UNIVERSITÀ DEGLI STUDI DI MILANO

FACOLTÀ DI MEDICINA E CHIRURGIA

PhD program in TRANSLATIONAL MEDICINE
XXXVII cycle

DEPARTMENT OF BIOMEDICAL SCIENCES FOR HEALTH

Multiple Myeloma Plasma Cell Profiling: Integrated Single-Cell RNA and VDJ Sequencing Approaches Uncover Novel Clinical Insights

PhD CANDIDATE:
Antonio MATERA
[MT: R13292]

TUTOR: Prof. Niccolò BOLLI

CO-TUTORS: Dott. Matteo Claudio DA VIÀ & Dott.ssa Francesca LAZZARONI

REVIEWER: Dott. Zewen Kelvin TUONG

COORDINATOR: Prof.ssa Chiarella SFORZA

A.A.
2023/2024

Abstract

Multiple Myeloma (MM) is a hematological malignancy characterized by the proliferation of antibody-secreting bone marrow (BM) plasma cells (PCs). It is preceded by asymptomatic stages defined as Monoclonal Gammopathy of Undetermined Significance (MGUS) or Smoldering MM (SMM). MM is characterized by complex clonal architectures and crucial tumor microenvironment (TME) interactions. As a matter of fact, disease progression is driven by the dysregulated interaction of the PC population with the tumor microenvironment (TME). The PC population comprises clonal PCs (cPCs) and polyclonal PCs (pPCs). Although the roles of immune and non-immune cells have been extensively studied, the disruption of pPCs and their role in TME dysregulation have not been deeply studied.
This thesis presents a cutting-edge investigation into the clonal and polyclonal landscape of MM, leveraging the power of single-cell RNA and VDJ sequencing (scRNAseq and scVDJseq) to provide a detailed exploration of both the genetic and transcriptomic landscapes of MM at single-cell resolution, with particular emphasis on pPCs, TME interactions and clonal evolution.

The first major outcome of this work is the validation of scVDJseq as an effective tool for identifying MM-specific VDJ rearrangements, enabling detailed clonal tracking with higher precision. Our results demonstrate that single-cell analysis could reach a better resolution, uncovering key VDJ rearrangements in cPCs compared with traditional bulk DNA sequencing. Specifically, scVDJseq identified clonotypes and new subclonal populations, providing a comprehensive view of clonal diversity within the BM of MM patients. This technology also enabled precise identification of pPCs at various disease stages, offering a new framework for understanding clonal evolution from early asymptomatic to advanced symptomatic disease.
Integrating with transcriptome data, we obtained crucial insights about pPCs. These exhibited distinct phenotypic changes, particularly in their interaction with the TME. Interestingly, we identified a progressive loss of key immune markers in pPCs across disease stages. This suggests that these cells may be underestimated in immune surveillance and the immune system's ability to recognize and respond to malignant cells. Moreover, the data showed that pPCs from MM patients undergo significant transcriptomic dysregulation, marked by increased expression of autophagy-related and inflammatory pathways, a pattern not observed in healthy donors (HDs). These findings suggest that pPCs may be actively reshaped by the TME, potentially contributing to immune evasion and disease progression.

The dissertation also explores the phenomenon of immunoparesis, a hallmark of PC dyscrasia, in greater detail. Using scRNAseq data, we uncovered a unique gene signature in pPCs associated with immunoparesis, characterized by the upregulation of inflammatory and stress-related pathways. These findings point to a mechanistic link between the disruption of pPCs function and the failure of the immune system to surveil the tumor adequately, highlighting the complex interplay between pPCs, cPCs, and the TME. Additionally, our analysis revealed a significant depletion of pPCs in patients with immunoparesis, correlating with a reduction in immunoglobulin production and a more pro-inflammatory TME environment.
The key result is the identification of a "healthy plasma cell" (hPC) signature. This novel signature, derived from PCs in HDs, was progressively lost across MM disease stages, correlating with poorer clinical outcomes. Patients whose polyclonal cells retained this hPC signature had significantly better progression-free survival (PFS) and overall survival (OS), suggesting that the preservation of functional pPCs may serve as a protective factor against disease progression.
Building on these findings, a groundbreaking analysis of clonal evolution was conducted using the Dandelion tool for VDJ tracking. This tool allowed for precise mapping of clonal complexity and diversity as the disease progressed from asymptomatic to symptomatic stages. Suggesting that as MM progresses, clonal diversity expands, potentially contributing to treatment resistance.

In conclusion, this dissertation advances our understanding of MM by leveraging cutting-edge single-cell technologies to dissect the clonal architecture and the role of the polyclonal counterpart, which is often overlooked. The research not only provides a refined picture of MM clonal dynamics but also underscores the potential clinical value of tracking pPCs signatures. These findings open the door to new research directions, including the development of targeted therapies aimed at restoring immune surveillance and normal PC function, with the ultimate goal of improving patient outcomes in MM.

List of publications

Used for thesis:

The results of this thesis have been published in the following three scientific articles, referred to as Article I, Article II, and Article III.

- **ARTICLE I:** Single-Cell RNA Sequencing for the Detection of Clonotypic V(D)J Rearrangements in Multiple Myeloma

  Matera A∗, Marella A, Maeda A, Da Và MC, Lazzaroni F, Fabris S, Pioggia S, Porretti L, Colombo F, Torricelli F, Neri A, Taiana E, Fabbiano G, Traini V, Genuardi E, Drandi D, Bolli N, Lionetti M.

  Int J Mol Sci. 2022 Dec 10;23(24):15691. doi: 10.3390/ijms232415691. PMID: 36555330; PMCID: PMC9779610.

- **ARTICLE II:** Inference of genomic lesions from single-cell RNA-seq in myeloma improves functional intraclonal and interclonal analysis

  Lazzaroni F∗, **Matera A∗**, Marella A, Maeda A, Castellano G, Marchetti A, Fabris S, Pioggia S, Silvestris I, Ronchetti D, Lonati S, Fabbiano G, Traini V, Taiana E, Porretti L, Colombo F, De Magistris C, Scopetti M, Barbieri M, Pettine L, Torricelli F, Neri A, Passamonti F, Lionetti M, Da Và MC, Bolli N.

  Blood Adv. 2024 Aug 13;8(15):3972-3984. doi: 10.1182/bloodadvances.2023012409. PMID: 38830132; PMCID: PMC11331727.

- **ARTICLE III:** Genotypic identification of polyclonal plasma cells in plasma cell dyscrasias shows an aberrant single-cell phenotype with clinical implications

  Matteo Claudio Da Và∗, Francesca Lazzaroni∗, **Antonio Matera**, Alessio Marella, Akhiro Maeda, Claudio De Magistris, Loredana Pettine, Sonia Fabris, Stefania Pioggia, Alfredo Marchetti, Marzia Barbieri, Silvia Lonati, Alessandra Cattaneo, Marta Tornese, Margherita Scopetti, Nayyer Latifinavid, Giancarlo Castellano, Federica Torricelli, Antonino Neri, Cathelijne Fokkema, Tom Coupedo, Marta Lionetti, Francesco Passamonti, Niccolò Bolli. Under Revision

  bioRxiv 2024.05.26.595470; doi: https://doi.org/10.1101/2024.05.26.595470

∗ These authors contributed equally.

Other Publications:

- **MGUS and clonal hematopoiesis show unrelated clinical and biological trajectories in an older population cohort**

  Da Vià MC, Lionetti M, Marella A, **Matera A**, Travaglino E, Signaroldi E, Galbussera AA, Lucca U, Mandelli S, Riva E, Tettamanti M, Pettine L, Pompa A, Baldini L, Neri A, Della Porta MG, Bolli N.

  Blood Adv. 2022 Nov 8;6(21):5702-5706. Doi: 10.1182/bloodadvances.2021006498. PMID: 35390146; PMCID: PMC9618778.

- **Clonal hematopoiesis is clonally unrelated to multiple myeloma and associated with specific microenvironmental changes**

  Marta Lionetti∗, Margherita Scopetti∗, **Antonio Matera**, Akihiro Maeda, Alessio Marella, Francesca Lazzaroni, Giancarlo Castellano, Sonia Fabris, Stefania Pioggia, Silvia Lonati, Alfredo Marchetti, Alessandra Cattaneo, Marta Tornese, Antonino Neri, Claudia Leoni, Claudio De Magistris, Loredana Pettine, Valentina Traini, Ilaria Silvestris, Marzia Barbieri, Giuseppina Fabbiano, Domenica Ronchetti, Elisa Taiana, Matteo C. Da Vià, Francesco Passamonti, and Niccolò Bolli. Under Revision

∗ These authors contributed equally.

Abstract / Congresses:

- **XVII National Congress SIES – Rome 2022**
  "Caratterizzazione su singola cellula dell'architettura clonale e trascrizionale e interazione con il microambiente nel mieloma smoldering"
  Matteo Da Vià, Akihiro Maeda, Alessio Marella, **Antonio Matera**, Antonino Neri, Marta Lionetti, Niccolò Bolli

- **IV National Congress SOHO (SOciety of Hematologic Oncology) – Rome 2022**
  "Single cell RNA sequencing for detection of clonotypic V(D)J rearrangements detection in multiple myeloma"
  Antonio Matera, Alessio Marella, Akihiro Maeda, Matteo C Da Vià, Francesca Lazzaroni, Sonia Fabris, Stefania Pioggia, Antonino Neri, Niccolò Bolli, Marta Lionetti

- **National Congress Under 40 in Hematology – Milan 2022**
  "Single cell RNA sequencing for detection of clonotypic V(D)J rearrangements in multiple myeloma"
  Antonio Matera, Alessio Marella, Akihiro Maeda, Matteo C Da Vià, Francesca Lazzaroni, Sonia Fabris, Stefania Pioggia, Antonino Neri, Niccolò Boll, Marta Lionetti

- **National Congress Under 40 in Hematology – Milan 2023**
  "Inference of genomic data from single-cell RNAseq in asymptomatic multiple myeloma allows precise functional intra- and inter-clonal analysis"
  F. Lazzaroni & **A. Matera**, A. Marella, A. Maeda, A. Marchetti, S. Fabris, S. Pioggia, I. Silvestris, D. Ronchetti, S. Lonati, G. Fabbiano, V. Traini, E. Taiana, L. Porretti, F. Colombo, C. De MagistrisM. Scopetti, A. Pompa, A. Neri, F. Passamonti, M. Lionetti, M.C Da Vià, N. Bolli

- **XVIII National Congress SIES – Florence 2024**
  "Functional dissection of the single cell transcriptomic landscape of genotypically-identified polyclonal plasma cells in multiple myeloma microenvironment"
  M.C. Da Vià, F. Lazzaroni, **A. Matera**, A. Marella, A. Maeda, A. Marchetti, S. Fabris, S. Pioggia, I. Silvestris, D. Ronchetti, S. Lonati, G. Fabbiano, V. Traini, E. Taiana, G. Castellano, M. Barbieri, L. Pettine, A. Neri, M. Lionetti, F. Passamonti, N. Bolli

  "Investigating the functional and clinical impact of clonal hematopoiesis of indeterminate potential in multiple myeloma pathogenesis"
  M. Lionetti, M. Scopetti, **A. Matera**, A. Maeda, A. Marella, F. Lazzaroni, G. Castellano, S. Fabris, S. Pioggia, S. Lonati, A. Cattaneo, M. Tornese, C. De Magistris, L. Pettine, M.C. Da Vià, F. Passamonti, N. Bolli

- **5th European Myeloma Network meeting – Turin 2024**
  "Functional dissection of the single cell transcriptomic landscape of Genotypically-identified polyclonal plasma cells in multiple myeloma Microenvironment"
  F. Lazzaroni, M.C. Da Via, **A. Matera**, A. Marella, A. Maeda, A. Marchetti, S. Fabris, S. Pioggia, I. Silvestris, D. Ronchetti, S. Lonati, G. Fabbiano, V. Traini, E. Taiana, G. Castellano, M. Barbieri, L. Pettine, A. Neri, M. Lionetti, F. Passamonti, N. Bolli

- **29th European Hematology Association Congress – Madrid 2024**
  "Functional dissection and clinical relevance of the single cell transcriptomic landscape of genotypically-identified polyclonal plasma cells in multiple myeloma microenvironment"
  M.C. Da Via, F. Lazzaroni, **A. Matera**, A. Marella, A. Maeda, A. Marchetti, S. Fabris, S. Pioggia, I. Silvestris, D. Ronchetti, S. Lonati, G. Fabbiano, V. Traini, E. Taiana, G. Castellano, M. Barbieri, L. Pettine, A. Neri, M. Lionetti, F. Passamonti, N. Bolli

  "Investigating the functional and clinical impact of clonal hematopoiesis of indeterminate potential in multiple myeloma pathogenesis"
  M. Lionetti, M. Scopetti, **A. Matera**, A. Maeda, A. Marella, F. Lazzaroni, G. Castellano, S. Fabris, S. Pioggia, S. Lonati, A. Cattaneo, M. Tornese, C. De Magistris, L. Pettine, M.C. Da Via, F. Passamonti, N. Bolli

- **7th Translational Research Conference on MULTIPLE MYELOMA from European School of Hematology (ESH) – Mandelieu-La Napoule 2024**
  "Geno-transcriptomic profiling of Polyclonal Plasma Cells uncovers a unique single-cell phenotype in plasma cell dyscrasias with clinical implications"
  A. Matera, M.C. Da Via, F. Lazzaroni, A. Marella, A. Maeda, A. Marchetti, S. Fabris, S. Pioggia, I. Silvestris, D. Ronchetti, S. Lonati, G. Fabbiano, V. Traini, E. Taiana, G. Castellano, M. Barbieri, L. Pettine, A. Neri, M. Lionetti, F. Passamonti, N. Bolli

- **National Congress Under 40 in Hematology – Milan 2024**
  "Geno-transcriptomic profiling of Polyclonal Plasma Cells uncovers a unique single-cell phenotype in plasma cell dyscrasias with clinical implications"
  A. Matera, M.C. Da Via, F. Lazzaroni, A. Marella, A. Maeda, A. Marchetti, S. Fabris, S. Pioggia, I. Silvestris, D. Ronchetti, S. Lonati, G. Fabbiano, V. Traini, E. Taiana, G. Castellano, M. Barbieri, L. Pettine, A. Neri, M. Lionetti, F. Passamonti, N. Bolli

Awards

Winner of the call "**bando estivo INDACO 2024**".

Prize: 25000 hours of HPC computing and 1TB storage.

Name of the project: **MR**eMMaine**D**

Second-level master's course

Bioinformatics and Functional Genomics A.Y. 2022/23 - UNIMI

Dissertation**: "Single-Cell Multi-omic Inference of eGRN on Breast Cancer Tumor-Initiating Cells"**

Tutors: Dott. F. Nicassio (IIT) and Dott. M. J. Marzi (IIT)

List of Contents

Abbreviations **13**

Multiple Myeloma **15**

Epidemiology *17*

Multiple Myeloma Pathogenesis *18*

Tumor Heterogeneity and cancer evolution *23*

BM Microenvironment *27*

Minimal Residual Disease *29*

Single cell RNA sequencing **32**

scRNAseq - Application on Multiple Myeloma *33*
scRNAseq - Focus on Plasma Cells 34
scRNAseq – Focus on Tumor Microenvironment 35

Aims **37**

Materials and Methods **39**

Sample collection *39*

External data collection *39*

Sample processing *40*

Thawing Cells *40*

DNA-Based Molecular Analysis of IGHV Rearrangement *40*

5' scRNAseq and scVDJseq: library construction and sequencing *40*

scRNAseq and scVDJseq: Bioinformatic pipelines *41*
scRNAseq and scVDJseq: Pre-processing 42
scRNAseq and scVDJseq: Seurat Analysis 43
Polyclonal Plasma cells determinations 45
In-silico MRD Generation 46
scRNAseq and scVDJseq: Bioinformatic Tools 46
Identification of IGH Translocations 46
Single-Cell Copy Number Variation Analysis 47
Single-Cell Differential Expression Analysis 47
Cell-Cell Communication and Interaction Analysis 48
Single-cell HD Signature and Scoring 49
VDJ analysis – Dandelion 50

Results **51**

Identification of Clonal BCR Rearrangements in MM using scVDJseq [Article I] *51*

Analysis of BCR Diversity and Transcriptomic Markers in Single Cell [Article II] *52*

Inference of Genomic Properties from scRNAseq [Article II] *54*

Single-Cell Transcriptome of Polyclonal PCs across PCs dyscrasia [Article III] *57*

Polyclonal Plasma Cells from Patients Exhibit Corrupted Phenotype [Article III] *58*

Single-Cell Transcriptomic Insights into Immunoparesis [Article III] *59*

Unraveling the Interactome Landscape of Immunoparesis in Polyclonal Plasma Cells [Article III] *60*

Signature Analysis as a Predictor of Survival in Independent Datasets [Article III] 62
Clonal VDJ tracking [Manuscript in preparation] 63
Discussion **66**
References **75**
Figures **94**
Tables **123**

Abbreviations

5' scRNAseq - 5' Single-Cell RNA Sequencing
5' scVDJseq - 5' Single-Cell VDJ Sequencing
AID - Activation-Induced Cytidine Deaminase
ALL - Acute Lymphoblastic Leukemia
ASCT - Autologous Stem Cell Transplantation
BAM - Binary Alignment Map (file format for storing sequence data)
BCMA - B-Cell Maturation Antigen (or TNFRSF17)
BCR - B Cell Receptor
BME - Bone Marrow Environment
BMMCs - Bone Marrow Mononuclear Cells
BMSC - Bone Marrow Stromal Cells
CCA - Canonical Correlation Analysis
CD14 Mono - CD14 Monocytes
CD16 Mono - CD16 Monocytes
CDR3 - Complementarity-Determining Region 3
cDC2 - Conventional Dendritic Cell 2
CITE-seq - Cellular Indexing of Transcriptomes and Epitopes by Sequencing
CRAB - Calcium increase, Renal insufficiency, Anemia, Bone lesions
CNA - Copy Number Alteration
CNV - Copy Number Variation
cPC - Clonal Plasma Cells
CSC - Cancer Stem Cells
CSR - Class Switch Recombination
CyTOF - Cytometry by Time of Flight
DC - Dendritic Cells
DEG - Differential Expression Gene
DNAseq - DNA Sequencing
ER - Endoplasmic Reticulum
FBS - Fetal Bovine Serum
FISH - Fluorescence In Situ Hybridization
FUSCIA - Fusion Transcript Analysis Tool
GC - Germinal Center
GEO - Gene Expression Omnibus
GEX - Gene Expression
GI - Genomic Instability
GMP - Granulocyte-Monocyte Progenitor
GO - Gene Ontology
GSEA - Gene Set Enrichment Analysis
H&E - Hematoxylin and Eosin
HD - Healthy Donor
HMM - Hidden Markov Model
HPC - High-Performance Computing
hPC - Healthy Plasma Cells
HRD - Hyperdiploidy
HSC - Hematopoietic Stem Cells
IGH - Immunoglobulin Heavy Chain

IGK/L - Immunoglobulin kappa/lambda Chain
IGHV - Immunoglobulin Heavy Chain – V gene
iMSC - Inflammatory Mesenchymal Stromal Cells
IMWG - International Myeloma Working Group
IRCCS - Istituto di Ricovero e Cura a Carattere Scientifico
ISH - In Situ Hybridization
LMPP - Lymphoid-Primed Multipotent Progenitor
ECM - Extracellular Matrix
MHC - Major Histocompatibility Complex
MSigDB - Molecular Signature Database
MM - Multiple Myeloma
MMP - Matrix Metalloproteinases
MRD - Minimal Residual Disease
NGF - Next-Generation Flow (cytometry)
NGS - Next-Generation Sequencing
MGUS – Monoclonal Gammopathy of Undetermined Significance
NHEJ - Non-Homologous End Joining
NMD - Nonsense-Mediated Decay
NP - Non-Progressors
OS - Overall Survival
OSR - Overall Survival Rate
PC – Plasma Cells
PCA - Principal Component Analysis
PBS - Phosphate-Buffered Saline
PCL - Plasma Cell Leukemia
PFS – Progression Free Survival
pPC - Polyclonal Plasma Cells
QC - Quality Control
RNAseq - RNA Sequencing
RP - Rapid-Progressors
RPCA - Reciprocal Principal Component Analysis
RSS - Recombination Signal Sequences
SHM - Somatic Hypermutation
SLiM - Sixty percent plasma cells, Light chain ratio, MRI focal lesions
sMP - Serum Monoclonal Protein
SMM - Smoldering Multiple Myeloma
SNP - Single Nucleotide Polymorphism
TME - Tumor Microenvironment
UMAP - Uniform Manifold Approximation and Projection
UMI - Unique Molecular Identifier
UPR - Unfolded Protein Response
VDJ - Variable, Diversity, and Joining
WGS -Whole Genome Sequencing
WM - Waldenström macroglobulinemia

Multiple Myeloma

Cancer pathogenesis is believed to follow a universally similar process. Each cancer results from a mechanism akin to Darwinian evolution occurring within cell populations in the microenvironments of a multicellular organism's tissues. Much like the Darwinian evolution observed in the development of species, cancer development involves two fundamental processes: the continuous acquisition of genetic variations in cells through random mutations and natural selection acting on the resulting phenotypic diversity[1].

This multi-step nature of cancer development has been validated and extensively investigated in plasma cell (PC) dyscrasias, such as multiple myeloma (MM). Although numerous studies are currently underway, significant progress is still required.

MM is characterized by the malignant proliferation of PCs within the bone marrow (BM)[2] (**Figure I**).

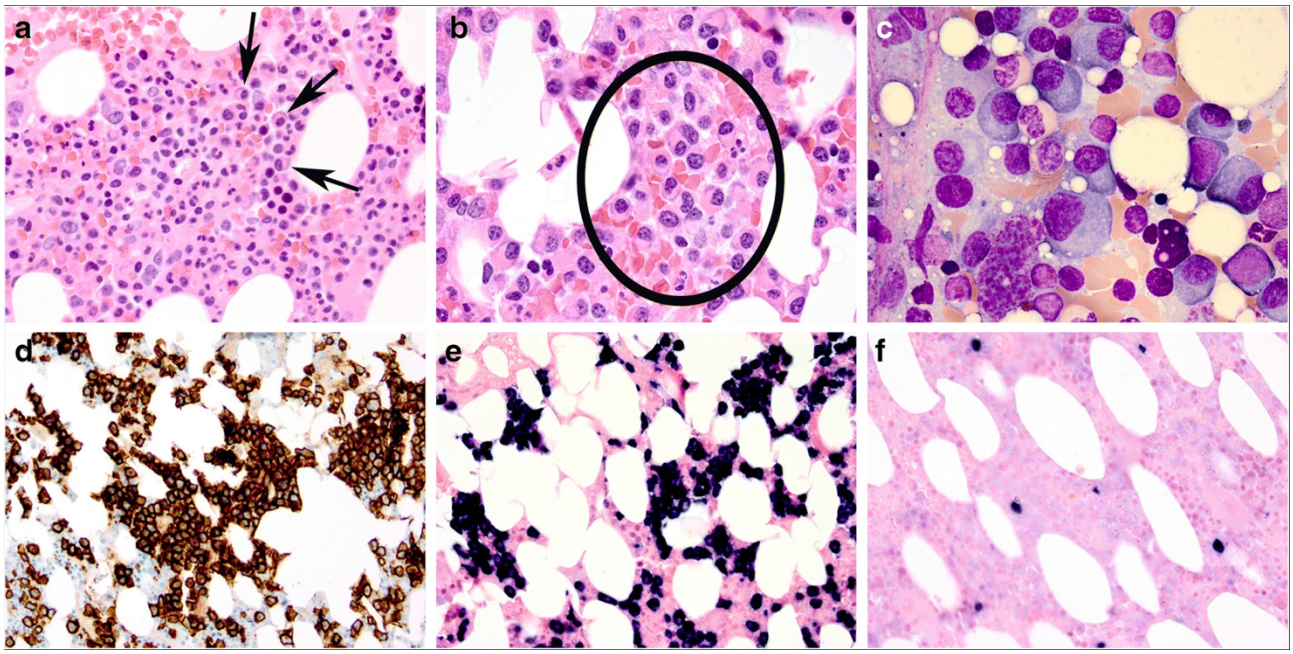

Figure I- Pathologic phenotype of PC dyscrasia cells involving BM biopsy and aspirate. ***a.*** *(H&E, ×200) arrows show scattered PCs present singly in some areas of BM and* ***b.*** *more conspicuous clusters in other areas (×400).* ***c.*** *(Wright-Giemsa, ×1000) The BM aspirate smear contains increased numbers of PCs with eccentric nuclei and prominent Golgi apparatus; some forms are atypical, with slight chromatin dispersal and subtle nuclear irregularities.* ***d.*** *(×200) Immunohistochemistry for CD138.* ***e.*** *(×200) Immunohistochemistry for kappa light chain over* ***f.*** *(×200) lambda light chain staining[3,4]. Image adapted from Palumbo, A., and Anderson, K. et al. (2011).*

The malignant PCs secrete a monoclonal, non-functional immunoglobulin referred to as M protein, which accumulates in the serum and/or urine. This accumulation results in organ damage and manifests the characteristic clinical symptoms of the disease.

MM typically progresses from an asymptomatic condition known as Monoclonal Gammopathy of Undetermined Significance (MGUS) to Smoldering MM (SMM) and eventually to symptomatic

MM[5,6]. Not all patients with MGUS or SMM will progress to symptomatic disease, but those who do often exhibit specific clinical features that signal the need for treatment. These features are identified using the CRAB[7] and SLiM[8] criteria, which are critical in determining the disease stage and guiding therapeutic decisions[9–11].

The CRAB criteria have traditionally been used to identify symptomatic MM, meaning the disease is active and requires immediate treatment. CRAB is an acronym representing four key clinical manifestations associated with the progression of myeloma[12] (**Table I**).

In 2014, the International Myeloma Working Group (IMWG) introduced the SLiM criteria, expanding upon the CRAB criteria to include patients at imminent risk of developing symptoms, even if they have not yet exhibited the classic CRAB features (**Table I**).

- MM: Patients who meet one or more CRAB and/or SLiM criteria and require immediate treatment.
- SMM: Patients who do not meet the CRAB or SLiM criteria but have elevated levels of M-protein in the blood and/or BM. These patients are closely monitored but do not require immediate treatment.
- MGUS: A benign condition in which clonal PC or presence of sMP, but there are no symptoms or organ damage (absence of CRAB and SLiM criteria). MGUS does not require treatment but does require regular monitoring for potential progression to MM or related diseases.

Laboratory diagnosis is done considering the combination of different approaches: biochemistry, immunohistochemistry and cytohistology.

MM accounts for 10% of hematological malignancies[8]. Despite advances in treatment that can extend patients' lives, MM remains an incurable PC disorder, and only a small group of patients achieving long-term disease remission MM exhibits significant clinical and genetic heterogeneity[13].

Approximately 1.5% of MGUS patients progress to MM annually, while SMM patients face a higher overall progression risk of 10% per year[14,15]. SMM, like MM, is a heterogeneous condition; some patients have over a 50% risk of progression within two years, while others exhibit a more MGUS-like, slower-growing disease[16].

Due to the high incidence of cases, there is considerable interest in elucidating the primary mechanisms of the disease to develop more personalized therapies and enhance overall survival rates (OSR). The advent of new drugs, alongside conventional chemotherapy and transplantation for eligible patients, has already improved therapeutic responses, quality of life, and extended remission periods[17]. Nevertheless, the relapse rate for MM patients remains significantly high.

Table I - Detailed CRAB and SLiM criteria

CRITERIA	INDICATOR	DETAILS	SIGNIFICANCE
CRAB criteria	**C** – Calcium increase	Serum calcium >11 mg/dL or above normal	Indicates bone resorption due to myeloma cell activity
	R - Renal insufficiency	Serum creatinine >2 mg/dL or creatinine clearance <40 mL/min	Renal impairment due to light chain deposition in kidneys
	A - Anemia	Hemoglobin <10 g/dL or 2 g/dL below normal	BM infiltration by myeloma cells leading to decreased RBC production
	B - Bone lesions	Presence of osteolytic lesions or fractures	Osteoclastic activity driven by tumor cells causing bone damage
SLiM Criteria	**S** - Sixty percent plasma cells	≥60% clonal PCs in BM	High risk of progression to symptomatic MM
	Li - Light chain ratio	Involved/uninvolved serum free light chain ratio ≥100 with involved light chain ≥100 mg/L	Indicates high production of light chains by pathological PCs
	M - MRI focal lesions	More than one focal lesion on MRI	Predicts disease progression due to BM infiltration

Epidemiology

MM is the second most common hematologic cancer, following non-Hodgkin lymphoma[18]. In the world, the highest incidence rate is seen in developed regions like the United States and Western Europe (**Figure II**). According to the latest Globocan data[19], in 2022, the global incidence of MM cases was 2.4 per 100,000 person-years. In Europe, the average incidence is 9.9 per 100,000 person-years, with Italy having the second-highest incidence at 10.5 per 100,000 person-years. MM is slightly more prevalent in males than females and is twice as common in African Americans compared to Caucasians[5]. The average age at diagnosis is around 65 years[20].

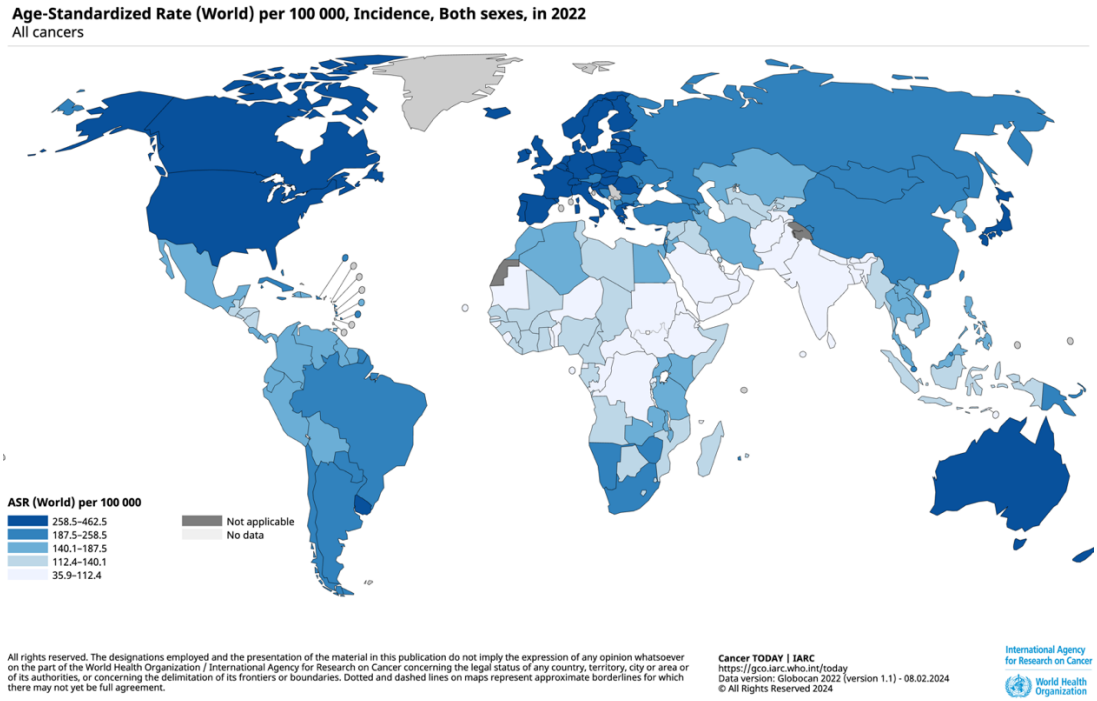

Figure II - Estimated incidence rate of MM in 2022. *Map obtained from GLOBOCAN 2022*

Multiple Myeloma Pathogenesis

PC formation is a vital process in the adaptive immune system, culminating in the production of high-affinity antibodies that provide targeted defense against pathogens. This complex process begins with the activation of naive B cells, typically occurring in the secondary lymphoid organs such as lymph nodes and the spleen. Here, naive B cells encounter specific antigens, usually presented by dendritic cells (DC), which bind to the B cell receptor (BCR), initiating the B cell activation cascade. This activation can be either T cell-dependent, involving help from T follicular helper cells, or T cell-independent, through direct antigen stimulation.

Naive B cells originate from hematopoietic stem cells (HSC) in the BM, where they undergo a series of genetic rearrangements to create a unique BCR. This diversity is achieved through the recombination of V (variable), D (diversity), and J (joining) gene segments in the immunoglobulin heavy chain (IGH) locus and V and J segments in the immunoglobulin light chain (IGL/K) locus. The V(D)J recombination process, facilitated by the recombination-activating genes (RAG1 and RAG2), is critical for generating the vast repertoire of BCRs necessary for recognizing a wide array of antigens[21]. During V(D)J recombination, the RAG enzymes introduce double-strand breaks at specific recombination signal sequences (RSS), flanking the V, D, and J gene segments. These segments are then joined together through a complex repair process involving the non-homologous end joining (NHEJ) pathway, resulting in the formation of a complete variable region exon. This exon

encodes the antigen-binding site of the BCR, which is crucial for the specificity and diversity of the antibody response[22].

Once activated, B cells migrate to the germinal centers (GC) of the lymphoid organs, where they undergo rapid proliferation and differentiation. Within the GC, two critical processes—somatic hypermutation (SHM) and class switch recombination (CSR)—refine the antibody response. SHM introduces point mutations at an extraordinarily high rate into the variable (V) regions of the immunoglobulin genes. Particularly in the complementarity-determining regions (CDRs), which directly interact with the antigen. These regions are crucial for the antibody's ability to bind to specific antigens with high affinity. CDR1 and CDR2 are found in the variable (V) region of a polypeptide chain, and CDR3 includes some of V, all of diversity (D, heavy chains only) and joining (J) regions. CDR3 is the most variable[23]. Mutations in these regions generate a diverse pool of B cells with varying affinities for the antigen. Selection mechanisms within the GC ensure that B cells expressing high-affinity BCRs are preferentially expanded and retained[24].

Concurrently, CSR modifies the constant (C) region of the antibody, altering its isotype while preserving antigen specificity. This isotype switching is crucial as it enables the immune system to deploy antibodies with different effector functions, such as switching from IgM to IgG, IgA, or IgE. These different isotypes can mediate distinct immune responses, enhancing the versatility and effectiveness of the immune defense[25].

A pivotal player in both SHM and CSR is the enzyme activation-induced cytidine deaminase (AID)[26]. AID deaminates cytosine residues in DNA, converting them to uracil. This enzymatic activity creates U-mismatches, which are processed differently depending on the context. In SHM, the mismatches are targeted by error-prone DNA repair pathways, resulting in mutations concentrated in the V regions of the Ig genes. In contrast, during CSR, these U-mismatches lead to double-strand breaks in the switch regions, facilitating recombination events that change the antibody isotype[27].

The GC reaction is a highly regulated environment where B cells undergo iterative cycles of mutation and selection. B cells that acquire beneficial mutations in their BCRs are selected for survival and proliferation, whereas those with deleterious mutations are eliminated. This Darwinian process of selection ensures the production of B cells with high-affinity antibodies.

Following SHM and CSR, B cells differentiate into either memory B cells or PCs. Memory B cells circulate in the body, providing a rapid and robust response upon re-exposure to the antigen. PCs, on the other hand, migrate to the BM or mucosal tissues, where they become specialized in the production and secretion of large amounts of antibodies. These antibodies are the primary effectors of the humoral immune response, neutralizing pathogens and marking them for destruction by other

components of the immune system. Overall, the formation of PCs from naive B cells is a multifaceted process involving intricate molecular mechanisms and precise regulatory controls. The roles of SHM and CSR, mediated by AID, are central to generating the antibody diversity and specificity necessary for effective immune protection. Understanding these processes not only elucidates normal immune function but also sheds light on the pathological mechanisms underlying hematological malignancies such as MM, where dysregulated PC proliferation leads to severe clinical consequences. Indeed, MM originates in the GC of a secondary follicle. Here, a mature B cell, referred to as the "cell of origin," undergoes transformation due to errors during CSR and/or SHM (**Figure III**)[28–30].

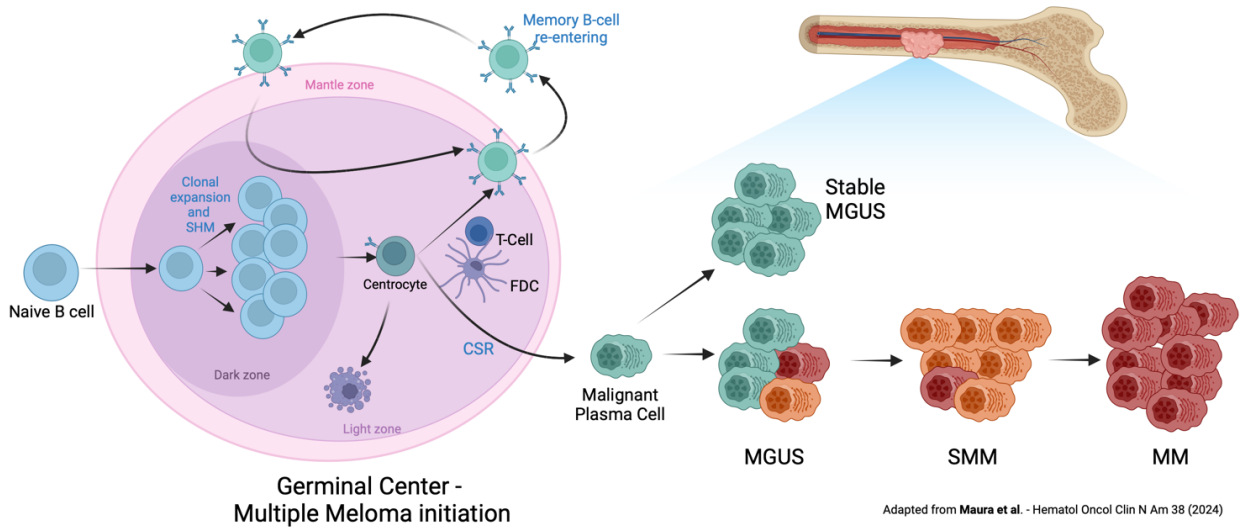

Figure III - Initiation of PC dyscrasia in the GC. *The genetic initiation events (chromosome translocations and hyperdiploidy) that occur in the GC are present in all stages of PC neoplasms. Subsequent genetic events cause the progression from MGUS to MM. Image adapted from Maura et al. – Hematol Onco Clin N Am(2024)*[30]

The enzyme AID is crucial for initiating CSR and SHM, and the study by Omori et al. revealed that PI3K signaling is necessary for the proper activation and function of AID and differentiation of B cells into PCs. This finding underscores the importance of PI3K in maintaining genomic stability during the immune response. Thus, aberrant PI3K signaling can lead to genomic instability (GI) through improper regulation of AID, resulting in IGH locus translocations and mutations that can confer proliferative advantages driving the development of MM[16,31]. Along with hyperdiploidy (HRD), translocations involving the IGH locus are supposed to be the initiating events that appear in the earliest stages of PC dyscrasia[16,17,32], including MGUS. HDR and translocations are nearly mutually exclusive events and can occur alongside disease progression, yet these are often not enough to develop MM; thus, asymptomatic patients can remain stable for many years[13,16,33,34]. Secondary genetic events can accumulate over time, resulting in the onset of symptomatic MM.

HRD refers to the trisomy of one or more odd chromosomes (3, 5, 7, 9, 11, 15, 19, and 21)[35]. The underlying reason for the selective duplication of these odd chromosomes remains unclear[36]. Those trisomies can result from abnormal missegregation of chromosomes during mitosis, though the exact mechanisms underlying these events are still unknown. Patients with HRD appear to have a better prognosis compared to those without HRD[37]. (**Table II**)

IGH translocations involve the locus at 14q32 and various oncogenic genes. Examples include cyclin D genes (CCND1, CCND2, and CCND3) in t(11;14), t(12;14), and t(6;14), respectively; the NSD2/MMSET gene in t(4;14); the MAF gene in t(14;16); and the MAFB gene in t(14;20)[38] (**Table II**).

This complex genetic asset significantly contributes to the tumor's intra-tumor and spatial heterogeneity[39].

Table II - Genetic events underlying the initiation and progression of myeloma plasma cell leukemia[40]

PRIMARY GENETIC EVENTS (% OF TUMOURS)	**SECONDARY GENETIC EVENTS (% OF TUMOURS)**
IGH translocations and genes affected t(4;14): FGFR3 and MMSET (11%) t(6;14): CCND3 (<1%) t(11:14): CCND1 (14%) t(14;16): MAF (3%) t(14;20): MAFB (1.5%) ***Hyperdiploidy (57%)*** Trisomies of chromosomes 3, 5, 7, 9, 11, 15, 19 and 21	***Molecular hallmarks*** Immortalization G1/S abnormality (CDKN2C, RB1 (3%), CCND1 (3%) and CDKN2A) Proliferation (NRAS (21%), KRAS (28%), BRAF (5%) and MYC (1%)) Resistance to apoptosis (PI3K and AKT) NF-κB pathway (TRAF3 (3%), CYLD (3%) and I-κB) Abnormal localization and bone disease (DKK1, FRZB and DNAH5 (8%)) Abnormal plasma cell differentiation (XBP1 (3%), BLIMP1 (also known as PRDM1) (6%) and IRF4 (5%)) Abnormal DNA repair (TP53 (6%), MRE11A (1%) and PARP1) RNA editing (DIS3 (13%), FAM46C (10%) and LRRK2 (5%)) Epigenetic abnormalities (KDM6A (also known as UTX) (10%), MLL (1%), MMSET (8%), HOXA9 and KDM6B) • Abnormal immune surveillance • Abnormal energy metabolism and ADME events **Deletions** 1p: CDKN2C, FAF1 and FAM46C (30%) 6q (33%) • 8p (25%) • 11q: BIRC2 and BIRC3 (7%) 13: RB1 and DIS3 (45%) • 14q: TRAF3 (38%) • 16q: CYLD and WWOX (35%) • 17p: TP53 (8%)
SECONDARY GENETIC EVENTS (% OF TUMOURS)	
Gains • 1q: CKS1B and ANP32E (40%) • 12p: LTBR • 17q: NIK ***Secondary translocations*** t(8;14): MYC Other non-IGH translocations ***Epigenetic events*** Global hypomethylation (MGUS to myeloma) Gene-specific hypermethylation (myeloma to plasma cell leukemia)	

Interestingly, the analysis of IGH translocation breakpoints has revealed that the t(4;14) predominantly involves the switch regions, indicating errors occurring during the CSR process. In contrast, the t(11;14) translocation is likely due to errors during the SHM process[41,42]. Breakpoint

analysis also suggested that approximately 20% of t(11;14) and t(14;16) translocations may occur within the BM during the maturation of IGH rearrangements[40]. (**Figure IV**)

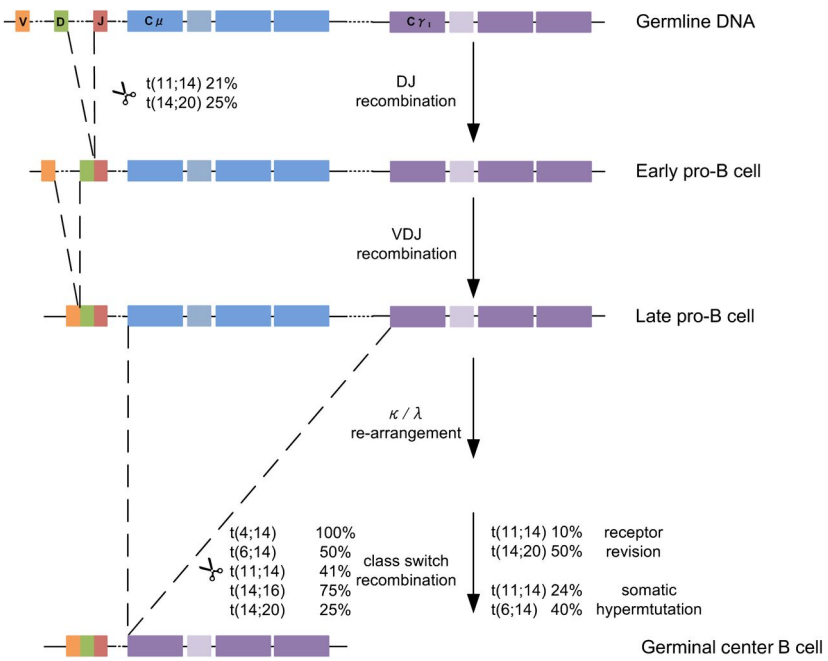

Figure IV - Translocation events in MM. *Translocations involve the IGH locus and all main partner chromosomes through CSR in the germinal center. t(11;14) and t(14;20) subgroups can also have different mechanisms of developing translocations through receptor revision in the germinal center or through DH-JH recombination at the early pro–B-cell stage[42]. Image adapted from Walker, B.A., et al. (2013).*

IGH translocations are slightly less common than HRD, with 40% to 50% of patients presenting IGH translocations[34,43]. The most frequent translocation is t(11;14)(q13;q32), leading to the overexpression of the CCND1 gene, which is more prevalent among MGUS patients[16]. The second most recurrent IGH translocation is t(4;14)(p16;q32). This translocation disrupts and upregulates two genes, FGFR3 and MMSET/WHSC1. Translocations could lead to a fusion of genes or to an upregulation of the involved oncogene, whose expression will inevitably be under the IGH enhancer[44,45].

As illustrated in Table II, other translocations are rarely seen in more than 3% of patients. Additionally, other IGH-independent translocations may play a role in MM pathogenesis. An important example is translocations involving the MYC locus, which are typically considered secondary events. Translocations between 8q24 (MYC) and IGH, IGK/IGL, XBP1, FAM46C, CCND1, and KRAS lead to aberrant MYC overexpression[46]. In this regard, it has been observed by Barwick et al. studies that translocations involving IGL locus were commonly associated with HDR and were present at diagnosis. In particular, 78% of IGL-MYC co-occur with HDR, and patients bearing IGL-MYC translocations experienced a poor prognosis independently of the HDR presence/absence[47].

The third major genomic aberration event in MM is deletions and gains. The most important and frequent are monosomy 13 (45% of patients)[48], loss of the short arm of chromosome 17 (where the tumor-suppressor gene TP53 is localized), duplication/amplification of the long arm of chromosome 1 (1q gains, 30% to 35% of patients), and deletion of chromosome 1p[49]. These last two recurrent copy number changes on chromosome 1 are associated with poorer survival.
The secondary hit can also involve several pathways; indeed, the disruption of DNA damage repair and replication-induced stress pathways are well-established mechanisms contributing to genomic instability (GI). Furthermore, non-coding RNA molecules are significantly dysregulated in myeloma and can influence key components of GI, suggesting additional mechanisms that may play a role in the development of the disease[50].
Epigenetic changes (such as alterations in DNA-hypomethylation and methylation/acetylation of histones) contribute to genomic plasticity and enhance the adaptability and prolonged survival of tumor cells[32].
Thus, it is evident that in myeloma, cytogenetic and molecular abnormalities fundamentally characterize the heterogeneous disease's biology. Detecting these abnormalities is crucial for risk stratification and determining patient prognosis. However, these are not suitable for monitoring the disease throughout its progression nor for determining the presence of post-therapy residual clonal cells. In fact, CNAs and mutations are not consistently present throughout disease progression.

Tumor Heterogeneity and cancer evolution

A further significant aspect of PC dyscrasia is its pronounced intra- and interclonal heterogeneity. The "Big Bang" model, from Sottoriva et al., posits that tumors predominantly grow as a single expansion event, generating numerous intermixed subclones without stringent selection pressures. This early divergence leads to significant intratumoral heterogeneity, observable in various subclonal mutations. Applying this to MM, it is obvious how early genetic alterations and clonal expansions contribute to the complex landscape of tumor heterogeneity observed in this malignancy[51].
In fact, MM exhibits characteristics that place it between hematological malignancies and solid tumors. This is reflected in the distinct mutational signatures found in MM, which share some similarities with both types of cancers. This unique position highlights the complex nature of MM and the diverse mutational processes contributing to its pathogenesis[52].
MM is also characterized by genetic diversity within and between tumor sites. Spatial heterogeneity refers to the variation in genetic and chromosomal aberrations across different regions (*foci*) of the tumor within the same patient. This phenomenon underscores the complexity of the disease, as

distinct clonal populations can evolve independently, leading to significant implications for prognosis and treatment[53]. As a consequence, disease characterization at diagnosis and relapse of the same patient demonstrated that the mutational profiles of clones at relapse can differ markedly from those at diagnosis[13,39,54,55].

Nearly all MM patients are preceded by the MGUS stage, even though MGUS is often not identified clinically. However, most BM PCs in MGUS exhibit malignant sub-clones responsible for disease progression[56]. This highlights the evolving genomic landscape of MGUS, which, despite being clinically indolent, can display significant clonal heterogeneity. This heterogeneity could drive further disease progression upon acquiring subsequent mutations, echoing the "Big Bang" model of cancer evolution[16].

MGUS consists of indolent clonal expansions, in which transformation events occur randomly, infrequently, and in bursts rather than through a slow, constant accumulation of genetic lesions. The progression rate of SMM is not constant: it is 10% per year in the first five years, then declines to 3% for the next five years, and 1% after ten years from diagnosis. This indicates a heterogeneous patient group within SMM, some with incipient MM and others with indolent MGUS-like disease despite a large tumor burden[16,57].

Bolli et al. identified two main patterns of clonal evolution to symptomatic MM: the "static progression model," where the malignant population is present at the SMM stage and grows symmetrically into MM, and the "spontaneous evolution model," where the genomic composition changes at progression due to the loss of subclones and/or acquisition of new ones driving progression. The static model typically progresses in less than a year, while spontaneous evolution takes longer, requiring new genetic lesions[57]. Da Vià et al. suggest that SMM cases can be divided into “MGUS-like” cases, requiring additional genomic events for progression, and “MM-like” cases, which already possess aggressive features and need time to accumulate sufficient burden to meet clinical criteria[16].

From the earliest stages of the disease, patients can develop different PC subclones originating from a common ancestral clone, which may display considerable molecular diversity[58]. This supports the theory that PCs begin to diverge early in their evolution. Under the influence of epigenetic changes or BM microenvironment (BME), these subclones can continue to accumulate spontaneous genetic mutations. This contributes to disease progression, fueled by clonal competition and the survival of the fittest clone[40]. Consequently, myeloma evolution is often not linear but rather follows a branching pattern with increasing genetic complexity[16,57,59]. As a result, a rare subclone present at the onset of

the disease may eventually become the dominant and most aggressive clone, or new subclones may emerge at the expense of others as the disease progresses[60] (**Figure V**).

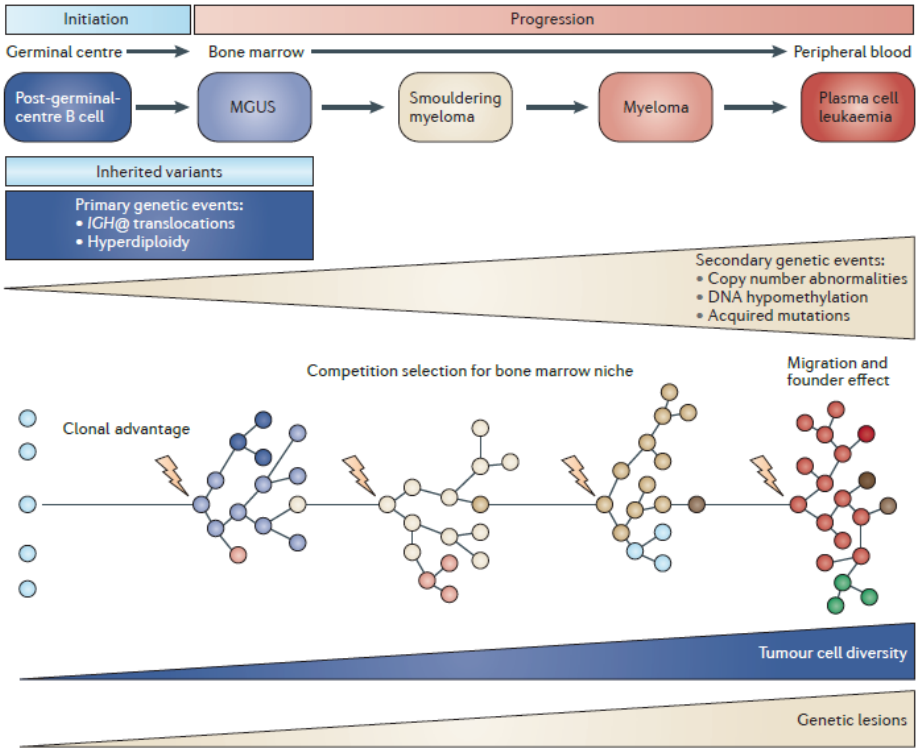

***Figure V - Initiation and progression of MM**. Transitioning through clinical states in myeloma involves acquiring genetic abnormalities that develop the disease's biological hallmarks. Eventually, the proliferative clone escapes the bone marrow, rapidly expands, and leads to patient death. Cells in this stage are significantly genetically altered. This phase can be seen as starting with a migration and founder effect [40]. Image adapted from Morgan et al. (2012).*

The concept of evolutionary trajectories refers to the path these genetic changes take as the disease advances. Each stage of progression is marked by specific genetic alterations that provide a selective advantage to the malignant cells, allowing them to expand and outcompete other clones. For instance, subclones can exhibit specific mutational processes, including the activity of APOBEC enzymes, which are associated with aggressive disease and poor outcomes[57]. This process results in a heterogeneous disease landscape with multiple subclones, each carrying different sets of mutations.

The existence of multiple clones requires highly sensitive techniques to detect and track them as the tumor progresses.

Due to advancements in sequencing methods, known collectively as next-generation sequencing (NGS) technologies, it has been possible to obtain a complete map of the genome and transcriptome with remarkable speed and detail. Years after the advent of these methods, we are moving towards an even more precise understanding of complex biological processes by considering the dynamic expression of the cellular genome across different spatial and temporal contexts. Achieving this next level of insight will require an enhanced level of resolution.

Using NGS, different research groups have shown the heterogeneity of the MM mutational landscape

and the considerable variability among patients at various stages of the disease (diagnosis, relapse)[13,61] and among the same tumor if sampled at different *loci*[53]. It has also been shown that there is a high frequency of nonsynonymous variants in MM and a progressive increase in non-silent mutations from asymptomatic to symptomatic conditions[12]. MM does not have specific driver events common to all patients. Only a few recurrent mutations have been identified in MM, with none occurring in more than 30% of patients (e.g., NRAS and KRAS)[13,62,63]. Additionally, various other mutations have been observed, but each appears in a small percentage of patients. In total, approximately 30 genes have been found to be recurrently mutated in at least 5% of patients. Importantly, many of these mutations are detected only at subclonal level, including genes thought to function as drivers (such as NRAS, KRAS, and BRAF)[63,64]. In MM, the most frequently mutated genes include KRAS, NRAS, TP53, BRAF, TRAF3, FAM46C, and DIS3. These mutations highlight the primary involvement of the tyrosine kinase receptor and NF-kB pathways. Mutations in NRAS, KRAS, and TP53 are well-established markers of advanced disease, and their prevalence was previously noted through conventional sequencing methods. Overall, these mutations are found in at least 50% of MM patients[40] (**Figure VI**).

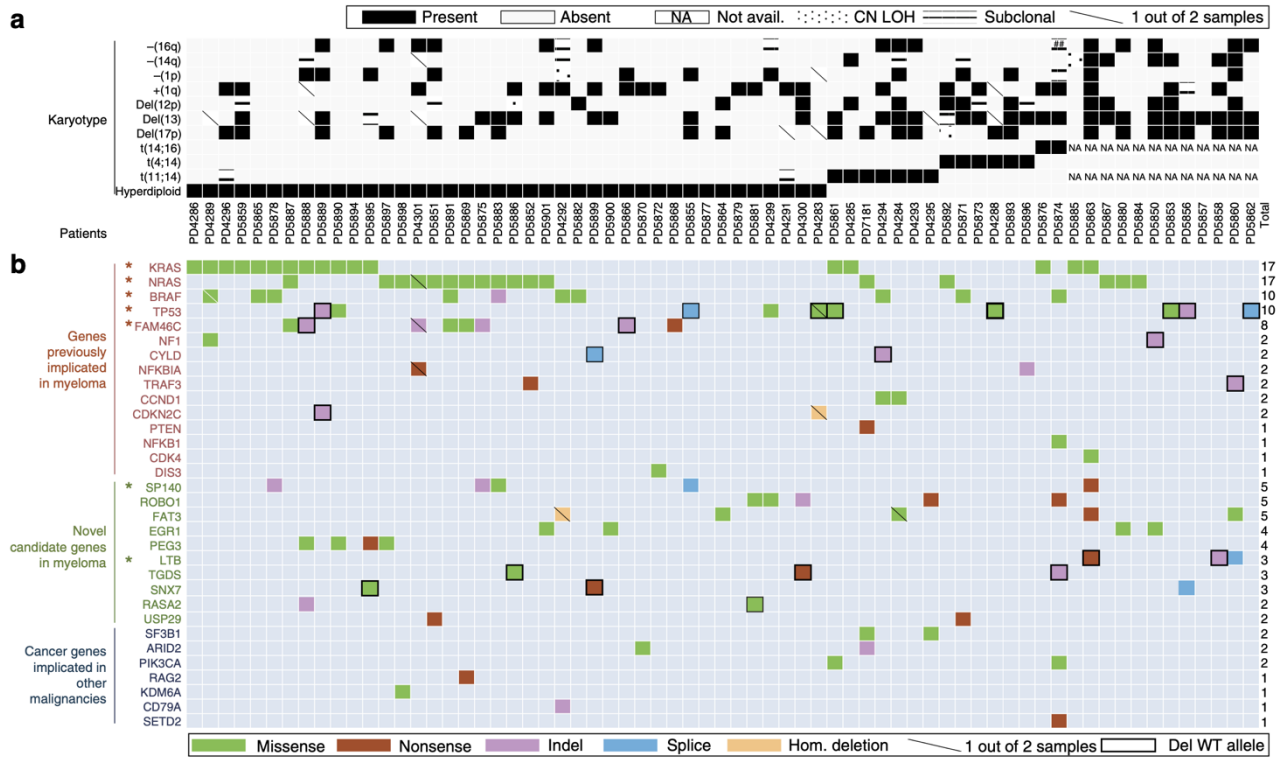

***Figure VI - Genetic alterations and gene mutations in MM**. Oncoplot shows the most relevant genomic lesions and frequently mutated genes. Patients are represented in columns. **a.** Karyotypic features; **b.** Recurrently gene mutation. Missense (green), nonsense (red), and splice-site (blue) substitutions are color-coded; indels are light purple, and homozygous deletions are ochre. For patients with serial samples, unique events are marked with a diagonal bar. Only one mutation per gene is plotted for each patient. If a mutation is linked to the deletion of the wild-type allele, the gene has a black contour. The 'Total' column shows the number of patients with at least one non-silent mutation per gene. Asterisks indicate genes with a significant recurrence rate (P-value < 0.02, false discovery rate < 10%)[64]. Image adapted from Bolli et al. (2014).*

BM Microenvironment

The BM niche is a sophisticated and dynamic microenvironment that plays an essential role in the regulation of hematopoietic stem cells (HSCs). This niche not only supports the maintenance and self-renewal of HSCs but also orchestrates their differentiation into various blood cell lineages, ensuring the continuous replenishment of the body's blood supply[65,66].

The cellular component includes both hematopoietic and non-hematopoietic cells. Hematopoietic cells comprise myeloid cells, T and B lymphocytes, Natural Killer (NK) cells, and osteoclasts. Non-hematopoietic cells consist of Bone Marrow Stromal Cells (BMSCs), mesenchymal stromal cells derived from BM, fibroblasts, osteoblasts, adipocytes, and endothelial cells. The non-cellular component consists of the extracellular matrix (ECM) and various cytokines, chemokines, growth factors, and exosomes that are secreted by the cellular component.

The ECM is made up of various proteins, such as fibronectin, laminin, and collagen, and can be reshaped by matrix metalloproteinases (MMPs)[66,67].

The BM niche provides a supportive and protective environment for HSCs. These niche cells interact closely with HSCs, providing crucial regulatory signals through both direct cell-cell contact and the secretion of soluble factors. These interactions are vital for maintaining the delicate balance between HSC quiescence, where cells remain in a dormant state, and activation, where they are prompted to divide and differentiate. In addition, Treg cells within create an immune-privileged environment, protecting HSCs from excessive inflammation[68].

Long-lived PCs require BM niches primarily for longevity support. After their generation in GC these immune cells migrate to the BM, where they can survive for years, continuously producing antibodies. Experimental models indicate that the BM contains multiple niches for long-lived plasma cells. In these niches, mesenchymal stromal cells (MSCs) and osteoblasts provide IL-6, while myeloid cells, including macrophages, neutrophils, and eosinophils, supply BCMA ligands like BAFF and APRIL. MSCs also play a role in retaining PC in the BM through interactions involving CXCL12–CXCR4 and VCAM1–α4β1 integrin[66].

In myeloma, tumor-associated inflammation enhances the support of malignant cells at the expense of normal PCs.

Myeloma PCs primarily grow and proliferate within the BM niche supporting the tumor development and survival. In advanced stages, however, clonal PCs can detach from the BM niche, leading to plasmacytomas or Plasma Cell Leukemia (PCL), characterized by free circulating tumor cells in the bloodstream[69]. This extravasation capacity from the BM environment (BME) increases the already present spatial heterogeneity of MM, as previously described[53]. The extravasation capacity acquired

by PCs is driven by mutual crosstalk with the BM niche[70,71]. The malignant PCs within the BM niche foster a supportive microenvironment through direct cell-to-cell contact and the secretion of soluble factors, including cytokines, chemokines, growth factors, exosomes, and miRNAs[72]. Therefore, understanding the mechanisms that myeloma cells use to leverage the BM niche for growth, survival, and subsequently colonize other sites is crucial. The typical PCS surface marker, CD138 (also known as syndecan 1), directly binds to ECM proteins like fibronectin and triggers the release of MMPs by BM stromal cells (BMSC), thereby promoting new blood vessel formation and tumor invasion.
Focusing on the non-hematopoietic cellular components, it is now understood that BMSCs directly bind to cPCs cells via the VLA-4 / VCAM interaction, activating pathways such as NF-Kβ[73]. This interaction leads to the secretion of proliferative, anti-apoptotic, and chemotactic cytokines like Interleukin (IL)-6, Stromal Cell-Derived Factor-1 (SDF1), and Insulin-like Growth Factor (IGF)-1, all of which support the adhesion, growth, and proliferation of myeloma cells. Notably, IL-6 enhances the secretion of Vascular Endothelial Growth Factor (VEGF) by cPCs, promoting angiogenesis within the bone marrow[35]. IL-6 also increases reliance on Mcl-1, an anti-apoptotic Bcl-2 molecule crucial for MM survival. In the absence of IL-6, B-cell activating factor (BAFF) and a proliferation-inducing ligand (APRIL) may aid in tumor survival, promoting cell growth and immunosuppression, another key mechanism for MM cell survival[73].
In MM, bone lesions are the result of an imbalance in bone turnover due to suppressed osteoblast activity (responsible for new bone deposition) and increased osteoclast activity (mediators of bone resorption). This imbalance is driven by the activation of several molecules and pathways, such as macrophage inflammatory protein-1-α (MIPα), NOTCH, IL-6, IL-17, and RANKL[73]. Moreover, evidence suggests that OCs protect myeloma cells from T cell responses through the expression of immunosuppressive molecules, including Programmed death-ligand 1 (PD-L1), galectin 9, herpesvirus entry mediator (HVEM), and CD200[74].
Adipocytes, which are fat storage cells abundantly present in the BM, may support malignant PCs growth through the secretion of leptin, adiponectin, and IL-6[35]. Lastly, MM exploits adhesion to the ECM via the activation of integrins (e.g., α4β1, β7, VLA-5, and β5 integrin CD56), enhancing angiogenesis, tumor invasion, and bone resorption. Endothelial cells, stimulated to produce VEGF, platelet-derived growth factor (PDGF), and other molecules, further promote neo-angiogenesis[73] (**Figure VII**).

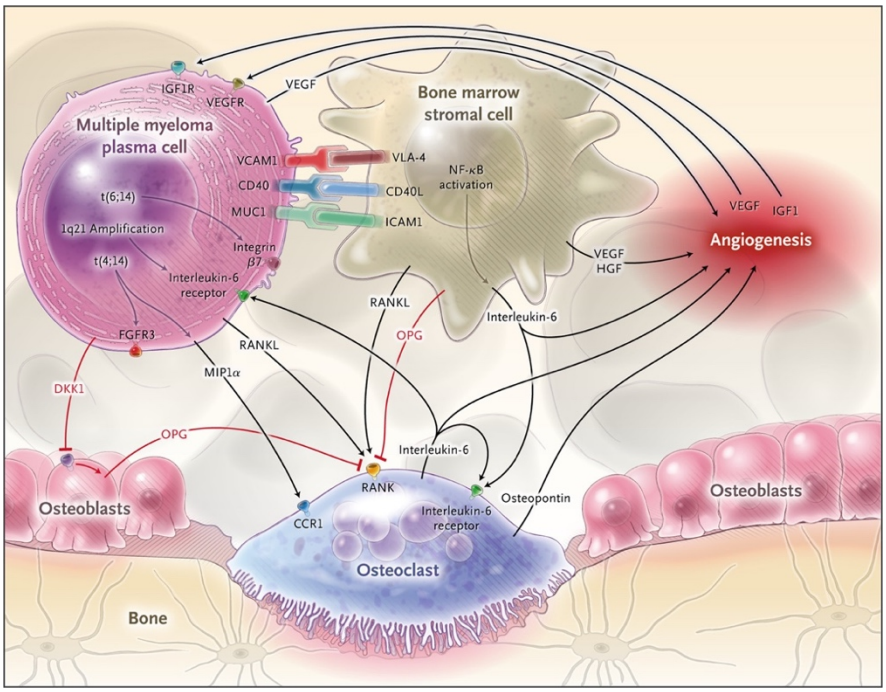

Figure VII - Summary of BME events. *cPCs exploit the BM microenvironment to grow and proliferate uncontrollably[2]. Image adapted from Palumbo et al. (2011).*

Under normal conditions, immune effector cells can be drivers of strong anti-tumor responses by recognizing and eliminating tumor cells based on the expression of "non-self" antigens with high specificity[75]. However, in patients with PC dyscrasias, "immunoparesis" is a common occurrence[76]. Immunoparesis is defined as reduced levels of one or more non-clonal Ig classes (polyclonal Ig). Its prevalence increases from MGUS to SMM and MM and holds prognostic value[76–78]. This condition arises due to changes induced by PCs in the BME[79]. Various immune effector cells are affected by this condition, including DCs, T cells, NK cells, and other cells, and this functional impairment is closely linked to the progression to clinical MM[80].

Minimal Residual Disease

The development of new therapeutics has significantly improved the treatment and response rate of MM. This advancement has created a demand for a deeper understanding of patient responses to these therapies.

Minimal Residual Disease (MRD) refers to the minute number of cancer cells that linger in a patient's body following treatment, which are often undetectable by conventional diagnostic methods. In the context of MM, MRD assessment has become a cornerstone for evaluating the depth of response to therapy and for prognosticating outcomes.

The concept of MRD has evolved significantly over the years, particularly in hematological malignancies. The International Myeloma Working Group (IMWG) has been instrumental in defining

and standardizing MRD criteria. According to the IMWG, MRD negativity is achieved when sensitive assays fail to detect cancer cells in bone marrow samples, indicating a profound response to treatment[81].

The significance of MRD in MM lies in its strong prognostic value. Numerous studies have shown that patients who achieve MRD negativity tend to have longer progression-free survival (PFS) and overall survival (OS) compared to those who remain MRD positive. For instance, Paiva et al. (2009) demonstrated that MRD negativity, as assessed by multiparameter flow cytometry, provided superior prognostic information compared to traditional morphological assessment of BM PCs[82]. In 2018, the European Society for Blood and Marrow Transplant Research (EBMT) introduced the first criteria for MRD assessment in myeloma, but there have been numerous updates since then due to challenges in accurately defining Complete Response (CR)[83]. Thus, the International Myeloma Working Group (IMGW) has set new standards to standardize the evaluation of treatment response. The criteria are fully shown in **Table IV**.

Table IV - IMGW criteria for MRD evaluation.

MRD criteria	**Description**
sCR (stringent complete response)	CR criteria + normal FLC + no cPCs in BM
CR (complete response)	No M-protein in serum/urine with IF, no plasmacytomas, <5% cPCs
VGPR (very good partial response)	M-protein in serum/urine with IF or >90%reduction serum M-protein + urine M-protein <100mg/24h
PR (partial response)	>50% reduction serum M-protein, >90% urine M-protein or <200mg/24h, >50% reduction of plasmocytoma
MR (minimal response)	>25% but <49% reduction serum M-protein, 50-89% reduction urine M-protein, 50% reduction of plasmocytoma
SD (stable disease)	No criteria for sCR, CR, VGPR, PR, MR or PD
PD (progressive disease)	Any >1 of: >25% enhancement of serum or urine M-protein, new lesions, >25% cPCs

Given that most patients relapse even after achieving CR, it became evident that more advanced methods are necessary to detect MRD at the lowest levels in order to identify patients at higher risk of relapse.

Currently, MRD can be assessed using two highly sensitive methods: eight-color Next-Generation Flow (NGF) cytometry and Next-Generation Sequencing (NGS)[84] (**Table V**).

NGS is one of the most sensitive methods for detecting MRD, capable of identifying one cancer cell among a million normal cells. This high sensitivity makes NGS a powerful tool in MRD assessment, highlighting its potential to guide therapeutic decisions and predict long-term outcomes[85]. Flow cytometry remains a standard technique for MRD detection, offering a balance between sensitivity and practicality. It involves analyzing BM samples to detect abnormal PCs. This method has been refined over the years, providing crucial prognostic information that helps tailor treatment strategies.

cPCs cells are identified by the expression of CD38 and CD138, and may also express markers such as CD56, CD28, and CD117. Compared to normal PCs, tumor cells are generally low in CD45 and negative for CD19, CD27, CD81, and exhibit restrictions to either the kappa or lambda immunoglobulin light chains[9]. Traditional flow cytometry assays have been replaced by the advanced NGF technique developed by the Euroflow consortium. This method uses two single eight-color tubes containing all necessary markers for PCs, allowing for the analysis of MM samples within 3-4 hours, processing up to 10^7 cells per sample, achieving a sensitivity of 10^{-5} to 10^{-6}, comparable to NGS assays[86]. Despite its advantages, there are challenges with using flow cytometry for MRD evaluation in MM, primarily due to sample processing and analysis variability between laboratories and the use of novel drugs like Daratumumab, which can mask the detection of PCs by targeting CD38. This issue has been addressed by using multiepitope anti-CD38 antibodies that bind to different sites than those targeted by the drug[87]. From a molecular perspective, clonal VDJ rearrangements are utilized for MRD detection. NGS technology can be used to sequence millions of reads to measure clonality in MRD with a sensitivity of 10^{-6}. Nevertheless, NGS is limited by high costs and the expertise required for the analysis[88].

Table V - Comparison of MRD testing with next-generation flow cytometry and next-generation sequencing. *Adapted from Yee et al. Hematology Am Soc Hematol Educ Program (2021)*

Characteristic	**Next-Generation Flow**	**Next-Generation Sequencing**
Requires baseline sample	No	Yes, to identify clone to track
Sensitivity	$<10^{-5}$	$<10^{-6}$
Sample requirement	$>5*10^6$ cells (more cells to increase sensitivity to 10^{-6})	$< 2*10^6$ cells
Sample processing	Requires fresh sample	Fresh or frozen
Assessment for hemodilution	Yes	No
Turnaround time	2 – 4 days	7 days

The IMWG guidelines for response definition indicate that a cut-off of 10^{-5} should be used to define MRD negativity, detectable by either NGS or NGF, with the choice of assay depending on the specific needs of each laboratory[81]. However, when measuring MM MRD, it is essential to consider the analytical performance of the assays and start with high-quality samples[89]. The quality of samples should be validated by assessing the degree of hemodilution to minimize the risk of peripheral blood cell contamination, which can alter cell population distribution and underestimate MRD burden. Reduced cellularity can also decrease the sensitivity of MRD assays, so maintaining high sample quality is crucial.

Achieving MRD negativity has profound implications for the management of the tumor. It serves as a benchmark for therapeutic efficacy and helps in making informed decisions about treatment intensity and duration. For example, patients who achieve MRD negativity may benefit from less intensive maintenance therapy, reducing the risk of adverse effects while maintaining disease control. However, current clinical approaches do not consider the achieving of MRD negativity, but its sustained negativity over time.

The ability to detect and monitor MRD provides a more sensitive measure of treatment response than traditional methods, potentially accelerating the approval of novel treatments.

Despite MRD detection is a powerful tool for managing MM, it is limited by issues of sensitivity, clonal heterogeneity, and the evolving resistance of residual cells. Even conceptually, there is a need to understand how to translate its measurement into clinical practice. The impact of cytogenetics on MRD's predictive value and the technical challenges involved in its accurate detection underscore the need for ongoing research and refinement of MRD-related strategies to better guide treatment and improve patient outcomes[90].

Single cell RNA sequencing

RNA sequencing (RNAseq) emerged as a groundbreaking and cost-effective approach for examining transcriptomic data in the late 2000s[91]. It overcomes all the problems related to microarrays by eliminating the need for prior sequencing knowledge, enabling direct RNA level measurements, detecting sequence variations, and allowing multiplexing[92]. However, RNAseq quantifies the entire transcriptome from a mixture of cells collected from a sample in one go, limiting the ability to capture cellular heterogeneity. This characteristic has led to it being termed “bulk RNAseq”. In 2009, Tang et al. addressed this limitation by isolating individual cells and performing RNA-seq on each cell, a method now known as single-cell RNA sequencing (scRNAseq)[93]. Initial applications of scRNAseq were constrained by the low number of cells that could be processed per experiment and the high

costs involved. Over the past decade, advancements in cell capture and sequencing efficiency have allowed for the analysis of millions of cells, while significantly reducing sequencing costs, thus facilitating widespread adoption of the technique[94]. scRNAseq captures the transcriptome of individual cells, offering a granular view of cellular states and transitions. This technology has revolutionized various fields of biology and medicine, revealing new dimensions in the study of complex tissues, rare cell populations, and tumor biology[95]. Tumors comprise a diverse array of cells, including cancer cells with varying degrees of differentiation, stromal cells, immune cells, and endothelial cells. Traditional bulk RNA sequencing often masks this heterogeneity, potentially overlooking critical biological signals. In contrast, scRNAseq allows for the dissection of tumors at unprecedented resolution, enabling the identification of distinct cell populations, subclones, and the interactions between tumor cells and their microenvironment[96].

One of the most significant contributions of scRNAseq in cancer research is the identification of cancer stem cells (CSCs) and their role in tumor progression, metastasis, and therapy resistance. By profiling individual cells within tumors, researchers can pinpoint unique transcriptomic signatures of CSCs, offering potential targets for novel therapies[97]. Furthermore, single-cell sequencing has been instrumental in mapping the immune landscape within tumors, revealing the diversity of immune cell types, their states of activation, and their interactions with tumor cells. This knowledge is crucial for developing immunotherapies and personalized medicine approaches[98]. Additionally, scRNAseq has been used to explore the evolutionary trajectories of cancer cells. By analyzing the transcriptomes of individual cells, researchers can trace the clonal evolution of tumors, identify key drivers of progression, and understand the mechanisms underlying therapeutic resistance[99,100].

scRNAseq - Application on Multiple Myeloma

In MM, single-cell transcriptome sequencing has been applied to dissect the clonal architecture of the disease, revealing the presence of subclonal populations that may drive relapse and drug resistance. By profiling individual myeloma cells, researchers identified specific gene expression patterns associated with aggressive disease and poor prognosis, peculiar cellular sub-populations, or findings intended to achieve a comprehensive understanding of the molecular mechanisms underlying MM[56,101,102].

Since the BM niche plays a critical role in supporting myeloma cell survival and mediating resistance to therapy, scRNAseq has been pivotal in understanding the interaction between PCs and the BME. Through single-cell analysis, it is possible to map the complex cellular interactions within the BM, identify key supportive cell types, and explore the signaling pathways that sustain myeloma growth[66].

Furthermore, this technology can shed light on the immune evasion mechanisms employed by clonal PCs, such as the expression of immune checkpoint molecules or BM cellular component shaping[103].

scRNAseq - Focus on Plasma Cells

Ledergor et al. (2018) were the pioneer work of scRNAseq studies on MM, investigating isolated $CD38^{+}CD138^{+}$ BM PCs from patients with PC dyscrasia and healthy donors (HDs). The study found limited heterogeneity among the control samples but significant heterogeneity in the malignant cells across different patients. The researchers identified shared transcriptional mechanisms among clusters of patients, highlighting the overexpression of MM key driver genes (CCND1, CCND2, FGFR3, FRZB).

The study also demonstrated that scRNAseq could accurately infer CNAs, revealing that the transcriptional state of clonal PCs is primarily influenced by trans-acting mechanisms rather than by direct DNA mutations. Additionally, the study uncovered three novel candidate genes: LAMP5, CDR1, and WFDC2. Notably, the investigation extended to residual malignant cells in patients after therapy, which shared the transcriptional profiles of the malignant cells present at diagnosis. Overall, the research offers a detailed insight into the molecular mechanisms of MM and highlights potential biomarkers for the disease's diagnosis and monitoring[56].

Liu et al.(2021) described a longitudinal study of 14 patients with PC dyscrasia coupling scRNAseq and 10x Genomics linked-read whole genome sequencing (10xWGS). By integrating data from PCs and B cells from HDs, authors traced the lineage from pre-B cells to mature B cells and eventually to normal or malignant PCs. This lineage mapping identified numerous genes, including well-known ones like XBP1. These findings suggest that single-cell transcriptome profiling of B cells and PCs could be utilized to trace the origin of MM. In addition, by combining mutation, CNV mapping, and gene expression clustering strategies, they discovered that specific PCs subpopulations remained consistent as the disease progressed from SMM to MM. However, as the condition advanced from MM to relapse, new subpopulations characterized by heightened genomic instability began to emerge[101].

Boiarsky et al.(2022) carried out an extensive study employing scRNAseq on PCs from 26 patients at different stages of MM as well as from 9 HDs. They performed labeling of normal and abnormal PCs in each sample revealed low tumor purity in early disease stages, implying that conclusions drawn from bulk studies of precursor conditions may be influenced by significant contamination from normal PCs. Asymptomatic patients already exhibit transcriptomic alterations similar to those seen in overt MM. Notably, they identified a gene signature present in normal PCs but uniformly lost at all stages of MM progression, with significant downregulation of CD27 compared to normal PCs as

early as the MGUS stage. This suggests that earlier studies might have been confounded by increasing tumor purity as the disease progresses. In addition to CD27, clonal PCs showed lower expression of other mature B cell markers like CD79A and reduced enrichment of immune pathways. These findings extend the characterization of abnormal versus normal PCs to the whole transcriptome and support the idea that the loss of B cell immune functionality is an early step in the development of tumor PCs[102].

scRNAseq – Focus on Tumor Microenvironment

Zavidij et al. (2020), in their study, collected 23 BM samples from patients with PC dyscrasia and 9 BM from HDs. Their goal was to uncover changes in the TME during the progression of MM. Using scRNAseq on TME cells revealed significant enrichment of T cells, $CD16^+$ monocytes, and NK cells, as well as a decrease in pDCs, neutrophils, $CD14^+$ monocytes, and progenitors, starting at the MGUS stage, indicating an early activation of the immune response. Interestingly, the study found that in patients with high NK-cell infiltration, these cells predominantly expressed CXCR4, whereas samples with fewer NK cells showed a shift toward CX3CR1 expression. This difference may indicate that MM might trigger NK cell egress into the circulation in some subsets, potentially as a mechanism of immune evasion. Analysis on T-cells revealed a notable population shift from a $GZMK^+$ state, which is associated with a less differentiated phenotype, to a more effector-like phenotype characterized by a $GZMB^+$ signature. Additionally, the authors observed an enrichment of an IFNα signature across various cell types as the disease progressed. Interestingly, while transcriptomic analysis indicated increased HLA-DR gene expression in $CD14^+$ monocytes, Cytometry by time of flight (CyTOF)[104] analysis contradicted this by showing a decrease in MHC class II presentation on the cell surface. Further cell culture experiments demonstrated that the upregulation of MARCH1/MARCHF1 expression was responsible for the internalization of MHC class II in $CD14^+$ monocytes during the later stages of the disease. The findings suggest that similar immune mechanisms may be present in the precursors of other cancers, and understanding the evolution of the myeloma microenvironment could lead to new therapeutic targets. By leveraging inter-patient heterogeneity, the study also aims to stratify patients by their risk of progression, creating opportunities for early intervention[103].
De Jong et al. (2021) conducted a study investigating the BM niche in both HDs and MM patients. They created a paired scRNAseq dataset of non-hematopoietic cells, $CD38^+$ cells, $CD38^-$ cells, and $CD138^+CD38^+$ populations. This analysis led to the identification of a MM-specific subpopulation: inflammatory mesenchymal stromal cells (iMSCs). The presence of these iMSCs was further validated using flow cytometry and in-situ hybridization, where they were found in close proximity

to malignant PCs and T-cells, suggesting potential interactions between these cell populations. Moreover, they discovered that the inflammatory signature in iMSCs could be induced by interactions with certain T and NK cell subsets, specifically $NCR1^{+}GZMB^{+}CX3CR1^{+}CD56^{dim}$ NK cells, $GZMB^{+}$ $CD8B^{+}$ effector T-cells, and $TNF^{+}FAS^{+}GZMK^{+}$ Tscm cells (T stem cell memory), all of which express TNF.

Further exploration of cell-cell interactions using CellPhoneDB revealed that interactions between iMSCs and clonal PCs, mediated through CCL2-CCR/CCR10 and IL6-IL6R pathways, could potentially promote tumor proliferation and modulate myeloid cell functionality[66]. Pilcher et al. (2023) leveraged scRNAseq to analyze 18 MM patients and their TME. The authors categorized patients into two groups based on progression-free survival (PFS): rapid-progressors (RP, PFS < 18 months) and non-progressors (NP, PFS > 4 years). By comparing immune subtypes between these groups, the researchers identified higher levels of $CD8^{+}$ exhausted T-cells and M2 macrophages, along with lower levels of $CD8^{+}$ effector T-cells and immature B-cells in the RP group. Additionally, they found an enrichment of IFN response in the RP group and TNF response in the NP group. Using CellChat for cell-cell interaction analysis, the study identified three signaling pathways—BAFF, CCL, and IL16—that were significantly associated with patient survival. Notably, CCL5 and CCL3 interactions with CCR1 from $CD8^{+}$ T-cells to monocytic cells were identified as negative effectors of patient survival in the RP group. The study also observed an increase in the abundance of BAFF within the TME of the RP group, which could potentially enhance the survival of clonal PCs and adversely affect patient outcomes. Conversely, IL16 expression within the TME was associated with improved patient survival outcomes[105].

In conclusion, this section provides a comprehensive overview of how single-cell technology has become widely adopted in the biological fields for the comprehensive study of all key players in pathological contexts. Specifically, there have been presented major studies that have contributed to the understanding of the complex mechanisms underlying the pathogenesis, progression and clinical outcomes of MM. Particular attention is given to the tumor cell: the plasma cell, and the immune counterpart or TME.

This dissertation focuses on the application of single-cell technology in myeloma, particularly on the novel insights it can provide beyond what is already established in the literature.

While firmly grounded in existing research, the thesis will explore gaps in current knowledge that could have significant clinical implications and contribute to advancing the understanding of the disease within the entire scientific community.

Aims

The clonal origin of MM from B cells ensures that all malignant PCs share identical immunoglobulin heavy and light chain variable regions. This unique feature makes it almost impossible for independent B-cell clones to generate identical CDR3 sequences. These tumor-specific sequences, which remain stable despite the dynamic nature of clonal evolution, serve as highly reliable biomarkers for detecting clonality, monitoring disease progression and MRD tracking[106,107].

Consequently, there is a growing interest in leveraging single-cell technologies to deeply characterize both tumor and immune cells in MM, whether in symptomatic or asymptomatic patients[56,101,103,108]. Given the high expression of tumor-specific VDJ sequences in bulk RNA from MM patients[109], the first aim of this dissertation is to evaluate the effectiveness of scVDJseq in pinpointing MM-specific CDR3 sequences at the single-cell level. This will involve a methodological comparison with clinical-grade VDJ diagnostics, adhering to the EuroClonality-NGS Working Group's standard operating procedures applied to DNA from purified BM PCs[110].

By integrating scRNAseq with scVDJseq, we can dissect cellular heterogeneity and clonal architecture throughout disease progression[56,111]. Given the complex geno-transcriptomic landscape of MM and its TME, the second chapter of this thesis aims to provide a roadmap for comprehensive molecular analysis of scRNAseq data, encompassing both malignant and genotypically normal (or polyclonal) PCs.

B-cell receptor (BCR) sequencing offers the potential to genetically distinguish between cPCs and pPCs. Preliminary studies have identified distinct transcriptomic signatures between these populations[102,112], but a purely transcriptomic approach falls short of accurately identifying pPCs. The multiomic approach of scVDJ and scRNA sequencing provides a powerful tool for unequivocally identifying pPCs at the single-cell genotypic level, enabling detailed computational studies of pPCs in both asymptomatic and symptomatic stages, and facilitating comparisons with healthy PCs (hPCs) from healthy donors (HD).

The TME plays a fundamental role in the pathogenesis of MM, where interactions between cPCs and pPCs and the TME drive disease progression[113,114]through pro-survival and anti-apoptotic signaling[66,115–117]. BM niche includes pPCs that often have been neglected in previous scRNAseq studies on MM, although they contribute to immune surveillance within the TME. Immune surveillance may prevent the progression of asymptomatic cPCs expansion[75,101,114,118]. Immunoparesis could be one of the mechanisms by which the TME evades immunosurveillance in myeloma, likely caused by the disruption of pPCs. To date, the complex interplay between tumoral

cells, pPCs, the TME, and immunoparesis remains poorly understood, yet it holds the potential to reveal critical insights with clinical implications.

The second goal of this dissertation is to utilize single-cell technologies to explore the behavior of genotypically defined pPCs across different disease stages, particularly their interactions with the TME delving into the context of immunoparesis. Furthermore, by analyzing hPCs through scRNAseq, we aim to assess the "functional fitness" of pPCs in patients with plasma cell dyscrasias.

Another crucial objective of this dissertation is to examine the clonal hierarchy and the evolutionary process of cPCs across disease progression through clonotypic VDJ rearrangements at a single-cell resolution.

In conclusion, the primary goals of this dissertation are summarized as follows:

1. **Evaluating the effectiveness of scVDJseq in identifying MM-specific VDJ rearrangements**, compared to DNA-based gold-standard diagnostic methods [Article I]
2. **Conducting a geno-transcriptomic analysis of scRNAseq data from cPCs, pPCs, and the TME**, including:
    a. Dissecting the clonotypic architecture of cPCs and pPCs using scVDJseq [Article II and Article III].
    b. Inferring genomic properties from single-cell transcriptomics [Article II].
    c. Unveiling phenotypic differences between hPCs and pPCs [Article III].
    d. Investigating the crosstalk between TME and cPCs/pPCs [Article III].
    e. Deriving single-cell signatures from hPCs [Article III].
3. **Track the clonal evolution across disease progression** [Manuscript in preparation]**.**

Materials and Methods

Sample collection

Table VI – Distribution of samples, patients and HDs used for analysis in the published articles

Article	Nº samples	Nº patients	Median Age Pts (range)	Nº HDs (Milan)	Nº HDs (External)	Median Age HDs (range)
I	19	18	63 y.o. (41 to 83)	-	-	-
II	5	5	83 y.o. (49 to 87)	-	-	-
III	46	41	62 y.o., (41 to 84)	1	17	52 y.o. (25 to 79)

This dissertation involved a cohort of 46 patients diagnosed with plasma cell (PC) dyscrasia who were admitted to the Hematology Department at Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico in Milan, Italy, between 2021 and 2023. The distribution of patients and samples in the published articles is described in *Table VI*. Samples used for *Article I* have been included also in *Article III*. Samples used for *Article II* were not shared in other publications. The patients were classified with MM, SMM, or MGUS based on the International Myeloma Working Group's (IMWG) revised criteria[8]. All study protocols adhered to the ethical principles outlined in the Declaration of Helsinki and received approval from the Institutional Review Board (Ethical Committee #553_2022). Informed written consent was obtained from all participants, and patient identities were anonymized before sample processing.
The median age of the patients at the time of sampling is reported in *Table VI*.

External data collection

In addition to the primary samples, the study of *Article III* incorporated data from 17 healthy donors (HDs). Of these, 3 HD samples were obtained and processed through collaboration with the Erasmus MC Cancer Institute in Rotterdam, The Netherlands. The remaining 14 HD samples were retrieved from publicly accessible databases. Specifically, 9 HD samples were sourced from the study by Boiarsky et al.[102], with data extracted from count-derived subsets (NCBI Gene Expression Omnibus (GEO) database accession number #GSE193531). Additionally, data from 5 HD samples described by Duan et al.[119] were used as controls, focusing exclusively on counts derived from sorted cell populations B ($CD19^+/CD138^+/CD38^+$) and D ($CD19^-/CD38^+/CD138^+$) (GEO accession number #GSE230705). The median age of the HDs at the time of sampling was 52 years, ranging from 25 to 79 years.

Sample processing

BM aspirates were first filtered using a 100 µm strainer to remove debris. Bone Marrow Mononuclear cells (BMMCs) were then isolated through density gradient centrifugation using Ficoll-Paque medium (Cytiva). Any remaining red blood cells were lysed using BulkLysis (Cytognos, Spain) for 15 minutes at 20°C, followed by a wash with PBS. From BMMCs, PCs were subsequently isolated via an immune-magnetic approach using the EasySep™ CD138 Positive Selection Kit II (STEMCELL Technologies). The number and viability of the positive ($CD138^+$) and negative (CD138-) fractions were assessed with the Trypan Blue exclusion assay, ensuring that only samples with viability greater than 80% were utilized for further applications.
2 $CD138^+$ samples (*Article II*) and 8 $CD138^-$ samples (*Article III*) were cryopreserved in liquid nitrogen, using Bambanker (NIPPON Genetics Europe) as the freezing solution. The remaining $CD138^+$ and $CD138^-$ fractions were freshly processed after isolation.

Thawing Cells

Cells were thawed according to standard protocols to ensure optimal viability. 10 samples stored in liquid nitrogen were quickly transferred to a 37°C water bath until fully thawed. The thawed contents were then immediately transferred into a Falcon tube containing pre-warmed RPMI medium (Gibco, Thermo Fisher, USA) supplemented with 10% FBS (Gibco, Thermo Fisher, USA), followed by a washing step to remove any residual cryoprotectant. The cell count and viability were assessed using the Trypan Blue exclusion assay.

DNA-Based Molecular Analysis of IGHV Rearrangement

Genomic DNA was extracted from 19 $CD138^+$ PC samples (*Article I*) using the AllPrep DNA/RNA/miRNA isolation kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Clonal IGHV rearrangements were identified by following the EuroClonality-NGS Working Group's standard operating procedure for two-step immunoglobulin NGS-based marker identification [Version 1.0 (11 June 2019)], utilizing the IGH V-J primer set. The resulting libraries were pooled and sequenced using 250 bp paired-end reads on a MiSeq instrument (Illumina), achieving an average of 31.96 Mbp per sample. The generated FASTQ files were uploaded to the Vidjil[120] web application for the analysis of high-throughput sequencing reads at the IGH locus.

5' scRNAseq and scVDJseq: library construction and sequencing

A total of 83 samples (52 $CD138^+$, 31 $CD138^-$) were processed for 5' scRNAseq and scVDJseq. Isolated cells from each sample were washed three times with 1X Phosphate-Buffered Saline (PBS, calcium and magnesium-free) containing 0.04% weight/volume FBS (Gibco). Cell counts were

manually determined using the Trypan Blue exclusion assay with a Bürker chamber. To achieve a target recovery of 9000 cells per sample, the volume of cells to load was calculated using the "Cell Suspension Volume Calculator Table" provided in the Chromium Next GEM Single Cell 5' v2 protocol (CG000330 Rev E).
Single-cell 5′ RNA-seq libraries were then generated using the Chromium Next GEM Chip K along with the Master Mix from the Chromium Next GEM Single Cell 5' Reagent Kits v2 (Dual Index) and Single Cell VDJ 5' Gel Beads, following the standard manufacturer protocols. In brief, cell suspensions were loaded onto the Single Cell K chip, and gel beads in emulsion (GEM) were created using the Chromium Controller.
After barcoded full-length cDNA amplification, the cDNA was purified using the SPRIselect Reagent and quantified with the PicoGreen™ dsDNA assay kit. Quality assessment was performed using High Sensitivity D5000 ScreenTape (Agilent) and High Sensitivity D5000 Reagents on the Agilent 4150 TapeStation. Illumina R2 sequences, along with P5, P7, i5, and i7 sample indexes, were added following standard protocol procedures. Simultaneously, the BCR Amplification Kit was used for V(D)J analysis. The quality of the libraries was assessed with High Sensitivity D5000 ScreenTape, and library concentrations were determined using PicoGreen™.
Finally, the generated libraries were pooled according to Illumina's specifications and sequenced on a NovaSeq 6000 platform (Illumina) using a 150 bp Paired-End protocol. Sequencing was targeted at approximately 100000 reads per cell for 5' gene expression libraries and 15000 reads per cell for V(D)J-enriched libraries, following the manufacturer's recommendations.

scRNAseq and scVDJseq: Bioinformatic pipelines

The bioinformatics pipeline for scRNAseq is outlined schematically below and is divided into three key phases: pre-processing, analysis with Seurat, and the application of specific tools for downstream analysis (**Figure VIII**). The entire process was carried out using a combination of programming languages, including Bash, R, and Python. High-performance computing (HPC) infrastructure was essential for handling computationally intensive tasks.
The pre-processing phase primarily involves the alignment of reads using the Cell Ranger software, which generates matrices for use in the subsequent Seurat analysis. After filtering and cleaning the single-cell data, the final phase involves applying a range of bioinformatics tools tailored to specific biological questions, such as cytogenetics, VDJ repertoire analysis, cell-cell interaction studies, pseudotime analysis, and more.

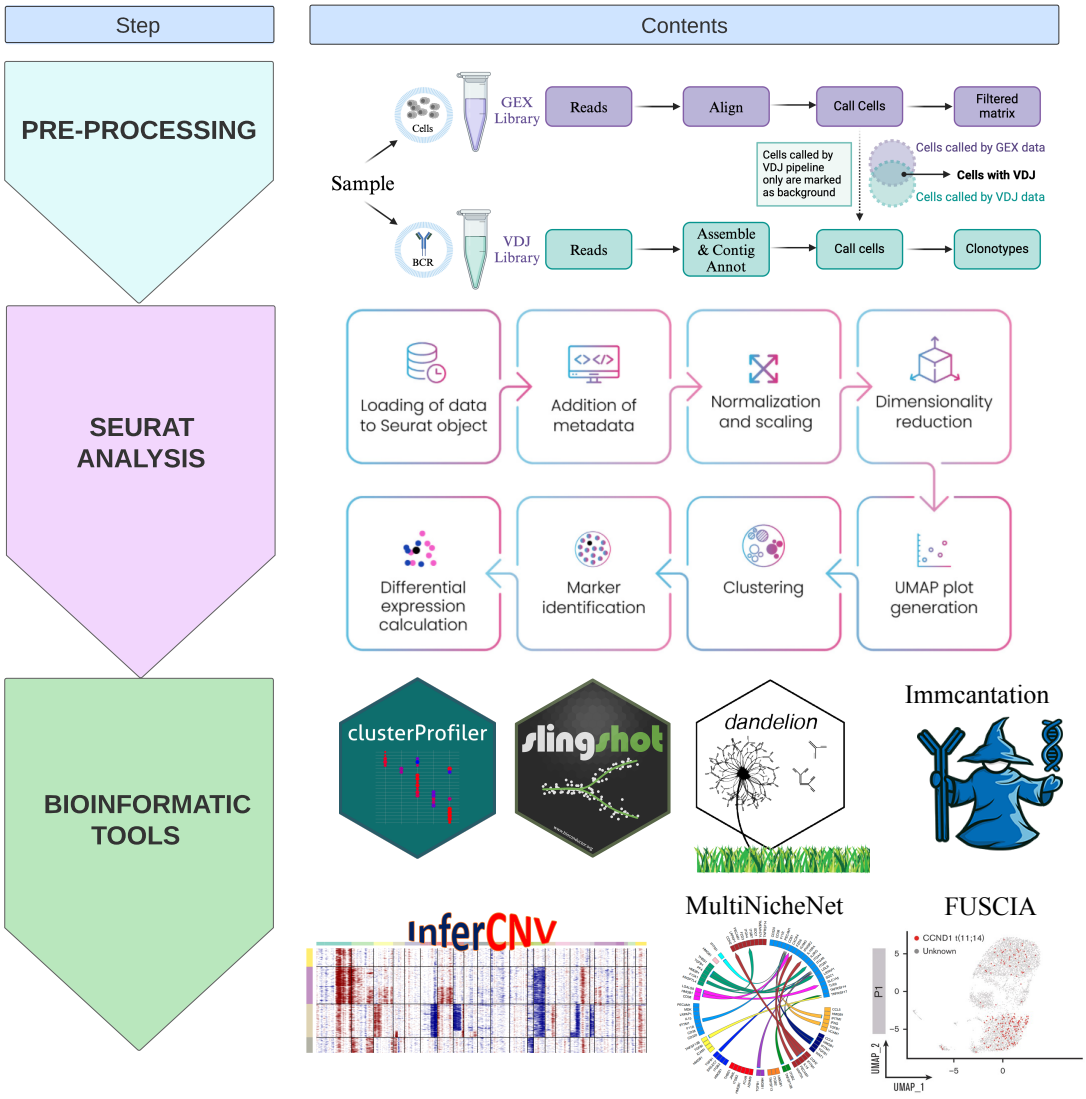

Figure VIII - Schematic representation of the bioinformatics pipeline divided into three main steps: *pre-processing with Cell Rager, Seurat analysis, and the use of bioinformatics tools such as clusterProfiler for GSEA, Slingshot for pseutotime analysis, Dandelion and Immcantation for VDJ analysis, InferCNV for CNA analysis, MultiNicheNet for cell-cell interactions and Fuscia for chimeric transcript detection.*

scRNAseq and scVDJseq: Pre-processing

The pre-processing phase begins with the FASTQ files generated from Illumina sequencing, which contain reads of the amplified cDNAs. These FASTQ files, both for gene expression (GEX) and VDJ data, are used as inputs for Cell Ranger, a software suite from 10x Genomics specifically designed to process data from the Chromium Single Cell Gene Expression platform (**Figure IX)**.

Cell Ranger offers various functionalities, each comprising multiple steps. For example, the "*cellranger count*" function is dedicated to single-cell GEX analysis, encompassing the entire workflow from preprocessing and alignment to quantification and quality control. Meanwhile, the "*cellranger vdj*" function is tailored for VDJ recombination analysis in immune cells, managing tasks

such as alignment, contig assembly, annotation, clonotype identification, and diversity analysis. Both of these pipelines are essential components of the Cell Ranger suite, and their main workflows are outlined in the diagram below. For this study, the Cell Ranger (v 5.0.0) pipeline (https://www.10xgenomics.com/) was employed, using the human genome reference GRCh38 (refdata-gex-GRCh38-2020-A and refdata-cellranger-vdj-GRCh38-alts-ensembl-5.0.0) for single-cell RNAseq and VDJ deconvolution, respectively. The "*counts*" and "*vdj*" functions were specifically used for these processes.

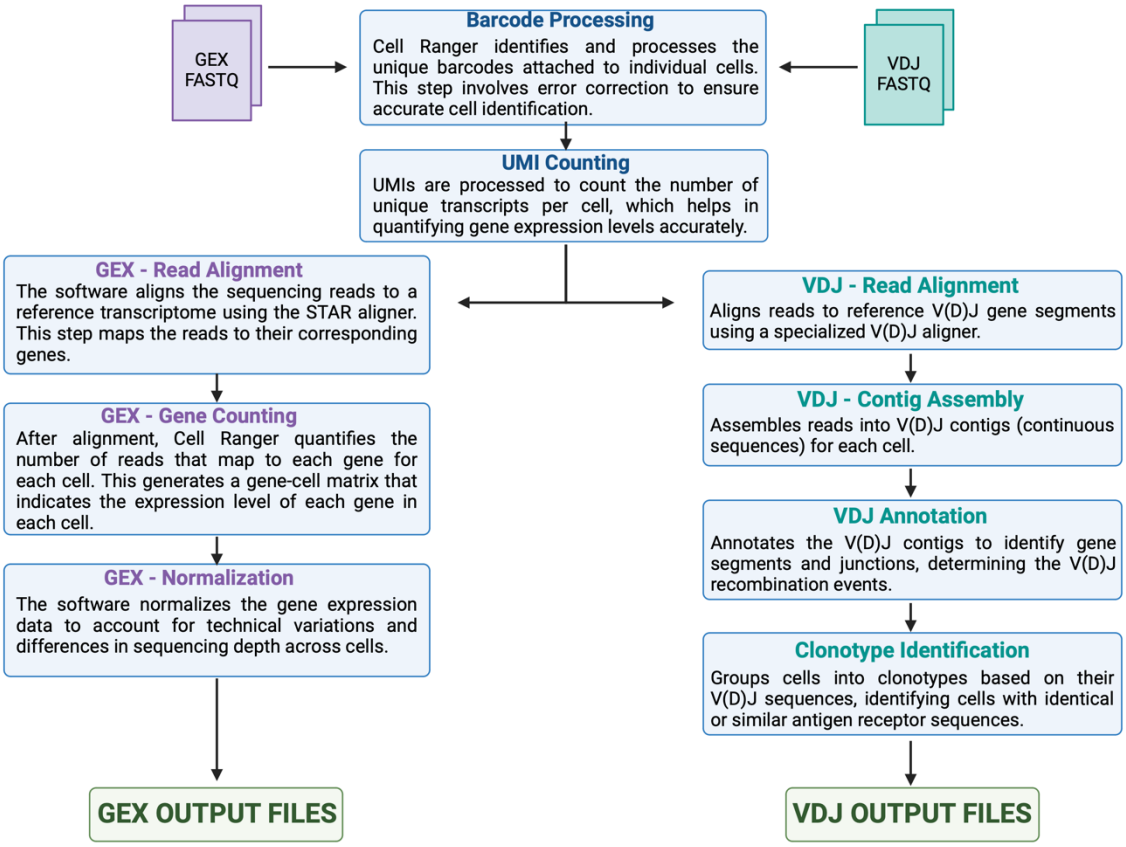

Figure IX *– Schematic representation and explanation of Cell Ranger pipeline for GEX and VDJ data.*

scRNAseq and scVDJseq: Seurat Analysis

A critical step in the scRNAseq pipeline is obtaining the count matrix, which serves as the foundation for exploring the underlying biology using tools like Seurat[121] and Scater[122] in R, as well as Scanpy[123] in Python.

While the overall analysis steps are similar across platforms, certain bioinformatics tools are optimized for specific programming languages, with Python offering advantages in parallelization and memory management over R in some cases. Seurat is an R package designed for the quality control, analysis, and exploration of scRNA data. The pipeline includes several essential steps that

are crucial for obtaining reliable results, with filtering and quality control (QC) being among the most critical. Cells to be filtered out include empty droplets, doublets, dead/dying cells, cells with a high proportion of unmapped reads, and cells expressing a low number of genes[124,125]. The careful selection of filtering parameters is essential, tailored to the experimental conditions[126]. For example, the percentage of mitochondrial genes is commonly used to identify compromised cells[127]. Setting appropriate thresholds is particularly important when working with freshly processed or frozen-thawed cells to avoid excluding valuable data.
Additional indicators of poor quality include low amounts of captured genetic material, few detected genes, high read counts, and low UMI counts. Significant variations among samples may necessitate repeating the experiment or excluding affected samples after thorough evaluation.
Doublet detection methods such as DoubletFinder[128] can be employed to identify potential doublets. DoubletFinder simulates doublets and identifies cells with a high number of artificial doublet neighbors as doublets. After normalization and scaling, datasets can be integrated using the "anchor-based" approach developed by the Satija lab, which matches shared cell types or states across datasets. This integrative analysis increases statistical power and accuracy in comparative studies. Different types of integration are available, such as canonical correlation analysis (CCA) and reciprocal PCA (RPCA). CCA is typically used when cell types are conserved but gene expression varies significantly across samples, while RPCA offers a more conservative approach. While integration can minimize batch effects, Harmony provides enhanced batch correction across samples[129].
Dimensionality reduction is essential for scRNAseq data, with Principal Component Analysis (PCA) and Uniform Manifold Approximation and Projection (UMAP)[130] being widely used for data visualization and clustering.
Seurat also offers tools for defining cell types and states using reference databases, and additional information, such as tag and VDJ data can be incorporated into the Seurat object.

In this study, the Seurat pipeline (v4.4.0) was used to analyze expression matrices. Low-quality cells were filtered out if they had mitochondrial gene content greater than 5% for freshly processed samples or greater than 10% for frozen-thawed samples, or if the number of detected genes was ≤ 200 or ≥ 3,000. To prevent confounding factors in cell clustering, immunoglobulin-related genes were excluded from downstream analyses. Doublets were filtered out using DoubletFinder[128] (v 2.0.4) with the core statistical parameters (nExp, pN and pK) determined automatically using recommended settings for each sample. Automated cell assignment was performed on filtered single cells, which were then projected onto a comprehensive Bone Marrow reference map (v4.4.0, https://satijalab.org/seurat/articles/multimodal_reference_mapping.html)[131]. Clonotype

information was incorporated for each cell by matching the barcodes from scVDJseq with those from scRNAseq, allowing for the classification of cells as harboring clonal or polyclonal clonotypes. Samples were grouped by clinical status (HDs, MGUS, SMM, and MM, and counts were normalized. Integration was performed using the *FindTransferAnchors* and *IntegrateData* functions. The integrated Seurat object was scaled, and the most variable features were identified using the *FindVariableFeatures* function. Principal components analysis was then conducted using the *runPCA* function.

Batch effect correction was applied to the principal components using Harmony[129] (v.1.2.0) with the *RunHarmony* function, using the sample origin as a correction variable. The harmonized matrix was subsequently used to plot UMAPs. For the polyclonal BM scRNAseq dataset, 30 principal components were used for dimensional reduction and cell clustering. Differential expression analysis across clinical stages was performed using the *FindMarkers* function from Seurat (v4.4.0). Cluster-specific markers were identified using the *FindAllMarkers* and *FindMarkers* functions. Violin plots comparing expression across cell types or clinical groups were generated using the *VlnPlot* function in R. All analyses used the Wilcoxon test with a p-value cutoff of 0.1, and figures were created using ggplot2 (v3.4.4) in R.

Polyclonal Plasma cell determinations

For *Article III*, to accurately identify and analyze a population of genotypically distinct pPCs, a rigorous multistep process was implemented:

1. **Dominant Clonotype Identification**: For each patient, the dominant clonotype was determined by identifying the recurrent VDJ rearrangement within the clonal population.
2. **Sequence Comparison**: The nucleotide sequence of the CDR3 region of each non-clonal clonotype (including both heavy and light chain rearrangements) was compared against the nucleotide sequence of the dominant VDJ rearrangement using BLAST analysis.
3. **Manual Inspection**: Cells showing a certain degree of homology in the BLAST analysis were manually reviewed. Cells with dubious homology (greater than 90% for both IGH and IGK/IGL chains) with the dominant clonotype were excluded. Additionally, cells sharing only the light chain rearrangement with the dominant clone were categorized as “light chain escape-derived clones” and were grouped within the clonal PCs.
4. **Phylogenetic Analysis**: Samples exhibiting a higher degree of homology between cell rearrangements underwent phylogenetic tree analysis using the Immcantation pipeline[132,133] to exclude any phylogenetic relationships between the dominant clonotype and other potential polyclonal PCs.

5. **inferCNV Analysis**: Each sample was subjected to InferCNV (https://github.com/broadinstitute/inferCNV, (see detailed methods below)[97] analysis to distinguish different copy number variant (CNV) backgrounds between clonal and polyclonal PC populations at the transcriptional level. The same healthy donor (HD) PCs were used as a normal control for all analyses.
6. **Final Classification**: A final list of cell barcodes for polyclonal and clonal PCs was generated for each sample and used for downstream analyses.

The polyclonal status of the HDs collected at the Hematology Department of Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico was determined using scVDJseq, specifically by confirming the absence of PC clones with a frequency higher than 5% of the total PC population.

In-silico MRD Generation

In *Article I*, to evaluate the performance of the immunoinformatics analysis, an in-silico dilution was conducted. This involved diluting reads from 5 barcodes associated with the dominant clonotype of a patient's sample into a pool of reads from 3,180 barcodes across 13 different samples.

scRNAseq and scVDJseq: Bioinformatic Tools

Identification of IGH Translocations

Single-cell data facilitate the revelation of cytogenetic information, which is particularly important in MM. To identify IGH translocations, the scRNAseq.bam files were used as input for FUSCIA[134], along with the start-end coordinates of the genes of interest. The analysis was conducted for *Article II* using the "discover_discordant_reads.py" function The focus was on investigating the primary MM translocations involving the IGH locus on chromosome 14 in MM, specifically t(4;14), t(6;14), t(8;14), t(11;14), t(14;16), and t(14;20). The genomic coordinates for these translocations are provided in **Table VII**. Bar plots were created using the Prism 5.0 software (GraphPad Prism).

***Table VII** – Genomic coordinates of genes used for FUSCIA analysis.*

GENE	CHROMOSOME	START	END
NSD2	4	1871393	1982207
FGFR3	4	1793293	1808872
CCND3	6	41934934	42050357
MYC	8	127735434	127742951
CCND1	11	69641156	69654474
IGH	14	105586437	106879844
MAF	16	79202624	79601952
MAFB	20	40685848	40689236

Single-Cell Copy Number Variation Analysis

CNAs are among the main features of MM. In scRNAseq, CNA can be inferred data using the inferCNV[97] R package (v 1.16.0, https://github.com/broadinstitute/inferCNV). This tool calculates gene expression intensities across genomic positions in malignant cells, comparing them to a reference set of cells. In *Article II* there were used contaminant TME cells found in the positive fractions of the sample used as reference cells. In *Article III* there were used cells from HD collected in Milan. The algorithm was executed with the following parameters: cutoff=0.1, HMM_type='i3', cluster_by_groups=TRUE, denoise=TRUE, and HMM=TRUE. To identify CNV-level subpopulations, the analysis_mode='subclusters' argument was used, allowing the partitioning of cells into groups with consistent CNV patterns.

Inference Analysis

Pseudotime analysis is fundamental for dissecting complex cellular development processes and understanding disease mechanisms by providing a temporal dimension to static single-cell data. Slingshot[135] emerges as a robust and versatile tool for detecting multiple developmental trajectories, offering deeper insights into cellular development and differentiation. In *Article II*, to enhance the quality of the data for trajectory analysis, a filtering step was applied to reduce dimensionality and eliminate uninformative genes. Only genes that were robustly expressed (at least 4 reads) in a minimum of 10 cells were retained as potential cell-type marker genes. Quantile normalization was then used to remove unwanted technical or biological artifacts. Dimensionality reduction was carried out using principal components analysis (PCA). Following the identification of clusters, the global structure of the underlying lineages was determined. Finally, the Slingshot[135] function was applied to infer the trajectory.

Single-Cell Differential Expression Analysis

Differentially expressed gene analysis and Gene Set Enrichment Analysis (GSEA) enable the dissection of changes in cellular processes by assessing collective alterations in gene groups, known as gene sets. A significant concentration of genes associated with a particular biological process indicates a strong correlation with the observed phenotype.
DEGs among various sample groups were identified using the *FindMarkers* function in Seurat, utilizing both DESeq2 and Wilcoxon tests. Gene set enrichment analysis (GSEA) on DEGs was conducted with the clusterProfiler[111] package (v4.10.1) using the *gseGO* function with the 'org.Hs.eg.db' database (v3.18.0) and the GSEA function with Hallmark gene sets from the Molecular Signature Database (MSigDB) (msigdbr, v7.5.1). Benjamini-Hochberg multiple test correction was

applied, with a p-value cutoff of 0.1 for all analyses. Figures were generated using the *DotPlot* function from the ggplot2 package (v3.5.1) in R.

Cell-Cell Communication and Interaction Analysis

Within the TME, complex cell-cell interactions shape intricate communication networks that significantly influence cellular states and functions. Although scRNAseq data do not directly provide cell proximity information, it is now possible to infer potential interactions between different cell populations using transcriptomic data. Multinichenet[136] , an R package relying on extensive cell-cell interaction databases, facilitates predictive analyses for differential cell-cell communication. Cell-cell communication analysis was conducted using the MultiNicheNet[136] package (version 1.0.3) in R to infer potential interactions between cell types based on the expression of ligands, receptors, and target genes within signaling pathways on samples in *Article III*. The MultiNicheNet analysis included 31 matched samples, integrating $CD138^+$ cells (comprising both polyclonal and clonal plasma cells) and $CD138^-$ cells from the tumor microenvironment (TME), totaling 171,928 cells. Comparisons were made between patients with and without immunoparesis, following standard criteria for immunoparesis assessment[76]. The default MultiNicheNet pipeline was applied (https://github.com/saeyslab/multinichenetr). To improve robustness and prevent cluster drop-outs due to low cell numbers, specific subclusters were merged into larger clusters for analysis. Specifically, "CD8 Effector_1" and "CD8 Effector_2" were merged into the "CD8_Effector cluster"; "CD8 Memory_1" and "CD8 Memory_2" were merged into the "CD8_Memory cluster"; "cDC2" and "prog_DC" were merged into the "DC cluster"; "prog_B 1" and "prog_B 2" were merged into the "prog_B cluster"; "HSC," "LMPP," and "GMP" were merged into the "HSC cluster"; "CD14 Mono" and "CD16 Mono" were merged into the "monocyte cluster"; and "CD56 bright NK" and "NK" were merged into the "NK cluster".

Clusters were included in the analysis only if a minimum of 4 cells were present in all samples per condition. As a result, γδ-T cells did not pass this filter and were excluded from the MultiNicheNet analysis. Key parameters were carefully selected: LogFC_threshold was set to 0.5, empirical_pval to TRUE, p_val_threshold to 0.05, p_val_adj to TRUE, top_n_target to 250, and prioritization weights were kept at default settings.

All cell-cell interaction analyses and visualizations were conducted by comparing clonal and polyclonal plasma cells with the entire TME. Clonal and polyclonal plasma cells were prioritized as sender cells, with the TME as receiver cells and vice versa. The top 75 ligands per condition were selected for chord plot visualizations, and prioritized interactions were depicted using ligand-receptor pseudobulk expression product plots and ligand activity plots.

Single-cell HD Signature and Scoring

In *Article III*, the HD gene signature was defined through differential gene expression analysis using the "*wilcoxauc*" function in R, comparing HDs with MGUS, SMM, and MM. For each comparison, only genes that were significantly upregulated in HDs (Benjamini-Hochberg adjusted p-value < 0.05, log fold change > 0.1) were retained. The intersection of deregulated genes across each comparison was then calculated using the ComplexHeatmap (v2.18.0) function make_comb_mat in "distinct" mode, capturing only the genes specifically deregulated in each contrast. The gene signature activity in single cells was estimated using the *AddModuleScore* function from Seurat (v4.4.0) and visualized with *ComplexHeatmap*. The genes used for "hPC signature" scoring are listed in **Table VIII.**

In *Article III*, "immunoparesis signature" scoring, the genes are listed in **Table VIII**. The activity of the HD gene expression signature for each sample was then estimated in the publicly available MMRF bulk RNAseq dataset (IA17 release, https://research.themmrf.org), which includes bulk RNAseq data from the CD138^{+} fraction derived from MM patients. Survival analysis was performed using the survival (v3.5-8) and survminer (v0.4.9) packages.

Table VIII – Genes related to healthy PC signature and immunoparesis signature

Signature	Genes
Healthy PC	"BCL2L11," "CCDC85B," "COBLL1," "DNAAF1," "EIF2AK3," "FNBP1," "HCST," "HIST1H1C," "JSRP1," "KHDRBS1," "NEAT1," "RHOH," "RPL10," "RPL10A," "RPL15," "RPL21," "RPLP2," "RPS16," "RPS24," "UBALD2," "UBE2J1," "UQCRB," "ACADM," "ANXA6," "ATM," "AURKAIP1," "CALM3," "COA3," "COX7B," "CPNE5," "CTSH," "CYC1," "DNAJC1," "DNAJC19," "DUSP22," "EIF2AK1," "FOXO3," "GADD45GIP1," "GFER," "GMPPA," "HSPB11," "IFI27L2," "IKBIP," "IRF2," "ISOC2," "KLHL5," "MANEA," "MDH1," "METTL23," "MPHOSPH8," "MRPL20," "MRPS34," "MRPS36," "NDUFA8," "NDUFB2," "NDUFB4," "NDUFV2," "OST4," "PARP1," "PLEK," "PNOC," "POLR2F," "PPCDC," "PPM1K," "PRDX3," "PRKAR1A," "PRR34-AS1," "PSMA4," "PSMA5," "PSMB9," "PTPN6," "PYCARD," "RPL22L1," "RSRP1," "SAR1B," "SERF2," "SMCHD1," "SPATS2," "SYTL1," "TIMM13," "TXNDC15," "UCP2," "UQCC2," "UQCRQ," "ZBP1," "ZBTB8OS," "ZNF706," "ZNHIT1," "ZRANB2," "AC103591.3," "AC233755.2," "AMPD1," "AP1S2," "ARMC10," "ASNSD1," "ATXN10," "BNIP3," "BRK1," "C15orf61," "CADPS2," "CAMTA1," "CCDC88C," "CMTM7," "CSRNP1," "CTHRC1," "CTSW," "DDT," "DNAJB11," "DSTN," "EIF4G2," "GLO1," "GLS," "HMGN2," "HSP90B1," "IFT27," "ITGA6," "ITGA8," "LYZ," "MOXD1," "MRPL33," "NDUFA5," "NME3," "NUS1," "PARVB," "PDIA4," "PGRMC2," "PHF14," "PLA2G2D," "PSMA3-AS1," "PTK2B," "RAB11FIP1," "RBBP6," "SEC22B," "SEC24D," "SEC61G," "SF3B5," "SNHG9," "SOD1," "SPG21," "SRP9," "SRSF5," "SRSF7," "TMED2," "TMEM50B," "TRA2A," "UFM1," "Z93241.1," "ZFP36".
Immunoparesis	"BTG1," "CCND1," "CIRBP," "EEF1A1," "IRF1," "MTRNR2L8," "NOP53," "PNRC1," "RPL12," "RPS14"

VDJ analysis – Dandelion

Detailed VDJ repertoire analyses were conducted through the Dandelion[132,137] tool (v0.3.8) following documentation at https://sc-dandelion.readthedocs.io/en/latest/ website.
Input files for Dandelion are *filtered_contig_annotation.csv* and *all_contig.fasta* from Cell Ranger vdj and *filtered_feature_bc_matrix.h5* from Cell Ranger GEX.
The pre-processing step is aimed at the VDJ gene reannotation and mutation quantification. VDJ reannotation was performed using IgBLASTN and formatted into the AIRR standard with Change-O's MakeDB.py script[133]. The pipeline was enhanced by running BLASTN in parallel and selecting the best alignments. Heavy chain IG V gene alleles were reassigned using TIgGER[138]. IG constant region assignments were refined by aligning against a curated set of highly specific C gene sequences. Mutations were quantified using SHazaM's *observedMutations* function[133]. Everything was run using a Singularity container on an HPC cluster.
Then, VDJ data were integrated with GEX data using Scanpy[123] (v.1.4.6), and then filtered. Post-processing steps include VDJ clustering (based on BCR sequence similarity) and network generation. All plots were generated in Scanpy following Dandelion documentation.
The BCR chain status is presented in **Table IX.**

Table IX – Glossary of possible BCR chain status

Chain Status	Description
Single pair	The B-cell has exactly **one matched pair** of heavy (IGH) and light (IGK or IGL) chains. This indicates a typical, complete BCR.
Orphan VDJ	Only the **heavy chain (IGH)** is detected, without a corresponding light chain (IGK or IGL).
Orphan VJ	Only the **light chain (IGK or IGL)** is detected, without a corresponding heavy chain (IGH).
Extra pair	The cell has **more than one complete pair** of BCR chains (e.g., multiple IGH+IGK or IGH+IGL pairs), rare
Extra pair-exception	Similar to "Extra pair", but flagged as an **exception**, possibly due to specific biological phenomena like a naive B-cell expressing both IgM and IgD heavy chains.
No contig	No BCR chains (neither heavy nor light) are detected in the cell.

Results

Identification of Clonal BCR Rearrangements in MM using scVDJseq [Article I]

Here we analyzed 19 samples derived from 18 patients affected by MGUS (n=6), SMM (n=11), and MM (n=1), discrepancy between the number of samples and patients arises from one patient being sampled twice, during both asymptomatic and symptomatic stages. $CD138^+$ BM PCs were purified from each sample using immuno-magnetic bead separation, followed by scRNAseq and DNA extraction.

DNA analysis of these PCs was performed according to the Euro Clonality-NGS Working Group's standard procedure for two-step Ig NGS-based marker identification [Version 1.0 (11 June 2019)][110]. The Vidjil tool[120] successfully identified VDJ rearranged sequences within the IGH locus of dominant clonotypes across all samples.

For scRNAseq, a total of 69828 unique cell barcodes were obtained. Among these, 8806 barcodes were not associated with cells expressing VDJ transcripts (Figure 1a). Seurat's automatic cell assignment identified these cells as components of the TME, including T cells, monocytes, and NK cells (Figure 1b). This indicates unwanted contamination of the positive selection of $CD138^+$ cells, leading to their initial exclusion from further analysis.

The remaining 61022 cells expressed a VDJ transcript, with an average of 3212 cells per sample (ranging from 606–7379) (Figure 2a). A dominant clonotype (characterized by the over-represented VDJ rearrangement in each PCs sample) was easily identified in each sample using scVDJseq. On average, 83.65% of VDJ rearranged cells (ranging from 49.79–99.44% per sample) belonged to the dominant clonotype (Figure 2b). These 19 clonotypes accounted for 53935 of these 61022 barcodes (representing 88.39% of the total). The IGHV3 family was the most abundant, with IGHV3-43 being the most represented gene and IGHJ4 the prevalent IGHJ family (Figure 3a and Figure 3b, respectively).

The VDJ rearranged sequences of dominant clonotypes identified by the two methods were identical in 17/19 samples (Figure 4a, Table 1). Relative clonal fractions were heterogeneous, very moderately correlated inter-assay (Pearson's product-moment correlation: $R^2 = 0.19$; p-value = 0.08), and constantly higher in scRNAseq analysis (Figure 4b).

Dominant CDR3 sequences identified by scRNAseq were confirmed to match exactly with those determined by amplicon-based NGS from bulk DNA, demonstrating the reliability of the single-cell

approach. However, in 6 out of 17 cases, the V gene identified as the closest match to the patient-specific rearrangement differed between the Vidjil and CellRanger algorithms (Table 1). These discrepancies can be attributed to several factors. Vidjil and CellRanger use different references databases (IMGT[139] and Ensembl[140] respectively). This can result into slight alignment differences between the two algorithms, which are more likely given the low homology of these genes with the germline repertoire due to SHM.

Two patients exhibited discordant rearrangements between DNA-based NGS and scRNAseq. In patient sample PLC-14, V(D)J analysis of bulk DNA revealed a single clonotype with an unproductive rearrangement (IGHV3-23*01/IGHD3-9*01/IGHJ5*02) due to stop codons, which appeared in 78% of the reads. In contrast, scRNAseq identified a productive rearrangement involving the IGHV1-18 and IGHJ4 genes in 98.08% of the barcodes. These V(D)J sequences were exclusive to their respective methods, with no crossover detected.

The absence of the unproductive rearrangement in scRNAseq results is consistent with the expected degradation of transcripts containing premature stop codons via the nonsense-mediated decay (NMD) pathway. Conversely, the productive rearrangement detected by scRNAseq likely evaded detection by the Euro Clonality-NGS protocol due to a mutation in the J gene. This mutation likely caused a mismatch at the 3′ end of the reverse primers used in the first-step PCR, which is known to impede PCR priming (Figure 5).

Furthermore, bulk DNA analysis failed to identify any productive clonotype in patient sample PLC-16, detecting only a dominant unproductive rearrangement (IGHV3-11*05/IGHD3-3*01/IGHJ5*02) in 29.97% of the reads. In contrast, scRNAseq revealed a dominant light chain-only clonotype (IGKV1-33/IGKJ3) in 79% of the barcodes because this tool enables the identification of both heavy and light chains. DNA assay was performed with IGH primers only. This finding was consistent with serum immunofixation electrophoresis results for this patient, which indicated an IGK myeloma.

Analysis of BCR Diversity and Transcriptomic Markers in Single Cell [Article II]

The application of scVDJseq has proven particularly effective in providing robust and accurate data for deriving BCR sequences at the single-cell level. A significant advantage of this approach lies in its ability to integrate BCR information with the transcriptome of individual cells, offering an unprecedented level of insight.

In this study, 5 BM aspirate samples (SMM n=4, MM n=1) were collected. They were first selected for CD138 obtaining the positive ($CD138^{+}$) and negative ($CD138^{-}$) fractions. Subsequently,

scRNAseq with BCR profiling was performed. Two of the five samples were previously frozen, thawed, and processed as described. Collectively, 33274 cells were sequenced. Following the bioinformatic pipeline, standard QC (Figure 6a), sample integration and Harmony for batch correction in Seurat were performed (Figure 6b). The final integrated dataset included 21326 cells (Figure 6c). UMAP projection showed that cells from individual patients predominantly formed distinct clusters of cells, only few of them clustered together (Figure 7a). Automated cell assignment identified 20465 plasmablasts (95.96%), the remaining cells were CD14/CD16 monocytes, T cells, B cells, NK cells, megakaryocytes progenitors (prog Mk), hematopoietic stem cells (HSC), red blood cells progenitors (prog RBC), granulocyte/monocyte progenitor cells (GMP), and dendritic cells (DC) (Figure 7b). Immuno-magnetic bead enrichment demonstrated purity levels between 95% and 99%, except for patient P4, whose clusters (C4 and C5) consisted mainly of CD14/CD16 monocytes and T cells, indicating lower purity (Figure 7c-d). Notably, distinct patterns of differential marker genes were observed in PCs from patients, highlighting the heterogeneity within the cohort (Figure 8a). Violin plots in Figure 8b illustrate the expression of selected PC marker genes, where CCND1 and CCND3 were upregulated in samples P1 and P2, respectively, associated with t(11;14) and t(6;14), as confirmed by fluorescence in situ hybridization (FISH). Additionally, CCND3 and CCND2 expression was observed across various patients and genetic subgroups. ITGB7, a well-known high-risk marker for MM, showed high expression in most samples.

Leveraging scVDJseq, PCs were carefully examined, identifying cPCs and pPCs (Figure 9a). The data indicate that $CD138^{+}$ fraction samples from patients with SMM and MM can contain a significant proportion of pPCs, ranging from 0.51% to 2.62%.

The analysis of clonal versus polyclonal PCs in this cohort revealed common patterns, such as increased GPRC5D expression and reduced CD27 marker expression in neoplastic PCs, along with patient-specific levels of ITGB7 integrin in polyclonal PCs. Additionally, CD19 and CD99 expression specifically identified polyclonal PCs, whereas clonal PCs were defined by frizzled related protein (FRZB) and other cyclin genes[56] as previously reported (Figure 9b).

The detailed BCR analysis of clonal and polyclonal cells (Figure 9c, 10a, 10c) revealed surprising findings. Patients P2, P3 and P5 were characterized by one dominant clonotype and pPCs (Figure 9c). Whereas, in patients P1 and P4, two clonal BCR clonotypes were discovered (Figures 10a and 10c). Patient P1, t(11;14) confirmed by FISH, exhibits clonotype 1 in 45.11% and clonotype 2 in 52.45% of PCs. The two were related clonotypes, in fact, clonotype 1 displayed clonal VDJ rearrangements

in both heavy and light chains, whereas clonotype 2 had the same light-chain sequence but no IGH rearrangement, indicating a MM characterized by "light-chain escape". Transcriptomic analysis indicated that both clonotypes overexpressed CCND1 (Figure 10b), suggesting they both carried the t(11;14) and they were subclones representative of the same disease. Interestingly, only clonotype 1 showed higher levels of CD20 and MYC expression. Conversely, retrospective analysis identified that P4 had two different sMP, IgA/k and IgM/k, indicating that the clonotypes were associated with different neoplastic conditions: SMM and Waldenström's macroglobulinemia (WM), respectively. Indeed, scVDJseq confirmed the two clonotypes. Clonotype 1 overexpressed typical MM genes, such as CCND2. (Figure 10d). Obtaining information about the clonal status of each single cell through the scVDJseq suggests that the study of cPCs and pPCs can be refined and improved, adding resolution to the assessment of subclonality. The underlying biological insights of subclonal clustering would not have been uncovered without BCR analysis.

Inference of Genomic Properties from scRNAseq [Article II]

The feasibility of using scRNAseq data to identify IGH translocations and CNAs at the single-cell level was explored. For mapping IGH translocations, the Fuscia tool was utilized. Fuscia detected few fusion events of varying nature across all patients, including cases with negative FISH results, indicating inherent baseline noise in the algorithm. However, a distinct signal was evident in FISH-translocated cases (Figure 11a). Fuscia was used to confirm that the dominant BCR clonotypes in sample P1, marked by overexpression of CCND1, originated from a cell with t(11;14) translocation. A fusion event was observed in 2.5% of the overall cells, with 89% belonging to clonotype 1 and 11% to clonotype 2 (Figure 11b). Additionally, IGH-CCND3 fusion events were identified in sample P2, which exhibited high CCND3 expression, indicating a t(6;14) translocation (Figure 11d). Fuscia demonstrated higher specificity than sensitivity, and the actual number of translocated cells is likely underestimated. Nonetheless, it shows promise for qualitatively assessing the presence of IGH translocations at the sample level.
InferCNV was employed to detect CNAs (Figure 11e). The expected pattern of HRD was found in non-translocated samples, with patient P2 exhibiting gains of chromosomes (chrs) 3, 11, 19, and 22 and losses of chrs 1p, 13, and 14. Patient P3 showed gains of chrs 1q, 11, 15, 18, 19, and 22 and losses of chr 1p. The most complex profile was observed in P5, characterized by gains of chrs 3, 5, 7, 9, 15, 19, and 21, along with a clonal 17p loss and chr 11 gain, which was linked to high CCND1 expression without t(11;14). Furthermore, not all trisomies were present in all PCs within each sample; for

instance, the chr 19 gain was missing in a fraction of clonal PCs in P2 and P5. These findings suggest that analyzing clonal V(D)J rearrangements in conjunction with detecting fusion transcripts as surrogates for IGH translocations and characterizing CNAs could be a viable approach for accurately defining the World Health Organization (WHO) MM subtype. This method also facilitates the identification of small, rare subclones, particularly in rare cases where IGH translocations and HRD coexist.

Together, scVDJseq and CNA analysis enhanced the ability to depict the subclonal heterogeneity. In sample P1, the one with two clonotypes exhibiting "light-chain escape" (Figure 12a, Table 2), common CNAs across the two clonotypes were detected. Chr 1p deletion and gains of chr 1q and chr 22 are shared between clonotypes. However, significant differences were also found within the tumoral PC compartment, thus across the two clonotypes (Figure 12b). Subclone 1 displayed gains of chr 9 and losses of chr 12, 13, and 19, along with an apparent gain of chr 14q. Subclone 2 was marked by gains of chr 11q and chr 19p and losses of chr 5 and chr 17. When mapping cells of these CNA-defined subclones onto the UMAP space, there was a perfect correlation between CNA subclones and BCR clonotypic rearrangements (Figure 10a, 12c). This indicates that the light-chain clone underwent significant evolution from the ancestral clone, extending far beyond the mere loss of the IGH locus, but acquires new CNAs.

Transcriptomic analysis, such as bulk RNAseq, alone is undoubtedly a limited approach considering that it cannot clearly explain the different UMAP space distribution of the 2 subclones. However, the inferred genotypic analysis using scRNAseq provided a clear biological rationale. The high expression of the IGH in clonotype 1 explained the apparent chr 14q gain in subclone 1. In contrast, subclone 2 showed low IGH locus expression but high light chain expression (Figure 12a-d). Both subclones exhibited overexpression of CCND1, with subclone 2 showing greater levels, likely due to a chr 11 gain (Figure 12b,e).

Subclone 1 is characterized by overexpression of CD20 (MS4A1) and MYC, driven by a t(8;14) translocation, which was absent in subclone 2, indicating significant functional differences (Figure 11c). Gene Ontology (GO) enrichment analysis revealed that subclone 2 had a significant over-representation of processes related to the Major Histocompatibility Complex (MHC) and protein processing in the endoplasmic reticulum (Figure 12f).

Clonotype 1 and clonotype 2 are clearly sequential stages in the evolution of the same ancestral clone, because of the presence of shared CNAs (1p loss and 1q amp), and distinct CNAs among clonotypes indicate that these two clonotypes have a common ancestor and additional genomic alterations may have driven the evolution of the clonotype 2. An RNA trajectory analysis was applied to determine

if cell states could reflect this clonal progression. The pseudotime analysis illustrated a trajectory beginning with cells containing the full immunoglobulin rearrangement and progressing to cells with IGH loss (Figure 12g-l), reinforcing the concept of sequential evolution among the MM subclones. The spatial distribution of the two clonotypes, along with the varying expression levels of IGH, IGK and CCND1 genes, further supported this evolutionary sequence.
Similar geno-transcriptomic analyses were conducted on patient P4, who was characterized by the co-existence of two distinct tumor clones: Waldenström macroglobulinemia (WM) and SMM. BM biopsy revealed dual infiltration by cPCs and clonal lymphoplasmacytoid cells. To study the lymphoid compartment, scRNAseq analysis was performed on both $CD138^+$ and $CD138^-$ fractions. scVDJseq analysis identified two populations with completely different BCR profiles (Table 3), each expressing different IGH isotypes, specifically IgA for SMM and IgM for WM (Figure 13a). InferCNV on both clonotypes revealed three distinct CNA subclones (Figure 13b): subclone 1 corresponding to the MM condition, and subclones 2 and 3 corresponding respectively to the B cells and PCs of the WM clone. MM PCs (subclone 1) exhibited amp1q, amp4, del13, and del16q, which are characteristic of MM. Both WM B cells and PCs shared del1p, but WM B cells also showed unique abnormalities such as amplifications of chr 6 and chr 2q, and deletions of 7q (Figure 13b). This suggests an independent clonal origin for the two neoplasms, with a shared ancestor for WM cells followed by cell type-specific evolution in B cells and PCs.
Expression levels of clonal Ig in the different subclones were consistent with their association with WM and MM diseases (Figure 13c), showing perfect correlation with clonotypes identified by scVDJseq analysis. BCR clonotypes and CNV subclones were projected back onto the UMAP containing all P4 cells (both $CD138^+$ and $CD138^-$ fractions), PCs were distinctly separated based on their clonal origin, while WM memory B cells clustered together (Figure 13d-e). Additionally, cells from clonotype 2 were separated based on their phenotype into B cells (subclone 2) and PCs (subclone 3).
Comparative gene expression analyses revealed significant upregulation of SDC1, GPRC5D, ITGB7, LAMP5, CCND2, and MYC in MM PCs, as expected by a MM transcriptomic profile. Conversely, WM PCs showed elevated expression of CD27, CD79A, and CD79B, indicative of a more physiological phenotype and highlighting the differences from clonal MM PCs. Subclone 2 (WM B cells) specifically overexpressed CD19, MS4A1, and CCND3, typical B cell markers. Additionally, both subclone 2 and subclone 3 (WM) shared upregulation of CD79A, CD79B, CXCR4, CD99, and CD27, though at varying levels (Figure 13f).
GO enrichment revealed that WM PCs were enriched in pathways related to IgM processing, while

MM PCs were characterized by pathways associated with proliferation and the microenvironment Figure 13g). WM PCs also showed enrichment for pathways involving the endoplasmic reticulum, consistent with their role in antibody production. In contrast, WM B cells exhibited upregulation of pathways involved in B-cell maturation through somatic immunoglobulin differentiation (Figure 13h).

Overall, the integrative geno-transcriptomic analysis of scVDJseq, scRNAseq, CNAs, and structural variants offers a powerful tool for dissecting the genomic architecture of tumors at the single-cell level. By doing so, it not only provides a clearer picture of the clonal heterogeneity within the tumor but also captures the downstream transcriptional, and therefore functional, consequences of these genomic alterations. This approach is essential for unraveling how specific genetic changes shape tumor behavior and drive its evolution, offering new insights into the dynamics of malignant progression and uncovering critical differential gene expressions tied to distinct genomic profiles.

Single-Cell Transcriptome of Polyclonal PCs across PC Dyscrasia [Article III]

In previous chapters, as in much of the scientific literature, the focus has predominantly been on the molecular features of clonal aspects of PCs, often neglecting the polyclonal PCs counterpart in MM. The study of pPCs can offer significant advantages in achieving a more comprehensive understanding of the disease.

Therefore, the following chapters will delve deeply into the concept of pPCs and their role in myeloma pathogenesis.

A total of 46 BM aspirates were collected from patients with PC dyscrasia (7 MGUS, 16 SMM, and 23 MM). Additionally, 1 BM aspirate from HD was collected. All BM aspirates were processed obtaining $CD138^+$ and $CD138^-$ fractions, then processed for scRNAseq and scVDJseq. Additional 17 HDs were added to the single-cell dataset (3 from collaborators and 14 from publicly available sources[102,119]. All HDs samples were used as control. Leveraging scVDJseq, genotypic identification of clonal and polyclonal PC (pPCs) was performed. The polyclonality status of PCs was defined by the absence of identity and hierarchical relationship between the heavy and light chain BCR sequence of pPCs and the clonal counterpart within the same patient. Cells sharing the same light chain or showing significant identity with the clonal sequence were excluded from further analysis. inferCNV[141] was used for CNAs analysis as confirmation of the absence of clonal CNAs in the pPCs of the same sample. For the bioinformatic analysis, standard quality control filters were applied

(Figure 14) totaling 213074 PCs (32704 from HDs, 15882 from MGUS, 55409 from SMM, and 109079 from MM) (Figure 15a). Clonotype analysis on PCs revealed 170357 cells (79.95%) assigned as clonal and 42717 (20.04%) assigned as polyclonal. Of note, frequency of pPCs was much higher in MGUS samples (19.53%) and decreased in SMM (5.93%) and MM samples (3.31%). (Figure 15a-b). From UMAP projection, a notable admixture of pPCs and cPCs was observed, suggesting that transcriptomics alone may not be sufficient to precisely differentiate pPCs from cPCs (Figure 15c). Additionally, MM samples generally appeared more distinct from asymptomatic and HD (Figure 15d). A comprehensive analysis of canonical PC marker genes corroborated literature findings, such as the frequent loss of CD27[102,142] in cPCs and consistent expression of CD38, CD138, and TNFRSF17 in PCs. Known oncogenes like NSD2, CCND1/2, and MAF were exclusively expressed in cPCs belonging to patients. Of note, GPRC5D expression was generally higher in cPCs than in pPCs, nearly absent in HDs (Figure 15e-f). Conversely, TNFRSF17 (the encoding gene for BCMA) did not show great differences in RNA expression levels between clonal and polyclonal PC, rather across disease stages (Figure 15g-h). These data could explain the clinical implications for the risk of infections after immunotherapies with anti-GPRC5D T-cell or anti-BCMA treatments[143]. Interestingly, there have been found markers typically associated with cPCs, such as IGTB7 or CCND3, slightly expressed by pPCs from patients as well as from HDs (Figure 15i). These preliminary findings emphasize the need to fill the gaps to better understand the tumor biology underneath PCs in MM. Indeed, cPCs exhibit expected expression patterns of canonical marker genes, but pPCs do not exhibit a peculiar transcriptomic profile different enough to distinguish them from cPCs. Thus, the clonotypic identification of clonal and polyclonal PCs in patients with PC dyscrasias is crucial to properly define the tumoral and the residual healthy PC population.

Polyclonal Plasma Cells from Patients Exhibit Corrupted Phenotype [Article III]

To better characterize the polyclonal counterpart, all the cPCs were removed from the dataset focusing on pPCs belonging to patients and HDs. UMAP representation displays a slight divergence between patient and HD but still a significant overlap of cells across all disease stages (Figure 16a). Differentially expressed genes across clinical stages revealed that patient-derived pPCs exhibited a distinct expression pattern compared to those from HDs (Figure 16b). Basically, pPCs from patients shared patterns of gene expression among different clinical stages (MGUS, SMM and MM), although pPCs from HDs were characterized by private gene expression patterns. HDs overexpressed genes such as FKBP2, SELK, GNB2L1 and DUSP1 which are primarily involved in protein folding and

protection against oxidative stress[144–146]. In contrast, patient-derived pPCs showed upregulation of genes typically associated with inflammation and stress conditions, including the autophagy receptor SQSTM1[147,148], the growth inhibitor GADD45B, and the VIM gene, which is often upregulated during epithelial to mesenchymal transition[149] (Figure 16c). pPCs (as PC from HDs) did not show gene sets implicated in the myeloma pathogenesis, however certain genes, such as CTSB, CTSD, and OPTN —associated with autophagy— were found to be overexpressed in pPCs from patients but not in those from HDs.

Remarkably, it was found that only pPCs from SMM and MM expressed higher levels of surface integrin ICAM1 compared to MGUS or HDs, suggesting that the interaction pPC-TME shapes as the disease progresses (Figure 16c). Subsequently, differential expressed genes (DEG) were investigated to assess pathway activity by comparing PCs from HDs with pPCs from patients in one-to-one contrast (HD vs MGUS, HD vs SMM, HD vs MM). Results showed a trend toward upregulation of inflammatory pathways (including hallmark TNF-α and IFN-α response, and hallmark oxidative phosphorylation) in pPCs from pathological samples. Additionally, the most pronounced differences were observed between MGUS and SMM/MM (Figure 17a). In conclusion, these findings reveal a progressive increase in inflammatory status from HDs to MM (Figure 17b).

Single-Cell Transcriptomic Insights into Immunoparesis [Article III]

Immunoparesis, a hallmark of PC dyscrasia, is characterized by the reduced production of normal immunoglobulins. This immunosuppressive state heightens the risk of infections and is often associated with more advanced disease and poorer prognosis[150–152]. It is linked to disease progression in asymptomatic stages[76,153] and reduced survival in multiple myeloma (MM) due to infection-related complications. Diagnosing immunoparesis involves measuring the levels of immunoglobulins uninvolved in the disease.

In this study, patients were clinically assessed for immunoparesis to examine the transcriptional profiles of pPCs in samples with and without immunoparesis. Excluding healthy donors, pPCs from 46 samples were analyzed, with immunoparesis present in 38 of them. At first glance, a noticeable decline in the frequency of pPCs was observed in patients with immunoparesis (Figure 18a-b), which may partly account for the reduced production of normal immunoglobulins.

Compared to immunoparesis-negative patients, pPCs from immunoparesis-positive exhibited a more deregulated transcriptome. This latter displays upregulation of genes related to the interferon pathway (e.g., IRF1, MX1, and ISG15), the Unfolded Protein Response (UPR) (e.g., ERN1), and genes involved in cell proliferation and apoptosis (e.g., H3F3B and NOP53).

pPCs from immunoparesis-negative patients showed upregulation of genes associated with NF-KB signaling (e.g., BIRC3), cell adhesion (e.g., CXCR4, SDC1, and PECAM1), autophagy (e.g., FOS, CST3), and anti-apoptotic pathways (e.g., DYNLL1) (Figure 18c).
To further elucidate the functional implications of these gene expression changes, pathway analyses were conducted by comparing the immunoparesis status within a subset of 23 symptomatic patients. In the presence of immunoparesis, pPCs were found to be functionally enriched for interferon-related pathways (hallmark IFN-α and hallmark IFN-γ) and oxidative phosphorylation (Figure 18d). These findings underscore how immunoparesis, often an overlooked clinical feature, can profoundly alter the expression profiles of pPCs, establishing a clear correlation between the presence of a pro-inflammatory phenotype and decreased immunoglobulin production.

Unraveling the Interactome Landscape of Immunoparesis in Polyclonal Plasma Cells [Article III]

MM progression is influenced by the TME, where the functional properties of clonal and polyclonal PCs are deeply dependent on their interactions with other cells in the bone marrow niche. scRNAseq provides a powerful tool to study and predict these cell-cell interactions by leveraging the transcriptome of each individual cell. In this study, 31 paired $CD138^+$ and $CD138^-$ cell fractions were analyzed by combining single-cell transcriptome data with clonal VDJ repertoire profiling. A total of 171928 cells were integrated into a single dataset, allowing for an in-depth exploration of cell composition and interaction dynamics using the MultiNicheNet pipeline[136]. This advanced bioinformatic tool infers cell-cell interactions based on ligand-receptor activity and downstream pathway activation within the cell. The analysis specifically focused on 102194 cPCs, 7447 pPCs, and 62287 cells from the TME (Table 4, Figure 19a).
In the context of MM, studying immunoparesis within the TME is critical as it provides insight into how immune suppression contributes to disease progression and resistance to therapies. Understanding these interactions could guide the development of more effective immunotherapies and improve patient outcomes[154]. Thus, clinical data related to immunoparesis were integrated, resulting in the inclusion of 148006 cells from 23 samples with immunoparesis (89037 cPCs, 3923 pPCs, and 55046 TME cells) and 23922 cells from 8 immunoparesis-negative samples (13157 cPCs, 3524 pPCs, and 7241 TME cells) (Figure 19b). The TME consisted of various cell types, including 2980 HSCs, 1019 DCs, 15261 myeloid cells, 6589 NK cells, 32966 T cells, and 3472 B cells. No significant differences were observed in the relative composition of TME populations between the two conditions (Table 4, Figure 19b). However, a significant depletion of pPCs was observed in

immunoparesis-positive patients (padjust = 0.004) (Figure 19c).
The MultiNicheNet tool was employed to explore potential functional differences across immunoparesis-positive and immunoparesis-negative interactomes. This tool allows specific cell types to be designated as either senders or receivers of signals.

The analysis initially focused on the immunoparesis-negative setting. By comparing predicted ligand-receptor interactions and analyzing downstream targets, it was predicted that pPCs actively send signals to various cell compartments, including other pPCs, CD8 effector cells, CD8 memory cells, and progenitor RBC cells (Figures 20a, 21, 22). Noteworthy interactions included the CD38-PECAM1 cell interaction, which is known to play a role in cell migration regulation, and a range of cell adhesion interactions involving ICAM2, ICAM3, and ITGB2 (T lymphocytes), all of which are associated with promoting cell survival processes[155–159]. Additionally, cPCs and pPCs were predicted to receive signals from myeloid cells (e.g., DCs, HSCs, monocytes, pDC, Mk progenitor cells) and lymphocytes (e.g., CD8 effector, CD8 memory, NK, progenitor B, and MAIT cells).
In the immunoparesis-negative scenario, immune and inflammatory responses were driven by specific interactions, such as THBS-1 (in monocytes and Mk progenitor cells) with SDC1 (in pPCs) and ITGA6 (in pPCs), as well as HMGB-1 (in HSCs, NK, pDCs, MAIT, CD8 cells, progenitor B cells, and monocytes) with SDC1 (in pPCs). The CCL5 ligand-receptor interaction between T cells and PCs, known to be involved in tumor cell migration[160], was also identified. Downstream, genes related to autophagy and ER stress regulation (e.g., BNIP3, MT1X, MT2A)[161,162] were upregulated, along with regulators of B cell survival pathways (e.g., CD74 and ITGA6)[163] (Figure 25a). These results reveal the active involvement of pPCs in sustaining a pro-inflammatory TME in the immunoparesis-negative cell landscape, where pPCs, despite the tumoral context, maintain their functional roles through autophagy, ER stress regulation, and activation of B cell survival pathways.

In contrast, in the immunoparesis-positive condition, cPCs became the main contributors to TME interactome, with a significant reduction in pPCs interactions (Figures 20b, 23, 24). These differences were not due to the depletion of pPCs, as the differential expression analyses in MultiNicheNet were normalized for cell count.
In this scenario, both clonal and polyclonal PCs were characterized by inflammatory signaling mediated by the IFITM1-CD81 interaction (pPCs-pPCs, pPCs-CD8 effector/memory cells, pPCs-Treg), which is involved in the IFN-gamma response pathway[164,165]. Notably, the interactions between the myeloid compartment and pPCs were significantly depleted, with DCs and pDCs found to support pPC survival through the BAFF-BCMA axis[166–168]. Additionally, the TGFB2 interacts with

TGFB receptors (TGFBR1, TGFBR2, TGFBR3) in cPCs and CD8 T cells, along with the TGFB1-ITGB1 interaction in MAIT and NK cells with pPCs, suggesting an immunosuppressive role of PCs against T-cell compartment.

In the immunoparesis setting, IL-15 signaling plays a critical role. In fact, immunoparesis-negative samples showed IL-15 interacting with IL2RG (or CD132) from pPCs as sender towards PCs (both clonal and polyclonal) as receiver. Conversely, in immunoparesis-positive samples, IL-15 was signaled from clonal counterparts towards PCs (both clonal and polyclonal) through the IL15RA receptor[169–172], but only in this case downstream signaling involving the activation of pro-inflammatory pathways. Additionally, downstream target genes related to inflammation and interferon pathways (e.g., IFI16, IFIT3, PLSCR1) were upregulated, along with negative regulators of the NF-kB pathway, such as TNFSF10. In contrast, genes related to autophagy and B cell maintenance (e.g., CSTH, BNIP3, MTX1, MT2A) were downregulated[102,161] (Figure 25b).

The conclusions drawn from the comparisons exhibit somewhat similar inflammation in both cases. However, there are key distinctions between the two groups analyzed.

For the Immunoparesis-positive group, inflammatory pathways, particularly the IFN pathways, are significantly more upregulated. This is accompanied by pro-apoptotic signals, such as the upregulation of TNFSF10. Additionally, the reduction in the interaction with the myeloid compartment, could potentially reflect a diminished ability to present antigens.

In contrast, the immunoparesis-negative group shows a lower degree of inflammation, but still present, which is expected since these are patients with PC dyscrasia. However, they exhibit a higher activation of autophagy, a mechanism known to play a role in the maintenance of pPCs[173]. Furthermore, there is increased interaction with the myeloid compartment, with these cells seemingly sending activating rather than inhibitory signals to CD8+ T cells.

Signature Analysis as a Predictor of Survival in Independent Datasets [Article III]

Residual PCs have been shown to correlate with prognosis in MM patients[56,102]. The data presented here demonstrate that even pPCs in MM patients can be significantly dysregulated and differ markedly from PCs in HDs. Leveraging this unique dataset, a "healthy PC" (hPC) signature was generated by identifying differentially expressed genes in PCs from HDs compared to pPCs from MGUS, SMM, and MM patients, respectively. This hPC signature, which consists of 148 genes (Figure 26a, see details in Material and Method), corrects for the dysregulated expression observed

in pPCs from patients and includes genes exclusively expressed in hPCs. Additionally, an immunoparesis-specific signature composed of 10 deregulated genes was derived by comparing immunoparesis-negative and immunoparesis-positive patients across both asymptomatic and symptomatic subsets (Figure 26b, see details in Material and Method). When the hPC signature was applied to the cohort, it showed that it was most enriched in HDs and progressively less represented in pPCs from MGUS, SMM, and MM cases (Figure 26c upper panel). In contrast, the immunoparesis signature displayed an opposite trend, with increased enrichment from MGUS to MM (Figure 26c lower panel). Moreover, the inverse correlation of the hPC signature with the immunoparesis signature in the pPC dataset suggested that cases enriched for abnormal pPCs have a reduced capacity to produce polyclonal immunoglobulins and may exhibit a more compromised phenotype compared to healthy PCs (R: -0.29, p.value < 2.2e-16, Figure 26d).

To assess the clinical impact of hPC signature enrichment, it was applied to the symptomatic MM cohort. It was hypothesized that in bulk RNAseq data generated from $CD138^{+}$ selected cells, a portion of the signal could be derived from pPCs. RNAseq data from the CoMMpass (https://research.themmrf.org/) study were interrogated, and patients were ranked based on the fractional representation of the hPC signature. A categorical distinction based on whether the hPC signature was enriched above or below the median clearly separated patients by their survival outcomes. Specifically, the median progression-free survival (PFS) was 3.52 years [95% confidence interval, 3.13-4.21 years] in patients above the median and 2.16 years [95% confidence interval, 1.94-2.75 years] in patients below the median (p.value = 0.0006, log-rank test, Figure 26e). Median overall survival (OS) was not reached in patients above the median but was 5.5 years [4.68-not reached years] in patients below the median (p.value < 0.0001, log-rank test, Figure 26f).
These results suggest that the hPC score may serve as a surrogate marker for residual functional "polyclonality" within cPCs, indicating that the preservation of a normal polyclonal population impacts patient survival. This could potentially explain the observed correlation between immunoparesis and prognosis[174].

Clonal VDJ tracking [Manuscript in preparation]

Clonal evolution plays a critical role in the progression of MM. Understanding this evolution is crucial, particularly through the lens of VDJ recombination. Tracking VDJ rearrangements provides key insights into the behavior of malignant PCs clones over time. By focusing on the clonal architecture of the BCR repertoire, it becomes possible to identify patterns of clonal expansion, persistence, or suppression that correlate with disease progression.

By applying the Dandelion[132,137,175] tool, it becomes possible to determine eventual clonal changes across disease progression at the single-cell level. This tool has the ability to detect and analyze clonotypes of BCR sequences. This is achieved through a rigorous comparison of CDR3 regions, V and J gene usage, and CDR3 length, enabling to gain a deeper understanding of immune diversity and specificity.

Moreover, Dandelion excels in constructing clonal networks by using the minimum spanning trees algorithm, which elucidates the relationships between cells based on the similarity of their BCR sequences. This visualization of clonal networks provides crucial insights into clonal expansion and the overall diversity of the immune repertoire.

Dandelion allows the integration of immune repertoire analysis with GEX profiles from scRNAseq data. This multimodal approach is invaluable in providing a holistic view of immune cell function and VDJ genotype, thereby facilitating the discovery of new biological insights.

Preliminary data have been obtained by applying the Dandelion pipeline to 8 samples derived from 4 patients, with each patient providing samples at both asymptomatic and symptomatic stages. These time points were then directly compared within each patient to observe clonal evolution over the course of the disease.

Dandelion re-analyzes VDJ gene sequences by comparing them to reference sequences from the IMGT database[176], using the igblastn tool[177] for accurate annotation. The improvements are: (1) reliable results for D/J gene calls; (2) detect "*multi-J mapping*" contigs and (3) retain contigs missing V gene calls, which were normally filtered out. This results in an accurate definition of the VDJ chain status (Table IX, see in Material and Method) that can be associated to a cell when scVDJseq is integrated with scRNAseq.

This enables the UMAP space representation of the chain status. Of the 41 patients analyzed in *Article III,* 4 progressed to symptomatic MM, their UMAP projection (Figure 27) displayed a variety of chain status. The most abundant one is *single pair* chain which means a typical, complete BCR with exactly one matched pair of heavy (IGH) and light (IGK or IGL) chains. Cells flagged as *No_contigs*, could be linked to different scenarios. They might be contaminating cells within the $CD138^{+}$ fraction, meaning they don't belong to the B-cell lineage and therefore lack a BCR. Alternatively, they could be dropout events on PCs where the VDJ sequence isn't detected.

Further analysis was conducted without *No_contig* cells.

Since clonal PCs share the same CDR3 amino acid sequence, we wondered whether the tumor expansion, and the transition from asymptomatic to symptomatic stages, might lead to changes in the clonal VDJ sequence. To explore this, we used Dandelion to create clone networks.

Dandelion's approach to defining clonotypes and analyzing diversity through networks follows these key steps, which apply to both heavy- and light-chain contigs: (1) matching V and J gene usage, (2) identical junctional CDR3 amino acid length, and (3) CDR3 sequence similarity. For BCRs, it used 85% nucleotide sequence identity at the CDR3 junction. To build VDJ networks, the minimum-spanning tree algorithm creates graphs where edges represent the shortest total edit distance between neighboring cells, defined as number of mutational events in the VDJ nucleotide sequence. Instead, nodes represent clones (cells bearing a specific clonotype). Cells with a pairwise edit distance of zero (cells with identical BCR) are connected into the graph generating a bigger node that includes two or more clones. The clone network could be useful in inferring the relationships between BCR clonality, somatic hypermutation and diversity across samples and disease stages.
In the MM context, we applied Dandelion's clonal network analysis on the 4 patients' progression (Figure 28). Nodes, or clonotypes, without edges could be considered polyclonal cells and they surround the central nodes without connecting to them. On the contrary, clonal cells are localized at the center and can display a more or less dense network structure according to the mutational events occurred in the VDJ sequence. The more branched the clones are, the more mutational events have occurred, most likely due to SHM.
Of note, as the disease progresses, there is generally an observed increase in network complexity, marked by a higher number of nodes and connections (Figure 28). These nodes, while sharing the same CDR3 amino acid sequence, have nucleotide sequences that are sufficiently different to create dense branching. This phenomenon could be linked to tumor progression. However, further studies with larger sample sizes are needed to confirm this. It would be particularly valuable to identify the mutations responsible for the diversity between clones, and more importantly, to develop a mutational burden score derived from SHM. Such a score could be used to compare clones at the diagnosis of asymptomatic MM with those from symptomatic cases.

Discussion

The analysis of the data presented in this dissertation provides new and significant insights into the clonal and genetic complexity of PCs in MM and related disorders. Through the integration of scVDJseq technology with single-cell transcriptomic analysis, the clonal and polyclonal composition of PCs, as well as their evolutionary dynamics and interactions with the TME, were elucidated. As described in *Article I*, the high concordance observed between scVDJseq and DNA-based NGS in identifying dominant tumor cell-specific VDJ rearrangement supports the reliability of single-cell approaches in capturing the clonal landscape. Moreover, single-cell approach proved more effective than the conventional method based on bulk DNA VDJ sequencing, overcoming limitations where the conventional method failed. The discrepancies identified in a subset of samples between the two methods highlight the importance of method selection and the potential influence of technical factors, such as primer specificity and sequence homology, on the detection of clonal rearrangements. scVDJseq effectively mitigated the issue of false negatives arising from DNA mispriming, a common limitation in conventional DNA-based VDJ genotyping, which typically requires substantial amounts of starting DNA and numerous PCRs—conditions that may not always be feasible. Furthermore, the single-cell approach can process a few thousand cells, providing an advantage in scenarios with limited DNA availability. Integration of scRNAseq with single-cell BCR sequence enables the dissection of the transcriptome associated with a specific VDJ rearrangement, offering an unparalleled depth of information.

However, it remains a costly and resource-intensive technique, limiting its application primarily to research settings. To enhance the utility of scRNAseq in MM, it is suggested to combine a 5′ gene expression library with a VDJ-enriched library derived from the same cells' amplified cDNA. This approach minimizes the need for additional input material, reduces wet-lab and sequencing costs, and simplifies data analysis.

It is noteworthy that, in MM, where VDJ rearrangements are typically productive, immunoglobulins marker screening via scRNAseq does not face the limitations encountered in other B cell neoplasms, such as acute lymphoblastic leukemia (ALL)[178]. In ALL, the majority of rearrangements are unproductive, making RNAseq, which targets productive transcripts, less comprehensive. While scVDJseq offers the unique advantage of providing single-cell resolution, enabling detailed insights into clonal diversity and immune repertoire at an unprecedented level, it currently faces significant limitations. Costs, aside, one of the main drawbacks of scVDJseq lies in its relatively limited sensitivity due to the small number of cells analyzed per sample. This contrasts with the highly

sensitive, standardized techniques used in the context of MRD detection in MM such as next-generation flow cytometry (NGF) and next-generation sequencing (NGS), which can detect MRD down to 10^-5 or even 10^-6 cells[179,180]. scVDJseq it is not yet competitive with the high-sensitivity detection required for routine MRD monitoring in clinical settings. However, validating the effectiveness and sensitivity of this cutting-edge approach is essential to dissect clonality in each sample enabling a comprehensive geno-transcriptomic analysis through single cell technology on cPCs, pPCs, as well as TME. A deeper exploration of tumor heterogeneity and the functional properties of cPCs and pPCs could lead to improved patient outcomes.

In fact, one of the aims of this dissertation involves the dissection of the clonotypic architecture of cPCs and pPCs by combining clonal analysis with transcriptomic data to enhance the classification of MM subtypes and improve understanding of subclonal architecture.

The results presented proposed the possibility of enhancing the effectiveness of the scRNAseq approach in uncovering the tangled heterogeneity of PCs in MM, emphasizing purely single-cell data interpretation. Several studies have investigated the transcriptomic consequences of specific genomic alterations, seeking to understand their impact on disease progression. The nature of scRNAseq facilitates such analyses by providing a higher resolution of cellular diversity[102,112,181,182]. This technology enables the identification of distinct subclones with varying geno-transcriptomic profiles, offering a more refined understanding of clonal heterogeneity and potentially leading to more targeted therapeutic strategies. However, the potential benefits of integrating genomic and transcriptomic data at the single-cell level in MM have yet to be fully realized.

The approach outlined in this dissertation is characterized by its robustness and flexibility; indeed it successfully provided equivalent results in the identification of clonal VDJ compared with the clinical grade NGS approach adding an impressive amount of data coupling VDJ genotype with the transcriptome. This enhances the accuracy of distinguishing between clonal and polyclonal PCs, a critical factor for FISH and bulk sequencing, particularly in the early stages of MM where polyclonal cells are more prevalent and might cause “disease dilution”, leading to inaccurate results. Tirier et al. demonstrate how detecting CNAs in scRNAseq can better define specific subclones, such as those involving chr 1q amplification and others[183]. This dissertation shows that integrating VDJ genotype and transcriptomic data provides a more detailed and accurate view of clonal structure, emphasizing that scVDJseq can play a pivotal role in dissecting subclonal populations more precisely. In fact, this approach allowed for the classification of all cases into WHO genetic subtypes directly from raw scRNAseq data, without requiring additional sample preparation. The CNA analysis

supported previous findings that trisomies in HRD cases are not acquired simultaneously[6,184], as not all samples showed the complete set of chromosome trisomy in each cell. Evidence supports that combining inferred genomic properties and transcriptomic data enhanced the resolution of subclone clustering in the UMAP space, revealing complexities not fully explained by transcriptomic data alone.

Without the VDJ integration, it becomes challenging to unequivocally define clonotypes[56] at single-cell resolution, particularly in complex cases involving multiple clonotypes or light chain escape events, which is critical in MM due to its well-known genomic complexity.

For instance, results presented in *Article II* showed two noteworthy case report. In one sample (P1), a subclone evolved through the loss of a heavy chain and the acquisition of further aneuploidies. These findings highlight the capability of scRNAseq to detect light-chain-only subclones, which are often missed by traditional biochemical methods, particularly during asymptomatic stages.

In another sample (P4), BCR analysis confirmed the presence of two distinct clonal populations: one related to SMM and another associated with WM, each with unique VDJ sequences. Notably, while the SMM clonal cells were exclusively PCs, the WM clonal cells included both B cells and PCs, reflecting the disease biology. Interestingly, the transcriptomic profile of WM PCs was found to be more similar to pPCs than to MM PCs. Conversely, the differences between WM B cells and PCs emphasized the importance of antigen-receptor pathways in B cells and immunoglobulin production in PCs. For this case report, these insights could have significant clinical implications, suggesting that different therapeutic approaches might be needed for treating WM, with proteasome inhibitors potentially being more effective against PCs and BTK inhibitors more effective against B cells.

Both Boiarsky et al. and Ding et al**.** provide great insights into the clonal evolution and disease progression in MM[101,102]. Their work underscores the importance of tracking how clones evolve over time. In *Article III,* we demonstrate that an integrative approach could partially provide a clearer understanding of disease progression and clonal selection, aspects that are harder to track with scRNAseq alone.

Nevertheless, additional translational applications of the single-cell technology extend beyond these findings already discussed. As a matter of fact, the characterization of the TME is crucial for predicting responses to new immunotherapies, as well as the polyclonal component of PCs in MM disease. To better achieve the pivotal aims extensively described before, further analysis on a larger dataset has been performed.

The architecture of the TME has long been recognized as a critical factor in the initiation and progression of PC dyscrasias.

Immune changes in the early stages of the disease, such as increased NK cells and modulated T cell compartment, may affect disease progression and treatment response[103,185,186]. As well as iMSCs interaction with immune cells, promoting tumor survival and immune modulation[66]. Although pPCs are a significant component of the TME, they have not been considerably studied, despite their clear biological and clinical relevance. So far, researchers focused attention more on cPCs which is the real tumor component. This gap in research is partly due to technical challenges; for instance, pPCs cannot be easily sorted from cPCs using magnetic bead isolation or flow sorting techniques.

In response to these challenges, scVDJseq accurately identifies pPCs in patients with PC dyscrasias. The process of polyclonal selection was also manually inspected at the CDR3 sequence level to avoid any contaminant in the pPCs dataset. The robustness of this method was confirmed through CNA analysis, which revealed clonal aneuploidies exclusively in the cPCs population, further distinguishing them from pPCs.

As confirmed by previous findings, relying solely on transcriptomic data may not be sufficient to accurately distinguish pPCs from cPCs. Even after excluding immunoglobulin genes —highly expressed, cell-specific genes that contribute significantly to transcriptomic diversity— transcriptome sequencing alone might not effectively differentiate clonal from polyclonal counterpart[56,102]. This was highlighted by data shown in results, indicating that pPCs themselves may exhibit phenotypic abnormalities in myeloma-related conditions (*Article III*). This observation differs from previous studies[102] of Boiarsky et al., which used marker gene signature to dissect tumoral and polyclonal PCs. Although marker genes were generally effective in differentiating clonal from polyclonal PCs, the patient's variability suggests a need for more extensive studies to better understand the functional properties of polyclonal PCs.

The gap can be attributed, at least in part, to the innovative methodology employed in this research, such as the BCR sequencing at the single-cell level.

Furthermore, results clearly demonstrate that pPCs in patients revealed differences in the expression of canonical marker genes compared to cPCs. Notably, it was observed that the expression of GPRC5D was significantly lower in pPCs than in cPCs. On the contrary, the BCMA expression did not differ significantly. These data would be of high clinical relevance, suggesting that treatments targeting GPRC5D might selectively spare pPCs, potentially explaining the lower incidence of hypogammaglobulinemia and infections in patients treated with these therapies, compared to those treated with anti-BCMA ones[187].

The analysis presented uncovered a markedly dysregulated transcriptome in pPCs belonging to

patients when compared to PCs from HDs. In this context, pPCs from patients exhibited upregulation of markers associated with inflammation, autophagy, and interferon pathways like IFITM1, SQTM1, CTSB, CTSD, and OPTN[164,165], and ICAM1[188–192]. Interestingly, the upregulation of these pathways intensified from asymptomatic to symptomatic patients. Autophagy of PCs in MM is extremely important because it maintains cellular homeostasis under extreme conditions[173], such as those induced by oxidative and proteotoxic stress due to the high production and secretion of immunoglobulins[173,193,194]. Therefore, this process could be particularly amplified in the residual pPCs patients with PC dyscrasias.

Single-cell analysis, then, focused on a more translational impact. In this regard, there have been evaluated the presence or absence of immunoparesis in patients affected by MGUS, SMM and MM. Indeed, immunoparesis is frequently linked to more advanced stages of the disease and a worse prognosis, primarily because of complications arising from infections[76,153].

Our findings reveal a distinct molecular profile in patients with immunoparesis, characterized by the upregulation of several genes pPCs linked to pivotal immune and cellular pathways. Notably, there have been observed the activation of both IFN-alpha and IFN-gamma signaling cascades, which play a crucial role in modulating immune system activity. The genes IRF1, MX1, and ISG15, known to be integral components of inflammatory pathways, suggest a heightened state of immune activation[195].

In parallel, data presented here highlight the upregulation of genes involved in the UPR, such as ERN1, indicating a stress response within the endoplasmic reticulum[196]. Additionally, the upregulation of signaling pathways related to cell proliferation/apoptosis, including genes such as H3F3B and NOP53, points to a complex balance between survival and death signals in these cells. This dual activation may underlie the compromised immune function in patients with immunoparesis, where an overactive proliferative signal is counterbalanced by mechanisms of cell death[197,198].

Similar results have been obtained when looking at the predicted interactome between PCs and the TME within individual sample exploiting the MultiNicheNet tool. Results obtained from MultiNicheNet are consistent with a TME primed for inflammation in immunoparesis-positive patients. The most noticeable finding concerns a marked reduction in the interactions between pPCs and other immune cells within the TME. As already anticipated, the few interactions between pPCs - TME exhibited pro-inflammatory ligand-receptors. Specifically, interactions involving IFITM1 and CD81[164,165], along with the activation of the TGFB2 - TGFBR2/TGFBR3 axis, appear to be primarily driven by INF pathways. Of course, the aforementioned interactions may influence the survival and

function of genotypically normal PCs, including their production of IgM and IgG[199], making them phenotypically aberrant. These findings are corroborated by a decreased expression of autophagy-related genes, coupled with an increase in pro-apoptotic gene expression. Collectively, those could be one of the mechanisms underlying the corruption of pPCs, which contribute to immunoparesis in patients affected by MM.

Previous studies[200–202] opened to speculation about a TME that may exhibit a reduced competence in antigen presentation and immunosurveillance in immunoparesis condition, further exacerbating the already challenging relationships between antibody production, immune function, and tumor progression.

The results obtained so far in this dissertation clearly create a conundrum about the starting point of the inflammation status of the TME. The widely described disease pathogenesis and data presented here suggested that tumoral PCs, through their disrupted interactions with pPCs and TME, might be a contributing factor. However, it is also reasonable that the perturbed and inflamed TME itself plays a primary role in the development of PC dyscrasias, or somehow facilitates the process. Disrupted TME may interfere with the normal transcriptional programs of BM PCs, thereby promoting the expansion of pre-existing cPCs that harbor initiating genomic lesions—cells whose proliferative potential had not been previously blossomed. The concept that cPCs can progress to overt MM without acquiring new genomic mutations, potentially driven by microenvironmental stimuli, is well-supported[203]. These analyses offer new insights into the selective pressures exerted by the TME on pPCs, shedding light on this process for the first time.

Recent studies in solid tumors have highlighted the important role of B cells in immune surveillance, suggesting that their depletion can contribute to immune evasion and tumor progression[202,204–207]. In MM, a similar mechanism may be at play, where the reduction of polyclonal repertoire could limit the immune system's ability to detect and control malignant cells. This loss of B cell diversity may impair antitumor immunity, allowing the tumor to evade immune detection and promoting disease progression.

To this extent, an "hPC signature" (healthy PC signature) was identified when comparing PCs from HDs with pPCs from patients, rather than with the broader population of $CD138^+$ cells. This hPC signature was found to decrease progressively across different disease stages and showed an inverse correlation with the immunoparesis signature. Moreover, analysis of RNA sequencing data from the CoMMpass dataset revealed that patients with enrichment of this HD signature above the median had improved PFS and overall survival OS. These findings highlight the critical role of maintaining a functional polyclonal population among BM PCs, suggesting that patients with a preserved and

functional immune system may benefit from a better control of residual disease and effective treatment response.
As a matter of fact, the efficacy of immunomodulatory agents, such as daratumumab and IMiDs, highlights the therapeutic importance of restoring immune surveillance in MM.
This comprehensive approach and the resulting data provide new insights into the complex biology of pPCs within the TME, highlighting their potential role in the progression of PC dyscrasias and offering implications for more targeted therapeutic strategies.
In summary, this study introduces an innovative analytical approach to explore the transcriptional landscape of pPCs in PC dyscrasias. The results presented suggest that pPCs are significant contributors of TME, and their disruption has profound biological and clinical implications in MM. Understanding these dynamics could have important implications for the research and treatment of MM.
Another key aspect not to overlook is the clonal part of the tumor, to understand what happens to the cPCs during disease progression and, consequently, after pharmacological treatment. Dandelion[132,137,175] is a powerful tool that allows precise tracking of VDJ sequences by creating networks between clones based on sequence similarity.
The results obtained with this technology are promising and could reveal important insights into the biology of MM. Preliminary findings suggest that as tumor burden increases in patients, moving from an asymptomatic to a symptomatic phase, there is a trend of expansion and increased complexity in malignant clones bearing the same CDR3 amino acid sequence. Only in one case (SMM6) does this trend appears to be reversed. Further validation is ongoing for this case to determine whether this phenomenon may be related to the patient's specific lesion.
The clonal BCR complexity observed in the results is reflected by the number of nodes in the network generated by Dandelion, which, in turn, is driven by differences in the VDJ nucleotide sequences. This might be the result of the clonal BCR being re-exposed to new SHM events, leading to further clonal expansion. In this case, it might be possible to develop a score related to SHM events, which could correlate with the tumor burden in the patient.
However, these results are still preliminary and require more in-depth analysis. The small number of samples analyzed does not yet allow for the validation of these hypotheses. The next step will be to combine transcriptomic data with these clonal tracking efforts, linking them to the nucleotide sequences.
Dandelion can be employed not only to study MM progression but also to trace back the entire clonal history, from the asymptomatic phase to MRD and potentially to relapse. During MRD, the traditional

single-cell approach may not be sensitive enough compared to current MRD assessment methods used in clinical settings. However, alternative combined approaches can provide valuable biological insights into the residual clones that remain after treatment. One such approach could be using CITE-seq (Cellular Indexing of Transcriptomes and Epitopes by Sequencing) (with the antibody panel currently employed in EuroFlow[208,209]) on previously sorted BM B-lineage compartement. By having the clonal VDJ sequence (both amino acid and nucleotide) from the time of diagnosis, it would be possible to track the malignant clone even in MRD conditions. This would reveal significant insight by enhancing our understanding of the disease and its progression over time.

The findings of this thesis demonstrate that the research aims have been successfully accomplished, contributing valuable insights into the biology of MM and establishing a solid foundation for future investigations.

Firstly, the study confirmed the effectiveness and sensitivity of scVDJseq in identifying MM-specific VDJ rearrangements, showing strong concordance with DNA-based diagnostic methods. This underscores the potential of scVDJseq in generating valuable biological and translational insights through transcriptome-based genotyping of BCR sequences.

The comprehensive geno-transcriptomic analysis of scRNAseq data from hPCs, cPCs, pPCs, and TME has provided a nuanced understanding of the clonotypic architecture, genomic characteristics, and phenotypic distinctions between hPCs and pPCs. Moreover, the investigation into the interactions between the TME and PCs in patients has unveiled critical insights into the crosstalk driving MM progression.

The analysis of evolving VDJ sequences with the Dandelion tool could reveal critical mechanisms underlying clonal selection and treatment resistance, making it a vital aspect of studying disease dynamics and informing therapeutic strategies.

However, while these achievements mark a significant advancement, they also highlight the need for further research. Expanding the patient cohort and increasing the number of time points per patients will be crucial for a more comprehensive evaluation of MM progression and especially for the intrinsic limitations of single-cell technologies. Indeed, insufficient sampling affects the ability to detect key subpopulations, accurately profile clonal diversity, and reduces the robustness of statistical inferences, emphasizing the need for careful experimental design and computational strategies to mitigate these challenges. These limitations are further exacerbated in MRD detection context. One of the way to minimize the issue is, precisely, to increase the number of samples and expand the cohort. Additionally, employing more sophisticated tools and methodologies will enhance the

understanding of the underlying biology and potentially lead to new therapeutic strategies. In conclusion, this work not only meets the initial research goals but also sets the stage for future studies that will delve deeper into the complexities of MM to improve patient outcomes.

References

1. Stratton, M.R., Campbell, P.J., and Futreal, P.A. (2009). The cancer genome. Nature *458*, 719–724. https://doi.org/10.1038/nature07943.

2. Palumbo, A., and Anderson, K. (2011). Multiple myeloma. N Engl J Med *364*, 1046–1060. https://doi.org/10.1056/NEJMra1011442.

3. Fitzpatrick, M.J., Nardi, V., and Sohani, A.R. (2022). Plasma cell myeloma: role of histopathology, immunophenotyping, and genetic testing. Skeletal Radiol *51*, 17–30. https://doi.org/10.1007/s00256-021-03754-3.

4. Raje, N.S., Steele, D.J.R., Lawrimore, T.M., Johri, A.M., and Sohani, A.R. (2011). Case records of the Massachusetts General Hospital: Case 29-2011: A 66-year-old woman with cardiac and renal failure. N Engl J Med *365*, 1129–1138. https://doi.org/10.1056/NEJMcpc1100925.

5. Landgren, O., and Weiss, B.M. (2009). Patterns of monoclonal gammopathy of undetermined significance and multiple myeloma in various ethnic/racial groups: support for genetic factors in pathogenesis. Leukemia *23*, 1691–1697. https://doi.org/10.1038/leu.2009.134.

6. Maura, F., Bolli, N., Angelopoulos, N., Dawson, K.J., Leongamornlert, D., Martincorena, I., Mitchell, T.J., Fullam, A., Gonzalez, S., Szalat, R., et al. (2019). Genomic landscape and chronological reconstruction of driver events in multiple myeloma. Nat Commun *10*, 3835. https://doi.org/10.1038/s41467-019-11680-1.

7. Consensus recommendations for the uniform reporting of clinical trials: report of the International Myeloma Workshop Consensus Panel 1 | Blood | American Society of Hematology https://ashpublications.org/blood/article/117/18/4691/21394/Consensus-recommendations-for-the-uniform.

8. Rajkumar, S.V., Dimopoulos, M.A., Palumbo, A., Blade, J., Merlini, G., Mateos, M.-V., Kumar, S., Hillengass, J., Kastritis, E., Richardson, P., et al. (2014). International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. Lancet Oncol *15*, e538-548. https://doi.org/10.1016/S1470-2045(14)70442-5.

9. Gozzetti, A., Candi, V., Papini, G., and Bocchia, M. (2014). Therapeutic Advancements in Multiple Myeloma. Front Oncol *4*, 241. https://doi.org/10.3389/fonc.2014.00241.

10. Eslick, R., and Talaulikar, D. (2020). Multiple myeloma: From diagnosis to treatment. Australian Family Physician *42*, 684–688. https://doi.org/10.3316/informit.616996133287652.

11. Gau, Y.-C., Yeh, T.-J., Hsu, C.-M., Hsiao, S.Y., and Hsiao, H.-H. (2022). Pathogenesis and Treatment of Myeloma-Related Bone Disease. International Journal of Molecular Sciences *23*, 3112. https://doi.org/10.3390/ijms23063112.

12. Walker, B.A., Wardell, C.P., Melchor, L., Brioli, A., Johnson, D.C., Kaiser, M.F., Mirabella, F., Lopez-Corral, L., Humphray, S., Murray, L., et al. (2014). Intraclonal heterogeneity is a critical early event in the development of myeloma and precedes the development of clinical symptoms. Leukemia *28*, 384–390. https://doi.org/10.1038/leu.2013.199.

13. Bolli, N., Avet-Loiseau, H., Wedge, D.C., Van Loo, P., Alexandrov, L.B., Martincorena, I., Dawson, K.J., Iorio, F., Nik-Zainal, S., Bignell, G.R., et al. (2014). Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nat Commun *5*, 2997. https://doi.org/10.1038/ncomms3997.

14. Kyle, R.A., Remstein, E.D., Therneau, T.M., Dispenzieri, A., Kurtin, P.J., Hodnefield, J.M., Larson, D.R., Plevak, M.F., Jelinek, D.F., Fonseca, R., et al. (2007). Clinical Course and Prognosis of Smoldering (Asymptomatic) Multiple Myeloma. N Engl J Med *356*, 2582–2590. https://doi.org/10.1056/NEJMoa070389.

15. Kyle, R.A., Larson, D.R., Therneau, T.M., Dispenzieri, A., Kumar, S., Cerhan, J.R., and Rajkumar, S.V. (2018). Long-Term Follow-up of Monoclonal Gammopathy of Undetermined Significance. N Engl J Med *378*, 241–249. https://doi.org/10.1056/NEJMoa1709974.

16. Da Vià, M.C., Zicchedda, B., Maeda, A., Bagnoli, F., Perrone, G., and Bolli, N. (2020). A Journey Through Myeloma Evolution: From the Normal Plasma Cell to Disease Complexity. Hemasphere *4*, e502. https://doi.org/10.1097/HS9.0000000000000502.

17. Bonello, F., Cetani, G., Bertamini, L., Gay, F., and Larocca, A. (2019). Moving Toward Continuous Therapy in Multiple Myeloma. Clinical Hematology International *1*, 189–200. https://doi.org/10.2991/chi.d.191101.001.

18. Siegel, R.L., Miller, K.D., Fuchs, H.E., and Jemal, A. (2021). Cancer Statistics, 2021. CA Cancer J Clin *71*, 7–33. https://doi.org/10.3322/caac.21654.

19. Cancer Today https://gco.iarc.who.int/today/.

20. Kyle, R.A., Gertz, M.A., Witzig, T.E., Lust, J.A., Lacy, M.Q., Dispenzieri, A., Fonseca, R., Rajkumar, S.V., Offord, J.R., Larson, D.R., et al. (2003). Review of 1027 Patients With Newly Diagnosed Multiple Myeloma. Mayo Clinic Proceedings *78*, 21–33. https://doi.org/10.4065/78.1.21.

21. Schatz, D.G., and Swanson, P.C. (2011). V(D)J Recombination: Mechanisms of Initiation. Annual Review of Genetics *45*, 167–202. https://doi.org/10.1146/annurev-genet-110410-132552.

22. Gellert, M. (2002). V(D)J Recombination: RAG Proteins, Repair Factors, and Regulation*. Annual Review of Biochemistry *71*, 101–132. https://doi.org/10.1146/annurev.biochem.71.090501.150203.

23. Gabrielli, E., Pericolini, E., Cenci, E., Ortelli, F., Magliani, W., Ciociola, T., Bistoni, F., Conti, S., Vecchiarelli, A., and Polonelli, L. (2009). Antibody Complementarity-Determining Regions (CDRs): A Bridge between Adaptive and Innate Immunity. PLoS One *4*, e8187. https://doi.org/10.1371/journal.pone.0008187.

24. Noia, J.M.D., and Neuberger, M.S. (2007). Molecular Mechanisms of Antibody Somatic Hypermutation. Annual Review of Biochemistry *76*, 1–22. https://doi.org/10.1146/annurev.biochem.76.061705.090740.

25. Stavnezer, J., and Schrader, C.E. (2014). IgH Chain Class Switch Recombination: Mechanism and Regulation. The Journal of Immunology *193*, 5370–5378. https://doi.org/10.4049/jimmunol.1401849.

26. Okazaki, I., Yoshikawa, K., Kinoshita, K., Muramatsu, M., Nagaoka, H., and Honjo, T. (2003). Activation-Induced Cytidine Deaminase Links Class Switch Recombination and Somatic Hypermutation. Annals of the New York Academy of Sciences *987*, 1–8. https://doi.org/10.1111/j.1749-6632.2003.tb06027.x.

27. Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T. (2000). Class Switch Recombination and Hypermutation Require Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme. Cell *102*, 553–563. https://doi.org/10.1016/S0092-8674(00)00078-7.

28. Barwick, B.G., Gupta, V.A., Vertino, P.M., and Boise, L.H. (2019). Cell of Origin and Genetic Alterations in the Pathogenesis of Multiple Myeloma. Front Immunol *10*, 1121. https://doi.org/10.3389/fimmu.2019.01121.

29. González, D., van der Burg, M., García-Sanz, R., Fenton, J.A., Langerak, A.W., González, M., van Dongen, J.J.M., San Miguel, J.F., and Morgan, G.J. (2007). Immunoglobulin gene rearrangements and the pathogenesis of multiple myeloma. Blood *110*, 3112–3121. https://doi.org/10.1182/blood-2007-02-069625.

30. Maura, F., and Bergsagel, P.L. (2024). Molecular Pathogenesis of Multiple Myeloma: Clinical Implications. Hematology/Oncology Clinics of North America *38*, 267–279. https://doi.org/10.1016/j.hoc.2023.12.010.

31. Omori, S.A., Cato, M.H., Anzelon-Mills, A., Puri, K.D., Shapiro-Shelef, M., Calame, K., and Rickert, R.C. (2006). Regulation of Class-Switch Recombination and Plasma Cell Differentiation by Phosphatidylinositol 3-Kinase Signaling. Immunity *25*, 545–557. https://doi.org/10.1016/j.immuni.2006.08.015.

32. Bolli, N., Martinelli, G., and Cerchione, C. (2020). The molecular pathogenesis of multiple myeloma. Hematol Rep *12*, 9054. https://doi.org/10.4081/hr.2020.9054.

33. Fonseca, R., Barlogie, B., Bataille, R., Bastard, C., Bergsagel, P.L., Chesi, M., Davies, F.E., Drach, J., Greipp, P.R., Kirsch, I.R., et al. (2004). Genetics and Cytogenetics of Multiple Myeloma: A Workshop Report. Cancer Research *64*, 1546–1558. https://doi.org/10.1158/0008-5472.CAN-03-2876.

34. Bergsagel, P.L., Chesi, M., Nardini, E., Brents, L.A., Kirby, S.L., and Kuehl, W.M. (1996). Promiscuous translocations into immunoglobulin heavy chain switch regions in multiple myeloma. Proceedings of the National Academy of Sciences *93*, 13931–13936. https://doi.org/10.1073/pnas.93.24.13931.

35. Brigle, K., and Rogers, B. (2017). Pathobiology and Diagnosis of Multiple Myeloma. Semin Oncol Nurs *33*, 225–236. https://doi.org/10.1016/j.soncn.2017.05.012.

36. Avet-Loiseau, H., Li, C., Magrangeas, F., Gouraud, W., Charbonnel, C., Harousseau, J.-L., Attal, M., Marit, G., Mathiot, C., Facon, T., et al. (2009). Prognostic Significance of Copy-

Number Alterations in Multiple Myeloma. JCO *27*, 4585–4590. https://doi.org/10.1200/JCO.2008.20.6136.

37. Kumar, S., Fonseca, R., Ketterling, R.P., Dispenzieri, A., Lacy, M.Q., Gertz, M.A., Hayman, S.R., Buadi, F.K., Dingli, D., Knudson, R.A., et al. (2012). Trisomies in multiple myeloma: impact on survival in patients with high-risk cytogenetics. Blood *119*, 2100–2105. https://doi.org/10.1182/blood-2011-11-390658.

38. Tian, E., Sawyer, J.R., Heuck, C.J., Zhang, Q., van Rhee, F., Barlogie, B., and Epstein, J. (2014). In Multiple Myeloma, 14q32 Translocations Are Non-Random Chromosomal Fusions Driving High Expression Levels of the Respective Partner Genes. Genes Chromosomes Cancer *53*, 549–557. https://doi.org/10.1002/gcc.22165.

39. Keats, J.J., Chesi, M., Egan, J.B., Garbitt, V.M., Palmer, S.E., Braggio, E., Van Wier, S., Blackburn, P.R., Baker, A.S., Dispenzieri, A., et al. (2012). Clonal competition with alternating dominance in multiple myeloma. Blood *120*, 1067–1076. https://doi.org/10.1182/blood-2012-01-405985.

40. Morgan, G.J., Walker, B.A., and Davies, F.E. (2012). The genetic architecture of multiple myeloma. Nat Rev Cancer *12*, 335–348. https://doi.org/10.1038/nrc3257.

41. Kuehl, W.M., and Bergsagel, P.L. (2005). Early Genetic Events Provide the Basis for a Clinical Classification of Multiple Myeloma. Hematology *2005*, 346–352. https://doi.org/10.1182/asheducation-2005.1.346.

42. Walker, B.A., Wardell, C.P., Johnson, D.C., Kaiser, M.F., Begum, D.B., Dahir, N.B., Ross, F.M., Davies, F.E., Gonzalez, D., and Morgan, G.J. (2013). Characterization of *IGH* locus breakpoints in multiple myeloma indicates a subset of translocations appear to occur in pregerminal center B cells. Blood *121*, 3413–3419. https://doi.org/10.1182/blood-2012-12-471888.

43. Avet-Loiseau, H., Li, J.Y., Facon, T., Brigaudeau, C., Morineau, N., Maloisel, F., Rapp, M.J., Talmant, P., Trimoreau, F., Jaccard, A., et al. (1998). High incidence of translocations t(11;14)(q13;q32) and t(4;14)(p16;q32) in patients with plasma cell malignancies. Cancer Res *58*, 5640–5645.

44. Chesi, M., Nardini, E., Lim, R.S.C., Smith, K.D., Kuehl, W.M., and Bergsagel, P.L. (1998). The t(4;14) Translocation in Myeloma Dysregulates Both *FGFR3*and a Novel Gene, *MMSET*, Resulting in IgH/MMSET Hybrid Transcripts. Blood *92*, 3025–3034. https://doi.org/10.1182/blood.V92.9.3025.

45. Chesi, M., Bergsagel, P.L., Shonukan, O.O., Martelli, M.L., Brents, L.A., Chen, T., Schröck, E., Ried, T., and Kuehl, W.M. (1998). Frequent Dysregulation of the c-*maf* Proto-Oncogene at 16q23 by Translocation to an Ig Locus in Multiple Myeloma. Blood *91*, 4457–4463. https://doi.org/10.1182/blood.V91.12.4457.

46. Walker, B.A., Wardell, C.P., Brioli, A., Boyle, E., Kaiser, M.F., Begum, D.B., Dahir, N.B., Johnson, D.C., Ross, F.M., Davies, F.E., et al. (2014). Translocations at 8q24 juxtapose MYC with genes that harbor superenhancers resulting in overexpression and poor prognosis in myeloma patients. Blood Cancer Journal *4*, e191–e191. https://doi.org/10.1038/bcj.2014.13.

47. Barwick, B.G., Neri, P., Bahlis, N.J., Nooka, A.K., Dhodapkar, M.V., Jaye, D.L., Hofmeister, C.C., Kaufman, J.L., Gupta, V.A., Auclair, D., et al. (2019). Multiple myeloma immunoglobulin lambda translocations portend poor prognosis. Nat Commun *10*, 1911. https://doi.org/10.1038/s41467-019-09555-6.

48. Avet-Loiseau, H., Attal, M., Moreau, P., Charbonnel, C., Garban, F., Hulin, C., Leyvraz, S., Michallet, M., Yakoub-Agha, I., Garderet, L., et al. (2007). Genetic abnormalities and survival in multiple myeloma: the experience of the Intergroupe Francophone du Myélome. Blood *109*, 3489–3495. https://doi.org/10.1182/blood-2006-08-040410.

49. Hanamura, I., Stewart, J.P., Huang, Y., Zhan, F., Santra, M., Sawyer, J.R., Hollmig, K., Zangarri, M., Pineda-Roman, M., van Rhee, F., et al. (2006). Frequent gain of chromosome band 1q21 in plasma-cell dyscrasias detected by fluorescence in situ hybridization: incidence increases from MGUS to relapsed myeloma and is related to prognosis and disease progression following tandem stem-cell transplantation. Blood *108*, 1724–1732. https://doi.org/10.1182/blood-2006-03-009910.

50. Saitoh, T., and Oda, T. (2021). DNA Damage Response in Multiple Myeloma: The Role of the Tumor Microenvironment. Cancers *13*, 504. https://doi.org/10.3390/cancers13030504.

51. Sottoriva, A., Kang, H., Ma, Z., Graham, T.A., Salomon, M.P., Zhao, J., Marjoram, P., Siegmund, K., Press, M.F., Shibata, D., et al. (2015). A Big Bang model of human colorectal tumor growth. Nat Genet *47*, 209–216. https://doi.org/10.1038/ng.3214.

52. Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A.J.R., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Børresen-Dale, A.-L., et al. (2013). Signatures of mutational processes in human cancer. Nature *500*, 415–421. https://doi.org/10.1038/nature12477.

53. Rasche, L., Chavan, S.S., Stephens, O.W., Patel, P.H., Tytarenko, R., Ashby, C., Bauer, M., Stein, C., Deshpande, S., Wardell, C., et al. (2017). Spatial genomic heterogeneity in multiple myeloma revealed by multi-region sequencing. Nat Commun *8*, 268. https://doi.org/10.1038/s41467-017-00296-y.

54. Moreau, P., and Minvielle, S. (2013). Multiple myeloma: so much progress, but so many unsolved questions. Haematologica *98*, 487–489. https://doi.org/10.3324/haematol.2013.083592.

55. Egan, J.B., Shi, C.-X., Tembe, W., Christoforides, A., Kurdoglu, A., Sinari, S., Middha, S., Asmann, Y., Schmidt, J., Braggio, E., et al. (2012). Whole-genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution, and clonal tides. Blood *120*, 1060–1066. https://doi.org/10.1182/blood-2012-01-405977.

56. Ledergor, G., Weiner, A., Zada, M., Wang, S.-Y., Cohen, Y.C., Gatt, M.E., Snir, N., Magen, H., Koren-Michowitz, M., Herzog-Tzarfati, K., et al. (2018). Single cell dissection of plasma cell heterogeneity in symptomatic and asymptomatic myeloma. Nat Med *24*, 1867–1876. https://doi.org/10.1038/s41591-018-0269-2.

57. Bolli, N., Maura, F., Minvielle, S., Gloznik, D., Szalat, R., Fullam, A., Martincorena, I., Dawson, K.J., Samur, M.K., Zamora, J., et al. (2018). Genomic patterns of progression in smoldering multiple myeloma. Nat Commun *9*, 3363. https://doi.org/10.1038/s41467-018-05058-y.

58. Maura, F., Rustad, E.H., Boyle, E.M., and Morgan, G.J. (2020). Reconstructing the evolutionary history of multiple myeloma. Best Pract Res Clin Haematol *33*, 101145. https://doi.org/10.1016/j.beha.2020.101145.

59. Robiou du Pont, S., Cleynen, A., Fontan, C., Attal, M., Munshi, N., Corre, J., and Avet-Loiseau, H. (2017). Genomics of Multiple Myeloma. JCO *35*, 963–967. https://doi.org/10.1200/JCO.2016.70.6705.

60. Landau, H.J., Yellapantula, V., Diamond, B.T., Rustad, E.H., Maclachlan, K.H., Gundem, G., Medina-Martinez, J., Ossa, J.A., Levine, M.F., Zhou, Y., et al. (2020). Accelerated single cell seeding in relapsed multiple myeloma. Nat Commun *11*, 3617. https://doi.org/10.1038/s41467-020-17459-z.

61. Broyl, A., Hose, D., Lokhorst, H., de Knegt, Y., Peeters, J., Jauch, A., Bertsch, U., Buijs, A., Stevens-Kroef, M., Beverloo, H.B., et al. (2010). Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. Blood *116*, 2543–2553. https://doi.org/10.1182/blood-2009-12-261032.

62. Chapman, M.A., Lawrence, M.S., Keats, J.J., Cibulskis, K., Sougnez, C., Schinzel, A.C., Harview, C.L., Brunet, J.-P., Ahmann, G.J., Adli, M., et al. (2011). Initial genome sequencing and analysis of multiple myeloma. Nature *471*, 467–472. https://doi.org/10.1038/nature09837.

63. Lohr, J.G., Stojanov, P., Carter, S.L., Cruz-Gordillo, P., Lawrence, M.S., Auclair, D., Sougnez, C., Knoechel, B., Gould, J., Saksena, G., et al. (2014). Widespread Genetic Heterogeneity in Multiple Myeloma: Implications for Targeted Therapy. Cancer Cell *25*, 91–101. https://doi.org/10.1016/j.ccr.2013.12.015.

64. Bolli, N., Avet-Loiseau, H., Wedge, D.C., Van Loo, P., Alexandrov, L.B., Martincorena, I., Dawson, K.J., Iorio, F., Nik-Zainal, S., Bignell, G.R., et al. (2014). Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nature Communications *5*, 2997. https://doi.org/10.1038/ncomms3997.

65. Naik, S., Larsen, S.B., Cowley, C.J., and Fuchs, E. (2018). Two to Tango: Dialog between Immunity and Stem Cells in Health and Disease. Cell *175*, 908–920. https://doi.org/10.1016/j.cell.2018.08.071.

66. de Jong, M.M.E., Kellermayer, Z., Papazian, N., Tahri, S., Hofste op Bruinink, D., Hoogenboezem, R., Sanders, M.A., van de Woestijne, P.C., Bos, P.K., Khandanpour, C., et al. (2021). The multiple myeloma microenvironment is defined by an inflammatory stromal cell landscape. Nat Immunol *22*, 769–780. https://doi.org/10.1038/s41590-021-00931-3.

67. Kawano, Y., Moschetta, M., Manier, S., Glavey, S., Görgün, G.T., Roccaro, A.M., Anderson, K.C., and Ghobrial, I.M. (2015). Targeting the bone marrow microenvironment in multiple myeloma. Immunol Rev *263*, 160–172. https://doi.org/10.1111/imr.12233.

68. Fujisaki, J., Wu, J., Carlson, A.L., Silberstein, L., Putheti, P., Larocca, R., Gao, W., Saito, T.I., Celso, C.L., Tsuyuzaki, H., et al. (2011). In vivo imaging of Treg cells providing immune privilege to the haematopoietic stem-cell niche. Nature *474*, 216–219. https://doi.org/10.1038/nature10160.

69. Paiva, B., Paino, T., Sayagues, J.-M., Garayoa, M., San-Segundo, L., Martín, M., Mota, I., Sanchez, M.-L., Bárcena, P., Aires-Mejia, I., et al. (2013). Detailed characterization of multiple myeloma circulating tumor cells shows unique phenotypic, cytogenetic, functional, and circadian distribution profile. Blood *122*, 3591–3598. https://doi.org/10.1182/blood-2013-06-510453.

70. Solimando, A.G., Da Via', M.C., Borrelli, P., Leone, P., Di Lernia, G., Tabares Gaviria, P., Brandl, A., Pedone, G.L., Rauert-Wunderlich, H., Lapa, C., et al. (2018). Central Function for JAM-a in Multiple Myeloma Patients with Extramedullary Disease. Blood *132*, 4455. https://doi.org/10.1182/blood-2018-99-117021.

71. Forster, S., and Radpour, R. (2022). Molecular Impact of the Tumor Microenvironment on Multiple Myeloma Dissemination and Extramedullary Disease. Front. Oncol. *12*, 941437. https://doi.org/10.3389/fonc.2022.941437.

72. Manier, S., Sacco, A., Leleu, X., Ghobrial, I.M., and Roccaro, A.M. (2012). Bone Marrow Microenvironment in Multiple Myeloma Progression. J Biomed Biotechnol *2012*, 157496. https://doi.org/10.1155/2012/157496.

73. Moser-Katz, T., Joseph, N.S., Dhodapkar, M.V., Lee, K.P., and Boise, L.H. (2021). Game of Bones: How Myeloma Manipulates Its Microenvironment. Front Oncol *10*, 625199. https://doi.org/10.3389/fonc.2020.625199.

74. An, G., Acharya, C., Feng, X., Wen, K., Zhong, M., Zhang, L., Munshi, N.C., Qiu, L., Tai, Y.-T., and Anderson, K.C. (2016). Osteoclasts promote immune suppressive microenvironment in multiple myeloma: therapeutic implication. Blood *128*, 1590–1603. https://doi.org/10.1182/blood-2016-03-707547.

75. Dunn, G.P., Bruce, A.T., Ikeda, H., Old, L.J., and Schreiber, R.D. (2002). Cancer immunoediting: from immunosurveillance to tumor escape. Nat Immunol *3*, 991–998. https://doi.org/10.1038/ni1102-991.

76. Pérez-Persona, E., Vidriales, M.-B., Mateo, G., García-Sanz, R., Mateos, M.-V., de Coca, A.G., Galende, J., Martín-Nuñez, G., Alonso, J.M., de las Heras, N., et al. (2007). New criteria to identify risk of progression in monoclonal gammopathy of uncertain significance and smoldering multiple myeloma based on multiparameter flow cytometry analysis of bone marrow plasma cells. Blood *110*, 2586–2592. https://doi.org/10.1182/blood-2007-05-088443.

77. Cherry, B.M., Costello, R., Zingone, A., Burris, J., Korde, N., Manasanch, E., Kwok, M., Annunziata, C., Roschewski, M.J., Engels, E.A., et al. (2013). Immunoparesis and monoclonal gammopathy of undetermined significance are disassociated in advanced age. Am J Hematol *88*, 89–92. https://doi.org/10.1002/ajh.23355.

78. Landgren, O., Hofmann, J.N., McShane, C.M., Santo, L., Hultcrantz, M., Korde, N., Mailankody, S., Kazandjian, D., Murata, K., Thoren, K., et al. (2019). Association of Immune Marker Changes With Progression of Monoclonal Gammopathy of Undetermined Significance to Multiple Myeloma. JAMA Oncology *5*, 1293–1301. https://doi.org/10.1001/jamaoncol.2019.1568.

79. Pratt, G., Goodyear, O., and Moss, P. (2007). Immunodeficiency and immunotherapy in multiple myeloma. British Journal of Haematology *138*, 563–579. https://doi.org/10.1111/j.1365-2141.2007.06705.x.

80. Dhodapkar, M.V., Geller, M.D., Chang, D.H., Shimizu, K., Fujii, S.-I., Dhodapkar, K.M., and Krasovsky, J. (2003). A Reversible Defect in Natural Killer T Cell Function Characterizes the Progression of Premalignant to Malignant Multiple Myeloma. J Exp Med *197*, 1667–1676. https://doi.org/10.1084/jem.20021650.

81. Kumar, S., Paiva, B., Anderson, K.C., Durie, B., Landgren, O., Moreau, P., Munshi, N., Lonial, S., Bladé, J., Mateos, M.-V., et al. (2016). International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. The Lancet Oncology *17*, e328–e346. https://doi.org/10.1016/S1470-2045(16)30206-6.

82. Paiva, B., Vidriales, M.-B., Pérez, J.J., Mateo, G., Montalbán, M.A., Mateos, M.V., Bladé, J., Lahuerta, J.J., Orfao, A., San Miguel, J.F., et al. (2009). Multiparameter flow cytometry quantification of bone marrow plasma cells at diagnosis provides more prognostic information than morphological assessment in myeloma patients. Haematologica *94*, 1599–1602. https://doi.org/10.3324/haematol.2009.009100.

83. Garderet, L., D'Souza, A., Jacobs, P., van Biezen, A., Schönland, S., Kroeger, N., Morris, C., and Hari, P. (2017). Response Assessment in Myeloma: Practical Manual on Consistent Reporting in an Era of Dramatic Therapeutic Advances. Biology of Blood and Marrow Transplantation *23*, 1193–1202. https://doi.org/10.1016/j.bbmt.2017.03.009.

84. Medina, A., Puig, N., Flores-Montero, J., Jimenez, C., Sarasquete, M.-E., Garcia-Alvarez, M., Prieto-Conde, I., Chillon, C., Alcoceba, M., Gutierrez, N.C., et al. (2020). Comparison of next-generation sequencing (NGS) and next-generation flow (NGF) for minimal residual disease (MRD) assessment in multiple myeloma. Blood Cancer J. *10*, 1–10. https://doi.org/10.1038/s41408-020-00377-0.

85. Perrot, A., Lauwers-Cances, V., Corre, J., Robillard, N., Hulin, C., Chretien, M.-L., Dejoie, T., Maheo, S., Stoppa, A.-M., Pegourie, B., et al. (2018). Minimal residual disease negativity using deep sequencing is a major prognostic factor in multiple myeloma. Blood *132*, 2456–2464. https://doi.org/10.1182/blood-2018-06-858613.

86. Flores-Montero, J., Sanoja-Flores, L., Paiva, B., Puig, N., García-Sánchez, O., Böttcher, S., van der Velden, V.H.J., Pérez-Morán, J.-J., Vidriales, M.-B., García-Sanz, R., et al. (2017). Next Generation Flow for highly sensitive and standardized detection of minimal residual disease in multiple myeloma. Leukemia *31*, 2094–2103. https://doi.org/10.1038/leu.2017.29.

87. Oberle, A., Brandt, A., Alawi, M., Langebrake, C., Janjetovic, S., Wolschke, C., Schütze, K., Bannas, P., Kröger, N., Koch-Nolte, F., et al. (2017). Long-term CD38 saturation by daratumumab interferes with diagnostic myeloma cell detection. Haematologica *102*, e368–e370. https://doi.org/10.3324/haematol.2017.169235.

88. Gozzetti, A., Raspadori, D., Bacchiarri, F., Sicuranza, A., Pacelli, P., Ferrigno, I., Tocci, D., and Bocchia, M. (2020). Minimal Residual Disease in Multiple Myeloma: State of the Art and Applications in Clinical Practice. Journal of Personalized Medicine *10*, 120. https://doi.org/10.3390/jpm10030120.

89. Costa, L.J., Derman, B.A., Bal, S., Sidana, S., Chhabra, S., Silbermann, R., Ye, J.C., Cook, G., Cornell, R.F., Holstein, S.A., et al. (2021). International harmonization in performing and reporting minimal residual disease assessment in multiple myeloma trials. Leukemia *35*, 18–30. https://doi.org/10.1038/s41375-020-01012-4.

90. Goicoechea, I., Puig, N., Cedena, M.-T., Burgos, L., Cordón, L., Vidriales, M.-B., Flores-Montero, J., Gutierrez, N.C., Calasanz, M.-J., Ramos, M.-L.M., et al. (2021). Deep MRD profiling defines outcome and unveils different modes of treatment resistance in standard- and high-risk myeloma. Blood *137*, 49–60. https://doi.org/10.1182/blood.2020006731.

91. Stark, R., Grzelak, M., and Hadfield, J. (2019). RNA sequencing: the teenage years. Nat Rev Genet *20*, 631–656. https://doi.org/10.1038/s41576-019-0150-2.

92. Weber, A.P.M. (2015). Discovering New Biology through Sequencing of RNA. Plant Physiology *169*, 1524–1531. https://doi.org/10.1104/pp.15.01081.

93. Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods *6*, 377–382. https://doi.org/10.1038/nmeth.1315.

94. Svensson, V., Vento-Tormo, R., and Teichmann, S.A. (2018). Exponential scaling of single-cell RNA-seq in the past decade. Nat Protoc *13*, 599–604. https://doi.org/10.1038/nprot.2017.149.

95. Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. Nat Commun *8*, 14049. https://doi.org/10.1038/ncomms14049.

96. Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189–196. https://doi.org/10.1126/science.aad0501.

97. Patel, A.P., Tirosh, I., Trombetta, J.J., Shalek, A.K., Gillespie, S.M., Wakimoto, H., Cahill, D.P., Nahed, B.V., Curry, W.T., Martuza, R.L., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science *344*, 1396–1401. https://doi.org/10.1126/science.1254257.

98. Azizi, E., Carr, A.J., Plitas, G., Cornish, A.E., Konopacki, C., Prabhakaran, S., Nainys, J., Wu, K., Kiseliovas, V., Setty, M., et al. (2018). Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. Cell *174*, 1293-1308.e36. https://doi.org/10.1016/j.cell.2018.05.060.

99. Bowes, A.L., Tarabichi, M., Pillay, N., and Van Loo, P. (2022). Leveraging single-cell sequencing to unravel intratumour heterogeneity and tumour evolution in human cancers. J Pathol *257*, 466–478. https://doi.org/10.1002/path.5914.

100. Yeo, A.T., Rawal, S., Delcuze, B., Christofides, A., Atayde, A., Strauss, L., Balaj, L., Rogers, V.A., Uhlmann, E.J., Varma, H., et al. (2022). Single-cell RNA sequencing reveals evolution of immune landscape during glioblastoma progression. Nat Immunol *23*, 971–984. https://doi.org/10.1038/s41590-022-01215-0.

101. Liu, R., Gao, Q., Foltz, S.M., Fowles, J.S., Yao, L., Wang, J.T., Cao, S., Sun, H., Wendl, M.C., Sethuraman, S., et al. (2021). Co-evolution of tumor and immune cells during progression of multiple myeloma. Nat Commun *12*, 2559. https://doi.org/10.1038/s41467-021-22804-x.

102. Boiarsky, R., Haradhvala, N.J., Alberge, J.-B., Sklavenitis-Pistofidis, R., Mouhieddine, T.H., Zavidij, O., Shih, M.-C., Firer, D., Miller, M., El-Khoury, H., et al. (2022). Single cell characterization of myeloma and its precursor conditions reveals transcriptional signatures of early tumorigenesis. Nat Commun *13*, 7040. https://doi.org/10.1038/s41467-022-33944-z.

103. Zavidij, O., Haradhvala, N.J., Mouhieddine, T.H., Sklavenitis-Pistofidis, R., Cai, S., Reidy, M., Rahmat, M., Flaifel, A., Ferland, B., Su, N.K., et al. (2020). Single-cell RNA sequencing reveals compromised immune microenvironment in precursor stages of multiple myeloma. Nat Cancer *1*, 493–506. https://doi.org/10.1038/s43018-020-0053-3.

104. Tanner, S.D., Bandura, D.R., Ornatsky, O., Baranov, V.I., Nitz, M., and Winnik, M.A. (2008). Flow cytometer with mass spectrometer detection for massively multiplexed single-cell biomarker assay. Pure and Applied Chemistry *80*, 2627–2641. https://doi.org/10.1351/pac200880122627.

105. Pilcher, W., Thomas, B.E., Bhasin, S.S., Jayasinghe, R.G., Yao, L., Gonzalez-Kozlova, E., Dasari, S., Kim-Schulze, S., Rahman, A., Patton, J., et al. (2023). Cross center single-cell RNA sequencing study of the immune microenvironment in rapid progressing multiple myeloma. NPJ Genom Med *8*, 3. https://doi.org/10.1038/s41525-022-00340-x.

106. Rustad, E.H., Misund, K., Bernard, E., Coward, E., Yellapantula, V.D., Hultcrantz, M., Ho, C., Kazandjian, D., Korde, N., Mailankody, S., et al. (2019). Stability and uniqueness of clonal immunoglobulin CDR3 sequences for MRD tracking in multiple myeloma. Am J Hematol *94*, 1364–1373. https://doi.org/10.1002/ajh.25641.

107. Puig, N., Conde, I., Jiménez, C., Sarasquete, M.E., Balanzategui, A., Alcoceba, M., Quintero, J., Chillón, M.C., Sebastián, E., Corral, R., et al. (2015). The predominant myeloma clone at diagnosis, CDR3 defined, is constantly detectable across all stages of disease evolution. Leukemia *29*, 1435–1437. https://doi.org/10.1038/leu.2015.7.

108. Frede, J., Anand, P., Sotudeh, N., Pinto, R.A., Nair, M.S., Stuart, H., Yee, A.J., Vijaykumar, T., Waldschmidt, J.M., Potdar, S., et al. (2021). Dynamic transcriptional reprogramming leads to immunotherapeutic vulnerabilities in myeloma. Nat Cell Biol *23*, 1199–1211. https://doi.org/10.1038/s41556-021-00766-y.

109. Rashid, N.U., Sperling, A.S., Bolli, N., Wedge, D.C., Van Loo, P., Tai, Y.-T., Shammas, M.A., Fulciniti, M., Samur, M.K., Richardson, P.G., et al. (2014). Differential and limited expression of mutant alleles in multiple myeloma. Blood *124*, 3110–3117. https://doi.org/10.1182/blood-2014-04-569327.

110. Brüggemann, M., Kotrová, M., Knecht, H., Bartram, J., Boudjogrha, M., Bystry, V., Fazio, G., Froňková, E., Giraud, M., Grioni, A., et al. (2019). Standardized next-generation sequencing of immunoglobulin and T-cell receptor gene recombinations for MRD marker identification in acute lymphoblastic leukaemia; a EuroClonality-NGS validation study. Leukemia *33*, 2241–2253. https://doi.org/10.1038/s41375-019-0496-7.

111. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation *2*, 100141. https://doi.org/10.1016/j.xinn.2021.100141.

112. Dang, M., Han, G., Lee, H.C., Patel, K.K., Becnel, M.R., Wang, R., Thomas, S.K., Hao, D., Chu, Y., Weber, D.M., et al. (2023). Single cell clonotypic and transcriptional evolution of multiple myeloma precursor disease. Cancer Cell *41*, 1032-1047.e4. https://doi.org/10.1016/j.ccell.2023.05.007.

113. Bianchi, G., and Munshi, N.C. (2015). Pathogenesis beyond the cancer clone(s) in multiple myeloma. Blood *125*, 3049–3058. https://doi.org/10.1182/blood-2014-11-568881.

114. Botta, C., Perez, C., Larrayoz, M., Puig, N., Cedena, M.-T., Termini, R., Goicoechea, I., Rodriguez, S., Zabaleta, A., Lopez, A., et al. (2023). Large T cell clones expressing immune checkpoints increase during multiple myeloma evolution and predict treatment resistance. Nat Commun *14*, 5825. https://doi.org/10.1038/s41467-023-41562-6.

115. Hideshima, T., Mitsiades, C., Tonon, G., Richardson, P.G., and Anderson, K.C. (2007). Understanding multiple myeloma pathogenesis in the bone marrow to identify new therapeutic targets. Nat Rev Cancer *7*, 585–598. https://doi.org/10.1038/nrc2189.

116. de Jong, M.M.E., Fokkema, C., Papazian, N., Czeti, Á., Appelman, M.K., Vermeulen, M., van Heusden, T., Hoogenboezem, R.M., van Beek, G., Tahri, S., et al. (2024). An IL-1β-driven neutrophil–stromal cell axis fosters a BAFF-rich protumor microenvironment in individuals with multiple myeloma. Nat Immunol *25*, 820–833. https://doi.org/10.1038/s41590-024-01808-x.

117. Solimando, A.G., Vià, M.C.D., Leone, P., Borrelli, P., Croci, G.A., Tabares, P., Brandl, A., Lernia, G.D., Bianchi, F.P., Tafuri, S., et al. (2021). Halting the vicious cycle within the multiple myeloma ecosystem: blocking JAM-A on bone marrow endothelial cells restores angiogenic homeostasis and suppresses tumor progression. Haematologica *106*, 1943–1956. https://doi.org/10.3324/haematol.2019.239913.

118. Nakamura, K., Smyth, M.J., and Martinet, L. (2020). Cancer immunoediting and immune dysregulation in multiple myeloma. Blood *136*, 2731–2740. https://doi.org/10.1182/blood.2020006540.

119. Duan, M., Nguyen, D.C., Joyner, C.J., Saney, C.L., Tipton, C.M., Andrews, J., Lonial, S., Kim, C., Hentenaar, I., Kosters, A., et al. (2023). Understanding heterogeneity of human bone marrow plasma cell maturation and survival pathways by single-cell analyses. Cell Reports *42*. https://doi.org/10.1016/j.celrep.2023.112682.

120. Duez, M., Giraud, M., Herbert, R., Rocher, T., Salson, M., and Thonier, F. (2016). Vidjil: A Web Platform for Analysis of High-Throughput Repertoire Sequencing. PLoS One *11*, e0166126. https://doi.org/10.1371/journal.pone.0166126.

121. Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat Biotechnol *33*, 495–502. https://doi.org/10.1038/nbt.3192.

122. McCarthy, D.J., Campbell, K.R., Lun, A.T.L., and Wills, Q.F. (2017). Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics *33*, 1179–1186. https://doi.org/10.1093/bioinformatics/btw777.

123. Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biology *19*, 15. https://doi.org/10.1186/s13059-017-1382-0.

124. Stegle, O., Teichmann, S.A., and Marioni, J.C. (2015). Computational and analytical challenges in single-cell transcriptomics. Nat Rev Genet *16*, 133–145. https://doi.org/10.1038/nrg3833.

125. Ilicic, T., Kim, J.K., Kolodziejczyk, A.A., Bagger, F.O., McCarthy, D.J., Marioni, J.C., and Teichmann, S.A. (2016). Classification of low quality cells from single-cell RNA-seq data. Genome Biology *17*, 29. https://doi.org/10.1186/s13059-016-0888-1.

126. AlJanahi, A.A., Danielsen, M., and Dunbar, C.E. (2018). An Introduction to the Analysis of Single-Cell RNA-Sequencing Data. Molecular Therapy - Methods & Clinical Development *10*, 189–196. https://doi.org/10.1016/j.omtm.2018.07.003.

127. Luecken, M.D., and Theis, F.J. (2019). Current best practices in single-cell RNA-seq analysis: a tutorial. Molecular Systems Biology *15*, e8746. https://doi.org/10.15252/msb.20188746.

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

129. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods *16*, 1289–1296. https://doi.org/10.1038/s41592-019-0619-0.

130. Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2019). Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol *37*, 38–44. https://doi.org/10.1038/nbt.4314.

131. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573-3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

132. Suo, C., Polanski, K., Dann, E., Lindeboom, R.G.H., Vilarrasa-Blasi, R., Vento-Tormo, R., Haniffa, M., Meyer, K.B., Dratva, L.M., Tuong, Z.K., et al. (2024). Dandelion uses the single-cell adaptive immune receptor repertoire to explore lymphocyte developmental origins. Nat Biotechnol *42*, 40–51. https://doi.org/10.1038/s41587-023-01734-7.

133. Gupta, N.T., Vander Heiden, J.A., Uduman, M., Gadala-Maria, D., Yaari, G., and Kleinstein, S.H. (2015). Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. Bioinformatics *31*, 3356–3358. https://doi.org/10.1093/bioinformatics/btv359.

134. Foltz, S.M., Gao, Q., Yoon, C.J., Sun, H., Yao, L., Li, Y., Jayasinghe, R.G., Cao, S., King, J., Kohnen, D.R., et al. (2020). Evolution and structure of clinically relevant gene fusions in multiple myeloma. Nat Commun *11*, 2666. https://doi.org/10.1038/s41467-020-16434-y.

135. Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics *19*, 477. https://doi.org/10.1186/s12864-018-4772-0.

136. Browaeys, R., Gilis, J., Sang-Aram, C., Bleser, P.D., Hoste, L., Tavernier, S., Lambrechts, D., Seurinck, R., and Saeys, Y. (2023). MultiNicheNet: a flexible framework for differential cell-cell communication analysis from multi-sample multi-condition single-cell transcriptomics data. Preprint at bioRxiv, https://doi.org/10.1101/2023.06.13.544751 https://doi.org/10.1101/2023.06.13.544751.

137. Stephenson, E., Reynolds, G., Botting, R.A., Calero-Nieto, F.J., Morgan, M.D., Tuong, Z.K., Bach, K., Sungnak, W., Worlock, K.B., Yoshida, M., et al. (2021). Single-cell multi-omics analysis of the immune response in COVID-19. Nat Med *27*, 904–916. https://doi.org/10.1038/s41591-021-01329-2.

138. Gadala-Maria, D., Yaari, G., Uduman, M., and Kleinstein, S.H. (2015). Automated analysis of high-throughput B-cell sequencing data reveals a high frequency of novel immunoglobulin V gene segment alleles. Proceedings of the National Academy of Sciences *112*, E862–E870. https://doi.org/10.1073/pnas.1417683112.

139. Lefranc, M.-P., Giudicelli, V., Ginestoux, C., Bodmer, J., Müller, W., Bontrop, R., Lemaitre, M., Malik, A., Barbié, V., and Chaume, D. IMGT, the international ImMunoGeneTics database.

140. Harrison, P.W., Amode, M.R., Austine-Orimoloye, O., Azov, A.G., Barba, M., Barnes, I., Becker, A., Bennett, R., Berry, A., Bhai, J., et al. (2024). Ensembl 2024. Nucleic Acids Research *52*, D891–D899. https://doi.org/10.1093/nar/gkad1049.

141. broadinstitute/infercnv (2024). (Broad Institute).

142. Jung, J., Choe, J., Li, L., and Choi, Y.S. (2000). Regulation of CD27 expression in the course of germinal center B cell differentiation: the pivotal role of IL-10. European Journal of Immunology *30*, 2437–2443. https://doi.org/10.1002/1521-4141(2000)30:8<2437::AID-IMMU2437>3.0.CO;2-M.

143. Lee, H., Ahn, S., Maity, R., Leblay, N., Ziccheddu, B., Truger, M., Chojnacka, M., Cirrincione, A., Durante, M., Tilmont, R., et al. (2023). Mechanisms of antigen escape from BCMA- or GPRC5D-targeted immunotherapies in multiple myeloma. Nat Med *29*, 2295–2306. https://doi.org/10.1038/s41591-023-02491-5.

144. Hoefner, C., Bryde, T.H., Pihl, C., Tiedemann, S.N., Bresson, S.E., Hotiana, H.A., Khilji, M.S., Santos, T.D., Puglia, M., Pisano, P., et al. (2023). FK506-Binding Protein 2 Participates in Proinsulin Folding. Biomolecules *13*, 152. https://doi.org/10.3390/biom13010152.

145. Liu, Y.-X., Wang, J., Guo, J., Wu, J., Lieberman, H.B., and Yin, Y. (2008). DUSP1 Is Controlled by p53 during the Cellular Response to Oxidative Stress. Molecular Cancer Research *6*, 624–633. https://doi.org/10.1158/1541-7786.MCR-07-2019.

146. Chen, H.-F., Chuang, H.-C., and Tan, T.-H. (2019). Regulation of Dual-Specificity Phosphatase (DUSP) Ubiquitination and Protein Stability. International Journal of Molecular Sciences *20*, 2668. https://doi.org/10.3390/ijms20112668.

147. Strippoli, R., Niayesh-Mehr, R., Adelipour, M., Khosravi, A., Cordani, M., Zarrabi, A., and Allameh, A. (2024). Contribution of Autophagy to Epithelial Mesenchymal Transition Induction during Cancer Progression. Cancers *16*, 807. https://doi.org/10.3390/cancers16040807.

148. Chen, H.-T., Liu, H., Mao, M.-J., Tan, Y., Mo, X.-Q., Meng, X.-J., Cao, M.-T., Zhong, C.-Y., Liu, Y., Shan, H., et al. (2019). Crosstalk between autophagy and epithelial-mesenchymal transition and its application in cancer therapy. Molecular Cancer *18*, 101. https://doi.org/10.1186/s12943-019-1030-2.

149. Usman, S., Waseem, N.H., Nguyen, T.K.N., Mohsin, S., Jamal, A., Teh, M.-T., and Waseem, A. (2021). Vimentin Is at the Heart of Epithelial Mesenchymal Transition (EMT) Mediated Metastasis. Cancers *13*, 4985. https://doi.org/10.3390/cancers13194985.

150. Sørrig, R., Klausen, T.W., Salomo, M., Vangsted, A.J., Frølund, U.C., Andersen, K.T., Klostergaard, A., Helleberg, C., Pedersen, R.S., Pedersen, P.T., et al. (2017). Immunoparesis in newly diagnosed Multiple Myeloma patients: Effects on overall survival and progression free survival in the Danish population. PLOS ONE *12*, e0188988. https://doi.org/10.1371/journal.pone.0188988.

151. Chakraborty, R., Rybicki, L., Nakashima, M., Dean, R.M., Faiman, B.M., Samaras, C.J., Rosko, N., Valent, J., and Anwer, F. (2019). Depth of Pre-Treatment Immunoparesis Is an Important Prognostic Factor for Survival in Relapsed Multiple Myeloma. Blood *134*, 4359. https://doi.org/10.1182/blood-2019-123861.

152. Lakhwani, S., Rosinol Dachs, L., Puig, N., Pico Picos, M.A., Medina-González, L., Martínez-López, J., Paiva, B., Oriol, A., Rios, R., Blanchard, M.J., et al. (2022). Recovery of Uninvolved Heavy/Light Chain Pair Immunoparesis during Maintenance Is an Independent Prognostic Factor for Multiple Myeloma and Could Complement Prognosis Value of Minimal Residual Disease. Blood *140*, 2101–2103. https://doi.org/10.1182/blood-2022-167247.

153. Sigurdardottir, E.E., Turesson, I., Lund, S.H., Lindqvist, E.K., Mailankody, S., Korde, N., Björkholm, M., Landgren, O., and Kristinsson, S.Y. (2015). The Role of Diagnosis and Clinical Follow-up of Monoclonal Gammopathy of Undetermined Significance on Survival in Multiple Myeloma. JAMA Oncology *1*, 168–174. https://doi.org/10.1001/jamaoncol.2015.23.

154. Chahin, M., Branham, Z., Fox, A., Leurinda, C., and Keruakous, A.R. (2022). Clinical Considerations for Immunoparesis in Multiple Myeloma. Cancers (Basel) *14*, 2278. https://doi.org/10.3390/cancers14092278.

155. Amici, S.A., Young, N.A., Narvaez-Miranda, J., Jablonski, K.A., Arcos, J., Rosas, L., Papenfuss, T.L., Torrelles, J.B., Jarjour, W.N., and Guerau-de-Arellano, M. (2018). CD38 Is Robustly Induced in Human Macrophages and Monocytes in Inflammatory Conditions. Front. Immunol. *9*. https://doi.org/10.3389/fimmu.2018.01593.

156. Newman, P.J., and Newman, D.K. (2003). Signal Transduction Pathways Mediated by PECAM-1. Arteriosclerosis, Thrombosis, and Vascular Biology *23*, 953–964. https://doi.org/10.1161/01.ATV.0000071347.69358.D9.

157. Andreata, F., Clément, M., Benson, R.A., Hadchouel, J., Procopio, E., Even, G., Vorbe, J., Benadda, S., Ollivier, V., Ho-Tin-Noe, B., et al. (2023). CD31 signaling promotes the

detachment at the uropod of extravasating neutrophils allowing their migration to sites of inflammation. eLife *12*, e84752. https://doi.org/10.7554/eLife.84752.

158. Woodfin, A., Voisin, M.-B., and Nourshargh, S. (2007). PECAM-1: A Multi-Functional Molecule in Inflammation and Vascular Biology. Arteriosclerosis, Thrombosis, and Vascular Biology *27*, 2514–2523. https://doi.org/10.1161/ATVBAHA.107.151456.

159. Piedra-Quintero, Z.L., Wilson, Z., Nava, P., and Guerau-de-Arellano, M. (2020). CD38: An Immunomodulatory Molecule in Inflammation and Autoimmunity. Front Immunol *11*, 597959. https://doi.org/10.3389/fimmu.2020.597959.

160. Chen, C., Yu, W., Alikarami, F., Qiu, Q., Chen, C., Flournoy, J., Gao, P., Uzun, Y., Fang, L., Davenport, J.W., et al. (2022). Single-cell multiomics reveals increased plasticity, resistant populations, and stem-cell–like blasts in *KMT2A* -rearranged leukemia. Blood *139*, 2198–2211. https://doi.org/10.1182/blood.2021013442.

161. Le Guerroué, F., Eck, F., Jung, J., Starzetz, T., Mittelbronn, M., Kaulich, M., and Behrends, C. (2017). Autophagosomal Content Profiling Reveals an LC3C-Dependent Piecemeal Mitophagy Pathway. Mol Cell *68*, 786-796.e6. https://doi.org/10.1016/j.molcel.2017.10.029.

162. Ullio, C., Brunk, U.T., Urani, C., Melchioretto, P., Bonelli, G., Baccino, F.M., and Autelli, R. (2015). Autophagy of metallothioneins prevents TNF-induced oxidative stress and toxicity in hepatoma cells. Autophagy *11*, 2184–2198. https://doi.org/10.1080/15548627.2015.1106662.

163. Gore, Y., Starlets, D., Maharshak, N., Becker-Herman, S., Kaneyuki, U., Leng, L., Bucala, R., and Shachar, I. (2008). Macrophage migration inhibitory factor induces B cell survival by activation of a CD74-CD44 receptor complex. J Biol Chem *283*, 2784–2792. https://doi.org/10.1074/jbc.M703265200.

164. Narayana, S.K., Helbig, K.J., McCartney, E.M., Eyre, N.S., Bull, R.A., Eltahla, A., Lloyd, A.R., and Beard, M.R. (2015). The Interferon-induced Transmembrane Proteins, IFITM1, IFITM2, and IFITM3 Inhibit Hepatitis C Virus Entry. J Biol Chem *290*, 25946–25959. https://doi.org/10.1074/jbc.M115.657346.

165. Wilkins, C., Woodward, J., Lau, D.T.-Y., Barnes, A., Joyce, M., McFarlane, N., McKeating, J.A., Tyrrell, D.L., and Gale Jr., M. (2013). IFITM1 is a tight junction protein that inhibits hepatitis C virus entry. Hepatology *57*, 461–469. https://doi.org/10.1002/hep.26066.

166. Shah, U.A., and Mailankody, S. (2020). CAR T and CAR NK cells in multiple myeloma: Expanding the targets. Best Practice & Research Clinical Haematology *33*, 101141. https://doi.org/10.1016/j.beha.2020.101141.

167. Hengeveld, P.J., and Kersten, M.J. (2015). B-cell activating factor in the pathophysiology of multiple myeloma: a target for therapy? Blood Cancer Journal *5*, e282–e282. https://doi.org/10.1038/bcj.2015.3.

168. Novak, A.J., Darce, J.R., Arendt, B.K., Harder, B., Henderson, K., Kindsvogel, W., Gross, J.A., Greipp, P.R., and Jelinek, D.F. (2004). Expression of BCMA, TACI, and BAFF-R in multiple myeloma: a mechanism for growth and survival. Blood *103*, 689–694. https://doi.org/10.1182/blood-2003-06-2043.

169. Tinhofer, I., Marschitz, I., Henn, T., Egle, A., and Greil, R. (2000). Expression of functional interleukin-15 receptor and autocrine production of interleukin-15 as mechanisms of tumor propagation in multiple myeloma. Blood *95*, 610–618.

170. Sindaco, P., Pandey, H., Isabelle, C., Chakravarti, N., Brammer, J.E., Porcu, P., and Mishra, A. (2023). The role of interleukin-15 in the development and treatment of hematological malignancies. Front. Immunol. *14*. https://doi.org/10.3389/fimmu.2023.1141208.

171. Waldmann, T.A., Miljkovic, M.D., and Conlon, K.C. (2019). Interleukin-15 (dys)regulation of lymphoid homeostasis: Implications for therapy of autoimmunity and cancer. Journal of Experimental Medicine *217*, e20191062. https://doi.org/10.1084/jem.20191062.

172. Wang, Y., Zhang, Y., Yi, P., Dong, W., Nalin, A.P., Zhang, J., Zhu, Z., Chen, L., Benson, D.M., Mundy-Bosse, B.L., et al. (2019). The IL-15–AKT–XBP1s signaling pathway contributes to effector functions and survival in human NK cells. Nat Immunol *20*, 10–17. https://doi.org/10.1038/s41590-018-0265-1.

173. Oliva, L., and Cenci, S. (2014). Autophagy in Plasma Cell Pathophysiology. Front. Immunol. *5*. https://doi.org/10.3389/fimmu.2014.00103.

174. Kastritis, E., Zagouri, F., Symeonidis, A., Roussou, M., Sioni, A., Pouli, A., Delimpasi, S., Katodritou, E., Michalis, E., Michael, M., et al. (2014). Preserved levels of uninvolved immunoglobulins are independently associated with favorable outcome in patients with symptomatic multiple myeloma. Leukemia *28*, 2075–2079. https://doi.org/10.1038/leu.2014.110.

175. Irac, S.E., Soon, M.S.F., Borcherding, N., and Tuong, Z.K. (2024). Single-cell immune repertoire analysis. Nat Methods *21*, 777–792. https://doi.org/10.1038/s41592-024-02243-4.

176. Lefranc, M.P., Giudicelli, V., Ginestoux, C., Bodmer, J., Müller, W., Bontrop, R., Lemaitre, M., Malik, A., Barbié, V., and Chaume, D. (1999). IMGT, the international ImMunoGeneTics database. Nucleic Acids Res *27*, 209–212. https://doi.org/10.1093/nar/27.1.209.

177. Ye, J., Ma, N., Madden, T.L., and Ostell, J.M. (2013). IgBLAST: an immunoglobulin variable domain sequence analysis tool. Nucleic Acids Res *41*, W34-40. https://doi.org/10.1093/nar/gkt382.

178. van der Velden, V.H.J., Bastian, L., Brüggemann, M., Hartmann, A.M., and Darzentas, N. (2022). Immunoglobulin/T-Cell Receptor Gene Rearrangement Analysis Using RNA-Seq. Methods Mol Biol *2453*, 61–77. https://doi.org/10.1007/978-1-0716-2115-8_4.

179. Bertamini, L., D'Agostino, M., and Gay, F. (2021). MRD Assessment in Multiple Myeloma: Progress and Challenges. Curr Hematol Malig Rep *16*, 162–171. https://doi.org/10.1007/s11899-021-00633-5.

180. Yee, A.J., and Raje, N. (2021). Minimal residual disease in multiple myeloma: why, when, where. Hematology *2021*, 37–45. https://doi.org/10.1182/hematology.2021000230.

181. Zicchedda, B., Da Vià, M.C., Lionetti, M., Maeda, A., Morlupi, S., Dugo, M., Todoerti, K., Oliva, S., D'Agostino, M., Corradini, P., et al. (2021). Functional Impact of Genomic

Complexity on the Transcriptome of Multiple Myeloma. Clinical Cancer Research *27*, 6479–6490. https://doi.org/10.1158/1078-0432.CCR-20-4366.

182. Bustoros, M., Anand, S., Sklavenitis-Pistofidis, R., Redd, R., Boyle, E.M., Zhitomirsky, B., Dunford, A.J., Tai, Y.-T., Chavda, S.J., Boehner, C., et al. (2022). Genetic subtypes of smoldering multiple myeloma are associated with distinct pathogenic phenotypes and clinical outcomes. Nat Commun *13*, 3449. https://doi.org/10.1038/s41467-022-30694-w.

183. Tirier, S.M., Mallm, J.-P., Steiger, S., Poos, A.M., Awwad, M.H.S., Giesen, N., Casiraghi, N., Susak, H., Bauer, K., Baumann, A., et al. (2021). Subclone-specific microenvironmental impact and drug response in refractory multiple myeloma revealed by single-cell transcriptomics. Nat Commun *12*, 6960. https://doi.org/10.1038/s41467-021-26951-z.

184. Gao, Y., Ni, X., Guo, H., Su, Z., Ba, Y., Tong, Z., Guo, Z., Yao, X., Chen, X., Yin, J., et al. (2017). Single-cell sequencing deciphers a convergent evolution of copy number alterations from primary to circulating tumor cells. Genome Res. *27*, 1312–1322. https://doi.org/10.1101/gr.216788.116.

185. Schinke, C., and Weinhold, N. (2023). The Immune Microenvironment in Multiple Myeloma Progression at a Single-cell Level. Hemasphere *7*, e894. https://doi.org/10.1097/HS9.0000000000000894.

186. Schinke, C., Poos, A.M., Bauer, M., John, L., Johnson, S., Deshpande, S., Carrillo, L., Alapat, D., Rasche, L., Thanendrarajan, S., et al. (2022). Characterizing the role of the immune microenvironment in multiple myeloma progression at a single-cell level. Blood Advances *6*, 5873–5883. https://doi.org/10.1182/bloodadvances.2022007217.

187. Rodríguez-Otero, P., Schinke, C.D., Chari, A., Lipe, B., Lavi, N., Rasche, L., Vishwamitra, D., Skerget, S., Verona, R., Ma, X., et al. (2023). Analysis of infections and parameters of humoral immunity in patients (pts) with relapsed/refractory multiple myeloma (RRMM) treated with talquetamab (tal) monotherapy in MonumenTAL-1. JCO *41*, 8020–8020. https://doi.org/10.1200/JCO.2023.41.16_suppl.8020.

188. Dragoni, S., Hudson, N., Kenny, B.-A., Burgoyne, T., McKenzie, J.A., Gill, Y., Blaber, R., Futter, C.E., Adamson, P., Greenwood, J., et al. (2017). Endothelial MAPKs Direct ICAM-1 Signaling to Divergent Inflammatory Functions. The Journal of Immunology *198*, 4074–4085. https://doi.org/10.4049/jimmunol.1600823.

189. Cavallaro, U., and Dejana, E. (2011). Adhesion molecule signalling: not always a sticky business. Nat Rev Mol Cell Biol *12*, 189–197. https://doi.org/10.1038/nrm3068.

190. Scott, D.W., Tolbert, C.E., Graham, D.M., Wittchen, E., Bear, J.E., and Burridge, K. (2015). N-glycosylation controls the function of junctional adhesion molecule-A. MBoC *26*, 3205–3214. https://doi.org/10.1091/mbc.e14-12-1604.

191. Gorina, R., Lyck, R., Vestweber, D., and Engelhardt, B. (2014). β2 integrin-mediated crawling on endothelial ICAM-1 and ICAM-2 is a prerequisite for transcellular neutrophil diapedesis across the inflamed blood-brain barrier. J Immunol *192*, 324–337. https://doi.org/10.4049/jimmunol.1300858.

192. Bui, T.M., Wiesolek, H.L., and Sumagin, R. (2020). ICAM-1: A master regulator of cellular responses in inflammation, injury resolution, and tumorigenesis. J Leukoc Biol *108*, 787–799. https://doi.org/10.1002/JLB.2MR0220-549R.

193. Pengo, N., and Cenci, S. (2013). The role of autophagy in plasma cell ontogenesis. Autophagy *9*, 942–944. https://doi.org/10.4161/auto.24399.

194. Pengo, N., Scolari, M., Oliva, L., Milan, E., Mainoldi, F., Raimondi, A., Fagioli, C., Merlini, A., Mariani, E., Pasqualetto, E., et al. (2013). Plasma cells require autophagy for sustainable immunoglobulin production. Nat Immunol *14*, 298–305. https://doi.org/10.1038/ni.2524.

195. Ivashkiv, L.B., and Donlin, L.T. (2014). Regulation of type I interferon responses. Nat Rev Immunol *14*, 36–49. https://doi.org/10.1038/nri3581.

196. Hetz, C., and Saxena, S. (2017). ER stress and the unfolded protein response in neurodegeneration. Nat Rev Neurol *13*, 477–491. https://doi.org/10.1038/nrneurol.2017.99.

197. Fulda, S. (2015). Targeting apoptosis for anticancer therapy. Seminars in Cancer Biology *31*, 84–88. https://doi.org/10.1016/j.semcancer.2014.05.002.

198. Elmore, S. (2007). Apoptosis: A Review of Programmed Cell Death. Toxicol Pathol *35*, 495–516. https://doi.org/10.1080/01926230701320337.

199. Vazquez, M.I., Catalan-Dibene, J., and Zlotnik, A. (2015). B cells responses and cytokine production are regulated by their immune microenvironment. Cytokine *74*, 318–326. https://doi.org/10.1016/j.cyto.2015.02.007.

200. Rashid, H.-O., Yadav, R.K., Kim, H.-R., and Chae, H.-J. (2015). ER stress: Autophagy induction, inhibition and selection. Autophagy *11*, 1956–1977. https://doi.org/10.1080/15548627.2015.1091141.

201. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.

202. Sammut, S.-J., Galson, J.D., Minter, R., Sun, B., Chin, S.-F., De Mattos-Arruda, L., Finch, D.K., Schätzle, S., Dias, J., Rueda, O.M., et al. (2024). Predictability of B cell clonal persistence and immunosurveillance in breast cancer. Nat Immunol *25*, 916–924. https://doi.org/10.1038/s41590-024-01821-0.

203. Das, R., Strowig, T., Verma, R., Koduru, S., Hafemann, A., Hopf, S., Kocoglu, M.H., Borsotti, C., Zhang, L., Branagan, A., et al. (2016). Microenvironment-dependent growth of preneoplastic and malignant plasma cells in humanized mice. Nat Med *22*, 1351–1357. https://doi.org/10.1038/nm.4202.

204. Bashford-Rogers, R.J.M., Nicolaou, K.A., Bartram, J., Goulden, N.J., Loizou, L., Koumas, L., Chi, J., Hubank, M., Kellam, P., Costeas, P.A., et al. (2016). Eye on the B-ALL: B-cell receptor repertoires reveal persistence of numerous B-lymphoblastic leukemia subclones from diagnosis to relapse. Leukemia *30*, 2312–2321. https://doi.org/10.1038/leu.2016.142.

205. Amoretti, M., Amsler, C., Bonomi, G., Bouchta, A., Bowe, P., Carraro, C., Cesar, C.L., Charlton, M., Collier, M.J.T., Doser, M., et al. (2002). Production and detection of cold antihydrogen atoms. Nature *419*, 456–459. https://doi.org/10.1038/nature01096.

206. Yang, Y., Chen, X., Pan, J., Ning, H., Zhang, Y., Bo, Y., Ren, X., Li, J., Qin, S., Wang, D., et al. (2024). Pan-cancer single-cell dissection reveals phenotypically distinct B cell subtypes. Cell *187*, 4790-4811.e22. https://doi.org/10.1016/j.cell.2024.06.038.

207. Russo, M., Chen, M., Mariella, E., Peng, H., Rehman, S.K., Sancho, E., Sogari, A., Toh, T.S., Balaban, N.Q., Batlle, E., et al. (2024). Cancer drug-tolerant persister cells: from biological questions to clinical opportunities. Nat Rev Cancer, 1–24. https://doi.org/10.1038/s41568-024-00737-z.

208. Turner, R., Kalff, A., Bergin, K., Gorniak, M., Fleming, S., and Spencer, A. (2022). The Utility of Euroflow MRD Assessment in Real-World Multiple Myeloma Practice. Front Oncol *12*, 820605. https://doi.org/10.3389/fonc.2022.820605.

209. Kalina, T., Flores-Montero, J., van der Velden, V.H.J., Martin-Ayuso, M., Böttcher, S., Ritgen, M., Almeida, J., Lhermitte, L., Asnafi, V., Mendonça, A., et al. (2012). EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. Leukemia *26*, 1986–2010. https://doi.org/10.1038/leu.2012.122.

Figures

Figure 1

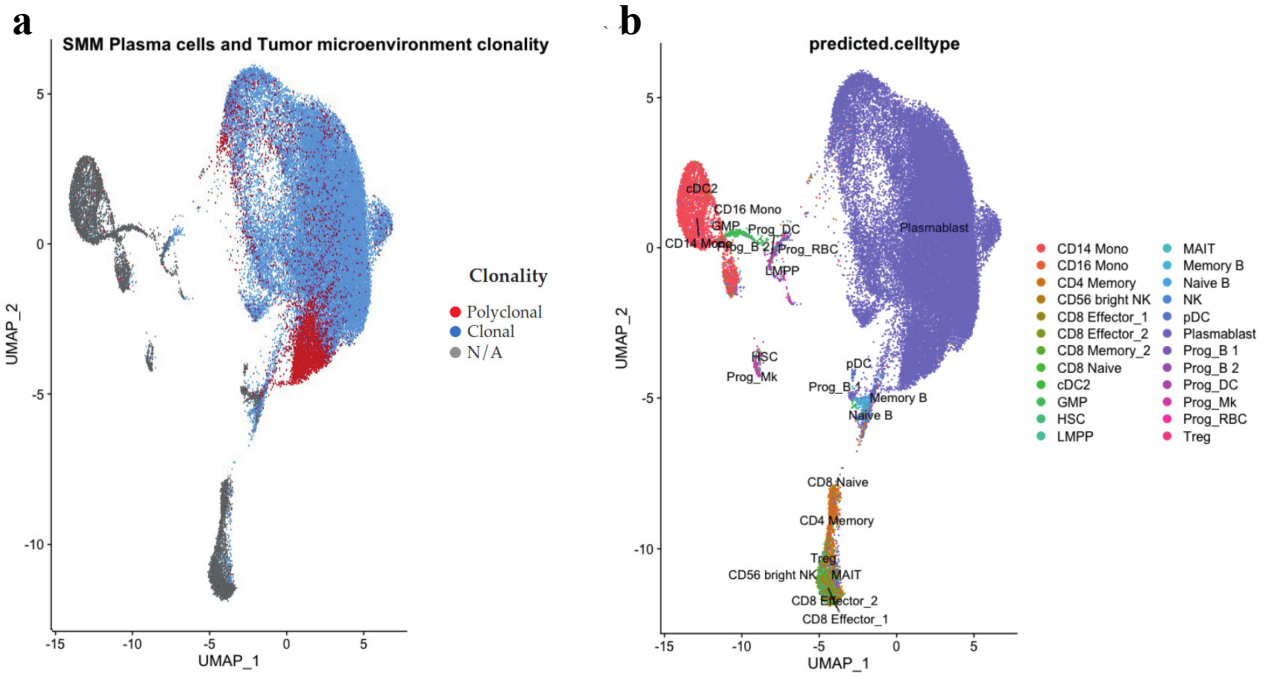

Figure 1 - UMAP projection of all cells from samples obtained after CD138 positive selection. ***a.*** *light-blue dots represent cells associated with the dominant clonotype, red dots cells associated with polyclonal VDJ, and grey dots cells without VDJ rearrangement.* ***b.*** *cells are color-coded according to transcriptome-based cell assignment as indicated in the legend.* *Image adapted from Matera et al. (2022).*

Figure 2

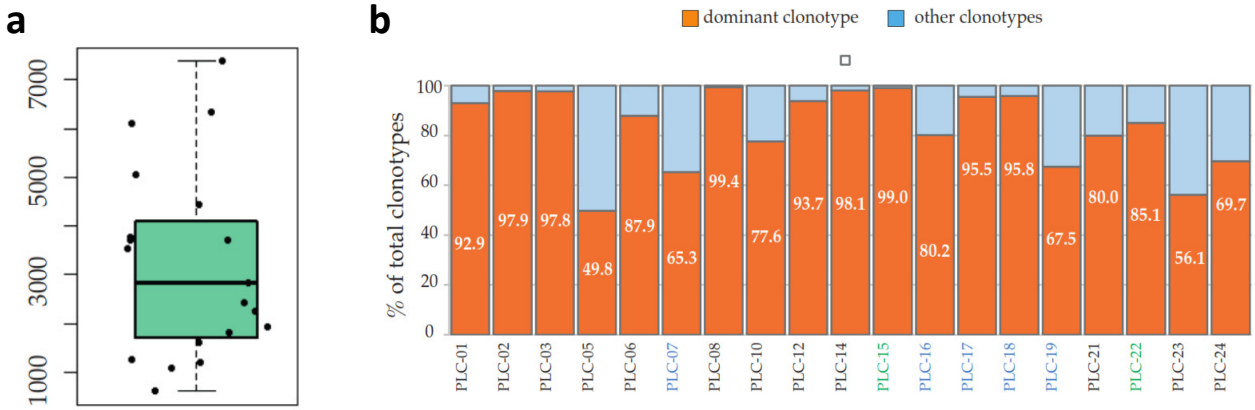

Figure 2 – Distribution of VDJ rearrangements. ***a.*** *Distribution of the number of cells expressing V(D)J transcripts in the 19 patients' samples.* ***b.*** *Percentage of cells associated with the dominant clonotype and other polyclonal VDJ rearrangement in the 19 patients' samples. On the x-axis, IDs of MGUS samples are in blue, IDs of SMM samples are in black and IDs of MM samples are in green.* *Image adapted from Matera et al. (2022).*

Figure 3

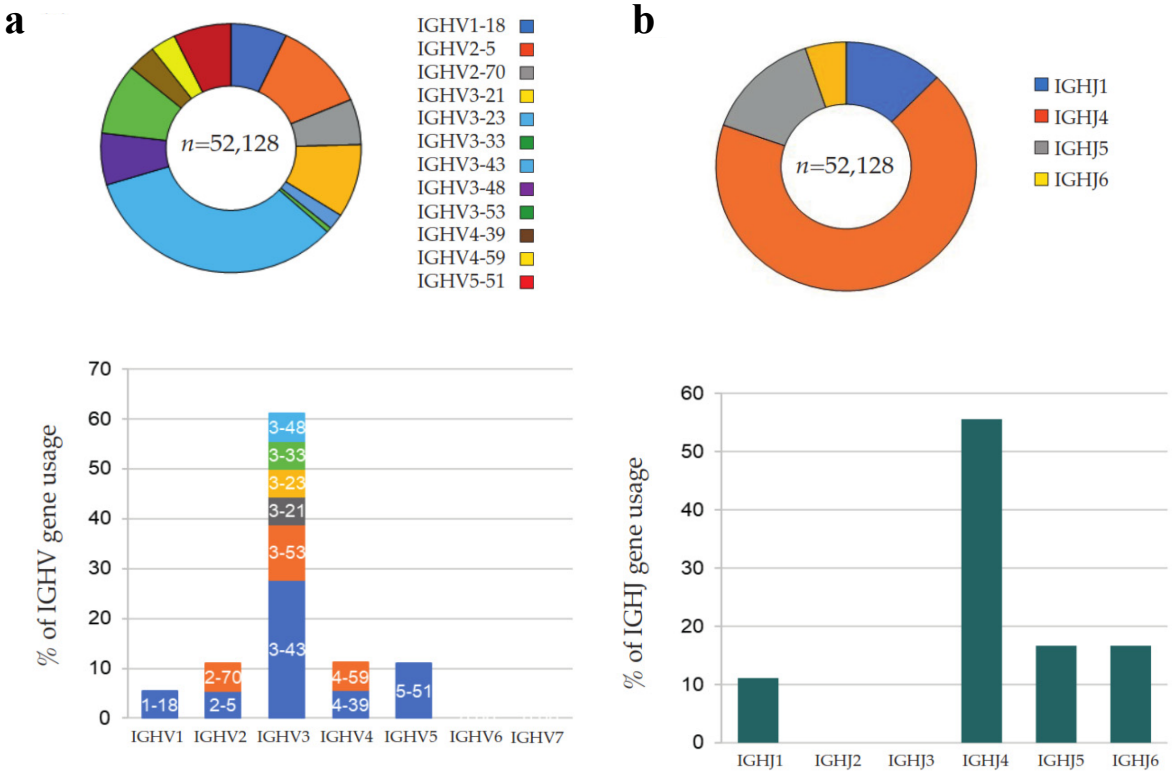

Figure 3 – VH and JH genes representation among 18/19 samples with clonal heavy chain rearrangements. ***a.*** *VH and* ***b.*** *JH gene usage. The histograms display the sample-level rate of representation of each IGH variable region gene family on total of clonal rearrangements. Within each V-family, discrete bands represent each of the individual genes, as indicated. Above each histogram, the representativeness of relative IGH genes/families is plotted at single-cell level, as number of associated cellular barcodes.* *Image adapted from Matera et al. (2022).*

Figure 4

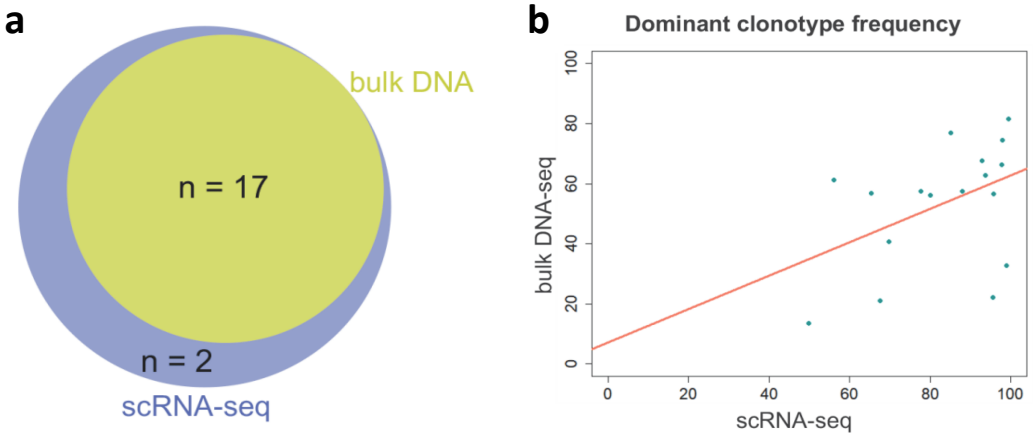

Figure 4 – Comparison of bulk DNA-seq and scRNAseq results. ***a.*** *Venn diagram illustrating the number of unique and shared clonotypes identified by scRNAseq (blue circle) and bulk DNA (yellow circle) analyses.* ***b.*** *Scatterplot displaying the percentage frequencies of the 17 dominant clonotypes identified in (**a**), with scRNA-seq frequencies on the x-axis and bulk DNAseq frequencies on the y-axis.* *Image adapted from Matera et al. (2022).*

Figure 5

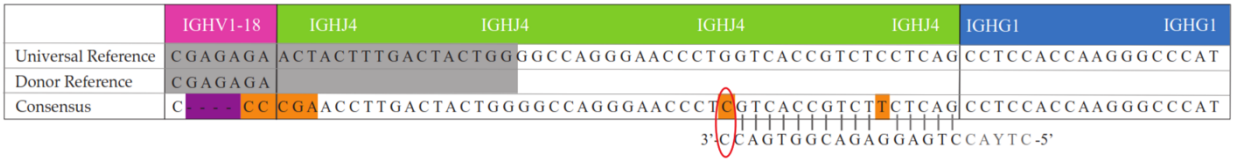

Figure 5 - Alignment analysis between EuroClonality-NGS reverse primers and dominant clonotype sequence. *Alignment of the DNA sequences from EuroClonality-NGS protocol's reverse primers, IGH-J-A-1 and IGH-J-A-2 (written in 3'-to-5' direction), with the rearranged sequence of the dominant clonotype ("Consensus") as displayed by Loupe V(D)J browser for patient sample PLC-14. Nucleotides highlighted in orange in the Consensus sequence represent mutations compared to the germline IGHJ4 gene ("Universal Reference"). The gray background indicates regions where the sequences align with the reference. The purple background corresponds to a nucleotide deletion. The red oval highlights the mutated nucleotide that causes misannealing with the 3' end of both reverse primers used in the first-step PCR, thereby preventing PCR priming. Nucleotides in gray in the primer sequences belong to the intron separating the J segment of the rearranged V-region and the C-region sequence. This intron is removed after transcription by RNA splicing, joining the V-region exon to the C-region sequence, and thus this portion of the reverse primers does not match the consensus sequence of the mature RNA transcript.* *Image adapted from Matera et al. (2022).*

Figure 6

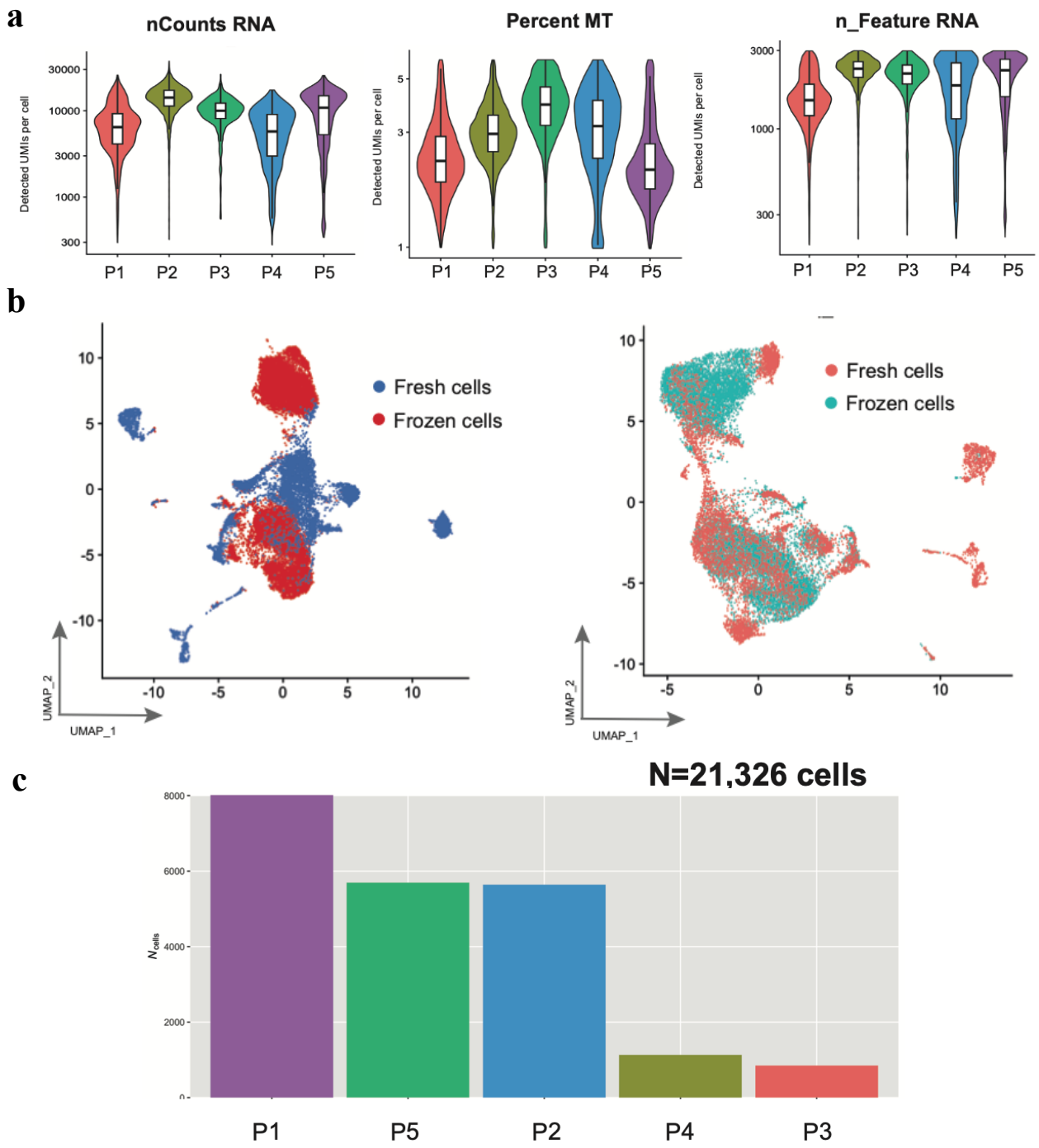

Figure 6 - QC of scRNAseq data. ***a.*** *QC metrics displayed as violin plot of quality metrics, as nCounts_RNA, percent MT and n-feature RNA per sample.* ***b.*** *UMAP dimensions reduction of integrated five samples before and after batch effect correction by Harmony. Colored by fresh vs. frozen processing.* ***c.*** *Bar plot showing number of cells by patient.* *Image adapted from Lazzaroni & Matera et al. (2024).*

Figure 7

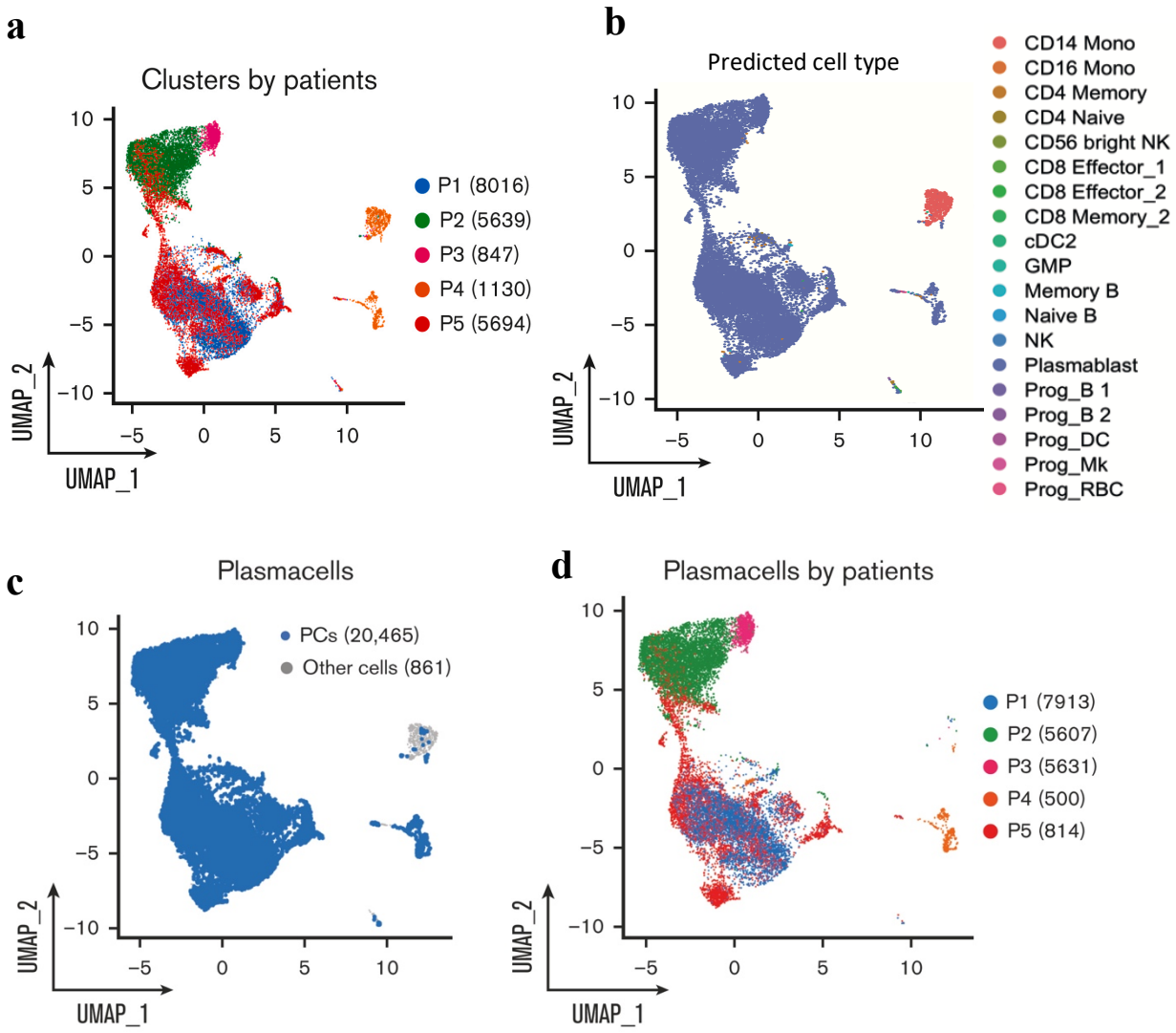

Figure 7 – UMAP representations of integrated samples. ***a.*** *UMAP dimensions reduction clustered by patients.* ***b.*** *UMAP visualization of cell type of the integrated dataset.* ***c.*** *UMAP shows PCs (blue) after cell assignment.* ***d.*** *UMAP clustering of PCs by patients.* Image adapted from Lazzaroni & Matera et al. (2024).

Figure 8

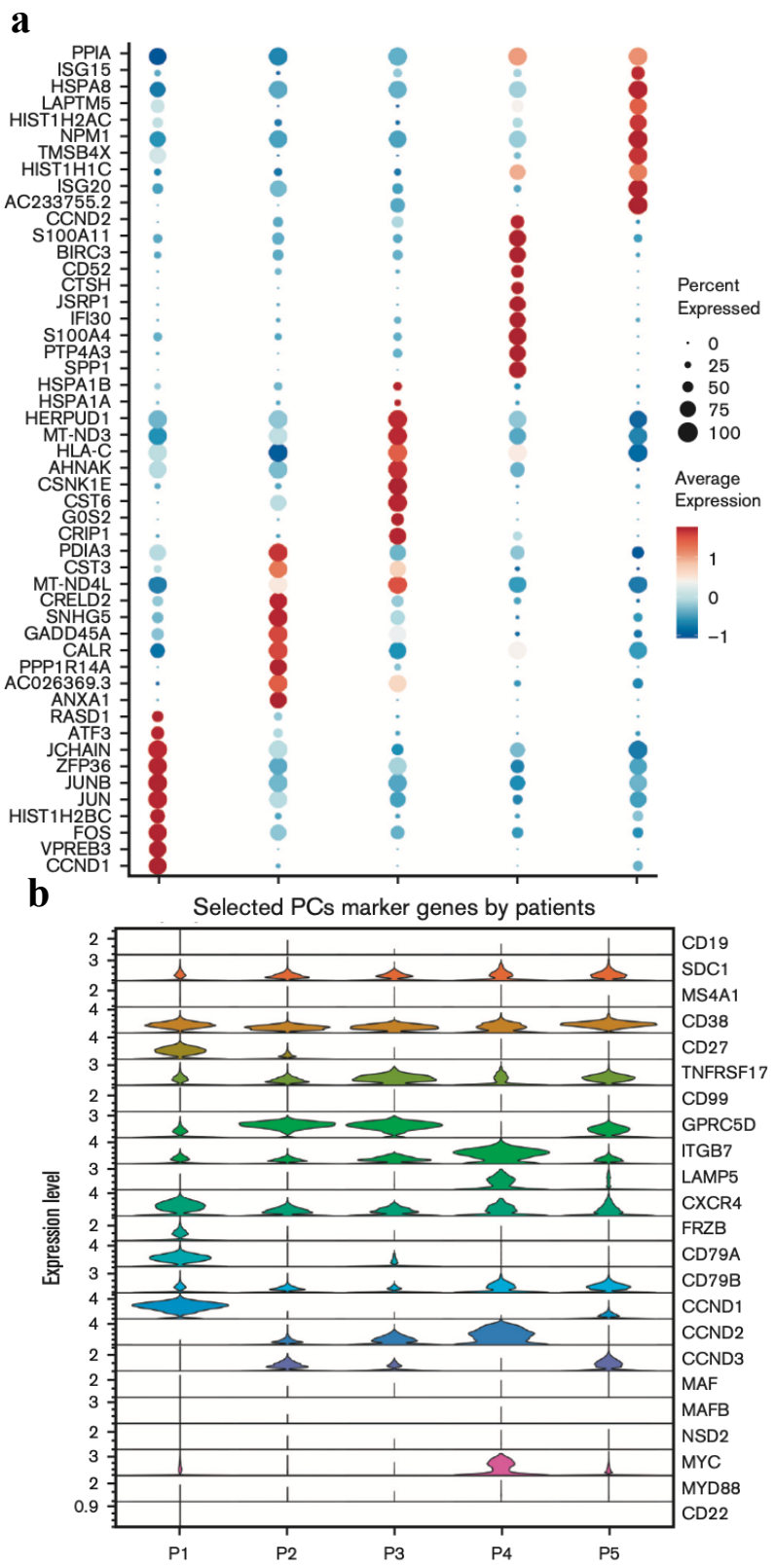

Figure 8 – Differential marker genes analysis. ***a.*** *Dot plot representing top 10 marker genes (y-axis) specific for each patient (x-axis).* ***b.*** *Violin plot showing expression levels (y-axis) of 23 selected PCs marker genes across the 5 patients (x-axis).* *Image adapted from Lazzaroni & Matera et al. (2024).*

Figure 9

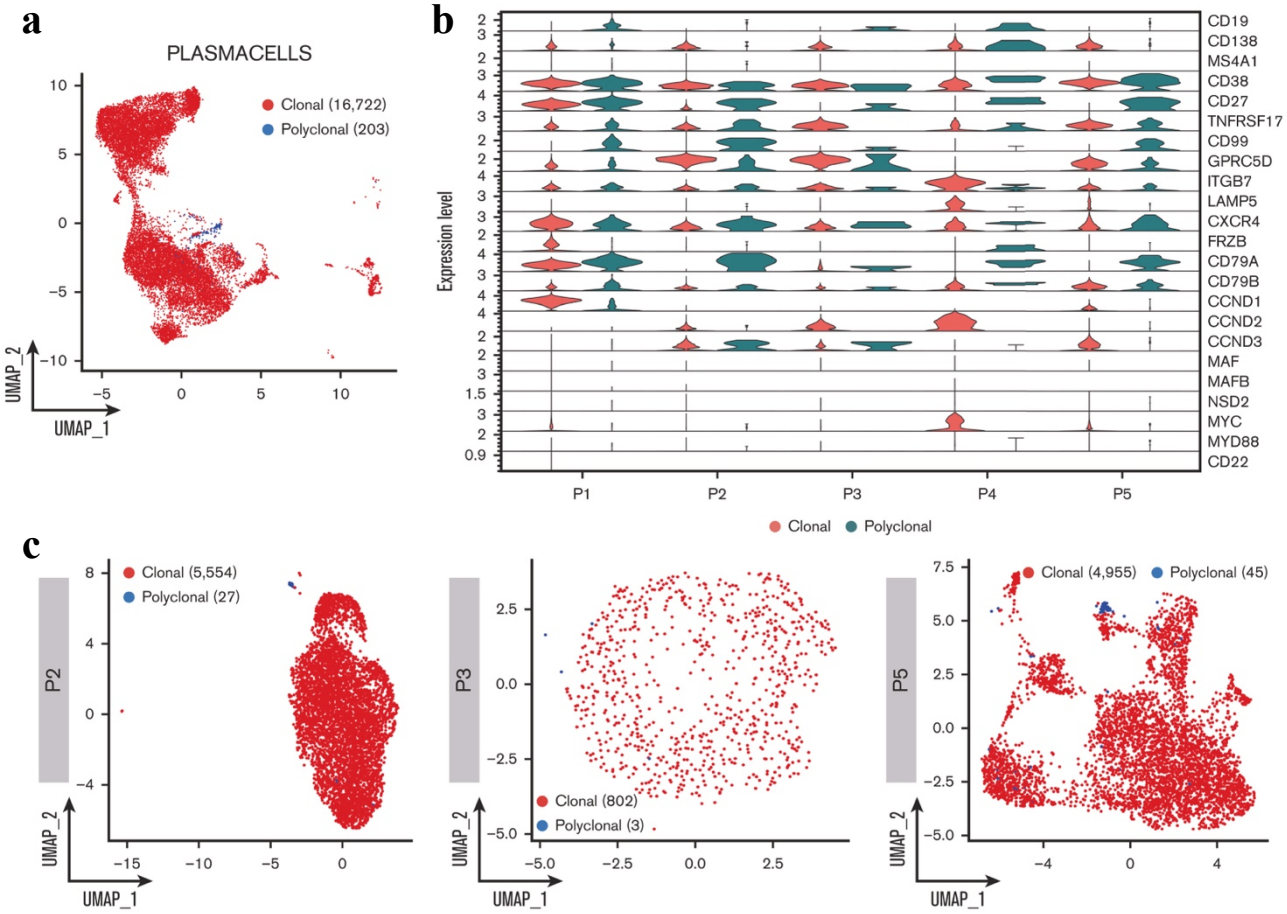

Figure 9 - Clonotype analysis through combination of scRNA and scVDJ. ***a.*** *UMAP dimensions reduction of integrated PC from all 5 patients showing scVDJ profiling, cPCs in red and pPCs blue dots.* ***b.*** *Violin plot representing expression levels of selected marker genes in cPCs and pPCs across all patients.* ***c.*** *UMAP projections of BCR clonality status patients P2, P3 and P5; cPCs in red and pPCs blue dots.* *Image adapted from Lazzaroni & Matera et al. (2024).*

Figure 10

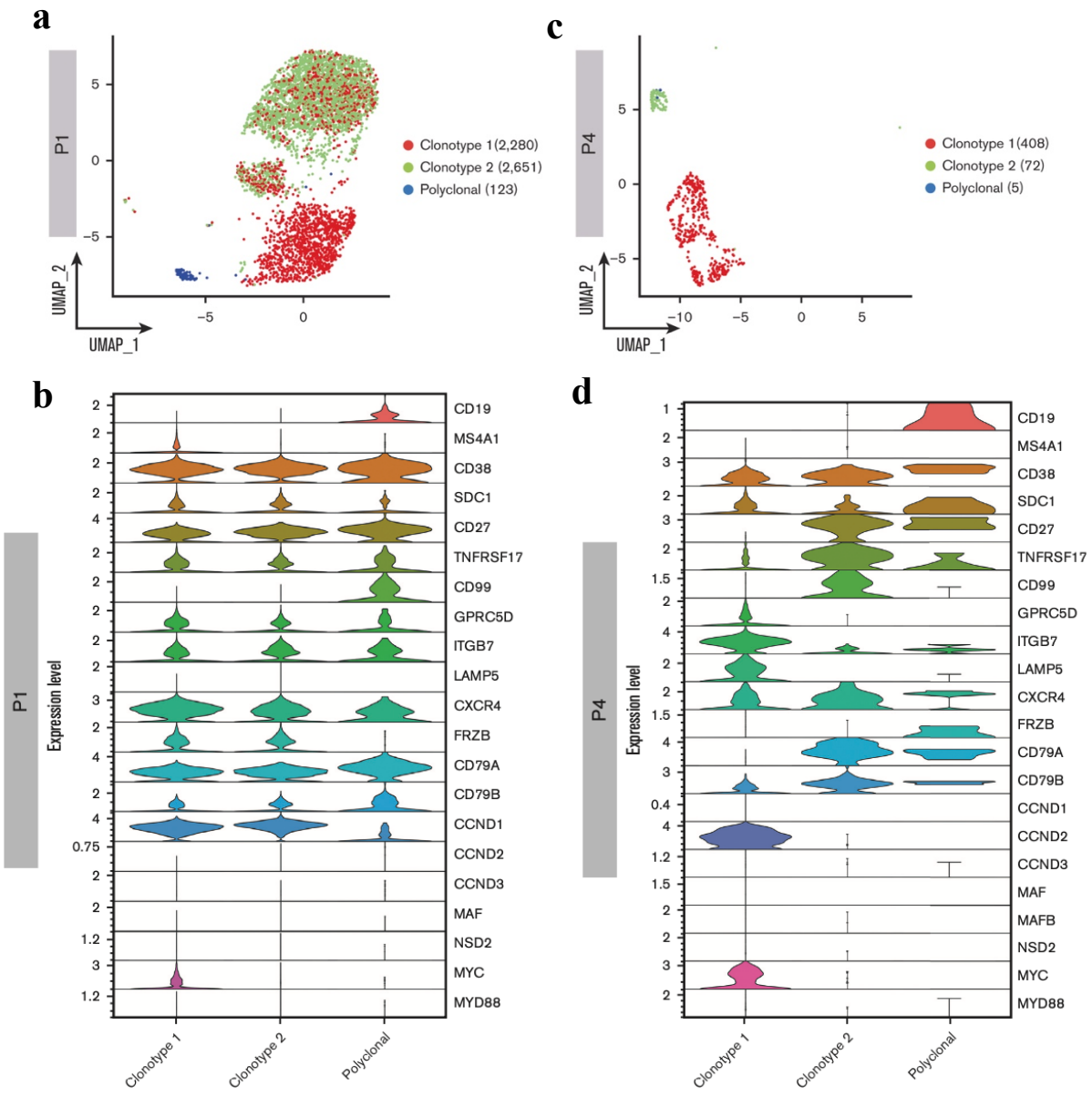

Figure 10 – scVDJseq analysis reveals two dominant clonotypes. ***a. and c****. UMAP projections of BCR clonality status patients P1 and P4; cPCs bearing clonotype 1 in red, cPCs bearing clonotype 2 in green and pPCs blue.* ***b. and d.*** *Violin plot representing expression levels of selected marker genes of PCs across the clonotypes of P1 and P4 patients, respectively.* *Image adapted from Lazzaroni & Matera et al. (2024).*

Figure 11

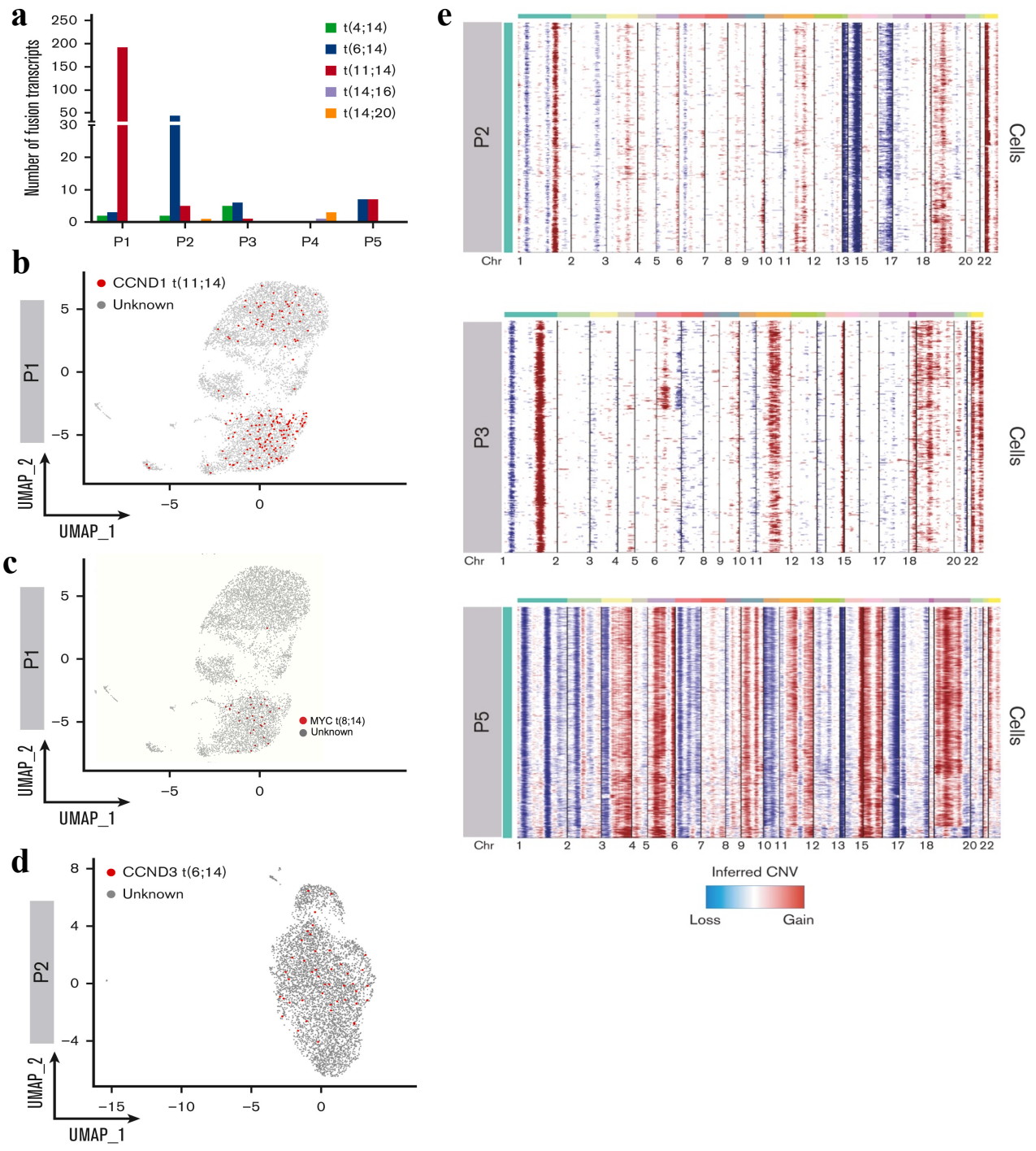

Figure 11 - Inference of genomic properties from scRNA-seq. ***a.*** *Bar chart viewing the quantification of Fuscia fusion transcript events across patients and translocation types (x-axis).* ***b-c.*** *t(11;14) and t(8;14) fusion transcripts mapped on P1 UMAP.* ***d.*** *t(6;14) fusion transcripts mapped on P2 UMAP. Cells bearing fusion transcripts are highlighted by red dots.* ***e.*** *CNAs from inferCNV analysis of patients P2, P3, and P5, characterized by one dominant clonotype. Image adapted from Lazzaroni & Matera et al. (2024).*

Figure 12

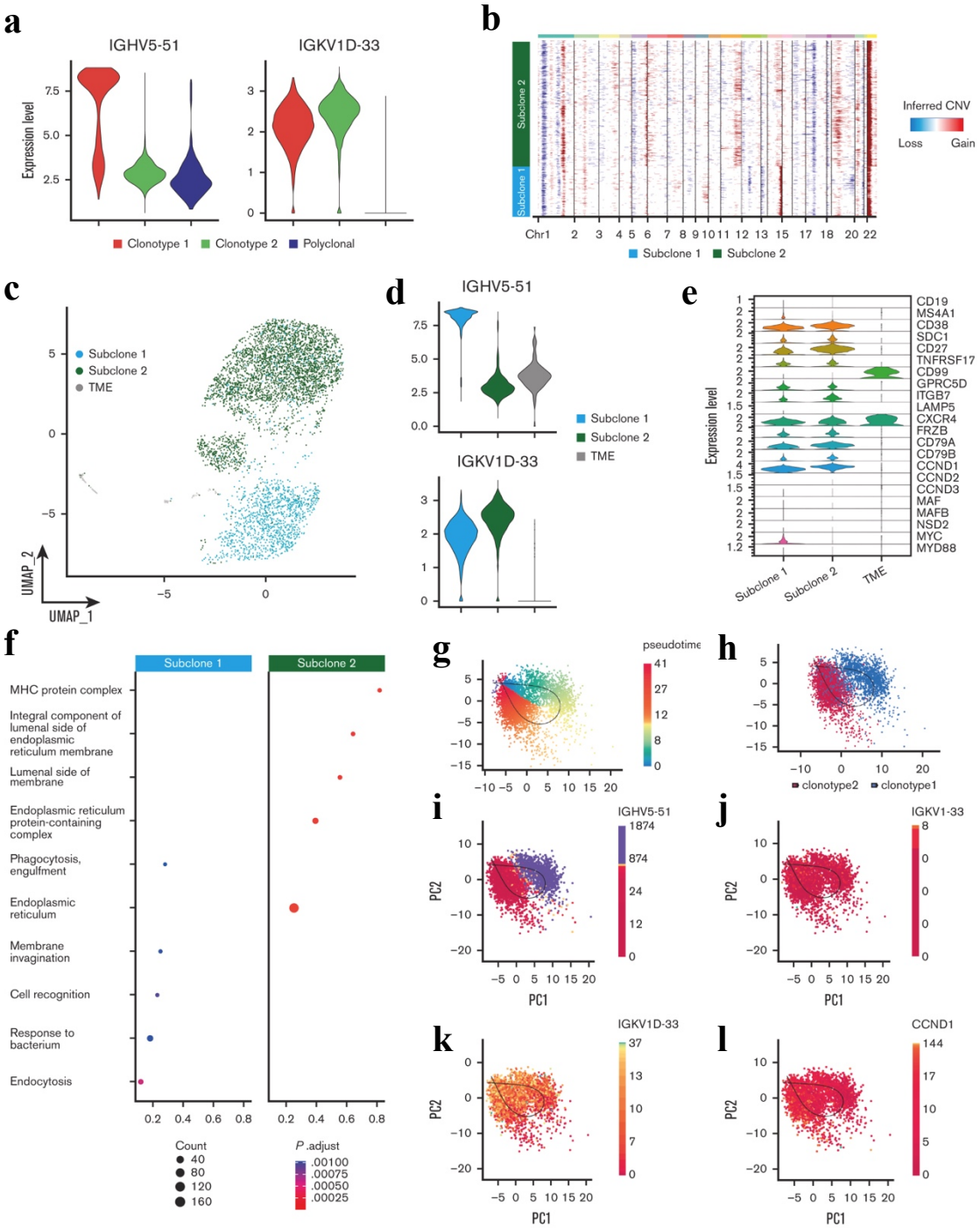
a
IGHV5-51
IGKV1D-33
Expression level
Clonotype 1
Clonotype 2
Polyclonal
b
Subclone 2
Subclone 1
Inferred CNV
Loss
Gain
Chr1
Subclone 1
Subclone 2
c
Subclone 1
Subclone 2
TME
UMAP_2
UMAP_1
d
IGHV5-51
IGKV1D-33
Subclone 1
Subclone 2
TME
e
Expression level
CD19
MS4A1
CD38
SDC1
CD27
TNFRSF17
CD99
GPRC5D
ITGB7
LAMP5
CXCR4
FRZB
CD79A
CD79B
CCND1
CCND2
CCND3
MAF
MAFB
NSD2
MYC
MYD88
Subclone 1
Subclone 2
TME
f
Subclone 1
Subclone 2
MHC protein complex
Integral component of lumenal side of endoplasmic reticulum membrane
Lumenal side of membrane
Endoplasmic reticulum protein-containing complex
Phagocytosis, engulfment
Endoplasmic reticulum
Membrane invagination
Cell recognition
Response to bacterium
Endocytosis
Count
40
80
120
160
P.adjust
.00100
.00075
.00050
.00025
g
pseudotime
h
clonotype2
clonotype1
i
IGHV5-51
PC2
PC1
j
IGKV1-33
PC2
PC1
k
IGKV1D-33
PC2
PC1
l
CCND1
PC2
PC1

Figure 12 – Subclone characterization of P1. ***a.*** *Violin plots illustrating the expression of IGH and IGK in two BCR clonotypes from sample P1. Clonotype 1 is shown in red, clonotype 2 in green, and pPCs in blue.* ***b.*** *CNAs from inferCNV highlight subclonal differences between the two tumor clonotypes. Subclone 1 in light blue and subclone 2 in green.* ***c.*** *A UMAP displaying the distribution of subclones. TME is indicated in gray.* ***d.*** *Violin plots of IGH and IGK expression in each subclone.* ***e.*** *Stacked violin plot presenting the analysis of marker gene expression by subclones and TME.* ***f.*** *Gene Ontology (GO) analysis comparing the transcriptional profiles of CNAs subclones.* ***g-k.*** *Trajectory analysis inferred by slingshot depicting pseudotime inference (**g**), clonotype origin (**h**), IGHV5-51 gene expression (**i**), IGKV1-33 gene expression (**j**), IGKV1D-33 gene expression (**k**), and CCND1 gene expression (**l**).* *Image adapted from Lazzaroni & Matera et al. (2024).*

Figure 13

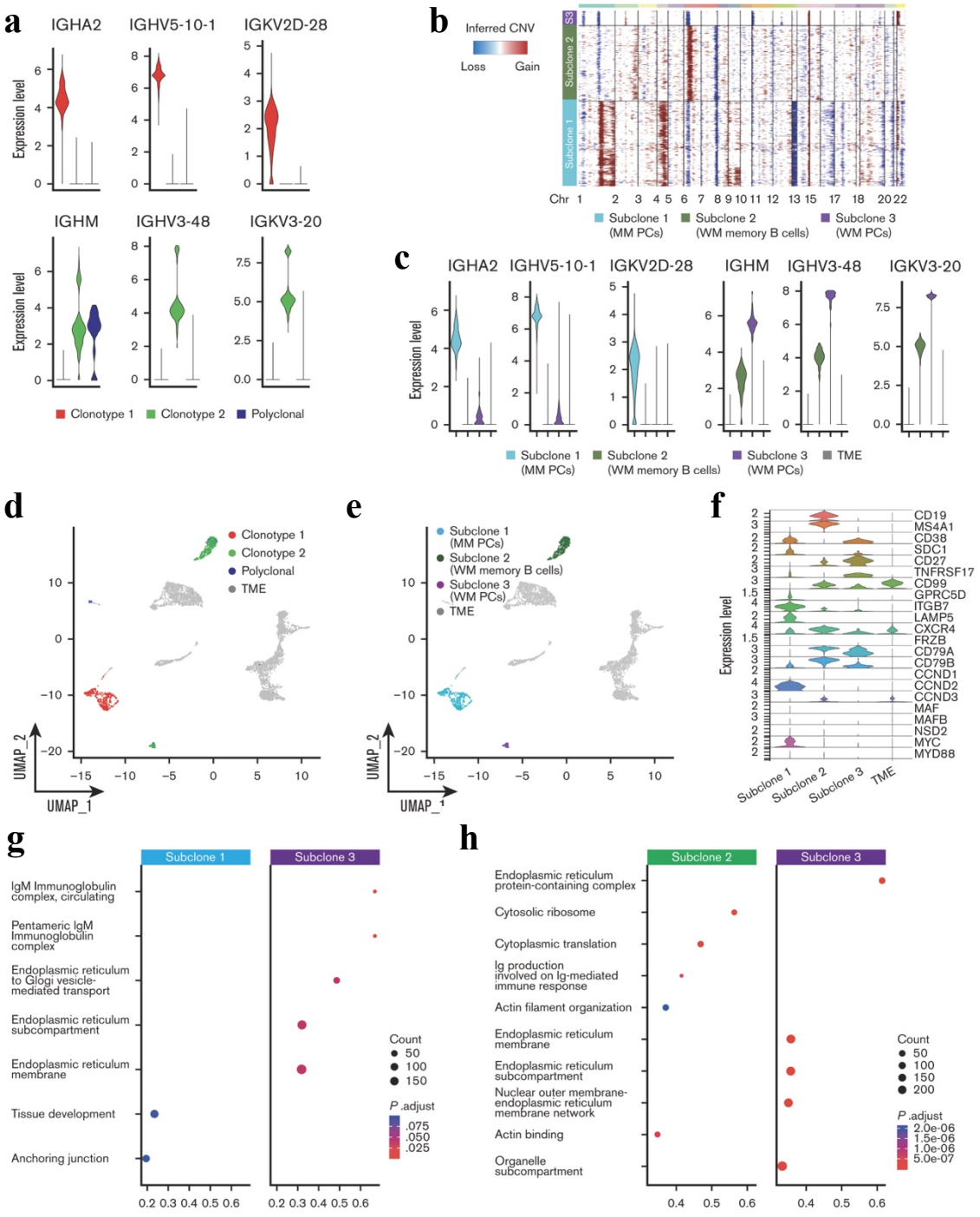
a
IGHA2
IGHV5-10-1
IGKV2D-28
IGHM
IGHV3-48
IGKV3-20
Expression level
Clonotype 1
Clonotype 2
Polyclonal
b
Inferred CNV
Loss
Gain
Subclone 3
Subclone 2
Subclone 1
Chr 1 2 3 4 5 6 7 8 9 10 11 12 13 15 17 18 20 22
Subclone 1 (MM PCs)
Subclone 2 (WM memory B cells)
Subclone 3 (WM PCs)
c
IGHA2
IGHV5-10-1
IGKV2D-28
IGHM
IGHV3-48
IGKV3-20
Expression level
Subclone 1 (MM PCs)
Subclone 2 (WM memory B cells)
Subclone 3 (WM PCs)
TME
d
Clonotype 1
Clonotype 2
Polyclonal
TME
UMAP_2
UMAP_1
e
Subclone 1 (MM PCs)
Subclone 2 (WM memory B cells)
Subclone 3 (WM PCs)
TME
UMAP_2
UMAP_1
f
Expression level
CD19
MS4A1
CD38
SDC1
CD27
TNFRSF17
CD99
GPRC5D
ITGB7
LAMP5
CXCR4
FRZB
CD79A
CD79B
CCND1
CCND2
CCND3
MAF
MAFB
NSD2
MYC
MYD88
Subclone 1
Subclone 2
Subclone 3
TME
g
Subclone 1
Subclone 3
IgM Immunoglobulin complex, circulating
Pentameric IgM Immunoglobulin complex
Endoplasmic reticulum to Golgi vesicle-mediated transport
Endoplasmic reticulum subcompartment
Endoplasmic reticulum membrane
Tissue development
Anchoring junction
Count
50
100
150
P.adjust
.075
.050
.025
h
Subclone 2
Subclone 3
Endoplasmic reticulum protein-containing complex
Cytosolic ribosome
Cytoplasmic translation
Ig production involved on Ig-mediated immune response
Actin filament organization
Endoplasmic reticulum membrane
Endoplasmic reticulum subcompartment
Nuclear outer membrane-endoplasmic reticulum membrane network
Actin binding
Organelle subcompartment
Count
50
100
150
200
P.adjust
2.0e-06
1.5e-06
1.0e-06
5.0e-07

Figure 13 – Subclone characterization of P4. ***a.*** *Violin plots illustrating the expression of isotype, IGH and IGK in two BCR clonotypes from sample P4, highlighting MM and WM cells. Clonotype 1 in red, clonotype 2 in green, and pPCs in blue.* ***b.*** *CNAs from inferCNV highlight subclonal differences between the three tumor clonotypes: subclone 1 (MM PCs) in light blue, subclone 2 (WM memory B cells) in green, subclone 3 (WM PCs) in violet.* ***c.*** *Violin plots of isotype, IGH and IGK expression in each MM and WM subclone and TME represented in gray.* ***d-e.*** *UMAPs displaying the concordance in distribution of dominant clonotypes with MM and WM subclones from inferCNV analysis. TME is indicated in gray.* ***f.*** *Stacked violin plot presenting the analysis of marker gene expression by subclones and TME.* ***g-h.*** *Gene Ontology (GO) analysis comparing the transcriptional profiles of MM PCs versus WM PCs (**g**) and WM memory B cells versus WM PCs (**h**).* *Image adapted from Lazzaroni & Matera et al. (2024).*

Figure 14

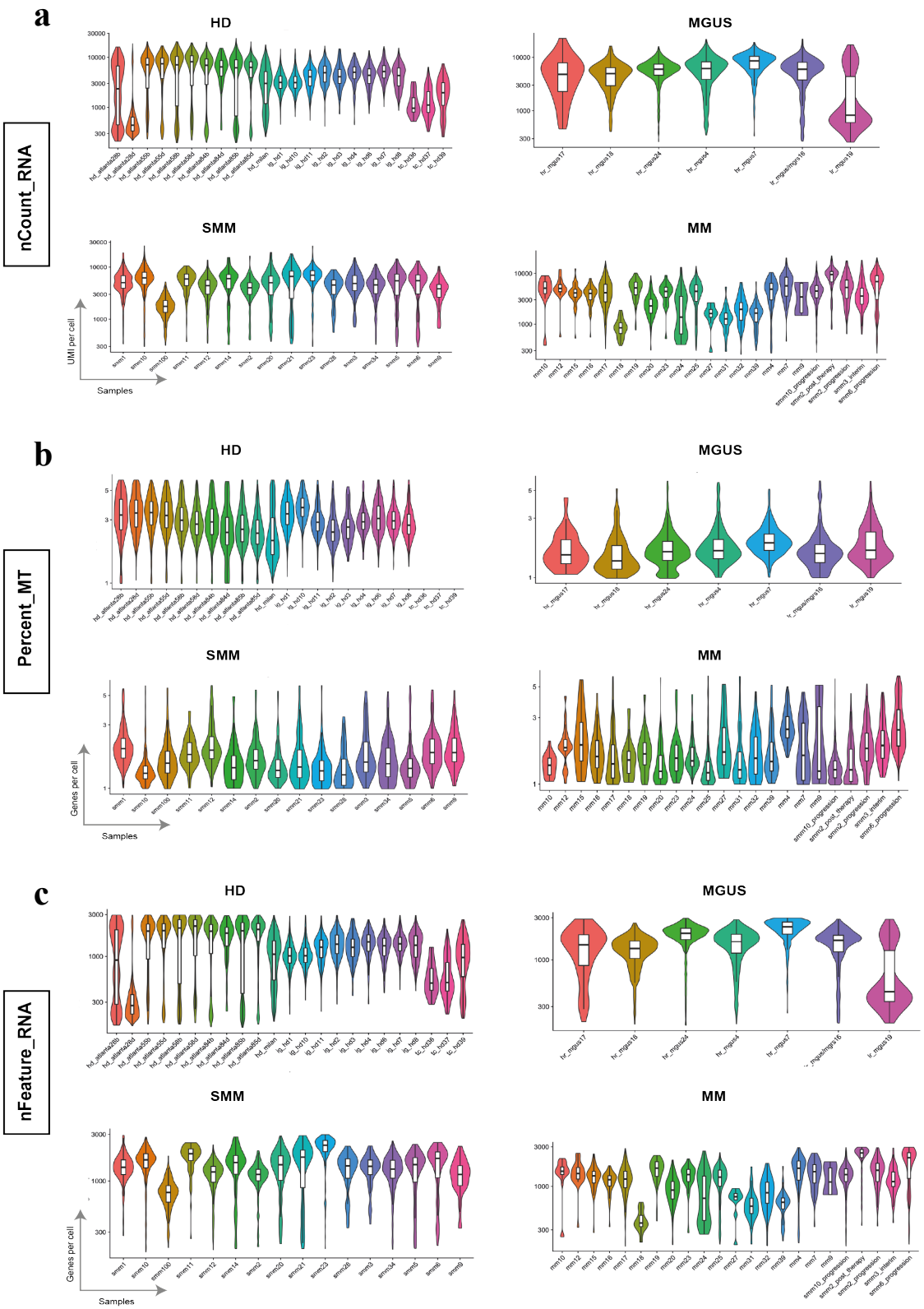

Figure 14 - QC of scRNAseq data*. QC metrics displayed as violin plot of quality metrics, as* ***a.*** *nCounts_RNA,* ***b.*** *Percent_MT and* ***c.*** *nFeature_RNA distribution per sample divided by HD, MGUS, SMM and MM. Image adapted from Da Vià et al. (2024).*

Figure 15

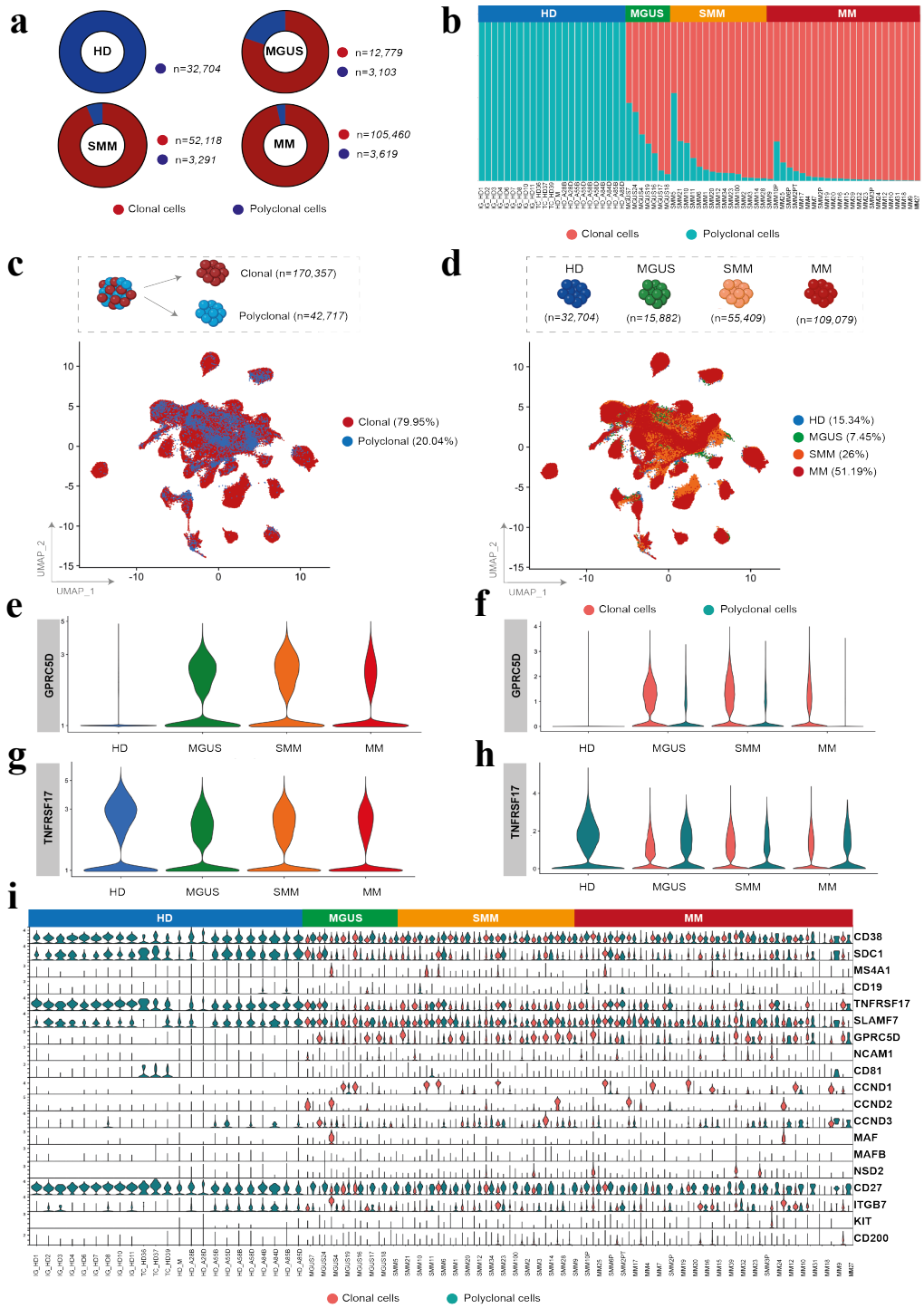

Figure 15 - Clonal and polyclonal PCs in patients affected by PC dyscrasias and healthy PCs. ***a.*** *Donut charts depict the distribution of clonal and polyclonal cells in PC dyscrasia and HD, categorized based on clinical stratification (HD: blue; MGUS: green; SMM: orange; MM: red). The legend includes the number of cells in each category.* ***b.*** *Bar charts illustrate the proportions of cPCs (red) and pPCs (blue) in each sample.* ***c.*** *UMAP projection shows all clonal (red) and polyclonal (blue) PCs, with numerical values and respective proportions and percentage.* ***d.*** *UMAP projection shows the clinical stratification of samples, with numerical values and percentages for each category reported in the legend.* ***e.*** *Violin plots display the distribution of gene expression commonly upregulated in PC dyscrasias among patients.* Image adapted from Da Vià et al. (2024).

Figure 16

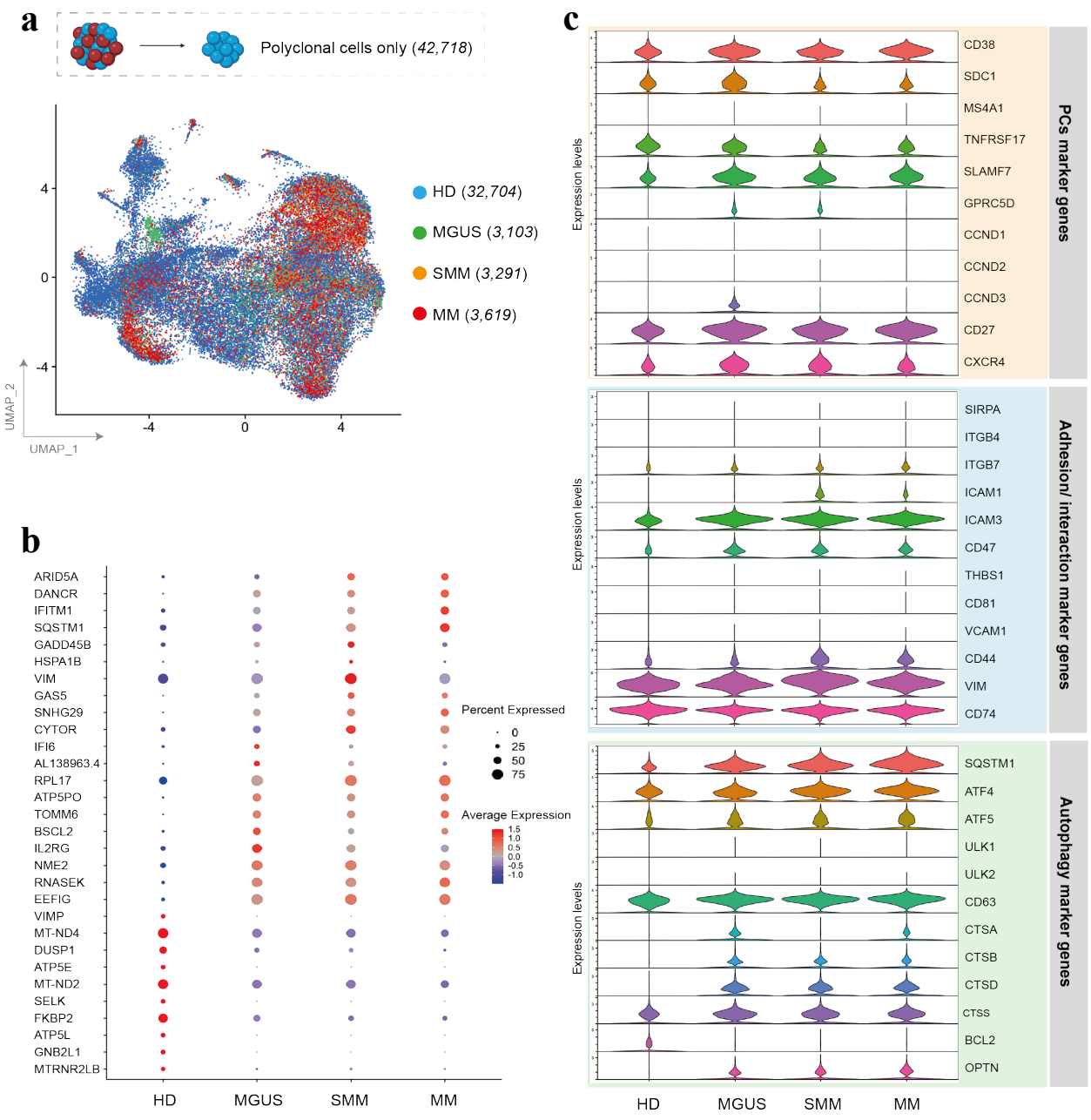

Figure 16 - Polyclonal cell selection and marker genes expression by clinical stages. ***a.*** *UMAP projection of scRNAseq data of pPCs, colors representing clinical status. Numerical values for each category are provided in the legend.* ***b.*** *Dot plot representing the top ten marker genes that distinguish each clinical status. On x-axis the clinical category, while on y-axis genes. Dot size corresponds to the number of cells in the category expressing the gene of interest, while shade correlates with the level of expression.* ***c.*** *Violin plot showing expression of selected gene signatures: PC marker genes, adhesion/interaction and autophagy across all the clinical status.*
Image adapted from Da Vià et al. (2024).

Figure 17

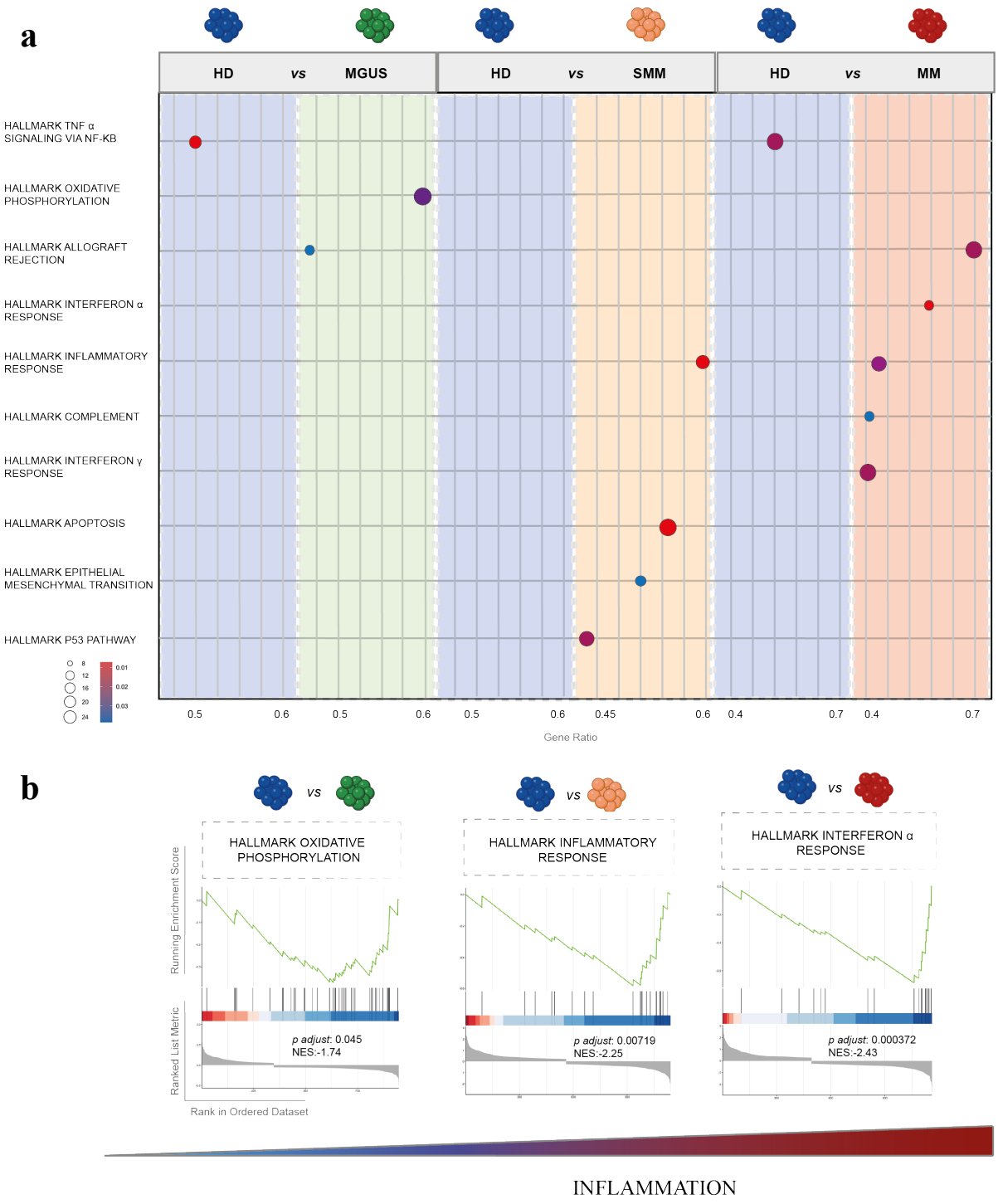

Figure 17 - Transcriptional hallmarks of pPC across status. ***a.*** *Illustration showing hallmarks enrichment profiles of pPCs, derived from 1:1 DE analyses by Wilcoxon test. Dot size indicates the number of cells expressing the gene-set of interest, while shade reflects the expression level.* ***b.*** *Gene set enrichment analysis (GSEA) conducted on genes ranked by their contribution to hallmark pathways, including oxidative phosphorylation, inflammatory response and interferon alpha response.* *Image adapted from Da Vià et al. (2024).*

Figure 18

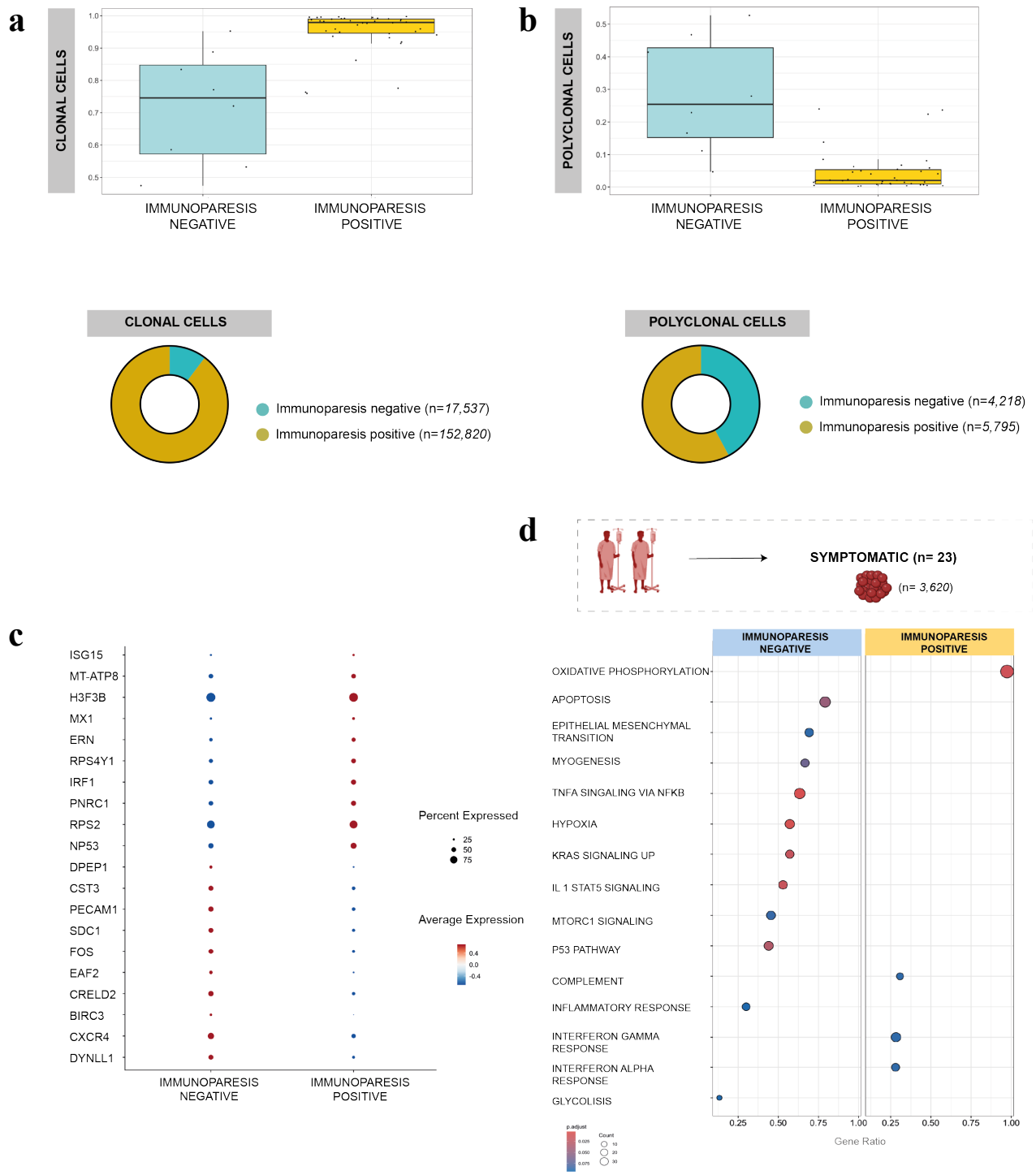

Figure 18 – PC transcriptome in immunoparesis setting. **a. and b.** Box plot representing patients' cPCs **(a)** and pPCs **(b)** distribution across immunoparesis negative (light blue) or positive (yellow) patients. Respective donut plot showing the proportion of immunoparesis negative or positive in patients' cPCs **(a)** and pPCs **(b)** and their respective number of cells. **c.** *A dot plot displays the top ten marker genes that differentiate between patients who are immunoparesis-negative and those who are immunoparesis-positive. On x-axis the clinical category, while on y-axis the genes. Dot size indicates the number of cells in each category that express the gene, and the color intensity reflects the expression level.* ***d.*** *DE analysis compares immunoparesis-negative and immunoparesis-positive patients within the symptomatic patient subset, utilizing the Wilcoxon rank sum test with p-value adjustment through the Benjamini-Hochberg correction.* Image adapted from Da Vià et al. (2024).

Figure 19

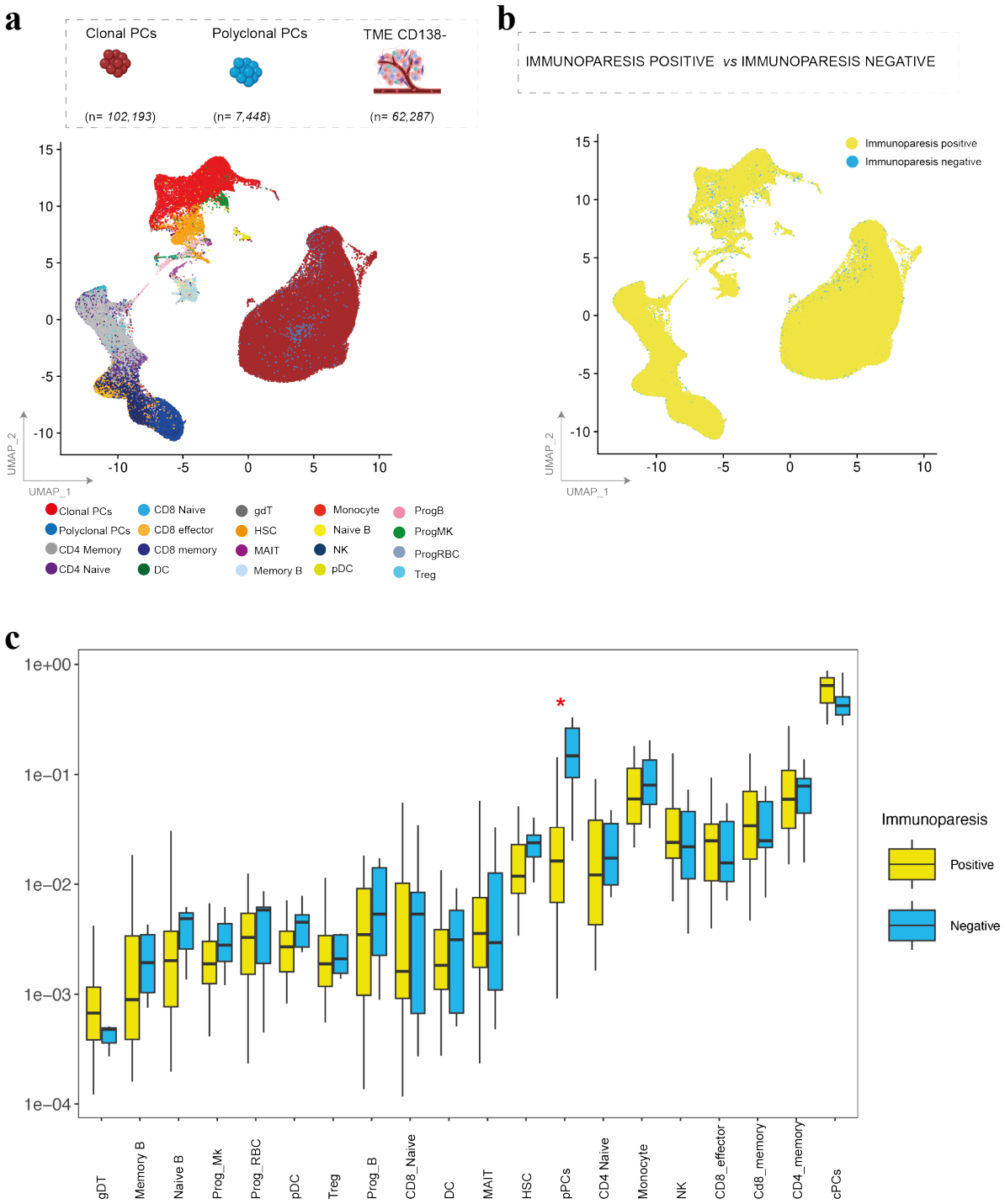

Figure 19 - Integration of the TME dataset highlighting immunoparesis. ***a.*** *UMAP projection of PC integrated with TME, with colors representing different TME cell types.* ***b.*** *UMAP projection of integrated PC and TME, colored based on immunoparesis status.* ***c.*** *A box plot shows the relative abundance of cells in the integrated dataset (TME and PCs). Cell types on the x-axis, relative frequency on a logarithmic scale on y-axis. The red asterisk (*) above the "pPCs" label denotes a padj value of 0.004, calculated using the Wilcoxon test with Benjamini-Hochberg correction.* *Image adapted from Da Vià et al. (2024).*

Figure 20

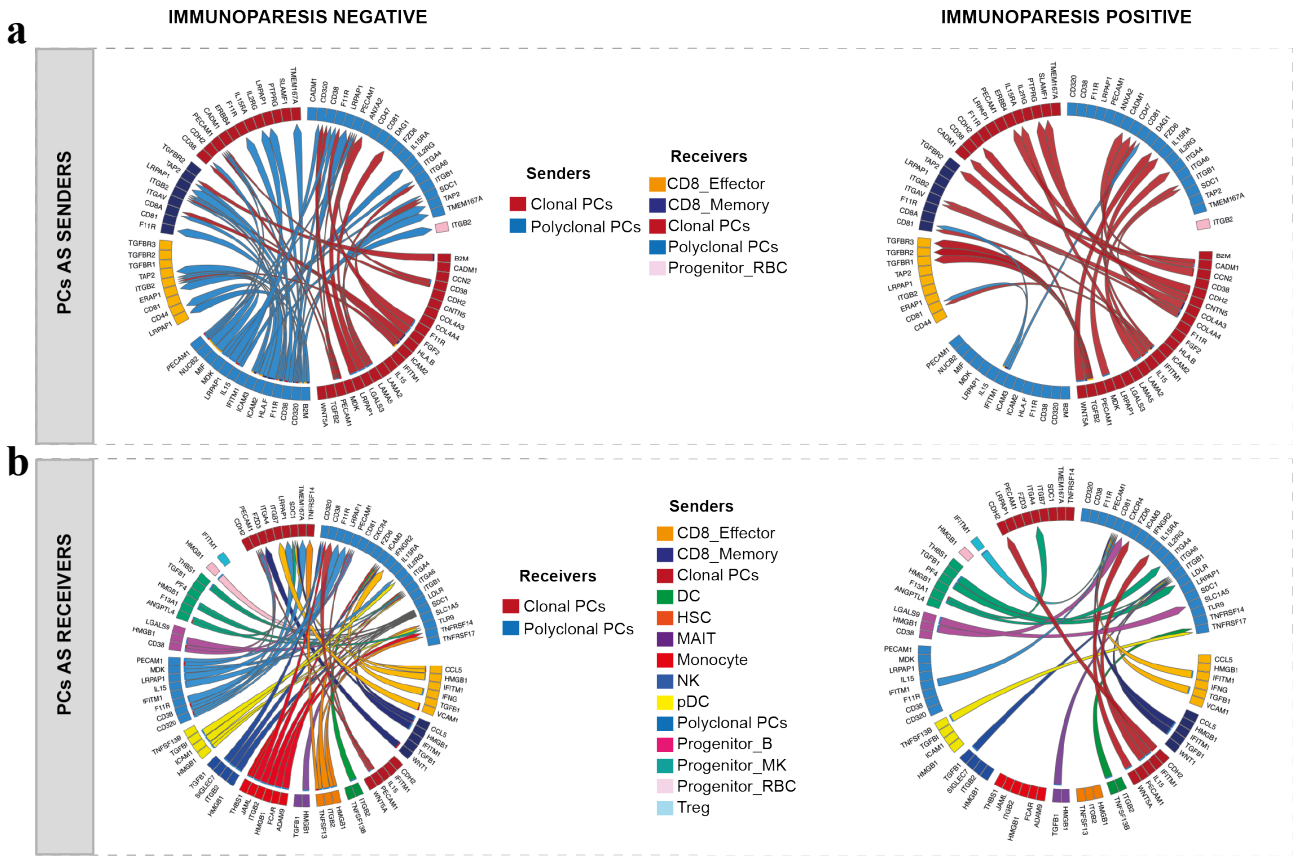

Figure 20 - Predicted interactions between PCs and TME. *Chord diagrams depict ligand-receptor interactions between pPCs (blue), cPCs (red), and the TME, categorized by immunoparesis clinical status. The ribbon arrows show the direction of communication from the sender to the receiver populations, with the arrow color representing the specific sender cell type expressing the ligand. The upper (**a**) panel shows PCs as the sender, while the lower panel (**b**) displays PCs as the receiver.*
Image adapted from Da Vià et al. (2024).

Figure 21

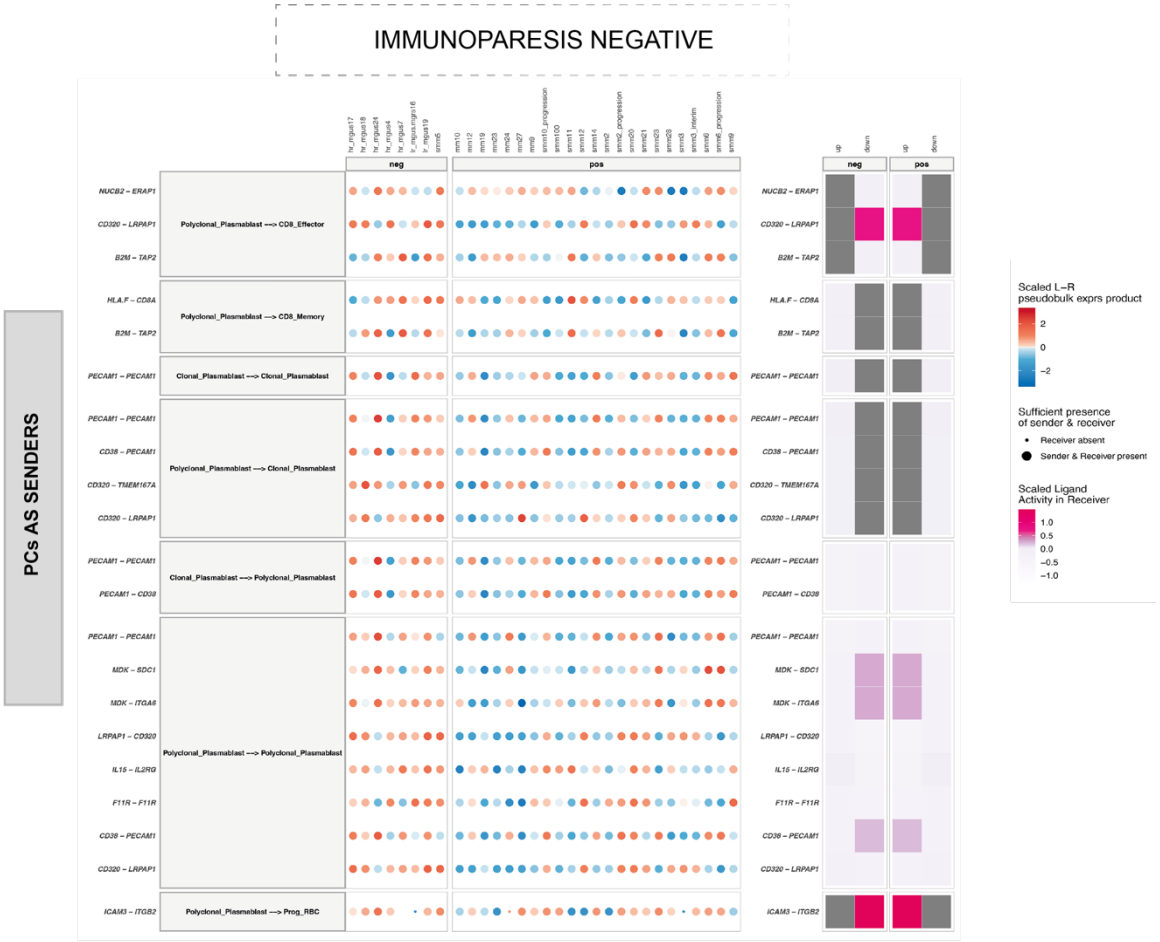

Figure 21 - Scaled ligand-receptor pseudobulk products and ligand activity in PCs as sender under immunoparesis-negative conditions. *Bubble plots show the scaled pseudobulk expression of ligand-receptor pairs specific to the "immunoparesis-negative" condition. Each row represents ligand-receptor interactions, with the sender and receiver cell types indicated accordingly. PCs were set as senders. The scaled ligand-receptor pseudobulk expression is detailed in the legend. The size of the dots reflects whether a sample had sufficient cells (≥ 4) to be included in the differential expression (DE) analysis. On the right, scaled ligand activity values, calculated from the DE genes of each cell type and represented as normalized z-scores, are shown for each receiver cell type. Higher values suggest greater enrichment of target genes for a particular ligand among the up- or downregulated genes in the "immunoparesis-negative" or "positive" groups. Gray indicates no DE genes (up or down) for those cell types due to the pipeline parameters.* *Image adapted from Da Vià et al. (2024).*

Figure 22

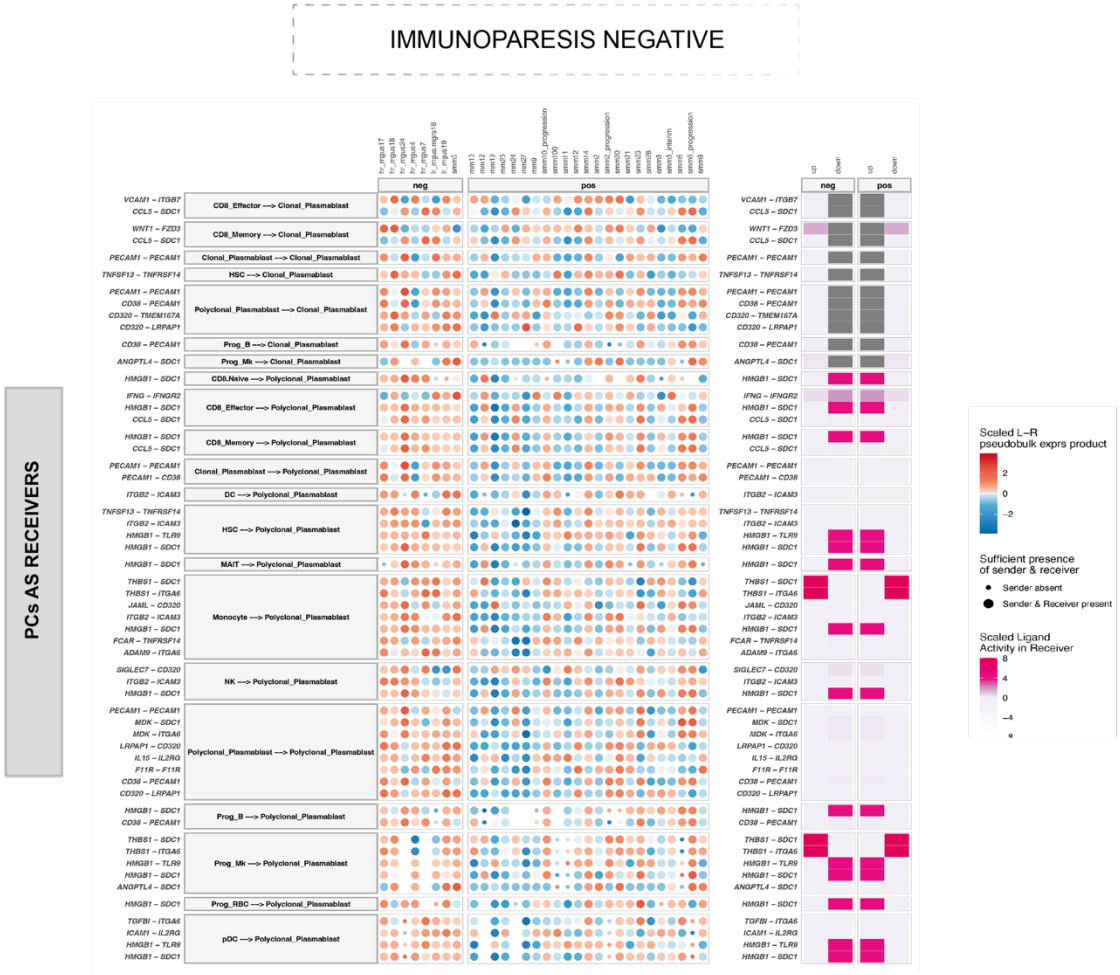

Figure 22 - Scaled ligand-receptor pseudobulk products and ligand activity in PCs as receiver under immunoparesis-negative conditions. *Bubble plots show the scaled pseudobulk expression of ligand-receptor pairs specific to the "immunoparesis-negative" condition. Each row represents ligand-receptor interactions, with the sender and receiver cell types indicated accordingly. PCs were set as receivers. The scaled ligand-receptor pseudobulk expression is detailed in the legend. The size of the dots reflects whether a sample had sufficient cells (≥ 4) to be included in the differential expression (DE) analysis. On the right, scaled ligand activity values, calculated from the DE genes of each cell type and represented as normalized z-scores, are shown for each receiver cell type. Higher values suggest greater enrichment of target genes for a particular ligand among the up- or downregulated genes in the "immunoparesis-negative" or "positive" groups. Gray indicates no DE genes (up or down) for those cell types due to the pipeline parameters.* *Image adapted from Da Vià et al. (2024).*

Figure 23

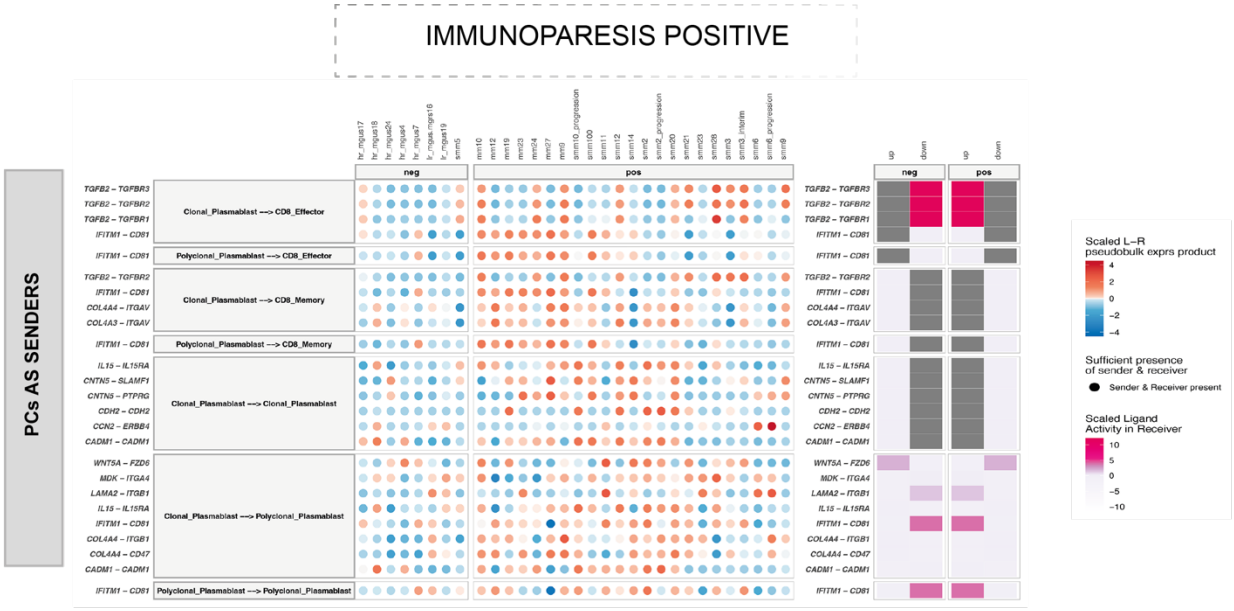

Figure 21 - Scaled ligand-receptor pseudobulk products and ligand activity in PCs as sender under immunoparesis-positive conditions. *Bubble plots show the scaled pseudobulk expression of ligand-receptor pairs specific to the "immunoparesis-positive" condition. Each row represents ligand-receptor interactions, with the sender and receiver cell types indicated accordingly. PCs were set as senders. The scaled ligand-receptor pseudobulk expression is detailed in the legend. The size of the dots reflects whether a sample had sufficient cells (≥ 4) to be included in the differential expression (DE) analysis. On the right, scaled ligand activity values, calculated from the DE genes of each cell type and represented as normalized z-scores, are shown for each receiver cell type. Higher values suggest greater enrichment of target genes for a particular ligand among the up- or downregulated genes in the "immunoparesis-negative" or "positive" groups. Gray indicates no DE genes (up or down) for those cell types due to the pipeline parameters. Image adapted from Da Vià et al. (2024).*

Figure 24

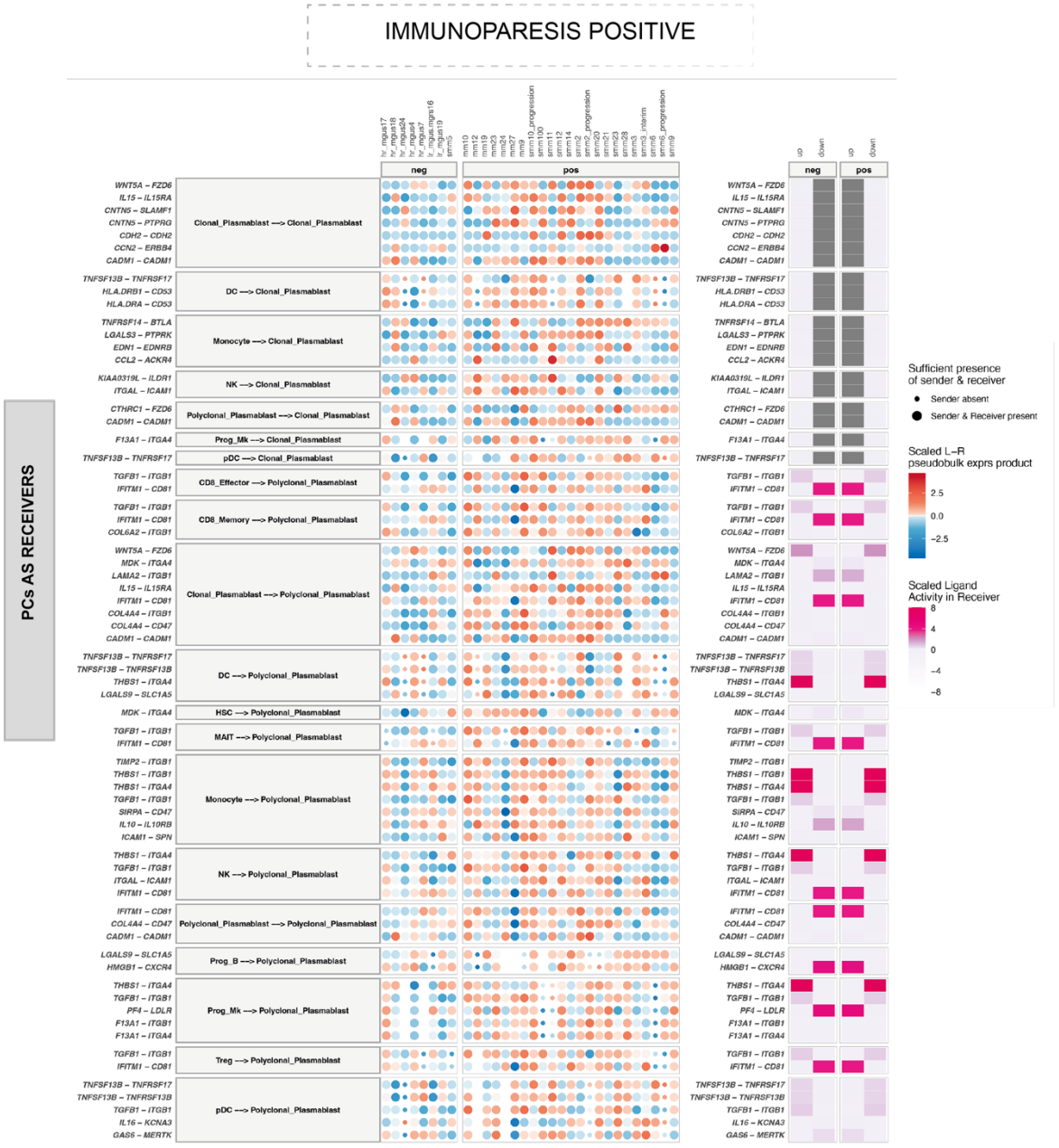

Figure 24 - Scaled ligand-receptor pseudobulk products and ligand activity in PCs as receiver under immunoparesis-positive conditions. *Bubble plots show the scaled pseudobulk expression of ligand-receptor pairs specific to the "immunoparesis-positive" condition. Each row represents ligand-receptor interactions, with the sender and receiver cell types indicated accordingly. PCs were set as receivers. The scaled ligand-receptor pseudobulk expression is detailed in the legend. The size of the dots reflects whether a sample had sufficient cells (≥ 4) to be included in the differential expression (DE) analysis. On the right, scaled ligand activity values, calculated from the DE genes of each cell type and represented as normalized z-scores, are shown for each receiver cell type. Higher values suggest greater enrichment of target genes for a particular ligand among the up- or downregulated genes in the "immunoparesis-negative" or "positive" groups. Gray indicates no DE genes (up or down) for those cell types due to the pipeline parameters.* *Image adapted from Da Vià et al. (2024).*

Figure 25

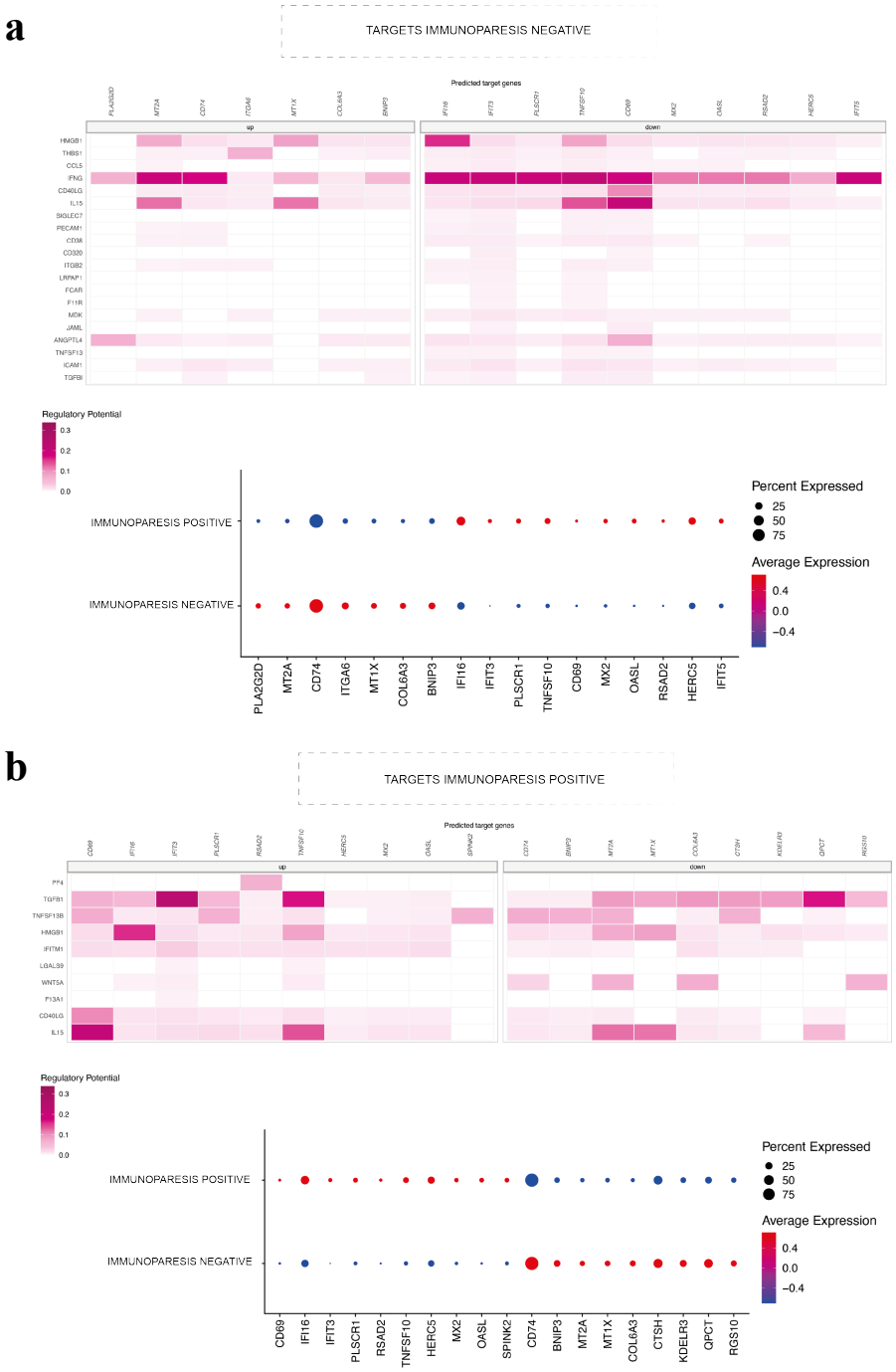

Figure 25 - Predicted ligand-target interactions and average expression levels of target genes in pPCs of immunoparesis-negative and -positive patients. *Heatmap displays the predicted ligand-target connections and the regulatory potential of target genes driving the ligand activity predictions with immunoparesis-negative (**a – upper panel**) and immunoparesis-positive (**b – upper panel**) prioritization. The columns represent the top target genes, while the rows correspond to the ligands that may contribute to the ligand activity prediction for each interaction. Dot plot illustrates the average expression levels of target genes that are either upregulated or downregulated in the predicted ligand-target interaction analysis. Upregulated genes are highlighted in red, while downregulated genes are shown in blue. These genes are identified from pPCs ligand-receptor-target (L-R-T) activity analyses under immunoparesis-negative (**a – lower panel**) and immunoparesis-positive (**b – lower panel**) conditions. Image adapted from Da Vià et al. (2024).*

Figure 26

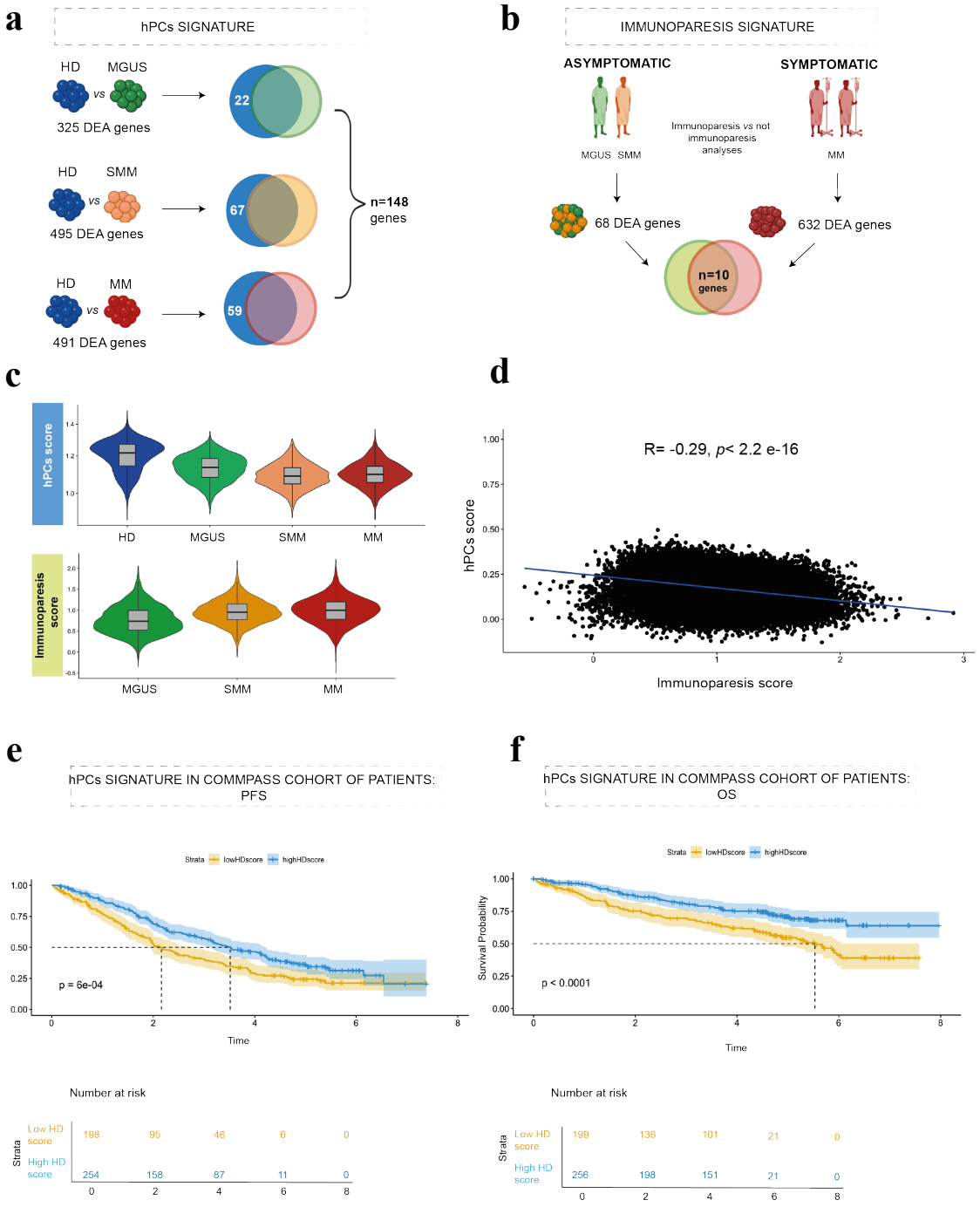
a
hPCs SIGNATURE
HD
MGUS
vs
325 DEA genes
22
HD
SMM
vs
495 DEA genes
67
HD
MM
vs
491 DEA genes
59
n=148 genes
b
IMMUNOPARESIS SIGNATURE
ASYMPTOMATIC
SYMPTOMATIC
MGUS SMM
Immunoparesis vs not immunoparesis analyses
MM
68 DEA genes
632 DEA genes
n=10 genes
c
hPCs score
HD
MGUS
SMM
MM
Immunoparesis score
MGUS
SMM
MM
d
R= -0.29, p< 2.2 e-16
hPCs score
Immunoparesis score
e
hPCs SIGNATURE IN COMMPASS COHORT OF PATIENTS: PFS
Strata
lowHDscore
highHDscore
Survival Probability
p = 6e-04
Time
Number at risk
Strata
Low HD score
198
95
46
6
0
High HD score
254
158
87
11
0
f
hPCs SIGNATURE IN COMMPASS COHORT OF PATIENTS: OS
Strata
lowHDscore
highHDscore
Survival Probability
p < 0.0001
Time
Number at risk
Strata
Low HD score
199
136
101
21
0
High HD score
256
198
151
21
0

Figure 26 – scRNAseq signatures can predict in independent datasets. ***a.*** *Schematic overview of "hPCs" signature set-up through DE analysis between HDs and patients with MGUS, SMM, and MM. For each comparison, genes significantly upregulated in HDs (B-H padj < 0.05, LFC > 0.1) were selected. Venn diagrams illustrate the overlap of deregulated genes across the different comparisons.* ***b.*** *Schematic representation of immunoparesis signature defined through DE analysis comparing immunoparesis-negative and immunoparesis-positive patients. Venn diagram shows the overlap of deregulated genes* ***c.*** *Violin plots display the inference of hPCs and immunoparesis signatures within the patient cohort.* ***d.*** *Scatterplot shows correlation between hPCs and immunoparesis signatures, with R: -0.29 and a p-value < 2.2e-16.* ***e. and f.*** *The clinical relevance of the hPCs signature assessed on CoMMpass dataset. Kaplan-Meier curves PFS (e) OS (f) for the patient cohort.* *Image adapted from Da Vià et al. (2024).*

Figure 27

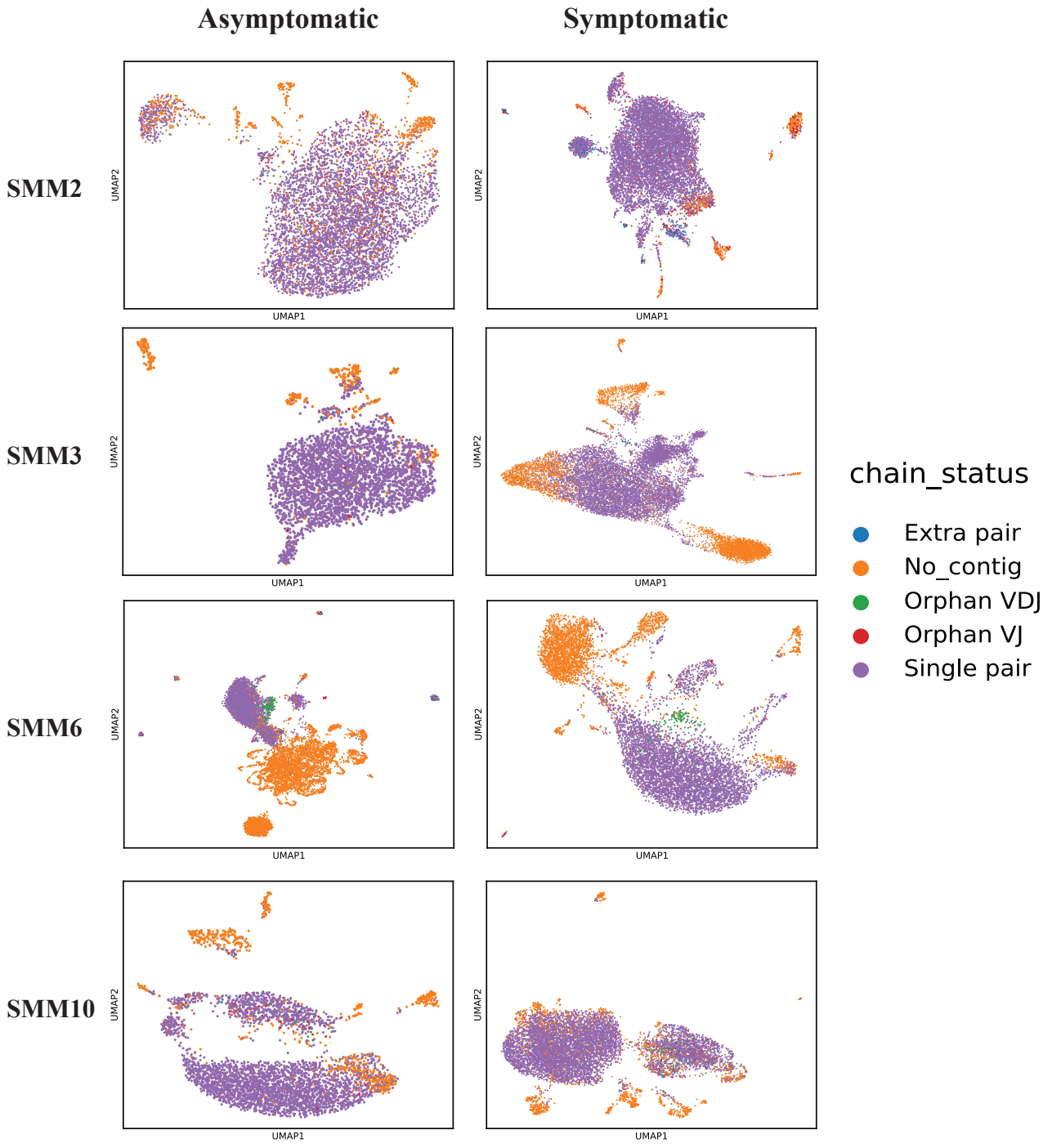

Figure 27 – VDJ chain status. *UMAP projection of cells belonging to 4 progressed patients (row) at diagnosis of asymptomatic MM (left column) and symptomatic MM (right column). Cells are colored by VDJ chain status (details in Material and Method). Generated through Dandelion.*

Figure 28

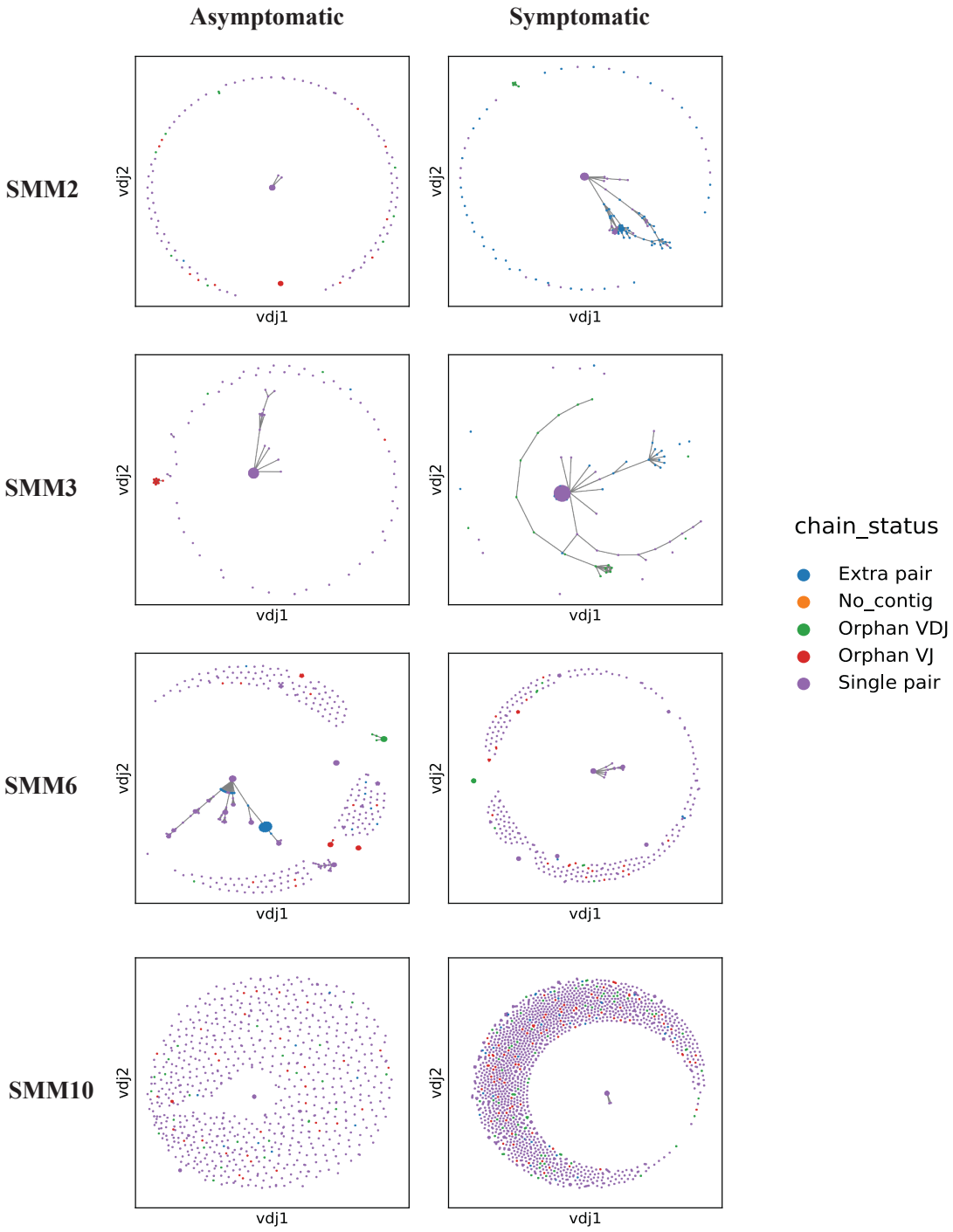

Figure 28 – VDJ network and chain status. *VDJ network generated through Dandelion. of cells belonging to 4 progressed patients (row) at diagnosis of asymptomatic MM (left column) and symptomatic MM (right column). Cells are colored by VDJ chain status (details in Material and Method). Nodes represent one or more clones with the same nucleotide CDR3 sequence. Edges and their length represent how much differences there are between two nodes in terms of nucleotide CDR3 sequence.*

Tables

Table 1 - Comparison of VDJ clonal rearrangements at the IGH locus identified by scRNAseq and amplicon-based NGS from bulk DNA in the 17/19 samples with matching results. In green samples from the same patient: PLC-06 sampled at SMM, PLC-22 sampled at MM.

Patient ID	scRNA-Seq			Amplicon-Based NGS from Bulk DNA			Common CDR3
	IGHV	IGHJ	Proportion	IGHV	IGHJ	Proportion	
PLC-01	IGHV3-48	IGHJ4	92.90%	IGHV3-30*09	IGHJ4*02	67.55%	CARDSYEDYVYW
PLC-02	IGHV3-21	IGHJ4	97.93%	IGHV3-21*01	IGHJ4*02	74.39%	CARYQLDAVAGKWGHYFDYW
PLC-03	IGHV3-43	IGHJ5	97.77%	IGHV3-9*01	IGHJ4*02	66.23%	CAKARLPLVGGLDSW
PLC-05	IGHV3-43	IGHJ6	49.79%	IGHV3-9*01	IGHJ6*02	13.62%	CTRVIGSGASCYDCYYHGMDVW
PLC-06	IGHV3-53	IGHJ5	87.91%	IGHV3-53*01	IGHJ4*02	57.58%	CARGLTAPGFPLDSW
PLC-07	IGHV3-23	IGHJ6	65.26%	IGHV3-23*01	IGHJ6*01	56.92%	CAKGRADCTDGVCYRRYGMDVW
PLC-08	IGHV3-43	IGHJ4	99.44%	IGHV3-43*01	IGHJ4*02	81.44%	CVKGQGGYTYGGFDCW
PLC-10	IGHV5-51	IGHJ4	77.63%	IGHV5-51*01	IGHJ4*02	57.39%	CARTNWPYYFDHW
PLC-12	IGHV3-43	IGHJ4	93.73%	IGHV3-9*01	IGHJ4*02	62.86%	CARDRYQLIIYYFDRW
PLC-15	IGHV3-43	IGHJ4	98.97%	IGHV3-9*01	IGHJ4*02	32.77%	CAKDVRYGYGSTQSAGFDYW
PLC-17	IGHV4-39	IGHJ4	95.53%	IGHV4-39*07	IGHJ4*02	22.10%	CARDKTTMTFSSPIFDYW
PLC-18	IGHV2-5	IGHJ1	95.78%	IGHV2-5*02	IGHJ1*01	56.58%	CAHSGSMWSGYAGTEYFQHW
PLC-19	IGHV4-59	IGHJ4	67.50%	IGHV4-59*01	IGHJ4*02	21.08%	CARAGDYDLLLLDYW
PLC-21	IGHV5-51	IGHJ6	80.02%	IGHV5-51*03	IGHJ6*03	56.06%	CARLPQGGYYYMDVW
PLC-22	IGHV3-53	IGHJ5	85.10%	IGHV3-53*01	IGHJ4*02	77.01%	CARGLTAPGFPLDSW
PLC-23	IGHV3-33	IGHJ1	56.11%	IGHV3-30-3*02	IGHJ1*01	61.13%	CAFAIGADGEYFQHW
PLC-24	IGHV2-70	IGHJ4	69.68%	IGHV2-70*01	IGHJ4*02	40.67%	CARGASETQVAMSTAELYFFDSW

Table 2 - Clonotype 1 V(D)J genes and CDR3 sequences of P1 and P4

Patient	Clonotype 1				
	Gene C	Gene V	Gene J	CDR3	
P1	IGHG1	IGHV5-51	IGHJ4	TGTGCGAGAGGCGTAGTGGCTACGATTACCGCTCTTGACTACTGG	CARGVVATITALDYW
	IGKC	IGKV1D-33	IGKJ3	TGTCAACAGTATGATAATCTCCCTATATTCACTTTC	CQQYDNLPIFTF
P4	IGHA2	IGHV5-10-1	IGHJ6	TGCATGCAAAATCTACAAACTCCTGGGTTT	CATSSYDAAHCSGPTCHHYYMDVW
	IGKC	IGKV2D-28	IGKJ2	TGTGCGACCTCCTCCTACGATGCCGCCCATTGTAGTGGTCCCACCTGCCACCACTACTACATGGATGTCTGG	CMQNLQTPGF

Table 3 - Clonotype 2 V(D)J genes and CDR3 sequences of P1 and P4

Patient	Clonotype 2				
	Gene C	Gene V	Gene J	CDR3	
P1	-	-	-	-	-
	IGKC	IGKV1D-33	IGKJ3	TGTCAACAGTATGATAATCTCCCTATATTC ACTTTC	CQQYDNLPIFTF
P4	IGHM	IGHV3-48	IGKJ6	TGTGCGAGAAGTCCTTATCACTATTACGGT CTGGACGTCTGG	CARSPYHYYGLDVW
	IGKC	IGKV3-20	IGKJ1	TGTCAGCACTATGGTAGTTCACCTATGTGG ACGTTC	CQHYGSSPMWTF

Table 4 - Number of cell types and immunoparesis status across 31 samples used for cell-cell interaction. Pt1

PTS	Immunoparesis	CD4 Mem	CD4 Naive	CD8 Naive	CD8 Eff	CD8 Mem	cPCs	DC	gdT	HSC	MAIT
MGUS17	NEG	288	84	17	24	54	768	12	1	85	1
MGUS18	NEG	104	69	5	56	50	5541	4	0	149	4
MGUS24	NEG	249	33	37	45	94	1785	2	2	51	14
MGUS4	NEG	222	124	12	101	187	1045	24	0	71	49
MGUS7	NEG	166	11	50	79	76	404	1	0	15	16
MGUS16	NEG	29	29	1	16	50	1586	7	0	56	3
MGUS19	NEG	275	85	1	137	290	1650	22	1	110	9
MM10	POS	847	724	323	292	272	3818	21	21	88	58
MM12	POS	337	290	80	309	131	4141	3	4	35	35
MM19	POS	1121	318	191	321	375	2966	47	52	355	71
MM23	POS	1071	189	83	681	612	2584	16	3	135	143
MM24	POS	165	18	1	422	921	6624	7	0	4	2
MM27	POS	2231	52	20	83	738	3930	6	27	66	631
MM9	POS	1137	299	45	476	124	2185	32	5	60	79
SMM10P	POS	742	58	2	283	645	5043	9	0	117	56
SMM100	POS	223	35	10	89	184	5216	32	2	363	25
SMM11	POS	592	80	119	52	82	406	3	9	29	92
SMM12	POS	226	81	7	227	105	4908	9	8	74	41
SMM14	POS	89	11	6	118	25	3275	9	2	122	33
SMM2	POS	221	19	0	149	440	4354	10	2	38	10
SMM2P	POS	111	23	9	29	177	6403	8	3	25	13
SMM20	POS	448	106	0	115	231	3597	15	4	148	13
SMM21	POS	247	126	25	31	347	637	30	1	79	2
SMM23	POS	448	96	21	155	466	3860	11	1	97	13
SMM28	POS	262	142	8	28	524	4011	0	7	38	18
SMM3	POS	88	34	2	15	15	2533	14	3	43	1
SMM3P	POS	927	473	101	302	119	7683	61	8	319	55
SMM5	NEG	105	44	8	27	25	378	4	0	25	42
SMM6	POS	102	10	1	14	47	1962	3	0	55	4
SMM6P	POS	215	6	0	52	56	2681	1	0	43	13
SMM9	POS	236	64	5	77	99	6220	22	1	85	2

Table 5 - Number of cell types and immunoparesis status across 31 samples used for cell-cell interaction. Pt2

PTS	Immunoparesis	Mem B	Monocyte	Naive B	NK	pDC	pPCs	Prog B	Prog Mk	Prog RBC	Treg
MGUS17	NEG	9	428	13	152	12	52	36	13	5	39
MGUS18	NEG	5	213	9	104	16	236	8	8	10	12
MGUS24	NEG	0	207	0	120	3	1262	15	0	23	6
MGUS4	NEG	3	168	0	124	12	405	35	5	16	9
MGUS7	NEG	0	195	0	66	4	355	4	0	0	2
MGUS16	NEG	0	121	0	8	10	316	2	5	1	8
MGUS19	NEG	12	504	18	15	19	477	28	13	23	9
MM10	POS	60	171	159	784	123	21	97	16	26	6
MM12	POS	51	124	51	19	0	23	1	8	6	0
MM19	POS	2	370	27	463	26	53	69	27	47	14
MM23	POS	1263	782	254	461	11	24	13	20	5	16
MM24	POS	29	28	0	273	0	52	0	6	2	6
MM27	POS	41	1327	27	1510	40	10	0	48	91	33
MM9	POS	16	189	1	204	75	53	3	21	25	58
SMM10P	POS	3	1975	18	215	24	1562	38	16	33	28
SMM100	POS	3	791	0	51	8	225	0	3	14	30
SMM11	POS	1	254	2	333	6	65	4	2	5	4
SMM12	POS	123	401	22	148	7	261	7	10	7	9
SMM14	POS	0	319	0	79	16	57	23	8	23	15
SMM2	POS	6	319	0	51	16	97	20	12	8	8
SMM2P	POS	5	247	0	26	3	146	1	43	40	6
SMM20	POS	3	906	4	113	18	210	2	13	10	32
SMM21	POS	2	293	2	126	16	201	21	15	28	9
SMM23	POS	1	609	2	230	10	190	23	5	15	5
SMM28	POS	1	895	3	148	15	66	5	16	39	23
SMM3	POS	1	208	9	138	12	52	29	4	1	4
SMM3P	POS	27	650	363	461	31	62	218	15	47	27
SMM5	NEG	0	127	0	20	10	421	21	6	11	2
SMM6	POS	1	87	8	62	6	184	28	2	9	3
SMM6P	POS	6	211	4	58	3	237	31	13	13	2
SMM9	POS	0	279	6	27	22	72	14	13	894	13