT markers.

Another persistent challenge is the observed loss of CN5–CN2 association in ET/MF compared to healthy bone marrow, which we interpret as impaired immunosurveillance. CN5 is enriched for DCs, B, and T cells. If the lymphocytes in CN2 differ functionally from those in CN5, e.g., through exhaustion or overactivation, they may drive inflammation rather than resolving aberrant Osteo-MSC activity. Future analyses should compare gene expression of lymphoid cells across CN2 and CN5.

Gene expression further supports the functional divergence of Osteo-MSCs across fibrosis. These cells showed the highest number of DEGs between grades (Fig. 4.53). In line with recent findings implicating Wnt-dependant NCAM1+ cells in fibrosis [56], Osteo-MSCs in MF-1/2 downregulate multiple Wnt inhibitors (SFRP4, APCDD1, NTN4; Fig. 4.54). Widespread SFRP4 downregulation across other cell types at higher grades suggests a global shift toward pro-fibrotic Wnt signalling in fibrotic MF.

Despite CN9's MF-0 localisation, its potential role in initiating fibrosis remains open. We observed global upregulation of SNAI1 in MF endothelial cells and megakaryocytes relative to normal BM, but not within fibrosis grades. If CN9 accounts for most of these cells, differences between CNs across grades may be diluted. CN-specific expression analysis, at a cell type level, would be valuable.

Interestingly, CCL5, a chemokine upregulated after endothelial injury, was highest at MF-0 in endothelial cells, megakaryocytes, macrophages, and T cells, declining with fibrosis. CCL5 has been linked to inflammation and megakaryocyte expansion [126, 127]. Chronic inflammation has been proposed to displace CAR cells, leaving endothelial cells exposed and susceptible to EndMT [56]. Together, elevated CCL5 in MF-0 and global SNAI1 upregulation suggest MF-0 may represent a pre-fibrotic endothelial injury state.

Whether CN9 represents an EndMT-prone region remains to be seen. However, given its spatial expansion and de-enrichment from immunoregulatory CN5, it remains a strong candidate for future investigation. Understanding whether specific megakaryocyte phenotypes associate with this transition could further clarify fibrosis initiation.

4.7 Gene Expression and Cytomorphology

4.7.1 Introduction

The spatial transcriptomic analysis presented above has proven valuable for generating new hypotheses grounded in spatial context and for validating existing findings from the literature. However, it comes with limitations. Our gene expression analysis is constrained by the small panel size and a limited number of samples, particularly problematic given the heterogeneity observed in MF. Additionally, spatial transcriptomics is currently too costly for routine use in every patient or integration into the clinical workflow.

To fully leverage the Xenium dataset, we must corroborate and extend our findings through complementary approaches. Some hypotheses will require follow-up with in vitro experimentation. However, another asset is the H&E-stained slide generated alongside every Xenium output. Every patient with MPN undergoes a biopsy with an H&E section assessed by a pathology team. These slides, spanning both current and archival cases, can be linked to clinical metadata, including outcomes. This provides a potential bridge for translating spatial transcriptomic findings into the clinic by identifying matching features in H&E.

For this to be effective, regions identified as relevant in the spatial transcriptomic analysis must have morphological correlates visible in H&E. Two major strategies emerge from the analysis: (1) identifying morphological correlates in H&E for the CNs identified via CellCharter, and (2) mapping H&E morphologies to specific cell types defined in the spatial transcriptomic data.

In this context, I highlight an approach developed in the host laboratory for the detection, segmentation, and phenotyping of megakaryocytes in H&E sections [17]. By registering the H&E with the Xenium output, we can relate gene expression directly to cell morphology at the single-cell level.

Atypical megakaryocytes are a well-recognised feature across MPNs. Subtype-specific descriptions exist, for instance, staghorn nuclei in ET, and bulbous or cloud-like nuclei in MF, yet the transcriptional profiles of these morphologies remain unclear. By linking megakaryocyte morphology to gene expression, we may uncover whether specific phenotypes are associated with particular biological programs.

In this thesis, I have identified several key megakaryocyte-related questions for further study:

- Do newly identified megakaryocytes in the spatial data alter our understanding of megakaryocyte spatial distribution in MPNs previously built from H&E analysis?
- Can we link transcriptional profiles to each megakaryocyte phenotype to better interpret the increased prevalence of phenotype 8 in ET and MF?
- Are there morphological correlates for the spatial relationship between megakaryocytes and Osteo-MSCs observed at higher fibrosis grades?
- Can we identify morphological correlates of megakaryocytes in CN9, a neighbourhood spatially expanded in MF and ET and uncoupled from CN5?
- Are there morphological correlates for megakaryocytes potentially involved in endothelial-to-mesenchymal transition?

While resolving all these questions lies beyond the scope of this thesis, I will demonstrate below how gene expression can be linked to megakaryocyte cytomorphology using the registered H&E, laying the groundwork for future investigation.

4.7.2 Methodology

Megakaryocyte Phenotypes

I phenotyped all detected megakaryocytes using a method previously developed in the host laboratory, as described in Sirinukunwattana et al. [17]. Briefly, an autoencoder was trained on segmented megakaryocytes from 42 reactive and 30 ET samples to extract morphological features, generating a 128-dimensional feature vector for each cell. These vectors were then organised using a self-organising map (10×10 grid), producing 100 groups of morphologically similar megakaryocytes, which were subsequently refined using Markov clustering. Crucially, these phenotypes are derived solely from morphological features and are independent of any gene expression data (Fig. 4.55).

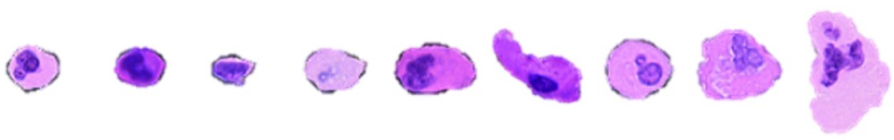


Figure 4.55: Representative image of the megakaryocyte phenotypes from 0 (far left) to 8 (far right)

Differential Gene Expression

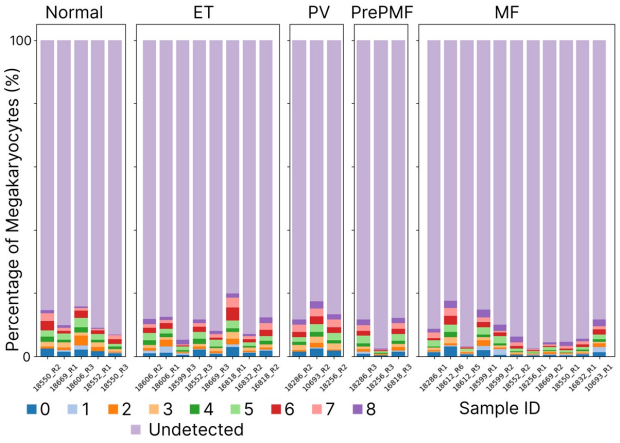
Megakaryocytes were classified into phenotypes 0–8, along with an “undetected” group. Phenotypes 0–8 were segmented directly from the registered H&E, whereas the undetected group consisted of megakaryocytes identified through clustering of Xenium gene expression profiles (Fig. 3.32). Many more megakaryocytes were detected via gene expression than through morphological segmentation (Fig. 4.56a). To prevent the undetected group from dominating comparisons, these cells were downsampled within each sample to match the number of cells in the next most abundant phenotypic group. For each pairwise comparison, megakaryocytes were grouped into phenotype of interest e.g. 0 and all other megakaryocytes were designated "Other". Gene expression was then aggregated by sample and the binary groups ("phenotype of interest" or "other") for pseudobulk analysis. DESeq2 was run (~ sample_id + group_x_vs_all) taking into account variability from sample ID. This process was repeated 10 times, and the mean adjusted p-value and mean average logFC were calculated for each gene that was significant (<0.05) in every run. For each comparison, I then selected the top 5 upregulated and top 5 downregulated genes, provided that the expression change was greater than |0.5| and the average adjusted p-value was <0.05. The heatmap is colored by mean avglogFC for that gene, and stars denote if the gene was significantly different for that comparison (*** < 0.001, ** < 0.01, * < 0.05).

4.7.3 Results

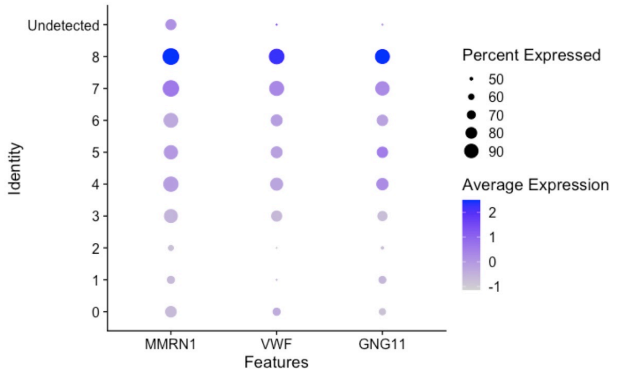
Megakaryocyte Markers across Phenotypes

As previously noted, clustering based on gene expression in the Xenium data revealed substantially more megakaryocytes than were identified either by manual review or by our prior H&E-based detection algorithm. This increase was consistent across all conditions (Fig. 4.56a). As shown in the previous chapter, these "undetected" megakaryocytes form a distinct cluster defined by canonical megakaryocyte markers, including MMRN1, VWF, and GNG11. We also examined the expression of these markers across the phenotyped megakaryocytes. Notably, 80% of the undetected megakaryocytes express MMRN1, and the expression levels of VWF and GNG11 are comparable to those observed in smaller phenotyped megakaryocytes previously identified based on cytomorphology. This provides strong support that the megakaryocytes

missed by the H&E-based detection model are indeed true megakaryocytes, within the scope of our gene panel (Fig. 4.56b).



(a) Proportion of phenotyped megakaryocytes detected by our H&E detection algorithm compared to those identified only by gene expression.



(b) Marker expression in megakaryocytes across all conditions and samples of key markers by phenotype status (untransformed expression).

Figure 4.56: Summary of phenotyped vs unphenotyped megakaryocytes and associated marker expression patterns.

Differental Gene Expression by Phenotype

We hypothesised that smaller megakaryocytes may be more immature, potentially explaining their lower expression of VWF and GNG11. To test this, we performed differential expression analysis across phenotypes. Because the undetected group of megakaryocytes is much larger than any individual phenotyped group, direct comparisons would be skewed. To mitigate this, we downsampled the undetected group to match the size of the next largest phenotype within each sample and performed bootstrapping over 10 iterations, as described above. Although the megakaryocyte groups were defined purely by morphology, we observed clear differences in gene expression between phenotypes. This supports the notion that cytomorphological features are linked to underlying transcriptional differences (Fig. 4.57).

Some phenotypes, such as 4 and 6, do not exhibit significant differential expression compared to other megakaryocytes (Fig. 4.57). This may be due to the limited gene panel, but it also suggests these groups could potentially be merged with other phenotypes. In this way, gene expression profiling can help refine morphologically defined phenotypes.

Despite this, several phenotypes show distinct transcriptional signatures. Proliferation-associated genes (TOP2A, CCNB2, MKI67) are significantly upregulated in undetected, phenotype 0, and

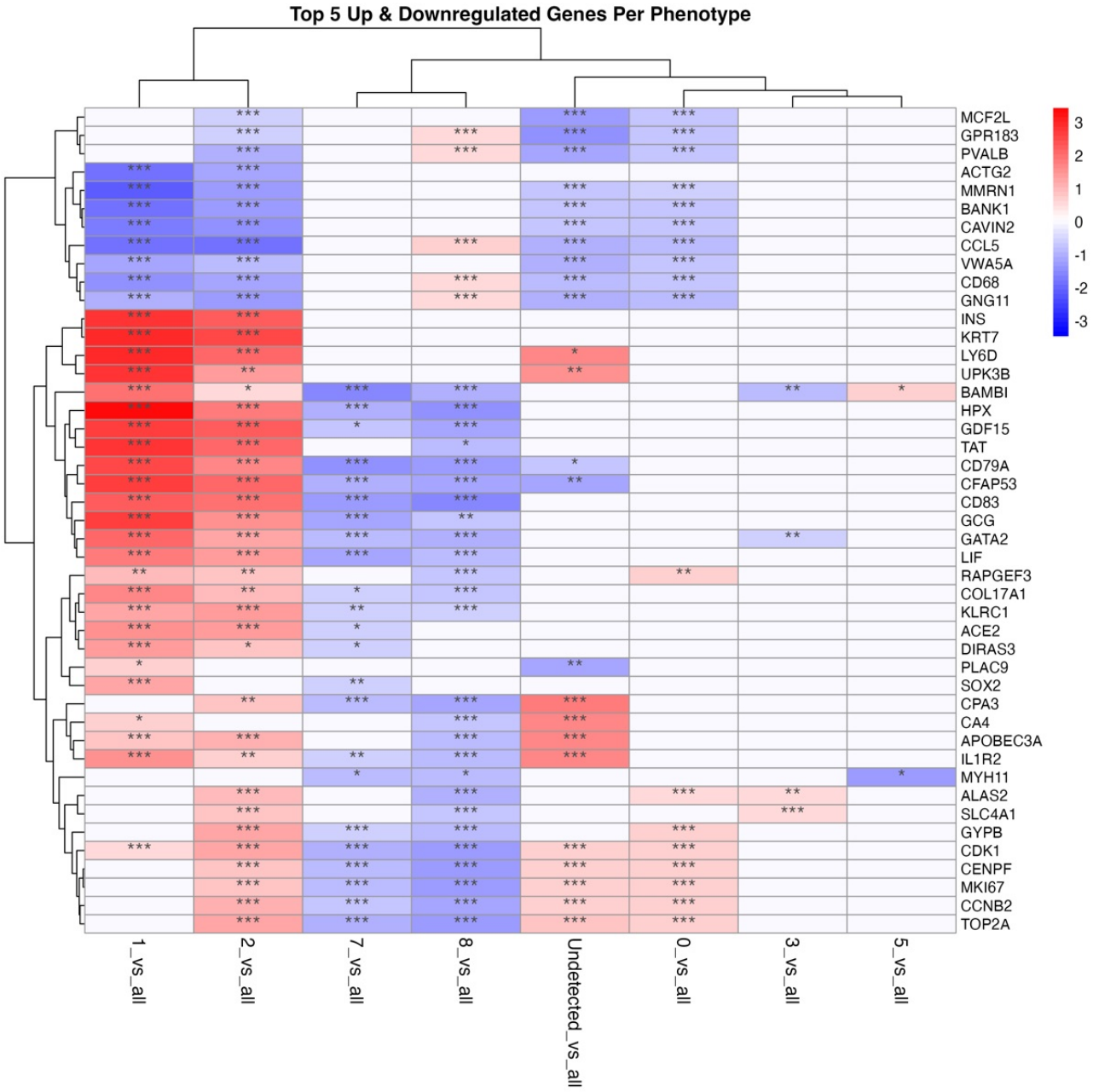


Figure 4.57: Megakaryocytes were extracted and grouped by phenotype (0-8) and undetected. Within each of the 30 samples, undetected was downsampled to the same size as the next largest phenotypic group within that sample. For each pairwise comparison, megakaryocytes were grouped into phenotype of interest and other. Gene expression was then aggregated by sample and the binary phenotype (phenotype of interest or other) for pseudobulk analysis. DESeq2 was run (~ sample_id + group_x_vs_all) taking into account variability from sample ID. This process was repeated 10 times, and the mean adjusted p-value and mean avglogFC were calculated for each gene that was significant (<0.05) in every run. For each comparison, I then selected the top 5 upregulated and top 5 downregulated genes, provided that the expression change was greater than |0.5| and the average adjusted p-value was <0.05. The heatmap is colored by mean avglogFC for that gene, and stars denote if the gene was significantly different for that comparison (p values: *** < 0.001, ** < 0.01, * < 0.05).

phenotype 2 megakaryocytes, cells that are typically smaller or indistinct in H&E. These genes are downregulated in larger phenotypes, such as 7 and 8, supporting the hypothesis that smaller megakaryocytes are less mature. CDK1 follows a similar pattern but is also upregulated in phenotype 1. In addition, GATA2, a regulator of megakaryocytic differentiation, is significantly upregulated in phenotypes 1 and 2 and downregulated in phenotypes 3, 7, and 8. Together, these data confirm that smaller morphologies are likely immature or differentiating (Fig. 4.57).

Phenotype 8, one of the largest, shows significant upregulation of CCL5, a gene linked to megakaryocyte expansion under inflammatory conditions [125, 126]. We previously associated increased CCL5 expression in megakaryocytes in MF-0. Phenotype 8 is expanded in ET and MF, mirroring the spatial expansion of CN9, a cellular neighbourhood enriched for HSPCs, megakaryocytes, and endothelial cells, which is also enriched in MF-0 and de-enriched with CN5. This suggests that phenotype 8 may be enriched within CN9, potentially providing a potential morphological correlate of CN9 for further investigation in archival H&E datasets, although this would need to be confirmed.

4.7.4 Discussion

In this analysis, we associate predefined megakaryocyte cytomorphological phenotypes with gene expression. Differential expression analysis reveals that genes related to proliferation and megakaryocyte differentiation are enriched in smaller phenotypes, consistent with the expectation that these cells are less mature. This suggests that, by identifying these phenotypes in H&E trephines, we can localise regions enriched for less differentiated megakaryocytes.

While our gene panel is limited to 377 markers, it is important to note that if a gene does not impact morphology itself, its association will depend on co-expression with other genes that do affect phenotype, potentially weakening its link to a morphological correlate. Despite this limitation, the association of CCL5 with phenotype 8, the largest megakaryocyte phenotype, is notable. CCL5 has been linked to endothelial injury and megakaryocyte expansion [90, 125, 126]. Future work should explore whether morphological phenotypes, such as phenotype 8, map onto spatially defined cellular neighbourhoods (CNs) identified in earlier analyses.

4.8 Limitations

Whilst the work above provides valuable insights into the spatial organisation of the healthy and MPN bone marrow microenvironment, there are several limitations.

The most significant limitation is the restricted gene panel. The current panel is enriched for pan-cancer markers, many of which are not specific to haematopoietic or stromal cell types in the bone marrow. This affects our ability to confidently assign cell lineages and explore key biological processes. For example, HSPCs were defined by co-expression of CD34 and KIT, but with higher-resolution transcriptomic data, we could distinguish true HSCs from progenitors, enabling trajectory analysis and more precise niche identification. The limited panel also impacts the interpretation of other findings. Although we identified a distinct cluster of megakaryocytes via gene expression, many of which are not visible in H&E, a broader panel including additional lineage markers would increase confidence in their classification. Similarly, while we identify Osteo-MSCs, we are unable to resolve specific subtypes such as NCAM-1+ MSCs or distinguish osteoblasts from osteoclasts, two functionally distinct osteolineage populations. Additionally, the panel limits our capacity to study critical processes such as chronic inflammation and fibrosis. While we highlight observations related to EndMT and Wnt signalling, deeper pathway-level analysis will require expanded gene coverage. Future work with broader panels or whole transcriptome data would allow further investigation into inflammatory pathways (e.g., TGF-beta signalling) and their spatial context in the tissue.

Our conclusions about spatial organisation also rest on decisions within the CellCharter analysis. While we tested multiple clustering solutions, the final choice of 10 clusters was a biologically informed decision. Moreover, CellCharter was run across all samples rather than within individual conditions to ensure comparability of CNs. However, this approach may obscure condition-specific spatial patterns and limit the resolution of MPN-specific or healthy bone marrow-specific CNs.

The analysis of fibrosis relied on the CIF map, which assigns a fibrosis score to each tile of the H&E image. Although the CIF map is not computed directly from the reticulin-stained section, it has been shown to correlate well. We also relied on accurate image registration between the CIF map and DAPI-stained Xenium images. Manual inspection confirmed high-quality alignment, with at most half a nucleus shift. While this is not a major concern in this dataset,

image registration remains a general limitation in the incorporation of morphology into spatial analyses.

An important finding in our study was the heterogeneity within MF samples. In particular, we observed expansion of CN8 at the expense of CN2 in a subset of MF cases. Although this heterogeneity is biologically interesting, it complicates analysis design. Since we have not yet linked these spatial patterns to clinical outcomes, all MF samples were analysed together. However, this could obscure meaningful subtype-specific associations and should be explored further in future work.

Finally, the last section of this thesis connects spatial transcriptomic data to standard imaging used in clinical workflows by relating gene expression to megakaryocyte morphological phenotypes. While this demonstrates proof-of-principle, the biological relevance of all nine defined phenotypes remains uncertain; some of these morphological groups could be merged.

4.9 Future Work

In this thesis, I have made several key observations about the spatial organisation of the bone marrow and the potential clinical significance of CellCharter-defined cellular neighbourhoods (CNs) identified using Xenium spatial transcriptomics. Given the high cost and limited scalability of spatial transcriptomics in clinical settings, a critical next step is to leverage the registered CN boundaries to infer CN identity from morphological features visible in routine H&E-stained bone marrow biopsies.

A priority for future work is to identify CN2 and CN9 in H&E images, as both may act as precursors or, CN2, a cellular correlate of fibrosis. Tracking their spatial dynamics could offer a practical means to monitor disease progression or therapeutic response. Similarly, identifying CN5 (lymphoid aggregates), or CN1, CN3, and CN4, which may represent preserved healthy marrow architecture, could inform prognosis. The two central challenges are whether these CNs can be reliably inferred from morphology alone, and whether shifts in CN distribution correlate with clinical outcomes in larger patient cohorts.

Another important direction is to investigate the heterogeneity observed among MF samples. Preliminary findings suggest that expansion of CN8 is associated with decreased MF-0 and

increased MF-2 fibrosis grades. This trajectory may be visible in archival H&E or reticulin sections and could point to distinct subtypes of fibrosis progression. Future work should assess whether these patterns are reproducible, linked to different prognoses, and associated with identifiable morphological correlates of CN8 or CN2. Such studies would help test the hypothesis that spatial heterogeneity in fibrosis reflects underlying differences in cellular composition.

5

Future Work

5.1 Summary of Contributions

Chapter 3: Preparing the Xenium Dataset

This chapter details a novel pipeline for processing Xenium spatial transcriptomic data of the bone marrow across all MPN subtypes and healthy controls. Machine learning outputs and fibrosis maps are integrated into the dataset to enhance spatial context.

Innovations include custom methods for incorporating irregular cell morphologies (e.g., megakaryocytes) into segmentation, and a Voronoi-based graph construction that preserves neighbourhood structure even when dealing with non-uniform cell morphologies. Cell type annotation is improved using multi-layered segmentations and scRNA-seq references, including a stromal-enriched dataset. Macrostructures like adipocytes, bone, and vasculature are also incorporated.

Chapter 4: Spatial Transcriptomics Analysis of the MPN Bone Marrow

This chapter presents the first spatial transcriptomic analysis of the bone marrow across all MPN subtypes. Ten distinct cellular neighbourhoods (CNs) were identified, many aligning with known histological regions. CN9, enriched in megakaryocytes, endothelial cells, and HSPCs, expanded in ET and MF and may represent a precursor fibrotic niche. CN2, linked to high

fibrosis scores and spatial shifts from the bone wth disease progression, also emerged as a potential early fibrotic microenvironment.

Stable CNs (CN1, CN3, CN4) maintained spatial architecture across disease states and resisted spatial expansion, indicating a conserved marrow organisation previously undescribed. CN8 expansion in MF correlated with a fibrotic pattern enriched for MF-2, potentially marking a late-stage fibrotic pathway. High fibrosis grades showed novel spatial associations between osteo-MSCs and megakaryocytes, and DC and T cells, along with downregulation of Wnt inhibitors in osteo-MSCs at higher fibrosis grades, suggesting recapitulation of a NCAM1+ profibrotic subtype.

Megakaryocyte morphology was linked to transcriptional state, with smaller, immature cells being more proliferative. These findings support the integration of spatial transcriptomics into clinical workflows and fibrosis monitoring.

5.2 Future Work

Chapter 3: Preparing the Xenium Dataset

The functions for megakaryocyte integration into Xenium segmentations and for graph construction that better captures neighbours of irregularly shaped cells will be released as a standalone package. This tool will be valuable for other groups working with bone marrow spatial transcriptomic data who seek accurate incorporation of megakaryocytes and biologically meaningful graph representations.

Chapter 4: Spatial Transcriptomics Analysis of the MPN Bone Marrow

In this thesis, I have made several key observations about the spatial organisation of the bone marrow and the potential clinical significance of CellCharter-defined cellular neighbourhoods (CNs) identified using Xenium spatial transcriptomics. Given the high cost and limited scalability of spatial transcriptomics in clinical settings, a critical next step is to leverage the registered CN boundaries to infer CN identity from morphological features visible in routine H&E-stained bone marrow biopsies.

A priority for future work is to identify CN2 and CN9 in H&E images, as both may act as precursors or, CN2, a cellular correlate of fibrosis. Tracking their spatial dynamics could offer a

practical means to monitor disease progression or therapeutic response. Similarly, identifying CN5 (lymphoid aggregates), or CN1, CN3, and CN4, which may represent preserved healthy marrow architecture, could inform prognosis. The two central challenges are whether these CNs can be reliably inferred from morphology alone, and whether shifts in CN distribution correlate with clinical outcomes in larger patient cohorts.

Another important direction is to investigate the heterogeneity observed among MF samples. Preliminary findings suggest that expansion of CN8 is associated with decreased MF-0 and increased MF-2 fibrosis grades. This trajectory may be visible in archival H&E or reticulin sections and could point to distinct subtypes of fibrosis progression. Future work should assess whether these patterns are reproducible, linked to different prognoses, and associated with identifiable morphological correlates of CN8 or CN2. Such studies would help test the hypothesis that spatial heterogeneity in fibrosis reflects underlying differences in cellular composition.

References

[1] Pan Zhang et al. “The physical microenvironment of hematopoietic stem cells and its emerging roles in engineering applications”. In: *Stem Cell Research & Therapy* 10 (2019), p. 327.

[2] Nicholas Williams et al. “Life histories of myeloproliferative neoplasms inferred from phylogenies”. en. In: *Nature* 602 (2022), pp. 162–168.

[3] Domenico Penna. “New Horizons in Myeloproliferative Neoplasms Treatment: A Review of Current and Future Therapeutic Options”. en. In: *Medicina* 57 (2021), p. 1181.

[4] Nabih Maslah et al. “Clonal architecture evolution in Myeloproliferative Neoplasms: from a driver mutation to a complex heterogeneous mutational and phenotypic landscape”. en. In: *Leukemia* 37 (2023), pp. 957–963.

[5] Marie Strickland, Lynn Quek, and Bethan Psaila. “The immune landscape in BCR-ABL negative myeloproliferative neoplasms: inflammation, infections and opportunities for immunotherapy”. en. In: *British Journal of Haematology* n/a ().

[6] Tiziano Barbui et al. “The 2016 WHO classification and diagnostic criteria for myeloproliferative neoplasms: document summary and in-depth discussion”. en. In: *Blood Cancer Journal* 8 (2018), pp. 1–11.

[7] Adam J. Mead and Ann Mullally. “Myeloproliferative neoplasm stem cells”. In: *Blood* 129 (2017), pp. 1607–1616.

[8] So-Woon Kim, Jin Roh, and Chan-Sik Park. “Immunohistochemistry for Pathologists: Protocols, Pitfalls, and Tips”. In: *Journal of Pathology and Translational Medicine* 50 (2016), pp. 411–418.

[9] Rosalin A. Cooper et al. “Spatial transcriptomic approaches for characterising the bone marrow landscape: pitfalls and potential”. en. In: *Leukemia* 39 (2025), pp. 291–295.

[10] Michael Angelo et al. “Multiplexed ion beam imaging of human breast tumors”. en. In: *Nature Medicine* 20 (2014), pp. 436–442.

[11] Bernd Bodenmiller. “Multiplexed Epitope-Based Tissue Imaging for Discovery and Healthcare Applications”. en. In: *Cell Systems* 2 (2016), pp. 225–238.

[12] J Thiele and H M Kvasnicka. “Diagnostic differentiation of essential thrombocythaemia from thrombocythaemias associated with chronic idiopathic myelofibrosis by discriminate analysis of bone marrow features - a clinicopathological study on 272 patients”. en. In: (), p. 10.

[13] Jürgen Thiele et al. “Essential thrombocythemia versus early primary myelofibrosis: a multicenter study to validate the WHO classification”. eng. In: *Blood* 117 (2011), pp. 5710–5718.

[14] Ann Brinch Madelung et al. “World Health Organization-defined classification of myeloproliferative neoplasms: morphological reproducibility and clinical correlations–the Danish experience”. eng. In: *American Journal of Hematology* 88 (2013), pp. 1012–1016.

[15] Hans Michael Kvasnicka et al. “European LeukemiaNet study on the reproducibility of bone marrow features in masked polycythemia vera and differentiation from essential thrombocythemia”. en. In: *American Journal of Hematology* 92 (2017), pp. 1062–1067.

[16] Ann B. Madelung et al. “WHO classification 2008 of myeloproliferative neoplasms: a workshop learning effect – the Danish experience”. en. In: *APMIS* 123 (2015), pp. 787–792.

[17] Korsuk Sirinukunwattana et al. “Artificial intelligence–based morphological fingerprinting of megakaryocytes: a new tool for assessing disease in MPN patients”. In: *Blood Advances* 4 (2020), pp. 3284–3294.

[18] Milica Vukotić et al. “Inhibition of proinflammatory signaling impairs fibrosis of bone marrow mesenchymal stromal cells in myeloproliferative neoplasms”. en. In: *Experimental & Molecular Medicine* (2022), pp. 1–12.

[19] Jacob Grinfeld et al. “Classification and Personalized Prognosis in Myeloproliferative Neoplasms”. In: *New England Journal of Medicine* 379 (2018), pp. 1416–1430.

[20] Edoardo Peroni et al. “Spatial-transcriptomic profiling: a new lens for understanding myelofibrosis pathophysiology”. en. In: *Cell Communication and Signaling* 22 (2024), p. 510.

[21] Sabrina Cordua et al. “Prevalence and phenotypes of JAK2 V617F and calreticulin mutations in a Danish general population”. In: *Blood* 134 (2019), pp. 469–479.

[22] Nikolaos Sousos et al. “In utero origin of myelofibrosis presenting in adult monozygotic twins”. en. In: *Nature Medicine* 28 (2022), pp. 1207–1211.

[23] Debra Van Egeren et al. “Reconstructing the Lineage Histories and Differentiation Trajectories of Individual Cancer Cells in Myeloproliferative Neoplasms”. In: *Cell Stem Cell* 28 (2021), 514–523.e9.

[24] Jennifer Mary O’Sullivan, Adam J. Mead, and Bethan Psaila. “Single-cell methods in myeloproliferative neoplasms: old questions, new technologies”. In: *Blood* 141 (2023), pp. 380–390.

[25] Simón Méndez-Ferrer et al. “Bone marrow niches in haematological malignancies”. en. In: *Nature Reviews Cancer* 20 (2020), pp. 285–298.

[26] Maria Kleppe et al. “JAK-STAT Pathway Activation in Malignant and Non-Malignant Cells Contributes to MPN Pathogenesis and Therapeutic Response”. In: *Cancer discovery* 5 (2015), pp. 316–331.

[27] Ana Luísa Pereira, Serena Galli, and César Nombela-Arrieta. "Bone marrow niches for hematopoietic stem cells". en. In: *HemaSphere* 8 (2024), e133.

[28] Caroline Busch, Kudzai Nyamondo, and Helen Wheadon. "Complexities of modeling the bone marrow microenvironment to facilitate hematopoietic research". In: *Experimental Hematology* 135 (2024), p. 104233.

[29] I. L. Boueya et al. "Endothelial heterogeneity in bone marrow: insights across development, adult life and leukemia". en. In: *Leukemia* 39 (2025), pp. 8–24.

[30] Raymond K. H. Yip et al. "Towards deciphering the bone marrow microenvironment with spatial multi-omics". In: *Seminars in Cell & Developmental Biology* 167 (2025), pp. 10–21.

[31] Sandra Pinho and Paul S. Frenette. "Haematopoietic stem cell activity and interactions with the niche". en. In: *Nature Reviews Molecular Cell Biology* 20 (2019), pp. 303–320.

[32] Susann Winter et al. "Clonal hematopoiesis and its impact on the aging osteo-hematopoietic niche". en. In: *Leukemia* 38 (2024), pp. 936–946.

[33] Matthew J. Kim, Rodrigo J. Valderrábano, and Joy Y. Wu. "Osteoblast Lineage Support of Hematopoiesis in Health and Disease". In: *Journal of Bone and Mineral Research* 37 (2022), pp. 1823–1842.

[34] Shovik Bandyopadhyay et al. "Mapping the cellular biogeography of human bone marrow niches using single-cell transcriptomics and proteomic imaging". English. In: *Cell* 187 (2024), 3120–3140.e29.

[35] John Mascarenhas et al. "Biological drivers of clinical phenotype in myelofibrosis". en. In: *Leukemia* 37 (2023), pp. 255–264.

[36] Taylor B. Collins et al. "Altered erythropoiesis via JAK2 and ASXL1 mutations in myeloproliferative neoplasms". eng. In: *Experimental Hematology* 132 (2024), p. 104178.

[37] Elodie Grockowiak et al. "Different niches for stem cells carrying the same oncogenic driver affect pathogenesis and therapy response in myeloproliferative neoplasms". en. In: *Nature Cancer* 4 (2023), pp. 1193–1209.

[38] Andrew L. Young et al. "Spatial Mapping of Hematopoietic Clones in Human Bone Marrow". eng. In: *Blood Cancer Discovery* 5 (2024), pp. 153–163.

[39] Daniel A. C. Fisher et al. "Cytokine production in myelofibrosis exhibits differential responsiveness to JAK-STAT, MAP kinase, and NFκB signaling". en. In: *Leukemia* 33 (2019), pp. 1978–1995.

[40] Mohammed Ferdous-Ur Rahman et al. "Interleukin-1 contributes to clonal expansion and progression of bone marrow fibrosis in JAK2V617F-induced myeloproliferative neoplasm". en. In: *Nature Communications* 13 (2022), p. 5347.

[41] Shivam Rai et al. "Inhibition of interleukin-1β reduces myelofibrosis and osteosclerosis in mice with JAK2-V617F driven myeloproliferative neoplasm". en. In: *Nature Communications* 13 (2022), p. 5346.

[42] Abdallah Abou Zahr et al. "Bone marrow fibrosis in myelofibrosis: pathogenesis, prognosis and targeted strategies". In: *Haematologica* 101 (2016), pp. 660–671.

[43] Francesco Passamonti and Barbara Mora. "Myelofibrosis". In: *Blood* 141 (2023), pp. 1954–1970.

[44] Andrea Duminuco et al. "Myelofibrosis and Survival Prognostic Models: A Journey between Past and Future". en. In: *Journal of Clinical Medicine* 12 (2023), p. 2188.

[45] Kanjaksha Ghosh et al. "Fibrosis and bone marrow: understanding causation and pathobiology". en. In: *Journal of Translational Medicine* 21 (2023), p. 703.

[46] Hosuk Ryou et al. "Continuous Indexing of Fibrosis (CIF): improving the assessment and classification of MPN patients". en. In: *Leukemia* 37 (2023), pp. 348–358.

[47] Hélène FE Gleitz, Adam Benabid, and Rebekka K. Schneider. "Still a burning question: the interplay between inflammation and fibrosis in myeloproliferative neoplasms". In: *Current Opinion in Hematology* 28 (2021), pp. 364–371.

[48] Hélène F. E. Gleitz et al. "Increased CXCL4 expression in hematopoietic cells links inflammation and progression of bone marrow fibrosis in MPN". In: *Blood* 136 (2020), pp. 2051–2064.

[49] Kishor K. Sivaraj et al. "Endothelial LATS2 is a suppressor of bone marrow fibrosis". en. In: *Nature Cardiovascular Research* 3 (2024), pp. 951–969.

[50] Daniele Capitanio et al. "Proteomic screening identifies PF4/Cxcl4 as a critical driver of myelofibrosis". en. In: *Leukemia* 38 (2024), pp. 1971–1984.

[51] Rong Li et al. "A proinflammatory stem cell niche drives myelofibrosis through a targetable galectin-1 axis". In: *Science Translational Medicine* 16 (2024), eadj7552.

[52] Yijun Li, Stefan Stanojevic, and Lana X. Garmire. *Emerging Artificial Intelligence Applications in Spatial Transcriptomics Analysis*. Tech. rep. arXiv, 2022.

[53] Nils B. Leimkühler et al. "Heterogeneous bone-marrow stromal progenitors drive myelofibrosis via a druggable alarmin axis". In: *Cell Stem Cell* 28 (2021), 637–652.e8.

[54] Bethan Psaila et al. "Single-Cell Analyses Reveal Megakaryocyte-Biased Hematopoiesis in Myelofibrosis and Identify Mutant Clone-Specific Targets". en. In: *Molecular Cell* 78 (2020), 477–492.e8.

[55] Helene Gleitz et al. "2003 – PHYSICALLY INTERACTING CELL SEQUENCING IN BONE MARROW FIBROSIS REVEALS DISTINCT CELLULAR NICHES IN FIBROSIS". In: *Experimental Hematology*. SUPP:EXPHEM Abstracts 2024 137 (2024), p. 104560.

[56] Bella Banjanin et al. *Wnt-dependent spatiotemporal reprogramming of bone marrow niches drives fibrosis*. en. 2025.

[57] Hosuk Ryou et al. “Reticulin-Free Quantitation of Bone Marrow Fibrosis in MPNs: Utility and Applications”. en. In: *eJHaem* 6 (2025), e70005.

[58] Anna M Sozanska et al. “Entering the era of spatial transcriptomics: opportunities and challenges for pathology”. In: *Diagnostic Histopathology* 31 (2025), pp. 257–266.

[59] Alex To, Zou Yu, and Ryohichi Sugimura. “Recent advancement in the spatial immuno-oncology”. In: *Seminars in Cell & Developmental Biology* 166 (2025), pp. 22–28.

[60] Alvaro Gomariz et al. “Quantitative spatial analysis of haematopoiesis-regulating stromal cells in the bone marrow microenvironment by 3D microscopy”. en. In: *Nature Communications* 9 (2018), p. 2532.

[61] Chiara Baccin et al. “Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization”. en. In: *Nature Cell Biology* 22 (2020), pp. 38–48.

[62] Xue Xiao et al. “Spatial transcriptomic interrogation of the murine bone marrow signaling landscape”. en. In: *Bone Research* 11 (2023), pp. 1–13.

[63] Gerald Coulis et al. “Single-cell and spatial transcriptomics identify a macrophage population associated with skeletal muscle fibrosis”. In: *Science Advances* 9 (2023), eadd9984.

[64] Laura Sudupe et al. *Bone Marrow Spatial Transcriptomics Reveals a Myeloma Cell Architecture with Dysfunctional T-Cell Distribution, Neutrophil Traps, and Inflammatory Signaling*. en. 2024.

[65] Robert Stanley et al. “Single-Cell Spatial Transcriptomics Identifies Lymphoid Aggregates with Location-Specific Alterations of Adaptive Immune Cells in Human MDS Bone Marrow”. In: *Blood*. 66th ASH Annual Meeting Abstracts 144 (2024), p. 817.

[66] Christian M. Schürch et al. “Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front”. en. In: *Cell* 182 (2020), 1341–1359.e19.

[67] Enes Dasdemir et al. *Spatial Transcriptomics Reveals Inflammation and Trans-differentiation States of Acute Myeloid Leukemia in Extramedullary and Medullary Tissues*. en. 2024.

[68] Mara John et al. “Spatial transcriptomics reveals profound subclonal heterogeneity and T-cell dysfunction in extramedullary myeloma”. eng. In: *Blood* 144 (2024), pp. 2121–2135.

[69] Gege Gui et al. *Single Cell Spatial Transcriptomics Reveals Immunotherapy-Driven Bone Marrow Niche Remodeling in AML*. en. 2025.

[70] Neashan Mathavan et al. “Spatial transcriptomics in bone mechanomics: Exploring the mechanoregulation of fracture healing in the era of spatial omics”. In: *Science Advances* 11 (2025), eadp8496.

[71] Itziar Cenzano et al. *Single-cell transcriptional profiling reveals a novel RAB13+ endothelial subpopulation and profibrotic mesenchymal cells in the aged human bone marrow*. en. 2025.

[72] Weiqiang Lin et al. “Mapping the spatial atlas of the human bone tissue integrating spatial and single-cell transcriptomics”. In: *Nucleic Acids Research* 53 (2025), gkae1298.

[73] Sandro Bräunig et al. “Three-dimensional spatial mapping of the human hematopoietic microenvironment in healthy and diseased bone marrow”. en. In: *Cytometry Part A* 103 (2023), pp. 763–776.

[74] Research Communities by Springer Nature. *Niche heterogeneity influences the outcomes in myeloproliferative neoplasms*. en. 2023.

[75] Eleftherios Zormpas et al. “Mapping the transcriptome: Realizing the full potential of spatial data analysis”. en. In: *Cell* 186 (2023), pp. 5677–5689.

[76] Huw D. Summers, John W. Wills, and Paul Rees. “Spatial statistics is a comprehensive tool for quantifying cell neighbor relationships and biological processes via tissue image analysis”. English. In: *Cell Reports Methods* 2 (2022).

[77] Britta Velten and Oliver Stegle. “Principles and challenges of modeling temporal and spatial omics data”. en. In: *Nature Methods* 20 (2023), pp. 1462–1474.

[78] Anjali Rao et al. “Exploring tissue architecture using spatial transcriptomics”. en. In: *Nature* 596 (2021), pp. 211–220.

[79] Rafael dos Santos Peixoto et al. “Characterizing cell-type spatial relationships across length scales in spatially resolved omics data”. en. In: *Nature Communications* 16 (2025), p. 350.

[80] Gunsagar S. Gulati et al. “Profiling cell identity and tissue architecture with single-cell and spatial transcriptomics”. en. In: *Nature Reviews Molecular Cell Biology* 26 (2025), pp. 11–31.

[81] Benjamin L. Walker et al. “Deciphering tissue structure and function using spatial transcriptomics”. en. In: *Communications Biology* 5 (2022), pp. 1–10.

[82] Zhiyuan Yuan et al. “Benchmarking spatial clustering methods with spatially resolved transcriptomics data”. en. In: *Nature Methods* 21 (2024), pp. 712–722.

[83] Yunfei Hu et al. “Benchmarking clustering, alignment, and integration methods for spatial transcriptomics”. In: *Genome Biology* 25 (2024), p. 212.

[84] Chengwei Zhong, Kok Siong Ang, and Jinmiao Chen. “Interpretable spatially aware dimension reduction of spatial transcriptomics with STAMP”. en. In: *Nature Methods* 21 (2024), pp. 2072–2083.

[85] Boyi Guo et al. “Integrating Spatially-Resolved Transcriptomics Data Across Tissues and Individuals: Challenges and Opportunities”. en. In: *Small Methods* n/a (), p. 2401194.

[86] Jingyang Qian et al. “Identification and characterization of cell niches in tissue from spatial omics data at single-cell resolution”. en. In: *Nature Communications* 16 (2025), p. 1693.

[87] Uthsav Chitra et al. “Mapping the topography of spatial gene expression with interpretable deep learning”. en. In: *Nature Methods* 22 (2025), pp. 298–309.

[88] Doron Haviv et al. “The covariance environment defines cellular niches for spatial inference”. en. In: *Nature Biotechnology* 43 (2025), pp. 269–280.

[89] Martin Emons et al. *pasta: Pattern Analysis for Spatial Omics Data*. 2025.

[90] Xin Gao, Jing Zhang, and Owen J. Tamplin. “The aging hematopoietic stem cell niche: a mini review”. English. In: *Frontiers in Hematology* 4 (2025).

[91] Amanda Janesick et al. “High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis”. en. In: *Nature Communications* 14 (2023), p. 8353.

[92] Quentin Blampey et al. “Sopa: a technology-invariant pipeline for analyses of image-based spatial omics”. en. In: *Nature Communications* 15 (2024), p. 4981.

[93] Lidan Wu, Joseph M. Beechem, and Patrick Danaher. *FastReseg: using transcript locations to refine image-based cell segmentation results in spatial transcriptomics*. en. 2024.

[94] Sergio Marco Salas et al. “Optimizing Xenium In Situ data utility by quality assessment and best-practice analysis workflows”. en. In: *Nature Methods* (2025), pp. 1–11.

[95] Karolyn A. Oetjen et al. “Human bone marrow assessment by single-cell RNA sequencing, mass cytometry, and flow cytometry”. en. In: *JCI Insight* 3 (2018).

[96] Jeffrey M. Granja et al. “Single-cell multiomic analysis identifies regulatory programs in mixed-phenotype acute leukemia”. en. In: *Nature Biotechnology* 37 (2019), pp. 1458–1465.

[97] Dvir Aran et al. “Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage”. en. In: *Nature Immunology* 20 (2019), pp. 163–172.

[98] Arezou Rahimi et al. “DOT: a flexible multi-objective optimization framework for transferring features across single-cell and spatial omics”. en. In: *Nature Communications* 15 (2024), p. 4994.

[99] Rita Sarkis et al. “MarrowQuant 2.0: A Digital Pathology Workflow Assisting Bone Marrow Evaluation in Experimental and Clinical Hematology”. In: *Modern Pathology* 36 (2023), p. 100088.

[100] Robert Amezquita et al. *Orchestrating Single-Cell Analysis with Bioconductor*.

[101] Jia-Xing Sun et al. “SNAI1, an endothelial-mesenchymal transition transcription factor, promotes the early phase of ocular neovascularization”. eng. In: *Angiogenesis* 21 (2018), pp. 635–652.

[102] Qiaofeng Chen, Meiai Wang, and Shanpeng Wu. “The lncRNA MCF2L-AS1 controls osteogenic differentiation by regulating miR-33a”. eng. In: *Cell Cycle (Georgetown, Tex.)* 19 (2020), pp. 1059–1065.

[103] Anja Saalbach and Ulf Anderegg. “Thy-1: more than a marker for mesenchymal stromal cells”. en. In: *The FASEB Journal* 33 (2019), pp. 6689–6696.

[104] Christopher Mazzeo et al. “JAK2V617F mutant endothelial cells promote neoplastic hematopoiesis in a mixed vascular microenvironment”. In: *Blood Cells, Molecules, and Diseases* 90 (2021), p. 102585.

[105] Sébastien Malinge. “SNAIL trail in myeloid malignancies”. In: *Blood* 136 (2020), pp. 920–921.

[106] Lucas Wadley and Angela Fleischman. “Primary myelofibrosis progression: a game of cellular telephone”. In: *Haematologica* 110 (2025), pp. 822–824.

[107] Tatsuki Uchiyama et al. “The role of growth differentiation factor 15 in the pathogenesis of primary myelofibrosis”. en. In: *Cancer Medicine* 4 (2015), pp. 1558–1572.

[108] Drenka Trivanović et al. “Molecular Deconvolution of Bone Marrow Adipose Tissue Interactions with Malignant Hematopoiesis: Potential for New Therapy Development”. en. In: *Current Osteoporosis Reports* 22 (2024), pp. 367–377.

[109] Ruiying Chen, Roland Baron, and Francesca Gori. “Sfrp4 and the Biology of Cortical Bone”. en. In: *Current Osteoporosis Reports* 20 (2022), pp. 153–161.

[110] Rebekka Schneider and Ralf H. Adams. “Endothelial cells modulate bone marrow fibrosis”. en. In: *Nature Cardiovascular Research* 3 (2024), pp. 895–896.

[111] Benedetta Gaia Erba et al. “Endothelial-to-Mesenchymal Transition in Bone Marrow and Spleen of Primary Myelofibrosis”. In: *The American Journal of Pathology* 187 (2017), pp. 1879–1892.

[112] Maria K. Andersen et al. “Spatial transcriptomics reveals strong association between SFRP4 and extracellular matrix remodeling in prostate cancer”. en. In: *Communications Biology* 7 (2024), pp. 1–15.

[113] Marco Varrone et al. “CellCharter reveals spatial cell niches associated with tissue remodeling and cell plasticity”. en. In: *Nature Genetics* 56 (2024), pp. 74–84.

[114] Zhuorui Xie et al. “Gene Set Knowledge Discovery with Enrichr”. en. In: *Current Protocols* 1 (2021), e90.

[115] Xuechun Zhao et al. “Effect of dermatopontin on osteogenic differentiation of periodontal ligament stem cells”. In: *Gene* 858 (2023), p. 147185.

[116] Ashley C. Kramer et al. “An irradiated marrow niche reveals a small noncollagenous protein mediator of homing, dermatopontin”. In: *Blood Advances* 5 (2021), pp. 3609–3622.

[117] Lixin Lin et al. “BAMBI inhibits skin fibrosis in keloid through suppressing TGF-β1-induced hypernomic fibroblast cell proliferation and excessive accumulation of collagen I”. In: *International Journal of Clinical and Experimental Medicine* 8 (2015), pp. 13227–13234.

[118] Abdullah O. Khan et al. “Human Bone Marrow Organoids for Disease Modeling, Discovery, and Validation of Therapeutic Targets in Hematologic Malignancies”. In: *Cancer Discovery* 13 (2023), pp. 364–385.

[119] Han Yang et al. “The risk variant rs11836367 contributes to breast cancer onset and metastasis by attenuating Wnt signaling via regulating NTN4 expression”. In: *Science Advances* 8 (2022), eabn3509.

[120] Yuichiro Enoki et al. “Netrin-4 derived from murine vascular endothelial cells inhibits osteoclast differentiation in vitro and prevents bone loss in vivo”. In: *FEBS Letters* 588 (2014), pp. 2262–2269.

[121] Nicole K. H. Yiew et al. “A novel role for the Wnt inhibitor APCDD1 in adipocyte differentiation: Implications for diet-induced obesity”. English. In: *Journal of Biological Chemistry* 292 (2017), pp. 6312–6324.

[122] Jialin Li et al. “VSIG4 inhibits proinflammatory macrophage activation by reprogramming mitochondrial pyruvate metabolism”. en. In: *Nature Communications* 8 (2017), p. 1322.

[123] Xiaoyong Huang et al. “VSIG4 mediates transcriptional inhibition of Nlrp3 and Il-1β in macrophages”. In: *Science Advances* 5 (2019), eaau7426.

[124] Mohammad Jafar Sharifi et al. “Immune-dysregulation harnessing in myeloid neoplasms”. en. In: *Cancer Medicine* 13 (2024), e70152.

[125] Maria N. Barrachina et al. “Acute CCL5 Exposure Promotes Megakaryopoiesis through Non-Canonical Pathways”. In: *Blood*. 66th ASH Annual Meeting Abstracts 144 (2024), p. 1171.

[126] Leila Noetzli et al. “CCL5 Enhances Megakaryocyte Differentiation and Development”. In: *Blood* 132 (2018), p. 522.

[127] Zhongwen Zhang et al. “Chemokine Receptor 5, a Double-Edged Sword in Metabolic Syndrome and Cardiovascular Disease”. English. In: *Frontiers in Pharmacology* 11 (2020).