HARVARD
Kenneth C. Griffin

GRADUATE SCHOOL
OF ARTS AND SCIENCES

THESIS ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Division of Medical Sciences

Committee on Virology

have examined a dissertation entitled

Deciphering Virus-Mediated Remodeling of the Lymphoma Tumor Microenvironment Through Spatial Multi-Omics

presented by Yao Yu Yeo

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature: James DeCaprio
James DeCaprio (Jan 7, 2026 16:54:48 EST)

Typed Name: Dr. James DeCaprio

Signature: Catherine Wu
Catherine Wu (Jan 7, 2026 16:58:10 EST)

Typed Name: Dr. Catherine Wu

Signature: Shira Weingarten-Gabbay
Shira Weingarten-Gabbay (Jan 7, 2026 17:25:58 EST)

Typed Name: Dr. Shira Weingarten-Gabbay

Signature: Evan Newell
Evan Newell (Jan 7, 2026 17:11:47 EST)

Typed Name: Dr. Evan Newell

Date: January 07, 2026

Deciphering Virus-Mediated Remodeling of the Lymphoma Tumor Microenvironment Through Spatial Multi-Omics

A DISSERTATION PRESENTED
BY
YAO YU YEO
TO
THE DEPARTMENT OF MEDICAL SCIENCES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE SUBJECT OF
VIROLOGY

HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS
JANUARY 2026

©2026 – Yao Yu Yeo
all rights reserved.

Dissertation advisor: Dr. Sizun Jiang Yao Yu Yeo

Deciphering Virus-Mediated Remodeling of the Lymphoma Tumor Microenvironment Through Spatial Multi-Omics

Abstract

Oncogenic viruses are associated with ~10% of cancers and play key roles in tumor development, yet how they remodel the tumor microenvironment to drive immune evasion and shape treatment responses remains poorly understood. Consequently, viral status remains mostly diagnostic rather than prognostic, leading to identical treatments against virus-associated cancers, despite growing evidence of different treatment responses. This dissertation addresses this gap by focusing on Epstein-Barr Virus (EBV), the prototypical human oncogenic virus that infects >90% of the population and is associated with several aggressive lymphoid malignancies. These include classic Hodgkin Lymphoma (cHL) and Diffuse Large B-cell Lymphoma (DLBCL), with EBV positivity in ~25% and 5–15% cases respectively.

Through spatial proteomics and transcriptomics, EBV was found to be major factor impairing antitumor immunity by spatially reorganizing the tumor microenvironment. In cHL, EBV-positive tumors were distinctly enriched in cytotoxic CD8 T-cells surrounding malignant cells that retained MHC Class I, yet these T-cells were reprogrammed into terminal exhaustion through direct engagement with malignant cells, a process associated with the viral oncogene LMP1. In DLBCL, EBV disrupted a macrophage:CD4:tumor immunomodulatory axis leading to down-regulated MHC Class II and suppressed CD4 T-cell activity, a process also linked to LMP1. In parallel, several technological advancements were made: IN-DEPTH, a same-slide spatial multi-omics method enabling simultaneous proteomics and transcriptomics profiling; SGCC, a computational tool developed by colleagues at Ohio State University that uses graph Fourier transform to detect spatially coordinated gene-regulatory programs between cell populations; and the incorporation of proximity labeling to detect molecular interactions *in situ*.

The findings presented herein demonstrate how EBV exploits distinct immunoevasive mechanisms across lymphomas, revealing therapeutic vulnerabilities overlooked by current virus-agnostic treatment approaches. The methodological innovations further expand the spatial biology toolkit with broadly applicable experimental and computational strategies. By integrating virology, immunology, cancer biology, and spatial biology, this dissertation also represents a framework for future systems-level investigations across diverse contexts.

Contents

Title Page i

Copyright ii

Abstract iii

Contents iv

Dedication vi

Acknowledgments vii

0 Introduction 1

1 Viruses and Cancers: Entwined Maladies of Subversion and Survival 4

2 Charting New Dimensions: The Rise of Spatial Biology 11

3 A New Frontier: Integrated Multimodal Profiling on a Single Slide 21
- 3.1 Development of IN-DEPTH to enable same-slide spatial multi-omics 22
- 3.2 Development of SGCC for integrative spatial multi-omics analysis 31
- 3.3 Incorporation of proximity labeling to spatial proteomics 35
- 3.4 Improvements to streamline antibody panel development 37

4 Reprogramming the Dysfunctional Immune Landscape of Lymphomas 40
- 4.1 EBV and T-cell dysfunction in the cHL tumor microenvironment 44
- 4.2 EBV and T-cell dysfunction in the DLBCL tumor microenvironment 56

5 Decoding the Mechanistic Logic Behind Lymphoma Remodeling 64
- 5.1 CD8:MHC Class I engagement and T-cell terminal exhaustion in cHL 65
- 5.2 Dysregulated macrophage:tumor:CD4 interaction axis in DLBCL 80

6 Conclusion **86**
6.1 Biological and technological implications 87
6.2 Future directions 95

Appendix A Information on Tumor Specimens **103**

References **112**

To my loved ones, my constant light across the voyage of life.

Acknowledgments

My heartfelt appreciation goes to these people:

- For mentoring my doctoral studies: Sizun Jiang, PhD
- For enabling the studies presented herein: Huaying Qiu & Yuzhou Chang, PhD & Stephanie Pei Tung Yiu, PhD & Yunhao Bai, PhD & Bokai Zhu, PhD
- For guidance on biostatistics and computer language: Huaying Qiu & Yuzhou Chang, PhD
- For spearheading the development of SGCC: Yuzhou Chang, PhD & Qin Ma, PhD
- For histopathological consultation: Scott Rodig, MD PhD & Richard Burack, MD PhD
- For serving on my dissertation advisory committee: Jim DeCaprio, PhD & Alex Shalek, PhD & Margaret Shipp, MD
- For serving on my dissertation examination committee: Jim DeCaprio, PhD & Catherine Wu, MD & Shira Weingarten-Gabbay, PhD & Evan Newell, PhD
- For supporting my doctoral studies: Albert J Ryan Foundation
- For being part of the memorable journey: Huaying Qiu & Yuzhou Chang, PhD & Stephanie Pei Tung Yiu, PhD & Hendrik Michel & Jia Le Lee, PhD & Precious Cramer, PhD & Hongbo Wang, PhD & Lindsay Parmelee & Chi Ngai Chan, PhD & Guanrui Liao, PhD & Shuli Luo, MD & Wenrui Wu, MD & Johanna Schaffenrath, PhD & Pierre Stéphan, MD PhD & Andrew Ma & Addison Deisher & Jason Yeung & Nagendra Dhanikonda & Yang Wang & Clara Kummerer & Yuchen Wang, PhD & Qi Yang, PhD & Fabio Iannelli, PhD
- For being part of the greatest cohort the Harvard Virology program has ever seen: Wanyu Li, PhD & Jessica Chen, PhD & Alexios Staikos, PhD & Tatum Sass, PhD & Sadie Antine, PhD & Victoria Rosado, PhD & Colin Mann, PhD & Sanjana Kulkarni, PhD
- For keeping the Harvard Virology program afloat: Aaron Schmidt, PhD & Ben Gewurz, MD PhD & David Knipe, PhD & Maria Bollinger
- For helping me get this far in my academic journey: Hector Aguilar-Carreño, PhD & David Buchholz, PhD & Bruce Ganem, PhD & Laura Schoenle, PhD & Kai Leng Low, PhD & Tuo Yeong Lee, PhD
- For always being by my side: June & Toes

And last but not least, to my family for always supporting my endeavors through thick and thin.

0

Introduction

Cancer is among the leading causes of deaths worldwide, with approximately one in four people developing cancer during their lifetime [101]. A substantial amount of research has thus been devoted to uncovering the etiology and developing treatment strategies for cancer. Over recent decades, viruses have been increasingly recognized as major contributors to cancer development, with an estimated 10% of new cases linked to viral infections [154]. For these virus-associated cancers, viral infections

are present in anywhere from 1–100% of cases depending on the cancer type[154]. Despite this, current clinical practices treat these cancers with identical approaches regardless of viral status, which remains primarily diagnostic and carries limited prognostic value.

A key reason for this disconnect is that much of our understanding of viral oncogenesis remains focused at the molecular scale, examining virus-host interactions primarily within the cancer cell. While such knowledge is essential, cancer is not merely a hyperproliferative cell. It is a highly organized ecosystem of malignant, immune, and stromal cells organized within a structured tumor microenvironment that collectively promote tumor growth. This complexity underpins the hallmarks of cancer, such as evading immune destruction and sustaining proliferative signaling[81,82,80]. To generate insights that bridge molecular findings with clinical features, a systems-level dissection of the tumor microenvironment *in situ*, comparing cohorts of virus-positive and virus-negative tissue samples, is required. Until recently, this has been impeded by the lack of technologies capable of enabling high-resolution, spatially-resolved analysis of the tumor microenvironment.

Technological advances over the recent years have led to the emergence of spatial biology[26]. These approaches enable the simultaneous detection of thousands of biomolecules within the native tissue context, a significant advancement beyond traditional low-parameter profiling approaches, while circumventing the limitations of dissociative assays. As a result, these advances have enabled a shift from molecular to systems-level understanding of cancer biology, as it is now possible to thoroughly dissect the tissue architecture and reveal the spatially coordinated cellular programs that govern tumor development. Importantly, this provides a unique opportunity to investigate virus-mediated remodeling of the tumor microenvironment, a step towards developing virus-informed prognostics and treatment strategies for cancer.

This dissertation details how Epstein–Barr virus, the prototypical human oncogenic virus, remodels the tumor microenvironment in lymphoid malignancies through innovative spatial proteomics and transcriptomics approaches. It also introduces experimental and computational innovations that

extend the potential of spatial biology to diverse disease contexts. Together, the biological and technological advances presented herein establish a framework for future spatial biology studies and motivate the development of virus-informed strategies for cancer treatment.

Gentlemen, it is the microbes who will have the last word.

Louis Pasteur

1

Viruses and Cancers: Entwined Maladies of Subversion and Survival

The idea that viruses can cause cancer was a radical notion over a century ago, as cancers were formally thought to arise solely from genetic mutations. It was not until 1911, when Peyton Rous demonstrated that a chicken sarcoma could be transmitted using a cell- and bacteria-free tumor

Virus	Cancer	% virus-positive patients
EBV	Classic Hodgkin Lymphoma	~25%
	Diffuse Large B-Cell Lymphoma	~10-15%
	Primary Central Nervous System Lymphoma	~5%
	Nasopharyngeal Carcinoma	~90%
	Gastric Cancer	~10%
HPV	Cervical Cancer	>95%
	Oropharyngeal Cancer	~60-70%
	Anal Cancer	~90%
	Vaginal Cancer	~70%
	Penile Cancer	~50%
HBV	Hepatocellular Carcinoma (HCC)	~50-65%
HCV	Hepatocellular Carcinoma (HCC)	~20-30%
	Non-Hodgkin Lymphoma	~15%
HTLV-1	Adult T-Cell Leukemia / Lymphoma	>95%
KSHV	Kaposi's Sarcoma	>95%
MCPyV	Merkel Cell Carcinoma	~80%

Figure 1.1: The seven human oncogenic viruses and examples of associated cancers. Note that the percent of patients with viral infections can significantly differ based on geography and immune status, and these numbers shown here reflect a global estimate. This dissertation will focus on EBV in classic Hodgkin Lymphoma and diffuse large B-cell lymphoma, as highlighted in red.

extract[169]. The causative agent, eventually known as Rous Sarcoma virus, earned him the Nobel Prize in Physiology or Medicine in 1966, marking the longest interval between a scientific discovery and its Nobel recognition. This seminal finding established a precedent for infectious carcinogenesis and transformed the cancer biology field, eventually paving way for the recognition of human oncogenic viruses.

Since then, viruses were found to contribute substantially to the global cancer burden. It is estimated that around 10% of all human cancers are attributable to viral infections[154]. To date, seven viruses are known to be oncogenic in humans **(Figure 1.1)**: Epstein–Barr virus (EBV), human papillomavirus (HPV), hepatitis B virus (HBV), hepatitis C virus (HCV), human T-cell lymphotropic virus type 1 (HTLV-1), Kaposi's sarcoma-associated herpesvirus (KSHV), and Merkel cell polyomavirus (MCV). Viruses, being obligate parasites, have evolved diverse sophisticated machineries to counter-

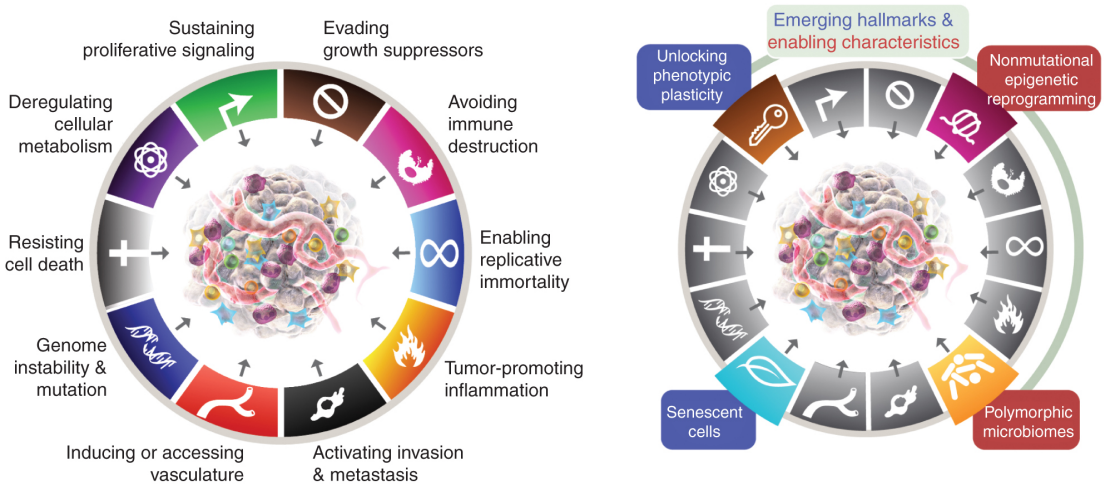


Figure 1.2: The hallmarkers of cancer. Last updated in 2022, this framework was developed to distill the complexity of cancer into unifying biological principles. It currently encompasses ten core hallmarks (left), with emerging features (right) continuing to refine this model as research advances. Cancer cells acquire these hallmarks progressively over time, each contributing to their ability to survive, proliferate, and reshape the tumor ecosystem. Figure from Hananan, 2022[80].

act host antiviral responses and ensure their persistence. For oncogenic viruses, each is equipped with distinct mechanisms of oncogenesis that contribute to the hallmarks of cancer, ultimately subverting control by host immune responses to ensure survival of the infected malignant cell — and in turn, themselves.

Cancer is best understood not as a disease of a rogue cell, but as a complex ecosystem of multiple interacting cell types that actively shape a tumor microenvironment, where cancer hallmarks are progressively acquired as normal cells evolve towards a malignant neoplasm[81,82,80] **(Figure 1.2)**. Importantly, oncogenic viruses express viral oncogenes that can directly contribute to these hallmarks. For example, to enable cellular replicative immortality, EBV latent membrane protein 1 (LMP1) constitutively activates NF-κB signaling[198], and HPV E6 and E7 proteins degrade p53 and Rb respectively[171,25]. Notably, virus-positive cancer cells generally harbor lower mutational burdens than their virus-negative counterparts[139]. This supports the notion that viruses can directly participate in tumorigenesis, such as subverting cellular checkpoints and promoting survival of the malignant clone.

Although our understanding of cancer has been transformed and now viewed as a cohesive biological entity, relatively little effort has been devoted to contextualizing oncogenic viruses within this framework. The molecular functions of viral oncogenes may have received much research attention, but they are largely described from *in-vitro* studies; any links to clinical diseases are usually associative or speculative. Consequently, viral infection status in clinical settings is typically considered only from a diagnostic standpoint, and virus-associated cancers continue to be treated the same way as their virus-negative counterparts. To address this gap, this dissertation focuses on B-cell lymphomas associated with EBV, the first discovered human oncogenic virus.

EBV was co-discovered in 1964 by Anthony Epstein, Bert Achong, and Yvonne Barr during their investigation of Burkitt lymphoma[64]. EBV is a ubiquitous gammaherpesvirus infecting >90% of the global population, yet paradoxically is associated with only a relatively small number of B and epithelial cell malignancies. These include classic Hodgkin lymphoma (cHL), Diffuse Large B-Cell Lymphoma (DLBCL), Burkitt lymphoma, nasopharyngeal carcinoma, and gastric carcinoma[176], among others. In addition, EBV is also the causative agent of infectious mononucleosis[89] and has recently been identified as a major risk factor for multiple sclerosis[20]. Despite the consensus that EBV is an oncogenic virus, these inconsistencies have fueled the long-standing controversy on whether EBV is a casual driver or a passenger in tumorigenesis.

At the cellular level, EBV can readily transform quiescent B-cells into immortalized lymphoblastoid cell lines *in vitro*[150], reflecting its oncogenic potential. *In vivo*, EBV infection is accompanied by the establishment of distinct latency programs, each defined by the expression of different viral latent oncogenes and associated with particular malignancies[105,137] **(Figure 1.3)**. For example, EBV-positive cHL is proposed to exist in latency stage II marked by the expression of EBV nuclear antigen 1 (EBNA1), LMP1, and LMP2, whereas EBV-positive DLBCL is thought to exist in both latency stages II and III, the latter stage encompassing a broad expression of EBNA and LMP oncogenes. Although these latency stages were traditionally viewed as rigid categories, emerging evidence suggests

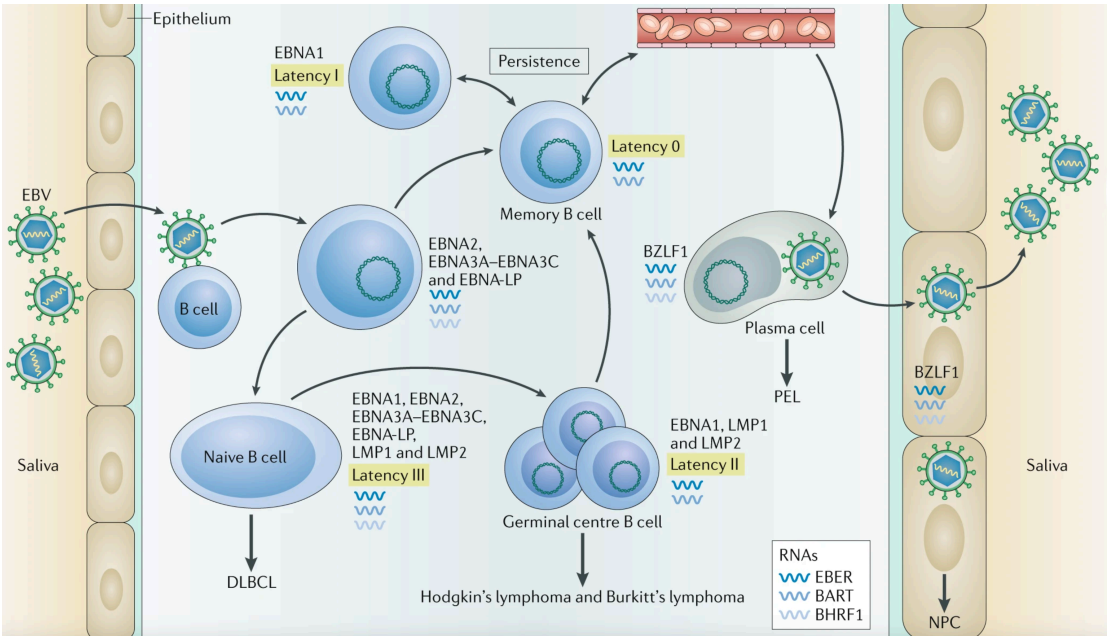


Figure 1.3: Proposed model of EBV latency stages and associated cancers during its infection cycle. Figure from Münz, 2019 [137].

substantial heterogeneity, with lytic and latent genes expressed beyond their canonical contexts [16,207]. Furthermore, a considerable number of these viral genes contribute to the cancer hallmarks, demonstrating the oncogenic properties of EBV **(Figure 1.4)**. This challenges the classical model of EBV latency and points to a more active and context-dependent role for EBV in tumorigenesis, with dynamic latency programs that enable evasion of host immunosurveillance to maintain long-term persistence. This also underscores the need for single-cell profiling of EBV viral oncogene expression in malignant cells within the native tumor microenvironment.

Far less is known about the role of EBV in shaping the tumor microenvironment. EBV has been linked to the secretion of immunomodulatory cytokines and expression of immune checkpoints across various lymphomas [41,168,192]. For example, EBV-positive malignant cells often upregulate interferon-related genes and overexpress PD-L1, two mechanisms that can induce T-cell exhaustion [202]. These processes can have major implications for remodeling the EBV-positive tumor microenvironment to

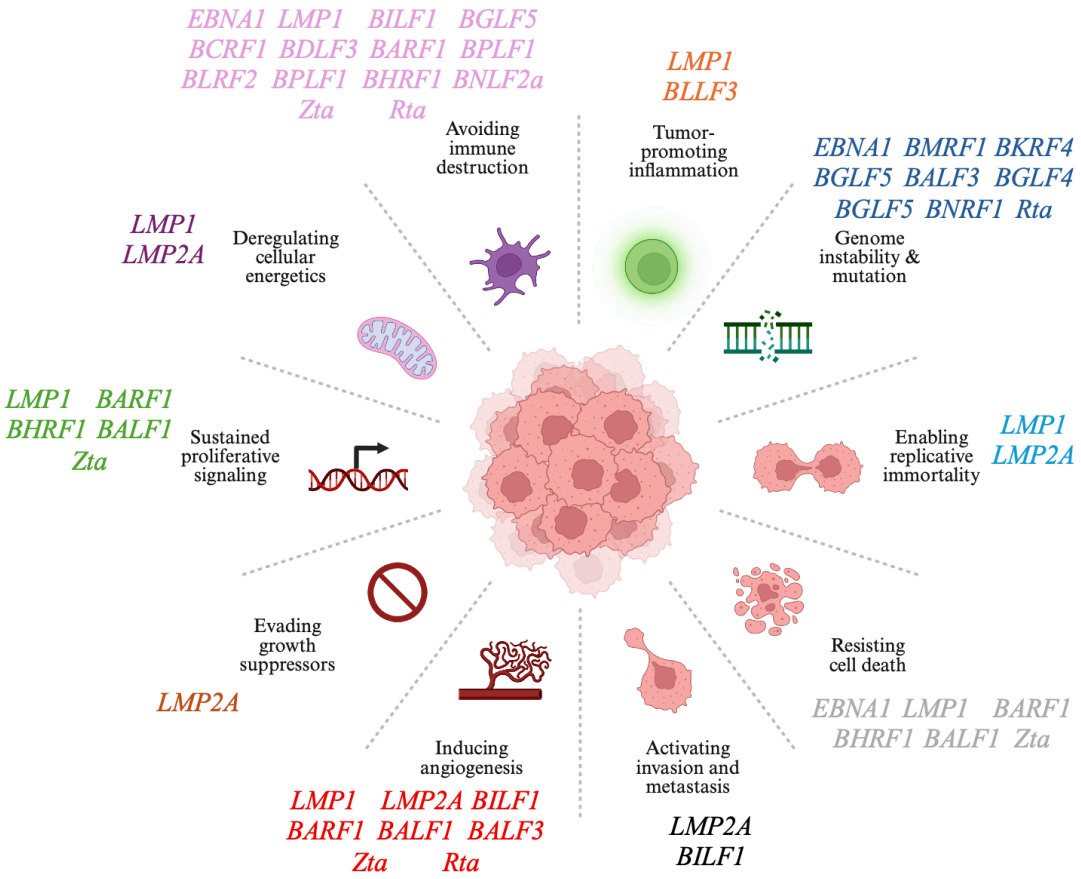


Figure 1.4: EBV encodes numerous genes across its lytic and latent stages that can contribute to all aspects of the cancer hallmarks. Adapted from Mesri et al., 2014[130] & Dorothea et al., 2023[62].

ensure tumor survival and progression[5], yet they remain poorly understood to date. Attempts to address these gaps include using humanized mouse models to explore tumor-immune interactions and therapeutic interventions[93,135]. These systems offer limited insights as EBV infection is restricted to humans, reflective of a strong co-evolutionary relationship across millennia[210]. As a result, non-human hosts often fail to recapitulate the complexity of EBV biology and the human tumor microenvironment.

Dissecting the mechanisms of EBV-driven remodeling of the tumor microenvironment thus requires the study of human biopsy samples *in situ*. However, progress has been constrained by techno-

logical limitations. Conventional histological, immunohistochemical, and *in-situ* hybridization approaches detect only a few targets, whereas assays with higher throughput typically require sample dissociation that sacrifices spatial context, a critical feature of tumor microenvironment organization. This gap has hindered mechanistic insight of the EBV-positive tumor microenvironment and identification of potential therapeutic vulnerabilities, contributing to the lack of consideration for EBV status in clinical practice.

Addressing this challenge from both biological and technological perspectives forms the central motivation of this dissertation. Recent technological breakthroughs have heralded the dawn of spatial biology, making it possible to visualize biomolecules *in situ* at the transcriptome and proteome scales. These approaches provide an unprecedented opportunity to dissect how oncogenic viruses remodel the tumor microenvironment and influence disease outcomes.

Progress in science depends on new techniques, new discoveries, and new ideas, probably in that order.

Sydney Brenner

2

Charting New Dimensions: The Rise of Spatial Biology

A FERVENT BIOTECHNOLOGICAL REVOLUTION over the past decade marked the emergence of spatial biology, offering unprecedented methods for genome-wide profiling of biomolecules directly from intact tissues *in situ*. These technologies preserve native tissue contexts unlike traditional assays that

Spatial Transcriptomics			
Platform	**Plex**	**Modality**	**Reference**
MERFISH[a]	1k	Single-molecule FISH	[35]
Visium	Genome	Poly-A capture	[183]
Slide-seq[b]	Genome	Direct RNA capture	[163] [186]
GeoMx	18k	Probe-based capture	[129]
DBit-Seq	Genome	Direct RNA capture	[119]
CosMx	18k	Single-molecule FISH	[87] [107]
Xenium	5k	Single-molecule FISH	[100]
VisiumHD	18k	Probe-based capture	[145]

Spatial Proteomics			
Platform	**Plex**	**Modality**	**Reference**
MIBI	>40	ToF-SIMS	[6]
IMC	>40	ICP-MS	[73]
Opal[c]	7	Multiplexed IF	[182]
CODEX[d]	>40	Multiplexed IF	[74]
IBEX	>40	Multiplexed IF	[158]
MICS	>40	Multiplexed IF	[109]
Orion	16	Multiplexed IF	[117]
COMET	>40	Multiplexed IF	[161]

Figure 2.1: Commonly used spatial platforms. Although these platforms are all spatially resolved, they vary in their capabilities and can be tailored to answer different biological questions. This is not an exhaustive list, and new platforms continue to be developed today. FISH, fluorescence in situ hybridization; ICP-MS, inductively coupled plasma mass spectrometry; IF, immunofluorescence; ToF-SIMS, time-of-flight secondary ion mass spectrometry. [a]Commercial MERSCOPE released by Vizgen in 2022. [b]Slide-seqV2, with higher efficiency and spatial resolution, was released in 2022. [c]Also known as Mantra or Vectra Polaris prior to acquisition by Akoya, and then rebranded as Phenoptics in 2024. [d]Rebranded as PhenoCycler in 2022. For more information regarding each technology: MERFISH[35], Visium[183], Slide-seq[163,186], GeoMx[129], DBit-seq[119], CosMx[87,107], Xenium[100], VisiumHD[145], MIBI[6], IMC[73], Opal[182], CODEX[74], IBEX[158] MICS[109], Orion[117], COMET[161]. Adapted from Aliazis et al., 2025[3].

dissociate cells and obscure tissue architecture, thus enabling the study of nucleic acids and proteins in relation to cellular interaction and organization. They are also applicable to archival Formalin-Fixed Paraffin-Embedded (FFPE) tissue blocks, the gold standard of clinical tissue preservation, thus unlocking the ability to analyze decades of carefully curated biobank specimens. The exceptional impact has been broadly recognized, with *Nature Methods* naming spatial transcriptomics the "Method of the Year" in 2021[131] and extending to spatial proteomics in 2024[132]. These technologies are now indispensable for systems-level study of disease mechanisms.

The rapidly expanding field is currently fueled by the development of innovative chemistries and imaging approaches, leading to a diverse set of platforms that enable spatial transcriptomics and spatial proteomics profiling **(Figure 2.1)**. Each platform builds on prior limitations to optimize throughput, signal-to-noise, spatial resolution, and efficiency **(Figure 2.2)**, offering distinct advantages when selecting the most suitable platform for a given application. This dissertation focuses on the use of

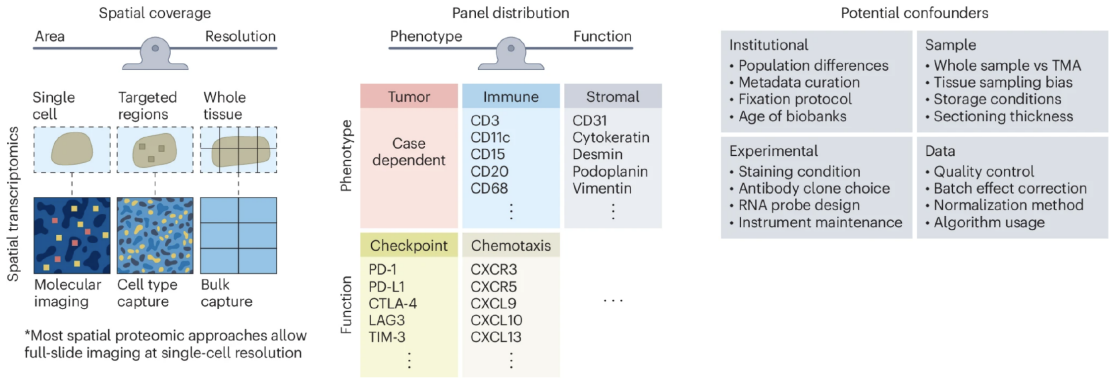


Figure 2.2: Key considerations for spatial omics studies. For spatial transcriptomics (left), while high-coverage single-cell analyses enable precise dissection of cell states, this often comes at the expense of whole-tissue insights and may lack the high signal-to-noise ratio that bulk approaches offer (left). For spatial proteomics (middle), a well-balanced panel incorporates both phenotyping and functional targets for a comprehensive investigation of tumor microenvironment constituents and cellular interactions within. As numerous experimental procedures occur before a tissue sample reaches the laboratory, spatial omics data are highly sensitive to experimental confounders across multiple levels (right); ensuring meticulous control over these factors is essential for producing high-quality and interpretable data that reliably capture the tumor's molecular landscape. TMA, tissue microarray. Figure from Aliazis et al., 2025[3].

four platforms, each with unique strengths and limitations: the GeoMx[129] and CosMx[87] for spatial transcriptomics, and the Multiplexed Ion Beam Imaging (MIBI)[6] and CO-DEtection by indeXing (CODEX)[74] for spatial proteomics.

The GeoMx Digital Spatial Profiler[129], introduced by Nanostring in 2020, provides a high degree of flexibility in defining spatial regions to profile. It relies on a small panel of antibodies ($\leq$3) to identify regions of interest, which are used to guide transcriptome capture through a panel of probes targeting >18,000 protein-coding transcripts. Each probe is conjugated to a unique barcode via a UV-photocleavable linker, where upon UV illumination, barcodes are released for next-generation sequencing to quantify the expression level of each gene within each capture region **(Figure 2.3)**. Limitations for the GeoMx include the ability for UV-cleaved barcodes to diffuse in aqueous media and contribute to nonspecific signals, as well as the lack of single-cell resolution. Nonetheless, from experience, the GeoMx retains robust signal-to-noise, making it ideal for macro-scale profiling to capture broad functional states across user-defined regions of interest. Note that while the GeoMx offers

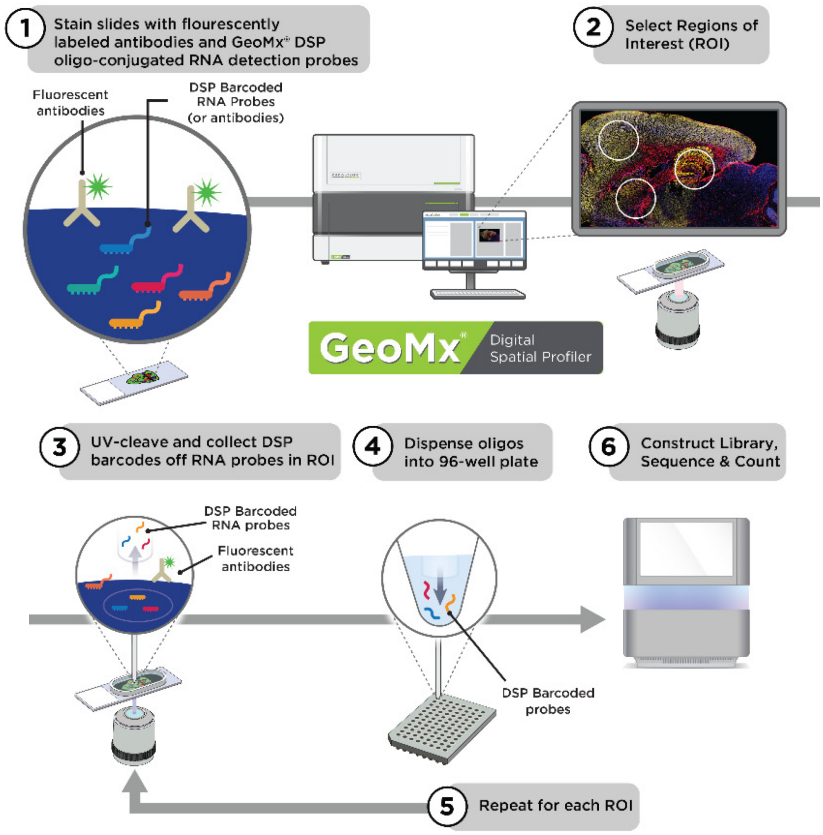
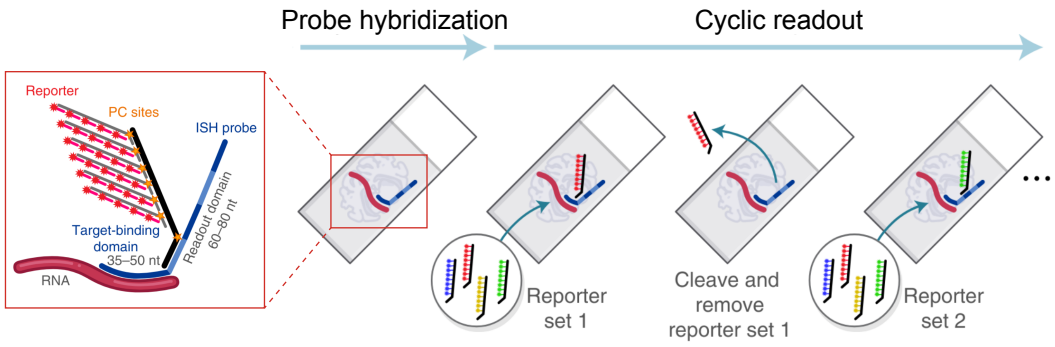


Figure 2.3: GeoMx experimental procedure. A tissue section is hybridized with RNA detection probes that contain unique oligonucleotide barcodes connected via UV-photocleavable linkers. Fluorophore-conjugated antibodies are also stained to visualize the tissue, from which regions of interest are selected for UV-cleavage, and the barcodes are collected for next-generation sequencing to profile each region's trancriptome. The protein-based readout uses the same approach by labeling antibodies with unique barcodes. Figure from Bergholtz et al., 2021 [17].

an orthogonal protein-based readout using barcoded antibodies instead of RNA-targeting probes, it was not used here in favor of single-cell spatial proteomics platforms.

The CosMx Spatial Molecular Imager [87], also introduced by Nanostring in 2022, advances the GeoMx by enabling imaging of individual RNA molecules. Individual cell boundaries are defined with an antibody cocktail, while RNAs are detected using probes labeled with combinatorial barcodes. Across iterative imaging cycles, each probe generates a unique sequence of fluorescent signals that, when demultiplexed, identifies the specific transcript. This combinatorial strategy dramatically ex-



Reporter Cycle

RNA spot numer	1	2	3	4	5	6	7	8	
1	ON	x	ON	x	x	ON	ON	x	Gene 1
2	x	x	ON	ON	ON	x	x	ON	Gene 2
3	x	ON	x	ON	ON	ON	x	x	Gene 3
4	ON	x	ON	x	x	ON	ON	x	Gene 1
5	ON	x	x	ON	x	x	x	x	N/A
6	x	ON	ON	ON	x	ON	x	x	Gene 4

Figure 2.4: CosMx experimental procedure. A tissue section is hybridized with RNA detection probes containing an extended RNA-specific readout domain that encodes n reporter cycles using a Hamming weight of 4. Sets of fluorescent reporters are then sequentially hybridized to these readout domains, imaged, and then photocleaved, enabling repeated rounds of signal detection. Each RNA punctum generates a sparse, 4-color barcode distributed across the n imaging cycles. The bottom table shows an example using 4 fluorophores across an 8-cycle design (theoretical coding space of 17,920): only RNA spots with exactly four "ON" cycles are decoded, and the specific combination of ON-cycle positions and fluorophore identities determines the target gene. Adapted from He et al., 2022[87].

pands detection capacity: by using m fluorophores across n cycles and limiting the readout to only 4 of the n cycles to improve specificity, the theoretical coding space is $({}^{n}\mathrm{C}_{4})\, m^{4}$. This enables a substantial number of transcripts to be resolved at single-molecule resolution **(Figure 2.4)**. However, this precision comes with significant trade-offs, with slower acquisition, smaller spatial coverage, and reduced signal-to-noise, especially when attempting high-plex or large-area profiling. In practice, the CosMx is most effective for dissecting specific functional states underpinning spatially coordinated processes within well-defined tissue regions.

The MIBI[6], developed by the lab of Garry Nolan at Stanford University in 2014, adapts princi-

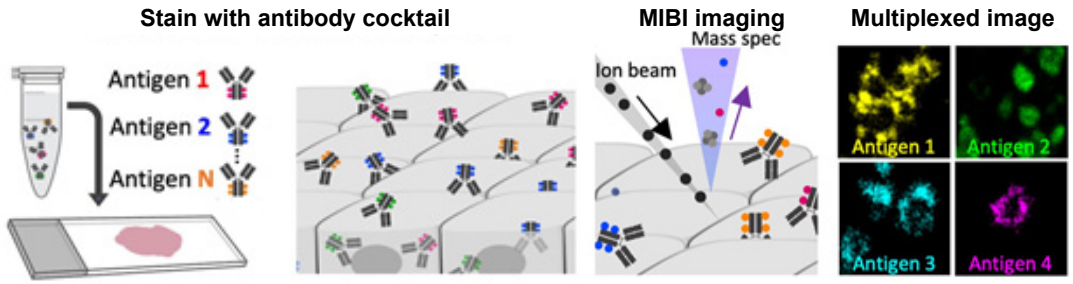


Figure 2.5: MIBI experimental procedure. A tissue section is stained with cocktail of antibodies, each conjugated with a unique metal isotope, which are detected by rastering the tissue pixel-by-pixel using primary charged ions. This process liberates secondary ions measurable by time-of-flight mass cytometry to generate an N-dimensional multiplexed image. Figure from Keren et al., 2019 [106].

ples from CYtometry by Time-Of-Flight (CyTOF) [15] to tissue imaging. The fundamental principle leverages on staining tissues with a panel of antibodies, each conjugated to a unique monoisotopic lanthanide reporter. A primary ion beam is then rasterized across the tissue pixel-by-pixel, releasing secondary ions that are detected via time-of-flight mass spectrometry [190] **(Figure 2.5)**. This approach provides high sensitivity and avoids tissue autofluorescence, producing spatially-resolved multiplexed protein maps at single-cell resolution. However, limitations include the availability of isotopes restricting antibody panels to ~40 markers, the destructive imaging process preventing reuse of tissues for further assays, and the moderate imaging resolution (~0.78 μm per pixel). Image processing pipelines for mass spectrometry imaging are also less mature compared to fluorescence-based workflows, resulting in slower data turnaround times. Nonetheless, the MIBI excels in visualizing highly autofluorescent tissues and in contexts where detecting low-abundance proteins is critical.

The CODEX [74], also from the Nolan lab in 2018, represents a major innovation in fluorescence-based imaging. It relies on staining tissues with a panel of antibodies, each conjugated to a unique oligonucleotide barcode, which are detected through iterative cycles of hybridization, imaging, and stripping of fluorophore-conjugated complementary oligonucleotides. This cyclic immunofluorescence approach generates individual image channels that, when stitched together, reconstructs the tissue architecture at a single-cell resolution [21] **(Figure 2.6)**. A key advantage of CODEX is its scalability:

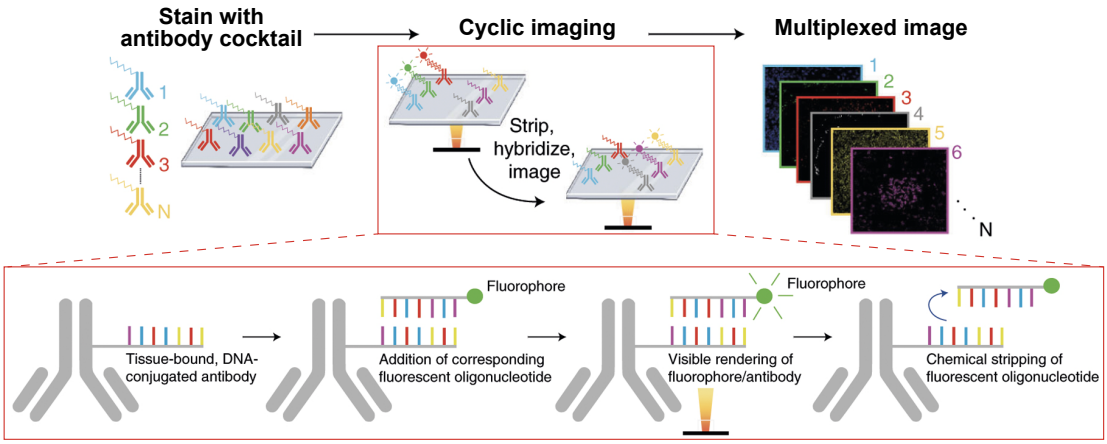


Figure 2.6: CODEX experimental procedure. A tissue section is stained with cocktail of antibodies, each conjugated with a unique oligonucleotide barcode, which are detected by hybridizing a fluorophore-conjugated complementary oligonucleotide. Through repeated hybridize-image-strip cycles, this process generates an N-dimensional multiplexed image. Adapted from Black et al., 2021[21].

the 20-nucleotide barcode design yields 4^{20} (~1 trillion) possible combinations, and thus constrains are mostly restricted to imaging logistics rather than chemistry: by using m channels to image an n-sized panel, a total of $\lceil n/m \rceil$ cycles is required, which may potentially degrade the quality of biomolecules on the tissue due to extensive microfluidics treatment. Nonetheless, the fluorescence imaging approach preserves tissue morphology and enables higher imaging resolutions (<0.6 μm per pixel) depending on the microscope used, which are major benefits over mass cytometry imaging approaches (MIBI[6], IMC[73]). However, challenges include tissue autofluorescence, point-spread function, and signal spillover that contribute to nonspecific signal, which can be partially alleviated through computational correction[12]. The CODEX combines scalability with streamlined workflows and, being nondestructive, offers potential for integration with complementary assays.

Spatial biology represents not only a technical advance but also a conceptual one, as spatial dimension is now recognized as an essential axis of biological function. For example, the proximity between malignant cells and exhausted immune cells within the tumor microenvironment can directly reveal cellular interactions driving tumor progression before the specific molecular pathways are identified.

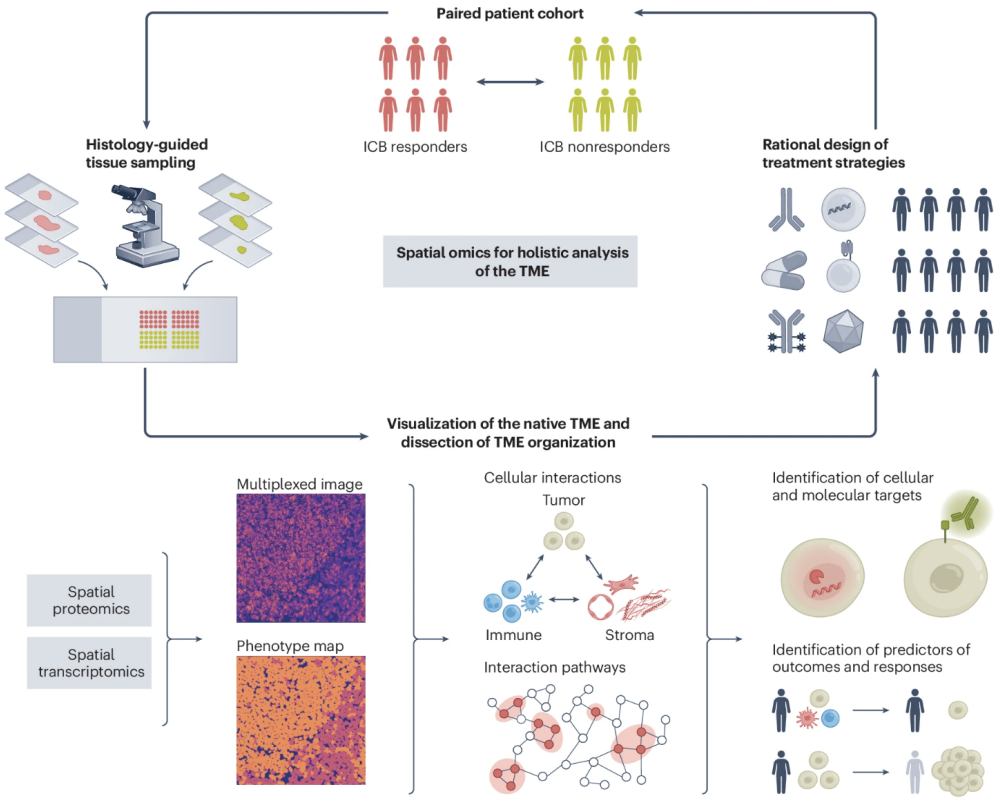


Figure 2.7: Spatial omics can advance our understanding of tumor microenvironmental responses and inform precision treatment approaches. By leveraging paired patient cohorts (e.g. Responders *vs.* Nonresponders, EBV-positive *vs.* EBV-negative), this establishes a functional link between molecular and tissue-level features based on clinical outcomes. These insights may guide rational design of therapeutics that help reshape the tumor for improved patient responses. ICB, immune checkpoint blockade; TME, tumor microenvironment. Figure from Aliazis et al., 2025[3].

This perspective has also transformed the approach towards tumor heterogeneity, a long-standing barrier in cancer biology, as heterogeneous signaling programs are often shaped by spatial relationships between distinct cell populations[208]. By contextualizing molecular programs within the intact tissue architecture, spatial biology enables a holistic view of the tumor microenvironment, guiding the rational design of therapeutic interventions aimed at reshaping it to improve clinical outcomes[3] **(Figure 2.7)**. This integrative perspective is increasingly recognized as a cornerstone of cancer research.

Each new advance in spatial biology enables the generation of increasingly large datasets that often

comprise genomic-scale molecular features measured across millions of cells. Managing this complexity has empowered a parallel revolution in computational biology that fortuitously coincide with the rapid rise of artificial intelligence. For example, supervised deep-learning models now enable segmentation of millions of cells within minutes [34] and accelerate the classification of cell types from weeks to days [173], self-supervised approaches can detect rare cell populations that are indistinguishable within a single modality [212], and foundation models are currently trained to identify features of tissue organization predictive of clinical outcomes [174]. These computational advances are no longer ancillary but constitute a discipline of their own, propelling spatial biology beyond descriptive mapping and towards mechanistic, predictive models of disease. This synergy has created an exceptionally fertile environment for transformative discoveries.

Important limitations remain despite the rapid advances in spatial biology. Spatial transcriptomics and proteomics remain largely separate modalities, constraining the depth of biological insights by limiting the number of biomolecules that can be detected simultaneously. Each platform has complementary strengths and weaknesses: spatial transcriptomics offer broad feature coverage and pathway-level insights but is limited by post-transcriptional processes and inconsistent RNA-protein correlations [51], whereas spatial proteomics directly capture molecular functions with high signal but lower multiplexing capacity. Integrating these technologies can overcome their individual shortcomings, address challenges in data cross-registration across different tissue sections, and enable significantly more comprehensive understanding of complex biological systems. Furthermore, analytic frameworks continue to lag behind experimental progress, with a lack of harmonized spatial multi-omics approaches to dissect coordinated biological programs that drive cell states and interactions. Proteomics and transcriptomics also dominate the field, with other modalities such as glycomics and metabolomics remaining largely in their infancy, further limiting the ability to capture the full biochemical complexity of tissues. Finally, current methods also provide static snapshots of dynamic processes, restricting insights about temporal interactions.

Overcoming these challenges is essential to address the complex question of how viruses remodel the tumor microenvironment. This motivated the introduction of a platform-agnostic, same-slide spatial multi-omics profiling approach, a new computational framework for integrated analysis of proteomics and transcriptomic features, and proximity labeling to directly measure molecular interactions *in situ*. Together, these advances establish a framework to interrogate virus–host interactions at an unprecedented resolution to reveal how oncogenic viruses remodel the tumor microenvironment, while being broadly applicable across diverse biological systems.

The important thing in science is not so much to obtain new facts as to discover new ways of thinking about them.

Lawrence Bragg

3

A New Frontier: Integrated Multimodal Profiling on a Single Slide

Current attempts at spatial multi-omics typically rely on serial tissue sections, with one section processed for transcriptomics and another for proteomics, due to the independence of these spatial technologies. This strategy is highly ineffective for several reasons. First, adjacent tissue sections

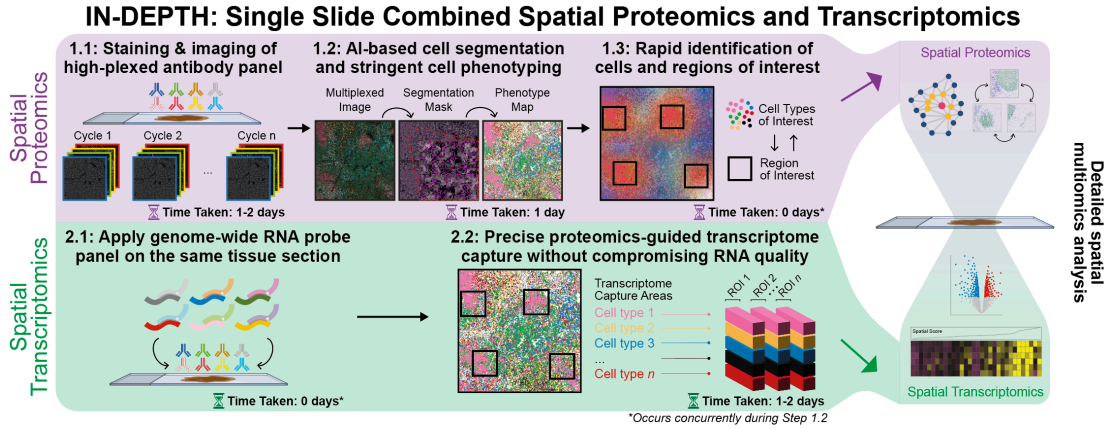


Figure 3.1: Schematic overview of IN-DEPTH, where spatial proteomics is used to guide targeted transcriptome capture on the same slide in a platform-agnostic manner. Figure from Yeo et al., 2024 [189].

differ morphologically, complicating the interpretation of spatially coordinated biological events. Second, the approach is time-consuming and resource-intensive, due to the cost and complexity of each assay. Third, data integration across separate tissue sections is non-trivial, as morphological differences and assay-specific biases hinder robust alignment and interpretation of multi-omics features. Fourth, batch effects accumulate exponentially with each added experiment, and can easily lead to inconsistent conclusions. Although there are efforts to combine protein and nucleic acid imaging [129,109,102,53], one modality is typically much lower in throughput than the other due to experimental and technical constraints. Consequently, analysis of spatial transcriptomics and proteomics datasets are typically performed independently, generating only associative biological relationships between orthogonal readouts, while true integrated analysis remains limited. The advancements introduced herein address these limitations and enable spatial biology to be revisited from a new perspective.

3.1 Development of IN-DEPTH to enable same-slide spatial multi-omics

IN-DEPTH (IN-situ DEtailed Phenotyping To High-resolution transcriptomics) [189] overcomes these limitations by enabling spatial proteomics and spatial transcriptomics capture on the same slide. Single-

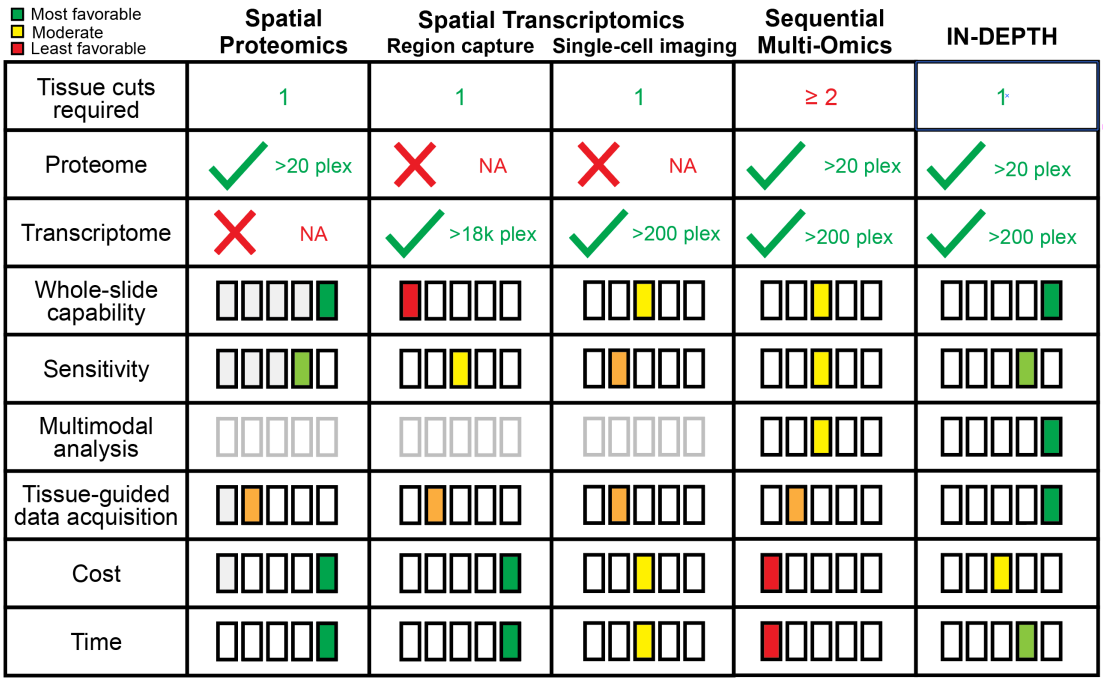


Figure 3.2: Comparison of IN-DEPTH with other existing spatial-omics approaches. Beyond offering same-slide multi-omics data, IN-DEPTH is significantly more efficient in cost, time, and resources, and thus is highly scalable. Figure from Yeo et al., 2024 [189].

cell spatial proteomics is performed first to enable rapid phenotyping of the tissue architecture, which in turn guides subsequent targeted spatial transcriptomics capture on the same slide at regions of interest identified from spatial proteomics **(Figure 3.1)**. Performing spatial proteomics first also avoids the waste of transcriptomics reads on tissue artifacts that are invisible to RNA assays alone. By combining both modalities in sequence, IN-DEPTH enables a streamlined and resource-efficient approach for multi-omics data capture within a single tissue section, reducing resource demands compared to whole-slide transcriptomics while ensuring that transcriptomics data is always biologically relevant **(Figure 3.2)**.

The success of IN-DEPTH required careful optimization to preserve molecular integrity and spatial context. Three key aspects were addressed:

1. *Molecular stability:* Spatial transcriptomics workflows often include chemical (acid/base) or enzymatic (protease) reverse-crosslinking steps to increase RNA accessibility. However, these treatments damage protein epitopes and can even degrade tissue morphology if used excessively, negatively impacting both immunofluorescence and histological staining. To mitigate this, spatial proteomics imaging was always performed first, followed by tissue fixation immediately after antibody staining. Additional fixation steps were also introduced after RNA probe hybridization, and reverse-crosslinking was only used sparingly. These adjustments ensured the stability of both proteins and RNA throughout the entire assay.

2. *Buffer and temperature compatibility:* Each spatial platform requires different buffers and conditions, risking morphological or osmotic artifacts if interchanged haphazardly. Consistent hydration in buffered saline (pH 7.0–7.4) within a controlled temperature range (4–40°C) was found to preserve tissue morphology across all platforms. Antigen retrieval was also standardized using a pH 9.0 buffer at 97°C for 10-20 minutes and performed using a temperature-controlled PT module, avoiding boiling or pressure-cooking approaches that often introduce batch-to-batch inconsistencies. Tissues were also equilibrated to room temperature before exposure to drastic temperature changes.

3. *Cross-platform image alignment:* Nuclear staining was used as the reference for image alignment, as it displayed the least variability when imaged on different platforms. Alignment across imaging modalities was reliably achieved using the Scale-Invariant Feature Transform (SIFT) algorithm [120] for feature detection and description, followed by brute-force matching across sections.

These careful optimizations enabled sequential spatial profiling on the same tissue section, establishing IN-DEPTH as a unified spatial multi-omics workflow designed to preserve tissue integrity across platforms. In brief, following antigen retrieval, tissue sections were blocked to minimize non-

specific interactions (e.g. donkey serum for antibody binding and sheared salmon sperm DNA for oligonucleotide-based readouts). A pre-titrated panel of conjugated antibodies tailored to the specific spatial proteomics platform was pooled into a single cocktail, filtered to remove aggregates, and applied on the tissue at either room temperature for 2 hours or 4°C overnight. After antibody staining, tissues were stabilized through a series of fixation steps (e.g. paraformaldehyde and/or methanol and/or bis-sulfosuccinimidyl suberate) to preserve molecular and morphological integrity prior to imaging. Spatial proteomics imaging was then performed, typically completing within 24-48 hours depending on the number of imaging cycles. Upon completion, slides were immediately removed from the imaging platform and exchanged into platform-compatible buffers (e.g. 1× PBS, 2× SSC) to prepare for transcriptomics processing. Where required, a mild and carefully optimized reverse-crosslinking step (e.g. protease digestion, high pH) would be applied, followed by hybridization with RNA-targeting probes for 16–24 hours under the recommended conditions. After post-hybridization washes, nuclear staining and broad morphology antibodies are applied to facilitate tissue visualization and selection of regions of interest. Spatially-resolved transcriptome capture would then be performed accordingly via the corresponding platform's chemistry. Finally, tissue sections were retained for H&E staining to assess tissue integrity and annotation accuracy by board-certified pathologists. While specific implementation details may vary across platforms, the overarching workflow principles remain largely consistent, provided tissue disruption is minimized. This sequential design thus enables protein-guided transcriptomic interrogation while maintaining spatial fidelity across both modalities.

The performance and generalizability of IN-DEPTH was systematically evaluated across multiple tissue and disease types using four spatial proteomics and four spatial transcriptomics platforms with proven records in research and clinical settings. The spatial proteomics platforms (CODEX[74], SignalStar[149], Polaris[182], Orion[117]) were selected for their whole-slide imaging capabilities, preservation of tissue integrity, and diversity of imaging methodologies including cyclic immunofluorescence, signal amplification, and spectral deconvolution. The spatial transcriptomics platforms (GeoMx[129],

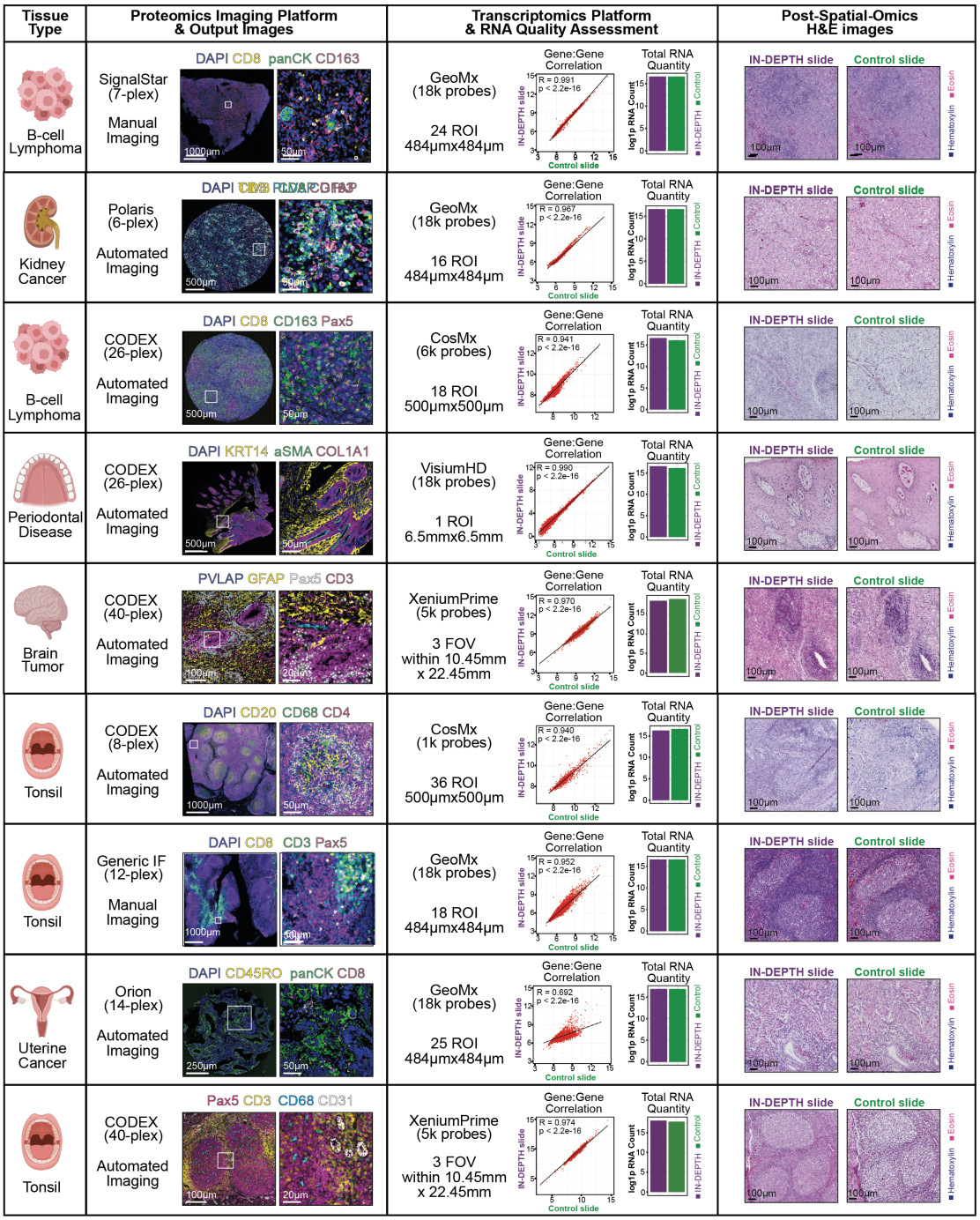


Figure 3.3: IN-DEPTH retains protein & RNA quality across various spatial modalities and tissue types. Adapted from Yeo et al., 2024 [189].

VisiumHD[145], CosMx[87], Xenium[100]) were selected as they represented different balances of spatial resolution, gene coverage, and imaging chemistry. Note that IN-DEPTH is incompatible with mass cytometry imaging approaches (MIBI[6], IMC[73]) due to their inherent destructive nature during the imaging process. Two adjacent sections were used for benchmarking: one processed with IN-DEPTH and the other with the corresponding transcriptomics platform alone as a control **(Figure 3.3)**. Both slides were subsequently stained with hematoxylin and eosin (H&E) for morphological assessment by board-certified pathologists. As expected, the quality of spatial proteomics staining and imaging was preserved across all cases. Importantly, spatial transcriptomics profiling demonstrated a strong positive correlation between IN-DEPTH and control slides for all platforms ($R > 0.9$) except for Orion-GeoMx ($R = 0.692$), and comparable amounts of genes were captured throughout. Non-binding control probes also remained consistently low, indicating minimal background. These results demonstrate that IN-DEPTH is robust, generalizable, and platform-agnostic. The CODEX–GeoMx combination was selected for further development: the CODEX provides rapid whole-slide imaging whereas the GeoMx provides genome-wide transcriptome profiling with precise regional selectivity, thus maximizing efficiency and cost-effectiveness when used together.

The ability for IN-DEPTH to identify biologically meaningful features was assessed using the CODEX-GeoMx implementation on tonsil tissue sections. A 12-plex antibody panel targeting key cell populations was applied for CODEX imaging, followed by genome-wide transcriptome capture on the GeoMx **(Figure 3.4)**. A point to note is that the default GeoMx workflow[129] relies on overlapping fluorescence signals from the three selected antibody markers to delineate cell populations, leading to a maximum of $2^3 = 8$ cell populations if the antibodies are not mutually exclusive. However, this approach is highly prone to misclassification of cell phenotypes due to intrinsic tissue autofluorescence that cannot be fully quenched. Upon further investigation, the GeoMx was found to offer the ability to incorporate customized segmentation masks to guide the instrument on the capture areas. Further exploration of this feature revealed that at least 30 customized masks can be accommodated

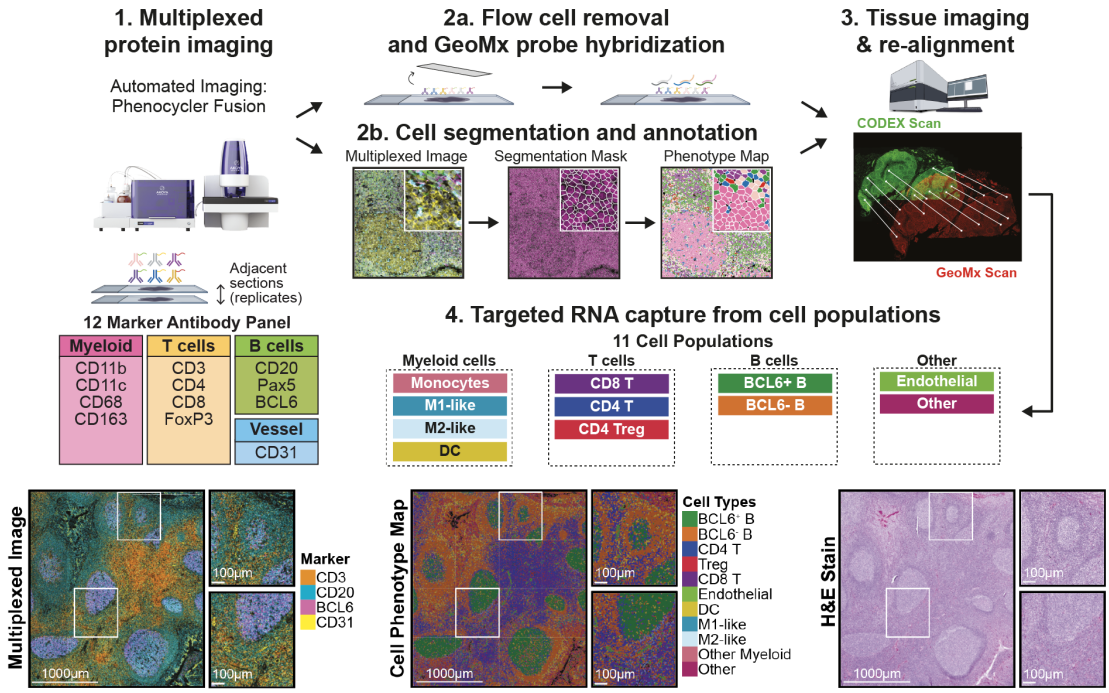


Figure 3.4: Application of CODEX-GeoMx on tonsil tissue sections to assess the accuracy of captured biological features. Cells that could not be classified using these markers were annotated as "Others". Adapted from Yeo et al., 2024[189].

within each region of interest. This feature was leveraged to its fullest capacity to circumvent the limitation of the default workflow, enabling the incorporation of single-cell segmentation and annotated cell populations from the CODEX images directly on the GeoMx. Cell phenotypes from spatial proteomics imaging were thus identified through the following steps:

1. *Cell segmentation:* Individual cell boundaries were identified using the deep learning-enabled cell segmentation algorithm Mesmer[77], as deep learning-based models consistently outperform traditional geometry-based techniques[34]. As Mesmer requires a membrane and nuclear channel, here, the membrane channel was generated by combining markers that collectively maximized the coverage of cell boundaries.

2. *Single-cell feature extraction:* For each marker, the pixel values within individual cell bound-

aries were summed and normalized by the cell area. The resulting cell-size scaled value was used as the expression value of a given marker.

3. *Clustering and annotation:* Extracted features were scaled to a standardized [0,1] range, and phenotyping markers were clustered using Phenograph [116]. The resulting clusters were annotated by examining the enriched markers, which were visualized through Z-score enrichment heatmaps and spatial overlays onto the original CODEX images. Clusters lacking clear enrichment patterns were re-clustered, and this iterative clustering and annotation process was repeated until no additional cell populations could be resolved.

4. *Confirmation:* Specificity and accuracy of cell type annotations were assessed by comparing with staining patterns on the original images and H&E-stained sections from the same or adjacent tissue section. Final confirmatory assessments were provided by board-certified pathologists prior to transcriptome capture.

Through this approach, a total of 11 cell populations were identified from the tonsil tissues, with the specificity and accuracy of cell phenotyping and transcriptome capture demonstrated through the expected enrichment of antibody markers and orthogonal gene signatures [127] within each respective population **(Figure 3.5)**. Importantly, this approach provides the additional benefit of capturing cell boundaries with greater precision compared to the vanilla GeoMx approach, thereby significantly improving transcriptome data quality. While it may be tempting to maximize the number of captured cell populations, this comes at the cost of reduced signal-to-noise due to smaller capture areas. For this reason, the CODEX–GeoMx approach is currently recommended to be limited to approximately 10 cell populations.

The spatial and functional dynamics of lymphocytes in the tonsillar tissue architecture was then examined by localizing CD4 T follicular helper (Tfh) cells despite the absence of Tfh-specific markers such as PD-1 or CXCR5. Tfh cells are known to migrate into B follicles during activation and mat-

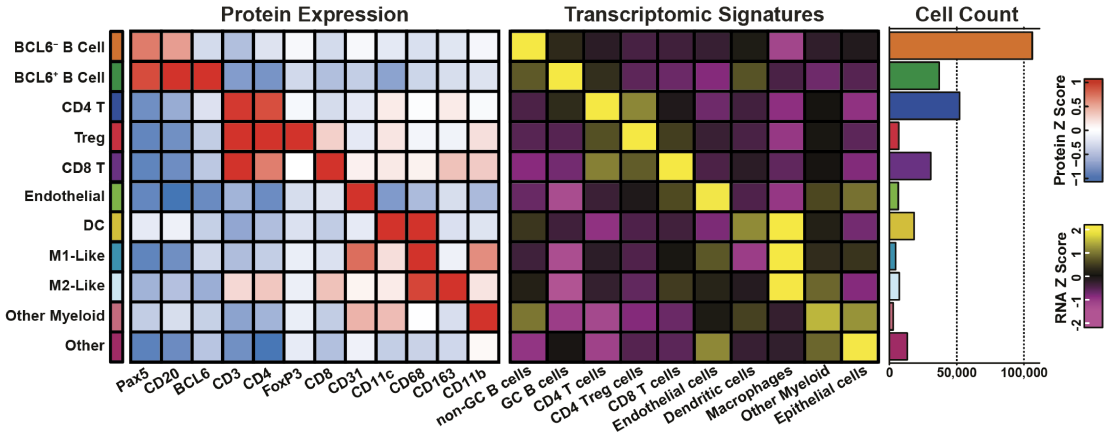


Figure 3.5: Heatmap of antibody expression level from CODEX data (left) and RNA signature enrichment from GeoMx data (middle) from each of the annotated cell population, as well as their relative abundance across the capture areas (right). Figure from Yeo et al., 2024 [189].

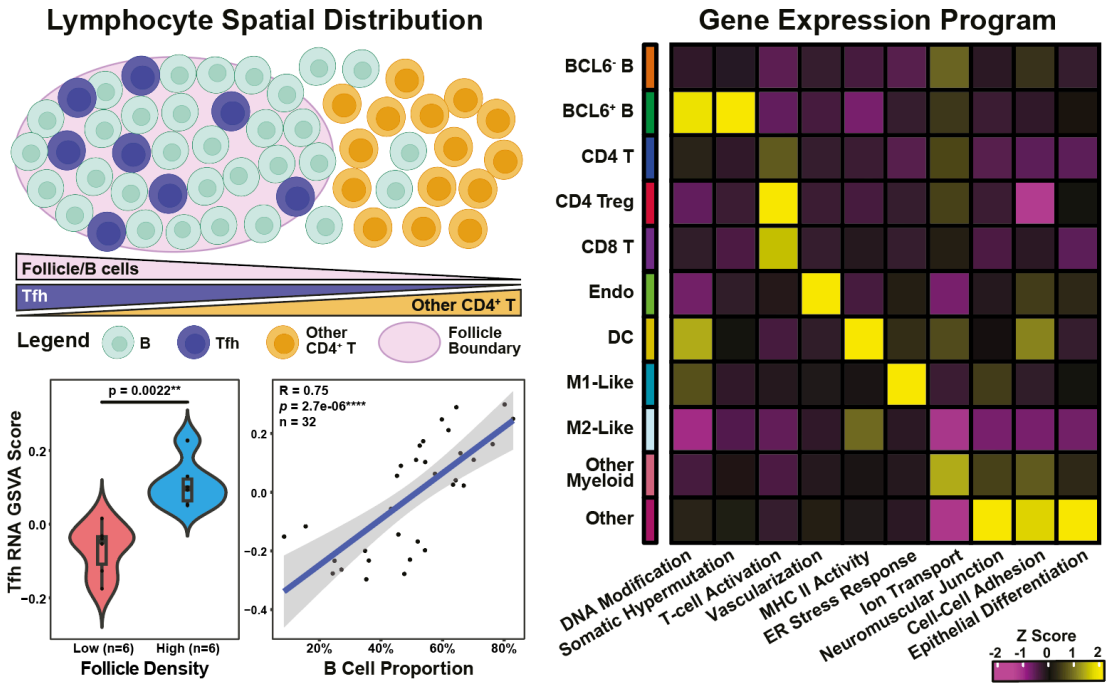


Figure 3.6: Capture of Tfh signatures within their correct spatial context (left) and biologically coherent factorized gene expression programs (right). Figure from Yeo et al., 2024 [189].

uration [127]. Consistent with this, Tfh gene signatures were significantly enriched in capture regions dense in follicles, and also positively correlated (R = 0.75) with increasing B-cell abundances across the capture regions, reflecting Tfh migration and maturation dynamics in the tonsil. Tissue-wide cell type–specific transcriptional programs were then characterized using consensus non-negative matrix factorization [113] to infer predominant gene expression programs for each cell type. Ten gene expression programs were identified and annotated with Gene Ontology Biological Process signatures [10], each exhibiting cell type–specific distributions aligning with their expected functional activities in the tonsil [127] **(Figure 3.6)**.

The ability to achieve reliable spatial and molecular profiling, from precise cell type identification to the capture of transitional states and tissue-wide transcriptional programs, establishes IN-DEPTH as a robust platform for deep multi-omics investigation of tissue biology. Beyond elucidating detailed cellular and molecular profiling within native spatial contexts, IN-DEPTH also enables the generation of data essential towards advancing computational methods for integrated spatial multi-omics analyses.

3.2 Development of SGCC for integrative spatial multi-omics analysis

Spectral Graph Cross-Correlation (SGCC) [189], a mathematical formulation developed by Yuzhou Chang that and built upon graph signal processing, provides a framework to quantitatively assess how the spatial organization of cell types relates to functional dynamics, and to dissect their coordinated changes. As any spatial omics feature (e.g. cell phenotype label) can be modeled as a graph signal [30], SGCC leverages the unbiased representation and interpretability of Graph Fourier transform to characterize the distributional relationships between pairs of cell phenotypes. The underlying graph, when defined as a lattice pixel graph or an irregular cell graph, allows Graph Fourier transform to project vertex-domain signals into the frequency domain via Fourier modes to yield interpretable

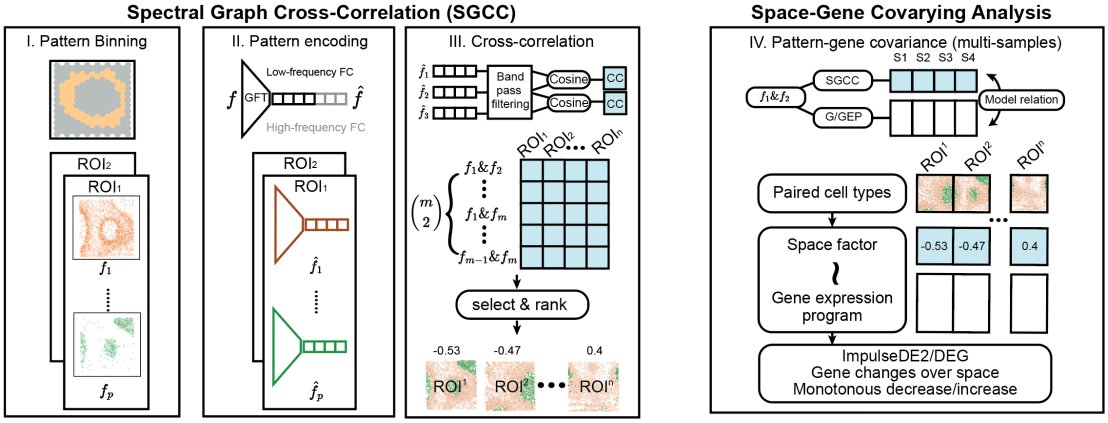


Figure 3.7: Schematic overview of the SGCC methodology developed by Yuzhou Chang from Qin Ma's lab at the Ohio State University. Single-cell patterns from spatial proteomics imaging are binned and encoded through graph Fourier Transform to generate low-frequency Fourier coefficients; this is followed by cross-correlation analysis to identify coordinated spatial patterns (left). The genes that co-vary with these spatial patterns are then integrated with models such as ImpulseDE2 and edgeR to link spatial factors with gene expression for functional analysis (right). Figure from Yeo et al., 2024 [189].

Fourier coefficients. The first k low-frequency modes capture spatially organized components of the graph signal [59,60], providing the basis for quantifying correlations between cell phenotypes in the frequency domain through the similarity of their k-bandlimited Fourier coefficients. Negative SGCC values indicate reduced spatial co-occurrence, whereas positive values reflect increased co-occurrence between cell populations. These values can then be tied to gene expression by identifying genes that co-vary with cellular spatial factors, which can be achieved with ImpulseDE2 [67] that models SGCC as a continuous covariate, or edgeR [162] that treats them as ordinal factors, thereby enabling the rapid identification of spatially dynamic genes **(Figure 3.7)**.

The performance of SGCC was benchmarked across a simulation of 80 datasets, each representing a 60 × 60 pixel graph with ring-like distributions of two cell phenotypes. These simulations varied in area and complementarity, reflecting both global and local patterns. SGCC scores increased under locally complementary patterns but decreased under globally complementary ones, indicating that SGCC effectively distinguishes between different spatial organizations. Another set of 80 simulated graphs further demonstrated that SGCC can also discriminate differences in area and spatial proximity

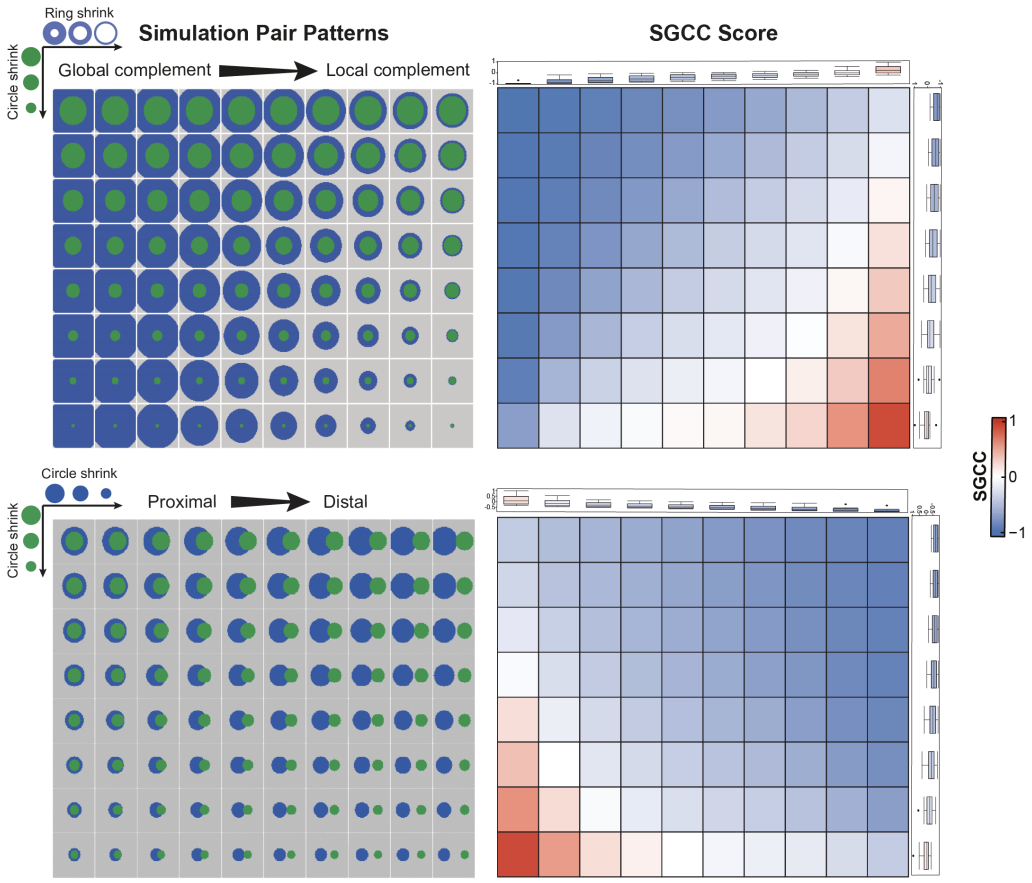


Figure 3.8: Systematic validation of SGCC using simulated spatial patterns, demonstrating its ability to detect transitions from global to local complement states (top) and from proximal to distal patterns (bottom). Adapted from Yeo et al., 2024 [189].

between cell phenotype patterns **(Figure 3.8)**. These highlight SGCC's potential as a sensitive tool for detecting spatial relationships across heterogeneous spatial contexts.

The applicability of SGCC on spatial multi-omics datasets was evaluated on the CODEX-GeoMx tonsil dataset, with the goal to resolve nuanced cell state transitions between CD4 T-cells and BCL6-positive B cells, key regulators of germinal center reactions [172]. Increasing SGCC scores between these cell populations reflected coordinated changes in tissue morphology and cellular organization, including elevated cytokine production, enhanced B-cell MHC Class II expression and chromatin reorgan-

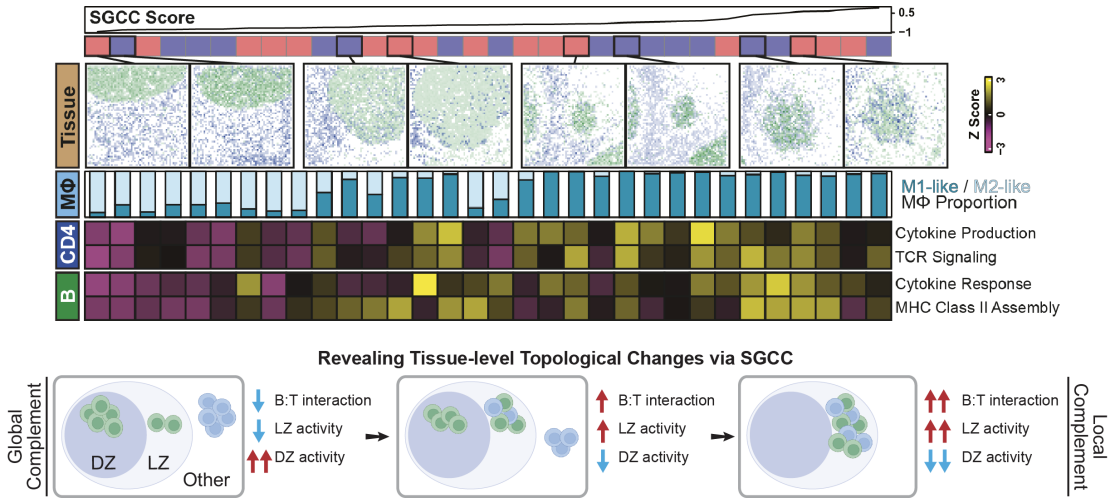


Figure 3.9: Assessment of CD4 T-cell and BCL6-positive B-cell dynamics based on SGCC scoring, showing coordinated changes in cellular spatial distribution, macrophage polarization states (M1/M2-like proportion), and gene expression programs (top). These results depicting the transitions in T-B cell interactions across the dark zone (DZ) and light zone (LZ) in germinal centers (bottom), demonstrating the ability for SGCC to to derive tissue-level organizational features. Adapted from Yeo et al., 2024 [189].

zation, and stronger T-cell receptor activation and translation initiation. Furthermore, macrophages, although not scored in SGCC, also showed an increasingly activated profile. These transitions coincided with a shift from global to local complementary spatial patterns, with low SGCC regions aligning with self-aggregation of T- and B-cells, and high SGCC regions aligning with B-T cell crosstalk akin to light zone interactions **(Figure 3.9)**. These results demonstrate the unique insights enabled by SGCC to reveal spatially coordinated transitions in cellular functional states beyond the capacity of either spatial modality alone.

While spatial relationships provide a valuable proxy for cellular interactions, many biological processes are ultimately governed by direct molecular interactions. For example, exertion of T-cell exhaustion through immune checkpoints involve receptor–ligand engagement and intracellular signaling cascades that occur at subcellular scales [201,202,11] and cannot be inferred from the distance between individual cells alone. Capturing these events *in situ* would therefore provide a complementary molec-

ular readout, enabling spatial methods to move beyond descriptive mapping toward mechanistic insights into molecular communications underlying spatially coordinated processes.

3.3 Incorporation of proximity labeling to spatial proteomics

Proximity labeling technologies enable the detection of protein–protein interactions directly within living cells and intact tissues by labeling the interacting partners with a measurable signal [157]. Among the most established methods is the Proximity Ligation Assay (PLA) [69] that consistently achieves high specificity and sensitivity in immunohistochemical and immunocytochemical approaches. PLA leverages pairs of oligonucleotide-conjugated antibodies that initiate a DNA amplification cycle only when the antibody pair simultaneously bind epitopes in close proximity (<40 nm). This double-recognition strategy, combined with rolling circle amplification of the hybridized oligonucleotide pair, produces a long DNA product that can be visualized by probes labeled with either a fluorophore or horseradish peroxidase. The resulting fluorescent puncta or chromogenic signal enables visualization of molecular interactions at a single-molecule resolution.

Despite the capability of PLA, it has not been integrated into high-dimensional spatial proteomics. Such incorporation offers the unique opportunity to bridge the gap between measuring the relative distance between cells and directly capturing the molecular interactions that underlie their communication. However, careful optimization was necessary, as PLA was developed long before the advent of high-dimensional spatial proteomics. First, PLA antibodies must be applied prior to staining the spatial proteomics antibody panel, as simultaneous application would cause interference with secondary anti-mouse and anti-rabbit reagents. Second, adjustments must be made to accommodate differences in assay buffers. Third, the detection strategy for PLA must be considered. Two complementary PLA detection strategies were optimized:

1. *Fluorophore-labeled probes:* PLA signals were imaged directly alongside a nuclear stain. Sig-

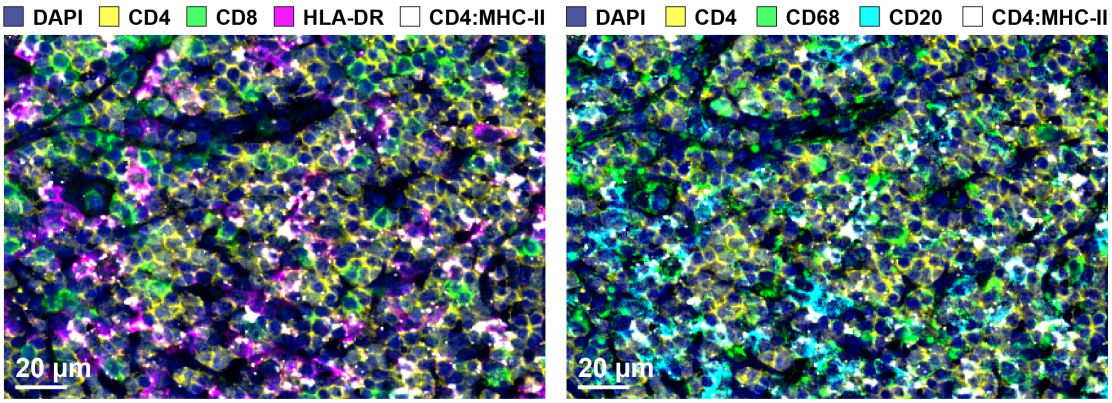


Figure 3.10: Proximity labeling on CODEX spatial proteomics. These representative images are part of a 30-marker panel applied on a human tonsil tissue, with PLA incorporated to detect CD4:MHC-II interactions. Notice the white punctum localized to HLA-DR and CD4 but not CD8 boundaries (left), which correspond to interactions with macrophages and B cells (right).

nals were then removed by incubating the tissue in a warm buffer (37-40°C) containing 90% dimethyl sulfoxide, which lowers the melting temperature and efficiently displaces the rolling circle amplification product along with the fluorophore. The PLA image will later be aligned to the spatial proteomics images via the cross-platform image alignment approach outlined above.

2. *Horseradish peroxidase-labeled probes:* PLA was coupled to tyramide signal amplification[22], depositing either a tyramine-conjugated hapten or DNA oligonucleotide barcode directly on the tissue[179,133]. The hapten can be detected by including the cognate antibody in the spatial proteomics panel, whereas the DNA barcode is directly compatible with the CODEX workflow. PLA signals will then be captured during spatial proteomics imaging.

Taken together, and building on the optimization principles established for IN-DEPTH, PLA was successfully integrated into the CODEX spatial proteomics workflow **(Figure 3.10)**[191], demonstrating the feasibility of coupling proximity labeling with spatial proteomics. These optimizations are also expected to integrate seamlessly across other spatial proteomics platforms.

3.4 Improvements to streamline antibody panel development

A key requirement for spatial biology is access to antibodies that are specific and compatible with tissue staining. While many commercial antibodies exist, their performance in FFPE tissues is often unreliable, as formalin fixation alters protein epitopes. Additionally, antibodies labeled as FFPE compatible frequently show inconsistent results, as antigen retrieval conditions vary across vendors. To address these challenges, a streamlined three-step workflow was developed:

1. *Immunohistochemical selection:* Commercial antibodies were screened on FFPE tissues under standardized antigen retrieval conditions as described above to ensure consistency, with reference tissues identified using the Human Protein Atlas resource[196]. To minimize non-specific staining, blocking and secondary antibodies were chosen from the same species. For poorly characterized targets (especially for viruses), paired control cell pellets were generated (e.g. virus-positive *vs.* virus-negative, overexpression *vs.* knockout), embedded into a supporting gel matrix, and processed into FFPE blocks. These FFPE cell pellets were used for the first-line assessment of antibody specificity prior to evaluation on actual FFPE tissues. Only antibodies that pass this immunohistochemical screen were advanced to conjugation.

2. *Antibody conjugation:* Spatial proteomics platforms rely on uniquely barcoded antibodies, but the conjugation process can destabilize antibody structure and compromise specificity. To mitigate risks, multiple conjugation strategies were evaluated:

    (a) *Maleimide chemistry:* Antibodies contain cysteine residues vital for forming stable disulfide bonds that stabilize their bimonoeric structure to maintain function. These cysteine residues can be reduced to expose sulfhydryl (thiol) groups, which can undergo nucleophilic attack on the electrophilic maleimide moiety linked to barcodes, forming stable thioether linkages[90]. While highly robust chemically, this approach requires antibody

reduction, which destabilizes its native structure and increases the risk of degradation and irreversible loss of specificity.

(b) *Click chemistry:* As antibodies are rich in primary amines such as lysine side chains, N-hydroxysuccinimide (NHS) ester can first be used to conjugate reactive azide groups to the antibody. This reaction proceeds efficiently under mildly alkaline conditions as lysine residues existing beyond the isoelectric point result in a significant proportion of deprotonated amines. These amine nucleophiles can then attack the electrophilic NHS ester to form stable amide bonds[90]. Afterwards, the azide groups can react with barcodes linked to dibenzocyclooctyne (DBCO) through the strain-promoted azide–alkyne cycloaddition (SPAAC) reaction that produces stable triazole linkages. A key benefit of this "click" reaction is its copper-free and bioorthogonal nature while proceeding rapidly under physiological conditions[180]. Although the SPAAC step is highly specific and minimally disruptive, antibody stability is still affected by the initial NHS ester coupling step.

(c) *Direct Fc conjugation:* As Protein A from *Staphylococcus aureus* has high affinity for the Fc region of antibodies, this property was exploited in the the oYo-link approach[97,96,95], which enables light-activated, site-specific conjugation to the Fc domain through an engineered antibody-binding protein. This protein, Protein Z, was derived from Protein A and incorporates a UV-active benzophenone cross-linker within its Fc-binding domain, as well as a functional moiety at its C-terminus. For the oYo-link method, barcodes are first conjugated to Protein Z (via maleimide or click chemistry) and subsequently photolinked to the Fc region of the antibody. While this avoids direct antibody destabilization, it is restricted to immunoglobulin G antibodies, yields lower degree of labeling that lead to weaker signals, and further limited by the potential degradation of Protein

Z after barcode conjugation.

3. *Imaging validation:* Conjugated antibodies were tested using the desired multiplexed imaging workflow to assess retention of epitope specificity. Ideally, the tissue type used would be the same as used for initial immunohistochemical staining to enable direct comparison. Successful antibodies were retained in the panel, while those failing validation were either re-conjugated using an alternative chemistry or replaced with a different clone, after which the workflow was repeated.

The antibody panels used herein were expanded upon on a published reference set [152] to cover additional relevant targets. Immunohistochemical and immunofluorescence stainings were performed on tissues generously provided by collaborators at the Dana-Farber Cancer Institute. Given the laborious and resource-demanding nature of antibody panel development, antibody clones and conjugation strategies were reported with each manuscript, thereby providing valuable resources for the pathology and spatial biology communities.

Altogether, these advances establish a framework where spatial proteomics and transcriptomics can not only be registered within the same tissue section using IN-DEPTH, but also integrated with computational approaches such as SGCC and molecular assays such as PLA to resolve spatial organization across multiple levels. This framework also serves as the foundation for the mechanistic dissection of how Epstein-Barr virus remodels the lymphoma tumor microenvironment.

All truths are easy to understand once they are discovered;

the point is to discover them.

Galileo Galilei

4

Reprogramming the Dysfunctional Immune Landscape of Lymphomas

Fatal diseases of the lymph nodes has been described for centuries but were usually attributed to tuberculosis, one of the deadliest infectious diseases known to mankind. In 1832, Thomas Hodgkin reported the clinical histories and postmortem findings of seven patients characterized by the unusual,

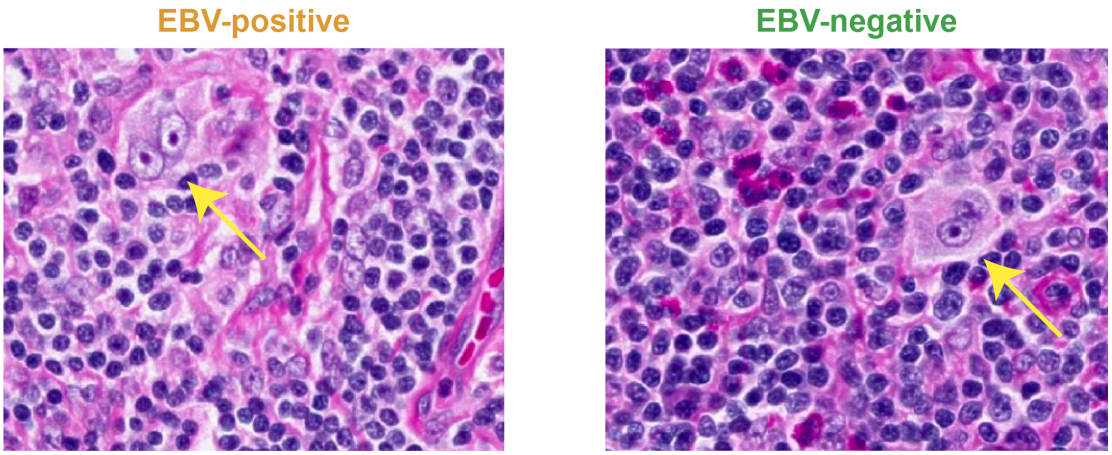


Figure 4.1: H&E staining of EBV-positive (left) and EBV-negative (right) cHL samples. Yellow arrows point to the distinctive giant multinucleated HRS cells which are morphologically indistinguishable by EBV status. Viral status is determined by in-situ hybridization of the viral non-coding RNA *EBER1*.

progressive enlargement of lymph nodes and spleens — but in patterns distinct from tuberculosis, syphilis, and other known causes of glandular swelling at the time[91]. Although initially overlooked, the condition was rediscovered in 1856 by Samuel Wilks[203], who provided further clinical descriptions and introduced the eponym "Hodgkin's Disease", thereby establishing classic Hodgkin Lymphoma (cHL) as a distinct entity.

cHL became a key subject of investigation over the next century, with breakthrough discoveries closely paired with technological advances. The microscopic identification of giant multinucleated malignant cells by Carl Sternberg in 1898[184] and Dorothy Reed in 1902[160], now termed Hodgkin and Reed–Sternberg (HRS) cells, established the pathognomonic hallmark of the disease **(Figure 4.1)**. Soon after, application of newly-discovered X-rays[167] led to the first documentation of tumor regressions in 1902[156], and by the 1950s, curative radiation therapies were available[151,103]. This was followed by the development of multi-drug chemotherapy regimens, from MOPP[55] (Mustard, vincristine (a.k.a. Oncovin), Procarbazine, Prednisone) in 1964 to ABVD[24] (doxorubicin (a.k.a. Adriamycin), Bleomycin, Vinblastine, Dacarbazine) in 1975, with the latter remaining the first-line treat-

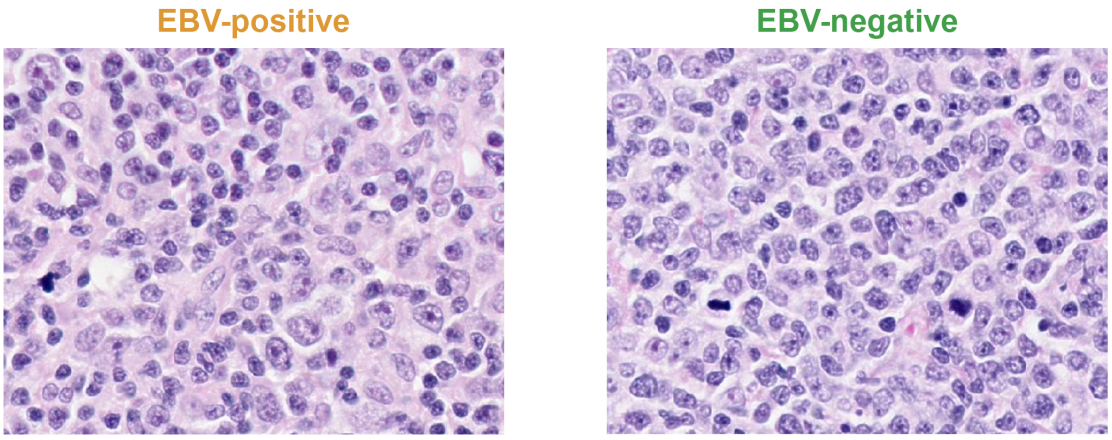


Figure 4.2: H&E staining of EBV-positive (left) and EBV-negative (right) DLBCL samples, each showing a characteristic diffuse sheet of atypical lymphoid cells. Malignant cells are not morphologically distinguishable based on EBV status like in cHL, but they can also be difficult to differentiate from non-malignant activated or germinal center B cells. Viral status is also determined by in-situ hybridization of *EBER1*.

ment today. Once a death sentence, cHL now stands among the greatest medical successes, with the majority of patients curable through a combination of modern treatments that include chemotherapy, radiotherapy, and immunotherapy.

The pioneering efforts in cHL research laid the foundation for the recognition of other lymphomas. Over the decades following the discovery of cHL, several additional lymphoid malignancies with distinct clinical and pathological features were identified, leading to the division of lymphomas into two broad categories: Hodgkin lymphoma and non-Hodgkin lymphoma. While this division is largely historical, it is still relevant due to their distinct pathobiology and clinical presentation. Among the non-Hodgkin lymphomas, the most common and clinically significant is Diffuse Large B-Cell Lymphoma (DLBCL) **(Figure 4.2)** that comprises ~30% of cases.

DLBCL is a highly heterogeneous lymphoma that encompasses diverse molecular subtypes that differ in pathogenesis and clinical outcomes[32]. Initially classified under broad morphological terms such as "reticulum cell sarcoma" or "diffuse histiocytic lymphoma"[159], DLBCL was eventually consolidated as a distinct entity through advances in genotyping, immunophenotyping, and classification

frameworks[85]. Its treatment history paralleled these advances, progressing from CHOP (Cyclophosphamide, doxorubicin (a.k.a. Hydroxydaunorubicin), vincristine (a.k.a. Oncovin), Prednisone) in the 1970s[128] to R-CHOP with the addition of the anti-CD20 antibody rituximab in the 1990s[46], which remains the modern standard of care. However, unlike cHL, DLBCL remains a formidable challenge, with outcomes that vary widely and a considerable fraction of patients experiencing relapse or refractory disease.

cHL and DLBCL are aggressive B-cell lymphomas when untreated, and both present unique immunopathological features. Their incidence are significantly higher in immunocompromised patients[181], and dysfunctional non-malignant immune cells constitute a substantial proportion of the tumor mass at >95% for cHL[47] and often exceeding 50% for DLBCL[204]. Therapeutically, cHL responds dramatically to PD-1 checkpoint blockade[7,9] while DLBCL is sensitive to immune-directed therapeutics including rituximab and CD3×CD20 bispecific antibodies[46,193] despite being poorly responsive to PD-1 blockade[8]. In parallel, PD-L1, a major inhibitor of T-cell effector activity by binding to PD-1, is overexpressed in cHL[166,29] and frequently upregulated in DLBCL[206]. These observations support the view that cHL and DLBCL are malignancies in which immune dysfunction, in part through sustained immunoevasion within an abnormal immune microenvironment, plays a central role in lymphomagenesis.

Notably, cHL and DLBCL are associated with EBV, which is present in ~25% of cHL and 5–15% of DLBCL cases[2]. Emerging evidence indicates that EBV status influences clinical outcomes, with EBV-positive cHL recently linked to a higher risk of relapse after chemotherapy[88] and EBV-positive DLBCL generally associated with poorer prognosis[121]. Furthermore, the incidence of cHL and DLBCL is markedly higher in the immunocompromised, where almost all HIV-associated cHL and a considerable proportion of AIDS-related DLBCL EBV-positive[27]. These findings led to the hypothesis that EBV contributes to spatial remodeling of the tumor microenvironment in cHL and DLBCL to augment immune dysfunction within.

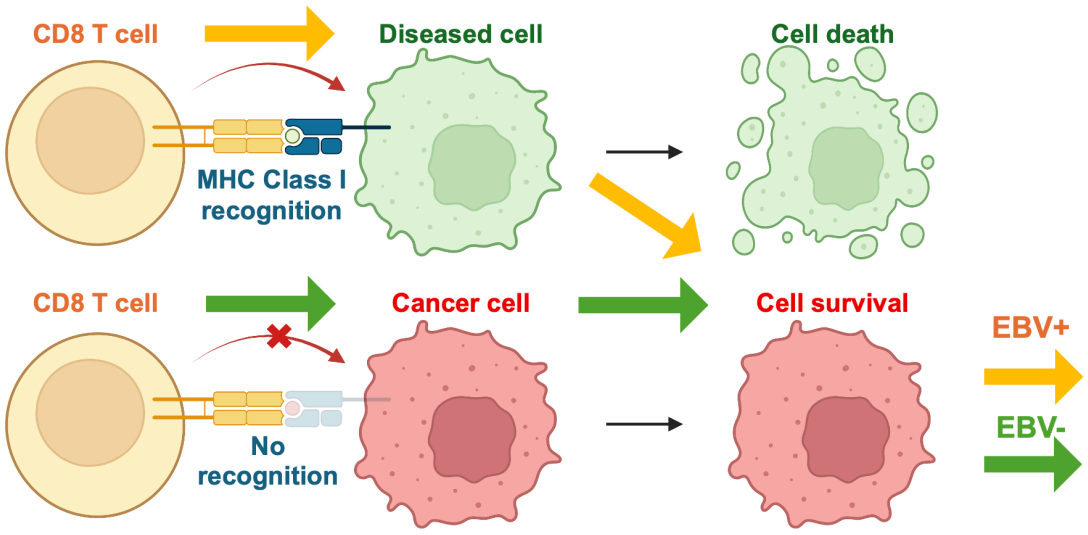


Figure 4.3: In many cancers, malignant cells downregulate MHC Class I to evade CD8 T-cell recognition. However, EBV-positive HRS cells frequently retain MHC Class I, suggesting a key role for EBV in reprogramming the cHL tumor microenvironment to effectively impair CD8 T-cell function.

4.1 EBV and T-cell dysfunction in the cHL tumor microenvironment

There remains a strong rationale to further understand the pathobiology of cHL, as 10–20% of patients continue to suffer from relapse or refractory disease even after first- and second-line therapies [1]. With a peak incidence in young adults (<35 years old)[47], cHL continues to cause a disproportionate loss of productive life years in affected individuals. These highlight a need to understand the mechanisms that support cHL progression, particularly the contribution of EBV.

A distinctive feature of EBV-positive HRS cells is the retention of MHC Class I expression, which is frequently absent in EBV-negative HRS cells due to the loss of beta-2 microglobulin (B2M) associated with higher mutational burdens [146,138,164,165,88]. This is consistent with the ability for EBV-positive HRS cells to present latent EBV oncogene antigens [115] and a purported increase in CD8 T-cell activation within the EBV-positive cHL tumor microenvironment [205]. The retention of MHC Class I is unusual, as its downregulation is frequently exploited by malignant cells across numerous different

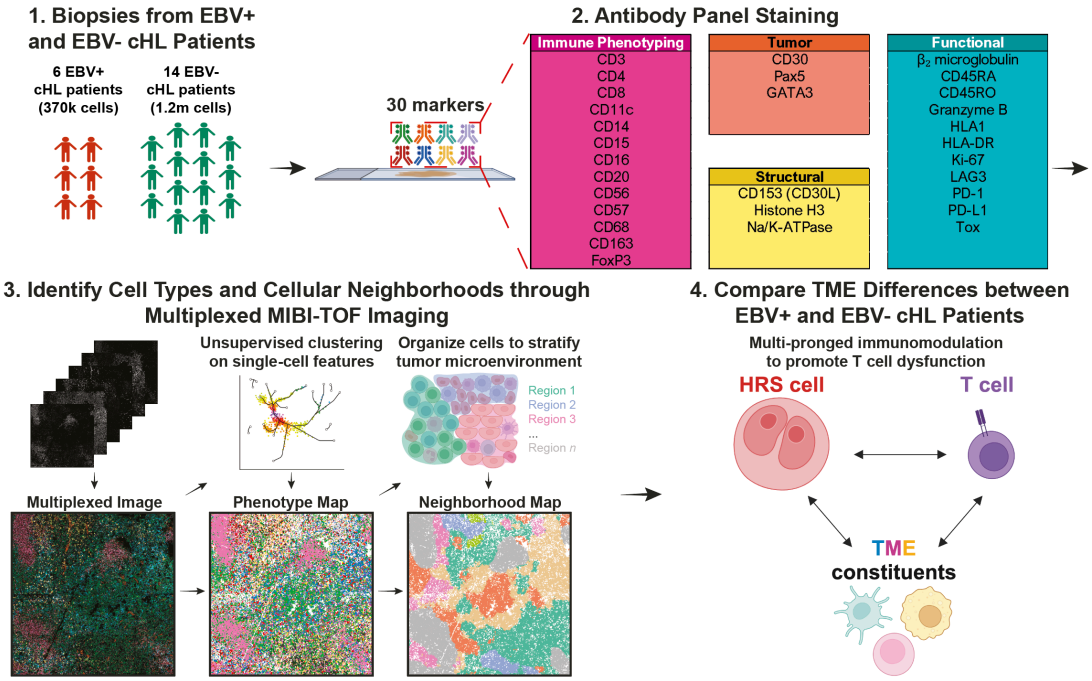


Figure 4.4: MIBI spatial proteomics workflow to dissect the EBV-positive and EBV-negative cHL tumor microenvironemnt. Biopsies from an initial cohort of 6 EBV-positive and 14 EBV-negative cHL patients were collected (1), stained with cocktail of 30 lanthanide-conjugated antibodies (2), imaged using the MIBI to generate multiplexed images that allowed for the identification of cell phenotypes and neighborhoods (3), where upon computational analysis illustrated key differences between EBV-positive and EBV-negative cHL (4). TME, tumor microenvironment. Figure from Yeo et al., 2024 [191].

cancers to evade CD8 T-cell recognition and impair responses to immunotherapies [56] **(Figure 4.3)**. This suggests that EBV-positive cHL must employ mechanisms to suppress CD8 T-cell responses, contributing to the immunosuppressed tumor microenvironment characteristic of cHL.

A panel of 30 antibody markers was developed to visualize the EBV-positive and EBV-negative cHL tumor microenvironment using the MIBI. This panel included key phenotyping and functional markers to enable a focused assessment of T-cell responses. Antibodies were meticulously titrated and optimized on reactive lymph nodes and cHL cases, then applied to archival FFPE biopsy samples from an initial cohort of 6 EBV-positive and 14 EBV-negative patients from the Dana-Farber Cancer Institute **(Figure 4.4)**. The panel captured diverse cell populations, including M2-like macrophages

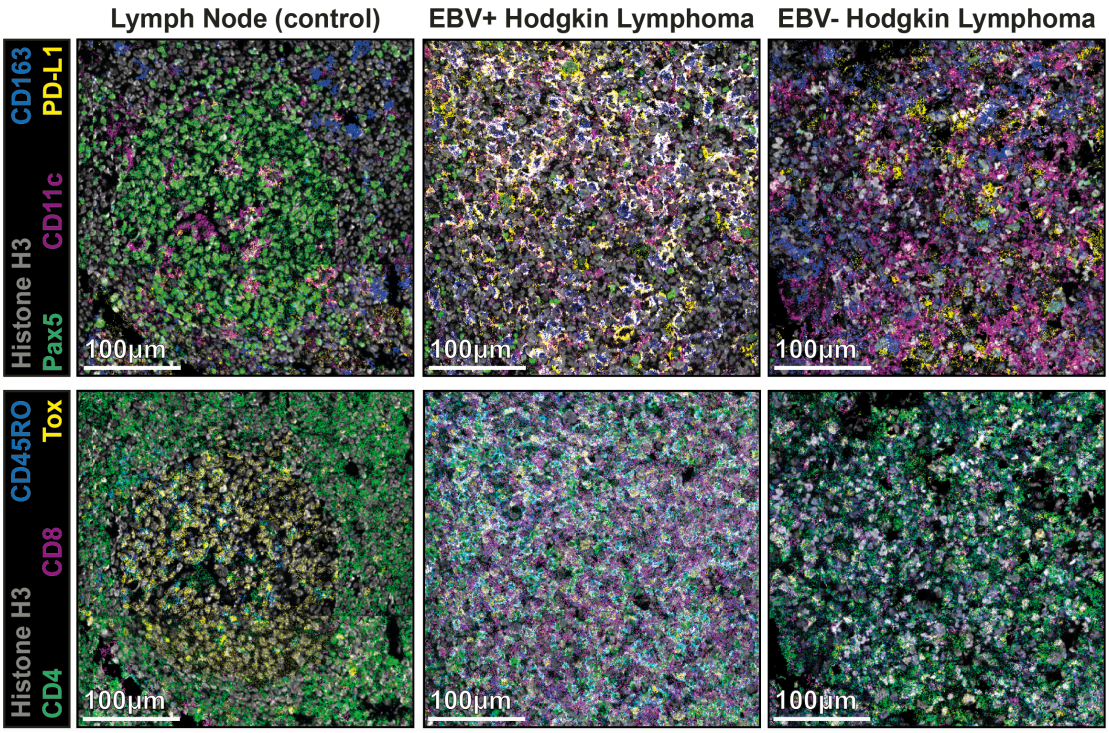


Figure 4.5: Representative MIBI images of lymph node (left), EBV-positive (middle), and EBV-negative cHL (right) tissues. Lymph nodes were used as a positive control to confirm antibody staining specificity as their intact spatial organization helps distinguish true marker expression from artifacts in the disorganized lymphoma tissue architecture. Figure from Yeo et al., 2024[191].

(CD163), B cells ($Pax5^{hi}$), HRS cells ($Pax5^{lo}$), dendritic cells (CD11c), PD-L1, and T-cell subsets (CD4, CD8) along with their memory (CD45RO) and dysfunctional (Tox) subpopulations **(Figure 4.5)**. In total, 41 stitched fields-of-view comprising 794 individual 400 × 400 µm tiles were imaged. Cells were subsequently segmented, annotated, and stringently validated as described previously.

A total of 13 cell populations were identified, with the specificity confirmed through the expected enrichment of antibody markers within each cell population and verified with board-certified pathologists. Although EBV-positive and EBV-negative cHL appear visually indistinguishable on H&E, their tumor microenvironments displayed strikingly distinct cellular profiles. EBV-positive tumors showed a significantly higher proportion of CD8 T-cells, whereas EBV-negative tumors were enriched

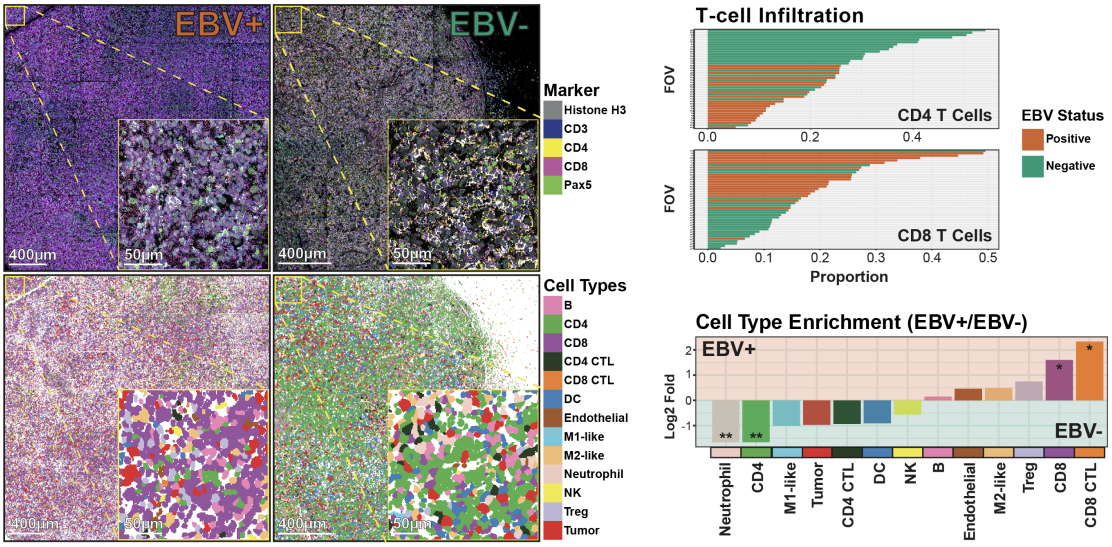


Figure 4.6: Contrasting cellular profiles between EBV-positive and EBV-negative cHL. The representative MIBI images (top left) and corresponding phenotype maps (bottom left) illustrate the visual EBV-linked differences. Further quantification of CD4 and CD8 T-cell proportions (top right) revealed consistent enrichment patterns associated with EBV status, supported by significant log-fold enrichment differences. Asterisks denote statistical significance (* $p < 0.05$, ** $p < 0.01$). Adapted from Yeo et al., 2024 [191].

in CD_4 T-cells, a pattern consistently observed across all 41 stitched fields-of-view. EBV-positive tumors also had a higher prevalence of immunosuppressive M2-like macrophages and Tregs, with a generally higher density of HRS cells in EBV-negative tumors **(Figure 4.6)**.

The abundance of CD8 T-cells, coupled with their ineffective tumoricidal function in the EBV-positive cHL tumor microenvironment, prompted an investigation into EBV-linked differences in antigen presentation **(Figure 4.7, left)**. MHC Class I (B2M, HLA1) was broadly upregulated across all cell types in EBV-positive tumors, a finding confirmed by visual inspection of raw images. In contrast, no significant difference was observed for MHC Class II (HLA-DR) expression. These features point to the need for EBV to actively suppress CD8 T-cell effector functions, and thus markers associated with T-cell functional regulation, including memory ($CD45RO$ [66]) and immune checkpoint regulators (Tox, Lag3, PD-1, PD-L1 [108,4,177,187]), were analyzed to assess the impact of EBV presence on T-cell dysfunction **(Figure 4.7, right)**. Note that although dysfunctional CD_4 and CD8

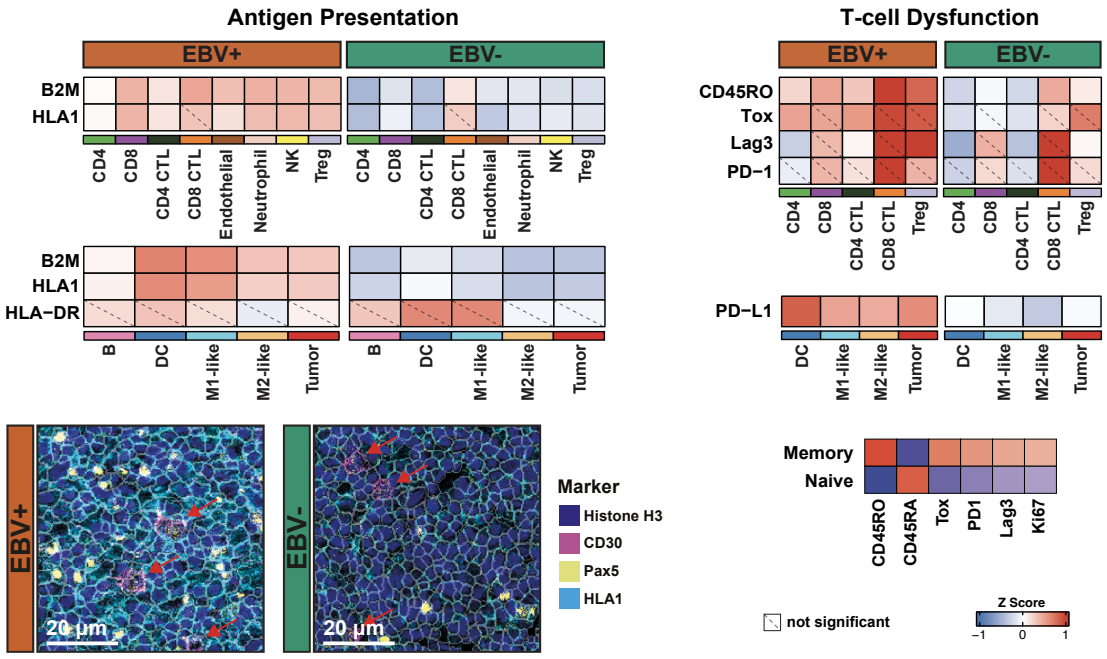


Figure 4.7: Differences in antigen presentation and T-cell functional marker expression between EBV-positive and EBV-negative cHL. In EBV-positive cHL, MHC Class I (B2M and HLA1) expression was higher across all cell types (top left) with no clear differences for MHC Class II (HLA-DR) (middle left). These findings were consistent with MIBI imaging, with a representative example (bottom left) showing HLA1 and markers for B cells ($Pax5^{hi}$) and HRS cells (CD30 and $Pax5^{lo}$, as indicated with red arrows). In parallel, T-cells displayed increased expression of dysfunction-associated markers (top right) as well as PD-L1 (middle right), with dysfunctional signatures higher in memory (CD45RO) than naïve (CD45RA) T-cells. Adapted from Yeo et al., 2024 [191].

T-cells exhibit distinct molecular profiles [48], both Lag3 and PD-1 are shared features of T-cell exhaustion [48,202,11]. An EBV-associated increase in memory and dysfunction markers on T-cells, as well as PD-L1 on HRS cells, dendritic cells, and macrophages, was observed. Further stratification of T-cells into memory and naïve populations (based on CD45RO and CD45RA) demonstrated an enrichment of dysfunction and proliferation (Ki67) markers in memory T-cells, implicating previously activated T-cells as a major source of dysfunction within the cHL tumor microenvironment.

These findings indicate that EBV amplifies immune dysfunction in the cHL tumor microenvironment by enriching immunosuppressive cell populations, increasing T-cell memory and dysfunction markers, and potentially enhancing PD-1/PD-L1 engagement between T-cells, HRS cells, and sur-

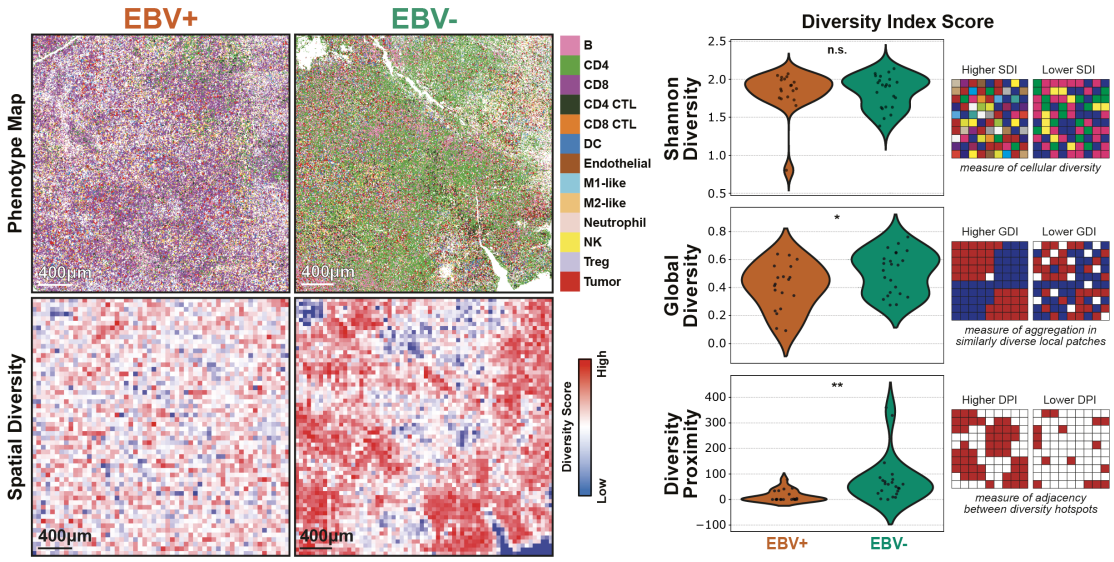


Figure 4.8: Visualization of annotated phenotype maps and their corresponding diversity maps in representative EBV-positive and EBV-negative cHL tissues (left), with three different spatial diversity index scores quantified (right). Asterisks denote statistical significance (* $p < 0.05$, ** $p < 0.01$), n.s. = not significant. Figure from Yeo et al., 2024 [191].

rounding immune constituents. The results support an EBV-dependent mechanism of suppressing T-cell effector functions through spatial reorganization and reprogramming of the cHL tumor microenvironment.

To compare the spatial organization between the EBV-positive and EBV-negative cHL tumor microenvironments, ecology-inspired diversity indices, including the Shannon [175] and MESA indices [57], were implemented **(Figure 4.8)**. Although there were no EBV-linked differences in the Shannon Diversity Index (measure of cellular diversity at the whole-tissue level), EBV-positive tumors were significantly lower in the MESA Diversity Proximity Index (measure of aggregation of similarly-diverse local patches) and the MESA Global Diversity Index (measure of adjacency between biodiversity hotspots). These data suggest that while EBV-positive and EBV-negative cHL harbored comparably similar diversity in cell composition, the EBV-positive tumor microenvironment was spatially organized to be less compartmentalized and structured, instead appearing more disrupted and homogenized with respect to local cellular diversity, as compared to the EBV-negative counterpart.

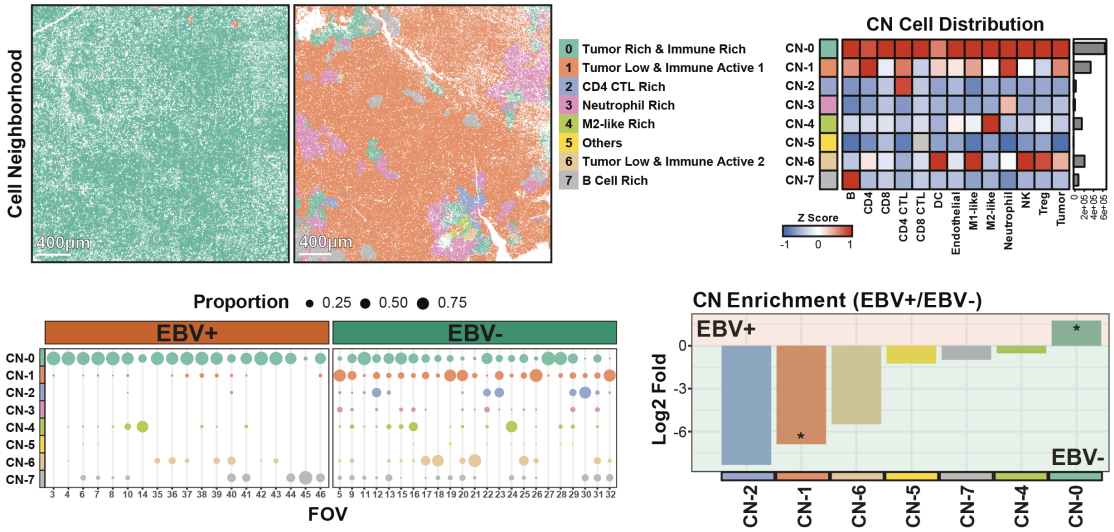


Figure 4.9: Visualization of the corresponding cellular neighborhood (CN) maps for representative EBV-positive and EBV-negative cHL tissues (top left), along with the enrichment heatmap of cell types across the 8 identified CNs (top right). Further quantification of cells assigned to each CN (bottom left) revealed contrasting enrichment patterns associated with EBV status, supported by significant log-fold enrichment differences (bottom right); note that CN3 was not shown due to its absence in EBV-positive cHL. Asterisks denote statistical significance (* $p < 0.05$). Adapted from Yeo et al., 2024[191].

The cHL tumor microenvironment was next compartmentalized into eight different cellular neighborhoods (CN) using the spatial Latent Dirichlet Allocation approach[39] **(Figure 4.9)**. HRS cells were predominantly localized within CN-0, CN-1, and CN-6, each associated with distinct T-cell contexts. Of key interest is CN-0 (Tumor Rich & Immune Rich), characterized by the high diversity of T and immune cell populations. The predominance of the highly heterogeneous CN-0 in EBV-positive tumors was also consistent with the spatial diversity index results, underscoring the robustness of the identified CNs. In contrast, EBV-negative tumors contained a broader mixture of CNs, including CN-1 (Tumor Low & Immune Active 1) and CN-6 (Tumor Low & Immune Active 2) that were relatively lower in T-cells. As CN-1 was uniformly present across all EBV-negative tumors unlike CN-6, subsequent analyses focused on CN-0 and CN-1 to evaluate the phenotypic and functional constituents that selectively shape tumor–immune reorganization in the EBV-positive tumor microenvironment.

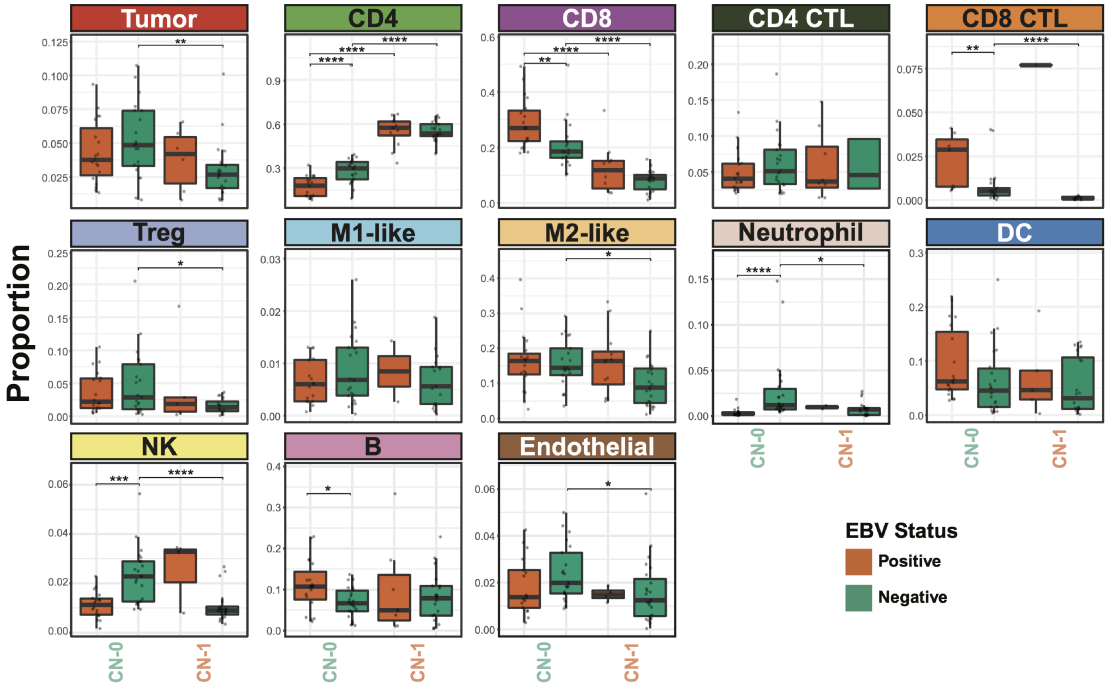


Figure 4.10: Relative proportion of the major immunomodulatory cell types within CN-0 and CN-1 and stratified by EBV status. Asterisks denote statistical significance (* p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001). Adapted from Yeo et al., 2024 [191].

Key EBV-linked differences in cellular composition were found in CN-0 and CN-1 **(Figure 4.10)**. In CN-0, EBV-positive tumors had significantly more CD8 T-cells that included cytotoxic subsets, whereas EBV-negative tumors were enriched for CD4 T-cells. There were also comparable abundances of HRS cells, regulatory T-cells, dendritic cells, macrophages, and endothelial cells between EBV-positive and EBV-negative tumors, although EBV-positive tumors showed higher B-cell and lower neutrophil and NK cell frequencies. In CN-1, cellular compositions were generally consistent between EBV-positive and EBV-negative tumors, with the exception of a modest increase in M2-like macrophages in EBV-positive tumors. A key observation is the relative stable abundance of HRS cells across CN-0 and CN-1, suggesting that functional differences of HRS cells, rather than their cell numbers, drive EBV-linked differences in immune diversity and T-cell dysfunction.

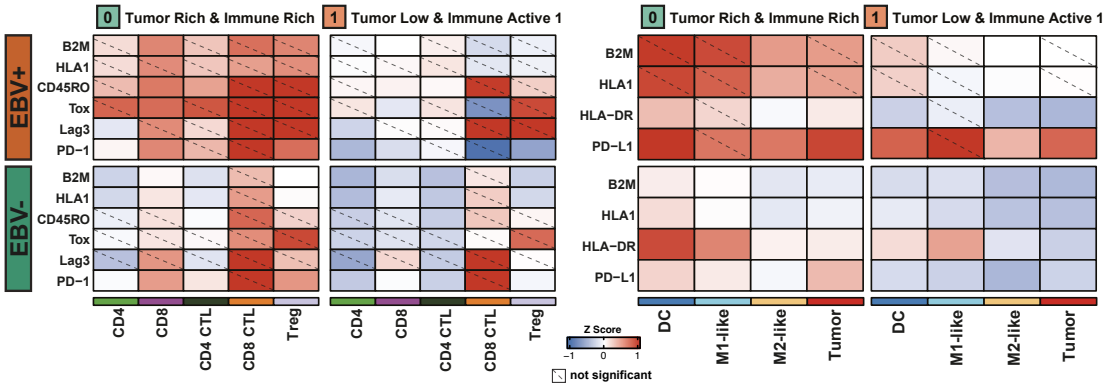


Figure 4.11: Differences in MHC Class I (B2M and HLA-1), MHC Class II (HLA-DR), memory T-cells (CD45RO), and T-cell dysfunction (Tox, Lag3, PD-1, PD-L1) expression stratified by CN and EBV status. Statistical comparisons were performed between CN-0 and CN-1 with each EBV strata. Adapted from Yeo et al., 2024[191].

Functional analysis within CN-0 and CN-1 supported and extended the tissue-level contrasts between the EBV-positive and EBV-negative tumor microenvironment **(Figure 4.11)**. EBV-positive tumors showed broad increases in MHC Class I (HLA1, B2M), heightened T-cell activation and dysfunction (CD45RO, Tox, Lag3, PD-1), and elevated PD-L1 on HRS cells, dendritic cells, and macrophages, with the most pronounced changes concentrated in CN-0. By localizing functional alterations to defined CNs, these findings demonstrated that EBV does not merely reshape the global immune composition, but also spatially reorganizes the tumor microenvironment into HRS-centered hubs that enforce T-cell dysfunction. The convergence of disordered tissue architecture, enhanced antigen presentation, and checkpoint upregulation highlights a distinct mode of immune dysfunction in the cHL tumor microenvironment, positioning CN-0 as a critical niche where EBV-positive HRS cells orchestrate immunosuppression.

The ability for HRS cells to modulate immune dysfunction within their immediate microenvironment implies that immune dysfunction in cHL is dependent on the proximity to HRS cells. To assess this, a "Tumor Score" metric was developed, which accounts for both the distance and density of nearby HRS cells. For each non-HRS cell, the "Tumor Score" was calculated as $\sum_{j \in \{k | d_{i,j} \leq r\}} \frac{1}{d_{i,j}}$,

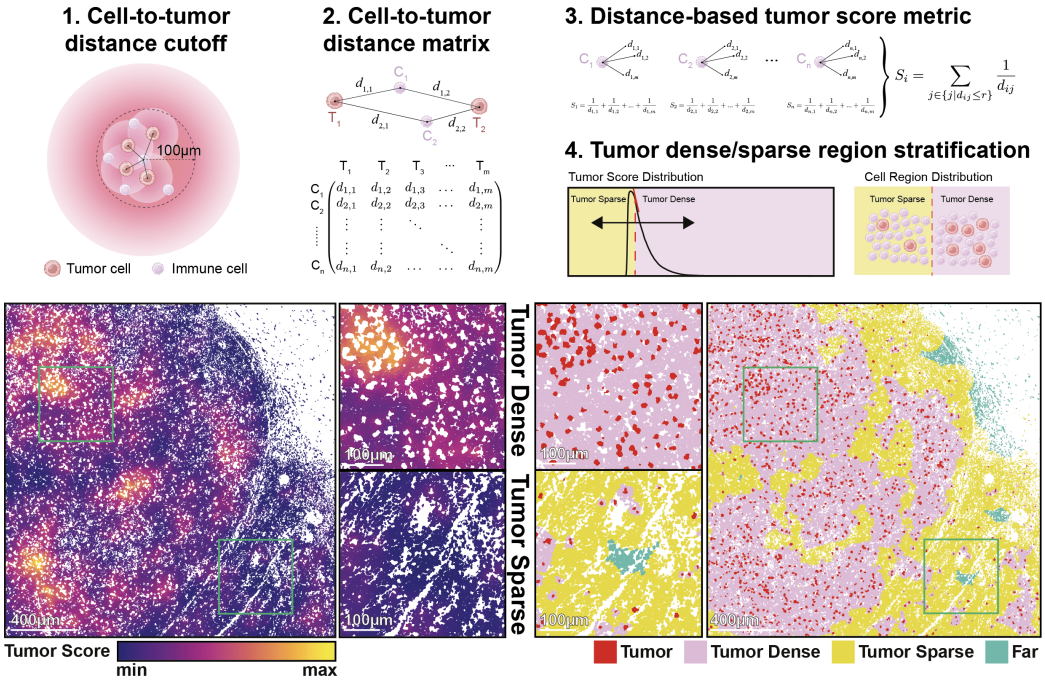


Figure 4.12: Stratification of cells into tumor-dense and tumor-sparse populations. The top panel details the workflow: cell-to-tumor distances for non-tumor cells within 100 μm from each HRS cell (1) were organized into a cell-to-tumor distance metric (2) and used to calculate the "tumor score" metric for each non-tumor cell (3); this score was then used to stratify cells into into tumor sparse and tumor dense regions based on the global tumor score distribution (4). The bottom panel shows a representative tumor score map (bottom left), along with the visualization of tumor dense and tumor sparse regions (bottom right); note that HRS cells have a white mask in the tumor score map because they do not have a tumor score. Adapted from Yeo et al., 2024[191].

where $d_{i,j}$ denoted the distance from a non-HRS cell i to a HRS cell j, and r was the search radius. Cells were then stratified into "Tumor Dense" or "Tumor Sparse" categories using a cutoff in the global "Tumor Score" distribution, defined as the tangent point with the steepest slope to the right of the distribution peak. Using a search radius of $r < 100$ μm (~10 cell diameters), this approach was applied uniformly across all samples, ensuring consistent classification of local tumor environments regardless of EBV status. As this metric effectively captured local variations in HRS cell density and proximity **(Figure 4.12)**, it enabled robust comparisons of "Tumor Dense" and "Tumor Sparse" regions across EBV-positive and EBV-negative tumors.

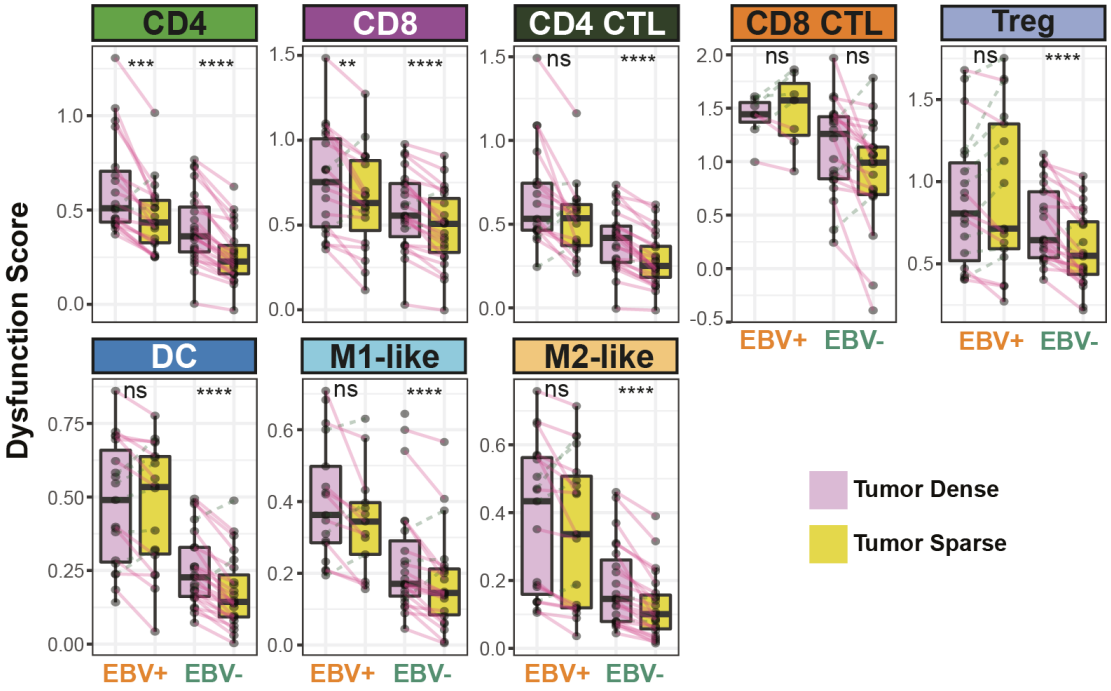


Figure 4.13: Comparison of dysfunction scores from the major immune cell compositions within tumor dense and tumor spare regions and stratified by EBV status. Asterisks denote statistical significance (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$). Adapted from Yeo et al., 2024[191].

Each immune cell was then assigned a "Dysfunction Score" by applying negative weights to markers of activation or proliferative potential (CD45RA, Ki-67) and positive weights to markers of T-cell activation and dysfunction (CD45RO, Tox, PD-1, Lag3, PD-L1). Quantification of "Dysfunction Scores" in "Tumor Dense" and "Tumor Sparse" regions revealed significantly higher dysfunctional in "Tumor Dense" regions across EBV-positive and EBV-negative tumors, and notably prevalent in both CD8 and CD4 T-cells. Moreover, these proximity-linked "Dysfunction Scores" were consistently higher in EBV-positive tumors compared to their EBV-negative counterparts **(Figure 4.13)**. These findings underscore a central role for HRS cells in modulating the functional states of surrounding immune constituents within the cHL tumor microenvironment, a mechanism exploited and markedly amplified by EBV.

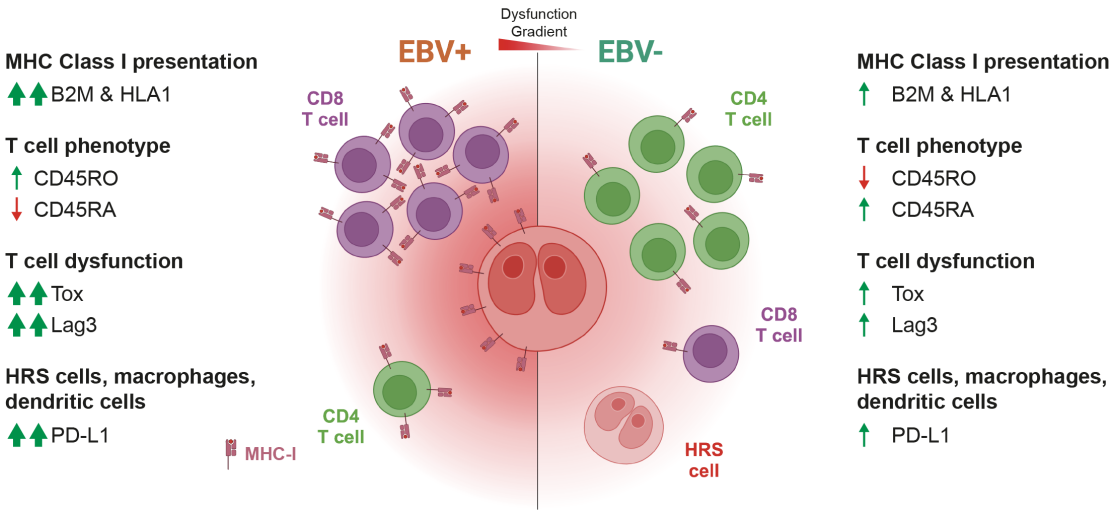


Figure 4.14: Cartoon depicting the key differences between EBV-positive and EBV-negative cHL tumor microenvironment. Figure from Yeo et al., 2024 [191].

Through a systematic dissection at increasing spatial resolutions, the cumulative spatial proteomics analyses support a model in which EBV reorganizes the cHL tumor microenvironment to amplify T-cell dysfunction in a manner distinct from EBV-negative cHL. EBV-positive cHL is characterized by a tissue-level increase in MHC Class I (HLA1, B2M) and predominance of dysfunctional memory CD8 T-cells, in contrast to the enrichment of CD4 T-cells and higher density of HRS cells observed in EBV-negative cHL. Beyond these global differences, HRS cells impose proximity-linked immunomodulation that augments T-cell dysfunction, a process exacerbated by EBV, in part through the elevated expression of dysfunction-associated markers that include Tox, Lag3, PD-1, and PD-L1 **(Figure 4.14)**. Taken together, these findings establish EBV-positive cHL as a model of how viruses can reprogram the tumor–immune interface, providing mechanistic insights supporting the broader principle that immune dysfunction is central to lymphomagenesis.

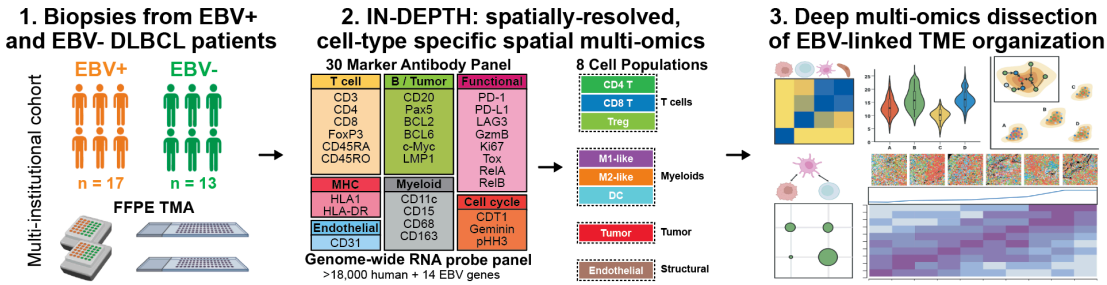


Figure 4.15: IN-DEPTH workflow to dissect the EBV-positive and EBV-negative DLBCL tumor microenvironment. Biopsies from a cohort of 17 EBV-positive and 13 EBV-negative DLBCL patients were collected (1), stained with a 30-marker antibody panel and hybridized with a genome-wide RNA probe panel spiked in with custom-designed probes targeting 14 EBV genes (2), where upon computational analysis illustrated key differences between EBV-positive and EBV-negative DLBCL. Figure from Yeo et al., 2024 [189].

4.2 EBV and T-cell dysfunction in the DLBCL tumor microenvironment

The profound elevation of T-cell dysfunction in the EBV-positive cHL tumor microenvironment highlighted the potential for EBV to similarly remodel the DLBCL tumor microenvironment. Beyond its association with poorer prognosis [121], EBV-positive DLBCL is highly unusual as it appears to have a distinct landscape of genetic aberrations, with many cases not aligning with established molecular subtypes [70]. These genetic distinctions are accompanied by a purported tolerogenic tumor microenvironment, including the upregulation of PD-L1 on malignant cells [33] and an increase in the ratio of M2/M1 macrophage gene signatures [104]. These features support a key role for EBV in remodeling the DLBCL tumor microenvironment to promote T-cell suppression.

For a comprehensive spatial mapping of the heterogenous DLBCL tumor microenvironment, the IN-DEPTH workflow (CODEX-GeoMx) was applied to archival FFPE biopsy samples from 17 EBV-positive and 13 EBV-negative patients across the Dana-Farber Cancer Institute and the University of Rochester Medical Center **(Figure 4.15)**. A separate panel of 30 antibody markers was first implemented and imaged on the CODEX, from which 8 distinct cell populations were identified and stringently verified as described previously. Note that all cells positive for CD20 and/or Pax5 were

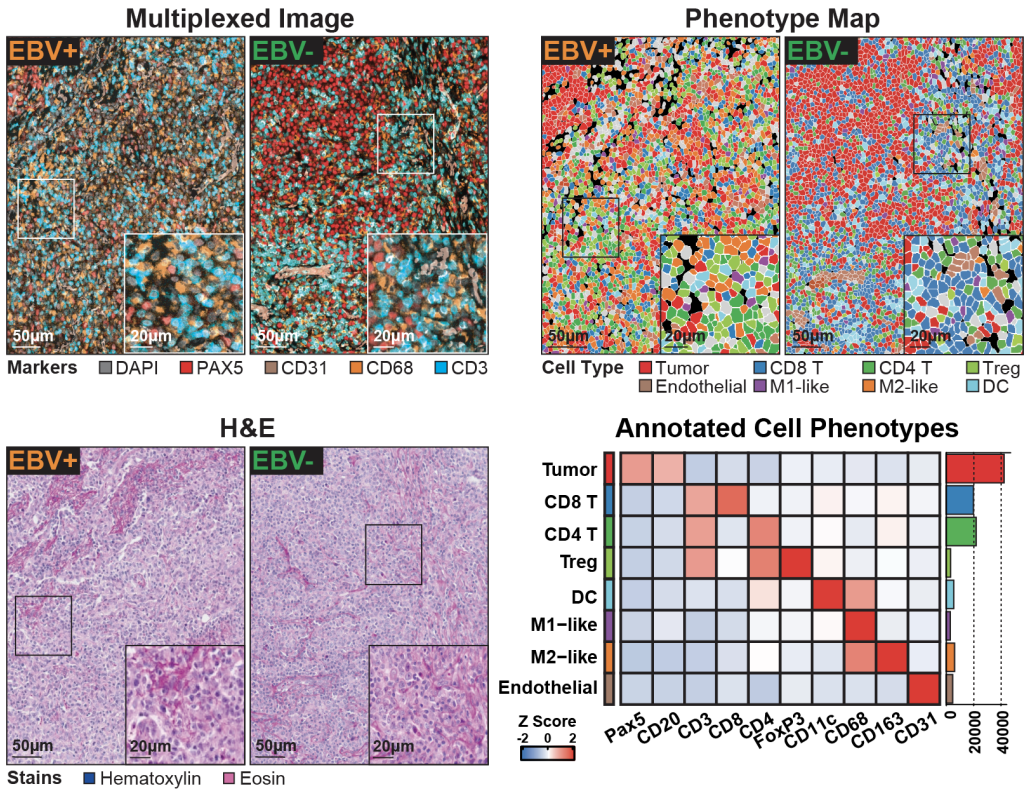


Figure 4.16: Representative CODEX multiplexed images (top left) with markers for nuclei (DAPI), B/tumor cells (Pax5), endothelial cells (CD31), macrophages (CD68), and T-cells (CD3), as well as the corresponding phenotype maps (top right) and H&E images (bottom left) of EBV-positive and EBV-negative DLBCL tissues. These are consistent with the heatmap showing antibody expression level and relative abundance of each annotated cell population (bottom right). While not shown here, a lymph node tissue was included during data acquisition as a positive control. Figure from Yeo et al., 2024 [189].

classified as malignant cells. The spatial coordinates of these 8 cell populations were then used to guide transcriptome capture on the GeoMx, using a genome-wide probe panel targeting >18,000 protein-coding genes supplemented with custom-designed probes for EBV transcripts. In total, 38 fields-of-view (one per patient) were captured. The accuracy of cell phenotyping and robustness of the IN-DEPTH workflow were demonstrated through the representative raw image, phenotype map, H&E staining, and the expected enrichment of antibody markers within each cell population, all of which were verified with board-certified pathologists **(Figure 4.16)**.

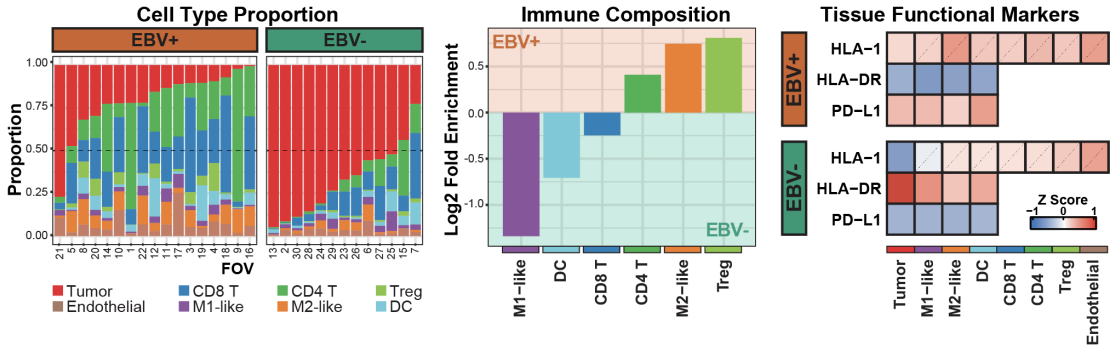


Figure 4.17: Contrasting immune profiles between EBV-positive and EBV-negative DLBCL. Relative proportions of annotated cell types across EBV-positive and EBV-negative tissues (tissues) reveal a considerable increase of immune cells in EBV-positive DLBCL. Further quantification reveal EBV-linked abundances of M1 *vs.* M2-like macrophages (middle) and tissue-level contrasts in MHC Class II and PD-L1 expression (right). Adapted from Yeo et al., 2024 [189].

Distinct cellular profiles were observed between EBV-positive and EBV-negative DLBCL tumor microenvironments. Notably, EBV-positive tumors were enriched in immune components, whereas EBV-negative tumors contained substantially higher proportions of malignant cells, supporting a major role for EBV in coordinating immune-mediated dysfunction in DLBCL. Within the immune compartment, EBV-positive tumors were characterized by an enrichment of immunosuppressive M2-like macrophages and regulatory T-cells, in contrast to more M1-like macrophages in EBV-negative tumors. Further tissue-level examination of antigen presentation and PD-L1 expression, key regulators of T-cell functional states, revealed no significant differences in MHC Class I (HLA-I) across cell populations. However, the EBV-positive DLBCL tumor microenvironment exhibited reduced MHC Class II and elevated PD-L1 expression across malignant cells, dendritic cells, and macrophages. These findings **(Figure 4.17)** point to a CD4 T-cell centered mechanism of dysfunction that may represent a distinct immunosuppressive program in EBV-positive DLBCL.

The functional state of T-cells were evaluated by quantifying T-cell dysfunction signatures [201,202,11] across protein (applying the "Dysfunction Score" metric defined previously) and RNA (incorporating *CTLA4*, *HAVCR2*, *LAG3*, *PDCD1*, *BTLA*, *TIGIT*, *CD160*, *CD244*, *ENTPD1*, and *VSIR* into a gene signature for Gene Set Variation Analysis (GSVA) [84] enrichment scoring) modalities. A consis-

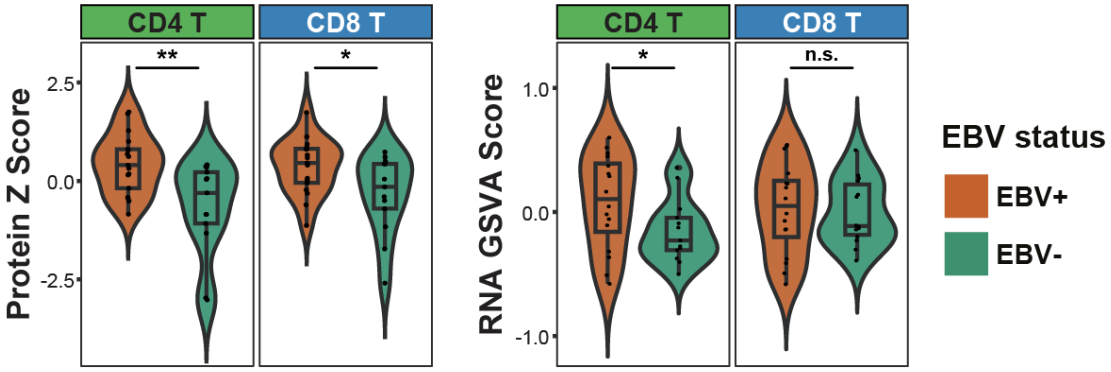


Figure 4.18: Comparison of CD4 and CD8 T-cell dysfunction scores calculated based on protein markers (left) and RNA signatures (right) between EBV-positive and EBV-negative DLBCL tissues. Asterisks denote statistical significance (* $p < 0.05$, ** $p < 0.01$), n.s. = not significant. Adapted from Yeo et al., 2024 [189].

tent increase in T-cell dysfunction was observed in EBV-positive DLBCL, with CD4 T-cells exhibiting significantly more pronounced effects than CD8 T-cells **(Figure 4.18)**. The orthogonal assessment at protein and RNA levels simultaneously also highlights the value of same-slide spatial multi-omics via IN-DEPTH for biological discovery and validation.

The paired increase in M2-like macrophages and CD4 T-cell dysfunction in the EBV-positive DLBCL tumor microenvironment suggests a role for macrophages in coordinating the increase in CD4 T-cell dysfunction. Tissue-level analysis of macrophage polarization, defined as the proportion of M2-like macrophages within each sample, revealed a clear bimodal distribution associated with EBV status, with EBV-positive tumors enriched in immunosuppressive M2-like macrophages while EBV-negative tumors containing more activatory M1-like macrophages. When tumors were stratified into M2-enriched and M1-enriched groups based on the intersection of their densities, the M2-enriched group exhibited significantly higher CD4 T-cell RNA dysfunction signature, with no corresponding differences in CD8 T-cells. These findings **(Figure 4.19)** support the notion that EBV-positive DLBCL leverages M2-like macrophages to enforce a CD4 T-cell focused axis of immune suppression. This approach also demonstrates how protein protein features can be used to guide the dissection of RNA signatures, another value of same-slide spatial multi-omics enabled by IN-DEPTH.

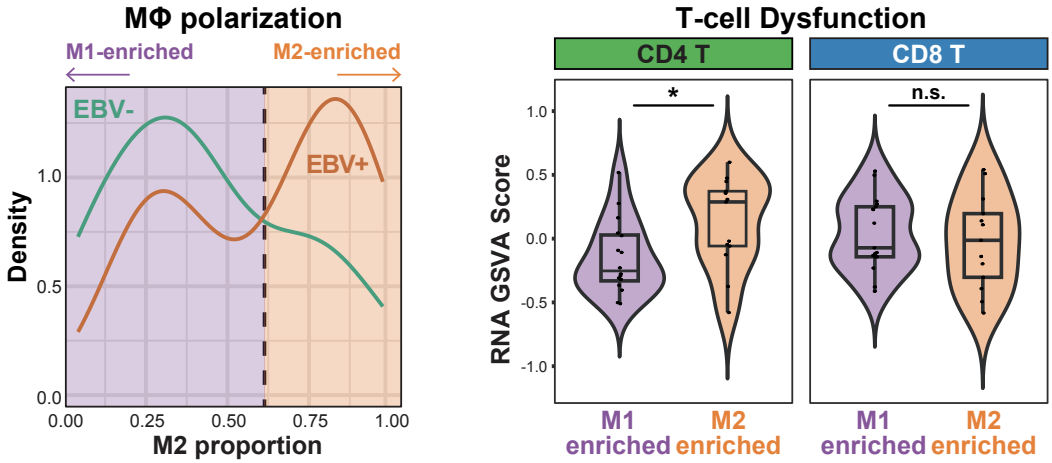


Figure 4.19: Relationship between macrophage polarization and T-cell dysfunction. Tissues were stratified into M1-enriched and M2-enriched populations based on the density of M2-like macrophages (left), from which the RNA signatures for CD4 and CD8 T-cell dysfunction were quantified (right). Asterisks denote statistical significance (* $p < 0.05$), n.s. = not significant. Adapted from Yeo et al., 2024 [189].

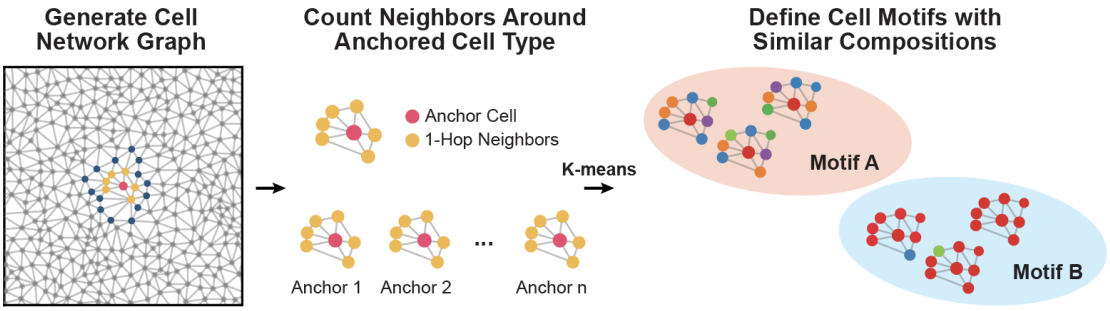


Figure 4.20: Schematic of the approach used to identify different cellular motifs through n-hop neighborhood analysis anchored on a cell type of interest. Here, CD4 T-cells were set as the anchor, from which the immediate 1-hop neighbors were identified and clustered. Figure from Yeo et al., 2024 [189].

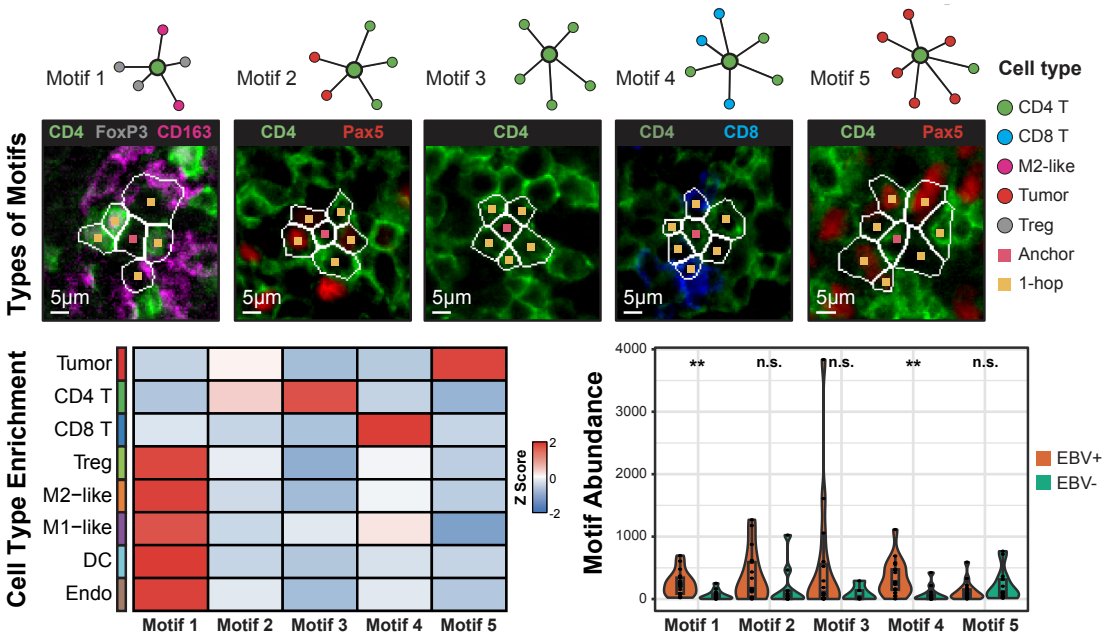


Figure 4.21: Five different CD4 T-cell anchored motifs were identified in DLBCL. A representative image of each motif is shown (top), which is consistent with the motif-stratified cell type enrichment heatmap (bottom left). EBV-linked differences in the abundance of motifs were only found in motifs 1 and 4 (bottom right). Asterisks denote statistical significance (** $p < 0.01$), n.s. = not significant. Adapted from Yeo et al., 2024 [189].

To assess how the spatial organization of the DLBCL tumor microenvironment contributes to CD4 T-cell dysfunction, the immediate neighborhoods of CD4 T-cells were analyzed using a network graph approach constructed through the following steps **(Figure 4.20)**:

1. *Setting the nodes:* Each node was represented using the spatial location of every cell in each tissue, which was represented by the centroid of its segmentation mask.

2. *Connecting the nodes:* Delaunay triangulation [54] was applied to the set of centroids, from which the simplices were used to construct the network graph, with the condition that two nodes were connected only if the Euclidean distance between them was ≤20 μm (~2 cell diameters).

Afterwards, for each node of interest (those corresponding to CD4 T-cells in this case), its directly adjacent nodes were identified and summarized into vectors representing the counts of each cell type,

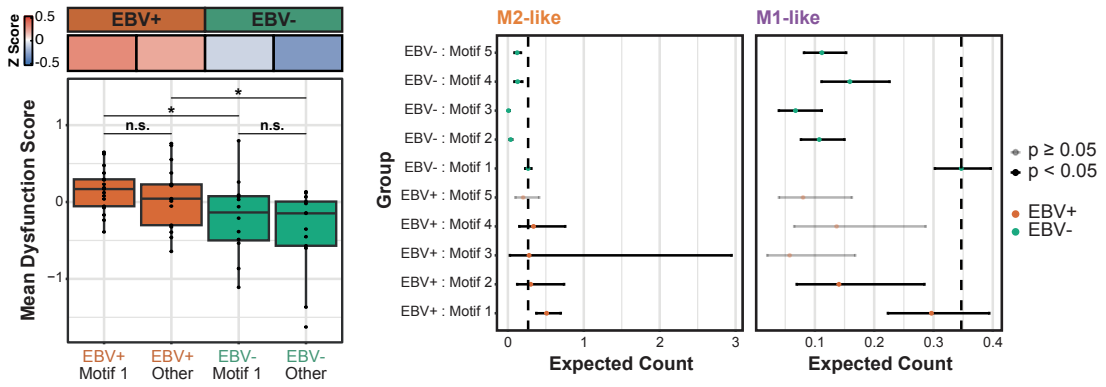


Figure 4.22: Motif-linked differences in T-cell functional state. Comparison of protein-level CD4 T-cell dysfunction between the immune-enriched motif 1 and all other motifs combined (left). Negative binomial regression results with the response variable being the count of M2-like or M1-like macrophages (right), with EBV- : Motif 1 set as the baseline measure. For each variable, the point estimate is represented by a solid dot, with its corresponding 95% confidence interval shown as the bars surrounding the dot. Asterisks denote statistical significance (* $p < 0.05$), n.s. = not significant. Adapted from Yeo et al., 2024 [189].

which were then clustered to define recurring spatial motifs around the cell type of interest. Here, five distinct CD4 T-cell anchored motifs were identified **(Figure 4.21)**. Among these, Motifs 1 and 4 were significantly more prevalent in EBV-positive tumors, while the other motifs showed no EBV-associated differences. Of particular interest was Motif-1 which was highly enriched in immune cells, including macrophages, regulatory T-cells, and dendritic cells, supporting the notion of immune-mediated T-cell dysfunction in EBV-positive DLBCL

Assessment of CD4 T-cell dysfunction signatures revealed a graded pattern across the motifs: highest in EBV-positive immune-enriched motifs (Motif-1) and decreasing progressively to the lowest in EBV-negative immune-deficient motifs (Motifs 2+3+4+5). Regression analysis further demonstrated an association between macrophages, motif composition, and EBV status, with EBV-positive samples having 1.91 times the expected count of M2-like macrophages compared to EBV-negative samples ($p < 0.05$, 95% CI [1.64, 2.25]) within a given motif, while the expected M1-like macrophage count was only 0.86 times ($p < 0.05$, 95% CI [0.74, 0.99]). These results **(Figure 4.22)** provide a spatial ba-

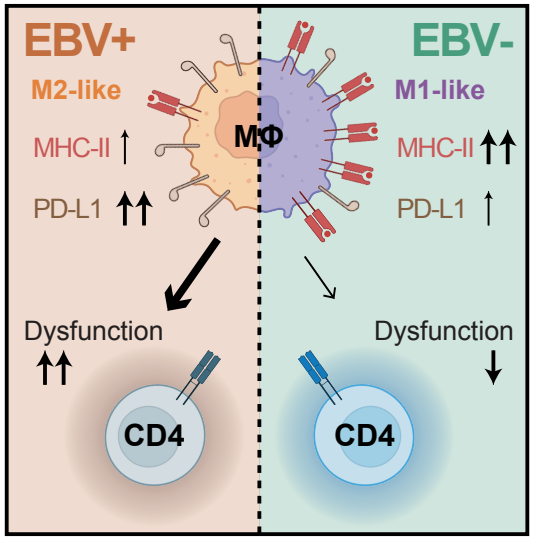


Figure 4.23: Cartoon depicting the key differences between EBV-positive and EBV-negative DLBCL tumor microenvironment. Figure from Yeo et al., 2024[191].

sis linking macrophages in immune-enriched niches to CD4 T-cell dysfunction, demonstrating a key role for EBV at driving spatial reprogramming of the DLBCL tumor microenvironment to promote CD4 T-cell dysfunction.

The cumulative analyses performed across protein and RNA modalities at increasing spatial resolutions, uniquely enabled by IN-DEPTH, support a model in which EBV reorganizes the DLBCL tumor microenvironment to promote CD4 T-cell dysfunction through a coordinated reduction of MHC Class II, elevation of PD-L1, and establishment of immune-rich niches dominated by immunosuppressive M2-polarized macrophages surrounding CD4 T-cells **(Figure 4.23)**. The collective findings across cHL[191] and DLBCL[189] further highlight EBV as a paradigm for how oncogenic viruses can reprogram the immune landscape of the lymphoma tumor microenvironment through distinct mechanisms that ultimately converge on immune dysfunction, a potential unifying principle of virus-driven tumorigenesis.

The hidden harmony is better than the obvious.

Heraclitus

5

Decoding the Mechanistic Logic Behind Lymphoma Remodeling

The essence of spatial organization extends beyond structural arrangement. It is a reflection of interlacing gene programs, signaling circuits, and communication networks that inscribe order onto tissues. In diseased states such as the tumor microenvironment, this organization becomes a lan-

guage through which cells communicate, adapt, and ultimately become misdirected. As demonstrated in the previous chapter, the oncogenic EBV weaves into this landscape by reprogramming malignant and immune phenotypes [191,189], key features that reshape the architecture of both cHL and DLBCL tumor microenvironments to promote T-cell dysfunction. However, deriving the underlying mechanisms that orchestrate such drastic tissue reorganization remains a rate-limiting step hindered by its inherent complexity. Through the experimental and computational innovations introduced in Chapter 3, these challenges can be addressed more efficiently, enabling deeper insights into the logic behind virus-mediated remodeling of the tumor microenvironment. Such approaches not only clarify how the tumor architecture was constructed, but also disentangle the hierarchical rules that govern spatial patterning under disease pressure.

5.1 CD8:MHC Class I engagement and T-cell terminal exhaustion in cHL

The substantial increase in T-cell dysfunction as a function of spatial proximity to HRS cells and EBV status in the cHL tumor microenvironment supports a principal role for HRS:T-cell interactions in driving immune dysfunction. To dissect the mechanisms underpinning this interaction, a customized GeoMx workflow was implemented to comprehensively assess the functional changes involved. A genome-wide probe panel targeting >18,000 protein-coding genes supplemented with custom-designed probes developed against EBV transcripts was used. The transcriptomes of six cell populations (Tumor-enriched, CD4 naïve, CD4 memory, CD8 naïve, CD8 memory, Others) were captured from cHL biopsies across 22 EBV-positive and 24 EBV-negative patients at the Dana-Farber Cancer Institute and the University of Rochester Medical Center. These six cell populations were defined through a combinatorial antibody staining strategy with CD3, CD8, and CD45RO, while referencing an adjacent slide with a CD30 immunohistochemical staining. Transcriptomics capture was also focused at regions densely populated with HRS cells **(Figure 5.1)**.

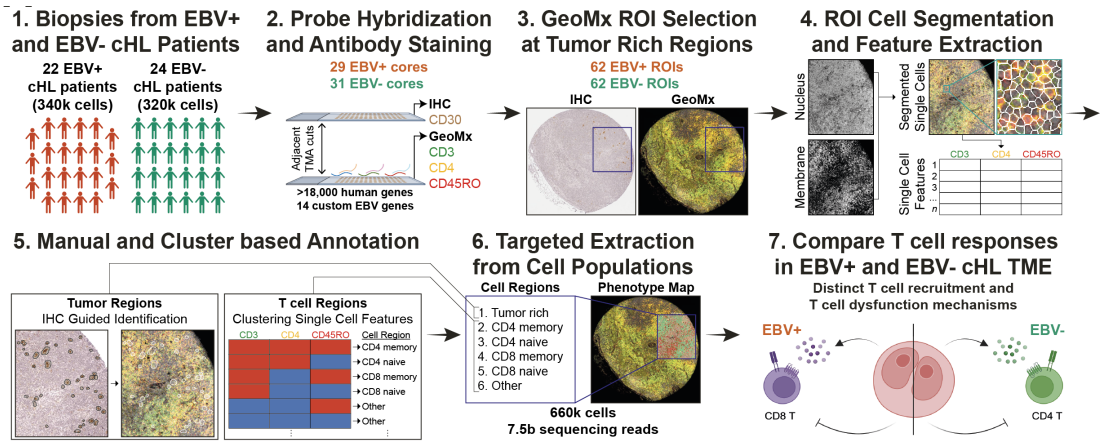


Figure 5.1: GeoMx spatial transcriptomics workflow to dissect mechanisms of T-cell dysfunction in the cHL tumor microenvironment. Biopsies from 22 EBV-positive and 24 EBV-negative cHL patients were assembled into tissue microarrays (1), from which one section was stained with anti-CD30 for immunohistochemical detection of HRS cells and the adjacent section for GeoMx profiling by hybridizing with an RNA probe panel containing >18,000 human protein-coding and 14 custom-designed EBV genes and staining with three fluorescent antibodies (2). This approach enabled the selection of 126 regions of interest at HRS-dense regions (3), extraction of single-cell marker expression (4), and annotation of six cell populations (5) for targeted transcriptome capture within (6) that enable the dissection of T-cell dysfunction programs (7). ROI, region-of-interest; TME, tumor microenvironment. Figure from Yeo et al., 2024 [191].

The accuracy of the GeoMx capture strategy was supported visually by the comparable staining pattern and annotated phenotype map as shown in the representative tissue images. The top differentially expressed genes across these cell populations were also consistent with their expected enrichment, including *TNFRSF8* (encoding CD30) and the EBV-encoded small RNAs *EBER1* and *EBER2* in the Tumor-enriched region, CD4 and CD8 T-cell specific transcripts (such as *CD3D*, *CD3E*, *CD4*, *CD7*, *CD8A*, *CD8B*) in their respective regions, cytotoxic granzyme and perforin transcripts associated with T-cell effector functions (including *PRF1*, *GZMK*, *GZMA*, *GZMH*) in memory T-cell regions, and immunoglobulin and B-cell transcripts (including *CD19*, *MS4A1* (encoding CD20), *CD79A*, *IGH*s, *IGL*s, *IGK*s) found in B-cells (under the Other region). EBV transcripts were also enriched in EBV-positive tumors, which included the expected EBV latency II transcripts *EBER1*, *EBER2*, and *LMP1*, but also other latent and lytic transcripts including *EBNA2*, *BALF1*, *BCRF1*, *BNLF2A*, *BNLF2B*, *BZLF1*, and *RPMS1*, consistent with bulk transcriptomics data on EBV-positive cHL samples re-

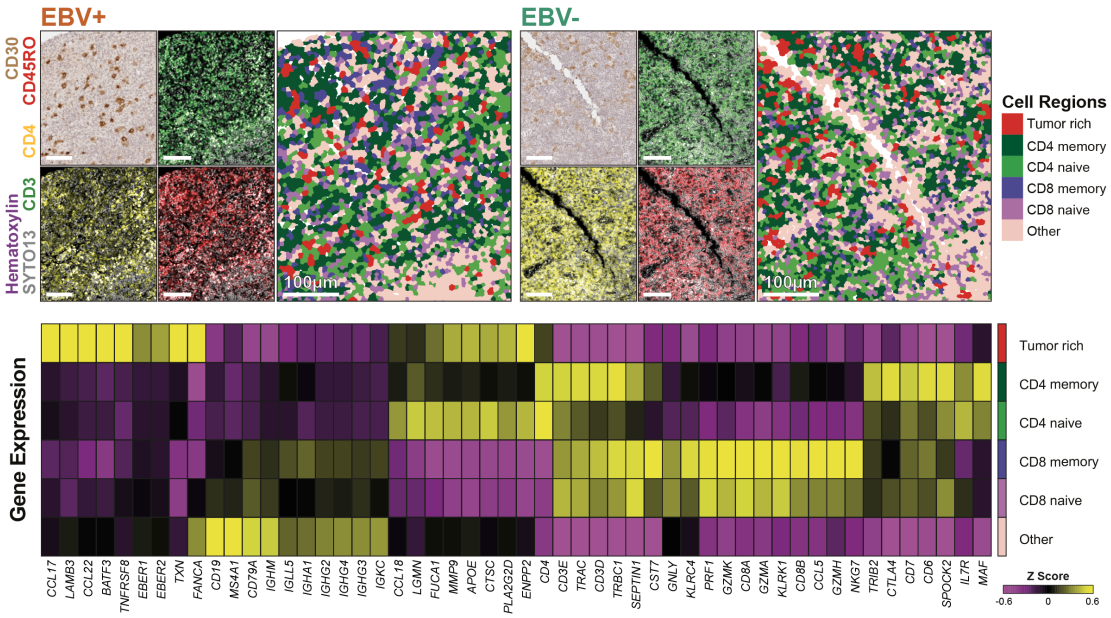


Figure 5.2: Validation of GeoMx transcriptome capture accuracy. The representative immunohistochemical staining and immunofluorescence images of EBV-positive and EBV-negative tissues were consistent with the final phenotype maps (top), contributing to the expected enrichment of cell-type specific transcripts in each cell population (bottom). Adapted from Yeo et al., 2024 [191].

ported in other studies [16]. These stringent validations **(Figure 5.2)** demonstrate that, despite relying on an adjacent CD30-stained section to manually define tumor-enriched regions due to the three-antibody limitation of the default GeoMx workflow, the resulting data possessed sufficient quality for downstream analyses.

Several features from the GeoMx spatial transcriptomics data corroborated findings from MIBI spatial proteomics described in the previous section. Differential gene expression between memory and naïve T-cells also revealed significant increase of *CD8A*, *CD8B*, and *LAG3* transcripts in memory subsets (*LAG3* is a major gene upregulated in both exhausted CD8 and CD4 T-cells [48]) uniquely in EBV-positive tumors, indicating a predominance of memory CD8 T-cells with elevated immune dysfunction. Consistent with this, EBV-positive tumors were similarly enriched in CD8 T-cells, in contrast to higher CD4 and HRS cell densities in EBV-negative tumors. Cytotoxic transcripts (in-

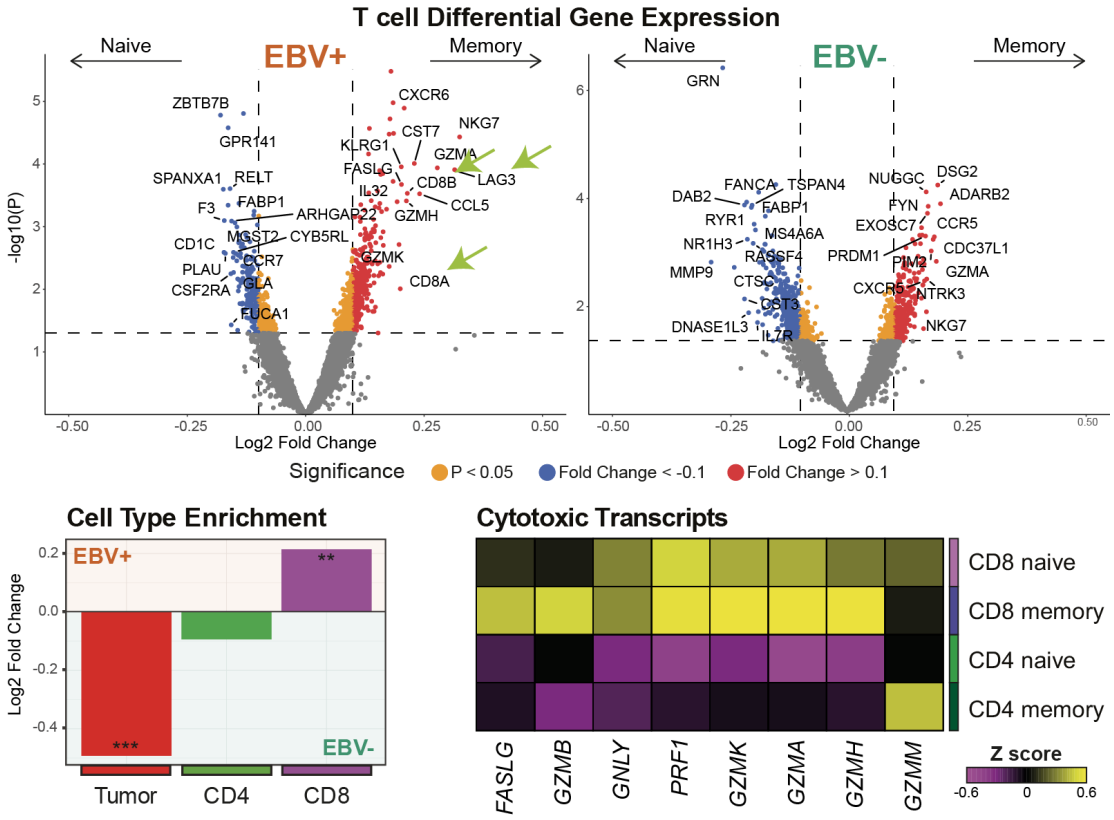


Figure 5.3: Distinctive features in GeoMx spatial transriptomics data were consistent with those in MIBI spatial proteomics. Volcano plots of differential gene expression in T-cells revealed increased expression of *CD8A* and *CD8B* (as well as *LAG3*) specifically in EBV-positive cHL as indicated in green arrows (top), which mirrored their relative cell abundances (bottom left). Cytotoxic genes were also higher in CD8 compared to CD4 T-cells and elevated in memory relative to naïve subsets within both lineages (bottom right). Asterisks denote statistical significance (** p < 0.01, *** p < 0.001). Adapted from Yeo et al., 2024 [191].

cluding *GZMA*, *GZMH*, and *GZMK*) were also found primarily in memory T-cell populations, in support of attempted effector functions within memory-differentiated T-cells in the cHL tumor microenvironment. These findings **(Figure 5.3)** reinforce the prevalence of memory CD8 T-cells in the EBV-positive cHL tumor microenvironment.

Further differential gene expression analysis of EBV-positive and EBV-negative Tumor-enriched regions revealed contrasting enrichment of *CXCL9* in EBV-positive tumors and *CCL17* in EBV-negative tumors, key chemokines that are known to regulate CD8 and CD4 T-cell homing into tumors [112].

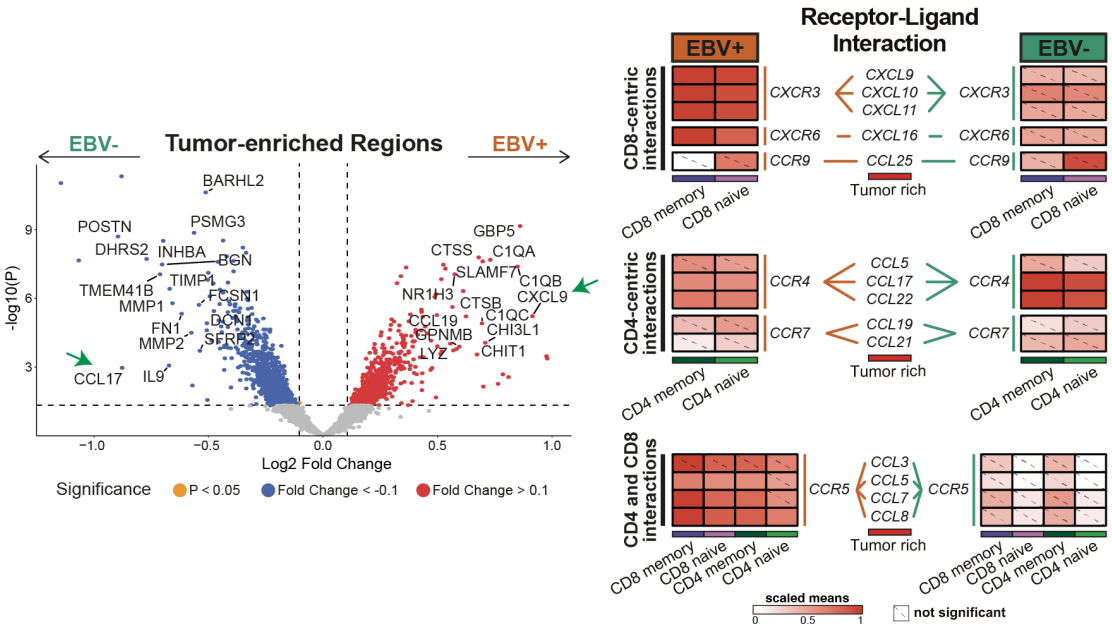


Figure 5.4: Contrasting T-cell homing dynamics between EBV-positive and EBV-negative cHL. Volcano plot of differential gene expression in EBV-positive *vs.* EBV-negative Tumor-enriched regions revealed contrasts in T-cell chemokine expression as indicated in green arrows (left). This was supported in a chemokine receptor-ligand analysis that indicate elevated CD8 T-cell recruitment in EBV-positive cHL in contrast to CD4 T-cell recruitment in EBV-negative cHL (right). Figure from Yeo et al., 2024[191].

To assess T-cell recruitment dynamics, receptor-ligand interaction analysis between Tumor-enriched and T-cell populations was performed using CellPhoneDB[63]. EBV-positive tumors displayed strong engagement of chemokine axes favoring CD8 T-cell recruitment, including *CXCR3–CXCL9/10/11* and *CXCR6–CXCL16*[112,153]. In contrast, CD4 T-cell favored engagements, such as the *CCR4–CCL17/22* axis[112], were present in both EBV-positive and EBV-negative microenvironments but at elevated levels in EBV-negative tumors. Additional engagement of the *CCR5–CCL5/7/8* T-cell homing axis[75] was also observed specifically in EBV-positive tumors. These findings **(Figure 5.4)** indicate that EBV shapes distinct chemokine signaling networks to skew lymphocyte recruitment towards CD8 T-cells in EBV-positive cHL.

The elevation of CD8 T-cell infiltration and MHC Class I expression indicate an active antiviral

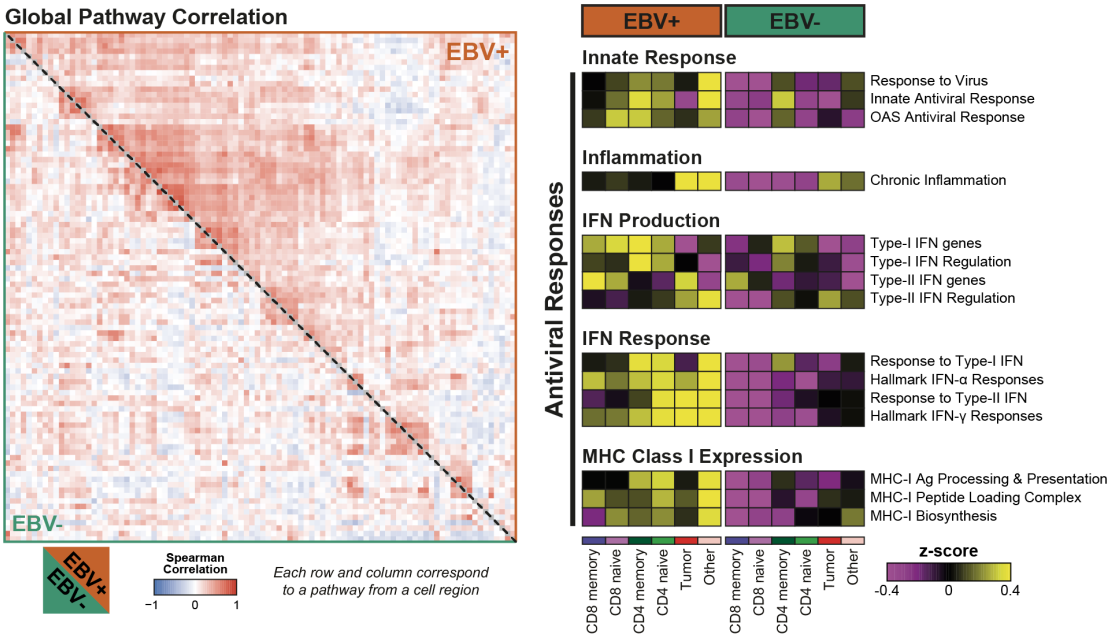
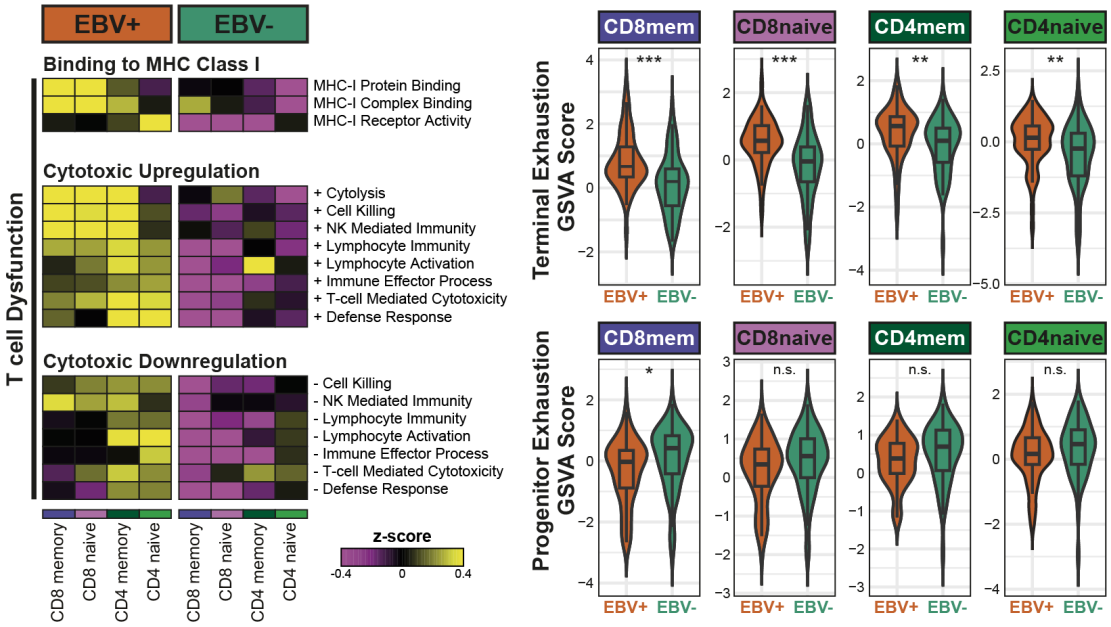


Figure 5.5: EBV-linked tissue-level contrasts in gene pathway activity. In the correlation heatmap (left), a considerably stronger relationship between inflammatory and T-cell functional signatures can be seen in EBV-positive cHL. Among these pathways include those associated with antiviral responses, which were consistently higher in EBV-negative tissues across all cell populations (right). Adapted from Yeo et al., 2024 [191].

and inflammatory response in the EBV-positive cHL tumor microenvironment. This was supported by a global correlation of Gene Ontology signatures [10] that broadly represent antiviral responses and T-cell dysfunction in EBV-positive tumors. For example, EBV-positive tumors had elevated immune signatures consistent with an antiviral state, including innate cellular immune responses, chronic inflammation, interferon signaling, all of which can enhance MHC Class I expression [83]. These findings **(Figure 5.5)** support a unique EBV-linked immunological landscape predisposed to T-cell dysfunction [48,202,11].

As T-cell dysfunction is a broad functional state that includes the progenitor to terminal exhaustion continuum [134], a careful dissection of these transcriptional programs in the cHL tumor microenvironment was imperative. Analysis of T-cell functional signatures showed that in EBV-positive tumors, they had increased expression of genes that promote MHC Class I binding and activate cytotoxic path-



Multi-variable regression of clinical factors to progenitor T cell exhaustion scores				
Term	Estimate	Standard Error	t value	p-value
(Intercept)	0.6286239	0.0254121	24.737	<2e-16****
Subtype:Nodular-Sclerosis	-0.0317717	0.016842	-1.886	0.061733
Subtype:Not otherwise specified	-0.0477028	0.0462891	-1.031	0.3049
Sex:M	0.0235794	0.0184444	1.278	0.203657
Sex:unknown	-0.0834367	0.0428976	-1.945	0.054192
Age	0.0007669	0.000455	1.685	0.094593
EBV	-0.0702777	0.0188236	-3.733	0.000294***
Source:Mass	-0.0186571	0.0224217	-0.832	0.407062

Multi-variable regression of clinical factors to terminal T cell exhaustion scores				
Term	Estimate	Standard Error	t value	p-value
(Intercept)	0.7627	0.02679	28.473	<2e-16****
Subtype:Nodular-Sclerosis	-0.007516	0.01775	-0.423	0.673
Subtype:Not otherwise specified	-0.04417	0.04879	-0.905	0.367
Sex:M	-0.01828	0.01944	-0.94	0.349
Sex:unknown	0.0306	0.04522	0.677	0.5
Age	8.39E-05	0.0004796	0.175	0.861
EBV	0.08129	0.01984	4.097	0.0000779****
Source:Mass	-0.03836	0.02363	-1.623	0.107

Figure 5.6: T-cells acquire distinctive dysfunction states in EBV-positive and EBV-negative cHL. Those in EBV-positive cHL exhibit elevated cytotoxic upregulation and downregulation pathways (top left) and had significantly higher terminally-exhausted and lower progenitor-exhausted signatures (top right) than those in EBV-negative cHL. These were supported by the multivariate regression analysis results (bottom); note that "Subtype: mixed cellularity", "Sex: F", and "Source: Lymph Node" were part of the metadata but not shown here as they were used as baseline measurements for their respective comparisons. Asterisks denote statistical significance (* $p < 0.01$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$), n.s. = not significant. Adapted from Yeo et al., 2024 [191].

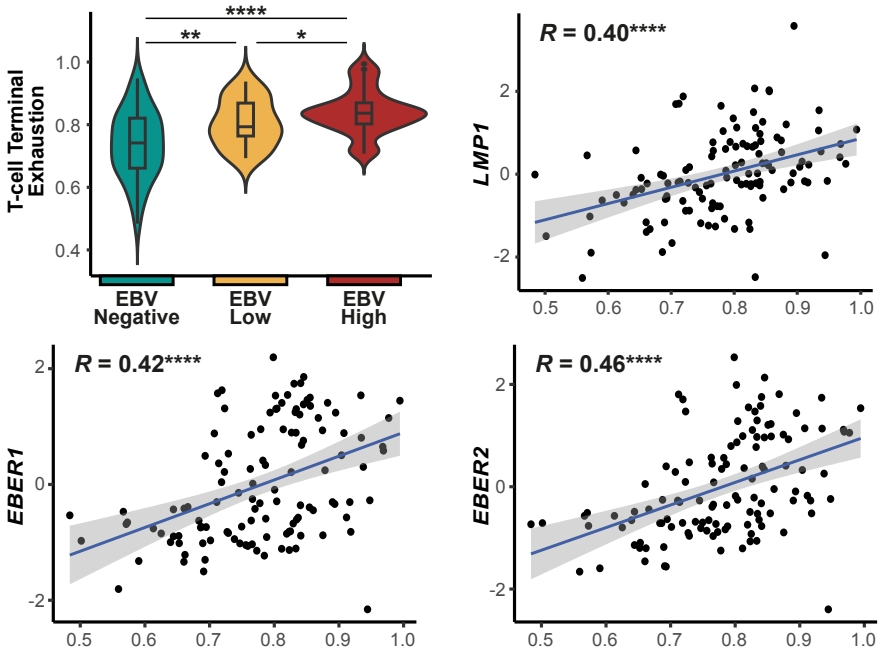


Figure 5.7: EBV oncogenes is linked to an increase in T-cell dysfunction. Higher overall expression of EBV transcripts genes is linked to an increase in T-cell terminal exhaustion signatures (top right), of which *LMP1*, *EBER1*, and *EBER2* showed the strongest positive correlations. Asterisks denote statistical significance (* p < 0.01, ** p < 0.01, **** p < 0.0001). Adapted from Yeo et al., 2024 [191].

ways, and also exhibited a concomitant rise in inhibitory pathways capable of suppressing cytotoxic responses. This was in sharp contrast with EBV-negative tumors, where both antigen recognition and cytotoxicity were relatively diminished. To resolve the unusual nature of these gene signatures, T-cell transcriptomes were compared against gene signatures that distinguish progenitor from terminally exhausted CD8 T-cells [134]. Note that while these states remain poorly characterized in CD4 T-cells, the same signatures were applied as a proxy. Across both CD8 and CD4 lineages, T-cells in EBV-positive tumors were significantly enriched in terminal exhaustion signatures, whereas those in EBV-negative tumors had a stronger progenitor profile. Regression analysis further confirmed EBV status as the primary clinical feature associated with terminal exhaustion **(Figure 5.6)**.

The strong association between T-cell dysfunction and EBV status prompted the identification of

responsible viral oncogenes. To verify the contribution of EBV to T-cell dysfunction, EBV-positive tumors were stratified into "EBV low" and "EBV high" categories based on the cumulative expression of EBV transcripts, which revealed a graded increase in T-cell terminal exhaustion signatures. Each EBV transcript in the Tumor-enriched region was next correlated to T-cell terminal exhaustion signatures, of which *LMP1*, one of the major EBV oncogenes[198], as well as *EBER1* and *EBER2*, the main diagnostic features of EBV status in lymphomas, showed the most robust positive correlations **(Figure 5.7)**. With LMP1 identified as a candidate viral oncogene promoting T-cell dysfunction in the EBV-positive cHL tumor microenvironment, the next goal was to evaluate the spatial relationship between EBV status, LMP1, and T-cell dysfunction. This required transcriptomics profiling at single-cell resolution, which was not possible with the GeoMx.

As a result, a CosMx single-cell spatial transcriptomics experiment encompassing 961 human and 6 EBV genes (which included *LMP1*) was next applied on cHL biopsies across 22 EBV-positive and 21 EBV-negative cHL patient samples from the Dana-Farber Cancer Institute and the University of Tübingen. Here, a total of 12 distinct cell populations were identified, which showed similar contrasts in CD8 and CD4 T-cell enrichment between EBV-positive and EBV-negative tumors, features consistently observed across the MIBI and GeoMx experiments. T-cell exhaustion (using *CTLA4*, *EOMES*, *HAVCR2*, *LAG3*, *PDCD1*, *TIGIT*, *TOX*, and *PRF1* as a gene signature for GSVA enrichment scoring) was also significantly higher in regions densely populated in HRS cells than in sparse regions only in EBV-positive tumors **(Figure 5.8)**. These results support the robustness of the current findings and indicate a pivotal role of EBV-positive HRS cells in modulating T-cell dysfunction in a proximity-dependent manner.

The contribution of *LMP1* in HRS cells was evident due to its association with elevated CD4 and CD8 T-cell exhaustion signatures, motivating the assessment of whether T-cell dysfunction was spatially organized around LMP1-expressing tumors. A correlation was performed between T-cell to HRS cell distance and T-cell exhaustion signature in EBV-positive tumors while similarly stratify-

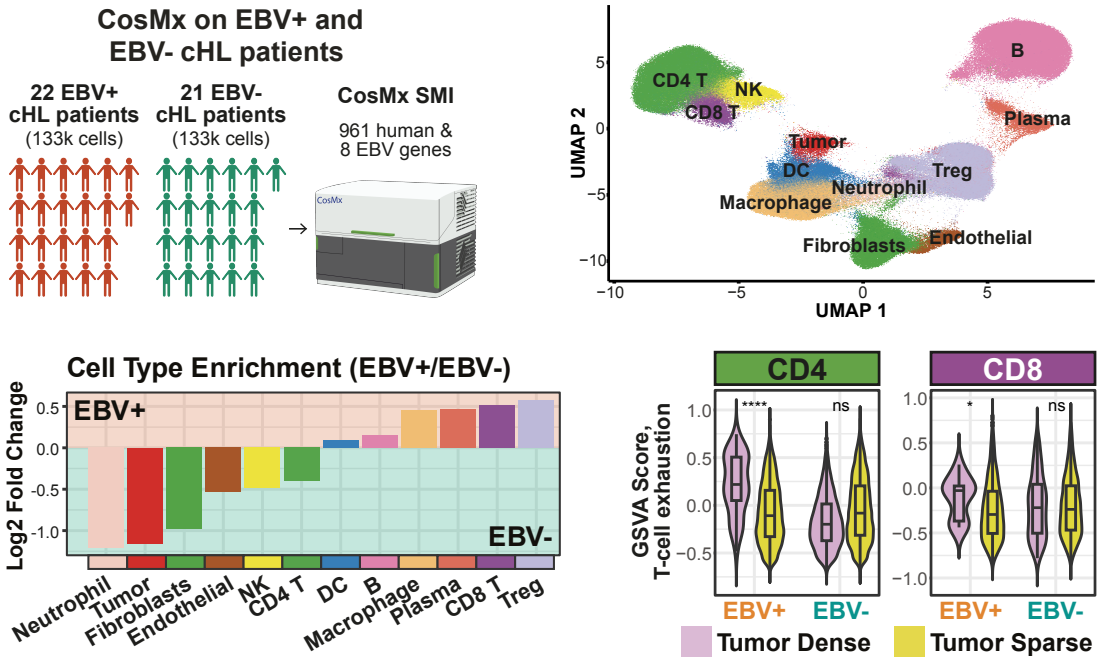


Figure 5.8: CosMx single-cell spatial transcriptomics workflow to evaluate the relationship between LMP1 and T-cell dysfunction. Biopsies from 22 EBV-positive and 21 EBV-negative cHL patients were profiled using an RNA probe panel containing 961 human and 8 custom-designed EBV genes (top left), from which 12 cell types were identified (top right). The relative cell abundances especially for CD4/CD8 T-cells (bottom left) and EBV-linked elevation of T-cell exhaustion signatures (bottom right) were consistent with findings from the MIBI and GeoMx data. Asterisks denote statistical significance (* $p < 0.01$, *** $p < 0.001$). Figure from Yeo et al., 2024 [191].

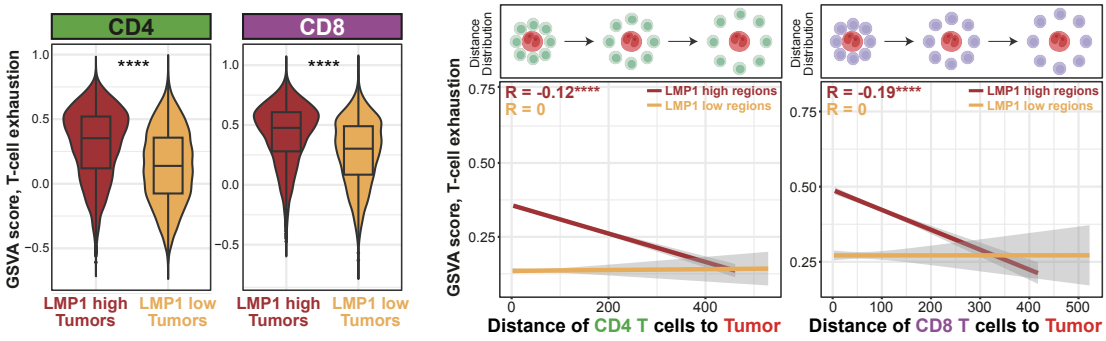


Figure 5.9: LMP1 promotes T-cell exhaustion in EBV-positive cHL in a spatially organized manner. Comparison of T-cell exhaustion signatures (left) and Spearman correlation of these signatures with the T-cell's distance from HRS cells (right) in EBV-positive tissues stratified by the expression level of LMP1 in HRS cells. Asterisks denote statistical significance (**** $p < 0.0001$). Figure from Yeo et al., 2024 [191].

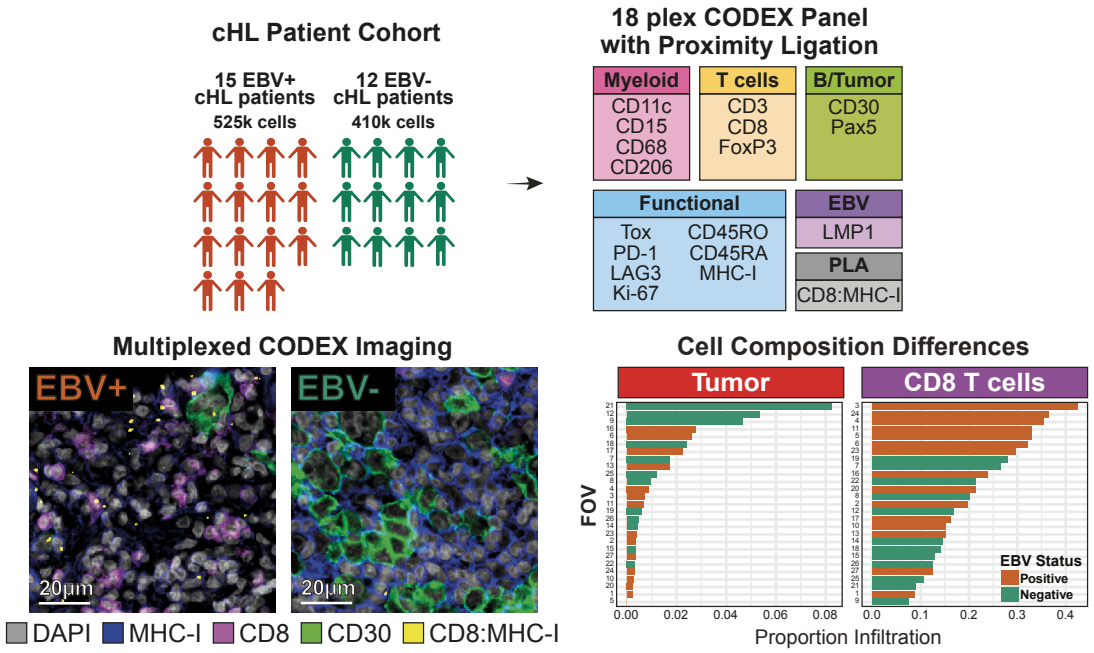


Figure 5.10: CODEX-PLA spatial proteomics workflow to assess the prevalence and functional consequence of CD8:HRS interactions. FFPE tissue microarrays containing biopsies from 15 EBV-positive and 12 EBV-negative cHL patients were subject to proximity ligation assay prior to staining with an 18-marker antibody panel (top). This enabled the detection of key cell phenotypes and proximity labeling puncta (bottom left), with the enrichment patterns of HRS and CD8 T-cells consistent across the previous datasets (bottom right). Adapted from Yeo et al., 2024 [191].

ing them into *LMP1*-high or *LMP1*-low regions. CD4 and CD8 T-cells were found to consistently show highest exhaustion signatures that negatively correlated with increasing distance from HRS cells only in *LMP1*-high regions **(Figure 5.9)**. This suggests that EBV-positive HRS cells skew T-cell dysfunction from a progenitor-like to a terminal-like exhausted state, a spatially-linked process tied to the expression of the LMP1 viral oncogene.

The predominance of memory-differentiated and highly dysfunctional CD8 T-cells spatially organized around EBV-positive HRS cells that retain MHC Class I expression led to the hypothesis that differences in CD8 T-cell engagement with HRS cells is a potential key mechanism that promotes T-cell dysfunction. To test this hypothesis, PLA (targeting CD8 and MHC Class I) was incorporated into a new CODEX panel containing key T-cell functional markers **(Figure 5.10)**. This CODEX-

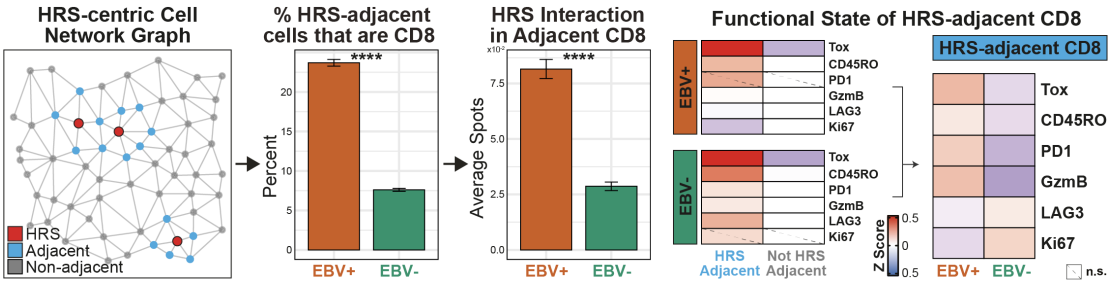


Figure 5.11: Increased CD8:HRS engagement in EBV-positive cHL augments CD8 T-cell dysfunction. A cell network graph (left) revealed significantly more CD8 T-cells adjacent to (middle left) and interacting with (middle right) HRS cells in EBV-positive cHL, which were linked to elevated T-cell dysfunction (right). Barplots show mean mean ± 1 standard error. Asterisks denote statistical significance (**** p < 0.0001). Figure from Yeo et al., 2024[191].

PLA panel was applied to another cohort of 15 EBV-positive and 12 EBV-negative patient biopsies from the University of Rochester Medical Center and the University of Tübingen to distinguish CD8:HRS interaction pairs *in situ*, thus allowing the contextualization of such interactions with T-cell dysfunction and provide a mechanistic basis underlying T-cell functional differences between the EBV-positive and EBV-negative cHL tumor microenvironment. Cells were similarly annotated and stringently validated as described previously. Once again, low HRS cell and high CD8 T-cell densities were observed in EBV-positive tumors, key features consistent across all datasets.

To determine if CD8:HRS engagement elevated T-cell dysfunction in the EBV-positive cHL tumor mircroenvironment, a cell network graph was built to identify the cells immediately adjacent to HRS cells. In EBV-positive tumors, a significantly higher proportion of CD8 T-cells were adjacent to HRS cells, and significantly more interaction events were detected between these CD8 T-cells with adjacent HRS cells. Further dissection of T-cell functional markers revealed that while HRS-adjacent CD8 T-cells were consistently more dysfunctional in both EBV-positive and EBV-negative tumors, this was markedly greater in EBV-positive tumors **(Figure 5.11)**. This demonstrated the ability for EBV to tune the CD8:HRS interaction axis and exacerbate T-cell dysfunction, in alignment of a more terminal-like exhaustion state observed across the spatial transcriptomics datasets. In support of this proposed mechanism, CD8 T-cells were found to require receptor engagement with MHC Class I to

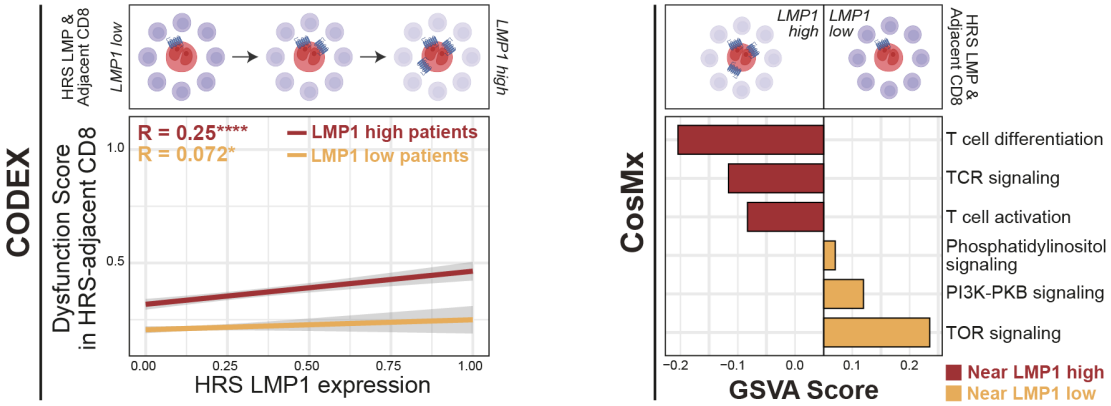


Figure 5.12: Abundance of LMP1-expressing HRS cells is positively correlated with CD8 T-cell dysfunction. Only in EBV-positive cHL patients with abundant LMP1-expressing HRS cells did CD8 T-cells show a positive correlation in dysfunction signatures in the CODEX spatial proteomics data (left), which was consistent with the pathway enrichments indicating stronger exhaustion when close to LMP1-high HRS cells in the CosMx spatial transcriptomics data (right). Asterisks denote statistical significance (* $p < 0.05$, **** $p < 0.0001$). Adapted from Yeo et al., 2024[191].

induce signaling cascades that drive them towards a terminally exhausted state in glioblastoma, resulting in poorer responses to immune checkpoint blockade[155].

As LMP1 appeared as a key viral factor promoting T-cell dysfunction in the EBV-positive cHL tumor microenvironment, the relationship between LMP1 expression on HRS cells and the dysfunction state of adjacent CD8 T-cells was further evaluated. Each HRS-adjacent CD8 T-cell was assigned a "Dysfunction Score" as described previously **(Figure 4.12)** and correlated with the expression of LMP1 on its corresponding HRS cell. In tumors with abundant LMP1-expressing HRS cells, a moderate positive correlation was observed, whereas no correlation was found in tumors with few LMP1-expressing HRS cells. These were orthogonally supported from the CosMx dataset, which revealed higher T-cell differentiation, receptor signaling, and activation pathways, along with reduced PI3K and TOR signaling — features consistent with a more terminally-exhausted phenotype[134,68] — in CD8 T-cells near LMP1-high HRS cells compared to those near LMP1-low HRS cells. These findings **(Figure 5.12)** highlight LMP1 as a key viral factor involved in functionally rewiring CD8 T-cells towards a more dysfunctional state in EBV-positive cHL.

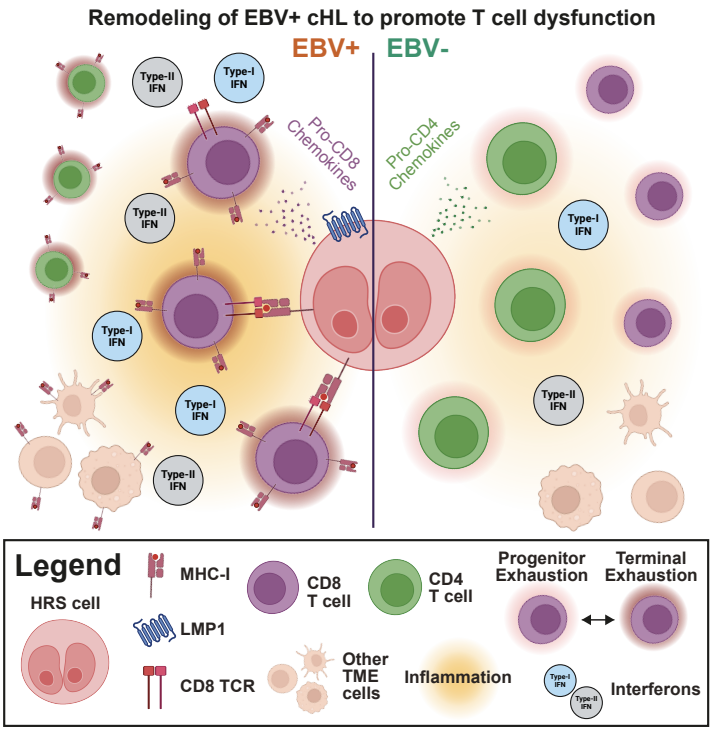


Figure 5.13: Cartoon depicting the proposed mechanism of T-cell dysfunction promoted by EBV in the cHL tumor microenvironment. Figure from Yeo et al., 2024[191].

The systematic multi-omics dissection presented here support a model in which EBV induces HRS cells, in part through the LMP1 viral oncogene, to condition their immediate microenvironment. This includes an altered expression of chemokines and immunomodulatory ligands to foster a CD8-enriched yet highly dysfunctional T-cell landscape. Chronic viral persistence further provokes inflammatory responses, including elevated interferon pathways and CD8:HRS engagement through MHC Class I, leading to an activated T-cell effector state that is limited in tumoricidal programs due to a terminally-exhausted functional state **(Figure 5.13)**. These findings explain the distinct spatial architecture of the EBV-positive cHL tumor microenvironment by illustrating the need for EBV to exploit spatial molecular interactions in order to reprogram both malignant and immune compartments and enforce a dysfunctional equilibrium necessary to support lymphomagenesis.

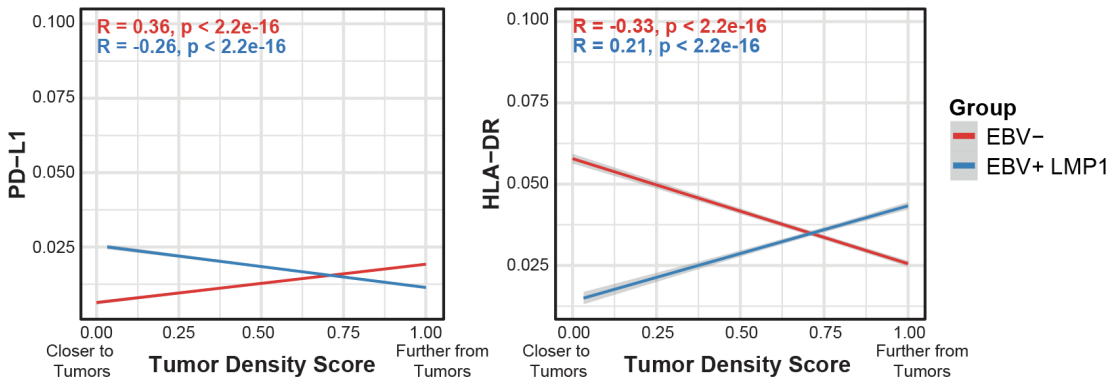


Figure 5.14: Contrasting macrophage PD-L1 and HLA-DR expression dynamics with increasing distance from tumor cells based on EBV status. The "Tumor density score" was the distance-based tumor score metric implemented previously for cHL. Adapted from Yeo et al., 2024 [189].

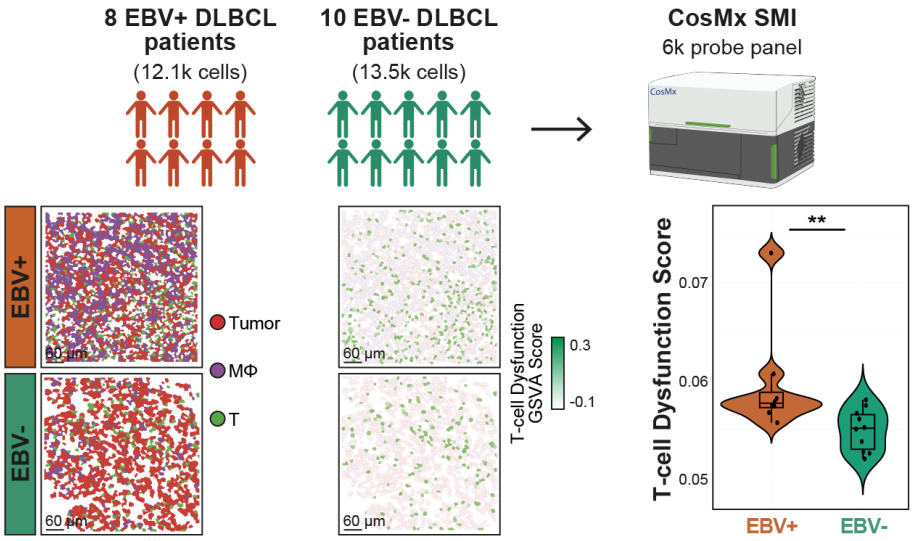


Figure 5.15: CosMx single-cell spatial transcriptomics workflow to verify tumor:macrophage:T-cell associations in EBV-positive DLBCL. Biopsies from 8 EBV-positive and 10 EBV-negative DLBCL patients were profiled using an RNA probe panel containing 6k human genes (top), from which malignant cells, macrophages, and T-cells were identified. EBV-positive tumors were relatively distributed in these cell populations as shown in the representative images (bottom left), with T-cells exhibiting a dynamic range of dysfunction (bottom right) that were significantly higher than EBV-negative tumors (bottom right). Asterisks denote statistical significance (** $p < 0.01$). Adapted from Yeo et al., 2024 [189].

5.2 Dysregulated macrophage:tumor:CD4 interaction axis in DLBCL

The increase of M2-like macrophages and elevated CD4 T-cell dysfunction within the abundant immune milieu associated with EBV status supports a key role of macrophage:T-cell interactions in coordinating immune dysfunction in the EBV-positive DLBCL tumor microenvironment. Further analysis showed that malignant cells appeared to influence macrophage functional states, as macrophages in EBV-negative tumors had lower PD-L1 and higher HLA-DR expression with increasing density of malignant cells, with both trends reversed in EBV-positive tumors with high expression of LMP1 **(Figure 5.14)**. To assess the prevalence of tumor:macrophage:T-cell association, a CosMx spatial transcriptomics experiment was applied on 8 EBV-positive and 10 EBV-negative DLBCL patient samples **(Figure 5.15)**. A homogeneous spatial organization between malignant cells, macrophages, and T-cells was observed in EBV-positive tumors and accompanied with elevated T-cell dysfunction signatures, in contrast to EBV-negative tumors that were dominated by malignant cells and had lower T-cell dysfunction signatures.

Given the need to contextualize CD4 T-cell dysfunction in relation to macrophages and tumor cells, the application of SGCC would provide an effective approach to uncover the gene-regulatory networks that underlie these interactions. First examining tumor-macrophage interactions **(Figure 5.16, top)**, the presence of EBV was confirmed through the expression of LMP1 viral oncoprotein and detection of viral transcripts in EBV-positive tumors. SGCC analysis revealed divergent immunomodulatory pathways in tumor cells and macrophages. EBV-negative tumor cells exhibited signatures that potentiate immune-activatory responses, as compared EBV-positive tumor cells that promote macrophage differentiation and chemotaxis. This was supported by the macrophage phenotype distribution and transcriptonal programs, with activatory signatures in EBV-positive and immunosuppressive signatures in EBV-negative DLBCL that align with the M1/M2-like contrasts. Next examining macrophage-CD4 interactions **(Figure 5.16, bottom)**, SGCC analysis revealed contrasting func-

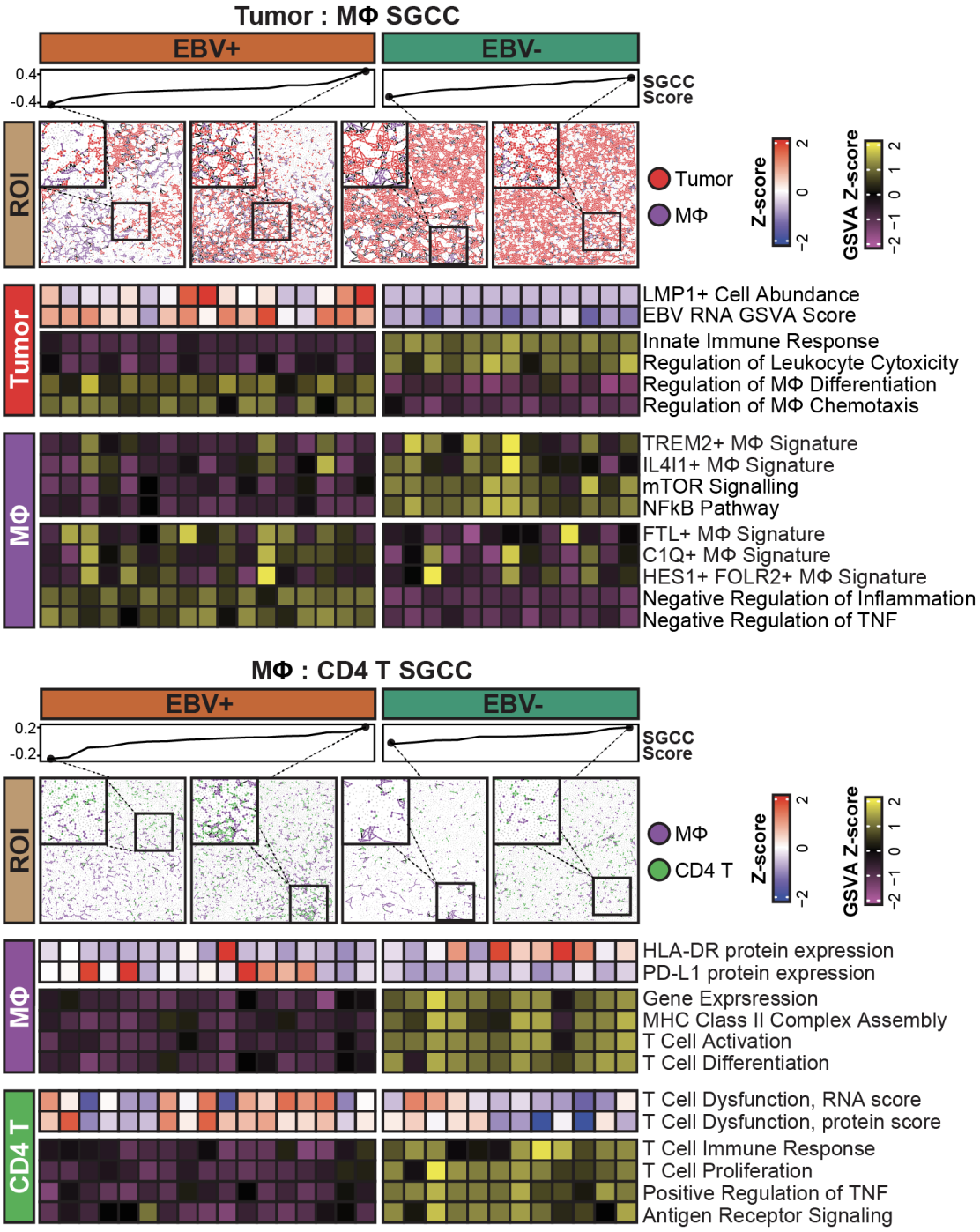


Figure 5.16: SGCC reveals coordinated spatial multimodal interactions and identifies EBV-linked differences in cell functional states along the tumor–macrophage–CD4 T-cell axis. Adapted from Yeo et al., 2024 [189].

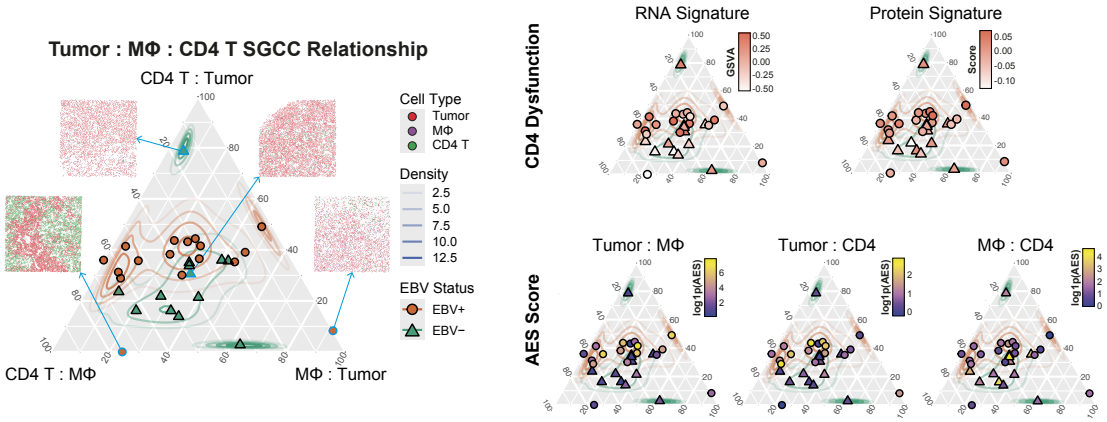


Figure 5.17: Ternary plots to visualize a three way relationship. Points located near the vertices indicate co-localization between the specific cell pair while forming a complementary structure with the third cell type, while those near the center signify co-localization across all three cell types (left). This approach contextualizes functional properties, such as CD4 T-cell dysfunction (top right) and adjacency enrichment statistic (bottom right), with their spatial relationships. AES, adjacency enrichment statistic. Adapted from Yeo et al., 2024[189].

tional states based on EBV infection status. Increasing SGCC was associated with HLA-DR protein expression and MHC Class II gene programs along with T-cell activatory signatures in EBV-negative DLBCL, which were all dampened in EBV-positive DLBCL that was associated with PD-L1 protein expression. This was consistent with elevated T-cell dysfunction signatures and dampened T-cell activatory pathways in EBV-positive DLBCL, with both trends reversed in EBV-negative DLBCL. These findings support a key spatial and immunomodulatory role for these cell populations in shaping the contrasting EBV-positive and EBV-negative tumor microenvironments.

Due to the complexity of tripartite spatial interactions, a ternary plot of SGCC scores between tumor:macrophage, macrophage:CD4, CD4:tumor cells was used **(Figure 5.17)** to better capture this complexity. T-cell dysfunction was found to peak upon maximal co-localization across all three cell types, supporting the relevance of this interaction axis in promoting CD4 T-cell dysfunction. To assess if there were preferential interactions between specific cell pairs, an adjacency enrichment statistic (AES) was applied[13], which measures the extent of which a pair of cell types was spatially adjacent

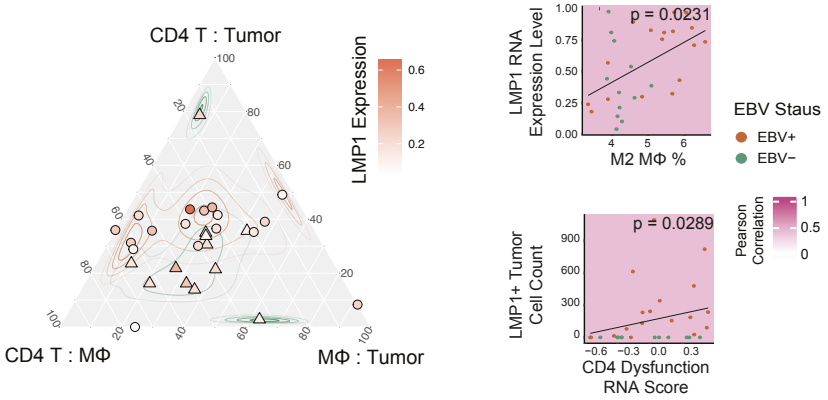


Figure 5.18: LMP1 is linked to immune dysfunction in EBV-positive DLBCL. Ternary plot of LMP1 expression (left) and Pearson correlation of LMP1 with macrophage proportions and T-cell signatures (right). Adapted from Yeo et al., 2024 [189].

beyond what would be expected from their abundances alone, with larger values indicating increased enrichment. AES analysis revealed preferential tumor-macrophage and tumor–CD4 interactions in EBV-positive DLBCL, supporting a model in which tumor-macrophage-CD4 crosstalk and immunosuppression predominates in EBV-positive DLBCL to promote CD4 T-cell dysfunction by limiting its activation. Conversely, macrophage–CD4 interactions were stronger in EBV-negative DLBCL, suggesting an increased prevalence of macrophage-mediated CD4 T-cell engagement to facilitate an immune-reactive tumor microenvironment with increased T-cell activation and functional immune responses.

As LMP1 was a major viral oncogene that was involved in coordinating a dysfunctional immune landscape in EBV-positive cHL, its involvement in the tumor–macrophage–CD4 T-cell dysfunction axis in DLBCL was assessed **(Figure 5.18)**. LMP1 showed the strongest signal at the center of the ternary plot, suggesting its functional relevance in this tripartite interaction. Its transcriptional activity in malignant cells positively correlated with the abundance of M2-like macrophages, consistent with prior spatial observations in which macrophages closer to malignant cells cells exhibited reduced HLA-DR and increased PD-L1 expression in EBV-positive DLBCL **(Figure 5.14)**. Addi-

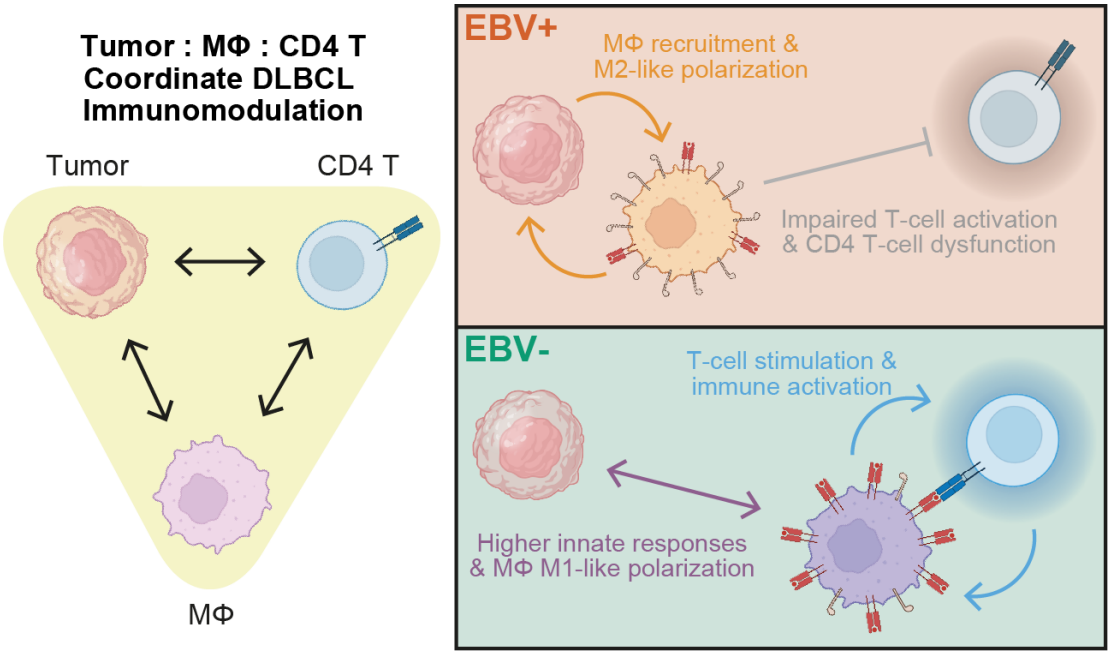


Figure 5.19: Cartoon depicting contrasts in the tumor-macrophage-CD4 interaction axis underlying the highly immunosuppressed EBV-positive *vs.* the less immunosuppressed EBV-negative DLBCL tumor microenvironment. Figure from Yeo et al., 2024 [189].

tionally, the abundance of LMP1-positive malignant cells was positively correlated with CD4 T-cell dysfunction. These findings support LMP1 as a viral factor that also promotes immune dysfunction in EBV-positive DLBCL.

By coupling IN-DEPTH with SGCC, the integrative multi-omics analyses presented here extend the macrophage:CD4 T-cell dysfunction mechanism **(Figure 4.23)** to reveal two distinct and spatially orchestrated cellular circuits in the DLBCL tumor microenvironment based on EBV status. In EBV-positive tissues, malignant cells preferentially associated with macrophages to condition an immunosuppressive microenvironment that impaired CD4 T-cell activation, whereas in EBV-negative cases, an increase in macrophage-CD4 interactions fostered a tumor microenvironment that favored T-cell activation **(Figure 5.19)**. These findings help explain the divergent spatial architectures of EBV-positive and EBV-negative DLBCL by revealing how EBV reprograms cellular interactions in the tumor mi-

croenvironment to enforce an immunosuppressive niche that supports malignant progression. More broadly, the mechanistic frameworks presented herein underscore the importance of integrating spatial and functional layers to decode how viruses rewire tissue ecosystems, offering critical insights for rational therapeutic targeting.

The important thing is not to stop questioning. Curiosity has its own reason for existence.

Albert Einstein

6

Conclusion

The central conclusion of this dissertation is that oncogenic viruses can profoundly remodel the tumor microenvironment to promote tumorigenesis, with immune dysfunction emerging as the common axis linking viral infection to malignancy. These findings yield insights that bridge fundamental biology and translational opportunity, and the methodological advancements established herein may potentially provide a foundation for the next generation of spatial biology.

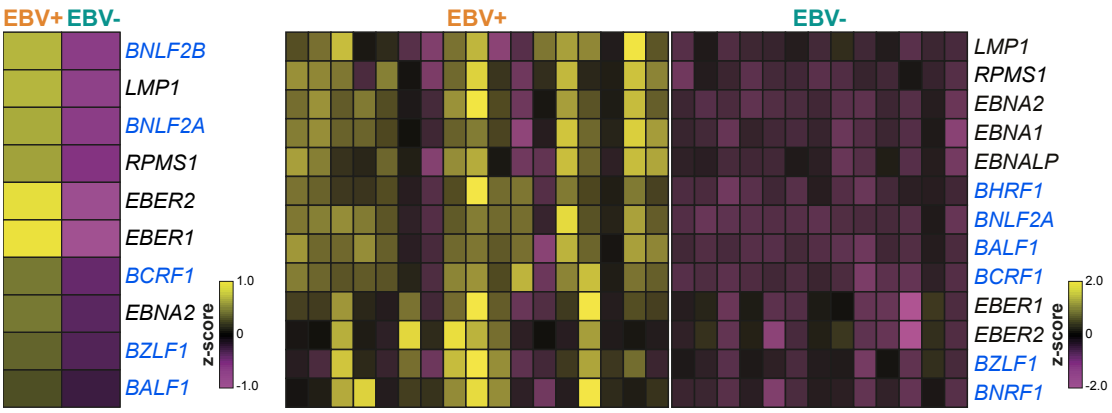


Figure 6.1: A diverse set of EBV lytic (blue) and latent (black) genes is expressed in EBV-positive cHL (left) and DLBCL (right). The cHL data was visualized on a bulk-level whereas the DLBCL data display variation across individual samples, and both data were captured using the GeoMx. Although EBV encodes many additional viral genes, probes were designed only for the subset shown here. All EBV-positive statuses were previously verified via *EBER1* in-situ hybridization. Data from Yeo et al., 2024 [191] and Yeo et al., 2024 [189].

6.1 Biological and technological implications

The study approach presented herein represents a substantive departure from traditional investigations of viral oncogenesis, which have largely examined virus–host interactions in isolation from the tumor ecosystem. By integrating viral biology within the native tumor microenvironment, this dissertation revealed several conceptual advances. Notably, the rigid EBV latency programs classically defined *in vitro* are insufficient to explain the diversity of viral transcriptional states observed *in situ*, as both lytic and latent transcripts were detected across EBV-positive cHL and DLBCL **(Figure 6.1)**, findings consistent with prior reports [16,45]. More broadly, the mere presence of an oncogenic virus could profoundly alter the tumor architecture, yielding a markedly more inflamed microenvironment with increased immune infiltration. Similar immune-inflamed phenotypes have recently been reported across several other virus-associated cancers, including EBV-positive nasopharyngeal carcinoma (NPC), gastric carcinoma, and lymphoepithelioma-like carcinoma of the lung, HPV-positive oropharyngeal squamous cell carcinoma and head and neck squamous cell carcinoma (HNSCC), and

MCPyV-posiitve merkel cell carcinoma (MCC), that were similarly enriched in CD8 T-cells and suppressive M2-like macrophages [38,195,79]. However, current data show that unlike EBV, viral lytic-like late genes are generally not expressed within the respective malignancies for HPV and MCPyV, with only a few viral latent-like early genes detected to date [52,61].

It is important to note that the detection of EBV lytic transcripts does not necessarily indicate productive lytic replication, as this requires confirmation that functional viral proteins were expressed in malignant cells. At the protein level, the expression of the immediate-early protein BZLF1, which mediates the transition from latency to lytic activation, has been detected by immunohistochemistry across several EBV-positive lymphomas including cHL, NPC, Burkitt lymphoma, and post-transplant lymphoproliferative disorder [148,44,141,86]. In EBV-positive DLBCL, immunohistochemical studies have further identified expression of the early lytic protein BMRF1 [45], which functions as the viral DNA polymerase processivity factor. In all cases, lytic protein expression was detected only in a minority of malignant cells and a subset of patients. Therefore, although several lytic transcripts have been detected, current data to date indicate that BZLF1 is the predominant lytic gene that has been detected at the protein level across EBV-positive lymphomas. These results are more consistent with abortive lytic activation rather than full viral replication, and abortive EBV infection has been shown to promote DLBCL development in humanized mouse models by enhancing the expression of BZLF1 [124]. Therefore, it is plausible that transient lytic activation followed by abortive termination contributes to lymphomagenesis, as EBV lytic genes can induce the expression of immunomodulatory cytokines, including IL10, TGFβ, and IFNγ [123]. This would shape a dysfunctional immune landscape within the tumor microenvironment while avoiding full-blown lytic activation that would otherwise trigger robust immune responses.

The association between immune infiltration and viral positivity highlights an important contrast with various cancers that are not associated with viruses. Several of them are often immunologically "cold" and characterized by sparse immune infiltration, where a major therapeutic objective is the re-

cruitment and activation of immune cells to convert the tumor microenvironment into an immunologically "hot" state[209]. However, in virus-associated cancers such as EBV-positive cHL, EBV-positive DLBCL, MCPyV-positive MCC, and HPV-positive HNSCC, immune infiltration is clearly not a rate-limiting step; tumors persist despite robust immune infiltrates due to pervasive dysfunction that include T-cell exhaustion[191,189,79]. These findings underscore a critical principle that the presence of inflammatory immune cells does not necessarily translate into effective immune control, and immune infiltration alone may not be a sufficient surrogate for functional antitumor immunity. This notion is reinforced by notable exceptions of non-viral malignancies, such as those with deficient DNA mismatch repair, that frequently exhibit high levels of immune infiltration driven by elevated mutational burden and neoantigen load[200]. Even in these immunologically "hot" settings, tumors frequently evade immune-mediated elimination through mechanisms that include immune checkpoint engagement, dysfunctional antigen presentation, and the establishment of immunosuppressive cellular and cytokine networks[143]. These observations highlight that effective antitumor immunity requires not only immune cell presence, but also sustained immune functionality within the tumor microenvironment.

Beyond recognizing an important association between viral status and immune infiltration in coordinating a tumorigenic tumor microenvironment, this dissertation firmly place spatial organization as a mechanistic determinant of immune dysfunction. This is a departure from traditional approaches that focus solely on the identity of specific genes and proteins of interest that are differentially expressed without considering their spatial relationship. In EBV-positive lymphomas, immune dysfunction was spatially structured, manifesting as gradients of T-cell exhaustion centered on HRS cells in cHL and as macrophage–CD4 T-cell niches associated with T-cell dysfunction in DLBCL[191,189]. The spatial axis is also an essential component towards understanding lower metastiatic risks in MCPyV-positive MCC, where the aggregation of CD8 T-cells with monocytes that overexpress PD-L1 represented a defining feature of the tumor invasive front[79]. Another notable example is HPV-positive

HNSCC, where traditional bulk sequencing analyses identified increased B-cell infiltration as an independent protective factor as compared to HPV-negative HNSCC [37]. Recent spatial profiling studies have refined this interpretation, revealing that the elevation of B-cells is driven by an increase in activating tertiary lymphoid structures that are situated in close proximity to the tumor core [125]. These findings demonstrate that spatial architecture fundamentally shapes biological interpretation, as distinct immune phenotypes emerge specifically from defined relationships rather than from the identity of specific genes or cell types alone.

The increase in T-cell dysfunction in both EBV-positive cHL and DLBCL supports the notion that immune exhaustion represents a central mechanism exploited by EBV to promote lymphomagenesis. This is reinforced by studies alluding to the same phenomenon across other EBV-associated lymphoid malignancies. For example, EBV-positive primary central nervous system lymphoma had higher expression of *TNF-α*, *CD68*, *CD163*, *PD-L1*, *PD-L2*, *LAG-3*, and *TIM-3* transcripts compared to EBV-negative counterparts, supporting a more inflammatory and immunosuppressive tumor microenvironment [71]. Furthermore, in the presence of human immunodeficiency virus (HIV) infection where the host is significantly immunocompromised, EBV positivity approaches ubiquity across lymphoma subtypes, further implicating EBV as a dominant driver of immune dysfunction when immune surveillance is weakened [49]. Extending beyond EBV, accumulating evidence from other virus-associated cancers supports a shared immunosuppressive paradigm. In MCPyV-positive MCC, tumors frequently exhibit dense CD8 T-cell infiltration spatially coupled to PD-L1 expressing monocytes, a spatial feature consistent with localized induction of T-cell exhaustion; a similar architecture was also described in HPV-positive HNSCC and linked to differential clinical responses [79]. In HPV-positive oropharyngeal squamous cell carcinoma, CD8 T-cells were linked to M2-like macrophages, regulatory T-cells, and malignant cells, a feature associated with longer disease-free survival [195] compared to HPV-negative subtypes. Collectively, these observations support immune dysfunction, particularly the induction of T-cell exhaustion within an inflammatory milieu, as a conserved feature of

viral oncogenesis.

Another central question concerns the diversity and nature of antigens that drive T-cell engagement and dysfunction in virus-associated cancers. In EBV-positive cHL, the significant increase in CD8:MHC Class I interactions raise the question of whether exhaustion is driven by viral and/or tumor-specific antigens. Previous studies have demonstrated the presence of EBV-specific antigens in EBV-positive cHL, yet the natural frequency of T-cells specific for these viral antigens is low[31,23]. This phenomenon has also been observed in other EBV-associated lymphomas, including DLBCL and PCNSL[147,28,114]. These prior findings implicate an important role for non-viral tumor antigens in EBV-positive lymphomas at coordinating T-cell responses, a notion supported by the ability for LMP1 to induce CD4 and CD8 T-cell activation and cytotoxic responses through inducing the expression of non-viral tumor antigens[42]. The presence of T-cells targeting both viral and tumor-specific antigens is also present across other virus-associated cancers, including HBV-positive hepatocellular carcinoma and HPV-positive cervical cancer[72,40,185]. As T-cell antigen specificity is tightly linked to downstream differentiation and functional state[78,144], it is therefore essential to delineate the diversity and relative contributions of viral and tumor-derived antigens in shaping T-cell effector states and dysfunction across virus-associated cancers.

While the induction of T-cell dysfunction appears to be conserved across virus-positive cancers, the clinical consequences of viral infection diverge markedly across disease contexts. In MCC and HNSCC, viral infection is associated with improved outcomes, in part due to enhanced metastatic control linked to spatial co-localization of activated CD8 T-cells with CXCL9+ PD-L1+ macrophages at the invasive front, in contrast to virus-negative tumors where cancer-associated fibroblasts accumulate at this boundary instead[79]. In contrast, EBV-associated cHL and DLBCL follow a distinct immunological logic, where EBV-positive cHL is characterized by terminally exhausted T-cells that are less responsive to immune checkpoint blockade[191], while EBV-positive DLBCL is associated with poorer clinical outcomes despite exhibiting an immunologically ”hotter” tumor microenvironment

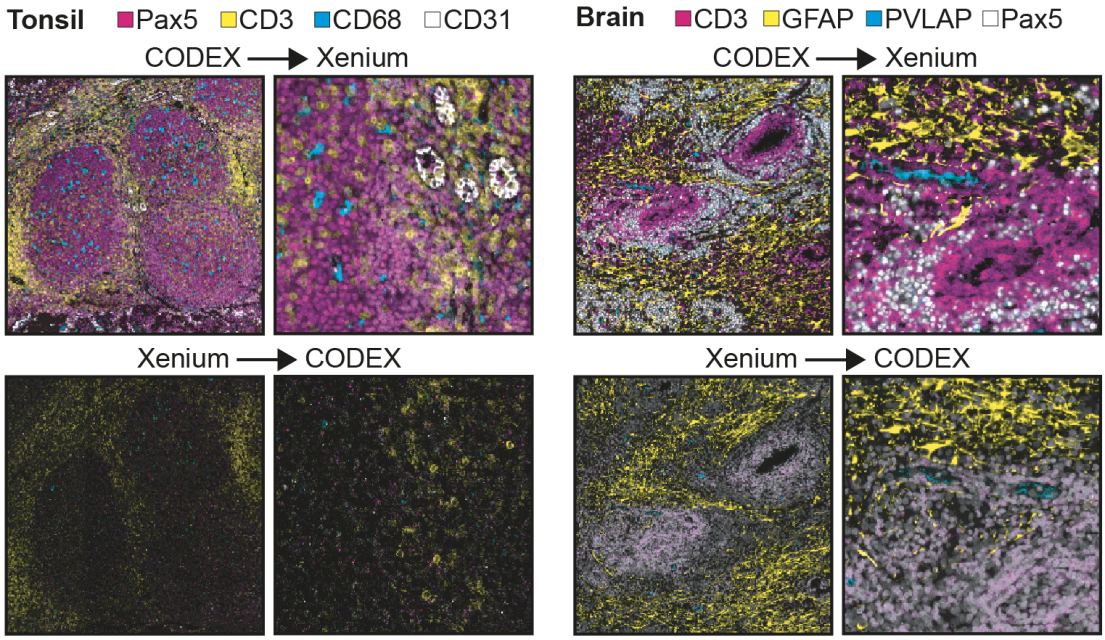


Figure 6.2: Comparison of CODEX imaging quality when applied before (top row) or after (bottom row) Xenium transcriptomics capture on tonsil (left column) and brain (right column) tissues. Notice the substantial loss in data quality after Xenium, and while not shown here, similar results were found for tissues stained after applying GeoMx or CosMx. Data from revision experiments for Yeo et al., 2024 [189].

than EBV-negative counterparts [189]. Although much remains to be understood across other virus-associated cancers, these emerging distinctions suggest that viral positivity does not uniformly predict disease outcome, and instead necessitates consideration of both virus and disease-specific therapeutic strategies.

From a technical perspective, the ability for same-slide spatial multi-omics profiling have recently been demonstrated by several commercial platforms, but most implementations prioritize spatial transcriptomics before proteomics. This ordering presents several practical limitations. The tissue preparation process for spatial transcriptomics frequently lead to irreversible degradation of protein epitopes that result in poor quality for spatial proteomics imaging **(Figure 6.2)**. Transcriptomic assays alone are also insufficient for detecting tissue artifacts or subtle morphological distortions, and given their high operating costs, the inability to pre-screen tissue quality can substantially inflate experimen-

tal cost and reduce biological yield. In parallel, emerging approaches that profile large antibody panels in a single staining step using pre-mixed antibody cocktails overlook the need for antibody-specific titration in FFPE tissues, and differences in binding kinetics and epitope accessibility will likely compromise quantitative accuracy and reproducibility. These considerations highlight the importance of developing practical workflows that retain data quality while maximizing efficiency, the key guiding principles during the design of IN-DEPTH, which are essential for generating robust and interpretable spatial multi-omics datasets.

As with any omics-scale study, an important consideration is whether it is sufficiently powered such that the findings presented reflect a true biological phenomenon rather than sampling artifacts. This concern is particularly salient for spatial omics, where the introduction of the spatial dimension dramatically expands the feature space beyond relative cell type abundances. These higher-order features include cell–cell adjacency networks and tissue neighborhoods, many of which are difficult to predefine a priori and are influenced by the local tissue context. Consequently, statistical power in spatial omics must also account for tissue architecture and entropy, field-of-view sampling strategy, and the spatial scale of the biological features being interrogated. Within this context, the experimental design of this dissertation is well positioned to meet the statistical demands of spatial analysis [13]. By integrating orthogonal spatial proteomics and transcriptomics datasets across multiple independent patient cohorts and leveraging both single-cell (MIBI, CODEX, CosMx) and region-level (GeoMx) platforms, the results presented in this dissertation achieved robust coverage of cellular composition and their spatial relationships. The cohort size and extensive samplings were sufficient to mitigate limitations observed in several spatial studies that rely on relatively few samples or limited fields of view within the chosen spatial technology, thereby increasing confidence in the reproducibility and relevance of the reported spatial relationships. Importantly, the study design applied herein align with emerging frameworks for spatial omics power analysis [13] that underlie optimal experimental design for robust interpretation of spatial features, a key oversight in several spatially-focused studies published

to date.

The incorporation of PLA into spatial proteomics, as demonstrated herein, enabled inference of receptor–ligand interactions *in situ* to provide dynamic dimension for spatial datasets that are otherwise static snapshots of archival tissues. However, while proximity labeling technologies including PLA has been widely applied to study functional dynamics and molecular interactions in tissues [157], its interpretation remains inherently constrained by the principles of proximity-based detection. The most notable limitation for PLA is the generation of false-positive signals, as PLA will label the molecular pair so long as it is within the assay's 40nm detection radius. The detected signal is also influenced by multiple technical factors including antibody affinity, epitope accessibility, steric hindrance between antibody pairs, and the efficiency of enzymatic ligation and rolling circle amplification, thus complicating quantitative comparisons across samples or conditions. While other proximity labeling technologies exist, including APEX, TurboID, and BioID [157], they are designed for *in-vitro* and *in-vivo* settings and not directly applicable for *in-situ* detection. These technical and biological constraints, together with the labor-intensive nature of PLA and current lack of robust multiplexing strategies, restrict its scalability for comprehensive discovery. As a result, the use of PLA in this dissertation is best interpreted as a targeted hypothesis-driven approach, where proximity signals were evaluated alongside orthogonal spatial proteomics measurements to support biological interpretation. Such integration further increases confidence in the relevance of detected proximity events, while demonstrating the capabilities of PLA when applied for functional validation rather than unbiased screening in complex tissues.

With the anticipated expansion of features detected in spatial biology, the analysis and representation of spatial omics data increasingly rely on graph-based methods, which provide a natural framework for encoding and modeling spatial relationships between cells and tissue structures [157]. Emerging algorithms [92,211,58,30] encompass a diverse set of strategies to produce latent embeddings of biological features that enable automatic segmentation of tissues into discrete spatial domains and identify spa-

tially variable gene expression patterns. While these approaches are highly effective and foundational for spatial analysis, they typically focus on resolving spatial structure by assigning unique local neighborhoods. The introduction of SGCC herein represents an advancement by modeling spatial features as a continuous property of tissue organization to identify how biological features co-vary across tissue space, rather than treating spatial structure as a collection of discrete regions that inform specific biological activities. This positions SGCC as a complementary analytical framework that links spatial structure with underlying biological processes across multiple spatial scales.

6.2 Future directions

In hematopathology, the presence of EBV in cHL is considered a heterogeneous and non-defining virological feature, as EBV is detected in only a subset of cases within an otherwise unified disease entity. In contrast, EBV-positive DLBCL is recognized as a distinct clinical and pathological entity, reflecting the association of EBV infection with a reproducible clinical, histological, and molecular characteristics. This dichotomy underscores a broader limitation of current lymphoma classification systems, which rely predominantly on morphological and molecular features centered on malignant cells[2]. The spatially-resolved patterns described in this dissertation, when integrated with established histopathological knowledge, could provide a framework to refine lymphoma classification using criteria that are currently unavailable or not widely scalable in routine diagnostic practice. This approach has the potential to improve prognostic stratification and inform therapeutic strategies that more precisely target the tumor microenvironment.

Direct clinical implications stem from the EBV-linked differences in T-cell functional states[191,189] as the severity of T-cell exhaustion is strongly correlated with poorer responses to checkpoint blockade, where T-cells becomes unresponsive upon reaching terminal exhaustion[134]. This provides a biological basis regarding the increased risk of relapse for EBV-positive cHL[88] and poorer prognosis for

EBV-positive DLBCL[121], which also underscores the importance of incorporating viral status into clinical strategies. Therefore, there is a strong rationale to re-analyze existing clinical trial datasets stratified by viral status to inform personalized immunotherapeutics, given the potential need for greater emphasis on strategies that restore T-cell tumoricidal activity in EBV-positive lymphomas. In parallel, defining the breadth and immunodominance of their antigenic repertoires represent a critical and complementary objective, as these features fundamentally shape T-cell engagement, immune escape, and therapeutic durability. The ability for viral oncogenes such as LMP1 to promote T-cell dysfunction and regulate the tumor antigenic landscape[42] further motivates potential virus-targeted therapeutics, a strategy that has shown favorable safety profile even when administered to vulnerable transplant patients[194].

Although these findings implicate EBV as a major clinical feature that promotes T-cell dysfunction in the lymphoma tumor microenvironment[191,189], insights remain limited to cHL and DLBCL. There are also very few studies comparing virus-positive and virus-negative cancers from a spatially-resolved perspective, with only MCC and HNSCC being the most prominent examples to date[79,125]. A comprehensive understanding of the principles behind virus-mediated remodeling of the tumor microenvironment will require extending such studies to other cancers and across the six additional human oncogenic viruses. The functional states of immune populations beyond CD4 and CD8 T-cells will also require further dissection to fully delineate the landscape of immune dysfunction. However, compared to T-cells, the spectrum of characterized dysfunctional states in other immune cell types remains more limited. Nonetheless, the spatial multi-modal approaches implemented here provide opportunities to understand viral and cancer dynamics, paving way towards an integrative atlas of virus-driven reprogramming of tumor–immune-stromal interactions in the tumor microenvironment.

The variable expression of viral oncogenes in malignant cells, exemplified by the inconsistent and dynamic range of LMP1 expression in EBV-positive HRS cells **(Figure 6.3)**[191], reframes the con-

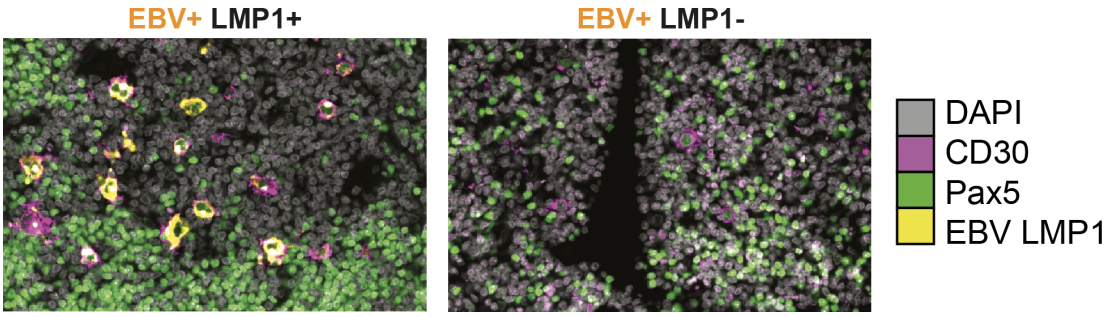


Figure 6.3: LMP1 is among the principal oncogenic genes of EBV and is often assumed to be expressed in all infected cells. However, although all HRS cells in EBV-positive cHL express *EBER1*, only a subset expressed LMP1 in EBV-positive tissues that contain LMP1-positive HRS cells (left), and there are also EBV-positive tissues in which LMP1 is absent (right). Data from Yeo et al., 2024 [191].

cept of tumor heterogeneity by illustrating that functional and phenotypic diversity may arise independently of genetic divergence. This perspective challenges the classical view of malignant cells as a uniform clonal population [142,76], underscoring the notion that a clonal origin does not preclude such diversity, particularly at the intersection of viral latency programs and host oncogenesis. This added layer of complexity is especially relevant for EBV-associated lymphomas, where multiple viral latent and lytic genes were detected, with heterogeneous expression patterns both within and across patients **(Figure 6.1)**. These observations highlight a critical but unmet need for validated antibodies and sensitive detection probes targeting viral oncogenes, which would enable more precise delineation of viral phenotypes *in situ*. Such reagents would also be particularly valuable for contextualizing the body of *in-vitro* literature on EBV oncogenesis with specific states of EBV-associated disease by linking viral gene expression dynamics to functional consequences; for example, by clarifying how abortive lytic activation rewires malignant cell signaling to augment immunosuppression within the tumor microenvironment.

Despite the resource intensive nature of antibody panel development, new chemistries are continuously being optimized to improve labeling efficiency and retention of antibody specificity. In parallel, proximity labeling technologies such as PLA that provide a complementary avenue to in-

fer molecular interactions *in situ* remains limited in their utility due to technical constrains and a lack of multiplexing capacity. Incremental improvements aimed at improving multiplexing[122] and leveraging smaller affinity reagents such as nanobodies[14] currently are being made to enhance the specificity and interpretability of these measurements. Furthermore, antibody-independent spatial proteomics are emerging such as deep visual proteomics[136], which combines laser microdissection with mass spectrometry–based proteome profiling of the tissue section. While this approach enables unbiased proteome coverage, its capacity for robust cell type identification remains limited due to the complexity and contextual ambiguity of protein expression across diverse cell populations. Ultimately, future advancements should implement hybrid strategies that integrate the high specificity of antibody-based detection and interaction-aware proximity assays with the depth and breadth of untargeted proteomics, which is potentially enabled through IN-DEPTH[189]. Such convergence would harmonize precision and scalability, thus providing a practical approach towards targeted and comprehensive spatial profiling.

A critical gap remains in translating biological and clinical insights from spatial datasets into actionable therapeutic strategies. This is because spatial biology primarily captures emergent functional interactions rather than establishes causality or directionality, and thus does not map directly onto discrete druggable targets. Moreover, spatially-informed insights are derived from static snapshots of the disease, which limits inference of the temporal dynamics and causative interactions that govern therapeutic response. Additional barriers include patient-to-patient heterogeneity, batch effects, limited cross-cohort validation of spatial features, and practical constraints related to scalability and data standardization. Collectively, these challenges underscore the need for strategies that enable rapid evaluation of candidate interventions, which can inform therapeutic development while serving as a functional validation of potential biological mechanisms. A major obstacle to this effort is the lack of animal models that faithfully recapitulate disease biology, particularly for human oncogenic viruses that are highly adapted to the human host. This limitation calls for new experimental systems for

preclinical testing, including organoid[43], organ-on-chip[19], and patient-derived tumor fragment platforms[197], which provide physiologically relevant and tractable models of disease. While challenging to develop, these systems are amenable to controlled perturbation and are compatible with both spatial and traditional multi-omics profiling.

As experimental platforms continue to mature, new analytical frameworks that integrate multi-dimensional data will be essential for extracting mechanistic insights. The developments in this dissertation, by combining spatial proteomics and transcriptomics, have provided high-resolution insights into the biology of virus-associated cancers. With continued technological progress, such approaches are poised to deliver increasingly mechanistic and comprehensive insights into disease processes. Future directions include the expansion of spatial multi-omics to encompass additional molecular modalities, such as glycomics and metabolomics, to enable a more complete understanding of post-translational regulation and metabolic states that govern cell-cell communication. In parallel, the development of *in-situ* multiplexed proximity labeling would enable a comprehensive view of molecular interactions and signaling networks that coordinate spatial reorganization during disease progression. These advances may be empowered by the optimization principles of IN-DEPTH, which emphasize the retention of tissue morphology and preservation of molecular features, and the mathematical formulation of SGCC, which provides a feature-agnostic representation of biological features that is applicable across diverse contexts.

While SGCC provides a robust framework for identifying spatially coordinated biological processes, its current implementation primarily emphasizes smooth low-frequency graph Fourier modes. This design parallels related graph spectral methods, as low-frequency components are generally considered to be the most reflective of coordinated biological processes[30,94]. However, observations from other spatially structured biological systems suggest that medium and high-frequency Fourier modes can encode distinct classes of functional biological phenomena characterized by more localized and rapidly varying spatial patterns; for instance, in functional neuroimaging where high-frequency com-

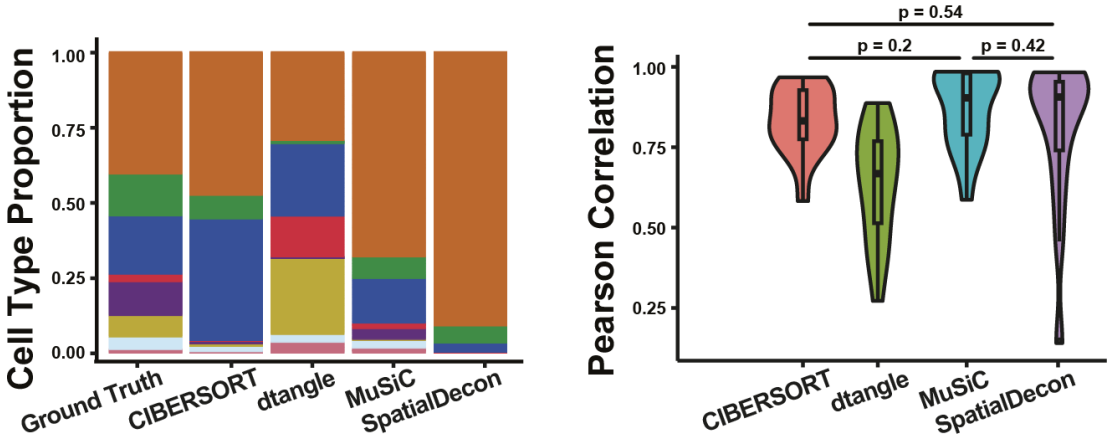


Figure 6.4: Comparison of cell type deconvolution methods using spatial transcriptomics data from tonsil tissues profiled with CODEX–GeoMx IN-DEPTH. The evaluated algorithms were generally inaccurate especially when resolving granular cellular populations (left) and produced inconsistent estimates across methods (right). Figure from Yeo et al., 2024[189].

ponents could capture transient brain activities in response to external stimuli[94]. In the context of viral and cancer immunology, high-frequency components may reflect transient features including localized immune activation around infected or malignant cells, as well as specific immune niches driven by local cytokine gradients. Future extensions of SGCC that incorporate multiple frequencies could provide a more comprehensive view of tissue organization by integrating global spatial patterns with localized functional processes.

The vast amount of high-quality data generated in this dissertation will also enable future computational advancements. For example, the application of IN-DEPTH on tonsil tissues using the CODEX-GeoMx implementation provided ground-truth transcriptomic capture of precisely defined cell populations, which can refine bulk deconvolution methods. Initial benchmarking on four popular deconvolution algorithms (CIBERSORT[140], dtangle[99], MuSic[199], SpatialDecon[50]) revealed inconsistent results when estimating the relative abundances across all cell populations, highlighting a need for improvement **(Figure 6.4)**. Additionally, the spatial coordination of T-cell dysfunction in EBV-positive cHL and DLBCL suggests that pseudotime trajectory concepts from single-cell RNA se-

quencing could be adapted into a "spatial pseudotime" framework, enabling dynamic reconstruction of immune dysfunction as a process that evolves within the tissue. Such approaches would further emphasize spatial context as an essential axis of biological function rather than treating it as an orthogonal feature appended to discrete molecular measurements, a concept that is now beginning to be explored [178,110,98].

Finally, the cumulative efforts of this dissertation are directed towards a long-term vision of a unified view of the tissue architecture in order to fully understand the biological mechanisms that lead to disease. Such a vision would be realized through three-dimensional live imaging across multiple biological modalities while remaining perturbable at all levels. At present, spatial technologies remain largely limited by their inherently two-dimensional and static nature. Early progress toward this goal includes three-dimensional spatial proteomics [18] and spatial transcriptomics approaches [65], emerging technologies for live-cell imaging [111] and transcriptomics [36], and *in vivo* CRISPR screening methods [170]. In parallel, rapid advancements in computational biology are providing a complementary path forward. The emergence of general-purpose foundation models trained on large-scale spatial datasets offers a powerful framework for learning shared and transferable biological representations that capture tissue organization, cell-cell interactions, and molecular signatures. For example, CellSAM, a cell segmentation foundation model, has addressed a critical bottleneck in spatial data analysis by achieving near human-level performance across biological systems ranging from prokaryotic to eukaryotic model organisms, and across multiple imaging modalities [126]. Building on this paradigm, recent spatial proteomics foundation models such as KRONOS [173] and EVA [118] demonstrate can rapidly distill complex imaging datasets into meaningful spatial representations to enable tasks such as cell phenotyping, neighborhood identification, and patient stratification without extensive manual intervention. Together with virtual cell simulations [188] and other emerging generative modeling approaches, these developments may ultimately support predictive and mechanistic hypotheses that bridge static observations with dynamic biological processes.

Achieving this vision will be very challenging, but as history has shown, progress in biology requires daring ambition. After all, success will only come to those who dare to dream, have the courage to pursue those dreams, and the determination to realize them.

A

Information on Tumor Specimens

The following tissue samples were used in this study and graciously provided by Dana-Farber Cancer Institute (IRB# 2016P002769, 2014P001026), University of Rochester Medical Center (IRB# STUDY1159), and University of Tübingen. Each row represents a unique patient, and multiple samples may have been used per patient. The metadata was used for batch effect correction for spatial transcriptomics analysis. Detailed information can be found in the corresponding articles [191,189].

Table A.1: Metadata for the Dana-Farber Cancer Institute (DFCI) cohort for MIBI in the EBV cHL study.

Institution	Sex	Histologic Subtype	EBV
DFCI	M	Mixed Cellularity	+
DFCI	M	Mixed Cellularity	-
DFCI	F	Mixed Cellularity	+
DFCI	F	Mixed Cellularity	-
DFCI	M	Mixed Cellularity	+
DFCI	M	Mixed Cellularity	-
DFCI	F	Mixed Cellularity	+
DFCI	M	Mixed Cellularity	-
DFCI	F	Mixed Cellularity	-
DFCI	M	Mixed Cellularity	-
DFCI	F	Nodular Sclerosis	-
DFCI	M	Nodular Sclerosis	-
DFCI	F	Nodular Sclerosis	-
DFCI	M	Nodular Sclerosis	-
DFCI	F	Nodular Sclerosis	-
DFCI	F	Nodular Sclerosis	-
DFCI	F	Nodular Sclerosis	-
DFCI	M	Nodular Sclerosis	-
DFCI	M	Mixed Cellularity	+
DFCI	M	Mixed Cellularity	+

Table A.2: Metadata for the Dana-Farber Cancer Institute (DFCI) cohort for GeoMx in the EBV cHL study.

Institution	Sex	Histologic Subtype	Age	Biopsy Site	EBV
DFCI	M	Nodular Sclerosis	22	Mass (left neck)	-
DFCI	M	Nodular Sclerosis	25	LN (station 4R)	+
DFCI	M	Mixed Cellularity	60	LN (left neck level 5)	+
DFCI	M	Mixed Cellularity	47	Mass (submental)	-
DFCI	M	Mixed Cellularity	73	LN (station 7)	-
DFCI	F	Mixed Cellularity	56	LN (right cervical)	+
DFCI	M	Mixed Cellularity	21	LN (left neck)	-
DFCI	Unknown	Mixed Cellularity	32	Mass (right neck)	-
DFCI	F	Mixed Cellularity	18	LN (right axillary)	-
DFCI	M	Nodular Sclerosis	27	LN (deep cervical)	+
DFCI	F	Nodular Sclerosis	29	Mass (left neck)	-
DFCI	M	Nodular Sclerosis	41	Mass (right supraclavicular)	-
DFCI	M	Nodular Sclerosis	52	LN (station 4R)	+
DFCI	M	Nodular Sclerosis	19	Mass (left neck)	-
DFCI	M	Nodular Sclerosis	50	LN (deep cervical)	+
DFCI	M	Nodular Sclerosis	71	LN (left cervical)	+
DFCI	F	Nodular Sclerosis	62	LN (right inguinal)	-
DFCI	F	Nodular Sclerosis	22	Mass (sternal)	-
DFCI	Unknown	Nodular Sclerosis	46	LN (unknown)	+
DFCI	M	Nodular Sclerosis	44	LN (right inguinal)	-
DFCI	M	Mixed Cellularity	66	LN (left supraclavicular)	+
DFCI	M	Nodular Sclerosis	33	LN (level 3)	-
DFCI	M	Mixed Cellularity	65	LN (right axillary)	+

Table A.3: Metadata for the University of Rochester Medical Center cohort for GeoMx in the EBV cHL study. NOS, not otherwise specified.

Institution	Sex	Histologic Subtype	Age	Biopsy Site	EBV
Rochester	F	Mixed Cellularity	25	LN (retroperitoneal)	+
Rochester	M	Nodular Sclerosis	31	LN (supraclavicular)	-
Rochester	M	Nodular Sclerosis	23	LN (cervical; level 2)	-
Rochester	M	Nodular Sclerosis	24	LN (supraclavicular)	-
Rochester	M	Mixed Cellularity	18	Mass (neck)	+
Rochester	M	Nodular Sclerosis	69	LN (cervical)	-
Rochester	F	Nodular Sclerosis	27	LN (axillary)	-
Rochester	F	Nodular Sclerosis	25	LN (cervical; level 5)	-
Rochester	M	Mixed Cellularity	38	LN (scalene)	-
Rochester	F	Mixed Cellularity	22	LN (cervical; level 5)	-
Rochester	F	Mixed Cellularity	37	LN (cervical; level 6)	+
Rochester	M	NOS	53	LN (axillary)	+
Rochester	M	Nodular Sclerosis	62	LN (periaortic)	+
Rochester	M	Mixed Cellularity	24	LN (axillary)	+
Rochester	F	Nodular Sclerosis	61	LN (inguinal)	+
Rochester	F	Nodular Sclerosis	23	LN (cervical, level 1b)	-
Rochester	F	Nodular Sclerosis	85	LN (axillary)	-
Rochester	M	Nodular Sclerosis	54	LN (supraclavicular)	-
Rochester	M	Nodular Sclerosis	66	LN (paraaortic)	+
Rochester	M	Mixed Cellularity	80	LN (axillary)	+
Rochester	M	Mixed Cellularity	22	LN (cervical; level 2)	+
Rochester	M	Mixed Cellularity	27	LN (submandibular)	+
Rochester	F	NOS	52	LN (axillary)	+

Table A.4: Metadata for the Dana-Farber Cancer Institute (DFCI) cohort for CosMx in the EBV cHL study.

Institution	Sex	Histologic Subtype	Age	Biopsy Site	EBV
DFCI	M	Nodular Sclerosis	25	LN (station 4R)	+
DFCI	M	Nodular Sclerosis	27	LN (deep cervical)	+
DFCI	M	Nodular Sclerosis	52	LN (station 4R)	+
DFCI	M	Nodular Sclerosis	45	LN (subcarinal level 7)	+
DFCI	M	Nodular Sclerosis	71	LN (left cervical)	+
DFCI	Unknown	Nodular Sclerosis	46	LN (unknown)	+
DFCI	M	Nodular Sclerosis	22	Mass (left neck)	-
DFCI	F	Nodular Sclerosis	29	Mass (left neck)	-
DFCI	M	Nodular Sclerosis	41	Mass (right supraclavicular)	-
DFCI	M	Nodular Sclerosis	19	Mass (left neck)	-
DFCI	F	Nodular Sclerosis	22	Mass (sternal)	-
DFCI	M	Nodular Sclerosis	44	LN (right inguinal)	-
DFCI	M	Nodular Sclerosis	33	LN (level 3)	-
DFCI	Unknown	Mixed Cellularity	32	Mass (right neck)	-
DFCI	F	Nodular Sclerosis	38	Nodule (right lung)	+
DFCI	M	Mixed Cellularity	65	LN (right axillary)	+
DFCI	M	Mixed Cellularity	60	LN (left neck level 5)	+
DFCI	F	Mixed Cellularity	27	LN (left cervical)	+
DFCI	F	Mixed Cellularity	56	LN (right cervical)	+
DFCI	M	Mixed Cellularity	47	Mass (submental)	-
DFCI	M	Mixed Cellularity	21	LN (left neck)	-

Table A.5: Metadata for the University of Tübingen cohort for CosMx in the EBV cHL study. NOS, not otherwise specified.

Institution	Histologic Subtype	EBV
Tübingen	Lymphocyte Rich	-
Tübingen	Mixed Cellularity	+
Tübingen	Lymphocyte Rich	+
Tübingen	Nodular Sclerosis	-
Tübingen	Nodular Sclerosis	-
Tübingen	Mixed Cellularity	+
Tübingen	Nodular Sclerosis	-
Tübingen	NOS	+
Tübingen	Nodular Sclerosis	-
Tübingen	Nodular Sclerosis	-
Tübingen	Nodular Sclerosis	-
Tübingen	Nodular Sclerosis	-
Tübingen	Nodular Sclerosis	-
Tübingen	Mixed Cellularity	+
Tübingen	Nodular Sclerosis	-
Tübingen	Mixed Cellularity	+
Tübingen	NOS	-
Tübingen	Mixed Cellularity	+
Tübingen	NOS	+
Tübingen	Mixed Cellularity	+
Tübingen	Mixed Cellularity	+
Tübingen	Mixed Cellularity	+

Table A.6: Metadata for the University of Rochester Medical Center cohort for CODEX-PLA in the EBV cHL study. NOS, not otherwise specified.

Institution	Histologic Subtype	EBV
Rochester	Nodular Sclerosis	+
Rochester	Mixed Cellularity	+
Rochester	Mixed Cellularity	+
Rochester	Mixed Cellularity	+
Rochester	NOS	+
Rochester	Nodular Sclerosis	+
Rochester	Nodular Sclerosis	-
Rochester	Nodular Sclerosis	-
Rochester	Nodular Sclerosis	-
Rochester	NOS	+
Rochester	Mixed Cellularity	+
Rochester	Nodular Sclerosis	-

Table A.7: Metadata for the University of Tübingen cohort for CODEX-PLA in the EBV cHL study. NOS, not otherwise specified.

Institution	Histologic Subtype	EBV
Tübingen	Lymphocyte Rich	+
Tübingen	Lymphocyte Rich	-
Tübingen	Lymphocyte Rich	-
Tübingen	NOS	+
Tübingen	Mixed Cellularity	+
Tübingen	Nodular Sclerosis	-
Tübingen	Nodular Sclerosis	-
Tübingen	Lymphocyte Rich	+
Tübingen	Nodular Sclerosis	-
Tübingen	Nodular Sclerosis	-
Tübingen	Mixed Cellularity	+
Tübingen	Mixed Cellularity	+
Tübingen	Nodular Sclerosis	-
Tübingen	Nodular Sclerosis	-
Tübingen	NOS	+

Table A.8: Metadata for the Dana-Farber Cancer Institute (DFCI) cohort for CODEX-GeoMx in the IN-DEPTH study.

Institution	Sex	Age	Biopsy Site	EBV
DFCI	M	52	Mass (neck)	+
DFCI	M	27	SMBWL (ileocecum)	+
DFCI	F	78	LN (cervical)	+
DFCI	M	40	LN (axillary)	+
DFCI	M	69	LN (neck)	+
DFCI	M	67	LN (cervical)	+
DFCI	M	75	LN (axillary)	-
DFCI	M	29	Mass (mesenteric)	-
DFCI	F	59	Mass (chest wall)	-
DFCI	M	56	LN (neck)	-
DFCI	M	43	LN (arm, soft tissue)	-
DFCI	F	68	Axillary (nasopharynx)	-
DFCI	M	77	LN (cervical)	-
DFCI	M	87	Mass (arm)	-

Table A.9: Metadata for the University of Rochester Medical Center cohort for CODEX-GeoMx in the IN-DEPTH study.

Institution	Sex	Age	Biopsy Site	EBV
Rochester	F	21	LN (cervical)	+
Rochester	M	78	LN (cervical)	-
Rochester	M	66	LN (cervical)	+
Rochester	M	63	SMBWL	+
Rochester	F	39	LN (cervical)	+
Rochester	F	60	Spleen	-
Rochester	M	64	LN (inguinal)	-
Rochester	F	64	Lung	+
Rochester	M	58	LN (axillary)	+
Rochester	M	69	Mass (anal)	+
Rochester	M	70	LN (inguinal)	+
Rochester	F	52	LN (axillary)	+
Rochester	M	63	LN (mesenteric)	-
Rochester	M	62	LN (inguinal)	+
Rochester	F	64	LN (axillary)	-
Rochester	M	64	SMBWL (colon)	+

Table A.10: Metadata for the University of Tübingen cohort for CosMx in the IN-DEPTH study.

Institution	Comments	EBV
Tübingen	In part Hodgkin-like morphology	+
Tübingen	With Hodgkin-like morphology, latency type 2	+
Tübingen	Non-germinal center type	+
Tübingen	Not otherwise specified	+
Tübingen	Not otherwise specified	+
Tübingen	/	+
Tübingen	With latency type 2	+
Tübingen	Not otherwise specified	+
Tübingen	/	-
Tübingen	/	-
Tübingen	/	-
Tübingen	/	-
Tübingen	/	-
Tübingen	/	-
Tübingen	/	-
Tübingen	/	-
Tübingen	/	-
Tübingen	/	-

References

[1] Al-Hadidi, S. A. & Lee, H. J. (2021). Checkpoint inhibition therapy in transplant-ineligible relapsed or refractory classic hodgkin lymphoma. *JCO Oncology Practice*, 17(2), 64–71.

[2] Alaggio, R., Amador, C., Anagnostopoulos, I., Attygalle, A. D., Araujo, I. B. d. O., Berti, E., Bhagat, G., Borges, A. M., Boyer, D., Calaminici, M., et al. (2022). The 5th edition of the world health organization classification of haematolymphoid tumours: lymphoid neoplasms. *Leukemia*, 36(7), 1720–1748.

[3] Aliazis, K., Christofides, A., Shah, R., **Yeo, Y. Y.**, Jiang, S., Charest, A., & Boussiotis, V. A. (2025). The tumor microenvironment's role in the response to immune checkpoint blockade. *Nature Cancer*, 6(6), 924–937.

[4] Anderson, A. C., Joller, N., & Kuchroo, V. K. (2016). Lag-3, tim-3, and tigit: Co-inhibitory receptors with specialized functions in immune regulation. *Immunity*, 44(5), 989–1004.

[5] Anderson, N. M. & Simon, M. C. (2020). The tumor microenvironment. *Current Biology*, 30(16), R921–R925.

[6] Angelo, M., Bendall, S. C., Finck, R., Hale, M. B., Hitzman, C., Borowsky, A. D., Levenson, R. M., Lowe, J. B., Liu, S. D., Zhao, S., et al. (2014). Multiplexed ion beam imaging of human breast tumors. *Nature Medicine*, 20(4), 436–442.

[7] Ansell, S. M., Lesokhin, A. M., Borrello, I., Halwani, A., Scott, E. C., Gutierrez, M., Schuster, S. J., Millenson, M. M., Cattry, D., Freeman, G. J., et al. (2015). Pd-1 blockade with nivolumab in relapsed or refractory hodgkin's lymphoma. *New England Journal of Medicine*, 372(4), 311–319.

[8] Ansell, S. M., Minnema, M. C., Johnson, P., Timmerman, J. M., Armand, P., Shipp, M. A., Rodig, S. J., Ligon, A. H., Roemer, M. G., Reddy, N., et al. (2019). Nivolumab for relapsed/refractory diffuse large b-cell lymphoma in patients ineligible for or having failed autologous transplantation: a single-arm, phase ii study. *Journal of Clinical Oncology*, 37(6), 481–489.

[9] Armand, P., Shipp, M. A., Ribrag, V., Michot, J.-M., Zinzani, P. L., Kuruvilla, J., Snyder, E. S., Ricart, A. D., Balakumaran, A., Rose, S., et al. (2016). Programmed death-1 blockade

with pembrolizumab in patients with classical hodgkin lymphoma after brentuximab vedotin failure. *Journal of Clinical Oncology*, 34(31), 3733–3739.

[10] Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., et al. (2000). Gene ontology: tool for the unification of biology. *Nature Genetics*, 25(1), 25–29.

[11] Baessler, A. & Vignali, D. A. (2024). T cell exhaustion. *Annual Review of Immunology*, 42(1), 179–206.

[12] Bai, Y., Zhu, B., Rovira-Clave, X., Chen, H., Markovic, M., Chan, C. N., Su, T.-H., McIlwain, D. R., Estes, J. D., Keren, L., et al. (2021). Adjacent cell marker lateral spillover compensation and reinforcement for multiplexed images. *Frontiers in Immunology*, 12, 652631.

[13] Baker, E. A., Schapiro, D., Dumitrascu, B., Vickovic, S., & Regev, A. (2023). In silico tissue generation and power analysis for spatial omics. *Nature Methods*, 20(3), 424–431.

[14] Baldinotti, R., Pauzin, F. P., Fevang, H., Ishizuka, Y., & Bramham, C. R. (2025). A nanobody-based proximity ligation assay detects constitutive and stimulus-regulated native arc/arg3. 1 oligomers in hippocampal neuronal dendrites. *Molecular Neurobiology*, 62(4), 3973–3990.

[15] Bandura, D. R., Baranov, V. I., Ornatsky, O. I., Antonov, A., Kinach, R., Lou, X., Pavlov, S., Vorobiev, S., Dick, J. E., & Tanner, S. D. (2009). Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Analytical Chemistry*, 81(16), 6813–6822.

[16] Bayda, N., Tilloy, V., Chaunavel, A., Bahri, R., Halabi, M. A., Feuillard, J., Jaccard, A., & Ranger-Rogez, S. (2021). Comprehensive epstein-barr virus transcriptome by rna-sequencing in angioimmunoblastic t cell lymphoma (aitl) and other lymphomas. *Cancers*, 13(4), 610.

[17] Bergholtz, H., Carter, J. M., Cesano, A., Cheang, M. C. U., Church, S. E., Divakar, P., Fuhrman, C. A., Goel, S., Gong, J., Guerriero, J. L., et al. (2021). Best practices for spatial profiling for breast cancer research with the geomx® digital spatial profiler. *Cancers*, 13(17), 4456.

[18] Bhatia, H. S., Brunner, A.-D., Öztürk, F., Kapoor, S., Rong, Z., Mai, H., Thielert, M., Ali, M., Al-Maskari, R., Paetzold, J. C., et al. (2022). Spatial proteomics in three-dimensional intact specimens. *Cell*, 185(26), 5040–5058.

[19] Bhatia, S. N. & Ingber, D. E. (2014). Microfluidic organs-on-chips. *Nature Biotechnology*, 32(8), 760–772.

[20] Bjornevik, K., Cortese, M., Healy, B. C., Kuhle, J., Mina, M. J., Leng, Y., Elledge, S. J., Niebuhr, D. W., Scher, A. I., Munger, K. L., et al. (2022). Longitudinal analysis reveals high prevalence of epstein-barr virus associated with multiple sclerosis. *Science*, 375(6578), 296–301.

[21] Black, S., Phillips, D., Hickey, J. W., Kennedy-Darling, J., Venkataraaman, V. G., Samusik, N., Goltsev, Y., Schürch, C. M., & Nolan, G. P. (2021). Codex multiplexed tissue imaging with dna-conjugated antibodies. *Nature Protocols*, 16(8), 3802–3835.

[22] Bobrow, M. N., Harris, T. D., Shaughnessy, K. J., & Litt, G. J. (1989). Catalyzed reporter deposition, a novel method of signal amplification application to immunoassays. *Journal of Immunological Methods*, 125(1-2), 279–285.

[23] Bollard, C. M., Gottschalk, S., Torrano, V., Diouf, O., Ku, S., Hazrat, Y., Carrum, G., Ramos, C., Fayad, L., Shpall, E. J., et al. (2014). Sustained complete responses in patients with lymphoma receiving autologous cytotoxic t lymphocytes targeting epstein-barr virus latent membrane proteins. *Journal of Clinical Oncology*, 32(8), 798–808.

[24] Bonadonna, G., Zucali, R., Monfardini, S., de Lena, M., & Uslenghi, C. (1975). Combination chemotherapy of hodgkin's disease with adriamycin, bleomycin, vinblastine, and imidazole carboxamide versus mopp. *Cancer*, 36(1), 252–259.

[25] Boyer, S. N., Wazer, D. E., & Band, V. (1996). E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. *Cancer Research*, 56(20), 4620–4624.

[26] Bressan, D., Battistoni, G., & Hannon, G. J. (2023). The dawn of spatial omics. *Science*, 381(6657), eabq4964.

[27] Carbone, A., Vaccher, E., Gloghini, A., Pantanowitz, L., Abayomi, A., De Paoli, P., & Franceschi, S. (2014). Diagnosis and management of lymphomas and other cancers in hiv-infected patients. *Nature Reviews Clinical Oncology*, 11(4), 223–238.

[28] Cárdenas, D., Vélez, G., Orfao, A., Herrera, M., Solano, J., Olaya, M., Uribe, A., Saavedra, C., Duarte, M., Rodríguez, M., et al. (2015). Epstein–barr virus-specific cd8+ t lymphocytes from diffuse large b cell lymphoma patients are functionally impaired. *Clinical & Experimental Immunology*, 182(2), 173–183.

[29] Carey, C. D., Gusenleitner, D., Lipschitz, M., Roemer, M. G., Stack, E. C., Gjini, E., Hu, X., Redd, R., Freeman, G. J., Neuberg, D., et al. (2017). Topological analysis reveals a pd-l1-associated microenvironmental niche for reed-sternberg cells in hodgkin lymphoma. *Blood*, 130(22), 2420–2430.

[30] Chang, Y., Liu, J., Jiang, Y., Ma, A., **Yeo, Y. Y.**, Guo, Q., McNutt, M., Krull, J. E., Rodig, S. J., Barouch, D. H., et al. (2024). Graph fourier transform for spatial omics representation and analyses of complex organs. *Nature Communications*, 15(1), 7467.

[31] Chapman, A. L., Rickinson, A. B., Thomas, W. A., Jarrett, R. F., Crocker, J., & Lee, S. P. (2001). Epstein-barr virus-specific cytotoxic t lymphocyte responses in the blood and tumor

site of hodgkin's disease patients: implications for a t-cell-based therapy. *Cancer Research*, 61(16), 6219–6226.

[32] Chapuy, B., Stewart, C., Dunford, A. J., Kim, J., Kamburov, A., Redd, R. A., Lawrence, M. S., Roemer, M. G., Li, A. J., Ziepert, M., et al. (2018). Molecular subtypes of diffuse large b cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nature Medicine*, 24(5), 679–690.

[33] Chen, B. J., Chapuy, B., Ouyang, J., Sun, H. H., Roemer, M. G., Xu, M. L., Yu, H., Fletcher, C. D., Freeman, G. J., Shipp, M. A., et al. (2013). Pd-l1 expression is characteristic of a subset of aggressive b-cell lymphomas and virus-associated malignancies. *Clinical Cancer Research*, 19(13), 3462–3473.

[34] Chen, H. & Murphy, R. F. (2023). Evaluation of cell segmentation methods without reference segmentations. *Molecular Biology of the Cell*, 34(6), ar50.

[35] Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S., & Zhuang, X. (2015). Spatially resolved, highly multiplexed rna profiling in single cells. *Science*, 348(6233), aaa6090.

[36] Chen, W., Guillaume-Gentil, O., Rainer, P. Y., Gäbelein, C. G., Saelens, W., Gardeux, V., Klaeger, A., Dainese, R., Zachara, M., Zambelli, T., et al. (2022). Live-seq enables temporal transcriptomic recording of single cells. *Nature*, 608(7924), 733–740.

[37] Chen, X., Yan, B., Lou, H., Shen, Z., Tong, F., Zhai, A., Wei, L., & Zhang, F. (2018). Immunological network analysis in hpv associated head and neck squamous cancer and implications for disease prognosis. *Molecular Immunology*, 96, 28–36.

[38] Chen, Y., Ouyang, D., Wang, Y., Pan, Q., Zhao, J., Chen, H., Yang, X., Tang, Y., Wang, Q., Li, Y., et al. (2024). Ebv promotes tcr-t-cell therapy resistance by inducing cd163+ m2 macrophage polarization and mmp9 secretion. *Journal for Immunotherapy of Cancer*, 12(6), e008375.

[39] Chen, Z., Soifer, I., Hilton, H., Keren, L., & Jojic, V. (2020). Modeling multiplexed images with spatial-lda reveals novel tissue microenvironments. *Journal of Computational Biology*, 27(8), 1204–1218.

[40] Cheng, Y., Gunasegaran, B., Singh, H. D., Dutertre, C.-A., Loh, C. Y., Lim, J. Q., Crawford, J. C., Lee, H. K., Zhang, X., Lee, B., et al. (2021). Non-terminally exhausted tumor-resident memory hbv-specific t cell responses correlate with relapse-free survival in hepatocellular carcinoma. *Immunity*, 54(8), 1825–1840.

[41] Cho, J., Kim, S. J., Park, W.-Y., Kim, J., Woo, J., Kim, G., Yoon, S. E., Ko, Y. H., & Kim, W. S. (2020). Immune subtyping of extranodal nk/t-cell lymphoma: a new biomarker and an immune shift during disease progression. *Modern Pathology*, 33(4), 603–615.

[42] Choi, I.-K., Wang, Z., Ke, Q., Hong, M., Paul Jr, D. W., Fernandes, S. M., Hu, Z., Stevens, J., Guleria, I., Kim, H.-J., et al. (2021). Mechanism of ebv inducing anti-tumour immunity and its therapeutic use. *Nature*, 590(7844), 157–162.

[43] Clevers, H. (2016). Modeling development and disease with organoids. *Cell*, 165(7), 1586–1597.

[44] Cochet, C., Martel-Renoir, D., Grunewald, V., Bosq, J., Cochet, G., Schwaab, G., Bernaudin, J.-F., & Joab, I. (1993). Expression of the epstein-barr virus immediate early gene, bzlf1, in nasopharyngeal carcinoma tumor cells. *Virology*, 197(1), 358–365.

[45] Cohen, M., Vistarop, A. G., Huaman, F., Narbaitz, M., Metrebian, F., De Matteo, E., Preciado, M. V., & Chabay, P. A. (2018). Epstein-barr virus lytic cycle involvement in diffuse large b cell lymphoma. *Hematological Oncology*, 36(1), 98–103.

[46] Coiffier, B., Lepage, E., Brière, J., Herbrecht, R., Tilly, H., Bouabdallah, R., Morel, P., Van Den Neste, E., Salles, G., Gaulard, P., et al. (2002). Chop chemotherapy plus rituximab compared with chop alone in elderly patients with diffuse large-b-cell lymphoma. *New England Journal of Medicine*, 346(4), 235–242.

[47] Connors, J. M., Cozen, W., Steidl, C., Carbone, A., Hoppe, R. T., Flechtner, H.-H., & Bartlett, N. L. (2020). Hodgkin lymphoma. *Nature Reviews Disease Primers*, 6(1), 61.

[48] Crawford, A., Angelosanto, J. M., Kao, C., Doering, T. A., Odorizzi, P. M., Barnett, B. E., & Wherry, E. J. (2014). Molecular and transcriptional basis of cd4+ t cell dysfunction during chronic infection. *Immunity*, 40(2), 289–302.

[49] da Silva, S. R. & de Oliveira, D. E. (2011). Hiv, ebv and kshv: viral cooperation in the pathogenesis of human malignancies. *Cancer Letters*, 305(2), 175–185.

[50] Danaher, P., Kim, Y., Nelson, B., Griswold, M., Yang, Z., Piazza, E., & Beechem, J. M. (2022). Advances in mixed cell deconvolution enable quantification of cell types in spatial transcriptomic data. *Nature Communications*, 13(1), 385.

[51] de Sousa Abreu, R., Penalva, L. O., Marcotte, E. M., & Vogel, C. (2009). Global signatures of protein and mrna expression levels. *Molecular BioSystems*, 5(12), 1512–1526.

[52] DeCaprio, J. A. (2017). Merkel cell polyomavirus and merkel cell carcinoma. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1732), 20160276.

[53] Deisher, A., **Yeo, Y. Y.**, & Jiang, S. (2022). Combined protein and nucleic acid staining in tissues with panini. *STAR Protocols*, 3(3), 101663.

[54] Delaunay, B. (1934). Sur la sphère vide. a la mémoire de georges voronoï. *Известия Российской академии наук. Серия математическая*, (6), 793–800.

[55] DeVita Jr, V. T., Serpick, A. A., & Carbone, P. P. (1970). Combination chemotherapy in the treatment of advanced hodgkin's disease. *Annals of Internal Medicine*, 73(6), 881–895.

[56] Dhatchinamoorthy, K., Colbert, J. D., & Rock, K. L. (2021). Cancer immune evasion through loss of mhc class i antigen presentation. *Frontiers in Immunology*, 12, 636568.

[57] Ding, D. Y., Tang, Z., Zhu, B., Ren, H., Shalek, A. K., Tibshirani, R., & Nolan, G. P. (2025). Quantitative characterization of tissue states using multiomics and ecological spatial analysis. *Nature Genetics*, (pp. 1–12).

[58] Dong, K. & Zhang, S. (2022). Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder. *Nature Communications*, 13(1), 1739.

[59] Dong, X., Thanou, D., Frossard, P., & Vandergheynst, P. (2016). Learning laplacian matrix in smooth graph signal representations. *IEEE Transactions on Signal Processing*, 64(23), 6160–6173.

[60] Dong, X., Thanou, D., Rabbat, M., & Frossard, P. (2019). Learning graphs from data: A signal representation perspective. *IEEE Signal Processing Magazine*, 36(3), 44–63.

[61] Doorbar, J., Quint, W., Banks, L., Bravo, I. G., Stoler, M., Broker, T. R., & Stanley, M. A. (2012). The biology and life-cycle of human papillomaviruses. *Vaccine*, 30, F55–F70.

[62] Dorothea, M., Xie, J., Yiu, S. P. T., & Chiang, A. K. S. (2023). Contribution of epstein–barr virus lytic proteins to cancer hallmarks and implications from other oncoviruses. *Cancers*, 15(7), 2120.

[63] Efremova, M., Vento-Tormo, M., Teichmann, S. A., & Vento-Tormo, R. (2020). Cellphonedb: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nature Protocols*, 15(4), 1484–1506.

[64] Epstein, M. A., Achong, B. G., & Barr, Y. M. (1964). Virus particles in cultured lymphoblasts from burkitt's lymphoma. *The Lancet*, 283(7335), 702–703.

[65] Fang, R., Halpern, A., Rahman, M. M., Huang, Z., Lei, Z., Hell, S. J., Dulac, C., & Zhuang, X. (2024). Three-dimensional single-cell transcriptome imaging of thick tissues. *Elife*, 12, RP90029.

[66] Farber, D. L., Yudanin, N. A., & Restifo, N. P. (2014). Human memory t cells: generation, compartmentalization and homeostasis. *Nature Reviews Immunology*, 14(1), 24–35.

[67] Fischer, D. S., Theis, F. J., & Yosef, N. (2018). Impulse model-based differential expression analysis of time course sequencing data. *Nucleic Acids Research*, 46(20), e119–e119.

[68] Franco, F., Jaccard, A., Romero, P., Yu, Y.-R., & Ho, P.-C. (2020). Metabolic and epigenetic regulation of t-cell exhaustion. *Nature Metabolism*, 2(10), 1001–1012.

[69] Fredriksson, S., Gullberg, M., Jarvius, J., Olsson, C., Pietras, K., Gústafsdóttir, S. M., Östman, A., & Landegren, U. (2002). Protein detection using proximity-dependent dna ligation assays. *Nature Biotechnology*, 20(5), 473–477.

[70] Frontzek, F., Staiger, A. M., Wullenkord, R., Grau, M., Zapukhlyak, M., Kurz, K. S., Horn, H., Erdmann, T., Fend, F., Richter, J., et al. (2023). Molecular profiling of ebv associated diffuse large b-cell lymphoma. *Leukemia*, 37(3), 670–679.

[71] Gandhi, M. K., Hoang, T., Law, S. C., Brosda, S., O'Rourke, K., Tobin, J. W., Vari, F., Murigneux, V., Fink, L., Gunawardana, J., et al. (2021). Ebv-associated primary cns lymphoma occurring after immunosuppression is a distinct immunobiological entity. *Blood*, 137(11), 1468–1477.

[72] Gehring, A. J., Ho, Z. Z., Tan, A. T., Aung, M. O., Lee, K. H., Tan, K. C., Lim, S. G., & Bertoletti, A. (2009). Profile of tumor antigen-specific cd8 t cells in patients with hepatitis b virus-related hepatocellular carcinoma. *Gastroenterology*, 137(2), 682–690.

[73] Giesen, C., Wang, H. A., Schapiro, D., Zivanovic, N., Jacobs, A., Hattendorf, B., Schüffler, P. J., Grolimund, D., Buhmann, J. M., Brandt, S., et al. (2014). Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nature Methods*, 11(4), 417–422.

[74] Goltsev, Y., Samusik, N., Kennedy-Darling, J., Bhate, S., Hale, M., Vazquez, G., Black, S., & Nolan, G. P. (2018). Deep profiling of mouse splenic architecture with codex multiplexed imaging. *Cell*, 174(4), 968–981.

[75] González-Martín, A., Gómez, L., Lustgarten, J., Mira, E., & Mañes, S. (2011). Maximal t cell–mediated antitumor responses rely upon ccr5 expression in both cd4+ and cd8+ t cells. *Cancer Research*, 71(16), 5455–5466.

[76] Greaves, M. & Maley, C. C. (2012). Clonal evolution in cancer. *Nature*, 481(7381), 306–313.

[77] Greenwald, N. F., Miller, G., Moen, E., Kong, A., Kagel, A., Dougherty, T., Fullaway, C. C., McIntosh, B. J., Leow, K. X., Schwartz, M. S., et al. (2022). Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nature Biotechnology*, 40(4), 555–565.

[78] Hadrup, S. R. & Newell, E. W. (2017). Determining t-cell specificity to understand and treat disease. *Nature Biomedical Engineering*, 1(10), 784–795.

[79] Haist, M., Matusiak, M., Tan, Y., Zimmer, S., Stege, H., Kempchen, T. N., Mitschke, S., Chu, P., Weidenthaler-Barth, B., Barlow, G. L., et al. (2025). Spatially organized inflammatory myeloid-cd8+ t cell aggregates linked to merkel-cell polyomavirus driven reorganization of the tumor microenvironment. *bioRxiv*.

[80] Hanahan, D. (2022). Hallmarks of cancer: new dimensions. *Cancer Discovery*, 12(1), 31–46.

[81] Hanahan, D. & Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*, 100(1), 57–70.

[82] Hanahan, D. & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell*, 144(5), 646–674.

[83] Hansen, T. H. & Bouvier, M. (2009). Mhc class i antigen presentation: learning from viral evasion strategies. *Nature Reviews Immunolology*, 9(7), 503–513.

[84] Hänzelmann, S., Castelo, R., & Guinney, J. (2013). Gsva: gene set variation analysis for microarray and rna-seq data. *BMC Bioinformatics*, 14(1), 7.

[85] Harris, N., Jaffe, E., Diebold, J., Flandrin, G., Muller-Hermelink, H., & Vardiman, J. (2000). Lymphoma classification–from controversy to consensus: the real and who classification of lymphoid neoplasms. *Annals of oncology*, 11, S3–S10.

[86] Hartlage, A. S., Liu, T., Patton, J. T., Garman, S. L., Zhang, X., Kurt, H., Lozanski, G., Lustberg, M. E., Caligiuri, M. A., & Baiocchi, R. A. (2015). The epstein–barr virus lytic protein bzlf1 as a candidate target antigen for vaccine development. *Cancer Immunology Research*, 3(7), 787–794.

[87] He, S., Bhatt, R., Brown, C., Brown, E. A., Buhr, D. L., Chantranuvatana, K., Danaher, P., Dunaway, D., Garrison, R. G., Geiss, G., et al. (2022). High-plex imaging of rna and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. *Nature Biotechnology*, 40(12), 1794–1806.

[88] Heger, J.-M., Mammadova, L., Mattlener, J., Sobesky, S., Cirillo, M., Altmüller, J., Kirst, E., Reinke, S., Klapper, W., Bröckelmann, P. J., et al. (2024). Circulating tumor dna sequencing for biologic classification and individualized risk stratification in patients with hodgkin lymphoma. *Journal of Clinical Oncology*, 42(35), 4218–4230.

[89] Henle, G., Henle, W., & Diehl, V. (1968). Relation of burkitt's tumor-associated herpes-ytpe virus to infectious mononucleosis. *Proceedings of the National Academy of Sciences*, 59(1), 94–101.

[90] Hermanson, G. T. (2013). *Bioconjugate Techniques*. Academic Press.

[91] Hodgkin, T. (1832). On some morbid appearances of the absorbent glands and spleen. *Medical Chirurgical Society Transactions*, 17, 68–114.

[92] Hu, J., Li, X., Coleman, K., Schroeder, A., Ma, N., Irwin, D. J., Lee, E. B., Shinohara, R. T., & Li, M. (2021). Spagcn: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nature Methods*, 18(11), 1342–1351.

[93] Huang, S. & Yasuda, T. (2021). Pathologically relevant mouse models for epstein–barr virus–associated b cell lymphoma. *Frontiers in Immunology*, 12, 639844.

[94] Huang, W., Goldsberry, L., Wymbs, N. F., Grafton, S. T., Bassett, D. S., & Ribeiro, A. (2016). Graph frequency analysis of brain signals. *IEEE Journal of Selected Topics in Signal Processing*, 10(7), 1189–1203.

[95] Hui, J. Z., Tamsen, S., Song, Y., & Tsourkas, A. (2015). Lasic: light activated site-specific conjugation of native iggs. *Bioconjugate Chemistry*, 26(8), 1456–1460.

[96] Hui, J. Z. & Tsourkas, A. (2014). Optimization of photoactive protein z for fast and efficient site-specific conjugation of native igg. *Bioconjugate Chemistry*, 25(9), 1709–1719.

[97] Hui, J. Z., Zaki, A. A., Cheng, Z., Popik, V., Zhang, H., Luning Prak, E. T., & Tsourkas, A. (2014). Facile method for the site-specific, covalent attachment of full-length igg onto nanoparticles. *Small*, 10(16), 3354–3363.

[98] Huizing, G.-J., Samaran, J., Capocefalo, D., Audit, A., Peyré, G., & Cantini, L. (2025). Stories: learning cell fate landscapes from spatial transcriptomics using optimal transport. *Nature Methods*, (pp. 1–10).

[99] Hunt, G. J., Freytag, S., Bahlo, M., & Gagnon-Bartsch, J. A. (2019). Dtangle: accurate and robust cell type deconvolution. *Bioinformatics*, 35(12), 2093–2099.

[100] Janesick, A., Shelansky, R., Gottscho, A. D., Wagner, F., Williams, S. R., Rouault, M., Beliakoff, G., Morrison, C. A., Oliveira, M. F., Sicherman, J. T., et al. (2023). High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. *Nature Communications*, 14(1), 8353.

[101] Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., & Forman, D. (2011). Global cancer statistics. *CA: A Cancer Journal for Clinicians*, 61(2), 69–90.

[102] Jiang, S., Chan, C. N., Rovira-Clavé, X., Chen, H., Bai, Y., Zhu, B., McCaffrey, E., Greenwald, N. F., Liu, C., Barlow, G. L., et al. (2022). Combined protein and nucleic acid imaging reveals virus-dependent b cell and macrophage immunosuppression of tissue microenvironments. *Immunity*, 55(6), 1118–1134.

[103] Kaplan, H. S. (1962). The radical radiotherapy of regionally localized hodgkin's disease. *Radiology*, 78(4), 553–561.

[104] Keane, C., Tobin, J., Gunawardana, J., Francis, S., Gifford, G., Gabrielli, S., Gill, A., Stevenson, W., Talaulikar, D., Gould, C., et al. (2019). The tumour microenvironment is immunotolerogenic and a principal determinant of patient outcome in ebv-positive diffuse large b-cell lymphoma. *European Journal of Haematology*, 103(3), 200–207.

[105] Kempkes, B. & Robertson, E. S. (2015). Epstein-barr virus latency: current and future perspectives. *Current Opinion in Virology*, 14, 138–144.

[106] Keren, L., Bosse, M., Thompson, S., Risom, T., Vijayaragavan, K., McCaffrey, E., Marquez, D., Angoshtari, R., Greenwald, N. F., Fienberg, H., et al. (2019). Mibi-tof: A multiplexed imaging platform relates cellular phenotypes and tissue structure. *Science Advances*, 5(10), eaax5851.

[107] Khafizov, R., Piazza, E., Cui, Y., Patrick, M., Metzger, E., McGuire, D., Dunaway, D., Danaher, P., Hoang, M. L., Grootsky, A., et al. (2024). Sub-cellular imaging of the entire protein-coding human transcriptome (18933-plex) on ffpe tissue using spatial molecular imaging. *bioRxiv*.

[108] Khan, O., Giles, J. R., McDonald, S., Manne, S., Ngiow, S. F., Patel, K. P., Werner, M. T., Huang, A. C., Alexander, K. A., Wu, J. E., & et al. (2019). Tox transcriptionally and epigenetically programs cd8+ t cell exhaustion. *Nature*, 571(7764), 211–218.

[109] Kinkhabwala, A., Herbel, C., Pankratz, J., Yushchenko, D. A., Rüberg, S., Praveen, P., Reiß, S., Rodriguez, F. C., Schäfer, D., Kollet, J., et al. (2022). Macsima imaging cyclic staining (mics) technology reveals combinatorial target pairs for car t cell treatment of solid tumors. *Scientific Reports*, 12(1), 1911.

[110] Klein, D., Palla, G., Lange, M., Klein, M., Piran, Z., Gander, M., Meng-Papaxanthos, L., Sterr, M., Saber, L., Jing, C., et al. (2025). Mapping cells through time and space with moscot. *Nature*, 638(8052), 1065–1075.

[111] Ko, J., Wilkovitsch, M., Oh, J., Kohler, R. H., Bolli, E., Pittet, M. J., Vinegoni, C., Sykes, D. B., Mikula, H., Weissleder, R., et al. (2022). Spatiotemporal multiplexed immunofluorescence imaging of living cells and tissues with bioorthogonal cycling of fluorescent probes. *Nature Biotechnology*, 40(11), 1654–1662.

[112] Kohli, K., Pillarisetty, V. G., & Kim, T. S. (2022). Key chemokines direct migration of immune cells in solid tumors. *Cancer Gene Therapy*, 29(1), 10–21.

[113] Kotliar, D., Veres, A., Nagy, M. A., Tabrizi, S., Hodis, E., Melton, D. A., & Sabeti, P. C. (2019). Identifying gene expression programs of cell-type identity and cellular activity with single-cell rna-seq. *Elife*, 8, e43803.

[114] Law, S. C., Hoang, T., O'Rourke, K., Tobin, J. W., Gunawardana, J., Loo-Oey, D., Bednarska, K., de Long, L. M., Sabdia, M. B., Hapgood, G., et al. (2021). Successful treatment of epstein–barr virus–associated primary central nervous system lymphoma due to post-transplantation lymphoproliferative disorder, with ibrutinib and third-party epstein–barr virus–specific t cells. *American Journal of Transplantation*, 21(10), 3465–3471.

[115] Lee, S. P., Constandinou, C. M., Thomas, W. A., Croom-Carter, D., Blake, N. W., Murray, P. G., Crocker, J., & Rickinson, A. B. (1998). Antigen presenting phenotype of hodgkin reed-sternberg cells: analysis of the hla class i processing pathway and the effects of interleukin-10 on epstein-barr virus-specific cytotoxic t-cell recognition. *Blood*, 92(3), 1020–1030.

[116] Levine, J. H., Simonds, E. F., Bendall, S. C., Davis, K. L., El-ad, D. A., Tadmor, M. D., Litvin, O., Fienberg, H. G., Jager, A., Zunder, E. R., et al. (2015). Data-driven phenotypic dissection of aml reveals progenitor-like cells that correlate with prognosis. *Cell*, 162(1), 184–197.

[117] Lin, J.-R., Chen, Y.-A., Campton, D., Cooper, J., Coy, S., Yapp, C., Tefft, J. B., McCarty, E., Ligon, K. L., Rodig, S. J., et al. (2023). High-plex immunofluorescence imaging and traditional histology of the same tissue section for discovering image-based biomarkers. *Nature Cancer*, 4(7), 1036–1052.

[118] Liu, Y., Sharma, R., Bieniosek, M., Kang, A., Wu, E., Chou, P., Li, I., Rahim, M., Bauer, E., Ji, R., et al. (2025). Modeling patient tissues at molecular resolution with eva. *bioRxiv*, (pp. 2025–12).

[119] Liu, Y., Yang, M., Deng, Y., Su, G., Enninful, A., Guo, C. C., Tebaldi, T., Zhang, D., Kim, D., Bai, Z., et al. (2020). High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. *Cell*, 183(6), 1665–1681.

[120] Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, volume 2 (pp. 1150–1157).: IEEE.

[121] Lu, T.-X., Liang, J.-H., Miao, Y., Fan, L., Wang, L., Qu, X.-Y., Cao, L., Gong, Q.-X., Wang, Z., Zhang, Z.-H., et al. (2015). Epstein-barr virus positive diffuse large b-cell lymphoma predict poor outcome, regardless of the age. *Scientific Reports*, 5(1), 12168.

[122] Lundberg, M., Thorsen, S. B., Assarsson, E., Villablanca, A., Tran, B., Gee, N., Knowles, M., Nielsen, B. S., Couto, E. G., Martin, R., et al. (2011). Multiplexed homogeneous proximity ligation assays for high-throughput protein biomarker research in serological material. *Molecular & Cellular Proteomics*, 10(4), M110.004978.

[123] Ma, S.-D., Hegde, S., Young, K. H., Sullivan, R., Rajesh, D., Zhou, Y., Jankowska-Gan, E., Burlingham, W. J., Sun, X., Gulley, M. L., et al. (2011). A new model of epstein-barr virus infection reveals an important role for early lytic viral protein expression in the development of lymphomas. *Journal of Virology*, 85(1), 165–177.

[124] Ma, S.-D., Yu, X., Mertz, J. E., Gumperz, J. E., Reinheim, E., Zhou, Y., Tang, W., Burlingham, W. J., Gulley, M. L., & Kenney, S. C. (2012). An epstein-barr virus (ebv) mutant with enhanced bzlf1 expression causes lymphomas with abortive lytic ebv infection in a humanized mouse model. *Journal of Virology*, 86(15), 7976–7987.

[125] Markovits, E., Klymyshyn, D., Froumine, R., Zong, H., Mints, M., Mahadevan, S., Bloom, K., Bates, J., Thomas, G. J., Diehl, L., et al. (2025). Spatial profiling of hpv-stratified head and neck squamous cell carcinoma reveals distinct immune niches and microenvironmental architectures. *Journal of Translational Medicine*, 23(1), 1304.

[126] Marks, M., Israel, U., Dilip, R., Li, Q., Yu, C., Laubscher, E., Iqbal, A., Pradhan, E., Ates, A., Abt, M., et al. (2025). Cellsam: a foundation model for cell segmentation. *Nature Methods*, (pp. 1–9).

[127] Massoni-Badosa, R., Aguilar-Fernández, S., Nieto, J. C., Soler-Vila, P., Elosua-Bayes, M., Marchese, D., Kulis, M., Vilas-Zornoza, A., Bühler, M. M., Rashmi, S., et al. (2024). An atlas of cells in the human tonsil. *Immunity*, 57(2), 379–399.

[128] McKelvey, E. M., Gottlieb, J. A., Wilson, H. E., Haut, A., Talley, R. W., Stephens, R., Lane, M., Gamble, J. F., Jones, S. E., Grozea, P. N., et al. (1976). Hydroxyldaunomycin (adriamycin) combination chemotherapy in malignant lymphoma. *Cancer*, 38(4), 1484–1493.

[129] Merritt, C. R., Ong, G. T., Church, S. E., Barker, K., Danaher, P., Geiss, G., Hoang, M., Jung, J., Liang, Y., McKay-Fleisch, J., et al. (2020). Multiplex digital spatial profiling of proteins and rna in fixed tissue. *Nature Biotechnology*, 38(5), 586–599.

[130] Mesri, E. A., Feitelson, M. A., & Munger, K. (2014). Human viral oncogenesis: a cancer hallmarks analysis. *Cell Host & Microbe*, 15(3), 266–282.

[131] Methods, N. (2021). Method of the year: spatially resolved transcriptomics. *Nature Methods*, 18(1), 9–14.

[132] Methods, N. (2024). Method of the year 2024: spatial proteomics. *Nature Methods*, 21(12), 2195–2196.

[133] Michel, H. A., McCallum, P., Wu, W., Le Lee, J., Luo, S., Chan, C. N., Schaffenrath, J., **Yeo, Y. Y.**, Yiu, S. P. T., Wang, Y., et al. (2025). Pasta: Versatile tyramine-oligonucleotide amplification for multi-modal spatial biology. *bioRxiv*.

[134] Miller, B., Sen, D., Abosy, R., Bi, K., Virkud, Y., LaFleur, M., Yates, K., Lako, A., Felt, K., Naik, G., & et al. (2019). Subsets of exhausted cd8+ t cells differentially mediate tumor control and respond to checkpoint blockade. *Nature Immunology*, 20(3), 326–336.

[135] Muenz, C. (2025). Recent advances in animal models of lymphomagenesis caused by human γ-herpesviruses. *Current Opinion in Virology*, 71, 101461.

[136] Mund, A., Coscia, F., Kriston, A., Hollandi, R., Kovács, F., Brunner, A.-D., Migh, E., Schweizer, L., Santos, A., Bzorek, M., et al. (2022). Deep visual proteomics defines single-cell identity and heterogeneity. *Nature Biotechnology*, 40(8), 1231–1240.

[137] Münz, C. (2019). Latency and lytic replication in epstein–barr virus-associated oncogenesis. *Nature Reviews Microbiology*, 17(11), 691–700.

[138] Murray, P., Constandinou, C., Crocker, J., Young, L., & Ambinder, R. (1998). Analysis of major histocompatibility complex class i, tap expression, and lmp2 epitope sequence in epstein-barr virus-positive hodgkin's disease. *Blood*, 92(7), 2477–2483.

[139] Nam, Y., Gomez, K., Reynier, J.-B., Khamnei, C., Aitken, M., Zheng, V., Lhakhang, T., Casula, M., Palmieri, G., Cossu, A., et al. (2025). Genomic landscape of virus-associated cancers. *Nature Communications*, 16(1), 5887.

[140] Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., Hoang, C. D., Diehn, M., & Alizadeh, A. A. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods*, 12(5), 453–457.

[141] Niedobitek, G., Agathanggelou, A., Rowe, M., Jones, E., Jones, D., Turyaguma, P., Oryema, J., Wright, D., & Young, L. (1995). Heterogeneous expression of epstein-barr virus latent proteins in endemic burkitt's lymphoma. *Blood*, 86(2), 659–665.

[142] Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science*, 194(4260), 23–28.

[143] Nowicki, T. S., Hu-Lieskovan, S., & Ribas, A. (2018). Mechanisms of resistance to pd-1 and pd-l1 blockade. *The Cancer Journal*, 24(1), 47–53.

[144] Oliveira, G. & Wu, C. J. (2023). Dynamics and specificities of t cells in cancer immunotherapy. *Nature Reviews Cancer*, 23(5), 295–316.

[145] Oliveira, M. F. d., Romero, J. P., Chung, M., Williams, S. R., Gottscho, A. D., Gupta, A., Pilipauskas, S. E., Mohabbat, S., Raman, N., Sukovich, D. J., et al. (2025). High-definition spatial transcriptomic profiling of immune cell populations in colorectal cancer. *Nature Genetics*, 57(6), 78–82.

[146] Oudejans, J., Jiwa, N., Kummer, J., Horstman, A., Vos, W., Baak, J., Kluin, P., Valk, P. v. d., Walboomers, J., & Meijer, C. (1996). Analysis of major histocompatibility complex class i expression on reed-sternberg cells in relation to the cytotoxic t-cell response in epstein-barr virus-positive and -negative hodgkin's disease. *Blood*, 87(9), 3844–3851.

[147] Ouyang, Q., Wagner, W. M., Walter, S., Müller, C. A., Wikby, A., Aubert, G., Klatt, T., Stevanovic, S., Dodi, T., & Pawelec, G. (2003). An age-related increase in the number of cd8+ t cells carrying receptors for an immunodominant epstein–barr virus (ebv) epitope is counteracted by a decreased frequency of their antigen-specific responsiveness. *Mechanisms of Ageing and Development*, 124(4), 477–485.

[148] Pallesen, G., Sandvej, K., Hamilton-Dutoit, S. J., Rowe, M., & Young, L. (1991). Activation of epstein-barr virus replication in hodgkin and reed-sternberg cells. *Blood*, 78(5), 1162–1165.

[149] Papalegis, D., Tkachev, S., Vu, L., & Klein, S. (2023). Signalstar™ is a novel multiplex ihc technology that demonstrates flexibility and reproducibility. *Journal for ImmunoTherapy of Cancer*.

[150] Pattengale, P., Smith, R., & Gerber, P. (1973). Selective transformation of b lymphocytes by eb virus. *The Lancet*, 302(7820), 93–94.

[151] Peters, M. V. (1950). A study of survivals in hodgkin's disease treated radiologically. *American Journal of Roentgenology and Radium Therapy*, 63.

[152] Phillips, D., Schürch, C. M., Khodadoust, M. S., Kim, Y. H., Nolan, G. P., & Jiang, S. (2021). Highly multiplexed phenotyping of immunoregulatory proteins in the tumor microenvironment by codex tissue imaging. *Frontiers in Immunology*, 12, 687673.

[153] Pilato, M., Kfuri-Rubens, R., Pruessmann, J., Ozga, A., Messemaker, M., Cadilha, B., Sivakumar, R., Cianciaruso, C., Warner, R., Marangoni, F., & et al. (2021). Cxcr6 positions cytotoxic t cells to receive critical survival signals in the tumor microenvironment. *Cell*, 184(17), 4512–4530.e22.

[154] Plummer, M., de Martel, C., Vignat, J., Ferlay, J., Bray, F., & Franceschi, S. (2016). Global burden of cancers attributable to infections in 2012: a synthetic analysis. *The Lancet Global Health*, 4(9), e609–e616.

[155] Polania, J. W., Hoyt-Miggelbrink, A., Tomaszewski, W. H., Wachsmuth, L. P., Lorrey, S. J., Wilkinson, D. S., Lerner, E., Woroniecka, K., Finlay, J. B., Ayasoufi, K., et al. (2025). Antigen presentation by tumor-associated macrophages drives t cells from a progenitor exhaustion state to terminal exhaustion. *Immunity*, 58(1), 232–246.

[156] Pusey, W. A. (1902). Cases of sarcoma and of hodgkin's disease treated by exposures to x-rays-a preliminary report. *Journal of the American Medical Association*, 38(3), 166–169.

[157] Qin, W., Cho, K. F., Cavanagh, P. E., & Ting, A. Y. (2021). Deciphering molecular interactions by proximity labeling. *Nature Methods*, 18(2), 133–143.

[158] Radtke, A. J., Kandov, E., Lowekamp, B., Speranza, E., Chu, C. J., Gola, A., Thakur, N., Shih, R., Yao, L., Yaniv, Z. R., et al. (2020). Ibex: A versatile multiplex optical imaging approach for deep phenotyping and spatial analysis of cells in complex tissues. *Proceedings of the National Academy of Sciences*, 117(52), 33455–33465.

[159] Rappaport, H. (1966). *Tumors of the hematopoietic system*. Armed Forces Institute of Pathology.

[160] Reed, D. M. (1902). On the pathological changes in hodgkin's disease, with especial reference to its relation to tuberculosis. *Johns Hopkins Hospital Report*, 10, 133–196.

[161] Rivest, F., Eroglu, D., Pelz, B., Kowal, J., Kehren, A., Navikas, V., Procopio, M. G., Bordignon, P., Pérès, E., Ammann, M., et al. (2023). Fully automated sequential immunofluorescence (seqif) for hyperplex spatial proteomics. *Scientific Reports*, 13(1), 16994.

[162] Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edger: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1), 139–140.

[163] Rodriques, S. G., Stickels, R. R., Goeva, A., Martin, C. A., Murray, E., Vanderburg, C. R., Welch, J., Chen, L. M., Chen, F., & Macosko, E. Z. (2019). Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science*, 363(6434), 1463–1467.

[164] Roemer, M., Advani, R., Redd, R., Pinkus, G., Natkunam, Y., Ligon, A., Connelly, C., Pak, C., Carey, C., Daadi, S., & et al. (2016a). Classical hodgkin lymphoma with reduced β2m/mhc class i expression is associated with inferior outcome independent of 9p24.1 status. *Cancer Immunology Research*, 4(11), 910–916.

[165] Roemer, M., Redd, R., Cader, F., Pak, C., Abdelrahman, S., Ouyang, J., Sasse, S., Younes, A., Fanale, M., Santoro, A., & et al. (2018). Major histocompatibility complex class ii and programmed death ligand 1 expression predict outcome after programmed death 1 blockade in classic hodgkin lymphoma. *Journal of Clinical Oncology*, 36(10), 942–950.

[166] Roemer, M. G., Advani, R. H., Ligon, A. H., Natkunam, Y., Redd, R. A., Homer, H., Connelly, C. F., Sun, H. H., Daadi, S. E., Freeman, G. J., et al. (2016b). Pd-l1 and pd-l2 genetic alterations define classical hodgkin lymphoma and predict outcome. *Journal of Clinical Oncology*, 34(23), 2690–2697.

[167] Röntgen, W. C. (1896). On a new kind of rays. *Science*, 3(59), 227–231.

[168] Ross, A. M., Leahy, C. I., Neylon, F., Steigerova, J., Flodr, P., Navratilova, M., Urbankova, H., Vrzalikova, K., Mundo, L., Lazzi, S., et al. (2023). Epstein–barr virus and the pathogenesis of diffuse large b-cell lymphoma. *Life*, 13(2), 521.

[169] Rous, P. (1911). A sarcoma of the fowl transmissible by an agent separable from the tumor cells. *Journal of Experimental Medicine*, 13(4), 397.

[170] Saunders, R. A., Allen, W. E., Pan, X., Sandhu, J., Lu, J., Lau, T. K., Smolyar, K., Sullivan, Z. A., Dulac, C., Weissman, J. S., et al. (2025). Perturb-multimodal: A platform for pooled genetic screens with imaging and sequencing in intact mammalian tissue. *Cell*, 188(17), 4790–4809.

[171] Scheffner, M., Werness, B. A., Huibregtse, J. M., Levine, A. J., & Howley, P. M. (1990). The e6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell*, 63(6), 1129–1136.

[172] Schwickert, T. A., Victora, G. D., Fooksman, D. R., Kamphorst, A. O., Mugnier, M. R., Gitlin, A. D., Dustin, M. L., & Nussenzweig, M. C. (2011). A dynamic t cell–limited checkpoint regulates affinity-dependent b cell entry into the germinal center. *Journal of Experimental Medicine*, 208(6), 1243–1252.

[173] Shaban, M., Bai, Y., Qiu, H., Mao, S., Yeung, J., **Yeo, Y. Y.**, Shanmugam, V., Chen, H., Zhu, B., Weirather, J. L., et al. (2024). Maps: pathologist-level cell type annotation from tissue images through machine learning. *Nature Communications*, 15(1), 28.

[174] Shaban, M., Chang, Y., Qiu, H., **Yeo, Y. Y.**, Song, A. H., Jaume, G., Wang, Y., Weishaupt, L. L., Ding, T., Vaidya, A., et al. (2025). A foundation model for spatial proteomics. *arXiv*.

[175] Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379–423.

[176] Shannon-Lowe, C. & Rickinson, A. (2019). The global landscape of ebv-associated tumors. *Frontiers in Oncology*, 9, 713.

[177] Sharpe, A. & Pauken, K. (2018). The diverse functions of the pd1 inhibitory pathway. *Nature Reviews Immunology*, 18(3), 153–167.

[178] Shen, X., Zuo, L., Ye, Z., Yuan, Z., Huang, K., Li, Z., Yu, Q., Zou, X., Wei, X., Xu, P., et al. (2025). Inferring cell trajectories of spatial transcriptomics via optimal transport analysis. *Cell Systems*, 16(2), 101194.

[179] Simonson, P. D., Valencia, I., & Patel, S. S. (2022). Tyramide-conjugated dna barcodes enable signal amplification for multiparametric codex imaging. *Communications Biology*, 5(1), 627.

[180] Sletten, E. M. & Bertozzi, C. R. (2009). Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angewandte Chemie International Edition*, 48(38), 6974–6998.

[181] Spina, M., Marcus, R., Kassam, S., & Tirelli, U. (2013). Lymphoma in the immunosuppressed. In *Lymphoma: Pathology, Diagnosis, and Treatment* (pp. 273–285). Cambridge University Press.

[182] Stack, E. C., Wang, C., Roman, K. A., & Hoyt, C. C. (2014). Multiplexed immunohistochemistry, imaging, and quantitation: a review, with an assessment of tyramide signal amplification, multispectral imaging and multiplex analysis. *Methods*, 70(1), 46–58.

[183] Ståhl, P. L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J. O., Huss, M., et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, 353(6294), 78–82.

[184] Sternberg, C. (1898). Über eine eigenartige unter dem bilde der pseudoleukamie verlaufende tuberculose des lymphatischen apparates. *Ztschr Heilk*, 19, 21–90.

[185] Stevanović, S., Pasetto, A., Helman, S. R., Gartner, J. J., Prickett, T. D., Howie, B., Robins, H. S., Robbins, P. F., Klebanoff, C. A., Rosenberg, S. A., et al. (2017). Landscape of immunogenic tumor antigens in successful immunotherapy of virally induced epithelial cancer. *Science*, 356(6334), 200–205.

[186] Stickels, R. R., Murray, E., Kumar, P., Li, J., Marshall, J. L., Di Bella, D. J., Arlotta, P., Macosko, E. Z., & Chen, F. (2021). Highly sensitive spatial transcriptomics at near-cellular resolution with slide-seqv2. *Nature Biotechnology*, 39(3), 313–319.

[187] Sun, C., Mezzadra, R., & Schumacher, T. N. (2018). Regulation and function of the pd-l1 checkpoint. *Immunity*, 48(3), 434–452.

[188] Tang, X., Yu, Z., Chen, J., Cui, Y., Shao, D., Wang, W., Wu, F., Zhuang, Y., Shi, W., Huang, Z., et al. (2025). Cellforge: Agentic design of virtual cell models. *arXiv*.

[189] **Yeo, Y. Y.**, Chang, Y., Qiu, H., Yiu, S. P. T., Michel, H. A., Wu, W., Jin, X., Kure, S., Parmelee, L., Luo, S., et al. (2024a). Same-slide spatial multi-omics integration reveals tumor virus-linked spatial reorganization of the tumor microenvironment. *bioRxiv*.

[190] **Yeo, Y. Y.**, Cramer, P., Deisher, A., Bai, Y., Zhu, B., Yeo, W.-J., Shipp, M. A., Rodig, S. J., & Jiang, S. (2024b). A hitchhiker's guide to high-dimensional tissue imaging with multiplexed ion beam imaging. In *Methods in Cell Biology*, volume 186 (pp. 213–231). Elsevier.

[191] **Yeo, Y. Y.**, Qiu, H., Bai, Y., Zhu, B., Chang, Y., Yeung, J., Michel, H. A., Wright, K., Shaban, M., Sadigh, S., et al. (2024c). Epstein-barr virus orchestrates spatial reorganization and immunomodulation within the classic hodgkin lymphoma tumor microenvironment. *bioRxiv*.

[192] Toh, J., Reitsma, A. J., Tajima, T., Younes, S. F., Ezeiruaku, C., Jenkins, K. C., Peña, J. K., Zhao, S., Wang, X., Lee, E. Y., et al. (2024). Multi-modal analysis reveals tumor and immune features distinguishing ebv-positive and ebv-negative post-transplant lymphoproliferative disorders. *Cell Reports Medicine*, 5(12).

[193] Tomarchio, V., Rigacci, L., & Annibali, O. (2025). Cd20 xcd3 bispecific antibodies for large b cell lymphoma: latest updates from 2024 ash annual meeting. *Journal of Hematology & Oncology*, 18(1), 46.

[194] Toner, K., McCann, C. D., & Bollard, C. M. (2024). Applications of cell therapy in the treatment of virus-associated cancers. *Nature Reviews Clinical Oncology*, 21(10), 709–724.

[195] Tosi, A., Parisatto, B., Menegaldo, A., Spinato, G., Guido, M., Del Mistro, A., Bussani, R., Zanconati, F., Tofanelli, M., Tirelli, G., et al. (2022). The immune microenvironment of hpv-positive and hpv-negative oropharyngeal squamous cell carcinoma: a multiparametric quantitative and spatial analysis unveils a rationale to target treatment-naïve tumors with immune checkpoint inhibitors. *Journal of Experimental & Clinical Cancer Research*, 41(1), 279.

[196] Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Tissue-based map of the human proteome. *Science*, 347(6220), 1260419.

[197] Voabil, P., de Bruijn, M., Roelofsen, L. M., Hendriks, S. H., Brokamp, S., van den Braber, M., Broeks, A., Sanders, J., Herzig, P., Zippelius, A., et al. (2021). An ex vivo tumor fragment platform to dissect response to pd-1 blockade in cancer. *Nature Medicine*, 27(7), 1250–1261.

[198] Wang, L. W., Jiang, S., & Gewurz, B. E. (2017). Epstein-barr virus lmp1-mediated oncogenicity. *Journal of Virology*, 91(21), 10–1128.

[199] Wang, X., Park, J., Susztak, K., Zhang, N. R., & Li, M. (2019). Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. *Nature Communications*, 10(1), 380.

[200] Westcott, P. M., Muyas, F., Hauck, H., Smith, O. C., Sacks, N. J., Ely, Z. A., Jaeger, A. M., Rideout III, W. M., Zhang, D., Bhutkar, A., et al. (2023). Mismatch repair deficiency is not sufficient to elicit tumor immunogenicity. *Nature Genetics*, 55(10), 1686–1695.

[201] Wherry, E. J. (2011). T cell exhaustion. *Nature Immunology*, 12(6), 492–499.

[202] Wherry, E. J. & Kurachi, M. (2015). Molecular and cellular insights into t cell exhaustion. *Nature Reviews Immunology*, 15(8), 486–499.

[203] Wilks, S. (1856). Enlargement of the lymphatic glands and spleen (or, hodgkin's disease) with remarks. *Guy's Hospital Reports*, 11, 56–67.

[204] Wright, K. T., Weirather, J. L., Jiang, S., Kao, K. Z., Sigal, Y., Giobbie-Hurder, A., Shipp, M. A., & Rodig, S. J. (2023). Diffuse large b-cell lymphomas have spatially defined, tumor immune microenvironments revealed by high-parameter imaging. *Blood Advances*, 7(16), 4633–4646.

[205] Wu, R., Sattarzadeh, A., Rutgers, B., Diepstra, A., Berg, A. v. d., & Visser, L. (2016). The microenvironment of classical hodgkin lymphoma: heterogeneity by epstein–barr virus presence and location within the tumor. *Blood Cancer Journal*, 6(5), e417.

[206] Xu-Monette, Z. Y., Zhou, J., & Young, K. H. (2018). Pd-1 expression and clinical pd-1 blockade in b-cell lymphomas. *Blood*, 131(1), 68–83.

[207] Yap, L. F., Wong, A. K. C., Paterson, I. C., & Young, L. S. (2022). Functional implications of epstein-barr virus lytic genes in carcinogenesis. *Cancers*, 14(23), 5780.

[208] Yuan, Y. (2016). Spatial heterogeneity in the tumor microenvironment. *Cold Spring Harbor Perspectives in Medicine*, 6(8), a026583.

[209] Zhang, J., Huang, D., Saw, P. E., & Song, E. (2022). Turning cold tumors hot: from molecular mechanisms to clinical applications. *Trends in Immunology*, 43(7), 523–545.

[210] Zhang, X., Chen, Y., Liang, J., Yang, Y., Chen, H., Chen, Z., Li, M., Chen, S., Chen, T., He, H., et al. (2025). Out-of-africa migration and clonal expansion of a recombinant epstein-barr

virus drives frequent nasopharyngeal carcinoma in southern china. *National Science Review*, 12(4), nwae438.

[211] Zhao, E., Stone, M. R., Ren, X., Guenthoer, J., Smythe, K. S., Pulliam, T., Williams, S. R., Uytingco, C. R., Taylor, S. E., Nghiem, P., et al. (2021). Spatial transcriptomics at subspot resolution with bayesspace. *Nature Biotechnology*, 39(11), 1375–1384.

[212] Zhu, B., Gao, S., Chen, S., Wang, Y., Yeung, J., Bai, Y., Huang, A. Y., **Yeo, Y. Y.**, Liao, G., Mao, S., et al. (2025). Celllens enables cross-domain information fusion for enhanced cell population delineation in single-cell spatial omics data. *Nature Immunology*, 26(6), 963–974.

This dissertation was typeset using LaTeX, originally developed by Leslie Lamport and based on Donald Knuth's TeX. The body text is set in 11 point Egenolff-Berner Garamond, a revival of Claude Garamont's humanist typeface. A template that can be used to format a PhD dissertation with this look and feel has been released under the permissive MIT (X11) license, and can be found online at github.com/suchow/Dissertate or from its author, Jordan Suchow, at suchow@post.harvard.edu.