

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

MULTI-SCALE TUMOR ECOSYSTEM MODELING REVEALS SPATIAL AND SYSTEMIC DETERMINANTS OF THERAPEUTIC RESPONSE

Permalink

<https://escholarship.org/uc/item/4hh707c0>

ISBN

9798186188117

Author

mesrizadeh, zahra

Publication Date

2026-08-05

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

MULTI-SCALE TUMOR ECOSYSTEM MODELING REVEALS SPATIAL AND
SYSTEMIC DETERMINANTS OF THERAPEUTIC RESPONSE

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Bioengineering

by

Zahra Mesrizadeh

Committee in charge:

Shankar Subramaniam, Chair
Adam J. Engler
Prashant Mali
Kavitha Mukund
Sonia Ramamoorthy
Victoria Seewaldt

Copyright

Zahra Mesrizadeh, 2026

All rights reserved.

The Dissertation of Zahra Mesrizadeh is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2026

DEDICATION

To love and hope.

TABLE OF CONTENTS

DISSERTATION APPROVAL PAGE	iii
DEDICATION.....	iv
TABLE OF CONTENTS.....	v
LIST OF FIGURES	vii
LIST OF TABLES.....	x
ACKNOWLEDGEMENTS.....	xi
VITA.....	xiii
ABSTRACT OF THE DISSERTATION	xiv
Chapter 1 Introduction	1
REFERENCES.....	9
Chapter 2 A novel subtyping method for tripe-negative breast cancer with implications for prognosis and therapy	11
2.1 Abstract.....	11
2.2 Introduction.....	11
2.3 Patients and Methods	13
2.4 Results	16
2.5 Discussion	28
2.6 Supplemental Materials.....	32
REFERENCES.....	39
Chapter 3 Spatial context of cellular heterogeneity in triple-negative breast cancer reveals actionable subtypes	46
3.1 Abstract.....	46
3.2 Introduction.....	46
3.3 Patients and Methods	48
3.4 Results	52

3.5 Discussion	63
3.6 Supplemental Materials.....	67
REFERENCES.....	71
Chapter 4 Differential peripheral immune dynamics underlie therapeutic response to chemotherapy and chemo-immunotherapy in triple-negative breast cancer	76
4.1 Abstract.....	76
4.2 Introduction.....	77
4.3 Methods.....	79
4.4 Results	84
4.5 Discussion	118
4.6 Supplemental Materials.....	123
REFERENCES.....	131
Chapter 5 A Spatially resolved tumor-microenvironment architecture defines therapeutic response in unresectable hepatoblastoma.....	136
5.1 Abstract.....	136
5.2 Introduction.....	137
5.3 Patients and Methods	138
5.4 Results	147
5.5 Discussion	163
5.6 Supplemental Materials.....	166
REFERENCES.....	174
Chapter 6 Conclusions.....	179
6.1 Summary of findings	179
6.2 Conceptual contribution.....	180
6.3 Translational implications	181
6.4 Limitations and future directions.....	181

LIST OF FIGURES

Figure 1.1: Conceptual framework for multi-scale tumor ecosystem analysis.	8
Figure 2.1: TNBC subtyping and characterization based on cell type marker expression and functional features.	19
Figure 2.2: TNBC subtyping and characterization based on cell type marker expression and functional features, Continued.	20
Figure 2.3: Differentially expressed genes and subtype-specific networks in LTR TNBC.	25
Figure 2.4: Prognostic relevance and subtype classification robustness of TNBC subtypes.	28
Figure 3.1: Spatial transcriptomic characterization of TNBC subtypes.	54
Figure 3.2: Spatial transcriptomic characterization of TNBC subtypes, Continued.	55
Figure 3.3: Integrated spatial characterization of tumor and immune ecosystems across TNBC subtypes.	58
Figure 3.4: Integrated spatial characterization of tumor and immune ecosystems across TNBC subtypes, Continued.	59
Figure 3.5: Spatially resolved regulatory networks across TNBC subtypes and invasive regions.	62
Figure 4.1: Comparative immune landscape of blood samples under different treatment conditions.	87
Figure 4.2: Comparative immune landscape of blood samples under different treatment conditions, Continued.	88
Figure 4.3: Comprehensive landscape of CD4⁺ T cell subsets in blood samples and their distribution across conditions.	91
Figure 4.4: Comprehensive landscape of CD4⁺ T cell subsets in blood samples and their distribution across conditions, Continued.	92
Figure 4.5: Comprehensive landscape of CD8⁺ T cell subsets in blood samples and their distribution across conditions.	96
Figure 4.6: Comprehensive landscape of CD8⁺ T cell subsets in blood samples and their distribution across conditions, Continued.	97
Figure 4.7: Comprehensive landscape of B cell subsets in blood samples and their distribution across conditions.	100
Figure 4.8: Comprehensive landscape of B cell subsets in blood samples and their distribution across conditions, Continued.	101
Figure 4.9: Landscape and dynamics of NK cell subsets in blood across treatment conditions.	103

Figure 4.10: Landscape and dynamics of NK cell subsets in blood across treatment conditions, Continued.	104
Figure 4.11: Landscape and dynamics of monocyte subsets in blood across treatment conditions.	108
Figure 4.12: Comprehensive landscape of CD4+ and CD8+ T cell subsets in tissue samples and their distribution across conditions.	110
Figure 4.13: Comprehensive landscape of CD4+ and CD8+ T cell subsets in tissue samples and their distribution across conditions, Continued.	111
Figure 4.14: Protein-protein interaction networks of treatment- and response-specific mechanisms in tissue.	113
Figure 4.15: Protein-protein interaction networks of treatment- and response-specific mechanisms in blood.	114
Figure 4.16: Protein-protein interaction networks of treatment- and response-specific mechanisms in blood, Continued.	115
Figure 5.1: Integrative spatial transcriptomics workflow and cross-platform mapping of tumor architecture in relapse-prone unresectable hepatoblastoma.	148
Figure 5.2: Spatially resolved tumor ecosystems across hepatoblastoma clinical groups.	150
Figure 5.3: Spatially resolved tumor ecosystems across hepatoblastoma clinical groups, Continued.	151
Figure 5.4: Spatial neighborhood enrichment analysis identifies cell-state co-localization patterns across hepatoblastoma clinical groups.	158
Figure 5.5: Spatially resolved signaling programs define relapse-associated tumor architecture and prognostic signatures.	160
Figure 5.6: Spatially resolved signaling programs define relapse-associated tumor architecture and prognostic signatures, Continued.	161
Figure 5.7: Distinct spatial architectures and signaling programs define relapse-associated and non-relapse hepatoblastoma microenvironments.	162
Supplementary Figure 2.1: Additional characterization of TNBC subtype composition and cellular deconvolution.	33
Supplementary Figure 2.1: Additional characterization of TNBC subtype composition and cellular deconvolution, Continued.	34
Supplementary Figure 2.2: Functional enrichment and nuclear receptor-associated programs across TNBC subtypes.	35
Supplementary Figure 2.3: ESR1-associated transcriptional and histologic features in CC3 TNBC.	36
Supplementary Figure 2.3: ESR1-associated transcriptional and histologic features in CC3 TNBC, Continued.	37

Supplementary Figure 2.4: Prognostic biomarkers and single-cell validation across TNBC subtypes.....	38
Supplementary Figure 3.1: Additional spatial characterization of TNBC subtype ecosystems.....	67
Supplementary Figure 3.2: Regional EcoTyper cell-state distributions across TNBC subtypes.....	68
Supplementary Figure 3.2: Regional EcoTyper cell-state distributions across TNBC subtypes, Continued.....	69
Supplementary Figure 3.3: Additional spatial regulatory network analyses across TNBC regions.....	70
Supplementary Figure 4.1: Comparative immune landscape of tissue samples under different treatment conditions.	124
Supplementary Figure 4.1: Comparative immune landscape of tissue samples under different treatment conditions, Continued.	125
Supplementary Figure 4.2: Comprehensive landscape of CD8⁺ T cell subsets in tissue samples and their distribution across conditions.....	126
Supplementary Figure 4.3: Comprehensive landscape of CD8⁺ T cell subsets in tissue samples and their distribution across conditions.....	128
Supplementary Figure 4.3: Comprehensive landscape of CD8⁺ T cell subsets in tissue samples and their distribution across conditions, Continued.....	129
Supplementary Figure 4.4: Peripheral immune signatures align with tumor immune programs and jointly associate with pCR.....	130
Supplementary Figure 5.1: Spatial deconvolution reveals distinct cellular organization and transcriptional structure across clinical groups.....	168
Supplementary Figure 5.1: Spatial deconvolution reveals distinct cellular organization and transcriptional structure across clinical groups, Continued.....	169
Supplementary Figure 5.2: Additional spatial and cellular characterization across hepatoblastoma clinical groups.	170
Supplementary Figure 5.2: Additional spatial and cellular characterization across hepatoblastoma clinical groups, Continued.	171
Supplementary Figure 5.3: Distinct tumor and TME regulatory networks define clinical group-specific microenvironmental programs.....	172
Supplementary Figure 5.3: Distinct tumor and TME regulatory networks define clinical group-specific microenvironmental programs, Continued.....	173

LIST OF TABLES

Supplementary Table 2.1: List of all the cell type-markers.....	32
Supplementary Table 4.1: Patient demographics.....	123
Supplementary Table 5.1: Patient demographics.....	166
Supplementary Table 5.2: Tumor-level clinical characteristics	167

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Professor Shankar Subramaniam, for his mentorship, guidance, and unwavering support throughout my PhD. His vision, scientific insight, and encouragement to think beyond conventional boundaries have profoundly shaped my approach to research and problem-solving.

I am grateful to the members of my dissertation committee for their valuable feedback, thoughtful discussions, and support throughout this work. Their perspectives have strengthened both the rigor and clarity of this dissertation. I would like to thank my collaborators and colleagues, whose expertise and collaboration made this work possible. I am especially grateful to Dr. Rakesh Sindhi and Dr. Mylarappa Ningappa for their clinical insight, collaboration, and contributions that significantly enriched the hepatoblastoma research.

I also gratefully acknowledge Wellcome Leap for providing the opportunity to travel and collaborate with international scientists. I am especially thankful to Jason Swedlow for his leadership and support, which made these collaborative experiences possible and broadened my scientific perspective.

I am thankful to my lab mates and friends for creating an environment of intellectual curiosity, support, and camaraderie. The shared challenges and successes have made this journey both meaningful and enjoyable. I am deeply grateful to my friends and family for their unconditional love, patience, and belief in me. Their support has been the foundation that allowed me to pursue this path with confidence and resilience.

Finally, I acknowledge all the patients whose samples and data made this research possible. Their contributions are the foundation of this work and a constant reminder of its purpose.

Chapter 2 and Chapter 3, are currently being prepared for submission for publication of the material. Zahra Mesrizadeh, et al. The dissertation author was the primary researcher and author of this material.

Chapter 4, is currently under review at ImmunoHorizons. Zahra Mesrizadeh, et al. The dissertation author was the primary researcher and author of this material.

Chapter 5, is currently being prepared for submission for publication of the material. Zahra Mesrizadeh, et al. The dissertation author was the primary researcher and author of this material.

VITA

2020 Bachelor of Science in Biomedical Engineering, California State University Long Beach

2026 Doctor of Philosophy in Bioengineering, University of California San Diego

PUBLICATIONS

Spatially distinct cellular and molecular landscapes define prognosis in triple-negative breast cancer

Zahra Mesrizadeh, et al. (co-author). **Cell Reports**

Immunogenomic diversity of triple-negative breast cancers in obese and non-obese black and white women

Zahra Mesrizadeh, et al. (co-author). **npj Breast Cancer**

Peripheral immune dynamics predict therapeutic response in triple-negative breast cancer

Zahra Mesrizadeh, et al. **Under review (ImmunoHorizons)**

Spatial context of cellular heterogeneity in triple negative breast cancer reveals actionable subtypes

Zahra Mesrizadeh, et al. **In preparation**

Spatially resolved tumor-microenvironment architecture defines therapeutic response in unresectable hepatoblastoma

Zahra Mesrizadeh, et al. **In preparation**

ABSTRACT OF THE DISSERTATION

MULTI-SCALE TUMOR ECOSYSTEM MODELING REVEALS SPATIAL AND SYSTEMIC DETERMINANTS OF THERAPEUTIC RESPONSE

by

Zahra Mesrizadeh

Doctor of Philosophy in Bioengineering

University of California San Diego, 2026

Professor Shankar Subramaniam, Chair

Advances in high-throughput sequencing and spatial profiling technologies have enabled unprecedented characterization of tumor biology; however, most approaches interrogate tumor-intrinsic programs, the tumor microenvironment, and systemic immunity in isolation, limiting mechanistic understanding of therapeutic response. In this dissertation, we develop an integrative, multi-scale framework to define tumor ecosystem states that capture the coordinated interactions between malignant cells, stromal compartments, and immune populations across molecular,

spatial, and systemic levels. Using bulk RNA sequencing of triple-negative breast cancer (TNBC) tumors, we identify three reproducible tumor ecosystem states (subtypes) defined by distinct combinations of epithelial, immune, and stromal programs. These subtypes are associated with differential survival outcomes and are governed by subtype-specific regulatory networks, including interferon-driven immune activation, stromal fibrosis and hypoxia, and transcriptional plasticity linked to nuclear receptor signaling, epithelial-to-mesenchymal transition, and metabolic reprogramming. Notably, we identify a TNBC subtype characterized by elevated estrogen receptor (ESR1) mRNA expression despite the absence of protein expression, associated with inflammatory and lipid signaling pathways and recurrent sequence variants, suggesting a non-canonical regulatory mechanism. To resolve how these subtypes are organized within tissue, we integrate spatial transcriptomics and multiplexed imaging, revealing that tumor-intrinsic and microenvironmental programs are spatially compartmentalized across epithelial and stromal regions. These spatial architectures govern immune engagement, stromal interactions, and invasive behavior. Complementary analysis of peripheral blood mononuclear cells further reveals that coordinated adaptive and innate immune dynamics distinguish therapeutic response, with pre-existing immune priming and effector activity associated with response, and persistent immune dysregulation and myeloid remodeling underlying resistance. Finally, we extend this framework to hepatoblastoma, demonstrating that therapeutic response is encoded in spatial tumor-microenvironment architectures, where relapse-prone tumors exhibit centralized, stromal-dominated niches driven by macrophage-fibroblast signaling, in contrast to immune-engaged, spatially distributed organization in non-relapse disease. Collectively, this work establishes a unified model in which tumor behavior emerges from spatially and systemically coordinated

ecosystem states. These integrative frameworks provide a foundation for biomarker development and rational therapeutic stratification across heterogeneous solid tumors.

Chapter 1 Introduction

Solid tumors remain a leading cause of mortality worldwide, with breast and liver cancers among the most lethal tumor types (1). Despite advances in genomic profiling and therapeutic development, clinical outcomes remain highly variable among patients with the same histopathologic diagnosis (2,3). Differences in survival, chemotherapy sensitivity, and responsiveness to immunotherapy are frequently observed, underscoring that tumor behavior cannot be explained by malignant cell genetics alone. Cancer is increasingly understood as a dynamic and adaptive disease, characterized by plasticity and heterogeneity that evolve across genetic, phenotypic, and pathological levels. Consequently, understanding the sources of tumor and tumor microenvironment heterogeneity is essential for improving prognostic stratification and therapeutic decision-making.

Cancer has long been conceptualized as a disease driven by the accumulation of somatic mutations, a framework that has yielded important insights into oncogenic drivers and enabled molecular classification of histologically similar tumors. In breast cancer, for example, alterations in *TP53*, *PIK3CA*, and *ERBB2* define major clinical subtypes (4). However, genetic alterations alone fail to account for the substantial clinical heterogeneity observed within genetically similar tumors. Selection pressures in cancer act on phenotypes rather than genotypes, and tumor cells can adopt diverse functional states that are not directly encoded by DNA sequence. These observations have shifted attention toward non-genetic sources of heterogeneity, including epigenetic regulation and cellular plasticity, as critical contributors to disease progression and therapeutic resistance (4).

The tumor microenvironment (TME) has emerged as a major determinant of solid tumor growth, metastatic progression, and therapeutic response. Tumors exist within a structured and dynamic ecosystem composed of diverse immune populations, cancer-associated fibroblasts

(CAFs), endothelial cells, pericytes, and tissue-specific stromal elements, all embedded within a remodeled and vascularized extracellular matrix (5). Through direct cell-cell interactions and paracrine signaling, malignant and non-malignant compartments jointly regulate tumor behavior.

The inseparability of tumor-intrinsic and microenvironmental signals was evident in early transcriptomic subtyping efforts. For example, the original Lehmann classification of triple-negative breast cancer identified immune- and stromal-enriched subtypes from bulk RNA sequencing data, reflecting composite transcriptional programs derived from multiple cellular sources (6). Subsequent revisions of this framework removed these categories to focus on epithelial-intrinsic states (7), not because the microenvironment was biologically irrelevant, but because bulk profiling lacked the resolution to disentangle cellular contributions. This limitation underscores the need for analytic approaches capable of resolving tumor-intrinsic programs alongside microenvironmental phenotypes.

This challenge is particularly evident in tumor-immune interactions, where immune cells are often present within tumors despite ongoing tumor progression. Although cytotoxic innate and adaptive immune cells can recognize and eliminate malignant cells, tumors frequently progress to clinically apparent disease. This progression reflects the ability of cancer cells to acquire immune-evasive strategies that parallel mechanisms of peripheral immune tolerance, thereby constraining effective anti-tumor immunity. The dynamic interplay between neoplastic cells and the immune microenvironment shapes tumor evolution through a process often described as immunoediting. During early stages of tumor development, cytotoxic lymphocytes, including CD8⁺ T cells and natural killer cells, can eliminate highly immunogenic cancer cells. This selective pressure favors the outgrowth of less immunogenic or immune-evasive variants that persist and expand. In later stages, immune populations can adopt tumor-promoting roles, with myeloid cells and macrophages

supporting invasion, stromal remodeling, and metastatic dissemination through inflammatory and growth factor signaling (8,9).

Bulk tumor profiling approaches have enabled sensitive detection of molecular biomarkers and informed diagnostic, prognostic, and therapeutic decision-making across cancer types (10). However, these methods inherently measure population-averaged signals, conflating malignant and non-malignant contributions and obscuring the cellular origins of key pathways. In addition, bulk profiling lacks spatial resolution, preventing assessment of how molecular programs are distributed across tumor and microenvironmental compartments. Single-cell RNA sequencing addresses cellular heterogeneity by resolving transcriptional states at the level of individual cells, yet this dissociative approach disrupts tissue architecture, introduces sampling biases, and remains resource-intensive, limiting clinical scalability. The importance of spatial context in cancer biology has been recognized since the earliest histopathological studies, beginning with Johannes Müller's 19th-century microscopic examinations of tumors, which demonstrated that malignant progression and tissue organization are intrinsically linked (11,12). Subsequent histological analyses established that immune infiltration, stromal remodeling, and invasive behavior are spatially patterned. However, conventional histopathology lacks molecular resolution. Together, these limitations underscore a clinical gap that existing approaches cannot simultaneously capture molecular programs, cellular identity, and spatial organization, constraining mechanistic understanding of immune-tumor interactions.

Spatially resolved multiplexed molecular assays have recently emerged as a solution to the long-standing disconnect between comprehensive molecular profiling and preserved tissue architecture. These technologies enable high-plex measurement of RNA and protein targets directly within intact clinical specimens, allowing molecular programs to be interpreted in their

native spatial context. Unlike dissociative single-cell approaches, spatial profiling preserves cellular organization and microenvironmental structure enabling interrogation of tumor-immune interface, stromal barriers, and spatially restricted signaling niches.

While spatial profiling resolves immune-tumor interactions within tissue, anti-tumor immunity is coordinated at the systemic level, where circulating immune populations influence treatment response, disease progression, and recurrence. Systemic immune states both shape and reflect the capacity of the host to mount effective anti-tumor responses, and their dynamics evolve during disease progression and therapeutic exposure. Peripheral immune profiling enables longitudinal assessment of these changes, providing insight into immune activation, suppression, and reprogramming over time. Although circulating immune biomarkers have demonstrated predictive value across multiple cancer types, their interpretation remains limited without mechanistic understanding of how systemic immune states interact with tumor-intrinsic and microenvironmental programs. Studying systemic immunity alongside tumor and spatial profiling therefore provides a more comprehensive view of tumor ecosystem dynamics during therapy.

These challenges are particularly evident in triple-negative breast cancer (TNBC), a clinically aggressive and molecularly heterogeneous breast cancer subtype that lacks targeted therapeutic options and exhibits highly variable immune engagement and treatment response (13–16). TNBC tumors display marked diversity in transcriptional programs, microenvironmental composition, and spatial organization, making them a compelling model for studying tumor-immune ecosystem dynamics. Consequently, multiple efforts have sought to stratify TNBC into biologically and clinically meaningful subtypes, primarily using transcriptomic profiling approaches.

Early efforts to stratify TNBC relied on transcriptomic profiling to capture tumor-intrinsic heterogeneity. A study by Lehmann and colleagues (6) classified TNBC into six transcriptional subtypes (BL1, BL2, M, MSL, IM, and LAR) based on gene expression patterns, providing one of the first systematic frameworks linking molecular heterogeneity to potential therapeutic vulnerabilities. These subtypes reflected distinct biological programs, including cell cycle and DNA damage signaling in basal-like tumors, androgen receptor-driven signaling in LAR tumors, mesenchymal and epithelial-mesenchymal transition (EMT) programs in M/MSL tumors, and immune-associated transcriptional signatures in IM tumors. Subsequent refinement of this framework reduced the classification to four primary subtypes (BL1, BL2, LAR, and M), as immune and stromal signatures were recognized to arise largely from the tumor microenvironment rather than epithelial tumor cells (7). While these efforts substantially advanced understanding of TNBC heterogeneity and informed therapeutic hypotheses, they also highlighted the challenges of disentangling tumor-intrinsic and microenvironmental signals using bulk profiling alone.

Building on these early efforts, additional classification systems sought to refine TNBC heterogeneity by incorporating immune signaling, stromal features, and clinicopathologic characteristics. Burstein and colleagues (17) identified four subgroups (LAR, BLIS, BLIA, and MES), emphasizing immune activation and suppression as key axes of variability. Similarly, the Jezequel (18) framework integrated molecular signaling, immune phenotypes, and clinical outcomes to define subtypes associated with androgen signaling, macrophage-dominant microenvironments, and immune-reactive basal tumors. The FUSCC classification (19) further incorporated non-coding RNA profiles to capture regulatory heterogeneity. Collectively, these studies underscored that TNBC subtypes are shaped not only by epithelial transcriptional programs but also by immune and stromal contributions.

Despite their contributions, existing TNBC classification frameworks share several limitations. Most rely on bulk transcriptomic data, which conflates tumor-intrinsic and microenvironmental signals and obscures spatial organization within tumors. Immune subtypes are often inferred indirectly, without direct measurement of cellular states or spatial interactions. In addition, many frameworks were derived from limited cohorts and lack consistent external validation across treatment contexts. As a result, current subtyping approaches provide limited insight into how tumor-intrinsic programs, immune dynamics, and spatial organization jointly shape therapeutic response. These gaps motivate the need for integrative frameworks that combine bulk, spatial, and systemic immune profiling to more accurately define tumor ecosystem states.

While these questions are particularly evident in TNBC, the broader challenge of understanding tumor-immune ecosystem dynamics extend across solid malignancies. Pediatric tumor, including hepatoblastoma, exhibit similarly complex interactions between malignant cells, stromal compartments, and immune populations that influence disease progression and therapeutic response (20,21). Hepatoblastoma, the most common primary liver malignancy in children, demonstrates substantial heterogeneity in tumor architecture, immune infiltration, and clinical outcomes, despite relatively constrained genomic landscape (22,23). Notably, patients with clinically similar disease can experience markedly different outcomes, including relapse following treatment. These features suggest that non-genetic determinant, including spatial organization and microenvironmental context, may play a critical role in shaping disease behavior. Studying hepatoblastoma therefore provides an opportunity to examine whether an ecosystem oriented analytical approach, linking transcriptional programs to spatial cellular composition and microenvironmental architecture, as applied in TNBC, can reveal clinically meaningful patterns in a biologically distinct solid tumor.

Taken together, these observations highlight the need for approaches that move beyond tumor-intrinsic molecular profiling toward integrative frameworks capable of capturing tumor-microenvironment interactions across spatial and systemic scales (**Figure 1.1**). This dissertation addresses this need by developing and applying multi-scale analytical strategies to delineate tumor ecosystem states and their relationship to therapeutic response. By integrating bulk transcriptomic profiling, spatially resolved molecular measurements, and systemic immune characterization, this work aims to define how tumor-intrinsic programs, microenvironmental organization, and circulating immune dynamics collectively shape disease behavior. Within TNBC, this framework is used to refine transcriptional subtyping by incorporating spatial context and immune architecture, and to examine how systemic immune responses evolve during therapy. These analyses seek to clarify how distinct ecosystem states influence treatment sensitivity and resistance. Extending beyond TNBC, spatial transcriptomic characterization of hepatoblastoma is used as an independent test of this integrative approach, evaluating whether similar relationships between tumor organization, immune dynamics, and clinical features are observed in a developmentally and biologically distinct malignancy.

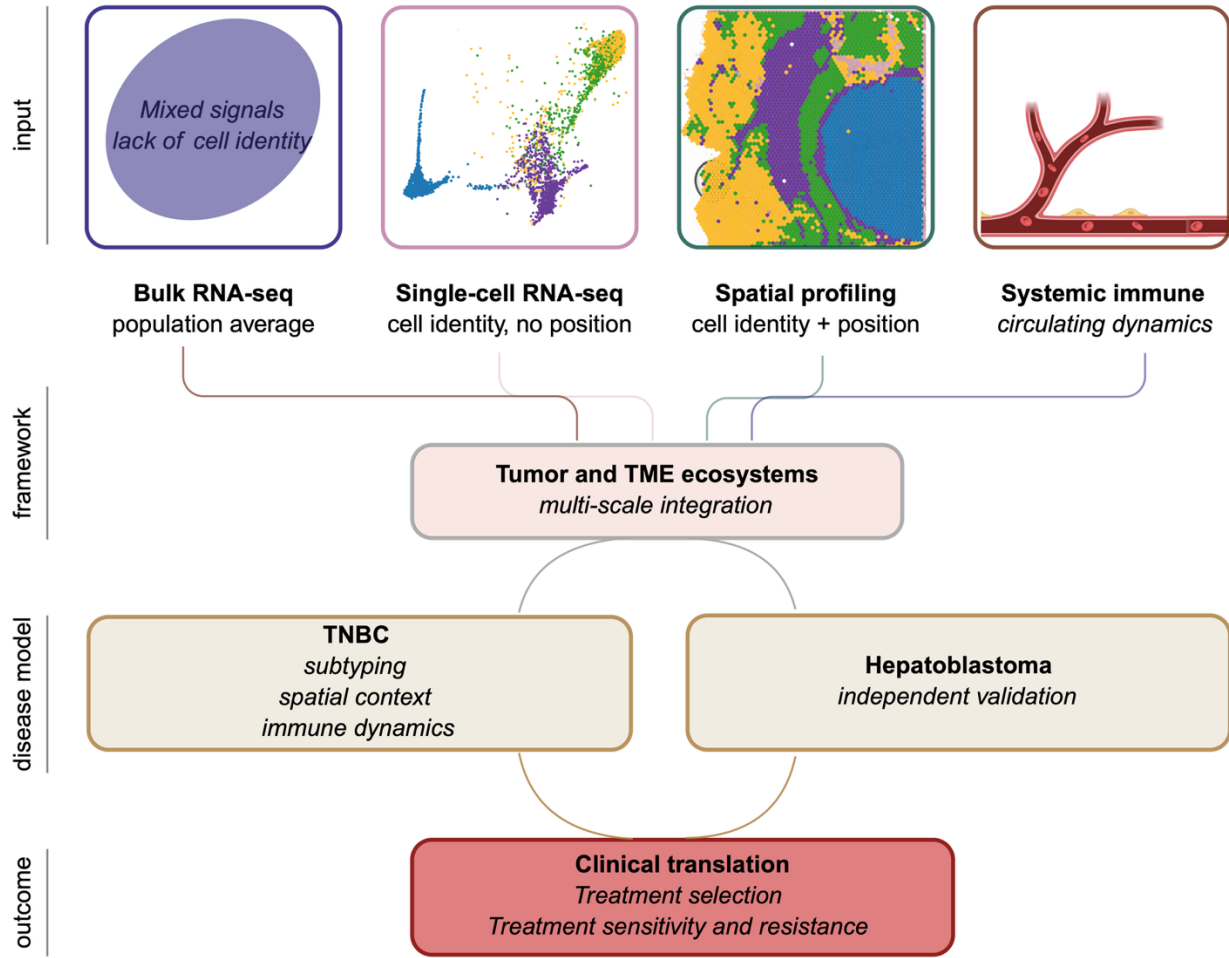


Figure 1.1: Conceptual framework for multi-scale tumor ecosystem analysis.

Schematic overview of the integrative framework developed in this dissertation. Traditional profiling approaches capture distinct aspects of tumor biology, including bulk RNA-seq (population-averaged signals), single-cell RNA-seq (cell identity without spatial context), spatial profiling (cell identity with spatial organization), and systemic immune profiling (circulating immune dynamics). These modalities are integrated to define tumor and tumor microenvironment (TME) ecosystem states across molecular, spatial, and systemic scales. This framework is applied to triple-negative breast cancer (TNBC) to characterize subtype-specific spatial and immune dynamics and extended to hepatoblastoma as an independent validation model. Ultimately, this multi-scale integration links tumor ecosystem organization to clinical translation, enabling improved treatment selection and understanding of therapeutic sensitivity and resistance.

REFERENCES

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*. 2024;74(3):229–63. doi:10.3322/caac.21834
2. Marusyk A, Janiszewska M, Polyak K. Intratumor Heterogeneity: The Rosetta Stone of Therapy Resistance. *Cancer Cell*. 2020 Apr 13;37(4):471–84. doi:10.1016/j.ccell.2020.03.007
3. Zhang A, Miao K, Sun H, Deng CX. Tumor heterogeneity reshapes the tumor microenvironment to influence drug resistance. *International Journal of Biological Sciences*. 2022 Apr 24;18(7):3019–33. doi:10.7150/ijbs.72534
4. Santarpia L, Bottai G, Kelly CM, Győrffy B, Székely B, Pusztai L. Deciphering and Targeting Oncogenic Mutations and Pathways in Breast Cancer. *Oncologist*. 2016 Sep;21(9):1063–78. doi:10.1634/theoncologist.2015-0369
5. de Visser KE, Joyce JA. The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. *Cancer Cell*. 2023 Mar 13;41(3):374–403. doi:10.1016/j.ccell.2023.02.016
6. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, Pietersen JA. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest*. 2011 Jul;121(7):2750–67. doi:10.1172/JCI45014
7. Lehmann BD, Jovanović B, Chen X, Estrada MV, Johnson KN, Shyr Y, Moses HL, Sanders ME, Pietersen JA. Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PLoS One*. 2016;11(6):e0157368. doi:10.1371/journal.pone.0157368
8. El-Kenawi A, Hänggi K, Ruffell B. The Immune Microenvironment and Cancer Metastasis. *Cold Spring Harb Perspect Med*. 2020 Apr;10(4):a037424. doi:10.1101/cshperspect.a037424
9. Gonzalez H, Hagerling C, Werb Z. Roles of the immune system in cancer: from tumor initiation to metastatic progression. *Genes Dev*. 2018 Oct 1;32(19–20):1267–84. doi:10.1101/gad.314617.118
10. Smith KL, Isaacs C. BRCA Mutation Testing in Determining Breast Cancer Therapy. *Cancer J*. 2011 Nov;17(6):492–9. doi:10.1097/PPO.0b013e318238f579
11. Haggard HW, Smith GM. Johannes Müller and the Modern Conception of Cancer. *Yale J Biol Med*. 1938 May;10(5):419.b1-436.
12. Gong D, Arbesfeld-Qiu JM, Perrault E, Bae JW, Hwang WL. Spatial oncology: Translating contextual biology to the clinic. *Cancer Cell*. 2024 Oct 14;42(10):1653–75. doi:10.1016/j.ccell.2024.09.001

13. Derakhshan F, Reis-Filho JS. Pathogenesis of Triple-Negative Breast Cancer. *Annu Rev Pathol.* 2022 Jan 24;17:181–204. doi:10.1146/annurev-pathol-042420-093238
14. Garrido-Castro AC, Lin NU, Polyak K. Insights into Molecular Classifications of Triple-Negative Breast Cancer: Improving Patient Selection for Treatment. *Cancer Discov.* 2019 Feb 8;9(2):176–98. doi:10.1158/2159-8290.CD-18-1177
15. de Ruijter TC, Veeck J, de Hoon JPJ, van Engeland M, Tjan-Heijnen VC. Characteristics of triple-negative breast cancer. *J Cancer Res Clin Oncol.* 2011 Feb;137(2):183–92. doi:10.1007/s00432-010-0957-x
16. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. *N Engl J Med.* 2010 Nov 11;363(20):1938–48. doi:10.1056/NEJMra1001389
17. Burstein MD, Tsimelzon A, Poage GM, Covington KR, Contreras A, Fuqua SAW, Savage MI, Osborne CK, Hilsenbeck SG, Chang JC, Mills GB, Lau CC, Brown PH. Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer. *Clin Cancer Res.* 2015 Apr 1;21(7):1688–98. doi:10.1158/1078-0432.CCR-14-0432
18. Jézéquel P, Kerdraon O, Hondermarck H, Guérin-Charbonnel C, Lasla H, Gouraud W, Canon JL, Gombos A, Dalenc F, Delaloge S, Lemonnier J, Loussouarn D, Verrière V, Campone M. Identification of three subtypes of triple-negative breast cancer with potential therapeutic implications. *Breast Cancer Res.* 2019;21:65. doi:10.1186/s13058-019-1148-6
19. Zhao S, Ma D, Xiao Y, Li X, Ma J, Zhang H, Xu X, Lv H, Jiang W, Yang W, Jiang Y, Zhang Q, Shao Z. Molecular Subtyping of Triple-Negative Breast Cancers by Immunohistochemistry: Molecular Basis and Clinical Relevance. *Oncologist.* 2020 Oct 1;25(10):e1481–91. doi:10.1634/theoncologist.2019-0982
20. Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discov.* 2022 Jan 12;12(1):31–46. doi:10.1158/2159-8290.CD-21-1059
21. Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, Coussens LM, Gabrilovich DI, Ostrand-Rosenberg S, Hedrick CC, Vonderheide RH, Pittet MJ, Jain RK, Zou W, Howcroft TK, Woodhouse EC, Weinberg RA, Krummel MF. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med.* 2018 May;24(5):541–50. doi:10.1038/s41591-018-0014-x
22. Song H, Bucher S, Rosenberg K, Tsui M, Burhan D, Hoffman D, Cho SJ, Rangaswami A, Breese M, Leung S, Ventura MVP, Sweet-Cordero EA, Huang FW, Nijagal A, Wang B. Single-cell analysis of hepatoblastoma identifies tumor signatures that predict chemotherapy susceptibility using patient-specific tumor spheroids. *Nat Commun.* 2022 Aug 25;13(1):4878. doi:10.1038/s41467-022-32473-z
23. Wang Y, Xiang X, Chen H, Zhou L, Chen S, Zhang G, Liu X, Ren X, Liu J, Kuang M, Jiang J, She J, Zhang Z, Xue R, Jiang H, Wang J, Peng S. Intratumoral erythroblastic islands restrain anti-tumor immunity in hepatoblastoma. *CR Med.* 2023 May 16;4(5). doi:10.1016/j.xcrm.2023.101044

Chapter 2 A novel subtyping method for tripe-negative breast cancer with implications for prognosis and therapy

2.1 Abstract

The biological heterogeneity of triple-negative breast cancer (TNBC) poses significant challenges for diagnosis, prognosis, and treatment. While prior TNBC subtype classifications exist, they are not widely used clinically. Here, we aimed to subtype TNBC based on transcriptomic profiles using cell type and state heterogeneity in tumor tissue from 250 pre-treatment women (127 African-American and 123 European-American). We identified three major subtypes and three distinct groups exhibiting unique cell-type composition and mechanisms: Subtype-1 immune signaling/T-cell response; Subtype-2 pro-fibrotic and immune cold; Subtype-3 fatty acid and nuclear receptor signaling. Subtype-1 showed potential responsiveness to immunotherapy, while Subtypes-2 and 3 suggested alternative therapeutic targets. In Subtype-3, which contained a patient group with high *ESR1*, (but not high ER α protein expression) we identified putative mutations in the gene that are unique to these patients. This framework provides a path toward personalized TNBC treatment and is accessible through a user-friendly RShiny application for clinical use.

2.2 Introduction

Triple-negative breast cancer (TNBC) is a biologically heterogeneous and aggressive subtype of breast cancer (BC) defined by <1% protein expression of estrogen receptor (ER α), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) (1–4). TNBC occurs with higher frequency in young, parous women, particularly those who self-report as African-American (AA)/ Black/ African, and in women harboring germline *BRCA1* mutations (5). Despite major progress in defining TNBC biology, a universally adopted subtyping framework

that meaningfully guides clinical decision-making remains elusive. Existing transcriptomic subtypes—such as basal-like (BL), immunomodulatory (IM), mesenchymal (M), and luminal androgen receptor (LAR)—provide informative biological insights (5–9), but are not typically used in clinical practice, except for androgen receptor (AR) targeting (10–13). Furthermore, traditional approaches prioritize clustering based on global gene-expression programs, overlooking the cell heterogeneity of the tissue which dictates the intricate regulatory interactions between malignant epithelial cells and the surrounding tumor microenvironment (TME).

Despite these limitations, bulk RNA-seq remains a robust and scalable approach for defining global transcriptional structure across large patient cohorts. Its ability to capture dominant gene expression programs enables the identification of reproducible subtype architectures and facilitates downstream association with clinical outcomes. Importantly, variation in cellular composition itself is a major source of biological heterogeneity that is embedded within bulk measurements and can be leveraged to refine subtype definitions.

In this study, we propose that tumor tissue heterogeneity, reflected in distinct combinations of epithelial, immune, and stromal cell populations, defines biologically meaningful TNBC subtypes. To investigate this, we develop a novel cell type-driven subtyping approach using tissue-specific gene expression in pre-treatment biopsies of 250 TNBCs from 127 AA and 123 EA women. We utilize breast tissue cell type-specific markers and clustering approaches to identify TNBC subtypes. Using our approach, we identify nine groups within three subtypes for our TNBC samples. Each subtype shows distinct cell type and cell state heterogeneity, prognosis, and survival and is associated with distinct functional mechanisms alluding to the subtype-specific response to targeted therapies.

2.3 Patients and Methods

Sample description and RNAseq data processing

As described in Hossain *et al.* (14), 3188 women with TNBC from the Louisiana Tumor Registry (LTR) were selected, of which, 253 TNBC pre-treatment samples were subject to RNA-sequencing (see Supplementary Table 1-2 from Hossain *et al.*, and GSE268851). We note here that we used 250 (127 AA, and 123 EA) samples amongst 253 samples presented in the companion paper after initial processing. These 250 samples are henceforth referred to as LTR TNBC. 227/250 samples were further subject to diversity analysis (genetic admixture testing) using 163 ancestry-informative single nucleotide polymorphisms, as described in Pakstis *et al.* (15). Germline genetic testing was not performed in this cohort, and the study protocol and institutional review board approvals did not permit additional germline analyses beyond the ancestry-informative marker panel.

The quality of raw sequencing data was evaluated using *fastqc v0.11.9* and summarized for downstream analysis using *multiQC*. RNA read mapping and alignment were performed on human reference genome GRCh38.p13_2021v using *Rsubread v2.10.5*, to produce a count matrix containing 44,715 genes and 250 samples, which were filtered for low-count genes. The resulting data matrix containing 23,264 genes across 250 samples was used for all downstream analysis. Count matrix normalization and differential analysis were performed using *DEseq2 v1.36.0* in R, as outlined in their vignette (16). Differentially expressed genes (DEGs) for every comparison were called at a Benjamini-Hochberg false adjusted $p < 0.05$.

Z-scores for each gene were computed, and genes with a standard deviation (SD) less than ± 2 were filtered to eliminate genes with low biological significance. A total of 18,308 genes passed our significance threshold, and their normalized expression was used as input for visualizations. Heatmaps were generated using the *pheatmap v1.0.12* library in R. The colors on

the heatmap represent the normalized expression count scaled row-wise using the built-in scale function available through *pheatmap*.

Collective Cell Type Markers and subtype Discovery

Cell type markers from twenty-two cell types expected to play a role in breast cancer tumors and tumor microenvironments were identified in the literature (**Supplementary Table 2.1**). Group-specific clusters were detected using a hierarchically clustered tree with the dynamic tree cut algorithm at a cut height of 1.3. *ConsensusClusterPlus* (17) was additionally employed to identify metagroups in an unsupervised manner, with 1000 iterations and the optimal number of clusters (subtypes) (k) at 3, determined using the cumulative distribution function (CDF) curve. To identify gene markers uniquely defining each subtype, we converted the z-scores to percentile values using the *pnorm* function in *stats v4.2.0* library in R, adjusting for an exact mean and SD of each sample. Unique subtype gene markers were defined as genes expressed in the top 20% in at least 50% of the subtype's population. *VennDiagram v1.7.3* and *ggvenn v.0.1.10* were used for all overlap visualizations.

Enrichment, visualization, and reference network generation

Gene set enrichment analysis (GSEA) was performed using the MSigDB Hallmark collection (category H) obtained via the *msigdb* (R package v25.1.1). Differentially expressed genes for each contrast (CC1, CC2, CC3) were ranked using a composite score defined as the sign of the log2 fold change multiplied by the -log10 adjusted P value. Ranked gene lists were analyzed using the *fgsea* algorithm (R package v1.22.0) with minimum and maximum gene-set sizes of 15 and 500, respectively. For each subtype, normalized enrichment scores (NES) and adjusted P values were computed, and significant pathways were defined as those with FDR < 0.05. For optimized visualization the terms have been shortened and selected pathways were shown to avoid

repetition. Additional GO and functional enrichments were performed using *clusterProfiler* (18) or *Enrichr* (19) in R. A high confidence (>0.7) STRING database v12.0 (20) human protein-protein interactions (PPI) containing 15,588 proteins and 236,930 interactions. Additionally, TF-target data downloaded from TRRUST-db (21) was mapped onto the PPI interaction network to generate a custom PPI/TF network. Bulk-seq derived subtype-specific PPI/TF-target networks were extracted using subtype-specific DEGs resulting in the CC1-specific network containing 470 proteins with 1,140 interactions, the CC2-specific network containing 276 proteins and 557 interactions, and the CC3-specific network containing 859 proteins and 4,564 interactions. All networks were constructed, annotated, and analyzed in Cytoscape (22).

Survival Analysis

Overall survival (OS) was compared using Kaplan Meier plots and log-rank tests using *survival* v3.7-0 and *survminer* v0.4.9 in R. OS was defined as the time from diagnosis to death from any cause.

Variant Calling

Single nucleotide variants (SNVs) and indels were called in all samples from CC1, CC2, and CC3 subtypes, focusing on the ESR1 chromosome locus using *FreeBayes* v1.3.5. Variants were filtered based on quality (QUAL >20) and depth (INFO/DP >100). The filtered variants were annotated using snpXplorer to identify their functional consequences.

RShiny App

The subtyping approach developed in this study is deployed as an interactive RShiny app for prospective samples and is publicly available via [<https://www.tnbcworkbench.org/shiny/TNBCsubtyping>].

Validation Strategy

TCGA-BRCA raw count data for 846 tumors was downloaded using TCGAbiolinks in December 2022 and used for validation of our patient subtyping approach, using the unique markers identified. Samples were characterized into Luminal A (n = 586), Luminal B (n = 107), and TNBC (n = 153) based on their IHC/FISH receptor status. Single-cell TNBC data was accumulated from publicly available studies (Xu *et al.* (23), Bassez *et al.* (24)) with total of 47 pre-treatment TNBC patients. Normalized gene expression data for testing datasets were correlated (Pearson) to the centroids of the unique gene markers for each TNBC subtype identified above. Each testing sample was assigned to the TNBC subtype with the highest correlation.

2.4 Results

TNBC subtyping based on tissue heterogeneity defined by cell types and states

We analyzed bulk RNA-seq data from pre-treatment TNBC biopsies of 250 racially diverse women from Louisiana Tumor Registry (LTR) (123 EA, 127 AA) (See Methods and **Figure 2.1a** for processing workflow). Benchmarking with the TNBCtype classifier (Lehmann *et al.* (6)) revealed that 38 tumors were excluded due to modest ESR1 transcript levels despite confirmed ER-negativity by IHC (ER <1%), illustrating known limitations of microarray-derived thresholds when applied to RNA-seq (**Supplementary Figure 2.1a**). The remaining tumors displayed substantial transcriptional heterogeneity. To resolve this complexity, we designed and implemented a cell type and cell state-based subtyping framework integrating 261 markers representing 22 breast-associated cell types (**Figure 2.1b; Supplementary Table 2.1**) (25–35). Unsupervised consensus clustering (CC) of these markers revealed 3 metagroups (CC1-CC3, subtypes), and hierarchical clustering (HC) identified 8 distinct clusters (HC1-HC8, hereafter referred to as groups): CC1 included HC1, part of HC2, and HC3; CC2 comprised HC4-HC6; and CC3 encompassed HC7-HC8, and a subset of HC2 (henceforth HC9) (**Figure 2.1c**). For the

remainder of this study, we focus on the CC subtypes, as they capture the major biological variation, whereas the HC groups primarily reflect intra-subtype heterogeneity. Demographic variables, including race/ethnicity and age, were not significantly associated with CC subtypes, indicating that the transcriptionally defined clusters are not driven by basic patient demographics. Obesity was enriched across the entire cohort (consistent with known TNBC epidemiology) but did not differ between subtypes. Only CC3 showed a significant enrichment of stage 1 tumors, suggesting intrinsic biological differences that may influence tumor detectability or growth patterns (**Supplementary Figure 2.1b**). To define subtype-specific cellular profiles, we applied xCell and EPIC (36,37) deconvolution and identified CC1 as immune-rich, CC2 as fibroblast-dominant, and CC3 enriched for adipocyte, fibroblast, and innate-immune cells (**Supplementary Figure 2.1c, d**). The distinct cellular compositions across CC1-CC3 motivated deeper analyses to resolve the functional programs that differentiate these subtypes.

Subtype-defining cell-type markers aid functional characterization across TNBC subtype

To characterize the functional programs of the cell type-based subtypes, we derived subtype-specific signatures using markers identified through HC (see Methods; **Figure 2.1d**). These markers provide structure feature space from which CC1-CC3 differences can be functionally resolved. To further resolve cellular functional states, we applied EcoTyper (38,39), which identified distinct transcriptional ecosystems across subtypes, linking cell-type and -state diversity to specific activation, exhaustion, and stromal remodeling programs (**Supplementary Figure 2.2a**). CC1 exhibited strong immune and stress-response signatures consistent with its immune-rich composition. Subtype-defining markers highlighted T-cell activation/differentiation, hypoxia/stress regulators, and macrophage/ECM-remodeling signatures were prominent. EcoTyper identified exhausted and naïve/central-memory CD4⁺ T-cell states, activated B cells,

pro-migratory CAF states, and a pro-inflammatory tumor state, defining an immune-engaged but functionally exhausted microenvironment (**Figure 2.1d; Supplementary Figure 2.2a**). CC2 was enriched for ECM and fibroblast-associated programs, suggesting a stroma/fibrotic dominant phenotype. ECM components, stromal remodeling, innate-immune mediators, and endothelial signatures defined this subtype. EcoTyper identified tumor-associated (TA) endothelial states, pro-migratory CAF, CAF1 and tumor cells marked by basal-like and pro-angiogenic states, defining an aggressive, immune-cold, fibrotic-like microenvironment (40) (**Figure 2.1d; Supplementary Figure 2.2a**). CC3 showed prominent metabolic reprogramming and fatty-acid signaling (FAS), accompanied by WNT regulators, nuclear receptors, and macrophage/ECM-associated genes. EcoTyper identified tumor cells in a normal-enriched state, CAF1/CAF2 states, and TA endothelial cells, supporting a metabolically flexible and stromal-interacting phenotype (**Figure 2.1d; Supplementary Figure 2.2a**). Together, these results define CC1 as immune-engaged, CC2 as stromal-fibrotic/immune-cold, and CC3 as metabolically reprogrammed, each representing a distinct tumor-microenvironment axis independent of TNBCtype classifications. To identify the regulatory programs driving these subtype-specific phenotypes, we next examined subtype-specific transcriptional networks.

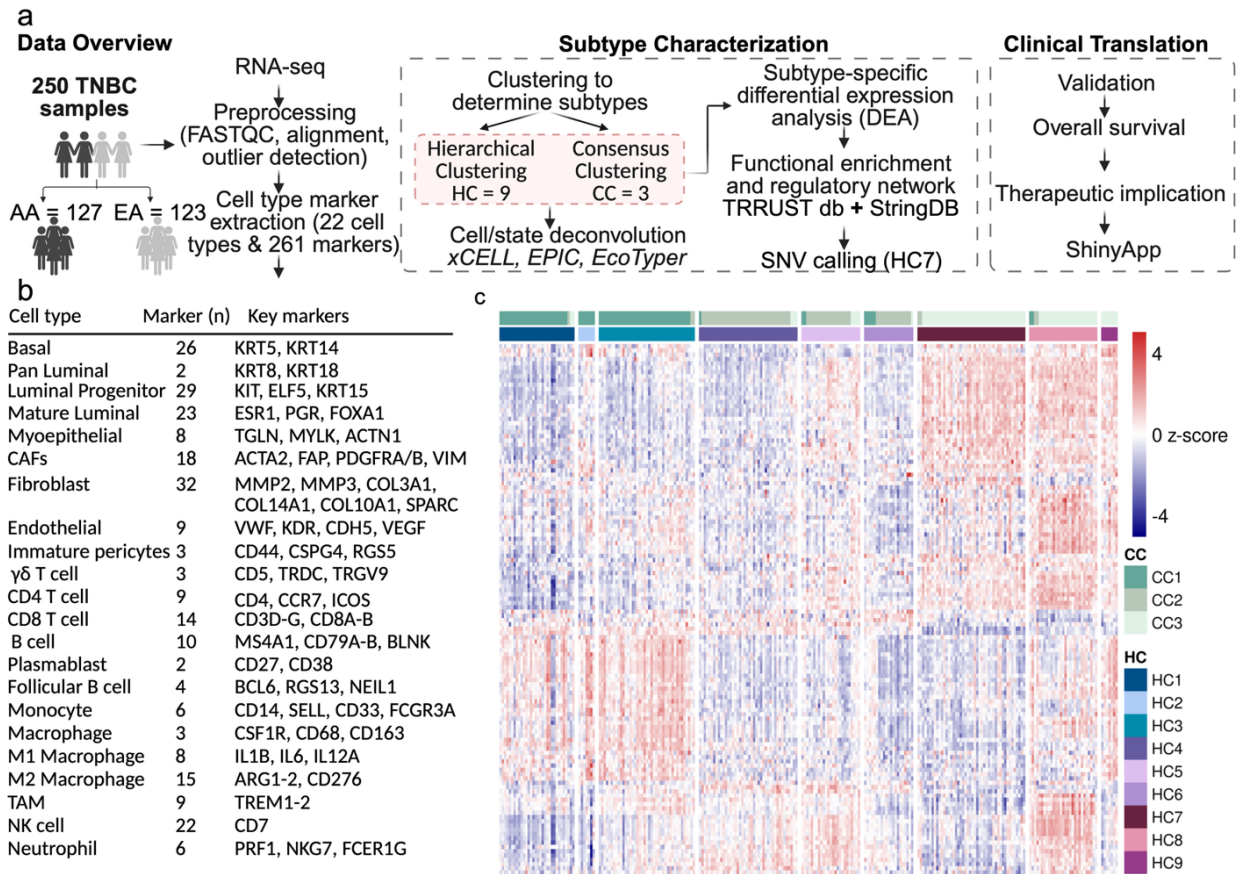


Figure 2.1: TNBC subtyping and characterization based on cell type marker expression and functional features.

a, Overview of the Bulk RNA-seq workflow: from patient to data analysis. Bulk RNA-seq analysis was conducted on fresh-frozen surgical samples from 250 TNBC patients. Bioinformatic analyses was carried to define cell-type based TNBC subtypes. **b**, The table presents tumor and tumor microenvironment cell types utilized for subtyping LTR TNBCs, along with the number and example of marker genes for each cell type. The complete list of markers is provided in **Supplementary Table 1.1**. **c**, The heatmap illustrates relative gene expression (Scaled Normalized Values) of cell type markers within each subtype (ConsensusCluster) and group (HierarchicalCluster) among women with TNBC. This subtyping system categorizes TNBCs into three distinct subtypes and nine groups based on cell type marker expressions. The visualization highlights the heterogeneity between subtypes and within groups, emphasizing the significance of cell type markers.

d

Functional Characterization		Unique Markers	Cellular Composition (xCELL and EPIC)	Cellular states (EcoTyper)		
CC1	Regulation of immune cell differentiation Increased stress markers Wound healing-like T cell mediated immune response	CD2, CD3D, CD3E, CD3G, IFNG, TIGIT, CTLA4, FOXP3, ICOS, IL32, CD38, CD79A, IGLC3, STAT1	CD4 T cell B cells CAFs Tumor cells	CD4 T cells T exhausted Naive/central memory B cells Activated	CAFs Pro-migratory	Epithelial cells Pro-inflammatory Basal-like
CC2	Hypoxic environment, increased inflammation Epigenetic suppression of immunity Early regulation of transcription Extracellular matrix enriched	COL10A1	Endothelial cells CAFs Tumor cells	Endothelial cells TA	CAFs Pro-migratory CAF1 CAF2	Epithelial cells Basal-like Proangiogenic Metabolic
CC3	Nuclear receptor signaling, Fatty acid signaling Epithelial-to mesenchymal transition (EMT) Immune suppression Nuclear receptor signaling and inflammation	CD36, ACKR1, CLDN5, TCF4, ZEB1, KLF4, TWIST2, TBX2, MEG3, IGF1, PDK4, PPARG, CAV1, ALDH1A1, LEPR, CLMP, DPT, IGFBP6, DCN, ZEB2, COL14A1, EPAS1, FOS, VWF, CDH5, CD33, MRC1, LYVE1	Endothelial cells CAFs Adipocytes Tumor cells	Endothelial cells Tumor associated	CAFs CAF1(TA) CAF2 (normal)	Epithelial cells Normal-enriched

Figure 2.2: TNBC subtyping and characterization based on cell type marker expression and functional features, Continued.

d, This table encapsulates the major functional features associated with TNBC subtypes and groups identified in the LTR TNBC samples, based on their cell type-associated gene expression. Functional enrichment was performed based on the increased cell type genes in each subtype and group to unveil underlying mechanisms. (From right to left panel: functional mechanism associated with each subtype, functional mechanism associated with each group, unique cell type markers in each subtype, combined results from xCELL and EPIC cellular decomposition, and results from EcoTyper cell state analysis are presented). Arrow colors correspond to the group color scheme from the heatmap in **c**.

Subtype-specific transcriptional and regulatory programs in TNBC

To define the molecular programs distinguishing each TNBC subtype, we performed differential expression (DEA) and network analyses. For each subtype, DEGs were identified relative to all remaining subtypes, and these subtype-specific DEGs ($FDR \leq 0.05$) were mapped onto curated protein-protein interaction (PPI) and transcriptional factor-target (TF-target) regulatory networks (**Figure 2.2a, c, d-e; Supplementary Figure 2.2b**; See Methods). CC1 exhibited strong immune activation governed by *NFKB1*, *STAT1*, *IRF1*, and *IRF7* transcriptional factors, accompanied by suppression of ECM and fatty acid metabolism. This signature reflects

an active NFKB1-dominant inflammatory and immune state, integrating antigen-presentation, interferon signaling (*CXCL9-11*, *IRF1*), and proliferative programs, consistent with an immunogenic, danger-sensing microenvironment (**Figure 2.2a**). EcoTyper confirmed enrichment of exhausted CD4⁺/CD8⁺ T cells (**Figure 2.2b**), and network analysis showed the first neighbors of inhibitory pathways (*ICOS*, *TIGIT*) were linked to cytotoxicity and antigen processing in CC1 (**Figure 2.2c**), defining an immune-engaged but exhausted phenotype potentially responsive to checkpoint blockade. CC2 tumors were defined by ECM-centered remodeling, p53-associated stress signaling, non-canonical WNT activation, and hypoxia responses. Network analysis identified a *TP53-BRCA1-AXIN1* axis indicative of an active DNA-damage response and restrained canonical WNT, whereas upregulation of *WNT5B*, *EGFR*, and *HIF1A* suggest parallel activation of non-canonical WNT pathways, growth-factor signaling, and hypoxia-driven metabolic adaptation. *TP53* induced *IGFBP3* linked to fibronectin-anchored ECM hub (*FNI*), indicating stromal fibrosis, and EcoTyper further defined mesenchymal/fibroblast cell types into pro-migratory and tumor-associated CAFs (**Figure 2.2d**; **Supplementary Figure 2.2a**). Collectively, these features suggest a transcriptional phenotype of stromal fibrosis and developmental plasticity, in line with the immunologically cold, fibrotic, and aggressive phenotype previously described (**Figure 2.2d**) (41). CC3 were defined by enriched EMT (*SMAD3*), inflammation (*IL6*), and FAS (*PPARG*), while downregulated *MAPK1* which was connected to above mentioned pathways. Within this cluster, *SMAD3* was associated with a canonical TGFβ-EMT axis linked to *ZEB1* and *VIM*, which aligned with AP-1 factors (*JUN*, *FOS*) and *IL6* to establish a *TGFβ-IL6* inflammatory EMT loop (42,43). *SMAD3* also linked to *PPARG*, *CAV1*, and *FABP4*, indicating lipid remodeling and metabolic adaptation. AP-1 linked to *ESR1* upregulation, which connected to *PGR* and *NR4A1*, suggesting aberrant, stress-induced NR activity within a

mesenchymal context. Together, these features identify CC3 as a transcriptionally plastic, TGF β -inflammatory EMT subtype with metabolic reprogramming and non-canonical steroid receptor signaling (**Figure 2.2e**).

Given the enrichment of nuclear receptor (NR)-associated transcriptional programs in the CC3 subtype and the availability of drugs targeting NRs, we further assessed the known 48 human NR (44) expression across our subtypes. We identified 23 NRs, including the entire NR4A subgroup, which showed increased expression in our CC3 subtype (**Supplementary Figure 2.2c**). The NR-centered PPI network revealed a multi-module structure linking nuclear receptors to transcriptional, metabolic, and immune regulators. Immediate-early factors (*FOS*, *FOSB*, *JUN*, *ATF3*, *EGR1/2/3*) and metabolic partners (*PPARG*, *ADIPOQ*, *CD36*, *IGF1/2*, *FABP4*) formed the dominant active hub, indicating preserved stress-responsive and adipogenic NR programs. In contrast, several major modules were selectively downregulated, including epigenetic regulators (*EZH2*, *DNMT1*, *HDAC2*, *HMGAI*, *SMARCC1*), DNA damage–repair proteins (*BRCA1*, *BARD1*, *TP53*, *XRCC6*, *PRKDC*), and chaperone/cofactor machinery (*HSP90AB1*, *HSPA8/4*, *STI1*, *FKBP4*, *PTGES3*, *PHB2*). Interferon and myeloid signaling nodes (*STAT1*, *FCGR1A*) were also reduced, indicating dampened IFN-STAT activity. Collectively, the network reflects a shift away from classical chromatin, repair, and interferon programs toward a stress-responsive, metabolic NR signaling state (**Supplementary Figure 2.2d**). These results underscore the subtype’s potential vulnerability to small molecule inhibitors targeting these NRs, providing a promising avenue for therapeutic development.

All TNBCs analyzed were required to have <1% ER expression by IHC. Despite this, transcript mapping of CC3 showed increased expression of *ESR1*, *ESR2*, and *PGR* mRNA (ESRmRNA^{hi}) (**Supplementary Figure 2.3a**). Within the CC3 subtype, 35 TNBCs fell within the

HC7 group, and to confirm the transcriptional uniqueness of HC7 TNBCs from luminal BC, we 1) compared ESR1 expression across different subtypes and groups in our samples, 2) assessed expression profiles of ESR1 and ESR2 downstream targets in HC7, and 3) compared common features between HC7 and luminal-A BC from TCGA (henceforth referred to as LumA). We observed a higher expression level of ESR1 in CC3 and particularly in HC7, while being lower than the luminal subtype (**Supplementary Figure 2.3b**). We identified that the ESR1 and ESR2 targets were not exclusively associated with CC3 and/or HC7 which indicates the heterogeneous distribution of ESR1 and ESR2 transcriptional activity across our samples (**Supplementary Figure 2.3c**). Lastly, we observed that the functional enrichment of unique DEGs identified in HC7 (vs. LumA) highlighted the role of the arachidonic acid (ArA) pathway (**Supplementary Figure 2.3d**), an essential fatty acid required for eicosanoid production (prostaglandins, leukotrienes) and associated with metastasis and pro-inflammation in various cancers including TNBC (45,46). Collectively, these findings suggest that CC3 represents a TNBC subtype characterized by increased ESR1 mRNA -but not protein- expression with connection to inflammation and FA/eicosanoid signaling based on the constructed network in **Figure 2.2e**.

Considering the observed upregulation of ESR1 within our samples, we explored if these samples exhibit unique mutations. We identified two single nucleotide variations (SNVs) on chr6 unique to the CC3 subtype: a synonymous codon changes at position 151880740 (T-C substitution) and a splice donor SNV at position 151880773 on chr6 (T-G substitution) (See Methods). The synonymous SNV was present in 30/35 ESRmRNA^{hi} samples, and the remaining 5 samples showed a splice donor SNV (**Figure 2.2f**). While synonymous SNVs do not result in an amino acid change in the protein they encode, they can affect the corresponding RNA expression level. Splice donors can alter the normal splicing process, potentially leading to the inclusion or

exclusion of exons in the mature mRNA influencing protein expression. Further experimental investigations are necessary to understand the role of ESR1 mutations in its mRNA expression in TNBC.

These bulk-derived subtype-specific programs do not resolve how tumor and microenvironmental signals are spatially organized within tissue. Addressing this limitation requires spatially resolved analysis, which is the focus of the next chapter. We sought to determine how these networks are deployed in situ across epithelial and stromal compartments. This was achieved by integration of spatial transcriptomics using GeoMx DSP with spatially resolved proteomics using Orion multiplexed imaging to map subtype-defining programs and cell states within the tumor microenvironment.

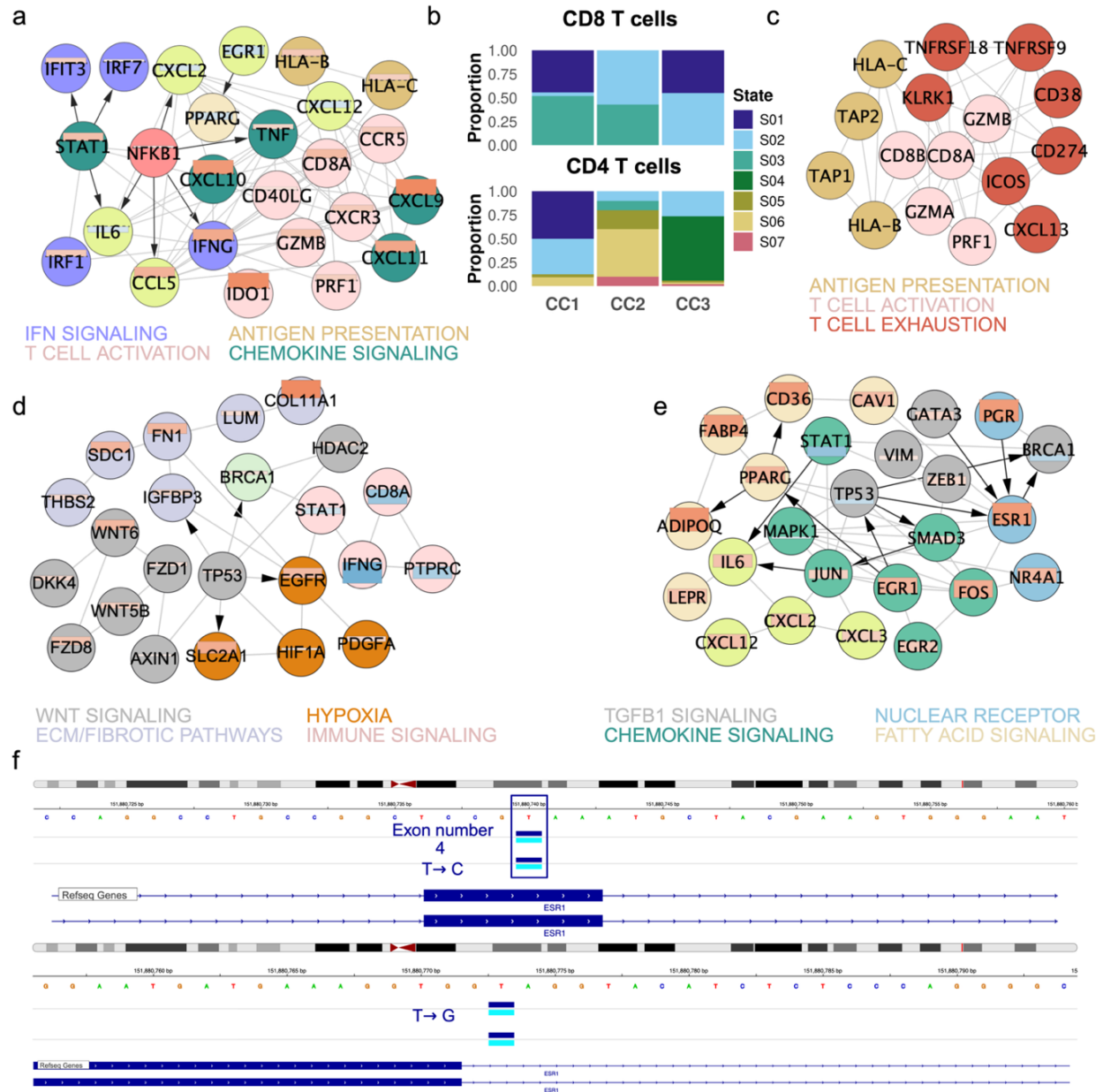


Figure 2.3: Differentially expressed genes and subtype-specific networks in LTR TNBC.

a, A CC1 specific network highlights immune response pathways. **b**, Normalized distribution of T-cell states identified using EcoTyper across subtypes. **c** CC1 specific network highlights exhausted markers and their first neighbors. **d**, A CC2 specific network highlights the significance of extracellular matrix. **e** A CC3 specific network highlights the role of fatty acid, pro-inflammation, and nuclear receptor signaling pathways. The edges present protein-protein interactions, with arrows indicating transcription to target relationships. Nodes of each subnetwork are colored based on biological mechanism represented in the network, embedded within nodes a heat-bar based on the log2FC value between subtypes: dark blue (log2FC of -2) to white (0) to dark red (2). **f**, Representative IGV snapshot showing SNVs on chr6 locations in ESRmRNA^{hi} samples (2 examples for each SNV are shown).

Prognostic association and cross-dataset validation of TNBC subtypes

To assess clinical relevance, we next examined overall survival (OS) for each TNBC subtype among patients with stage I-III and younger than 63 years ($n_{Total}=130$, and $n_{AA}=70$ and $n_{EA}=60$). CC1 was associated with favorable outcomes, whereas CC2 exhibited markedly worse prognosis ($p < 0.05$; **Figure 2.3a**). Subtype-specific prognostic markers further revealed distinct survival associations across CC1-CC3. In CC1, *CD244* and *SLC4A2* were associated with favorable and unfavorable outcomes, respectively. In CC2, *GNLY* and *IKZF3* correlated with better survival, while *MIR1262* and *TYR* were linked to poorer prognosis. Among patients classified as CC3, we identified several genes including *PRF1*, *TNFRSF8*, *NCRI*, *IL18RAP*, *TRGV2*, *TRGV3*, *IL2RB*, *CD79B*, *CARD11*, *NCF1*, *NCF1C*, *AIM2*, *FGD2*, *SIGLEC10*, *SLC4A2*, *KCNA5*, *FBN2*, and *TYR* which are associated with favorable prognosis. In contrast, genes including *SLFN12L*, *GFII*, *IKZF3*, *RNF128*, and *CRTAM* are associated with unfavorable prognosis (**Supplementary Figure 2.4**). Interestingly, *SLFN12L* over-expression in TNBC is associated with sensitivity to chemotherapy (47) and luminal markers with race-inferred biological differences in TNBC (48). However, both *GFII* (49) and *CRTAM* expression have been associated with better prognosis in TNBC indicating further investigation of these biomarkers in the context of CC3 TNBC is necessary to elucidate their role in disease progression.

Cross-platform validation confirmed the robustness and biological specificity of the TNBC subtypes across molecular and cellular levels. The centroid-based classification model reassigned TCGA TNBC cohort ($n = 156$) to their corresponding subtypes (**Figure 2.3b**), demonstrating reproducibility of the CC1-CC3 framework in an external cohort. Single-cell transcriptomic profiling of 47 treatment-naïve TNBC (24,29) tumors independently recapitulated the three ecosystems, revealing subtype-concordant cellular programs: CC1 reflected an immune-inflamed

but exhausted microenvironment dominated by T and myeloid lineages; CC2 exhibited epithelial-fibroblast proliferative states characteristic of an immune-cold, stromal-rich niche; and CC3 showed hormone-like epithelial plasticity and metabolic remodeling with limited immune engagement. Together, these analyses validate the CC1-CC3 classification as a reproducible and mechanistically distinct framework that integrates tumor-intrinsic, stromal, and immune axes of TNBC heterogeneity. To further test clinical relevance, we applied the subtype classifier to the I-SPY2 neoadjuvant trial. While subtype alone did not predict pCR, a significant subtype-treatment interaction was observed in the pembrolizumab arm ($p = 0.027$), with CC1 showing enhanced benefit and CC2 exhibiting attenuated response to immune checkpoint blockade. Recent studies have shown that pCR showed IFN-signaling activation while pNR showed increased pro-angiogenic signaling (50). Together, these findings suggest that distinct TNBC ecosystem states exhibit differential therapeutic sensitivities linked to coordinated epithelial-microenvironment programs.

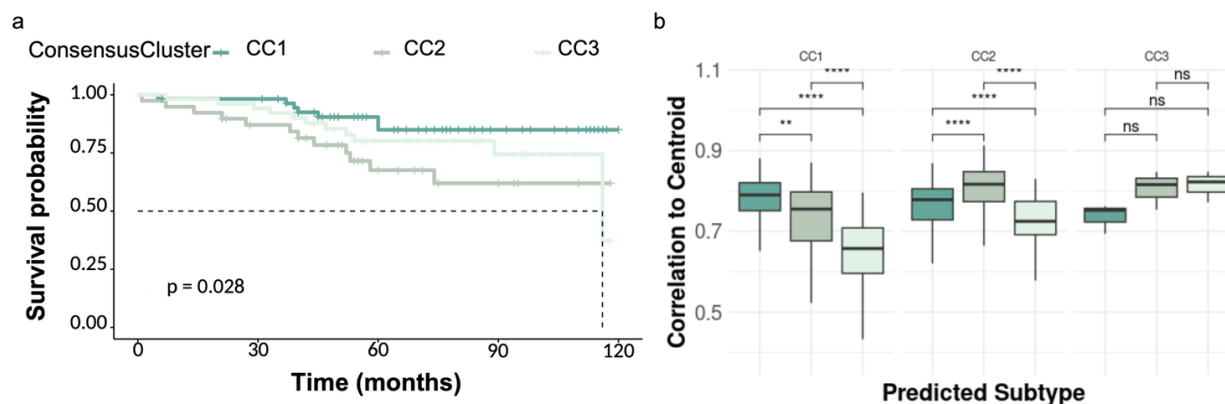


Figure 2.4: Prognostic relevance and subtype classification robustness of TNBC subtypes.

a, Kaplan-Meier survival analysis stratified by TNBC subtypes (CC1, CC2, CC3) derived from consensus clustering. CC1 tumors show improved overall survival compared to CC2 and CC3, with significant differences across groups (log-rank test, $p = 0.028$). Tick marks indicate censored observations. **b**, Correlation of samples to subtype centroids across predicted subtype assignments. Boxplots show the distribution of correlation coefficients between individual samples and subtype-specific centroids, grouped by predicted subtype (CC1, CC2, CC3). Samples classified as a given subtype exhibit highest correlation to their corresponding centroid, demonstrating robustness of subtype assignment. Statistical significance between groups was assessed (Wilcoxon rank-sum test); **** $p < 0.0001$, ** $p < 0.01$, ns = not significant.

2.5 Discussion

Existing TNBC classification frameworks, including those proposed by Lehmann *et al.* (6), and Burstein *et al.* (8), have provided valuable insights into transcriptional heterogeneity. However, these approaches are limited by technical variability, intra-subtype heterogeneity, and challenges in reproducibility. Here, we hypothesized that tumor heterogeneity is fundamentally reflected in underlying cellular composition, and that incorporating cell type-specific markers would enable more biologically grounded subtype definitions. To this end, we developed a novel subtyping approach utilizing 22 breast-specific cell type markers in 250 women with TNBC - with nearly equal distribution of self-reported AA and EA women. This approach uncovered three subtypes with distinct inferred cell types, cell states, functional-modules, and survival outcomes, opening avenues for tailored and targeted therapy.

Tumors pose a significant challenge to the immune system, often disrupting the delicate balance between immune activation and suppression, influencing outcomes and prognosis. We identified the CC1 subtype as an immunologically active subtype, characterized by significant immune exhaustion. Furthermore, we discovered *CD244*, an immunoregulatory receptor expressed on various immune cell types, as a favorable prognostic biomarker. *CD244* has been reported to play dual roles in anti-tumor immunity (51) with evidence for sustaining exhaustion (52) and as a marker for tumor cytolytic activity (53) -suggesting its complexity as a therapeutic target. The CC2 subtype was significantly associated with the enrichment of ECM components as well as an immunologically “cold” microenvironment. In recent years, studies have highlighted the importance of TME and cell-matrix interactions in understanding tumor development, progression, and metastasis as well as treatment responses (54–58). Hammerl *et al.* (59) identified *COL10A1* as a crucial driver of immune exclusion in tumors (identified as DEG in CC2) alluding to the idea that the observed immune exclusion seen in these patients could be associated with the *COL10A1*. Studies have demonstrated that ECM-hypoxia crosstalk leads to increased BC aggressiveness (60), and inflammatory mediators released from TNBC among other factors are involved in upstream regulatory expression of HIF-1 α (61). Consistent with this, we observed an increased expression of hypoxic signatures within the CC2 subtype compared to all other groups.

The CC3 subtype showed enhanced expression of genes involved in FA metabolism and NRS. NRs have a well-defined ligand-binding domain and are amenable to small-molecule intervention (62). NR4A NRs play a diverse role in homeostasis, proliferation, cell migration, apoptosis, metabolite metabolism, DNA repair, glucose utilization, and tumorigenesis (63). Muscat *et al.* identified NR4A NRs to be associated with both ER+ and ER- BCs compared to normal breast tissue (64). Additionally, Zhou *et al.*, showed that *NR4A1* is involved in inducing

EMT in a TGFB-dependent manner, and migration and invasion in an inflammation-dependent manner in breast cancers (65). Further experiments are necessary to understand the underlying role of NR4A NRs in CC3 TNBC and to utilize these markers to improve patients' survival outcomes. Furthermore, we observed a set of steroid NRs with increased expression in CC3 as well as ESR1 was tightly connected to the upregulation of inflammatory genes as well as FA/eicosanoid signaling. These results alongside ESR1-specific SNVs located on CC3 TNBCs suggest a significant role of NRs in the formation and progression of the CC3 subtype of TNBCs, and further experiments to validate these results are essential. The possible role of ArA metabolites, including eicosanoids, prostaglandins, and leukotrienes, deserves further investigation. *PGE2*, an abundant ArA metabolite, is one of the most potent immune suppressants in the TME (66). Tumors of this subgroup may respond to ArA metabolism inhibitors, such as phospholipase A2 (PLA2) inhibitors or cyclo-oxygenase (COX) inhibitors, including NSAIDs. Interestingly, the Miele group and others have reported that NSAID sulindac has activity in immune-competent preclinical models of TNBC, where it enhances tumor immunity (67). Collectively, these findings demonstrate that bulk RNA-seq, when interpreted through a cell-type informed framework, can resolve biologically meaningful tumor ecosystem states that integrate tumor-intrinsic and microenvironmental signals.

To enable application of our subtyping framework, we developed a user-friendly RShiny App for prospective TNBC classification based on tissue heterogeneity and underlying biological programs. The robustness of this approach is supported by cross-dataset validation, including reproducible subtype assignments, and subtype-specific therapeutic responses observed in the I-SPY2 trial. These results highlight the potential clinical utility of this framework for stratifying patients and informing therapeutic strategies, particularly in the context of immunotherapy. To ensure the robustness and clinical relevance of our approach, both retrospective and prospective

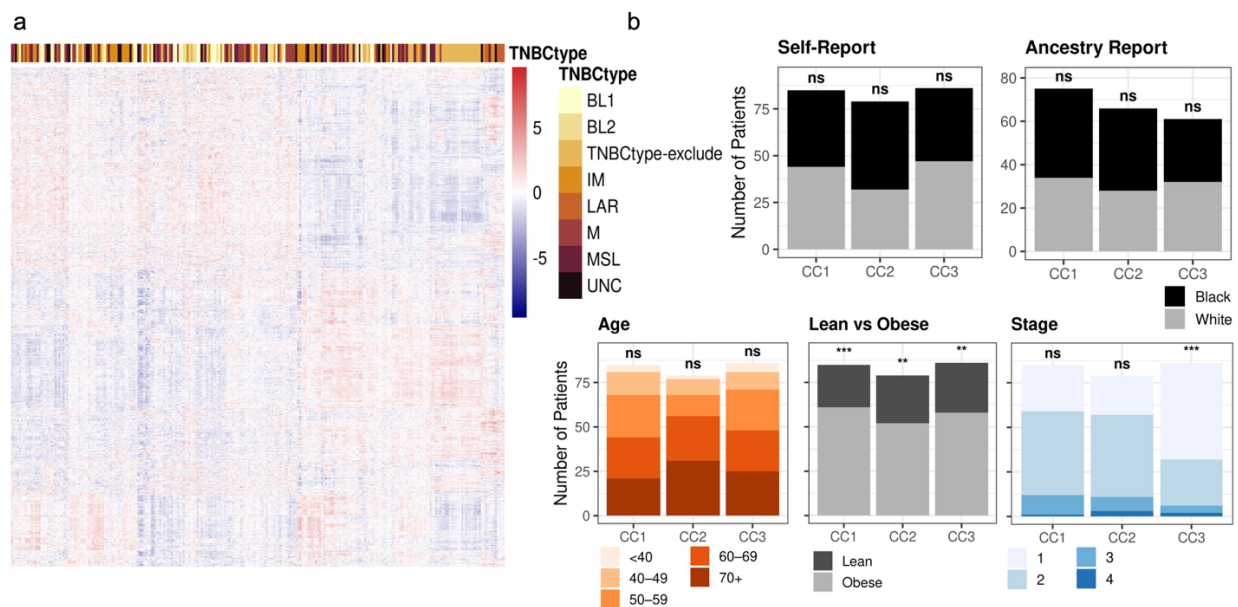
studies are necessary to validate its predictive power for treatment response and survival outcomes in diverse TNBC populations. With further validation, this system could significantly improve personalized treatment strategies for TNBC patients. However, bulk transcriptomic analysis cannot resolve how these programs are spatially organized within tumor tissue, motivating the spatial analyses presented in the following chapter.

This Chapter is being prepared for publication and a version of it is in preprint at DOI: <https://doi.org/10.1101/2025.07.04.663242>. Please refer to the online manuscript for detailed Supplementary Materials.

2.6 Supplemental Materials

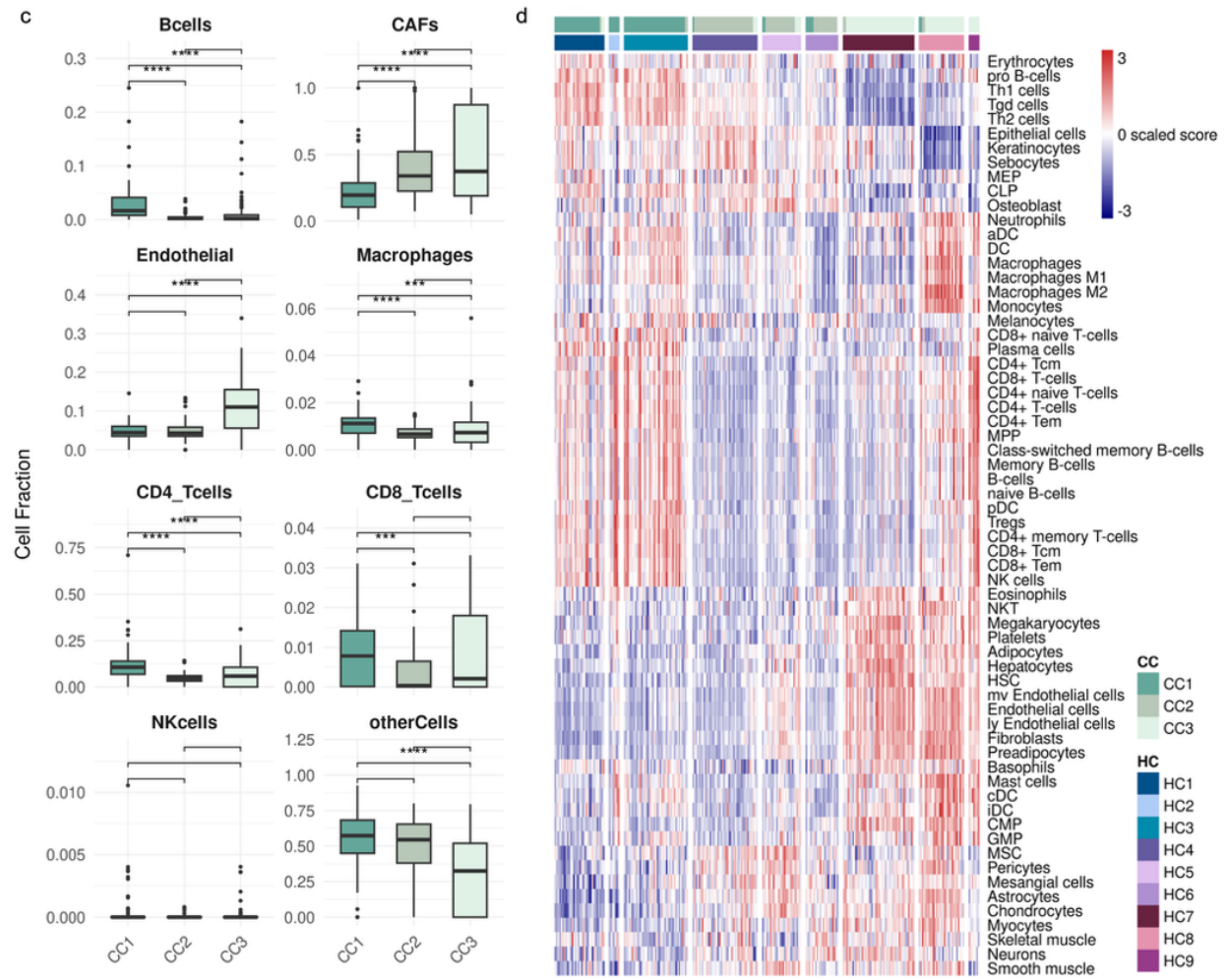
Supplementary Table 2.1: List of all the cell type-markers

Cell Type	Markers
Basal	NDUFA4L2, CD36, VIM, ACKR1, IL8, MECOM, CLDN5, TCF4, ZEB1, NOTCH4, DKK3, KRT14, ITGA6, KRT5, TP63, KRT17, MME, SOX2, NANOG, KLF4, IRX4, MEF2C, SLUG, ERG2, TWIST2, TBX2
Pan Luminal	KRT8, KRT18
Luminal Progenitor	SFRP1, CXXC5, SLP1, ANXA1, RARRES1, KLK5, KRT15, SCGB2A1, CALML5, GLYATL2, TOP2A, NUSAP1, UBE2C, TPX2, SPC25, MKI67, CDK1, CENPF, CCNA2, MEG3, IGF1, PTGDS, KIT, ALDH1A3, EHF, ELF5, CYP24A1, LBP, SOX10
Mature Luminal	ESR1, TNFSF11, PGR, FOXA1, TBX3, PDK4, XBP1, STC2, RUNX1, BATF, ANKRD30A, PIP, MUCL1, TAT, TSPAN8, PRLR, NFE2L2, CHD1, MYC, LMO4, MYB, PR, CITED1
Myoepithelial	TAGLN, MYLK, ACTG2, ACTN1, CALD1, MYL9, RP63, PPARG
CAFs	ACTA2, PDGFRA, PDGFRB, PDPN, THY1, COL1A1, CAV1, ITGB1, S100A4, AIFM2, BSG, CD74, C5AR2, CD70, ALDH1A1, LEPR, TAGLN, FAP, DAN
Fibroblast	MMP2, APOD, CLMP, MMP3, DPT, IGFBP6, DCN, LUM, SNAI1, ZEB2, POSTN, COL3A1, SPARC, COL5A2, COL14A1, COL10A1, WISP2, CILP, SFRP4, COMP, NFAT, HIF1A, EPAS1, FLI1, FOS, JUN, EGR, ATF3, GADD45, HSP1A1, DNAJB
Endothelial	VWF, KDR, CDH5, VCAM1, VEGF, KLF4, NOS3
Immature pericytes	CD44, CSPG4, RGS5
B cell	MS4A1, CD79A, CD79B, BLNK, IGHD,IGHM, IGHA1, IGHA2, IGKV, IGLC3
Plasmablast	CD27, CD38
Follicular B cell	MEF2B, BCL6, RGS13, NEIL1
Monocyte	CD14, FCGR3A, HLA-DRA, TNFRSF1B, CD33, SELL
Macrophage	CSF1R, CD68, CD163
M1 Macrophage	IL12A, TNF, IL6, IL1B, IL23A, NFKB1, STAT1, IRF5
M2 Macrophage	ARG1, ARG2, STAT6, MMP14, MMP9, CD276, FN1, MRC1, CCL13, CCL18, LYVE1, PDCD1LG2, TGFB2, IL10, CCL2
TAM	TREM1, TREM2, CD1D, ITGAM, VCAM1, VTCN1, STAT3, HLA-DRB1, MSR1
NK cell	CD7, PRF1, NKG7, GZMA, GZMB, GZMK, ACTB, ARPC3, CFL1, CST7, KIR2DL1, KIR2DL3, KIR3DL1, KIR3DL2, KLRC2, IL32, GZMH, XCL1, KLRC1, KLRB1, SREG, FCER1G
$\gamma\delta$ T cell	CD5, TRDC, TRGV9
CD4 T cell	CD4, CCR7, LEF1, TCF7, CTLA4, FOXP3, GATA3, IL2R, ICOS
CD8 T cell	PTPRC, CD2, CD3D, CD3E, CD3G, CD8A, CD8B, CCL5, IFNG, CCL3, CCL4, PDCD1, LAG3, TIGIT
Neutrophil	CD11A, CD11B, ITGAL, CD55, CD16, CD10



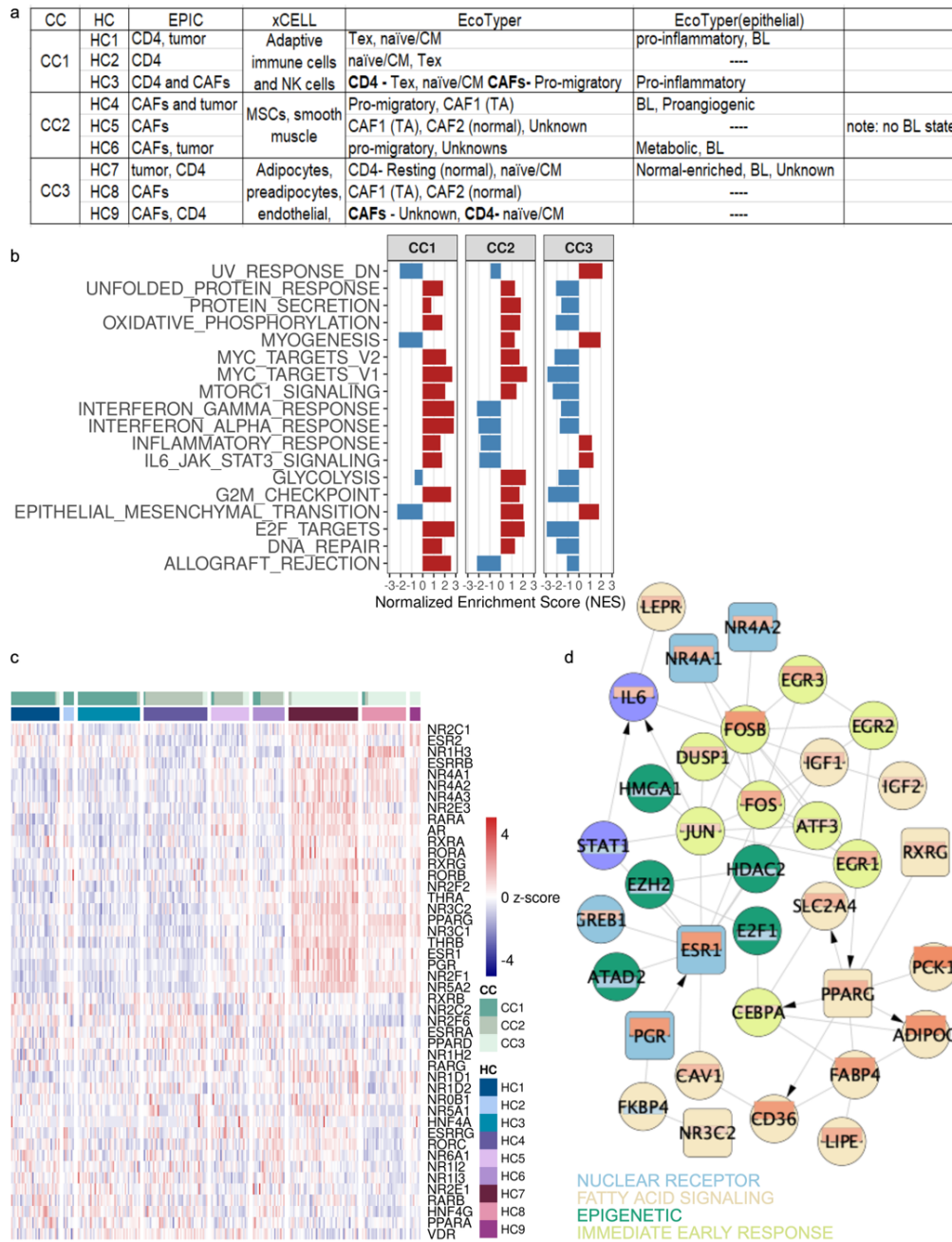
Supplementary Figure 2.1: Additional characterization of TNBC subtype composition and cellular deconvolution.

a, The heatmap illustrates the differentially expressed genes for each TNBCtype, as outlined in the Lehmann *et al.* (6) study. The accompanying bar plot within the heatmap delineates the 7 TNBCtypes, further including the "Lehmann TNBCtype-excluded" group as described in the companion manuscript. Using the Lehmann TNBCtype classification, TNBCs were stratified into 6 subtypes based on significant genes associated with each subtype. **b**, Features bar plots showing the proportion of SRR, AA and EA who have >60% Central/West African and Western/Central European ancestry, respectively, Age, Obesity status, and stage.



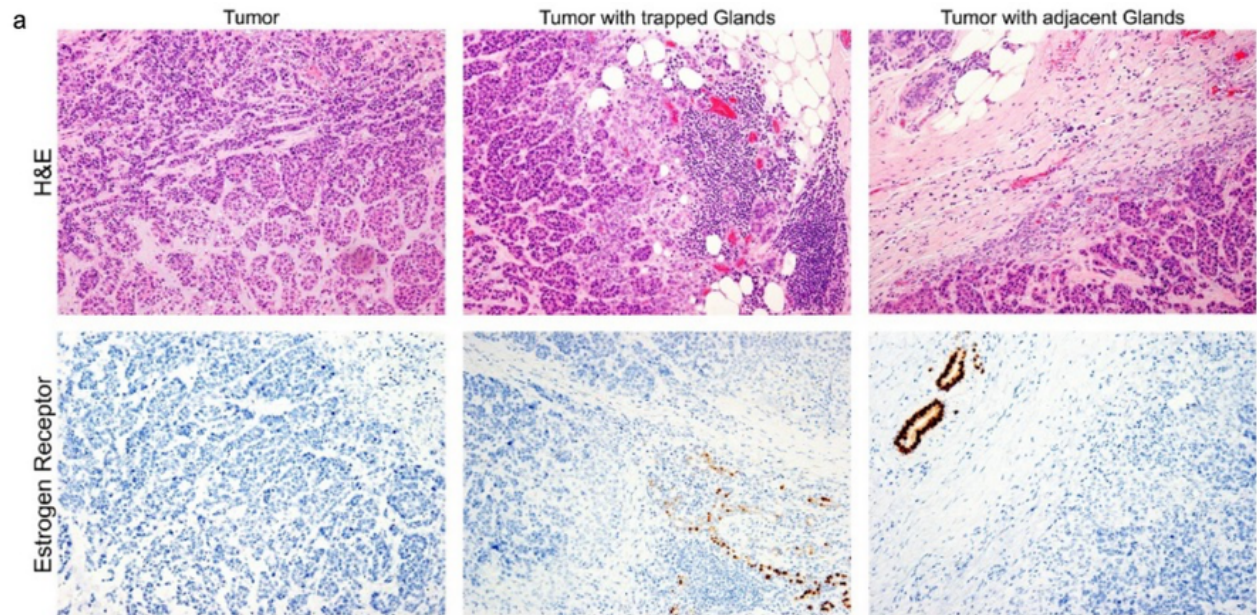
Supplementary Figure 2.2: Additional characterization of TNBC subtype composition and cellular deconvolution, Continued.

c, A barplot shows the proportion of different EPIC cell types in each sample. **d**, A heatmap is displaying the relative enrichment score for each of the xCELL 64 cell types across the samples in each subtype (ConsensusCluster) and group (HierarchicalCluster) in the LTR TNBCs.



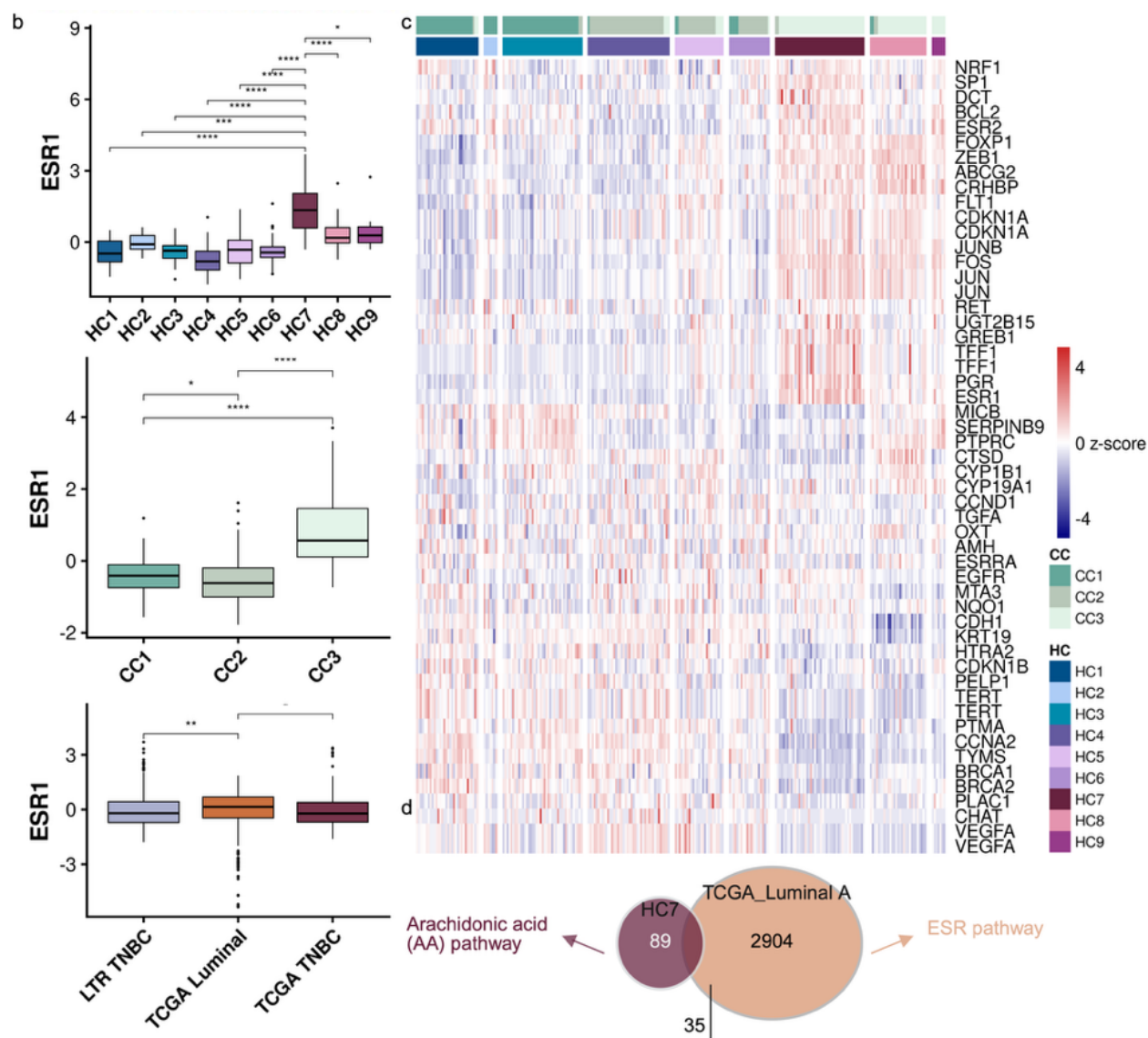
Supplementary Figure 2.3: Functional enrichment and nuclear receptor-associated programs across TNBC subtypes.

a, This table encapsulates the results from xCELL, EPIC, and EcoTyper for each of the subtypes. **b**, GSEA of HALLMARK gene sets, for ranked differential expression genes in each subtype compared to the others. **c**, A heatmap displays the relative gene expression of 48 members of the human nuclear receptors in each subtype (ConsensusCluster) and group (HierarchicalCluster) in the LTR TNBCs. **d**, A CC3-specific network captures nuclear receptors and its neighboring DEGs. Embedded within nodes a heat-bar based on the log2FC value between subtypes: dark blue (log2FC of -2) to white (0) to dark red (2). The edges represent protein-protein interactions, with arrow indicating transcription to target relationships. The square nodes indicate nuclear receptor markers.



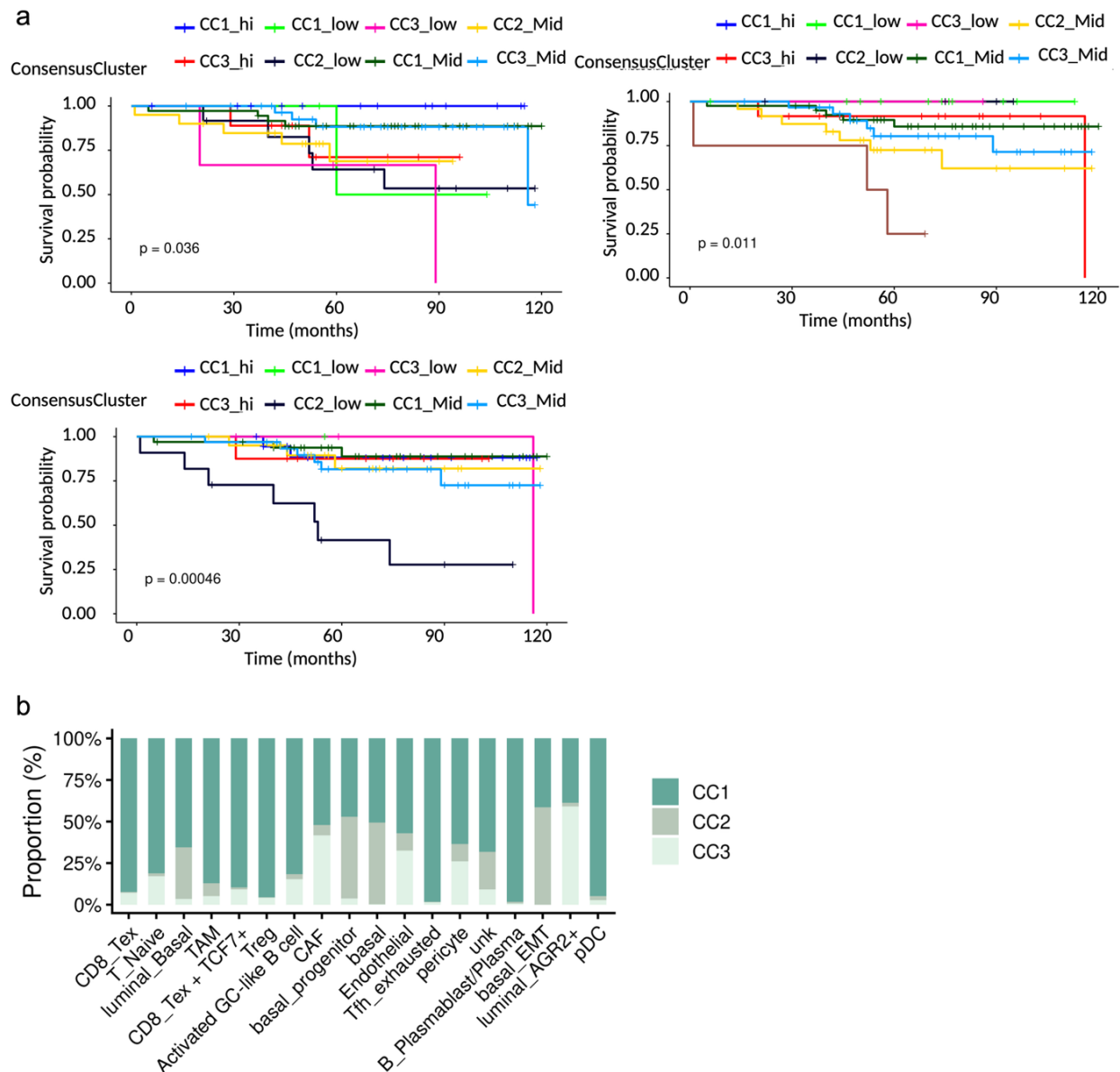
Supplementary Figure 2.4: ESR1-associated transcriptional and histologic features in CC3 TNBC.

a, Three areas of the tumor are shown. Areas of solid tumors are characterized by nests of poorly differentiated epithelial cells, with hyperchromatic nuclei; in these areas, expression of Estrogen Receptor is completely negative (Left Panels). Towards the periphery of the tumor, a few normal breast glands are trapped by neoplastic cells, showing significant lymphocytic infiltration; in these areas, the tumor remains negative and the trapped glands show moderate nuclear expression of ER (middle panels). On the edge, the tumor cells are separated from normal glands, which show robust nuclear labeling of ER (right panels). All panels' original magnification is 200x.



Supplementary Figure 2.5: ESR1-associated transcriptional and histologic features in CC3 TNBC, Continued.

b, Boxplots show the scaled expression of ESR1 in groups and subtypes in the LTR TNBCs, and LTR TNBCs compared to luminal and TNBCs in the TCGA dataset. **c**, A heatmap displays the relative gene expression of downstream targets of ESR1 and ESR2 in each subtype (ConsensusCluster) and group (HierarchicalCluster) in the LTR TNBCs. **d**, Identifying the overlap between genes significantly upregulated in luminal A vs. TNBCs (from TCGA) and genes significantly upregulated in HC7 vs. all others (from LTR TNBCs).



Supplementary Figure 2.6: Prognostic biomarkers and single-cell validation across TNBC subtypes.

a, The overall survival associated with CD244, MIR1626, and SLFN12L, respectively, for each subtype in women under 63 years old with TNBC. **b**, Proportional composition of CC1, CC2, and CC3 across 18 cell subtypes identified by scRNA-seq, highlighting subtype-specific ecosystem enrichment.

REFERENCES

1. Derakhshan F, Reis-Filho JS. Pathogenesis of Triple-Negative Breast Cancer. *Annu Rev Pathol.* 2022 Jan 24;17:181–204. doi:10.1146/annurev-pathol-042420-093238
2. Garrido-Castro AC, Lin NU, Polyak K. Insights into Molecular Classifications of Triple-Negative Breast Cancer: Improving Patient Selection for Treatment. *Cancer Discov.* 2019 Feb 8;9(2):176–98. doi:10.1158/2159-8290.CD-18-1177
3. de Ruijter TC, Veeck J, de Hoon JPJ, van Engeland M, Tjan-Heijnen VC. Characteristics of triple-negative breast cancer. *J Cancer Res Clin Oncol.* 2011 Feb;137(2):183–92. doi:10.1007/s00432-010-0957-x
4. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. *N Engl J Med.* 2010 Nov 11;363(20):1938–48. doi:10.1056/NEJMra1001389
5. Martini R, Delpe P, Chu TR, Arora K, Lord B, Verma A, Bedi D, Karanam B, Elhussin I, Chen Y, Gebregzabher E, Oppong JK, Adjei EK, Jibril Suleiman A, Awuah B, Muleta MB, Abebe E, Kyei I, Aitpillah FS, Adinku MO, Ankomah K, Osei-Bonsu EB, Chitale DA, Bensenhaver JM, Nathanson DS, Jackson L, Petersen LF, Proctor E, Stonaker B, Gyan KK, Gibbs LD, Monojlovic Z, Kittles RA, White J, Yates CC, Manne U, Gardner K, Mongan N, Cheng E, Ginter P, Hoda S, Elemento O, Robine N, Sboner A, Carpten JD, Newman L, Davis MB. African Ancestry-Associated Gene Expression Profiles in Triple-Negative Breast Cancer Underlie Altered Tumor Biology and Clinical Outcome in Women of African Descent. *Cancer Discov.* 2022 Nov 2;12(11):2530–51. doi:10.1158/2159-8290.CD-22-0138
6. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, Pietenpol JA. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *Journal of Clinical Investigation.* 2011;121(7):2750–67. doi:10.1172/jci45014
7. Lehmann BD, Jovanović B, Chen X, Estrada MV, Johnson KN, Shyr Y, Moses HL, Sanders ME, Pietenpol JA. Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PLoS One.* 2016;11(6):e0157368. doi:10.1371/journal.pone.0157368
8. Burstein MD, Tsimelzon A, Poage GM, Covington KR, Contreras A, Fuqua SAW, Savage MI, Osborne CK, Hilsenbeck SG, Chang JC, Mills GB, Lau CC, Brown PH. Comprehensive Genomic Analysis Identifies Novel Subtypes and Targets of Triple-Negative Breast Cancer. *Clinical Cancer Research.* 2014;21(7):1688–98. doi:10.1158/1078-0432.ccr-14-0432
9. Roelands J, Mall R, Almeer H, Thomas R, Mohamed MG, Bedri S, Al-Bader SB, Junejo K, Ziv E, Sayaman RW, Kuppen PJK, Bedognetti D, Hendrickx W, Decock J. Ancestry-associated transcriptomic profiles of breast cancer in patients of African, Arab, and European ancestry. *npj Breast Cancer.* 2021 Feb 8;7(1):10. doi:10.1038/s41523-021-00215-x
10. Gucalp A, Tolaney S, Isakoff SJ, Ingle JN, Liu MC, Carey LA, Blackwell K, Rugo H, Nabell L, Forero A, Stearns V, Doane AS, Danso M, Moynahan ME, Momen LF, Gonzalez JM,

- Akhtar A, Giri DD, Patil S, Feigin KN, Hudis CA, Traina TA, Translational Breast Cancer Research Consortium (TBCRC 011). Phase II trial of bicalutamide in patients with androgen receptor-positive, estrogen receptor-negative metastatic Breast Cancer. *Clin Cancer Res*. 2013 Oct 1;19(19):5505–12. doi:10.1158/1078-0432.CCR-12-3327
11. Bonnefoi H, Grellety T, Tredan O, Saghatchian M, Dalenc F, Mailliez A, L'Haridon T, Cottu P, Abadie-Lacourtoisie S, You B, Mousseau M, Dauba J, Del Piano F, Desmoulins I, Coussy F, Madranges N, Grenier J, Bidard FC, Proudhon C, MacGrogan G, Orsini C, Pulido M, Gonçalves A. A phase II trial of abiraterone acetate plus prednisone in patients with triple-negative androgen receptor positive locally advanced or metastatic breast cancer (UCBG 12-1). *Ann Oncol*. 2016 May;27(5):812–8. doi:10.1093/annonc/mdw067
 12. Traina TA, Miller K, Yardley DA, Eakle J, Schwartzberg LS, O'Shaughnessy J, Gradishar W, Schmid P, Winer E, Kelly C, Nanda R, Gucalp A, Awada A, Garcia-Estevez L, Trudeau ME, Steinberg J, Uppal H, Tudor IC, Peterson A, Cortes J. Enzalutamide for the Treatment of Androgen Receptor-Expressing Triple-Negative Breast Cancer. *J Clin Oncol*. 2018 Mar 20;36(9):884–90. doi:10.1200/JCO.2016.71.3495
 13. Barton VN, D'Amato NC, Gordon MA, Lind HT, Spoelstra NS, Babbs BL, Heinz RE, Elias A, Jedlicka P, Jacobsen BM, Richer JK. Multiple molecular subtypes of triple-negative breast cancer critically rely on androgen receptor and respond to enzalutamide in vivo. *Mol Cancer Ther*. 2015 Mar;14(3):769–78. doi:10.1158/1535-7163.MCT-14-0926
 14. Hossain F, Danos D, Zabaleta J, Wu XC, Del Valle L, Hicks C, Wu J, Tomsic J, Neuhausen S, Ding YC, Mukund K, Mesrizadeh Z, Subramaniam S, Ochoa A, Seewaldt V, Miele L. Immunogenomic diversity of triple-negative breast cancers in obese and non-obese black and white women. *npj Breast Cancer*. 2025 Nov 12;11(1):126. doi:10.1038/s41523-025-00836-6
 15. Pakstis AJ, Speed WC, Fang R, Hyland FCL, Furtado MR, Kidd JR, Kidd KK. SNPs for a universal individual identification panel. *Human Genetics*. 2009;127(3):315–24. doi:10.1007/s00439-009-0771-1
 16. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*. 2014 Dec;15(12):550. doi:10.1186/s13059-014-0550-8
 17. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010 Jun;26(12):1572–3. doi:10.1093/bioinformatics/btq170
 18. Yu G, Wang LG, Han Y, He QY. Clusterprofiler: an r package for comparing biological themes among gene clusters. *OMICS: A Journal of Integrative Biology*. 2012 May;16(5):284–7. doi:10.1089/omi.2011.0118
 19. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, Ma'ayan A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*. 2016 Jul;44(W1):W90–7. doi:10.1093/nar/gkw377

20. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S, Bork P, Jensen LJ, von Mering C. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 2023 Jan 6;51(D1):D638–46. doi:10.1093/nar/gkac1000
21. Han H, Cho JW, Lee S, Yun A, Kim H, Bae D, Yang S, Kim CY, Lee M, Kim E, Lee S, Kang B, Jeong D, Kim Y, Jeon HN, Jung H, Nam S, Chung M, Kim JH, Lee I. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Research.* 2018 Jan;46(D1):D380–6. doi:10.1093/nar/gkx1013
22. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research.* 2003 Nov;13(11):2498–504. doi:10.1101/gr.1239303
23. Xu L, Saunders K, Huang SP, Knutsdottir H, Martinez-Algarin K, Terrazas I, Chen K, McArthur HM, Maués J, Hodgdon C, Reddy SM, Roussos Torres ET, Xu L, Chan IS. A comprehensive single-cell breast tumor atlas defines epithelial and immune heterogeneity and interactions predicting anti-PD-1 therapy response. *Cell Rep Med.* 2024 Apr 12;5(5):101511. doi:10.1016/j.xcrm.2024.101511
24. Bassez A, Vos H, Van Dyck L, Floris G, Arijs I, Desmedt C, Boeckx B, Vanden Bempt M, Nevelsteen I, Lambein K, Punie K, Neven P, Garg AD, Wildiers H, Qian J, Smeets A, Lambrechts D. A single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. *Nat Med.* 2021 May;27(5):820–32. doi:10.1038/s41591-021-01323-8
25. Nguyen QH, Pervolarakis N, Blake K, Ma D, Davis RT, James N, Phung AT, Willey E, Kumar R, Jabart E, Driver I, Rock J, Goga A, Khan SA, Lawson DA, Werb Z, Kessenbrock K. Profiling human breast epithelial cells using single cell RNA sequencing identifies cell diversity. *Nature Communications.* 2018;9(1). doi:10.1038/s41467-018-04334-1
26. Karaayvaz M, Cristea S, Gillespie SM, Patel AP, Mylvaganam R, Luo CC, Specht MC, Bernstein BE, Michor F, Ellisen LW. Unravelling subclonal heterogeneity and aggressive disease states in TNBC through single-cell RNA-seq. *Nature Communications.* 2018;9(1). doi:10.1038/s41467-018-06052-0
27. Han C, Liu T, Yin R. Biomarkers for cancer-associated fibroblasts. *Biomarker Research.* 2020;8(64). doi:10.1186/s40364-020-00245-w
28. Zheng S, Zou Y, Tang Y, Yang A, Liang JY, Wu L, Tian W, Xiao W, Xie X, Yang L, Xie J, Wei W, Xie X. Landscape of cancer-associated fibroblasts identifies the secreted biglycan as a protumor and immunosuppressive factor in triple-negative breast cancer. *OncoImmunology.* 2022;11(1). doi:10.1080/2162402x.2021.2020984
29. Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, Andersson A, Thennavan A, Wang C, Torpy JR, Bartonicek N, Wang T, Larsson L, Kaczorowski D, Weisenfeld NI, Uytingco CR, Chew JG, Bent ZW, Chan CL, Gnanasambandapillai V, Dutertre CA, Gluch L, Hui MN, Beith J, Parker A, Robbins E, Segara D, Cooper C, Mak C, Chan B, Warriar S, Ginhoux F, Millar E, Powell JE, Williams SR, Liu XS, O'Toole S, Lim E, Lundeberg J, Perou CM, Swarbrick A. A

single-cell and spatially resolved atlas of human breast cancers. *Nature Genetics*. 2021;53(9):1334–47. doi:10.1038/s41588-021-00911-1

30. Mukund K, Nayak P, Ashokkumar C, Rao S, Almeda J, Betancourt-Garcia MM, Sindhi R, Subramaniam S. Immune Response in Severe and Non-Severe Coronavirus Disease 2019 (COVID-19) Infection: A Mechanistic Landscape. *Frontiers in Immunology*. 2021;12:738073. doi:10.3389/fimmu.2021.738073

31. Xiong D, Wang Y, You M. A gene expression signature of TREM2hi macrophages and $\gamma\delta$ T cells predicts immunotherapy response. *Nature Communications*. 2020;11(1). doi:10.1038/s41467-020-18546-x

32. Bhat-Nakshatri P, Gao H, Sheng L, McGuire PC, Xuei X, Wan J, Liu Y, Althouse SK, Colter A, Sandusky G, Storniolo AM, Nakshatri H. A single-cell atlas of the healthy breast tissues reveals clinically relevant clusters of breast epithelial cells. *Cell Rep Med*. 2021 Mar 16;2(3):100219. doi:10.1016/j.xcrm.2021.100219

33. Pal B, Chen Y, Vaillant F, Capaldo BD, Joyce R, Song X, Bryant VL, Penington JS, Di Stefano L, Tubau Ribera N, Wilcox S, Mann GB, kConFab, Papenfuss AT, Lindeman GJ, Smyth GK, Visvader JE. A single-cell RNA expression atlas of normal, preneoplastic and tumorigenic states in the human breast. *EMBO J*. 2021 Jun 1;40(11):e107333. doi:10.15252/embj.2020107333

34. Costa A, Kieffer Y, Scholer-Dahirel A, Pelon F, Bourachot B, Cardon M, Sirven P, Magagna I, Fuhrmann L, Bernard C, Bonneau C, Kondratova M, Kuperstein I, Zinovyev A, Givel AM, Parrini MC, Soumelis V, Vincent-Salomon A, Mechta-Grigoriou F. Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer Cell*. 2018 Mar 12;33(3):463–479.e10. doi:10.1016/j.ccell.2018.01.011

35. Lin Y, Xu J, Lan H. Tumor-associated macrophages in tumor metastasis: biological roles and clinical therapeutic applications. *J Hematol Oncol*. 2019 Jul 12;12(1):76. doi:10.1186/s13045-019-0760-3

36. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biology*. 2017;18(1). doi:10.1186/s13059-017-1349-1

37. Racle J, Gfeller D. EPIC: A Tool to Estimate the Proportions of Different Cell Types from Bulk Gene Expression Data. *Bioinformatics for Cancer Immunotherapy*. 2020;2120:233–48. doi:10.1007/978-1-0716-0327-7_17

38. Steen CB, Luca BA, Esfahani MS, Azizi A, Sworder BJ, Nabet BY, Kurtz DM, Liu CL, Khameneh F, Advani RH, Natkunam Y, Myklebust JH, Diehn M, Gentles AJ, Newman AM, Alizadeh AA. The landscape of tumor cell states and ecosystems in diffuse large B cell lymphoma. *Cancer Cell*. 2021 Oct 11;39(10):1422–1437.e10. doi:10.1016/j.ccell.2021.08.011

39. Luca BA, Steen CB, Matusiak M, Azizi A, Varma S, Zhu C, Przybyl J, Espín-Pérez A, Diehn M, Alizadeh AA, van de Rijn M, Gentles AJ, Newman AM. Atlas of clinically distinct cell states and ecosystems across human solid tumors. *Cell*. 2021 Oct 14;184(21):5482–5496.e28. doi:10.1016/j.cell.2021.09.014

40. Boulter L, Bullock E, Mabruk Z, Brunton VG. The fibrotic and immune microenvironments as targetable drivers of metastasis. *Br J Cancer*. 2021 Jan;124(1):27–36. doi:10.1038/s41416-020-01172-1
41. Jiang S, Zhang M, Zhang Y, Zhou W, Zhu T, Ruan Q, Chen H, Fang J, Zhou F, Sun J, Yang X. WNT5B governs the phenotype of basal-like breast cancer by activating WNT signaling. *Cell Commun Signal*. 2019 Aug 28;17:109. doi:10.1186/s12964-019-0419-2
42. Dominguez C, David JM, Palena C. Epithelial-mesenchymal transition and inflammation at the site of the primary tumor. *Seminars in Cancer Biology*. 2017 Dec 1;Mitochondria in Cancer47:177–84. doi:10.1016/j.semcancer.2017.08.002
43. Xu J, Lamouille S, Derynck R. TGF- β -induced epithelial to mesenchymal transition. *Cell Res*. 2009 Feb;19(2):156–72. doi:10.1038/cr.2009.5
44. Frigo DE, Bondesson M, Williams C. Nuclear receptors: from molecular mechanisms to therapeutics. *Essays in Biochemistry*. 2021 Nov 26;65(6):847–56. doi:10.1042/EBC20210020
45. Borin TF, Angara K, Rashid MH, Achyut BR, Arbab AS. Arachidonic Acid Metabolite as a Novel Therapeutic Target in Breast Cancer Metastasis. *Int J Mol Sci*. 2017 Dec 8;18(12):2661. doi:10.3390/ijms18122661
46. Borin TF, Shankar A, Angara K, Rashid MH, Jain M, Iskander A, Ara R, Lebedyeva I, Korkaya H, Achyut BR, Arbab AS. HET0016 decreases lung metastasis from breast cancer in immune-competent mouse model. *PLOS ONE*. 2017 Jun 13;12(6):e0178830. doi:10.1371/journal.pone.0178830
47. ELSAYED AAR, AL-MARSOUMMI S, VOMHOF-DEKREY EE, BASSON MD. Slfn12 over-expression sensitizes triple negative breast cancer cells to chemotherapy drugs and radiotherapy. *Cancer Genomics & Proteomics*. 2022 May;19(3):328–38. doi:10.21873/cgp.20323
48. Singhal SK, Al-Marsoumami S, Vomhof-DeKrey EE, Lauckner B, Beyer T, Basson MD. Schlafen 12 slows tnbc tumor growth, induces luminal markers, and predicts favorable survival. *Cancers*. 2023 Jan;15(2):402. doi:10.3390/cancers15020402
49. Ashour N, Angulo JC, González-Corpas A, Orea MJ, Lobo MVT, Colomer R, Colás B, Esteller M, Ropero S. Epigenetic regulation of gfi1 in endocrine-related cancers: a role regulating tumor growth. *International Journal of Molecular Sciences*. 2020 Jun;21(13):4687. doi:10.3390/ijms21134687
50. Donati B, Reggiani F, Torricelli F, Santandrea G, Rossi T, Bisagni A, Gasparini E, Neri A, Cortesi L, Ferrari G, Bisagni G, Ragazzi M, Ciarrocchi A. Spatial Distribution of Immune Cells Drives Resistance to Neoadjuvant Chemotherapy in Triple-Negative Breast Cancer. *Cancer Immunol Res*. 2024 Jan 3;12(1):120–34. doi:10.1158/2326-6066.CIR-23-0076
51. Sun L, Gang X, Li Z, Zhao X, Zhou T, Zhang S, Wang G. Advances in understanding the roles of cd244 (Slamf4) in immune regulation and associated diseases. *Frontiers in Immunology*. 2021 Mar;12. doi:10.3389/fimmu.2021.648182

52. Wherry EJ, Kurachi M. Molecular and cellular insights into T cell exhaustion. *Nature Reviews Immunology*. 2015 Aug;15(8):486–99. doi:10.1038/nri3862
53. Cheng J, Ding X, Xu S, Zhu B, Jia Q. Gene expression profiling identified TP53MutPIK3CAWild as a potential biomarker for patients with triple-negative breast cancer treated with immune checkpoint inhibitors. *Oncology Letters*. 2020 Apr;19(4):2817–24. doi:10.3892/ol.2020.11381
54. Tian Wei-hong, Qin G, Jia M, Li W, Cai W, Wang H, Zhao Y, Bao X, W W, E. Dale Abel, Shao Q. Hierarchical transcriptional network governing heterogeneous T cell exhaustion and its implications for immune checkpoint blockade. *Frontiers in Immunology*. 2023;14(14). doi:10.3389/fimmu.2023.1198551
55. Liverani C, De Vita A, Spadazzi C, Miserocchi G, Cocchi C, Bongiovanni A, De Lucia A, La Manna F, Fabbri F, Tebaldi M, Amadori D, Tasciotti E, Martinelli G, Mercatali L, Ibrahim T. Lineage-specific mechanisms and drivers of breast cancer chemoresistance revealed by 3D biomimetic culture. *Molecular Oncology*. 2021;16(4):921–39. doi:10.1002/1878-0261.13037
56. Arcucci A, Ruocco MR, Granato G, Sacco AM, Montagnani S. Cancer: an oxidative crosstalk between solid tumor cells and cancer associated fibroblasts. *BioMed Research International*. 2016;2016:1–7. doi:10.1155/2016/4502846
57. Mollah F, Varamini P. Overcoming therapy resistance and relapse in tnbc: emerging technologies to target breast cancer-associated fibroblasts. *Biomedicines*. 2021 Dec 15;9(12):1921. doi:10.3390/biomedicines9121921
58. Tan Q, Xu L, Zhang J, Ning L, Jiang Y, He T, Luo J, Chen J, Lv Q, Yang X, Xie H. Breast cancer cells interact with tumor-derived extracellular matrix in a molecular subtype-specific manner. *Biomaterials Advances*. 2023;146(213301):213301. doi:10.1016/j.bioadv.2023.213301
59. Hammerl D, Martens JWM, Timmermans M, Smid M, Trapman-Jansen AM, Foekens R, Isaeva OI, Voorwerk L, Balcioglu HE, Wijers R, Nederlof I, Salgado R, Horlings H, Kok M, Debets R. Spatial immunophenotypes predict response to anti-PD1 treatment and capture distinct paths of T cell evasion in triple negative breast cancer. *Nature Communications*. 2021;12(1). doi:10.1038/s41467-021-25962-0
60. Yeh YH, Hsiao HF, Yeh YC, Chen TW, Li TK. Inflammatory interferon activates HIF-1 α -mediated epithelial-to-mesenchymal transition via PI3K/AKT/mTOR pathway. *Journal of Experimental & Clinical Cancer Research*. 2018;37(1). doi:10.1186/s13046-018-0730-6
61. Liu Q, Guan C, Liu C, Li H, Wu J, Sun C. Targeting hypoxia-inducible factor-1 α : A new strategy for triple-negative breast cancer therapy. *Biomedicine & Pharmacotherapy*. 2022;156(113861):113861. doi:10.1016/j.biopha.2022.113861
62. Zhao L, Zhou S, Gustafsson JÅ. Nuclear receptors: recent drug discovery for cancer therapies. *Endocrine Reviews*. 2019 Mar 14. doi:10.1210/er.2018-00222

63. Yousefi H, Fong J, Alahari SK. Nr4a family genes: a review of comprehensive prognostic and gene expression profile analysis in breast cancer. *Front Oncol.* 2022 Apr 25;12:777824. doi:10.3389/fonc.2022.777824
64. Muscat GEO, Eriksson NA, Byth K, Loi S, Graham D, Jindal S, Davis MJ, Clyne C, Funder JW, Simpson ER, Ragan MA, Kuczek E, Fuller PJ, Tilley WD, Leedman PJ, Clarke CL. Research resource: nuclear receptors as transcriptome: discriminant and prognostic value in breast cancer. *Molecular Endocrinology.* 2013 Feb;27(2):350–65. doi:10.1210/me.2012-1265
65. Zhou F, Drabsch Y, Dekker TJA, De Vinuesa AG, Li Y, Hawinkels LJAC, Sheppard KA, Goumans MJ, Luwor RB, De Vries CJ, Mesker WE, Tollenaar RAEM, Devilee P, Lu CX, Zhu H, Zhang L, Dijke PT. Nuclear receptor NR4A1 promotes breast cancer invasion and metastasis by activating TGF- β signalling. *Nat Commun.* 2014 Mar 3;5(1):3388. doi:10.1038/ncomms4388
66. Wang D, DuBois RN. The Role of Prostaglandin E 2 in Tumor-Associated Immunosuppression. *Trends in Molecular Medicine.* 2016 Jan;22(1):1–3. doi:10.1016/j.molmed.2015.11.003
67. Hossain F, Ucar DA, Monticone G, Ran Y, Majumder S, Larter K, Luu H, Wyczzechowska D, Heidari S, Xu K, Shanthalingam S, Matossian M, Xi Y, Burow M, Collins-Burow B, Del Valle L, Hicks C, Zabaleta J, Golde T, Osborne B, Miele L. Sulindac sulfide as a non-immune suppressive γ -secretase modulator to target triple-negative breast cancer. *Frontiers in Immunology.* 2023 Oct;14:1244159. doi:10.3389/fimmu.2023.1244159

Chapter 3 Spatial context of cellular heterogeneity in triple-negative breast cancer reveals actionable subtypes

3.1 Abstract

Triple-negative breast cancer (TNBC) is an aggressive, heterogeneous disease with limited targeted treatment options. Leveraging bulk and spatial transcriptomics on a large patient cohort, we resolved the spatio-cellular heterogeneity of the TNBC ecosystem, identifying three distinct subtypes requiring tailored therapeutic strategies. Our analyses define: (1) an immune-exhausted subtype primed for immune checkpoint therapy (ICT); (2) a highly fibrotic/hypoxic subtype; and (3) a novel adipocyte-rich subtype, which includes a subset with ESR1-mutant genotypes displaying transcriptional activity without protein expression. The unique markers associated with the subtypes also provide the basis for subtyping and targeted treatments. Validated by immunohistochemistry (IHC) and Orion spatial proteomics, these findings substantiate the complex spatial architecture of the TNBC microenvironment. Our study provides a robust framework for personalized medicine, highlighting that targeting specific spatio-cellular architectures is critical to improving clinical outcomes.

3.2 Introduction

While transcriptomic subtyping frameworks have provided important insights into triple-negative breast cancer (TNBC) heterogeneity, they are inherently limited by their reliance on bulk gene expression, which averages signals across malignant and non-malignant compartments. As a result, these approaches do not resolve the spatial organization of tumor and microenvironmental cell states that critically shape tumor behavior and therapeutic response. Here, we extend a bulk-

derived, cell-type-informed subtyping framework by integrating spatial transcriptomics to localize and validate subtype-defining programs in situ.

TNBC progression is driven not only by epithelial-intrinsic mechanisms but also by spatially organized stromal and immune niches that dictate selective pressures on tumor evolution (1,2). The "tumor ecosystem" concept emphasizes that TNBC functions as an interacting, adaptive system composed of malignant epithelial, stromal, and immune cell states that collectively dictate phenotype and therapeutic response (3). Whereas classical molecular subtyping captures global transcriptomic states, cell-type-informed subtyping is necessary to delineate the underlying tumor ecosystem. Integrating these complementary layers allows for a unified, mechanistically interpretable model of TNBC heterogeneity. While the TME is a central determinant of TNBC progression, prognosis, and therapeutic response, its precise contribution to subtype identity remains incompletely understood. Bulk RNA-seq has historically enabled major advances in TNBC classification, but it averages signals across malignant and non-malignant compartments, masking spatial complexity. Spatial transcriptomics overcomes this limitation by resolving gene expression within anatomically defined tumor and TME regions, providing a mechanistic link between whole-tumor molecular programs and the spatially organized cellular ecosystems that drive TNBC behavior. Recent spatial transcriptomic studies in breast cancer and TNBC have begun to resolve region-specific immune and stromal niches and their association with tumor behavior (4–8). However, these efforts largely focus on spatial patterning or local cell-cell interactions, rather than establishing a unified subtyping framework that integrates epithelial and TME programs across patients.

Unlike prior TNBC spatial studies focused primarily on local niche composition or cell-cell interactions, our framework integrates bulk transcriptomics, spatial transcriptomics, cell-state

analysis, and regulatory network inference to define conserved tumor ecosystem subtypes across patients. To address the critical knowledge gaps in triple-negative breast cancer (TNBC) heterogeneity, we developed a cell-type-driven subtyping framework incorporating tumor-TME interactions. Analyzing gene expression profiles from 250 pre-treatment TNBC biopsies (127 AA, 123 EA), we show that cellular heterogeneity defines three major subtypes: immune-rich with exhausted T-cells, highly fibrotic immune-deficient, and adipocytic subtypes, the last of which features a novel ESR1-mutant subclass. Examination of spatial biology in the tumor tissue through spatial transcriptomics, multiplexed cell assays, and IHC validate the subtyping and provides mechanism-driven phenotypes for each subtype. The spatial organization of cellular states and signaling programs across invasive and non-invasive tumor regions provide insights into the subtype-specific spatial programs and how TME shapes tumor prognosis. These findings offer unique, actionable therapeutic targets and enhance prognostic insights, providing a framework for precision medicine in TNBC.

3.3 Patients and Methods

Spatial transcriptomic profiling of TNBC using the GeoMx Digital Spatial Profiler

Spatial profiling was performed on 75 pre-treatment FFPE tumors using GeoMx-DSPTM (DSP; NanoString Technologies). Tissue sections (5 μ m) were stained with H&E and hybridized with the GeoMx Whole Transcriptome Atlas (HuWTA) probe set (~18,000 genes). Slides were stained with GeoMx Solid Tumor TME and Nuclear Morphology kits, and circular ROIs (300 μ m; ≥ 3 per region) were selected at tumor center, tumor edge, and non-invasive foci. Automated segmentation defined PanCK⁺ epithelial and PanCK⁻ stromal compartments. UV-mediated oligo release was guided by digital micromirror device (DMD) masks, and oligos were collected per AOI for library preparation using GeoMx SeqCode primers. Libraries were pooled, purified

(AMPure XP), assessed by Bioanalyzer, and sequenced on Illumina NextSeq2000 (2×27 bp, ~100 reads/μm²). FASTQs were processed into DCC files using GeoMxNGSPipeline.

Quality control and normalization were performed using GeomxTools v3.0.1. Samples were filtered based on recommended criteria (min nuclei = 40, min area = 1,000 μm², LOQ = 2), retaining genes detected above background in ≥5% of AOIs. AOIs with <5% detectable genes were removed. Q3 normalization as recommended by GeoMx after gene filtering and transformed into log values for downstream analysis, when necessary (log_q). The comparison of interest dictated whether p-values or adjusted p-values (FDR) were chosen for thresholding and identifying differentially expressed genes. Differential analysis was performed using a linear mixed model. Comparisons of interest once again dictated whether an LMM was performed with a random slope (across slides) or with a random slope and intercept (across-slide with nested random intercepts).

Differential analysis using Linear Mixed Models

Differential expression analysis was performed using linear mixed-effects models (LMMs) implemented through the GeoMxTools framework, following the recommendations outlined in the GeoMxTools vignette. LMMs are well-suited for spatial transcriptomic data because they explicitly account for the hierarchical structure of GeoMx profiling, where multiple regions of interest (ROIs) obtained from the same slide or patient are not independent observations. For this study, we employed an across-slide LMM design with nested random intercepts, appropriate for comparisons spanning multiple slides and patients. This approach models biological and technical structure by incorporating patient-level and slide-level random effects while testing fixed effects associated with the biological comparison of interest (subtypes CC1, CC2, CC3 and regions center, edge, isolated foci). This model structure allowed us to adjust for both repeated sampling per tissue block and variability introduced by slide processing. Analyses were performed separately for

PanCK⁺ epithelial and PanCK⁻ TME segments and further stratified by spatial region. For invasive analyses, ROIs were subset into center (C) and edge (E), whereas non-invasive analyses used isolated foci (F) and subsetted into atypia (A) and adjacent normal (J) epithelial compartments and their matched TMEs.

Orion Analysis

Consecutive tissue sections were cut from eight (2 CC1, 3 CC2, 3 CC3) of the same tissue blocks profiled by GeoMx and were processed using the Orion multiplex immunofluorescence protocol (Akoya Biosciences) with a custom 20-marker panel. Orion images were processed and analyzed using QuPath v0.5.1. Regions were manually annotated on each image using square annotations matched to the diameter of the GeoMx ROIs (300 μ m), allowing corresponding GeoMx ROIs to be manually registered and classified as Center, Edge, or Isolated Foci regions. Cell segmentation was performed using the QuPath cell detection workflow with default settings, and intensity features were calculated for each biomarker stain channel using the *add intensity features* function. For each channel, cell-level mean intensities were modeled using a two-component Gaussian mixture model. The intersection point between the means of the two fitted distributions was used as the positivity threshold, and cells with mean intensity values above this threshold were labeled positive for the corresponding biomarker. Marker positivity calls were then combined to assign cell phenotypes, with T cells defined as CD3e⁺, Exhausted T cells as CD3e⁺PD-1⁺, Cytotoxic T cells as CD3e⁺GranzymeB⁺, NKT cells as CD3e⁺CD56⁺, NK cells as CD56⁺CD3e⁻, and tumor-associated macrophages (TAM) as CD163⁺. The number of cells assigned to each phenotype was quantified within each individual region and compared between regions using the Wilcoxon test.

Enrichment, visualization, and reference network generation

Gene set enrichment analysis (GSEA) was performed using the MSigDB Hallmark collection (category H) obtained via the *msigdb* (R package v25.1.1). Differentially expressed genes for each contrast (CC1, CC2, CC3) were ranked using a composite score defined as the sign of the log2 fold change multiplied by the -log10 adjusted P value. Ranked gene lists were analyzed using the *fgsea* algorithm (R package v1.22.0) with minimum and maximum gene-set sizes of 15 and 500, respectively. For each subtype, normalized enrichment scores (NES) and adjusted P values were computed, and significant pathways were defined as those with FDR < 0.05. For optimized visualization the terms have been shortened and selected pathways were shown to avoid repetition. Additional GO and functional enrichments were performed using *clusterProfiler* (38) or *Enrichr* (39) in R. A high confidence (>0.7) STRING database v12.0 (40) human protein-protein interactions (PPI) containing 15,588 proteins and 236,930 interactions. Additionally, TF-target data downloaded from TRRUST-db (41) was mapped onto the PPI interaction network to generate a custom PPI/TF network. Bulk-seq derived subtype-specific PPI/TF-target networks were extracted using subtype-specific DEGs resulting in the CC1-specific network containing 470 proteins with 1,140 interactions, the CC2-specific network containing 276 proteins and 557 interactions, and the CC3-specific network containing 859 proteins and 4,564 interactions. GeoMx derived subtype- and region-specific PPI/TF-target networks resulted in CC1 Center 237 proteins and 534 interactions, CC1 Edge 466 proteins and 1495 interactions, CC2 Center 53 proteins and 22 interactions, CC2 Edge 328 proteins and 508 interactions, CC3 Center 230 proteins and 236 interactions, CC3 Edge 520 proteins and 1299 interactions. These networks were further customized to capture the functional mechanisms underlying each subtype by selecting all DEGs,

then reducing to the TF-targets and GSEA enrichment associations. All networks were constructed, annotated, and analyzed in Cytoscape (42).

Spatial cell-type deconvolution using SpatialDecon and a TNBC-specific reference matrix

To improve cell-type deconvolution of PanCK⁻ segments, a TNBC-specific reference matrix was constructed using scRNA-seq data from Kumar *et al.* (43) 100,064 cells were downloaded, and 30,719 non-epithelial TNBC cells were retained. Seven major cell types (B cells, T cells, Myeloid, Endothelial, PVL, Plasmablasts, Fibroblast/Mesenchymal) were used to build a reference profile using *SpatialDecon v1.1* (minCellNum=10, scalingFactor=1, minGenes=100). Deconvolution was performed using *RunSpatialDecon*, yielding cell abundance estimates per AOI.

EcoTyper

TPM-normalized GeoMx expression (before Q3 normalization) for each segment was input to EcoTyper to recover cell states using the *ecotyper_recovery_bulk.R* pipeline. Cell-state abundances were multiplied by SpatialDecon-estimated cell-type fractions to produce state-weighted abundance scores. State annotations followed the original EcoTyper publication (44).

3.4 Results

Subtype-specific spatial architecture of TNBC

To map subtype-specific programs within tissue architecture, we profiled 76 TNBC tumors (20 CC1, 19 CC2, 37 CC3) using GeoMx® DSP across 1,597 segmented regions of interest (ROIs) (**Figure 3.1a**). These were then parsed into areas of illumination (AOI) PanCK⁺ (epithelial) and PanCK⁻ (TME) segments depicting Center (C), invasive edge (E), or distant isolated foci (F) tumor regions chosen by a trained pathologist. Across all subtypes, morphological annotation revealed that regions C and E were primarily composed of invasive epithelia, whereas region F consisted

of non-invasive epithelia, including atypia (A) and adjacent normal (J) (**Figure 3.1b, c; Supplementary Figure 3.1a**). We identified validated DEGs between PanCK⁺ and PanCK⁻ segments, with tumor ROIs showing high cytokeratin enrichment and TME ROIs showing enrichment for immune/ECM genes (FDR < 0.05); (**Supplementary Figure 3.1b**). To characterize the TME, we applied SpatialDecon (9) to determine cell-type composition (see Methods). Consistent with the high stroma content in TNBC, analysis across all subtypes revealed a dominant fibroblast/mesenchymal population (**Figure 3.1d; Supplementary Figure 3.2a**). Notably, spatial analysis revealed distinct subtype characteristics. CC1 showed T-cell, B-cell, and myeloid enrichment, particularly at the edge, matching the immune-engaged, interferon/NFkB-driven phenotype from bulk data. Conversely, CC2 and CC3 featured perivascular-like (PVL) populations, with myeloid lineages enriched at foci, indicating distinct stromal and innate immune organization within mesenchymal and fibrotic subtypes (**Figure 3.1d; Supplementary Figure 3.2b**). Because tumor phenotypes arise from context-dependent functional states within organized ecosystems rather than cell-type identity alone, we used GeoMx to analyze how key signaling pathways—interferon/NFkB, WNT/fibrotic-like, TGFB-IL6 EMT, and NRS metabolic programs—are spatially organized across tumor/TME compartments and across center, edge, and foci regions.

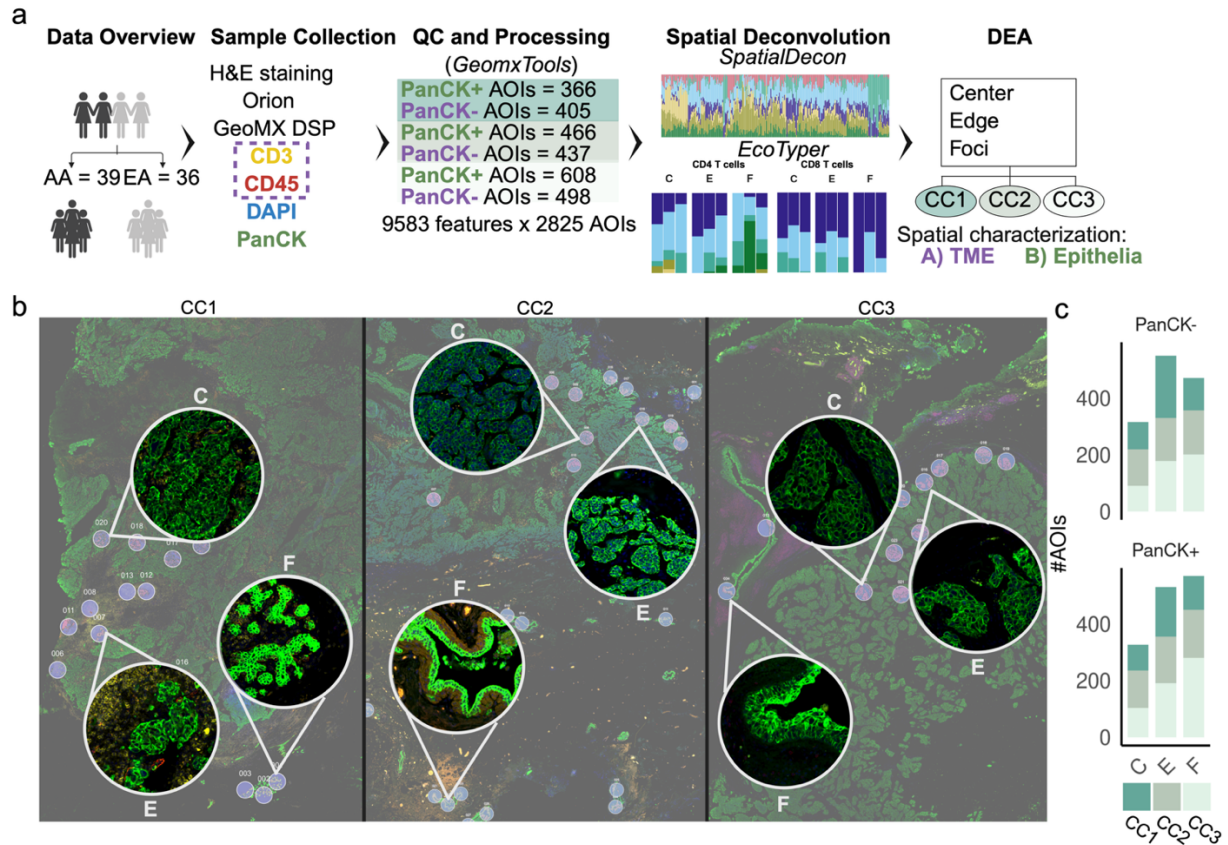


Figure 3.1: Spatial transcriptomic characterization of TNBC subtypes.

a, Overview of the ST GeoMx workflow: from patient to data analysis. ST analysis was conducted on fresh-frozen surgical samples from 75 TNBC samples, H&E stained, IF stained for Pan-cytokeratin and immune (CD3/CD45), and regions of interest were selected with varying epithelial morphology from different spatial regions. Bioinformatic analyses was carried to characterize the spatial distribution of cellular composition and signaling programs for epithelial and TME separately. **b**, IF imaging examples for each TNBC subtypes. ROIs annotated for differing epithelial morphology and spatial regions. ROIs were selected by a pathologist/trained clinician at the tumor interface (E), within the center of the tumor (C), and if available, in regions distant from the center and edge (F, isolated foci). **c**, bar plots showing distribution of AOIs for each region across TNBC subtypes.

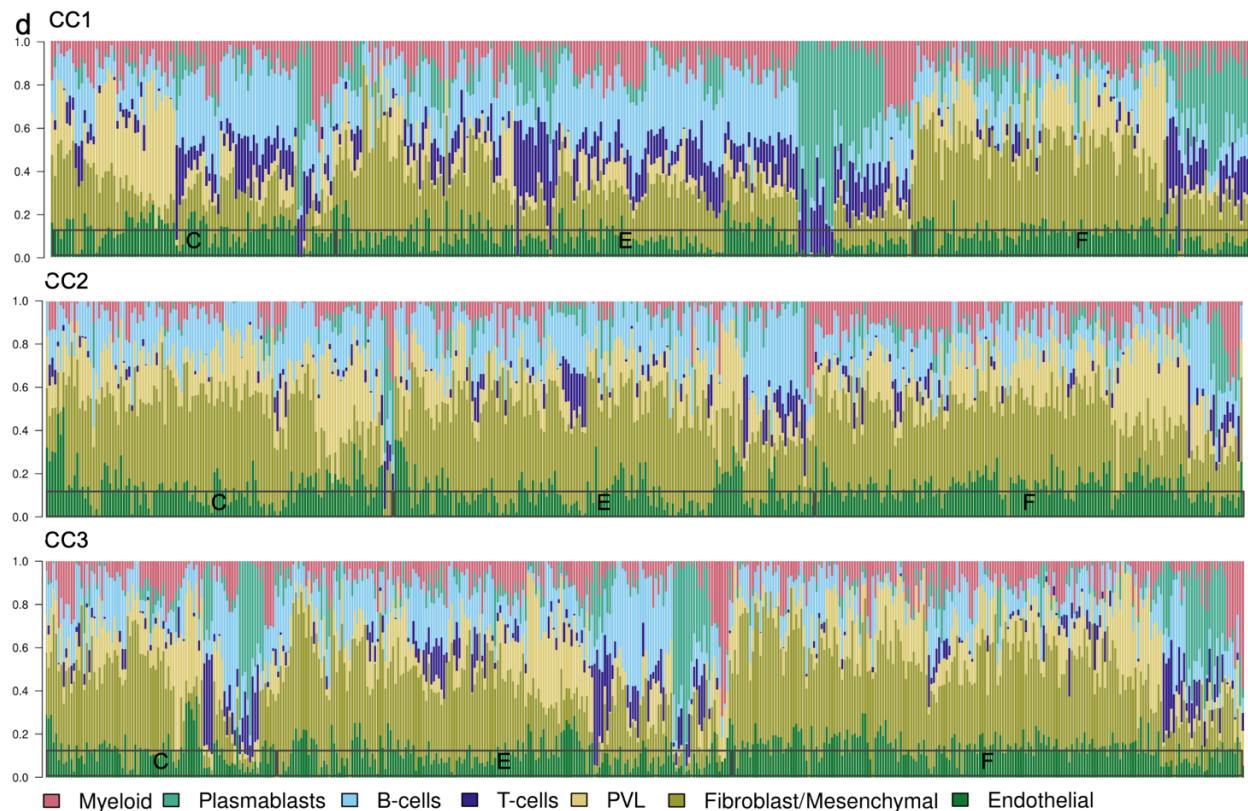


Figure 3.2: Spatial transcriptomic characterization of TNBC subtypes, Continued.

d, The normalized frequency of immune cell types deconvoluted across CC1-CC3 subtypes in PanCK- AOIs generated using SpatialDecon and a custom reference profile matrix. We estimated the abundance of 7 major immune cell types (T-cells, B- cells, plasmablasts, myeloid, PVLs and fibroblast (mesenchymal cell types) and endothelial cells).

Spatial organization of cellular states and signaling across invasive and non-invasive regions

To further refine subtype-specific ecosystems, we applied EcoTyper to both the center and edge of invasive ROIs, enabling a systematic characterization of cell-state compositions across tumor and TME compartments. Immune-inflamed tumors (CC1) showed a distinct polarization, with the center dominated by proliferative myeloid and pro-migratory fibroblasts (*SLC29A4*⁺), while the invasive edge was enriched with exhausted/effector CD4⁺ and CD8⁺ T cells, naïve and activated B cells, cDC1 dendritic cells, classical monocytes, and M1-like macrophages. In contrast, immune-cold tumors (CC2) exhibited a uniform TME rich in CAFs, M2-like foam macrophages, and endothelial cells across both areas. The metabolic subtype (CC3) exhibited a mixed TME with M2 foam-like myeloid cells and mixed fibroblast populations centrally, while the edge featured a

heterogeneous, normal-enriched immune-stromal ecosystem (**Figure 3.2a**). These TME profiles coincided with distinct epithelial transcriptional states: CC1 (pro-inflammatory), CC2 (basal-like), and CC3 (*AGR2*⁺) states, , consistent with coordinated epithelial-microenvironment organization (**Figure 3.2b**). Finally, Orion imaging validated these architectures, confirming high immune infiltration in CC1 and exclusion in CC2, with strong spatial mapping to bulk-assigned classes (**Figure 3.2c-f**).

To understand how distinct immune-stromal architectures are associated with distinct signaling programs, we conducted gene-set enrichment analysis (GSEA) on both epithelial and TME regions (**Figure 3.2g**). Analysis of invasive ROIs revealed distinct pathway activations by subtype (FDR < 0.05). CC1 demonstrated an immune-visible, proliferative phenotype, characterized by high IFNG/IFN α signaling, antigen presentation, and cell-cycle progression in both tumor and TME compartments, with prominent JAK-STAT signaling at the invasive edge indicating active immune cross-talk. Conversely, CC2 represented an immune-cold, fibrotic-like state, showing downregulation of immune activity and enrichment of EMT, angiogenesis, myogenesis, and TGFB signaling. CC3 was characterized by downregulated cell-cycle programs in both compartments, with specific KRAS pathway upregulation within the TME center (**Figure 3.2g**).

Next, we investigated whether the regulatory programs observed in invasive regions emerge earlier by analyzing non-invasive ROI foci, including both the epithelial (A and J) and paired TME compartments. EcoTyper analysis showed conserved cellular composition across non-invasive TME subtypes, characterized by normal-enriched fibroblasts, endothelial cells, and heterogeneous monocytes, representing a less segregated environment than the invasive front. In the epithelial compartment, cellular states converged on *AGR2*⁺ (CC1, CC3) or mixed normal-

enriched and *AGR2*⁺ (CC2) states, suggesting a shared pre-invasive phenotype (**Figure 3.2b**). Despite this structural convergence, GSEA revealed distinct, subtype-specific signaling differences (FDR < 0.05). Non-invasive CC1 showed immune-related enrichment (Notch, IFN α), CC2 exhibited a fibro-inflammatory, immune-suppressed state (TGFB, PI3K), and CC3 showed metabolic stress adaptation (WNT, ROS, UPR) (**Supplementary Figure 3.3a, b**). Together, these findings suggest that functional signaling programs, rather than cellular composition alone, are critical determinants of subtype-specific ecosystem organization and may emerge prior to overt invasion.

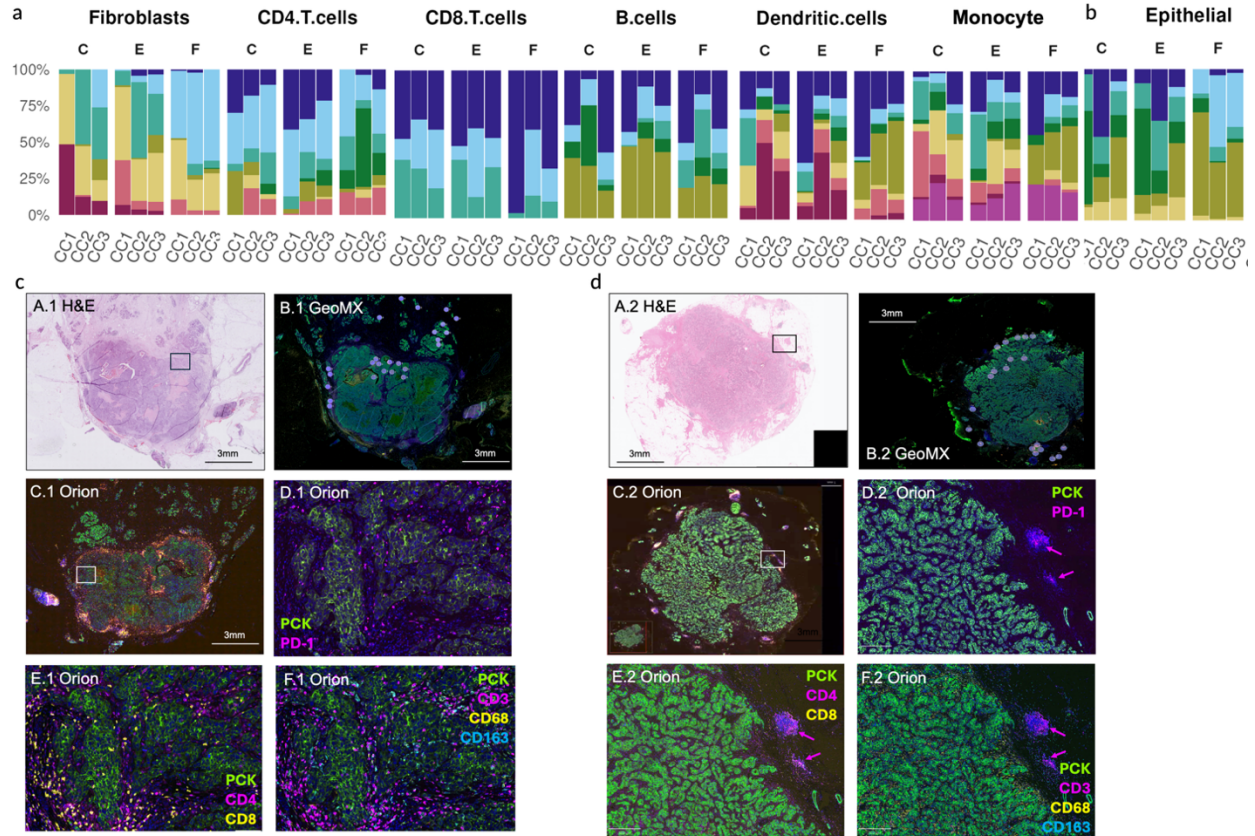


Figure 3.3: Integrated spatial characterization of tumor and immune ecosystems across TNBC subtypes.

a, b, Normalized distribution of cell states identified using EcoTyper across subtypes and separated by regions (center (C), edge (E), and isolated foci (F)). **c, d,** Representative Orion staining of CC1 and CC2 TNBC, respectively: **A** 5 um sections are stained with Hematoxylin and Eosin (**H&E**). **B** is the corresponding GeoMx image of **A**. **C-F** are Orion Analysis of representative CC1 and CC2 TNBC. In **C**, samples are stained with the Orion panel as outlined in materials and methods and is provided for spatial orientation; the white box illustrates the area detailed in **D, E**, and **F**. **D** depicts co-staining with pan-cytokeratin (**PCK**) and the immune exhaustion marker, PD-1 (**PD-1**). **E** illustrates co-staining with PCK and the T-cell markers CD4 (**CD4**) and CD8 (**CD8**). **F** depicts co-staining with a) PCK, b) the mature T-cell marker CD3 (**CD3**), that is critical for antigen recognition and initiation of the immune response, c) CD68 (**CD68**) a marker expressed on the surface and within macrophages, monocytes, and granulocytes as well as tumor-associated macrophages, and d) the macrophage/monocyte scavenger marker CD163 that can also appear on M2-polarized macrophage (**CD163**). CC1 image 1 is an immunologically active TNBC with abundant cells that stain positively for CD4, CD8, and CD3; some cells stain positively for PD-1, but many do not. In contrast, CC2 image 2, contains small islands of cells peripheral to the tumor that stain positively for CD4 (**E.2**) and CD3 (**F.2**) but also stain strongly for PD-1 (**D.2**).

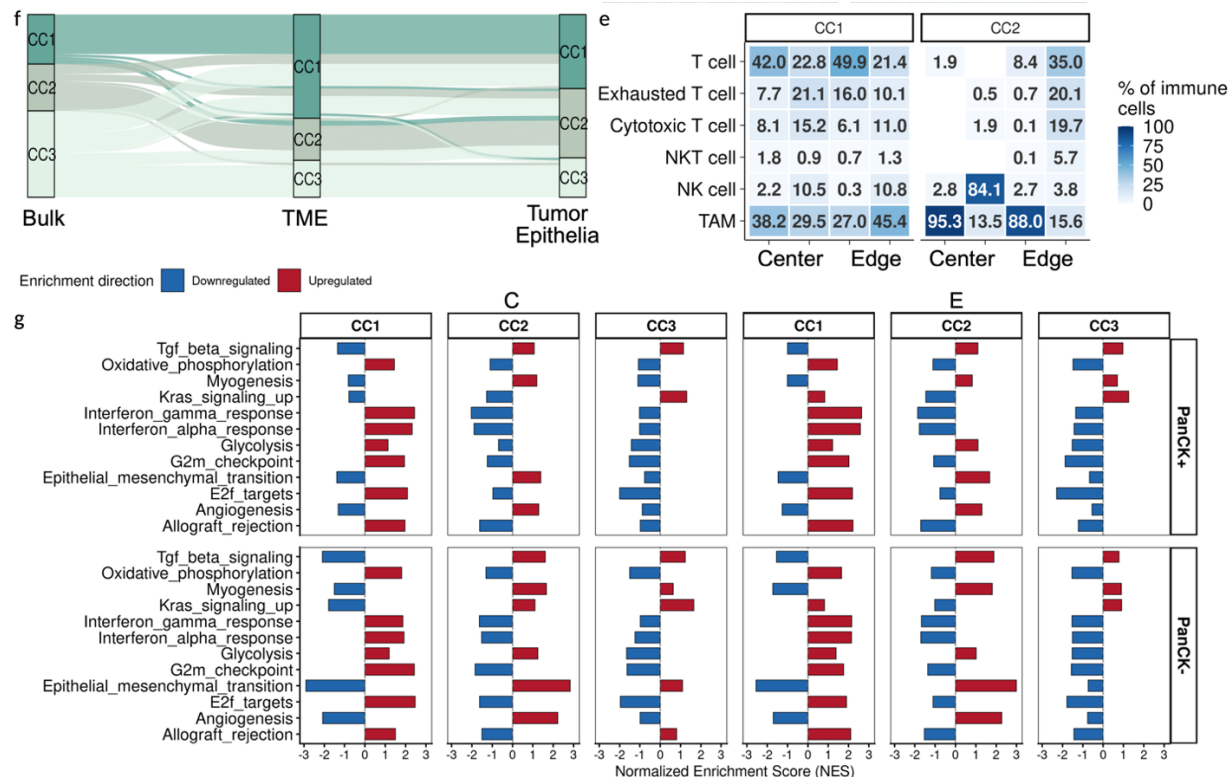


Figure 3.4: Integrated spatial characterization of tumor and immune ecosystems across TNBC subtypes, Continued.

e, Heatmap showing the proportion of immune cell populations (% of total immune cells) across tumor regions (Center and Edge) in CC1 (P1, P2) and CC2 (P3, P4) profiled by Orion multiplex imaging. Cell types were classified based on marker co-expression: T cells (CD3e⁺), Exhausted T cells (CD3e⁺PD-1⁺), Cytotoxic T cells (CD3e⁺GranzymeB⁺), NKT cells (CD3e⁺CD56⁺), NK cells (CD56⁺CD3e⁻), and tumor-associated macrophages (TAM; CD163⁺). Color scale represents percentage of immune cells per region, ranging from 0% (white) to 100% (dark blue). **f**, Contribution of tumor and TME compartments to the TNBC cellular subtypes. The alluvial plot shows the distribution of TNBC subtypes from global Bulk RNA-seq, ST tumor and TME pseudobulks. **g**, Molecular characterization of invasive tumors (top) and TME (bottom) compartments across TNBC subtypes. The most relevant and significant molecular features from Hallmark are illustrated. FDRs are based on the Wilcoxon rank-sum test. FDR < 0.05 shown; blue = negative association, red = positive association.

Spatial coupling of regulatory programs within the tumor ecosystem across TNBC

subtypes

Spatially resolved analyses revealed that subtype-defining epithelial programs were consistently coupled with complementary immune and stromal states across invasive regions, suggesting coordinated ecosystem-level organization. To define molecular programs associated

with invasive regions, we performed subtype- and region-specific DEGs (FDR < 0.05) and PPI-TF-target network analysis (See Methods). While these spatially resolved circuits recapitulated bulk-derived CC programs (**Figure 2.2**), they now resolve how tumor epithelial and TME signaling states are spatially coordinated across invasive center-edge regions. Across subtypes, epithelial compartments consistently encode the dominant regulatory programs that are associated with and reinforced by TME signaling states. Specifically, the *IFN-STAT1-NFKB* inflammatory axis in CC1, ECM-fibrotic remodeling in CC2, and *MAPK/WNT/TGFB*-driven metabolic plasticity in CC3 are all anchored in epithelial programs.

In the immune-inflamed CC1 subtype, tumor cells at the center and invasive edge display a potent IFN-associated regulatory program, with network analysis identifying an *IRF1-CEBPB-E2F1-DNMT1* axis linking interferon response genes to cell-cycle regulators (**Figure 3.3a**). At the invasive edge, an epithelial-intrinsic NFKB/TNF module, involving *STAT1/RELA/TRAF2* activation, was associated with elevated *CCL5* and *CXCL9* expression. This epithelial chemokine output aligned with *CXCR3/CCR1* signaling within the TME, suggesting coordinated immune recruitment and epithelial-immune communication (10) (**Figure 3.3b**). Concurrent upregulation of MHC-I and MHC-II by epithelial compartments highlights their role as active antigen-presenting and organizing agents rather than passive immune targets. CC2 was characterized by a low-proliferative, stromal-dominant core enriched for collagen-associated ECM-EMT programs, with ECM and TGFB-associated signaling primarily within the TME compartment and reduced proliferative programs within epithelial regions (**Figure 3.3b**). In contrast, the invasive edge exhibited compartment-specific signaling rewiring characterized by TME-enriched *FN1-TGFBRI-TGFB*-associated matrix remodeling together with reduced epithelial antigen-presentation and NFKB-associated programs, consistent with a fibro-inflammatory and immune-

restricted invasive interface (11) (**Figure 3.3b**). In CC3, tumor epithelial cells exhibited a unique program of signaling rewiring and metabolic plasticity. At the center, epithelial cells exhibited stress-responsive non-canonical MAPK rewiring, including reduced *MAPK1* expression together with increased *MAP3K1* (12,13) and *IL1R* upregulation (14) (**Figure 3.3b**). This was accompanied by increased *RASA1* and *EPHB4*, suggesting vascular-associated signaling (15). Furthermore, the convergence of WNT (activated *SVL2*, *AXIN2*) and TGFB (induced *TGFBR2* in the TME) was consistent with coordinated epithelial-stromal plasticity, aligned with bulk RNA data (**Figure 3.3b**). At the invasive edge, stress-adaptive MAPK signaling persisted through MAP3K1-associated connectivity with *SPI* and *HRAS* despite reduced *MAPK1* expression. Concurrently, a nuclear receptor-associated metabolic state emerged, defined by epithelial upregulation of *ESRRA* and *NR1P1* and induction of fatty acid metabolism genes (*FABP1*, *CPT1A*). Complementary TME rewiring was also observed, including reduced *RELA* expression alongside increased *AKT3/PI3KR3* (16) and decreased *HSP90AA1*, consistent with reprogrammed PI3K/AKT-HIF signaling states (**Figure 3.3b**). Together, these findings define CC3 as a metabolically adaptive and stress-responsive TNBC ecosystem characterized by coordinated epithelial-TME rewiring and non-canonical NR activity despite ER negativity.

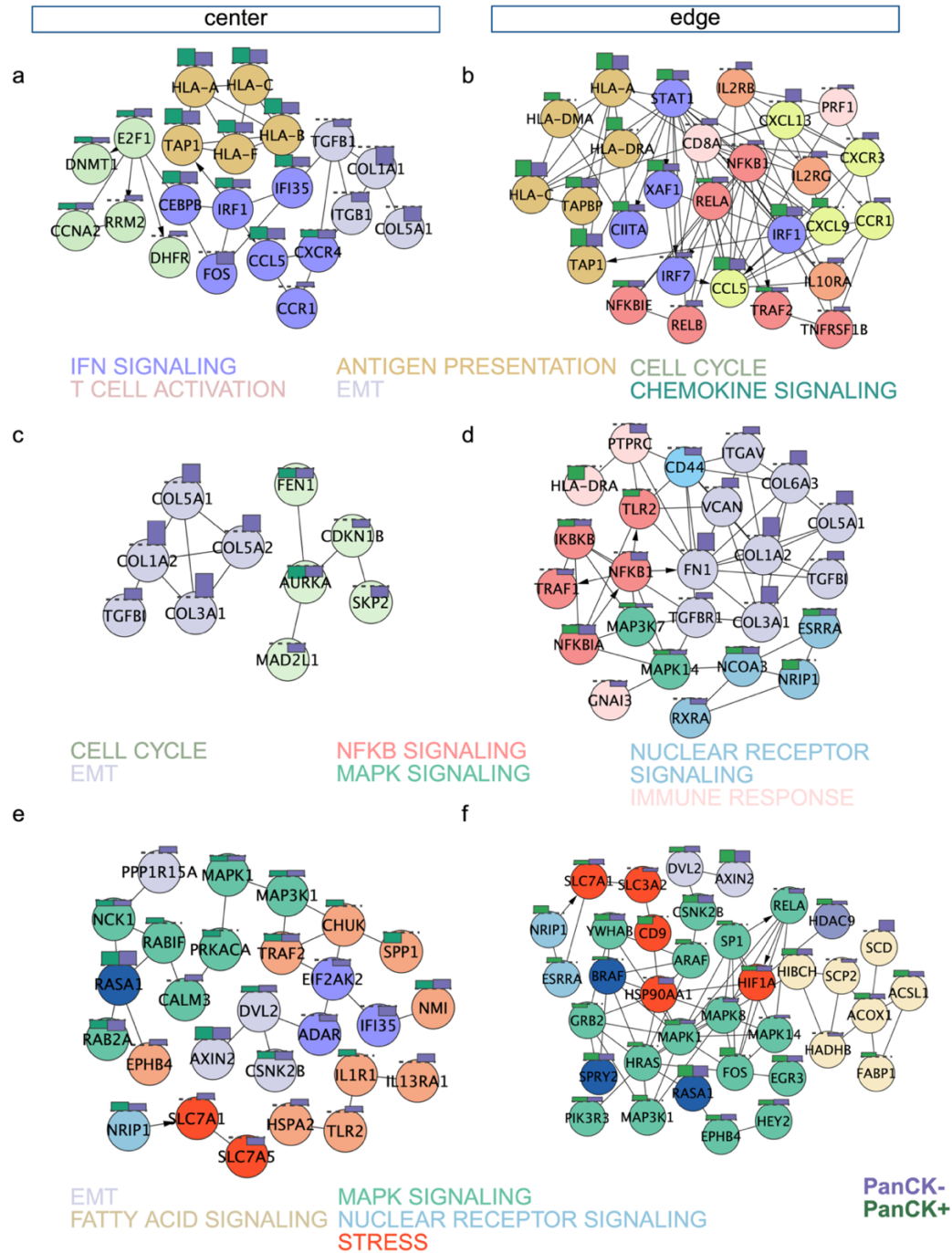


Figure 3.5: Spatially resolved regulatory networks across TNBC subtypes and invasive regions.

a, b, A CC1 specific network highlights immune response pathways in center and edge. **c, d,** A CC2 specific network highlights the significance of extracellular matrix and EMT in the center and edge. **e, f,** A CC3 specific network highlights the role of fatty acid, pro-inflammation, and nuclear receptor signaling pathways. The edges present protein-protein interactions, with arrows indicating transcription to target relationships. Nodes of each subnetwork are colored based on biological mechanism represented in the network, with bar plot inside nodes colored based on epithelial (green) and TME (purple) and the height based on log2FC.

3.5 Discussion

Our findings support a model in which TNBC subtypes represents conserved spatial ecosystem states arising from coordinated epithelial, immune, and stromal programs rather than tumor-intrinsic transcriptional states alone. By integrating bulk transcriptomics, spatial transcriptomics, multiplexed imaging, cell-state analysis, and regulatory network inference, we identified three reproducible TNBC ecosystems (CC1-CC3) with distinct spatial architectures, signaling programs, and therapeutic associations. These subtypes demonstrate that specific tumor-TME interactions shape epithelial plasticity within invasive regions (17). Unlike prior TNBC classification systems that primarily define tumor-intrinsic molecular states, our framework resolves subtype identity coordinated tumor-microenvironment organization across invasive and non-invasive regions.

A central finding of this study is that subtype-defining epithelial programs were consistently spatially coupled with complementary immune and stromal states. Across subtypes, epithelial compartments encoded dominant regulatory programs that aligned with distinct TME signaling architectures, suggesting that TNBC phenotypes emerge through coordinated ecosystem-level interactions. Immune-associated IFN/NF κ B signaling in CC1, ECM-fibrotic remodeling in CC2, and stress-adaptive metabolic rewiring in CC3 were all maintained across spatially organized invasive regions, supporting the concept that tumor behavior is shaped by integrated epithelial-microenvironment states rather than isolated cellular compartments.

Our detailed analysis identifies the CC1 subtype as immunologically active but chronically stimulated, where spatial epithelial-immune interactions drive simultaneous immune visibility and exhaustion. Transcriptomics revealed high IRF1/NF κ B activity and identified CD244 as a complex, dual-role prognostic marker associated with both cytolytic activity (18) and exhaustion

(19–21). CC1 tumors exhibit an "immune-hot" center, characterized by heightened MHC-I processing, IRF1-driven signaling, and elevated cell cycle activity. Meanwhile, elevated DNMT1 suggests that epigenetic mechanisms constrain full immune activation, indicating a potential target for therapeutic reversing of immune evasion, similar to strategies in TNBC to restore immunogenicity (22–24). At the invasive edge, CC1 presents a complex, heterogeneous immune-stromal network. Spatial analyses indicate a compartment-specific role for NFkB: driving immune activation within the TME and facilitating epithelial-immune mimicry via IFNG-induced chemokines in the epithelia (25). Notably, the *CXCL9-CXCR3* axis is prominently featured, reinforcing the potential for this pathway as a prognostic biomarker for favorable prognosis and therapeutic intervention in breast cancer (10,26–29). The improved pembrolizumab response observed in the I-SPY2 clinical trial further supports the translational relevance of the inflammatory yet epigenetically constrained CC1 ecosystem and its potential susceptibility to immune checkpoint blockade.

Our study shows that CC2 is an immunologically cold, fibroblast-dominated TNBC subtype defined by dense ECM organization, hypoxia-associated signaling, and stromal plasticity, in contrasts to the immune-engaged CC1 subtype. Instead of supporting proliferative epithelial growth, this subtype prioritizes structural reinforcement, which aligns with aggressive, treatment-resistant TNBC phenotype (30–34). Spatial analyses demonstrated a collagen-dense tumor core together with a more dynamic and invasive edge. These findings align with previous studies suggesting that ECM-dense regions may restrict immune infiltration (30), while the invasive edges retain stromal plasticity. Early TGFB activation was associated with aggressive, pro-fibrotic environment, limiting immune infiltration and may contribute to the attenuated pembrolizumab benefit observed in the I-SPY2 trial. These findings suggest that ECM remodeling and TGFB-

associated signaling represent candidate therapeutic vulnerabilities for overcoming immune resistance within this subtype.

CC3 represents a distinct immune-excluded TNBC ecosystem in which both tumor and TME compartments preferentially exhibit stress-adaptive and metabolically flexible programs over proliferative signaling. At the molecular level, this subtype shifts away from canonical ERK signaling toward JNK/p38 pathways, allowing the tumor to endure inflammatory and metabolic pressure (35,36). Bulk RNA-seq identified a distinct NRS marked by ESR1/PGR-associated DEGs, suggesting a luminal epithelial plasticity despite ER negativity (37). Spatial profiling further clarified that this apparent luminal-like signature was associated with metabolic adaptation rather than canonical hormone signaling. Enrichment of *ESRRA*, *NR1P1*, and FA metabolism programs suggests a coordinated metabolic state supporting mitochondrial resilience and lipid utilization. At the invasive edge, these metabolic pathways converged with WNT and TGFB-associated signaling and were associated with epithelial plasticity and stromal remodeling. Together, these findings support CC3 as a metabolically adaptive and stress-responsive ecosystem distinct from the fibrotic architecture of CC2 and nominate nuclear receptor, metabolic, and WNT/TGFB-associated pathways as candidate therapeutic vulnerabilities.

Although TNBC studies frequently focus on invasive regions, analyses of non-invasive compartments provided complementary insights into early subtype-associated ecosystem organization. We found that non-invasive regions display subtype-specific signaling programs, suggesting detectable signatures prior to overt invasion. Notably, CC1 regions demonstrated enrichment of IFN-associated signaling and CXCL13 expression, suggesting early immune-associated activation programs. Conversely, CC2 exhibited TME-restricted TGFB activation, suggestive of early stromal crosstalk, while CC3 displayed early metabolic reprogramming

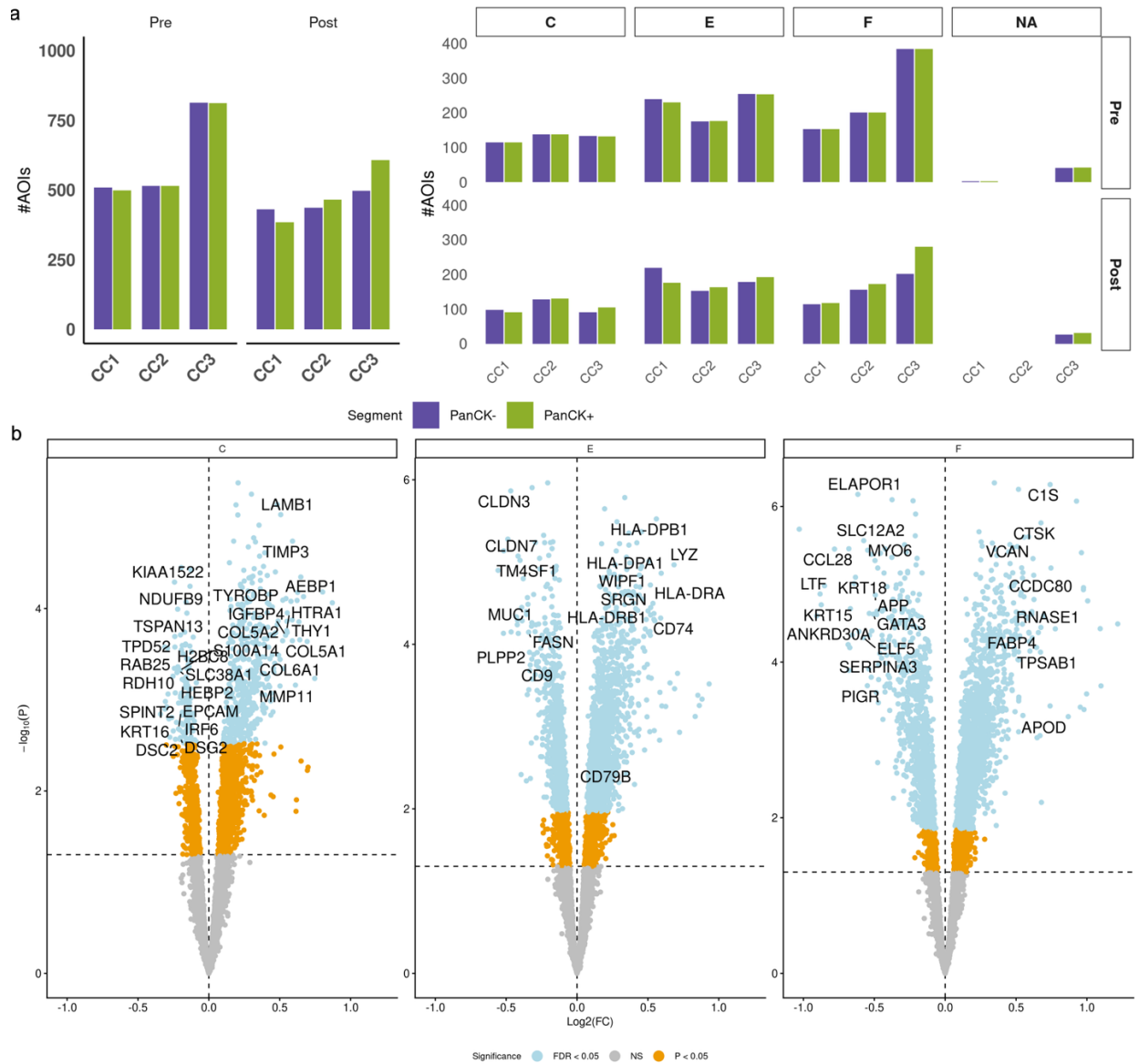
through WNT and adipogenesis signatures. Collectively, these findings suggest that coordinated ecosystem-level organization may emerge prior to histologic invasion and support the potential relevance of non-invasive regions for early therapeutic intervention.

Overall, this study provides a framework for understanding TNBC as a spatially organized ecosystem disease. By integrating tumor-intrinsic and microenvironmental program across spatial compartments, our findings suggest that ecosystem-level organization may provide clinically actionable information beyond conventional bulk transcriptomic classification alone. These results support the development of subtype-specific therapeutic strategies targeting coordinated epithelial-microenvironment states and establish a foundation for spatially informed precision oncology approaches in TNBC.

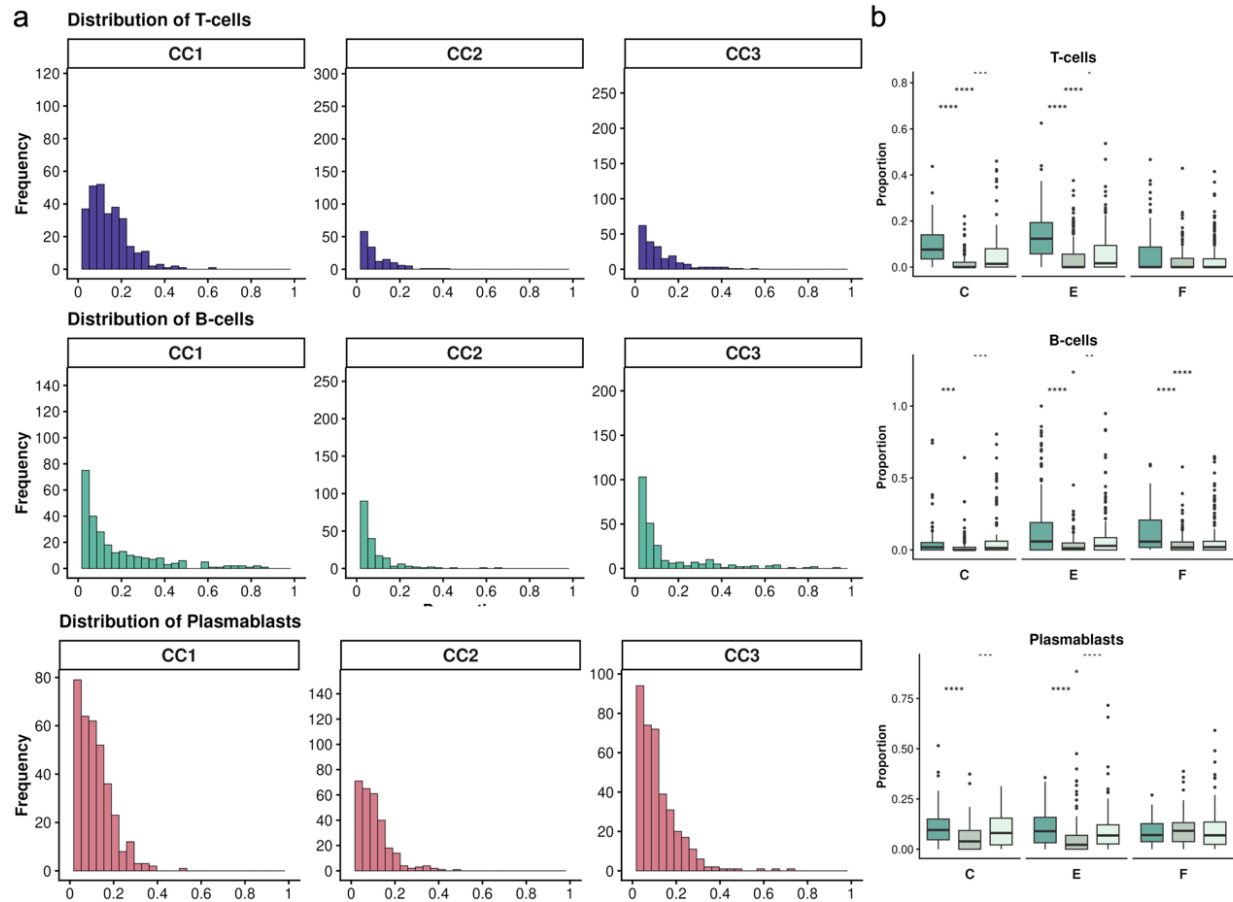
While the preceding chapters focused on tumor-intrinsic and spatially organized microenvironmental ecosystems within TNBC tissue, anti-tumor immunity also operates at the systemic level. Peripheral immune populations dynamically respond to therapy and may provide complementary insight into treatment-associated immune remodeling that cannot be captured from tissue alone. To characterize these systemic immune dynamics, we next analyzed longitudinal PBMC profiles from patients receiving chemotherapy or chemo-immunotherapy.

This Chapter is being prepared for publication and a version of it is in preprint at DOI: <https://doi.org/10.1101/2025.07.04.663242>. Please refer to the online manuscript for detailed Supplementary Materials.

3.6 Supplemental Materials

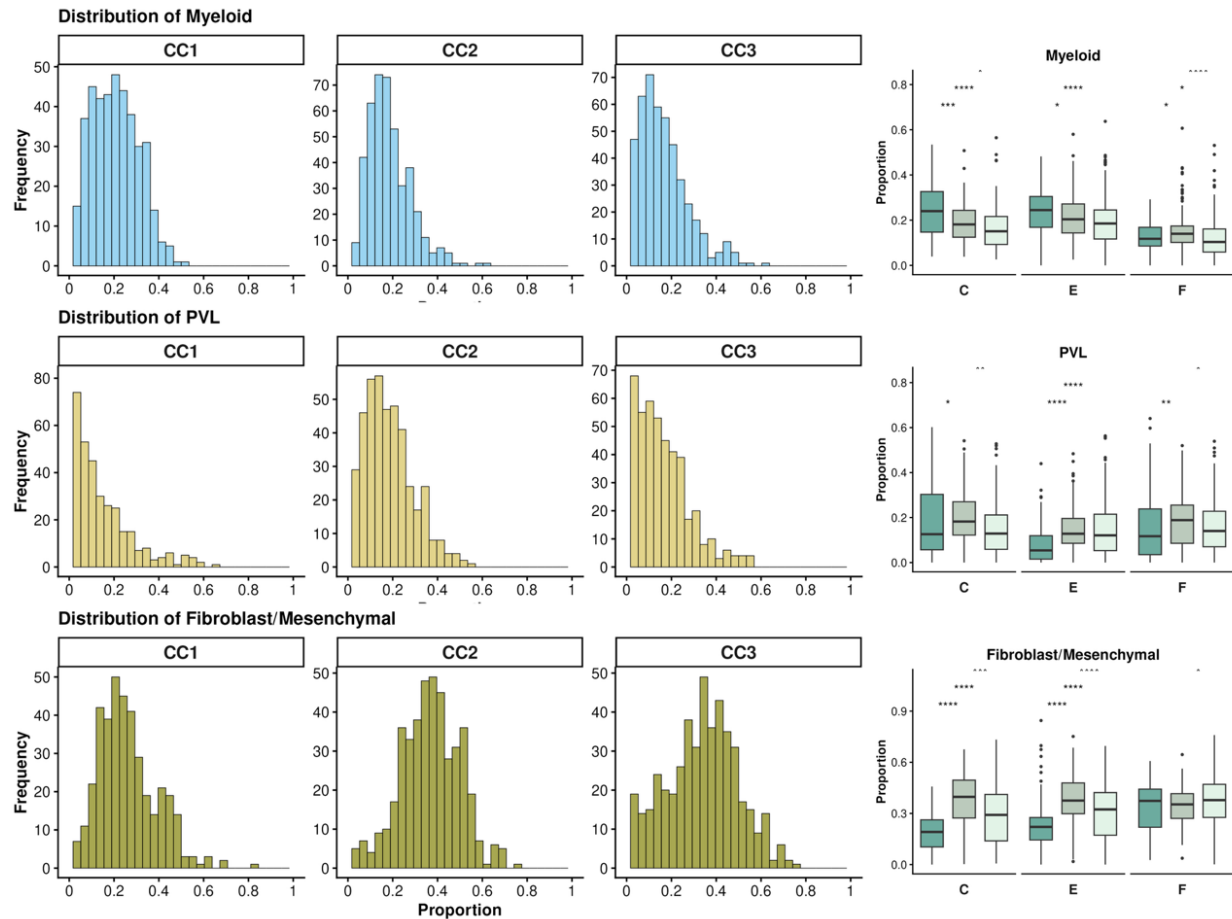


Supplementary Figure 3.1: Additional spatial characterization of TNBC subtype ecosystems.
a, The distribution of AOIs pre-processing (left/top panels) and post-processing (right/bottom panels) for CC1-CC3 TNBC samples. **b**, Differential expression captured for each region (center (c), edge (e), isolated foci (F)) in PanCK- with respect to PanCK+. Top genes are highlighted.



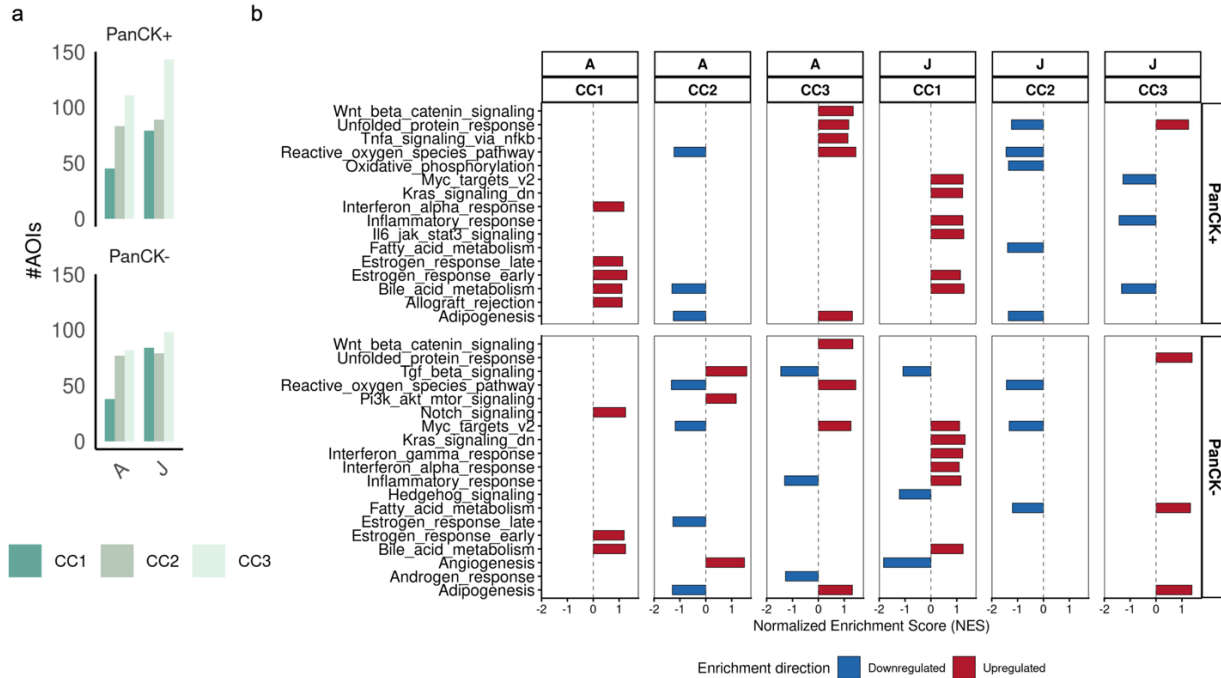
Supplementary Figure 3.2: Regional EcoTyper cell-state distributions across TNBC subtypes.

a, The panels capture the cell type frequencies of the T-cells, B- cells, Plasmablasts, Myeloid cells, PVLs and mesenchymal and fibroblasts in CC1 – CC3 (as annotated) in PanCK- AOIs. **b**, comparison of the proportion of each cell type within AOIs captured in each spatial region between TNBC subtypes.



Supplementary Figure 3.3: Regional EcoTyper cell-state distributions across TNBC subtypes, Continued.

a, The panels capture the cell type frequencies of the T-cells, B- cells, Plasmablasts, Myeloid cells, PVLs and mesenchymal and fibroblasts in CC1 – CC3 (as annotated) in PanCK- AOIs. **b**, comparison of the proportion of each cell type within AOIs captured in each spatial region between TNBC subtypes.



Supplementary Figure 3.4: Additional spatial regulatory network analyses across TNBC regions.

a, bar plots showing distribution of AOIs for each region across TNBC subtypes and epithelial morphology (atypia (A) and adjacent normal (J)). **b**, Molecular characterization of non-invasive tumors (top) and TME (bottom) compartments across TNBC subtypes and epithelial morphology. The most relevant and significant molecular features from Hallmark are illustrated. FDRs are based on the Wilcoxon rank-sum test. $FDR < 0.05$ shown; blue = negative association, red = positive association.

REFERENCES

1. Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discov.* 2022 Jan 12;12(1):31–46. doi:10.1158/2159-8290.CD-21-1059
2. Kvokačková B, Remšík J, Jolly MK, Souček K. Phenotypic Heterogeneity of Triple-Negative Breast Cancer Mediated by Epithelial–Mesenchymal Plasticity. *Cancers (Basel).* 2021 May 2;13(9):2188. doi:10.3390/cancers13092188
3. Chen X, Song E. The theory of tumor ecosystem. *Cancer Commun (Lond).* 2022 Jun 1;42(7):587–608. doi:10.1002/cac2.12316
4. Wang X, Venet D, Lifrange F, Larsimont D, Rediti M, Stenbeck L, Dupont F, Rouas G, Garcia AJ, Craciun L, Buisseret L, Ignatiadis M, Carausu M, Bhalla N, Masarapu Y, Villacampa EG, Franzén L, Saarenpää S, Kvastad L, Thrane K, Lundeberg J, Rothé F, Sotiriou C. Spatial transcriptomics reveals substantial heterogeneity in triple-negative breast cancer with potential clinical implications. *Nat Commun.* 2024 Nov 26;15(1):10232. doi:10.1038/s41467-024-54145-w
5. Wu SZ, Al-Eryani G, Roden DL, Junankar S, Harvey K, Andersson A, Thennavan A, Wang C, Torpy JR, Bartonicek N, Wang T, Larsson L, Kaczorowski D, Weisenfeld NI, Uytingco CR, Chew JG, Bent ZW, Chan CL, Gnanasambandapillai V, Dutertre CA, Gluch L, Hui MN, Beith J, Parker A, Robbins E, Segara D, Cooper C, Mak C, Chan B, Warriar S, Ginhoux F, Millar E, Powell JE, Williams SR, Liu XS, O'Toole S, Lim E, Lundeberg J, Perou CM, Swarbrick A. A single-cell and spatially resolved atlas of human breast cancers. *Nat Genet.* 2021 Sep;53(9):1334–47. doi:10.1038/s41588-021-00911-1
6. Bassez A, Vos H, Van Dyck L, Floris G, Arijs I, Desmedt C, Boeckx B, Vanden Bempt M, Nevelsteen I, Lambein K, Punie K, Neven P, Garg AD, Wildiers H, Qian J, Smeets A, Lambrechts D. A single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. *Nat Med.* 2021 May;27(5):820–32. doi:10.1038/s41591-021-01323-8
7. Allam M, Cai S, Coskun AF. Multiplex bioimaging of single-cell spatial profiles for precision cancer diagnostics and therapeutics. *npj Precis Onc.* 2020 May 1;4(1):11. doi:10.1038/s41698-020-0114-1
8. Seo ES, Park S, Cho EY, Lee JE, Jung HH, Hyeon J, An S, Kim SW, Shin J, Ahn JS, Park YH, Im YH, Kim H, Lee S, Park WY, Kim JY. Spatial and genomic profiling of residual breast cancer after neoadjuvant chemotherapy unveil divergent fates for each breast cancer subtype. *Cell Rep Med.* 2025 Jun 17;6(6):102164. doi:10.1016/j.xcrm.2025.102164
9. Danaher P, Kim Y, Nelson B, Griswold M, Yang Z, Piazza E, Beechem JM. Advances in mixed cell deconvolution enable quantification of cell types in spatial transcriptomic data. *Nat Commun.* 2022 Jan 19;13(1):385. doi:10.1038/s41467-022-28020-5
10. Humblin E, Kamphorst AO. CXCR3-CXCL9: It's All in the Tumor. *Immunity.* 2019 Jun 18;50(6):1347–9. doi:10.1016/j.immuni.2019.05.013

11. Ding JH, Xiao Y, Zhao S, Xu Y, Xiao YL, Shao ZM, Jiang YZ, Di GH. Integrated analysis reveals the molecular features of fibrosis in triple-negative breast cancer. *Molecular Therapy - Oncolytics*. 2022 Mar 17;24:624–35. doi:10.1016/j.omto.2022.02.003
12. Lappano R, Talia M, Cirillo F, Rigiracciolo DC, Scordamaglia D, Guzzi R, Miglietta AM, De Francesco EM, Belfiore A, Sims AH, Maggiolini M. The IL1 β -IL1R signaling is involved in the stimulatory effects triggered by hypoxia in breast cancer cells and cancer-associated fibroblasts (CAFs). *J Exp Clin Cancer Res*. 2020 Aug 10;39(1):153. doi:10.1186/s13046-020-01667-y
13. Hao M, Huang B, Wu R, Peng Z, Luo KQ. The Interaction between Macrophages and Triple-negative Breast Cancer Cells Induces ROS-Mediated Interleukin 1 α Expression to Enhance Tumorigenesis and Metastasis. *Advanced Science*. 2023;10(29):2302857. doi:10.1002/advs.202302857
14. Wang X, Zhao L, Song X, Wu X, Krishnamurthy S, Semba T, Shao S, Knafl M, Coffey LW, Alexander A, Vines A, Bopparaju S, Woodward WA, Chu R, Zhang J, Yam C, Loo LWM, Nasrazadani A, Huong LP, Woodman SE, Futreal A, Tripathy D, Ueno NT. Genomic and transcriptomic analyses identify distinctive features of triple-negative inflammatory breast cancer. *npj Precis Onc*. 2024 Nov 18;8(1):265. doi:10.1038/s41698-024-00729-0
15. Chen D, Van der Ent MA, Lartey NL, King PD. EPHB4-RASA1-Mediated Negative Regulation of Ras-MAPK Signaling in the Vasculature: Implications for the Treatment of EPHB4- and RASA1-Related Vascular Anomalies in Humans. *Pharmaceuticals (Basel)*. 2023 Jan 23;16(2):165. doi:10.3390/ph16020165
16. Chin YR, Yoshida T, Marusyk A, Beck AH, Polyak K, Toker A. TARGETING AKT3 SIGNALING IN TRIPLE NEGATIVE BREAST CANCER. *Cancer Res*. 2014 Feb 1;74(3):964–73. doi:10.1158/0008-5472.CAN-13-2175
17. Gong R, Huang Y, Wang X, Chen X, Tian Z, Ren H. Epithelial cells mimic immune cells: a novel path toward tumor immunotherapy. *Cancer Biol Med*. 2021 Nov 15;18(4):937–9. doi:10.20892/j.issn.2095-3941.2020.0406
18. NF- κ B in inflammation and cancer | Cellular & Molecular Immunology [Internet]. [cited 2026 May 2]. Available from: <https://www.nature.com/articles/s41423-025-01310-w>
19. Sun L, Gang X, Li Z, Zhao X, Zhou T, Zhang S, Wang G. Advances in Understanding the Roles of CD244 (SLAMF4) in Immune Regulation and Associated Diseases. *Front Immunol*. 2021;12:648182. doi:10.3389/fimmu.2021.648182
20. Wherry EJ, Kurachi M. Molecular and cellular insights into T cell exhaustion. *Nat Rev Immunol*. 2015 Aug;15(8):486–99. doi:10.1038/nri3862
21. Cheng J, Ding X, Xu S, Zhu B, Jia Q. Gene expression profiling identified TP53MutPIK3CAWild as a potential biomarker for patients with triple-negative breast cancer treated with immune checkpoint inhibitors. *Oncol Lett*. 2020 Apr;19(4):2817–24. doi:10.3892/ol.2020.11381

22. Fan W, Li W, Li L, Qin M, Mao C, Yuan Z, Wang P, Chu B, Jiang Y. Bifunctional HDAC and DNMT inhibitor induces viral mimicry activates the innate immune response in triple-negative breast cancer. *European Journal of Pharmaceutical Sciences*. 2024 Jun 1;197:106767. doi:10.1016/j.ejps.2024.106767
23. Wang R, Dong X, Zhang X, Liao J, Cui W, Li W. Exploring viral mimicry combined with epigenetics and tumor immunity: new perspectives in cancer therapy. *Int J Biol Sci*. 2025 Jan 6;21(3):958–73. doi:10.7150/ijbs.103877
24. Wong KK. DNMT1: A key drug target in triple-negative breast cancer. *Semin Cancer Biol*. 2021 Jul;72:198–213. doi:10.1016/j.semcancer.2020.05.010
25. Ding JH, Xiao Y, Zhao S, Xu Y, Xiao YL, Shao ZM, Jiang YZ, Di GH. Integrated analysis reveals the molecular features of fibrosis in triple-negative breast cancer. *Molecular Therapy - Oncolytics*. 2022 Mar 17;24:624–35. doi:10.1016/j.omto.2022.02.003
26. Cao X, Song Y, Wu H, Ren X, Sun Q, Liang Z. Quantifying spatial CXCL9 distribution with image analysis predicts improved prognosis of triple-negative breast cancer. *Front Genet*. 2024 Jun 18;15:1421573. doi:10.3389/fgene.2024.1421573
27. Seitz S, Dreyer TF, Stange C, Steiger K, Bräuer R, Scheutz L, Multhoff G, Weichert W, Kiechle M, Magdolen V, Bronger H. CXCL9 inhibits tumour growth and drives anti-PD-L1 therapy in ovarian cancer. *Br J Cancer*. 2022 Jun;126(10):1470–80. doi:10.1038/s41416-022-01763-0
28. Lv Y, Lv D, Lv X, Xing P, Zhang J, Zhang Y. Immune Cell Infiltration-Based Characterization of Triple-Negative Breast Cancer Predicts Prognosis and Chemotherapy Response Markers. *Front Genet*. 2021 Mar 19;12. doi:10.3389/fgene.2021.616469
29. Wang X, Zhang Y, Wang S, Ni H, Zhao P, Chen G, Xu B, Yuan L. The role of CXCR3 and its ligands in cancer. *Front Oncol*. 2022 Nov 21;12. doi:10.3389/fonc.2022.1022688
30. Hammerl D, Martens JWM, Timmermans M, Smid M, Trapman-Jansen AM, Foekens R, Isaeva OI, Voorwerk L, Balcioglu HE, Wijers R, Nederlof I, Salgado R, Horlings H, Kok M, Debets R. Spatial immunophenotypes predict response to anti-PD1 treatment and capture distinct paths of T cell evasion in triple negative breast cancer. *Nat Commun*. 2021 Sep 27;12(1):5668. doi:10.1038/s41467-021-25962-0
31. Liverani C, De Vita A, Spadazzi C, Miserocchi G, Cocchi C, Bongiovanni A, De Lucia A, La Manna F, Fabbri F, Tebaldi M, Amadori D, Tasciotti E, Martinelli G, Mercatali L, Ibrahim T. Lineage-specific mechanisms and drivers of breast cancer chemoresistance revealed by 3D biomimetic culture. *Mol Oncol*. 2022 Feb;16(4):921–39. doi:10.1002/1878-0261.13037
32. Arcucci, A., Ruocco, M. R., Granato, G., Sacco, A. M. & Montagnani, S. Cancer: an oxidative crosstalk between solid tumor cells and cancer associated fibroblasts. *BioMed Res. Int*. 2016, 1–7 (2016). - Google Search [Internet]. [cited 2026 May 2]. Available from: <https://onlinelibrary.wiley.com/doi/10.1155/2016/4502846>

33. Mollah F, Varamini P. Overcoming Therapy Resistance and Relapse in TNBC: Emerging Technologies to Target Breast Cancer-Associated Fibroblasts. *Biomedicines*. 2021 Dec 15;9(12):1921. doi:10.3390/biomedicines9121921
34. Tan Q, Xu L, Zhang J, Ning L, Jiang Y, He T, Luo J, Chen J, Lv Q, Yang X, Xie H. Breast cancer cells interact with tumor-derived extracellular matrix in a molecular subtype-specific manner. *Biomater Adv*. 2023 Mar;146:213301. doi:10.1016/j.bioadv.2023.213301
35. Pham TT, Angus SP, Johnson GL. MAP3K1: Genomic Alterations in Cancer and Function in Promoting Cell Survival or Apoptosis. *Genes Cancer*. 2013 Nov;4(11–12):419–26. doi:10.1177/1947601913513950
36. Loesch M, Chen G. The p38 MAPK stress pathway as a tumor suppressor or more? *Front Biosci*. 2008 May 1;13:3581–93. doi:10.2741/2951
37. Hossain F, Danos D, Zabaleta J, Wu XC, Del Valle L, Hicks C, Wu J, Tomsic J, Neuhausen S, Ding YC, Mukund K, Mesrizadeh Z, Subramaniam S, Ochoa A, Seewaldt V, Miele L. Immunogenomic diversity of triple-negative breast cancers in obese and non-obese black and white women. *NPJ Breast Cancer*. 2025 Nov 12;11(1):126. doi:10.1038/s41523-025-00836-6
38. Yu G, Wang LG, Han Y, He QY. Clusterprofiler: an r package for comparing biological themes among gene clusters. *OMICS: A Journal of Integrative Biology*. 2012 May;16(5):284–7. doi:10.1089/omi.2011.0118
39. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, Ma'ayan A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*. 2016 Jul;44(W1):W90–7. doi:10.1093/nar/gkw377
40. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S, Bork P, Jensen LJ, von Mering C. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res*. 2023 Jan 6;51(D1):D638–46. doi:10.1093/nar/gkac1000
41. Han H, Cho JW, Lee S, Yun A, Kim H, Bae D, Yang S, Kim CY, Lee M, Kim E, Lee S, Kang B, Jeong D, Kim Y, Jeon HN, Jung H, Nam S, Chung M, Kim JH, Lee I. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Research*. 2018 Jan;46(D1):D380–6. doi:10.1093/nar/gkx1013
42. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*. 2003 Nov;13(11):2498–504. doi:10.1101/gr.1239303
43. Kumar T, Nee K, Wei R, He S, Nguyen QH, Bai S, Blake K, Pein M, Gong Y, Sei E, Hu M, Casasent AK, Thennavan A, Li J, Tran T, Chen K, Nilges B, Kashikar N, Braubach O, Ben Cheikh B, Nikulina N, Chen H, Teshome M, Menegaz B, Javaid H, Nagi C, Montalvan J, Lev T, Mallya S, Tifrea DF, Edwards R, Lin E, Parajuli R, Hanson S, Winocour S, Thompson A, Lim B,

Lawson DA, Kessenbrock K, Navin N. A spatially resolved single-cell genomic atlas of the adult human breast. *Nature*. 2023 Aug;620(7972):181–91. doi:10.1038/s41586-023-06252-9

44. Luca BA, Steen CB, Matusiak M, Azizi A, Varma S, Zhu C, Przybyl J, Espín-Pérez A, Diehn M, Alizadeh AA, van de Rijn M, Gentles AJ, Newman AM. Atlas of clinically distinct cell states and ecosystems across human solid tumors. *Cell*. 2021 Oct 14;184(21):5482-5496.e28. doi:10.1016/j.cell.2021.09.014

Chapter 4 Differential peripheral immune dynamics underlie therapeutic response to chemotherapy and chemo-immunotherapy in triple-negative breast cancer

4.1 Abstract

Triple-negative breast cancer (TNBC) remains the most aggressive breast cancer subtype, with limited treatment options and variable response to immune checkpoint inhibitors. While tumor-infiltrating lymphocytes have been extensively studied, the integration of system-level peripheral immune dynamics with mechanistic immune regulation underlying therapeutic response and resistance remain poorly defined. Here, we integrate systems-level immune state modeling with pathway-level mechanistic inference to analyze single-cell RNA sequencing of peripheral blood mononuclear cells from advanced TNBC patients treated with paclitaxel alone (chemotherapy) or in combination with anti-PD-L1 antibody atezolizumab (combination). This framework leverages treatment arm, longitudinal sampling, and clinical response to resolve coordinated immune programs across lymphoid and myeloid compartments. Using this approach, we identified distinct treatment- and response-specific immune states in pre- and post-treatment. Chemotherapy responders displayed pre-treatment adaptive immune priming, whereas combination therapy responders exhibited pre-existing effector T cell activity coupled with tumor tissue PD-L1 expression. In contrast, chemotherapy non-responders developed persistent post-treatment immune dysregulation in regulatory and terminal effector programs, while combination therapy non-responders demonstrated maladaptive remodeling of adaptive and innate lymphoid compartments, including dysfunctional NK and metabolically reprogrammed myeloid populations. Across both regimens, pathways involving protein translation, metabolic adaptation, and stress signaling emerged as critical modulators of response. These findings suggest that coordinated adaptive–innate immune dynamics underlie therapeutic efficacy, whereas systemic

immune exhaustion and myeloid immunoregulation lead to resistance. Projection of these peripheral immune programs onto independent I-SPY2 showed concordant associations with tumor immune phenotypes and pathological complete response, supporting generalizability of the identified systemic immune states. Our study demonstrates the utility of an integrative systems-level approach for linking peripheral immune state organization with mechanistic insights, informing immune response and resistance in TNBC.

4.2 Introduction

Triple-negative breast cancer (TNBC) is the most aggressive subtype of breast cancer, characterized by early relapse, poor prognosis, and an absence of targeted therapy. Despite the use of standard chemotherapy, approximately one-third of patients with stage II/III TNBC develop distant metastases within 2-3 years of diagnosis (1,2). The addition of immune checkpoint inhibitors (ICIs) to chemotherapy have improved survival in several malignancies with high tumor mutational burdens (3). This prompted investigation into their efficacy in TNBC, a relatively immunogenic subtype of breast cancer, enriched for tumor-infiltrating lymphocytes (TILs). In the IMpassion130 trial, combination of atezolizumab with Nab-paclitaxel improved progression-free survival (PFS) in PD-L1⁺ advanced TNBC patients (median PFS: 7.5 vs. 5.0 months), leading to accelerated regulatory approval (4,5). However, follow-up studies, including IMpassion131, which paired atezolizumab with paclitaxel, and a separate international phase III trial (ALEXANDRA/IMpassion030), resulted in responders and non-responders, even in PD-L1⁺ patients (2,6). These mixed outcomes underscore a critical challenge that only a subset of TNBC patients achieve a durable response, highlighting the need to better define the system-level immune mechanisms that differentiate therapeutic response and resistance.

While most efforts to date have focused on local immune responses within the tumor

microenvironment (TME), effective antitumor immunity requires continuous interplay between the TME and the peripheral immune system. Circulating immune cells reflect systemic immune competence and, importantly, offer accessible and dynamic biomarkers of treatment response and disease progression. To date, peripheral immune profiling in breast cancer has largely relied on broad metrics such as neutrophil-to-lymphocyte (NLR) or lymphocyte-to-monocyte ratios (LMR) (7,8), or limited immunophenotyping of major immune subsets using next-generation sequencing technologies such as single-cell RNA sequencing (scRNA-seq) (1). Emerging studies in TNBC suggest that systemic immune dysregulation is a defining feature. Early-stage patients exhibit reduced circulating B cells and elevated regulatory T cells compared to healthy controls (9), while patients with advanced disease demonstrate widespread immunosuppression marked by diminished CD4⁺ T cells, expansion of monocytes, and dysfunctional effector populations (10). These studies have largely lacked integrative, systems-level analyses capable of resolving the mechanistic basis of systemic immune response and resistance, particularly after chemotherapy versus combination therapy. Large neoadjuvant trials such as I-SPY2 have generated clinically annotated transcriptomic datasets linking systemic immune signaling to treatment response but these analyses largely rely on bulk immune signatures and do not resolve high-resolution immune subset states or coordinated peripheral immune programs (11,12).

Here, we apply an integrative systems-level immune state analysis to scRNA-seq data from peripheral blood mononuclear cells (PBMC/blood) and breast tissue with advanced TNBC treated with paclitaxel alone or in combination with anti-PD-L1 antibody atezolizumab (ATZ) (combination therapy) (13). Our analysis highlights the transcriptional programs and immune cell states associated with therapeutic response (in responders and non-responders) and treatment regimens (chemo- and combination therapy). Within peripheral immune cells (blood),

chemotherapy responders exhibited pre-treatment adaptive immune priming, whereas combination therapy responders showed pre-existing effector adaptive immunity with innate lymphoid activation. Chemotherapy non-responders developed persistent post-treatment immune dysregulation in regulatory and terminal effector programs, while combination therapy non-responders showed post-treatment remodeling in adaptive and innate lymphoid compartments. Together, our study delineates coordinated systemic immune programs associated with therapeutic efficacy and resistance, providing insight into peripheral regulation in TNBC.

4.3 Methods

Data overview

Published scRNA-seq data from Zhang *et al.* (GEO accession: GSE169246) which included 489,490 cells from 22 advanced TNBC patients, spanning 75 tissue (primary/metastatic) and blood samples were downloaded. The samples were collected pre-treatment, 4 weeks post-treatment, and at progression (**Supplementary Table 4.1**). Patient P002 was excluded as indicated in Zhang *et al.* (7). The final dataset used for all downstream analyses included tissue samples (n = 15) and blood samples (n = 20). Patients had received either paclitaxel alone (chemotherapy; tissue n = 7, blood n = 10) or paclitaxel plus atezolizumab (combination therapy; tissue n = 8, blood n = 10), with documented treatment response available for most cases. Clinical response, including responder and non-responder, tumor-infiltrating lymphocytes (TIL), and PD-L1 scores, were downloaded from the original study.

Data filtering, scaling and dimensionality reduction

Data was filtered by excluding cells based on the following criteria: more than 10% mitochondria read, less than 500 expressed genes, less than 500 detected transcripts, and a complexity (log10 genes/UMI) of less than 80%. Following this QC, 432,837 cells remained for

downstream analysis. Tissue-derived and PBMC samples were processed independently using Seurat (v.5.3.0), resulting in two data objects comprising 143,085 cells (tissue) and 289,752 cells (PBMC) across 27085 features/genes. Data scaling and dimensionality reduction was performed per vignette, with the optimal dimensions for clustering identified using *ElbowPlot* (tissue nPCs = 30 and PBMC nPCs= 31). The cells were clustered using the Louvain algorithm and visualized using UMAP.

Detecting cell types (Cluster annotation)

Cell type identities for each cluster detected was established using a threefold strategy: (1) SingleR with multiple reference datasets (DatabaseImmuneCellExpressionData, MonacoImmuneData, NovershternHeamtopoieticData), (2) differential expression testing via Seurat's FindAllMarkers (Wilcoxon, logFC > 0.25, min.pct > 0.25, Bonferroni < 0.05), and (3) canonical marker expression and legacy knowledge. Markers were visualized using feature and bubble plots (top DEGs), enabling lineage assignment and transcriptional state assessment. Clustering yielded distinct immune populations in tissue samples, including 15 T cell clusters, 5 B cell clusters, 2 NK cell clusters, and 8 myeloid clusters. PBMC samples revealed 13 T cell clusters, 4 B cell clusters, 5 NK cell clusters, and 9 myeloid clusters.

Profiling T cell subsets

From the dataset, CD4⁺ and CD8⁺ T cells were extracted from both tissue and PBMC compartments for higher-resolution analysis. In tissue, a total of 90,286 CD4⁺/CD8⁺ T cells (comprising 3 initial CD4⁺ clusters and 12 initial CD8⁺ clusters) were subsampled and re-clustered into 7 CD4⁺ and 14 CD8⁺ transcriptionally and functionally distinct subsets. In PBMC, 120,213 CD4⁺/CD8⁺ T cells (from 9 initial CD4⁺ clusters and 14 initial CD8⁺ clusters) were subsampled and re-clustered into 6 CD4⁺ and 15 CD8⁺ distinct subsets. Subsampling was performed to ensure

balanced representation of patient samples and avoid bias from uneven cell recovery. Cell type annotation was guided by established lineage and functional markers, including canonical naïve, central memory, effector, cytotoxic, proliferative, and regulatory T cell signatures (**Supplementary Table 4.2**) (1,2,23,24). To assess the transcriptional characteristics of various T cell subsets, we evaluated curated functional scores (e.g., exhaustion score, cytotoxicity score, anergy score), which were established based on a predefined set of genes obtained from a published pan-cancer T cell study (23), using gene module scoring (*AddModuleScore* function) in Seurat. Each score was calculated as the average expression level of the respective gene set.

Profiling B cell subsets

33,880 B cells (4 original clusters) were subset from the original data object, re-clustered, and grouped into 10 functionally distinct B cell states (plus 2 undefined), based on markers and pathway enrichment related to BCR signaling, MHC-II expression, IFN response, AP-1 transcriptional regulation, and metabolic/stress programs (**Supplementary Table 4.2**) (1).

Profiling NK cell subsets

33,880 B cells (4 original clusters) were subset from the original data object, re-clustered, and grouped into 10 functionally distinct B cell states (plus 2 undefined), based on markers and pathway enrichment related to BCR signaling, MHC-II expression, IFN response, AP-1 transcriptional regulation, and metabolic/stress programs (**Supplementary Table 4.2**) (1).

Profiling Myeloid subsets

In tissue, 17,866 myeloid cells (8 initial clusters) were subset and re-clustered into 23 transcriptionally distinct subsets. In PBMC, 52,949 myeloid cells (9 initial clusters) were subset and re-clustered into 9 subsets. Annotation identified classical (CD14⁺) and non-classical (CD16⁺) monocyte populations, with further subdivision guided by transcriptional heterogeneity and

functional gene programs, including inflammatory signaling, stress responses, and metabolic pathways (**Supplementary Table 4.2**).

Trajectory inference of CD8⁺ T cells

To reconstruct differentiation trajectories of CD8⁺ T cells, we used Monocle 3 (v.1.3.7) (25–29). After dimensionality reduction with UMAP and clustering, we constructed cell trajectories by learning principal graphs using the *learn_graph* function with *use_partition* = FALSE to preserve global connectivity across all subsets. Default parameters were used, as these yielded biologically coherent trajectories, with naïve T cells at the origin and expected branching toward effector and exhausted states. Pseudotime was assigned using *order_cells*, anchoring the root node based on naïve marker expression.

Functional enrichment and PPI/TF network construction

Functional enrichment was identified via functional annotation clustering available through *clusterProfiler* (30) or *Enrichr* (31) in R. Gene Ontology (biological process and molecular function) pathway analyses was performed using *clusterProfiler*'s *EnrichGO* function. A network of high confidence (>0.70) human protein-protein interaction (PPI) was downloaded from STRING database v12 (32), containing 15,588 proteins and 236,930 interactions. Additionally, transcription factor (TF)-target data downloaded from TRRUST-db (33) was mapped onto the PPI interaction network to generate a custom PPI/TF network. Therapy-, response, and timepoint-specific PPI/TF subnetworks were generated using DEGs from adaptive and innate immune subsets described earlier.

In blood these included:

- Adaptive Chemo Responder (Pre-treatment): 746 proteins, 8,108 interactions
- Innate Chemo Responder (Pre-treatment): 986 proteins, 6,876 interactions

- Adaptive Chemo Non-responder (Post-treatment): 176 proteins, 4,078 interactions
- Innate Chemo Non-responder (Post-treatment): 232 proteins, 3,560 interactions
- Adaptive Combo Responder (Pre-treatment): 361 proteins, 3,560 interactions
- Innate Combo Responder (Post-treatment): 2,089 proteins, 13,919 interactions
- Adaptive Combo Non-responder (Post-treatment): 103 proteins, 1,178 interactions
- Innate Combo Non-responder (Post-treatment): 731 proteins, 5,101 interactions

In tissue, these included:

- Adaptive Chemo Responder (Pre-treatment): 211 proteins, 1,139 interactions
- Adaptive Chemo Responder (Post-treatment): 41 proteins, 274 interactions
- Adaptive Chemo Non-responder (Post-treatment): 239 proteins, 3,280 interactions
- Adaptive Combo Responder (Pre-treatment): 166 proteins, 2,579 interactions
- Adaptive Combo Responder (Post-treatment): 38 proteins, 101 interactions
- Adaptive Combo Non-responder (Pre-treatment): 162 proteins, 2,066 interactions

Each network was refined by selecting DEGs, identifying associated TFs, and extracting their 1-step PPI neighbors. All networks were visualized, annotated, and analyzed using Cytoscape (34).

Statistical modeling of blood-tissue expression relationships

We used linear mixed-effects models (LMMs) to quantify the relationship between blood and tissue gene expression across immune subsets. Models were fitted in R using lme4 package. For each gene, expression in tissue was modeled as a function of blood expression, immune cell subsets, and their interaction, with a random intercept for the patient. Models were fit separately for pre- and post-treatment data using restricted maximum likelihood (REML).

Model performance was summarized by marginal and conditional R^2 values using the performance package. To assess robustness, we performed leave-one-patient-out analyses by iteratively removing one patient, refitting the model, and recording (i) changes in REML criterion values, (ii) shifts in fixed-effect estimates, and (iii) variance components. The fixed-effect regression coefficient (β) quantifies the expected change in tissue gene expression per one-unit increase in blood gene expression, adjusted for immune cell subset composition.

External validation using I-SPY2 peripheral transcriptomic data

ISPY2 clinical trial raw data was downloaded from GSE194040, which includes pretreatment tumor gene expression profiles and clinical annotations for patients enrolled across multiple neoadjuvant treatment arms (11). Analyses were restricted to the paclitaxel arm to ensure treatment comparability with the discovery cohort. Gene expression data generated on Agilent microarray platforms (GPL20078) were downloaded from GEO. Probe-level measurements were mapped to gene symbols using platform annotations and collapsed to gene-level expression using median aggregation. Genes with zero variance across samples were excluded. Peripheral immune subset gene signatures derived from PBMC profiling in this study were projected onto the I-SPY2 cohort using single-sample gene set enrichment analysis (ssGSEA) implemented via the GSVA framework, generating per-sample immune program scores. Associations between immune subset scores and pathological complete response (pCR) were evaluated using logistic regression. Multivariate relationships were assessed using LASSO-regularized logistic regression with cross-validation, and model robustness was evaluated using permutation testing.

4.4 Results

4.4.1 Identification of key mechanisms driving peripheral blood immune cell dynamics across treatment and response

Overall, 143,085 and 289,752 high-quality cells across four major cell types were identified after rigorous quality control across tissue and blood, respectively (see Methods). To systematically assess treatment-associated immune alterations, we analyzed changes in circulating immune cell frequencies post- versus pre-treatment across each treatment and response group. Given biased patient-specific clustering of immune cell types in tissue samples - except for T cells - we focused most of our dynamic analysis on blood (**Figure 4.1a, b; Supplementary Figure 4.1a, b**). In chemotherapy-treated patients, responders (R) exhibited a post-treatment decrease in B cells, T cells, and myeloid cells, alongside an increase in NK cells (**Figure 4.1c**). In contrast, non-responders (NR) showed a global increase across all immune cell types following treatment. In the combination therapy group, both responders and non-responders displayed increased immune cell populations post-treatment, except B cells, which decreased in both groups (**Figure 4.1c**).

To resolve immune programs associated with these frequency shifts, we next stratified major immune lineages into functional sub-clusters based on gene expression programs (e.g., activation, exhaustion, migration) using known immunological signatures (**Figure 4.1d**) (1,2). We focused on subtypes that were (1) representative across patients and (2) differentially enriched by treatment response. All comparisons were made between post- and pre-treatment samples within each group. Although stromal TIL levels and peripheral immune cell counts often only moderately correlate in TNBC patients (3), declining TILs following chemotherapy have been linked to treatment response, while persistently high post-treatment TILs can lead to poor outcomes (4). Furthermore, the predictive value of TILs and PD-L1 expression in informing ICI efficacy is well-established (5). In this dataset, analysis of response-specific TIL levels demonstrated that chemotherapy-treated patients, irrespective of response, exhibited consistently low TIL levels (**Figure 3.1e**; See Methods). Notably, chemotherapy responders included samples from both

primary breast and metastatic lesions (stage IV), whereas chemotherapy non-responders were exclusively derived from primary breast tumors with stage II–III disease (**Supplementary Figure 4.1d**). When integrating peripheral immune profiling, functional subsets defined from PBMC decreased only in responders post-treatment in chemotherapy, consistent with systemic immune activation and potential redistribution toward tumor or other tissues (**Figure 4.1e**). In combination therapy, responders, predominantly from lymph node lesions (stage IV), also exhibited post-treatment decreases in circulating functional subsets, particularly in high TIL and PD-L1⁺ tumors, further suggesting immune localization (**Figure 4.1d, e; Supplementary Figure 4.1d**; see Methods). In contrast, combination non-responders displayed post-treatment increases in peripheral immune subsets despite low TIL levels, indicating a disconnect between systemic activation and intratumoral recruitment, consistent with immune dysfunction or exclusion (**Figure 4.1e; Supplementary Figure 4.1d**). Using linear mixed-effects modeling, we demonstrated that peripheral blood gene expression robustly predicted tissue-level expression (pre-treatment $\beta = 0.49$, $t = 44.9$; post-treatment $\beta = 0.57$, $t = 54.6$, $p < 0.001$), with significant interactions indicating that this relationship varies across immune functional subsets (**Supplementary Figure 4.1e**). These results support the hypothesis that systemic immune signatures reflect and are associated with tumor immune states in a subset-dependent manner. To further understand these dynamics, we examined changes within the lymphoid and myeloid subsets across treatment and response groups.

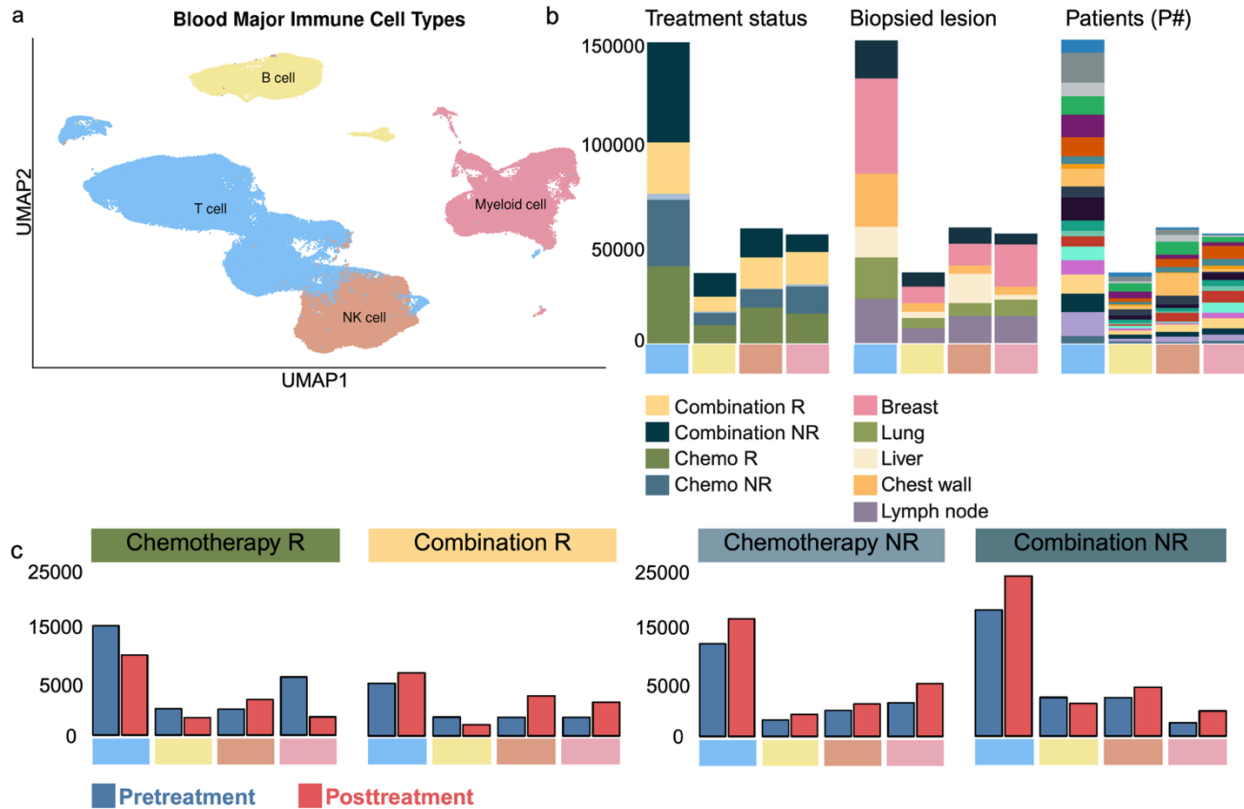


Figure 4.1: Comparative immune landscape of blood samples under different treatment conditions.

a, UMAP plots depict major immune cell types in blood, highlighting T cells (blue), B cells (yellow), monocytes (pink), and NK cells (brown). **b**, Compositions of major immune cell types in treatment and response groups, biopsied lesions, and individual patients. **c**, Bar plots quantifying immune cell frequencies, illustrating treatment-associated shifts in immune cell distribution.

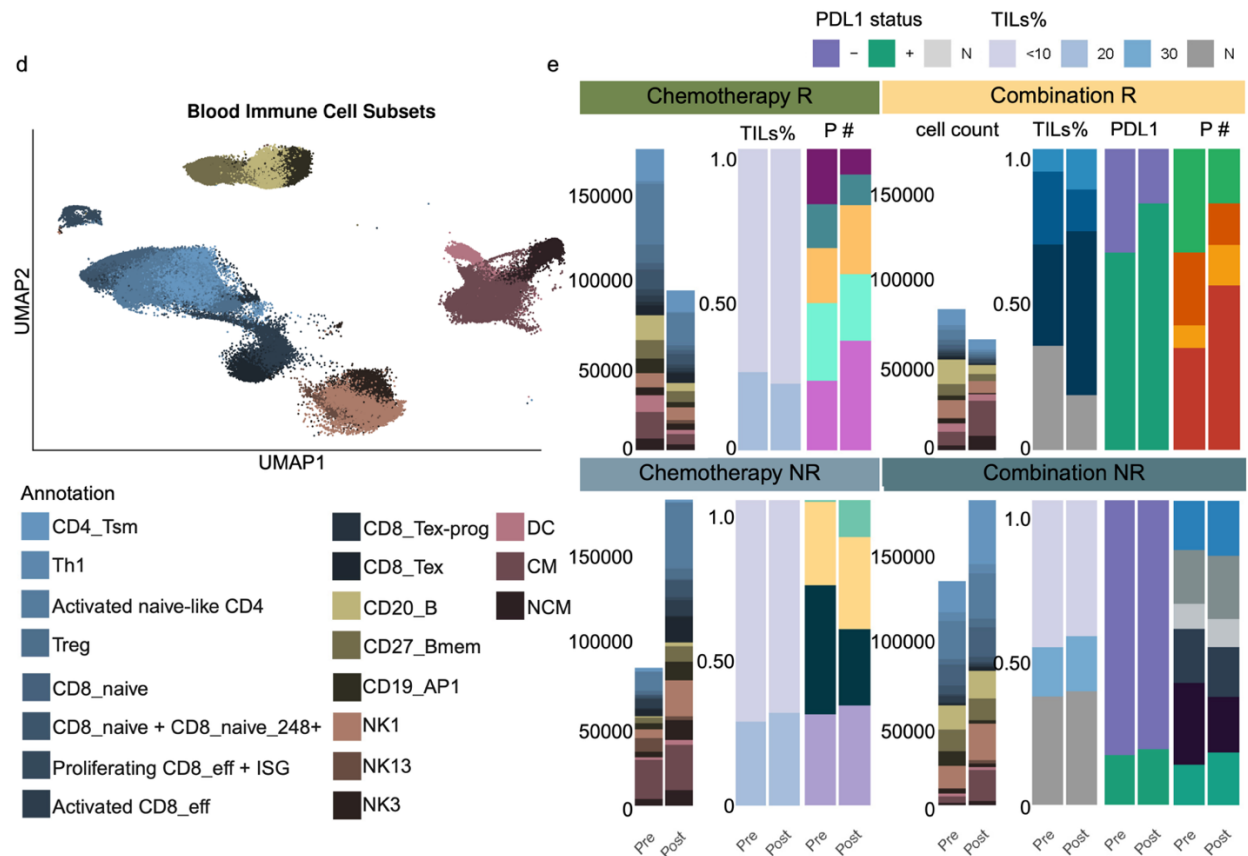


Figure 4.2: Comparative immune landscape of blood samples under different treatment conditions, Continued.

d, UMAP plot depicting blood immune subsets, with colors denoting functionally annotated subsets. **e**, Integrated bar plots showing frequencies of functional immune subsets, proportions of tumor-infiltrating lymphocytes (TILs), PD-L1⁺ (PDL) cells, and patient group IDs (P#), stratified by responder (R) and non-responder (NR) status. This panel links peripheral immune subset dynamics with intratumoral immune contexture.

4.4.2 Lymphoid Compartments

The lymphoid compartment showed distinct composition and mechanistic patterns across treatments and responses. In chemotherapy responders, pre-treatment priming of adaptive immunity was a key feature, whereas post-treatment persistence of effector/exhaustion states characterized resistance, suggesting these patients might benefit from immunotherapy. In combination therapy, pre-treatment effector/exhaustion signatures were associated with response, while resistance was linked to dysregulated adaptive immunity.

4.4.2.1 CD4 T cell subsets: pre-treatment immune priming characterizes responders

CD4⁺ T cell subsets coordinate adaptive immune responses through helper, regulatory, and memory programs that can either support antitumor immunity or promote immune suppression within the TME. Of the six CD4⁺ T cell clusters identified (see Methods, **Supplementary Table 4.2**), four functional clusters (Activated naïve-like CD4, regulatory CD4 (Treg), T-helper 1 (Th1), and T stem-like memory (CD4_Tsm) that were representative across patients were selected for detailed analysis (**Figure 4.2a, b**). In the chemotherapy group, responders exhibited a pre-treatment increase in activated naïve-like CD4 with increased chemokine receptor, IFN response, adhesion, and metabolic scores. Differential analysis further showed enrichment of interferon-driven inflammatory programs and antigen processing, consistent with enhanced immune priming (**Figure 4.2c-f**) (6). Notably, this aligns with prior observations by Zhang *et al.* (7,8) where they identified a naïve CD4 cluster (CD4 Tn_CCR7) as predictive of response to both PTX and Nab-PTX. CD4_Tsm subsets also showed increased stress response and glycolysis scores, suggesting metabolically active memory-like cells responding to sustained immune stimulation (**Figure 4.2c-f**). Treg were also increased pre-treatment with stress response, metabolic programs, and pro-apoptotic scores, consistent with regulatory activation before treatment (**Figure 4.2c, f**). In the combination therapy group, responders also exhibited a pre-treatment increase in activated naïve-like CD4 with elevated lipid metabolism scores alongside CD4_Tsm subsets enriched for stress response, OXPHOS, glycolysis, and pro-apoptotic scores, mirroring the metabolically active immune priming signatures observed in the chemotherapy groups (**Figure 4.2c**).

3.4.2.2 CD4 T cell subsets: metabolic dysregulation characterizes non-responders

In contrast with responders, non-responders lacked evidence of pre-treatment immune priming. In chemotherapy non-responders, treatment induced expansion of activated naïve-like

CD4 with increased IFN response, OXPHOS, glycolysis, and pro-apoptotic scores, alongside Treg enriched for OXPHOS score, indicating a shift towards regulatory and pro-tumor activation (**Figure 4.2c-e**). Differential analysis further showed enrichment of translational machinery in both CD4 subsets, supporting a regulatory and metabolically active but potentially ineffective immune state (**Figure 4.2f**). Notably, while our observations are from peripheral blood, they align with findings by Zhang *et al.* who reported increased T_{reg} frequencies post-chemotherapy in tumor tissue non-responders, suggesting a systemic and intratumoral response associated with treatment resistance (7,8). In combination therapy, non-responders, post-treatment CD4_Tsm expanded with high cytotoxicity, IFN response, glycolysis, and pro-apoptotic scores alongside enrichment for translational programs. Activated naïve-like CD4 exhibited high cytotoxicity, cytokine receptor, adhesion, and metabolic scores (**Figure 4.2c-e**) while differential analysis demonstrated enrichment of cytokine signaling and unfolded protein response (UPR) pathways, indicative of cellular stress response rather than effective immune priming (**Figure 4.2f**). Increased frequencies of T_{reg} with elevated IFN response scores, together with Th1 cells enriched for IFN response and anti-apoptotic scores, reflected an imbalanced and potentially dysfunctional CD4 T cell response (**Figure 4.2c-e**).

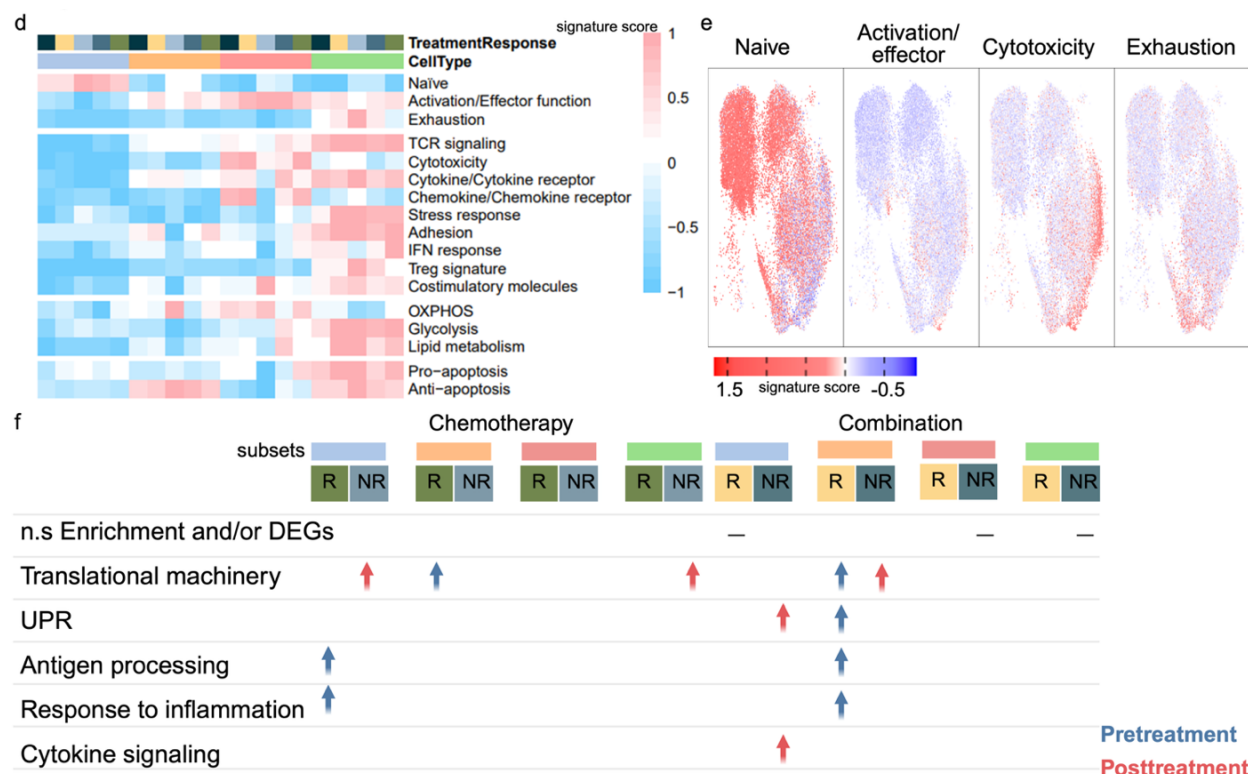


Figure 4.4: Comprehensive landscape of CD4⁺ T cell subsets in blood samples and their distribution across conditions, Continued.

d, Heatmap of 16 curated gene signature scores across CD4⁺ T cell clusters (23). **e**, Expression of four representative gene signatures selected from **d**. **f**, Table showing enriched pathways or functional gene signatures associated with CD4⁺ T cell sub-clusters, stratified by condition (blue: enriched in pre-treatment; red: enriched in post-treatment). This Figure highlights the transcriptional heterogeneity of CD4⁺ T cells in blood, their patient-specific variability, and dynamic changes in sub-cluster composition and function in response to treatment. Pathway enrichment analyses point to key functional programs shaping peripheral CD4⁺ T cell dynamics.

4.4.2.3 CD8 T cell subsets: pre-treatment naïve and effector expansion in chemotherapy and terminal differentiation in combination therapy characterize responders

CD8⁺ T cells mediate cytotoxic antitumor immunity, while chronic stimulation can drive exhaustion and impaired effector function. Of the fifteen CD8⁺ T cell clusters identified (see Methods, **Supplementary Table 4.2**), seven functional clusters (CD8_naïve, CD8_naïve + CD8_naïve_248⁺, effector cells (CD8_eff), Activated CD8_eff, exhausted T cells (CD8_Tex), Proliferating CD8_eff + interferon stimulating genes (ISG), CD8_Tex-prog) that had cell contributions from all patients were selected for detailed analysis (**Figure 4.3a, b**). In the

chemotherapy group, responders exhibited a pre-treatment increase in several naïve clusters including CD8_naïve enriched for cytotoxicity and chemokine receptor scores, and CD8_naïve + CD8_naïve_248+ enriched for cytotoxicity, IFN response, OXPHOS, and glycolysis scores, alongside differential enrichment in translational and antigen processing programs (**Figure 4c-g, Supplementary Table 4.2**). The effector CD8⁺ subsets, including CD8_eff, activated CD8_eff, proliferating CD8_eff + ISG, and CD8_Tex-prog, were elevated, highlighting the association of transitional and cytotoxic states with chemotherapy response (**Figure 4.3c**). CD8_eff cells showed elevated TCR signaling, glycolysis, and pro-apoptotic scores (**Figure 4.3d, e**). Differential analysis further confirmed upregulation of translational machinery, cell migration, cytoskeleton remodeling, and OXPHOS (**Figure 4.3g**). Activated CD8_eff were enriched for naïve, adhesion, IFN response, OXPHOS, and glycolysis scores (**Figure 4.3d, e**). Proliferating CD8_eff + ISG subsets were marked by high stress, OXPHOS, and metabolic scores, and CD8_Tex-prog displayed increased adhesion scores (**Figure 4.3d, e**). Expansion of effector and transitional CD8⁺ T cells with heightened TCR signaling and metabolic activity is consistent with antigen-experienced cytotoxic immune programs during chemotherapy, a pattern consistent with established models with metabolic reprogramming in activated CD8⁺ T cells (9). In combination therapy responders, CD8⁺ T cell dynamics showed distinct features compared to chemotherapy alone. Responders displayed increased frequencies of CD8_naïve with high naïve score, alongside elevated effector subsets (CD8_eff, activated CD8_eff, and CD8_Tex), indicating a functionally engaged and terminally differentiated cytotoxic program (**Figure 4.3c-e**). While CD8_eff cells showed increased activation/effector, exhaustion, TCR signaling, cytotoxicity, cytokine receptor, metabolic, and pro-apoptotic scores (**Figure 4.3d, e**). Differential analysis revealed similar enrichment in cell migration and cytoskeletal remodeling pathways, with combination therapy

uniquely adding inflammatory response programs (**Figure 4.3g**). Notably, CD8_Tex cells in the combination setting exhibited elevated OXPHOS scores (**Figure 4.3d, e**). These systemic profiles are consistent with previously reported intratumoral CD8⁺CXCL13⁺ exhausted T cells (7), suggesting a coordinated peripheral and tissue immune engagement.

4.4.2.4 CD8 T cell subsets: terminal differentiation and dysregulated cytotoxic programs in non-responders

In chemotherapy non-responders, treatment induced expansion of CD8_naïve + CD8_naïve_248⁺, characterized by high naïve scores (**Figure 4.3c-e**) and enrichment in translational pathways (**Figure 4.3g**). Additionally, effector and terminal subsets including CD8_eff, activated CD8_eff, proliferating CD8_eff + ISG, and CD8_Tex were expanded, reflecting a skew toward cytotoxic and fully differentiated exhausted phenotypes (**Figure 4.3c**). These subsets showed strong enrichment for cytotoxicity, chemokine receptor, OXPHOS, and glycolysis scores, while proliferating CD8_eff + ISG increased IFN response and OXPHOS scores (**Figure 4.3d, e**). Differential analysis showed enrichment of translational machinery associated with Activated CD8_eff and CD8_Tex (**Figure 4.3g**). In the combination therapy group, non-responders exhibited a distinct pattern of immune remodeling compared to chemotherapy. In pre-treatment, CD8_naïve + CD8_naïve_248⁺ subsets were elevated and exhibited high energy scores (**Figure 4.3c-e**) and enrichment in translational pathways (**Figure 4.3g**). Additionally, activated CD8_eff subsets showed high activation, chemokine receptors, IFN response scores, suggesting dampened effector function prior to treatment (**Figure 4.3c-e**). Post-treatment increases in CD8_naïve cells with high naïve scores, CD8_eff subsets enriched for activation, cytotoxicity, cytokine and chemokine receptors, and adhesion scores, and CD8_Tex cytotoxicity and adhesion scores (**Figure 4.3c-e**). Differential analysis confirmed strong enrichment of cytolysis-associated

genes in effector subsets (**Figure 4g**). These results indicated ineffective or dysregulated T cell activation in the combination non-responder group.

Pseudo time analysis of CD8⁺ T cells revealed two main trajectories: (1) from naïve through activated effector states toward transitional populations (Tex-prog, CD8_eff+ISG), and (2) toward fully differentiated exhausted (Tex) and cytotoxic effector subsets (**Figure 4.3f**). In the chemotherapy group, cells were primarily enriched in the transitional branch, whereas combination therapy cells predominantly occupied the terminally differentiated branch, reflecting more complete effector maturation under combination therapy. Notably, the apparent differences between responders and non-responders within each treatment group were largely explained by temporal dynamics, with responders enriched in these states at pre-treatment and non-responders at post-treatment.

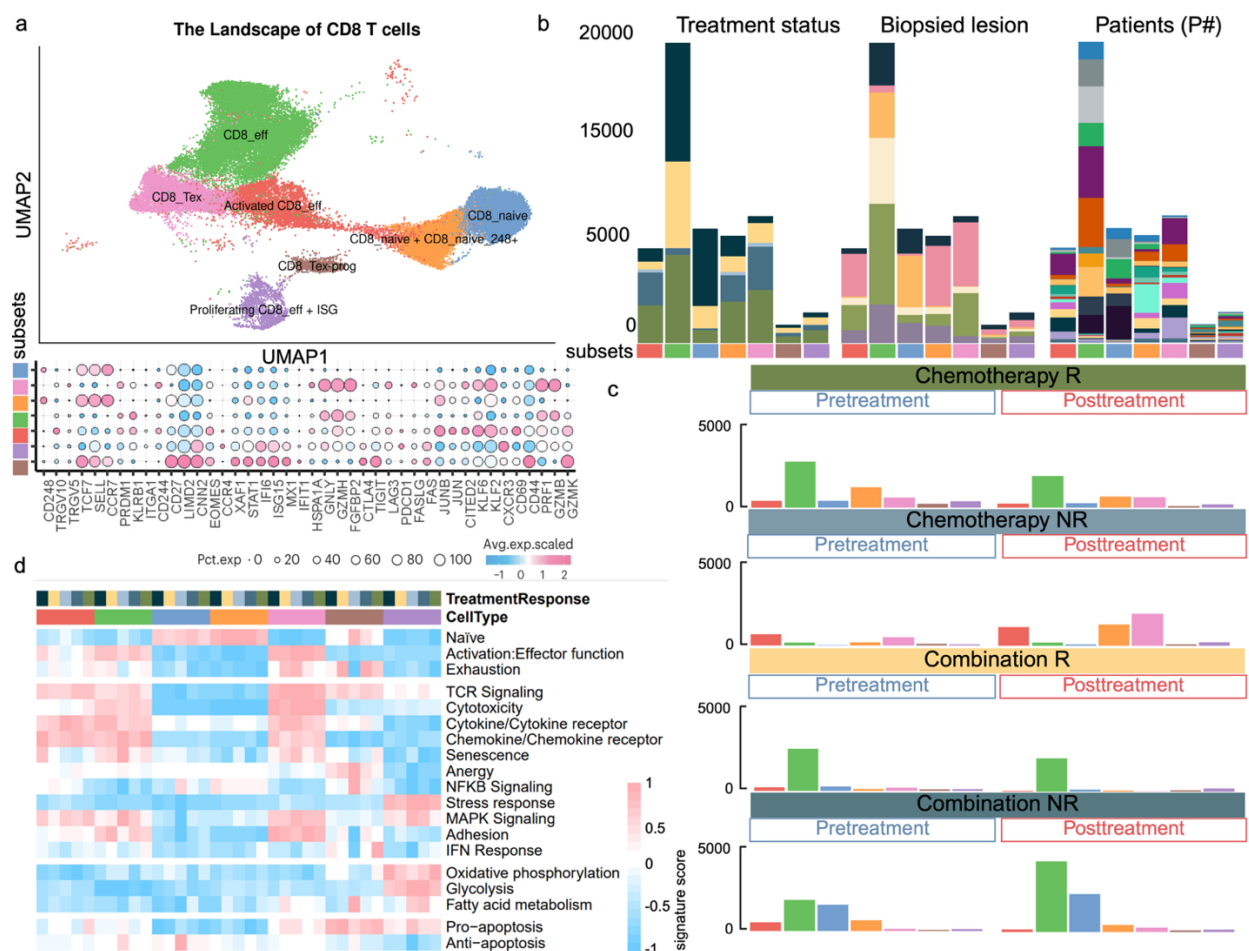


Figure 4.5: Comprehensive landscape of CD8⁺ T cell subsets in blood samples and their distribution across conditions.

a, UMAP representation of CD8⁺ T cell sub-clusters in blood, colored by distinct functional states (top). Bubble plot showing representative marker gene expression across clusters (bottom). Additional marker genes are listed in **Supplementary Table 3.2**. **b**, Stacked bar plots of CD8⁺ T cell sub-clusters across treatment and response groups, biopsied lesions, and individual patients, using color annotations consistent with **Figure 3.1b**. **c**, Bar plots comparing the composition of CD8⁺ T cell sub-clusters in blood before and after treatment across multiple samples, colored by distinct functional states. **d**, Heat map illustrating the expression of 19 curated gene signatures across CD8⁺ T cell clusters.

treatment PPI network analysis of these populations revealed enrichment of ribosomal and translational regulators, stress proteins, and BCR signaling components (*CD79A/B*, *PTPN6*), indicating a transcriptionally and translationally engaged B cell state in chemotherapy responders (**Figure 4.4e**). Notably, while these results are derived from peripheral blood, they align with prior findings demonstrating that B cells within the tumor microenvironment were predictive of response to chemotherapy: decreasing in paclitaxel-treated responders but increasing in those treated with nab-paclitaxel plus atezolizumab (8). Together, these data suggest that B cell activation signatures in circulation and tissue are associated with treatment responsiveness. In combination therapy responder, pre-treatment B cell populations displayed distinct dynamics compared to chemotherapy. CD20⁺ B cells were enriched for translational machinery, B cell activation, and antigen processing, while CD27⁺ memory B cells were enriched for cytoskeleton remodeling and antigen processing, and CD19⁺_AP1 were elevated in pre-treatment but lacked significant mechanism enrichment, suggesting differential B cell dynamics compared to chemotherapy (**Figure 4.4d**).

4.4.2.6 B cells: stress-regulated and dysregulated programs in non-responders

In chemotherapy non-responders, treatment induced expansion of B cell populations, including CD20⁺ B cells, CD19⁺_AP1, and CD27⁺ memory B cells, which were enriched for translational machinery (**Figure 4.4c, d**). PPI network analysis highlighted ribosomal and translational regulators, stress- and apoptosis-related proteins, and interferon-inducible genes. Immune checkpoint and regulatory-associated genes, including *CD69*, *CXCR4*, *JUN*, *IRF8*, and *ZBTB16*, were also notable, suggesting a multifaceted stress-regulatory B cell response (**Figure 4.4e**). In combination therapy non-responders, pre-treatment frequencies of B cell subsets including CD20⁺ cells enriched for translational programs, CD19⁺_AP1, and CD27⁺ memory B

cells were elevated. Among these, only CD19⁺_AP1 showed a higher frequency relative to post-treatment, indicating selective persistence of specific B cell populations (**Figure 4.4c, d**).

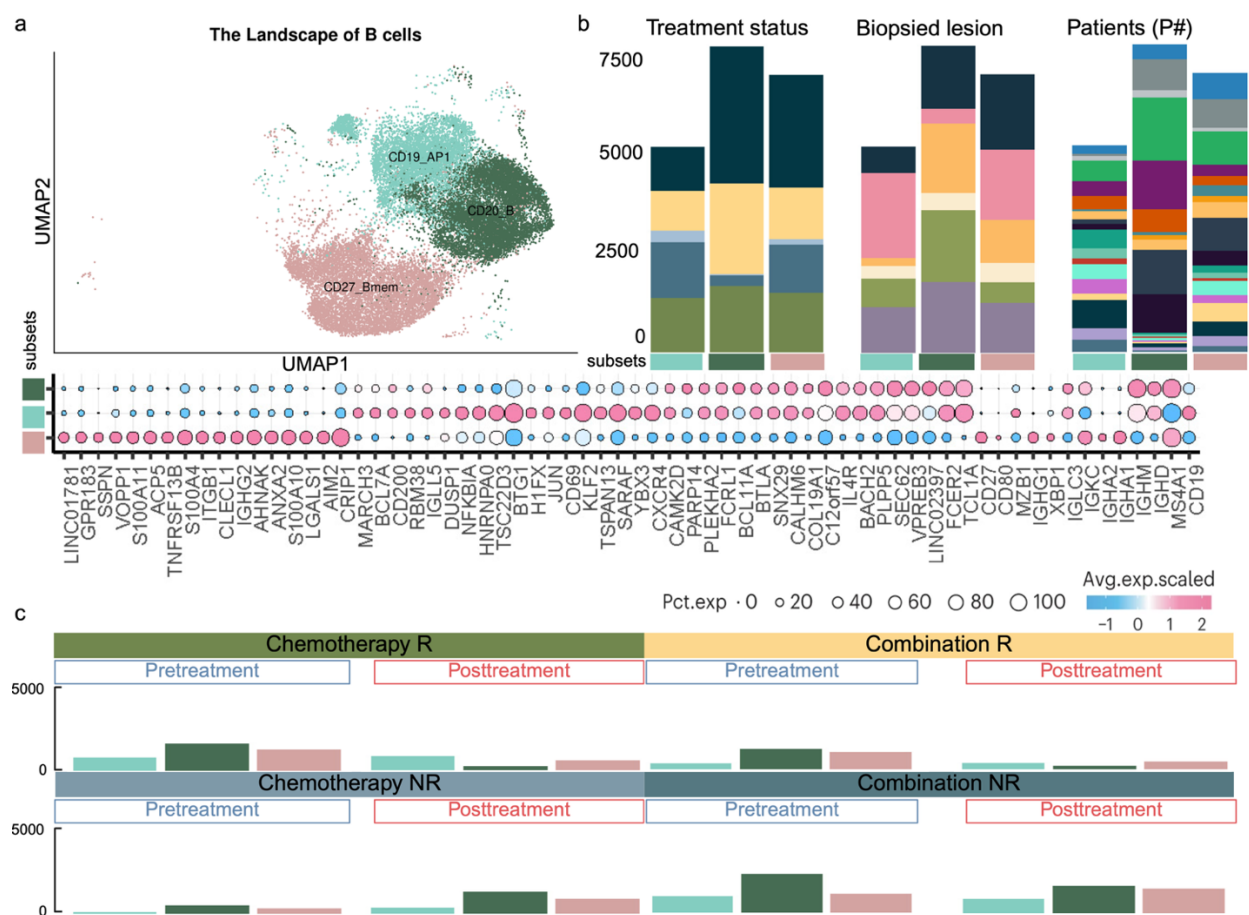


Figure 4.7: Comprehensive landscape of B cell subsets in blood samples and their distribution across conditions.

a, UMAP representation of B cell sub-clusters in blood, colored by distinct functional states. The bubble plot shows representative marker gene expression across clusters. Additional marker genes are listed in **Supplementary Table 4.2**. **b**, Stacked bar plots of B cell sub-clusters across treatment and response groups, biopsied lesions, and individual patients, using color annotations consistent with **Figure 4.1b**. **c**, Bar plots comparing the composition of B cell sub-clusters in blood before and after treatment across multiple samples, colored by functional states.

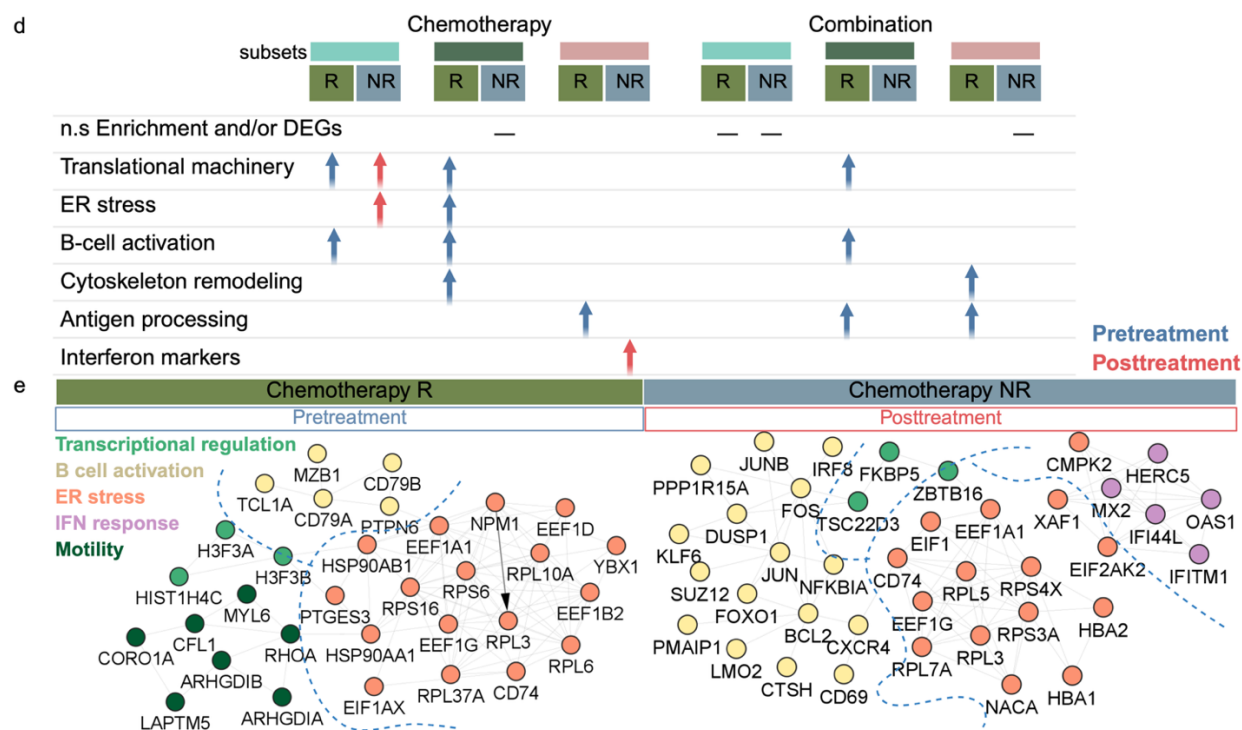


Figure 4.8: Comprehensive landscape of B cell subsets in blood samples and their distribution across conditions, Continued.

d, Table summarizing enriched pathways or functional signatures associated with B cell sub-clusters, stratified by condition. Blue: enriched in pre-treatment; Red: enriched in post-treatment.

e, PPI networks highlight distinct mechanisms enriched in the Chemo-R pre-treatment (right) and Chemo-NR post-treatment (left) groups. This Figure illustrates the heterogeneity of B cells in peripheral blood, highlighting patient-specific variation and dynamic changes in sub-cluster composition and functional states in response to treatment. Enrichment and network analyses underscore key signaling pathways associated with B cell activity across clinical conditions.

3.4.2.7 Innate lymphoid cells: translational enrichment in responders versus exhausted phenotypes in non-responders

NK cells provide rapid innate cytotoxic responses against tumor cells and contribute to immune activation through inflammatory cytokine production. Of the thirteen NK cell clusters identified (see Methods, **Supplementary Table 4.2**), three functional clusters (NK1, NK3, NK8) that were representative across patients were selected for detailed analysis (**Figure 4.5a, b**). In chemotherapy responders, NK1 cells exhibited elevated translational activity and cytotoxic markers pre-treatment, suggesting an innate readiness that parallels CD8⁺ T cell

activation (**Figure 4.5c, d**). This coordinated priming reflects classical NK cell function, part of the broader family of group 1 innate lymphoid cells (ILC1), which are known to provide rapid effector responses and support adaptive immunity (13). In contrast, combination therapy responders showed enhanced innate engagement following treatment, with expansion of NK1 cells with elevated translational and inflammatory signaling, indicative of heightened NK cell activation and functional differentiation (**Figure 4.5c, d**).

In chemotherapy non-responders, treatment induced an increase in NK1 cells enriched for translational machinery, inflammatory response, and cytotoxicity/exhaustion markers, while NK3 subsets were enriched for translational activity (**Figure 4.5c, d**). Collectively, these patterns suggest that chemotherapy in non-responders promotes immune subset expansion, but skewed toward a dysregulated, pro-inflammatory, and exhausted phenotype lacking coordinated effector function (14). Similarly, in combination therapy non-responders, pre-treatment NK3 cells exhibited cytoplasmic translation enrichment and inhibitory markers (*KIR2DL3*, *KLRG1*), consistent with an exhausted phenotype (**Figure 4.5c, d**). Following treatment, NK1 cells showed increased ER stress, inflammatory response, and cytotoxic/exhaustion signatures, paralleling the chemotherapy non-responder profile (**Figure 4.5c-e**).

These findings highlight distinct patterns of immune dysregulation between treatment groups. In chemotherapy non-responders, CD8⁺ T cell diversity expanded but lacked coordinated functional engagement with an elevated exhaustion phenotype. In contrast, combination-treated non-responders exhibited more defined CD8⁺ activation but were marked by CD4⁺ stress responses, suggesting dysfunctional immunity. This aligns with observations of a pan-cancer CD4⁺ T cell stress–response state linked to immunotherapy resistance (15). Notably, translationally

active CD20⁺ and CD27⁺ B cell subsets remained elevated post-treatment, but lacked functional enrichment, suggesting limited contribution to effective immunity. Both treatments showed elevated exhaustion levels.

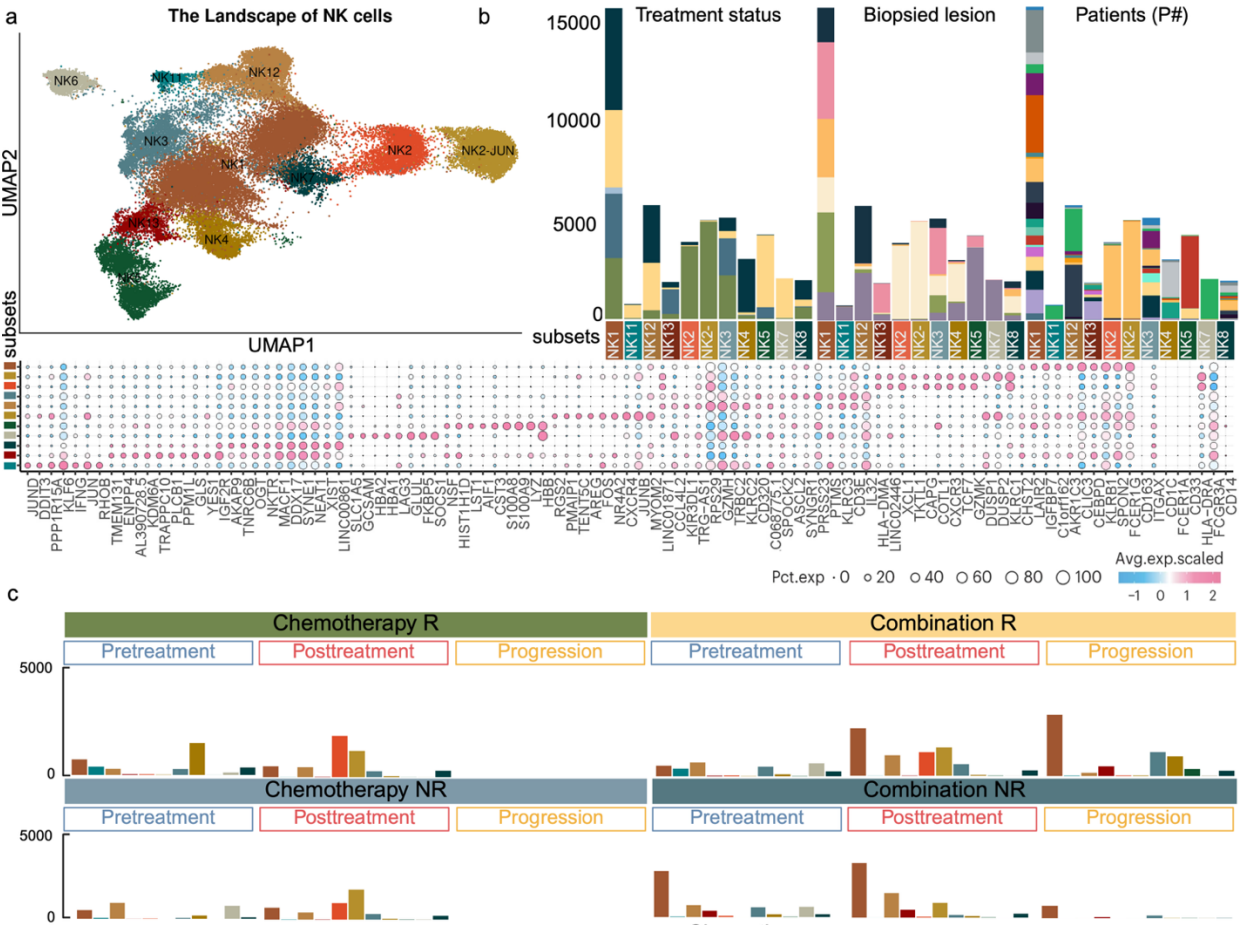


Figure 4.9: Landscape and dynamics of NK cell subsets in blood across treatment conditions. **a**, UMAP of NK cell sub-clusters in blood, colored by functional states. Bubble plot shows representative marker gene expression across clusters. Additional marker genes are listed in **Supplementary Table 3.2**. **b**, Stacked bar plots showing NK sub-cluster distributions across treatment groups, response status, biopsied lesions, and individual patients, using color annotations consistent with **Figure 3.1b**. **c**, Bar plots comparing NK sub-cluster compositions before and after treatment.

4.4.3.1 Divergent Timing of Myeloid Immune Programs in Responders to Chemotherapy and Combination Therapy

Myeloid populations coordinate antigen presentation, inflammatory signaling, and tissue remodeling, but can also acquire immunosuppressive and pro-tumorigenic phenotypes. Of the nine myeloid clusters identified (see Methods, **Supplementary Table 4.2**), three functional clusters (classical monocyte (CM), non-classical monocyte (NCM), dendritic cells (DC)) that were representative across patients were selected for detailed analysis (**Figure 4.6a, b**). In the chemotherapy group, pre-treatment samples from responders showed increased frequencies of CMs enriched for translational machinery, and DCs enriched for translational machinery, ER stress, OXPHOS, and interferon signaling. This transcriptional profile is consistent with a metabolically active, antigen-presenting phenotype (**Figure 4.6c, d**). This suggests that pre-existing immune activation, particularly metabolically active and antigen-presenting myeloid cells, is necessary for an effective response to chemotherapy (6). Together with adaptive immune priming, this suggests an interplay between the adaptive and innate immunity, priming the microenvironment for tumor clearance. However, no significant expansion of myeloid cells was observed following chemotherapy treatment, similar to the lack of expansion observed in lymphoid cell types. In the combination therapy group, baseline samples from responders showed increased DC infiltration, with the transcriptional landscape showing enrichment for translational machinery, stress, and inflammatory response (**Figure 4.6c, d**). Following treatment, responders exhibited an expansion of CM enriched for chromatin remodeling, suggesting a memory-like phenotype with a pro-inflammatory transcriptional program (**Figure 4.6c, d**). NCMs were enriched for cytotoxic, interferon-responsive, and migratory phenotype, including *ITGA4*, *TLN1*, *MSN*, *CORO1B* markers (**Figure 4.6c, d**). These findings in parallel with reduced adaptive immunity are consistent

with robust immune remodeling following anti-PD-L1 therapy.

4.4.3.2 Myeloid compartment in non-responders exhibits inflammatory skewing and immune dysfunction

Pre-treatment, non-responder samples across both treatment groups showed a lack of higher myeloid presence than post-treatment. Following chemotherapy, non-responders exhibited an increased frequency of CMs enriched for cytoplasmic translation, inflammatory signaling, and wound-healing pathways (*PF4*, *PF4VI*, *SPARC*, *TGFB1*), indicative of a pro-tumorigenic and potentially immunosuppressive state. NCMs showed upregulation of inflammatory and type I IFN response programs, while DCs exhibited elevated cytoplasmic translation and OXPHOS, reflecting heightened metabolic activity (**Figure 4.6c, d**). Post-treatment, non-responders displayed a myeloid compartment dominated by metabolically active, pro-tumorigenic programs, in line with evidence that tumor metabolic reprogramming can induce suppressive myeloid phenotypes (16,17). Concurrently, the lymphoid compartment exhibited increased diversity, with features suggestive of T cell dysfunction/exhaustion. This pattern is consistent with delayed and ineffective immune activation, potentially reflecting tumor-driven immune dysregulation. A similar escape mechanism has been described in the context of tumor-mediated disruption of antigen presentation of T cell exhaustion as a mode of cancer immune evasion (18). In the combination group, post-treatment CMs upregulated cytoplasmic translation, OXPHOS, type I IFN signaling, and immunoregulatory genes (*SOCS3*, *BCL3*), suggesting activation skewed towards dysregulation rather than effective antitumor immunity (19) (**Figure 4.6c, d**). NCMs showed elevated cytoplasmic translation but lacked hallmark inflammatory or cytotoxic gene expression (**Figure 4.6c, d**), indicating a metabolically primed state without clear evidence of functional activation. Such metabolic reprogramming without effector function is a recognized

pattern in tumor-associated myeloid cells, where translation-driven activation contributes to suppression rather than immune effector activity (20). In parallel, the lymphoid compartment exhibited an increase in terminally exhausted T cells, further reinforcing a state of ineffective immune activation in non-responders.

4.4.4 Tissue CD4⁺/CD8⁺ T cells dynamics diverge in non-responders

Of the seven CD4⁺ and fourteen CD8⁺ T cell clusters identified (see Methods, Supplementary Table S2), five CD4⁺ and four CD8⁺ functional clusters that were representative across patients were selected for detailed analysis (**Figure 4.7a, b, d, e**). Pre-treatment, both chemotherapy and combination therapy responders showed enrichment of adaptive immune subsets, including CD4_naive, Tfh + Tex-prog, T_{reg}, CD8_cytotoxic, CD8_Tex + MKI67, and CD8_naive-like + ISG populations, indicating a pre-existing, antigen-experienced T cell compartment poised for activation (**Figure 4.7c, f**). These observations align with Zhang *et al.* (7,8), that identified subsets: naïve T cells, Tfh-CXCR5, and Tex-MKI67 subsets, enriched pre-treatment in PTX responders but appeared post-treatment in Nab-PTX responders. Following treatment, both groups showed increased activated naïve-like CD4⁺ T cells, while combination therapy showed increased activated CD8-eff and chemotherapy showed CD4-CM (**Figure 4.7c, f**). Note that Zhang *et al.* (7,8) showed that activated naïve-like CD4 was associated with combination post-treatment in PTX, while pre-treatment in Nab-PTX. In non-responders, pre-treatment profiles diverged by treatment. Combination therapy non-responders showed enrichment of CD4_naive, CD4_CM, Treg, activated CD8_eff, and CD8_cytotoxic subsets (**Figure 4.7c, f**). Chemotherapy non-responders instead exhibited broad T cell expansion post-treatment (**Figure 4.7c, f**), consistent with Zhang *et al.*'s (7,8) observation of post-chemotherapy T_{reg} increases.

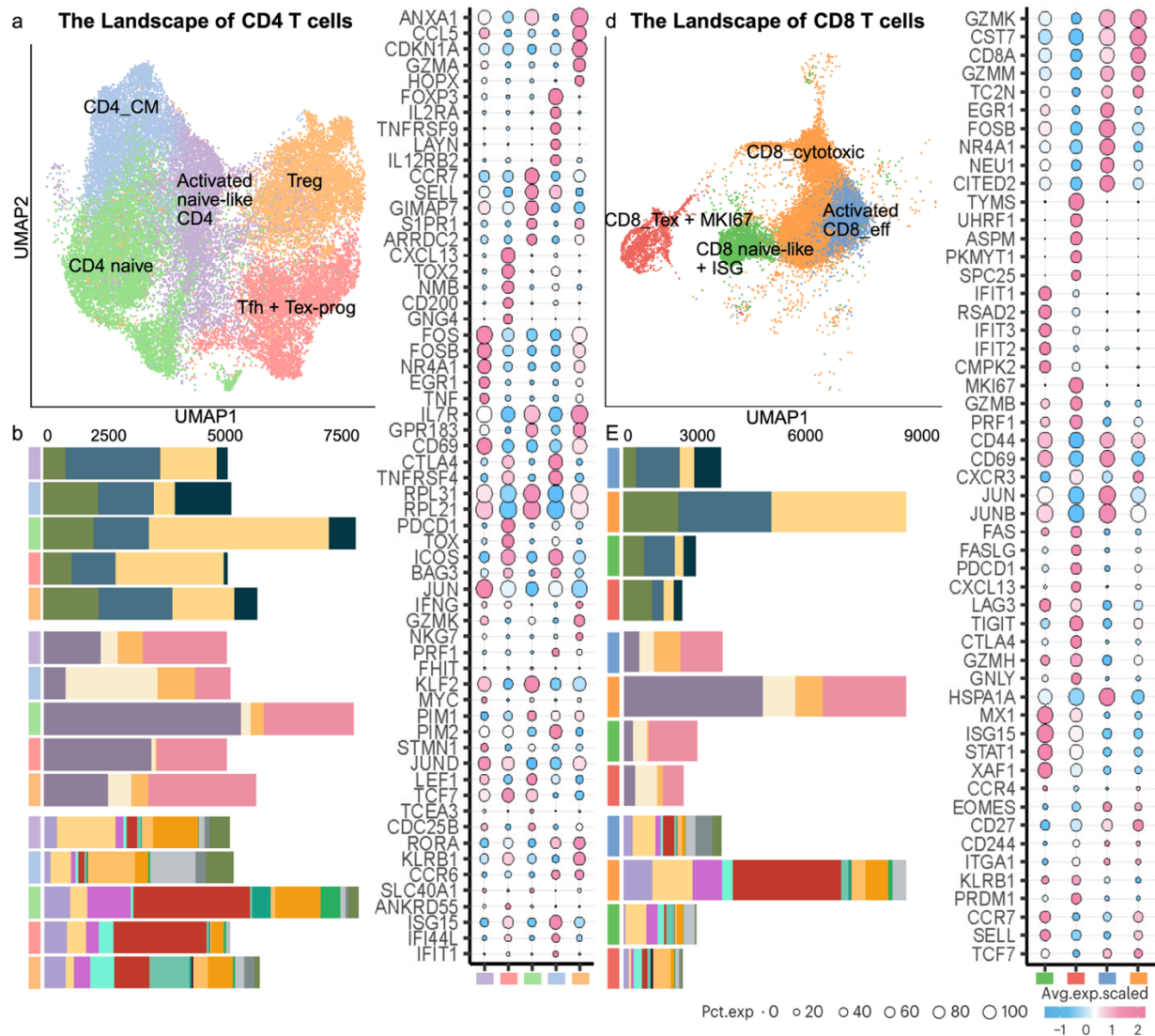


Figure 4.12: Comprehensive landscape of CD4+ and CD8+ T cell subsets in tissue samples and their distribution across conditions.

a-d, UMAP representation of CD4+ and CD8+ T cell sub-clusters in tissue, respectively, colored by distinct functional states. The bubble plot shows representative marker gene expression; additional markers are listed in **Supplementary Table 4.2**. **b-e**, Stacked bar plots showing CD4+ and CD8+ sub-cluster distributions, respectively, across treatment groups, response status, biopsied lesions, and individual patients, using color annotations consistent with **Figure 4.1B**.

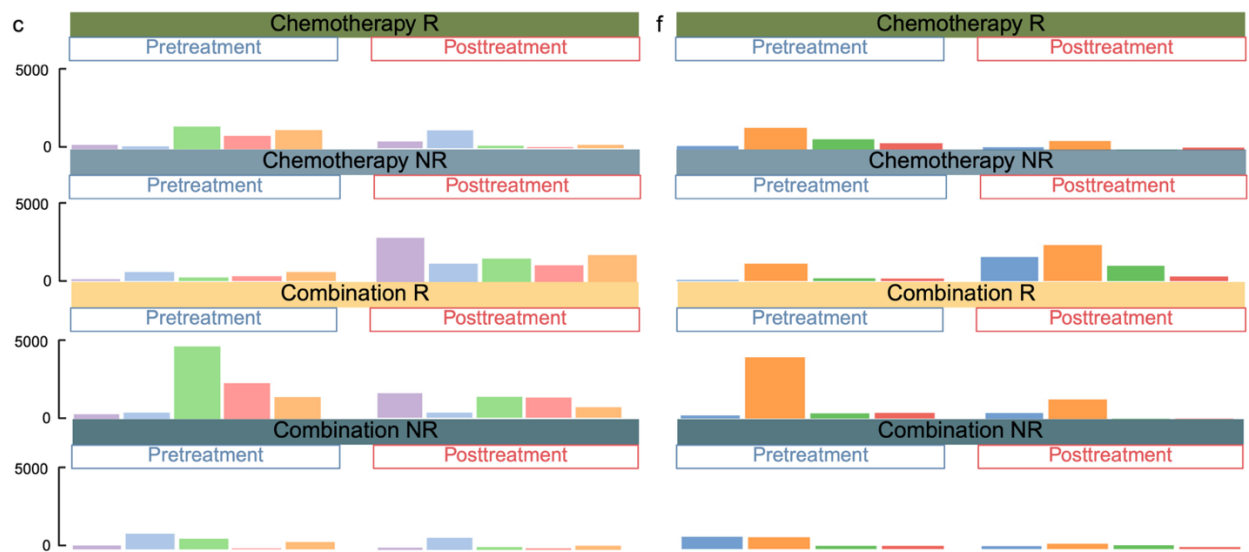


Figure 4.13: Comprehensive landscape of CD4⁺ and CD8⁺ T cell subsets in tissue samples and their distribution across conditions, Continued.

c-f, Bar plots comparing T cell sub-cluster compositions before and after treatment.

4.4.5 Coordinated systemic and tumor T cell programs distinguish chemotherapy and combination therapy responders

In chemotherapy responders, tissue PPI network analysis revealed enrichment of IFN-driven signaling, antigen processing, and strong exhaustion markers (**Figure 4.8a**). Blood profiling showed upregulation of T cell priming/activation states and early activation of CD8⁺ T cells (**Figure 4.9a**). This concordance suggests a sequential immune mobilization model, in which activated T cells detected in the blood are associated with intratumoral cytotoxic and exhaustion programs. In combination therapy responders, tissue networks highlighted ribosomal proteins, metabolic enzymes, and TCR activation markers, consistent with sustained antigen-driven effector function supported by high biosynthetic and metabolic capacity (**Figure 4.8b**). Pre-treatment blood profiles mirrored this state, showing more fully differentiated and engaged adaptive immunity with active TCR signaling and cytotoxicity responses (**Figure 4.9b**), indicating parallel activation of systemic and intratumoral T cell compartments at treatment onset. Notably, in combination therapy, post-treatment blood shifted toward expansion of innate immune subsets

with JAK-STAT, TNF, and IFN pathway activation. These observations are consistent with systemic innate immune re-engagement that may follow effective adaptive tumor clearance, even if not captured in matched tissue samples due to lack of representative innate immunity across patients (**Figure 4.9b**).

In chemotherapy non-responders, post-treatment showed tissue enrichment for antigen-driven activation and effector differentiation, but also stress-response and immunoregulatory programs (*TNF*, *TNFSF10*, *IL6ST*, *NFKB1A*, *SOCS3*, *ZC3HAV1*) (**Figure 4.8c**). Blood analysis revealed parallel induction of effector, cytotoxicity, and exhaustion markers, reinforcing the interpretation that transcriptionally primed T cells remain functionally constrained (**Figure 4.9c**). In tissue, pre-treatment combination therapy non-responders displayed PPI networks combining antigen-driven activation, cytotoxic function, and metabolic adaptation, but these were accompanied by immunoregulatory and stress-response signatures (*TNFAIP3*, *BIRC3*, *TGFB1*, *NR3C1*, *NFKB1A*) and chemokine/cytokine modulators (*CCR5*, *CXCR4*, *CCL4*) (**Figure 4.8d**), suggesting functional suppression within an immunosuppressive microenvironment (21). This suppression was mirrored in blood, where reduced effector signatures were observed (**Figure 4**). Across both therapies, non-responders demonstrated post-treatment expansion of innate immune subsets in blood, characterized by oxidative stress, metabolic dysfunction, and immunosuppressive signaling. PPI analysis indicated upregulation of antioxidant and inflammatory regulators (*SOD2*, *PRDX*, *HMOX1*) alongside *NFKB1A* and *PTGS2*, with stronger induction in combination non-responders, suggesting unresolved inflammation and dysfunctional innate immune re-engagement (**Figure 4.9c, d**).

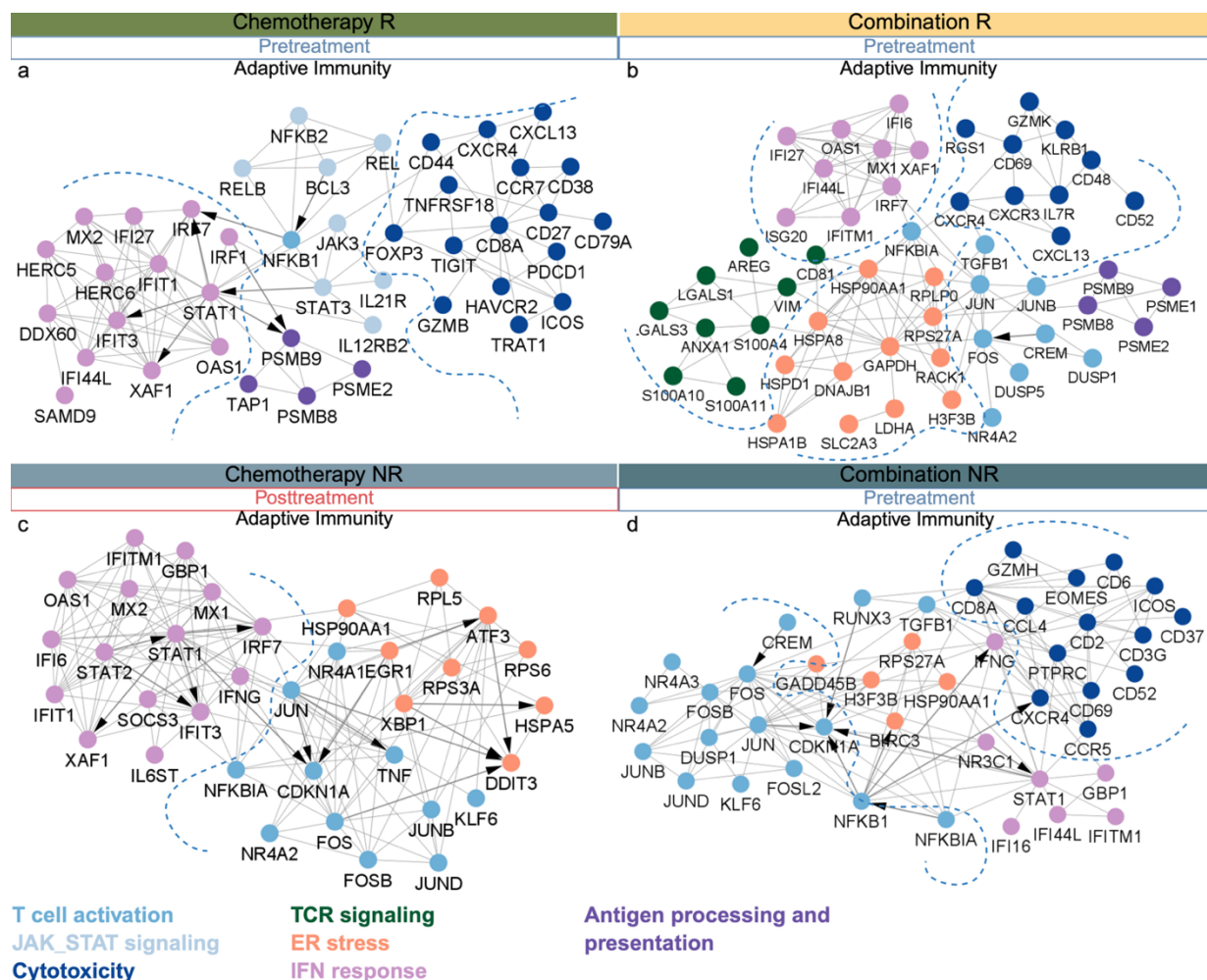


Figure 4.14: Protein-protein interaction networks of treatment- and response-specific mechanisms in tissue.

PPI networks were constructed from differentially expressed genes identified in each treatment and response group when comparing post- versus pre-treatment samples (see Methods). **a**, Chemotherapy responders (pre-treatment): modules enriched in T cell activation, antigen processing, interferon signaling, and cytotoxicity. **b**, Combination responders (pre-treatment): modules enriched in interferon signaling, T cell activation, ER stress, and TCR signaling. **c**, Chemotherapy non-responders (post-treatment): modules enriched in T cell activation, ER stress, and interferon response. **d**, Combination non-responders (pre-treatment): modules enriched in T cell activation, interferon signaling, ER stress, and cytotoxicity. Node color denotes functional modules (see legend) and arrows indicate transcription-target relationships. Together, these networks reveal condition-specific immune regulatory circuits and highlight potential therapeutic targets.

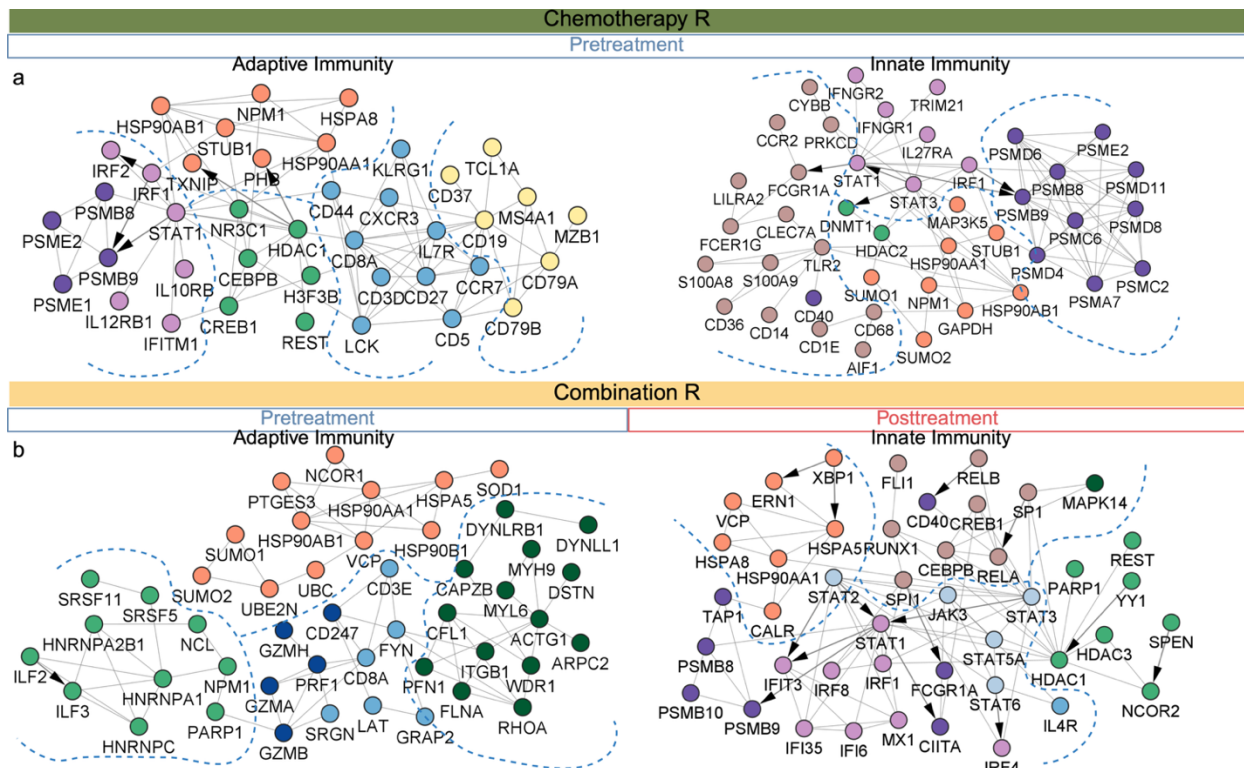


Figure 4.15: Protein-protein interaction networks of treatment- and response-specific mechanisms in blood.

PPI networks were constructed from differentially expressed genes identified in each treatment and response group when comparing post- versus pre-treatment samples (see Methods). **a**, Chemotherapy responders (pre-treatment): adaptive immunity (left panel) modules enriched in T cell activation, antigen processing, interferon signaling, and transcriptional regulation. Innate immunity (right panel) modules enriched in myeloid activation, interferon signaling, ER stress and antigen processing. **b**, Combination responders (pre-treatment): adaptive (left panel) modules enriched in T cell activation, cytotoxicity, ER stress, and transcriptional regulation. Combination responders (post-treatment): innate immunity (right panel) modules enriched in myeloid activation, interferon signaling, ER stress, and transcriptional regulation.

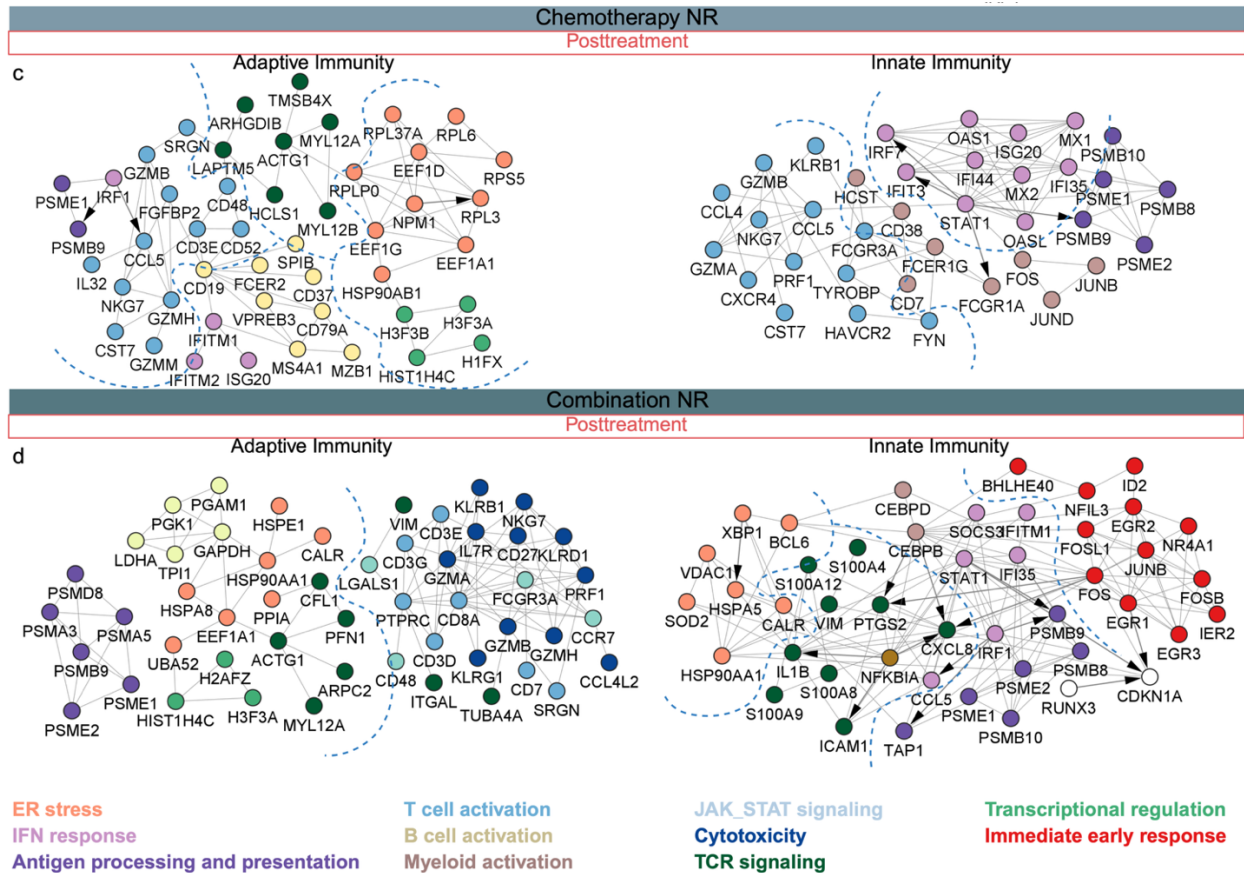


Figure 4.16: Protein-protein interaction networks of treatment- and response-specific mechanisms in blood, Continued.

PPI networks were constructed from differentially expressed genes identified in each treatment and response group when comparing post- versus pre-treatment samples (see Methods). **c**, Chemotherapy non-responders (post-treatment): adaptive immunity (left panel) modules enriched in T- and B-cell activation, TCR signaling, ER stress, and interferon response. Innate immunity (right panel), myeloid activation, JAK-STAT signaling, antigen processing, and interferon signaling. **d**, Combination non-responders (post-treatment): adaptive immunity (left panel) modules enriched in T cell activation, interferon signaling, ER stress, and cytotoxicity. Innate immunity (right panel) modules enriched in motility, antigen presentation, interferon signaling, ER stress and immediate early response. Node color denotes functional modules (see legend) and arrows indicate transcription–target relationships. Together, these networks reveal condition-specific immune regulatory circuits and highlight potential therapeutic targets.

4.4.6 Patient-specific immune dynamics

While group-level analyses revealed broad immune differences between responders and non-responders, patient-specific analyses uncovered nuanced immune adaptations shaped by metastatic site and therapeutic response. In chemotherapy responders, patients P020 (breast), P022 (chest wall metastasis), and P013 (liver metastasis) exhibited distinct myeloid and NK cell

dynamics (**Figure 4.5b and 4.6b**). In blood, pre-treatment increase in CM_IFIT cells was consistently observed across these patients, reflecting a shared inflammatory signature (**Figure 4.6c**). Interestingly, in P013, treatment was followed by an increase of NK2 and at progression, increased by NK2_JUN, both involved in processing and presenting antigens, suggesting a dynamic NK cell response (**Figure 4.5c-e**). Tissue profiling of P013 revealed a sequential immune trajectory. Pre-treatment pro-inflammatory monocytes transitioned into pro-inflammatory macrophages post-treatment, followed by increased NK1, NK2, TAM + ISG, macrophages, and macrophage + LYVE1 (Supplementary Figure S2A, C) (Supplementary Figure S3A, C). These populations displayed elevated expression of inflammatory and cytokine response genes, emphasizing the role of innate immunity in mediating chemotherapy response in metastatic disease (22) (**Supplementary Figure 4.2 and 4.3d**). Notably, pre-treatment adaptive immune engagement, characterized by primed CD4⁺ T cells and diverse CD8⁺ effector subsets, preceded the expansion of innate immune populations, suggesting a coordinated interplay in which early adaptive activation primes subsequent innate immune responses during treatment and disease progression.

In chemotherapy non-responders, blood samples from patients P018, P023, P024, and P025 (all breast) showed a post-treatment increase in CM_IFIT, aligning with the inflammatory profile observed at the group level (**Figure 4.6c**). In combination therapy responders, patients P019 and P007 (LN metastases) exhibited distinct immune dynamics. In blood, both patients showed post-treatment increases in NK5 and NK11, subsets associated with NK activation and antigen processing, consistent with group-level observations and highlighting innate immune activation following combination therapy (**Figure 4.6c, d**). This aligns with group-level findings, reinforcing the role of innate lymphoid immune activation following combination therapy. At

progression, NK6, NK12, and macrophage populations were elevated, indicating ongoing immune adaptation (**Figure 4.6c**). In contrast, combination therapy non-responders, including P005 and P017 (both with liver metastases), showed distinct pre-treatment immune features. In PBMC analysis, NK4 emerged in a patient-specific manner, present pre-treatment in P005 and post-treatment in P017, suggesting differential NK cell adaptation (**Figure 4.5b, c**). Tissue profiling of P017 revealed pre-treatment enrichment of NK5, characterized by cytoplasmic translation and cytoskeleton organization (**Supplementary Figure 4.2b, c**), alongside CMs characterized with early immune activation, which were replaced post-treatment by macrophage + HS3ST2 expressing immunosuppressive markers (**Supplementary Figure 4.3b, c**). P005 showed NK3 cells enriched for cytoplasmic translation pre-treatment (**Supplementary Figure 4.2b, c**). Additionally, P016 (chest wall metastasis) showed increased NK12 at baseline, with NK7 expressing cytoskeleton organization and checkpoint markers (**Supplementary Figure 4.2b, c**), and emergence of CM_AP1 emerging at progression in PBMC (**Figure 4.6b**). Tissue samples revealed early macrophage + CCL3 enrichment, suggesting a pre-existing myeloid-driven immune phenotype (**Supplementary Figure 4.3b, c**). Collectively, these findings suggest that innate immune subsets with immunosuppressive features are associated with poor response to combination therapy.

Together, these findings highlight the heterogeneity of immune responses across patients and therapeutic regimens. While adaptive immunity appears to play a critical role in initiating response, the dynamic engagement or suppression of innate immune subsets during treatment and progression may be key to sustaining or modulating therapeutic efficacy. These results underscore the importance of incorporating patient-specific immune profiling to refine stratification strategies and inform the development of individualized therapeutic approaches.

4.4.7 Extension to I-SPY2 clinical trial data

To contextualize these findings against publicly available datasets, we leveraged scores derived from the I-SPY2 clinical trial (11). Correlation analysis demonstrated that NK1, Treg, and activated/exhausted CD8 programs positively aligned with tumor chemokine, interferon, and T/B-cell modules, while proliferating ISG-CD8 programs showed consistent negative association, indicating concordance between systemic immune states and established tumor immune phenotypes (**Supplementary Figure 4.4a**). To maintain treatment consistency, analyses were restricted to the pretreatment paclitaxel arm. Univariate logistic regression showed directional but underpowered associations with pCR. In contrast, multivariable LASSO modeling identified a coordinated peripheral immune signature associated with response, negative contributions from Activated naïve-like CD4 and dendritic programs, and positive contributions from NK, naïve CD8, and Treg states, suggesting composite systemic immune readiness may distinguish responders (**Supplementary Figure 4.4b**). Permutation testing indicated that the multivariate model performed better than random expectation but did not reach statistical significance ($p = 0.128$), consistent with limited cohort size and correlated immune features.

4.5 Discussion

This study provides a systematic characterization of immune cell populations and their associated mechanisms, revealing shared and distinct immunological features underlying treatment response and resistance in advanced TNBC patients receiving chemotherapy or chemo-immunotherapy. Peripheral immune dynamics differed by treatment and response, revealing therapy-specific immunological states with predictive and mechanistic interpretability. Chemotherapy responders had higher pre-treatment immune cell counts than combination therapy responders, followed by post-treatment depletion consistent with known cytotoxic effects (41, 42).

In non-responders, combination therapy patients showed higher pre-treatment counts than chemo-only, but immune levels remained elevated post-treatment across groups. These findings are concordant with recent I-SPY2 peripheral blood reporting that therapeutic response is linked to systemic immune activation and cytotoxic remodeling in chemo-immunotherapy (12).

Across both treatment groups, our results highlight the importance of the dynamic coordination between adaptive and innate immune programs in modulating treatment outcomes. Enrichment of pathways including cytoplasmic translation, stress response, inflammatory signaling, cytoskeletal remodeling and migration, and metabolic programs including OXPHOS were observed in both responders and non-responders, underscoring their association with treatment outcomes. Post-transcriptional mechanisms, including the translation of mRNA, are critical contributors to rapid adaptive changes in the cellular proteome during immune responses (35). For instance, cytosolic protein translation is essential to increase glucose metabolism and degranulation capacity upon TCR activation, regulating the effector functions of CTLs and influencing the formation of memory cells (42, 43). Additionally, reprogramming cellular protein synthesis in innate immunity plays a pivotal role in inflammatory responses to pathogens and tumorigenesis by enhancing DNA-dependent cGAS activation, which drives the expression of ISGs (36).

Chronic antigen exposure can induce PD-L1 expression in tumor cells (likely via *IFNG* secreted from T cells) (37), and sustain PD-1 signaling in T cells, leading to an epigenetically imprinted exhaustion program (38–40). The presence of more differentiated T cells in combination therapy responders reflects this exhaustion phenotype and suggests higher tumor PD-L1 levels, potentially enhancing sensitivity to anti-PD-L1 therapy. This was supported by our observation that in responders tumors in combination were PD-L1⁺, while in non-responders they were PD-

L1⁻. In contrast, chemotherapy responders showed less differentiated and more plastic T cells pre-treatment, consistent with lower PD-L1 expression and increased chemotherapy responsiveness. Combination therapy induced increased innate immune engagement, potentially in response to increased antigenic debris from tumor cell death. Such innate activation is critical for sustaining T cell memory and effector function, linking antigen presentation and cytokine signaling to adaptive immune maintenance (41). Innate immune reactivation post combination therapy may serve as a surrogate marker of response. Overall, these results provide a system-level context for previous studies showing that pre-treatment immune cell diversity plays a crucial role in treatment outcomes (42–44).

In contrast, non-responders exhibited a myeloid compartment skewed toward an immunoregulatory, metabolically reprogrammed state, paralleled by increased dysfunctional/exhausted lymphoid populations across both treatments. Although canonical myeloid-derived suppressor cells (MDSC) markers such as *ARG1* were not prominently expressed, *IL1B* was present among DEGs, and coupled with upregulation of stress-adaptive and metabolic genes, this profile is consistent with an early-stage or atypical MDSC-like phenotype (45). This immunoregulatory shift may be associated with CD4⁺ T cell polarization toward suppressive (T_{reg}) or inflammatory (Th17) fates, as reported for metabolically reprogrammed human myeloid subsets (46). Such metabolically primed myeloid cells have been shown to suppress T- and NK cell activity or promote chronic inflammation, consistent with prior studies implicating monocyte and stromal metabolic reprogramming via cGAS–STING as a driver of therapy resistance (47). These findings highlight the plasticity of tumor-associated myeloid cells and the emerging importance of immunometabolism in shaping pro-tumoral functions (17,45).

Dysregulated NK cell responses emerged as a converging axis of resistance, with increased

expression of *DUSP1*, *CD81*, and cytotoxic stress markers. Prior studies in TNBC have linked persistent NK activation with poor prognosis, with inflammatory NK signatures being increasingly recognized as pro-metastatic in solid tumors (46,48–50). Together, these data support a model in which adaptive-innate immune interplay in the peripheral blood is a marker and correlate of therapeutic outcome. Responders characterized by coordinated activation of lymphoid and metabolically competent myeloid populations, whereas non-responders display myeloid immunoregulation and lymphoid dysfunction or exhaustion. These results further highlight the potential of peripheral immune profiling to guide response prediction, resistance monitoring, and immunomodulatory strategies in TNBC.

In conclusion, our single-cell profiling of paired blood and tumor samples reveals that systemic and intratumoral immune remodeling distinguishes therapeutic responders from non-responders in TNBC. Responders exhibit coordinated activation of adaptive and innate immune programs, whereas non-responders display metabolic reprogramming and immune dysfunction across compartments. The implications of our study include potential development of blood-based biomarkers associated with treatment response, informing patient-specific therapeutic strategies, and informing the rational design of more precise combination therapies to increase response to treatment in TNBC patients.

Although these analyses were performed in TNBC, the broader question of how spatially organized tumor ecosystems shape therapeutic response extends across solid malignancies. To evaluate whether the ecosystem-oriented framework developed in TNBC generalizes beyond breast cancer, we next applied spatial and cellular ecosystem modeling to hepatoblastoma, a biologically distinct pediatric malignancy characterized by marked architectural heterogeneity despite a relatively constrained genomic landscape.

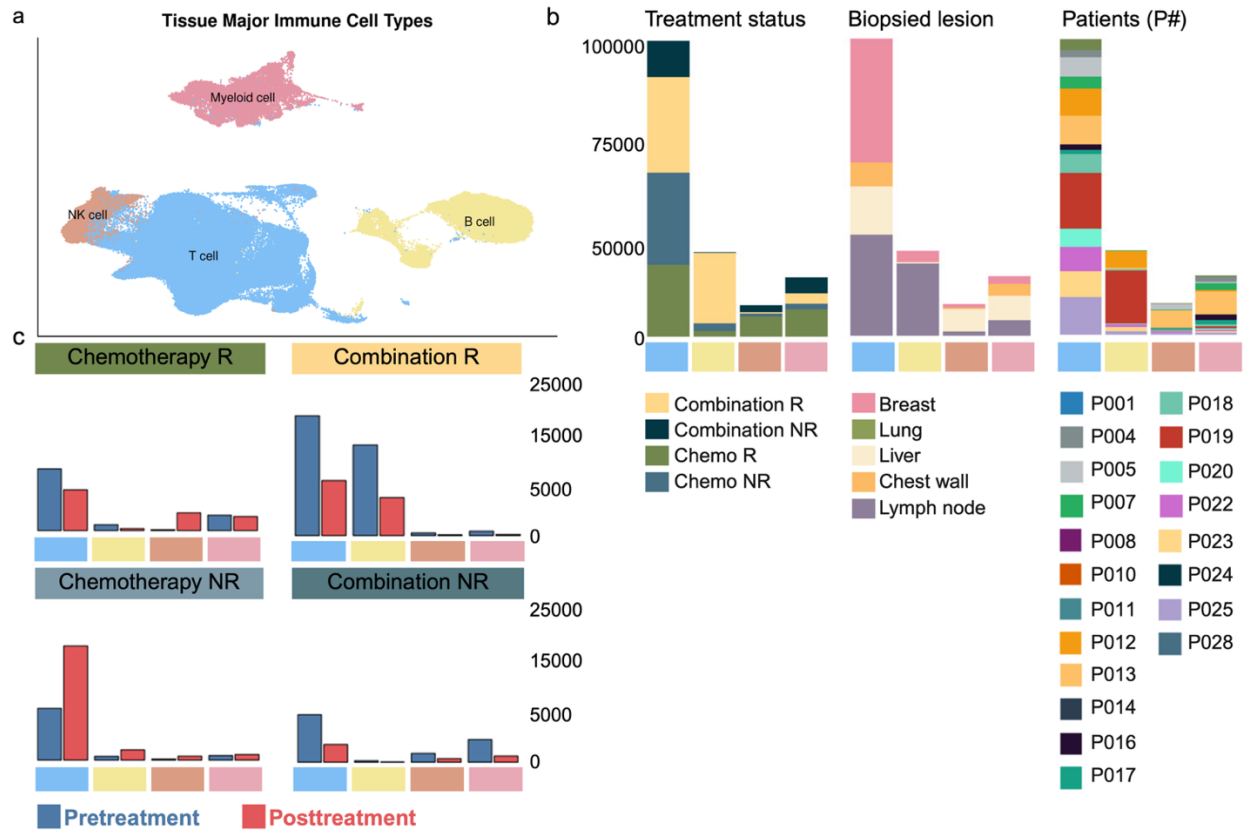
This Chapter is under review for publication and a version of it is in preprint at DOI: <https://doi.org/10.64898/2026.04.28.721416>. Please refer to the online manuscript for detailed Supplementary Materials.

4.6 Supplemental Materials

Supplementary Table 4.1: Patient demographics.

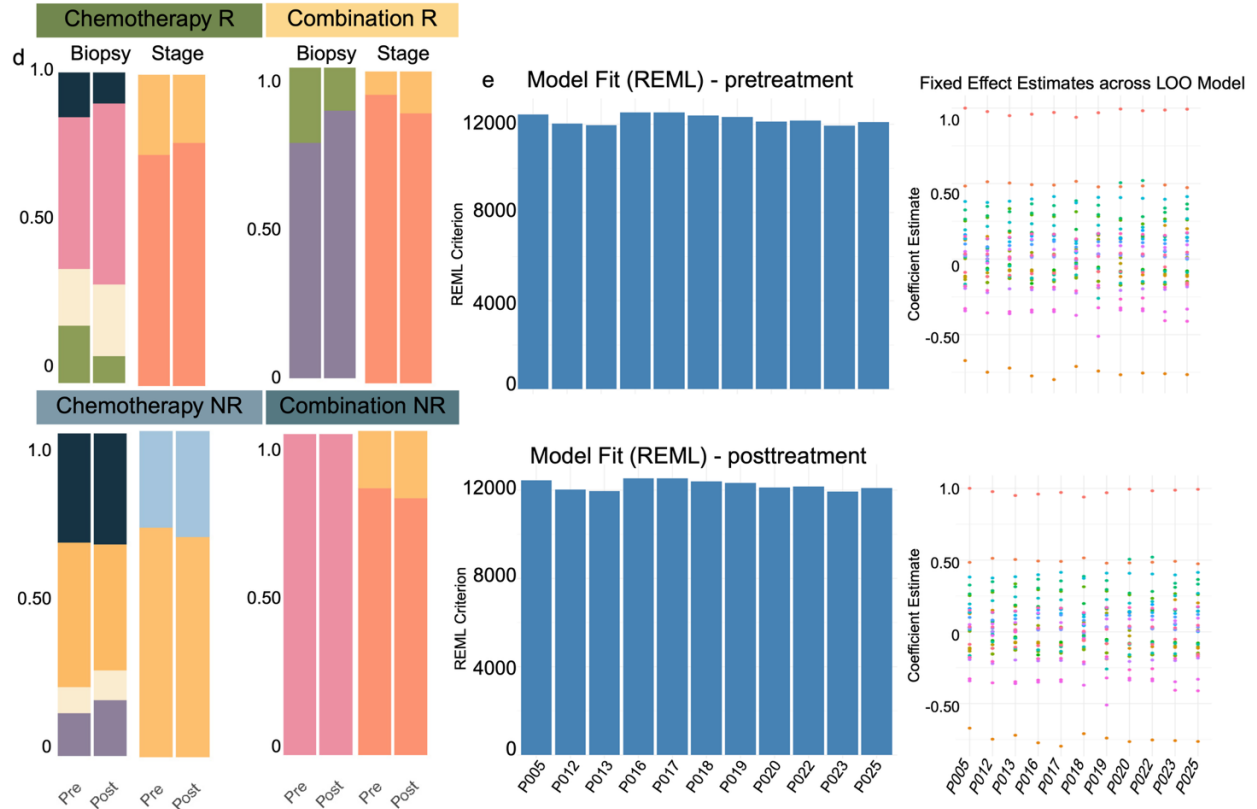
Treatment	Patient_ID	Biopsied lesion	Tumor			Blood			Clinical efficacy [#]
			Pre	Post	Prog	Pre	Post	Prog	
combination therapy	P019	Lymph Node	Y	Y		Y	Y		PR
	P010	Lung	Y	N		Y	Y	Y	PR
	P012	Lymph Node	Y	Y		Y	Y		PR
	P007	Lymph Node	Y	N	Y	Y	Y	Y	PR
	P017	Lymph Node	Y	Y		Y	Y		SD
	P001	-	N	N		Y	Y		SD
	P002	Chest Wall	Y	Y		Y	Y	Y	SD
	P014	-	N	N		Y	Y		SD
	P004	Chest Wall	Y	N		Y	Y	Y	SD
	P005	Liver	Y	Y		Y	Y		SD
	P016	Chest Wall	Y	Y		Y	Y	Y	SD
chemotherapy	P022	Breast	Y	Y		Y	Y		PR
	P011	-	N	N		Y	Y		PR
	P020	Breast	Y	Y		Y	Y		PR
	P008	Lung	N	N		Y	Y	Y	PR
	P013	Liver	Y	Y	Y	Y	Y	Y	PR
	P025	Breast	Y	Y		Y	Y		SD
	P018	Breast	Y	Y		Y	Y		SD
	P023	Breast	Y	Y		Y	Y		SD
	P024	Breast	N	N		Y	Y		PD
	P003	Chest Wall	N	Y		N	N		PD
	P028	-	N	N		Y	N		-

responder/regression = CR/PR | non-responder/progression = SD/PD



Supplementary Figure 4.1: Comparative immune landscape of tissue samples under different treatment conditions.

a, UMAP plots depict major immune cell types in tissue, highlighting T cells (blue), B cells (yellow), monocytes (pink), and NK cells (brown). **b**, Compositions of major immune cell types in treatment and response groups, biopsied lesions, and individual patients. **c**, Bar plots quantifying immune cell frequencies, illustrating treatment-associated shifts in immune cell distribution.

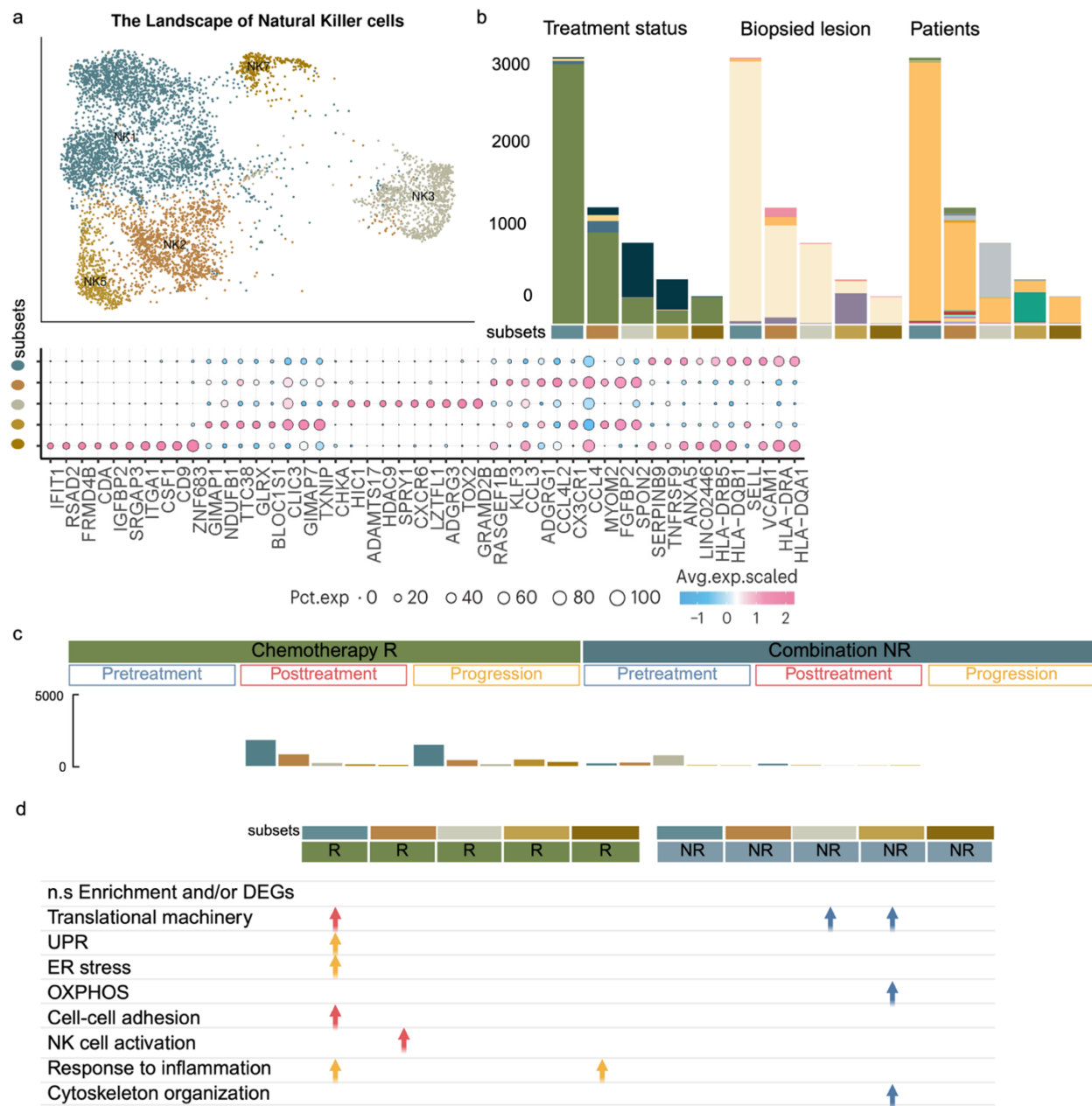


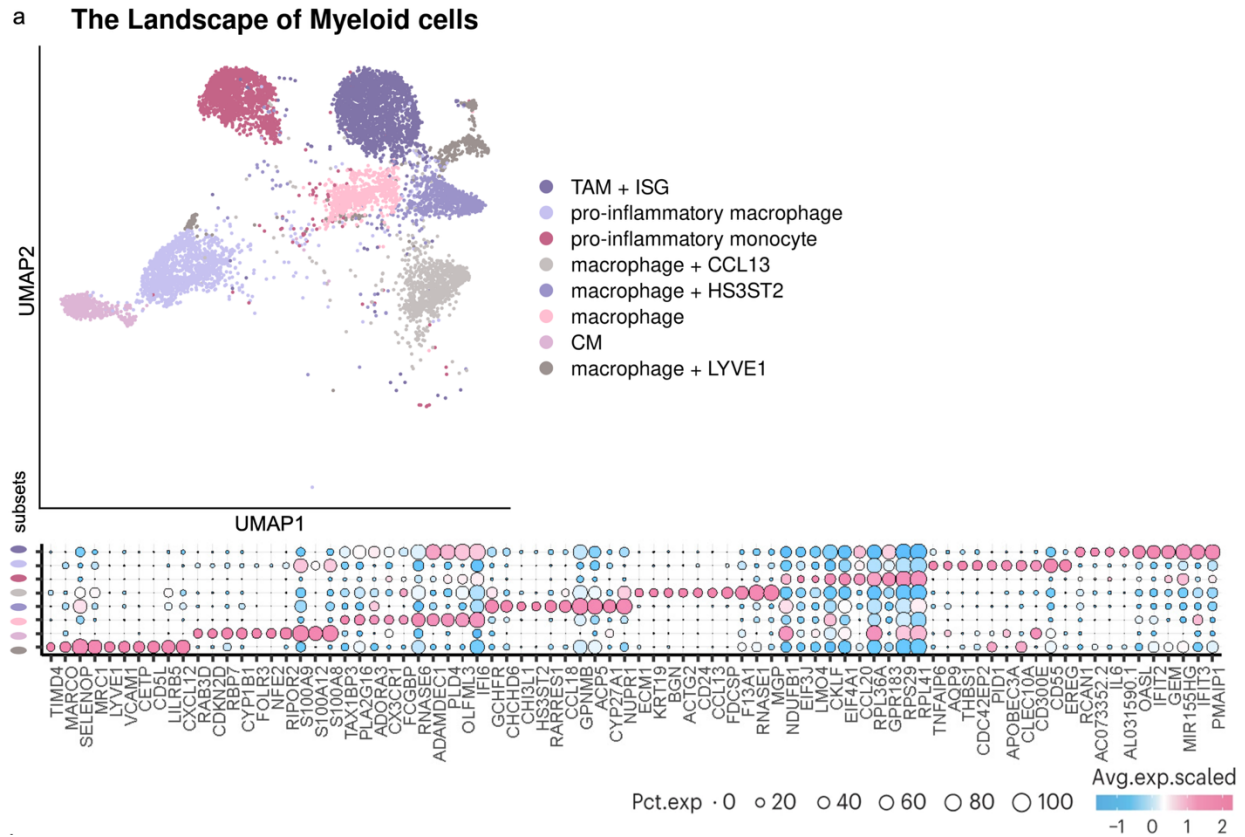
Supplementary Figure 4.2: Comparative immune landscape of tissue samples under different treatment conditions, Continued.

d, Relative distribution of biopsy and stage classifications across chemotherapy responders (Chemotherapy R), combination therapy responders (Combination R), chemotherapy non-responders (Chemotherapy NR), and combination therapy non-responders (Combination NR), shown separately for pre-treatment and post-treatment samples. **e**, Regression model sensitivity analysis showing the impact of removing each patient on model fit (REML) (top left, pretreatment; bottom left, posttreatment) and stability of fixed-effect estimates (β) across leave-one-out iterations (top right, pretreatment; bottom right, posttreatment).

Supplementary Figure 4.3: Comprehensive landscape of CD8⁺ T cell subsets in tissue samples and their distribution across conditions.

a, UMAP representation of CD8⁺ T cell sub-clusters in tissue, colored by distinct functional states. **b**, UMAP plots showing patient-specific distributions of CD8⁺ T cell sub-clusters across multiple individuals. **c**, Stacked bar plots depicting the relative abundance of CD8⁺ T cell sub-clusters in tissue across different conditions. **d**, Bar plots comparing the composition of CD8⁺ T cell sub-clusters in tissue before and after treatment across multiple samples, colored by distinct functional states. **f**, Table representation of enriched pathways or functional signatures associated with CD8⁺ T cell sub-clusters, stratified by condition. Blue: enriched in pretreatment. Red: enriched in posttreatment. This figure illustrates the heterogeneity of CD8⁺ T cells in tissue, their patient-specific variations, and dynamic changes in sub-cluster composition and functional states in response to treatment. Enrichment analyses highlight key pathways involved in CD4⁺ T cell dynamics in the peripheral blood compartment.





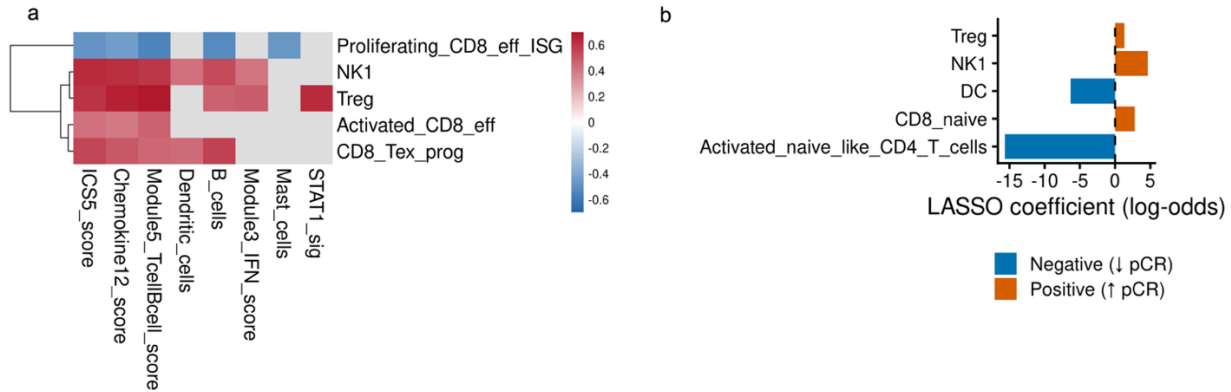
Supplementary Figure 4.4: Comprehensive landscape of CD8⁺ T cell subsets in tissue samples and their distribution across conditions.

a, UMAP representation of CD8⁺ T cell sub-clusters in tissue, colored by distinct functional states.



Supplementary Figure 4.5: Comprehensive landscape of CD8⁺ T cell subsets in tissue samples and their distribution across conditions, Continued.

b, UMAP plots showing patient-specific distributions of CD8⁺ T cell sub-clusters across multiple individuals. **c**, Stacked bar plots depicting the relative abundance of CD8⁺ T cell sub-clusters in tissue across different conditions. **d**, Bar plots comparing the composition of CD8⁺ T cell sub-clusters in tissue before and after treatment across multiple samples, colored by distinct functional states. **f**, Table representation of enriched pathways or functional signatures associated with CD8⁺ T cell sub-clusters, stratified by condition. Blue: enriched in pretreatment. Red: enriched in posttreatment. This figure illustrates the heterogeneity of CD8⁺ T cells in tissue, their patient-specific variations, and dynamic changes in sub-cluster composition and functional states in response to treatment. Enrichment analyses highlight key pathways involved in CD4⁺ T cell dynamics in the peripheral blood compartment.



Supplementary Figure 4.6: Peripheral immune signatures align with tumor immune programs and jointly associate with pCR.

a, Correlation heatmap between peripheral immune cell subset scores derived from PBMC profiling in this study and tumor immune module scores from the I-SPY2 cohort. Tumor scores represent previously defined chemokine, interferon/STAT1, and T/B-cell immune programs, while rows denote blood-derived immune subsets identified from our PBMC analysis. Positive correlations indicate concordant systemic tumor immune activation states, whereas negative correlations highlight opposing proliferative/ISG-biased CD8 programs. **b**, Multivariate LASSO logistic regression was applied to baseline PBMC-derived immune subset scores and I-SPY2 pCR outcomes to identify a coordinated peripheral immune signature associated with response. Bars represent non-zero coefficients at $\lambda_{\min}$ from cross-validated glmnet models. Positive coefficients (orange) indicate subsets positively associated with pCR, whereas negative coefficients (blue) indicate inverse associations. Coefficients are shown on the log-odds scale; the dashed line denotes no association.

REFERENCES

1. Mukund K, Nayak P, Ashokkumar C, Rao S, Almeda J, Betancourt-Garcia MM, Sindhi R, Subramaniam S. Immune Response in Severe and Non-Severe Coronavirus Disease 2019 (COVID-19) Infection: A Mechanistic Landscape. *Front Immunol*. 2021 Oct 13;12. doi:10.3389/fimmu.2021.738073
2. Jiang W, He Y, He W, Wu G, Zhou X, Sheng Q, Zhong W, Lu Y, Ding Y, Lu Q, Ye F, Hua H. Exhausted CD8+T Cells in the Tumor Immune Microenvironment: New Pathways to Therapy. *Front Immunol*. 2021 Feb 2;11. doi:10.3389/fimmu.2020.622509
3. Leon-Ferre RA, Whitaker KR, Suman VJ, Hoskin T, Giridhar KV, Moore RM, Al-Jarrad A, McLaughlin SA, Northfelt DW, Hunt KN, Connors AL, Moyer A, Carter JM, Kalari K, Weinshilboum R, Wang L, Ingle JN, Knutson KL, Ansell SM, Boughey JC, Goetz MP, Villasboas JC. Pre-treatment peripheral blood immunophenotyping and response to neoadjuvant chemotherapy in operable breast cancer. *Breast Cancer Research*. 2024 Jun 10;26(1):97. doi:10.1186/s13058-024-01848-z
4. Giacchetti S, Fauchoux L, Gardair C, Cuvier C, de Roquancourt A, Campedel L, Groheux D, de Bazelaire C, Lehmann-Che J, Miquel C, Cahen Doidy L, Amellou M, Madelaine I, Reyat F, Someil L, Hocini H, Hennequin C, Teixeira L, Espié M, Chevret S, Soumelis V, Hamy AS. Negative Relationship between Post-Treatment Stromal Tumor-Infiltrating Lymphocyte (TIL) and Survival in Triple-Negative Breast Cancer Patients Treated with Dose-Dense Dose-Intense NeoAdjuvant Chemotherapy. *Cancers (Basel)*. 2022 Mar 4;14(5):1331. doi:10.3390/cancers14051331
5. Obeagu EI, Obeagu GU. Lymphocyte infiltration in breast cancer: A promising prognostic indicator. *Medicine*. 2024 Dec 6;103(49):e40845. doi:10.1097/MD.00000000000040845
6. Tilsed CM, Principe N, Kidman J, Chin WL, Orozco Morales ML, Zemek RM, Chee J, Islam R, Fear VS, Forbes C, Aston WJ, Jansen M, Chopra A, Lassmann T, Nowak AK, Fisher SA, Lake RA, Lesterhuis WJ. CD4+ T cells drive an inflammatory, TNF- α /IFN-rich tumor microenvironment responsive to chemotherapy. *Cell Reports*. 2022 Dec 27;41(13):111874. doi:10.1016/j.celrep.2022.111874
7. Zhang Y, Chen H, Mo H, Hu X, Gao R, Zhao Y, Liu B, Niu L, Sun X, Yu X, Wang Y, Chang Q, Gong T, Guan X, Hu T, Qian T, Xu B, Ma F, Zhang Z, Liu Z. Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. *Cancer Cell*. 2021 Dec 13;39(12):1578-1593.e8. doi:10.1016/j.ccell.2021.09.010
8. Zhang Y, Chen H, Mo H, Zhao N, Sun X, Liu B, Gao R, Xu B, Zhang Z, Liu Z, Ma F. Distinct cellular mechanisms underlie chemotherapies and PD-L1 blockade combinations in triple-negative breast cancer. *Cancer Cell*. 2025 Mar 10;43(3):446-463.e7. doi:10.1016/j.ccell.2025.01.007
9. Reina-Campos M, Scharping NE, Goldrath AW. CD8+ T cell metabolism in infection and cancer. *Nat Rev Immunol*. 2021 Nov;21(11):718–38. doi:10.1038/s41577-021-00537-8

10. Wang V, Pober JS, Manes TD. Transendothelial migration of human B cells: chemokine versus antigen. *J Immunol.* 2023 Sep 15;211(6):923–31. doi:10.4049/jimmunol.2200887
11. Wolf DM, Yau C, Wulfkuhle J, Brown-Swigart L, Gallagher RI, Lee PRE, Zhu Z, Magbanua MJ, Sayaman R, O’Grady N, Basu A, Delson A, Coppé JP, Lu R, Braun J, Asare SM, Sit L, Matthews JB, Perlmutter J, Hylton N, Liu MC, Pohlmann P, Symmans WF, Rugo HS, Isaacs C, DeMichele AM, Yee D, Berry DA, Pusztai L, Petricoin EF, Hirst GL, Esserman LJ, Veer LJ van ’t. Redefining breast cancer subtypes to guide treatment prioritization and maximize response: Predictive biomarkers across 10 cancer therapies. *Cancer Cell.* 2022 Jun 13;40(6):609-623.e6. doi:10.1016/j.ccell.2022.05.005
12. Balko J, Sun X, Hanna A, Ocampo A, Taylor B, Marshall J, Steele J, Axelrod M, Wescott E, Opalenik S, DeMichele A, Esserman L, Liu M, Nanda R, Wolf D, Swigart LB, Hirst G, Van ’t Veer L, Xu Y. Peripheral blood gene expression signatures of systemic immunity predict tumor microenvironment biology and therapeutic response in breast cancer. *Res Sq.* 2025 Jul 7;rs.3.rs-6926989. doi:10.21203/rs.3.rs-6926989/v1
13. Mojić M, Radulović N, Bobić S, Radenković S, Cruz-Adalia A, Stojanović I. Innate lymphoid cells: Dual roles and therapeutic opportunities in breast cancer. *Immunology Letters.* 2025 Dec 1;276:107051. doi:10.1016/j.imlet.2025.107051
14. Ozga AJ, Chow MT, Luster AD. Chemokines and the immune response to cancer. *Immunity.* 2021 May 11;54(5):859–74. doi:10.1016/j.immuni.2021.01.012
15. Stress response in tumor-infiltrating T cells is linked to immunotherapy resistance. *Nat Med.* 2023 Jun;29(6):1336–7. doi:10.1038/s41591-023-02370-z
16. Li J, Bolyard C, Xin G, Li Z. Targeting Metabolic Pathways of Myeloid Cells Improves Cancer Immunotherapy. *Front Cell Dev Biol.* 2021 Dec 20;9. doi:10.3389/fcell.2021.747863
17. Porta C, Sica A, Riboldi E. Tumor-associated myeloid cells: new understandings on their metabolic regulation and their influence in cancer immunotherapy. *The FEBS Journal.* 2018;285(4):717–33. doi:10.1111/febs.14288
18. Puleo J, Polyak K. A Darwinian perspective on tumor immune evasion. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer.* 2022 Jan 1;1877(1):188671. doi:10.1016/j.bbcan.2021.188671
19. Yu H, Liu Y, McFarland BC, Deshane JS, Hurst DR, Ponnazhagan S, Benveniste EN, Qin H. SOCS3 Deficiency in Myeloid Cells Promotes Tumor Development: Involvement of STAT3 Activation and Myeloid-Derived Suppressor Cells. *Cancer Immunol Res.* 2015 Jul;3(7):727–40. doi:10.1158/2326-6066.CIR-15-0004
20. Cheng H, Zheng Y. Advances in macrophage and T cell metabolic reprogramming and immunotherapy in the tumor microenvironment. *PeerJ.* 2024 Jan 15;12:e16825. doi:10.7717/peerj.16825

21. Contento RL, Molon B, Boularan C, Pozzan T, Manes S, Marullo S, Viola A. CXCR4–CCR5: A couple modulating T cell functions. *Proceedings of the National Academy of Sciences*. 2008 Jul 22;105(29):10101–6. doi:10.1073/pnas.0804286105
22. Liu M, Qian M, Sun W, Sun X, Sun Y, Yu M, Tang X, Mao X, Sun C, Qi Q, Zhang W, Ling P, Pang Z, Li W, Pan H, Wang S, Zhou W. Immunosuppressive microenvironment of liver restrains chemotherapeutic efficacy in triple-negative breast cancer. *J Immunother Cancer*. 2025 Mar 6;13(3):e010871. doi:10.1136/jitc-2024-010871
23. Chu Y, Dai E, Li Y, Han G, Pei G, Ingram DR, Thakkar K, Qin JJ, Dang M, Le X, Hu C, Deng Q, Sinjab A, Gupta P, Wang R, Hao D, Peng F, Yan X, Liu Y, Song S, Zhang S, Heymach JV, Reuben A, Elamin YY, Pizzi MP, Lu Y, Lazcano R, Hu J, Li M, Curran M, Futreal A, Maitra A, Jazaeri AA, Ajani JA, Swanton C, Cheng XD, Abbas HA, Gillison M, Bhat K, Lazar AJ, Green M, Litchfield K, Kadara H, Yee C, Wang L. Pan-cancer T cell atlas links a cellular stress response state to immunotherapy resistance. *Nat Med*. 2023 Jun;29(6):1550–62. doi:10.1038/s41591-023-02371-y
24. Zhu X, Zhu J. CD4 T Helper Cell Subsets and Related Human Immunological Disorders. *Int J Mol Sci*. 2020 Oct 28;21(21):8011. doi:10.3390/ijms21218011
25. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*. 2014 Apr;32(4):381–6. doi:10.1038/nbt.2859
26. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, Trapnell C. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. 2017 Oct;14(10):979–82. doi:10.1038/nmeth.4402
27. Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, Trapnell C, Shendure J. The single-cell transcriptional landscape of mammalian organogenesis. *Nature*. 2019 Feb;566(7745):496–502. doi:10.1038/s41586-019-0969-x
28. Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep*. 2019 Mar 26;9(1):5233. doi:10.1038/s41598-019-41695-z
29. Levine JH, Simonds EF, Bendall SC, Davis KL, Amir E ad D, Tadmor MD, Litvin O, Fienberg HG, Jager A, Zunder ER, Finck R, Gedman AL, Radtke I, Downing JR, Pe’er D, Nolan GP. Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell*. 2015 Jul 2;162(1):184–97. doi:10.1016/j.cell.2015.05.047
30. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology*. 2012 May;16(5):284–7. doi:10.1089/omi.2011.0118
31. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW,

Ma'ayan A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 2016 Jul 8;44(W1):W90-97. doi:10.1093/nar/gkw377

32. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering C von. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research.* 2019 Jan;47(D1):D607–13. doi:10.1093/nar/gky1131

33. Han H, Cho JW, Lee S, Yun A, Kim H, Bae D, Yang S, Kim CY, Lee M, Kim E, Lee S, Kang B, Jeong D, Kim Y, Jeon HN, Jung H, Nam S, Chung M, Kim JH, Lee I. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res.* 2018 Jan 4;46(D1):D380–6. doi:10.1093/nar/gkx1013

34. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research.* 2003 Nov;13(11):2498–504. doi:10.1101/gr.1239303

35. Piccirillo CA, Bjur E, Topisirovic I, Sonenberg N, Larsson O. Translational control of immune responses: from transcripts to translatores. *Nat Immunol.* 2014 Jun;15(6):503–11. doi:10.1038/ni.2891

36. Wan L, Juszkievicz S, Blears D, Bajpe PK, Han Z, Faull P, Mitter R, Stewart A, Snijders AP, Hegde RS, Svejstrup JQ. Translation stress and collided ribosomes are co-activators of cGAS. *Mol Cell.* 2021 Jul 1;81(13):2808-2822.e10. doi:10.1016/j.molcel.2021.05.018

37. Zheng H, Siddharth S, Parida S, Wu X, Sharma D. Tumor Microenvironment: Key Players in Triple Negative Breast Cancer Immunomodulation. *Cancers.* 2021 Jan;13(13):3357. doi:10.3390/cancers13133357

38. Ribas A, Wolchok JD. Cancer immunotherapy using checkpoint blockade. *Science.* 2018 Mar 23;359(6382):1350–5. doi:10.1126/science.aar4060

39. Philip M, Schietinger A. CD8+ T cell differentiation and dysfunction in cancer. *Nat Rev Immunol.* 2022 Apr;22(4):209–23. doi:10.1038/s41577-021-00574-3

40. Wherry EJ, Kurachi M. Molecular and cellular insights into T cell exhaustion. *Nat Rev Immunol.* 2015 Aug;15(8):486–99. doi:10.1038/nri3862

41. Yi M, Li T, Niu M, Mei Q, Zhao B, Chu Q, Dai Z, Wu K. Exploiting innate immunity for cancer immunotherapy. *Molecular Cancer.* 2023 Nov 27;22(1):187. doi:10.1186/s12943-023-01885-w

42. Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, Lickley LA, Rawlinson E, Sun P, Narod SA. Triple-Negative Breast Cancer: Clinical Features and Patterns of Recurrence. *Clin Cancer Res.* 2007 Aug 1;13(15):4429–34. doi:10.1158/1078-0432.CCR-06-3045

43. Wang XQ, Danenberg E, Huang CS, Egle D, Callari M, Bermejo B, Dugo M, Zamagni C, Thill M, Anton A, Zambelli S, Russo S, Ciruelos EM, Greil R, Györfy B, Semiglazov V, Colleoni M, Kelly CM, Mariani G, Del Mastro L, Biasi O, Seitz RS, Valagussa P, Viale G, Gianni L, Bianchini G, Ali HR. Spatial predictors of immunotherapy response in triple-negative breast cancer. *Nature*. 2023 Sep;621(7980):868–76. doi:10.1038/s41586-023-06498-3
44. Danenberg E, Bardwell H, Zanutelli VRT, Provenzano E, Chin SF, Rueda OM, Green A, Rakha E, Aparicio S, Ellis IO, Bodenmiller B, Caldas C, Ali HR. Breast tumor microenvironment structures are associated with genomic features and clinical outcome. *Nat Genet*. 2022 May;54(5):660–9. doi:10.1038/s41588-022-01041-y
45. Lasser SA, Ozbay Kurt FG, Arkhypov I, Utikal J, Umansky V. Myeloid-derived suppressor cells in cancer and cancer therapy. *Nat Rev Clin Oncol*. 2024 Feb;21(2):147–64. doi:10.1038/s41571-023-00846-y
46. Hoechst B, Gamrekelashvili J, Manns MP, Greten TF, Korangy F. Plasticity of human Th17 cells and iTregs is orchestrated by different subsets of myeloid cells. *Blood*. 2011 Jun 16;117(24):6532–41. doi:10.1182/blood-2010-11-317321
47. Hu A, Sun L, Lin H, Liao Y, Yang H, Mao Y. Harnessing innate immune pathways for therapeutic advancement in cancer. *Sig Transduct Target Ther*. 2024 Mar 25;9(1):68. doi:10.1038/s41392-024-01765-9
48. Thacker G, Henry S, Nandi A, Debnath R, Singh S, Nayak A, Susnik B, Boone MM, Zhang Q, Kesmodel SB, Gumber S, Das GM, Kambayashi T, Dos Santos CO, Chakrabarti R. Immature natural killer cells promote progression of triple-negative breast cancer. *Sci Transl Med*. 2023 Mar 8;15(686):eabl4414. doi:10.1126/scitranslmed.abl4414
49. Imai K, Matsuyama S, Miyake S, Suga K, Nakachi K. Natural cytotoxic activity of peripheral-blood lymphocytes and cancer incidence: an 11-year follow-up study of a general population. *The Lancet*. 2000 Nov 25;356(9244):1795–9. doi:10.1016/S0140-6736(00)03231-1
50. Xu L, Saunders K, Huang SP, Knutsdottir H, Martinez-Algarin K, Terrazas I, Chen K, McArthur HM, Maués J, Hodgdon C, Reddy SM, Roussos Torres ET, Xu L, Chan IS. A comprehensive single-cell breast tumor atlas defines epithelial and immune heterogeneity and interactions predicting anti-PD-1 therapy response. *Cell Rep Med*. 2024 Apr 12;5(5):101511. doi:10.1016/j.xcrm.2024.101511

Chapter 5 A Spatially resolved tumor-microenvironment architecture defines therapeutic response in unresectable hepatoblastoma

5.1 Abstract

Hepatoblastoma exhibits variable therapeutic response that is not fully explained by current risk stratification. While prior studies have characterized tumor composition, the role of spatial organization in determining outcome remains unclear. We integrated spatial transcriptomics using the 10x Genomics Visium platform ($n = 11$ tumors) with single-cell spatial molecular imaging using NanoString CosMx ($n = 3$ tumors) across clinically defined groups (relapse-prone unresectable, non-relapse unresectable, and resectable hepatoblastoma). Spatial deconvolution, module scoring, spatial co-localization, and cell-cell interaction analyses were used to define region-level transcriptional programs and their underlying cellular organization. We identified three spatial architectures associated with clinical outcome. Relapse-prone tumors exhibited a centralized, TME-dominated organization enriched for progenitor-like, cycling, and stress-responsive tumor cell states co-localized with SPP1⁺ TAMs and CAFs, with enrichment of hypoxia response, extracellular matrix remodeling, and Midkine (MDK-LRP1) signaling. Non-relapse tumors displayed a distributed architecture with compartmentalized tumor-immune interactions, increased immune engagement, and preserved epithelial differentiation, including inflammatory hepatocyte-like states. Resection samples showed reduced tumor cell state heterogeneity and more uniform TME-immune organization. Integration of Visium and CosMx data enabled mapping of region-level transcriptional programs to single-cell spatial organization, demonstrating that therapeutic response in hepatoblastoma is encoded in spatially organized tumor-microenvironment interactions.

5.2 Introduction

Hepatoblastoma is the most common primary liver malignancy in children, typically arising before the age of four and frequently associated with developmental anomalies (1–3). While multimodal treatment strategies combining chemotherapy and surgical resection have improved survival rates, a subset of patients with unresectable tumors requiring liver transplantation experience relapse, which remains frequently fatal (4,5). Current risk stratification frameworks, including the PRETEXT staging system, histological classification, and serum alpha-fetoprotein (AFP) levels, provide important clinical guidance but remain insufficient to fully explain variability in therapeutic response.

Unlike many adult cancers, hepatoblastoma exhibits a relatively low mutational burden (6,7), suggesting that tumor heterogeneity is driven less by genetic alterations and more by dysregulated developmental programs. Indeed, HB tumors recapitulate features of fetal liver development, with distinct epithelial states ranging from differentiated hepatocyte-like populations to progenitor and mesenchymal-like programs. These transcriptional states are governed by key developmental pathways, including Wnt/ β -catenin, Notch, and Hedgehog signaling, and are associated with differences in proliferation, differentiation, and clinical outcome (8–12). In parallel, the tumor microenvironment (TME) of hepatoblastoma is increasingly recognized as a critical regulator of tumor progression and therapeutic response (13). Despite the success of immune checkpoint blockade (ICB) therapies in adult cancers, their efficacy in pediatric tumors, including hepatoblastoma, has been limited, potentially due to low tumor mutational burden and reduced T-cell infiltration characteristic of pediatric malignancies (14). These features, combined with the embryonic origin of hepatoblastoma, suggest a fundamentally distinct and potentially immune-restricted tumor microenvironment. While prior studies have begun to characterize

hepatoblastoma immune composition using bulk and single-cell approaches (13,15,16), a comprehensive understanding of its spatially organized tumor microenvironment, particularly in the context of chemotherapy, remains lacking.

Here, we integrate region-level spatial transcriptomics with single-cell spatial molecular imaging across clinically defined groups (unresectable (relapse-prone and non-relapse), and resectable hepatoblastoma samples) to define the spatial organization of hepatoblastoma following chemotherapy. By combining Visium-based deconvolution with CosMx single-cell spatial profiling, we characterize how tumor cell states, stromal populations, and immune cells are organized into distinct spatial architectures across clinical outcomes. We show that in unresectable group, relapse-prone tumors are characterized by a macrophage-fibroblast-centered regulatory structure associated with hypoxia-adaptive tumor states and Midkine signaling, whereas non-relapse tumors exhibit distributed tumor-immune interactions and preserved differentiation hierarchies. Together, these findings demonstrate that therapeutic response in unresectable hepatoblastoma is encoded in spatial organization of cell-type and -state composition.

5.3 Patients and Methods

Human Subjects and Clinical Outcomes

All studies were conducted with informed consent and approval from the University of Pittsburgh Institutional Review Board (IRB #0405628). Twelve patients with hepatoblastoma who received treatment between 2007 and 2019, were included in this study, with a median age at liver transplantation or primary tumor resection of 3.1 years (range, 0.99–6.45 years); 10/12 were male (**Supplementary Table 5.1**). Ten patients underwent liver transplantation, while two underwent primary tumor resection. Among the hepatoblastoma samples analyzed, two developed post-transplant recurrence (relapse) and resulted in death from progressive disease at 0 and

18 months, respectively, while ten remained non-relapsed at last follow-up (median, 121 months) (**Supplementary Table 5.1-4.2**). PRETEXT stage at diagnosis was available for eleven tumors and ranged from stages I to IV, with one tumor lacking diagnostic imaging at presentation (**Supplementary Table 5.1 and 5.3**). Relapse status was determined by longitudinal radiologic follow-up and post-treatment histopathologic evaluation. Post-treatment pathologic assessment showed variable responses to chemotherapy, ranging from extensive tumor necrosis to residual viable disease, with diverse histologic and metastatic features observed (**Supplementary Table 5.2-5.3**).

Spatial Transcriptomics of Hepatoblastoma Tumors Using the Visium Platform

Spatial transcriptomic profiling was performed on formalin-fixed, paraffin-embedded (FFPE) liver tumor tissue sections from twelve children with hepatoblastoma (**Supplementary Table 5.1**) using the Visium Spatial Gene Expression platform and the FFPE Gene Expression Starter Kit for Human transcriptome (10x Genomics, Pleasanton, CA). FFPE tissue sections were placed onto the 6.5 mm × 6.5 mm capture area of Visium slides containing approximately 5,000 spatially barcoded spots. Each spot comprised oligonucleotides including a 16-nucleotide spatial barcode for positional identification, a 30-nucleotide poly(dT) sequence for mRNA capture, a 12-nucleotide unique molecular identifier (UMI), and an Illumina Read 1 sequencing adapter. Tissue sections were deparaffinized and stained with hematoxylin and eosin, followed by high-resolution imaging to enable spatial alignment of transcriptomic data. Crosslinks were reversed to release RNA, and transcriptome-wide probe sets were hybridized to target mRNA molecules. Probe pairs bound to RNA were ligated, after which tissues were treated with RNase and permeabilized to release ligated products for capture on the Visium slide. Captured products were extended via a polymerase reaction to incorporate complementary sequences of the spatial

barcode and UMI. Spatially barcoded probe products were released and PCR-amplified using sample-specific indexing primers. Final libraries were sequenced to a depth of approximately 25,000 raw reads per capture spot containing tissue using a NovaSeq 6000 system at the UPMC Genome Center. Sequencing data were aligned to histologic images using Space Ranger software (10x Genomics, Pleasanton, CA) and visualized with Loupe Browser. Each capture spot measured 55 μm in diameter with 100 μm center-to-center spacing, corresponding to an estimated average of 5–10 cells per spot, depending on tissue cellularity. Of the twelve tumors processed for Visium analysis, one failed quality-control criteria and was excluded from downstream spatial transcriptomic analyses.

Single Cell Spatial Molecular Imaging of Hepatoblastoma Tumors Using CosMx SMI

Spatial molecular imaging (SMI) was performed on FFPE liver tumor tissue sections from five children with hepatoblastoma (**Supplementary Table 5.1**) using the CosMx™ Spatial Molecular Imager (NanoString Technologies, Bruker; Seattle, WA) and the Human Universal Cell Characterization of 6000 genes panel (part no. 121500041), following the manufacturer's protocol (MAN-10184-02). Tissue sections were cut at 5 μm thickness, mounted on Superfrost Plus slides, and baked overnight at 60 °C. Sections were deparaffinized, subjected to antigen retrieval at 95–100 °C for 15 minutes, and permeabilized with Proteinase K (3 $\mu\text{g}/\text{mL}$) at 40 °C for 30 minutes. Fiducial markers were applied to facilitate image registration, followed by post-fixation in neutral buffered formalin and blocking with N-hydroxysuccinimide (NHS) acetate. In situ hybridization was performed by overnight incubation at 37 °C with RNA-specific probe sets targeting approximately 6,000 transcripts. After stringent washes, nuclei were stained with DAPI, and cell-segmentation markers including CD298/ β 2-microglobulin (B2M), pan-cytokeratin (PanCK), and CD45 were applied. Prepared slides were imaged, and iterative cycles of branched fluorescent probe hybridization enabled highly multiplexed RNA detection at single-cell resolution. Images

were processed and decoded using the AtoMx™ Spatial Informatics Platform, and the resulting spatial expression data were exported for downstream analysis using open-source bioinformatics tools.

Visium data preprocessing and quality control

Twelve tissue sections were processed using the Visium spatial transcriptomics platform (10x Genomics, Pleasanton, CA). Downstream analysis of Visium data was performed using the Seurat package (v4.3.0) (17). One sample was excluded due to poor quality, resulting in a final dataset of 11 samples. Each sample was independently normalized using Seurat's SCTransform (SCT) method (18) and subsequently merged into a combined dataset comprising 16,481 features across 51,016 spots. Quality control filtering was then applied to the merged dataset, retaining spots with a total Unique Molecular Identifier (UMI) count between 500 and 50,000, more than 500 detected genes, and less than 10% hemoglobin gene content. This filtering yielded a final dataset of 16,481 features and 49,505 spatial spots. To identify genes with the highest expression across the spatial transcriptomics dataset, we extracted the raw count matrix from the Spatial assay within the merged Seurat object. Counts were normalized by dividing each gene count by the total number of counts per spot to account for sequencing depth. The top 20 most highly expressed genes were ranked by total normalized expression. The top ALB gene was liver-specific, highly abundant transcripts such as albumin, were excluded from downstream analysis to reduce tissue-specific expression bias. After this exclusion, the dataset consisted of 16,480 features and 49,505 spots. Dimensionality reduction and clustering were performed on the SCT-normalized data. Principal component analysis (PCA) was conducted, followed by construction of a shared nearest neighbor (SNN) graph using the top principal components. Louvain-based clustering was applied to identify transcriptionally distinct spatial domains, and Uniform Manifold Approximation and Projection (UMAP) was used for two-dimensional visualization of the clustering results.

Visium data feature Selection and deconvolution

For deconvolution, STdeconvolve (19) was applied to each of the 11 samples individually. STdeconvolve uses Latent Dirichlet Allocation (LDA) framework to model each spot as a mixture of latent transcriptional programs (“topics”), where each topic represents a gene expression distribution that may correspond to a cell type or cell state. Feature selection was performed by removing genes expressed in fewer than 5% or more than 95% of pixels. Overdispersed genes- defined as genes exhibiting higher variability than expected given their mean expression- were identified using a generalized additive model with five basis functions and an adjusted p-value threshold of 0.05. The 1,000 most significantly overdispersed genes were retained for model fitting. LDA models were trained on each dataset using a range of topic numbers (K) from 2 to 10. For each sample, the model with $K = 5$ was selected as it achieved the lowest perplexity (a measure of model fit reflecting how well the model predicts unseen data) while avoiding overfitting to rare cell types, defined as deconvolved cell types with mean spot proportions less than 0.05. Cell-type-specific marker genes were identified from the LDA output by calculating log2 fold changes in expression between each deconvolved cell type and all others. Genes with log2 fold change greater than 1 and high expression levels were retained as top markers (**Supplementary Table 5.4**). Each deconvolved transcriptional profile, along with its inferred spot-level inferred topic proportions, was mapped back to the merged Seurat object for downstream analysis. Topic assignment was performed using the STdeconvolve method. t-SNE was used to visualize the structure of the deconvolved transcriptional profiles, where visually defined groupings were used to guide topic assignment based on the highest correlation with their transcriptional profiles. These topic assignments were then mapped onto Seurat-defined clusters. Using marker gene profiles from both STdeconvolve and differentially expressed genes (DEGs), each deconvolved cell type was

annotated into major biological regions: tumor, stroma, fibrosis, immune mixtures, and hepatocyte. These regional annotations were further validated by a pathologist based on histological inspection of H&E images.

Visium data enrichment and visualization

Region-specific differential expression gene (DEG) analysis was then performed by comparing each annotated region to all other regions within the same sample. Differentially expressed genes were identified using Seurat's FindAllMarkers function with the Wilcoxon rank-sum test (parameters: $\logFC > 0.25$, $\min.pct > 0.25$, adjusted p-value < 0.05). Multiple testing correction was performed using the Benjamini–Hochberg false discovery rate (FDR) method. Functional enrichment analysis of DEGs was conducted using the clusterProfiler (20) and Enrichr (21) packages in R. Gene Ontology (GO) analyses for biological processes and molecular functions were performed using the enrichGO function. To explore molecular interactions, a high-confidence human protein-protein interaction (PPI) network (confidence score > 0.85) was obtained from the STRING database v11.5 (22), containing 12,402 proteins and 139,384 interactions. Transcription factor (TF)-target gene relationships from the TRRUST database (23) were integrated with the PPI network to generate a custom TF-PPI network for visualizing regulatory programs associated with region-specific DEGs. All networks were constructed, annotated, and analyzed in Cytoscape (24).

CosMx data preprocessing and quality control

Single-cell spatial transcriptomic data were generated using CosMx™ Spatial Molecular Imager (NanoString Technologies, Bruker; Seattle, WA) platform. Raw decoded data, including per-cell gene expression matrices, spatial coordinates, and cell-level metadata, were obtained from the AtoMx analysis pipeline (NanoString Technologies, Bruker; Seattle, WA), which performs

automated cell segmentation and transcript assignment based on imaging-derived morphology and transcript localization. Downstream analysis of CosMx data was performed using the Seurat package (v5.4) (25). Decoded data were imported into R using NanoStringNCTools package and converted into Seurat objects. Three CosMx datasets were processed and merged for initial quality control and metadata harmonization. Cell-level annotations, including field-of-view (FOV) identifiers and patient labels, were assigned based on NanoString metadata and external sample manifests. Cells lacking patient annotation were excluded before downstream analysis. Quality control was performed using cell-level metrics, including nCount_RNA, nFeature_RNA, and percent.NegPrb. Cells were retained if they satisfied the following criteria: nCount between 500-10,000, nFeature_RNA between 400-5,000, and percent.NegPrb $\leq 10\%$. Negative control probes and system control probes were separated into independent assays and excluded from downstream expression analysis. This filtering yielded a final dataset of 6,175 features and 348,350 single-cells. To preserve biologically meaningful spatial heterogeneity and avoid cross-sample integration artifacts, all downstream analyses were performed independently for each patient. Cells were analyzed separately for each patient and normalized using SCTransform with 3,000 variable features per sample. For each patient, PCA was performed, and the top 20 principal components were used to construct a shared nearest neighbor graph. Clustering was performed using the Louvain algorithm across multiple resolutions (0.4-0.6), with resolution 0.5 selected based on cluster stability. UMAP was used to visualize patient-specific cellular structure.

CosMx data cell type annotation

Cell clusters were annotated independently for each patient using a combination of 1) Canonical marker gene expression, 2) differential expression analysis (FindAllMarkers), and 3) spatial localization within tissue. Marker panels included epithelial (e.g., EPCAM, AFP), immune

(e.g., *PTPRC*, *CD3D*, *CD68*), stromal (e.g., *COL1A1*, *DCN*), and endothelial (e.g., *VWF*, *CDH5*). Annotations were refined iteratively to ensure consistency across transcriptional and spatial features.

CosMx spatial reconstruction

Cell centroid coordinates provided by the CosMx platform were used to reconstruct spatial organization. Coordinates were incorporated into Seurat as field-of-view (FOV) objects, enabling visualization of cell-type distributions within tissue architecture.

CosMx data differential expression and pathway analysis

Differential gene expression analysis was performed independently for each patient and annotated cell population using Seurat's FindAllMarkers function. Functional enrichment analysis was conducted using MSigDB (26) gene sets (C5 category) and fGSEA (27), with ranked gene lists defined by signed $-\log_{10}$ adjusted p-values.

Cell-cell interaction and neighborhood analysis

Intercellular signaling was inferred using the CellChat (28) framework, which models communication probability based on the expression of curated ligand-receptor pairs. For each patient-specific dataset, normalized expression data and annotated cell identities were used as input. Communication probabilities were computed using the default CellChat probabilistic model, which accounts for the average expression of ligands in sender cell populations and receptors in receiver populations. Statistical significance of inferred interactions was assessed through permutation testing, with multiple testing correction performed using Benjamini-Hochberg false discovery rate (FDR) method. High-confidence interactions (adjusted $p < 0.01$, top 10% interaction strength) were retained for downstream analysis.

To incorporate spatial context, cell-cell proximity relationships were modeled using Giotto (v1.0.4) (29). For each patient, spatial coordinates derived from CosMx cell centroids were used to construct spatial networks. A Delaunay triangulation-based graph was generated to define adjacency relationships between cells, enabling unbiased identification of neighboring cells independent of predefined distance thresholds. This approach captures local tissue architecture while preserving connectivity across irregular spatial distributions. From the Delaunay network, spatial neighborhoods were defined as first-order adjacent cells (direct graph neighbors). Cell-type-specific neighborhood enrichment was quantified using Giotto's cellProximityEnrichment function, which compares observed cell-cell adjacency frequencies against randomized spatial permutations to identify statistically enriched or depleted interactions.

Integration of Visium and CosMx data

To relate region-level transcriptional programs identified in Visium data to single-cell spatial organization, we employed a marker-guided annotation strategy. Region-specific gene signatures derived from STdeconvolve analysis of Visium data described above were used to define transcriptional programs associated with tumor, TME, and hepatocyte-enriched regions. These marker sets were then evaluated within the CosMx single-cell dataset to identify spatial regions exhibiting concordant transcriptional profiles. Cells in the CosMx dataset were scored using these Visium-derived marker gene set via AddModuleScore function, enabling assignment of region-associated transcriptional states. Following region annotation, the cellular composition and heterogeneity within each region were characterized using CosMx-defined cell types and spatial organization. This approach enabled resolution of the cell populations and interaction networks underlying Visium-defined spatial domains at single-cell resolution.

Data Availability

Visium and CosMx SMI data will be provided in the published format.

5.4 Results

Integration of Visium sequencing and spatial molecular imaging analysis

To characterize spatial organization of hepatoblastoma following chemotherapy, we performed spatial transcriptomics (ST) analysis using the Visium platform (10x Genomics, Pleasanton, CA) on eleven tumor sections representing distinct clinical groups unresectable (relapse-prone and non-relapse), and resectable hepatoblastoma samples (**Supplementary Table 5.1**). Histological annotation of H&E-stained slides delineated distinct tissue regions across section (**Figure 5.1a-b; Supplementary Table 5.2**). Three representative samples, one per clinical group, were selected for CosMx spatial molecular imaging (SMI; NanoString Technologies, Bruker; Seattle, WA) profiling with multiplex panel targeting ~6000 genes to capture canonical spatial architectures and to resolve the cellular composition underlying Visium-defined regions at single-cell resolution (**Supplementary Table 5.1**). Here, cell types refer to canonical lineage-defined populations, whereas cell states denote context-dependent transcriptional programs within these populations, such as activation, stress, or differentiation states.

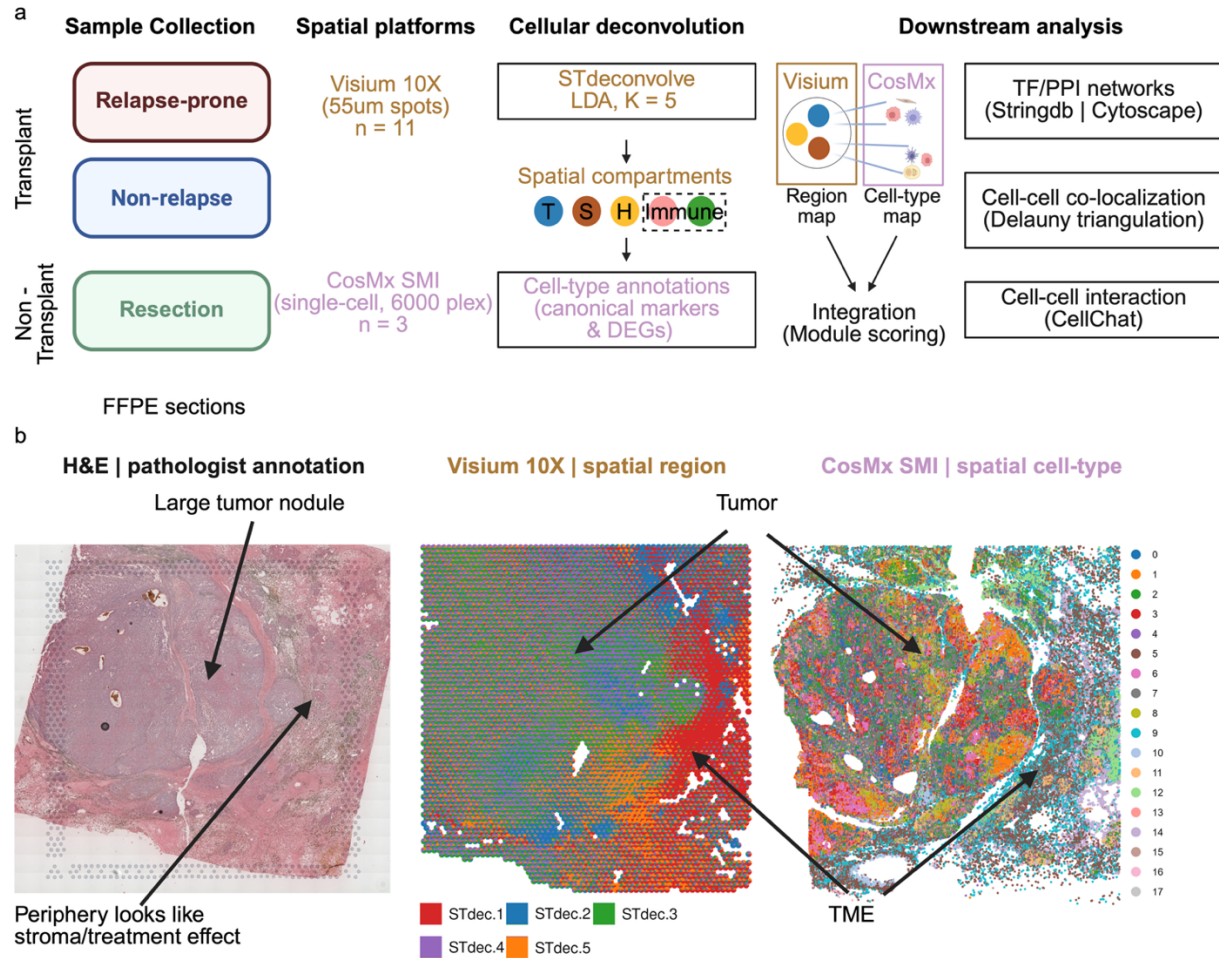


Figure 5.1: Integrative spatial transcriptomics workflow and cross-platform mapping of tumor architecture in relapse-prone unresectable hepatoblastoma.

a, Schematic overview of the study design. FFPE tumor sections from relapse-prone and non-relapsing unresectable HB, and HB cured with resection were profiled using Visium 10X ($n = 11$) and CosMx SMI ($n = 3$). Reference-free deconvolution of Visium data using STdeconvolve (LDA, $K = 5$) identified spatial compartments/regions, including tumor (T), stroma (S), hepatocyte-like (H), and immune-associated regions. Cell-type annotations were refined using canonical markers and differentially expressed genes (DEGs). Integration of Visium region-level and CosMx single-cell maps was performed via module scoring. Downstream analyses included cell-cell communication (CellChat), spatial co-localization (Delauny triangulation), and transcription factor/protein-protein interaction (TF/PPI) network analysis (Stringdb/Cytoscape). **b**, Representative spatial correspondence across modalities. Left: H&E-stained section with pathologist annotation highlighting a large tumor nodule and peripheral stroma/treatment-associated regions. Middle: Visium-derived spatial compartments (STdec clusters) delineating tumor and surrounding microenvironment. Right: CosMx single-cell map showing high-resolution cell-type distribution, distinguishing tumor and TME compartments. Together, these data demonstrate concordance between histopathology, region-level transcriptomics, and single-cell spatial profiling.

Following quality control filtering, 32,960 features across 49,505 spatial spots from Visium and 6,175 features across 348,350 segmented cells from CosMx were retained for downstream analyses (**Figure 5.2a-b**). Each Visium spot (55 μ m diameter) captures transcripts from multiple cells (typically 5-50), necessitating deconvolution to infer cellular composition. To address this, we applied a reference-free deconvolution algorithm STdeconvolve and defined cell types (STdec) based on highly overdispersed genes (**Figure 5.2c**; see Methods). In parallel, unsupervised clustering identified transcriptional profiles independent of deconvolution (**Figure 5.2d**). Integration of these approaches with histological annotations enabled assignment of major spatial compartments, including tumor, stroma, hepatocyte, and immune-enriched regions (**Figure 4.2e**; see Methods; **Supplementary Table 5.4** and **Supplementary Figure 5.1 and 5.2a**). To link these region-level programs to single-cell resolution, we projected them onto CosMx data using module scoring, enabling resolution of the cellular state underlying each spatial compartment (**Supplementary Table 5.5**). This integration establishes a multi-scale framework linking regional transcriptional programs with single-cell spatial organization.

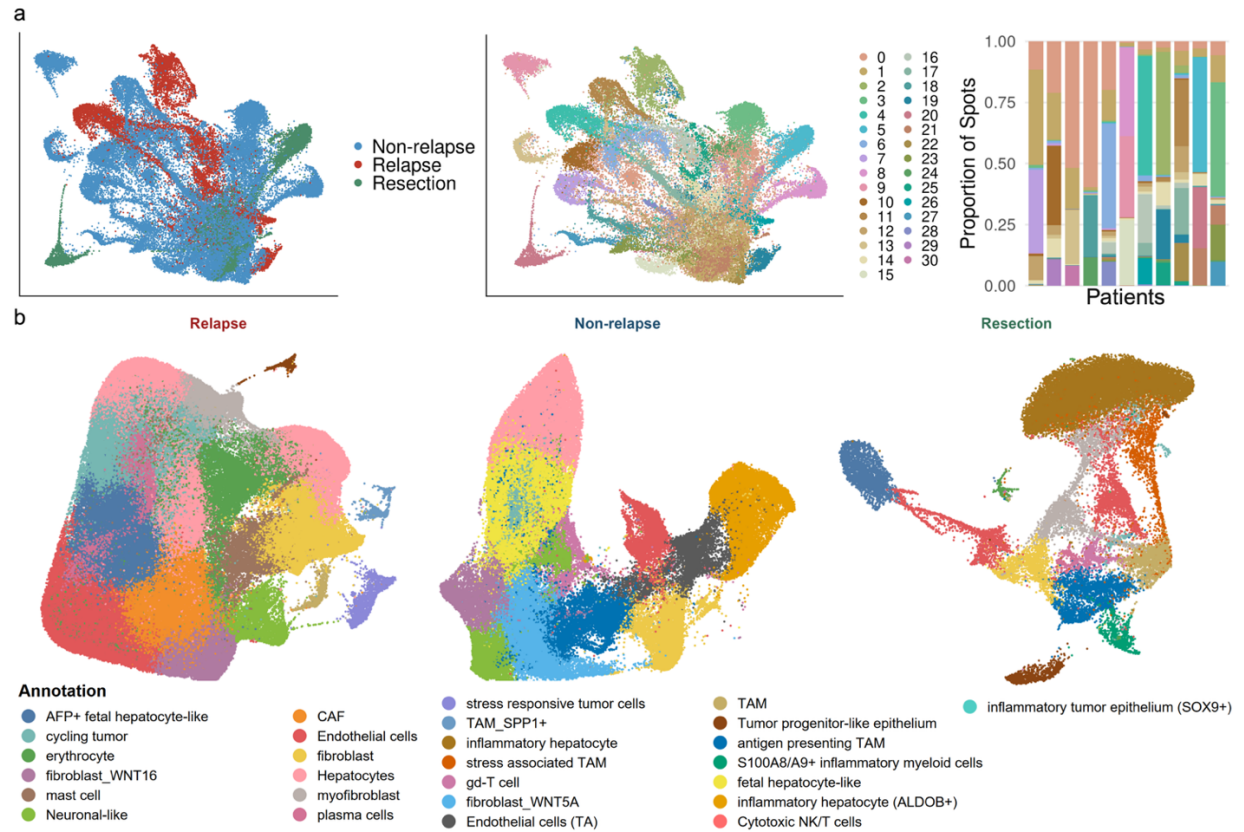


Figure 5.2: Spatially resolved tumor ecosystems across hepatoblastoma clinical groups.

a, Integrated low-dimensional embedding of Visium spots across all samples colored by clinical group (left; relapse, non-relapse, resection) and Seurat clusters (middle), revealing shared and distinct transcriptional programs. Right: stacked bar plot showing the proportion of Seurat clusters per patient, highlighting inter-patient heterogeneity and differential cluster enrichment across clinical groups. **b**, High-resolution cell-state landscapes derived from CosMx SMI for representative samples from each clinical group (relapse, non-relapse, resection). Cells are annotated based on canonical markers and DEGs, capturing tumor, stromal, immune, hepatocyte-associated states.

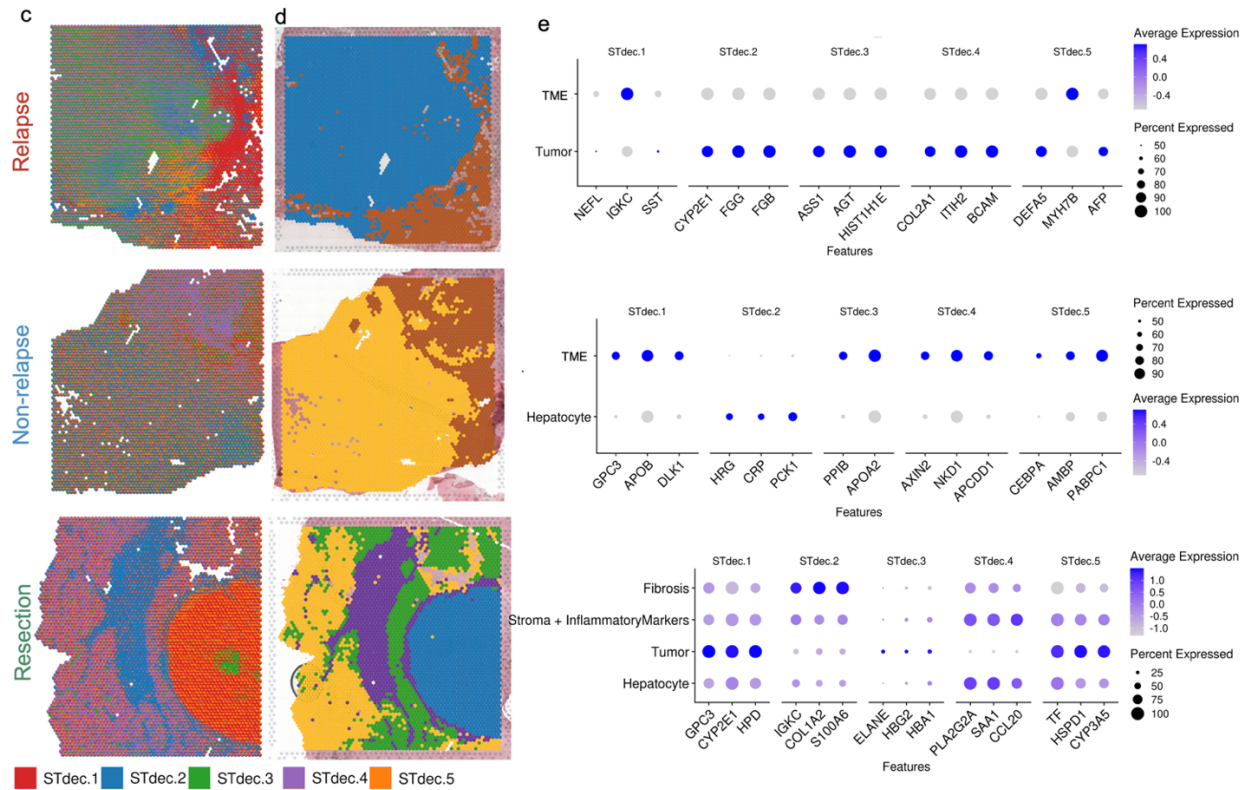


Figure 5.3: Spatially resolved tumor ecosystems across hepatoblastoma clinical groups, Continued.

c, Visium spatial maps of STdec topics ($K = 5$, STdec1 through STdec5) for representative sections from relapse, non-relapse, and resection samples, illustrating region-level compartmentalization of tumor and microenvironmental programs. **d**, Corresponding region-level annotations derived from Seurat cluster profiling and marker-based interpretation, delineating tumor, stroma, hepatocyte-like, and inflammatory compartments across clinical groups. **e**, Dot plots showing top 3 marker gene enrichment across STdeconvolve topics ($K = 5$) for each clinical group. Dot size represents the percentage of spots expressing each gene, and color indicates average expression. These profiles define the biological identity of spatial compartments, including tumor-associated programs, hepatocyte-like states, fibrosis/stromal signatures, and inflammatory or immune-associated features.

Spatial architecture differs between relapse-prone, non-relapse, and resection hepatoblastoma samples

Relapse tumors exhibit a TME-dominated and spatially coupled architecture. Mapping CosMx-derived cellular states to Visium-defined regions revealed marked differences in spatial organization across clinical groups (**Figure 5.3a**; **Supplementary Figure 5.2b-d**). Importantly, this integration enabled resolution of cellular states not captured by STdec markers, refining the

cellular composition of spatial regions. To derive mechanistic insights into the regulatory drivers of region-specific transcriptional programs, we integrated differential expression with protein-protein interaction (PPI) and TF-target network analyses, and annotated variants associated with SNV/CNV to assess their functional contribution.

Within post-chemotherapy unresectable samples, relapse-prone tumors exhibited high tumor cell-type and -state heterogeneity, with tumor regions composed of progenitor-like, stress-responsive, AFP⁺ fetal hepatocyte-like, and cycling tumor cells embedded within a microenvironment enriched for SPP1⁺ tumor-associated macrophages (TAMs), fibroblasts, hepatocytes, plasma cells, and myofibroblasts. Furthermore, TME regions were enriched for plasma cells, myofibroblasts, mast cells, erythrocytes, and hepatocyte-like populations, with fewer tumor cell states (**Supplementary Figure 5.2b**). CosMx further resolved additional populations including neuronal-like cells (TME regions), endothelial cells (tumor and TME regions), fibroblast_WNT16 populations, and CAFs within tumor regions that were not identified by STdec markers, indicating unresolved layers of spatial heterogeneity. Region-specific differential expression analysis from Visium data showed that tumor regions were enriched for proliferative, MAPK, and Wnt signaling programs, whereas TME regions were enriched for extracellular matrix remodeling, interferon response, and immune regulatory pathways (**Supplementary Figure 5.3a**). Collectively, these features are consistent with aggressive tumor behavior where a TME-dominated regulatory environment supports tumor plasticity and heterogeneity.

In contrast, non-relapse tumors display a more distributed and immune-engaged spatial organization. These tumors displayed a more integrated and less tumor-dominated architecture, characterized by prominent hepatocyte-associated programs and broader immune engagement. Tumor regions contained cycling tumor cells, hepatocyte-like states, fetal hepatocyte-like

populations, neuronal-like cells, and fibroblast populations including WNT16⁺ and WNT5A⁺ fibroblasts. While hepatocyte regions were enriched for canonical hepatocytes, inflammatory hepatocytes (ALDOB⁺), cycling tumor cells, and fetal hepatocyte-like populations. Several immune and stromal populations not captured by STdec markers including antigen-presenting TAMs, fibroblasts, endothelial cells, and cytotoxic NK/T cells were distributed across regions rather than confined to discrete regions. Additional cell-types and -states, including inflammatory tumor cells (SOX9⁺) and $\gamma\delta$ T cells, were specifically associated with hepatocyte regions and were likewise not identified by STdec markers (**Supplementary Figure 5.2c**). The coexistence of tumor and hepatocyte transcriptional programs within these regions suggests active tumor remodeling toward hepatocyte-like states. Region-specific differential expression analysis from Visium data showed that tumor regions were enriched for vascular-ECM signaling, metabolic and stress-response pathways, and signal transduction programs, whereas TME regions exhibited enrichment of fibroblast/myeloid signaling, immune regulation, and interferon response (**Supplementary Figure 5.3b**). Together, these findings support a model of distributed tumor-immune crosstalk and hepatocyte-associated tumor remodeling, consistent with a more immune-engaged and constrained tumor ecosystem.

Resectable tumors exhibit a more structured and stabilized tumor microenvironment. In contrast to unresectable samples, these tumors represent a spatially organized and constrained architecture, although CosMx profiling depth was limited by technical factors. Resectable samples showed reduced tumor cell-type and state heterogeneity relative to relapse-prone tumors and increased immune diversity compared to non-relapse tumors. Tumor regions contained AFP⁺ fetal hepatocyte-like tumor cells, while fibrotic regions were enriched for fibroblasts and plasma cells. Hepatocyte and TME regions contained inflammatory hepatocytes, myofibroblasts, stress-

associated TAMs, *S100A8/A9* inflammatory myeloid cells, and erythrocytes. Furthermore, antigen-presenting TAMs and endothelial cells were distributed across multiple regions, suggesting active immune response. Additional cell populations including TAMs, tumor progenitor-like epithelial cells, $\gamma\delta$ T cells, and cycling tumor cells were primarily associated with hepatocyte and stroma regions but were not captured by STdec markers (**Supplementary Figure 5.2d**). Region-specific differential expression analysis from Visium data showed that tumor regions were enriched for epithelial proliferation, lipid and cholesterol biosynthesis, and Wnt/ β -catenin signaling, alongside regulators of epithelial identity and transcriptional plasticity, consistent with a metabolically active, Wnt-driven epithelial state with transcriptional reprogramming and adaptive plasticity. In contrast, TME regions were enriched for macrophage differentiation, interferon response, and tissue remodeling programs (**Supplementary Figure 5.3c**). Together with the reduced burden of SNV/CNV-associated DEGs, these findings suggest a tumor ecosystem in which cellular adaptation is driven primarily by transcriptional and microenvironmental programs rather than genomic instability. Collectively, these findings define three distinct spatial organizational states: a TME-dominated, centrally regulated architecture in relapse-prone tumors, a distributed and immune-engaged architecture in non-relapse tumors, and a stabilized, constrained architecture in resection samples.

Spatial proximity analysis defines functionally distinct tumor neighborhoods across hepatoblastoma clinical groups

To define higher-order spatial organization beyond cell-type composition, we analyzed cell-cell co-localization using Delaunay triangulation-based networks. This approach identified recurrent cellular neighborhoods across clinical groups, defined by spatial proximity of specific cell types. To characterize the functional programs associated with these neighborhoods, we

performed pathway enrichment analysis on differentially expressed genes ($\text{FDR} \leq 0.01$, $\log_2\text{FC} \geq 0.05$) of co-localized cell types.

Relapse-prone tumors are organized around tightly coupled tumor-TME-immune architectures. In these tumor regions, spatial co-localization analysis revealed a dominant cellular neighborhood in which tumor cells (progenitor-like and stress-responsive cell states), fibroblasts, myofibroblasts, and macrophages (SPP1⁺ TAMs and TAMs) co-localized, forming a TME-dominated regulatory environment. Within this neighborhood, tumor cells were enriched for hypoxia-response and metabolic reprogramming, consistent with adaptive stress states, while TME populations contributed to extracellular matrix remodeling, migration, and Wnt/MAPK activity. Macrophages exhibited inflammatory and antigen presentation signatures, and co-localized broadly with multiple cell-types and states, consistent with a hub-like regulatory role in coordinating TME-tumor interactions. In TME regions, co-localization of myofibroblasts, neuronal-like, and macrophages defined a secondary neighborhood enriched for extracellular matrix organization and inflammatory signaling, indicating compartmentalized TME remodeling supporting tumor progression. These spatially defined neighborhoods further resolve the cellular basis of the TME-dominated regulatory environment. Collectively, these findings support a model in which relapse-prone tumors are organized around a macrophage-fibroblast regulatory circuit that promotes hypoxia-adaptive tumor plasticity and therapy-resistant states (**Figure 5.3a**).

Non-relapse tumors exhibit compartmentalized and functionally distinct cellular neighborhoods. These tumors revealed multiple co-existing neighborhoods, reflecting localized tumor-immune crosstalk and compartmentalized microenvironmental regulation. In tumor regions, two functionally distinct neighborhoods were observed: (i) cycling tumor cell states co-localized with fetal hepatocyte-like states, forming a metabolically regulated cellular neighborhood, and (ii)

cytotoxic NK/T cells co-localized with fibroblast, defining a pro-inflammatory, ECM remodeling neighborhood consistent with immune-engagement. Similarly, TME regions were organized into distinct cellular neighborhoods. Endothelial cells co-localized with inflammatory hepatocyte, forming an angiogenic and metabolically active neighborhood, while fibroblast co-localized with inflammatory epithelium and cytotoxic NK/T cells, exhibiting stress-response and inflammatory signaling (**Figure 5.3b**). This spatial organization parallels the underlying transcriptional programs and highlights localized tumor-immune crosstalk and compartmentalized microenvironmental regulation.

Resection tumors display more uniform and recurrent spatial interactions. In contrast to unresectable tumors, resection tumors represent a distinct spatial organization characterized by reduced tumor heterogeneity and increased uniformity of cellular interactions. Resection samples display homogeneous and recurrent cellular neighborhoods, characterized by consistent co-localization of TME and immune populations across tissue regions. Dominant cellular neighborhoods involve fibroblasts, antigen-presenting macrophages, and tumor progenitor-like cells, alongside adaptive immune populations, indicating a structured and stabilized tumor-microenvironment architecture (**Figure 5.3c**). This recurrent pattern reflects the coordinated transcriptional and microenvironmental programs observed at the regional level, indicating a spatially structured and functionally coordinated TME-immune network with reduced heterogeneity of tumor-associated cell states.

Distinct cell-cell signaling programs underlie these spatially organized cellular neighborhoods. To define the functional communication underlying these cellular neighborhoods, we performed ligand-receptor interaction analysis. Ligand-receptor analysis revealed distinct dominant signaling programs across clinical groups, with Midkine signaling enriched in relapse-

prone tumors, primarily involving interactions between endothelial cells, CAFs, and AFP⁺ fetal hepatocyte-like tumor cell states, whereas collagen-mediated signaling predominated in non-relapse and resection samples, involving fibroblasts, myofibroblasts, and immune populations **(Figure 5.4a)**. Collectively, these findings demonstrate that spatial proximity and functional communication represent complementary layers of tumor organization, together defining the regulatory architecture of hepatoblastoma across distinct clinical contexts.

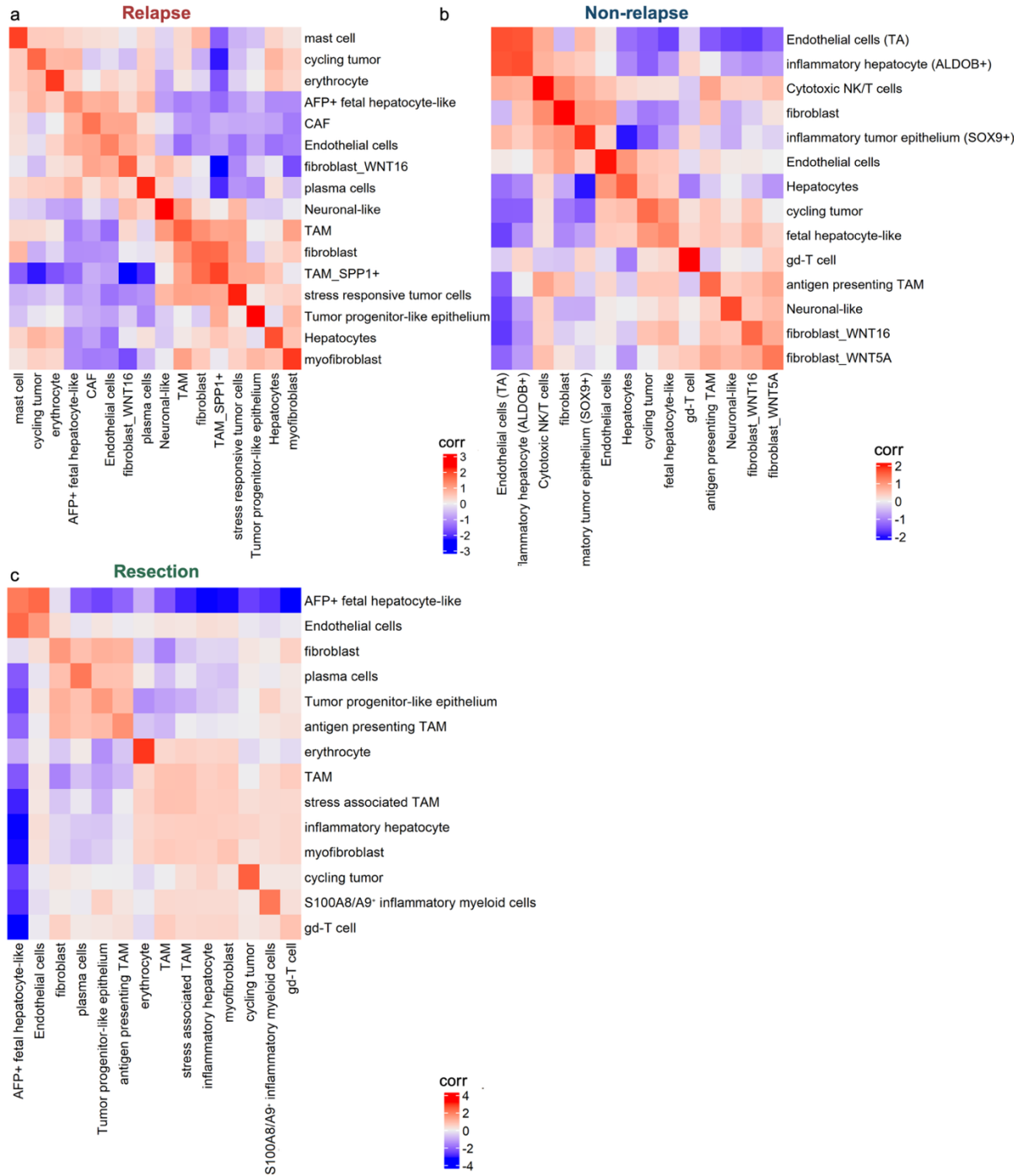


Figure 5.4: Spatial neighborhood enrichment analysis identifies cell-state co-localization patterns across hepatoblastoma clinical groups.

a-c, Heatmaps showing pairwise cell proximity enrichment scores computed from CosMx single-cell spatial data using Delaunay triangulation and permutation-based testing ($n = 1000$ simulations with FDR adjustment) for **a**, relapse **b**, non-relapse, and **c**, resection samples. Rows and columns represent annotated cell states. Positive values (red) denote cell-state pairs found in spatial proximity more often than expected by chance, whereas negative values (blue) denote spatial separation.

Validation of relapse-associated tumor progenitor signature across independent cohort

We next assessed whether relapse-associated spatial niche translate into a clinically relevant gene signature. To validate this, we derived a gene signature capturing the tumor-intrinsic transcriptional state of the relapse-associated microenvironment. To identify the optimal cellular source, we first compared the discriminative performance of DEG sets from each tumor niche relapse-associated cell types in the discovery cohort. Tumor progenitor-like DEGs showed the strongest discrimination (effect size = 0.569, reflecting magnitude of separation between groups, $p = 0.073$), while neuronal-like, fibroblast, and stress-responsive DEG sets showed no signal (effect size ≤ 0.071 , $p \geq 0.909$; **Figure 5.4b**), indicating tumor progenitor-like cells as the signature source. Consistent with this, liver progenitor populations have been reported to expand following neoadjuvant chemotherapy and are enriched in relapse and metastatic settings (30).

The final signature was constructed by integrating tumor progenitor DEGs, pathway-enriched genes from EnrichR analysis, and ligand-receptor interaction genes from high-confidence cell-cell communication analysis ($p < 0.01$, top 90th percentile), particularly among CAFs and AFP⁺ fetal hepatocyte-like cell states. This integration improved discrimination beyond DEGs alone, yielding a 68-gene signature with effect size 0.640 ($p = 0.036$; **Figure 5.4a**). Six genes including *FGFR*, *FGFR4*, *CXCL12*, *AGTR1*, *LRP1*, and *LAMC1* were supported by both differential expression and active intercellular signaling, representing a high-confidence core. In the discovery cohort ($n = 11$, Visium), the signature was significantly enriched in relapse versus non-relapse patients ($p = 0.036$, median effect = 0.163), with both relapse patients scoring above the non-relapse median (**Figure 5.4c**). Spatial visualization revealed focal enrichment at tumor-TME boundaries (**Figure 5.4d**). Applied to an independent external cohort (31) (GSE261958, CytAssist FFPE, $n = 7$), patients with poor outcome scored consistently higher than survivors ($p =$

0.057, median effect = 0.063), supporting directional consistency across independent platforms and cohorts (**Figure 5.4c**).

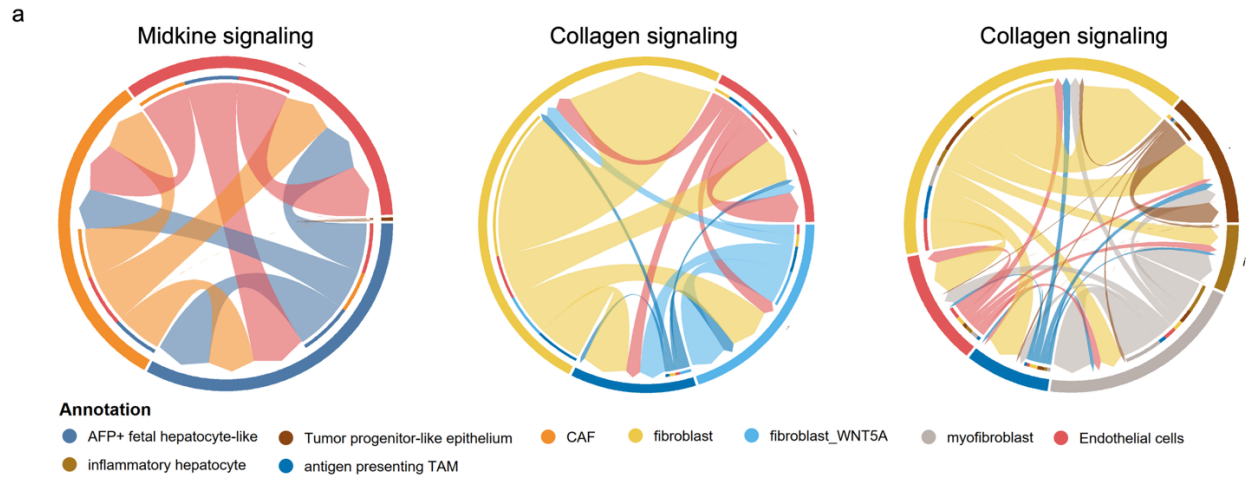


Figure 5.5: Spatially resolved signaling programs define relapse-associated tumor architecture and prognostic signatures.

a, Ligand-receptor interaction analysis of spatially defined cellular neighborhoods reveals distinct dominant signaling programs across clinical groups. Relapse-prone tumors are characterized by enriched Midkine signaling, primarily involving interactions between endothelial cells, cancer-associated fibroblasts (CAFs), and AFP⁺ fetal hepatocyte-like and tumor progenitor-like epithelial states. In contrast, non-relapse and resection samples exhibit dominant collagen-mediated signaling, driven by fibroblast, myofibroblast, and immune-associated interactions. Networks represent aggregated interaction strengths inferred using CellChat.

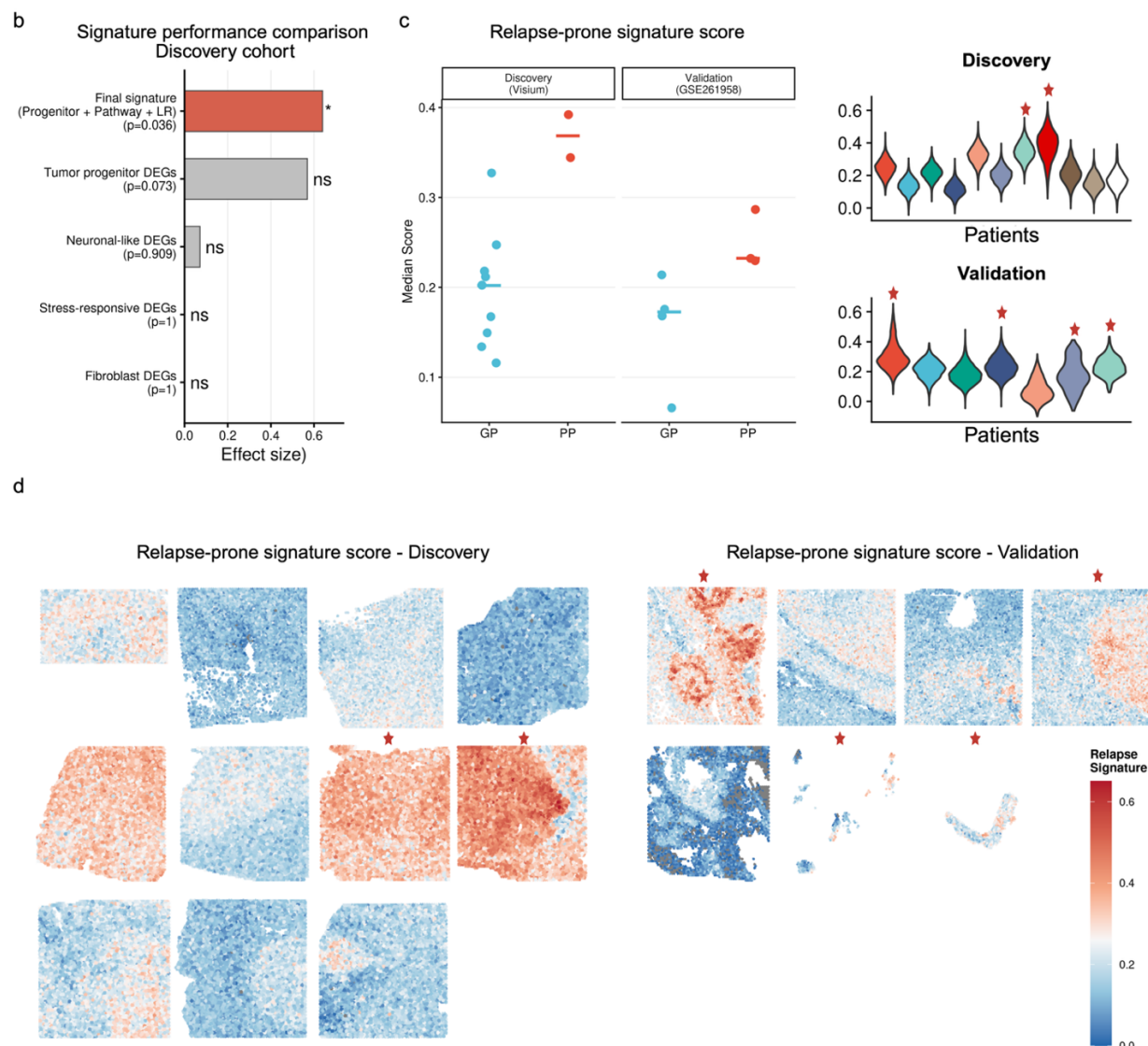


Figure 5.6: Spatially resolved signaling programs define relapse-associated tumor architecture and prognostic signatures, Continued.

b, Comparative performance of candidate gene signature for relapse prognostic in the discovery cohort. The final composite signature (tumor progenitor + pathway + LR features) demonstrates improved effect size relative to individual gene sets, indicating that integration of spatial and signaling features enhances prognostic power. **c**, Relapse-prone signature scores across discovery (Visium) and validation cohorts. Relapse-associated samples exhibited significantly higher signature scores compared to non-relapse samples, demonstrating robustness across independent datasets. Violin plots show per-patient distribution, red asterisks indicate relapse samples. **d**, Spatial distribution of relapse-prone signature scores tissue sections in discovery and validation cohorts. High-scoring regions localize to distinct spatial niches, red asterisks indicate relapse samples.

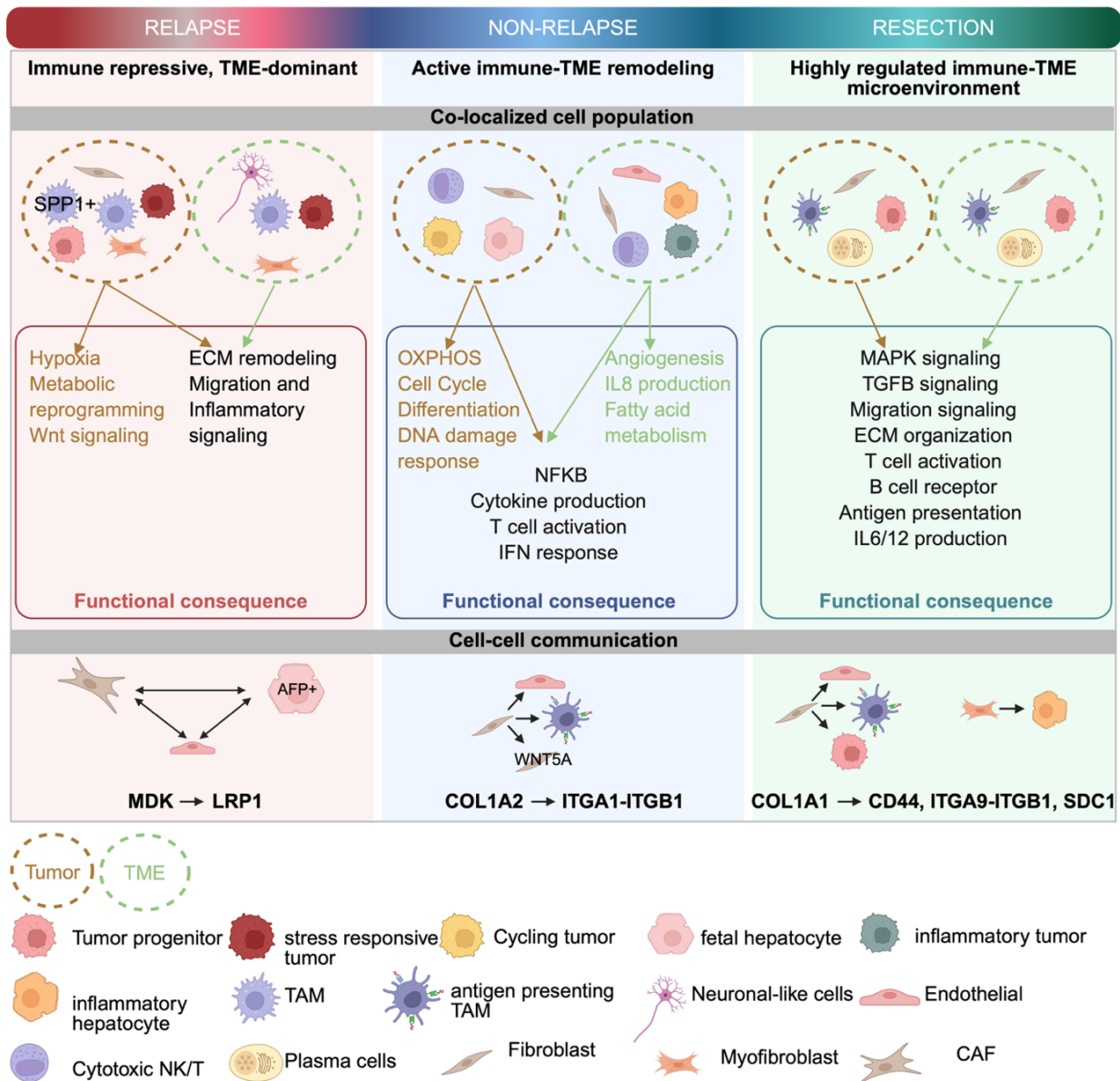


Figure 5.7: Distinct spatial architectures and signaling programs define relapse-associated and non-relapse hepatoblastoma microenvironments.

Schematic representation of spatially resolved tumor ecosystems across clinical groups. Relapse-prone tumors exhibit an immune-repressive, TME-dominant microenvironment characterized by co-localization of tumor progenitor-like cells with CAFs and TAMs, associated with hypoxia, metabolic reprogramming, WNT signaling, and Midkine-mediated interactions. Non-relapse tumors display active immune-TME remodeling with coordinated tumor, immune, and stromal interactions enriched for oxidative phosphorylation, cell cycle activity, inflammatory signaling, and collagen-mediated communication. Resection samples exhibit a highly regulated immune-TME microenvironment with organized immune-stromal interactions enriched for MAPK and TGFb signaling, ECM organization, adaptive immune responses, and collagen-driven signaling.

5.5 Discussion

Therapeutic response in hepatoblastoma is reflected in the spatial architecture of the tumor ecosystem rather than cellular composition alone. Using integrated regional and single-cell spatial transcriptomics, we demonstrate that post-chemotherapy hepatoblastoma is reorganized into distinct regulatory architectures that differ systematically across clinical outcomes. These architectures represent therapy-remodeled tumor ecosystems in which tumor plasticity, TME organization, and immune engagement are coordinately restructured. These patterns can be broadly categorized into constrained (resection), distributed (non-relapse), and centralized (relapse) architectures.

Myeloid cells are known to play a central role in hepatoblastoma and other solid tumors, particularly in shaping tumor-stroma interactions and therapeutic response (13,30,32). A defining feature of the relapse-prone microenvironment in our study is the spatial enrichment of SPP1⁺ TAMs within a TME-dominated neighborhood co-localizing with CAFs and AFP⁺ fetal hepatocyte-like tumor cells. SPP1⁺ TAMs have been broadly associated with poor prognosis and linked to hypoxia, TGF β signaling, and immune exclusion, with CXCL9:SPP1 macrophage polarization emerging as a key determinant of outcome across cancers (33–35). However, prior studies have primarily focused on macrophage abundance and transcriptional state. Here, we show that SPP1⁺ TAMs are centrally positioned within TME, where they co-localize with fibroblasts and tumor cells and coordinate their interactions following chemotherapy. Within this architecture, TME populations, including stromal and endothelial populations show strong interactions with AFP⁺ tumor cell states, driven in part by Midkine signaling via MDK-LRP1 ligand-receptor interaction, consistent with coordinated regulation of tumor plasticity and immune evasion (31). Beyond Midkine signaling, the relapse-associated transcriptional program is enriched for

pathways linked to chemokine signaling, extracellular matrix remodeling, angiogenesis, and receptor tyrosine kinase activation, as reflected by genes such as *CXCL12*, *LAMC1*, *AGTR1*, *LRP1*, and *FGFR4*. Notably, *FGFR4* has been associated with proliferative, chemotherapy-resistant hepatoblastoma tumor cell states and is enhanced through fibroblast-tumor signaling interactions, supporting a role for FGF-FGFR signaling as a stromal-reinforced growth axis in this spatial neighborhood (15,36–39). These findings align with prior work demonstrating roles for SPP1⁺ TAMs in promoting tumor stemness, metabolic dysregulation, and immune suppression (40), as well as TAM-CAF interactions in desmoplastic remodeling (41). Importantly, our data extend these observations to hepatoblastoma and demonstrate that these processes are spatially organized, highlighting a coordinated stromal-immune regulatory architecture associated with therapy-resistant tumor states.

Hepatoblastoma prognosis is closely linked to its inherent histological heterogeneity, with specific epithelial cell states—including high-risk small cell undifferentiated populations—defining clinical behavior (41,42). Correspondingly, bulk transcriptomic analysis reveals that epithelial gene expression programs drive tumor risk stratification (9). Our study defines a distinct, context-dependent landscape of hepatoblastoma cell states: relapse-prone tumors are enriched for progenitor-like, cycling, and stress-responsive cells, while non-relapse cases exhibit more differentiated, inflammatory epithelial states. Furthermore, resection samples show reduced heterogeneity with a predominance of fetal hepatocyte-like and progenitor-like states. Spatial analysis reveals that the tumor microenvironment (TME), specifically immune and stromal components, actively shapes these tumor cell compositions. Non-relapse tumors maintain a distinct differentiation hierarchy, comprising inflammatory hepatocyte-like cells and inflammatory tumor progenitors. While previous studies show that treatment shifts hepatoblastoma toward hepatocytic

differentiation, specifically enriching HNF1A/HNF4A programs (29), our data indicates that these inflammatory hepatocyte-like states also harbor mature signatures alongside acute-phase, stress-adaptive responses. These states localize to regions with treatment-associated necrosis and fibrosis (**Supplementary Table 5.2**). Conversely, inflammatory tumor progenitor-like populations retain developmental signaling programs, such as Wnt and Notch pathway activity, highlighting the persistence of regulated tumor plasticity. Spatially, these epithelial states exist within microenvironments enriched for antigen-presenting macrophages, cytotoxic lymphocytes, and stromal populations.

Limitations of this study include a small, predominantly post-treatment sample size, which may bias results toward chemo-resistant states and underrepresent pre-treatment heterogeneity. While our spatial transcriptomics approach offers multi-scale insights, further studies with large cohorts and functional validation will be necessary to confirm the causal role of these architectures in driving therapeutic resistance. Looking ahead, advanced single-nucleus profiling will allow for broader analysis. In conclusion, we show that post-chemotherapy hepatoblastoma is organized into distinct spatial regulatory states link to clinical outcome. These findings position spatially organized tumor-microenvironment interactions as central determinants of therapeutic response and highlight microenvironmental signaling networks as targets of overcoming treatment resistance.

This Chapter is being prepared for publication. Please refer to the online manuscript for detailed Supplementary Materials.

5.6 Supplemental Materials

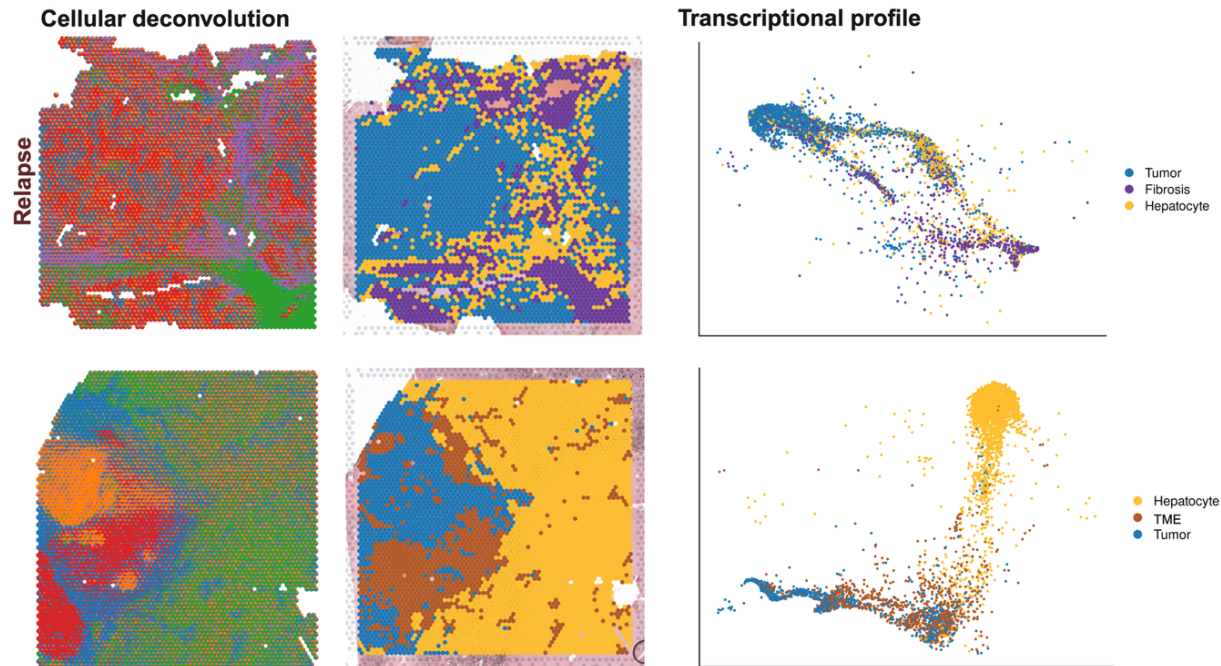
Supplementary Table 5.1: Patient demographics.

Sample ID	Cos Mx	Gender	Race	Liver Transplant (LT) or resection (R)	Age at LT or R (years)	Relapse status	Alive after LT or R	Time between transplant (Tx) and relapse (Days)	Time between relapse and death (Days)	Chemo/Roadmap
HB_3781		M	W	LT	3.1	No	147			C5V
HB_4433		M	W	LT	2.57	No	158			C5VD
HB_7859		M	W	LT	4.18	No	95			VIT, C5VD
HB_1895	Yes	M	W	LT	3.01	No	149			CD
HB_600		F	W	LT	0.99	No	162			C5V
HB_838		M	W	LT	6.45	No	174			C5VD
HB_4942	Yes	M	W	LT	3.75	Yes		57	135	CD, VI, CaE
HB_1364		M	W	R	2.13	No	29			CD, CaE
HB_9714		M	W	LT	6.1	Yes		61	15	VIT, C5VD
HB_6488		M	W	LT	1.4	No	36			VIT, C5VD
HB_7826	Yes	M	LATIN	R	3.66	No	34			C5VD, VI
HB_8277		F	W	LT	2.33	No	34			C5VD

C – cisplatin, D – doxorubicin, 5 - 5-Fluorouracil, I – Irinotecan, V – Vincristine, T – temsirolimus, Ca – Carboplatin, E – Etoposide, If - ifosfamide

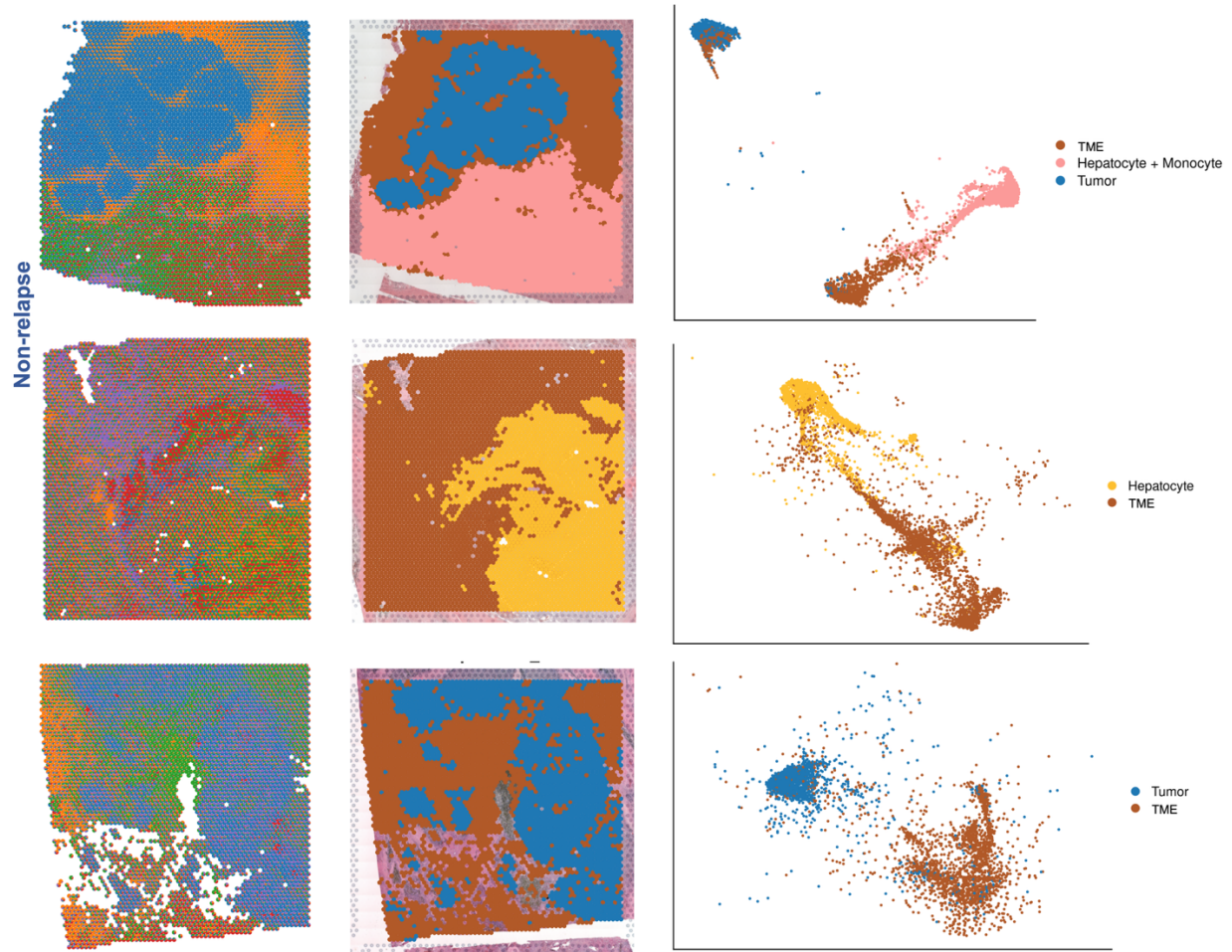
Supplementary Table 5.2: Tumor-level clinical characteristics

Sample ID	Histology (summary)	PRETEXT stage
HB_3781	Mixed fetal/embryonal/mesenchymal	NA
HB_4433	Fetal/embryonal + bone	III
HB_7859	Fetal (crowded)	IV
HB_1895	Mixed epithelial	I
HB_600	Fetal	II
HB_838	Epithelial, predominantly fetal with small foci of crowded fetal	III
HB_4942	Multinodular epithelial	IV
HB_1364	Not specified	III
HB_9714	HB + HCC-like features	III
HB_6488	Crowded fetal (predominant)	II
HB_7826	Crowded/pleomorphic fetal	IV
HB_8277	Mixed epithelial–mesenchymal	III



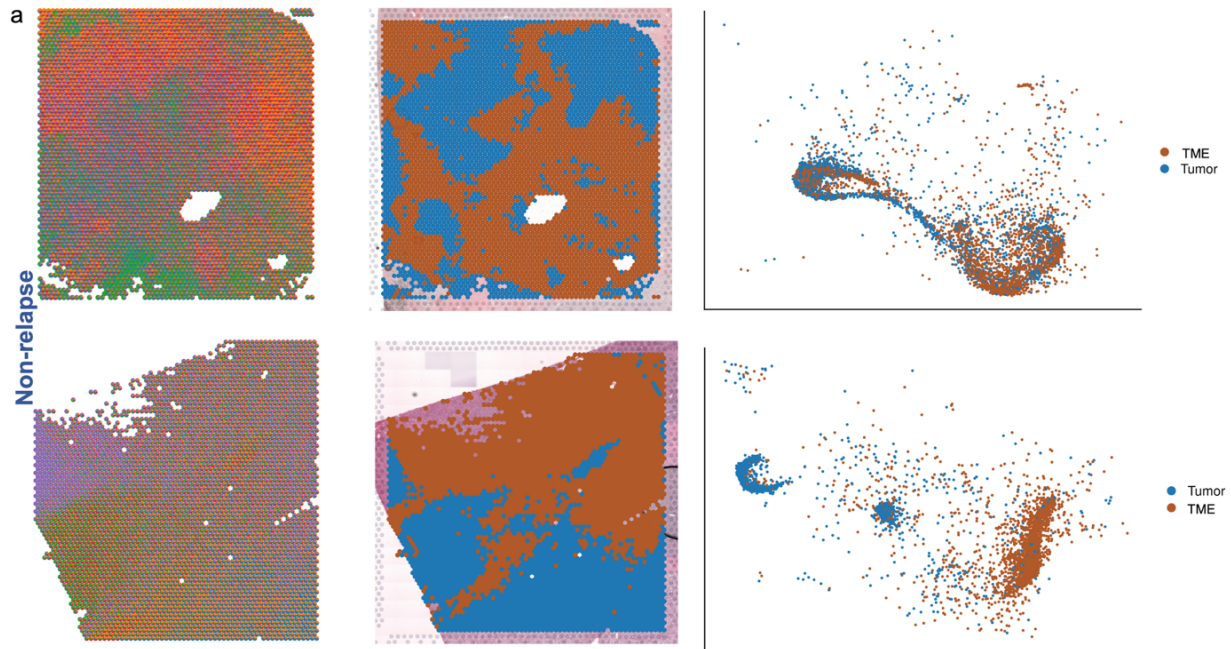
Supplementary Figure 5.1: Spatial deconvolution reveals distinct cellular organization and transcriptional structure across clinical groups.

Left: Visium spatial maps of STdec topics ($K=5$) for representative sections from relapse and non-relapse samples, illustrating region-level compartmentalization of tumor and microenvironmental programs. Middle: Corresponding region-level annotations derived from Seurat cluster profiling and marker-based interpretation, delineating tumor, stroma, hepatocyte-like, and inflammatory compartments across clinical groups. Right: Low-dimensional embedding of transcriptional profiles for the same samples, colored by dominant cell-type programs (tumor, stromal/fibrotic, hepatocyte-like). Relapse sample showed overlapping and less resolved transcriptional states, consistent with increased heterogeneity and cellular plasticity. In contrast, non-relapse samples exhibit more distinct clustering of cell states, reflecting structured cellular identities and compartmentalization.



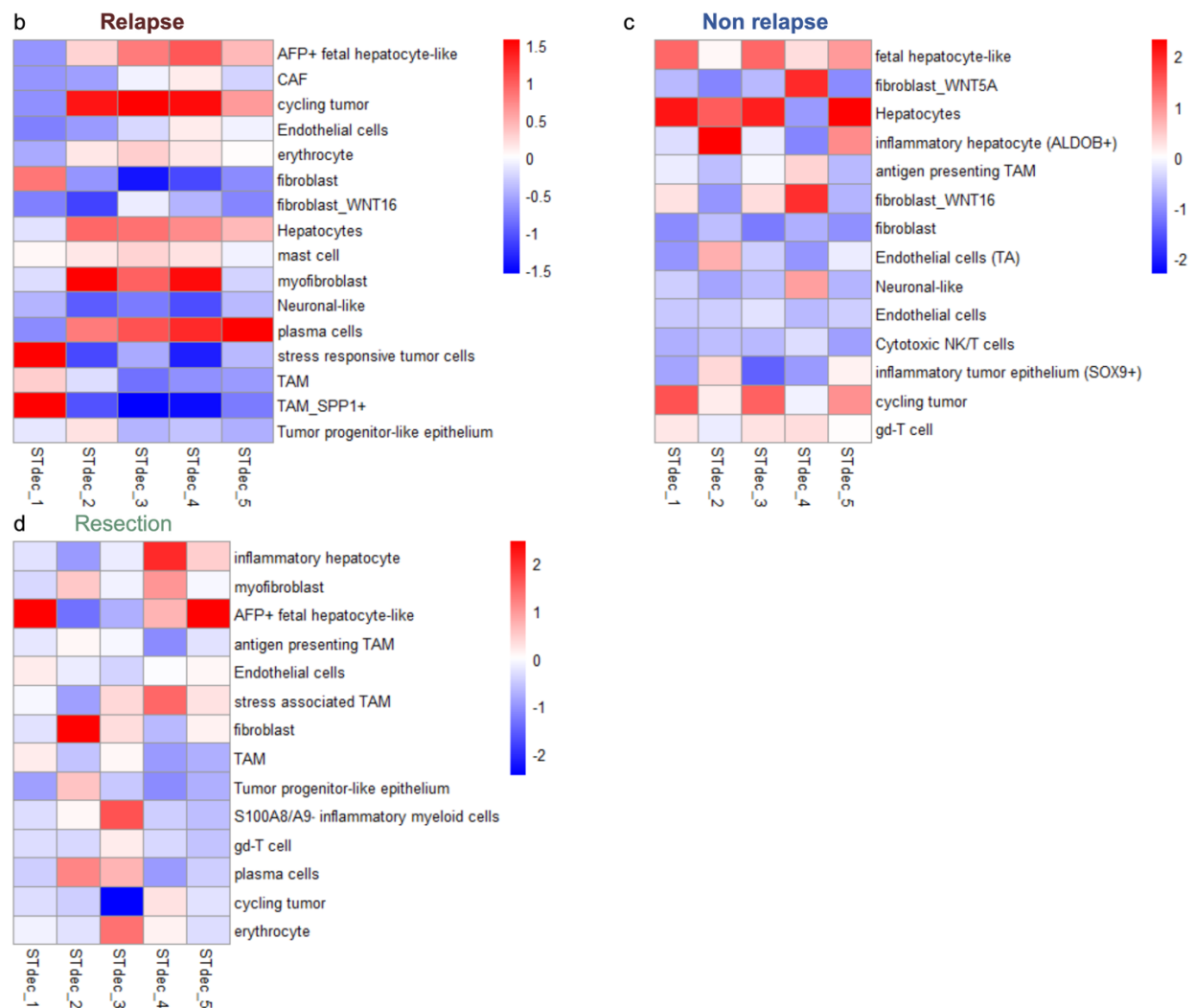
Supplementary Figure 5.2: Spatial deconvolution reveals distinct cellular organization and transcriptional structure across clinical groups, Continued.

Left: Visium spatial maps of STdec topics ($K=5$) for representative sections from relapse and non-relapse samples, illustrating region-level compartmentalization of tumor and microenvironmental programs. Middle: Corresponding region-level annotations derived from Seurat cluster profiling and marker-based interpretation, delineating tumor, stroma, hepatocyte-like, and inflammatory compartments across clinical groups. Right: Low-dimensional embedding of transcriptional profiles for the same samples, colored by dominant cell-type programs (tumor, stromal/fibrotic, hepatocyte-like). Relapse sample showed overlapping and less resolved transcriptional states, consistent with increased heterogeneity and cellular plasticity. In contrast, non-relapse samples exhibit more distinct clustering of cell states, reflecting structured cellular identities and compartmentalization.



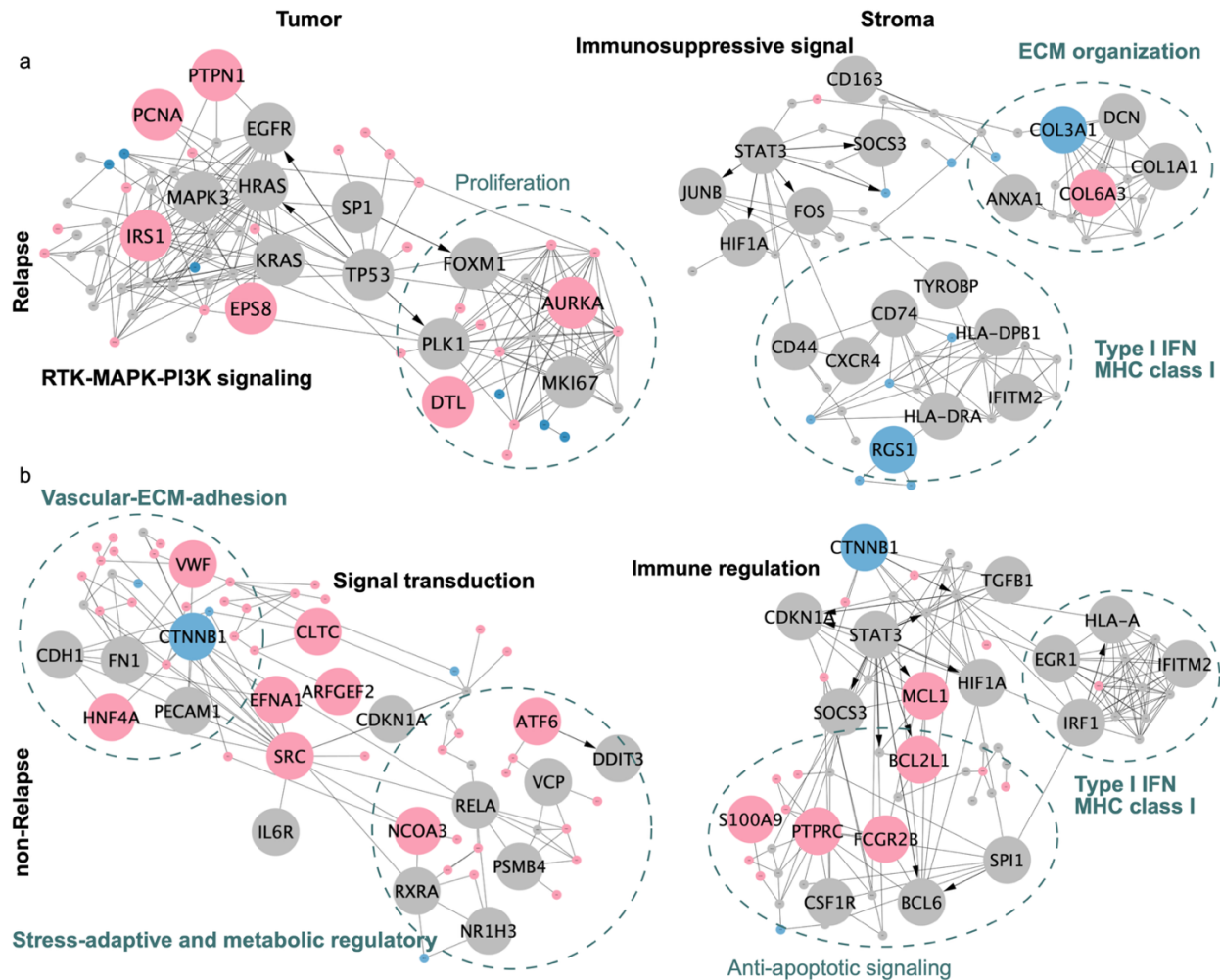
Supplementary Figure 5.3: Additional spatial and cellular characterization across hepatoblastoma clinical groups.

a, continued from **Supplementary Figure 5.1. b-d**, Cross-platform mapping of Visium-derived STdeconvolve topics onto CosMx-defined cell-types for relapse **b**, non-relapse **c**, and resection **d** samples. Heatmaps show enrichment scores obtained by projecting STdeconvolve topic signatures onto single-cell CosMx profiles, linking spatially defined transcriptional programs to underlying cellular identities. Values represent scaled enrichment, highlighting associations with tumor, stromal, hepatocyte-like, and immune cell populations.



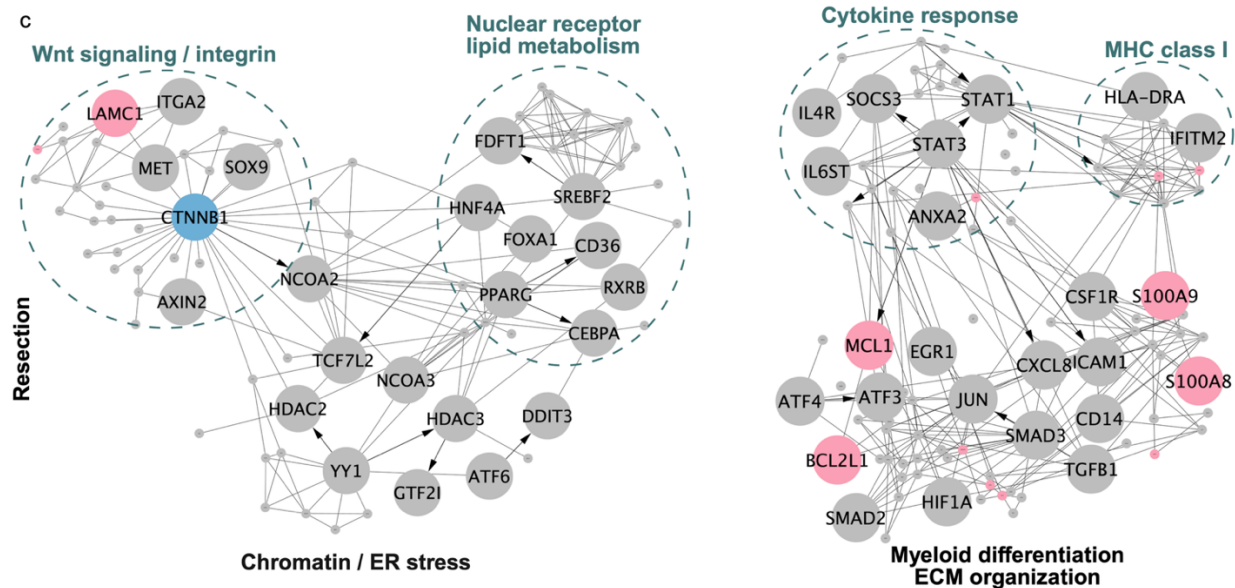
Supplementary Figure 5.4: Additional spatial and cellular characterization across hepatoblastoma clinical groups, Continued.

b-d, Cross-platform mapping of Visium-derived STdeconvolve topics onto CosMx-defined cell-types for relapse **b**, non-relapse **c**, and resection **d** samples. Heatmaps show enrichment scores obtained by projecting STdeconvolve topic signatures onto single-cell CosMx profiles, linking spatially defined transcriptional programs to underlying cellular identities. Values represent scaled enrichment, highlighting associations with tumor, stromal, hepatocyte-like, and immune cell populations.



Supplementary Figure 5.5: Distinct tumor and TME regulatory networks define clinical group-specific microenvironmental programs.

a-c, Protein-protein interaction (PPI) networks constructed from differentially expressed genes in tumor (left) and TME (right) compartments for relapse **a**, non-relapse **b**, and resection **c** samples. Nodes represent genes and edges indicate high-confidence interactions; highlighted clusters denote functionally enriched modules. In relapse **a**, tumor networks are dominated by RTK-MAPK-PI3K signaling and proliferative programs, while TME networks exhibit immunosuppressive signaling, extracellular matrix (ECM) organization, and Type I interferon/MHC class I associated modules. In non-relapse **b**, tumor networks are enriched for vascular-ECM adhesion and stress-adaptive signaling, whereas TME networks display immune regulatory, anti-apoptotic, and cytokine response programs, alongside Type I IFN/MHC class I activity.



Supplementary Figure 5.6: Distinct tumor and TME regulatory networks define clinical group-specific microenvironmental programs, Continued.

a-c, Protein-protein interaction (PPI) networks constructed from differentially expressed genes in tumor (left) and TME (right) compartments for relapse **a**, non-relapse **b**, and resection **c** samples. Nodes represent genes and edges indicate high-confidence interactions; highlighted clusters denote functionally enriched modules. In resection **c**, tumor networks highlight Wnt/integrin signaling, lipid metabolism, and chromatin/ER stress pathways, while TME networks show coordinated cytokine signaling, MHC class I activity, and myeloid differentiation coupled with ECM organization. Pink and blue nodes represent SNV/CNV.

REFERENCES

1. Ranganathan S, Lopez-Terrada D, Alaggio R. Hepatoblastoma and Pediatric Hepatocellular Carcinoma: An Update. *Pediatr Dev Pathol.* 2020 Apr 1;23(2):79–95. doi:10.1177/1093526619875228
2. Pio L, O'Neill AF, Woodley H, Murphy AJ, Tiao G, Franchi-Abella S, Fresneau B, Watanabe K, Alaggio R, Lopez-Terrada D, Hiyama E, Branchereau S. Hepatoblastoma. *Nat Rev Dis Primers.* 2025 May 22;11(1):36. doi:10.1038/s41572-025-00620-7
3. Dasgupta P, Henshaw C, Youlden DR, Aitken JF, Sullivan A, Irving H, Baade PD. Global trends in incidence rates of childhood liver cancers: A systematic review and meta-analysis. *Paediatric and Perinatal Epidemiology.* 2020;34(5):609–17. doi:10.1111/ppe.12671
4. Espinoza AF, Patel KR, Shetty PB, Whitlock RS, Sumazin P, Yu X, Sarabia SF, Urbicain M, Heczey A, Masand P, Woodfield SE, López-Terrada DH, Vasudevan SA. Navigating relapsed hepatoblastoma: Predictive factors and surgical treatment strategy. *Cancer Medicine.* 2023;12(23):21270–8. doi:10.1002/cam4.6705
5. Semeraro M, Branchereau S, Maibach R, Zsiros J, Casanova M, Brock P, Domerg C, Aronson DC, Zimmermann A, Laithier V, Childs M, Roebuck D, Perilongo G, Czauderna P, Brugieres L. Relapses in hepatoblastoma patients: Clinical characteristics and outcome – Experience of the International Childhood Liver Tumour Strategy Group (SIOPEL). *European Journal of Cancer.* 2013 Mar 1;49(4):915–22. doi:10.1016/j.ejca.2012.10.003
6. Gröbner SN, Worst BC, Weischenfeldt J, Buchhalter I, Kleinheinz K, Rudneva VA, Johann PD, Balasubramanian GP, Segura-Wang M, Brabetz S, Bender S, Hutter B, Sturm D, Pfaff E, Hübschmann D, Zipprich G, Heinold M, Eils J, Lawrenz C, Erkek S, Lambo S, Waszak S, Blattmann C, Borkhardt A, Kuhlen M, Eggert A, Fulda S, Gessler M, Wegert J, Kappler R, Baumhoer D, Burdach S, Kirschner-Schwabe R, Kontny U, Kulozik AE, Lohmann D, Hettmer S, Eckert C, Bielack S, Nathrath M, Niemeyer C, Richter GH, Schulte J, Siebert R, Westermann F, Molenaar JJ, Vassal G, Witt H, Burkhardt B, Kratz CP, Witt O, van Tilburg CM, Kramm CM, Fleischhack G, Dirksen U, Rutkowski S, Frühwald M, von Hoff K, Wolf S, Klingebiel T, Koscielniak E, Landgraf P, Koster J, Resnick AC, Zhang J, Liu Y, Zhou X, Waanders AJ, Zwijnenburg DA, Raman P, Brors B, Weber UD, Northcott PA, Pajtler KW, Kool M, Piro RM, Korbel JO, Matthias Schlesner, Eils R, Jones DTW, Lichter P, Chavez L, Zapatka M, Pfister SM. The landscape of genomic alterations across childhood cancers. *Nature.* 2018 Mar;555(7696):321–7. doi:10.1038/nature25480
7. Eichenmüller M, Trippel F, Kreuder M, Beck A, Schwarzmayer T, Häberle B, Cairo S, Leuschner I, von Schweinitz D, Strom TM, Kappler R. The genomic landscape of hepatoblastoma and their progenies with HCC-like features. *Journal of Hepatology.* 2014 Dec 1;61(6):1312–20. doi:10.1016/j.jhep.2014.08.009
8. Kanawa M, Hiyama E, Kawashima K, Hiyama K, Ikeda K, Morihara N, Kurihara S, Fukazawa T, Ueda Y. Gene expression profiling in hepatoblastoma cases of the Japanese Study Group for Pediatric Liver Tumors-2 (JPLT-2) trial. *European journal of Molecular Cancer.* 2019 Feb 12;2018(1):1–8. doi:10.31487/j.EJMC.2018.01.003

9. Cairo S, Armengol C, De Reyniès A, Wei Y, Thomas E, Renard CA, Goga A, Balakrishnan A, Semeraro M, Gresh L, Pontoglio M, Strick-Marchand H, Levillayer F, Nouet Y, Rickman D, Gauthier F, Branchereau S, Brugières L, Laithier V, Bouvier R, Boman F, Basso G, Michiels JF, Hofman P, Arbez-Gindre F, Jouan H, Rousselet-Chapeau MC, Berrebi D, Marcellin L, Plenat F, Zachar D, Joubert M, Selves J, Pasquier D, Bioulac-Sage P, Grotzer M, Childs M, Fabre M, Buendia MA. Hepatic stem-like phenotype and interplay of Wnt/beta-catenin and Myc signaling in aggressive childhood liver cancer. *Cancer Cell*. 2008 Dec 9;14(6):471–84. doi:10.1016/j.ccr.2008.11.002
10. Sumazin P, Chen Y, Treviño LR, Sarabia SF, Hampton OA, Patel K, Mistretta TA, Zorman B, Thompson P, Heczey A, Comerford S, Wheeler DA, Chintagumpala M, Meyers R, Rakheja D, Finegold MJ, Tomlinson G, Parsons DW, López-Terrada D. Genomic analysis of hepatoblastoma identifies distinct molecular and prognostic subgroups. *Hepatology*. 2017 Jan 1;65(1):104–21. doi:10.1002/hep.28888
11. Hirsch TZ, Pilet J, Morcrette G, Roehrig A, Monteiro BJE, Molina L, Bayard Q, Trépo E, Meunier L, Caruso S, Renault V, Deleuze JF, Fresneau B, Chardot C, Gonzales E, Jacquemin E, Guerin F, Fabre M, Aerts I, Taque S, Laithier V, Branchereau S, Guettier C, Brugières L, Rebouissou S, Letouzé E, Zucman-Rossi J. Integrated Genomic Analysis Identifies Driver Genes and Cisplatin-Resistant Progenitor Phenotype in Pediatric Liver Cancer. *Cancer Discov*. 2021 Oct 1;11(10):2524–43. doi:10.1158/2159-8290.CD-20-1809
12. Kluiver TA, Lu Y, Schubert SA, Kraaijer LJ, Ringnalda F, Lijnzaad P, DeMartino J, Megchelenbrink WL, Amo-Addae V, Eising S, de Faria FW, Münter D, van de Wetering M, Kerl K, Duiker E, van den Heuvel MC, de Meijer VE, de Kleine RH, Molenaar JJ, Margaritis T, Stunnenberg HG, de Krijger RR, Zsiros J, Clevers H, Peng WC. Divergent WNT signaling and drug sensitivity profiles within hepatoblastoma tumors and organoids. *Nat Commun*. 2024 Nov 20;15(1):8576. doi:10.1038/s41467-024-52757-w
13. Wang Y, Xiang X, Chen H, Zhou L, Chen S, Zhang G, Liu X, Ren X, Liu J, Kuang M, Jiang J, She J, Zhang Z, Xue R, Jiang H, Wang J, Peng S. Intratumoral erythroblastic islands restrain anti-tumor immunity in hepatoblastoma. *CR Med*. 2023 May 16;4(5). doi:10.1016/j.xcrm.2023.101044
14. Tang R, Qi SQ, Zhang T, Pan ZB, Xu JH. Progress in the study of therapeutic strategies for hepatoblastoma in children. *World J Gastrointest Oncol*. 2025 Jun 15;17(6):107700. doi:10.4251/wjgo.v17.i6.107700
15. Chen D, Zhan Y, Zhou E, Chen H, Dai S, Yang R, Ye C, Wang J, Li Y, Lu Y, Fang Y, Chen L, Yao W, Li K, Zheng S, Dong K, Wang J, Dong R. Single-cell profiles delineate immune cell remodeling and enhanced tumor-fibroblast interaction of hepatoblastoma after chemotherapy. *CR Med*. 2025 Oct 21;6(10). doi:10.1016/j.xcrm.2025.102401
16. Ma M, Jin C, Dong Q. Intratumoral Heterogeneity and Immune Microenvironment in Hepatoblastoma Revealed by Single-Cell RNA Sequencing. *J Cell Mol Med*. 2025 Mar 18;29(6):e70482. doi:10.1111/jcmm.70482

17. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, Hoffman P, Stoeckius M, Papalexi E, Mimitou EP, Jain J, Srivastava A, Stuart T, Fleming LM, Yeung B, Rogers AJ, McElrath JM, Blish CA, Gottardo R, Smibert P, Satija R. Integrated analysis of multimodal single-cell data. *Cell*. 2021 Jun 24;184(13):3573-3587.e29. doi:10.1016/j.cell.2021.04.048
18. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol*. 2019 Dec 23;20(1):296. doi:10.1186/s13059-019-1874-1
19. Miller BF, Huang F, Atta L, Sahoo A, Fan J. Reference-free cell type deconvolution of multi-cellular pixel-resolution spatially resolved transcriptomics data. *Nat Commun*. 2022 Apr 29;13(1):2339. doi:10.1038/s41467-022-30033-z
20. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology*. 2012 May;16(5):284–7. doi:10.1089/omi.2011.0118
21. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, Ma'ayan A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 2016 Jul 8;44(W1):W90-97. doi:10.1093/nar/gkw377
22. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering C von. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019 Jan 8;47(D1):D607–13. doi:10.1093/nar/gky1131
23. Han H, Cho JW, Lee S, Yun A, Kim H, Bae D, Yang S, Kim CY, Lee M, Kim E, Lee S, Kang B, Jeong D, Kim Y, Jeon HN, Jung H, Nam S, Chung M, Kim JH, Lee I. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res*. 2018 Jan 4;46(D1):D380–6. doi:10.1093/nar/gkx1013
24. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003 Nov;13(11):2498–504. doi:10.1101/gr.1239303
25. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, Srivastava A, Molla G, Madad S, Fernandez-Granda C, Satija R. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol*. 2024 Feb;42(2):293–304. doi:10.1038/s41587-023-01767-y
26. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015 Dec 23;1(6):417–25. doi:10.1016/j.cels.2015.12.004

27. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis [Internet]. *bioRxiv*; 2021 [cited 2026 Apr 27]. p. 060012. Available from: <https://www.biorxiv.org/content/10.1101/060012v3> doi:10.1101/060012
28. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, Myung P, Plikus MV, Nie Q. Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. 2021 Feb 17;12(1):1088. doi:10.1038/s41467-021-21246-9
29. Dries R, Zhu Q, Dong R, Eng CHL, Li H, Liu K, Fu Y, Zhao T, Sarkar A, Bao F, George RE, Pierson N, Cai L, Yuan GC. Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol*. 2021 Mar 8;22(1):78. doi:10.1186/s13059-021-02286-2
30. Hirsch TZ, Pilet J, Morcrette G, Roehrig A, Monteiro BJ, Molina L, Bayard Q, Trépo E, Meunier L, Caruso S, Renault V, Deleuze JF, Fresneau B, Chardot C, Gonzales E, Jacquemin E, Guerin F, Fabre M, Aerts I, Taque S, Laithier V, Branchereau S, Guettier C, Brugières L, Rebouissou S, Letouzé E, Zucman-Rossi J. Integrated genomic analysis identifies driver genes and cisplatin-resistant progenitor phenotype in pediatric liver cancer. *Cancer Discov*. 2021 Oct;11(10):2524–43. doi:10.1158/2159-8290.CD-20-1809
31. Münter D, de Faria FW, Richter M, Aranda-Pardos I, Hotfilder M, Walter C, Paga E, Inserte C, Albert TK, Roy R, Rahman S, Riedel NC, Müller V, Pascher A, Wiebe K, Schmid I, Vokuhl C, Winkler B, Jüttner E, Vieth S, Mücke U, Kluiver TA, Peng WC, Rossig C, Schlué J, Madadi-Sanjani O, Sandmann S, Hartmann W, A-Gonzalez N, Soehnlein O, Kerl K. Multiomic analysis uncovers a continuous spectrum of differentiation and Wnt-MDK-driven immune evasion in hepatoblastoma. *J Hepatol*. 2025 Aug;83(2):367–82. doi:10.1016/j.jhep.2025.01.031
32. Single-cell analysis of hepatoblastoma identifies tumor signatures that predict chemotherapy susceptibility using patient-specific tumor spheroids | *Nature Communications* [Internet]. [cited 2026 Apr 28]. Available from: <https://www.nature.com/articles/s41467-022-32473-z>
33. Bill R, Wirapati P, Messemaker M, Roh W, Zitti B, Duval F, Kiss M, Park JC, Saal TM, Hoelzl J, Tarussio D, Benedetti F, Tissot S, Kandalaft L, Varrone M, Ciriello G, McKee TA, Monnier Y, Mermoud M, Blaum EM, Gushterova I, Gonye ALK, Hacohen N, Getz G, Mempel TR, Klein AM, Weissleder R, Faquin WC, Sadow PM, Lin D, Pai SI, Sade-Feldman M, Pittet MJ. CXCL9:SPP1 macrophage polarity identifies a network of cellular programs that control human cancers. *Science*. 2023 Aug 4;381(6657):515–24. doi:10.1126/science.ade2292
34. Xu WX, Ye YM, Chen JL, Guo XY, Li C, Luo J, Lu LB, Chen X. SPP1+ tumor-associated macrophages define a high-risk subgroup and inform personalized therapy in hepatocellular carcinoma. *Front Oncol*. 2025 Jul 1;15. doi:10.3389/fonc.2025.1606195
35. Wang J, Wang Y, Liu Y, Yang R. SPP1+ macrophages in tumor immunosuppression: mechanisms and therapeutic implications. *Front Immunol*. 16:1711015. doi:10.3389/fimmu.2025.1711015
36. Ye G, Qin Y, Wang S, Pan D, Xu S, Wu C, Wang X, Wang J, Ye H, Shen H. Lamc1 promotes the Warburg effect in hepatocellular carcinoma cells by regulating PKM2 expression

through AKT pathway. *Cancer Biol Ther.* 2019 Feb 12;20(5):711–9. doi:10.1080/15384047.2018.1564558

37. Overexpression of LAMC1 predicts poor prognosis and enhances tumor cell invasion and migration in hepatocellular carcinoma - PMC [Internet]. [cited 2026 Apr 28]. Available from: <https://pmc.ncbi.nlm.nih.gov/articles/PMC5604451/>

38. Bondoc A, Glaser K, Jin K, Lake C, Cairo S, Geller J, Tiao G, Aronow B. Identification of distinct tumor cell populations and key genetic mechanisms through single cell sequencing in hepatoblastoma. *Commun Biol.* 2021 Sep 8;4(1):1049. doi:10.1038/s42003-021-02562-8

39. Gadaleta RM, Moschetta A. Dark and bright side of targeting fibroblast growth factor receptor 4 in the liver. *Journal of Hepatology.* 2021 Dec 1;75(6):1440–51. doi:10.1016/j.jhep.2021.07.029

40. Wang Y, Wang Q, Tao S, Li H, Zhang X, Xia Y, Wang Y, Yang C, Sui C. Identification of SPP1+ macrophages in promoting cancer stemness via vitronectin and CCL15 signals crosstalk in liver cancer. *Cancer Letters.* 2024 Nov 1;604:217199. doi:10.1016/j.canlet.2024.217199

41. Qi J, Sun H, Zhang Y, Wang Z, Xun Z, Li Z, Ding X, Bao R, Hong L, Jia W, Fang F, Liu H, Chen L, Zhong J, Zou D, Liu L, Han L, Ginhoux F, Liu Y, Ye Y, Su B. Single-cell and spatial analysis reveal interaction of FAP+ fibroblasts and SPP1+ macrophages in colorectal cancer. *Nat Commun.* 2022 Apr 1;13(1):1742. doi:10.1038/s41467-022-29366-6

42. De Ioris M, Brugieres L, Zimmermann A, Keeling J, Brock P, Maibach R, Pritchard J, Shafford L, Zsiros J, Czauderna P, Perilongo G. Hepatoblastoma with a low serum alpha-fetoprotein level at diagnosis: The SIOPEL group experience. *European Journal of Cancer.* 2008 Mar 1;44(4):545–50. doi:10.1016/j.ejca.2007.11.022

43. Predictive power of pretreatment prognostic factors in children with hepatoblastoma: A report from the Children's Oncology Group - Meyers - 2009 - *Pediatric Blood & Cancer* - Wiley Online Library [Internet]. [cited 2026 Apr 28]. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/pbc.22088?casa_token=ZQL-EV7JI5AAAAAA%3AQmqYexyn81oCh6zdxWPx8dazRD_XYM7SUdqnEJBhatnBj5aFTBGX72PoFS4D35ayr0rNbgh7ZyoNZdQ

Chapter 6 Conclusions

6.1 Summary of findings

This dissertation developed and applied a multi-scale framework to characterize tumor ecosystem states across molecular, spatial, and systemic levels. By integrating bulk transcriptomics, spatially resolved profiling, and multiplexed imaging, and systemic immune analysis, this work demonstrates that therapeutic response emerges from coordinated interactions between malignant epithelial cells, stromal populations, and immune compartments rather than from tumor-intrinsic programs alone.

In TNBC, cell-type informed transcriptomic profiling identified three biologically distinct ecosystem states associated with differential immune engagement, stromal remodeling, metabolic adaptation, and clinical outcome. CC1 was characterized by interferon-driven immune activation and T-cell engagement, CC2 by extracellular matrix remodeling and immune exclusion, and CC3 by transcriptional plasticity, inflammatory signaling, and metabolic reprogramming linked to nuclear receptor activity. These findings demonstrate that bulk transcriptomic heterogeneity reflects coordinated tumor-microenvironment organization rather than epithelial states. Spatial profiling further revealed that these subtype-specific programs are not uniformly distributed throughout tissue but are spatially compartmentalized across epithelial, TME, and invasive regions. Distinct spatial architectures governed immune infiltration, stromal interactions, and regulatory signaling within tumors. Importantly, non-invasive regions exhibited subtype-associated regulatory programs that recapitulated key features of invasive disease, suggesting that critical ecosystem states are established prior to overt invasion.

Complementary analysis of systemic immunity demonstrated that therapeutic response is additionally associated with coordinated peripheral immune dynamics. Longitudinal profiling of

peripheral blood mononuclear cells identified distinct adaptive and innate immune states associated with response and resistance to chemotherapy and chemo-immunotherapy. These findings establish that tumor-associated immune remodeling extends beyond local tissue architecture and can be captured systematically during therapy.

Finally, this framework was extended to hepatoblastoma to evaluate whether ecosystem-oriented spatial modeling generalizes beyond TNBC. Spatial transcriptomic analyses demonstrated that relapse-prone tumors are characterized by centralized stromal and macrophage-associated niches, whereas non-relapse tumors exhibit more distributed and immune-engaged spatial architectures. Together, these findings support a model in which therapeutic response is encoded within coordinated tumor-microenvironment ecosystems across multiple biological scales.

6.2 Conceptual contribution

Collectively, this dissertation advances a conceptual framework in which tumors are understood as spatially and systemically organized ecosystems rather than isolated collections of malignant cells. Across TNBC and hepatoblastoma, tumor behavior emerged from coordinated interactions between epithelial, immune, stromal, and vascular compartments, whose organization varied across spatial regions and therapeutic contexts. Importantly, tumor cells themselves mirrored and potentially reinforced surrounding microenvironmental states, suggesting bidirectional interactions between malignant and non-malignant compartments. Spatial organization therefore represents a biologically meaningful regulatory layer, where immune signaling, stromal remodeling, epithelial plasticity, and systemic immune dynamics collectively shape therapeutic response. By integrating molecular, spatial, and systemic profiling, this

dissertation establishes a multi-scale framework for understanding cancer as an adaptive ecosystem across heterogeneous solid tumors.

6.3 Translational implications

The findings presented in this dissertation also have several translational implications. Across TNBC subtypes, distinct molecular and cellular biomarkers associated with prognosis, immune engagement, stromal remodeling, and metabolic adaptation were identified, highlighting potential opportunities for patient stratification and therapeutic targeting. In parallel, systemic immune profiling identified peripheral immune programs associated with therapeutic response and resistance, demonstrating that circulating immune states can reflect clinically relevant tumor-associated biology. Notably, concordance between tissue-associated and peripheral immune dynamics suggests the potential for minimally invasive blood-based monitoring during therapy. Extending this framework to hepatoblastoma, relapse-prone spatial architectures and prognostic gene signatures were identified that may serve as biomarker of therapy resistance following therapy. Collectively, these approaches support the development of more precise therapeutic strategies that account for both tumor-intrinsic and microenvironmental determinants of response. More broadly, this work highlights the importance of integrating multi-modal datasets to capture the full complexity of tumor biology in clinically actionable way.

6.4 Limitations and future directions

Several limitations should be acknowledged. While these studies integrated multiple molecular and spatial modalities, many analyses were retrospective and performed across independent patient cohorts. Future studies incorporating fully matched longitudinal tissue, spatial, and peripheral immune datasets will be important for resolving causal relationships between local and systemic ecosystem dynamics. In addition, many regulatory interactions identified in this work

were computationally inferred and require functional validation to establish mechanistic roles in therapeutic response and disease progression.

Nevertheless, the rapid emergence of high-dimensional molecular and spatial technologies presents an unprecedented opportunity to transform cancer biology. Advances in spatial transcriptomics, multiplexed imaging, single-cell sequencing, and longitudinal immune profiling are generating increasingly comprehensive representation of tumor ecosystems across scales. At the same time, rapid progress in artificial intelligence (AI) and machine learning provides new computational frameworks capable of integrating these complex datasets into predictive and mechanistically interpretable models. Importantly, while technologies such as spatial transcriptomics provide high-resolution characterization of tumor ecosystems, routine histopathology remains the most widely available representation of tumor architecture in clinical practice. Future AI-driven approaches may therefore use molecular and spatial ecosystem states as training framework to determine whether histologic organization captured by H&E slides can predict immune exclusion, stromal remodeling, relapse-associated architectures, and therapeutic response. In this context, H&E imaging may serve as a scalable bridge between tissue morphology, spatial biology, and clinically actionable ecosystem-level prediction. Ultimately, integrating molecular, spatial, systemic, and histologic information may provide a foundation for AI-enabled cancer biology, where tumors are modeled as dynamic ecosystems rather than isolated malignant cell population.