

© Copyright 2022

Jennifer Eng

Quantifying the Single-Cell Spatial Composition of the Breast Tumor Microenvironment

Jennifer Eng

A dissertation

submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

Oregon Health and Science University

2022

Reading Committee:

Sandra Rugonyi, PhD, Chair

Summer Gibbs, PhD
Sara Courtneidge, PhD
Koei Chin, MD, PhD
Young Hwan Chang, PhD
Pepper Schedin, PhD

Program Authorized to Offer Degree:

Biomedical Engineering

Oregon Health and Science University

Abstract

Quantifying the Single-Cell Spatial Composition of the Breast Tumor Microenvironment

Jennifer Eng

Chair of the Supervisory Committee:
Sandra Rugonyi, PhD
Biomedical Engineering

The breast tumor microenvironment is a complex ecosystem of epithelial, immune, endothelial, fibroblast and adipose cells. Cellular phenotypes and spatial localizations within the microenvironment are linked to patient prognosis and response to therapy. We established a protocol for multiplex imaging, cyclic immunofluorescence (CyCIF), and an image processing pipeline, mplexable, for quantitative single-cell spatial analysis of breast tumors and other tissues. We developed improved methods for signal removal, antibody specificity, background correction and batch normalization in CyCIF. We used our experimental and analytical methods for characterizing tumor heterogeneity and spatial relationships in diverse tissue samples including patient biopsies, surgical resections, tissue microarrays, cell lines and mouse models. We then generated CyCIF data and processed publicly available imaging mass cytometry and multiplex ion beam imaging data for biomarker discovery in large clinical cohorts of breast cancer patients. Our work reveals prognostic cell types and spatial arrangements in breast cancer as well as conserved tumor-stromal associations across breast cancer subtypes. Our CyCIF protocol, image analysis tools and atlases of breast cancer tissues all contribute to deepening our understanding of the tumor microenvironment in order to improve patient stratification and inform therapeutic strategies.

TABLE OF CONTENTS

CONTENTS

List of Figures	vii
List of Tables	ix
Chapter 1. Breast cancer biomarker discovery using multiplex imaging	13
1.1 Breast cancer incidence and clinical management.....	13
1.1.1 Luminal ER+ Standard of Care	13
1.1.2 TNBC and HER2+ Standard of Care.....	14
1.2 Targeted therapies in breast cancer.....	14
1.3 Prognostic and predictive biomarkers in breast cancer.....	16
1.4 Biomarkers in the breast tumor microenvironment	16
1.5 TILs interact with tumor characteristics and treatment in ER+ disease	17
1.6 TILs prognostic and predictive in HER2+ disease	19
1.7 TILs delineate good prognosis TNBC	19
1.8 Immune Subsets as Biomarkers	20
1.9 Significance of Fibroblasts and ECM Components of the Tumor Microenvironment.	21
1.10 Biomarker Discovery with Multiplex Imaging.....	23
1.11 Thesis Outline	23
Chapter 2. Cyclic Multiplexed-Immunofluorescence (cmIF), a Highly Multiplexed Method for Single-Cell Analysis	25
2.1 Abstract	27

2.2	Introduction.....	27
2.2.1	Introduction.....	27
2.2.2	Background.....	29
2.2.3	Biomarkers.....	33
2.3	Materials	40
2.3.1	Tissue Processing.....	40
2.3.2	Blocking and Staining.....	40
2.3.3	Antibody Labeling (Optional).....	40
2.3.4	Image Acquisition.....	40
2.3.5	Image Processing and Analysis	40
2.4	Methods.....	41
2.4.1	Overview	41
2.4.2	Staining	41
2.4.3	Image Processing	45
2.4.4	Analysis.....	48
2.4.5	Protocols	52
2.5	Notes	55
2.6	Conclusions.....	56
2.7	Acknowledgments.....	56
Chapter 3. A framework for multiplex imaging optimization and reproducible analysis		57
3.1	Abstract	58
3.2	Introduction.....	58
3.3	Results and Discussion	59

3.3.1	Comparison to Standard IF	60
3.3.2	Tissue Loss Characterization	61
3.3.3	Quenching Condition Assessment	62
3.3.4	Antibody Order Evaluation and Improvement	63
3.3.5	Autofluorescence Dynamics and Corrections.....	65
3.3.6	Reproducibility and Batch Normalization	66
3.4	Methods.....	69
3.4.1	Tissue and Cell Button Preparation	69
3.4.2	Cyclic Immunostaining.....	70
3.4.3	Pre-quenching to reduce autofluorescence	70
3.4.4	Primary Antibody Staining	71
3.4.5	Fluorescence Microscopy	71
3.4.6	Quenching of Fluorescence Signal for Cyclic Immunostaining.	71
3.4.7	Image Registration, Autofluorescence Removal and Segmentation.....	71
3.4.8	Data Analysis	72
3.4.9	Statistics and Reproducibility	74
3.4.10	Data Availability.....	74
3.4.11	Code Availability.....	75
3.4.12	Competing Interests	75
3.4.13	Acknowledgments.....	75
3.4.14	Author Contributions	76
Chapter 4. Use of cyclic immunofluorescence protocol, mplexable pipeline and novel analytical methods for characterizing tumor heterogeneity		77

4.1	Abstract	78
4.2	Overview	78
4.3	Example ONE: Cellular phenotypic heterogeneity and response to neoadjuvant HER2-targeting therapies	79
4.3.1	Introduction	79
4.3.2	Methods	79
4.3.3	Results	80
4.3.4	Discussion	83
4.4	Example TWO: Tumor evolution and spatial biomarker expression patterns in metastatic ER+ breast cancer	84
4.4.1	Introduction	84
4.4.2	Results	86
4.4.3	Discussion	89
4.5	Example THREE: A mouse model of TNBC recapitulates human disease subtypes and survival. Multiplex Imaging defines Tumor Microenvironment Subtypes in TNBC mouse model and Human TNBC.	89
4.5.1	Introduction	89
4.5.2	Methods	90
4.5.3	Results	91
4.5.4	Discussion	98
4.6	Conclusion	99
Chapter 5. Multiplatform analysis of Breast cancer Clinical cohorts Using mplexable		101
5.1	Abstract	102

5.2	Introduction.....	102
5.3	Methods.....	103
5.3.1	Patient samples.....	103
5.3.2	Imaging data generation and sources	103
5.3.3	Image Processing	103
5.3.4	Image QC	104
5.3.5	Single cell phenotyping.....	105
5.3.6	Statistical Analysis.....	106
5.4	Results.....	107
5.4.1	Imaging and Single-Cell Phenotyping.....	107
5.4.2	Cell type fractions are similar across platforms.....	113
5.4.3	Single cell phenotyping identifies prognostic tumor subtypes	113
5.4.4	Microenvironment subtypes associated with clinical subtype.....	114
5.4.5	T cells are an independent prognostic factor in TNBC.....	114
5.4.6	Tumor phenotypes correlate with stromal cell abundance and proximity	118
5.4.7	Lymphocyte-proliferating tumor proximity predicts survival in TNBC	119
5.5	Discussion	120
Chapter 6. Summary and Future Perspectives		125
6.1	Chapter 2: Getting Started with CyCIF and Identifying Challenges	125
6.2	Chapter 3: Making CyCIF Work Well.....	126
6.3	Chapter 4: Quantifying and Interpreting Tumor Heterogeneity	127
6.4	Chapter 5: The Path to Robust Biomarker Discovery	128
6.5	Future Perspectives	129

References	131
Appendix A: Supplementary Information Related to Chapter 5	147
Appendix B: Supplementary Information Related to Chapter 3.....	157

LIST OF FIGURES

Figure 2.1. The workflow of cyclic multiplexed-immunofluorescence (cmIF).	30
Figure 2.2. Antibody Validation in cmIF.....	43
Figure 2.3. Reproducibility in cmIF.	45
Figure 2.4. Visualization in cmIF.	47
Figure 2.5. Unsupervised clustering analysis in cmIF.....	49
Figure 2.6. Cell proximity analysis in cmIF.	51
Figure 3.1. Reproducible Generation, Processing and Analysis of Multiplex Images.	60
Figure 3.2. CycIF Signal to Background and Tissue Loss.	61
Figure 3.3. CyCIF Optimization: Quenching and Panel Order.	63
Figure 3.4. Autofluorescence Characterization in the CyCIF Protocol.	65
Figure 3.5. Reproducibility and Normalization of CyCIF Staining Intensity.....	67
Figure 4.1. Phenotypic single-cell diversity measured by CyCIF.	82
Figure 4.2. Monitoring tumor and stromal responses to therapy using CyCIF.	87
Figure 4.3. Additional CyCIF Analyses, Related to Figure 4.2.....	88
Figure 4.4. Tumor Microenvironment Subtypes Defined by Multiplex Imaging in Mouse and Human TNBC.	93
Figure 4.5. Unsupervised clustering and gating-based cell types.	94
Figure 4.6. Cell frequency-based subtyping in mouse TMA tissues.	95
Figure 4.7. Cell types determined by unsupervised clustering with the Leiden algorithm and human patient subtyping.	96
Figure 4.8. Comparison of Marker Expression Across Subtypes.	97
Figure 5.1. Breast Cancer Multiplex Imaging Datasets.	108
Figure 5.2. Unsupervised clustering defines cell types in CyCIF data.	109
Figure 5.3. Unsupervised clustering defines cell types in IMC data.	110
Figure 5.4. Unsupervised clustering defines cell types in MIBI data.	111
Figure 5.5. Cross-platform Cell Typing.....	112

Figure 5.6. Prognostic breast cancer subtypes defined by multiplex imaging.....	116
Figure 5.7. Immune infiltrate has prognostic value in TNBC breast cancer.	117
Figure 5.8. Tumor phenotypes correlate with stromal cell types and proximity.	123
Figure 5.9. Immune-proliferating tumor proximity predicts survival in TNBC.....	124
Figure 6.1. De-noising and Segmentation Optimization.	148
Figure 6.2. Mesmer versus Watershed Segmentation.....	149
Figure 6.3. ER Quality Control.....	150
Figure 6.4. Subtype and Single Variable Prognosis in the Combined Cohort.....	151
Figure 6.5. Stromal Subtypes.....	152
Figure 6.6. Determining Cell Abundance Cut-offs.....	153
Figure 6.7. Cell Type Correlation in Breast Cancer Tissues.	154
Figure 6.8. Epithelial Subtype versus Stromal Cell Abundance.....	155
Figure 6.9. Prognostic Value of Proximity in Separate Platforms.....	156

LIST OF TABLES

Table 2.1. Biomarkers	34
Table 2.2. Validated Antibodies	35
Table 3.3. Reproducible Image Analytics.....	74
Table 4.4. Cyclic immunofluorescence (CyCIF) gating strategy	80
Table 4.5. Cyclic immunofluorescence antibodies	84
Table 4.6. Code repositories	99

ACKNOWLEDGEMENTS

I would like to thank my mentors, Drs. Summer Gibbs and Koei Chin, for all of their positive support over the years. I would like to thank the members of my committee, Drs. Joe Gray, Sara Courtneidge, Young Hwan Chang and recently joining our party, Dr. Sandra Rugonyi. Thank you to Dr. Pepper Schedin for serving on my exam committee.

Thank you to all current and former members of the Chin, Gibbs, Chang and Gray labs for support, feedback and great memories. I would especially like to acknowledge the members of the Chin lab who contributed to this work: Ting Zheng for establishing antibody labelling protocols, Zhi Hu for staining and scanning many of the data sets, Dave Kilburn for image processing and Elmar Bucher for help with the mplexable pipeline.

Finally, I would like to thank my family, including Elliot Gray, for scientific feedback and helping out with the baby, and Sierra and Felix, my daughter and son, for helping me take breaks.

DEDICATION

Dedicated to my mom
Ruth Eng, dearly missed
by family, friends, neighbors

It is for you that I fight
Battling cancer
More time is what we wanted

Chapter 1. BREAST CANCER BIOMARKER DISCOVERY USING MULTIPLEX IMAGING

1.1 BREAST CANCER INCIDENCE AND CLINICAL MANAGEMENT

Breast cancer is the most common cancer worldwide¹. In the United States, it accounts for 30% of female cancers and 15% of cancer deaths, and is the leading cause of cancer deaths in women aged 20 – 59 years². Although breast cancer death rates are declining, the decline has slowed in the last decade, emphasizing a need for basic and clinical research to advance treatment options².

Clinically, breast cancer is divided into three subtypes by hormone receptor (estrogen receptor [ER] and progesterone receptor [PR]) and human-epidermal growth factor receptor (HER2) status: luminal ER-and/or PR-positive, HER2-positive and triple-negative breast cancer (TNBC)³. These categories are clinically useful for determining the standard of care treatment.

1.1.1 *Luminal ER+ Standard of Care*

Treatment for early stage ER+ breast cancer is hormonal therapy for 5 – 10 years, depending on clinical characteristics such as patient age, menopausal status, tumor stage, size and grade as well as nodal involvement or high proliferation rate (i.e., Ki-67)⁴. Common hormonal therapies are tamoxifen, a selective estrogen receptor modulator or aromatase inhibitors, which block the conversion of androgens to estrogen⁴.

Luminal ER+ breast cancer can be subdivided into good prognosis luminal A and poorer prognosis luminal B using gene expression profiling⁵. Molecular testing platforms such as OncotypeDX®, MammaPrint®, PAM-50 ROR® score, EndoPredict® and the Breast Cancer Index® are used to predict risk of recurrence within 1-5 years, and some are predictive beyond five years⁴. Alternatively, immunohistochemical staining with the proliferation marker Ki-67 can distinguish luminal B tumors, although use of Ki-67 as a predictive biomarker has been difficult to standardize and implement in the clinic⁴. This is due to visual scoring of Ki67 positivity, a marker for which even low values are clinically important⁶. Automated methods perform as well as manual methods for predicting patient prognosis, and can help standardize Ki67 scoring for use in routine diagnostics⁶.

Adjuvant chemotherapy is indicated for ER+ breast cancers with either high grade, four or more positive lymph nodes, low hormone receptor staining, high Ki-67 and extensive lymphovascular invasion⁴. Luminal A is recognized to be less responsive to chemotherapy, but luminal B disease is not automatically treated with chemotherapy, rather, chemotherapy is omitted for patients with low molecular testing scores⁴.

Neoadjuvant chemotherapy is generally not indicated for patients with luminal ER+ disease⁴. Endocrine neoadjuvant therapy without cytotoxic therapy is often recommended for “luminal-A like” breast cancer not suitable for breast conserving surgery at diagnosis⁴.

1.1.2 TNBC and HER2+ Standard of Care

TNBC is typically treated with anthracycline and taxane-based chemotherapy, or possibly a platinum-based regimen if the patient has a BRCA mutation⁴. HER2-positive, stage 2 disease is also treated with anthracycline and taxane-based chemotherapy with the addition of anti-HER2 therapy⁴.

Neoadjuvant chemotherapy (taxane- and/or anthracycline- based) is recommended for TN and HER2+ disease, with the addition of anti-HER2 therapies trastuzumab and pertuzumab for HER2+ disease and platinum based therapy for TNBC patients with BRCA mutation⁴.

1.2 TARGETED THERAPIES IN BREAST CANCER

There have been successes and failures with targeted therapies in breast cancer, and improved biomarkers could help predict success for patients. Until recently, chemotherapy was the standard of care treatment for patients with early-stage TNBC, as those patients lacked expression of ER and HER2 which are indications for endocrine therapy and anti-HER2 therapy, respectively. However, in July 2021, the Food and Drug Administration (FDA) approved pembrolizumab, an anti-PD1 immune checkpoint inhibitor, for high-risk, early-stage TNBC in combination with chemotherapy as neoadjuvant treatment, and then continued as a single agent as adjuvant treatment after surgery⁷. In the phase III KEYNOTE-522 trial patients receiving pembrolizumab in combination with chemotherapy had higher pathologic complete response (63%) and significantly higher event-free survival at 36 months (EFS; 84.5%) versus the chemotherapy only group (56% pCR, 76.8% EFS)⁸. Patients benefited from pembrolizumab regardless of whether their tumors stained positive for PD-L1, the ligand for PD-1⁹. While PD-L1 is a predictive biomarker in the advanced setting, it does not differentiate responders and non-responders in the early setting¹⁰.

Even more recently, in March 2022, the FDA approved olaparib, a PARP inhibitor, for the adjuvant treatment in patients with germline BRCA-mutated (gBRCAm) HER2-negative, high-risk early breast cancer who have been treated with neoadjuvant or adjuvant chemotherapy¹¹. Approval was based on the phase III OlympiA trial which enrolled patients with gBRCAm early breast cancer¹². The three-year invasive disease-free survival showed an absolute improvement of 8.8% in the olaparib versus placebo arm, while distant disease free survival showed a 7.1% absolute improvement¹². How to best integrate pembrolizumab with olaparib and accompanying biomarker development is an area of urgent clinical need¹⁰.

While anti-PD1 agents are becoming a new standard of care in TNBC, the saga of atezolizumab, an anti-PD-L1 immune checkpoint inhibitor, has generated controversy and illustrated the need for good predictive biomarkers (PD-L1 proved to be an inadequate one). Atezolizumab was granted FDA approval for use in metastatic TNBC based on the phase III Impassion 130 trial's observation of longer survival in patients receiving atezolizumab whose tumors were positive for PD-L1; however, results from a second trial, the Impassion131, showed no survival benefit in PD-L1 positive patients and the approval was withdrawn¹³⁻¹⁵. While it would be easy to conclude that PD-1 is simply a superior target in TNBC, a case report shows atezolizumab controlling metastatic TNBC that had progressed on pembrolizumab¹⁶. These recent developments in the landscape of TNBC treatment emphasize a need for predictive biomarker development for targeted therapies,

as well as further basic and translational studies to identify new treatments for TNBC that does not respond to PD-1 or PARP inhibitors.

Patients with early ER+ breast cancer recently received approval for a new target therapy in October 2021 when the FDA approved the CDK4/6 inhibitor abemaciclib in combination with endocrine therapy for adjuvant treatment of ER+, HER2- early breast cancer at high risk of recurrence¹⁷. The approval was supported by results from the phase III monarchE trial which evaluated adding two years of abemaciclib to five years of endocrine therapy in high risk ER+ patients¹⁸. They found that the combination showed an absolute improvement of 5.4% in invasive disease-free survival and of 4.2% in distant recurrence-free survival at three years follow-up¹⁸. Interestingly, patients with Ki-67 high and Ki-67 low tumors both showed benefit from CDK4/6 inhibition, which targets proliferation, underscoring the need to develop predictive biomarkers of response¹⁸.

In metastatic ER+ disease, CDK4/6 inhibitors have been approved in combination with endocrine therapy since 2015, where they have led to a median improvement of 8.8 months in progression free survival¹⁹. Clinicopathological factors fail to predict which patients will derive benefit from CDK4/6 inhibition, again underscoring the importance of developing biomarkers for these agents.

PIK3CA mutation occurs in approximately 40% of patients with ER+ disease, and can be targeted with the PI3K α -specific inhibitor alpelisib which was approved in ER+, HER2- metastatic breast cancer in 2019^{20,21}. The phase III SOLAR-1 trial combined alpelisib with fulvestrant, an estrogen receptor antagonist²¹. Patients with a PIK3CA mutation receiving the combination has a median progression free survival of 11 months (at 20 month follow up) versus 5.7 months in the control arm²¹. Areas of active investigation are whether PI3K inhibition can be used to overcome resistance in patients that progressed on CDK4/6 inhibitors and whether PI3K inhibition will play a role in TN and HER2+ treatment²². How PI3K α inhibition synergizes with immunotherapies is also an area of active investigation²³.

Many different immunotherapy strategies are being investigated in phase I clinical trials in breast cancer. Adoptive T-cell therapy with harvest, expansion and reinfusion of autologous tumor infiltrating lymphocytes (TILs) caused durable remission in a metastatic ER+ breast cancer patient²⁴. T cells with synthetic chimeric antigen receptors (CAR-T) cells are being tested in HER2+ and TNBC (NCT01837602, NCT02792114)²⁵. Immune checkpoint targets LAG3, TIGIT, and B7-H4 are being investigated in combination with anti-PD1 in phase I trials (NCT03250832, NCT03514121, NCT03628677)²⁵. Proinflammatory cytokines IL-6, IL-8 and TGF- β , which are associated with epithelial-to-mesenchymal transition (EMT), tumor progression, metastasis, therapy resistance, and poor outcomes, are also being targeted in phase I trials (NCT03135171, NCT02370238, NCT03524170, NCT03579472, NCT03620201)²⁵. Effector cells release IL-1 β which can drive tumor invasion, which is targeted in a phase I trial (NCT03742349)²⁵. Targeting of B-cells, tumor associated macrophages, and myeloid derived suppressor cells, which are thought to have pro-tumor roles, is currently being investigated (NCT02403271, NCT02393794, NCT02637531, NCT02265536, NCT01596751, NCT02435680, NCT02824575)²⁵. Beneficial cells such as natural killer (NK) cells are also treatment targets (NCT02627274, NCT03319459)²⁵. STING-activators, which can stimulate immune-cold tumors to generate an immune response, are of interest (NCT02675439, NCT03172936)²⁵. Clearly, the proliferation of immune targeted

therapies necessitates the need to characterize the heterogeneity in the breast tumor immune microenvironment and develop microenvironment biomarkers to aid in stratification of patients for different immunotherapy approaches.

1.3 PROGNOSTIC AND PREDICTIVE BIOMARKERS IN BREAST CANCER

As alluded to above, biomarkers can have prognostic and predictive value. Predictive biomarkers such as ER and HER2 allow selection of treatment; however, PD-L1 and Ki-67 have proven to be insufficient as predictive biomarkers in breast cancer for anti-PD-L1 and anti-CDK4/6 therapies, respectively^{13,18}. Prognostic biomarkers predict risk of disease recurrence and/or lethality, and are therefore used to make clinical decisions about escalation/de-escalation of therapy. For example, because luminal A patients have a favorable prognosis, the small benefit conferred by chemotherapy does not outweigh the risk⁴. In high-risk ER+ patients, however, the recurrence rate can be decreased with adjuvant chemotherapy²⁶. Therefore, an important clinical need is to select patients at high risk-of-recurrence to receive chemotherapy while patients with lower risk can choose to avoid unnecessary treatment²⁶.

The OncotypeDX 21-gene recurrence score (RS) assigns patients into high, intermediate, and low risk of recurrence categories²⁷. Four immunohistochemical markers: ER, PR, Ki-67 and HER2 can provide information similar to RS in a score designated IHC4²⁸. Alternatively, the risk-of-recurrence (ROR) score uses the 50-gene PAM50 signature used to classify the intrinsic breast cancer subtypes, weighting the proliferation genes and adding a tumor size variable.²⁹ The ROR score predicts risk of distant recurrence after endocrine therapy in patients with ER-positive, node-negative disease better than the RS, with fewer patients placed in the intermediate group compared to RS²⁶. Since the intermediate group does not benefit from chemotherapy, ROR better predicts chemotherapy response²⁶.

Another important clinical question is late recurrence, that is, recurrence 5 – 10 years after diagnosis, which remains a risk for women with ER+ disease³⁰. In women who received 5 years of hormone therapy (tamoxifen or anastrozole, an aromatase inhibitor) late recurrence is best predicted with the ROR score, with nodal status and tumor size also predictive³⁰. Such scores are important clinically for selecting patients to receive hormone therapy beyond 5 years, which has shown to significantly reduce the risk of recurrence in ER+ disease³¹.

1.4 BIOMARKERS IN THE BREAST TUMOR MICROENVIRONMENT

One extensively studied biomarker in the tumor microenvironment are tumor infiltrating lymphocytes (TILs), which are scored in hematoxylin and eosin (H&E) stained slides produced in the routine pathology workflow. TILs are a prognostic and potentially predictive biomarker for immunotherapy³². Standardized methods for evaluating TILs in H&E stained slides are validated³³. TILs include mononuclear lymphocytes and plasma cells, and are scored as intratumoral, or directly touching tumor cells, and stromal, or located dispersed in the stroma between tumor cells³³. In H&E sections, stromal TILs have been found to be a superior and more reproducible parameter than intratumoral TILs³³. Stromal TILs are scored by the percent of stromal area occupied by TILs (excluding the areas of the slide occupied by tumor nests)³³. Tertiary lymphoid structures are not scored, as they generally form in peritumoral areas and may be absent in core biopsies or heterogenous in full sections³³.

Many studies have investigated the prognostic and predictive value of TILs in H&E stained breast cancer tissues. Either core needle biopsies or full sections are scored, with the average TIL percentage over all the stromal area reported³³. Tissue microarrays (TMAs) have been used in CD8 and FoxP3 biomarker studies, and the results were concordant with other studies in the field^{34,35}. More evaluation of TILs in H&E stained TMAs is needed, but they may be a good option for rapid evaluation of large clinical cohorts³³. One study showed that a limited number of TMA cores (≤ 3) are needed to binarize patients in to high and low TILs, although a larger number of cores (≥ 11) are needed to accurately estimate the mean TIL abundance^{36,37}.

1.5 TILS INTERACT WITH TUMOR CHARACTERISTICS AND TREATMENT IN ER+ DISEASE

In ER+ breast cancer TILs have a neutral or negative prognostic value. In 1,079 node-positive, ER+ patients, TILs were not predictive of overall survival (OS) after adjuvant chemotherapy³⁸. In a cohort of 187 ER+ breast cancer patients who received adjuvant endocrine therapy but no neoadjuvant therapy, TILs were negatively associated with Oncotype DX recurrence score but not associated with OS³⁹. In a meta-analysis of six trials, Denkert et al found TILs were associated with shorter OS in ER+ HER2- tumors⁴⁰. Analysis of spatial heterogeneity of TILs found that in endocrine therapy treated, ER+ disease, immune abundance was not associated with recurrence free survival, however, high immune cell clustering was associated with poor 10-year, early (0-5 years) and late (5-10 years) recurrence-free survival⁴¹.

However, there is also evidence that TIL confer positive prognosis to some subsets of ER+ disease. In the ER+, Ki-67 high tumors, high-TILs conferred favorable DFS, although this was not statistically significant ($p = 0.48$); in contrast, unfavorable DFS was observed in the Ki67-low group ($p = 0.027$)⁴². In a separate study, tamoxifen-treated ER+ tumors in the highest proliferation tertile had significantly longer DFS when they had high immune gene expression⁴³. In the neoadjuvant setting, TILs were found to be predictive of higher pathologic complete response (pCR) in ER+ tumors treated with chemotherapy⁴⁰. Therefore, in proliferative ER+ tumors or when predicting pCR, TILs are a positive prognostic factor in ER+ breast cancer.

In contrast to neoadjuvant chemotherapy (NAC), evidence supports a negative impact of TILs in ER+ cancers treated with aromatase inhibitors. Evaluating TILs before and after neoadjuvant hormonal therapy found that increased TILs post letrozole, an aromatase inhibitor (AI), was significantly associated with non-responders⁴⁴. Another study found in tumors with low numbers of CD8 T cells (below the median), patients receiving exemestane, an AI, had longer DFS than those receiving tamoxifen in two separate cohorts⁴⁵. In the CD8-high patients, there was no difference in DFS between exemestane and tamoxifen treatment⁴⁵. One hypothesis is that CD8 T cells are a biomarker of luminal B tumors and therefore select the poorer prognosis tumors, whereas low TILs are more frequent in the lower mutational burden, more ER-dependent tumors⁴⁵. AI is preferred over tamoxifen for higher risk patients⁴, thus CD8 as a biomarker could simply correlate with, but not be mechanistically linked to benefit from AI. However, there is evidence that the effects of AI are dependent on immune suppression rather than activation⁴⁶.

Asano et al. introduced a biomarker combining residual cancer burden (RCB) after NAC with TILs, with RCB-TILs+ patients having minimal residual disease in the resection and TILs in the core needle biopsy. In ER+ breast cancer, high TILs alone were not associated with DFS, but being RCB-TILs-positive predicted significantly prolonged DFS.⁴⁷ In ER+ patients treated with endocrine therapy, patients who were not RCB-TILs positive had a high rate of recurrence⁴⁷. These data suggest that TILs should be considered in combination with tumor characteristics for better stratification of risk and prediction of response to therapy.

One tumor characteristic that is directly relevant to anti-tumor immunity is human leukocyte antigen HLA expression. In order for anti-tumor immune responses to be generated, the antigens must be presented with HLA molecules for recognition by T cells⁴⁸. HLA class I expression was found to be lower in ER+/HER2- breast cancer relative to other subtypes⁴⁹. HLA correlated increased TILs, and similar to TILs content, high expression of HLA was associated with pCR after chemotherapy but worse DFS in patients receiving adjuvant endocrine therapy⁴⁹. HLA and ER expression are inversely correlated in breast cancer samples, and gene expression of HLA and ESR1 are negative correlated in TCGA and the cancer cell line encyclopedia⁵⁰. Low HLA expression should make tumors more sensitive to NK cells, and in fact, NK cells derived from tumor draining lymph nodes preferentially lysed HLA class I low MCK-7 cell lines compared to triple-negative MBA-MB-231 cells⁵¹. However, a mechanism underlying the anti-correlation in ER and HLA has yet to be elucidated⁵². Antigen processing machinery, including β 2-microglobulin, transporter associated with antigen processing (TAP) 1 and calnexin, which are essential for antigen processing and loading onto HLA, had lower expression comparing breast lesions with brain metastases to those without⁵³. CD8 T cell infiltration was significantly higher in primary tumors without brain metastases and positively correlated with TAP1 expression. Furthermore, inhibition of TAP1 with silencing hairpin (sh) RNA increased spontaneous brain metastases in a mouse model of breast cancer⁵³. These results suggest that defects in antigen-processing machinery may increase the probability of metastasis through decreased immune surveillance.

The negative impact of TILs in ER+ breast cancer could be due to tumor genotype influencing the immune microenvironment. PIK3CA mutation occurs in approximately 40% of patients with ER+ disease and has been investigated for influence on the immune microenvironment. One study quantified CD8, CD4 and FoxP3+ cells in 563 early stage postmenopausal ER+ breast cancer patients randomized to tamoxifen or no treatment⁵⁴. CD8+ lymphocytes were more common in PIK3CA-mutated tumors, especially in the low-grade tumors, which are more likely to have PIK3CA mutation⁵⁴. Similar to other studies, within the relatively good prognosis low-grade tumors, high CD4 and CD8 were associated with poor outcome⁵⁴. Lymphocytes, especially FoxP3+ cells, were more common in tumors with increased PI3K pathway phosphorylation⁵⁴. If CD8 infiltration is consistently correlated with PI3KCA mutation, then CD8 correlation with OS may be simply explained by PI3K pathway activation, which has been implicated in endocrine resistance^{55,56}. However, the PI3K pathway has been mechanistically linked to immune suppression: B7-H1-mediated immune resistance is reduced by PI3K inhibition, is dependent on S6K1-mediated translational regulation of B7-H1⁵⁷. The pathway is also relevant to other breast cancer subtypes. In HER2+/ER+ breast cancer, PIK3CA mutation predicted lower pCR to HER2-targeted neoadjuvant therapy⁵⁸. PIK3CA alterations are also frequently seen in the luminal androgen receptor (AR) subtype of TNBC, with 40% of AR+ cases having mutation versus 4% of

AR-.⁵⁹ Further characterization of the immune microenvironment versus the genomic landscape of breast cancer will be important in discovering links between tumor cell-intrinsic and tumor cell-extrinsic mechanisms of drug response and resistance.

1.6 TILS PROGNOSTIC AND PREDICTIVE IN HER2+ DISEASE

HER2+ breast cancer accounts for 15-20% of all breast cancers⁶⁰. In HER2+ disease, immune infiltrate is a neutral or positive prognostic marker and predictive of response to therapy. In multiple studies, TILs were predictive of pCR following neoadjuvant chemotherapy^{40,61}. TILs also predicted response to adjuvant trastuzumab and chemotherapy⁶². In one cohort, TILs predicted response to neoadjuvant trastuzumab, an anti-HER2 antibody, but not lapatinib, a small molecule inhibitor of HER2 and EGFR⁶³. However, in another cohort, TILs predicted pCR following trastuzumab, lapatinib or both⁶⁴. A meta-analysis of five studies found that high TILs are associated with a higher rate of pCR irrespective of which neoadjuvant anti-HER2 agent and chemotherapy regimens are used⁶⁵.

The relationship between TILs and overall survival have been positive⁶⁴ or no relationship⁶³, depending on the patient cohort. In advanced HER2+ breast cancer, TILs were predictive of OS but not DFS⁶⁶. Conversely, a meta-analysis of six trials found that newly diagnosed HER2+ patients with high TILs had significantly longer DFS but not OS⁴⁰. Mechanistically, trastuzumab action has been shown to be immune-mediated. Trastuzumab binds HER2 and activates NK cells, increases HER2 uptake and processing by DCs and enhances the generation of CD8 T cells, essentially acting as “vaccine” against HER2 that results in T-cell and antibody immunity against HER2²⁵.

1.7 TILS DELINEATE GOOD PROGNOSIS TNBC

TNBC represents 15% of breast carcinomas⁶⁷. Multiple studies have established TILs as predictive of pCR and overall survival in TNBC^{38,40,68}. Gruosso et al. used spatial analysis to elaborate on the simple classification of high/low TILs, classifying the tumor microenvironment in TNBC into three groups based on CD8+ TIL spatial distribution⁶⁹. They define an immunoreactive TME as having high intratumoral CD8, type I IFN signatures and elevated expression of immune checkpoint molecules indoleamine 2,3-dioxygenase (IDO) and PD-L1, pro-inflammatory CD68+CD206- macrophages, which correlated with good outcomes⁶⁹. They also observed an immune-cold microenvironment with an absence of CD8+ T cells, fibrotic stroma, expression of immunosuppressive marker B7-H4 and poor outcomes, and a second poor outcome group with stromally restricted CD8, stromal PD-L1 and enrichment for signatures of cholesterol biosynthesis⁶⁹. These three “microenvironment subtypes” of TNBC have yet to be validated in additional cohorts and further characterized for drug vulnerabilities.

Spatial modelling found that hypothetical tumor-produced chemo-attractant and repellent, but not stromal barriers, could explain the distribution of CD8 T cells in 28 TNBC tumors⁷⁰. Similar characterization of the spatial organization of T cells and other immune cells in ER+ breast cancer has not been done⁵².

1.8 IMMUNE SUBSETS AS BIOMARKERS

More insight into the nature of anti-tumor immunity in breast cancer progression can be obtained from analysis of immune cell subsets. Ali et al. analyzed 11,000 breast tumor transcriptomes, using CIBERSORT to deconvolve 22 subsets of immune cells⁷¹. In ER+ disease, tumors lacking immune infiltration had intermediate prognosis and the presence of M0 macrophages was associated with poor prognosis. In ER- tumors, CD8+ T cells and activated memory T cells were associated with good outcome, and T follicular helper cells with response to neoadjuvant chemotherapy⁷¹. T regs, M0 and M2 macrophages were strongly associated with poor outcome, regardless of ER status⁷¹. A separate study found T follicular helper cells were predictive of pCR in a cohort treated with NAC, and they predicted survival in all breast cancer subtypes in cohort 794 breast cancer patients receiving no hormone or chemotherapy before or after surgery⁷². Intriguingly, two previously published immune gene signatures were also associated with good prognosis the ER+ patients in this untreated cohort, suggesting that the negative association between immune infiltrate and prognosis may not be observed in untreated ER+ breast cancer patients⁷². In a second CIBERSORT analysis of 7,270 primary breast tumors, M0 macrophages were associated with worse OS in ER+/HER2- tumors, and a high CD8+ T cell exhaustion signature was associated with shorter DFS in ER+ tumors⁷³. M1 macrophages were associated with pCR and longer DFS in ER+, and longer OS in ER+/HER2+⁷³.

FoxP3+ T regulatory cells (Tregs) are associated with poor survival in ER+ breast cancers that lack CD8+ T-cell infiltrate, but in HER2+/ER- breast cancers they are associated with better survival, especially in tumors infiltrated with CD8 T cells³⁴. More prognostic information may exist in the CD8/FoxP3 T cell ratio, which after neoadjuvant chemotherapy was found to be highly predictive of DFS and OS in HER2+ and HER2- breast cancer⁷⁴. B-regulatory cells (Bregs) are correlated with Tregs in breast cancer, and patients with Bregs have significantly shorter metastasis free survival than those with Tregs alone⁷⁵. A meta-analysis of 16 studies found tumor associated macrophages predict worse patient prognosis in breast cancer, and CD68 is a superior biomarker to CD163 or CD206 alone⁷⁶. Macrophage-colony stimulating factor, CSF, and its receptor, CSF1R, correlate with poor prognosis⁷⁷.

Correlations between tumor phenotype, microenvironment subtype and prognosis have been observed. An analysis of 10,000 tumors from TCGA across 33 cancers identifies six immune subtypes⁷⁸. Basal-like breast cancer, HER2 and luminal B cancers were 60, 49 and 46% INF- γ dominant subtype, respectively, with high CD8 T cell and M1 macrophage density⁷⁸. In this subtype, a higher lymphocyte expression signature improved survival⁷⁸. Separately, Miller et al found that within basal-like, HER2-enriched and luminal B tumors, an immune classifier could stratify patients into good and poor prognosis, but was not prognostic in Luminal A and claudin-low subtypes⁷⁹. Characterization of tumors that could be stratified by immune classifiers, termed “immune benefit enabled” (IBE) versus those that could not be stratified, termed “immune benefit disabled” (IBD), showed that IBE tumors expressed genes related to Th1-T cell polarization, INF γ and TNF α signaling, and cytotoxicity, while IBD tumors had higher TGF- β , an immunosuppressive cytokine⁷⁹. The emerging picture is one of breast cancer subtypes with reproducible differences not only in their epithelial compartment, but in the surrounding tumor microenvironment. Considering the breast cancer subtype and microenvironment context will likely yield the most prognostic and predictive information.

1.9 SIGNIFICANCE OF FIBROBLASTS AND ECM COMPONENTS OF THE TUMOR MICROENVIRONMENT

While much focus has been on immune components of the microenvironment, evidence implicates fibroblasts as important mediators of immune suppression and increased cancer growth, invasion and drug resistance. Quiescent or resting fibroblasts are non-proliferating spindle-shaped cells embedded in the extracellular matrix (ECM) containing type I collagen, laminin and fibronectin. In response to injury, fibroblasts become activated, gaining expression of α -smooth muscle actin (α SMA) and vimentin, becoming stellate in shape, and producing ECM components such as fibronectin. Under consistent stimulation, such as cancer lesions, fibroblasts may become irreversibly activated and acquire enhanced proliferation, ECM-remodeling, and immunomodulatory functions, a phenotype referred to as cancer-associated fibroblasts, or CAFs⁸⁰.

The CAFs present in tumors can have multiple sources, and heterogenous and non-fibroblast-specific marker expression, necessitating multiple markers as well as morphology and spatial distribution to be taken into account when identifying fibroblasts⁸⁰. CAFs may originate from activated tissue resident fibroblasts or bone-marrow derived mesenchymal stem cells (BM-MSCs), the latter which have been shown to promote metastasis in breast cancer mouse xenografts⁸¹. Additional sources of CAFs include endothelial cells, mesenchymal stem cells, liver and pancreatic stellate cells, and even epithelial cells⁸⁰. The heterogeneity of CAF origin and activation is reflected in overlapping and non-overlapping markers of fibroblast activation including vimentin, α SMA, podoplanin (PDPN), fibroblast activation protein (FAP), PDGF receptor- α (PDGFR α), PDGFR β , desmin and caveolin 1 (CAV1)⁸⁰.

A well characterized mechanism of tumor promotion by CAFs is via IL-1 β , a cytokine that is being targeted in phase I clinical trials in breast cancer²⁵. Immune cell derived IL-1 β has been shown to induce a proinflammatory gene signature in mouse mammary CAFs, a mechanism common to skin and pancreatic tumors that promoted macrophage recruitment, vascularization and tumor growth⁸². Recently, cancer-associated fibroblasts (CAFs) were demonstrated to produce IL-1 β themselves through sensing of damage-associated molecular patterns and activation of the NLRP3 inflammasome pathway⁸³. The CAF-derived IL-1 β recruited immune suppressive MDSCs and upregulated adhesion molecules on endothelial cells which facilitated tumor migration across the blood vessel endothelium, leading to increased lung metastases in a mouse breast cancer model⁸³.

CAFs have been shown to be positive regulators of cancer growth and invasion. They can support tumor growth directly, as in CAF secretion of stromal cell-derived factor (SDF-1) which stimulates breast tumor growth through the cognate receptor CXCR4 expressed on tumor cells, or indirectly, through SDF-1 mediated recruitment of endothelial cells⁸⁴. Mechanisms often involve ECM remodeling: CAFs from the MMTV-PyMT mouse model of breast cancer showed activation of the mechano-sensitive Yes-associated protein 1 (YAP), leading to matrix stiffening, a positive feedback loop that increases CAF activation and cancer cell invasion⁸⁵. In the aforementioned study, researchers observed fibronectin was increased in CAFs relative to normal fibroblasts and CAFs had enhanced matrix remodeling capabilities, producing thicker collagen fibers and stiff matrix which promoted the invasion of breast cancer cells⁸⁵. CAFs also alter the ECM in tumor-promoting ways through the production of ECM-degrading proteases such as matrix-metalloproteins (MMPs)⁸⁰. MMP3, a.k.a. stromelysin 1 is produced by CAFs and cleaves E-

cadherin, promoting epithelial-to-mesenchymal transition (EMT) and invasiveness in breast cancer cells⁸⁶. CAF derived MMP1 has similarly been linked to breast cancer invasion⁸⁷. Periostin (POSTN), a component of the extracellular matrix, is produced by stromal α SMA+VIM+ fibroblasts in primary tumor, recruiting WNT ligands that maintain a CD90+ CD24+ cancer-stem-cell phenotype that enables lung metastasis in the MMTV-PyMT mouse model of breast cancer⁸⁸. These studies establish a connection between CAF matrix-remodeling and aggressive EMT or stem-like phenotypes in breast cancer cells.

CAFs have been shown to metabolically support tumor cells. Metabolic coupling occurs between catabolic reactions, such as glycolysis, in fibroblasts and anabolic reactions, such as oxidative phosphorylation (OXPHOS), in cancer cells and occurs when CAFs encounter oxidative stress and hypoxia⁸⁹. CAV-1 loss in the stroma is a marker of this metabolic coupling, which features OXPHOS in the epithelial compartment and stromal glycolysis⁸⁹. Stromal expression of the gene GFPT2, which encodes for the rate limiting enzyme in the hexosamine biosynthesis pathway, enabled higher glucose uptake in lung adenocarcinoma and was prognostic⁹⁰. Links between CAF and cancer metabolism continue to be an area of active investigation.

CAFs can modulate the anti-tumor immune response in diverse ways. CAF-derived IL-4, IL-6 and IL-8 may induce immunosuppressive myeloid cell differentiation and CAF produced TGF- β is primarily immunosuppressive⁸⁰. Depletion of FAP+ CAFs, which also expressed α SMA and Collagen I, led to recruitment of intratumoral CD8+ T cells and immune control of tumor growth in a subcutaneous model of pancreatic ductal adenocarcinoma (PDAC)⁹¹. Conversely, deletion of α SMA+ myofibroblasts in PDAC mouse models led to invasive, undifferentiated tumors with increased hypoxia, EMT and reduced survival, while PDAC patients with fewer myofibroblasts also had reduced survival⁹². The myofibroblast-depleted mouse tumors had increased CD4+FoxP3+ T regs, were resistant to gemcitabine, but responded to anti-CTLA4 immunotherapy⁹², motivating strategies to deplete CAFs combined with immunotherapy⁹³. ECM can physically alter immune cell migration; it has been suggested that loose fibronectin and collagen regions promote T cell migration, while dense and aligned ECM structures inhibit T cell movement^{94,95}. The conflicting roles as CAFs mediating pro- and anti-tumor immune responses likely reflects their substantial heterogeneity and plasticity and the need to better characterize the correlation of CAF phenotypes with immune subsets in patient tumors.

Finally, CAFs can alter tumor phenotypes in ways that promote therapy resistance. A stroma-related gene signature characteristic of activated fibroblasts predicts resistance to neoadjuvant chemotherapy in ER-negative breast cancer patients⁹⁶. Sensitivity to radiotherapy is promoted by blocking B1-integrin adhesion of breast cancer cell lines to ECM, which suppressed Akt-mediated resistance⁹⁷. Adhesion of tumor cells to the ECM can drive programs for dormancy and EMT, which contribute to drug resistance⁹⁸. Characterization of the spatial proximity of tumor cells and ECM components is likely to reveal additional insights.

CAFs in some contexts can be tumor-restraining as opposed to tumor-promoting. In ER+, HER2-breast cancer, they take on more of a Sméagol than a Gollum character, as tumors with high numbers of stromal cells and ECM content, and lacking in immune infiltration make up a low risk group of cancers⁹⁹. Stromal content as a biomarker seems to be subtype specific, as increased stroma was associated with poor prognosis in TNBC and good prognosis in ER+ tumors^{100–102}. In

ER+ breast cancer, CD146-positive fibroblasts were shown to support ER expression in breast cancer cells and maintain sensitivity to tamoxifen.¹⁰³ On the other hand, in TNBC, a myofibroblastic CAF subset promotes an immunosuppressive environment through attraction and promotion of differentiation of FoxP3+ Tregs¹⁰⁴. Future studies must consider both fibroblast and tumor heterogeneity for improved understanding of the role of CAFs in breast cancer.

1.10 BIOMARKER DISCOVERY WITH MULTIPLEX IMAGING

Due to the diversity of cell types and states that are linked to patient prognosis and treatment response in breast cancer, multiparameter profiling of the tumor microenvironment is necessary. Comprehensive characterization of the ecosystems of cellular and extracellular components in breast cancer should improve prediction of therapy response and tumor progression. Moreover, spatial and morphological information is necessary to define cell types and distinguish clinically relevant spatial signatures in tissues. Therefore, multiplex imaging methods have been introduced for high-dimensional spatial proteomic profiling of the tumor microenvironment^{105–110}. For more discussion of these methods, see section 3.2. Multiplex imaging enables identification of multiparameter biomarkers for patient stratification, and furthermore, allow identification of cellular ecosystems and spatial arrangements that correlate to cancer aggressiveness and relapse. Such data not only aid in patient selection and stratification, but yield deeper understanding of the interacting biological systems underlying tumor progression.

Several studies in breast cancer have employed multiplex-ion beam imaging (MIBI) and imaging mass cytometry (IMC) to characterize breast cancer patient cohorts. Using the MIBI platform, Keren et al. found that TNBC patient tissues exhibited either compartmentalized or mixed tumor-immune spatial organization with compartmentalized tumors having better outcomes¹¹¹. Jackson et al. identified breast cancer subtypes based on single cell composition of tissues and further grouped tissues according to spatial communities of cells¹¹². The authors found many associations between single cell density and tumor or microenvironment subtypes and outcome, but only spatially defined tumor or microenvironment communities were predictive of overall survival independent of clinical variables¹¹². Ali et al. used IMC to analyze breast cancer patient tissues from the METABRIC cohort which also had genomic information¹¹³. They were able to compare the intrinsic breast cancer subtypes and genomic alterations to multiplex imaging phenotypes, finding that the luminal B subtype was defined by low cytokeratin (CK) and ER expression, β 2-microglobulin deletion and PD-L1 amplification were associated with the hypoxia-marker carbonic anhydrase IX and genomic instability correlated with proliferation, T cell and macrophage infiltration¹¹³. Finally, the authors calculated significant spatial neighbor interactions using permutation testing, which was found to be more prognostic than abundance alone for myofibroblasts with heterotypic interactions¹¹³. All three of the studies highlight the additional prognostic information gained from spatial organization of cells in the tumor microenvironment.

1.11 THESIS OUTLINE

This thesis describes work undertaken to establish a multiplex imaging protocol, image processing pipeline and multiplex imaging analysis tools which are then utilized for quantitative analysis of human disease and breast cancer model organisms. In **Chapter 2**, we establish our breast cancer antibody panel and protocol for cyclic multiplex immunofluorescence (cmIF, later CyCIF) staining of formalin-fixed, paraffin embedded (FFPE) tissues. In **Chapter 3** we develop tools for image

processing and reproducible analysis, which we demonstrate in optimization of signal removal, antibody specificity, background correction and batch normalization in CyCIF. In **Chapter 4** we develop additional novel analytical methods for characterizing tumor heterogeneity and spatial relationships in diverse tissue samples. In **Chapter 5**, we utilize our CyCIF platform for biomarker discovery in clinical cohorts of breast cancer patients, and also extend our tools to utilize publicly available multiplex ion beam imaging (MIBI) and imaging mass cytometry (IMC) datasets as validation cohorts for efficient discovery of robust biomarkers. Finally, we summarize our results and give perspective on future directions in **Chapter 6**.

Chapter 2. CYCLIC MULTIPLEXED-IMMUNOFLUORESCENCE (CMIF), A HIGHLY MULTIPLEXED METHOD FOR SINGLE-CELL ANALYSIS

As previously published in:

Eng J, Thibault G, Luoh SW, Gray JW, Chang YH, and Chin K. Cyclic Multiplexed-Immunofluorescence (cmIF), a Highly Multiplexed Method for Single-Cell Analysis. in Biomarkers for Immunotherapy of Cancer. Methods in Molecular Biology (eds. Thurin, M., Cesano, A. & Marincola, F. M.) 2055, 521–562 (Humana, 2020).¹¹⁰

2.1 ABSTRACT

Immunotherapy harnesses the power of the adaptive immune system and has revolutionized the field of oncotherapy, as novel therapeutic strategies have been introduced into clinical use. The development of immune checkpoint inhibitors has led to durable control of disease in a subset of advanced cancer patients, such as those with melanoma and non-small cell lung cancer. However, predicting patient responses to therapy remains a major challenge, due to the remarkable genomic, epigenetic, and microenvironmental heterogeneity present in each tumor. Breast cancer (BC) is the most common cancer in women, where hormone receptor-positive (HR+; estrogen receptor and/or progesterone receptor) BC comprises the majority (>50%) and has better prognosis, while a minority (<20%) are triple negative BC (TNBC), which has an aggressive phenotype. There is a clinical need to identify predictors of late recurrence in HR+ BC and predictors of immunotherapy outcomes in advanced TNBC. Tumor-infiltrating lymphocytes (TILs) have recently been shown to predict late recurrence in HR+, counter to the findings that TILs confer good prognosis in TNBC and human epidermal growth factor receptor 2 positive (HER2+) subtypes. Furthermore, the spatial arrangement of TILs also appears to have prognostic value, with dense clusters of immune cells predicting poor prognosis in HR+ and good prognosis in TNBC. Whether TIL clusters in different breast cancer subtypes represent the same or different landscapes of TILs is unknown and may have treatment implications for a significant portion of breast cancer patients. Current histopathological staining technology is not sufficient for characterizing the ensembles of TILs and their spatial patterns, in addition to tumor and microenvironmental heterogeneity. However, recent advances in cyclic immunofluorescence enable differentiation of the subsets based on TILs, tumor heterogeneity, and microenvironment composition between good and poor responders. A computational framework for understanding the importance of the spatial relationships between TILs and tumor cells in cancer tissues, which will allow for quantitative interpretation of cyclic immunostaining, is also under development. This chapter will explore the workflow for a newly developed cyclic multiplexed-immunofluorescence (cmIF) assay, which has been optimized for formalin-fixed, paraffin-embedded tissues and developed to process digital images for quantitative single-cell based spatial analysis of tumor heterogeneity and microenvironment, including immune cell composition.

2.2 INTRODUCTION

2.2.1 *Introduction*

2.2.1.1 Standard Practice in Histopathology

Hematoxylin and Eosin (H&E) staining of formalin-fixed, paraffin- embedded (FFPE) tissues in histopathology is most widely used for clinical diagnosis and biomedical research, and is often supplemented by chromogen development-based immunohistochemistry (IHC). Tissue architecture and cellular morphology in H&E-stained and protein expression in IHC-stained tissues are observable by bright-field microscopy, can be preserved for many years, and can be digitally archived through optical camera or Charge-Coupled Device (CCD) imaging systems. IHC-staining is typically a single channel colorimetric stain limited to cellular resolution, and is not suitable for colocalization analysis of multiple protein expressions.

2.2.1.2 Development of Immunofluorescence

Immunofluorescence (IF) is a specialized type of IHC that uses a fluorescent dye to visualize antibody binding. The localization of a protein in tissue and cells, as well as any variations thereof, provides insights into the function of this protein. In early 1950, Albert H. Coons and his colleague paved the way for the future of immunofluorescent staining by developing a fluorescein-labeled antibody and fluorescence microscopy technique(s) to identify antigens in tissue sections, followed by the development of indirect IF, which uses secondary antibodies conjugated to fluorescent dyes. This technique, which is no longer limited to protein detection, is now the most widely used immunofluorescent method. Indirect IF permits multiplexed tissue staining, as well as screening and titrations of patient sera with fluorescently labeled antibodies, just as in any other serologic method, and has contributed to a variety of biomedical research and clinical studies.

2.2.1.3 Tumor Infiltrating Lymphocytes (TILs)

TILs in cancer tissues can be examined with H&E staining¹¹⁴; however, better delineation of lymphocyte subsets and their functional status can be achieved by staining for various cell surface markers using a series of single channel IHCs onto serial tissue sections¹¹⁵. A larger set of biomarkers can be more efficiently interrogated by the highly multiplexed immunofluorescent procedure that enables linear quantification of multiple protein expression at subcellular resolution using a single tissue section^{107,116–118}. To examine the states of immune and tumor cells in cancer specimens, tissues are often dissociated into single cells for flow cytometry analysis, which can also be applied to the bulk analysis of genomic and epigenetic state. However, the context of the tissue architecture will be lost, and as a result, the spatial distribution of the microenvironmental landscape, including TILs and stromal compositions representing the tumor microenvironment state, cannot be analyzed.

2.2.1.4 Cyclic Multiplexed-Immunofluorescence (cmIF)

Recent advances in digital pathology and whole-tissue scanning technology have allowed the development of the cyclic multiplexed immunofluorescence (cmIF) workflow, which can be used to assess the histopathological architecture at a subcellular level and the interactions between tumor cells and their microenvironments in the tissue context. Highly multiplexed imaging will be used to achieve quantitative single-cell analysis of the spatial pattern of all cellular compositions, which will be classified into the categories of heterogeneous tumor cells, immune cells, fibroblast, endothelial cells, and other mesenchymal cells in the tumor microenvironment. The cmIF is a powerful, robust, quantifiable, and reproducible staining tool for highly multiplexed immunofluorescence imaging of tissues and cells. This method has been optimized to quantify the expression of 60 protein biomarkers and apply single cell-based multivariate analysis^{107,116}, all within the same section from an archive of formalin-fixed, paraffin-embedded (FFPE) tissues (Figure 2.1). Further analysis of the spatial patterns of immune effector cells and aggressively proliferating tumor cells beyond cellular composition, colocalization, and proximity assessment is the current subject of machine learning algorithm development, and will help identify spatially defined subpopulations and their associations with clinical outcomes.

2.2.2 *Background*

2.2.2.1 Cancer Immunotherapy

The modulation of the immune system for checkpoint inhibition through therapeutic monoclonal antibodies has rapidly revolutionized treatment strategies in various malignancies over the last decade. Several immunotherapeutic antibodies, namely, “immune checkpoint inhibitors” or “immune modulators,” have been approved by the Food and Drug Administration (FDA), and have led to outcomes with significantly better efficacy and patient tolerance than chemotherapy among advanced cancer patients, such as those with melanoma, lymphoma, or cancers in the lung, bladder, kidney, colon, and rectum. However, increased experience in the use of immunotherapy agents in clinical settings has revealed several limitations, such as treatment resistance and undesirable immunogenic reactions¹¹⁹.

2.2.2.2 Immunotherapy for Breast Cancer

Particularly in breast cancer (BC), a retrospective analysis of TILs by H&E in prospectively collected ~4000 tissues from clinical trials revealed the influence of host immunosurveillance on breast cancer biology while undergoing treatment⁴⁰. Analyses of past trial results have been generally consistent; TNBC and HER2+ BC are most strongly associated with better outcomes, while the association in HR+ BC is potentially limited by the substantial nature of luminal BC subtype. The expression of certain genes in luminal BC subtype—such as p53 mutation, and the amplification of FGFR and CCND1—drive tumor cell proliferation and genomic instability, resulting in aggressive cancer phenotype and resistance to treatment. Interestingly, high-grade luminal BC that achieved a pathological complete response (pCR) showed better long-term survival compared to grade 1 or 2 luminal tumors¹²⁰. In general, there are far fewer TILs in luminal BC that are of the HER2+ BC or TNBC subtypes, and the prognostic benefit of TILs has been shown to be quantitative in nature¹²¹. A number of clinical trials involving immune checkpoint inhibitors for locally advanced and metastatic BC patients are currently ongoing to determine whether immunotherapy provides a clinical benefit in BC treatment, and preliminary analyses of the trials have shown contradictory results between subtypes. The use of immune checkpoint inhibitors, both with and without cytotoxic agents, has shown promising results in TNBC, while the rate of response to immune checkpoint inhibitors is lower in HR+ BC, even when combined with molecular targeting agents^{65,122–126}. These results may reflect a fundamental difference in the landscape of tumor heterogeneity, TILs, angiogenesis, and stromal microenvironment between BC subtypes. The defined tumor area is divided into a stromal region and an enriched tumor cell region, typically referred to as “Stromal TILs” and “Intratumoral TILs” in most studies. Since stromal TILs are generally more abundant, they are readily assessed in H&E stained tissues, in contrast, intratumoral TILs are generally lower in number, more heterogeneous across samples, and less visible. A significant association with clinical outcomes has been established with stromal TILs¹¹⁵. A study of cytotoxic neoadjuvant setting for BC reported that both TIL types were independent predictors for pCR in TNBC, but only intratumoral TILs was an independent predictor in luminal BC. Interestingly, neither TIL type predicted pCR in HER2+ subtype¹²⁷.

Intrinsic gene expression analyses suggest that there are 4–6 molecular subtypes in TNBC, such as BL1, BL2, IM, M, MSL, and LAR. The BL1 subtype is characterized by elevated cell cycle and DNA damage response gene expression, while the BL2 subtype is enriched in growth factor signaling and myoepithelial markers. The IM subtype is composed of genes encoding immune antigens, cytokine, and core immune signal transduction pathways, and likely represents gene expression from both the tumor cells and infiltrating lymphocytes. Both M and MSL subtypes share elevated expression of genes that are involved in the epithelial-mesenchymal-transition and growth factor pathways, but only the MSL subtype shows decreased expression of the genes involved in proliferation and is associated with better outcome. The LAR subtype is characterized by luminal gene expression and is driven by the androgen receptor (AR). LAR, or the steroid receptor subtype conferred poor prognosis in TNBC. The studies identified TNBC cell lines representative of each subtype and demonstrated that they have differential sensitivity to the alkylating agent cisplatin, with BL1 cell lines displaying the greatest sensitivity^{128–130}. Another study identified four TNBC subtypes based on mRNA expression and DNA profiling: (1) LAR, (2) mesenchymal, (3) basal-like immune suppressive, and (4) basal-like immune active, in which immune-suppressive and active basal-like TNBC are associated with the worst and the best

prognosis, respectively¹³¹, suggesting that the basal-like immune suppressive subtype might be a target for immune checkpoint inhibition.

The development of a relevant TIL evaluation method is critical for stratification of the patient population into those who may respond to immune checkpoint blockade. Most studies measure both stromal and intratumoral TILs; however, stromal TILs are considered to be more representative and reproducible. Certainly, stromal TILs are generally more abundant, whereas intratumoral TILs are typically lower in number, more heterogeneous across samples, and less visible. Consequently, significant associations with clinicopathological parameters and disease outcomes have been better established for stromal TILs than for intratumoral TILs. The evaluation of TILs becomes increasingly relevant with the emerging role of immunotherapy in breast cancer. Early phase trials of checkpoint blockade show promising results; however, some patients may not derive any benefits from immunotherapies. This underlines the importance of the development of predictive biomarkers, which may enable further stratification of these BC subtypes into subclasses.

2.2.2.3 Immunotherapy for Prostate Cancer

Prostate cancer (PC) is one of the leading causes of cancer death in men, second only to lung cancer (<https://seer.cancer.gov/statistics/>), for which current standard care is based on surgical resection, radiation, and hormone therapy, including castration. Immunotherapy for advanced PC patients appears promising, as patients with melanoma and other cancers have been shown to respond to immune checkpoint inhibitors¹³², resulting in increased quality of life for those patients. The successful management of advanced PC, such as metastatic castrate-resistant prostate cancer (CRPC), requires an understanding of the complexity of tumor cells and their interactions with the microenvironment, particularly with TILs. Cancer vaccine therapy was first approved for PC, and more have become available in subsequent years, but most fail to meet treatment expectations. However, combining a checkpoint inhibitor with prostate cancer vaccines appears more beneficial for CRPC patients^{133,134}. Preliminary studies on new drugs that block the immune checkpoints in CRPC indicate that such immune checkpoint inhibition might be a solution for these patients. Synergistic combinations of immunotherapy with conventional cancer treatments may have tremendous potential¹³⁵.

2.2.2.4 Immunotherapy for Pancreatic Cancer

Contrary to the promising results of immune checkpoint inhibition among TNBC and CRPC patients, the majority of patients with pancreatic ductal adenocarcinoma (PDAC) have never benefited from immunotherapy in clinical trials^{136,137}. These unexpected outcomes are the result of the nature of PDAC: low immunogenicity and abundant stroma induce an immunosuppressive environment through activated cancer-associated fibroblasts and their TGF- β secretion. A new treatment strategy is required for patients with PDAC, such as the combination of checkpoint inhibitors and FOLFIRINOX, which may present some advantages in addition to anti-TGF- β targeting therapy^{138,139}. Therefore, it is necessary to develop new approaches to further stratify various cancer subtypes into those that will or will not respond to immune checkpoint inhibitors. Improvements in the clinical efficacy of immunomodulation for a majority of advanced cancer patients are critically needed.

2.2.3 Biomarkers

2.2.3.1 Overview

Cyclic IF, including our cmIF, is the most promising approach for highly multiplexed and single-cell-based proteomic analyses because it allows for the interrogation of the spatial distribution and interaction of tumor and microenvironment at the cellular and molecular levels using conventional fluorescence microscopy and inexpensive reagents. The quantitative and spatial analyses of intratumor heterogeneity and tumor microenvironment, which is composed of immune cells, fibroblasts, vasculature, and noncellular components in the stroma, is important for the characterization and classification of further subgroups of malignancy or responsiveness to therapeutic agents. Immunotherapy has emerged as a highly promising approach for the treatment of cancer patients. However, the clinical efficacy of immunotherapy is limited to a minority of cancer patients. Thus, new approaches to improve the efficacy of immunotherapies for most patients are critically needed. Strategies to improve the efficacy of immunotherapy rely on both an enhanced understanding of the tumor/immune interface at the cellular and molecular levels, and the ability to stratify and select appropriate patients for a specific immunotherapy agent or combination therapy. Optimized biomarker strategies could help elucidate both of these areas and allow cancer immunotherapy to be tailored to each individual patient's disease.

2.2.3.2 Tumor Heterogeneity

Tumor heterogeneity is one of the hallmarks of malignancy^{140,141}. Cancer tissues differ greatly between patients (intertumor heterogeneity), as well as within each individual tumor (intratumor heterogeneity)¹⁴², and such heterogeneity is caused by both intrinsic and extrinsic factors. For instance, intrinsic heterogeneity may develop from genomic instability at the malignant transition^{143,144} or due to epigenetic plasticity, and extrinsic signals from the tumor microenvironment can introduce additional heterogeneity to these tumor cells as a result of drug treatment-induced stressors¹⁴⁵. Different cell states can possess distinct drug sensitivities^{146–148}, making cell-state heterogeneity a challenge for therapeutic management of breast tumors. An additional challenge to therapeutic treatment is the inherent plasticity of tumor cell states^{149,150}. Cytotoxic and molecular targeted therapies have been shown to drive cells into drug tolerant persister cell states, which can survive drug pressure in a low-proliferative state, leading to incomplete response or recurrence^{145,151,152}. Early studies— particularly those investigating BC— defined tumor heterogeneity based on several characteristics of intratumor cell populations, including tumorigenicity, treatment resistance, and metastatic potential^{153–155}. Recent investigations support the notion that the TILs, stroma cells, parenchyma, and vasculature making up the tumor microenvironment all play important roles in tumor initiation, progression, and metastasis¹⁵⁶. It is important to understand the state of individual cells among the many tumor cells that may be undergoing differentiation (Keratin subtypes and CD44), hyper proliferation (PCNA and Ki67), evasion of growth repression (pS6RP and pRB), resisting cell death (53BP1 and Rad51), and other cell states, as represented by the many biomarkers listed in Table 2.1.

Table 2.1. Biomarkers

Class	Biomarker	Table 2.2	Target Biomarker Protein for cmIF
Cell State	Differentiation	1A	CK5, CK7, CK8, CK14, CK17, CK18, CK19, CD36, SLUG, SNAIL, TWIST, ZEB1
	Proliferation	1B	Ki-67, PCNA, pS6RP, p21, pRB
	Apoptosis	1C	cPARP, cCaspase3, p53, ATR, p53BP1, Rad51
	Chromatin remodeling	1D	Histone H3 (tri-methyl K27), Histone H3 (tri-methyl K4), Histone H3 (acetyl K9), Histone H3 (pS10), Histone H2A.X (pS139)
	Autophagy	1E	LC3B, SQSTM1/p62, Beclin-1
	Network reporters	1F	pERK, pAKT, pS6RP, pGSK3A/B, pMTORC1, p4EPB1, pEGFR, SMAD1, SMAD5, SMAD8
	Receptor/Junc.	1G	HER2, EGFR, ER, PgR, AR, IGF1R, CDH1, E-Cadherin
Architecture	Nucleus	2A	Lamin A/C, Fibrillarin, Lamin B1, Lamin B2
	Cytoskeleton	2B	Actin, Tubulin
	Organelles	2C	Tomm20, CoxIV, TNG46, Calreticulin, LAMP1, EEA1, RAB7, Catalase, RPS3, Pericentrin, CAV1, γ Tubulin
Immune	Lymphoid lineage	3A	CD3, CD4, CD8, CD20, CD27, CD38, CD39, CD138, FoxP3, PD-1, T-bet, RORgamma(t), GATA3, CTLA-4, CD3, CD45, CD45RO, Nkp46
	Myeloid lineage	3B	CD68, CD66a/b/c/d, CSF1R, CD163, MHC-II, PD-L1, CD56, CD11b, CD11c, CD141, CD83, CD206, Tryptase
Stroma	Stromal cells	4A	CD44, VIM, HSP47, CD34, CD90, FSA, α SMA, FAP, FSP-1, PDGFRa, PDGFRb
	Vasculature	4B	CD31, CD105, FLI1, vWF, Podoplanin, VEGF, LYVE1, CD34, Desmin, Fibrinogen, VE-cadherin
	MEPs*	4C	FN, Laminin, Collagen I, CXCL12, CTGF1, FGF6, BMP3, EGF1, JAG2L, KNG1, BMP2, BMP4, BMP6, DSG2, Collagen IV
Disease	Breast cancer	NA	HER2, ER, PgR, AR, Calponin, p63, GCDPF-15, Mammaglobin, p120 Catenin, panCK, CK5/6/7/8/14/17/18/19, MUC1/2/4/5AC/6, c-Kit, CD31, CD34, CK903, EMACEA, S100, Maspin, WT1, Synap, NSE, CD56, MGB, GCDPF-15
	Pancreatic cancer	NA	Synaptophysin, Chromogranin A, SMAD4, p16, β Catenin, CD10, S100P, EGFR, Ira, CK7/17/19/20, pVHL, Maspin, IMP-3, CAM5.2, MUC1/2/4/5AC/6, Claudin4/18, Mesothelin, Annexin A8, PSCA, CDX-2, Fascin
	Prostate cancer	NA	AR, PSA, c-Myc, p63, AMACR, ERG, NKX3.1, PTEN, PSAP, PSMA, P504S, P501S, p63, CK903, CK5/6, (HMWCK), PAX2, S100A1, Bcl2, MUC1/6, CK7/20, CAM502, panCK (AE1/AE3), CDX2, MSA, S100

Table 2.2. Validated Antibodies

Biomarker	Function	Table 2.1	Vendor (Catalog #)
Keratin	Cytoskeleton	1A	CST (3478)
CycD1	Cell cycle	1B	Abcam (AB190194)
Cyclin A2	Cell cycle (S/M)	1B	Abcam (AB38)
Ki67	Cell cycle	1B	CST (11882)
MCM6	Cell cycle (DNA replication)	1B	Santa Cruz (SC-9843)
p21 (Waf1/Cip1)	Cell cycle (Suppressor)	1B	CST (8587)
p27/Kip1	Cell cycle (Suppressor)	1B	Abcam (AB194234)
PCNA	Cell cycle	1B	CST (8580)
PDI	protein translation	1B	CST (5051)
pHistone H3	Mitosis	1B	CST (3475)
POL2A	Cell cycle	1B	Novus (NB200-598AF488)
pRB	Cell cycle	1B	CST (8957)
pRb (S807/811)	Cell cycle	1B	Santa Cruz (SC-16670)
pS6RP (235/236)	protein translation	1B	CST (4851)
pS6RP (240/244)	protein translation	1B	CST (5018)
pSQSTM(p62)	Cell cycle check point	1B	CST (88588)
S6RP	protein translation	1B	CST (6989)
Bax	Apoptosis	1C	BioLegend (633603)
Bcl2	Apoptosis	1C	BioLegend (658705)
cCasp3	Apoptosis	1C	R&D(IC835G-025)
pATM	Cell cycle check point	1C	Thermo (50-9046-41)
pCHK2 (Thr68)	Apoptosis	1C	CST (12812)
pSRC	Signaling transducer	1C	Thermo (12-9034-41)
Survivin	Anti-apoptosis	1C	CST (2866)
TP53	Cell cycle check point	1C	CST (5429)
gH2AX	DNA damage checkpoint	1D	CST (9720)
gH2AX	DNA damage checkpoint	1D	BioLegend (613407)
14--3-3	Signaling transducer	1F	Santa Cruz (SC-629-G)
AKT	Active PI3K signaling	1F	CST (5186)
cJUN	Transcription factor	1F	Abcam (AB193780)
cMYC	Signaling transducer	1F	Abcam (AB201780)
E2F1	Signaling transducer	1F	Abcam (AB208078)
FOXA2	Transcription factor	1F	Thermo (50-4778-80)
FOXO1a	Transcription factor	1F	Abcam (AB207244)
HSP90b	molecular chaperon	1F	Abcam (AB115641)

KAP1/TIF1 β	molecular chaperon	1F	BioLegend (619304)
MITF	Signaling transducer	1F	Abcam (AB201675)
mTOR	Active PI3K signaling	1F	CST (5048)
NFATc1 (NFAT2	Transcription factor	1F	BioLegend (649605)
NFkB	Transcription factor	1F	Abcam (AB190589)
Pax8	Transcription factor	1F	Abcam (AB191870)
pERK (T202/Y204)	Active MAPK signaling	1F	CST (14095)
p-Tyrosine (pTyr)	Signaling transducer	1F	CST (9415)
pmTOR	Active PI3K signaling	1F	Thermo (50-9718-41)
pNDRG1	Signaling transducer	1F	CST (6992)
pSTAT3	Transcription factor	1F	CST (4324), CST (9145)
S100a	Signaling transducer	1F	Abcam(ab207367)
S100A4	Signaling transducer	1F	Abcam (AB196168)
STAT3	Signaling transducer	1F	CST (14047)
Survivin	Signaling transducer	1F	CST (2810)
Ecad	Epithelial cell junction	1G	CST (3199)
E-Cadherin	Epithelial cell junction	1G	Abcam(ab201499)
EGFR	Receptor tyrosine kinase	1G	CST (5616), Abcam(ab192982)
EpCAM	Epithelial cell junction	1G	CST (5198), BioLegend (324205)
HER2	Oncoprotein	1G	Dako(A0485)
MET	Receptor tyrosine kinase	1G	CST (8494)
N-Cadherin	Epithelial cell junction	1G	BioLegend (350809)
NGFR	Receptor tyrosine kinase	1G	Abcam (AB195180)
pEGFR (Tyr1068)	Receptor tyrosine kinase	1G	CST (3777)
pERK	Active MAPK signaling	1G	CST (4370)
Lamin		2A	CST (8617)
NUP98	Nuclear pore complex	2A	CST (13393)
Actin	Cytoskeleton	2B	CST (8046)
GFAP	Cytoskeleton (neural cell)	2B	Thermo (41-9892-80)
MAP2	Cytoskeleton (neural cell)	2B	Thermo (41-9763-80)
TUBB3	Cytoskeleton (neural cell)	2B	BioLegend (657405)
Tubulin	Cytoskeleton	2B	CST (3624)
Vinculin	Cytoskeleton	2B	Thermo (41-9777-80)
COXIV	Mitochondria	2C	CST (7561)
LAMP2	Lysosomal membrane	2C	Abcam(ab25631)
TPCN2	Lysosomal membrane	2C	NOVUSBIO(NBP1-86923)
CD2	T-cell, NK cell	3A	Abcam(ab37212)

CD20	B-cell	3A	Thermo (50-0202-80), Dako(M0755)
CD3	T-cell	3A	Abcam(ab208514), Dako(A0452)
CD4	T-cell	3A	Thermo (41-2444-80), Dako(M7310)
CD45	Immune cell	3A	R&D(FAB1430P-100), BioLegend (304020)
CD45RO	Memory T Cell	3A	Thermo (41-0452-80), Dako(M0742)
CD8	T-cell	3A	Thermo (50-0008-80), Dako(M7103)
FOXP3	Regulatory T-cell	3A	Thermo (41-4777-80), BioLegend (320113)
PD1	Immune check point	3A	CST (86163), Abcam(ab137132)
CD11b (ITGAM)	Transmembrane receptor	3B	Abcam(ab133357)
CD11c (ITGAX)	Transmembrane receptor	3B	Thermo (41-9761-80)
MHCI	antigen presentation	3B	Abcam(ab199837)
MHCII	antigen presentation	3B	Abcam(ab201347)
PD-L1	Immune check point	3B	CST (15005), Abcam (AB213358)
αSMA	Cytoskeleton	4A	Thermo (50-9760-80)
FAP	Transmembrane receptor	4A	Thermo (BMS168)
β-Catenin	Cell adhesion	4A	CST (4627)
EDN1	Endothelial cell	4B	Abcam (AB2786)
VEGFR1	Endothelial cell	4B	Santa Cruz (SC-31173)
VEGFR2	Endothelial cell	4B	CST (12872)
Collagen IV	Basement membrane	4C	Thermo (51-9871-80)
Vimentin	Cytoskeleton	4C	CST (9855), BioLegend (677807)

2.2.3.3 TILs

TILs are thought to be an important player in prognosis and therapy. TILs have been reported to be positively associated with improved survival in TNBC and other cancers^{68,73,157–161}, while the rates of response to immune checkpoint inhibitors are lower in HR+ BC, even when combined with molecular targeting agents^{123,124}. A successful antitumor immune response requires several key steps, which are (1) the brief capture of tumor antigens by APCs, (2) presentation of tumor antigens to lymphocytes, (3) activation and expansion of CD4+ and CD8+ T lymphocytes, (4) direct cell-to-cell contact between activated lymphocytes and tumor cells, and (5) production of inflammatory cytokines (CD4+), cells that kill tumor cells (CD8+), B cells, natural killer (NK) cells, and macrophages^{162–164}. However, suppression of tumor growth by CD8+ cytotoxic T cells and Nkp46+ NK cells involves a complex series of direct and indirect interactions between immune cells. Negative regulation through immune check-point proteins— such as PD1, PD-L1, and FoxP3—occurs in 30–50% of breast cancers and limits CD4+ T cell functions to promote CD8+ T cell activation within tumor microenvironments¹⁶⁵. Additionally, the tumor cells and their microenvironment often manipulate the successful antitumor immune response through immune checkpoint proteins. The following biomarkers: CD4, CD8, PD1, FoxP3, CD20, and CD68, can be used to characterize the composition of immune cells in TILs as a minimum set of immune cell

markers. The analysis of proximity from those immune cells to tumor boundary, or directly to PCNA/Ki67 positive proliferating tumor cells enables mapping of the spatial organization of tumor cells, immune cells, and cells in the tumor microenvironment.

2.2.3.4 Angiogenesis

Angiogenesis is another hallmark of malignancy, and clinical benefits of antiangiogenic agents have been reported in colorectal cancer, non-small cell lung carcinoma^{166,167}, and renal cell carcinoma¹⁶⁸. Preclinical studies have confirmed that angiogenesis also plays a central role in breast cancer carcinogenesis and metastatic potential^{169–172}. However, the benefit of antiangiogenic therapy with bevacizumab or protein kinase inhibitors, such as sunitinib, has been controversial because most trials of antiangiogenic agents in breast cancer have reported improved response rate and progression-free survival, but no increase in overall survival compared to the patients who received chemotherapy alone¹⁷³. Quantitative measures to assess mature and immature angiogenesis by examining endothelial markers (CD31 and CD34) and their pericyte marker (α SMA), the spatial distribution of hypoxic tumor cells expressing HIF1a or GLUT1, and those cells' proximity to mature blood vessels, to proliferating tumor cells, and to each other are required for the stratification of responses to antiangiogenic therapy. The antiangiogenic control via bevacizumab or sunitinib enhances the efficacy of cytotoxic agents (nab-paclitaxel, anthracycline, and cyclophosphamide) against tumor cells, and also normalizes cancer-associated vasculature (i.e., vascular normalization), which reduces vascular hotspots, tumor microvessel density (MVD), and immature vessel distribution^{174–176}. The density, maturity, and formation of immature lymphatic vessels may also be involved in the mechanisms of tumor resistance to treatment in trials with bevacizumab, of which these vessels' synergistic or adverse effects are as of yet unclear. Therefore, characterization of lymphatic vessel distribution using lymphatic endothelial markers (PDPN and LYVE1) and α SMA is required. Increases in lymphatic vessel promotion in the S0800 trial may also induce regional lymph-node or distant organ metastasis through the lymphatic system¹⁷⁷.

2.2.3.5 Cancer-Associated Fibroblasts (CAFs)

Cancer-associated fibroblasts (CAFs) are the most frequent component of tumor stroma, especially in breast and pancreatic cancer^{178,179}. According to the seed-and-soil hypothesis, a permissive microenvironment (soil) facilitates growth of cancer cells (seed) in extended primary and metastatic sites, which interact with not only cancer cells, but also with TILs and endothelial cells in the tumor stroma¹⁸⁰. CAFs may induce therapeutic resistances against endocrine therapy¹⁸¹, chemotherapy¹⁸², and molecular target therapy¹⁸³. The tumor microenvironment has been demonstrated to promote breast cancer initiation, growth, migration, metastasis, and therapeutic resistances. CAFs, TAMs, EC, CAAs, leukocytes, etc. are critical components of tumor stroma, comprising the majority of tumor microenvironment and taking part in induction of breast cancer malignancy through various mechanisms¹⁸⁴. The characteristics of CAFs are not yet fully understood; they may potentially be transformed resident fibroblasts, epithelial-mesenchymal-transformed tumor cells, bone marrow-derived mesenchymal stem cells, or other cells in stroma, and should be characterized in order to identify the composition of the tumor microenvironment, and their individual roles therein. At present, no evidence suggests a single dominant origin of CAFs in the tumor microenvironment. Biomarkers considered acceptable for reliably identifying CAFs include α SMA, FSP-1, FAP, PDGFRa, PDGFRb, and vimentin^{179,185}. Some new related therapeutic targets, such as FAP, have been developed and are currently under preclinical

evaluation^{178,179}. These targets may potentially augment the existing therapeutic options for cancer treatment and prevention. Several studies confirmed the correlation between MVD and prognosis^{186,187}, inflammatory response¹⁸⁸, and immune suppression. Angiogenic factors such as VEGF released by tumor cells and CAFs inhibit dendritic cell (DC) maturation and proinflammatory cytokine production, and promote Treg infiltration. Tumor blood vessels are often functionally compromised and often do not express receptors required for T-cell extravasation into tissues, such as ICAM-1 and VCAM-1, which precludes effective infiltration of cytotoxic T-cells into tumors¹⁸⁹.

2.2.3.6 Disease Specific Biomarkers

Disease specific biomarkers for tumors and normal tissues in breast, pancreas, and prostate cancers are listed in the Table 2.1. The markers can be synergistically used with others for characterization, subtyping, and further classification of each cancer tissue^{190,191}. These biomarkers are frequently used for histopathological diagnosis in routine practices of breast, pancreas, and prostate pathology. In breast cancer pathology, it is important to characterize stromal invasion, columnar cell lesions, intraductal proliferations, papillary lesions, sclerosing lesions, spindle cell lesions, and prognostic and predictive factors, in addition to molecular phenotypes such as luminal A, B, TNBC (basal and claudin-low), and HER2. In pancreatic cancer pathology, the most widely used biomarkers for the diagnosis of ductal adenocarcinoma are pVHL, maspin, S100P, and IMP-3, and for the diagnosis of pancreatic endocrine neoplasm are synaptophysin, chromogranin, NSE, and beta-catenin. Poor prognostic markers in advanced pancreatic adenocarcinoma are overexpressions of DNMT1, HDAC1, CD166, and nuclear S100A6, and those in endocrine neoplasm are CK19, Ki67, CD99, topoisomerase IIa, and nuclear survivin. In prostate cancer, the recommended biomarkers to distinguish seminal vesicles from prostatic ductal adenocarcinoma and high-grade prostatic intraepithelial neoplasia are CK7, PAX2, and MUC6. CK7 and CK20 (epithelial markers) are negative in both normal and neoplastic prostatic epithelia, MUC6 is positive in seminal vesicles but negative in normal prostate and adenocarcinoma, and nuclear PAX2 is positive in seminal vesicles, negative in adenocarcinoma, and high-grade prostatic intraepithelial neoplasia. Cancers originating from these three organs are histopathologically and genetically more complex than others, which may be one of the reasons that results of immune checkpoint inhibition in clinical trials are mixed.

2.3 MATERIALS

2.3.1 *Tissue Processing*

1. Temperature control chamber or incubator.
2. Xylene (Laboratory Grade, Fisher Chemical, X4).
3. Ethanol 200 PROOF (Decon Lab).
4. Glass Coplin staining jars or staining station (Sakura-Tek).
5. Programmable pressure cooker (Biocare Medical Decloaking Chamber/Dako Pascal*).

*Discontinued.

2.3.2 *Blocking and Staining*

1. Bovine Serum albumin (Sigma, A7906).
2. Phosphate buffered saline, 10× solution (Fisher, BP399).
3. Normal goat, rabbit, mouse sera (Vector lab, S-1000 Thermo, 10410 and 08-6199).
4. Biomarkers and corresponding antibodies are in the Table 2.1 and Table 2.2.
5. Staining cover film (Grace Bio-Lab).
6. Cover glass #1.5 thickness, 24 × 30 mm and 24 × 50 mm (Corning Life Sciences, 2980-243 and 2980-245).
7. Antifade Mountant with DAPI (Thermo, S36938).

2.3.3 *Antibody Labeling (Optional)*

1. Labeling kit (100 µg scale) for Alexa Fluor 488, 555, 647, and 750 (Thermo, A20181, A20187, A20186, and S30046).
2. Labeling kit (10–20 µg scale) for Alexa Fluor 488, 555, and 647 (Thermo, A10468, A10470, and A10475).
3. For larger scale, NHS Ester form of Alexa Fluor 488, 555, 647, and 750 (Thermo, A20000, A20009, A20006, and A20011).
4. Melon™ Gel IgG Spin Purification Kit (Thermo, 45206).
5. Antibody purification by buffer exchange (Millipore, Amicon Ultra-0.5 Centrifugal Filter Unit [10 KDa], UFC501096).

2.3.4 *Image Acquisition*

1. Axio Imager.M2 equipped with motorized stage and filter turret (Zeiss, Germany).
2. Axio Scan.Z1 slide scan system (Zeiss, Germany).
3. Colibri7 LED with Cy7 line (Zeiss, Germany).
4. Filter sets for DAPI, GFP, Cy3, Cy5, and Cy7 (Zeiss, Germany).
5. Plan-Apochromat 20×/0.8 objective (Zeiss, Germany).
6. Plan Apochromat 40×/0.95 objective (Zeiss, Germany).
7. Orca Flash 4.0 V3 CMOS Camera (Hamamatsu Photonics, Japan).
8. Zen Blue edition imaging software (Zeiss, Germany).

2.3.5 *Image Processing and Analysis*

1. Zen Blue image processing software (Zeiss, Germany).

2. Fiji image processing Package with ImageJ (Open source: <https://fiji.sc/>).
3. CellProfiler cell image analysis software (Open source: <https://cellprofiler.org/>).
4. Matlab software (MathWorks, <https://www.mathworks.com/>).
5. Qi Tissue image processing and visualization software (Quantitative Imaging System, PA).
6. Windows work station (HP Z840 or Dell Precision 5820 with 8 core Xeon CPU, 16 GB Graphic board, and 256 GBRDIMM).

2.4 METHODS

2.4.1 Overview

In recent years, methods of simultaneously visualizing multiple biomarkers that have become popular in biomedical research, including brightfield multiplexing technology with sequential staining^{105,192,193} (see Note 1, section 2.5), fluorescent multiplexing with fluorophores and nanocrystals, and combination of unmixing technology using multispectral acquisition system to image enzymatic amplification of fluorescent signals¹⁹⁴. However, unlike previous methods, the newly developed method of cyclic multiplexed- immunofluorescence (cmIF) allows iterative rounds of multiplex immunofluorescence staining using directly labeled antibodies without the need for incubation with secondary antibodies or enzymatic amplification. Instead, cmIF only requires fluorophore quenching to achieve highly multiplexed staining of up to 60 protein biomarkers in a single tissue section^{107,116}. The protocol has been optimized for formalin-fixed paraffin-embedded (FFPE) tissue samples, and a workflow has been developed to process digital images for quantitative single-cell based spatial analysis. The cmIF assay may enter routine use in clinical settings as this technique becomes more widely available in the future. Some potential issues with the cmIF assay are that there is a limit to the number of biomarkers that can be interrogated, and that the procedure is both labor-intensive and time consuming. To overcome these disadvantages, a more robust method of cyclic IF using oligonucleotide-conjugated antibodies is under development (see Note 2, section 2.5).

2.4.2 Staining

2.4.2.1 Tissue Preparation

Tumor tissues are fixed in 10% neutral buffered formalin or 4% formaldehyde in PBS (pH 7). To preserve phosphorylated proteins and prevent tissue autolysis and degradation, the immediate fixation of the tissues with a refrigerated fixative is critical; however, it is also important to maintain cell morphology and tissue architecture over the course of cmIF protocol. The optimal fixation time varies between 6 and 48 hour(s), depending on the tissue size, staining methods, and antibodies being used to target proteins. However, a fixation step that lasts longer than 48 h is not recommended, as overfixation may present difficulties in antigen retrieval following staining. After fixation, the tissues are gently dehydrated in increasing grades of ethanol to minimize cell damage, processed through xylene to clear the tissues of ethanol, embedded in heated paraffin at 60 °C, and then hardened overnight to form tissue blocks. Finally, the tissues are sectioned at a thickness of 4–5 µm using a microtome, placed on slides that are silanized or otherwise coated with a suitable adhesive, and dried at room temperature. The slides may optionally be baked at 55–65 °C for 30–120 min in order to minimize tissue loss during the staining procedure. The sectioned tissues should be stored at 4 °C to prevent tissue oxidation, which may cause the loss of antigenicity.

in tissues. For long-term storage of sections without loss of antigenicity, the slides can be dipped in paraffin to cover the surface of the tissue sections (see Note 3, section 2.5).

2.4.2.2 Antigen Retrieval

The slides should be baked prior to the start of tissue processing in preparation for staining, if they have not been baked already. The paraffin must be completely removed from the tissue, which should then be rehydrated to allow the efficient bonding of antigens and antibodies during the staining steps. Antigen retrieval is necessary for successful immunohistochemical staining of the FFPE tissues, as fixation causes protein-protein cross-linking and masking the epitopes that are specific antibody binding targets. Heat-mediated antigen retrieval (by microwave, pressure cooker, or autoclave) in either citrate buffer (pH 6) or Tris EDTA buffer (pH 9) is the widely accepted standard procedure in most research and pathology laboratories. However, certain antibodies may require processing at lower temperatures, the use of buffers with higher or lower pH, or enzyme-mediated antigen retrieval with the use of proteinase K, trypsin, or pepsin. Antigen retrieval methods that involve enzymatic pretreatment are more difficult to control compared to methods of heat-mediated antigen retrieval. As a large number of antibodies (typically 20–40) will be used in cyclic immunofluorescence staining, heat-mediated antigen retrieval in citrate buffer or Tris EDTA buffer, or both buffers sequentially, is most recommended. A commercial buffer for antigen retrieval will also produce satisfactory staining results. The antigen retrieval process should only be applied once for cyclic immunofluorescence before beginning the staining procedure.

2.4.2.3 Antibody Selection

There are a number of commercially available primary antibodies that are directly conjugated to N-hydroxysuccinimide (NHS) ester fluorophores. These fluorophores are often Cyanine dyes or similar alternatives, which randomly label lysine residues on antibodies. The desired primary antibodies can be labeled in-house by using NHS-activated dyes or a commercial labeling kit. Several vendors provide a variety of protein- and antibody-labeling reagents, including Cy2, Cy3, Cy5, and, Alexa Fluor (AF). Hydrogen peroxide inactivation of fluorescence, as required for cyclic immunofluorescence, is difficult in certain fluorophores, such as FITC or AF546. The processes of validating antibodies for biomarker expression specificity and selecting fluorophores for labeling are the two most important steps in cmIF. The use of directly labeled antibodies provides greater flexibility in multiplex immunofluorescence, especially in cmIF, because there is greater freedom in selecting the most optimal combination of antibodies to examine. Additionally, directly labeled antibodies are suitable for quantitative analysis of target protein expression because there is no signal amplification in the procedure. The intensity of fluorescence has a linear relationship to the target protein's expression in stained tissues, so differences in protein expression can therefore be estimated by the difference between fluorescence intensities under constant staining and imaging conditions. However, the use of directly labeled antibodies may compromise the signal-to-noise ratio (SNR), so careful antibody selection for each target protein is one of the most important key factors to obtaining successful results from cmIF. Many labeling kits with Alexa Fluor and their analogs, including all components for antibody labeling, such as a purification column and buffer, are commercially available (see Subheading 2.3 in Materials and Note 4 in subheading 2.5). In-house antibody labeling with the desired fluorophore can also be done if it is not available on the market.

2.4.2.4 Antibody Validation

The validation of antibodies for specificity, reproducibility, and durability is an essential step for the cmIF (Figure 2.2). However, this is sometimes difficult, time-consuming, and costly. Directly labeled antibodies often compromise staining results compared to immunofluorescence staining with secondary antibodies, due to the smaller number of fluorophore molecules representing target proteins and the loss of antibody affinity to epitopes in proteins. These problems may be exacerbated when the selected antibodies are confirmation-specific or the target proteins contain fragile epitopes. Positive and negative control tissues and cell lines are also important for antibody validation, and knockout, knockdown, or knock-in cell lines may need to be established for this process. A variety of reference images for immunohistochemistry and immunofluorescence staining can be found in The Human Protein Atlas (<https://www.proteinatlas.org/humanproteome/tissue/breast>).

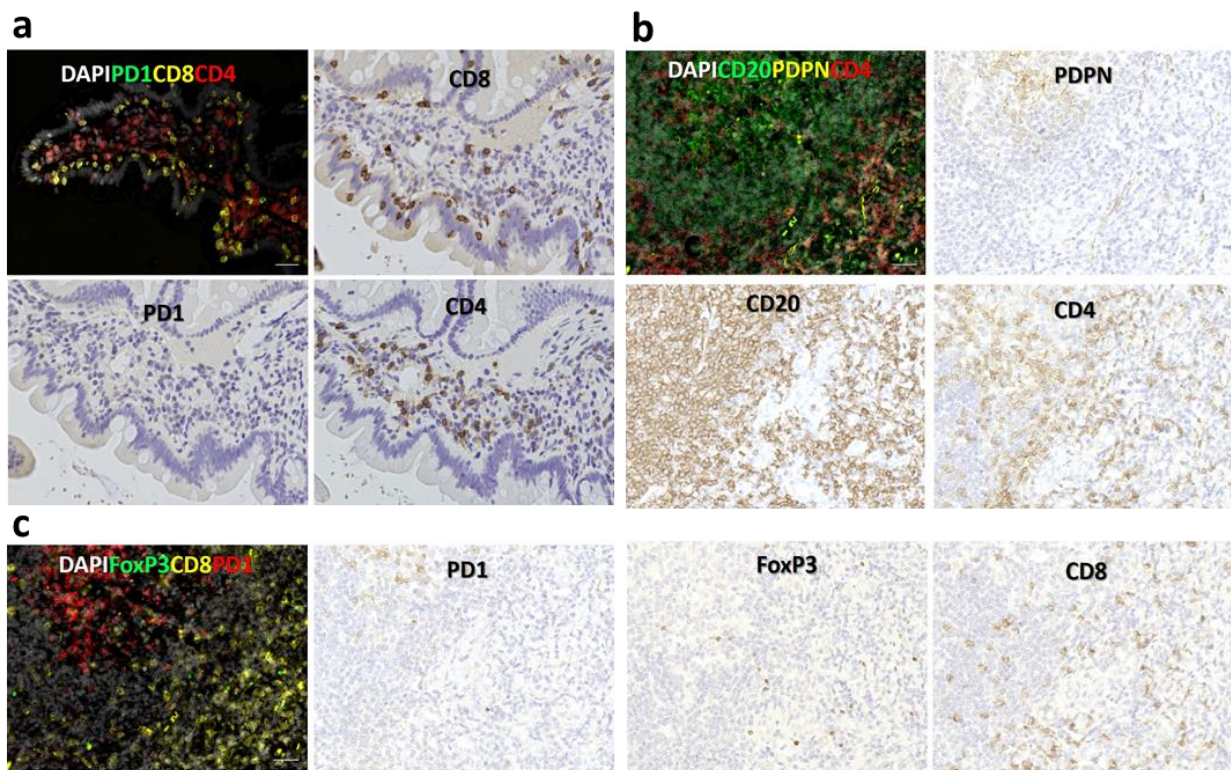


Figure 2.2. Antibody Validation in cmIF.

(a) Validation of antibodies for immune cells in cmIF and IHC using the same antibodies on adjacent normal jejunum tissues. The spatial distribution of CD4 T cells (red in cmIF image) differs from that of CD8 T cells (yellow cmIF). CD8 T cells disperse in intestinal villus epithelial cells, while CD4 T cells are present in the stroma region. PD1+ T cells are very rare in normal jejunum tissue (only one green cell is observed in cmIF, which has been confirmed in IHC as well). (b) Comparison of immune cell composition between cmIF and IHC, using the same antibodies on adjacent tonsil tissues. The spatial distribution of CD4 T cells (red in cmIF) and that of CD20 B cells (green in cmIF) appear mutually exclusive, in which CD4+ T cells are distributed in the peri-germinal center region and CD20+ B cells in the germinal center. PDPN+ T cells (yellow in cmIF) are dispersed throughout the tonsil tissue. (c) Comparison of immune cell composition between cmIF and IHC, using the same antibodies on adjacent tonsil tissues. The spatial distribution of CD8 T cells (yellow in cmIF image) shows a uniform pattern in the ROI, peri-germinal center region. PD1+ T cells (red in cmIF) and FoxP3+ Treg cells (green in cmIF) are dispersed throughout the tonsil tissue.

2.4.2.5 Staining and Acquisition

After antigen retrieval, the tissues are incubated in blocking solution with either bovine serum albumin alone or combined with other sera. The tissues are incubated in the directly conjugated primary antibodies for 4–16 h, washed in PBS, stained with DAPI for nucleus imaging, and mounted in antifade mounting media. Immunofluorescent microscopy relies on the detection of light emitted at specific wavelengths from each fluorophore directly bound to a primary antibody when excited by light of another specific wavelength. The emitted light is always of a longer wavelength than the excitation light. This difference between the band maxima of excitation and emission wavelengths, which is known as Stokes shift, is a unique property intrinsic to each individual fluorophore. Selection of an appropriate combination of fluorophores, target antigens, and primary antibodies, and maximizing spectra from target antigen channels with minimal bleeding spectra from others by setting multiband path microscopy filters are factors that are critical to effective cmIF. The optimal components for obtaining high quality digital images with high SNR and dynamic range for the regions of interest (or full tissue sections) from fluorescence microscopy systems, which are systems for which research and pathology laboratories most commonly use widefield epifluorescence microscopes, are an LED light source, a 20× objective with high numerical aperture, an X-Y motorized stage, a 16-bit grayscale CCD or CMOS camera, and multichannel tile-imaging capabilities with autofocus functionality. One of the most promising technologies that has emerged in the last several years for illumination in optical microscopy is the light-emitting diode (LED). Recent advances in illumination technology to expand the LED spectrum to include infrared red enable imaging of signals emitted by Cy7/ AF750 and similar dyes. This ability to use Cy7/AF750 makes 5-channel imaging possible. Most major microscope manufacturers generally make such high-end digital imaging systems available for purchase.

2.4.2.6 Fluorescence Inactivation

After image acquisition, the coverslip is gently removed in PBS and the fluorophores are chemically inactivated by horizontal incubation of slides under incandescence light. Following fluorophore inactivation, slides are washed in PBS. Tissue bleaching is then visually confirmed under a microscope, the next round of primary antibodies is applied, and the cycle of staining, imaging, and bleaching is repeated (Figure 2.3). The antibodies from previous rounds of staining are never stripped off from the tissues at any step.

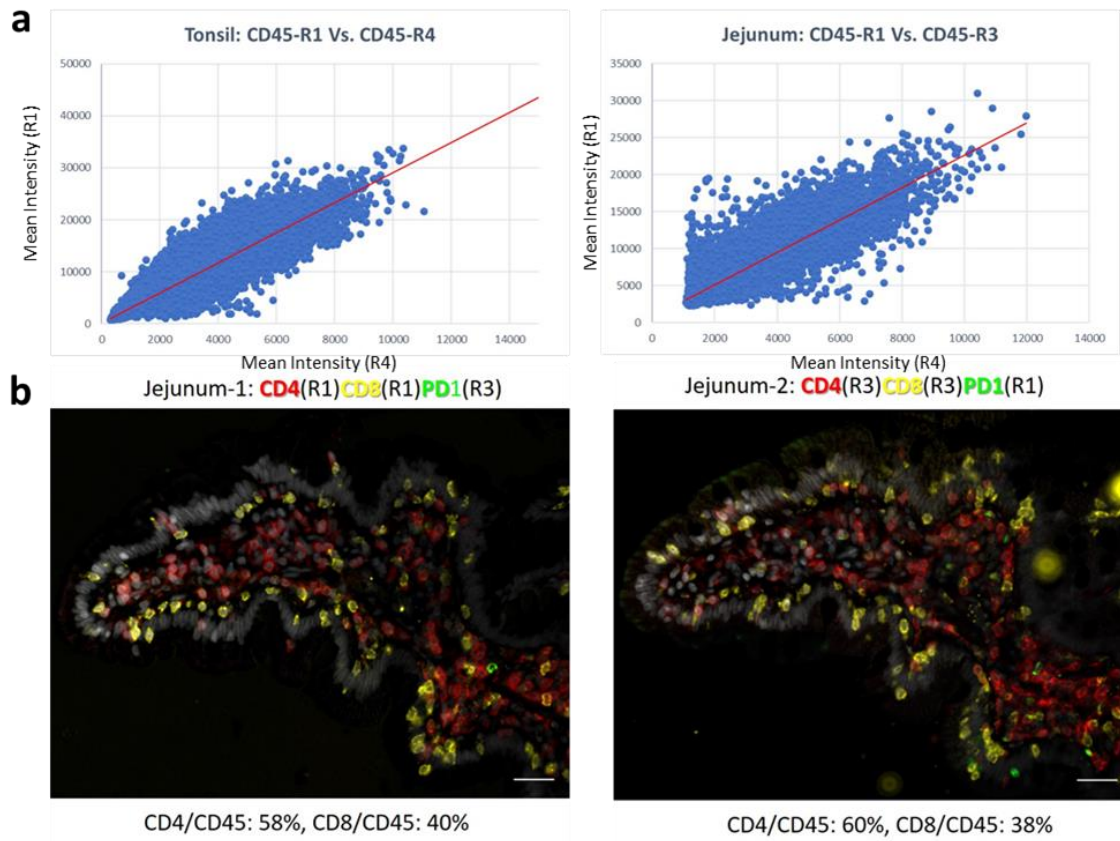


Figure 2.3. Reproducibility in cmIF.

(a) Pearson's correlation (CD45 between different rounds). CD45 was stained twice within the same tonsil tissues (Round1 and Round4) and the same normal jejunum tissues (Round1 and Round3). Scatter plots of mean intensity for each tissue show very high correlations (i.e., high reproducibility between two repeated CD45 stains in the same tissue). The Pearson's correlation coefficients for tonsil and jejunum tissues are 0.947 (209,369 cells) and 0.915 (75,879 cells), respectively. Mean intensity units are pixel intensity in 16-bit grayscale images. (b) Two adjacent normal jejunum tissues were each stained for CD45, CD4, CD8, and PD1 in differing orders. The proportions of CD4/CD45 and CD8/CD45 are 58% and 40% in Jejunum-1 tissue, and 60% and 38% in Jejunum-2 tissue, respectively, which demonstrate reproducibility of cellular composition across different rounds of cmIF with differing staining orders. Scale bar = 50 μ m.

2.4.3 Image Processing

2.4.3.1 Overview

The acquired images should be stored in single- or multichannel grayscale Tagged Image File Format (TIFF), which is a flexible and adaptable file format for handling images and data within a single file, and can include header tags (size, definition, etc.). Image files in TIFF format can be reviewed, edited, and analyzed in ImageJ, which is an open-source software program. The majority of manufacturer- provided imaging software included with microscopes or image acquisition cameras are capable of saving or exporting the acquired images in TIFF format. It is recommended that the TIFF images be organized by staining cycle, or "round" number (R), and the channel number (C) within each cycle (e.g., R1C1 is the nuclear DAPI image from the first cycle). For

single cell-based analysis using image data generated from cmIF, the image registration is the first step of the analysis process.

2.4.3.2 Registration

Registration transforms different sets of image data into a single coordinate system. Typically, all images are registered to the same coordinates in the first DAPI image (i.e., the first DAPI is the registration reference). The other channel images in the same cycle must be aligned with the corresponding DAPI image, regardless of whether they are saved as individual TIFFs or in a stacked TIFF file. The image registration functionality is available through multiple plugins (StackReg or MultiStackReg) for ImageJ [<http://brad-busse.net/downloads.html>], or from the Computer Vision System Toolbox in Matlab (MathWorks) such as “matchFeatures” and “estimateGeometricTransform.” These functions can be used to automatically identify nuclear features in each DAPI image, such as size, shape, and texture^{195,196} as well as to identify matching features between two DAPI images and find an affine transformation matrix to align those matched features. Once DAPI is registered, the transformation matrix is applied to the biomarker channel images in each respective round. Other algorithms for image registration, such as “wavelet based automatic registration” and “Deep learning-based registration” have been reported and may be of potential use^{197,198}. Once the images have been registered, overlays of biomarkers from any round can be visualized (Fig. 2.4a, b).

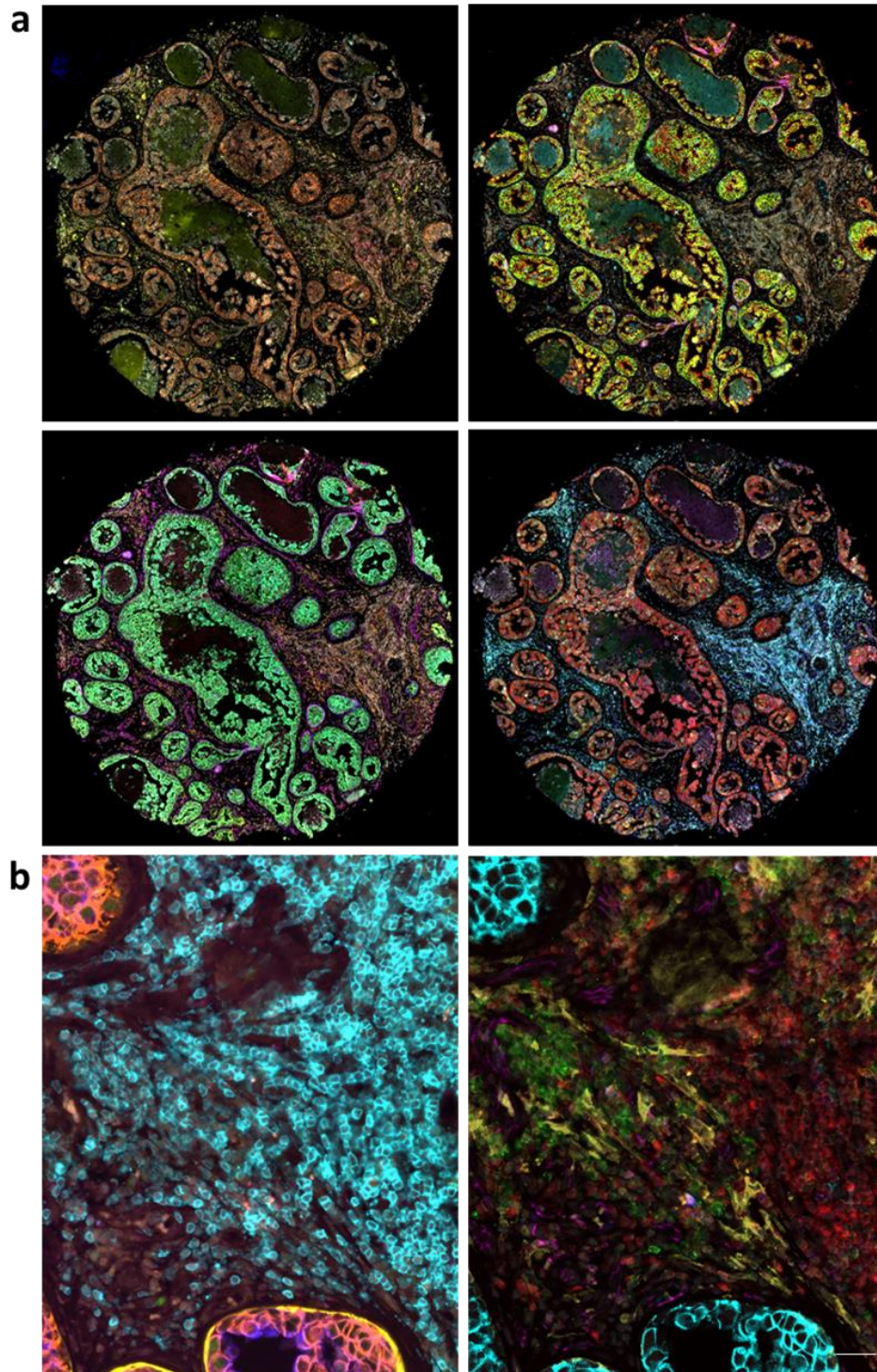


Figure 2.4. Visualization in cmIF.

(a) The four panels show four image overlays of the same tissue: HER2+ TMA Core #6 displayed with immune cell markers (upper left), subtyping markers (upper right), differentiation markers (bottom left), and proliferation markers (bottom right). The TMA was stained for 27 biomarkers through 8-Round cmIF. Biomarkers in each panel and their corresponding pseudocolors are the following: Immune: CD20 (Green), CD8 (Yellow), CD4 (Red), CD68 (navy), PD-1 (Purple), E-cad (Orange). Subtype: HER2 (Green), ER (Navy), CK19 (Red), CK14 (Purple), CD44 (Orange), E-Cad (Yellow). Differentiation: CK19 (Green), CK5 (Red), CK7 (Cyan), CK14 (Navy), CD44 (Orange), CK8

(Yellow), VIM (Purple). Proliferation: PCNA (Green), Ki67 (Red), HER2 (Yellow), H3K27me3 (Navy), CoxIV (Orange), CD45 (Cyan). (b) Two composition images at high magnification. HER2B Core #6—Immune cell-enriched region in tumor microenvironment. The left panel shows PCNA, Ki67, CK19, CK14, HER2, and ER positive cells each in different colors, and the right panel shows CD20, CD8, CD4, CD68, PD1, FoxP3, CD31, and PDPN positive cells. Immune cells are defined in cyan in the left panel (i.e., CD45), and their subset cells, such as CD20, CD8, and CD4, are in green, yellow, and red, respectively in the right panel. The two panels are of the same location and at the same magnification from Core #6 (The video is available at <https://youtu.be/KbBKAcqIMhI>). Scale bar = 50 μ m.

2.4.3.3 Segmentation

Segmentation is the next important step for feature extraction among individual cells. Each individual nucleus is detected and segmented using alternating sequential filters, OTSU thresholding, and seeded Watershed algorithms in CellProfiler¹⁹⁹ or ImageJ²⁰⁰, which can also be found in Matlab or Python²⁰¹. In the segmentation process, single and multiple nuclear DAPI images are denoised and locally flattened using an alternating sequential filter, which is a mathematical morphology operator, in which alternating transformations of opening and closing are sequentially applied with structuring elements of increasing size. The resulting images are used in OTSU thresholding, creating a binarized image of nuclear masks. The masks are eroded with a mathematical morphology operator to separate touching nuclei, and are then individually labeled to serve as seeds in a watershed algorithm applied to the gradient of a single or multiple DAPI images^{202,203}. The centripetal expansion equation is applied to the nuclear segmentation results to obtain intensity values in the cytoplasmic region and plot the absolute distance from each nuclear boundary. More accurate subcellular segmentation may require channel images targeting specific biomarkers, such as: pan-cadherin, sodium–potassium ATPase, and CD45 as biomarkers for the plasma membrane; pan-cytokeratin and S6 ribosomal protein as biomarkers of the cytoplasm; and Cox IV and TOMM20 marking the mitochondria. Biomarkers for subcellular segmentation may be selected according to one's experimental needs or research interests, and the same algorithms used for nuclear segmentation can be used in subcellular segmentation as well. Following single cell segmentation, moments of each biomarker are extracted from the subcellular component of the segmentation mask as appropriate for the known subcellular localization of the target proteins.

2.4.4 Analysis

2.4.4.1 Overview

Serial digitized cmIF images will be processed using the computational image analysis workflow described above to align and visualize several markers simultaneously. The biomarker information from each single cell in an FFPE section that has been iteratively stained with a series of antibodies is transformed into a large spreadsheet, which contains mean intensity, pixel number, and x-y coordinates, among a number of other extracted features. Since the antigens in protein biomarkers are detected based on fluorophore- conjugated antibodies, the measured fluorescence intensity in the image is considered representative of the antigen quantity.

2.4.4.2 Clustering

The clustering of all cells based on the expression level of each biomarker (i.e., mean intensity of each single cell) is a widely accepted measure for classifying cells with similar biological phenotypes, such as HR+ and Ki67+ tumor cells. Figure 2.5a shows the result of K-means (k=10) clustering with 40 biomarkers to classify all cells from four TMA cores into ten groups, such as CD45 positive immune cells and CAFs in stroma. Each group is coded by a unique color, and the population of clustered cells is plotted on the bar graph in the middle.

2.4.4.3 Spatial Mapping

Cells classified in the same group are mapped to their original locations in the TMA cores on the right (Figure 2.5a). Immune cells and CAFs, represented respectively as pink and red, are distributed to four cores. Figure 2.5b and c show the plot of each cell after reducing high dimensionality by T-distributed stochastic neighbor embedding (tSNE) algorithm, which indicates that each core tends to carry a unique set of sample-based tumor cells, as well as a small population of shared tumor and other cells, common across the four TMA cores (Figure 2.5b). The majority of HER2+, ER+, and CD45+ immune cells are all mutually exclusive. Ki67+ is expressed within, or in close proximity to, HER2+ and ER+ cells (Figure 2.5c).

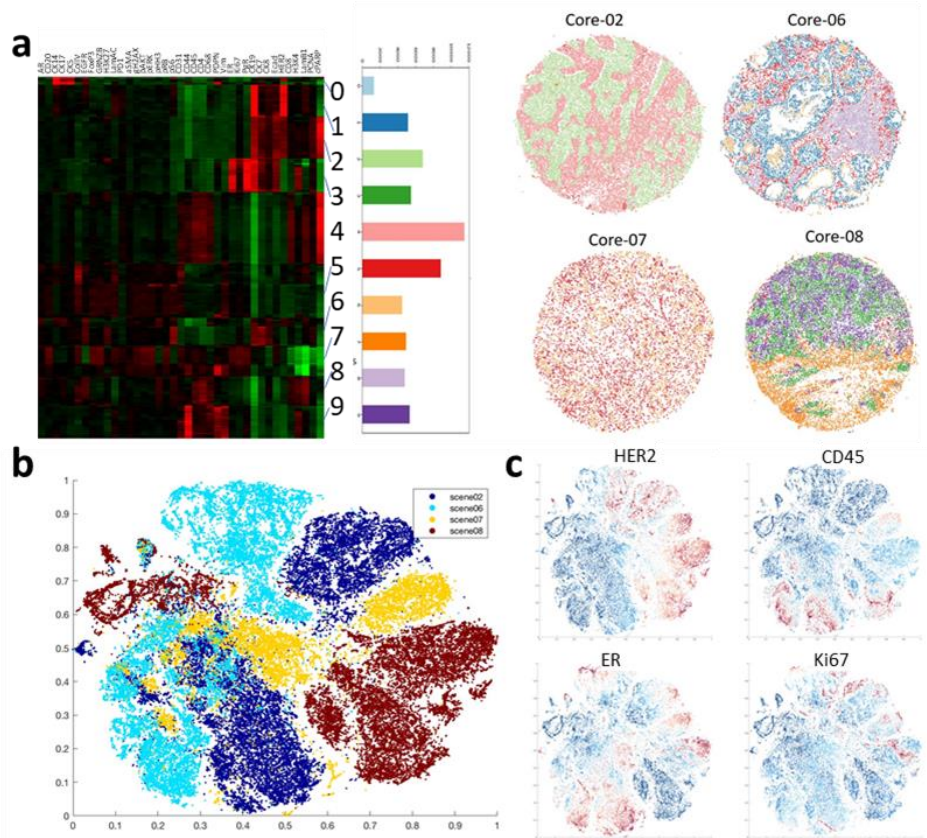


Figure 2.5. Unsupervised clustering analysis in cmIF.

(a) K-Means Clustering results are shown in the heat map on the left, representing four cores in Her2B-SD52 TMA. Ten classified cells are assigned in different colors, and the bar graph shows their respective proportions. Those cells are mapped into the original x-y coordinate with the same colors. (b, c) tSNE clustering and plotting for 4 cores in HER2+ TMA. Each dot represents one cell, and each core shows a different HER2 expression level. (c, left top). ER+ cells (Green) are dominantly present in Core #8, and express Ki67 (Orange). (c, left & right bottom). CD45+ immune cells (Pink) are present more in Cores #2 and #6, and less in Cores #7 and #8.

2.4.4.4 TILs Composition

To characterize immune contexture in each TMA core, the intensity information of DAPI, PCNA, Ki67, CD20, CD4, CD8, PD1, FoxP3, and CD31 are obtained from segmented single cell, which is then processed to analyze the immune cell composition and spatial information at a single-cell level. Tumor segmentation masks to distinguish intratumoral from peritumoral regions will be generated from the images of cell nuclei and Type I keratins (acidic, low MW) of the same tissue region as the tumor marker. Segmentation of tumor regions and other cell types will be validated

by visual assessment to exclude objects falsely annotated due to unspecific staining, background, or errors in annotation overlooked by the automated procedure described above.

2.4.4.5 Proximity Analysis

The distance from the tumor border and evaluated cellular components of the tumor microenvironment will be calculated by spatial proximity analysis. Extraction of spatial proximity and distance measurements will be processed using an automated workflow²⁰⁴. For instance, CD8+ T cell positional data can be used to measure distance from each individual T cell to the annotated tumor border. Using the Quickhull Algorithm for Convex Hulls, “dsearchn” function in MATLAB R2016, the shortest distance between T cell centroids and the boundary of the tumor mask can be measured to determine T cell distance to tumor boundaries. Similarly, the shortest distance between other cell types and the tumor border have been determined and visualized in Figure 2.6a, b. The spatial proximity analysis and visualization can reveal spatial interactions between infiltrating immune cell, tumor cells, and the vascular endothelium within BC subtypes, can identify tumor sub- populations with differences in proliferative activity and differentiation status, and can examine differences in tumor-immune cell proximity in BC.

2.4.4.6 Image Visualization and Analysis Software

OHSU Center for Spatial Systems Biomedicine (OCSSB) has collaborated with Quantitative Imaging Systems (Pittsburgh, PA) to develop a novel software package that seamlessly enables importing registered images and segmenting each nucleus and cell followed by feature extraction for analysis. The software (Qi Tissue, <http://www.qi-tissue.com>) uses Dynamic Visualization Engine (DyVE) for visualizing highly multiplexed fluorescence images (e.g., from 40+ channel cmIF) with multicolor rendering, blinking and height-map in 3D, which facilitate better understanding of biomarkers’ expression, colocalization and spatial distribution patterns. Tissues can be visualized as 3D landscapes where the height and visualization effects are modulated by image parameters of interest. The high speed of the image processing application provides us with instant access to the rich information available in images that are over 100 GB in size. The software package enables to store and retrieve combinations of channels, and export those visualizations in movie format.

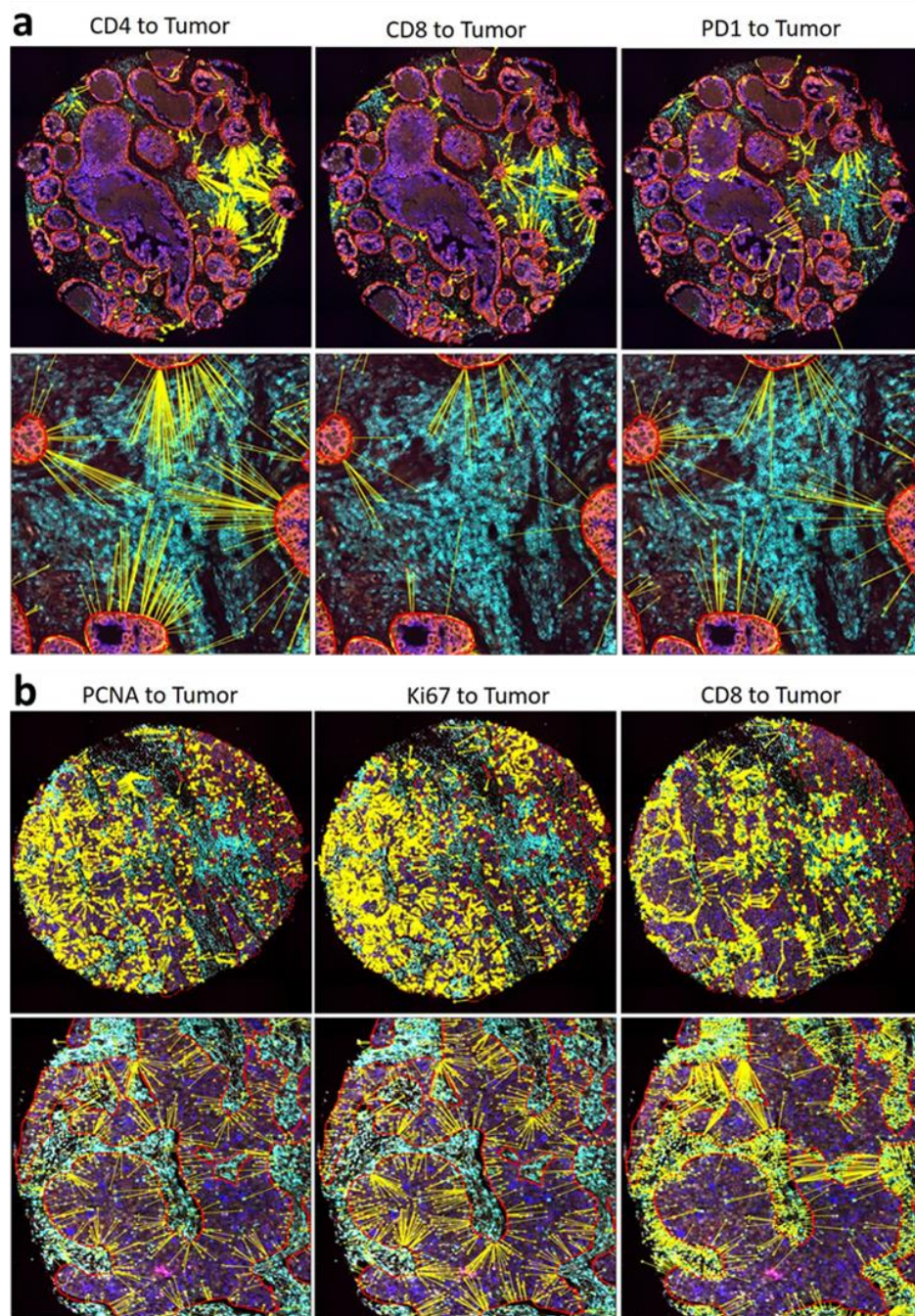


Figure 2.6. Cell proximity analysis in cmIF.

(a) He2B-SD52 TMA Core #6. Proximity analysis of immune cells (CD4, CD8, PD1 positive T cells) to the nearest tumor border in HER2+ breast cancer (Core 6). (b) HER2B-SD52 TMA Core #2. Proximity analysis of proliferating cells (PCNA and Ki67 positive) and CD8+ T cells to the nearest tumor border in HER2+ breast cancer.

2.4.5 *Protocols*

2.4.5.1 Cyclic IF Staining

1. Obtain 4–5 micron-thick sections on charged glass slides.
2. Bake slides overnight at 55 °C.
3. Bake slides for 30 min at 65 °C.
4. Using appropriate solvent resistant containers (such as glass Coplin jars, or Sakura Tissue-Tek containers) holding 40–250 mL of solvent depending on number of slides, place the slides through the following steps.
5. Deparaffinize slides:
 - (a) Place slides in xylene, three times for 5 min each.
 - (b) Place slides in 100% ethanol, two times for 5 min each.
 - (c) Place slides in 95% ethanol, two times for 2 min each.
 - (d) Place slides in 70% ethanol, two times for 2 min each.
 - (e) Place slides in dH₂O water, two times for 5 min each.
 - (f) Allow slides to air-dry (28 °C) for at least 5 min before proceeding.
6. Retrieve antigens:
 - (a) Programmable pressure cooker or heat-induced epitope- retrieval system (Biocare Medical Decloaking Chamber) to the following settings: • Set Point-1: 125 °C, 30 s. • Set Point-2: 90 °C, 30 s. • Set Point-3 (limit): 10 °C.
 - (b) Put 500 mL of dH₂O water in the pressure cooker, and center the heat shield in the bottom of the chamber.
 - (c) Fill three staining containers (use plastic Coplin jar for nine or fewer slides, or Tissue Tek staining container for 10–24 slides; do NOT use glass Coplin jars) with: • pH 6 Citrate buffer (10 mM citrate, Sigma C-1909), • dH₂O water, • Target Retrieval Solution, pH 9 (10X) (Agilent/Dako S2367).
 - (d) Place slides in the container filled with pH 6 buffer. [Fill blank slots with dummy slides to ensure even heating.]
 - (e) Place [the three] staining containers into the pressure cooker chamber.
 - (f) Power on the pressure cooker, and run through the two temperature settings programmed during Step 3.5.1.6(a).
 - (g) Release the pressure of the chamber when the program has ended, open the chamber lid, and turn off the power.
 - (h) Rinse slides in the hot dH₂O staining container.
 - (i) Place slides in the hot pH 9 Target Retrieval Solution staining container for 15 min.
 - (j) Rinse slides twice in room temperature dH₂O to cool.
7. Wash slides in PBS for 5 min.
8. Prequench samples to reduce autofluorescence:
 - (a) Add 10 mL quenching solution to each compartment of the plastic chamber.
 - 7.96 mL of ddH₂O
 - 1 mL of 10× PBS.
 - 40 µL of 5 M NaOH.
 - 1 mL of 30% H₂O₂—add last, right before using.

- (b) Place slides face-down in quenching solution. (c) Turn on incandescent lamp (60 W), and quench for 30 min. (d) Rinse in PBS, three times for 2 min. (e) Tap off excess PBS, and wipe around section.
9. Cover tissue section with blocking solution (10% NGS, 1% BSA, diluted in PBS).
 10. Allow blocking by covering tissue section with a plastic cover-slip and place in a humidity chamber for 30 min at room temperature.
 11. Remove blocking solution and coverslip by rinsing slide in PBS in a Coplin jar until cover slip floats off.
 12. Tap off excess PBS, and wipe around section.
 13. Prepare primary antibody solutions (5% NGS, 1% BSA, diluted in PBS).
 14. Prepare as many solutions as necessary to include all primary antibodies.
 15. Apply one primary antibody solution at a time, and repeat as many times as needed until all primary antibody solutions have been applied.
 - (a) Cover tissue section with primary antibody solution, and incubate in a humidity chamber protected from light for 2–4 h at room temperature, OR incubate overnight at 4 °C for stronger staining.
 - (b) Wash in PBS, three times for 5 min while protecting slides from light.
 - (c) Mount and scan slides.
 - (d) Quench tissue sections:
 - Immerse slides in PBS for up to 30 min to remove coverslip. (Allow coverslip to slide off naturally, without manually pushing or pulling it.)
 - Wash slides in PBS, three times for 2 min, to remove residual mounting media.
 - Add 10 mL quenching solution to each compartment of the plastic chamber.
 - 7.96 mL of ddH₂O.
 - 1 mL of 10× PBS.
 - 40 µL of 5 M NaOH.
 - 1 mL of 30% H₂O₂—add last, right before using.
 - Place slides face-down in quenching solution. • Turn on incandescent lamp (60W), and quench for 30 min.
 - Rinse in PBS, three times for 2 min.
 - Tap off excess PBS, and wipe around section.

2.4.5.2 Antibody Labeling

1. Remove sodium azide from antibody if necessary (i.e., if azide is greater than 0.02%), using a MilliporeAmicon Ultra 0.5 mL, 10 K spin column. (a) Add 100 µL PBS to the spin column, and centrifuge at 12.6 RCF for 2 min.
 - (b) Add antibody to column.
 - (c) Add PBS to bring the final volume of the solution to 500 µL, and at 12.6 RCF for 6 min.
 - (d) Discard flow-through. (e) Add 400 µL of PBS, and centrifuge at 12.6 RCF 6 min.
 - (f) Discard flow-through. (g) Add 400 µL of PBS, and centrifuge at 12.6 RCF 6 min. (h) Invert contents into a fresh 2 mL tube (provided with spin filters), and centrifuge at 3 RCF for 2 min.
2. Prepare dye as necessary.

- (a) Add DMSO to final concentration of 10 mM (e.g., AF750 and AF647, add 7.7 μ L, AF55 add 8 μ L, AF488 add 15.5 μ L) and pipette to mix.
3. Add antibody to 1 M NaHCO₃ at a 1:10 dilution, add PBS to bring the final volume to 100 μ L, and vortex briefly.
4. Add dye at a conjugation ratio of 1:10 to antibody solution, vortex briefly, and rock gently for 2 h, protected from light.
5. Exchange buffer to remove dye from antibody using Amicon Ultra 10KD spin column.
 - (a) Add 100 μ L PBS to the spin column, and centrifuge at 12.6 RCF for 2 min.
 - (b) Add antibody to column.
 - (c) Add PBS to bring the final volume of the solution to 500 μ L, and at 12.6 RCF for 6 min.
 - (e) Discard flow through.
 - (f) Add 400 μ L of PBS for a first wash and centrifuge at 12.6 RCF 6 min.
 - (g) Discard flow through.
 - (h) Add 400 μ L of PBS for a second wash, and centrifuge at 12.6 RCF 6 min.
 - (i) Invert contents into a fresh tube, and centrifuge at 3 RCF for 2 min.
6. Add 2% sodium azide to the final solution at a 1:100 dilution (resulting in a final sodium azide concentration of 0.02%).
7. Add 30% BSA (to the final solution at a 1:100 dilution (resulting in a final BSA concentration of 0.3%)).
8. Store final solution at 4 °C, protected from light.

2.4.5.3 Removal of BSA and Gelatin from Antibody Solution

Follow manufacturer's instructions for purification of IgG using Melon Gel IgG Spin Purification Kit (Thermo 45206).

2.4.5.4 Processing of Scanned Image for Analysis

1. All tiled images must be stitched in imaging software, then saved.
2. Multichannel images should be saved or exported in single- or multichannel Tiff format.
3. The images are subject to be registered in Matlab software.
4. The registered images can be overlaid in Zen Lite (Blue edition), ImageJ, and Qi Tissue software (see Note 5).
5. These software are available from following sites. Zen Lite (Blue edition, Zeiss): <https://www.zeiss.com/microscopy/us/products/microscope-software/zen-lite.html>. ImageJ and Plugins: <https://imagej.net/Fiji> as Fiji distribution package. QiTissue: contact Quantitative Imaging Systems ([http:// www.qi-tissue.com/](http://www.qi-tissue.com/)) to obtain a trial version or beta testing version.

2.4.5.5 Cellular Segmentation

1. Nuclear segmentation (Primary object) based on DNA DAPI image can be carried out using watershed and OTSU functions in ImageJ or CellProfiler.
2. Cellular segmentation (Secondary object) can be processed using nuclear segmentation results by expanding a fixed pixel distance.

2.4.5.6 Feature Extraction

1. Extract a common x and y coordinates in the corresponding primary and secondary objects as the centroid x and y positions.
2. Calculate mean intensity in the primary, the secondary and the region of the secondary excluding the primary object.
3. Spread sheets, including x, y, mean intensities in three regions for each cell, are created for analysis.

2.4.5.7 Analysis

1. Tumor cell composition and heterogeneity by thresholds and unsupervised clustering.
2. Immune cell composition and density by thresholds and KNN methods.
3. K-means clustering and spatial mapping of classified cells.
4. Nearest cell-to-cell, cell-to-blood vessel, and cell-to-tumor border proximity measurements.
5. Statistics for clinical associations.

2.5 NOTES

1. Chromogen-based multiplex immunohistochemistry. The Coussens laboratory at OHSU developed a related multiplex immunohistochemistry (mIHC) platform that enables deep auditing of immune contexture in situ in FFPE sections. Specifically, FFPE tissue sections are iteratively stained with panels of leukocyte lineage-specific and/or function-associated antibodies reflecting ~40 validated targets. An automated computational pipeline for image analysis enables classification of immune cell types using either image cytometry gating or sparse coding strategies [87, 94, 98]. Each stain is detected in a separate channel, and is imaged by digital scanning and bright-field microscopy, followed by chemical-discoloring and antibody stripping. This method allows for evaluation of mm of tissue and millions of cells in situ. As antigen detection is based on precipitated chromogen, measurements of the markers' expression are less quantitative compared to that of IF-based staining, but is potentially more adaptable for clinical applications. This platform has been applied to multiple clinical and preclinical studies to evaluate immune response to disease and/or therapy^{105,205–207}.
2. Next-generation Cyclic IF. A novel labeling and staining technology has been developed in Drs. Gibbs, Chin and Gray laboratories at OHSU in collaboration with Quantitative Imaging Systems, and uses validated primary antibody-conjugated oligonucleotides (Ab-oligos). In situ hybridization of complementary oligonucleotides labeled with fluorophores (imaging-oligos) facilitates biomarker detection for imaging by conventional fluorescent microscopy. A variety of complementary oligo configurations were validated to remove fluorescent signals from the stained tissue before subsequent staining cycles (2019 SPIE Conference Proceedings). Each Ab-oligo can only be applied once to any given tissue section at the beginning, followed by the cycles of hybridization and signal removal. The novel strategy may solve many of the potential issues currently present in cmIF, such as the time and labor costs required for repeated antibody incubation during each staining cycle. The technology is similar to the codetection by indexing (CODEX) strategy, which was developed in Dr. Nolan laboratory, and employed the combination of DNA barcodes, fluorescent dNTP analogs, and an in situ polymerization-based indexing procedure¹⁰⁸.

3. It is difficult to manage batch-to-batch and day-to-day variations in staining through mathematical methods, such as normalization for the results from multiple samples. A fundamental underlying issue is the inability to control tissue quality between samples within the same batch in the cmIF procedure. Variations in tissue sample quality likely depend on variations in tissue age, fixation time, leading time until tissue fixation, and laboratory technician skills. To overcome experimental or technical variabilities, a deep learning model has been developed, which enables direct translation of H&E images into realistic predictions of two commonly used prognostic markers (pan-CK and α SMA) and a nuclear counterstain DAPI¹⁹⁶. Since prediction happens in silico, the method is consistent and does not show the technical artifacts often appear in IF staining caused by improper fixation, improper reagents, or poor preservation. We envision that this technique could alleviate technical challenge and difficulties of cmIF imaging and may facilitate compatibility with clinical histopathology workflow.
4. The cmIF procedure uses primary or prelabeled antibodies from a large pool of commercial suppliers. The labeling protocol for primary antibodies is well established, and the labeling is inexpensive. The fluorophores used to label antibodies are relatively small molecules, so the potential adverse influence of fluorophores on the antibody's original affinity is minimal. The cost of all necessary chemicals and reagents for cmIF is very low except for that of the primary antibodies. The required imaging equipment and image processing for analysis are the same as those of other platforms.

2.6 CONCLUSIONS

The journey to establish a workflow for cmIF that enables the interrogation of the expression of 60 protein biomarkers in individual cells from a single FFPE section since the first time an advanced breast cancer tissue was successfully stained with nine antibodies over three cycles. The assay workflow of cmIF also includes analysis of the spatial distribution of classified cells and their proximity. The antibodies, especially in-house labeled antibodies, have been validated, as have the reproducibility and repeatability of cmIF using a number of adjacent tissues (most of the data for which is not included in this chapter). The cmIF now applies to the SMMART Clinical Trials²⁰⁸ at the OHSU Knight Cancer Institute for patients suffering from advanced breast, pancreas, or prostate cancer. The ultimate goal of this development is to improve cancer therapy strategies by understanding the mechanisms of cancer progression and drug resistance. We hope this chapter will disseminate our expertise of the cyclic IF procedure and analytics to the broader cancer research community in order to improve cancer patients' quality of life.

2.7 ACKNOWLEDGMENTS

The authors thank Ting Zheng and Lydia Grace Campbell for technical assistance and Yuki Chin for proofreading. The development of cmIF was supported by the funding from Prospect Creek Foundation, Susan G. Komen Foundation, OHSU Foundation, OHSU Center for Spatial Systems Biomedicine (OCSSB), and J.W.G acknowledges the support from NIH/NCI U54 CA209988, NIH/NCI U2C CA233280, and NCI SBIR 1R44CA224994-01. Y.H.C acknowledges the support from Biomedical Innovation Program (Oregon Clinical & Translational Research Institute).

Chapter 3. A FRAMEWORK FOR MULTIPLEX IMAGING OPTIMIZATION AND REPRODUCIBLE ANALYSIS

Published in Communications Biology as:

Eng J, Bucher E, Hu Z, Zheng T, Gibbs SL, Chin K and Gray JW. A framework for multiplex imaging optimization and reproducible analysis. *Commun. Biol.* (2022).

3.1 ABSTRACT

Multiplex imaging technologies are increasingly used for single-cell phenotyping and spatial characterization of tissues; however, transparent methods are needed for comparing the performance of platforms, protocols and analytical pipelines. We developed a python software, mplexable, for reproducible image processing and utilize Jupyter notebooks to share our optimization of signal removal, antibody specificity, background correction and batch normalization of the multiplex imaging with a focus on cyclic immunofluorescence (CyCIF). Our work both improves the CyCIF methodology and provides a framework for multiplexed image analytics that can be easily shared and reproduced.

3.2 INTRODUCTION

Multiplex imaging methods enable quantification of numerous proteins in tissues while retaining spatial and morphological information. In many contexts, dozens of targets must be labeled in an intact tissue section to identify key cell types and quantify their spatial relationships. Multiplex imaging technologies utilize various strategies to overcome the limitation of conventional immunofluorescence (IF) protocols that label a few markers per section. Five and seven-plex immunohistochemistry (IHC) have been achieved by fluorophore-conjugated tyramide deposited on the tissue^{209–211}. Twelve to 29-plex immunohistochemistry (IHC) is enabled with alcohol-soluble peroxidase substrate 3-amino-9-ethylcarbazole (AEC) detection combined with antibody stripping^{105,212} and 40-plex IF can be achieved with antibody elution in iterative indirect immunofluorescence imaging (4i)¹⁰⁶ and multiple interactive labeling by antibody neodeposition (MILAN)^{213,214}. Direct immunofluorescence using fluorophore-conjugated primary antibodies and chemical inactivation of fluorescent dyes enables detection of over 50 protein targets in a single tissue section, in cyclic immunofluorescence (CyCIF)^{107,110,118}, NeoGenomics’ MultiOmyx platform²¹⁵, and iterative bleaching extends multiplexity (IBEX)²¹⁶. Similarly, multi-epitope-ligand cartography (MELC), employs photo-inactivation of fluorescently labeled antibodies^{217,218}. Furthermore, fluorophore-conjugated DNA barcodes (i.e., oligonucleotides) facilitate multiplexing in co-detection by indexing (Akoya’s CODEX)²¹⁹, Ultivue’s InSituPlex²²⁰, Immuno-SABER²²¹, and Ab-oligo cyCIF²²². Imaging mass cytometry¹⁰⁹ and multiplex ion beam imaging¹¹¹ can detect over 40 antibodies conjugated to metal reporters by time-of-flight mass-spectrometry. The Nanostring GeoMx²²³ platform can detect >100 protein targets conjugated to oligonucleotide barcodes, although the data are not images, but spatially registered counts of released oligos²²⁴. All of these methods rely on specific labeling of proteins with antibodies, a process that must be validated by studies that establish specificity, sensitivity and reproducibility. Ideally this is accomplished via transparent analyses using well documented, open-source image analysis pipelines.

Herein, we quantitatively assess antibody labeling in one multiplex imaging platform, CyCIF^{107,110,118}, while varying antibody application strategies, fluorescence signal and tissue autofluorescence removal methods, and experimental batches. This resulted in several strategies to improve CyCIF, including optimization of aspects of chemical bleaching, autofluorescence correction, antibody application order and batch normalization. Herein, we build on protocol optimization efforts^{118,225} and share our image processing pipeline software, the free and open

source python library, mplexable (Figure 1a-b). Finally, we introduced a framework for reproducible image analysis and visualization. Our analytics comprise linked and executable code and data that enabled all figures and analytical results in our paper to be fully reproduced from the accompanying image data (<https://www.synapse.org/#!Synapse:syn23644107>), processed data and code (https://github.com/engjen/cycIF_Validation). We used Jupyter notebooks to link code, data, metadata and the computational environment in a machine- and human-readable document (Figure 3.1c, Table 3.3). Although we focused on validation of a single platform, our tools for image analysis facilitate quantitative and reproducible characterization of tissues by any multiplex imaging platform.

3.3 RESULTS AND DISCUSSION

Adopting the method of Lin *et al.*¹⁰⁷, we stained formalin-fixed, paraffin-embedded (FFPE) tissues using a cyclic process in which four proteins per cycle were labeled using a direct immunofluorescence (IF) protocol, imaged, and quenched of fluorescence signal (Figure 3.1a, [dx.doi.org/10.17504/protocols.io.23vggn6](https://doi.org/10.17504/protocols.io.23vggn6)). We typically labeled 40 – 60 proteins per slide before tissue and staining quality degraded (quantified in Figure 2.2e and Supplementary Figure 7 [Appendix B], respectively). We developed free and open-source tools to perform quality control on images and metadata and automate image processing from registration through single-cell segmentation and feature extraction (Figure 3.1b, <https://pypi.org/project/mplexable/>). Finally, we shared all analytics used to produce figures and conclusions as linked and executable code and data (Figure 3.1c). Jupyter notebooks serve as a human and machine-readable record that described the data, image analysis steps and computational environment, both enabling replication of our findings and linking data and metadata to facilitate use by others.

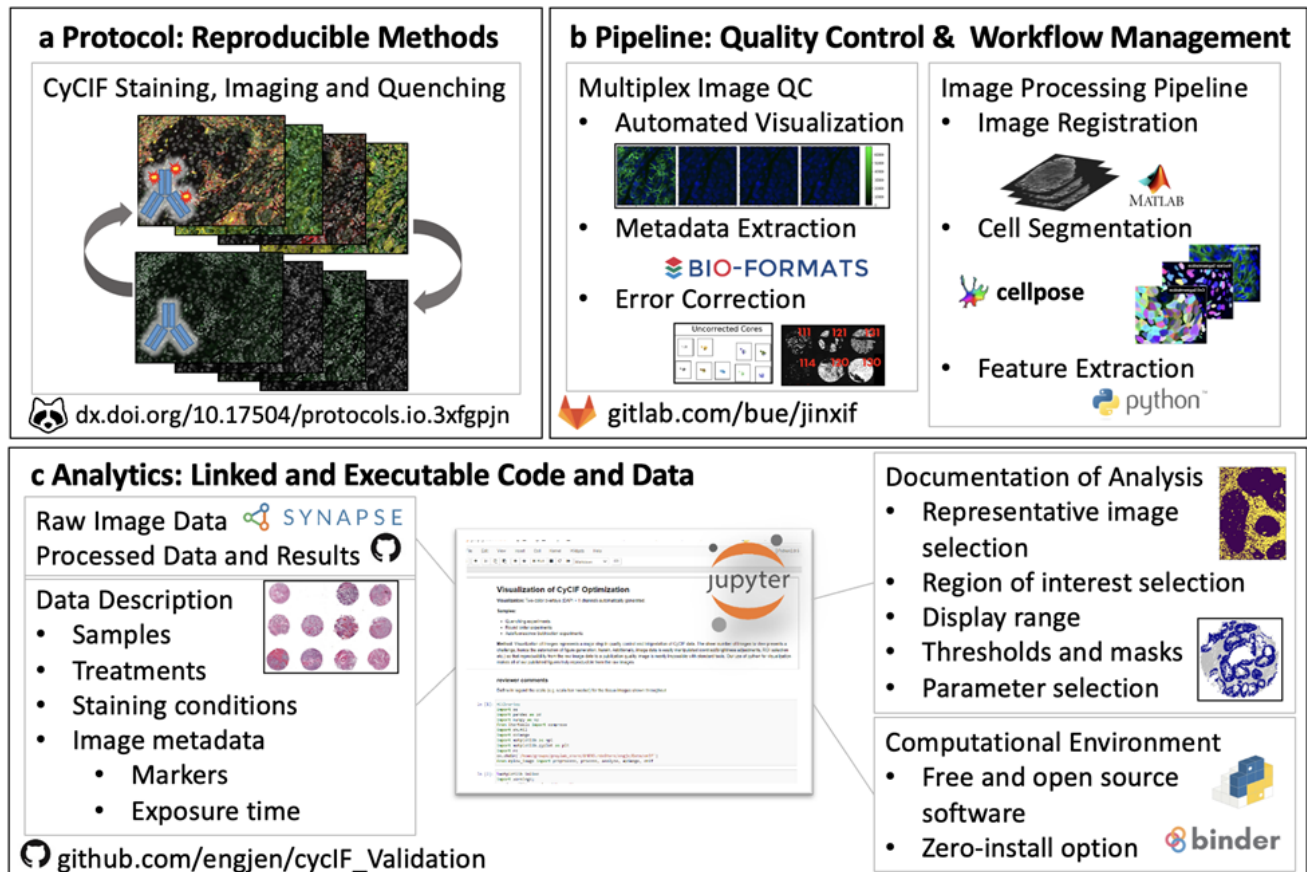


Figure 3.1. Reproducible Generation, Processing and Analysis of Multiplex Images.

a. CyCIF data is generated with a reproducible protocol (dx.doi.org/10.17504/protocols.io.3xfgpjn). b. Image processing pipeline integrates quality control and workflow management with in-house Python software, mplexable, facilitating data quality, scaling and reproducibility. c. Analysis, image visualization and figure generation are fully reproducible with shared linked and executable code and data. Raw and processed data, metadata and analysis are documented in Jupyter notebooks running our free and open-source software that provides a framework for reproducible image analysis.

3.3.1 Comparison to Standard IF

We benchmarked our CyCIF protocol against a direct IF protocol. CyCIF staining in normal and malignant breast tissue was compared to standard direct IF on adjacent tissue sections. Antibodies and staining conditions were selected that produced visually similar staining patterns in both CyCIF and standard IF staining (Figure 3.2a, Notebook 1-1, NB1-1). For quantification, we set an intensity threshold for each marker, with all pixels above the set threshold considered positive. All pixels at least 10 micrometers away from positive staining were considered background pixels (the 10 μ m gap was to exclude any influence of lateral bleed through around positive pixels; Figure 3.2b, Supplementary Figure 1 [Appendix B], NB2-1). If a marker was negative in a given tissue section, it was not analyzed in that tissue. Although the positive and background intensity varied between conditions, presumably due to the imprecisions introduced by manual pipetting, the relative signal-to-background ratio (SBR, see methods) of standard IF compared to CyCIF was near one (relative SBR: mean=0.96, standard error of the mean [SEM]=0.22) and highly correlated (Pearson R =0.89, p =1.6e-9, Figure 3.2c, Supplementary Figure 1 [Appendix B], NB2-2). We

assessed specificity between conditions by segmenting single cells and counting positive cells with a mean intensity above the set threshold. We found that the percentage of positive cells of standard IF relative to CyCIF was also near one (relative % positive: mean=0.96, SEM=0.18) and the fraction positive was highly correlated (Pearson $R=0.99$, $p=5.8e-24$, Figure 3.2c, Supplementary Figure 1c [Appendix B], NB2-2). We concluded that the SBR and specificity of antibody staining in the first five cycles of CyCIF is similar to standard IF, the cyclic process does not excessively impact tissue staining, and CyCIF offers the advantage of detecting increased marker combinations while utilizing a single slide.

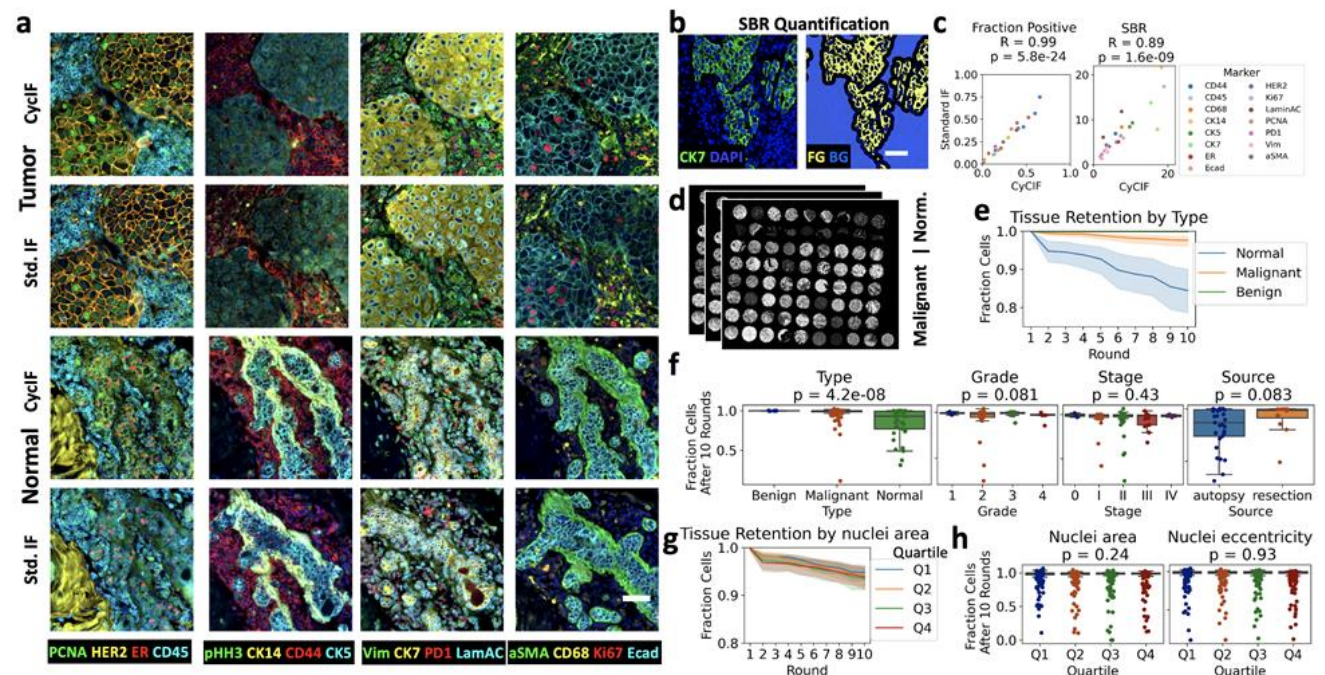


Figure 3.2. CyCIF Signal to Background and Tissue Loss.

a-c. Comparison of CyCIF staining to a standard IF protocol, in adjacent sections of HER2+ tumor and normal breast tissue. a. Multicolor visualizations are produced with our python code. b. Positive pixels, or foreground (FG), were determined by thresholding and negative pixels, or background (BG), were determined by selecting regions 10 μm away from positive regions. The signal-to-background ratio (SBR) is the mean intensity of FG/mean intensity of BG. c. Positive cell counts were determined by applying the thresholds from (b) to single cell mean intensities of segmented cells. Pearson correlation (R), was calculated for standard IF versus CyCIF fraction positive in tissue and SBR, $n=26$; R and p -value given in figure title. d. Three adjacent sections of a 72-core tissue microarray (TMA) containing normal and malignant tissues were repeatedly quenched to assess tissue loss during CyCIF. e. Tissue retention after each round of quenching. Error is 95% confidence interval of tissue retention at each round for normal, benign and malignant tissues, $n = 54, 6$ and 156 , respectively (18 normal, 2 benign and 52 malignant tissues $\times$ 3 replicate TMAs). f. Fraction of cells remaining after 10 rounds of CyCIF, separated by type, stage, grade and source. g. Tissue retention by nuclear area quartile; error is 95% confidence interval, $n=216$ cores. h. Fraction of cells remaining after 10 rounds of CyCIF, separated by nuclear area quartile and nuclear eccentricity quartile. In f, h, Kruskal-Wallis H-test was used to assess differences in tissue retention, $n = 216$ (72 cores $\times$ 3 replicate TMAs), except Source, which includes only normal tissue, $n=41$ cores. p -value shown in figure title. In a-b, scale bar = 50 μm .

3.3.2 Tissue Loss Characterization

The CyCIF process resulted in increased tissue loss compared to standard IF (Supplementary Figure 1a [Appendix B], NB2-1). In order to characterize the numbers and types of cells and tissues

that were susceptible to tissue loss during CyCIF, three adjacent sections of a 72-core tissue microarray (TMA) containing normal and malignant tissues were repeatedly quenched and imaged (Figure 3.2d). Slow but steady tissue loss was observed over 10 rounds of CyCIF, with 95% of cells remaining after 10 rounds (Figure 3.2e). We found that the 18 normal tissues suffered more tissue loss than the 52 malignant and 2 benign tissues (Figure 3.2e), and the difference in fraction of cells remaining after 10 rounds was significant (Figure 3.2f). We found no difference in number of cells lost in tissues separated by stage or grade, or by cells separated by their nuclear size and shape (which correlate with epithelial, immune, endothelial or mesenchymal cell types, Figure 3.2f, g, h). While all malignant tissues were obtained from autopsy, normal tissues obtained from surgical resections vs. autopsy showed a trend towards reduced tissue loss (Mann-Whitney U, $p=0.083$). This could reflect tissue processing variation, as smaller tissue size and longer time in fixation have been shown to reduce section detachment in immunohistochemistry²²⁶. There were significant differences in tissue loss in normal tissues from different organs and malignant tissues with different pathologies (Supplementary Figure 3 [Appendix B]). Our results suggest that tissue loss varied more between different tissues than between individual cell types within a given tissue.

3.3.3 Quenching Condition Assessment

Most tissue loss was observed during the quenching step, necessitating development of gentle yet effective conditions for complete signal removal. CyCIF fluorescence signal removal was accomplished by chemical quenching of fluorescent dyes using hydrogen peroxide (H_2O_2). Only certain dyes, including Alexa Fluor (AF)-488, AF555, AF647 and AF750, were quenchable using this method. Other dyes, including AF546 and fluorescein, were resistant to quenching, as previously reported²²⁵. Increasing H_2O_2 concentration from 3% to 4.5% or 6% in 20 mM sodium hydroxide did not improve quenching rate, and additional time in H_2O_2 also failed to completely eliminate strong signal (Figure 3.3a, b, d Supplementary Figure 4 [Appendix B], 5 NB3-1). Gentle heating with an incandescent light placed ~4 inches above the sample during quenching (see methods) appeared to increase the rate of H_2O_2 oxidation (estimated by heat generation, Supplementary Figure 5c [Appendix B]), resulting in complete signal removal (Figure 3.3c, d Supplementary Figure 5 a, b [Appendix B], NB3-1). Quantification of the unstained negative controls in all quenching conditions showed that mean tissue autofluorescence decreased during the first quench (AF488: mean 22% decrease, SEM=2.8%; AF555: mean 24% decrease, SEM=2%, AF647: mean 5% decrease, SEM=1.4%, $n=3$), a decrease not seen in an unquenched tissue that was repeatedly imaged (Supplementary Figure 5e [Appendix B], f, NB3-1). Therefore, we standardized quenching using 3% H_2O_2 for 30 minutes with incandescent light and added an additional round of quenching (i.e., pre-quenching before CyCIF staining) similar to others¹¹⁸, to reduce overall tissue autofluorescence by roughly 25% (Supplementary Figure 5e [Appendix B]).

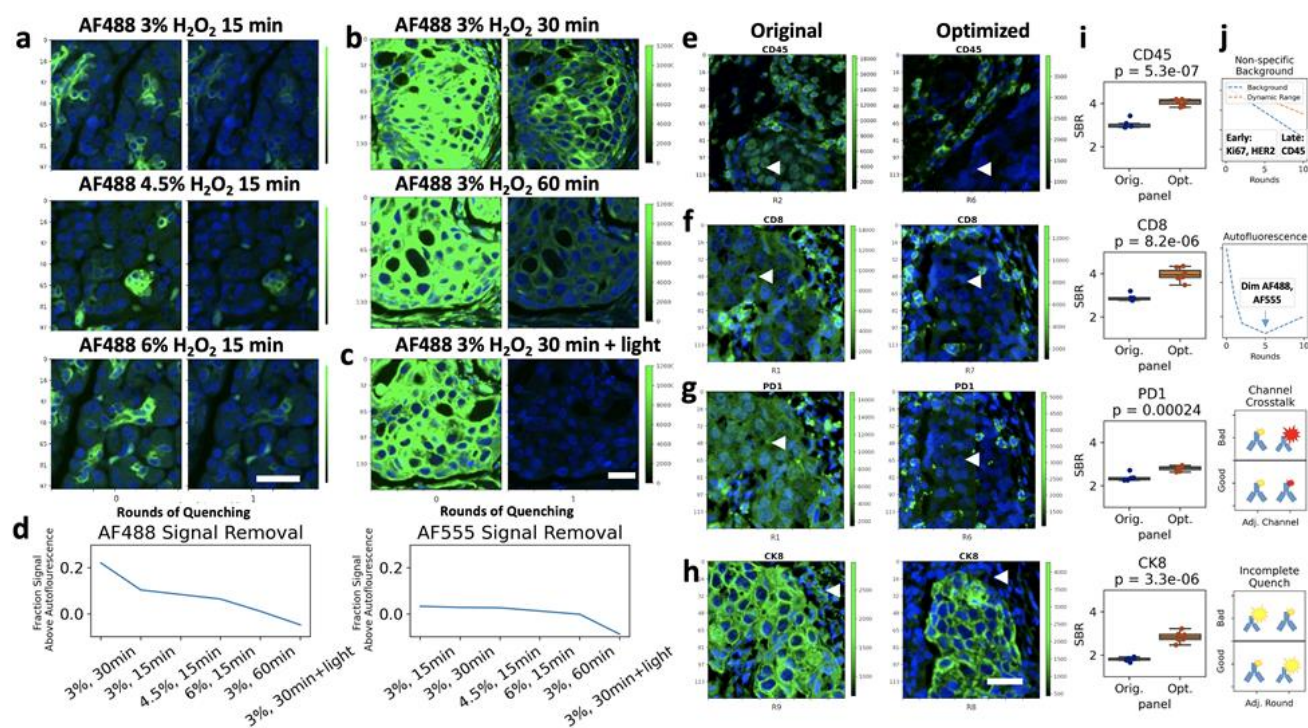


Figure 3.3. CyCIF Optimization: Quenching and Panel Order.

a-d. Quenching Optimization. a. Normal pancreas was stained with CK7-AF488 and imaged before and after fifteen minutes of quenching in 3% (top row), 4.5% (middle row) or 6% (bottom row) H₂O₂. b. Breast cancer tissue stained with CK7-AF488 and Vimentin-AF488 imaged before and after 30 (top) and 60 minutes of quenching (bottom). c. Tissue stained as in (b) and imaged before and after 30-minute quenching under incandescent light source. d. Mean AF488 (left) or AF555 (right) fluorescence intensity in tissue area of stained tissue relative to a blank autofluorescence control after one round of quenching with conditions shown in a-c. e - h. Panel order optimization, with representative images. i. Signal-to-background ratio (SBR) quantification was done by applying a threshold to find positive pixels, and manually selecting areas of non-specific background, e.g., tumor nests in e-g and stromal areas in h., n=6 areas per tissue. A t-test was used to assess significance; p-value shown in figure title. e. Non-specific nuclear staining in the AF750 channel is improved by moving antibody to a later round. CD45-AF750 showed higher SBR in round (R) 6 than in R2. f. Autofluorescence in the AF55 or AF488 channel is mitigated by moving antibody from R1 to a later round. CD8-AF555 had higher SBR in R7 than R1. g. Channel bleed through is mitigated by pairing two bright or two dim antibodies in adjacent channels, not a bright with a dim. PD1-AF647 shows bleed through from bright CK19-AF750 in R1, but not from CD45-AF750 in R6. h. Incomplete quenching is mitigated by moving markers resistant to quenching to later in the panel. Vimentin-AF488 resists quenching and is moved after cytokeratin staining, rather than before. j. Schematic of panel optimizations addressing background and dynamic range, autofluorescence, channel bleed-through and incomplete quenching. a-c & e-h. Blue = DAPI, Green = Stain, green colorbar = 16-bit grayscale intensity, Y-axis scale is in micrometers, scale bar = 30 μ m.

3.3.4 Antibody Order Evaluation and Improvement

Finally, we tested if the order of antibody application in the CyCIF panel influenced antibody sensitivity and specificity. We applied an 11-round, 44-antibody panel twice to the same TMA, for a total of 22 rounds (Supplementary Table 1 [Appendix B]). We visually and quantitatively compared the antibody staining pattern and single cell mean intensities of the images acquired after the first application of each antibody to the second (Supplementary Figure 6 [Appendix B], NB3-3). Of the 44 antibodies evaluated, 86% had a high Pearson correlation (>0.8) between first and second applications, but 81% had lower dynamic range and SBR on the second application (Supplementary Figure 7 [Appendix B], NB3-4). Effects of double-application were antibody and

epitope specific, as 28% antibodies showed similar dynamic range (<50% change) on the second application, and 62% still had an estimated SBR>1.5 on the second application, versus 78% on the first (Supplementary Figure 7 [Appendix B], NB3-4). Non-specific nuclear staining was observed for several antibodies conjugated to the AF750 fluorophore, and unexpectedly, it diminished on the second application (Supplementary Figure 6c [Appendix B], NB3-4). We tested several potential variables impacting stain quality in later rounds, including non-specific IgG interactions, non-specific fluorophore interactions, specific IgG interactions, and quenching effects. Out of all of these, the only condition that apparently reduced non-specific nuclear background was additional quenching (Supplementary Figure 8 [Appendix B]). Thus, we designed an improved panel order which positioned antibody-conjugates tending to have non-specific staining in a later round (Figure 3.3e), as well as addressed challenges of autofluorescence, channel bleed through and incomplete quenching (Figure 3.3f-j, Supplementary Table 2 [Appendix B]). We compared the original and optimized orders on near-adjacent slides cut from a TMA containing three tissues with positive staining for each marker. SBR quantification was completed by applying a threshold to each marker to find positive pixels, and manually selecting areas of non-specific background (e.g., tumor nests for immune markers and stromal areas for tumor markers). Our new panel order significantly increased SBR (Figure 3.3i, NB3-5). In earlier rounds we generally prioritized antibodies to antigens with varied expression levels such as human epidermal growth factor receptor 2 (HER2), E-cadherin and cytokeratins, which generally lose their dynamic range in later staining rounds. We placed immune cell markers in later rounds, as they have a binary (i.e., on/off) expression pattern and sometimes showed better SBR in later rounds, due to decreased background (Figure 3.3e, f, j). Markers expressed in the same cell type and/or subcellular location were placed in non-adjacent channels and rounds to allow detection artifacts from channel bleed through and any incomplete quenching. Bleed through was evident for weakly staining and strongly staining antibodies in adjacent channels (e.g., PD1-AF647 and CK19-AF750, Figure 3.3g, j), but was minimized by avoiding such combinations in the same round. Bleed through was not specific to our selected fluorophores and filter sets, as we observed FITC to PE bleed through in another multiplex imaging platform, which could also be mitigated by avoiding strongly staining and weakly staining antibodies in adjacent channels (Supplementary Figure 8 [Appendix B], NB3-6).

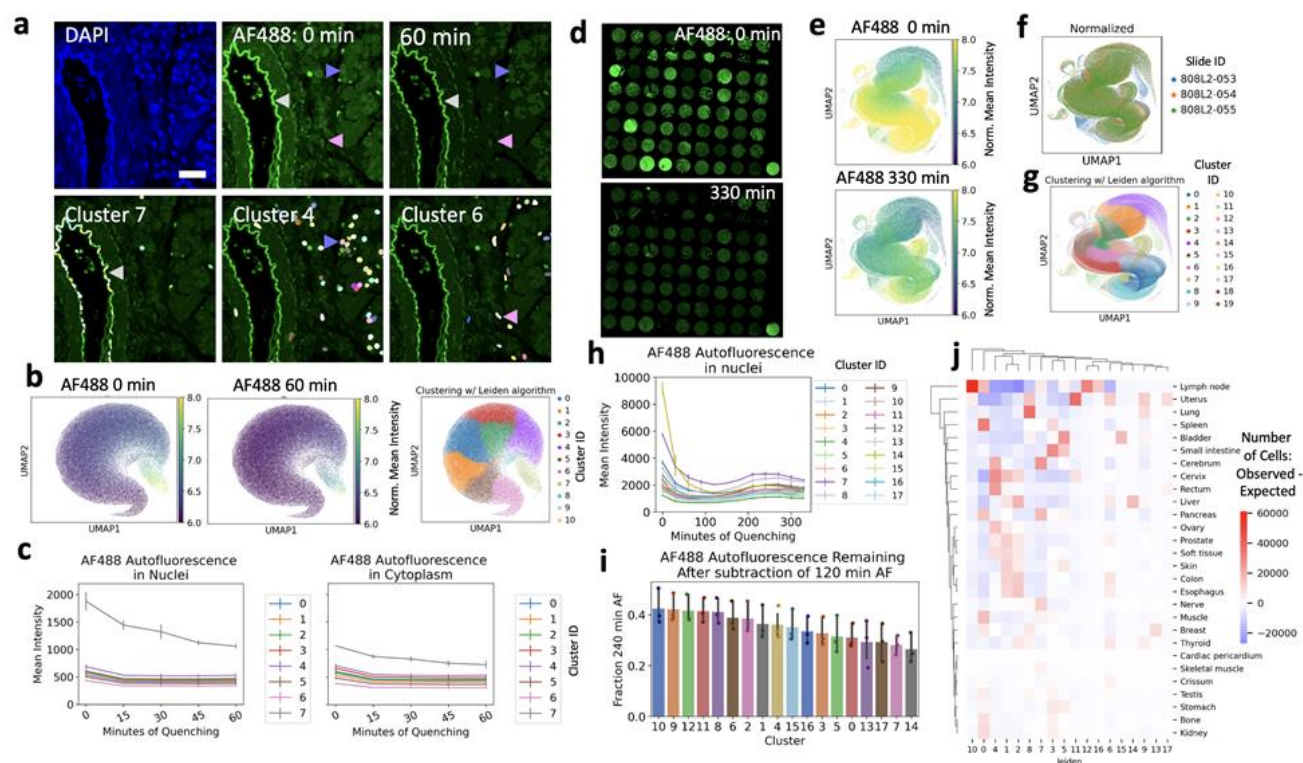


Figure 3.4. Autofluorescence Characterization in the CyCIF Protocol.

a. Top: DAPI nuclear staining and autofluorescence in the AF488 channel before and after 60 minutes of quenching. Arrowheads show areas of bright (gray), medium (purple) and dim (pink) autofluorescence. Bottom: Colored nuclei of clusters 7 (left), 4 (middle) and 6 (right) on the autofluorescence image, representing cells with bright (gray arrowhead), medium (purple arrowhead) and dim (pink arrowhead) autofluorescence, respectively. Scale bar = 30 μ m. b. UMAP projection of single cells based on autofluorescence, colored by AF488 intensity after 0 or 60 minutes of quenching (left and middle); Umap colored by unsupervised clustering results of the Leiden algorithm (right). c. Mean AF488 intensity of cells in each Leiden cluster over rounds of quenching; note bright (gray line) cells quench differently than medium (purple line) and dim (pink line) cells. Similar trend in nuclear (left) and cytoplasmic (right) autofluorescence, although nuclear is brighter. d. Autofluorescence analysis as in (a-c) was repeated on a 72-core tissue microarray (TMA). AF488 autofluorescence shown after 0 minutes (top) and 330 minutes of quenching (bottom). e. UMAP projection of single cells based on autofluorescence, colored by AF488 intensity after 0 (top) or 330 minutes of quenching (bottom). f. UMAP colored by batch. Three adjacent sections from TMA were repeatedly quenched and normalized by batch for analysis. g. UMAP colored by unsupervised clustering results of the Leiden algorithm. h. Mean AF488 intensity of cells in each Leiden cluster over rounds of quenching. Overall, minimum intensity was observed at 120 minutes and maximum intensity at 240 minutes. i. Subtracting the minimum intensity autofluorescence will avoid over-subtraction while still removing 60 - 70% of autofluorescence at 240 minutes. j. Heatmap showing observed number of cells per Leiden cluster – expected number of cells per cluster, by tissue, y-axis, illustrating different trends by tissue. Brightest autofluorescence, cluster 14 (yellow line, h), originates from liver tissue. Cluster 10, showing greatest fraction of autofluorescence remaining after subtraction (blue bar, i), originates primarily in lymph node. c, h and i. Error bars = standard error of the mean, n=3 slides.

3.3.5 Autofluorescence Dynamics and Corrections

Although pre-quenching reduced overall tissue autofluorescence by ~25% (Supplementary Figure 5e [Appendix B]), we sought to develop a method to computationally subtract remaining autofluorescence, which, in some tissue structures, was brighter than the biomarker staining. First, we characterized single-cell autofluorescence dynamics over rounds of CyCIF by quantifying

single-cell AF488 intensity in three unstained normal pancreas tissues that had been imaged and quenched repeatedly. We segmented the nuclei based on DAPI staining and expanded each nucleus by five pixels to capture the cytoplasm. Tissues were batch normalized and unsupervised clustering revealed groups of cells with comparable autofluorescence profiles over the experiment that localized to similar regions in each tissue (Supplementary Figure 9 [Appendix B], NB4-1). Overlaying the cells in each cluster on the tissue revealed cells in cluster seven comprised tissue structures with bright autofluorescence that covered a wide spectrum from DAPI to AF488 channels (Figure 3.4a). Quantification of mean cluster intensity over rounds of quenching showed that while other clusters' intensities stayed constant, the cluster seven intensity dropped linearly between one and four rounds of quenching (Figure 3.4c, NB3-1, NB4-1). To confirm these results, we collected two additional datasets, three adjacent sections from a 72-core TMA containing normal and tumor tissues (Figure 3.4d) and three adjacent sections from an 11-core TMA containing HER2+ breast cancer tissues. The two groups of slides were quenched for a total of 10 rounds, 330 minutes and 6 rounds, 210 minutes, respectively. We segmented cells to obtain single-cell autofluorescence values, batch normalized and clustered as described for the pancreas tissue (Figure 3.4e-g, Supplementary Figure 10 [Appendix B]). Although the AF488 autofluorescence initially declined in these datasets, as had been observed in the pancreas, surprisingly it increased in later rounds, globally across all clusters (Figure 3.4h). We did observe that the autofluorescence clusters correlated with one or more tissue types and pathologies (Figure 3.4j, Supplementary Figure 11 [Appendix B]), but they all showed the same trend, with a minimum intensity at three or four rounds of quenching (after pre-quenching, Figure 3.4h, Supplementary Figure 12 [Appendix B]). We saw similar trends in AF555 autofluorescence (Supplementary Figure 11, 12 [Appendix B]). Therefore, to computationally remove autofluorescence without over-subtracting, we recommend collecting baseline autofluorescence images after quenching at round 3 or 4, and subtracting these, scaled by exposure time, from AF488, AF555 and AF647 channels. Subtraction from the AF750 channel appeared unnecessary, given its minimal autofluorescence. This procedure will remove an additional 60 – 70% of autofluorescence from the brightest rounds (Figure 3.4i, Supplementary Figure 11, 12 [Appendix B]). Since different tissues and other multiplex imaging platforms may exhibit different autofluorescence dynamics, we implemented both the baseline algorithm and a scaled algorithm, which assumes linear increase or decrease in autofluorescence, for autofluorescence subtraction in mplexable. We found that the scaled algorithm, by linearly interpolating autofluorescence between background images collected at the beginning and end of staining, did reduce false positives for some markers in our panel (Supplementary Figure 11 [Appendix B]). Since the brightest areas of autofluorescence showed linear decrease across the first few rounds of CyCIF, the scaled algorithm was most advantageous in tissues with strong autofluorescence that are stained with only a few rounds of CyCIF. On the other hand, the baseline algorithm is simple, requiring collection of just one background image, and can remove the majority of autofluorescence background without over-subtraction in various tissue and experimental contexts.

3.3.6 *Reproducibility and Batch Normalization*

Following panel optimization, we assessed the reproducibility of CyCIF by analyses of three serial TMA sections stained with the same 20-antibody panel on different dates (Figure 3.5a). We used manual thresholding to quantify SBR, as in our standard versus CyCIF analysis (Supplementary Figure 13a, b [Appendix B], NB5-1). We compared our manual method of SBR calculation to the

use of intensity quantiles to estimate dynamic range¹¹⁸, and found that manual thresholds were more reliable when markers were negative, (e.g., HER2 in normal breast, Supplementary Figure 14 [Appendix B], NB5-2). With both methods, we found that the ratio of SBR between replicates for all 20 antibodies was close to one (relative SBR: mean=1.02, SEM=0.23, n=20, Figure 3.5b, Supplementary Figure 13c & 14 [Appendix B], NB5-2). Similarly, the difference in percent positive cells between each replicate was small (relative percent positive: mean=1.10, SEM=0.38, Supplementary Figure 13c [Appendix B], NB5-2). These data demonstrate reproducibility between batches, with the SBR of each antibody being stable, despite variations in raw intensity values (Supplementary Figure 9c [Appendix B], NB5-2). We used the same dataset to examine inter-patient variability on each TMA. The relative SBR between patients was often far from one (Supplementary Figure 13c [Appendix B], NB5-2), likely reflecting a combination of biological variability and pre-analytical factors that affect antibody performance in FFPE tissues.²²⁷

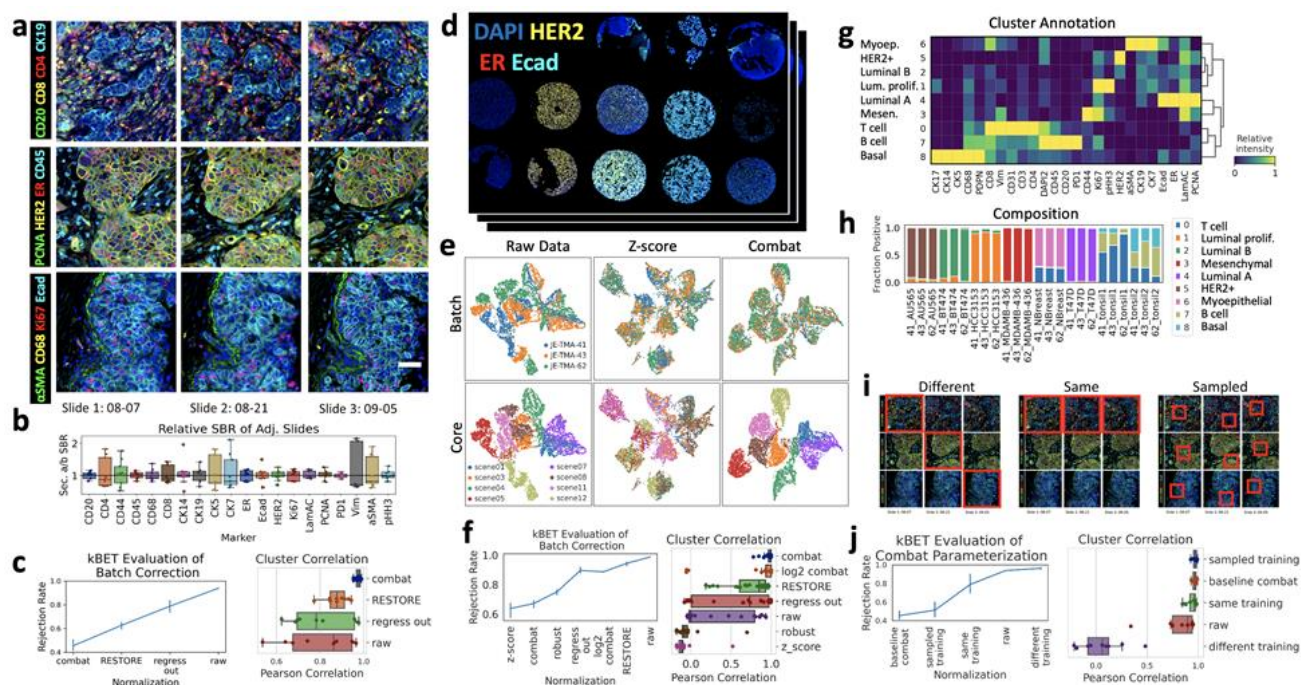


Figure 3.5. Reproducibility and Normalization of CyCIF Staining Intensity.

a. Representative images, generated with our python code, of 12 markers from three adjacent sections of a breast cancer TMA stained with a 20-marker CyCIF panel, scale bar = 50 μ m. b. Relative SBR of slide 1-2, 2-3 and 1-3 for each marker. c. Evaluation of batch correction with kBET (left, lower rejection rates indicating better batch correction, n=3 batches, 5400 cells) and correlation (Pearson, n=9) of Leiden cluster composition between replicate cores for different batch correction methods (right) d. Overview of breast cancer cell line and normal tissue TMA created to represent breast cancer subtypes. e. UMAP projection based on single-cell marker intensity; left to right: raw data, z-score normalization and ComBat normalization, colored by batch (top) and TMA core/cell line (bottom). f. Evaluation of batch correction with kBET (left, n=3 batches, 7200 cells) and cluster correlation (right, n=24) on cell line TMA. g. Heat map of relative mean intensity of each marker in the ComBat normalized, Leiden-clustered cell line TMA data, with annotation on left. h. Fraction of cells in each cluster from (g), showing similar composition of technical replicates and reflecting known normal tissue and cell line cell types. i. Schematic of different selection methods to parameterize the ComBat algorithm. j. Evaluation of ComBat parameterization with kBET (left, n=3 batches, 7200 cells) and cluster correlation (right, n=9). c, f and j. Error bars = standard error of the mean, n=3 slides.

Since we observed intensity variation between CyCIF replicates, batch normalization was required for downstream analytics such as cell classification by unsupervised clustering. To quantify the

effectiveness of batch correction methods on our serial TMA sections, we used the kBET algorithm²²⁸, which, given a k-nearest neighbor graph, compares batch-label distribution in random subsets of neighboring cells to the global batch-label distribution, thus quantifying how well mixed batches are in technical replicates (Figure 3.5a, c, Supplementary Figure 15 [Appendix B], NB5-3). Previously, we developed RESTORE, which utilizes mutually exclusive marker expression to predict a background threshold²²⁹. For normalization, we set all values below the RESTORE threshold to zero and scaled those above the threshold between 0 and 1. We compared RESTORE normalization to ComBat²³⁰ and mean-only correction (regress out), both implemented in scanpy²³¹. ComBat produced the best batch correction (kBET=0.45, SEM=0.05, n=3 slides, lower kBET rejection rates indicate better batch correction, with 0 being perfect mixing and 1 being no mixing of batches) with RESTORE also correcting some of the batch effect (kBET=0.62, SEM=0.03, Supplementary Figure 15 [Appendix B], NB5-3). Unsupervised Leiden clustering showed that both ComBat and RESTORE recovered known biology (Supplementary Figure 15 [Appendix B]). To quantify the stability of clustering across batches, we calculated the Pearson correlation of the positive fraction of each cluster between each section (n=9; 3 cores x 3 sections). ComBat normalization produced higher correlation with lower variance (indicating fewer cores with low correlation) than RESTORE normalization, but both were improvements over the raw data (ComBat mean Pearson correlation=0.97, SEM=0.01, RESTORE mean=0.88, SEM=0.05, raw mean=0.8, SEM=0.15, n=9, Figure 3.5c, Supplementary Figure 15 [Appendix B], NB5-4).

To further evaluate batch normalization methods, we created a TMA using FFPE breast cancer cell lines and normal tissues (Figure 3.5d). Importantly, the cell lines were of known clinical and intrinsic subtype²³², establishing a ground truth for clustering results (Supplementary Table 3 [Appendix B]). Three TMA sections were stained using the same antibody conjugates but different panel orders (Supplementary Table 4 [Appendix B]). Images were processed through registration, autofluorescence subtraction, single cell segmentation, and feature extraction. Due to different panel orders, raw data showed significant batch effects, with little mixing of cells from different TMAs in a UMAP projection based on single cell marker intensity (Figure 3.5e, NB5-5). Again, the kBET algorithm²²⁸ and replicates' cluster correlation were used to quantify the effectiveness of batch correction methods (Figure 3.5f, Supplementary Figure 16 [Appendix B], NB5-6). Thirty methods were evaluated on three random samples of 600 cells from each core in each replicate (Supplementary Figure 16 [Appendix B], NB5-5). Although Z-score normalization produced a reasonable kBET score (mean rejection rate=0.64, SEM=0.05, n=3), the UMAP did not show separation of each core, i.e. known breast cancer subtypes (Figure 3.5e, NB5-5). In contrast, the ComBat algorithm scored well with kBET (mean rejection rate=0.67, SEM=0.03, n=3), separated the cores in the UMAP, and clustering recovered known cell types within cell lines and normal tissues (Figure 3.5e, g, h). ComBat normalization applied without log2 transformation produced the most consistent clustering between replicate cores (mean Pearson correlation=0.97, SEM=0.06, 8 cores x 3 sections, n=24; Figure 3.5f, h, Supplementary Figure 16 [Appendix B], NB5-5). RESTORE normalization resulted in improved cluster correlation compared to raw data (mean=0.73, SEM=0.25 versus mean=0.30, SEM=0.41, n=24), and recovered known biology (Figure 3.5f, Supplementary Figure 17 [Appendix B], NB5-7); however, due to the lack of mutually exclusive marker pairs in cell lines, RESTORE did not perform as well as ComBat in this dataset (Figure 3.5f).

Since the ComBat normalization method²³⁰ scored highest in kBET batch correction and cluster correlation in multiple datasets while recovering biological features of interest, we tested its sensitivity to different input distributions for parameterization. Since ComBat standardizes the mean and variance across batches, the same control tissue must be present in each batch (Figure 3.5i, j, Supplementary Figure 16e [Appendix B], f). Sampling cells from three control tissues produced better kBET scores than using a single control tissue (sampled training with kBET=0.51, SEM=0.06 versus same training with kBET=0.79, SEM=0.07, n=3, Figure 3.5i, j, Supplementary Figure 16 [Appendix B], NB5-3). Cluster correlation varied with the resolution of clustering, indicating higher precision batch correction with sampling from three rather than a single tissue (with 9 clusters, (sampled training with mean=0.98, SEM=0.02 versus same training with mean=0.98, SEM=0.04; with 17 clusters, (sampled training with mean=0.97, SEM=0.02 versus same training with mean= 0.94, SEM=0.05, n=9 Figure 3.5i, j, Supplementary Figure 16f [Appendix B]). Therefore, one or more control tissues with similar cellular composition to the sample tissues should be included in each batch of CyCIF for proper ComBat normalization. In the absence of control tissues and given mutually-exclusive marker pairs, the RESTORE algorithm is a good alternative that produced high correlation in cluster composition in multiple replicate CyCIF experiments (mean correlation=0.88; 0.73). Additionally, in the absence of blank images for autofluorescence (AF) subtraction, RESTORE reduced the influence of AF on analytics (Supplementary Figure 15 [Appendix B]).

In summary, we performed extensive validation of the CyCIF multiplex imaging method to optimize fluorophore signal removal, antibody order and autofluorescence subtraction. We also quantified CyCIF similarity to standard IF and the reproducibility of staining, using this data to evaluate methods for batch normalization and cluster analysis to define biologically-relevant cell types. In parallel, we developed the open-source python library, mplexable, for reproducible image processing and analysis. We provided all of our code and data not only to document and reproduce our work, but to enable use and further community development of our analytics (https://github.com/engjen/cycIF_Validation). Our validation studies and computational tools will facilitate maturation of multiplex imaging methods towards quantitative, reproducible characterization of protein expression in intact tissue sections.

3.4 METHODS

3.4.1 Tissue and Cell Button Preparation

We purchased TMA tissue sections (BR1506, US Biomax, Derwood, MD). All human tissue was collected under HIPPA approved protocols with the highest ethical standards with the donor being informed completely and with their consent (OHSU Biobank IRB 4918). Tissue blocks and tissue microarrays were from archival tissues fixed with standard methods (Biomax, Derwood, MD and OHSU Biobank, Portland, OR). Cell lines were cultured in conditions following Neve *et al*²³². Cell line sources and culture media are as follows: AU565, ATCC, RPMI1640+10%FBS, BT474, ATCC, PMI1640+10%FBS, HCC1143, Adi Gazdar (now available through ATCC), RPMI1640+10%FBS, HCC3153, Adi Gazdar, RPMI1640+10%FBS, MDAMB436, ATCC, DMEM+10%FBS, T47D, ATCC, RPMI1640+10%FBS. FFPE cell buttons were prepared as previously described²³³. For each cell button, cells were scraped from two 15 cm plates and spun down (in 10 ml of culture media, no serum, for 4 min X 1000 rpm) in collodion-coated 15 ml glass

centrifuge tubes (Fisher #C408-500). Collodion bags containing pellets were gently removed from tube, tied with thread, and transferred to 10% buffered formalin on ice. Cell pellets in collodion bags were fixed overnight at 4° C in 10% buffered formalin, then transferred through graded ethanol series (30%, 50%, 70%) for one hour each, and processed and embedded in paraffin. For cell line TMA creation, 1 mm cores from normal breast and tonsil (OHSU Biolibrary), or FFPE cell buttons were punched and inserted into the recipient block with the Estigen manual tissue arrayer MTA-1.

3.4.2 *Cyclic Immunostaining*

Formalin-fixed paraffin-embedded (FFPE) human tissues were sectioned at 4-5 microns and mounted on positively charged slides (Tanner Adhesive Slides, Mercedes Medical, TNR WHT45AD). The slides were baked overnight at 55 °C (Robbin Scientific, Model 1000) and for an additional 30 minutes at 65 °C, (Clinical Scientific Equipment NO. 100). Tissues were deparaffinized and hydrated through xylenes and graded ethanol (EtOH) as follows: xylenes (3 x 5 min), 100% EtOH (2 x 5 min), 95% EtOH (2 x 2 min), 70% EtOH (2 x 2 min), and distilled and deionized water (ddH₂O, 2 x 5 min). Two step antigen retrieval was performed in a Decloaking Chamber (Biocare Medical, Pacheco, CA) using the following settings: set point 1 (SP1), 125 °C, 30 seconds; SP2: 90 °C, 30 seconds; SP limit: 10 °C variation. Briefly, slides were placed in decloaking chamber in a plastic Coplin jar containing citrate buffer, pH 6 (10 mM citrate, Sigma C-1909). Two additional polyethylene Coplin jars with buffer were placed in chamber to heat, which contained ddH₂O and 1x Target Retrieval Solution, pH 9 (Agilent S2367). The chamber was heated to 125 °C, held for 30 seconds (SP1), then cooled to 90 °C, 0 psi and held for 30 seconds (SP2). After the SP2 program was completed, the decloaking chamber was turned off, opened, and slides were dipped in the Coplin jar containing hot ddH₂O for ~1 second. Slides were then transferred to hot 1x Target Retrieval Solution pH 9. The lid was placed back on chamber, and slides remained in hot pH 9 buffer for 15 minutes. Following this two-step antigen retrieval, the tissues were washed in two brief changes of ddH₂O (~2 seconds) and then washed once for 5 minutes in 1x phosphate buffered saline (PBS), pH 7.4 (Fisher, BP39920).

3.4.3 *Pre-quenching to reduce autofluorescence*

Next, pre-quenching was performed on tissues to reduce tissue autofluorescence. Quenching solution containing 20 mM sodium hydroxide (NaOH) and 3% hydrogen peroxide (H₂O₂) in 1x PBS was freshly prepared from stock solutions of 5 M NaOH and 30% H₂O₂, and each slide was placed in 10 ml quenching solution. Slides were quenched face down on ~1 mm risers in a 4-well rectangular tissue culture dish (each well holds one slide), under incandescent light, for 30 minutes. Lamps with 60-Watt incandescent bulbs were positioned so the bulb was four inches above the 4-well dish. Placing slides in the outer two wells and leaving the center wells empty resulted in the temperature increasing from 23° C to 39° C over 30 minutes (see Supplementary Figure 5c [Appendix B]). Slides were then removed from the chamber with forceps and washed 3 x 2 min in 1x PBS. Sections were blocked in 10% normal goat serum (NGS, Vector S-1000) and 1% bovine serum albumin (BSA, Sigma A7906) in 1x PBS for 30 minutes at room temperature in a humidified chamber. Plastic cover slips (IHC World, IW-2601) were used to spread blocking solution evenly

across tissue. Tissues were soaked briefly in PBS in a Coplin jar to remove plastic coverslip, then washed 1 x 5 min in PBS. Coverslips (Corning; 2980-243 and 2980-245) were mounted in Slowfade Gold plus DAPI mounting media (Life Technologies, S36938). Pre-staining autofluorescence signal was acquired using a Zeiss Axioscan Z.1 (see imaging protocol below). After acquiring autofluorescence signal, slides were soaked in 1x PBS for 10 – 30 minutes in a glass Coplin jar, waiting until glass coverslip slid off without agitation.

3.4.4 *Primary Antibody Staining*

Primary antibodies were diluted in 5% NGS and 1% BSA in 1x PBS (see Supplementary Data 1) and applied overnight at 4 °C in a humidified chamber, covered with plastic coverslips (IHC World, IW-2601). Following overnight incubation, tissues were washed 3 x 10 min in 1x PBS, and coverslipped as described above. Antibody information is provided in Supplementary Data 1.

3.4.5 *Fluorescence Microscopy*

For standard versus CyCIF and reproducibility experiments, fluorescently stained slides were scanned on the Zeiss AxioScan.Z1 (Zeiss, Germany) with a Lumencor SpectraX-IR light source (Lumencor Inc., Beaverton, OR). The filter cubes used for image collection were DAPI (Semrock, LED-DAPI-A-000), AF488 (Zeiss 38 HE), AF555 (Zeiss 43 HE), AF647 (Zeiss 50) and Alexa Fluor 750 (AF750, Chroma 49007 ET Cy7). The exposure time was determined individually for each slide and stain, and the LED light intensity was fixed at 100%. Full tissue scans were taken with the 20x objective and stitching was performed in Zen Blue image acquisition software (Zeiss). After those initial experiments, a new light source was purchased and used for collecting quenching and round order data. For those experiments, a Colibri 7 light source (Zeiss), with the same filter cubes, except DAPI (Zeiss 96 HE), was used. The exposure time was determined individually for each slide and stain to ensure maximum dynamic range without saturation, and the LED light intensity was fixed at 10% (DAPI), 20% (AF488), and 50% (AF555, AF647, AF750). Full tissue scans were taken with the 20x objective (Plan-Apochromat 0.8NA WD=0.55, Zeiss) and stitching was performed.

3.4.6 *Quenching of Fluorescence Signal for Cyclic Immunostaining.*

After successful scanning, slides were soaked in 1x PBS for 10 – 30 minutes in a glass Coplin jar, waiting until glass coverslip slid off without agitation. Quenching was performed as described above, in the section of Pre-quenching to reduce autofluorescence. After removal from quenching solution, slides were washed 1 x 5 min in 1x PBS and subsequent rounds of primary antibodies were applied, diluted in blocking buffer as described in the section of Primary Antibody Staining.

3.4.7 *Image Registration, Autofluorescence Removal and Segmentation*

Scanned images were first split into separate scenes using the function Split Scenes (Write files) in Zeiss Zen Blue software (to uncheck “Include scene information in the Generated File name”). For the datasets used in this work, we did not apply flat field correction, although it may be applied in Zen using the function Shading Correction. Using the same software, each scene was then exported to 16-bit grayscale uncompressed TIFF using the function Image Export. Quality control

and metadata extraction were performed in python. TIFF images from each round of cyclic immunostaining were registered based on DAPI staining as follows. Image features were found with the detectSURFFeatures function, and automated feature matching was performed with the matchFeatures function, in Matlab (R2017B 9.3.0.713579, MathWorks, Natick, MA). Image registration was performed using Matlab's estimateGeometricTransform function and performing affine registration (scaling, rotation, translation). Although Matlab was used for the majority of registration in this work, we also successfully registered images using a python implementation. Both scripts are provided (Matlab: https://gitlab.com/engje/mplexable/-/blob/master/mplexable/src/template_registration_mscene.m and python: https://gitlab.com/engje/mplexable/-/blob/master/mplexable/src/template_registration_mscene.py).

Autofluorescence subtraction preceded segmentation. Images of unstained tissue were acquired in each channel, before and after staining. For each marker, background images were scaled linearly by exposure time and relative round, and subtracted using mplexable.

Deep learning based cell segmentation was performed with Cellpose, a generalist algorithm for cellular segmentation²³⁴. Cellpose was used to generate nuclear and cell masks by classifying pixels on the basis of a DAPI or E-cadherin antibody staining, respectively. The following parameters were used for Cellpose segmentation: for the cells, diameter=30 pixels, flow_threshold=0.6, min_size=113; for the nuclei, diameter=30, flow_threshold=0, min_size=28. Nuclei with no E-cadherin (Ecad) staining (i.e., non-epithelial cells) were expanded by 5 pixels (1.6 micrometers) to approximate the cytoplasm, based on the average measurement of immune cell cytoplasm in images. The cytoplasm was derived by subtracting the nuclei area from the cell segmentation result, or from the 5-pixel expansion result in the case of Ecad negative cells. The mean intensity of each subcellular region was extracted using mplexable. Watershed-algorithm-based cell segmentation was performed for some datasets, using in-house java software for the following operations. A z-projection of DAPI images from all rounds of staining was processed with the white top-hat algorithm that separates individual nuclear candidates from the background. Contours of nuclei were detected with the Prewitt operator, and single nuclei were segmented by applying a watershed algorithm to the nuclear contours, with top-hat candidates as seeds. Nuclear segmentation accuracy was improved by sorting nuclei based on expression of tumor cytokeratins or immune markers and using this information to set a maximum nuclear size for the watershed algorithm. If a cell was positive for cytokeratins, it was allowed to have a larger nucleus than cells that were negative for cytokeratins because it was assumed to be an epithelial cell. Cell segmentation was achieved by applying a watershed on the Ecad image with segmented nuclei as seeds, or by inflating the nuclei if the Ecad marker was negative. The resulting segmentation mask defined nuclear and cytoplasmic regions for each cell. Mean intensity used for downstream analysis was selected for each marker based on its biologically-relevant subcellular region (e.g., cytoplasm for CK19, nuclei for Ki67).

3.4.8 Data Analysis

For single cell analyses, single cell mean intensity was used for clustering, as in Figure 3.4 and 3.5. For percent positive calculation, as in Figure 3.2c (left), cells with a mean intensity above threshold were considered positive. Tissue retention was calculated in Figure 3.2e – h by thresholding DAPI using the Li algorithm²³⁵ and considering cells above DAPI threshold as retained in that round. For signal-to-background (SBR) calculations, mean intensity was integrated

across the entire slide or region of interest, as in Figure 3.2c, right, Figure 3.3, and Figure 3.5b. In fluorescence imaging, background adds to the signal of interest²³⁶, so SBR was calculated as *(mean intensity of positive pixels – mean intensity of negative pixels) / mean intensity of negative pixels* (NB2-1). For SBR quantification in Figure 3.2, thresholds were applied directly to image data (no single cell segmentation) and signal was taken as the mean pixel intensity above the threshold, while background was defined as the mean pixel intensity of pixels below the threshold and 30 pixels away (10 μ m) from the positive pixels to exclude the influence of lateral bleed through. Since the threshold directly determines the result, thresholds were used that selected a similar pixel pattern and area in adjacent sections. The same marker in adjacent sections was visualized side-by-side, and the respective thresholds were adjusted until the positive pixels were as equivalent as possible, which was estimated by eye. Therefore, although the threshold reflected the subjective decision of the researcher, it allowed comparison of similar pixels in replicates across adjacent sections. The masks that resulted from thresholding are provided for visual assessment of thresholds

(https://github.com/engjen/cycIF_Validation/blob/master/Extended_single_vs_cyclic.ipynb and https://github.com/engjen/cycIF_Validation/blob/master/Extended_Reproducibility_3TMA_Tissue.ipynb).

We tested whether the same threshold could be applied to different regions of the same slide by measuring correlation between two ROIs in normal breast and two ROIs in HER2+ breast tumor given the same threshold. The mean fluorescence intensity measured above threshold, and the intensity of background noise were highly correlated between ROIs (Supplementary Figure 2 [Appendix B]). Finally, we tested whether manual thresholds gave us a different answer for SBR calculations than estimating SBR at the 95th/5th quantile (Supplementary Figure 14 [Appendix B]). Dynamic range was estimated using the 4th and 99.5th quantile of mean intensity for markers in tissues with known positive staining. We compared different ranges for estimating dynamic range (5th to 98th percentile versus 5th to 99.9 percentile, see Supplementary Figure 14 [Appendix B]). We found that using a higher maximum did not change the dynamic range as much for common markers (e.g., CK7) but had more effect on rare markers (e.g., Ki67 and alpha-SMA). Therefore, we selected the 99.5th percentile as the maximum to reflect both common and rare markers' dynamic range. In cases where we did not set a manual threshold, SBR was estimated as the ratio of the (99.5th quantile – 4th quantile)/4th quantile of mean intensity.

For SBR quantification in Figure 3.3, we needed to calculate SBR in the presence of artifacts such as non-specific staining, bleed through, autofluorescence and incomplete quenching. We utilized the napari image viewer to overlay all of the markers, set thresholds and create masks. For positive signal, we set a threshold (recorded in our 20211007_napari.py script) that created a mask including all of the positive staining plus any bright artifacts. We then manually erased areas of the mask covering imaging or staining artifacts and saved the mask for future reproducibility. For background signal, we manually selected six regions of the image exhibiting background caused by the artifacts listed above, and again saved the mask. We then extracted the mean intensity of the positive and background areas of the image and used these to calculate SBR as described above. For F1 score calculation in Supplementary Figure 11 [Appendix B], we again used the napari image viewer to overlay staining, segmentation results, and positive cells based on manual thresholding. Based on the staining pattern and other marker's expression (e.g., membranous CD8 staining in CD45+ cells was a true CD8 positive), we manually annotated false negatives in three 2000 x 2000-pixel ROIs. False positives were any cell with AF488 autofluorescence >1024-pixel

intensity, true positives were cells above this threshold excluding false positives and true negatives were all other cells neither positive, false positive or false negative.

Normalization methods tested included transformations (raw, log2 or arcsinh), division by background signal, determined either with RESTORE, which requires mutually exclusive marker expression in different cell populations, or the 3rd quantile of background fluorescence measured in the reverse subcellular compartment of expected localization (e.g., cytoplasmic signal for nuclear markers), scaling methods (standard, min-max, max-abs, robust, quantile and power) which are implemented in the python library scikit-learn²³⁷ and batch correction algorithms regress_out and ComBat, implemented in the python library scanpy²³¹. For RESTORE normalization, the scikit-learn²³⁷ TruncatedSVD decomposition function was used to quantify the L-shaped distribution of marker pairs; in each core, markers were considered mutually exclusive if the function yielded an R-value above 0.5, or above 0.2 when no pairs reached the 0.5 cut-off. For global thresholds, a more stringent cutoff was used (for the presented TMA datasets 0.66 globally or 0.5 when no pairs reached the 0.66 cut-off. For each core or batch, this procedure generated data-driven mutually exclusive marker pairs. The selected mutually exclusive marker pairs were used to calculate RESTORE thresholds, and the median threshold produced by multiple marker pairs was selected for normalization. Mean intensity was normalized with division by local (per core) and global (per batch) thresholds. For the RESTORE scale method, cells with intensity below threshold were set to a random value between 0 and 0.02, while all cells with intensity above threshold were scaled to a range of 0.02 – 1 for each marker.

Following normalization, 7200 cells (cell lines) or 5400 cells (tissues) were randomly sampled (i.e., 600 cells per TMA core from each batch) and evaluated for batch effects using the kBET algorithm²²⁸ (n=3). UMAP visualization and graph-based Leiden clustering (resolution=0.6) was carried out using scanpy²³¹.

3.4.9 Statistics and Reproducibility

All visualizations and results can be fully reproduced from the raw images with the accompanying code and data, here https://github.com/engjen/cycIF_Validation. Statistical analyses were conducted in python using the scipy²³⁸ library. Replicates were defined as separate CyCIF experiments. Sample sizes are defined by the number of tissues stained, including tissue cores in TMAs.

3.4.10 Data Availability

All image data is available at synapse.org, at cycIF_Validation, syn23644107. The synapse platform requires registration. With a free account, the images are freely accessible. All other source data used to produce the graphs and figures is available here: https://github.com/engjen/cycIF_Validation, DOI: 10.5281/zenodo.6049278. (Table 3.3).

Table 3.3. Reproducible Image Analytics

. All figures and results can be fully reproduced with the accompanying Jupyter notebooks and data provided here: https://github.com/engjen/cycIF_Validation, DOI: 10.5281/zenodo.6049278

Table 1: Links to Analysis Notebooks	
Name	url
NB1-1	https://github.com/engjen/cycIF_Validation/blob/master/Multicolor_Image_Visualization.ipynb

NB2-1	https://github.com/engjen/cyclF_Validation/blob/master/Extended_single_vs_cyclic.ipynb
NB2-2	https://github.com/engjen/cyclF_Validation/blob/master/SinglevsCyclic_44290.ipynb
NB3-1	https://github.com/engjen/cyclF_Validation/blob/master/Quenching_analysis.ipynb
NB3-3	https://github.com/engjen/cyclF_Validation/blob/master/Image_Analysis_Visualization.ipynb
NB3-4	https://github.com/engjen/cyclF_Validation/blob/master/DoubleApplication_K157.ipynb
NB3-5	https://github.com/engjen/cyclF_Validation/blob/master/OrderOptimization_K154vsK175.ipynb
NB3-6	https://github.com/engjen/cyclF_Validation/blob/master/Macsima_clustering.ipynb
NB4-1	https://github.com/engjen/cyclF_Validation/blob/master/Quenching_Single_Cell.ipynb
NB5-1	https://github.com/engjen/cyclF_Validation/blob/master/Extended_Reproducibility_3TMA_Tissue.ipynb
NB5-2	https://github.com/engjen/cyclF_Validation/blob/master/TMAReplicates_analysis.ipynb
NB5-3	https://github.com/engjen/cyclF_Validation/blob/master/Normalization_testing_tissue.ipynb
NB5-4	https://github.com/engjen/cyclF_Validation/blob/master/Normalization_testing_HER2-N75.ipynb
NB5-5	https://github.com/engjen/cyclF_Validation/blob/master/Normalization_testing.ipynb
NB5-6	https://github.com/engjen/cyclF_Validation/blob/master/kBET.ipynb
NB5-7	https://github.com/engjen/cyclF_Validation/blob/master/RestoreNorm_scale.ipynb

3.4.11 Code Availability

The data, code and Jupyter notebooks to reproduce the analyses herein are at https://github.com/engjen/cyclF_Validation, DOI: 10.5281/zenodo.6049278. For image processing, we developed mplexable, available through the Python Package Index, <https://pypi.org/project/mplexable/>. An image processing tutorial with Zeiss Axioscan example images is available at <https://www.synapse.org/#!Synapse:syn26958265>. Additionally, we demonstrate processing Zeiss Axioscan, Akoya CODEX and Miltenyi MACSima prototype images in our pipeline example Jupyter notebooks, here: <https://gitlab.com/engje/mplexable/-/tree/master/jupyter>.

3.4.12 Competing Interests

J.W.G. has licensed technologies to Abbott Diagnostics; has ownership positions in Convergent Genomics, Health Technology Innovations, Zorro Bio, and PDX Pharmaceuticals; serves as a paid consultant to New Leaf Ventures; has received research support from Thermo Fisher Scientific (formerly FEI), Zeiss, Miltenyi Biotech, Cepheid (Danaher), Quantitative Imaging, Health Technology Innovations, and Micron Technologies; and owns stock in Abbott Diagnostics, AbbVie, Alphabet, Amazon, Amgen, Apple, General Electric, Gilead, Intel, Microsoft, Nvidia, and Zimmer Biomet. S.L.G. has received research support from Quantitative Imaging. The remaining authors declare no competing interests.

3.4.13 Acknowledgments

We would like to thank former lab members Dr. Moqing Liu, Aesoon Bensen and Lydia Grace Campbell as well as former interns Clayton Peltz and Cole Meyer for assistance with staining and imaging. We would also like to thank Damir Sudar for technical support related to image storage. We appreciate the sample scanning assistance of both Drs. Stefanie Kaech Petrie and Crystal Chaw at the OHSU Advance Light Microscopy Core. The OHSU Biolibrary and OHSU Histopathology Share Resource provided tissue blocks and histology services. We would like to thank Elliot Gray

for assistance with python and shell scripts. We would like to thank Drs. Andreas Bosio and Werner Muller for insightful discussions on the Miltenyi MACSima prototype image analysis. The Miltenyi MACSima Prototype instrument was provided by Miltenyi. Drs. Young Hwan Change and Erik Burlingame contributed ideas for data analysis and image visualization as well as code for RESTORE normalization and registration. Dr. Guillaume Thibault provided the code for watershed segmentation. This work was supported by funding from the Prospect Creek Foundation, Susan G. Komen Foundation, OHSU Foundation. Additionally, J.W.G acknowledges support from NIH/NCI U54 CA209988, NCI SBIR 1R44CA224994-01, NIH/NCI U2C CA233280. The OHSU Knight Cancer Institute Shared Resources are supported by the Knight Cancer Institute Cancer Center Support Grant (NIH/NCI P30CA69533).

3.4.14 *Author Contributions*

Jennifer Eng, Summer L. Gibbs, Koei Chin and Joe W. Gray conceived of the project.
J.E. and K.C. designed the antibody panel, performed experiments and processed image data.
J.E and Elmar Bucher developed free and open-source python tools and Jupyter notebooks for image processing, visualization and analysis.
Zhi Hu performed staining and imaging experiments.
Ting Zheng established quenching and antibody labeling protocols.
J.E. drafted the manuscript and prepared the figures.
S.L.G, K.C and J.W.G directed the project and edited the manuscript.

Chapter 4. USE OF CYCLIC IMMUNOFLUORESCENCE PROTOCOL, MPLEXABLE PIPELINE AND NOVEL ANALYTICAL METHODS FOR CHARACTERIZING TUMOR HETEROGENEITY

The methods and results sections presented herein were originally published in:

- Janiszewska, M, Stein, S, Metzger, O, **Eng, J**, Kingston, N, Harper, N, Rye, I, Alečković, M, Trinh, A, Murphy, K, Marangoni, E, Cristea, S, Winer, E, Krop, I, Russnes, H, Spellman, P, Bucher, E, Hu, Z, Chin, K, Gray, JW, Michor, F, Polyak, K. The impact of tumor epithelial and microenvironmental heterogeneity on treatment responses in HER2+ breast cancer. JCI Insight. 2021 Jun 8; 6 (11):e147617. PMID: 33886505.²³⁹
- Johnson, B, Creason, A, Stommel, J, Keck, J, Parmar, S, Betts, C, Blucher, A, Boniface, C, Bucher, E, Burlingame, E, Chin, K, **Eng, J**, Feiler, H, Kolodzie, A, Kong, B, Labrie, M, Leyshock, P, Mitri, S, Patterson, J, Riesterer, J, Sivagnanam, S, Sudar, D, Thibault, G, Zheng, C, Nan, X, Heiser, L, Spellman, P, Thomas, G, Demir, E, Chang, YH, Coussens, L, Guimaraes, A, Corless, C, Goecks, J, Bergan, R, Mitri, Z, Mills, G, Gray, JW. An omic and multidimensional spatial atlas from serial biopsies of an evolving metastatic breast cancer. Cell Rep Med. 2022 Feb 15;3(2):100525. doi: 10.1016/j.xcrm.2022.100525.²⁰⁸

And also included in the manuscript:

- Combined MYC Activation and PTEN Loss In mammary Epithelium Cells Models Heterogeneity in of Human Triple Negative Breast Cancer. [Doha et al. 2022, in preparation]

4.1 ABSTRACT

Our established cyclic immunofluorescence (CyCIF) protocol (Chapter 2) and image processing and analysis pipeline (Chapter 3) allowed application of our methods in a variety of collaborative projects evaluating the spatial biology of breast cancer and are described herein. First, we used CyCIF to profile tumor epithelial and microenvironmental heterogeneity in HER2+ breast cancer and link it to treatment response. Next, we used CyCIF to contribute to an omic and multidimensional spatial atlas from serial biopsies of an evolving metastatic breast cancer. Finally, we used CyCIF to analyze a mouse model of TNBC driven by combined MYC activation and PTEN loss (MYC;PTENfl) in mammary epithelial cells. Additionally, we used our image analysis pipeline to process and analyze publicly available multiplex ion beam imaging (MIBI) data from 40 human TNBC tissues to show how the MYC;PTENfl mouse models the heterogeneity of human TNBC. Together, these works demonstrate the value of our methods and tools to be applied to diverse samples, species and imaging modalities, capturing biology relevant to treatment response, tumor evolution, and model system validation.

4.2 OVERVIEW

Heterogeneity of tumors both within and between patients is a challenge pre-clinically and clinically, as it has proven difficult to measure, model and treat. We used CyCIF, our mplexable processing pipeline, and our reproducible analytics to examine heterogeneity in three contexts: a cohort of HER2+ breast cancer patients, serial biopsies from a single patient with metastatic ER+ breast cancer, and a mouse model of TNBC which we compared to human TNBC multiplex imaging data. We leverage our tools to address the following biological questions:

1. How does heterogeneity in pretreatment tumor epithelial and microenvironment composition predict response to neoadjuvant TDM-1 in HER+ breast cancer patients?
2. How does tumor composition evolve over time, treatment and metastatic sites and how does microenvironment cell proximity correlate to clinically relevant biomarkers such as proliferation and ER expression in the tumor?
3. In a mouse model of TNBC driven by MYC activation and PTEN loss, do we recapitulate tumor and microenvironment heterogeneity and prognosis of human disease?

Our work contributes to efforts to measure, model and control heterogeneity in order to improve patient stratification and treatment. There is an urgent clinical need to characterize tumor heterogeneity in different contexts. Specifically, our works examine tumor heterogeneity in the three clinical subtypes of breast cancer: HER2+, ER+ and TNBC. We will discuss our work on these three vital questions in cancer therapy through three proof of concept examples, explained in sequence.

4.3 EXAMPLE ONE: CELLULAR PHENOTYPIC HETEROGENEITY AND RESPONSE TO NEOADJUVANT HER2-TARGETING THERAPIES.

4.3.1 *Introduction*

HER2+ breast cancer affects approximately 15% of breast cancer patients and is currently treated with dual blockade of HER2 with trastuzumab and pertuzumab²⁴⁰. The antibody drug conjugate trastuzumab emtansine (T-DM1), which combines the anti-HER2 monoclonal antibody trastuzumab with a potent derivative of the anti-microtubule drug maytansine, is especially valuable in patients with residual disease following neoadjuvant therapy in which it significantly increased disease-free survival compared to trastuzumab alone in phase 3 trials (hazard ratio 0.50, $p < 0.001$)²⁴¹. However, T-DM1 caused more side effects than trastuzumab alone²⁴¹. Therefore, it is important to identify predictors of response to T-DM1 to aid in patient stratification for effective therapy. Furthermore, it is critical to elucidate mechanisms of resistance to T-DM1. Understanding mechanisms of resistance to HER2+ targeted therapy represents an urgent clinical need as relapsed HER2+ breast cancer has only a 36% 5-year survival rate²⁴⁰. Therefore, we analyzed pre-treatment biopsies from HER2+ patients treated with T-DM1 using CyCIF and examined predictors of response.

4.3.2 *Methods*

Cyclic Immunofluorescence staining of tumor tissue was completed as described in sections 3.4.1 to 3.4.6.

Cyclic IF quantification and analysis: Each image acquired during the CyCIF assay was registered based on DAPI features acquired from each round of staining¹⁹⁶. Cellpose, a generalist algorithm for cellular segmentation²³⁴, was used to generate nuclear and cell masks by classifying pixels on the basis of a DAPI or E-cadherin antibody staining, respectively. The following parameters were used for Cellpose segmentation: cells -diameter=30 pixels, flow_threshold=0.6, min_size=113; nuclei - diameter=30, flow_threshold=0, min_size=28. Nuclei with no E-cadherin staining (non-epithelial cells) were expanded by 5 pixels to approximate the cytoplasm. Subcellular regions (e.g., cytoplasm, membrane) were derived from nuclei and cell segmentation results and mean intensity of each subcellular region was extracted using in-house software [<https://gitlab.com/engje/cmIF> now mplexable]. Mean intensity used for downstream analysis was selected for each marker based on its biologically-relevant subcellular region (e.g., CK19 - cytoplasm, Ki67 - nuclei). Since the cell membrane is difficult to segment accurately, for membrane markers such as HER2, the brightest 25% of pixels in the segmented region were used for analysis. The last round DAPI image was used to filter out cells lost during each round of CyCIF staining.

For cell type determination and composition analysis, single-cell mean intensity was used, and threshold was assessed to generate a binary data set. Binary calls for each marker were combined into a gating strategy to define final cell types (Table 4.4). For spatial analysis, nuclear centroids were used to calculate distances between cells. For each cell type A–cell type B pair, the number of cell type B within 75 μm of each single cell of type A was counted, and the mean number of cell type B proximal to cell type A was calculated for each case. For heterogeneity analysis,

Shannon entropy of each case was calculated based on single-cell phenotype frequency of tumor receptor status expression (HER2+, ER+, HER2+ ER+, HER2– ER–) or immune cell types (B cell, CD4+ T cell, CD8+ T cell, and CD68+ macrophage)

Table 4.4. Cyclic immunofluorescence (CyCIF) gating strategy

Combinations of markers related to specific cell types were used to assign individual cells to a particular cell type.

CELL TYPE	GATING
Tumor	CD31-, any cytokeratin+
Luminal tumor	CK19, CK7 or CK8 +
Basal Tumor	CK5 or CK14 +
Mesenchymal Tumor	Vimentin+, cytokeratin +
Immune	CD31- and cytokeratin-; CD45, CD68, CD3 or CD4+
Macrophage	CD68+, CD31-, cytokeratin- (+/- GRNZB and PD1)
B cell	CD20+, CD45+, CD68-, CD31-, cytokeratin-
CD8 T cell	CD8+, CD45 or CD3+, CD20-, CD68-, CD31-, cytokeratin- (+/- GRNZB, PD1 and FoxP3)
CD4 T cell	CD4+, CD8-, CD20-, CD68-, CD31-, cytokeratin- (+/- GRNZB, PD1 and FoxP3)
Endothelial	CD31+
Stromal	CD31-, cytokeratin-, CD45-, CD3-, CD4-, CD68-
Stromal cell types	Defined above, (+/- αSMA, Vimentin, Podoplanin)
Proliferation	Ki67+

4.3.3 Results

To test whether phenotypic heterogeneity could be a better predictor of response to neoadjuvant therapy than metrics based on genetic features, we performed cyclic immunofluorescence (CyCIF) for 22 protein markers on consecutive slides of pretreatment samples from the T-DM1 cohort (Figure 4.1A). This allowed us to classify individual cells within each tissue sample as tumor, stroma, or immune subtype based on marker combinations (Table 4.4). Tumor cells were classified according to their expression of HER2 and ER into four categories (Figure 4.1, B and C). The frequency of these cell types in each case correlated with response to neoadjuvant therapy, with pCR cases having higher frequency of HER2+ cells and no-pCR cases having higher proportion of ER–HER2– cells (χ^2 test $p = 0.022$, Figure 4.1, B–D). Moreover, pCR was also associated with increased fraction of CD8+ T cells among tumor-infiltrating immune cells ($p = 0.016$, Figure 4.1D). Next, we asked if the proximity between distinct cell types could be predictive of response to neoadjuvant therapy (Figure 4.1E). We found that tumors with pCR were characterized by more frequent contacts between tumor cells and tumor cells of luminal HER2+ subtype ($p = 0.004$). Tumors with no-PCR had higher frequency of tumor cells neighboring GZM+ macrophages ($p = 0.023$).

The immune cells in tumors with no-pCR also had fewer contacts with vimentin+ stromal cells (non-tumor, nonimmune, nonendothelial cell gate, $p = 0.048$), potentially representing a mesenchymal and more migratory phenotype, than those in pCR cases. Moreover, immune cells in no-pCR cases had significantly ($p = 0.016$) fewer neighboring FoxP3+ CD8+ regulatory T cells compared with tumor tissues with pCR. Thus, our results suggest that resistance to combined T-

DM1 and pertuzumab treatment might be associated with impaired immune targeting and killing of tumor cells. Using the single-cell CyCIF quantitative data, we tested whether ITH for expression of HER2 and ER protein or the diversity of immune cells within a tumor could stratify the samples according to response to neoadjuvant treatment. Neither of the heterogeneity metrics showed a significant association with pCR (Figure 4.1F).

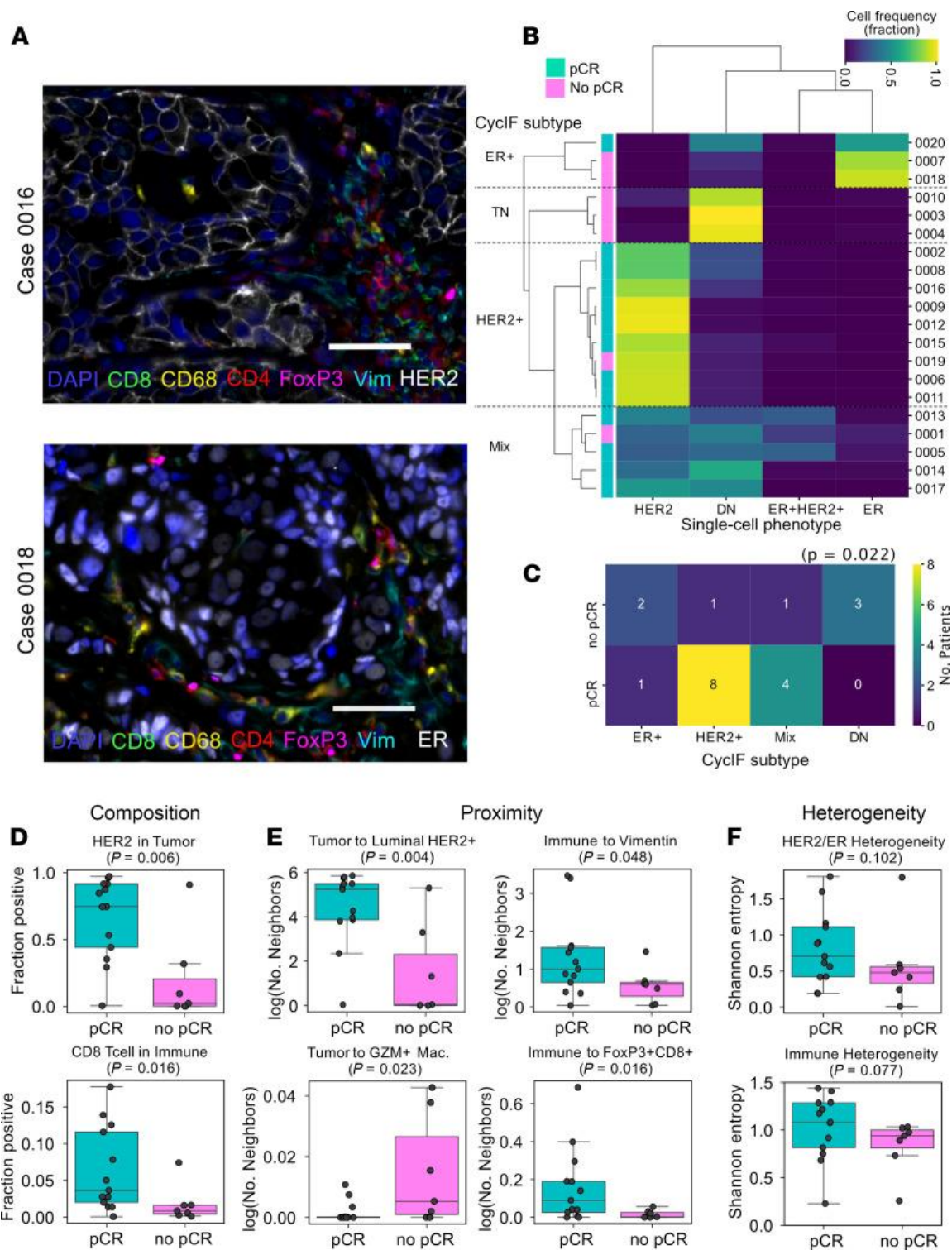


Figure 4.1. Phenotypic single-cell diversity measured by CyCIF.

(A) Representative CyCIF staining in HER2+ case 0016 (top) and ER+ case 0018 (bottom). Vim, Vimentin. Scale bars: 50 μ m. (B) Hierarchical clustering of samples based on HER2 and ER positivity in tumor cells. pCR, pathological complete response; DN, double-negative (ER- HER2-). (C) Association between CyCIF subtype identified in B and pCR. HER2+ and mixed tumors had significantly better response than DN or ER+ (χ^2 , $P = 0.022$). (D–F) Associations between tumor and immune measurements of composition (D), proximity (E), and heterogeneity (F) and pCR. Wilcoxon's rank-sum test P values are shown. GZM, granzyme B.

4.3.4 Discussion

Our group first examined genetic heterogeneity and found no relationship between genotype frequencies (i.e., HER2 amplification and PIK3CA mutation) or heterogeneity and pCR in two cohorts treated with HER2 targeted therapies²⁴². There was no significant correlation between HER2 expression by CyCIF and ERBB2 amplification measured by STAR-FISH, and we did not find significant trends in CyCIF data that were associated with the frequency of PIK3CA-mutant cells²⁴². This is consistent with IHC, as we also did not find significant associations between pretreatment genotype frequency and hormone receptor and HER2 status²⁴².

Interestingly, the association of CD8+ T cell frequency with response suggests that a pre-existing immune response is present in responders. It also agrees with the literature on TILs showing pCR associated with immune infiltrate in HER2+ breast cancer⁴⁰. A major mechanism of action of trastuzumab is antibody-dependent cellular cytotoxicity, ADCC, whereby target cells coated with antibody are lysed by effector cells, and our findings could imply that T-DM1 therapy and hence ADCC is more effective in immune-activated tissues. This suggests a strategy to stimulate an anti-tumor immune response in combination with HER2+ targeted therapies. Trials evaluating anti-PD-L1 and anti-PD1 immune checkpoint inhibitors in metastatic HER2+ breast cancer have been negative or showed very modest benefit of the combination²⁴⁰. More work needs to be done to describe and target immunosuppressive mechanisms of resistance to HER2 therapy.

Contrary to our original hypothesis, intratumor heterogeneity was not predictive of pCR, either at the genetic, tumor protein or tumor microenvironment level. This could be due to small sample size, as there was a trend ($p = 0.07$) for higher immune heterogeneity in responders versus non-responders. Also, higher tumor heterogeneity reflected in the diversity index correlated to larger populations of ERBB2-amplified cells in genetic data and greater numbers of HER2+ cells in CyCIF data. Quantification of heterogeneity using the Shannon diversity index only measures the diversity of cell types in the tissue, and therefore there is no straightforward link between heterogeneity and “good” or “bad” cell types. In these data, heterogeneity correlated with greater numbers of tumor cells expressing the drug target and greater fractions of CD4 and CD8 T cells, but that would not be the case in another data set. Therefore, heterogeneity must be precisely defined and caution must be used in comparing heterogeneity measures across datasets.

Our spatial analysis revealed an increase in homotypic contacts between tumor cells in responders, reflecting a phenotype of highly compartmentalized tumor and stromal areas in the tissue. This phenotype has also been observed in TNBC and linked to good outcomes¹¹¹. This compartmentalization could also explain the increased contacts between immune cells and vimentin-positive stromal cells in the tissue, as the tumor compartmentalization necessarily decreases tumor-immune contacts and generally increases immune-stromal contacts. However, due to the immune-modulatory roles of CAFs, the proximity of immune cells and vimentin+ CD45- CD31- putative CAFs is an interesting finding. While CAFs are generally thought to be immunosuppressive⁸⁰, deletion of α SMA+ myofibroblasts in pancreatic cancer lead to invasive, undifferentiated tumors with increased hypoxia, EMT and immune suppression as reflected in increased FoxP3+ Tregs⁹². The responders in this study had tumors with a well-differentiated,

compartmentalized tumor phenotype, higher levels of CD8 T-cell infiltration and increased contacts between immune cells and fibroblasts, fitting with a model of fibroblasts promoting tumor differentiation and immune activation. However, given the small sample sizes and the limited markers for CAF characterization, our findings should be verified with additional samples and deeper CAF profiling.

4.4 EXAMPLE TWO: TUMOR EVOLUTION AND SPATIAL BIOMARKER EXPRESSION PATTERNS IN METASTATIC ER+ BREAST CANCER

4.4.1 Introduction

Precision medicine has focused on predicting patient drug response based on information derived from diagnostic biopsies or primary tumor resections, however, some patients do not respond despite meeting selection criteria, and in the metastatic setting, durable control is difficult²⁴³. Critics of precision oncology have pointed to tumor evolution and intratumor heterogeneity as limiting the ability of molecularly targeted therapies to permanently control advanced disease²⁴³. However, understanding the mechanisms underlying failures in precision medicine is an important task which can lead to development of new treatment options²⁴⁴. Therefore, our group sought to characterize the evolution of an individual's tumor by generating an omic and multidimensional spatial (OMS) atlas from four serial biopsies of metastatic breast cancer during 3.5 years of therapy²⁰⁸. Our specific contribution was to generate multiplex imaging data using CyCIF, produce and share data with our mplexable pipeline and utilize our reproducible image analytics to characterize tumor tissue composition and spatial organization in the serial biopsies.

Methods

Cyclic Immunofluorescence staining of tumor tissue was completed as described in sections 3.4.1 to 3.4.6.

Table 4.5. Cyclic immunofluorescence antibodies

Biomarker	Vendor	Clone	Probe
α SMA	Santa Cruz	1A4	AF488
CD3	abcam	EP449E	AF750*
CD31	abcam	EPR3094	AF647
CD4	abcam	EPR6855	AF647
CD44	abcam	EPR1013Y	AF647
CD45	abcam	EP322Y	AF647/AF750*
CD68	Biolegend	KP1	AF555*
CD8	abcam	C8/468	AF555
CK14	abcam	LL002	AF555*
CK17	abcam	EP1623	AF488
CK19	Biolegend	A53-B/A2	AF750*
CK5	abcam	EP1601Y	AF750*
CK7	abcam	EPR1619Y	AF488/AF555

Ecad	abcam	EP700Y	AF750*
ER	abcam	EPR4097	AF647
HER2	Santa Cruz	3B5	AF555
Ki67	CST	D3B5	AF647
PCNA	CST	PC10	AF488
pHH3	CST	D2C8	AF488
Vimentin	CST	D21H3	AF488
Collagen I	abcam	EPR7785	AF750*
Collagen IV	ThermoFisher	1042	AF647
EGFR	CST	D38B1	AF555
CK8	abcam	EP1628Y	AF488
*Fluorophore labeling of purified antibody performed in-house			

Quantification and Analysis: Each image acquired was registered based on DAPI features acquired from each round of staining¹⁹⁶. Cellpose, a generalist algorithm for cellular segmentation, was used to generate nuclear and cell segmentation masks with a pre-trained neural network classifier²³⁴. Extracted single-cell features included centroids and mean intensity of each marker from its biologically relevant segmentation mask (e.g., Ecad_Cytoplasm, Ki67_Nuclei). The last round DAPI image was used to filter out cells lost during each round of cyclic immunofluorescence staining. For cell type determination and composition analysis, single cell mean intensities from each biopsy were batch corrected using the ComBat algorithm²³⁰. ComBat was used to adjust the mean and variance of fluorescence intensity on control tissue-microarrays (TMAs) that were stained with each biopsy, and the same adjustments were applied to the corresponding biopsies. Eighteen markers were selected for clustering; some markers were excluded due to tissue loss in the TMA controls. Principal component analysis was performed with scanpy to reduce dimensionality, and Umap (<https://arxiv.org/abs/1802.03426>) was run on the top 17 principal components to calculate a nearest neighbor graph based on the 30 nearest neighbors. Leiden clustering was performed on the nearest neighbor graph to define clustering-based cell types²⁴⁵. The Leiden clustering resolution of 0.5 was selected based on appropriate clustering of technical replicates in the control TMAs. Immune, endothelial, and stromal cells were identified by manual thresholding and gating. Endothelial cells were defined as CD31+, immune cells were either CD45+ or CD68+ and CD31-, and stromal cells were cytokeratin-, E-cadherin-, CD31-, CD45-, and CD68-. Tumor was defined as cytokeratin+, and proliferating cells were Ki67+. Cell segmentation borders of manually defined cell types were visualized on the images using napari software [napari.org]. To calculate distance to extracellular matrix proteins, a threshold was applied to create a pixel mask of positive staining. The distance from each nuclear centroid to the nearest mask pixel was measured. Cells were grouped into bins of 0–25 microns, 25–50 microns, and 50–75 microns from the mask, and the intensity distributions of cells (n = 2 to 32324 cells) were compared using ANOVA implemented in SciPy²³⁸. Significance was assigned to Bonferroni-corrected p values < 0.001.

4.4.2 Results

Tumor and stromal compositions and organizations of Bx1–Bx4 plus control biospecimens were assessed using CyCIF (Figure 4.2A)^{107,246}. This joint analysis revealed 17 tumor and stromal clusters (Figure 4.2B and Figure 4.3A–E). Three of the stromal clusters (11, 13, and 14) and five of the tumor clusters (0, 1, 5, 7, and 9) comprised major subpopulations in Bx2–Bx4. The three stromal clusters were identified as fibroblast-like cells that differed in levels of vimentin (VIM; cluster 11: 3.23 other clustered means, 13: 2.63, 14: 4.93). Endothelial cells (CD31) and macrophages (CD68) were excluded from cluster analysis because of loss of the normal breast and tonsil tissues needed for normalization during staining; their presence was confirmed using manual gating (Figure 4.3F). All tumor clusters expressed CK7/CK19 but different levels of ER, EGFR, and CK8 (Figure 4.2B). An additional proliferative cluster, cluster 16, appeared in Bx3 and Bx4, comprised of tumor and stromal cells expressing high levels of Ki67 (31.33) and/or PCNA (2.43). Spatial analyses indicated that tumor cells formed nests surrounded by immune, fibroblast, and endothelial cells as well as collagen I and collagen IV deposits. This was observed in all biopsies but was pronounced in Bx3. Quantitative analyses of nuclear ER and PCNA expression in Bx2–Bx4 as a function of distance to collagen I-rich tumor nest boundaries showed that cells expressing higher levels of ER and PCNA were closest to these boundaries and other stromal compositions (Bx2: mean ER intensity at 0–25 mm from collagen I = 780, 50–75 mm = 463, $p < 0.001$; Bx3: 0–25 mm = 1,058, 50–75 mm = 600, $p < 0.001$; Bx4: 0–25 mm = 1687, 50–75 mm = 1,105, $p < 0.001$. Bx2: mean PCNA intensity at 0–25 mm = 745, 50–75 mm = 218, $p = 0.17$; Bx3: 0–25 mm = 948, 50–75 mm = 567, $p < 0.001$; Bx4: 0–25 mm = 713, 50–75 mm = 406, $p < 0.001$; Figure 4.2C and Figure 4.3G); p values describe differences in mean intensities between distances.

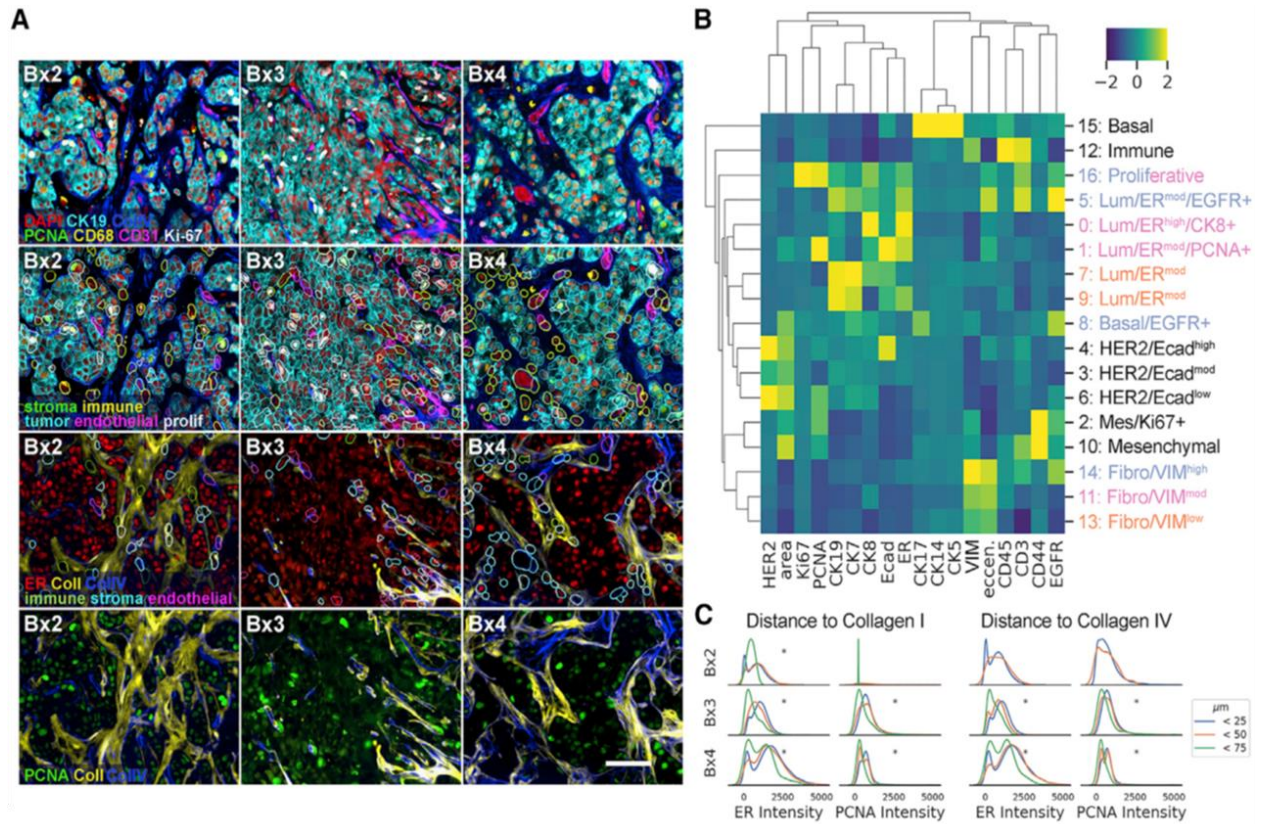


Figure 4.2. Monitoring tumor and stromal responses to therapy using CyCIF.

(A) Example images of antibody staining overlaid with segmentation borders, colored by cell type. Scale bar, 50 mm. (B) Heatmap of mean Z-scored intensity of unsupervised Leiden clustering (resolution, 0.45) on single-cell mean intensity of biopsies and control tissues and cell lines, with annotations on the right. Lum, luminal; Mes, mesenchymal; Fibro, fibroblast. Colored row labels indicate which biopsy was most dominant for each cluster: Bx2 (orange), Bx3 (blue), or Bx4 (pink). Cluster 16 is evenly split between Bx3 and Bx4. (C) Single-cell mean intensity distributions of ER and PCNA staining of cells 0–25, 25–50, and 50–75 mm from positive collagen staining. Asterisks indicate significant ($p < 0.001$) differences in mean intensity between distances (ANOVA).

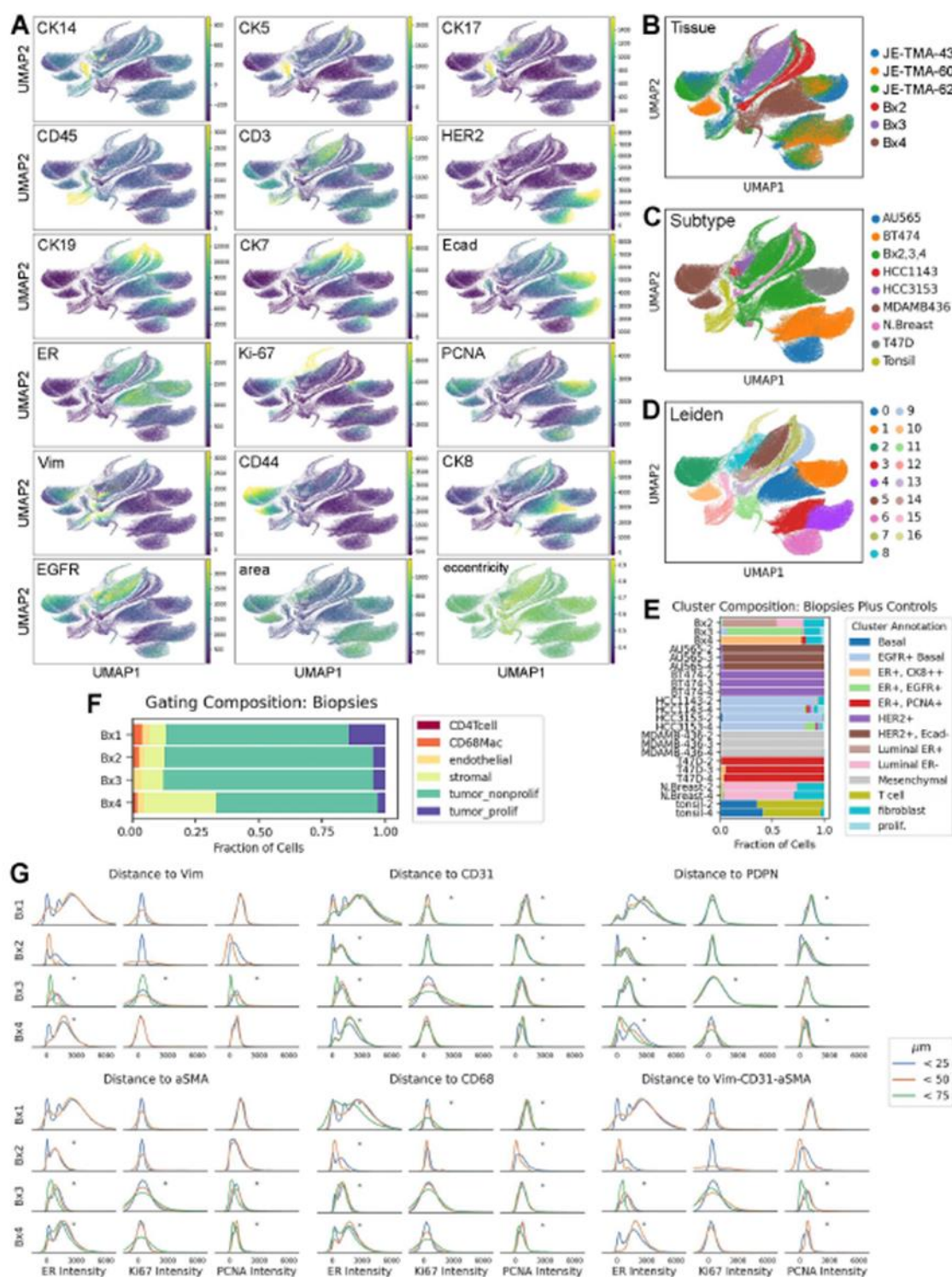


Figure 4.3. Additional CyCIF Analyses, Related to Figure 4.2.

(A) Umap projection of 18 features used for clustering, colored by feature. (B) Tissues included for cluster analysis. Controls for each biopsy are as follows. Bx2: JE-TMA-43, Bx3: JE-TMA-60, Bx4: JE-TMA-62. TMAs were used for normalization, so mixing of cells from all TMAs indicates successful normalization. (C) Umap colored by cell lines and normal tissues included in clustering. (D) Umap projection colored by Leiden clustering result. (E) Composition of biopsies and controls by annotated Leiden cluster. Number following cell line and tissue names indicates the biopsy

that was paired with those controls. (F) Composition of biopsies based on manual thresholding and gating (see methods). (G) Intensity of ER, Ki67, and PCNA at 0-25, 25-50, and 50-75 μm away from various markers. Vim = vimentin, in mesenchymal cells such as fibroblasts; αSMA = Alpha Smooth Muscle Actin expressed in fibroblasts, pericytes and myoepithelial cells; PDPN = podoplanin, expressed in fibroblasts and lymphatic vessels; CD31 = cluster of differentiation 31, expressed in endothelial cells; CD68 = Cluster of Differentiation 68, expressed by monocytes and macrophages; Vim-CD31- αSMA = pixels positive for any of vimentin, CD31, αSMA . Asterisks indicate significant ($p < 0.001$) difference in mean intensity between distances (ANOVA).

4.4.3 Discussion

Our CyCIF analysis revealed tumor and stromal changes across biopsies in progressive metastatic ER+ disease that was supported by other multi-omic assays on the same samples. Multiplex IHC was also completed on the same biopsies, and both platforms revealed decreased CD45+ immune infiltrate in Bx2 and decreased macrophages in Bx4. Gene expression profiling indicated a luminal A intrinsic subtype for biopsies 1, 2 and 4, which was reflected in luminal cytokeratins CK7, 8 and 19 and ER expression in CyCIF data from those samples. Bx3 was luminal B subtype, and we observed EGFR expression in Bx3, which is often associated with luminal B tumors²⁴⁷. Our CyCIF data revealed dense collagen I fibers in all biopsies (Figure 4.2A), which were shown to be associated with tamoxifen resistance in ER+ breast cancer patients²⁴⁸. Fibronectin, another ECM component, has been mechanistically linked to endocrine therapy resistance through binding its main receptor, $\beta 1$ -integrin, on tumor cells and activation of MAPK/ERK and PI3K/AKT pathways²⁴⁹. Interestingly, ER α colocalizes with $\beta 1$ -integrin at the membrane and fibronectin binding stabilizes ER and strengthens its transcriptional activity²⁵⁰. This mechanism might explain the significant proximity between ER-high tumor cells and ECM components in the CyCIF data. Similar analyses in larger cohorts are needed to validate these findings.

4.5 EXAMPLE THREE: A MOUSE MODEL OF TNBC RECAPITULATES HUMAN DISEASE SUBTYPES AND SURVIVAL. MULTIPLEX IMAGING DEFINES TUMOR MICROENVIRONMENT SUBTYPES IN TNBC MOUSE MODEL AND HUMAN TNBC.

4.5.1 Introduction

TNBC is a subtype of breast cancer characterized by lack of ER, PR and HER2 receptor expression which has been shown to exhibit substantial heterogeneity, both between patients^{251, 129} and within individual tumors²⁵². As discussed above, intratumor heterogeneity is a major constraint to durable control of disease. For mechanistic understanding and development of therapeutic strategies to control heterogeneity, pre-clinical models are needed that recapitulate human disease. Genetically engineered mouse models (GEMMs) offer a robust platform for elucidation of mechanisms of disease development, progression and drug response²⁵³. TNBC shows frequent MYC activation (60%) and PTEN deletion (40%) [in preparation, Doha 2022]. MYC expression predicts poor prognosis in primary breast cancer and is associated with outgrowth of metastatic lesions^{254, 255}. Therefore, our collaborators generated a mouse model with bi-allelic PTEN deletion and one MYC allele constitutively expressed at physiologically relevant levels under the beta-lactoglobulin

promoter specific to the mammary gland (ROSA-LSL-MYC; PTEN^{flox/flox}; Blg-CRE, abbreviated MYC;PTEN^{fl}). The MYC:PTEN^{fl} mouse tumors exhibited distinct histologies, termed stroma-rich (SR) and stroma-poor (SP), that recapitulated human histology and gene expression [in preparation, Doha 2022]. We analyzed mouse tissue composition and heterogeneity using CyCIF and compared our results to a human dataset generated with MIBI^{111,256}.

4.5.2 Methods

CyCIF staining of tumor tissue was completed as described in sections 3.4.1 to 3.4.6.

Image Processing: Quality control, image registration, autofluorescence subtraction, cell segmentation, matching of cell and nuclear masks, feature extraction and feature filtering were performed with our mplexable software²⁵⁷. Nuclei segmentation was run using the Cellpose²³⁴ algorithm on DAPI stained images, and cell segmentation was run using Ecad staining and 5-pixel expansion of Ecad- nuclear masks.

CyCIF Analysis: Code used for analysis is available at <https://github.com/engjen/MYC-PTENfl-mouse>. Single-cell mean intensity values of autofluorescence-subtracted images were selected from either nucleus or cytoplasm masks for each marker, depending on expected intracellular distribution. Areas of floating tissue that created bright imaging artifacts and air bubbles that created dark artifacts were manually circled using the napari image viewer and excluded. Background subtraction was performed on markers with high background: CD31, CD45, CD8, ColIV, FoxP3, CD103, CD11b and CD11c. Filtered, background subtracted data were clustered with scanpy²³¹. Two morphology features (nuclear area and nuclear eccentricity) and 22 markers were used for clustering. A Umap embedding was generated using 15 neighbors, and the Leiden algorithm was used for unsupervised clustering (resolution 0.6), resulting in 20 cell types. Inspection of cluster results in images revealed that 3 of the clusters were due to imaging artifacts and excluded. The remaining clusters were evaluated on the images and annotated.

We then performed manual gating to verify our annotated-cluster cell type calling. A threshold was set for each gating marker based on positive pixel patterns in images. Endothelial cells were defined as CD31⁺. Epithelial cells were positive for one or more of Ecad, EpCAM and CK5 and CD31⁻. Immune cells were CD45⁺ CD31⁻ and epithelial marker negative. Stromal cells were all non-endothelial, non-epithelial, non-immune segmented nuclei. Three cores had pMYC positive tumor cells negative for Ecad, EpCAM and CK5 (I11, G09 and H11). In these tissues, tumor cells were defined as any cells negative for all stromal and immune gating markers (CD31, CD45, Vim, aSMA, ColIV, ColIV, Gal3).

Cell frequencies were calculated for gated epithelial, immune and stromal cells in each tissue core. Endothelial cells were rare and added to “stromal” cells for subtyping. Tissues were clustered with the Leiden algorithm, resolution = 0.5, resulting in six subtypes. Subtypes were annotated as stroma-rich, immune-rich (cluster 2 and 4), stroma-rich, immune-poor (cluster 0) and stroma poor (cluster 1, 3 and 5). Chi-squared analysis was performed on cell frequency subtype versus histology subtype (i.e., stroma-rich and stroma-poor based on H&E evaluation). Most of the CyCIF stroma poor samples were also stroma poor by histology (6/9).

To compare expression of markers in each subtype, mean marker expression was calculated in each tissue and compartment (epithelial or stromal, i.e., immune, endothelial and non-immune stromal cells). Distributions were visualized as boxplots and the Kruskal Wallis H test of Mann Whitney U test, implemented in `scipy`²³⁸, were used to test for significant differences in the median expression of markers between groups. To compare detailed cell types (14 annotated cell types from Leiden clustering, described above) between subtypes, the frequency of each cell type in each subtype was calculated and displayed as a bar plot. Samples were also clustered hierarchically based on the z-score of cell abundances.

For human tissue samples, we obtained a publicly available multiplex ion beam imaging (MIBI) dataset²⁵⁶ at <https://github.com/aalokpatwa/rasp-mibi>. We used `scanpy` for single-cell clustering. Thirty-two markers were used for generating a Umap embedding with 15 neighbors. Unsupervised clustering was then performed with the Leiden algorithm (resolution = 0.6) resulting in 24 cell types. These were annotated as epithelial, immune or stromal. Inspection of cluster results versus images revealed that some clusters contained mixed immune and stromal cells. Therefore, epithelial clusters were used to define epithelial cells, and CD45 and CD31 were used to manually gate immune and endothelial cells within the stromal/immune clusters. Tissues were clustered on cell type frequencies with the Leiden algorithm, resolution = 0.1, resulting in three subtypes, annotated as stroma-rich, immune-rich (cluster 0), stroma-poor (cluster 1) and stroma-rich, immune-poor (cluster 2). Mean marker expression was calculate in each tissue and compartment, visualized and statistically evaluated as described for mouse data, above. For survival analysis, the `lifelines`²⁵⁸ python software was used. Kaplan-Meier estimates were generated and plotted for overall survival and recurrence-free survival. The log-rank test was used to test for significant survival differences between the subtypes.

4.5.3 Results

To determine how closely tumor microenvironment heterogeneity, cell type frequency and biomarker expression in our TNBC mouse model aligned with human TNBC, we performed cyclic immunofluorescence staining of a tissue microarray (TMA) from our TNBC mouse model, which included both MYC;PTENfl subtypes; SR and SP (Figure 4.4A). Indeed, SR tumors showed more stroma infiltration than SP (Figure 4.4A). Image processing and single-cell segmentation was performed with the `mplexable` software [<https://doi.org/10.1101/2021.11.29.470281>]. Gating on cell-type specific markers determined cell frequencies of epithelial, immune and stromal cells in each tissue, and showed SP as enriched with epithelial cells compare with SR (Figure 4.4B, Figure 4.5). For human tissue samples, we obtained a publicly available human TNBC multiplex ion beam imaging (MIBI) dataset²⁵⁶. We compared cell frequencies and expression of markers between our TNBC mouse model and human TNBC MIBI data.

To compare cell type frequency, human and mouse tissues were clustered with the Leiden algorithm, resulting in stroma poor, stroma-rich-immune-rich and stroma-rich-immune-poor subtypes (Figure 4.4C-D, and Figure 4.6, Figure 4.7). These subtypes were significantly associated with mouse histology subtypes (Figure 4.6G), in which stroma poor samples by cell type frequency were also SP by histology (6 out of 9) and SR histology are subdivided into immune rich and immune poor by cell type frequency. Importantly cell type distribution among these three subtypes was similar between human TNBC and mouse TNBC, with mean abundance of cell types in each

subtype is recapitulated in mouse compare with human tissues (Figure 4.4E). Interestingly, these subtypes are prognostic in human MIBI data, with stroma poor having worst survival compare with the two stroma rich subtypes, consistent with our TNBC mouse model (Figure 4.4F). However, these subtypes show no significant difference in recurrence-free survival in human MIBI data (Figure 4.7I).

Next, we compared marker expression in each subtype, and we found that subtypes significantly differ in mean expression of pan-immune, Treg, dendritic, endothelial and proliferation markers in the stroma (Figure 4.4F and G, P-values determined by Kruskal–Wallis H test) as well as other markers (Figure 4.8A and B). The distributions of markers among subtypes were similar in across human and mouse tissue (Figure 4.4G and H). In both human and mouse tissues, the stroma-rich-immune rich subtype had greater pan-immune CD45, T regulatory FoxP3 and dendritic CD11c, consistent with mIHC and gene expression analyses showing increased APC signatures in SR tumors. The MIBI panel included HLA-DR, an MHC Class II molecule expressed on antigen presenting cells, which was elevated in SR_IR tumors (Figure 4.8). MIBI data also showed higher stromal expression of immune checkpoint molecules PD1, PD-L1 and IDO in SR_IR tumors (Figure 4.8). In mouse SP tumors, the sparse stromal cells expressed more endothelial marker CD31 and fibroblast markers α SMA and Vimentin (Figure 4.4G, Figure 4.8). Human SP tumors also showed high CD31 expression in SP tumors (Figure 4.4H). Stromal proliferation measured by Ki67 expression differed between the subtypes in mouse and human, but there was no difference in epithelial Ki67 (Figure 4.4G-H, Figure 4.8). Interesting and consistent with the previous bulk RNA analysis, human and mouse SR_IR also showed higher expression of mesenchymal markers Vimentin and α SMA, higher nuclear eccentricity and lower expression of epithelial markers E-cadherin and EpCAM in epithelial cells compare with SP (Figure 4.8C and D). This indicated that mouse SR represented human TNBC mesenchymal subtype where mouse SP represent human non-mesenchymal TNBC subtype, at the molecular and imaging level. S100A6 was elevated in the SR_IR mouse epithelial cells, while β -catenin was elevated in the SR_IR human epithelial cells (Figure 4.8). Since previous work showed S100A6 promoting EMT through β -Catenin in pancreatic cancer cell lines²⁵⁹, the S100A6-high/ β -Catenin-high phenotype in mouse/human SR_IR tumors may reflect a shared mechanism of EMT. Comparison of marker expression in histology subtypes recapitulated the IHC staining, with higher CK5 and α SMA in the SR histology than the SP (Figure 4.8). Phospho-MYC was higher in the SP histology subtype epithelial cells, and showed a trend ($p < 0.1$) towards differential expression across the three cell type frequency subtypes. Unsupervised clustering of single cells in mouse TMA was performed with the Leiden algorithm resulted in 18 cell types (Figure 4.4I-K, Figure 4.5 and Figure 4.7A and B). The unsupervised clustering-based cell types captured the intratumoral heterogeneity observed in the images (Figure 4.4I, J). Mouse TMA tissues were hierarchically clustered based on the z-scored cell abundance of cell types, revealing heterogeneity within the stroma-poor subtype (Figure 4.4K). Most of the SR_IR tissues clustered together, driven by high abundance of fibroblasts and immune cells in the tissue (Figure 4.4K). SR_IR tumors also contained E-cadherin negative epithelial cells, consistent with a mesenchymal phenotype (Figure 4.4K). The SP tumors formed separate clusters with distinct enrichment of CK5-high or E-cadherin-high epithelial cells, which provides further evidence for SP tumors recapitulating multiple human TNBC subtypes (Figure 4.4K).

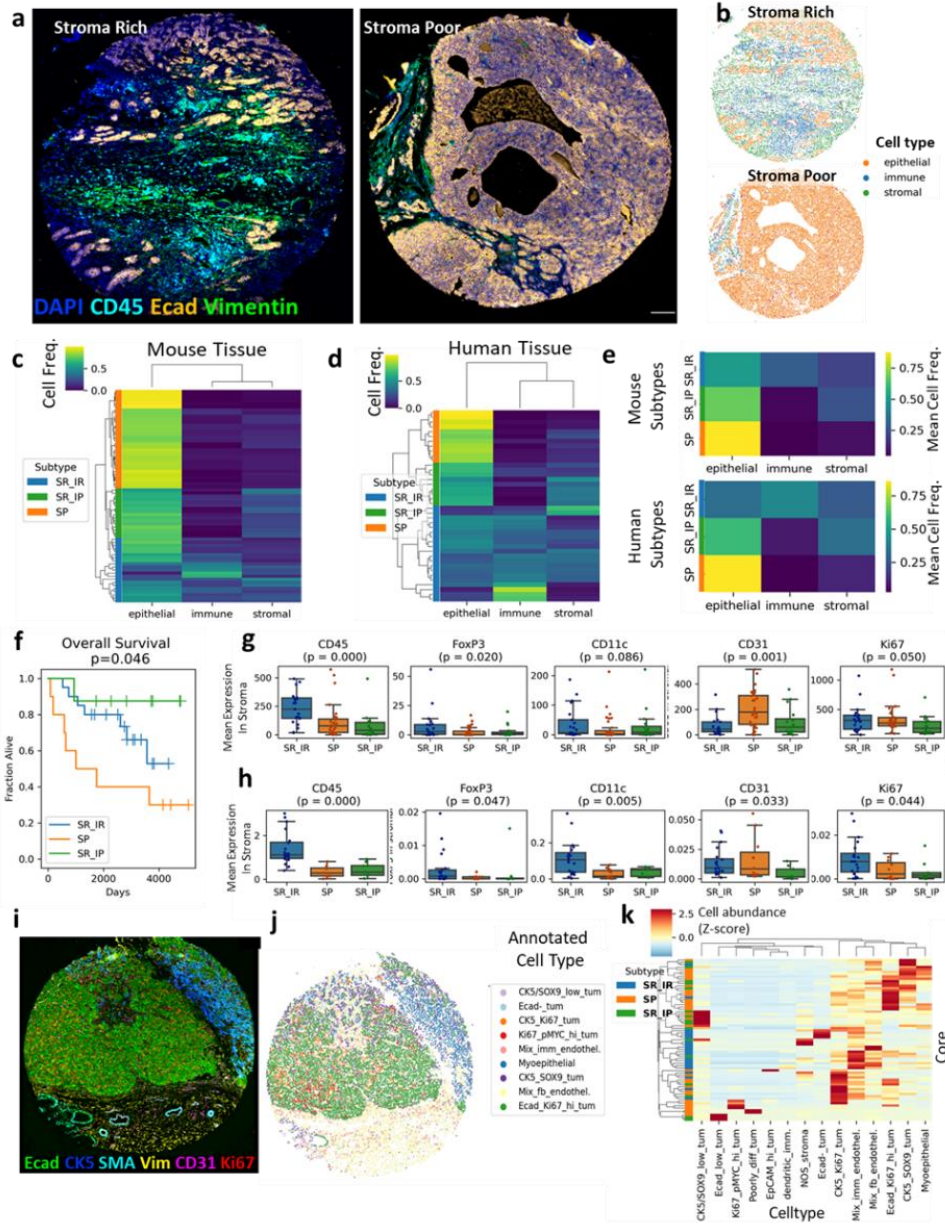


Figure 4.4. Tumor Microenvironment Subtypes Defined by Multiplex Imaging in Mouse and Human TNBC.

a. Cyclic immunofluorescence staining of representative tissue microarray (TMA) cores of stroma rich (left) and stroma poor (right) histology subtypes. Scale bar = 130 μ m. b. Cell type calling defined by gating on cores in (a). c. Heatmap of cell type frequency in each mouse tumor in TMA. d. Heatmap of cell type frequency in each region of interest in human tissues imaged with multiplex ion-beam imaging (MIBI). e. Leiden clustering of mouse (top) and human (bottom) samples resulted in three subtypes: stroma-poor (SP, orange), stroma-rich, immune-rich (SR_IR, blue) and stroma-rich-immune-poor (SR_IP, green) with similar mean abundance of cell types. f. Stroma-poor human disease has shorter overall survival than stroma-rich subtypes, Log-rank $p=0.046$. g. Mouse subtypes differ in expression of pan-immune, Treg, dendritic, endothelial and proliferation markers in the stroma. h. Human subtypes show similar patterns of expression of pan-immune, Treg, dendritic, endothelial and proliferation markers in the stroma. g-h. P-values determined by the Kruskal–Wallis H test. i. Six-color overlay shows intratumoral and stromal heterogeneity. j. Unsupervised-clustering based detailed cell types reflect the spatial heterogeneity observed in (i). k. Hierarchical clustering of mouse samples based on detailed cell types reveal heterogeneity in stroma-poor tumor epithelial state.

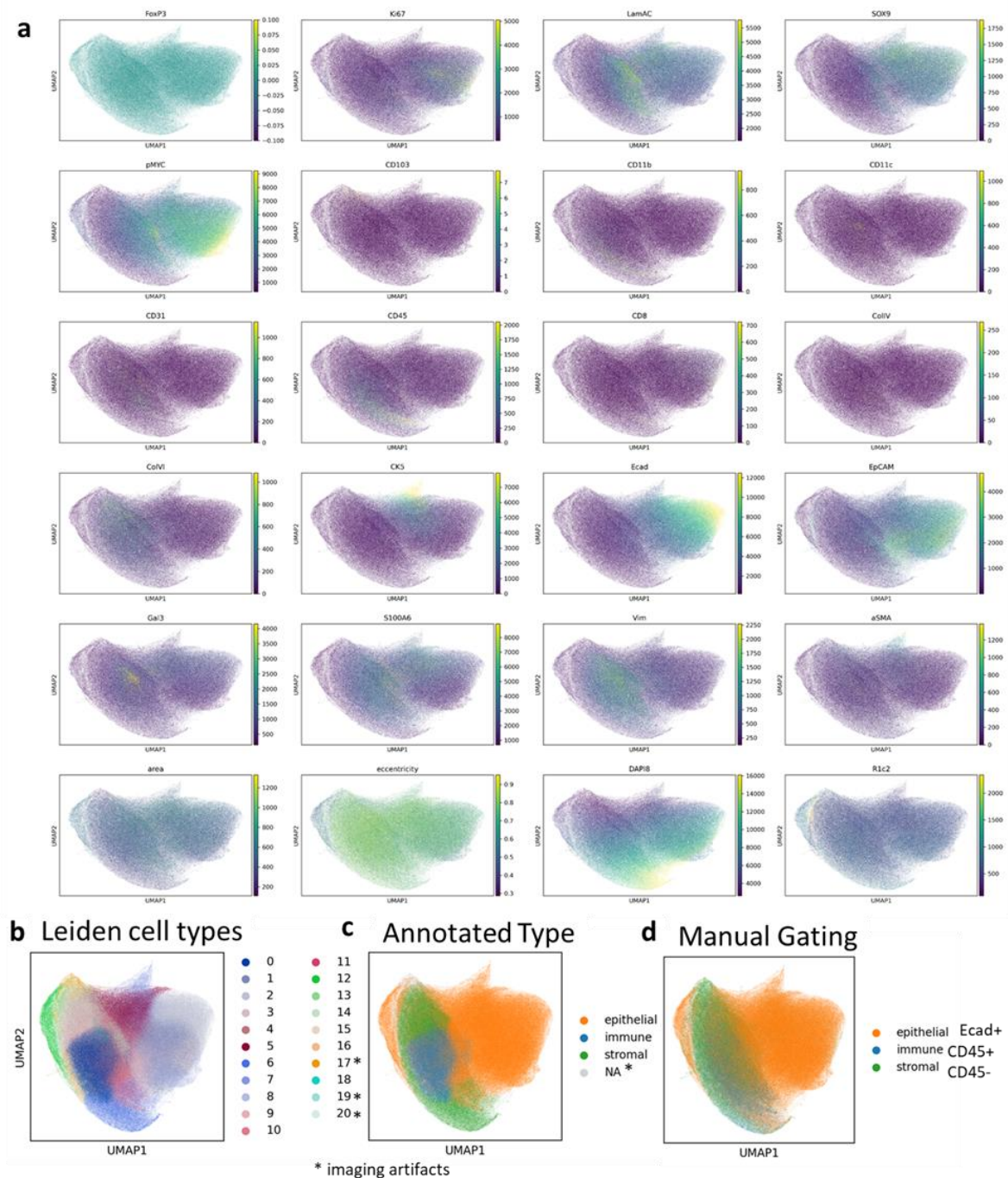


Figure 4.5. Unsupervised clustering and gating-based cell types.

a. UMAP projection of single cells in mouse TMA, colored by expression of each marker or feature used in unsupervised clustering. b. Unsupervised clustering with the Leiden algorithm resulted in 21 cell types, 18 of which were subsequently annotated and 3 of which were imaging artifacts. c. Annotated cell types were grouped into epithelial and stromal categories and projected back on the UMAP. d. Manual gating of epithelial and stromal cell types projected on the UMAP shows good agreement with unsupervised clustering based epithelial annotations in (c), but shows that unsupervised non-epithelial clusters in (c) are mixtures of CD45+ and CD45- cells.

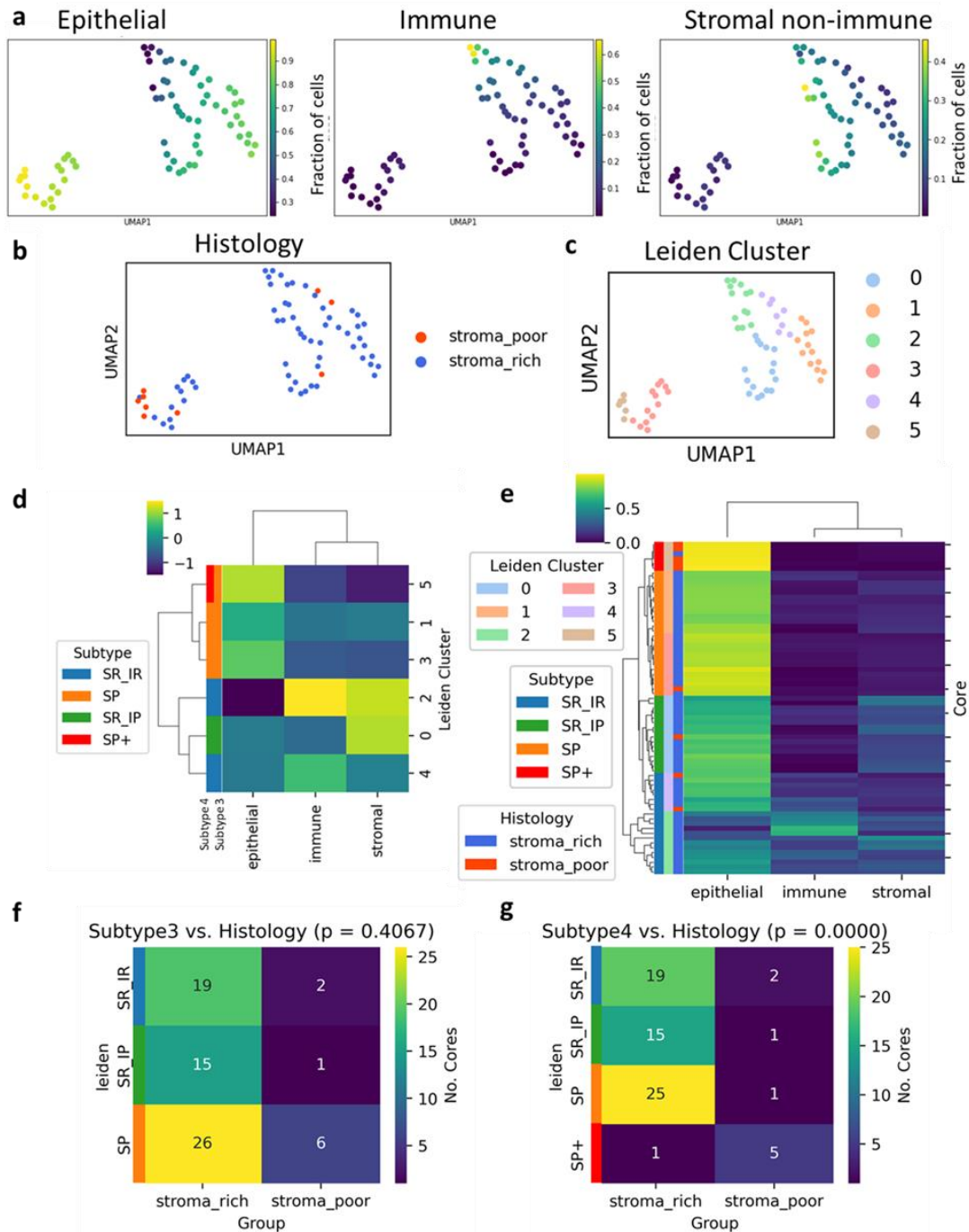


Figure 4.6. Cell frequency-based subtyping in mouse TMA tissues.

a. A k-nearest neighbors graph using 5 nearest neighbors was calculated to construct a UMAP embedding of each triple-negative mouse tumor in TMA, shown colored by frequency of epithelial, immune or non-immune stromal cells. b. Histology based subtypes projected on same UMAP in (a). c. Unsupervised clustering with the Leiden algorithm (resolution 0.5) resulted in six clusters. d. Hierarchical clustering of mean cell type frequency of six Leiden clusters from (c) produced 3 and 4 annotated subtypes: Very stroma poor (SP+), stroma poor (SP), stroma-rich, immune-rich (SR_IR) and stroma-rich, immune poor (SR_IP). e. Heatmap of cell type frequencies in mouse TMA tissues showing relationship between Leiden clusters, subtypes, and histology subtypes. f. Three subtypes based on cell type frequency are not significantly associated with histology subtypes, Chi-squared $p=0.4$. g. Adding a fourth subtype, very stroma-poor (SP+) shows significant association with histology subtypes, Chi-squared $p < 0.0001$.

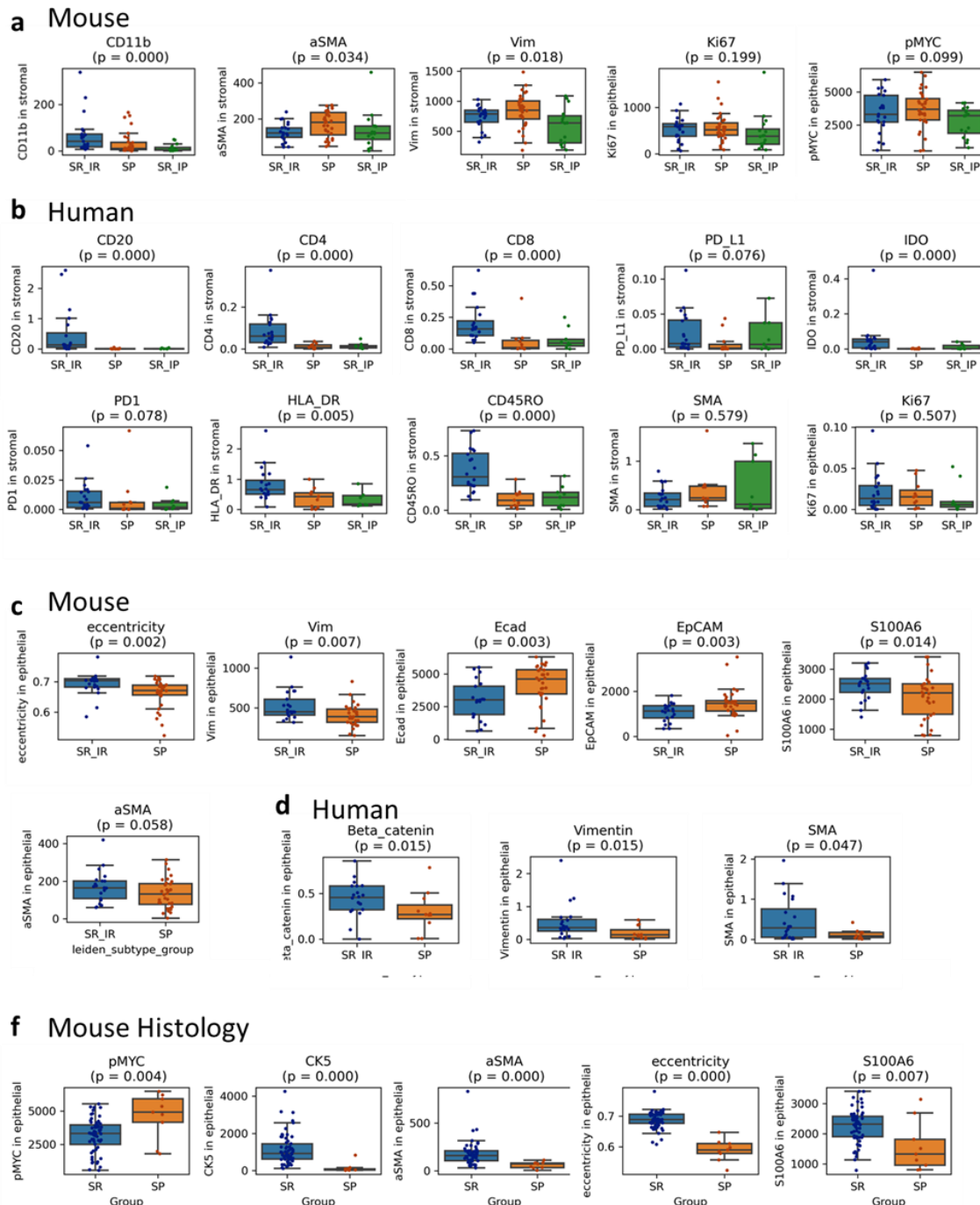


Figure 4.8. Comparison of Marker Expression Across Subtypes.

a. Mouse subtypes determined by cell type frequency differ in expression of myeloid-lineage marker CD11b, and activated fibroblast/pericyte marker alpha-SMA and mesenchymal marker Vimentin in stromal cells, but not in Ki67 or pMYC in epithelial cells. b. Human subtypes differ in expression of lymphoid, immune checkpoint, memory and antigen presentation markers, but not SMA in stromal cells or Ki67 in epithelial cells. a-b. P-values determined by the Kruskal–Wallis H test. c. Comparison of stroma-rich, immune rich (SP_IR) and stroma-poor (SP) subtypes show higher expression of mesenchymal markers in SR_IR epithelial cells. d. Human SR_IR also shows higher expression of mesenchymal markers in epithelial cells. f. Mouse histological subtypes show increased basal marker expression in stroma-rich (SR) and increased phospho-MYC expression in stroma-poor (SP). c, d, f. P-values determined by Mann–Whitney U test.

4.5.4 Discussion

Altogether, our multiplex imaging analysis indicate that our mouse model shares tumor microenvironment patterns that present in human TNBC. Limitations of our analysis include the limited number of overlapping markers (Ki67, CD45, CD11b, CD11c, FoxP3, CD31, Vimentin, α SMA, basal keratin 5/6) between the MIBI human data and CyCIF mouse data. Therefore, we classified tumors using epithelial, immune, endothelial and non-immune stroma content rather than matching specific cell types across platforms. However, these tumor classifications shared many characteristics. Both datasets had a stroma-poor (SP) tumor type, that in the human exhibited poorer overall survival than the stroma-rich (SR) tumors. In mice, the SP tumors took longer to develop (average 90 days) than the SR tumors (average 45 days). However, the SP tumors grew faster, taking 30 days to reach the 2 cm endpoint versus 45 days for SR. In humans, patients with SP tumors had shorter OS, but not RFS, which suggests no difference in the time for tumors to recur but faster growing tumors, similar to the mouse model.

Most of the markers that overlapped across CyCIF and MIBI platforms were specific to the stroma as opposed to the tumor. Therefore, we were able to establish similarities in mouse versus human TNBC stromal marker expression. The stroma-rich, immune-rich (SR_IR) tumors had significantly higher pan-immune CD45, regulatory T cell FoxP3 and dendritic cell CD11c expression in both human and mice subtypes. The SP tumors had significantly higher endothelial CD31 expression in the stroma, and the stroma-rich immune-poor (SR_IP) subtype had significantly lower Ki67 expression in the stroma in both human and mice subtypes. The lower proliferation in the SR_IP stroma may reflect a quiescent or resting fibroblast phenotype⁸⁰. The recapitulation of stromal phenotypes is important for modeling of breast cancer disease progression, as stromal gene expression predicts clinical outcome in breast cancer with higher accuracy than an FDA-approved 70 gene predictor²⁶⁰. In fact, the poor outcome individuals with a negative stromal gene expression signature showed increased CD31+ endothelial cells in the stroma, consistent with our mouse model and MIBI data analysis²⁶⁰.

With fewer epithelial-specific markers overlapping between the CyCIF and MIBI panels, image analysis was less conclusive regarding the recapitulation of human epithelial phenotypes in the mouse model. Based on gene expression and image analysis, the SR mice had a mesenchymal phenotype. Our imaging data showed loss of E-cadherin in a subset of SR_IR tumors (Figure 4.4 k), and decreased E-cadherin expression compared to SP tumors (Figure 4.8 c) consistent with E-cadherin loss during EMT²⁶¹. Human SP_IR tumors showed higher expression of β -catenin, a binding partner of E-cad, which accumulates in the cytoplasm and nuclei of tumor cells when the cadherin-catenin complex is disrupted, decreasing cell adhesion and increasing cancer invasiveness²⁶¹. Human and mouse tumors showed higher expression of mesenchymal markers Vimentin and α SMA in epithelial cells (Figure 4.8 c, d), suggesting EMT processes in both datasets, although examination of the images revealed that in human data, elevated mesenchymal marker expression was due to lateral signal spillover from fibroblasts in close proximity to tumor cells. α SMA+ myofibroblasts have been shown to drive EMT through paracrine secretion of TGF- β ²⁶², and in fact, mouse tumors showed increased stromal α SMA expression (Figure 4.8 a) and increased TGF- β signaling in the SR subtype [in preparation, Doha 2022]. Thus, our mouse model

may recapitulate features of tumor-stroma crosstalk that impact tumor progression. Finally, S100A6 was elevated in SR mouse tumors, and since S100A6 promotes EMT through β -Catenin in pancreatic cancer cell lines²⁵⁹, a similar mechanism could underlie the elevated β -catenin in human SR tumors. Analysis of additional human cohorts with E-cadherin, S100A6, α SMA and TGF- β could better link the MYC:PTENfl mouse model to phenotypes in human tissue.

4.6 CONCLUSION

The work presented in this chapter encompasses both the generation of new datasets as well as the introduction and sharing of analytical tools. Importantly, all the analytics can be reproduced in the Jupyter notebooks published along with the works presented (Table 4.6). The repositories below enable other researchers to reproduce our work as well as deeply understand the complex datasets behind the paper’s figures.

Table 4.6. Code repositories

Paper	Repository
Janiszewska, et al	https://github.com/mjaniszewska-lab/HER2_heterogeneity_2021/tree/main/cycIF
Johnson, et al	https://github.com/engjen/SMMART_BC
Doha, et al	https://github.com/engjen/MYC-PTENfl-mouse

Numerous image and data analysis challenges were addressed in each of these projects. In the first project, large datasets were successfully processed with our mplexable pipeline, demonstrating the scalable nature of our software. Additionally, multiple staining runs had to be combined into a master dataset and any batch effect accounted for. We addressed this challenge by processing control TMAs which were used as a reference for standardizing manual gating across batches.

We extended the concept of using control TMAs in the second project, where we utilized breast cancer cell line TMA controls to standardize mean and variance between staining batches using the ComBat algorithm. This method was based on the experiments done in Chapter 3, a framework for multiplex imaging optimization and reproducible analysis.

In the third project, we demonstrate the utility of our mplexable pipeline in analysis of a publicly available dataset generated with a different multiplex imaging modality, the MIBI platform. This platform uses secondary ion mass spectrometry to image antibodies tagged with isotopically pure elemental metal reporters in FFPE tissue sections¹¹¹. We used mplexable to de-noise, segment and extract features from the raw images, which we then analyzed with the same downstream analytical tools as CyCIF data. The ability to process and analyze data from multiple platforms greatly increases the utility of our tools.

In conclusion, our protocols, computational tools and analytics contributed to multiple collaborative projects aimed at characterizing breast cancer heterogeneity. Firstly, we examined factors contributing to T-DM1 response and resistance in HER2+ breast cancer patients. Secondly, we charted heterogeneity across serial biopsies from a patient with metastatic ER+ breast cancer and identified stromal cells proximal to tumor that significantly correlated with biomarker

expression. Thirdly, we characterized subtypes of tumors generated by MYC activation and PTEN deletion in a mouse model of TNBC and correlated their marker expression and survival outcomes to human TNBC subtypes. In these three collaborative projects, we demonstrate the power of our methods and analytics to draw biologically meaningful conclusions from multiplex imaging data. Specifically, our work contributes to efforts to predict patient response to therapy, characterize the complexity of tumor evolution during treatment, and model human disease heterogeneity using GEMMs that recapitulate human disease subtypes and outcomes. Our findings linked between tumor heterogeneity and treatment response, cell proximity and biomarker expression, and animal models and human disease phenotype and outcome. Together, these works exemplify the types of analyses enabled by our wet-lab and computational advances.

Chapter 5. MULTIPLATFORM ANALYSIS OF BREAST CANCER CLINICAL COHORTS USING MPLEXABLE

Manuscript in Preparation.

- Eng J, Bucher E, Hu Z, Gibbs S, Chin K. Multiplatform analysis of breast cancer clinical cohorts for robust biomarker discovery.

5.1 ABSTRACT

Various multiplex imaging platforms have been developed for high dimensional in-situ proteomic analysis of tissues. These imaging technologies show promise in describing biologically and clinically relevant single-cell spatial heterogeneity in tumors and other tissues. Until recently²⁶³, multiplex imaging analytical pipelines have been platform-specific, hindering cross-platform analyses that are useful for validation of methods and their findings. Herein, we show that our image processing pipeline, mplexable, can be utilized for processing and comparative analysis of data from disparate imaging technologies including cyclic immunofluorescence, imaging mass cytometry, and multiplex ion-beam imaging. We demonstrate similar single-cell phenotyping results across breast cancer patient cohorts imaged with these methods, and furthermore, show agreement in the prognostic value of our quantitative enumeration of cellular abundances and proximities across platforms. Our methods enable assembly of larger clinical cohorts from diverse platforms to aid in prognostic biomarker identification and validation.

5.2 INTRODUCTION

Breast cancer is the most common cancer worldwide¹. In the United States, it accounts for 30% of female cancers and 15% of cancer deaths, and is the leading cause of cancer deaths in women aged 20 – 59 years². Although breast cancer death rates are declining, the decline has slowed in the last decade, emphasizing a need for basic and clinical research to advance treatment options². Clinically, breast cancer is divided into three subtypes by hormone receptor (estrogen receptor [ER] and progesterone receptor [PR]) and human-epidermal growth factor receptor (HER2) status: luminal ER-and/or PR-positive, HER2-positive and triple-negative breast cancer (TNBC)³. These categories are clinically useful for determining the standard-of-care treatment. In addition to surgery and radiation, ER+ patients receive hormonal therapy, HER2+ patients receive anti-HER2 targeted agents, and TNBC patients are treated with chemotherapy⁴. However, recent advances in the treatment landscape of breast cancer, including the approval of immune checkpoint inhibitors for early TNBC⁷, have motivated the characterization of the breast tumor microenvironment for deeper understanding of immune-tumor interactions. Identification of biomarkers predicting breast cancer immunotherapy response is still an urgent clinical need³². Furthermore, the proliferation of phase I/II trials evaluating the efficacy of novel immune targeted therapies in breast cancer should be accompanied by biomarker development for maximum therapeutic efficacy²⁵.

Multiplex imaging methods enable quantification of dozens of biomarkers in a single tissue section with at sub-cellular resolution while retaining spatial context^{105–110}. Tissue structures such as tertiary lymphoid structures, identified with multiplex imaging, are predictive biomarkers of immunotherapy response in melanoma^{264,265}. Employment of these methods in breast cancer has revealed that single-cell spatial context has prognostic relevance and shows correlations with transcriptomic and genomic features of tumors^{111–113}. However, different imaging platforms can introduce different technical artifacts, therefore, identification of biomarkers in additional platforms can better validate biological and clinical findings. Furthermore, due to the number of steps in the processing pipeline and the number of variables and parameters involved, overfitting is an issue in biomedical imaging data, which can be addressed through use of a training cohort in which to fix analytical methods subsequently applied to a validation cohort²⁶⁶. Incorporation of

data from additional imaging platforms into biomarker studies enables reuse of valuable datasets as validation cohorts. Importantly, to facilitate reuse of data and reproducibility of findings, documentation of metadata, analysis protocols and code is essential, preferably using open-source software tools²⁶⁶. We developed an open-source python software, mplexable, for multiplex image processing and analysis, which we use herein to process and analyze clinical breast cancer cohorts imaged with three multiplex imaging platforms: cyclic immunofluorescence (CyCIF), imaging mass cytometry (IMC) and multiplex ion-beam imaging (MIBI)^{111,112,246}. In this proof-of-concept study, we identify prognostic single-cell spatial biomarkers common across the imaging platforms. As such imaging datasets become more widely available, tools such as ours can facilitate biomarker discovery with high accuracy, reliability and efficiency.

5.3 METHODS

5.3.1 Patient samples

Breast cancer tissue microarrays were graciously provided by Dr. Jennifer Pietenpol (JP). All samples were surgical (mastectomy or resections) collected at Vanderbilt with the same fixation protocol. JP-TMA1 had 131 cores of approximately 1.2 μm diameter, with duplicate cores from 19 TNBC, 8 HER2+ and 36 ER+ patients. It also included one control core each of inflamed appendix, colon cancer, muscle, pancreas and normal breast. Four of the TMA1 TNBC patients received neoadjuvant chemotherapy. Clinical outcome and clinicopathological information were available for TMA1. JP-TMA2 contained a single, slightly larger ($\sim 1.4\mu\text{m}$ diameter) cores from 39 triple-negative tumors and 1 ER+/HER2+ control core. Thirteen of the patients in TMA2 had neoadjuvant therapy. Clinical outcome data were not available for TMA2.

5.3.2 Imaging data generation and sources

CyCIF staining of tumor tissue was completed on JP-TMA1 and TMA2. The whole tissue core was imaged, as described in sections 3.4.1 to 3.4.6. MIBI imaging data was previously published by Keren et al.¹¹¹ and the images were downloaded from <https://mibi-share.ionpath.com/tracker/imageset> under the name “Keren et al., Triple Negative Breast Cancer.” Survival and recurrence data were obtained from a second publication by the same group²⁵⁶, and were downloaded from <https://github.com/aalokpatwa/rasp-mibi>. IMC imaging data were previously published by Jackson et al.¹¹² and images and clinical data were downloaded from <https://doi.org/10.5281/zenodo.3518284>.

5.3.3 Image Processing

CyCIF tiff images were registered, segmented and single-cell intensity as well as nuclear size and shape features were extracted as described in section 3.4.7. The CyCIF pipeline is available at <https://gitlab.com/engje/mplexable/-/tree/master/jupyter>. Nuclear and cell segmentation were conducted using the Cellpose algorithm, which showed visually superior performance on CyCIF data compared to a watershed algorithm (Figure 6.1a). Nuclear and cell segmentation masks were matched using mplexable, enabling subtraction of nuclear mask from cell mask to obtain segmentation of the cytoplasm (Figure 6.1b).

MIBI and IMC processing pipelines are available as Jupyter notebooks at https://github.com/engjen/cycIF_TMAs. MIBI and IMC images were downloaded as multipage OMETiffs. Hot pixels were detected by identifying pixels that were 10 standard deviations above a median filtered image with a 2x2 pixel kernel size. Hot pixels were set to the median filter values and resulting images were saved as tiff for downstream feature extraction. DNA channels were processed for cellular segmentation as follows. DNA images were rescaled between the 3rd and 1.5 x the 99.999 quantile. The gamma value was adjusted by 0.6 in MIBI data and 0.4 in IMC data to enhance dimly staining nuclei. A two-channel nuclear plus cytoplasm image was generated for cell segmentation. A mixture of markers was used to generate a composite image which stained the cytoplasm of the majority of cells. For MIBI data, the Beta catenin, Vimentin, CD45 and CD31 channels were combined into a maximum intensity projection cytoplasm image and the gamma value was adjusted by 0.6. For IMC data, E-cadherin, Vimentin, CD44 and CD45 were combined into a maximum intensity projection cytoplasm image and gamma adjusted by 0.4. Chamboelle total-variation denoising, implemented in *scipy*²³⁸, was used to smooth out pixelated nuclear and cytoplasmic projection images (weight=0.1, except weight=0.05 for IMC cytoplasm). All parameters were selected by testing segmentation results at <https://www.deepcell.org/predict> and [cellpose.org](https://www.cellpose.org) (Figure 6.1c-d). Skipping either nuclear or cytoplasmic Chamboelle total-variation de-noising resulted in failure of deep learning-based algorithms on the IMC data (Figure 6.1e). Mesmer segmentation²⁶⁷ performed better than Cellpose²³⁴ in IMC data due to improved detection of dim nuclei, likely due to the incorporation of cytoplasmic staining in the nuclear segmentation model (Figure 6.1d). This issue with Cellpose was not detected in CyCIF images, which had brighter DNA staining. For IMC and MIBI data, nuclear and cellular segmentation were performed on processed nuclear and nuclear + cytoplasm images using Mesmer²⁶⁷. Matching of cell IDs in the nuclear and cell masks was done with *mplexable*, with cell masks relabeled to match the ID of the nucleus which they had most overlap with. Cytoplasm masks were calculated by subtracting the nuclear mask from the matching cell mask. Single-cell mean intensity, nuclear size and shape features, and nuclear centroid coordinates were extracted with *scikit-image*²⁶⁸.

Mesmer segmentation of IMC data was compared to the watershed segmentation originally published by Jackson et al¹¹². The cell counts across the two methods had a Pearson correlation of 0.98 (Figure 6.2a). Visual examination of outliers revealed that Mesmer segmentation performed better in tissues with necrosis and high background noise in the DNA channel (Figure 6.2b-d).

5.3.4 Image QC

IMC data were collected from small ROIs (~600 μ m) within TMA cores and some samples annotated as ER+ did not show any ER+ staining in the ROI. Therefore, quality control was performed on ER images, a marker which has been noted to exhibit non-specific background staining on the IMC platform²⁶⁹. QC images of ER staining were generated and sorted in a blinded fashion into negative and positive for nuclear-specific-staining (Figure 6.3a). Only ROIs from ER+ patients that had ER positive QC or ROIs that came from ER negative patients and had ER negative QC were used for analysis. Our QC metric did not significantly enrich for tumors with different grade, PR status, TMA block, age of specimen, age of patient or tumor size (Figure 6.3c-d). There were no significant survival differences between QC positive and negative tumors from ER+, TNBC or ER+HER2+ patients (Figure 6.3e).

In the IMC dataset, additional QC steps included: necrotic regions were manually circled using the napari image viewer and excluded and bright aggregates in the CD3 channel were excluded with an upper threshold.

In the CyCIF data, imaging artifacts included autofluorescence (AF), non-specific background, floating tissue and tissue loss. Following our recommendations for AF subtraction in Chapter 3, background AF images were obtained half-way through CyCIF data collection and these images were scaled by exposure time and subtracted from the AF488, AF55 and AF647 channels. Feature extraction was performed on AF subtracted images. Areas of floating tissue, air bubbles or necrotic regions were manually circled using the napari image viewer and excluded. Background subtraction was performed by setting manual thresholds for selected markers and subtracting those values from extracted data. The PD1 antibody had bright aggregates that were excluded with an upper threshold. Tissue loss was detected by cells that lacked DAPI staining in the last round of imaging, and these cells were excluded.

In all three platforms, additional artifacts caused by floating tissue or imaging problems (e.g., dark or bright bands across IMC and MIBI images perhaps caused by problems with the rastering process) were detected through unsupervised clustering and clusters comprised of artifacts were removed, as these clusters showed atypical very bright or dim staining in many channels and formed distinct artifact clusters.

5.3.5 *Single cell phenotyping*

Cell types were defined in two ways: manual gating and unsupervised clustering. Unsupervised clustering was conducted using the scanpy²³¹ software. Single-cell mean intensity values were selected from either the nucleus or cytoplasm masks for each marker, depending on expected intracellular distribution. Beta-catenin, which can localize on the membrane, in the cytoplasm or in the nucleus, was selected from both nuclear and cytoplasmic masks for separate analysis. Since the CyCIF and IMC platforms had more marker and breast cancer subtype overlap than the MIBI panel (Figure 5.1, Figure 5.5), 20 matching markers were selected for clustering in these datasets, plus selected markers for immune, epithelial and fibroblast subsets (ColI, CD4, CD8 in CyCIF, and FN, panCK in IMC). For MIBI data, all available markers were used for clustering. Additionally, the nuclear area feature was used for clustering. Each marker was scaled by its standard deviation, without zero-centering, and clipped above 20 standard deviations. A UMAP embedding was generated using 30 k-nearest neighbors and the embedding was clustered using the Leiden community detection algorithm²⁴⁵. The Leiden resolution parameter was selected that resulted in 20 - 25 clusters. Each cluster was annotated and categorized as epithelial, endothelial, fibroblast, immune or stromal. Some clusters were comprised of multiple cell types, and these were manually split, for example, the CD44+ cluster was split into CD44+ tumor and CD44+ stroma based on manual gating results (described below).

We then performed manual gating to verify our annotated-cluster cell type calling. A threshold was set for each gating marker based on positive pixel patterns in images. Fibroblasts were positive for one or more of Vimentin, FN or ColI. Epithelial cells were positive for one or more of Ecad, Cytokeratins, or β -catenin. Endothelial cells were defined as CD31+. Immune cells were CD45+.

Stromal cells were all non-fibroblast, non-endothelial, non-epithelial, non-immune segmented nuclei. Jupyter notebooks recording single cell phenotyping pipelines are available here: https://github.com/engjen/cycIF_TMAs

5.3.6 Statistical Analysis

The CyCIF dataset was used as a training dataset to determine the quantile separating high and low abundance of each cell type that was most predictive of survival. For analysis of single variable association with recurrence-free (RFS) or overall survival (OS), each cell type was tested (log-rank test) for significant association with RFS and OS in CyCIF TNBC, ER+ and HER2+ samples. Mean fraction of cells from all ROIs/cores from each patient was calculated and binarized into high and low abundance based on tertile or median values within each subtype. The most prognostic cutoff value was selected for each cell type and for cell types having prognostic value ($\alpha < 0.05$) these cutoffs were applied in the two additional datasets. Since overall cell type fractions differed between platforms and subtypes (Figure 5.5), high and low values were determined relative to other samples from the same platform and subtype. High/low binarized abundances were combined for survival analysis on the whole cohort ($n=98$ TNBC and $n=163$ ER+). Kaplan-Meier analysis was used to estimate survival functions and the log-rank test was used to assess significant differences in survival. Multivariable Cox proportional hazards (CPH) modelling was used to combine cell abundance variables with patient age, tumor size and clinical stage. Fifty-five TNBC and 160 ER+ patients had these additional clinical parameters (additional clinical parameters were not included in the MIBI dataset).

Epithelial and stromal subtypes were determined by hierarchical clustering of patients based on the fraction of epithelial or stromal cell types within each compartment, respectively. Cell types representing greater than 2% of the total cell population in the respective tissue compartment were used for clustering. Unsupervised hierarchical clustering of patients was performed using the AgglomerativeClustering function in scikit-learn²³⁷, using a Euclidean distance metric and complete linkage. The number of clusters/subtypes were selected as follows. Based on gene expression studies of breast cancer subtypes, we expected four triple-negative subtypes, two luminal clusters, a HER2 enriched and a normal-like cluster, therefore we selected eight as the number of epithelial clusters^{5,129}. For stromal clusters, we selected the minimum number of clusters need to separate T cells from other clusters ($k=6$).

For analysis of tumor-stroma correlation, the fraction of cell type within the epithelial or non-epithelial stromal compartment (including fibroblast, immune, endothelial and other stroma) were calculated. Pearson correlation was used to assess significant correlations between cell frequencies within their respective tissue compartments.

For analysis of cellular neighbors, each cell's neighbors within a 25 μm radius (~ 1 epithelial cell diameter) were counted. Either the total number of neighbors or the number of neighbors normalized by the number of stromal neighbors for each epithelial cell were used for analysis. Survival analysis was done as described above, using the CyCIF dataset to determine significant prognostic cell-cell interactions and cutoff values. These were then verified in the two additional datasets. Spatial and survival analysis code is available here: https://github.com/engjen/cycIF_TMAs.

5.4 RESULTS

5.4.1 *Imaging and Single-Cell Phenotyping*

We generated CyCIF data from two tissue microarrays (TMAs) containing surgical breast cancer specimens and analyzed publicly available breast cancer data from IMC and MIBI platforms generated with similar antigen targets (Figure 5.1). The CyCIF data was comprised of 47 biomarkers imaged at a resolution of 0.325 μm per pixel in entire cores with 1.2 to 1.4 mm diameters, with 1-2 cores imaged per patient (Figure 5.1a-b, left). The IMC data included 35 biomarkers imaged in the largest square area contained within the TMA cores (from 0.6 - 0.8 mm diameter), at a resolution of 1 μm per pixel (Figure 5.1a-b, center)¹¹². The MIBI data included 36 biomarkers imaged in 0.8 x 0.8 mm square ROIs at a resolution of 0.5 μm per pixel (Figure 5.1a-b, right)¹¹¹.

We used the same methods to generate single-cell phenotypes via unsupervised clustering in each platform. In the CyCIF dataset, we generated a 30 k-nearest neighbors (k-NN) graph using a subset of 23 markers common between the CycIF and IMC datasets, plus nuclear area. Markers were scaled by the standard deviation (SD), clipped at 20 SDs, and nearest neighbors were calculated using scanpy²³¹. We visualized the k-NN using UMAP embedding, confirming good separation of lineage specific markers, CD31, endothelial, E-cadherin (Ecad), epithelial, Collagen I (ColI), extracellular matrix, Vimentin, mesenchymal cells including activated fibroblasts, and CD45, immune infiltrate (Figure 5.2a). The UMAP visualization showed good mixing of cells from different TMA sources and separation of tumor cells from different breast cancer subtypes (Figure 5.2b). We used the Leiden algorithm (implemented in scanpy²³¹) to cluster the k-NN graph, resulting in 23 cell type clusters (Figure 5.2c). The mean expression of each biomarker in each cluster was used to annotate cell types (e.g., CD8 T cell, CD4 T cell, luminal ER+ tumor) and lineages (i.e., endothelial, epithelial, fibroblast, immune and other stromal, Figure 5.2d). The most common cell types included luminal and luminal ER+ tumor, CD4 T cells, Vimentin+ fibroblasts and quiescent stroma (Figure 5.2d). To confirm our cell typing, we performed manual gating on lineage specific markers (Figure 5.2a). Gating and clustering-based cell lineages localized to similar areas of the UMAP and had 73% accuracy on a single cell level, as calculated using `metrics.accuracy_score` in `scikit-learn`²³⁷ (Figure 5.2e, f).

In the IMC dataset, we used the same method as in the CyCIF dataset to visualize marker expression, TMA, and subtype and cluster cell types (Figure 5.3a-c). We used 21 markers plus nuclear area for clustering, resulting in 24 cell types (Figure 5.3d). Upon annotation, we found the most common cell types were luminal, luminal ER+, and ER+ HER2+ tumor, Vimentin or fibronectin (FN)+ fibroblasts, quiescent stroma, and T cells (d). Clustering and gating-based cell lineages localized to similar areas of the UMAP and had 77% accuracy (Figure 5.3e-f).

The MIBI panel included more immune-specific markers than the other panels and had 15 markers similar to our CyCIF panel. In order to generate a dataset in which to audit deeper immune contexture, we clustered on all 33 markers plus nuclear area and eccentricity. Again, we saw good separation of lineage specific markers and Leiden clustering generated 22 cell types of which Luminal tumor, fibroblasts, T cells and quiescent stroma were the most common (Figure 5.4a-c).

Gating-based cell types showed good overlap with clustering cell types with 72% accuracy (Figure 5.4d-e).

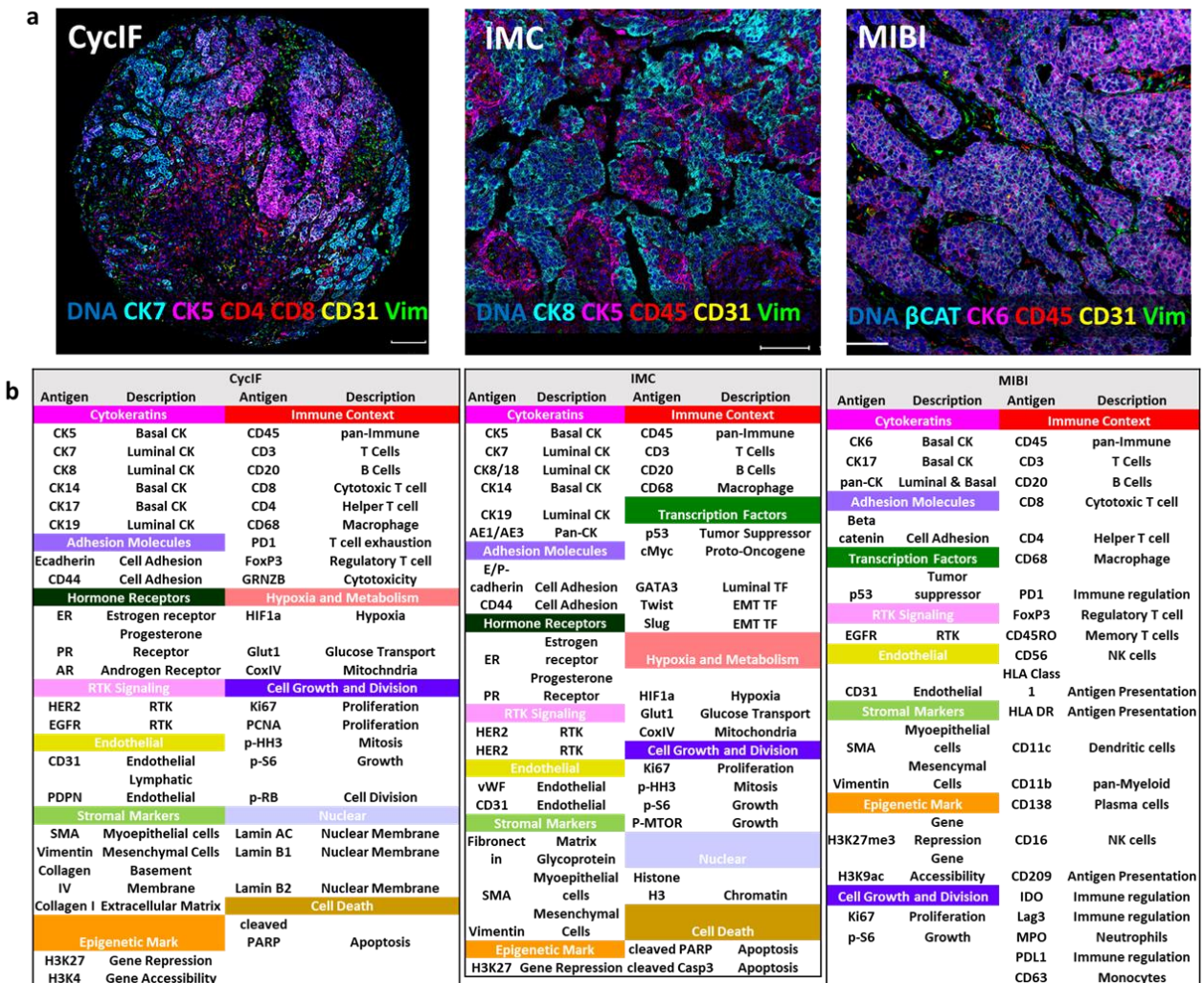


Figure 5.1. Breast Cancer Multiplex Imaging Datasets.

a. We generated CyCIF data from two breast cancer TMAs and obtained publicly available IMC and MIBI data from three total additional breast cancer TMAs. b. The antibody panels from the three datasets included markers for cytokeratins, adhesion molecules, hormone receptors, receptor tyrosine kinase (RTK) signaling, cell growth and division, endothelial, immune and stromal cells. CyCIF scale = 130 um, IMC scale = 100 um, MIBI scale = 96 um.

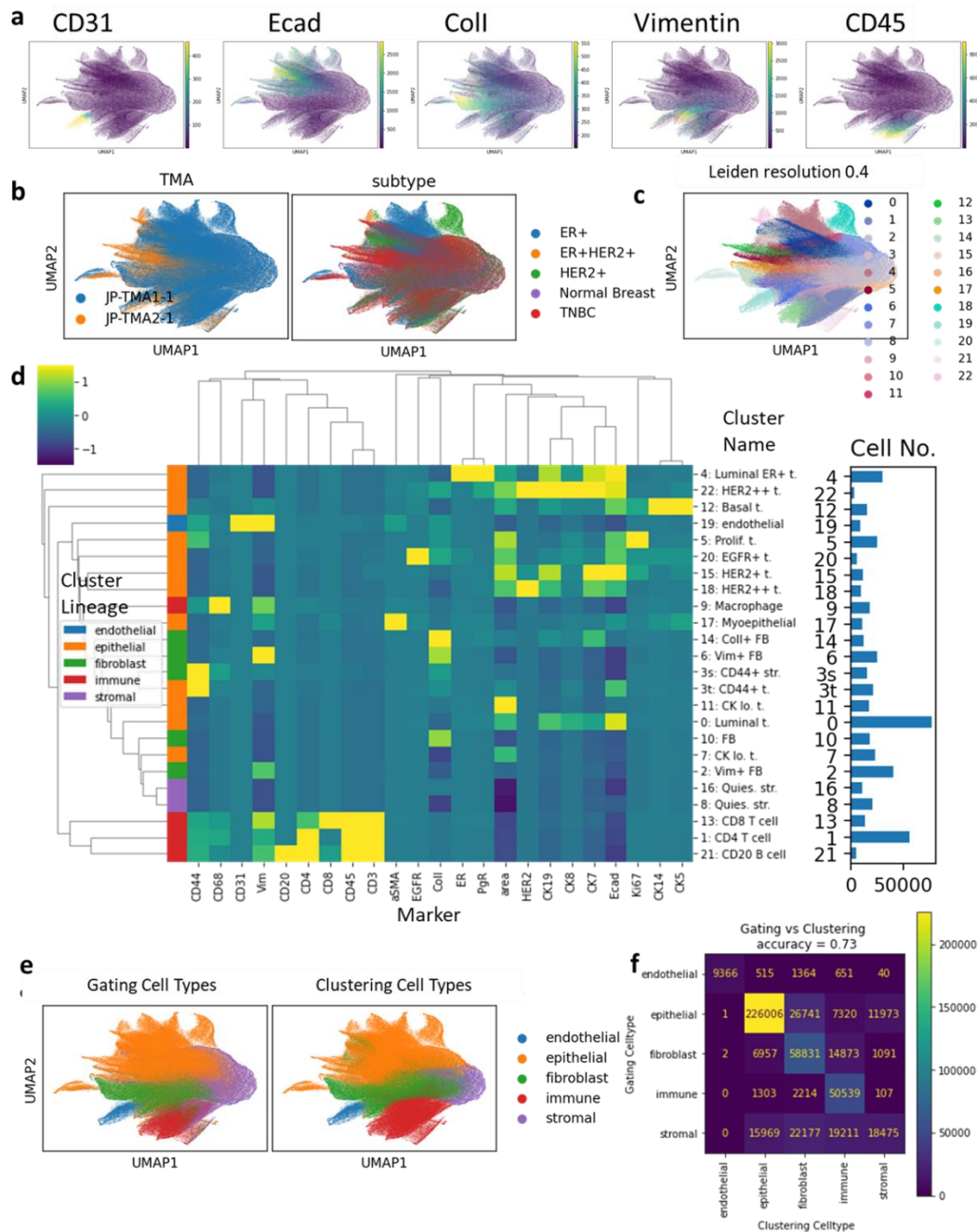


Figure 5.2. Unsupervised clustering defines cell types in CyCIF data.

a. Single-cell segmentation and feature extraction were completed using mplexable. A UMAP embedding was generated based on scaled single-cell mean intensity values (30 k-nearest neighbors). The UMAP is colored by cell lineage markers CD31, endothelial, E-cadherin (Ecad) epithelial, collagen I (ColI) and vimentin, fibroblast, and CD45, immune. b. Two TMAs (left) and five clinical subtypes (right) were clustered together for cell typing. c. Unsupervised clustering with the Leiden algorithm (resolution 0.4) resulted in 23 cell types. d. Heatmap of mean fluorescence intensity of each marker in each Leiden cluster from (c). Twenty-two markers and one morphology feature (nuclear area) were used for clustering. Clusters were annotated by lineage: endothelial, epithelial, fibroblast, immune or stromal (left color bar on heatmap) and named based on marker expression (right labels on heatmap). The number of cells in each cluster are plotted to the right (Cell No. bar plot), showing the abundance of each cell type in the whole dataset. e. Manual gating of the markers in (a) were used to determine cell lineages separately from the Leiden annotated cell lineages. f. Confusion matrix of gating-based versus clustering-based cell lineages shows 73% accuracy.

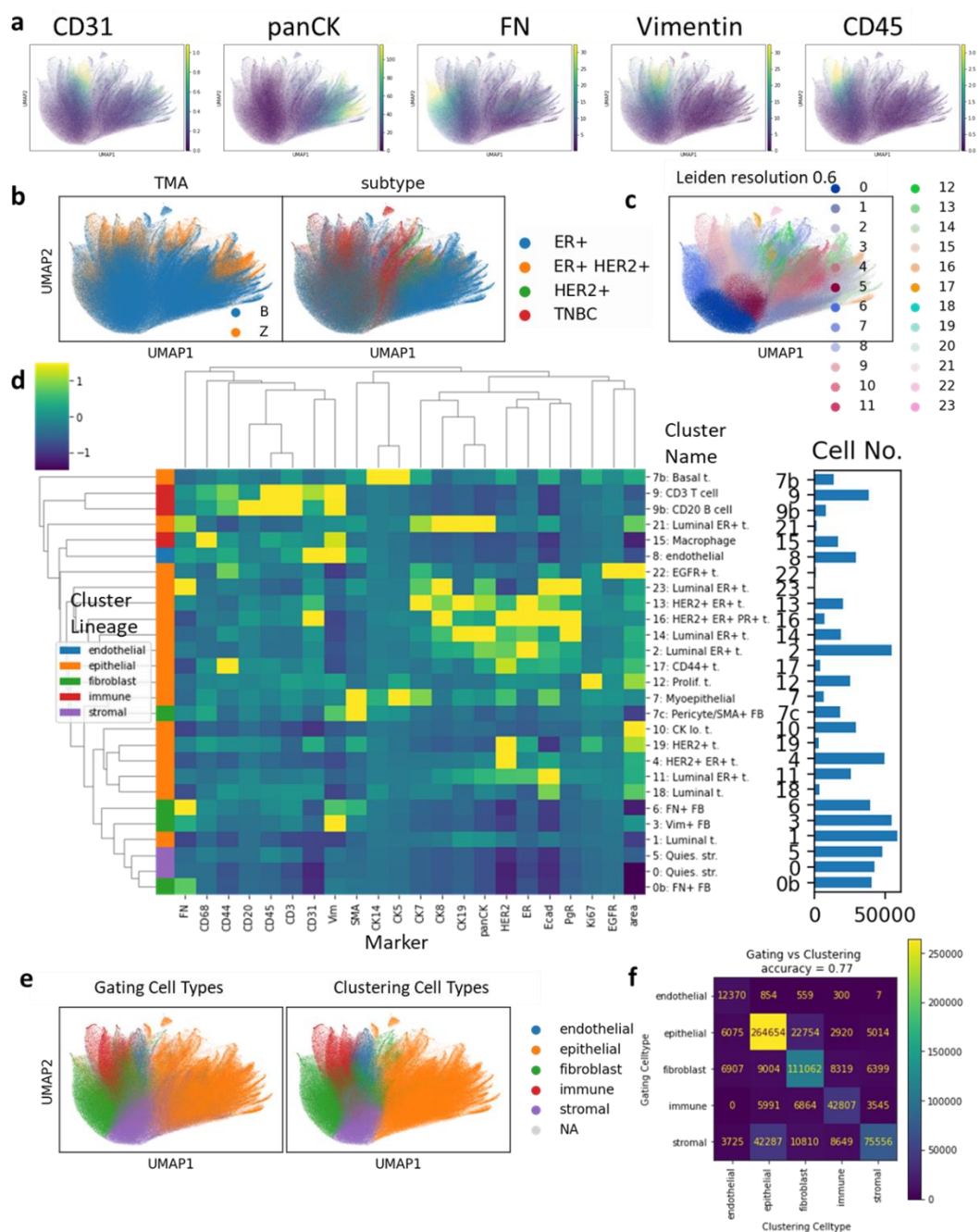


Figure 5.3. Unsupervised clustering defines cell types in IMC data.

a. Single-cell segmentation and feature extraction were completed using mplexable. A UMAP embedding was generated based on scaled single-cell mean intensity values (30 k-nearest neighbors). The UMAP is colored by cell lineage markers CD31, endothelial, pan-cytokeratin (panCK) epithelial, fibronectin (FN) and vimentin, fibroblast, and CD45, immune. b. Two TMAs (left) and four subtypes (right) were clustered together for cell typing. c. Unsupervised clustering with the Leiden algorithm (resolution 0.6) resulted in 25 cell types. d. Heatmap of mean fluorescence intensity of each marker in each Leiden cluster from (c). Twenty-one markers and one morphology feature (nuclear area) were used for clustering. Clusters were annotated by lineage: endothelial, epithelial, fibroblast, immune or stromal (left color bar on heatmap) and named based on marker expression (right labels on heatmap). The number of cells in each cluster are plotted to the right (Cell Count bar plot), showing the abundance of each cell type in the whole dataset. e. Manual gating of the markers in (a) were used to determine cell lineages separately from the Leiden annotated cell lineages. f. Confusion matrix of gating-based versus clustering-based cell lineages shows 76% accuracy.

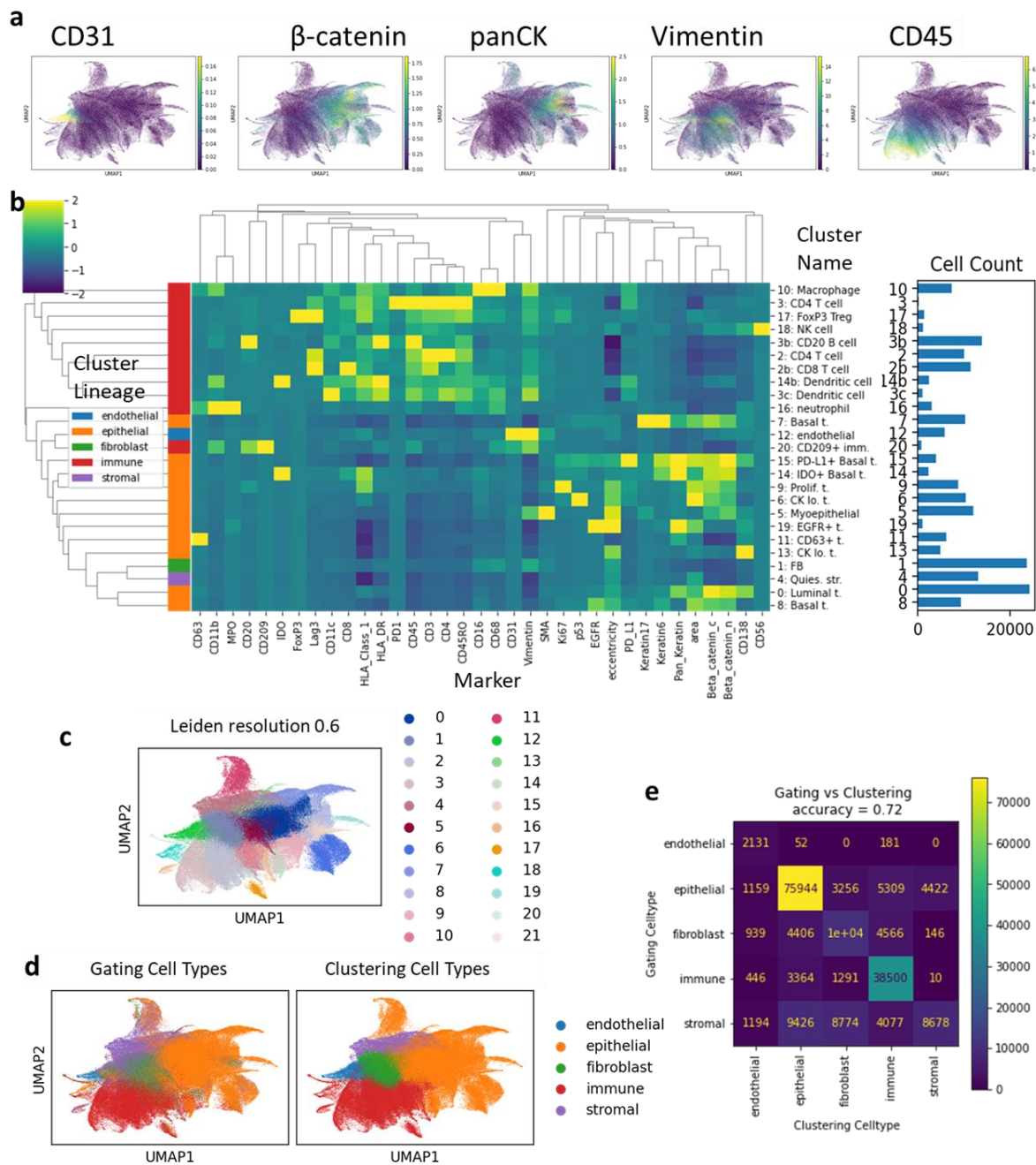


Figure 5.4. Unsupervised clustering defines cell types in MIBI data.

a. Single-cell segmentation and feature extraction were completed using with mplexable. A UMAP embedding was generated based on scaled single-cell mean intensity values (30 k-nearest neighbors). The UMAP is colored by cell lineage markers CD31, endothelial, β -catenin and pan-cytokeratin (panCK) epithelial, vimentin, fibroblast, and CD45, immune. b. Heatmap of mean fluorescence intensity of each marker in each Leiden cluster from (c). Thirty-three markers and two morphology features (nuclear area and nuclear eccentricity) were used for clustering. Clusters were annotated by lineage: endothelial, epithelial, fibroblast, immune or stromal (left color bar on heatmap) and named based on marker expression (right labels on heatmap). The number of cells in each cluster are plotted to the right (Cell Count bar plot), showing the abundance of each cell type in the whole dataset. c. Unsupervised clustering with the Leiden algorithm (resolution 0.6) resulted in 22 cell types. d. Manual gating of the markers in (a) were used to determine cell lineages separately from the Leiden annotated cell lineages. e. Confusion matrix of gating-based versus clustering-based cell lineages shows 72% accuracy.

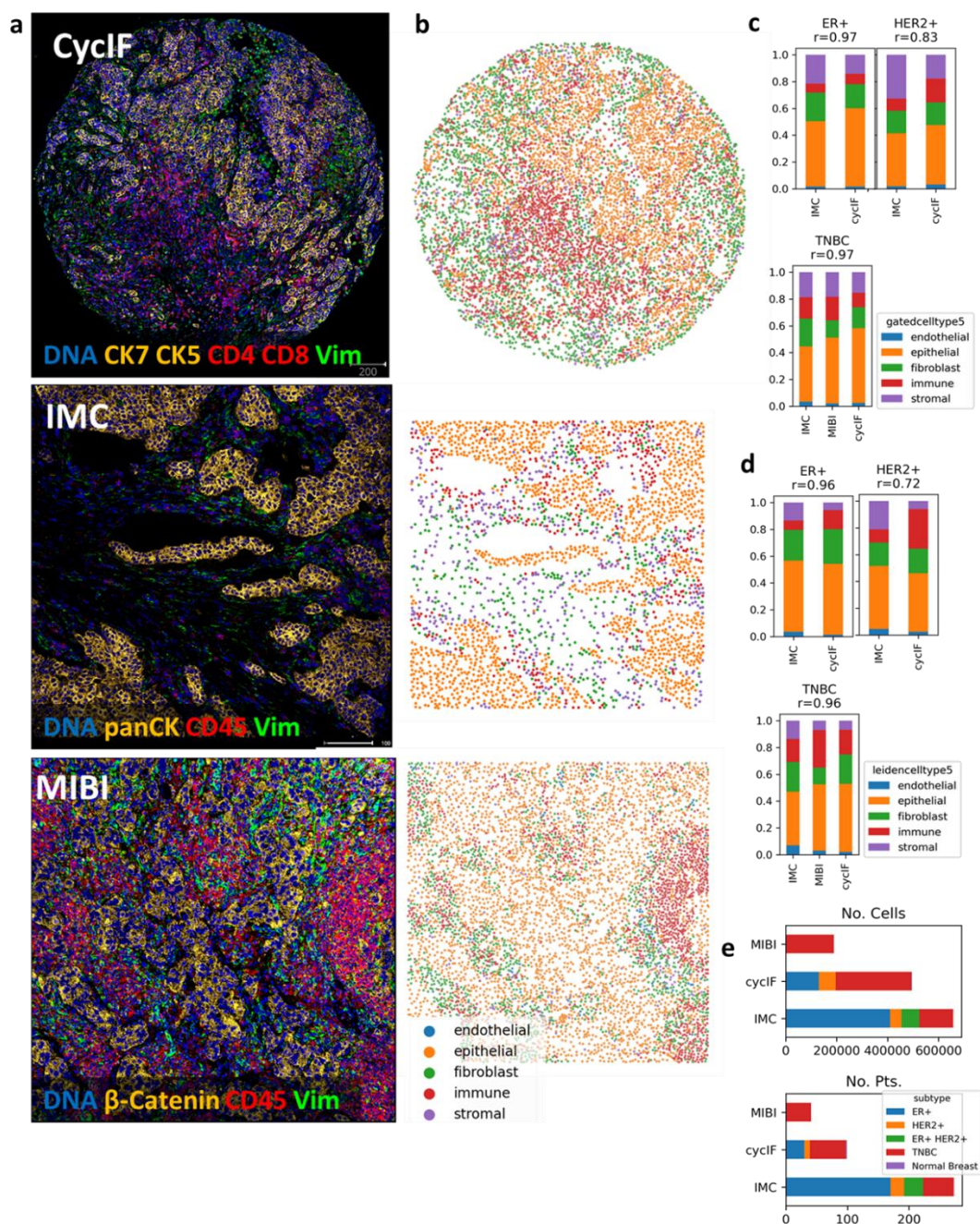


Figure 5.5. Cross-platform Cell Typing.

a. Representative images from the three multiplex imaging platforms showing epithelial (orange), immune (red) and fibroblast (green) markers. b. Gated cell types showing cell location and lineages: epithelial (orange), immune (red), fibroblast (green), endothelial (blue) and stromal (purple). c. Gated cell lineage totals per subtype (ER+, HER2+ and TNBC), per platform (IMC, CyCIF and MIBI). d. Clustered cell type lineage per subtype, per platform. c-d. Pearson's correlation between platforms shown in figure title ($r=0.xx$) e. Number of cells analyzed (top) and number of patients (bottom), colored by subtype, per platform.

5.4.2 *Cell type fractions are similar across platforms*

As our most common cell types were similar across platforms, we were encouraged that we could successfully combine the data for cross-platform analysis. We visually validated individual clusters, as well as cell lineages by inspecting the images. In all three platforms, cell lineage identity and spatial distribution matched the underlying imaging data (Figure 5.5a, b).

For further validation, we sought to quantitatively compare the fractions of the five main cell lineages. We calculated the fraction of cells in each lineage for each platform and breast cancer subtype. Both gating and clustering cell types showed high correlation (Pearson R= 0.97 gating and 0.96 clustering) across platforms for ER+ (n=30 CyCIF, 170 IMC) and TNBC (n=59 CyCIF, 50 IMC, 41 MIBI) while HER2+, which had a smaller number of samples (n=8 CyCIF, 22 IMC) had more variability between platforms (Figure 5.5c, d, e). We did note some platform specific bias; for example, IMC showed a smaller fraction of immune cells defined by clustering in all three subtypes (Figure 5.5c, d). Therefore, when setting high/low cutoffs for cell abundances, we calculated high and low relative to each platform and subtype, as opposed to the whole dataset. Since different antibody clones, probes and imaging systems were used, resulting in different signal-to-background ratios between platforms, even for the same target, we believe this is a necessary step to account for technical variability.

5.4.3 *Single cell phenotyping identifies prognostic tumor subtypes*

For subtyping based on single-cell phenotypes, we calculated the fraction of epithelial or non-epithelial stromal cells in the total epithelial or stromal cells in that sample, respectively. Fractions of epithelial cell types representing greater than 4% of the epithelial compartment were used to hierarchically cluster patients from all subtypes and platforms (Figure 5.6a). Given the five intrinsic breast cancer subtypes (Luminal A and B, Erb-B2+, Basal-like and normal-like)^{5,270} and the subsequent classification of four subtypes within TNBC/Basal-like tumors^{129,251}, we set k=8 to capture the previously characterized biology of breast cancer epithelial cells. The resulting eight subtypes included tissues enriched for luminal, cytokeratin low, luminal ER+, myoepithelial, basal, HER2+ ER+, and proliferating tumor cells, as well as a heterogeneous group not dominated by one phenotype (Figure 5.6a). Examination of images from tissues representing different subtypes showed the expected intensity differences in marker expression (Figure 5.6b). In the ER+ patients, the epithelial subtypes were prognostic (log-rank OS p=0.00004, n=160 patients), and Cox proportional hazards modelling (CPH) showed that the heterogeneous subtype 6 had a hazard ratio significantly > 1, indicating poor prognosis (Figure 5.6c). The CPH hazard ratios (HRs) were similarly ordered across the IMC and CyCIF platforms, with heterogeneous subtype 6 and CK low subtype 1 having HRs > 1, or poor prognosis, and Luminal ER+ subtype 1 and Luminal subtype 0 having HRs < 1, or good prognosis (Figure 5.6d). Since the Luminal ER+ subtype was associated with good prognosis, we also investigated Luminal ER+ tumor cell fraction as a single variable, and it was found to predict longer OS (log-rank p=0.036, n=162 patients, Figure 5.6e). In a multivariable CPH model including stage, patient age and tumor size, the fraction of Luminal ER+ tumor was no longer significantly associated with survival (p=0.18, Figure 5.6f). In TNBC patients, the epithelial subtypes had borderline prognostic significance (log-rank p=0.1; CPH p=0.051, n=98 patients, Figure 5.6g). However, the epithelial subtypes did not show a consistent order in hazard ratios obtained from CPH modelling across platforms in TNBC (Figure 6.4 a).

5.4.4 *Microenvironment subtypes associated with clinical subtype*

We then clustered patients based on the stromal cell type fraction in each tissue in a similar manner to the epithelial subtyping above, selecting $k=6$ for the number of clusters (Figure 6.5). Since the stromal phenotypes differed across platforms due to different markers, we clustered each platform separately, using the fractions of stromal cell types representing greater than 2% of the stromal compartment. The stromal subtypes were not prognostic, with the exception of the MIBI platform (log-rank=0.003, CPH $p=0.056$, $n=39$ patients Figure 6.5b-g). However, we observed significant correlation between stromal subtypes and clinical subtypes ER+ and TNBC. In the CyCIF cohort, ER+ patients had significantly more of the Vimentin+ stromal subtype 0 and less T cell and macrophage rich stroma (Chi-squared $p=0.098$, Bonferroni $p\text{-adj}<0.05$ for subtype 0, $n=89$ patients [ER+ and TNBC, including patients without clinical outcome], Figure 6.5e). In the IMC cohort, ER+ patients had more Vim+/FN+ fibroblast stromal subtype 0 and significantly less T cell stromal subtype 4 (Chi squared $p=0.002$, Bonferroni $p\text{-adj}<0.05$ for subtype 4, $n=220$ patients [ER+ and TNBC, including patients without clinical outcome] Figure 6.5h). Our characterization of stromal subtypes confirms the general understanding that ER+ breast cancer is immune-poor⁵², and also shows significant enrichment for Vimentin and FN+ fibroblasts in ER+ relative to TNBC.

5.4.5 *T cells are an independent prognostic factor in TNBC*

Next, we investigated the prognostic value of cell types within each clinical subtype. We used the CyCIF dataset as a training cohort to identify prognostic cell types and cut-offs for high or low abundance. We tested the 0.33, 0.5 and 0.66 quantile to binarize tissues into low and high cell abundance (Figure 6.6a-c). For any cell showing prognostic significance (log-rank $p<0.05$), we selected the cut-off with the lowest p -value for validation in the other cohorts (Figure 6.6d-e). The prognostic significance of T and B cells in TNBC was confirmed across multiple platforms (CyCIF $p=0.006$, $n=18$, IMC $p=0.068$, $n=41$, T cells; CyCIF $p=0.049$, $n=18$, MIBI $p=0.0069$, $n=39$ B cells, Figure 6.6). High macrophages ($p=0.043$), low endothelial cells ($p=0.018$) and high proliferating tumor ($p=0.031$) predicted survival in CyCIF TNBC samples ($n=18$), but were not significant in the other cohorts (Figure 6.6). In ER+ patients, high Luminal ER+ tumor (IMC, $p=0.034$, $n=132$), low HER2+ Ki67+ (IMC $p=0.058$, $n=132$) and low CK low tumor (CyCIF $p=0.067$, $n=30$) were associated with longer survival, but only in a single cohort (Figure 6.6).

We then combined all samples into a single cohort for multivariable CPH modelling with the clinical variables of stage, patient age and tumor size. In combined TNBC samples, high CD3, CD4 and CD8 T abundance were significantly associated with longer OS (log-rank $p=0.018$, $n=98$; $p=0.0048$ and 0.00023 , $n=57$ respectively Figure 5.7 a, b), and they remained significant in a multivariable model (CPH $p=0.01$, $n=55$, $p=0.049$ and 0.0071 , $n=15$, Figure 5.7c). High CD20 B cell abundance was associated with longer OS (log-rank $p=0.0013$, $n=98$) but it did not remain significant in a multivariate model (CPH $p=0.15$, $n=55$, Figure 5.7d). CD3 T cell abundance also predicted longer RFS in TNBC ($p=0.028$, $n=98$), which remained significant in the multivariable model ($p=0.026$, $n=55$, Figure 5.7e). CD20 had borderline predictive value for RFS (log-rank $p=0.054$, multivariable CPH $p=0.082$, $n=55$), while CD8 and CD4 T cells were not independently predictive of RFS in the combined CyCIF and MIBI cohorts ($n=15$, Figure 6.4 b - d).

In ER+ breast cancer, Luminal ER+ tumor predicted longer OS ($p=0.036$, $n=162$, Figure 5.6e) but not RFS (Figure 6.4e). Conversely, HER2+ Ki67+ tumor predicted shorter RFS in ER+ and was borderline predictive in HER2+, but did not predict OS (Figure 6.4 f, h). Tumor proliferation did not predict survival or recurrence in ER+ breast cancer (Figure 6.4 g). CD3 T cells did not predict survival in ER+ breast cancer, however, in the subtype 5, ER+ breast cancer having elevated HER2 positivity, there was a trend towards T cells predicting longer survival (log-rank $p=0.076$, $n=25$), a trend not seen in the Luminal ER+ or Luminal subtype (Figure 5.7f-g).

Overall, our approach to training on one platform and validating in another revealed CD3 T cells as robust biomarkers of OS and RFS in TNBC, even when combined with clinical variables. B lymphocytes also showed predictive value. Since large cohort studies have established the prognostic significance of lymphocytes in TNBC⁴⁰, our findings validated our approach for biomarker discovery.

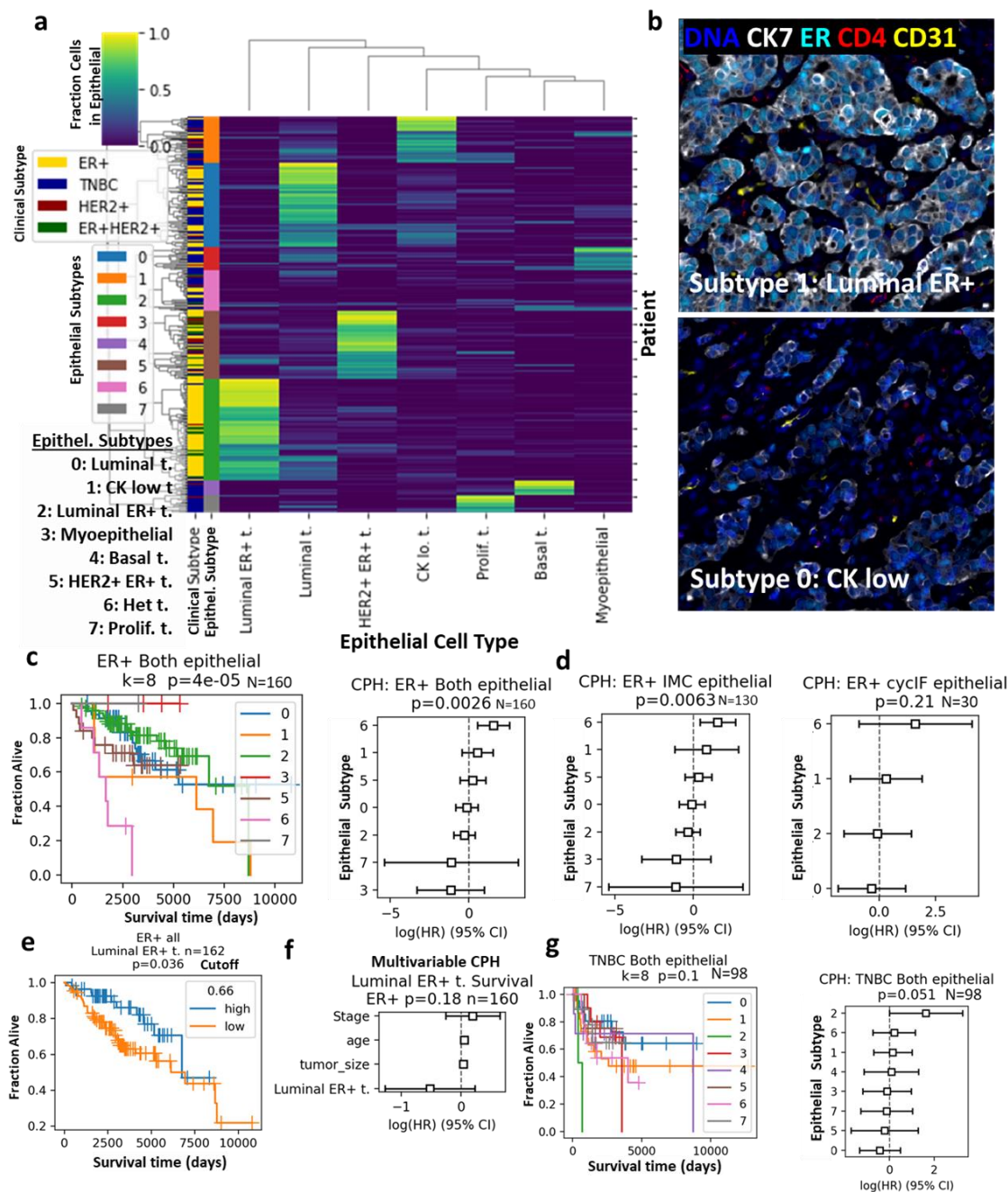


Figure 5.6. Prognostic breast cancer subtypes defined by multiplex imaging.

a. All patients were hierarchically clustered based on the fraction of the seven most common epithelial cell types in each patient's tissue. This resulted in eight epithelial (Epithel.) subtypes b. Representative CyCIF images of subtype 1 luminal ER+ tumor (top) and subtype 2 cytokeatin-low tumor (CK low, bottom). c. Kaplan-Meier curves (p-value from log-rank test) and Cox proportional hazard (CPH) models comparing OS in seven epithelial subtypes that were present in ER+ tumors. d. CPH models of epithelial subtypes in ER+ tumors for IMC platform (left) and CyCIF platform (right) e. Kaplan-Meier curves for Luminal ER+ tumor cells versus OS (highest tertile versus lower tertiles; left; log-rank p=0.036), and a multivariable CPH model adding patient age, tumor size and clinical stage (right). f. Kaplan-Meier and CPH modeling of eight epithelial subtypes in TNBC versus OS (log-rank p=0.1, CPH p=0.051). c-g. P-values given in figure title obtained through log-rank test for Kaplan-Meier analysis, and CPH modelling. N number of patients given in figure titles.

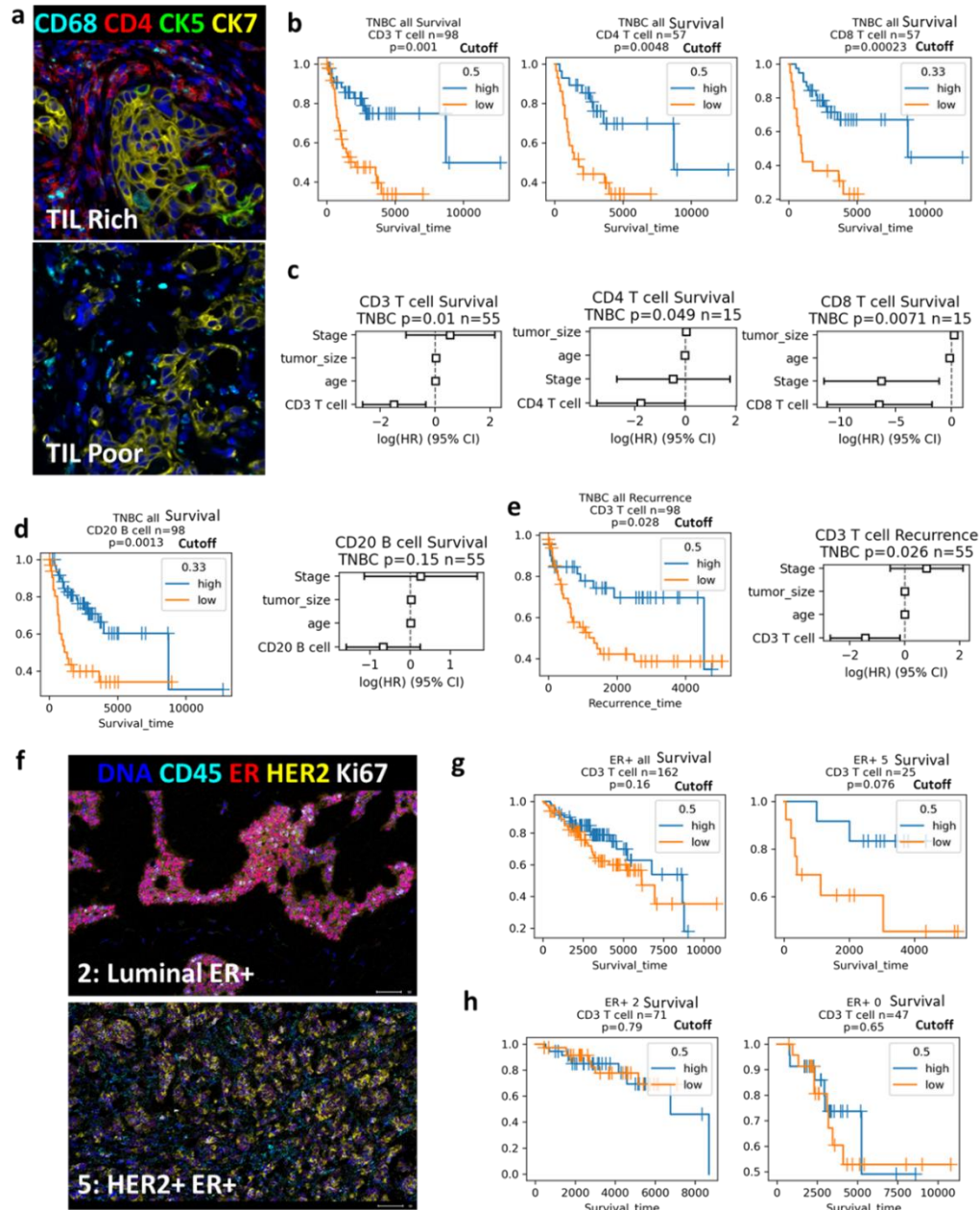


Figure 5.7. Immune infiltrate has prognostic value in TNBC breast cancer.

a. CyCIF images of tumor infiltrating lymphocyte (TIL)-rich (top) and TIL-poor (bottom) TNBC tissue. b. Kaplan-Meier survival curves based on TNBC patients' abundance of CD3 T cells (above the median) and CD8 (upper tertile). c. Multivariate CPH modeling adding patient age, tumor size and stage to CD3, CD4 and CD8 T cell high variable defined in (b). d. Kaplan-Meier curve based on TNBC patients' abundance of CD20 T cells (upper tertile, left), and a multivariable CPH model including CD20 high abundance and clinical variables (right). e. Kaplan-Meier curve based on TNBC patients' abundance of CD3 T cells (above the median) versus RFS (left), and a multivariable CPH model of RFS (right). f. ER+ patient tumors from the IMC dataset showing strong ER staining (top) versus elevated HER2 staining (bottom). g. Kaplan-Meier curve based on ER+ patients' abundance of CD3 T cell versus OS in all ER+ patients (left), and in ER+, HER2+ subtype 5 (right). h. Kaplan-Meier curves for CD3 T cell abundance versus OS in Luminal ER+ tumors (subtype 2), and CK low tumors (subtype 0). b,d,e,g,h. All Kaplan-Meier p-values obtained from the log-rank test, n given in figure titles, cutoff quantile for high/low indicated in figure, e.g. 0.5 indicates the median was used to binarize patients. c, d, e. CPH modelling p-values for cell type variable given in figure titles, n number of patients given in figure titles.

5.4.6 *Tumor phenotypes correlate with stromal cell abundance and proximity*

We hypothesized that there would be significant correlation between tumor and stromal cell types in tumor tissues and that these correlations could shed light on biologically and clinically relevant tumor-stroma crosstalk. First, we visualized a matrix of pairwise correlation between stromal and epithelial cell fractions across the three platforms (Figure 6.7). Epithelial cell types were anti-correlated with each other, indicating most tumors had just one main epithelial cell type. The exception was Luminal tumor was slightly correlated with CK low tumor, indicating mixing of these phenotypes within the same tissues. This conclusion is supported by our hierarchical clustering, which revealed luminal tumor as the main epithelial cell type likely to be found mixing with other phenotypes in the same tissue (Figure 5.6a).

Immune cells, including CD4 and CD8 T cells, B cells, dendritic cells, FoxP3 T regs, and macrophages, as well as endothelial cells, were correlated with each other (Figure 6.7). Vimentin+, FN+, and ColI+ fibroblasts, as well as quiescent stroma were anti-correlated with immune cells (Figure 6.7). In the combined correlation data, CD3 T cells anti-correlate more strongly with quiescent stroma, ColI+ and FN+ fibroblasts as opposed to Vimentin+ fibroblasts (Figure 6.7d). Neutrophils and NK cells, while only present in the MIBI dataset, appeared to be uncorrelated with T cells and other immune cells (Figure 6.7b). Together, the correlation data indicate mostly homogenous tumor phenotypes, and several distinct microenvironments, including immunoreactive (lymphocyte, dendritic cell and macrophage high), fibroblast-rich (at least two distinct populations, Vimentin high or ECM high), quiescent, and potentially an innate immune dominant neutrophil/NK cell microenvironment. The former three microenvironments were also evident in our stromal subtypes defined by hierarchical clustering (Figure 6.5).

In the IMC dataset, ER+ luminal tumor correlated with FN+ fibroblasts in the tissue ($r=0.36$, $p<0.00001$, $n=170$ patients, Figure 5.8a, b). Since fibronectin has been shown to stabilize ER through colocalization of ER with fibronectin's receptor, $\beta 1$ -integrin²⁵⁰, we hypothesized that fibronectin would be in close proximity to ER+ tumor cells. We calculated the number of neighbors within 15 μm , or roughly one cell diameter, of each cell, and confirmed that Luminal ER+ tumor cells had significantly more FN+ fibroblast neighbors than other tumor phenotypes (Kruskal-Wallis $p=0.00427$, $n=144$ patients, Figure 5.8c). Vimentin+ fibroblasts were correlated with Luminal ER+ tumor in IMC and cycIF datasets (Pearson $r = 0.19, 0.15$ Figure 6.7 e) but they were not significantly closer to Luminal ER+ tumor (Kruskal Wallis CyCIF $p=0.05$, IMC $p=0.045$ favoring proliferating tumor, Figure 6.7 f). Therefore, although both FN+ and Vimentin+ FB occur in ER+ tissues, only FN+ FB locate closer to ER+ tumor cells than other phenotypes.

Proliferating tumor was correlated with T cell abundance in ER+ samples in IMC but not CyCIF data (Pearson, $p=0.008$ and $p=0.36$, $n = 170$ and 30 patients), however, CD3 T cell neighbors were significantly higher around proliferating tumor cells in both platforms (Kruskal-Wallis, $p=0.0048$, $n=99$ IMC, $p=0.02$, $n=26$, cycIF, Figure 5.8 d). Endothelial cells had similarly increased proximity to proliferating tumor (K-W $p=0.0015$, $n=121$ IMC, $p=0.03$, $n=19$ CyCIF) although they did not correlate with abundance in both platforms ($p=0.01$, $n=170$ IMC, $p=0.69$, $n=30$ CyCIF, Figure 5.8 e). Quiescent stroma was anti-correlated with Luminal ER+ tumor and correlated with CK low tumor (Pearson $p<0.00001$, $n=170$, and $p=0.0013$, $n=50$ patients), and showed significantly higher and lower proximity, respectively, relative to those tumor phenotypes as well (K-W $p<0.00001$,

n=149, and p=0.003, n=40, Figure 5.8 f). In summary, most cell types with correlated abundances in the tissue were also closer, with the exception of CD3 T cells and endothelial cell proximity to proliferating tumor. Therefore, we hypothesized that T cell proximity to proliferating tumor could have unique prognostic significance, as this interaction would not be correlated with abundance.

To further confirm tumor and stromal cell type correlations, we investigated our tumor subtypes and found similar patterns on the subtype level, with the Luminal ER+ subtype significantly enriched for FN+FB, and the proliferating subtype significantly enriched for T cells (Tukey's HSD p=0.008 and p=0.002, Figure 6.8). Interestingly, we found that the heterogenous, poor prognosis subtype 6 was associated with significantly higher endothelial cells in the stroma in ER+ breast cancer (Tukey's HSD p=0.001, Figure 6.8).

In summary, we found notable correlations between tumor and stromal cell types and proximity. Our data support a mechanistic link between FN in close proximity to tumor cells supporting ER expression²⁵⁰. We also found correlations between tumor proliferation and T cells, and T cell proximity that could not be explained by abundance. This is interesting given previous work delineating proliferation as a biomarker of productive anti-tumor immunity^{43,79}.

5.4.7 *Lymphocyte-proliferating tumor proximity predicts survival in TNBC*

Finally, we sought to investigate the high proximity between proliferating tumor and T cells in relationship to prognosis. We calculated the number of T cells in a 25-um radius around tumor cells, which corresponds to one tumor cell diameter. In ER+ breast cancer, there were more T cell neighbors around proliferating tumor (Figure 5.8d), but this was not true in TNBC (Figure 5.9b). There was a correlation between T cells and proliferating tumor in the IMC cohort but not the other data sets (Figure 6.7 g), therefore, we do not expect higher T cell proximity simply due to higher abundance in proliferating versus non-proliferating tumors. However, a high number of CD3 T cell neighbors around proliferating tumor predicted longer overall and recurrence free survival in TNBC (log-rank p=0.00005 OS, 0.001 RFS, n=92, Figure 5.9 c, d), and the CD3 T cell neighbors of proliferating tumor remained predictive in a multivariable CPH model (p=0.024 OS, p=0.038 RFS, n=52, Figure 5.9 c, d). CD20 B cells were not closer to proliferating tumor than other tumor phenotypes (Figure 6.9 a), yet again, CD20 B cell neighbors around proliferating tumor cells were also predictive of longer survival (CPH multivariable p=0.035 OS, n=52, Figure 5.9 e, Figure 6.9 b), although their proximity to non-proliferating tumor also predicted OS in the multivariable model (CPH multivariable p=0.035, n=54, Figure 5.9 e, Figure 6.9 c), a trend seen across platforms (Figure 6.9 f, g). In contrast, CD3 T cell proximity to non-proliferating tumor was not prognostic in the multivariable model (log-rank p=0.02 OS, p=0.55 RFS; CPH multivariable p=0.08 OS, p=0.21 RFS, Figure 5.9 e, Figure 6.9 d). T cell to proliferating-tumor proximity showed significant prognostic value in all three cohorts independently (log-rank p=0.013 CyCIF, 0.016 MIBI and 0.042 IMC Figure 6.9 e). Therefore, CD3 T cell proximity specifically to proliferating, but not non-proliferating tumor is an independent predictor of overall survival.

We compared the performance of the lymphocyte-proliferating tumor biomarker to a previously published biomarker in TNBC, the immune-tumor mixing score, which Keren et. al. developed in the MIBI cohort¹¹¹. In order to calculate this metric, the authors first excluded immune-poor tissues with less than 250 immune cells, or six tissues from their cohort. We excluded tissues with less

than 1% immune cells in the patient tissues, which corresponded to excluding five MIBI tissues (our cohort was 39 as opposed to the original 41). Keren et al. defined mixing score as the number of immune-tumor interactions divided by the number of immune-immune interactions, with interactions defined as cells directly touching. We set our neighborhood radius to 25 μm which corresponds to one tumor cell diameter and captures touching cells. Keren et al. set a threshold for high mixing at 20/35 patients, or the 0.6 quantile. We set ours at the upper tertile, or 0.66. Using these parameters, we were able to reproduce the prognostic value of the mixing score in the MIBI cohort with the same p-value reported by the original authors (log-rank $p=0.03$, $n=37$ Figure 6.9 h). However, the mixing score was not prognostic in the other two cohorts (log-rank $p=.21$, $n=17$ cycIF, $p=0.46$ $n=31$ IMC, Figure 6.9 h). In the combined cohort, mixing score predicted shorter OS (log-rank $p=0.0092$, $n=85$), but it was not prognostic in the multivariable model (CPH $p=0.41$, $n=45$) and did not predict RFS as a single variable ($p=0.1$) or in the multivariable model ($p=0.39$, Figure 5.9f).

5.5 DISCUSSION

Our approach of combined analysis across multiple platforms shows the power of our methods for robust biomarker discovery. We developed a single-cell spatial metric of survival, lymphocyte-proliferating-tumor proximity, that was superior to a previously published biomarker developed in a smaller cohort on a single platform. We were able to incorporate analysis of two publicly available imaging datasets with our own CyCIF data for efficient discovery of robust biomarkers.

Using our image processing and analysis pipeline, we generated single cell phenotyping data from three multiplex imaging datasets. The advantage of using mplexable for image processing and feature extraction is the ability to match nuclear and cell segmentation results from separate deep-learning segmentation models and extract features from subcellular compartments such as the nucleus and cytoplasm (Figure 6.1b). We found that cell type fractions across cohorts from the same breast cancer subtype were highly correlated across platforms. Therefore, we combined the epithelial cells and identified eight subtypes with hierarchical clustering. We identified subtypes similar to the intrinsic breast cancer subtypes²⁷⁰, including a luminal ER+ luminal A-like group with good prognosis, a cytokeratin low group previously shown to share features with luminal B tumors¹¹³, and a HER2 enriched group. Triple-negative breast cancers also fell into categories similar to those defined by gene expression profiling¹²⁹, including a highly proliferative, basal-like 1 (BL1) group, a luminal androgen receptor (LAR)-like group with luminal differentiation states, and subtypes of basal/myoepithelial phenotype reminiscent of the basal-like 2 (BL2) group. Our proliferative subtype was correlated with immune infiltrate, which was in turn correlated with better prognosis, suggesting that the relatively good prognosis typically seen in proliferative BL1 TNBC patients may be linked to higher immunogenicity of the tumor. We also identified a heterogeneous group with low cytokeratin expression that may represent tumors with mesenchymal features. The ER+ tumors in this subtype had significantly worse prognosis, and subsequent analysis showed increased angiogenesis in these tissues. An EMT-program in breast cancer cells is linked to increased vascular endothelial growth factor A expression, increasing angiogenesis and the capacity for tumor initiation²⁷¹, a mechanism that could explain these correlated tumor and stromal phenotypes and their association with poor outcome.

Next, we identified significant correlations between stromal and tumor phenotypes. First, we found that clinical subtypes correlated with stromal subtypes, with ER+ tissue enriched for Vimentin+ fibroblasts and TNBC enriched for T cell and macrophage microenvironments. We then found correlations between our tumor subtypes and stromal cells, with Fibronectin (FN)+ fibroblasts correlated with the luminal ER+ subtype and T cells correlated with the proliferative subtype. We hypothesized that FN in close proximity to tumor cells could stabilize ER, and indeed we found significantly more FN+ fibroblasts surrounding luminal ER+ compared to other tumor phenotypes.

Next, we found T cells are an independent prognostic factor in TNBC, which agrees with the literature linking TILs to improved outcomes in TNBC patients⁴⁰. We also identified a group of ER+ HER2+ breast cancer patients who showed a trend towards survival benefit from high T cell infiltration. Finally, we examined the spatial relationship between T cells and proliferating tumor and revealed that lymphocyte-proliferating tumor proximity independently predicts survival in TNBC, while lymphocyte-non-proliferating tumor proximity does not.

Previous gene expression profiling studies link anti-tumor immunity and proliferation. Nagalla et al. found that immune signatures were prognostic solely in breast cancer patients with the highest proliferation gene expression⁴³. Subsequently, the same group showed that immune gene signatures were prognostic in highly proliferative basal-like, HER2-enriched and luminal B subtypes, but not those with low proliferation⁷⁹. In luminal A and claudin-low subtypes, immune gene signatures were not prognostic, regardless of proliferation⁷⁹. Indeed, our lymphocyte-proliferating tumor proximity biomarker did not stratify ER+ patients. Gene network analysis showed activation of TNF α /IFN γ signaling pathways in tumors with productive anti-tumor immunity and TGF- β , an immunosuppressive cytokine, was activated in tumors with unproductive anti-tumor immunity. TGF- β also has anti-proliferative effects and is associated with good outcome in ER+ breast cancer cohorts²⁷², suggesting that it could mechanistically link lower proliferation rates with immunosuppression. More recently, the group mentioned above showed that immune gene signatures were prognostic exclusively in tumor-mutation burden (TMB)-high breast cancer tumors^{273,274}. Thirty-eight percent of HER2 enriched and 37% of basal-like tumors had high TMB, while only 11.5% of luminal A tumors did²⁷⁴, explaining the poor immunogenicity of the latter subtype. Together, these data point to a model of high TMB correlating with high-proliferation and both linked to productive anti-tumor immunity. It had been hypothesized that oncogenes driving sustained proliferation also induce DNA replication stress, which generates genomic instability and would increase TMB²⁷⁵. In summary, TMB mechanistically links proliferation and anti-tumor immunity and we hypothesize its elevation in the good-prognosis tumors with high lymphocyte-proliferating tumor proximity.

Intriguingly, proliferation status added to PD-L1 status has been proposed as a biomarker to better stratify response to immune checkpoint blockade in non-small cell lung cancer (NSCLC) patients²⁷⁶. A strongly predictive biomarker is the tumor immunogenic signature (TIGS) derived from inflamed tumors with CD8 T cell infiltration into the tumor core²⁷⁷. A high TIGS score predicted response to immunotherapy in highly and moderately proliferative tumors compared to low proliferation tumors, and was able to identify responders in PD-L1 negative and TMB-low tumors²⁷⁷, thus functioning better than those two leading biomarkers for immunotherapy response. Since our lymphocyte-proliferating tumor biomarker accounts for both immune infiltration into the tumor core and proliferation status, it could also be a biomarker for immunotherapy response.

The limitations of our study include different antibody probes and imaging systems resulting in different signal-to-background ratios for biomarkers across platforms. Therefore, we relied on matching hand-annotated clusters across platforms for our integrated analysis. Since our annotations may not correspond to the same cell types in each platform, this introduces uncertainty. Some well-defined phenotypes, such as T cells and proliferating tumor, are relatively straightforward. However, variable performance of antibodies, such as anti-ER, for example, could lead to variability in the classification of phenotypes such as luminal ER+ versus luminal tumor across platforms. To correct for platform-specific bias in cell types, we binarized patients into high/low expression within each subtype and platform. However, such binarization may not reflect to underlying heterogeneity in quantitative biomarker abundance.

Another limitation of our study is the use of 1-2 TMA cores per patient for analysis. It has been shown that a limited number of TMA cores (≤ 3) are needed to binarize patients into high and low TILs, although a larger number of cores (≥ 11) are needed to accurately estimate the mean TIL abundance^{36,37}. Our survival analysis relied on binarizing patients; therefore, the use of TMAs may be appropriate in this context. However, our tumor-stroma correlation analysis, which relied on mean abundance, would be expected to suffer from noise introduced by sampling just 1-2 cores. Indeed, examination of correlation scatterplots shows relatively low, if significant, Pearson's R values (< 0.5 , Figure 5.8). Alternatively, our analysis of correlation between clinical and stromal subtypes (Figure 6.5) identified previously observed patterns such as the relatively immune-cold stroma of ER+ breast cancer⁵². Subtype analysis, which uses multiple cell types to define stromal subtypes, may be less subject to sampling noise than single-parameter correlations. Future studies of tumor-stroma correlation could analyze larger tissue areas and/or additional TMA cores to account for sampling error. The optimal tissue area for various composition and spatial analytics remains an open and important question in the field.

In conclusion, our single-cell spatial profiling of breast cancer tumors revealed an independently prognostic biomarker, lymphocyte-proliferating tumor proximity, that captures the biology underlying tumor immunogenicity, including tumor proliferation and immune infiltration into the tumor core. Future work will define the utility of this biomarker in predicting response to immunotherapy.

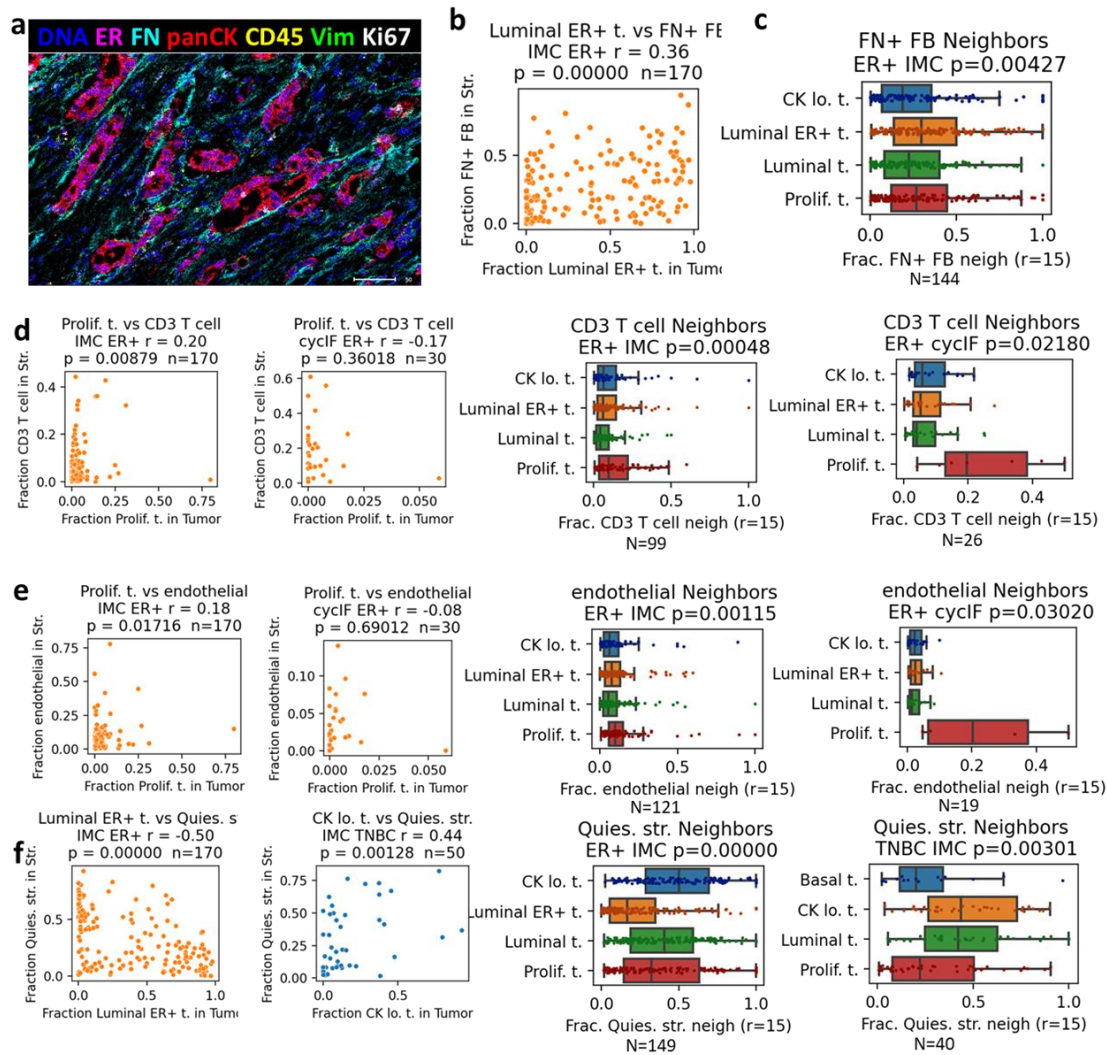


Figure 5.8. Tumor phenotypes correlate with stromal cell types and proximity.

a. IMC image showing strong ER and fibronectin (FN) staining in close proximity. b. Correlation between the fraction of tumor cells with a Luminal ER+ phenotype and the fraction of FN+ fibroblasts in the stroma (Pearson $R = 0.39$, $p < 0.00001$). c. Fraction of stromal neighbors that are FN+ FB relative to tumor phenotypes. d. Correlation in tissue and stromal neighbor fraction between T cells and tumor phenotypes. e. Correlation in tissue and stromal neighbor fraction between endothelial cells and tumor phenotypes. f. Correlation in tissue and stromal neighbor fraction between quiescent stroma and tumor phenotypes. b-f. All correlation r - and p -values are Pearson correlation, all boxplot p -values are Kruskal-Wallis H test for difference in medians. N number of patients included in figure title.

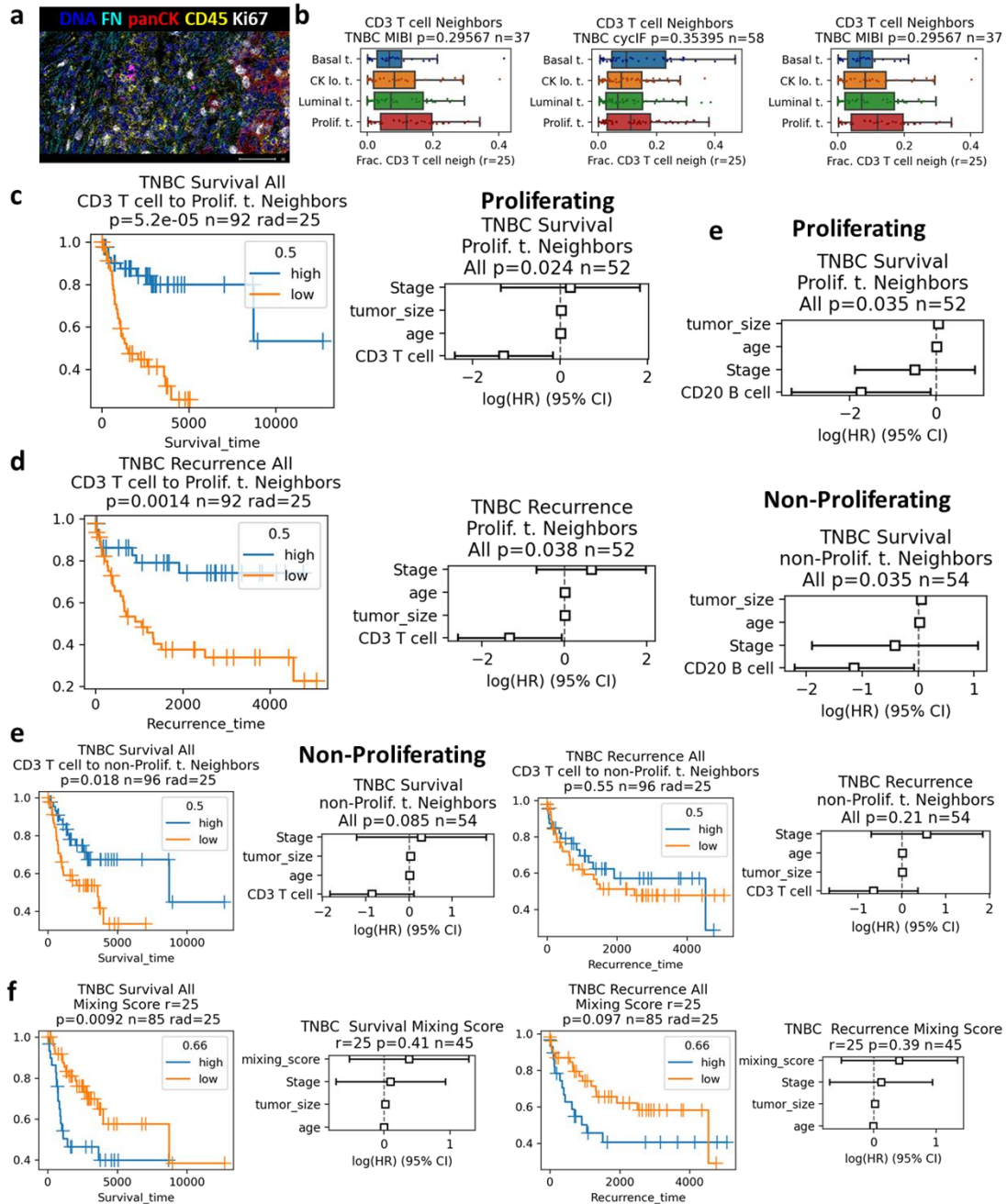


Figure 5.9. Immune-proliferating tumor proximity predicts survival in TNBC.

a. IMC image of highly proliferative tumor showing immune cells in proximity to tumor, scale = 50 μ m. b. Fraction of T cells as stromal neighbors relative to tumor phenotypes. c. T cells neighboring proliferating tumor versus OS. d. T cells neighboring proliferating tumor versus RFS. e. B cells neighboring proliferating tumor versus OS and RFS. e. T cells neighboring non-proliferating tumor versus OS and RFS. f. Tumor-immune mixing score versus OS and RFS. b-f. All Kaplan-Meier p-values are from the log-rank test; all boxplot p-values are Kruskal-Wallis H test for difference in medians. For all fraction of neighbors, high = above the median. N number of patients included in figure title.

Chapter 6. SUMMARY AND FUTURE PERSPECTIVES

Heterogeneous cell types and spatial patterns are linked to patient prognosis and treatment response in breast cancer, but a systems-level understanding of how these cell types interact necessitates multiparameter profiling of the tumor microenvironment. Comprehensive characterization of the cellular and extracellular components in breast cancer should improve prediction of therapy response and tumor progression. Multiplex imaging allows identification of cellular assemblages and spatial arrangements that correlate to cancer aggressiveness and relapse. Such data not only aid in therapy selection and patient stratification, but yield deeper understanding of the interacting biological systems underlying tumor progression.

This thesis describes work undertaken to optimize cyclic immunofluorescence, a multiplex imaging protocol, and create open-source software for image processing and quantitative analysis of human tissues and model organisms. In **Chapter 2**, we established our breast cancer antibody panel and protocol for cyclic multiplex immunofluorescence (cmIF, later CyCIF) staining of formalin-fixed, paraffin embedded (FFPE) tissues, as well as a manual pipeline for image analysis. Since the manual pipeline was highly labor intensive and error-prone, in **Chapter 3** we develop tools for scalable image processing and reproducible analysis. We also designed, undertook and analyzed experiments leading to optimization of signal removal, antibody specificity, background correction and batch normalization in CyCIF. These developments enabled us to reduce or computationally correct for imaging artifacts in subsequent quantitative analyses in **Chapters 4** and **5**. In **Chapter 4** we develop additional novel analytical methods for characterizing tumor heterogeneity and spatial relationships in diverse tissue samples. In **Chapter 5**, we utilize our CyCIF platform for biomarker discovery in clinical cohorts of breast cancer patients, and also extend our tools to utilize publicly available multiplex ion beam imaging (MIBI) and imaging mass cytometry (IMC) datasets as validation cohorts for efficient discovery of robust biomarkers.

6.1 CHAPTER 2: GETTING STARTED WITH CYCIF AND IDENTIFYING CHALLENGES

In our previously published methods chapter¹¹⁰, we established a protocol for CyCIF staining in FFPE tissues (Figure 2.1) by adapting a protocol developed in cell lines¹⁰⁷. We describe the basics of antibody validation, a laborious process as > 50% of prospective antibodies fail. We validated our staining against immunohistochemistry (Figure 2.2) and showed basic reproducibility through high correlation across adjacent tissue sections (Figure 2.3). We then demonstrate successful registration of separate rounds of staining for visualization of biomarkers (Figure 2.4) and extract single cell features to complete basic k-means clustering for cell type identification (Figure 2.5). A co-author, Dr. Chang, demonstrates TNSE visualizations and cell proximity analysis using the proprietary software Matlab (Figure 2.5 and Figure 2.6).

This early work revealed technical and computational issues that hindered scalable and artifact-free quantitative analysis. The major technical challenge was with image processing. Each dataset took weeks to months to process, and we had a 100% failure rate in producing error-free data. One

of the most time consuming and computationally expensive events was finding errors during downstream analysis that would necessitate going back and re-segmenting cells, wasting time and computational resources. If we had lacked an in-house high-speed computational cluster, as many researchers do, these types of mistakes would have incurred significant financial costs for computing time.

On the technical side, autofluorescence was a significant artifact that was often brighter than the staining we were attempting to image. Other technical challenges included incomplete quenching, channel bleed through, non-specific background staining, tissue loss, and batch effects. All of these challenges motivated the next phase of our work to address the two towers that challenged successful multiplex imaging: the Orthanc-like technical artifacts marring beautiful imaging data and the true enemy in Barad-dûr: the error-prone manual image processing pipeline.

6.2 CHAPTER 3: MAKING CYCIF WORK WELL

In our recently accepted paper²⁵⁷, we developed the open-source python software, mplexable, for reproducible, scalable processing and analysis of multiplex images (Figure 3.1). While there are other options for image processing pipelines^{263,278}, mplexable is unique in that it was developed from the perspective of a data generator who also analyzes data. Therefore, I had knowledge of the both requirements of automated pipelines (e.g., standardized folder structures and file names) as well as the reality of the data coming off the scanner (e.g., typos, re-scans, focus issues). As a result, mplexable included many tools for quality control, such as automated file name checking and correction, thumbnail generation and automated exposure time extraction, since errors are virtually guaranteed when copying a table of exposure time values from a notebook to a text document for the software to read. Beyond QC tools, mplexable introduced autofluorescence (AF) subtraction of images, automatic parallel processing of time-consuming registration, segmentation and feature extraction steps, and a novel algorithm for matching segmentation results from separate masks. This allows utilization of the best algorithms for nuclear and cell segmentation, which mplexable then matches and processes to extract numerous subcellular compartments including cell and nuclear membrane, nucleus and cytoplasm. A key feature of mplexable is automatically running parallel jobs using other software, enabling the use of the best tool or algorithm for each step in a mix and match approach that facilitates processing of different types of data, all while documenting processing steps for full reproducibility.

Our first demonstration of mplexable was the somewhat unglamorous but important processing of validation data. We processed large data sets quantifying signal-to-background and tissue loss in CyCIF (Figure 3.2). We found that CyCIF signal-to-background and cell type calling were highly correlated to data from a standard IF protocol (Figure 3.2) and that tissue loss was specific to tissue types and fixation as opposed to specific cell types (Figure 3.2). Next, we varied quenching conditions and antibody order to optimize signal removal and staining quality (Figure 3.3). As alluded to above, we addressed autofluorescence (AF) artifacts by collected three separate AF datasets, quantifying AF over rounds of CyCIF, and implementing methods for computational AF subtraction (Figure 3.4). Finally, we quantified batch effects in two separate CyCIF experiments each evaluating three adjacent TMA slides stained in different batches. We tested methods to normalize for batch effects without losing the biological signal (Figure 3.5).

Through the work undertaken in Chapter 3, we established an efficient, error- and artifact-reduced protocol and computational tool for scalable and reproducible quantitative image analysis. The utility of the methods and tools are demonstrated in applications covered in Chapters 4 and 5.

6.3 CHAPTER 4: QUANTIFYING AND INTERPRETING TUMOR HETEROGENEITY

This chapter highlights three collaborative projects in which we contributed CyCIF data and analysis to larger multi-omic studies. The first study investigated mechanisms of response and resistance to the HER2 targeted therapy T-DM1. The second study focused on monitoring tumor and stromal responses to therapy in a metastatic ER+ breast cancer patient. The third study was a comparative analysis of tumor tissues from a mouse model of TNBC relative to human TNBC.

In the first study, we hypothesized that intratumor heterogeneity would be a mechanism of resistance to neoadjuvant T-DM1 in HER2+ breast cancer patients. However, we found that the diversity of tumor phenotypes was not significantly different in responders and non-responders, but the fraction of HER2+ cells was higher in tumors that responded to therapy (Figure 4.1). We also showed that CD8 T cells were more abundant in responders (Figure 4.1). These findings suggest that increased drug target abundance (HER2) and a pre-existing immune response are present in responders. Our findings are in agreement with analyses of TILs predicting of pCR in patients treated with neoadjuvant trastuzumab, lapatinib (a small molecule inhibitor of HER2) and chemotherapy⁶⁵. Spatial analysis revealed differences in the immune-tumor architecture of responders, who had more tumor-tumor interactions and more immune-Vimentin+ fibroblast interaction (Figure 4.1). Based on this, we hypothesized that tumor-immune mixing could be a prognostic biomarker in HER2+ breast cancer as in TNBC¹¹¹, but our subsequent analyses in Chapter 5 did not support this hypothesis.

The limitations of this study were small sample size and use of pCR as an endpoint rather than survival. There were three interesting cases, one responder with an ER+ HER2- tumor, and two non-responders that had strong and moderate HER2 expression, respectively. We don't know if these patients would eventually respond or recur after adjuvant therapy, or if any of the correctly predicted responders/non-responders would flip categories. A potential follow-up study would be quantitative imaging of just HER2, CD8 and an epithelial marker such as panCK to get the same readout without doing CyCIF. This would enable us to evaluate a larger cohort with targeted quantitative imaging.

In the second study, we performed CyCIF on metastatic breast cancer biopsies that were also profiled by a suite of other omics methods. Our results agreed with other methods such as multiplex immunohistochemistry¹⁰⁵ and RNA seq, further validating the batch normalization methods we developed in Chapter 3. An interesting spatial finding was increased ER and PCNA expression in close proximity to extracellular matrix (ECM) proteins collagen I and collagen IV (Figure 4.2). We hypothesized that ECM, possibly fibronectin, mediated ER stability and increased expression. We followed up on this hypothesis in Chapter 5, and found it supported, with ER+ tumor cells having significantly more fibronectin positive fibroblast neighbors.

In the third study, we did a comparative analysis of mouse and human tissues in order to validate a mouse model of TNBC. We found that the MYC:PTENfl mouse mammary tumors shared many features with human TNBC, including an immune-rich subtype with mesenchymal tumor states and an aggressive, stroma-poor subtype with increased endothelial cells in the stroma that resembled poor prognosis human TNBC (Figure 4.4). Our characterization of the two datasets revealed potential biomarkers for EMT, including S100A6, which was elevated in mouse stroma-rich, mesenchymal subtype tumors, and β -catenin, which was elevated in human stroma-rich tumors. These would be interesting biomarkers for follow up studies with RNA seq data from which to derive the mesenchymal subtype gene expression signatures^{129,251}.

6.4 CHAPTER 5: THE PATH TO ROBUST BIOMARKER DISCOVERY

In the final chapter, we leveraged our mplexable pipeline to process not only our own CyCIF data but two publicly available breast cancer datasets, bringing the total number of patients in our analysis to 413, 296 with clinical outcome data. Given the small size of our HER2+ training set (8 patients), we focused most of our survival analyses on the 162 ER+ and 98 TNBC patients in our cohort. Using the CyCIF dataset as a training dataset, we investigated the prognostic value of cell types, tumor and stromal subtypes, and cell-cell proximity quantified as the number of neighboring cells within one cell diameter. We were then able to re-use publicly available data as validation cohorts to identify robust biomarkers that predicted survival independently of clinical variables. We identified CD3 T cells as a fraction of total cells in the tissue, and the number of CD3 T cells neighboring proliferating tumor cells as predicting OS and RFS in TNBC (Figure 5.7, Figure 5.9). We compared our biomarkers to a previously published biomarker, immune-tumor mixing score, and found that our biomarkers were prognostic in multivariable models with clinicopathological variables, whereas immune-tumor mixing score was not (Figure 5.9). Our identification of lymphocyte-proliferating tumor proximity as a prognostic biomarker is consistent a hypothesis linking tumor proliferation to DNA replication stress, tumor mutation burden and productive anti-tumor immunity^{43,79,274,275}. Future work should examine our biomarker in the context of immunotherapy response, since immune infiltration and tumor proliferation are linked to immune checkpoint response in non-small cell lung cancer²⁷⁷.

We were able to reproduce the prognostic value of the mixing score in the MIBI cohort in which it was developed, illustrating its significance in that cohort as well as the necessity of collecting additional cohorts. The heterogeneity of TNBC, especially, which is composed of multiple disease subtypes¹²⁹, likely contributes to additional noise when analyzing small cohorts of 20-40 patients which made up each separate cohort. In fact, our 18 patient CyCIF cohort had very significant prognostic cell types, such as endothelial cells and endothelial-to-tumor proximity, that did not generalize to the other two cohorts. The beauty of our approach is that we did not have to assemble a huge patient cohort or generate a huge amount of data, but could utilize what was already available in the community.

In ER+ breast cancer patients, although we had 162 patients total, we did not identify any independently prognostic biomarkers. Biomarkers such as Luminal ER+ tumor, cytokeratin low tumor and proliferation had some prognostic value, and correlate with previously described luminal A versus B differences in phenotypes¹¹³. However, these variables did not remain prognostic in the multivariable model with clinicopathological variables. Since risk stratification

with molecular testing and IHC panels is already part of clinical management of ER+ breast cancer, perhaps our more compelling findings relate to correlation of tumor phenotype and composition of the microenvironment.

Intriguingly, we found that a high-risk subtype of ER+ breast cancer had high tumor heterogeneity and significantly elevated CD31+ endothelial cells (Figure 6.8). Angiogenesis can be upregulated by Wnt/ β -catenin signaling in endocrine resistant breast cancers which also exhibit EMT²⁷⁹, a potential mechanism that would explain both the heterogenous phenotype lacking differentiated luminal tumor cells and the high angiogenesis. Another interesting finding was the elevated levels of FN+ fibroblasts in tissues with more luminal ER+ tumor cells (Figure 5.8). CD146+ fibroblasts were previously shown to maintain estrogen signaling and tamoxifen resistance in ER+ breast cancer¹⁰³, a mechanism that would explain the association of fibroblasts with the good prognosis luminal ER high tumors. Although we do not have CD146 in our panel, fibronectin itself has been mechanistically linked to ER stabilization and we showed high proximity between FN and ER+ tumor cells (Figure 5.8). These two findings suggest that tumor differentiation and de-differentiation/EMT are influenced by and/or influence the tumor microenvironment in ER+ breast cancer.

6.5 FUTURE PERSPECTIVES

As imaging systems mature, I believe the field of spatial analysis of proteomic data is on the cusp of an explosion of data. As this thesis was being written, a new breast cancer IMC dataset with 693 patients was published²⁶⁹. Soon the field will have the numbers of patients necessary for statistically powered biomarker discovery. The upcoming abundance of imaging data must be accompanied by scalable tools for quality control. Automated tools to assess image quality may become necessary as there will be simply too many samples for visual inspection of each image. Tools to assist the researcher in quickly identifying and viewing images with potential quality issues would also add efficiency and important quality checkpoints to analysis workflows.

Normalization of datasets may be another issue to revisit as cohorts grow in size. In my work, I avoided the challenge of normalizing without reference tissues by simply clustering the datasets separately and annotating cell types. However, more direct comparison would involve normalizing and combining the datasets, a huge challenge considering the different signal-to-background characteristics of different image platforms, antibodies and probes.

Spatial analysis in my study showed many correlations with simple abundance, an observation that may be related to the small ROIs imaged in the early data collection efforts. In ROIs less than half a millimeter, any cells that are imaged are already pretty close together. As imaging systems develop, scan areas are increasing, which should be advantageous for developing more informative spatial biomarkers. Another potential avenue for improved spatial statistics is to leverage tissue samples from routine pathology workflows, such as H&E, for whole tissue section spatial analysis which can be linked to deeper profiling with multiplex imaging panels on targeted regions of the tissue. Tools to help predict interesting ROIs from H&E or DNA scans could create synergy between deep and wide profiling of tissues.

Another way to increase the efficiency of multiplex imaging is to become smarter about targeted biomarker panels. As the technologies were developed, every antibody that would stick was thrown at the tissue. However, redundant, poor-quality markers or those with low prognostic value should be identified and eliminated from panels to increase efficiency. While I have just confidently predicted a glut of imaging data, most of the current studies include patients treated under standard of care. We need to start evaluating clinical trial cohorts, such as patients treated with immune checkpoint inhibitors, CDK4/6 inhibitors and PARP inhibitors. There is strong evidence that these targeted therapies have effects on both the tumor and microenvironment, and multiplex imaging will be a key tool in understanding the mechanism of action of new inhibitors and discovering biomarkers of response. We need ways to optimize small, nimble and informative panels to quickly profile trial cohorts. Alternatively, we need to develop quantitative imaging tools and algorithms that can quantitatively analyze existing H&E or IHC slides. Our simple quantitative enumeration of HER2+ breast tumor cells that predicted response to neoadjuvant T-DM1 could feasibly be done from IHC. Many of the tools developed for highly multiplexed analysis could potentially be repurposed for analysis of readily available histopathology data.

Finally, my own pet project would be unravelling the mystery of interactions between the immune system and aromatase inhibitors in ER+ breast cancer. My literature review revealed that TILs are a poor prognostic factor in endocrine therapy treated patients, but not necessarily in those given neoadjuvant chemotherapy, tamoxifen or left untreated. Considering that 75% of breast tumors express ER and the patients are treated with endocrine therapy, it is urgent to understand why patients with increased TILs are doing worse and if aromatase inhibitors are worsening prognosis. Additionally, immunotherapy trials have been concentrated in TNBC, but our analyses indicate there may be biomarkers, such as HER2, which stratify for immune-benefit in ER+ breast cancer patients. Personally, I would like to pursue more characterization of immunosuppressive mechanisms in ER+ breast cancer and understand the subtype from an immunological point of view.

In conclusion, after spending four years developing methods and tools for multiplex image analysis, I have come to realize that images are all around. The next steps will be thinking about ways to extract high quality, biologically meaningful data and findings from all different types of images, and use what we have learned from these early studies to inform the development of new panels and methods to extend the capabilities of multiplex imaging in an intelligent way for maximum clinical and translational impact.

REFERENCES

1. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* **71**, 209–249 (2021).
2. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2020. *CA. Cancer J. Clin.* **70**, 7–30 (2020).
3. Loibl, S., Poortmans, P., Morrow, M., Denkert, C. & Curigliano, G. Breast cancer. *Lancet* **397**, 1750–1769 (2021).
4. Coates, A. S. *et al.* Tailoring therapies — improving the management of early breast cancer : St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann. of Oncology* **26**, 1533–1546 (2015).
5. Sørlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10869–10874 (2001).
6. Klauschen, F. *et al.* Standardized Ki67 Diagnostics Using Automated Scoring--Clinical Validation in the GeparTrio Breast Cancer Study. *Clin. cancer Res. an Off. J. Am. Assoc. Cancer Res.* **21**, 3651–3657 (2015).
7. FDA. FDA approves pembrolizumab for high-risk early-stage triple-negative breast cancer. 20–22 (2021). Available at: <https://www.fda.gov/drugs/drug-approvals-and-databases/resources-information-approved-drugs>.
8. Schmid, P. *et al.* Event-free Survival with Pembrolizumab in Early Triple-Negative Breast Cancer. *N. Engl. J. Med.* **386**, 556–567 (2022).
9. Schmid, P. *et al.* Pembrolizumab for Early Triple-Negative Breast Cancer. *N. Engl. J. Med.* **382**, 810–821 (2020).
10. Tarantino, P. *et al.* Immunotherapy for early triple negative breast cancer: research agenda for the next decade. *npj Breast Cancer* **8**, (2022).
11. FDA. FDA approves olaparib for adjuvant treatment of high-risk early breast cancer. 11–12 (2022). Available at: <https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-olaparib-adjuvant-treatment-high-risk-early-breast-cancer>.
12. Tutt, A. N. J. *et al.* Adjuvant Olaparib for Patients with BRCA1 - or BRCA2 -Mutated Breast Cancer . *N. Engl. J. Med.* **384**, 2394–2405 (2021).
13. Miles, D. *et al.* Primary results from IMpassion131 , a double-blind , placebo-controlled , randomised phase III trial of fi rst-line paclitaxel with or without atezolizumab for unresectable locally advanced / metastatic triple-negative breast cancer. *Ann. Oncol.* **32**, (2021).
14. The ASCO Post Staff. Update on U . S . Indication for Atezolizumab in PD-L1 – Positive Metastatic Triple-Negative Breast Cancer. *ASCO Post* 2–5 (2021). Available at: <https://ascopost.com/issues/september-25-2021/update-on-us-indication-for-atezolizumab-in-pd-l1-positive-metastatic-triple-negative-breast-cancer>.
15. Van Wambeke, S. & Gyawali, B. Atezolizumab inMetastatic Triple-Negative Breast Cancer— No Contradiction in the Eyes ofa Dispassionate Observer. *JAMA* **7**, 1285–1286 (2021).
16. Feng, D. *et al.* Excellent Response to Atezolizumab After Clinically Defined Hyperprogression Upon Previous Treatment With Pembrolizumab in Metastatic Triple-

- Negative Breast Cancer: A Case Report and Review of the Literature. *Front. Immunol.* **12**, (2021).
17. FDA. FDA approves abemaciclib with endocrine therapy for early breast cancer. 2–3 (2021). Available at: <https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-abemaciclib-endocrine-therapy-early-breast-cancer>.
 18. Harbeck, N. *et al.* Adjuvant abemaciclib combined with endocrine therapy for high-risk early breast cancer : updated efficacy and Ki-67 analysis from the monarchE study. *Ann. Oncol.* **32**, (2021).
 19. Gao, J. J. *et al.* CDK4/6 inhibitor treatment for patients with hormone receptor-positive, HER2-negative, advanced or metastatic breast cancer: a US Food and Drug Administration pooled analysis. *Lancet Oncol.* **21**, 250–260 (2020).
 20. FDA. FDA approves alpelisib for metastatic breast cancer. 3–4 (2019). doi:10.31525/cmr-14804bd
 21. André, F. *et al.* Alpelisib for PIK3CA -Mutated, Hormone Receptor–Positive Advanced Breast Cancer . *N. Engl. J. Med.* **380**, 1929–1940 (2019).
 22. Verret, B., Cortes, J., Bachelot, T., Andre, F. & Arnedos, M. Efficacy of PI3K inhibitors in advanced breast cancer. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **30**, x12–x20 (2019).
 23. Borcoman, E. *et al.* Inhibition of PI3K pathway increases immune infiltrate in muscle-invasive bladder cancer. *Oncoimmunology* **8**, e1581556 (2019).
 24. Zacharakis, N. *et al.* Immune recognition of somatic mutations leading to complete durable regression in metastatic breast cancer. *Nat. Med.* **24**, 724–730 (2018).
 25. Gatti-Mays, M. E. *et al.* If we build it they will come: targeting the immune response to breast cancer. *npj Breast Cancer* **5**, (2019).
 26. Dowsett, M. *et al.* Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. *J. Clin. Oncol.* **31**, 2783–2790 (2013).
 27. Melisko, M. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *Women's Oncol. Rev.* **5**, 45–47 (2005).
 28. Cuzick, J. *et al.* Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the genomic health recurrence score in early breast cancer. *J. Clin. Oncol.* **29**, 4273–4278 (2011).
 29. Bernard, P. S. *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009).
 30. Sestak, I. *et al.* Factors predicting late recurrence for estrogen receptor-positive breast cancer. *J. Natl. Cancer Inst.* **105**, 1504–1511 (2013).
 31. Davies, C. *et al.* Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. *Lancet* **381**, 805–816 (2013).
 32. Rizzo, A. & Ricci, A. D. Biomarkers for breast cancer immunotherapy: PD-L1, TILs, and beyond. *Expert Opin. Investig. Drugs* 1–7 (2021). doi:10.1080/13543784.2022.2008354
 33. Salgado, R. *et al.* The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer : recommendations by an International TILs Working Group 2014. *Ann. Oncol.* 259–271 (2015). doi:10.1093/annonc/mdu450
 34. Liu, S. *et al.* Prognostic significance of FOXP3 + tumor-infiltrating lymphocytes in breast cancer depends on estrogen receptor and human epidermal growth factor receptor-2

- expression status and concurrent cytotoxic T-cell infiltration. 1–12 (2014).
35. Ali, H. R. *et al.* Association between CD8 + T-cell infiltration and breast cancer survival in 12 439 patients. 1536–1543 (2014). doi:10.1093/annonc/mdu191
 36. Lee, A. T. J. *et al.* The adequacy of tissue microarrays in the assessment of inter- and intra-tumoural heterogeneity of infiltrating lymphocyte burden in leiomyosarcoma. *Sci. Rep.* **9**, 14602 (2019).
 37. Liudahl, S. M. *et al.* Leukocyte heterogeneity in pancreatic ductal adenocarcinoma: Phenotypic and spatial features associated with clinical outcome. *Cancer Discov.* **11**, 2014–2031 (2021).
 38. Loi, S. *et al.* Prognostic and predictive value of tumor-infiltrating lymphocytes in a phase III randomized adjuvant breast cancer trial in node-positive breast cancer comparing the addition of docetaxel to doxorubicin with doxorubicin-based chemotherapy: BIG 02-98. *J. Clin. Oncol.* **31**, 860–867 (2013).
 39. Krishnamurti, U., Wetherilt, C. S., Yang, J., Peng, L. & Li, X. Tumor-infiltrating lymphocytes are significantly associated with better overall survival and disease-free survival in triple-negative but not estrogen receptor-positive breast cancers. *Hum. Pathol.* **64**, 7–12 (2017).
 40. Denkert, C. *et al.* Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet Oncol.* **19**, 40–50 (2018).
 41. Heindl, A. *et al.* Relevance of Spatial Heterogeneity of Immune Infiltration for Predicting Risk of Recurrence after Endocrine Therapy of ER+ Breast Cancer. *J. Natl. Cancer Inst.* **110**, 166–175 (2018).
 42. Y, F. *et al.* Prognostic significance of tumor- infiltrating lymphocytes may differ depending on Ki67 expression levels in estrogen receptor-positive/HER2-negative operated breast cancers. *Breast Cancer* **26**, 738–747 (2019).
 43. Nagalla, S. *et al.* Interactions between immunity, proliferation and molecular subtype in breast cancer prognosis. *Genome Biol.* **14**, (2013).
 44. Skriver, S. K., Jensen, M. B., Knoop, A. S., Ejlersen, B. & Laenkholm, A. V. Tumour-infiltrating lymphocytes and response to neoadjuvant letrozole in patients with early oestrogen receptor-positive breast cancer: Analysis from a nationwide phase II DBCG trial. *Breast Cancer Res.* **22**, 1–8 (2020).
 45. Blok, E. J. *et al.* Exploration of tumour-infiltrating lymphocytes as a predictive biomarker for adjuvant endocrine therapy in early breast cancer. *Breast Cancer Res. Treat.* **171**, 65–74 (2018).
 46. Dunbier, A. K. *et al.* Molecular Profiling of Aromatase Inhibitor – Treated Postmenopausal Breast Tumors Identifies Immune-Related Correlates of Resistance. 2775–2787 (2013). doi:10.1158/1078-0432.CCR-12-1000
 47. Asano, Y. *et al.* Prediction of survival after neoadjuvant chemotherapy for breast cancer by evaluation of tumor-infiltrating lymphocytes and residual cancer burden. *BMC Cancer* **17**, 1–10 (2017).
 48. Leone, P. *et al.* MHC class I antigen processing and presenting machinery: Organization, function, and defects in tumor cells. *J. Natl. Cancer Inst.* **105**, 1172–1187 (2013).
 49. Sinn, B. V. *et al.* Human leucocyte antigen class I in hormone receptor-positive, HER2-negative breast cancer: Association with response and survival after neoadjuvant chemotherapy. *Breast Cancer Res.* **21**, 1–9 (2019).

50. Lee, H. J. *et al.* Differential expression of major histocompatibility complex class I in subtypes of breast cancer is associated with estrogen receptor and interferon signaling. *Oncotarget* **7**, 30119–30132 (2016).
51. Frazao, A. *et al.* CD16 + NKG2A high natural killer cells infiltrate breast cancer-draining lymph nodes. *Cancer Immunol. Res.* **7**, 208–218 (2019).
52. Goldberg, J. *et al.* The Immunology of Hormone Receptor Positive Breast Cancer. *Front. Immunol.* **12**, 1–22 (2021).
53. Liu, Y. *et al.* Expression of antigen processing and presenting molecules in brain metastasis of breast cancer. *Cancer Immunol. Immunother.* **61**, 789–801 (2012).
54. Sobral-Leite, M. *et al.* Cancer-immune interactions in ER-positive breast cancers: PI3K pathway alterations and tumor-infiltrating lymphocytes. *Breast Cancer Res.* **21**, 1–12 (2019).
55. Tokunaga, E. *et al.* Activation of PI3K/Akt signaling and hormone resistance in breast cancer. *Breast cancer* **13**, 137–144 (2006).
56. Guerrero-Zotano, A., Mayer, I. A. & Arteaga, C. L. PI3K/AKT/mTOR: role in breast cancer progression, drug resistance, and treatment. *Cancer Metastasis Rev.* **35**, 515–524 (2016).
57. Crane, C. A. *et al.* PI(3) kinase is associated with a mechanism of immunoresistance in breast and prostate cancer. *Oncogene* **28**, 306–312 (2009).
58. Loibl, S. *et al.* PIK3CA mutations are associated with reduced pathological complete response rates in primary HER2-positive breast cancer: pooled analysis of 967 patients from five prospective trials investigating lapatinib and trastuzumab. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **27**, 1519–1525 (2016).
59. Lehmann, B. D. *et al.* PIK3CA mutations in androgen receptor-positive triple negative breast cancer confer sensitivity to the combination of PI3K and androgen receptor inhibitors. 1–14 (2014).
60. Loibl, S. & Gianni, L. HER2-positive breast cancer. *Lancet* **389**, 2415–2429 (2017).
61. Denkert, C. *et al.* Tumor-infiltrating lymphocytes and response to neoadjuvant chemotherapy with or without carboplatin in human epidermal growth factor receptor 2-positive and triple-negative primary breast cancers. *J. Clin. Oncol.* **33**, 983–991 (2015).
62. Loi, S. *et al.* Tumor infiltrating lymphocytes are prognostic in triple negative breast cancer and predictive for trastuzumab benefit in early breast cancer: Results from the FinHER trial. *Ann. Oncol.* **25**, 1544–1550 (2014).
63. Heppner, B. I. *et al.* Tumor-Infiltrating Lymphocytes : A Predictive and Prognostic Biomarker in Neoadjuvant-Treated HER2-Positive Breast Cancer. 5747–5755 (2016). doi:10.1158/1078-0432.CCR-15-2338
64. Salgado, R. *et al.* Tumor-infiltrating lymphocytes and associations with pathological complete response and event-free survival in HER2-positive early-stage breast cancer treated with lapatinib and trastuzumab: A secondary analysis of the NeoALTTO trial. *JAMA Oncol.* **1**, 448–455 (2015).
65. Solinas, C. *et al.* Tumor-infiltrating lymphocytes in patients with HER2-positive breast cancer treated with neoadjuvant chemotherapy plus trastuzumab , lapatinib or their combination : A meta-analysis of randomized controlled trials. *Cancer Treat. Rev.* **57**, 8–15 (2017).
66. Luen, S. J. *et al.* Tumour-infiltrating lymphocytes in advanced HER2-positive breast cancer treated with pertuzumab or placebo in addition to trastuzumab and docetaxel : a

- retrospective analysis of the CLEOPATRA study. *Lancet Oncol.* **18**, 52–62 (2017).
67. Denkert, C., Liedtke, C., Tutt, A. & von Minckwitz, G. Molecular alterations in triple-negative breast cancer—the road to new treatment strategies. *Lancet* **389**, 2430–2442 (2017).
 68. Denkert, C. *et al.* Tumor-associated lymphocytes as an independent predictor of response to neoadjuvant chemotherapy in breast cancer. *J. Clin. Oncol.* **28**, 105–113 (2010).
 69. Gruosso, T. *et al.* Spatially distinct tumor immune microenvironments stratify triple-negative breast cancers. *J. Clin. Invest.* **129**, 1785–1800 (2019).
 70. Li, X. *et al.* Infiltration of CD8 + T cells into tumor cell clusters in triple-negative breast cancer. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 3678–3687 (2019).
 71. Ali, H. R., Chlon, L., Pharoah, P. D. P., Markowitz, F. & Caldas, C. Patterns of Immune Infiltration in Breast Cancer and Their Clinical Implications: A Gene-Expression-Based Retrospective Study. *PLoS Med.* **13**, 1–24 (2016).
 72. Gu-trantien, C. *et al.* CD4 + follicular helper T cell infiltration predicts breast cancer survival. *J. Clin. Invest.* **123**, 2873–2892 (2013).
 73. Bense, R. D. *et al.* Relevance of tumor-infiltrating immune cell composition and functionality for disease outcome in breast cancer. *J. Natl. Cancer Inst.* **109**, 1–9 (2017).
 74. Ladoire, S. *et al.* In situ immune response after neoadjuvant chemotherapy for breast cancer predicts survival. *J. Pathol.* **224**, 389–400 (2011).
 75. Ishigami, E. *et al.* Coexistence of regulatory B cells and regulatory T cells in tumor-infiltrating lymphocyte aggregates is a prognostic factor in patients with breast cancer. *Breast Cancer* **26**, 180–189 (2019).
 76. Zhao, X. *et al.* Prognostic significance of tumor-associated macrophages in breast cancer : a meta-analysis of the literature. **8**, 30576–30586 (2017).
 77. Richardsen, E., Uglehus, R. D., Johnsen, S. H. & Busund, L. Macrophage-Colony Stimulating Factor (CSF1) Predicts Breast Cancer Progression and Mortality. **874**, 865–874 (2015).
 78. Thorsson, V. *et al.* The Immune Landscape of Cancer. *Immunity* **48**, 812-830.e14 (2018).
 79. Miller, L. D. *et al.* Immunogenic Subtypes of Breast Cancer Delineated by Gene Classifiers of Immune Responsiveness. *Cancer Immunol. Res.* **4**, 600–611 (2016).
 80. Kalluri, R. The biology and function of fibroblasts in cancer. *Nat. Rev. Cancer* **16**, 582–598 (2016).
 81. Karnoub, A. E. *et al.* Mesenchymal stem cells within tumour stroma promote breast cancer metastasis. *Nature* **449**, 557–563 (2007).
 82. Erez, N., Truitt, M., Olson, P. & Hanahan, D. Cancer-Associated Fibroblasts Are Activated in Incipient Neoplasia to Orchestrate Tumor-Promoting Inflammation in an NF- κ B-Dependent Manner. *Cancer Cell* **17**, 135–147 (2010).
 83. Ershaid, N. *et al.* NLRP3 inflammasome in fibroblasts links tissue damage with inflammation in breast cancer progression and metastasis. *Nat. Commun.* **10**, (2019).
 84. Orimo, A. *et al.* Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. *Cell* **121**, 335–348 (2005).
 85. Calvo, F. *et al.* Mechanotransduction and YAP-dependent matrix remodelling is required for the generation and maintenance of cancer-associated fibroblasts. *Nat. Cell Biol.* **15**, 637–646 (2013).
 86. Sternlicht, M. D. *et al.* The stromal proteinase MMP3/stromelysin-1 promotes mammary

- carcinogenesis. *Cell* **98**, 137–146 (1999).
87. Boire, A. *et al.* PAR1 is a matrix metalloprotease-1 receptor that promotes invasion and tumorigenesis of breast cancer cells. *Cell* **120**, 303–313 (2005).
 88. Malanchi, I. *et al.* Interactions between cancer stem cells and their niche govern metastatic colonization. *Nature* **481**, 85–91 (2012).
 89. Martinez-Outschoorn, U. E., Lisanti, M. P. & Sotgia, F. Catabolic cancer-associated fibroblasts transfer energy and biomass to anabolic cancer cells, fueling tumor growth. *Semin. Cancer Biol.* **25**, 47–60 (2014).
 90. Zhang, W. *et al.* GFPT2-expressing cancer-associated fibroblasts mediate metabolic reprogramming in human lung adenocarcinoma. *Cancer Res.* **78**, 3445–3457 (2018).
 91. Kraman, M. *et al.* Suppression of Antitumor Immunity by Stromal Cells Expressing Fibroblast Activation Protein- α . *Science* (80-.). **330**, 827–830 (2010).
 92. Özdemir, B. C. *et al.* Depletion of carcinoma-associated fibroblasts and fibrosis induces immunosuppression and accelerates pancreas cancer with reduced survival. *Cancer Cell* **25**, 719–734 (2014).
 93. Hanley, C. J. & Thomas, G. J. Targeting cancer associated fibroblasts to enhance immunotherapy: emerging strategies and future perspectives. *Oncotarget* **12**, 1427–1433 (2021).
 94. Peranzoni, E., Rivas, A., Bougherara, H., Salmon, H. & Donnadieu, E. Positive and negative influence of the matrix architecture on antitumor immune surveillance. *Cell. Mol. Life Sci.* **70**, 4431–4448 (2013).
 95. Salmon, H., Mami-chouaib, F., Salmon, H., Franciszkiewicz, K. & Damotte, D. Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *J. Clin. Invest.* **133**, 899–910 (2012).
 96. Farmer, P. *et al.* A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. *Nat. Med.* **15**, 68–74 (2009).
 97. Park, C. C., Zhang, H. J., Yao, E. S., Park, C. J. & Bissell, M. J. B1 Integrin Inhibition Dramatically Enhances Radiotherapy Efficacy in Human Breast Cancer Xenografts. *Cancer Res.* **68**, 4398–4405 (2008).
 98. Mitra, A., Mishra, L. & Li, S. EMT, CTCs and CSCs in tumor relapse and drug-resistance. *Oncotarget* **6**, 10697–10711 (2015).
 99. Dennison, J. B. *et al.* High intratumoral stromal content defines reactive breast cancer as a low-risk breast cancer subtype. *Clin. Cancer Res.* **22**, 5068–5078 (2016).
 100. Millar, E. K. A. *et al.* Tumour stroma ratio assessment using digital image analysis predicts survival in triple negative and luminal breast cancer. *Cancers (Basel)*. **12**, 1–14 (2020).
 101. Downey, C. L. *et al.* The prognostic significance of tumour-stroma ratio in oestrogen receptor-positive breast cancer. *Br. J. Cancer* **110**, 1744–1747 (2014).
 102. Dekker, T. J. A. *et al.* Prognostic significance of the tumor-stroma ratio: Validation study in node-negative premenopausal breast cancer patients from the EORTC perioperative chemotherapy (POP) trial (10854). *Breast Cancer Res. Treat.* **139**, 371–379 (2013).
 103. Brechbuhl, H. M. *et al.* Fibroblast subtypes regulate responsiveness of luminal breast cancer to estrogen. *Clin. Cancer Res.* **23**, 1710–1721 (2017).
 104. Costa, A. *et al.* Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer Cell* **33**, 463–479.e10 (2018).
 105. Tsujikawa, T. *et al.* Quantitative Multiplex Immunohistochemistry Reveals Myeloid-

- Inflamed Tumor-Immune Complexity Associated with Poor Prognosis. *Cell Rep.* **19**, 203–217 (2017).
106. Gut, G., Herrmann, M. D. & Pelkmans, L. Multiplexed protein maps link subcellular organization to cellular states. *Science* (80-.). **361**, (2018).
 107. Lin, J. R., Fallahi-Sichani, M. & Sorger, P. K. Highly multiplexed imaging of single cells using a high-throughput cyclic immunofluorescence method. *Nat. Commun.* **6**, 1–7 (2015).
 108. Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968-981.e15 (2018).
 109. Giesen, C. *et al.* Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11**, 417–422 (2014).
 110. Eng, J. *et al.* Cyclic Multiplexed-Immunofluorescence (cmIF), a Highly Multiplexed Method for Single-Cell Analysis. in *Biomarkers for Immunotherapy of Cancer. Methods in Molecular Biology* (eds. Thurin, M., Cesano, A. & Marincola, F. M.) **2055**, 521–562 (Humana, 2020).
 111. Keren, L. *et al.* A Structure Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam imaging. *Cell* **174**, 1373-1387.e19 (2018).
 112. Jackson, H. W. *et al.* The single-cell pathology landscape of breast cancer. *Nature* (2019). doi:10.1038/s41586-019-1876-x
 113. Ali, H. R. *et al.* Imaging mass cytometry and multiplatform genomics define the phenogenomic landscape of breast cancer. *Nat. Cancer* **1**, 163–175 (2020).
 114. Perez, E. A. *et al.* Association of stromal tumor-infiltrating lymphocytes with recurrence-free survival in the n9831 adjuvant trial in patients with early-stage HER2-positive breast cancer. *JAMA Oncol.* **2**, 56–64 (2016).
 115. Luen, S. J., Savas, P., Fox, S. B., Salgado, R. & Loi, S. Tumour-infiltrating lymphocytes and the emerging role of immunotherapy in breast cancer. *Pathology* **49**, 141–155 (2017).
 116. Gerdes, M. J. *et al.* Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11982–11987 (2013).
 117. McKinley, E. T. *et al.* Optimized multiplex immunofluorescence single-cell analysis reveals tuft cell heterogeneity. *JCI Insight* **2**, (2017).
 118. Lin, J.-R. *et al.* A simple open-source method for highly multiplexed imaging of single cells in tissues and tumours. *Elife* **7**, 151738 (2018).
 119. Chae, Y. K. *et al.* Current landscape and future of dual anti-CTLA4 and PD-1/PD-L1 blockade immunotherapy in cancer; lessons learned from clinical trials with melanoma and non-small cell lung cancer (NSCLC). *J. Immunother. Cancer* **6**, (2018).
 120. Cortazar, P. *et al.* Pathological complete response and long-term clinical benefit in breast cancer: The CTNeoBC pooled analysis. *Lancet* **384**, 164–172 (2014).
 121. Savas, P., Salgado, R., Denkert, C., Sotiriou, C. & Darcy, P. K. Clinical relevance of host immunity in breast cancer : from TILs to the clinic. *Nat. Publ. Gr.* **13**, 228–241 (2015).
 122. Adams, S. *et al.* Pembrolizumab monotherapy for previously treated metastatic triple-negative breast cancer: cohort A of the phase II KEYNOTE-086 study. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **30**, 397–404 (2019).
 123. Kelly, K. *et al.* Avelumab (MSB0010718C), an anti-PD-L1 antibody, in patients with metastatic or locally advanced solid tumors: assessment of safety and tolerability in a phase I, open-label expansion study. *J. Clin. Oncol.* **33**, 3044 (2015).
 124. Nanda, R. *et al.* Effect of Pembrolizumab Plus Neoadjuvant Chemotherapy on Pathologic

- Complete Response in Women with Early-Stage Breast Cancer: An Analysis of the Ongoing Phase 2 Adaptively Randomized I-SPY2 Trial. *JAMA Oncol.* **6**, 676–684 (2020).
125. Schmid, P. *et al.* Atezolizumab and Nab-Paclitaxel in Advanced Triple-Negative Breast Cancer. *N. Engl. J. Med.* **379**, 2108–2121 (2018).
 126. Rugo, H. S. *et al.* Safety and antitumor activity of pembrolizumab in patients with estrogen receptor–positive/human epidermal growth factor receptor 2–negative advanced breast cancer. *Clin. Cancer Res.* **24**, 2804–2811 (2018).
 127. Khoury, T., Nagrale, V., Opyrchal, M., Peng, X. & Yao, S. Prognostic Significance of Stromal Versus Intratumoral Infiltrating Lymphocytes in Different Subtypes of Breast Cancer Treated with Cytotoxic Neoadjuvant Chemotherapy. *Appl Immunohistochem Mol Morphol* (2018).
 128. Chen, X. *et al.* TNBCtype: A Subtyping Tool for Triple-Negative Breast Cancer. *Cancer informatics* **11**, 147–156 (2012).
 129. Lehmann, B. D. *et al.* Refinement of triple-negative breast cancer molecular subtypes: Implications for neoadjuvant chemotherapy selection. *PLoS One* **11**, (2016).
 130. Teschendorff, A. E., Miremadi, A., Pinder, S. E., Ellis, I. O. & Caldas, C. An immune response gene expression module identifies a good prognosis subtype in estrogen receptor negative breast cancer. *Genome Biol.* **8**, (2007).
 131. Burstein, M. D. *et al.* Comprehensive Genomic Analysis Identifies Novel Subtypes and Targets of Triple-negative Breast Cancer Analysis and interpretation of data: HHS Public Access. *Clin Cancer Res* **21**, 1688–1698 (2015).
 132. Brahmer, J. R. *et al.* Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **28**, 3167–3175 (2010).
 133. Kwon, E. D. *et al.* Ipilimumab versus placebo after radiotherapy in patients with metastatic castration-resistant prostate cancer that had progressed after docetaxel chemotherapy (CA184-043): a multicentre, randomised, double-blind, phase 3 trial. *Lancet. Oncol.* **15**, 700–712 (2014).
 134. Madan, R. A., Heery, C. R. & Gulley, J. L. Combination of vaccine and immune checkpoint inhibitor is safe with encouraging clinical activity. *Oncoimmunology* **1**, 1167–1168 (2012).
 135. Couzin-Frankel, J. Breakthrough of the year 2013. Cancer immunotherapy. *Science (New York, N.Y.)* **342**, 1432–1433 (2013).
 136. Patnaik, A. *et al.* Phase I Study of Pembrolizumab (MK-3475; Anti-PD-1 Monoclonal Antibody) in Patients with Advanced Solid Tumors. *Clin. cancer Res. an Off. J. Am. Assoc. Cancer Res.* **21**, 4286–4293 (2015).
 137. Royal, R. E. *et al.* Phase 2 trial of single agent Ipilimumab (anti-CTLA-4) for locally advanced or metastatic pancreatic adenocarcinoma. *J. Immunother.* **33**, 828–833 (2010).
 138. Shen, W. *et al.* TGF- β in pancreatic cancer initiation and progression: two sides of the same coin. *Cell Biosci.* **7**, 39 (2017).
 139. Haque, S. & Morris, J. C. Transforming growth factor- β : A therapeutic target for cancer. *Hum. Vaccin. Immunother.* **13**, 1741–1750 (2017).
 140. Weinberg, D. H. and R. A. The Hallmarks of Cancer. *Cell* **107**, 251–262 (2000).
 141. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: The next generation. *Cell* **144**, 646–674 (2011).
 142. Ellsworth, R. E., Blackburn, H. L., Shriver, C. D., Soon-Shiong, P. & Ellsworth, D. L.

- Molecular heterogeneity in breast cancer: State of the science and implications for patient care. *Semin. Cell Dev. Biol.* **64**, 65–72 (2017).
143. Chin, K. *et al.* In situ analyses of genome instability in breast cancer. *Nat. Genet.* **36**, 984–988 (2004).
 144. Ferguson, L. R. *et al.* Genomic instability in human cancer: Molecular insights and opportunities for therapeutic attack and prevention through diet and nutrition. *Semin. Cancer Biol.* **35 Suppl**, S5–S24 (2015).
 145. Risom, T. *et al.* Differentiation-state plasticity is a targetable resistance mechanism in basal-like breast cancer. *Nat. Commun.* **9**, 3815 (2018).
 146. Creighton, C. J. *et al.* Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13820–13825 (2009).
 147. M., H. L. *et al.* Subtype and pathway specific responses to anticancer compounds in breast cancer. *Proc. Natl. Acad. Sci.* **109**, 2724–2729 (2012).
 148. Lesniak, D. *et al.* Spontaneous epithelial-mesenchymal transition and resistance to HER-2-targeted therapies in HER-2-positive luminal breast cancer. *PLoS One* **8**, e71987 (2013).
 149. Gupta, P. B. *et al.* Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* **146**, 633–644 (2011).
 150. Chaffer, C. L. *et al.* Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 7950–7955 (2011).
 151. Goldman, A. *et al.* Temporally sequenced anticancer drugs overcome adaptive resistance by targeting a vulnerable chemotherapy-induced phenotypic transition. *Nat. Commun.* **6**, 6139 (2015).
 152. Sharma, S. V. *et al.* A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* **141**, 69–80 (2010).
 153. Fidler, I. J. & Kripke, M. L. Metastasis results from preexisting variant cells within a malignant tumor. *Science* **197**, 893–895 (1977).
 154. Fidler, I. J. Tumor heterogeneity and the biology of cancer invasion and metastasis. *Cancer Res.* **38**, 2651–2660 (1978).
 155. Miller, F. R., Miller, B. E. & Heppner, G. H. Characterization of metastatic heterogeneity among subpopulations of a single mouse mammary tumor: heterogeneity in phenotypic stability. *Invasion Metastasis* **3**, 22–31 (1983).
 156. Weber, C. E. & Kuo, P. C. The tumor microenvironment. *Surg. Oncol.* **21**, 172–177 (2012).
 157. Robert, C. *et al.* Pembrolizumab versus Ipilimumab in Advanced Melanoma. *N. Engl. J. Med.* **372**, 2521–2532 (2015).
 158. Brahmer, J. *et al.* Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. *N. Engl. J. Med.* **373**, 123–135 (2015).
 159. Ferris, R. L. *et al.* Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. *N. Engl. J. Med.* **375**, 1856–1867 (2016).
 160. Motzer, R. J. *et al.* Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. *N. Engl. J. Med.* **373**, 1803–1813 (2015).
 161. Ansell, S. M. *et al.* PD-1 Blockade with Nivolumab in Relapsed or Refractory Hodgkin's Lymphoma. *N. Engl. J. Med.* **372**, 311–319 (2015).
 162. Sarvaria, A., Madrigal, J. A. & Saudemont, A. B cell regulation in cancer and anti-tumor

- immunity. *Cell. Mol. Immunol.* **14**, 662–674 (2017).
163. Kumar, V., Patel, S., Tcyganov, E. & Gabrilovich, D. I. The Nature of Myeloid-Derived Suppressor Cells in the Tumor Microenvironment. *Trends Immunol.* **37**, 208–220 (2016).
 164. Marcus, A. *et al.* Recognition of tumors by the innate immune system and natural killer cells. *Adv. Immunol.* **122**, 91–128 (2014).
 165. Wimberly, H. *et al.* PD-L1 expression correlates with tumor-infiltrating lymphocytes and response to neoadjuvant chemotherapy in breast cancer. *Cancer Immunol Res* **3**, 326–332 (2015).
 166. Hurwitz, H. *et al.* Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N. Engl. J. Med.* **350**, 2335–2342 (2004).
 167. Sandler, A. *et al.* Paclitaxel-carboplatin alone or with bevacizumab for non-small-cell lung cancer. *N. Engl. J. Med.* **355**, 2542–2550 (2006).
 168. Motzer, R. J. *et al.* Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *N. Engl. J. Med.* **356**, 115–124 (2007).
 169. Brem, S. S., Gullino, P. M. & Medina, D. Angiogenesis: a marker for neoplastic transformation of mammary papillary hyperplasia. *Science* **195**, 880–882 (1977).
 170. Jensen, H. M., Chen, I., DeVault, M. R. & Lewis, A. E. Angiogenesis induced by ‘normal’ human breast tissue: a probable marker for precancer. *Science* **218**, 293–295 (1982).
 171. McLeskey, S. W. *et al.* Tumor growth of FGF or VEGF transfected MCF-7 breast carcinoma cells correlates with density of specific microvessels independent of the transfected angiogenic factor. *Am. J. Pathol.* **153**, 1993–2006 (1998).
 172. Weinstat-Saslow, D. L. *et al.* Transfection of thrombospondin 1 complementary DNA into a human breast carcinoma cell line reduces primary tumor growth, metastatic potential, and angiogenesis. *Cancer Res.* **54**, 6504–6511 (1994).
 173. Mackey, J. R. *et al.* Controlling angiogenesis in breast cancer: A systematic review of anti-angiogenic trials. *Cancer Treat. Rev.* **38**, 673–688 (2012).
 174. Kather, J. N. *et al.* Continuous representation of tumor microvessel density and detection of angiogenic hotspots in histological whole-slide images. *Oncotarget* **6**, 19163–19176 (2015).
 175. Weidner, N. Chapter 14. Measuring intratumoral microvessel density. *Methods Enzymol.* **444**, 305–323 (2008).
 176. Tolaney, S. M. *et al.* Role of vascular density and normalization in response to neoadjuvant bevacizumab and chemotherapy in breast cancer patients. *Proc. Natl. Acad. Sci.* **112**, 14325–14330 (2015).
 177. Zhang, S. *et al.* Intratumoral and peritumoral lymphatic vessel density both correlate with lymph node metastasis in breast cancer. *Sci. Rep.* **7**, 1–7 (2017).
 178. Kalluri, R. & Zeisberg, M. Fibroblasts in cancer. *Nat. Rev. Cancer* **6**, 392–401 (2006).
 179. Ostman, A. & Augsten, M. Cancer-associated fibroblasts and tumor growth--bystanders turning into key players. *Curr. Opin. Genet. Dev.* **19**, 67–73 (2009).
 180. Langley, R. R. & Fidler, I. J. The seed and soil hypothesis revisited--the role of tumor-stroma interactions in metastasis to different organs. *Int. J. cancer* **128**, 2527–2535 (2011).
 181. Martinez-Outschoorn, U. E. *et al.* Anti-estrogen resistance in breast cancer is induced by the tumor microenvironment and can be overcome by inhibiting mitochondrial function in epithelial cancer cells. *Cancer Biol. Ther.* **12**, 924–938 (2011).
 182. Loeffler, M., Krüger, J. A., Niethammer, A. G. & Reisfeld, R. A. Targeting tumor-

- associated fibroblasts improves cancer chemotherapy by increasing intratumoral drug uptake. *J. Clin. Invest.* **116**, 1955–1962 (2006).
183. Sun, Y. *et al.* Treatment-induced damage to the tumor microenvironment promotes prostate cancer therapy resistance through WNT16B. *Nat. Med.* **18**, 1359–1368 (2012).
 184. Mao, Y., Keller, E. T., Garfield, D. H., Shen, K. & Wang, J. Stromal cells in tumor microenvironment and breast cancer. 303–315 (2013). doi:10.1007/s10555-012-9415-3
 185. Zeisberg, E. M., Potenta, S., Xie, L., Zeisberg, M. & Kalluri, R. Discovery of endothelial to mesenchymal transition as a source for carcinoma-associated fibroblasts. *Cancer Res.* **67**, 10123–10128 (2007).
 186. Weidner, N. *et al.* Tumor angiogenesis: a new significant and independent prognostic indicator in early-stage breast carcinoma. *J. Natl. Cancer Inst.* **84**, 1875–1887 (1992).
 187. Gasparini, G. *et al.* Tumor microvessel density, p53 expression, tumor size, and peritumoral lymphatic vessel invasion are relevant prognostic markers in node-negative breast carcinoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **12**, 454–466 (1994).
 188. Murri, A. M. Al *et al.* The relationship between the systemic inflammatory response, tumour proliferative activity, T-lymphocytic and macrophage infiltration, microvessel density and survival in patients with primary operable breast cancer. *Br. J. Cancer* **99**, 1013–1019 (2008).
 189. Motz, G. T. & Coukos, G. The parallel lives of angiogenesis and immunosuppression : cancer and other tales. *Nat. Publ. Gr.* **11**, 702–711 (2011).
 190. Rekhman, N. & Bishop, J. A. *Quick Reference Handbook for Surgical Pathologists. Quick Reference Handbook for Surgical Pathologists* (2011). doi:10.1007/978-3-642-20086-1
 191. Lin, F. & Prichard, J. *Handbook of Practical Immunohistochemistry: Frequently Asked Questions.* (Springer, 2011).
 192. Setiadi, A. F. *et al.* Quantitative, architectural analysis of immune cell subsets in tumor-draining lymph nodes from breast cancer patients and healthy lymph nodes. *PLoS One* **5**, e12420 (2010).
 193. Levenson, R. M. & Mansfield, J. R. Multispectral imaging in biology and medicine: slices of life. *Cytometry. A* **69**, 748–758 (2006).
 194. Gorris, M. A. J. *et al.* Eight-Color Multiplex Immunohistochemistry for Simultaneous Detection of Multiple Immune Checkpoint Molecules within the Tumor Microenvironment. *J. Immunol.* **200**, 347–354 (2018).
 195. Bay, H., Tuytelaars, T. & Van Gool, L. SURF: Speeded up robust features. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **3951 LNCS**, 404–417 (2006).
 196. Chang, Y. H. *et al.* Deep learning based Nucleus Classification in pancreas histological images. *2017 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* (2017).
 197. Deubler, J. & Olivo, J.-C. A Wavelet-Based Multiresolution Method to Automatically Register Images. *J. Math. Imaging Vis.* **7**, 199–209 (1997).
 198. Yoo, I., Hildebrand, D. G. C., Tobin, W. F., Lee, W.-C. A. & Jeong, W.-K. ssEMnet: Serial-Section Electron Microscopy Image Registration Using a Spatial Transformer Network with Learned Features. in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (eds. Cardoso, M. J. *et al.*) 249–257 (Springer International Publishing, 2017).
 199. Carpenter, A. E. *et al.* CellProfiler: image analysis software for identifying and

- quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006).
200. Arena, E. T. *et al.* Quantitating the cell: turning images into numbers with ImageJ. *Wiley Interdiscip. Rev. Dev. Biol.* **6**, (2017).
 201. Wang, A., Yan, X. & Wei, Z. ImagePy: an open-source, Python-based and platform-independent software package for bioimage analysis. *Bioinformatics* **34**, 3238–3240 (2018).
 202. Tsujikawa, T. *et al.* Robust Cell Detection and Segmentation for Image Cytometry Reveal Th17 Cell Heterogeneity. *Cytometry. A* **95**, 389–398 (2019).
 203. Chang, Y. H. *et al.* Quantitative analysis of histological tissue image based on cytological profiles and spatial statistics. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.* **2016**, 1175–1178 (2016).
 204. Lane, R. S. *et al.* IFN γ -activated dermal lymphatic vessels inhibit cytotoxic T cells in melanoma and inflamed skin. *J. Exp. Med.* **215**, 3057–3074 (2018).
 205. Gunderson, A. J. *et al.* Bruton Tyrosine Kinase-Dependent Immune Cell Cross-talk Drives Pancreas Cancer. *Cancer Discov.* **6**, 270–285 (2016).
 206. Pennock, N. D. *et al.* Ibuprofen supports macrophage differentiation, T cell recruitment, and tumor suppression in a model of postpartum breast cancer. *J. Immunother. Cancer* **6**, 1–23 (2018).
 207. Gopalakrishnan, V. *et al.* Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* **359**, 97–103 (2018).
 208. Johnson, B. E. *et al.* An omic and multidimensional spatial atlas from serial biopsies of an evolving metastatic breast cancer. *Cell Reports Med.* **3**, 100525 (2022).
 209. Carstens, J. L. *et al.* Spatial computation of intratumoral T cells correlates with survival of patients with pancreatic cancer. *Nat. Commun.* **8**, 1–13 (2017).
 210. Viratham Pulsawatdi, A. *et al.* A robust multiplex immunofluorescence and digital pathology workflow for the characterisation of the tumour immune microenvironment. *Mol. Oncol.* **14**, 2384–2402 (2020).
 211. Zhang, W. *et al.* Fully automated 5-plex fluorescent immunohistochemistry with tyramide signal amplification and same species antibodies. *Lab. Investig.* **97**, 873–885 (2017).
 212. Banik, G. *et al.* High-dimensional multiplexed immunohistochemical characterization of immune contexture in human cancers. *Methods in Enzymology* **635**, (Elsevier Inc., 2020).
 213. Bolognesi, M. M. *et al.* Multiplex Staining by Sequential Immunostaining and Antibody Removal on Routine Tissue Sections. *J. Histochem. Cytochem.* **65**, 431–444 (2017).
 214. Cattoretto, G., Cattoretto, G., Bosisio, F. M., Marcelis, L. & Bolognesi, M. M. Multiple Interactive Labeling by Antibody Neodeposition (MILAN). *Protoc. Exch.* 1–12 (2018). doi:10.1038/protex.2018.106
 215. Gerdes, M. J., Sevinsky, C. J., Sood, A., Adak, S. & Bello, M. O. Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 11982–11987 (2013).
 216. Radtke, A. J. *et al.* IBEX: A versatile multiplex optical imaging approach for deep phenotyping and spatial analysis of cells in complex tissues. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 33455–33465 (2020).
 217. Schubert, W. *et al.* Analyzing proteome topology and function by automated multidimensional fluorescence microscopy. *Nat. Biotechnol.* **24**, 1270–1278 (2006).
 218. Pascual-Reguant, A. *et al.* Multiplexed histology analyses for the phenotypic and spatial characterization of human innate lymphoid cells. *Nat. Commun.* **12**, 1–15 (2021).

219. Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968-981.e15 (2018).
220. Manesse, M., Katir K. Patel, Bobrow, M. & Downing, S. R. The InSituPlex® Staining Method for Multiplexed Immunofluorescence Cell Phenotyping and Spatial Profiling of Tumor FFPE Samples. in *Biomarkers for Immunotherapy of Cancer. Methods in Molecular Biology* (eds. Thurin, M., Cesano, A. & Marincola, F. M.) **2055**, 625–626 (Humana, 2020).
221. Saka, S. K. *et al.* Immuno-SABER enables highly multiplexed and amplified protein imaging in tissues. *Nat. Biotechnol.* **37**, 1080–1090 (2019).
222. McMahon, N. P. *et al.* Oligonucleotide conjugated antibodies permit highly multiplexed immunofluorescence for future use in clinical histopathology. *J. Biomed. Opt.* **25**, 1 (2020).
223. Stewart, R. L., Matynia, A. P., Factor, R. E. & Varley, K. E. Spatially-resolved quantification of proteins in triple negative breast cancers reveals differences in the immune microenvironment associated with prognosis. *Sci. Rep.* **10**, 1–8 (2020).
224. McGinnis, L. M., Ibarra-Lopez, V., Rost, S. & Ziai, J. Clinical and research applications of multiplexed immunohistochemistry and in situ hybridization. *J. Pathol.* **254**, 405–417 (2021).
225. Du, Z. *et al.* Qualifying antibodies for image-based immune profiling and multiplexed tissue imaging. *Nat. Protoc.* **14**, 2900–2930 (2019).
226. Gambella, A. *et al.* Section detachment in immunohistochemistry: causes, troubleshooting, and problem-solving. *Histochem. Cell Biol.* **148**, 95–101 (2017).
227. Bass, B. P., Engel, K. B., Greytak, S. R. & Moore, H. M. A review of preanalytical factors affecting molecular, protein, and morphological analysis of Formalin-Fixed, Paraffin-Embedded (FFPE) tissue: How well do you know your FFPE specimen? *Arch. Pathol. Lab. Med.* **138**, 1520–1530 (2014).
228. Büttner, M., Miao, Z., Wolf, F. A., Teichmann, S. A. & Theis, F. J. A test metric for assessing single-cell RNA-seq batch correction. *Nat. Methods* **16**, 43–49 (2019).
229. Chang, Y. H. *et al.* RESTORE: Robust intEnSiTy nORmalization mEthod for multiplexed imaging. *Commun. Biol.* **3**, 1–9 (2020).
230. Johnson, W. E., Li, C. & Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **8**, 118–127 (2007).
231. Wolf, A., Angerer, P. & Theis, F. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, (2018).
232. Neve, R. M. *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**, 515–527 (2006).
233. Balassanian, R. *et al.* A superior method for cell block preparation for fine-needle aspiration biopsies. *Cancer Cytopathol.* **124**, 508–518 (2016).
234. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. 1–19 (2020). doi:10.1101/2020.02.02.931238
235. Li, C. H. & Tam, P. K. S. An iterative algorithm for minimum cross entropy thresholding. *Pattern Recognit. Lett.* **19**, 771–776 (1998).
236. Waters, J. C. Accuracy and precision in quantitative fluorescence microscopy. *J. Cell Biol.* **185**, 1135–1148 (2009).
237. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python Fabian. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

238. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
239. Janiszewska, M. *et al.* The impact of tumor epithelial and microenvironmental heterogeneity on treatment responses in HER2+ breast cancer. *JCI Insight* **6**, (2021).
240. Marti, J. L. G., Hyder, T., Nasrazadani, A. & Brufsky, A. M. The Evolving Landscape of HER2-Directed Breast Cancer Therapy. *Curr. Treat. Options Oncol.* **21**, (2020).
241. von Minckwitz, G. *et al.* Trastuzumab Emtansine for Residual Invasive HER2-Positive Breast Cancer. *N. Engl. J. Med.* **380**, 617–628 (2019).
242. Janiszewska, M. *et al.* The impact of tumor epithelial and microenvironmental heterogeneity on treatment responses in HER2+ breast cancer. *JCI Insight* **6**, (2021).
243. Tannock, I. & Hickman, J. Limits to Personalized Cancer Medicine. *N. Engl. J. Med.* **375**, 1289–1294 (2016).
244. Sánchez, N. S., Mills, G. B. & Shaw, K. R. M. Precision oncology: Neither a silver bullet nor a dream. *Pharmacogenomics* **18**, 1525–1539 (2017).
245. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 1–12 (2019).
246. Eng, J. *et al.* Cyclic Multiplexed-Immunofluorescence (cmIF), a Highly Multiplexed Method for Single-Cell Analysis. *Methods in Molecular Biology* **2055**, (2020).
247. Hoadley, K. A. *et al.* EGFR associated expression profiles vary with breast tumor subtype. *BMC Genomics* **8**, 1–19 (2007).
248. Jansen, M. P. H. M. *et al.* Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **23**, 732–740 (2005).
249. Pontiggia, O. *et al.* The tumor microenvironment modulates tamoxifen resistance in breast cancer: a role for soluble stromal factors and fibronectin through $\beta 1$ integrin. *Breast Cancer Res. Treat.* **133**, 459–471 (2012).
250. Sampayo, R. G. *et al.* Fibronectin rescues estrogen receptor α from lysosomal degradation in breast cancer cells. *J. Cell Biol.* **217**, 2777–2798 (2018).
251. Lehmann, B. D. B. *et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Invest.* **121**, 2750–2767 (2011).
252. Hinohara, K. & Polyak, K. Intratumoral Heterogeneity: More Than Just Mutations. *Trends Cell Biol.* **29**, 569–579 (2019).
253. Usary, J., Darr, D. B., Pfefferle, A. D. & Perou, C. M. Overview of genetically engineered mouse models of distinct breast cancer subtypes. *Curr. Protoc. Pharmacol.* **2016**, 14.38.1–14.38.11 (2016).
254. Horiuchi, D. *et al.* MYC pathway activation in triple-negative breast cancer is synthetic lethal with CDK inhibition. *J. Exp. Med.* **209**, 679–696 (2012).
255. Lawson, D. A. *et al.* Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. (2015). doi:10.1038/nature15260
256. Patwa, A. *et al.* Multiplexed imaging analysis of the tumor-immune microenvironment reveals predictors of outcome in triple-negative breast cancer. *Commun. Biol.* **4**, 1–14 (2021).
257. Eng, J. R. *et al.* A framework for multiplex imaging optimization and reproducible analysis. *Commun. Biol.* (2022).
258. Davidson-Pilon, C. lifelines: survival analysis in Python. *J. Open Source Softw.* **4**, 1317

- (2019).
259. Chen, X. *et al.* S100 calcium-binding protein A6 promotes epithelial-mesenchymal transition through β -catenin in pancreatic cancer cell line. *PLoS One* **10**, 1–14 (2015).
 260. Finak, G. *et al.* Stromal gene expression predicts clinical outcome in breast cancer. *Nat. Med.* **14**, 518–527 (2008).
 261. Kaszak, I. *et al.* Role of cadherins in cancer—a review. *Int. J. Mol. Sci.* **21**, 1–17 (2020).
 262. Yu, Y. *et al.* Cancer-associated fibroblasts induce epithelial-mesenchymal transition of breast cancer cells through paracrine TGF- β signalling. *Br. J. Cancer* **110**, 724–732 (2014).
 263. Schapiro, D. *et al.* MCMICRO: a scalable, modular image-processing pipeline for multiplexed tissue imaging. *Nat. Methods* **19**, 311–315 (2022).
 264. Cabrita, R. *et al.* Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature* **577**, 561–565 (2020).
 265. Helmink, B. A. *et al.* B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* **577**, 549–555 (2020).
 266. Parmar, C., Barry, J. D., Hosny, A., Quackenbush, J. & Aerts, H. J. W. L. Data Analysis Strategies in Medical Imaging. *Clin. Cancer Res.* **24**, 3492–3499 (2018).
 267. Greenwald, N. F. *et al.* Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nat. Biotechnol.* **40**, (2021).
 268. van der Walt, S. *et al.* scikit-image: image processing in Python. *PeerJ* **2**, e453 (2014).
 269. Danenberg, E. *et al.* Breast tumor microenvironment structures are associated with genomic features and clinical outcome. *Nat. Genet.* **18**, (2022).
 270. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
 271. Fantozzi, A. *et al.* VEGF-mediated angiogenesis links EMT-induced cancer stemness to tumor initiation. *Cancer Res.* **74**, 1566–1575 (2014).
 272. Sato, M. *et al.* An integrated genomic approach identifies persistent tumor suppressive effects of transforming growth factor- β in human breast cancer. *Breast Cancer Res.* **16**, 1–23 (2014).
 273. Lumley, T., Diehr, P., Emerson, S. & Chen, L. The Importance of the Normality Assumption in Large Public Health Data Sets. *Annu. Rev. Public Health* **23**, 151–169 (2002).
 274. Thomas, A. *et al.* Tumor mutational burden is a determinant of immune-mediated survival in breast cancer. *Oncoimmunology* **7**, 1–12 (2018).
 275. Macheret, M. & Halazonetis, T. D. DNA replication stress as a hallmark of cancer. *Annu. Rev. Pathol. Mech. Dis.* **10**, 425–448 (2015).
 276. Pabla, S. *et al.* Proliferative potential and resistance to immune checkpoint blockade in lung cancer patients. *J. Immunother. Cancer* **7**, 1–13 (2019).
 277. Pabla, S. *et al.* Integration of tumor inflammation, cell proliferation, and traditional biomarkers improves prediction of immunotherapy resistance and response. *Biomark. Res.* **9**, 1–11 (2021).
 278. Schapiro, D. *et al.* histoCAT : analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nat. Methods* **14**, (2017).
 279. Sakunrangsit, N. & Ketchart, W. Plumbagin inhibits cancer stem-like cells, angiogenesis and suppresses cell proliferation and invasion by targeting Wnt/ β -catenin pathway in

endocrine resistant breast cancer. *Pharmacol. Res.* **150**, 104517 (2019).

APPENDIX A: SUPPLEMENTARY INFORMATION RELATED TO CHAPTER 5

Eng J. et al. Multiplatform analysis of breast cancer clinical cohorts for robust biomarker discovery. In Preparation.

List of Supplementary Figures

Figure 6.1. De-noising and Segmentation Optimization.

Figure 6.2. Mesmer versus Watershed Segmentation.

Figure 6.3. ER Quality Control.

Figure 6.4. Subtype and Single Variable Prognosis in the Combined Cohort.

Figure 6.5. Stromal Subtypes.

Figure 6.6. Determining Cell Abundance Cut-offs.

Figure 6.7. Cell Type Correlation in Breast Cancer Tissues.

Figure 6.8. Epithelial Subtype versus Stromal Cell Abundance.

Figure 6.9. Prognostic Value of Proximity in Separate Platforms.

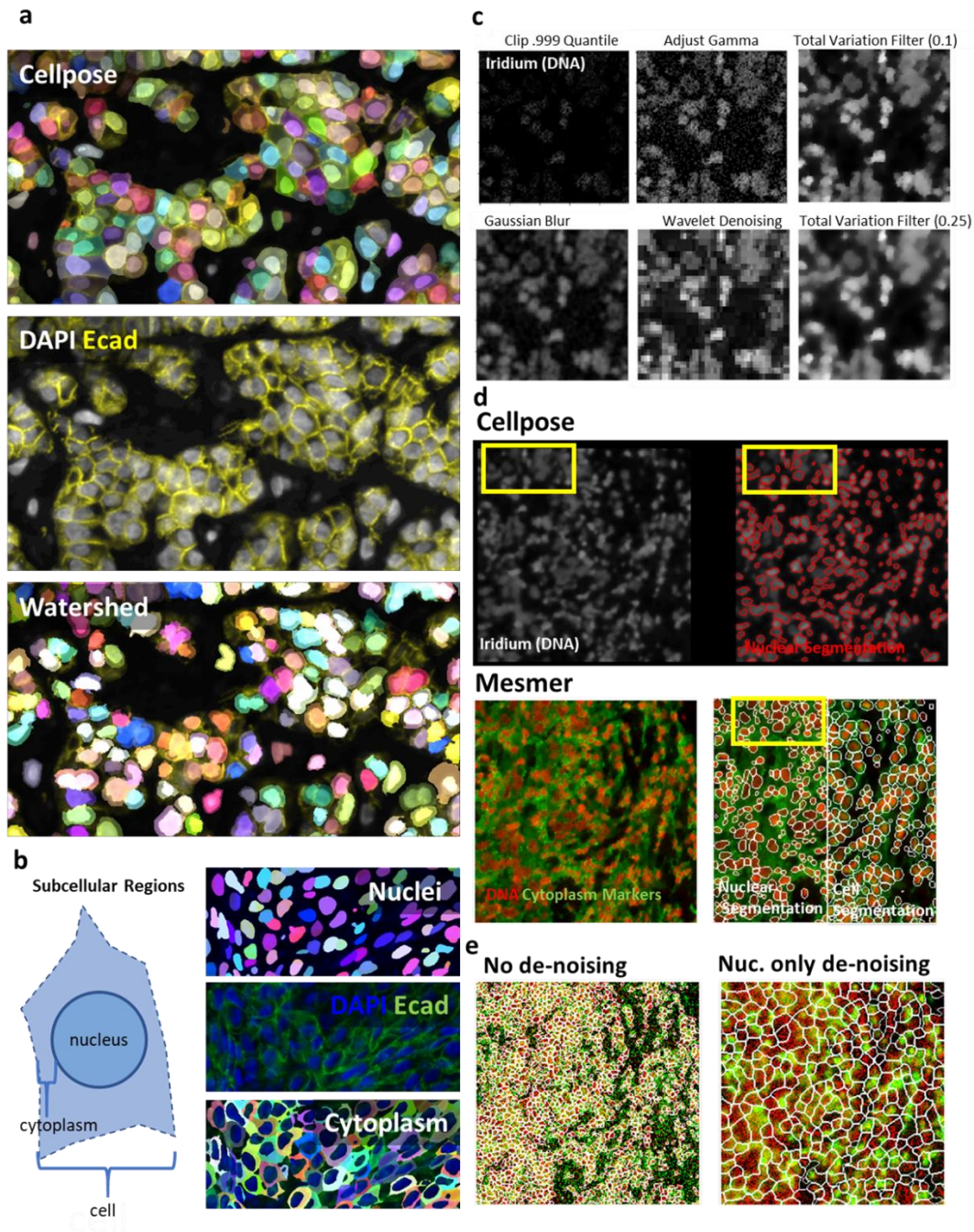


Figure 6.1. De-noising and Segmentation Optimization.

a. CyCIF Cellpose nuclear and cell segmentation (top) versus watershed segmentation (bottom). b. Construction of cytoplasm mask from cell and nuclear masks after matching with mplexable. c. Image processing steps tested for the IMC denoising pipeline. d. IMC Cellpose segmentation (top) versus Mesmer segmentation (bottom). e. IMC Mesmer segmentation of gamma-adjusted only image of ROI shown in (d) (left), and of de-noised nuclear channel with no cytoplasm denoising (right).

a

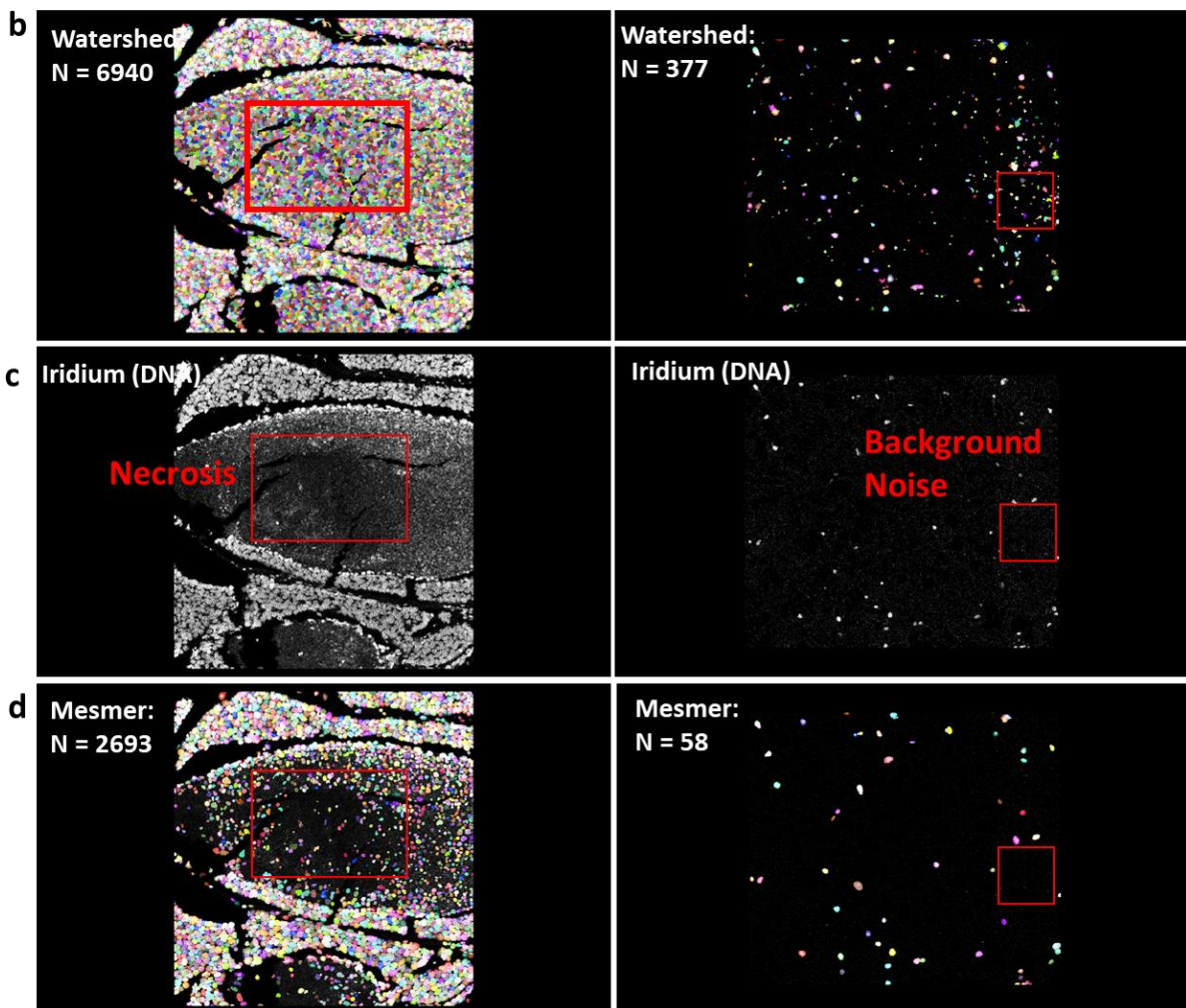
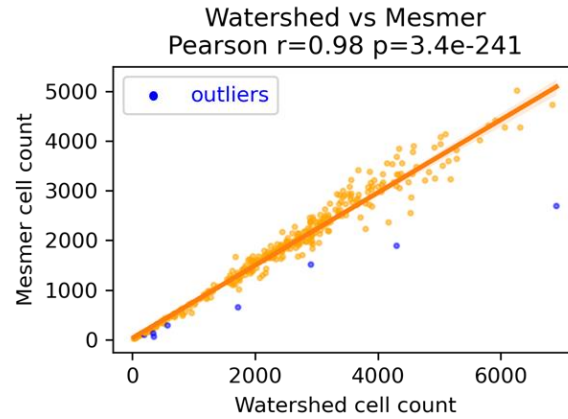


Figure 6.2. Mesmer versus Watershed Segmentation.

a. Pearson correlation between cell counts of watershed versus Mesmer segmentation in IMC tissues. Tissues with $>50\%$ change shown in blue. b. Watershed segmentation of two selected outliers. Red boxes indicate areas of over segmentation. c. DNA channel of selected outliers. d. Mesmer segmentation of selected outliers.

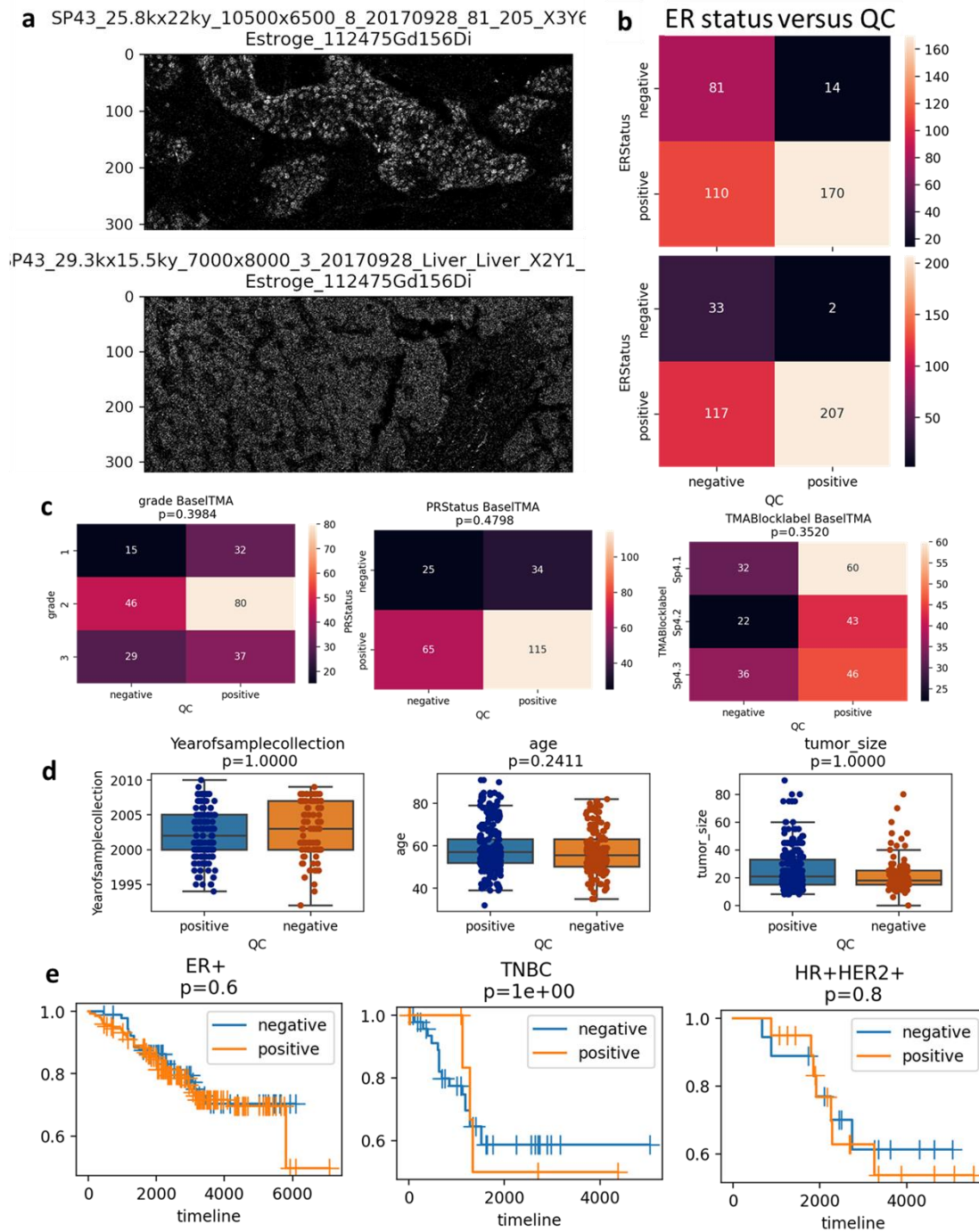


Figure 6.3. ER Quality Control.

a. Representative IMC images of estrogen receptor channel QC images sorted for positive (top) and negative (bottom) nuclear ER staining. b. Image ROIs classified by annotated ER status (y-axis) versus our QC call for ER positive or negative (x-axis) in two IMC TMAs. c. Grade, PR status and TMA block versus QC calls, p-values from Chi-squared analysis shown in figure title. b-c. N number of ROIs in each category annotated on heatmap cells. d. Year of sample collection, patient age and tumor size versus QC status of each ROI, p-values for Wilcoxon-rank sum test shown in figure title. e. Kaplan-Meier curves for OS versus QC status, p-values from log-rank test.

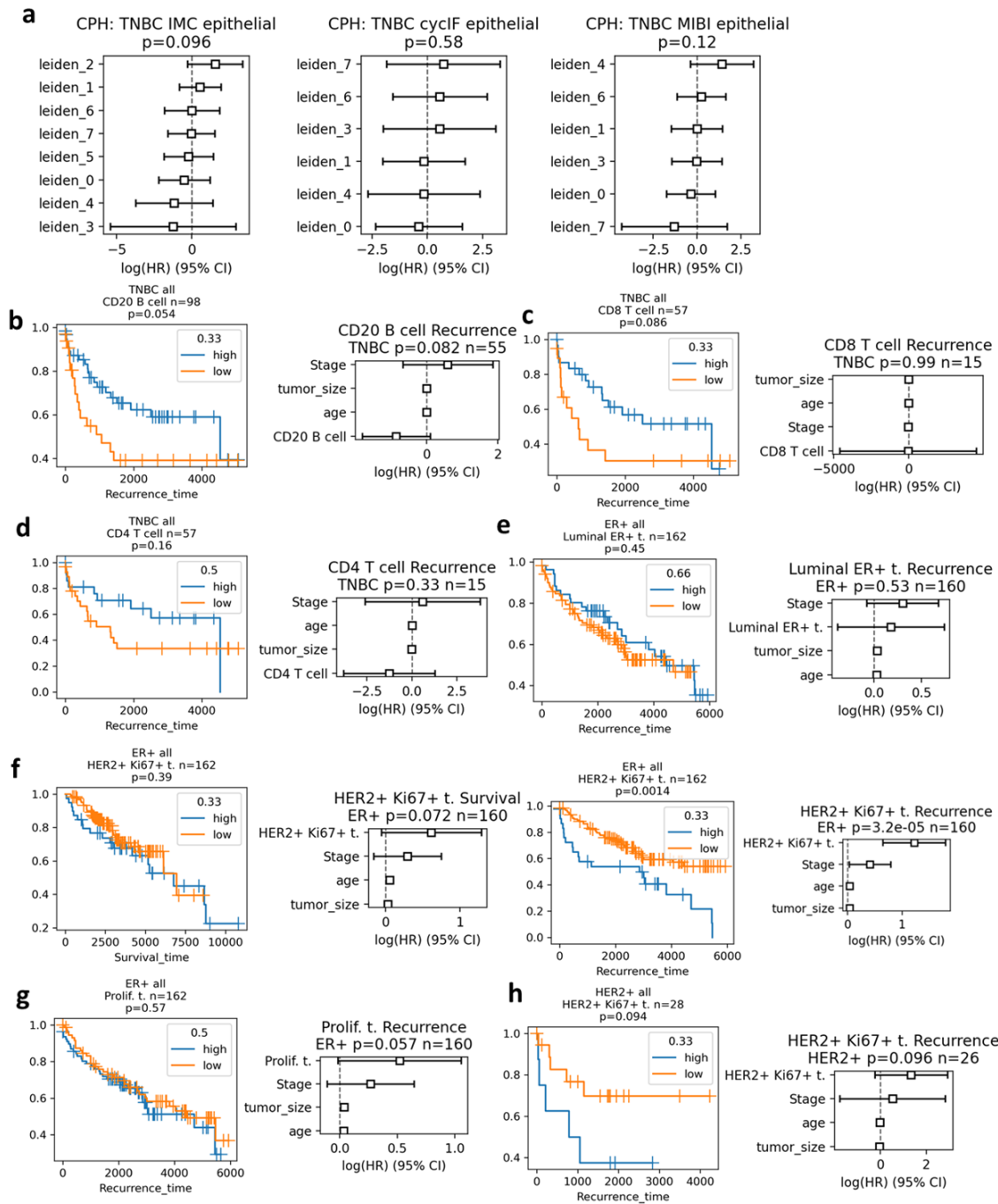


Figure 6.4. Subtype and Single Variable Prognosis in the Combined Cohort.

a. Cox proportional hazard (CPH) modelling of epithelial subtypes versus overall survival in TNBC, platform and p-value shown in title. b-h. Kaplan-Meier (K-M) and multivariable CPH modelling of high/low abundance of cell types versus OS and RFS. Cut-off values (tertiles or median) given in K-M legends. K-M p-values derived from log-rank test and given in figure titles, along with subtype and n number of patients. CPH p-value for cell type and patient n given in titles.

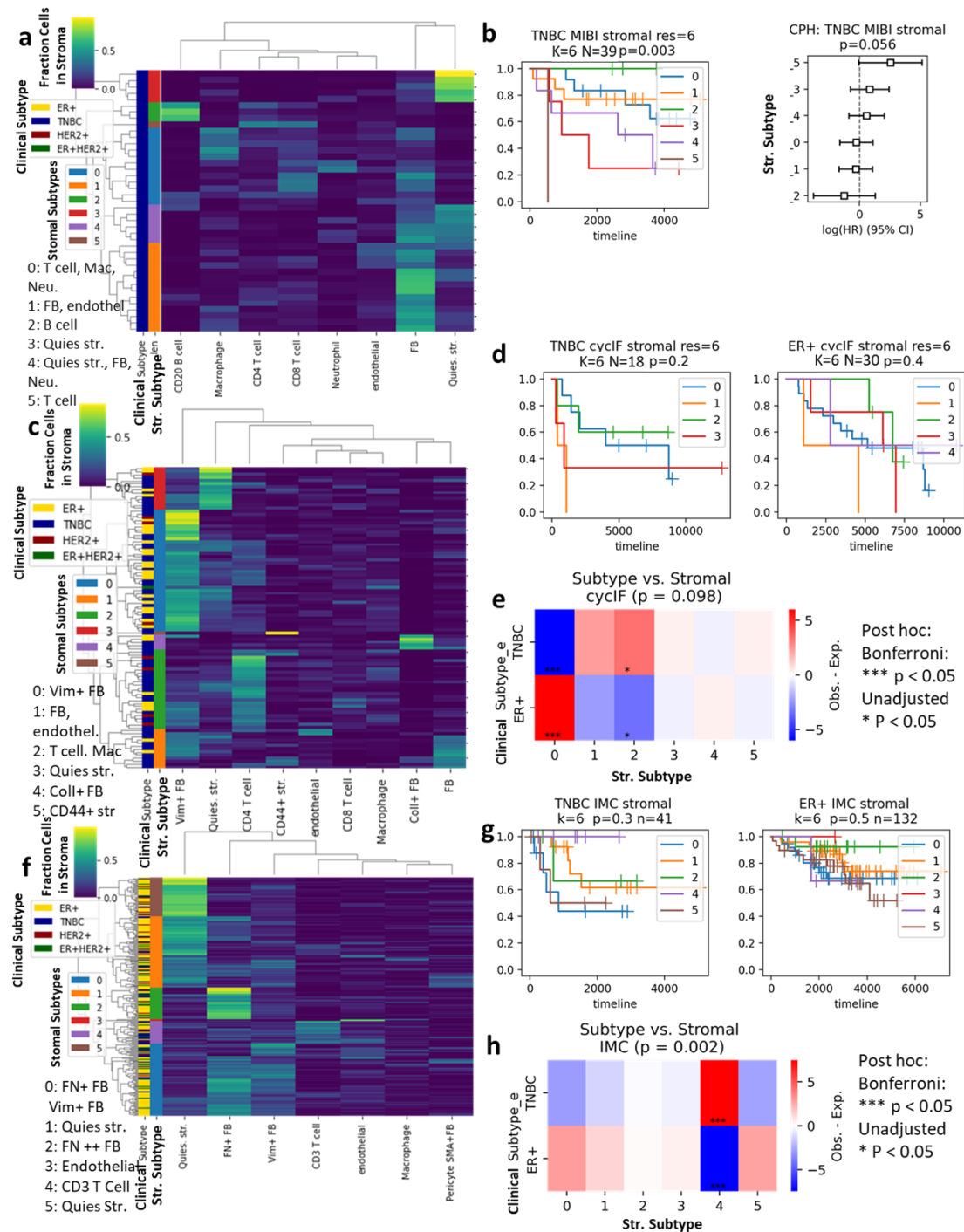


Figure 6.5. Stromal Subtypes.

a, c, f. All patient tissues from each platform were hierarchically clustered based on the fraction of the common stromal cell types (>2%) in all stromal cells, selecting k=6 stromal subtypes. b. MIBI Kaplan-Meier curves (p-value from log-rank test) and Cox proportional hazard (CPH) models comparing OS in stromal subtypes, n=39 patients. c, d. CyCIF Kaplan-Meier curves comparing OS in stromal subtypes (p-value from log-rank test) n=18 TNBC, 30 ER+ patients. f, g. IMC Kaplan-Meier curves comparing OS in stromal subtypes (p-value from log-rank test) n=41 TNBC, 132 ER+ patients. e, h. Observed minus expected number of patients for each clinical subtype versus stromal subtype for CyCIF (e) and IMC (h). Overall p-value given in title (Chi-squared) and pairwise Bonferroni adjusted (***) and unadjusted (*) p-values < 0.05 marked on heatmap cells.

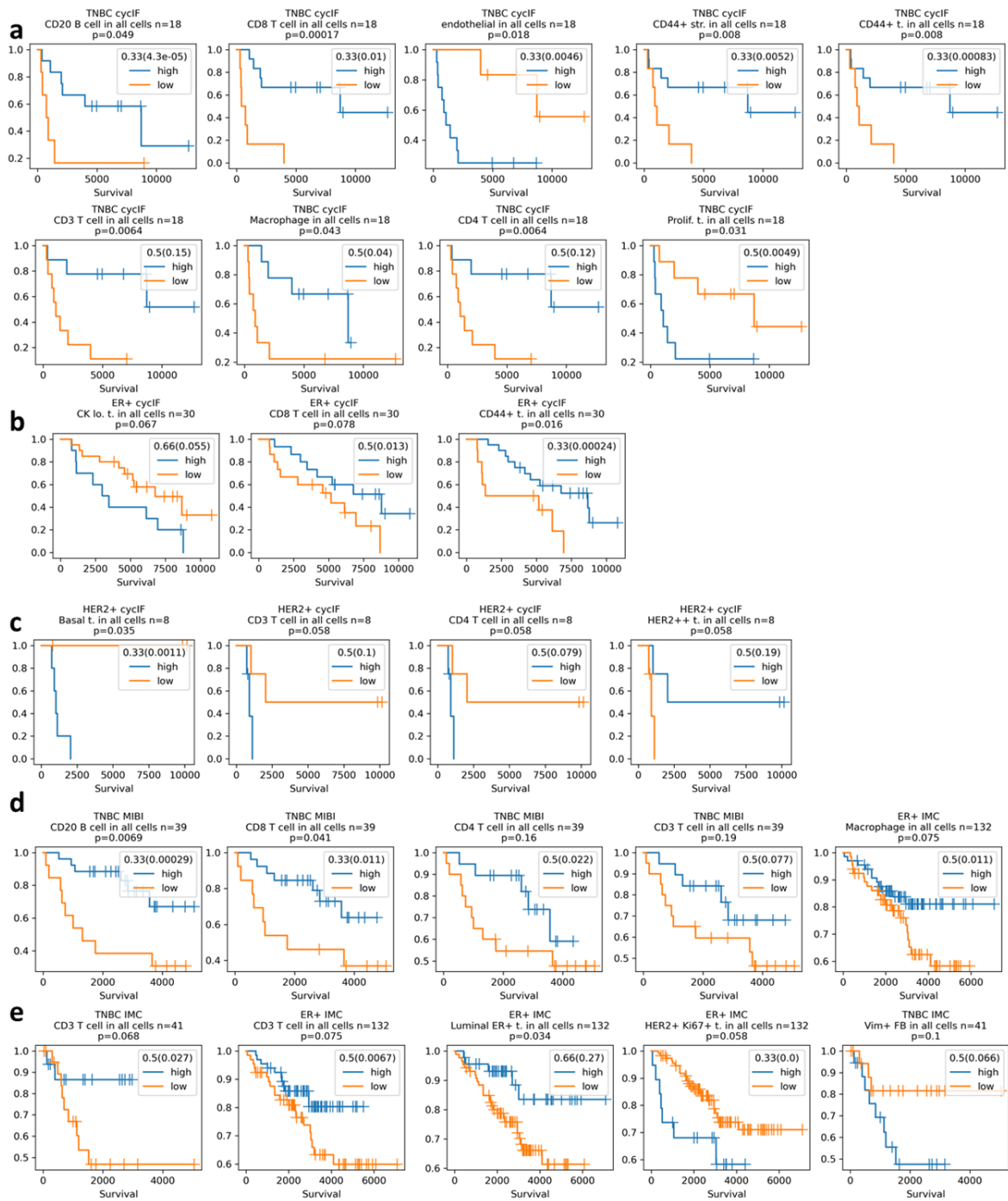


Figure 6.6. Determining Cell Abundance Cut-offs.

a-c . Kaplan-Meier (K-M) curves of high/low abundance of various cell types versus OS in CyCIF cohort. d-e. Kaplan-Meier (K-M) curves of high/low abundance of cell types versus OS in IMC and MIBI cohorts. a-e. Cut-off values (tertiles of median) given in K-M legends. K-M p-values derived from log-rank test and given in figure titles, along with subtype and n number of patients.

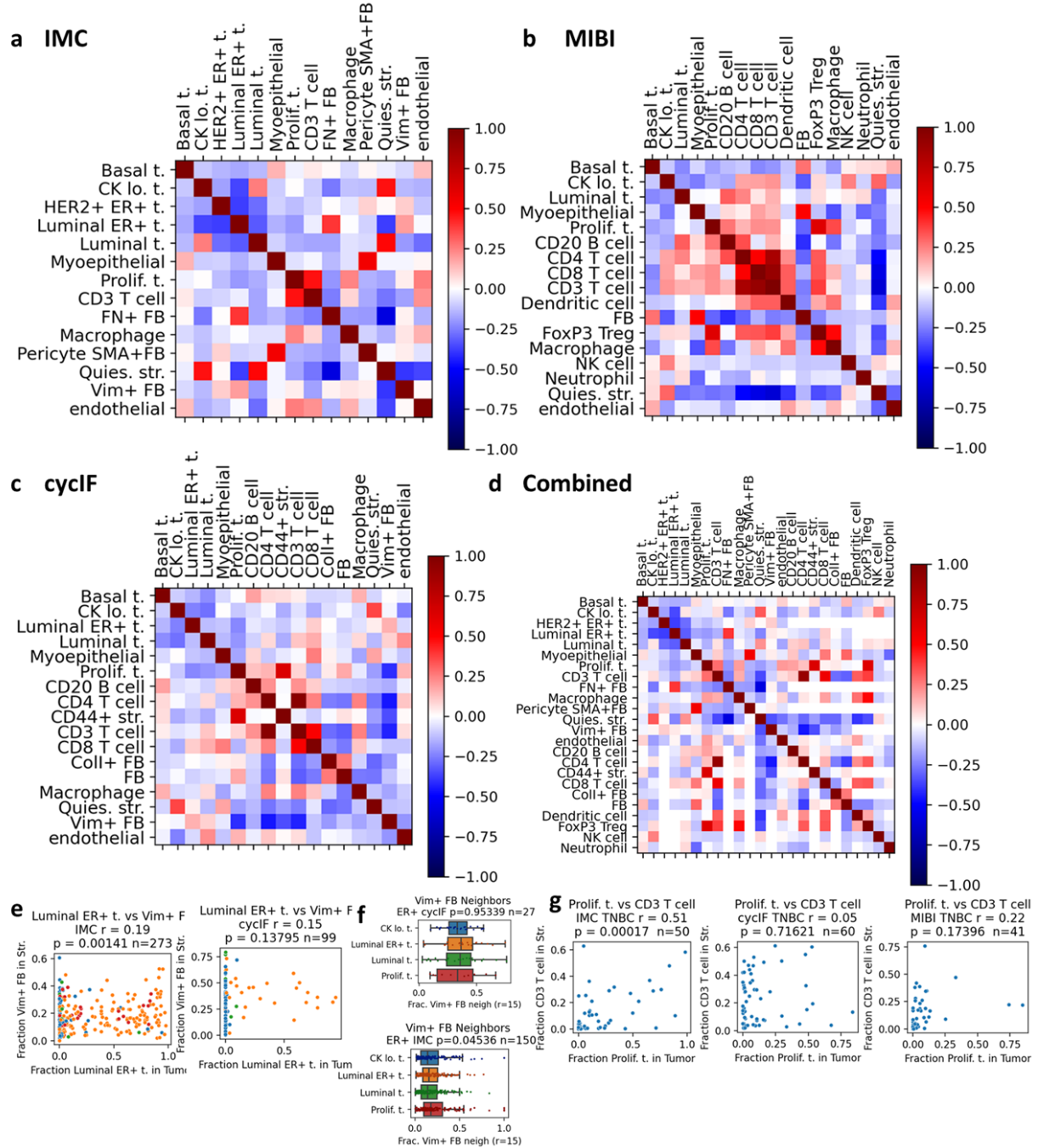


Figure 6.7. Cell Type Correlation in Breast Cancer Tissues.

a-d. Pairwise cell type Pearson correlation R values for tissues from IMC (a), MIBI (b), CyCIF (c), and three cohorts combined (d). e. Correlation in IMC (left) and CyCIF (right) between Vim+ FB and Luminal ER+ tumor in ER+ tumors. f. Stromal neighbor fraction of Vim+ FB versus tumor phenotypes. g. Correlation in IMC (left), CyCIF (center) and MIBI (right) between CD3 T cell and proliferating tumor in TNBC. e-g. All correlation r- and p-values are Pearson correlation, all boxplot p-values are Kruskal-Wallis H test for difference in medians. N number of patients given in figure title.

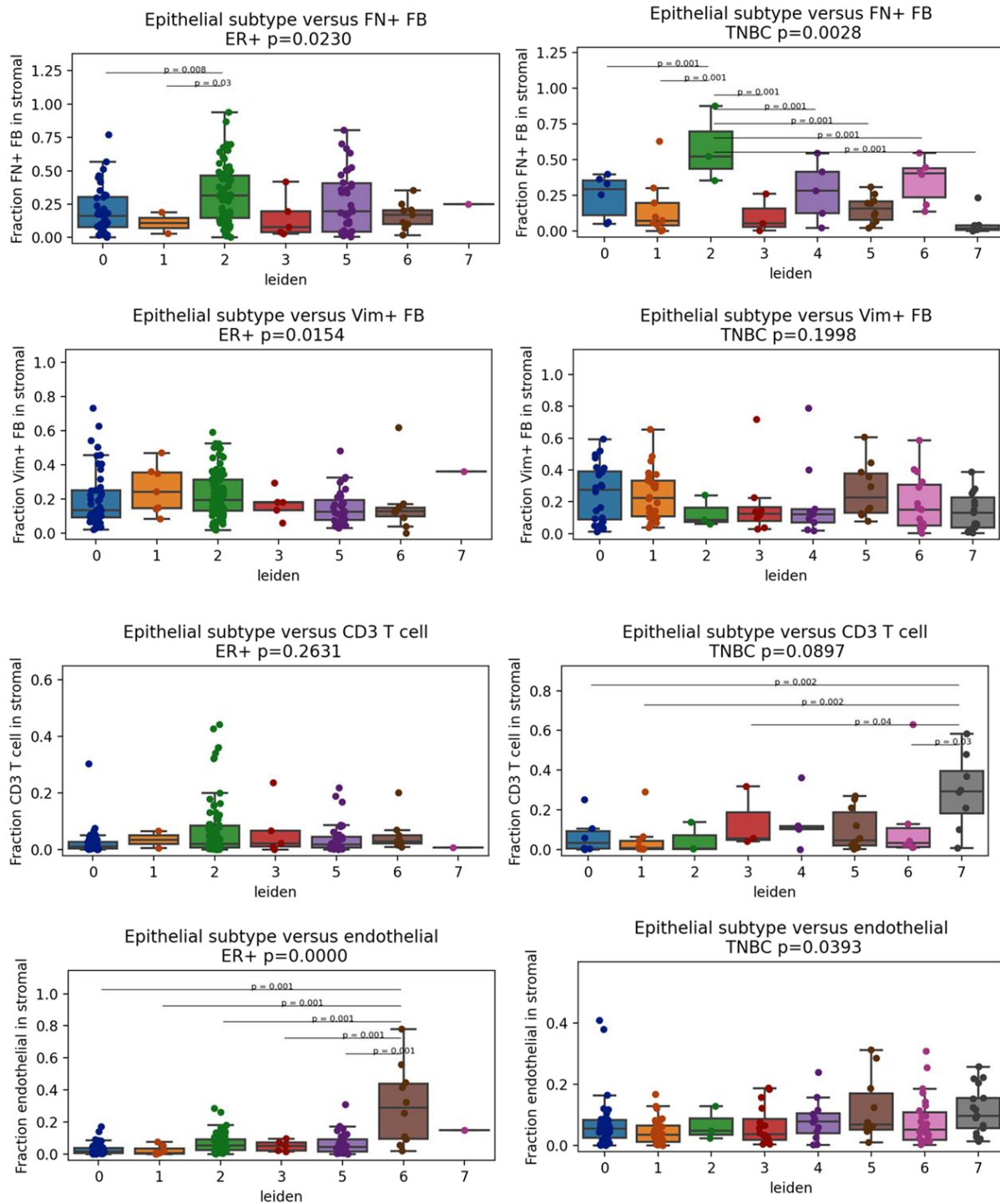


Figure 6.8. Epithelial Subtype versus Stromal Cell Abundance.

Eight epithelial subtypes versus stromal fractions of FN+ FB, Vim+ FB, CD3 T cell and endothelial cells for ER+ (right) and TNBC (left). Overall group difference p-value shown in title (Kruskal-Wallis) and pairwise corrected p-values shown on plot (Tukey's HSD). Each dot denotes a separate patient.

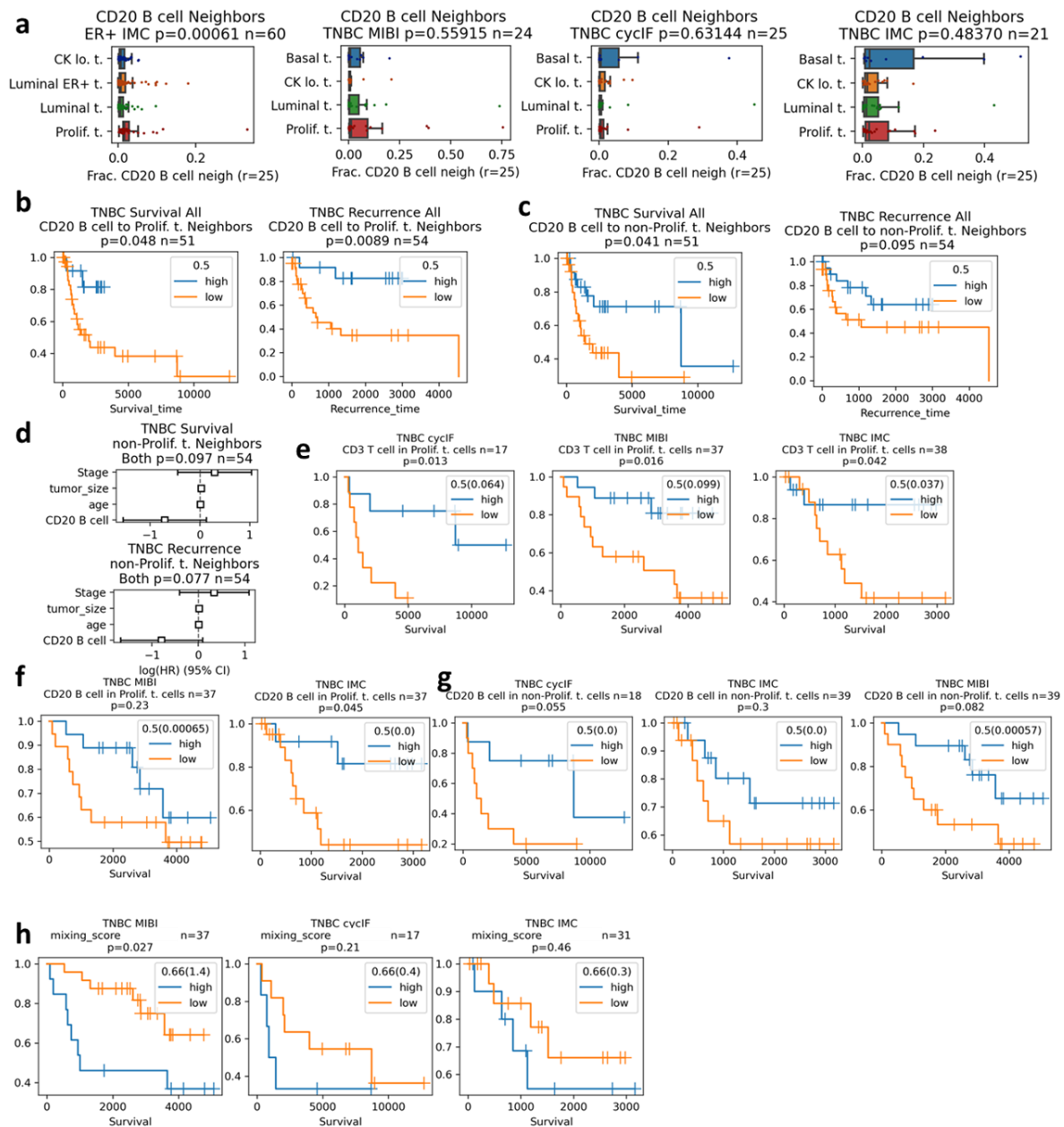


Figure 6.9. Prognostic Value of Proximity in Separate Platforms.

a. Fraction of CD20 B cells in stromal cells within a 25 μ m radius of tumor cell types; all boxplot p-values are Kruskal-Wallis H test for difference in medians, n number of patients given in figure title. b, e-g. Kaplan-Meier (K-M) curves of high/low abundance of cell neighbors within 25 μ m of proliferating or non-proliferating tumor cells versus survival. d. Multivariable CPH of high abundance B cell neighbors to non-proliferating tumor versus OS and RFS h. Mixing score versus OS in each platform. b, e-h. Cut-off values given in K-M legends (tertiles or median). K-M p-values derived from log-rank test and given in figure titles, along with subtype, platform ('All' indicates all three platforms) and n number of patients.

APPENDIX B: SUPPLEMENTARY INFORMATION RELATED TO CHAPTER 3

A framework for multiplex imaging optimization and reproducible analysis

Jennifer Eng^{1,2}, Elmar Bucher^{1,2}, Zhi Hu¹, Ting Zheng³, Summer L. Gibbs^{1,2}, Koei Chin^{1,2} and

Joe W. Gray^{1,2*}

List of Supplementary Figures and Tables

Supplementary Figure 1: Standard IF versus CyCIF in Normal Breast and HER2+ Tumor Tissue.

Supplementary Figure 2: Regional Variation in Threshold Results.

Supplementary Figure 3: Tissue Retention during CyCIF.

Supplementary Figure 4: 15-Minute Quenching in Different H₂O₂ Concentrations.

Supplementary Figure 5: Quenching under Different Time and Light Conditions.

Supplementary Figure 6: First versus second antibody application to same TMA tissue.

Supplementary Figure 7: Quantification of first versus second antibody application to same TMA tissue.

Supplementary Figure 8: Variables Impacting Round-Effect of Stain Quality and Bleed Through Observed in Miltenyi MACSima prototype Instrument.

Supplementary Figure 9: Single Cell Effect of Quenching in Pancreas Tissue.

Supplementary Figure 10: Single Cell Effect of Quenching in 72-core Tissue Microarray.

Supplementary Figure 11: Single Cell Annotation for Evaluation of Autofluorescence Subtraction.

Supplementary Figure 12: Single Cell Effect of Quenching in 11-core HER2+ Tumor Tissue Microarray.

Supplementary Figure 13: Replicates of Three HER2+ Tumor TMA Cores

Supplementary Figure 14: Comparison of Manual Versus Automated SBR

Supplementary Figure 15: kBET evaluation of batch correction methods in HER2+ breast cancer tissue.

Supplementary Figure 16: kBET evaluation of batch correction methods.

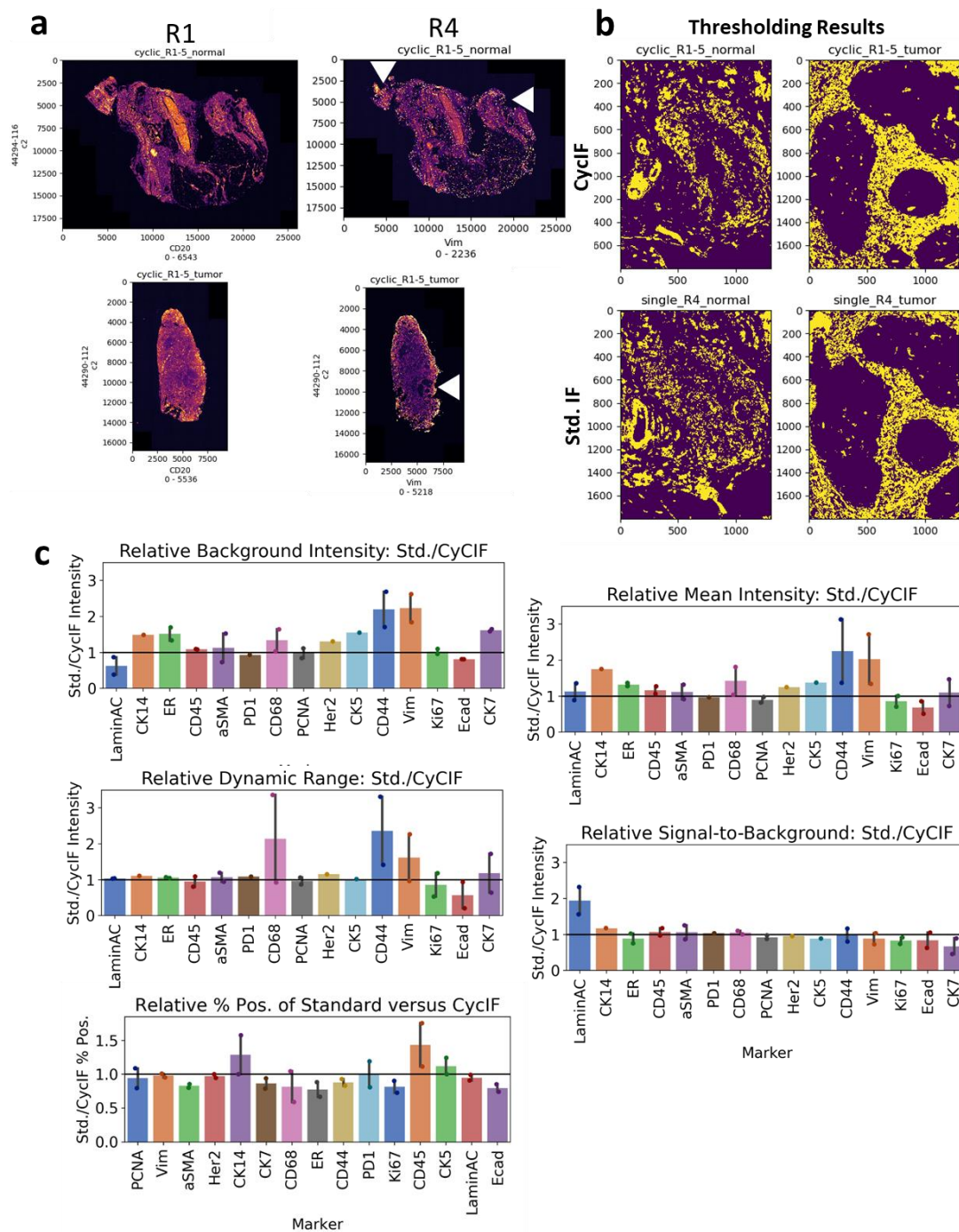
Supplementary Figure 17: Normalization and cluster annotation.

Supplementary Table 1: Double Application of Antibodies

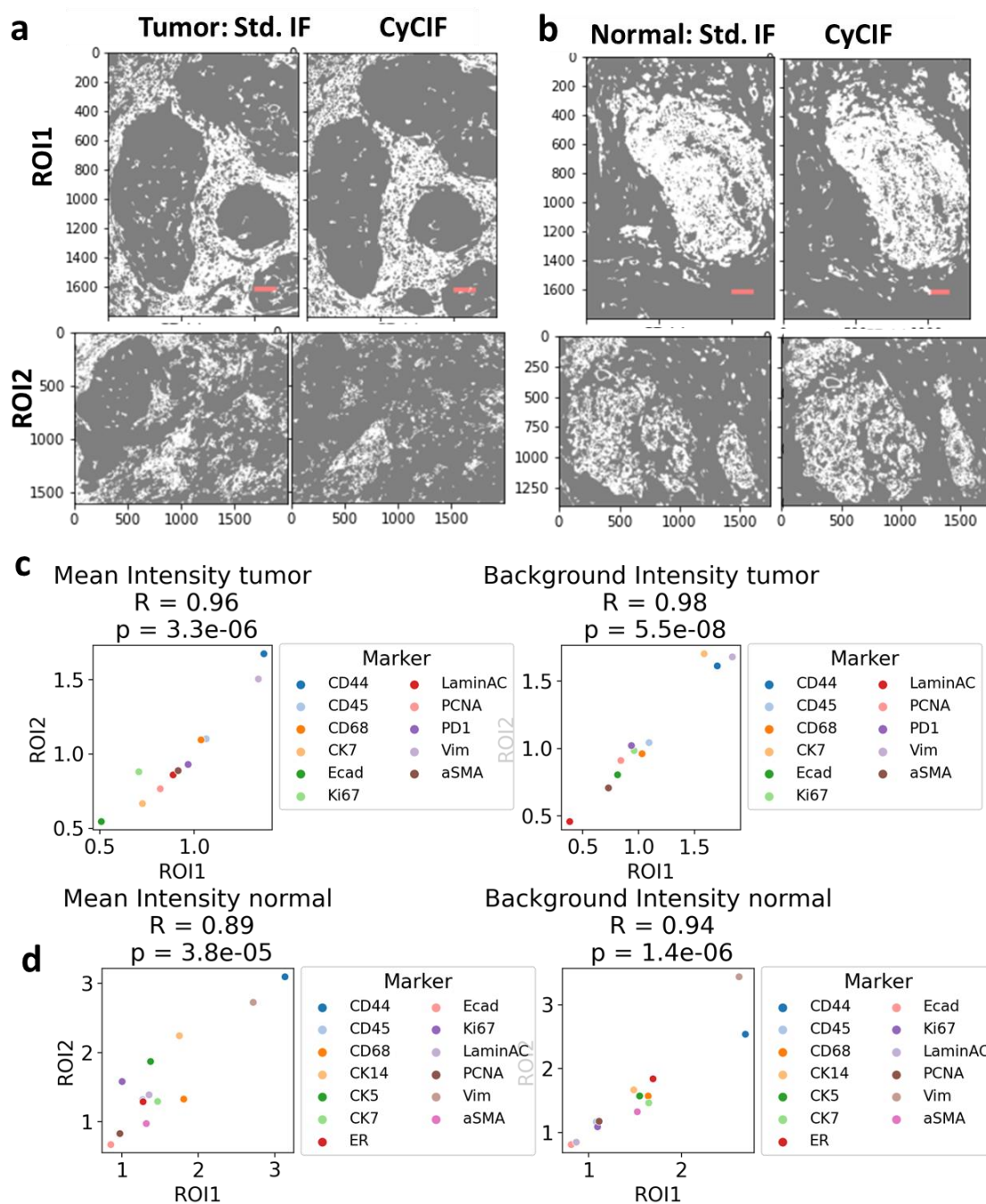
Supplementary Table 2: Antibody Order Optimization

Supplementary Table 3: Control TMA

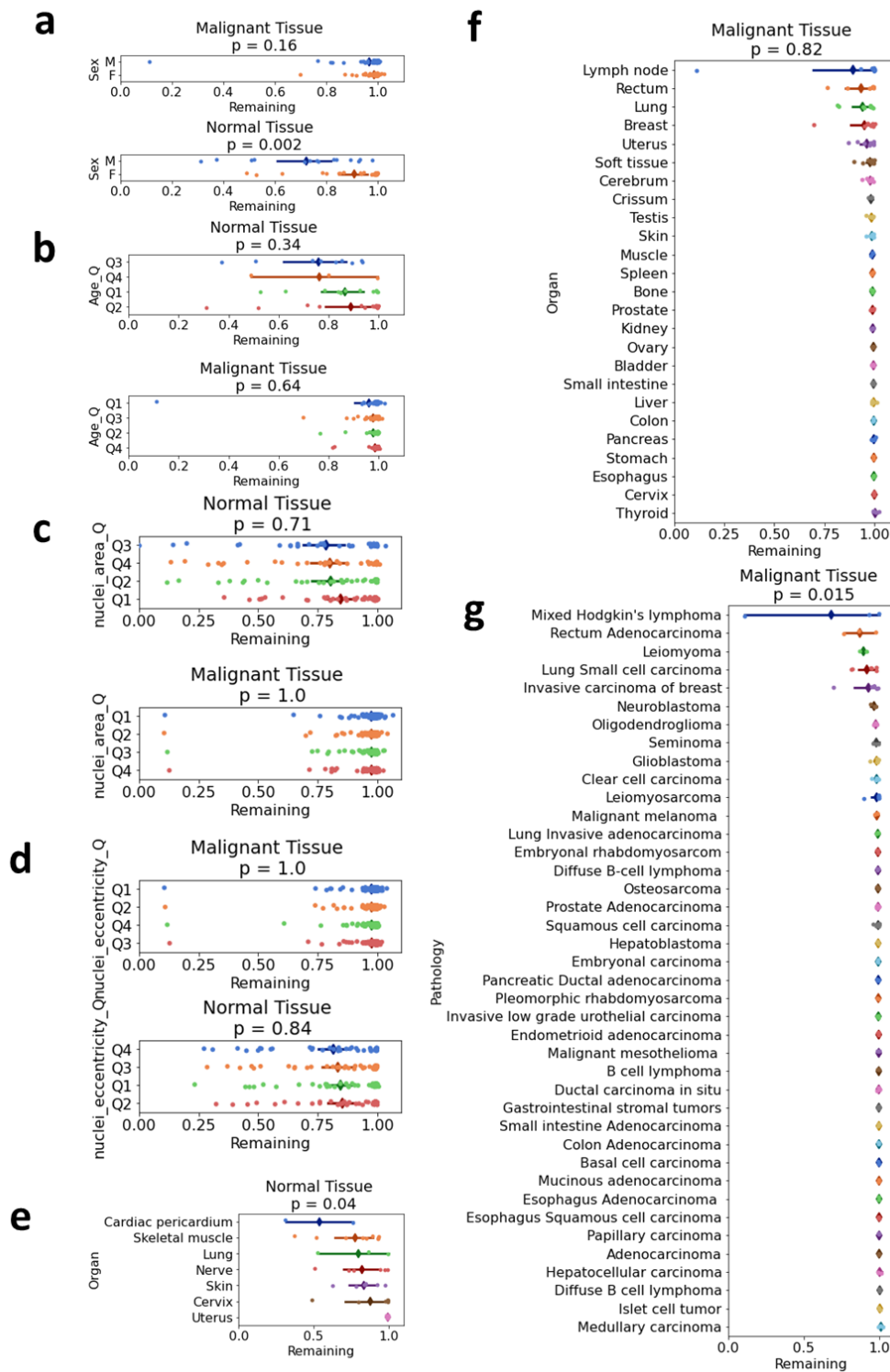
Supplementary Table 4: Antibody Order in Replicate TMAs



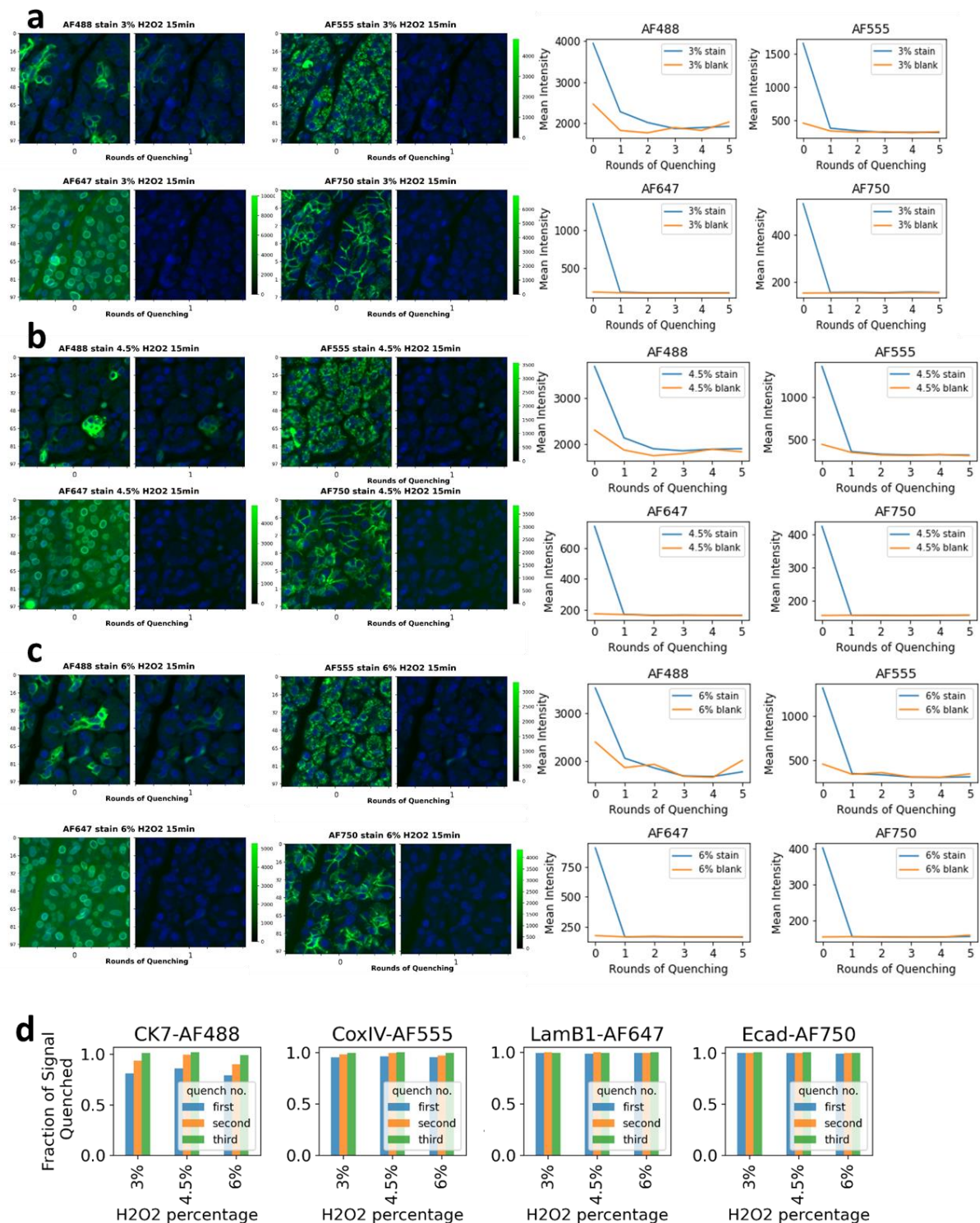
Supplementary Figure 1: Standard IF versus CycIF in Normal Breast and HER2+ Tumor Tissue. a. Tissue overview, display range in x-axis label. Round (R) 1 on left, versus R4 on right. Arrowheads show areas of tissue loss. b. Region-of-Interest (ROI) analyzed and mask applied for signal-to-background ratio (SBR) calculation. Yellow=Foreground pixels; background pixels are located 30 pixels (~10 μ m) away from foreground pixels. Top panels are CycIF ROIs and bottom panels are matched standard IF ROIs. c. Intensity, Background, Dynamic Range and SBR and Percent Positive quantification based on threshold in b. (additional markers' thresholds and visualizations here: https://github.com/engjen/cycIF_Validation)



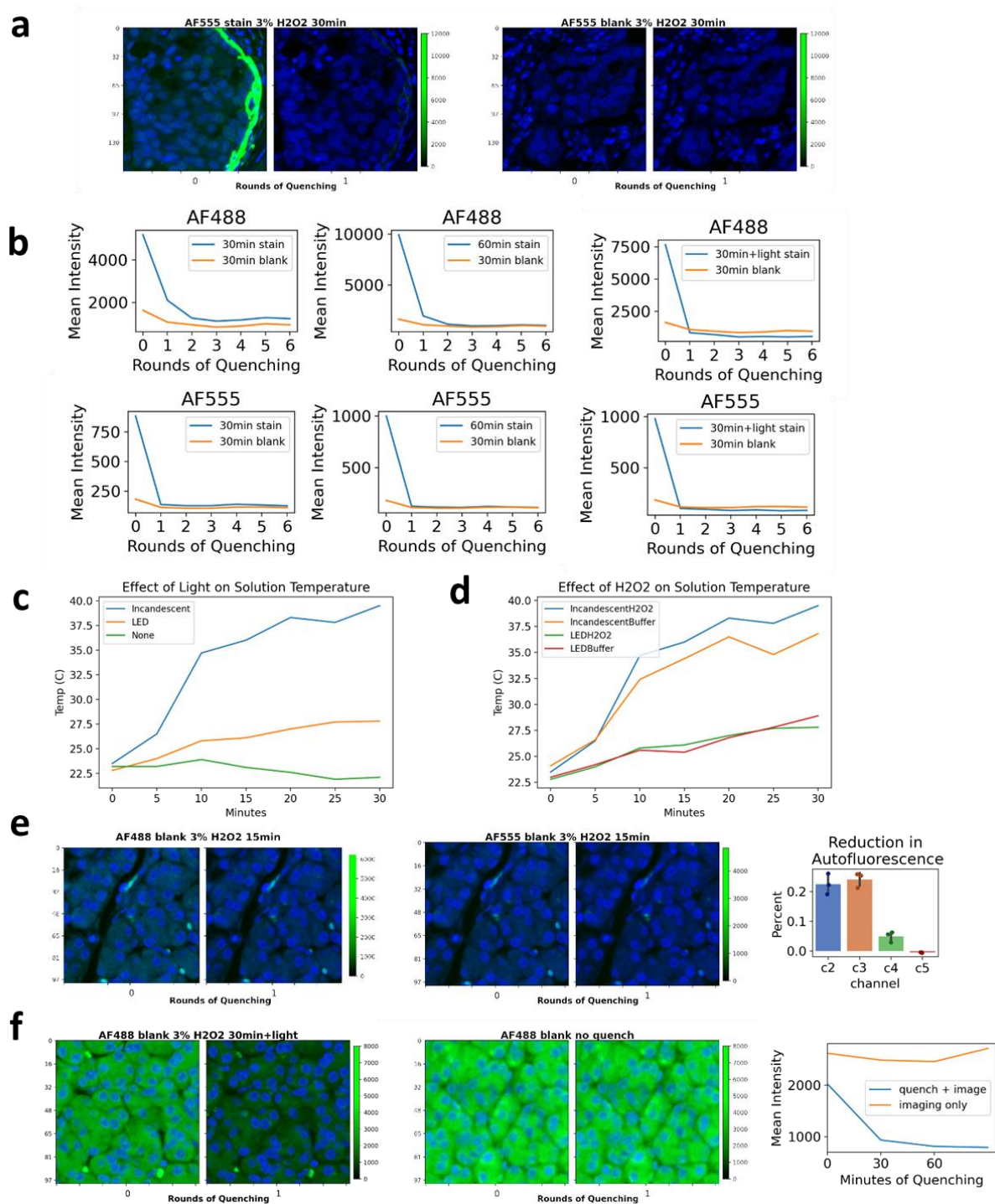
Supplementary Figure 2: Regional Variation in Threshold Results. a-b. The same threshold applied to CD44 staining in different ROIs in the tissue, a, tumor tissue and b, normal breast. Top panels show the ROI used for thresholding. Bottom panels show a separate ROI picked from corresponding areas in standard IF and CyCIF slides. Light areas are pixels above threshold, dark areas are pixels below threshold. Comparison of Std. versus CyCIF shows similar pixel patterns in both ROIs. Red scale bar = 50 μm . Axis labels show pixel coordinates; 1 pixel = 0.325 μm . c-d. Correlation of mean positive intensity (left) and background intensity (right) between ROI1 and ROI2. Pearson correlation (R) and p value shown in figure title. c. Correlation of markers expressed in tumor. d. Correlation of markers expressed in normal breast.



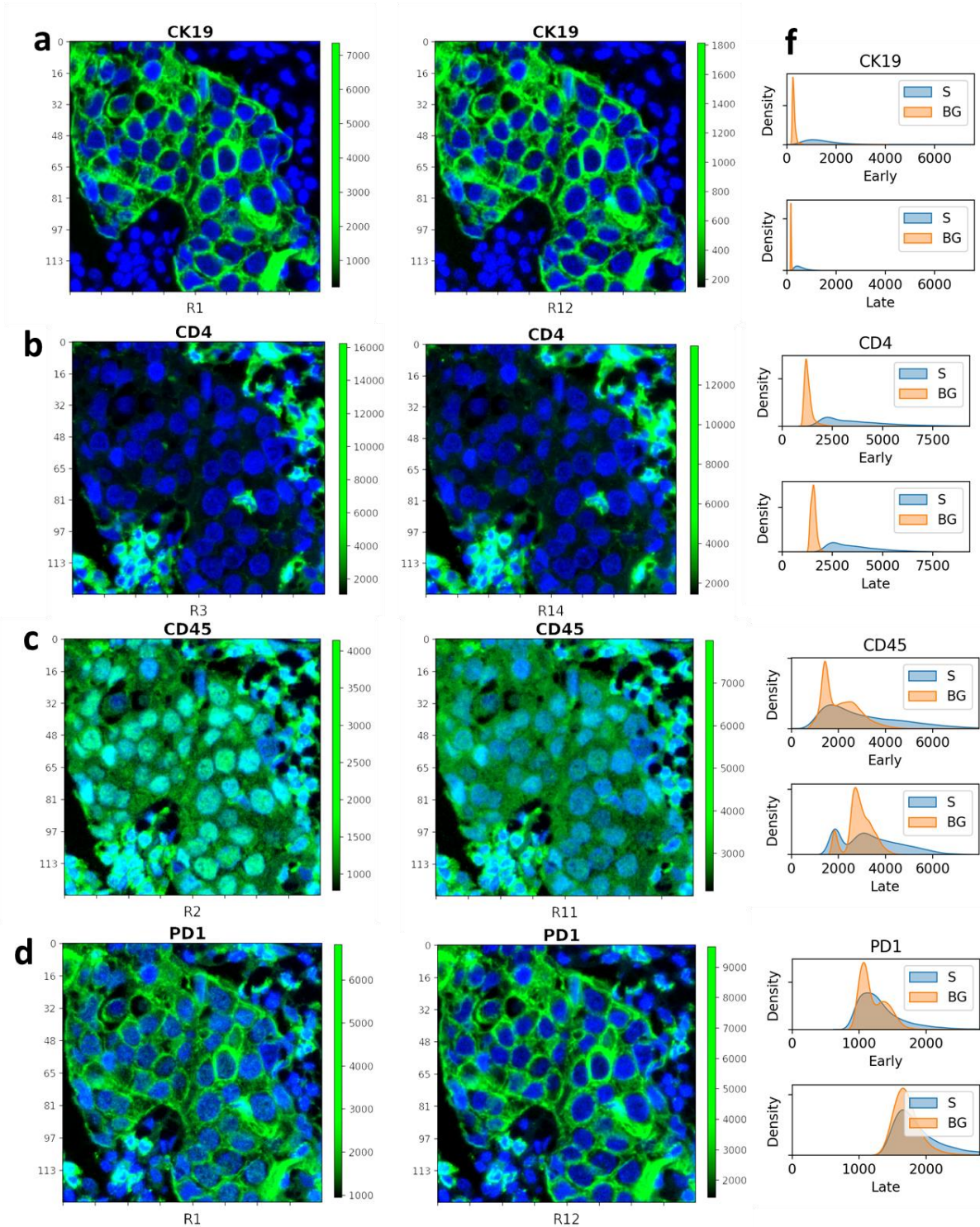
Supplementary Figure 3: Tissue Retention during CyCIF. a-g. X-axis shows fraction of cells remaining after 10 rounds of CyCIF. Y-axis shows variables and groups compared. ANOVA was used to assess significant differences; p-value in figure title. n = 156 malignant (top), n = 54 normal (bottom). a. Sex. b. Patient age, by quartile. c. Nuclear area, by quartile. d. Nuclear eccentricity, by quartile. e. Organ, normal tissue. f. Organ, malignant tissue. g. Pathology Diagnosis, malignant tissue.



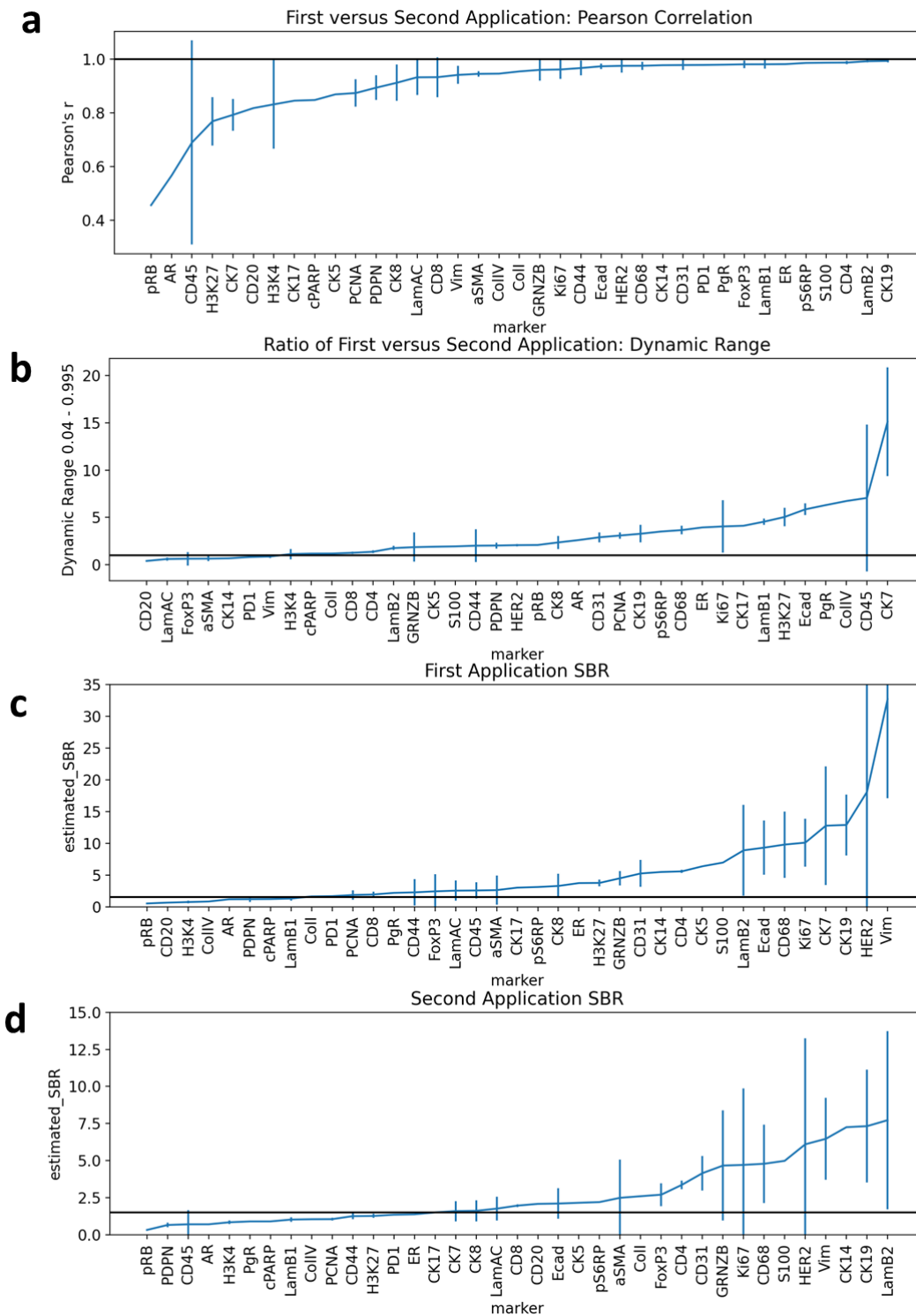
Supplementary Figure 4: 15-Minute Quenching in Different H₂O₂ Concentrations. a-c. Image panels show staining and after first quench, line plots quantify mean intensity in tissue area over 5 rounds of quenching (blue line). a. 3% H₂O₂ b. 4.5% H₂O₂ c. 6% H₂O₂. Negative controls for all concentrations H₂O₂ quantified in a-c (orange line). d. Fraction of signal quenched after first, second and third round of quenching. AF488 takes 3 rounds to quench, AF647 and AF750 only one round.



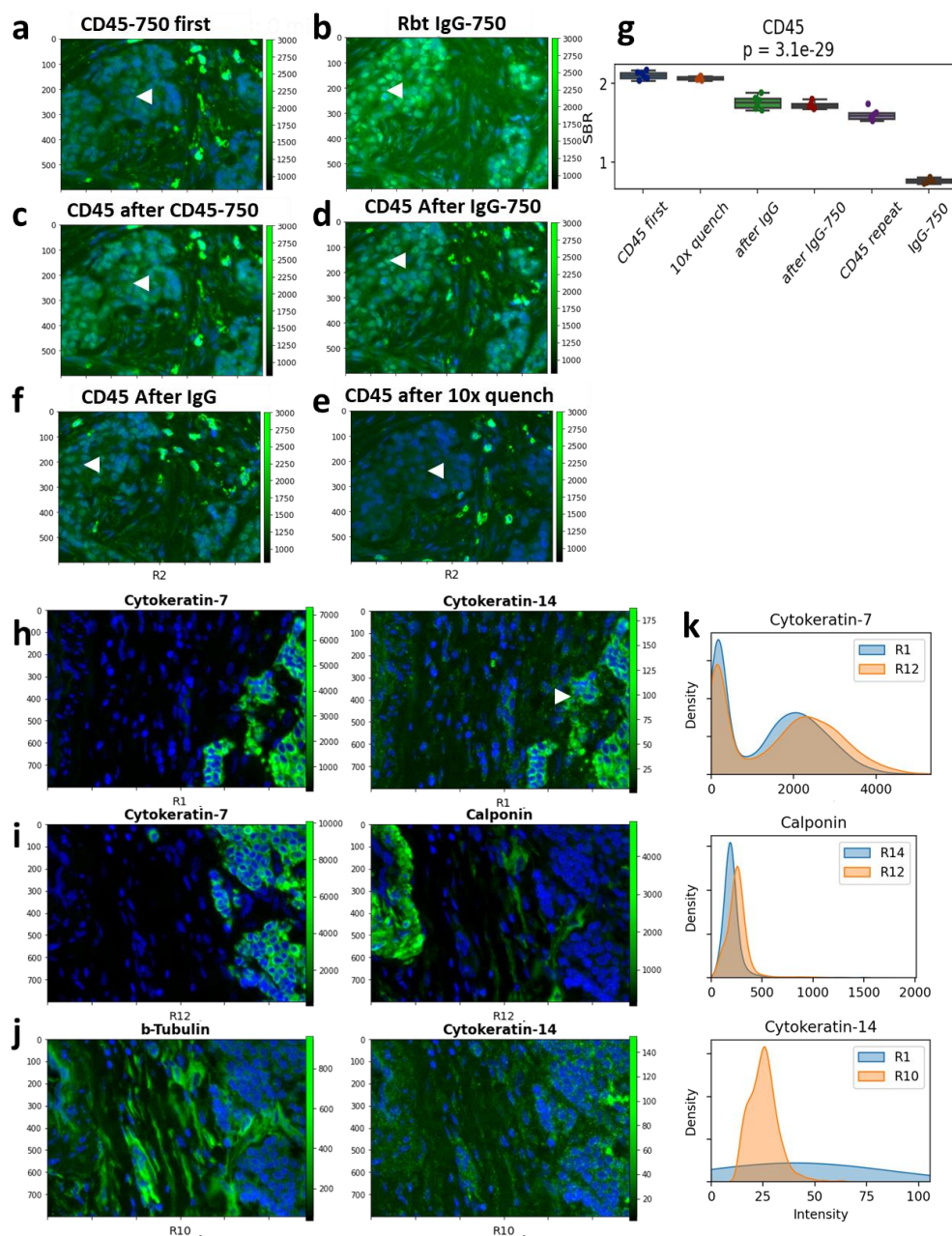
Supplementary Figure 5. Quenching under Different Time and Light Conditions. a-b. Quenching performed for 30-minutes, 60 minutes and 30 minutes-plus-light. a. Image panels show AF555 staining and after first quench (left) and blank control (right). b. Line plots quantify mean intensity in tissue area over 6 rounds of quenching, for AF488 (top) and AF555 (bottom), blue line = staining, orange line = blank control. c. Temperature of Quenching Solution under different light conditions. d. Temperature of quenching solution with and without H_2O_2 , under two light conditions. e. Negative controls in AF488 (left) and AF555 channel (middle), quantified on right, showing reduction in autofluorescence after quenching. f. Blank tissue quenched and imaged (left) versus imaged only (middle) shows quenching is necessary for decrease autofluorescence, seen in comparing blue and orange line in quantification (right).



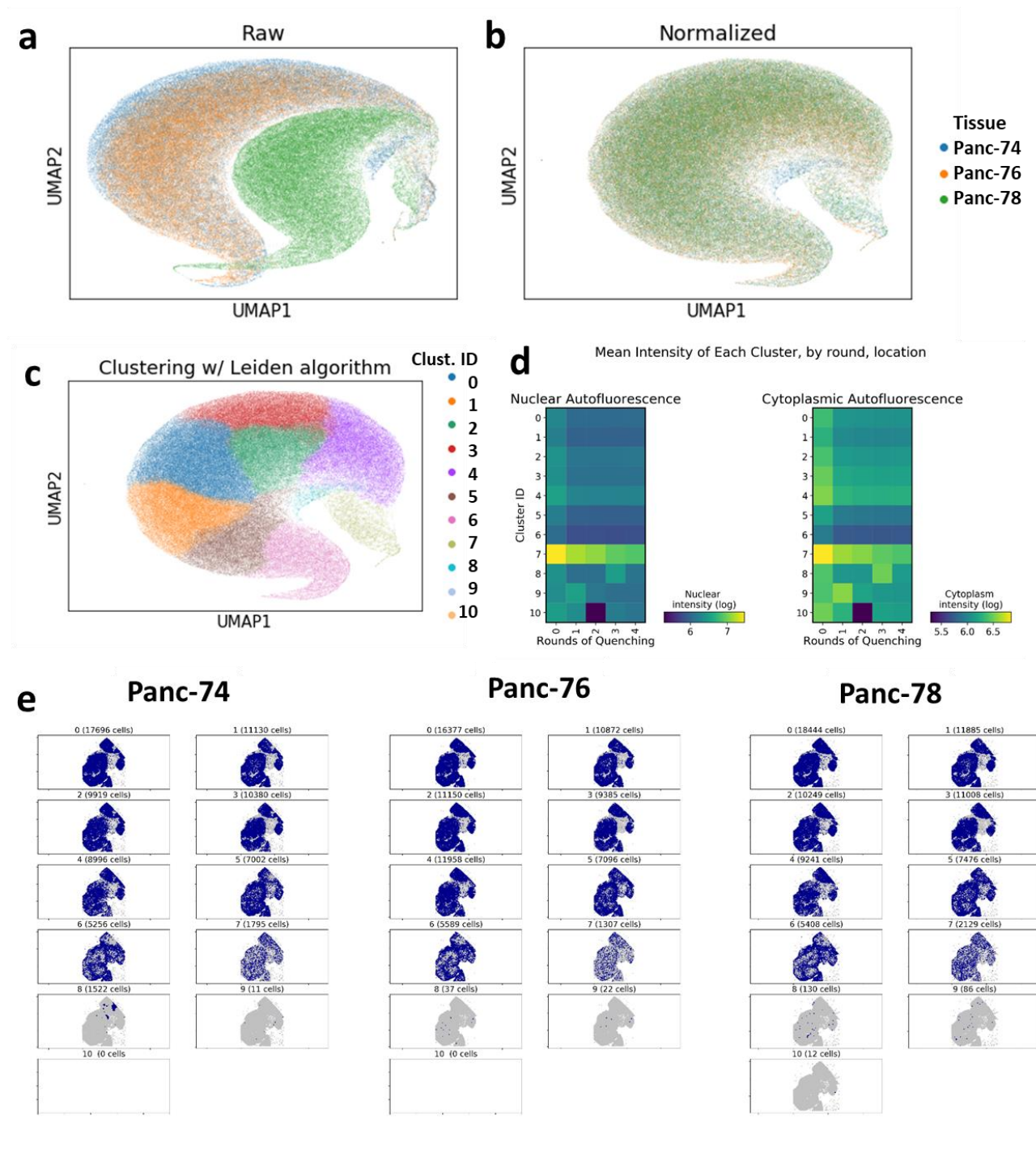
Supplementary Figure 6: First versus second antibody application to same TMA tissue. a. CK19 in R1 (first) and again in R12 (second) shows decrease in dynamic range. b. CD4 shows little decrease in dynamic range. c. CD45 shows improvement in non-specific nuclear staining on second application. d. PD1-AF647 has clear bleed through from CK19-AF750 (matching pattern in a), in both applications. f. Single cell mean intensity distribution of positive signal (S) or background (BG). Cell types were defined by thresholding and gating. (Signal = positive cells for respective marker. Background = Tumor cells for stromal markers, and vice versa. Additional markers' visualizations and figures here: https://github.com/engjen/cycIF_Validation)



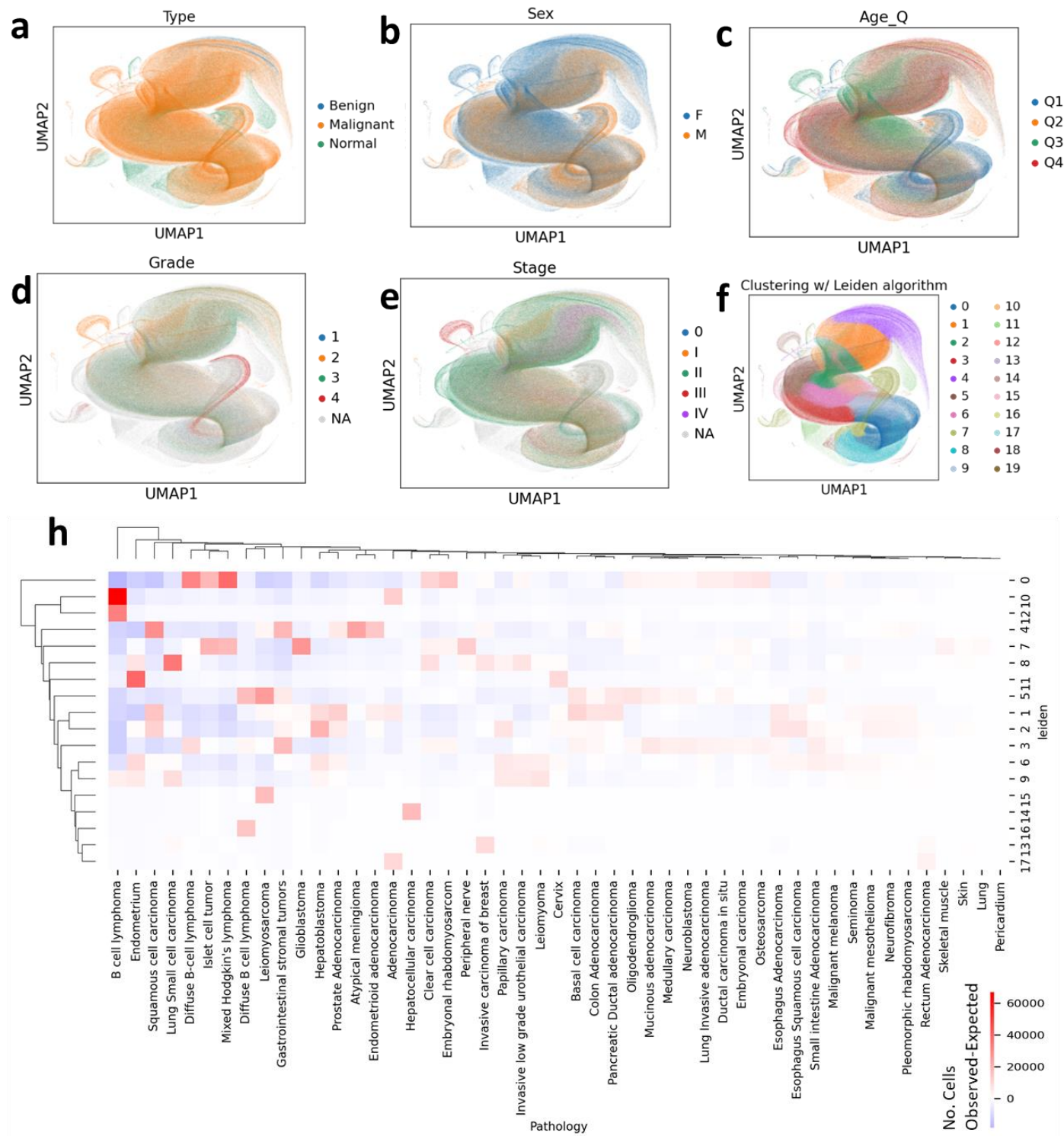
Supplementary Figure 7: Quantification of first versus second antibody application to same TMA tissue. a. Pearson Correlation of cells' mean intensity between first application and second application (Error bars S.E.M. $n=3$). b. Estimated dynamic range (dimmiest and brightest 5% of cells, $n=3$). c-d. Signal to Background ration based on thresholds. 29 of 37 antibodies had SBR above 1.5, while only 23 of 37 second application antibodies had SBR > 1.5 ($n=3$).



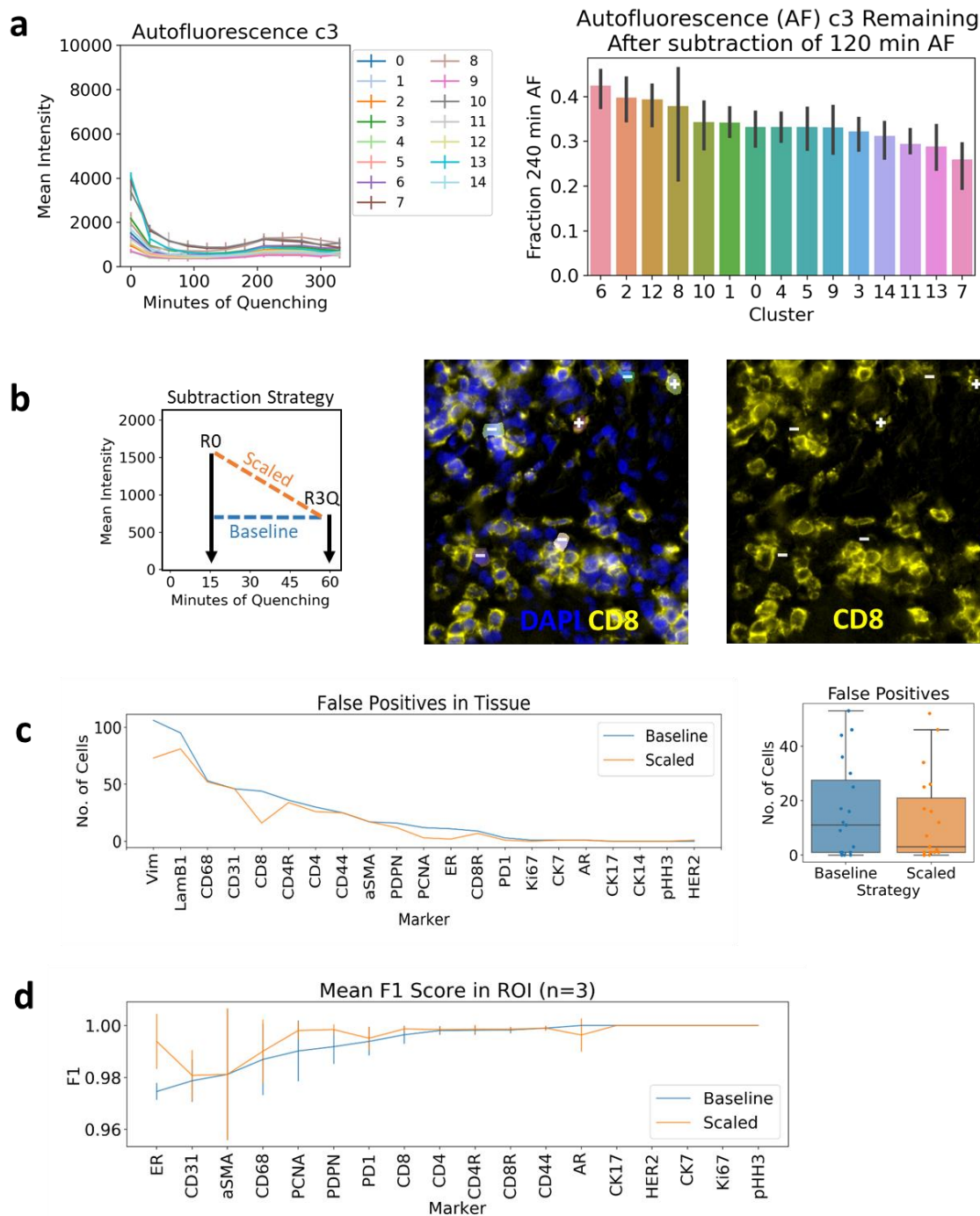
Supplementary Figure 8: Variables Impacting Round-Effect of Stain Quality and Bleed Through Observed in Miltenyi MACSima prototype Instrument. a. CD45-750 shows a certain level of non-specific nuclear accumulation (arrowhead). b. Rabbit-IgG antibody labelled with AF750 strongly accumulates in the nucleus or tumor cells (arrowhead). c. Applying CD45-750 again to the same tissue as in (a) increases nuclear background (arrowhead). d. Applying CD45-750 to the same tissue as in (b) increases nuclear background (arrowhead). e. Applying CD45-750 to a tissue previously blocked with unlabeled rabbit-IgG antibody increases nuclear background as well (arrowhead). f. Applying CD45-750 to a tissue previously quenched ten times decreases nuclear background, but also decreases intensity of CD45 staining. g. Signal-to-background quantification of different conditions in a-e. h. Bleed through example in round 1 (R1), from bright Cytokeratin-7-FITC to dim Cytokeratin-14-PE, bleed through shown with arrowhead. i-j. Panel optimization to reduce bleed through. Two bright markers, Cytokeratin-7-FITC and Calponin-PE are placed in R12 (i). Two dim markers, b-Tubulin-FITC and Cytokeratin-14-PE are placed in R10 (j). k. Histograms of single cell mean intensity before and after optimization show Cytokeratin-7 and Calponin are unchanged, and Cytokeratin-14 has reduced bleed through when paired with b-Tubulin in R10.



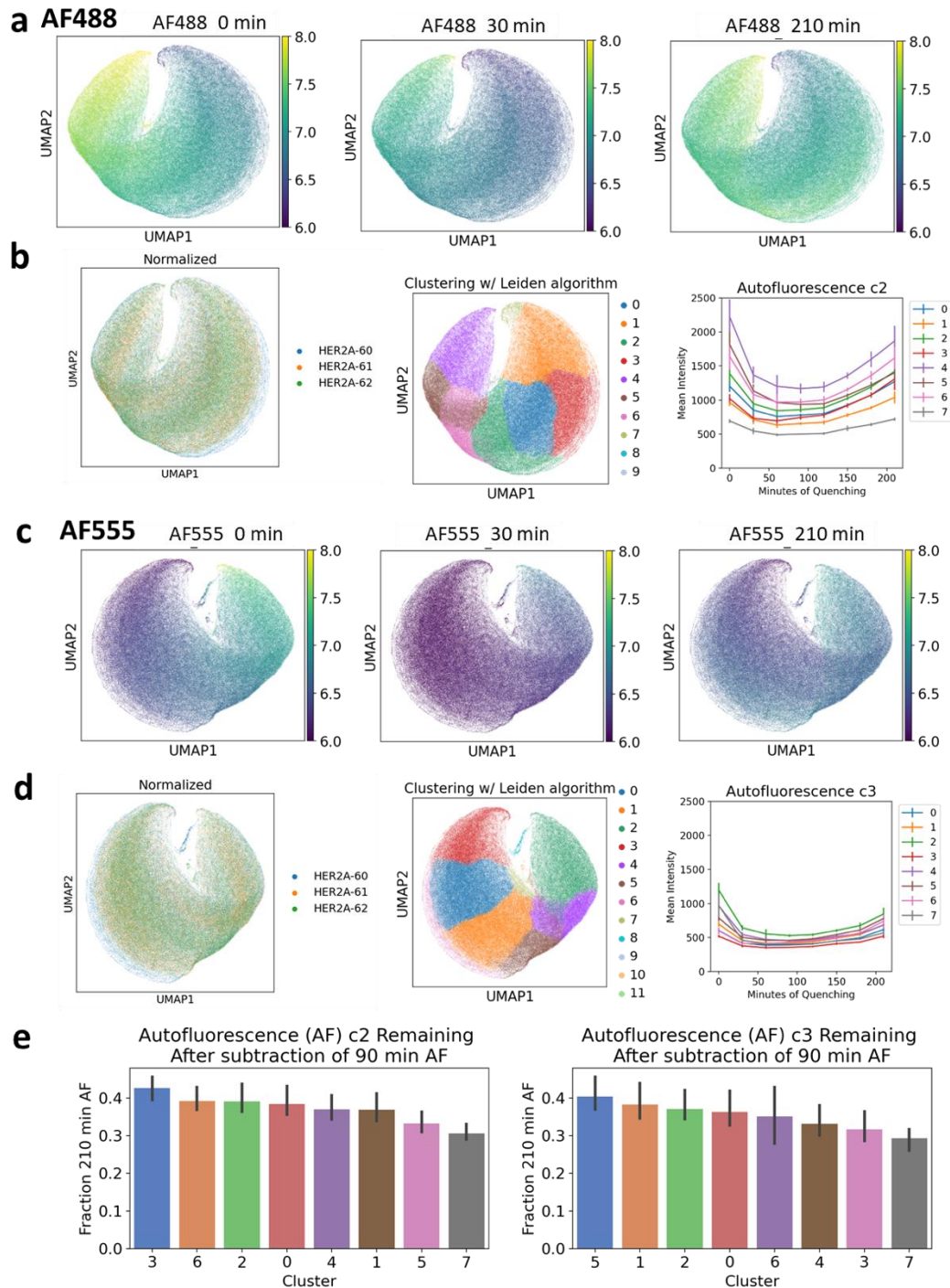
Supplementary Figure 9. Single Cell Effect of Quenching. a. UMAP of single-cell autofluorescence in the AF488 channel. b. UMAP of batch-normalized (combat algorithm) single-cell autofluorescence in the AF488 channel. c. Clustering with the Leiden Algorithm yields 10 clusters of similar intensity profiles. d. Mean Autofluorescence in each Leiden Cluster over 0 to 4 15-minute rounds of quenching, separated by subcellular localization. e. Spatial location of cells in each cluster, in each tissue (n=3). Clusters 1-7 show similar localization in all tissues, clusters 8-10 are imaging artifacts.



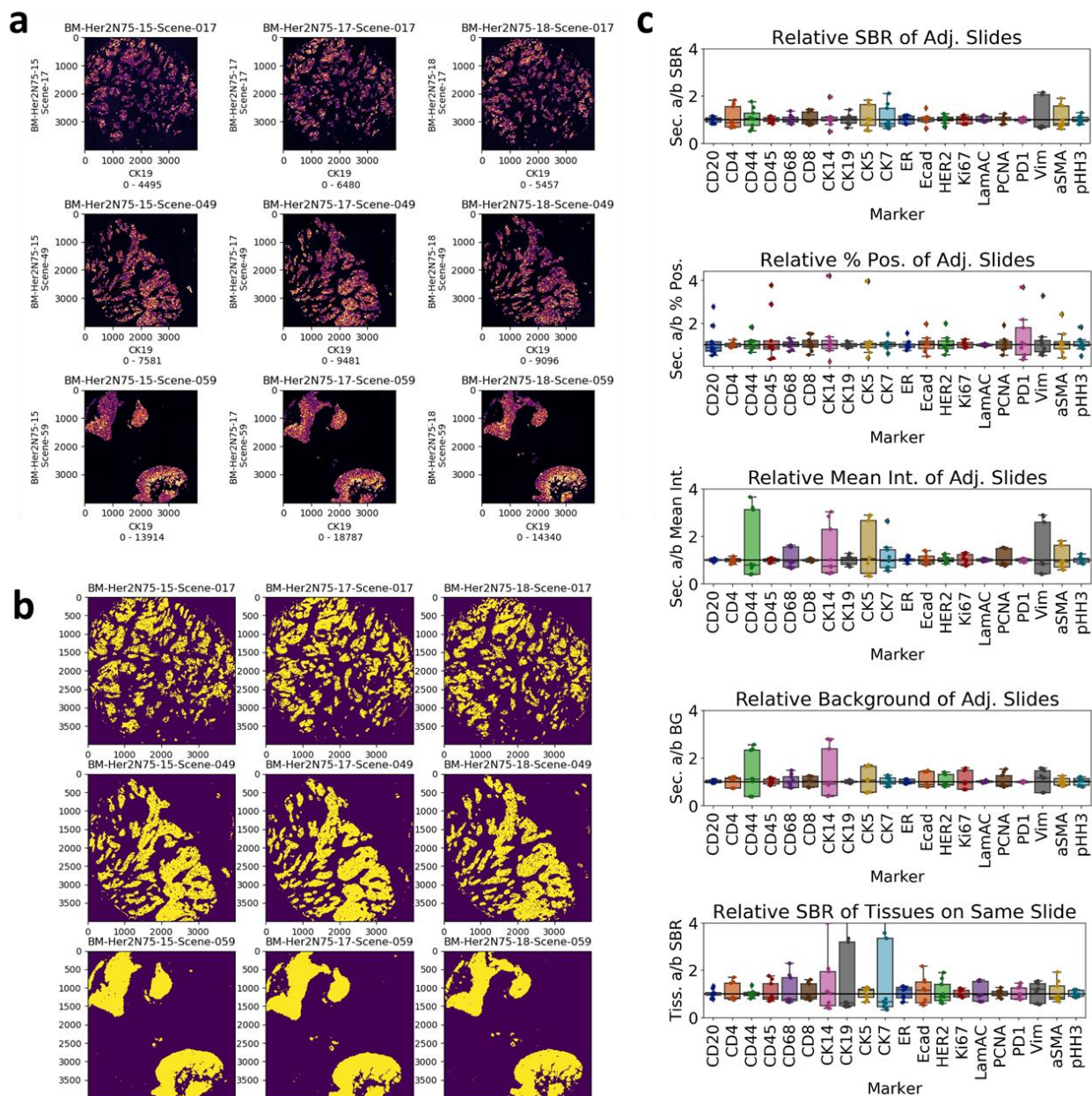
Supplementary Figure 10. Single Cell Effect of Quenching in 72-core Tissue Microarray. a-f. UMAP projection based on single-cell autofluorescence in the AF488 channel, colored by: a. Tissue type (normal, benign, malignant). b. Patient sex. c. Patient age, by quartile. d. Tumor grade. e. Tumor stage. f. Leiden cluster ID. h. Heatmap showing observed number of cells per Leiden cluster – expected number of cells per cluster, by pathology diagnosis, y-axis, illustrating different trends by tumor type.



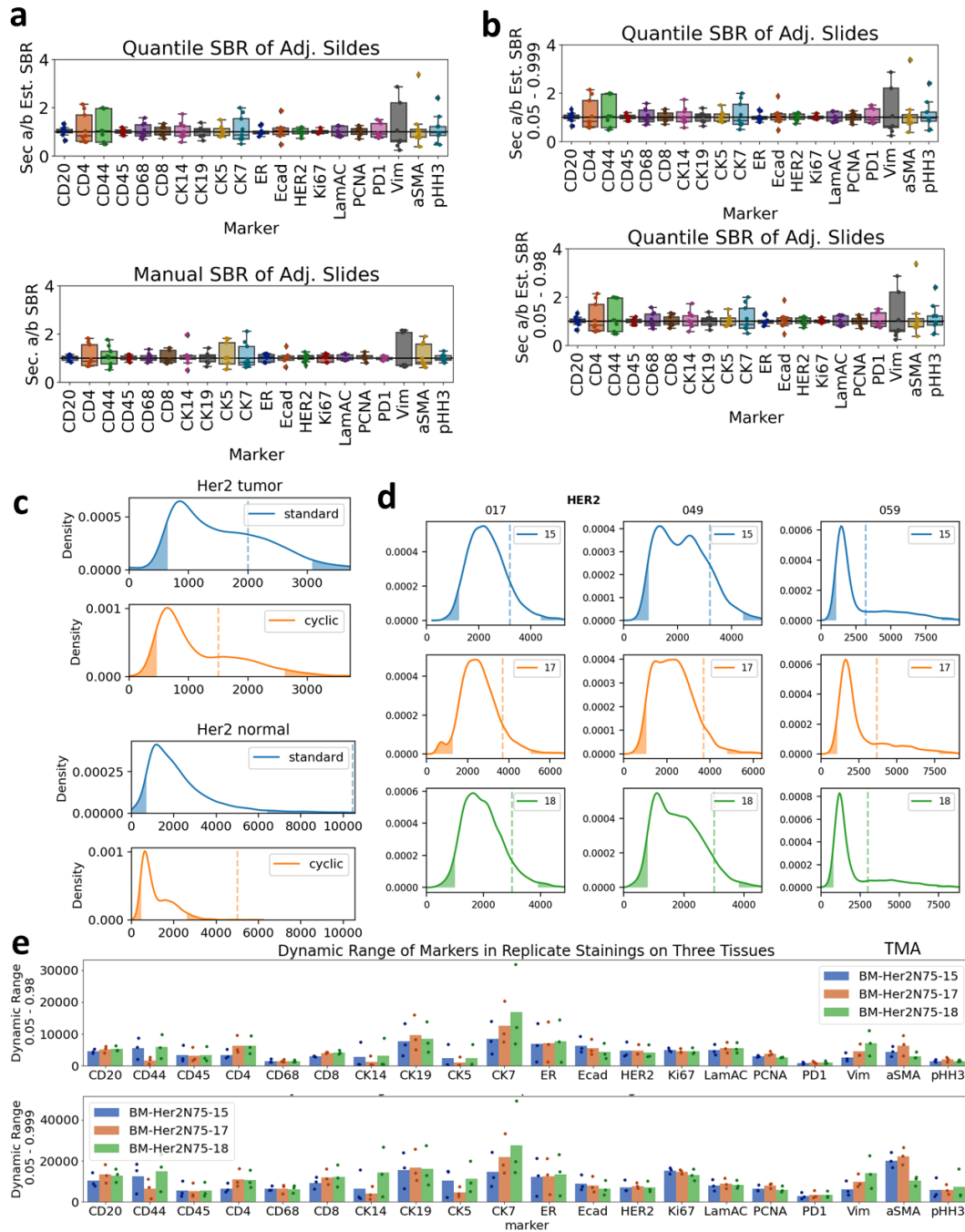
Supplementary Figure 11. Single Cell Annotation for Evaluation of Autofluorescence Subtraction. Mean AF555 intensity of cells in each Leiden cluster of 72-core TMA over rounds of quenching (left). Overall, minimum intensity was observed at 120 minutes and maximum intensity at 240 minutes. Subtracting the minimum intensity autofluorescence will avoid over-subtraction while still removing 60 - 70% of autofluorescence at 240 minutes (right). b. Schematic of autofluorescence subtraction strategies, which were applied in (c-d) to a HER2-positive breast cancer tissue (left). Example of CD8 positive/negative annotation in napari image viewer (+ and - overlaid on cell segmentation, center, and CD8 staining, right). c. Number of false positives for each marker (i.e. cells with AF488 autofluorescence > 1024). Boxplots show results of all markers, Mann-Whitney $p=0.3$. d. Based on single cell annotation in (b), F1 score (true positives)/(true pos + 0.5*(false pos. + false neg.)) of baseline versus scaled subtraction algorithms.



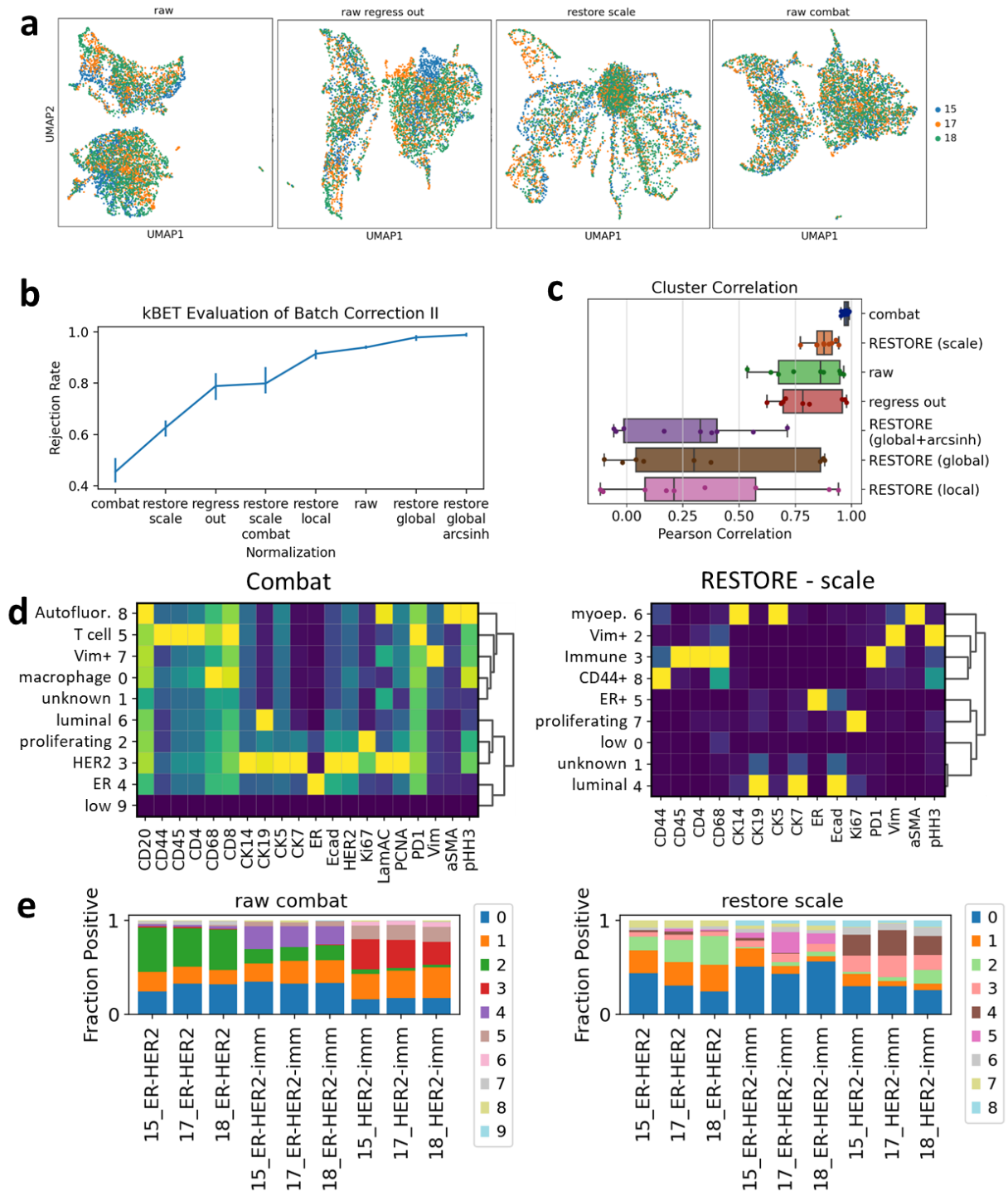
Supplementary Figure 12. Single Cell Effect of Quenching in 11-core HER2+ Tumor Tissue Microarray. a-b. UMAP projection based on single-cell autofluorescence in the AF488 channel. c-d. UMAP projection based on single-cell autofluorescence in the AF555 channel. a, c. Autofluorescence intensity at 0 (left), 30 (middle) and 210 minutes (right). b, d. UMAP colored by batch (left) - three adjacent sections from TMA were repeatedly quenched and normalized by batch for analysis; UMAP colored by unsupervised clustering results of the Leiden algorithm (middle); Mean autofluorescence intensity of cells in each Leiden cluster over rounds of quenching (right). Overall, minimum intensity was observed at 90 minutes and maximum intensity at 210 minutes. e. Subtracting the minimum intensity autofluorescence will avoid over-subtraction while still removing 60 - 70% of autofluorescence at 210 minutes. (n=3 TMAs x 10 tissues)



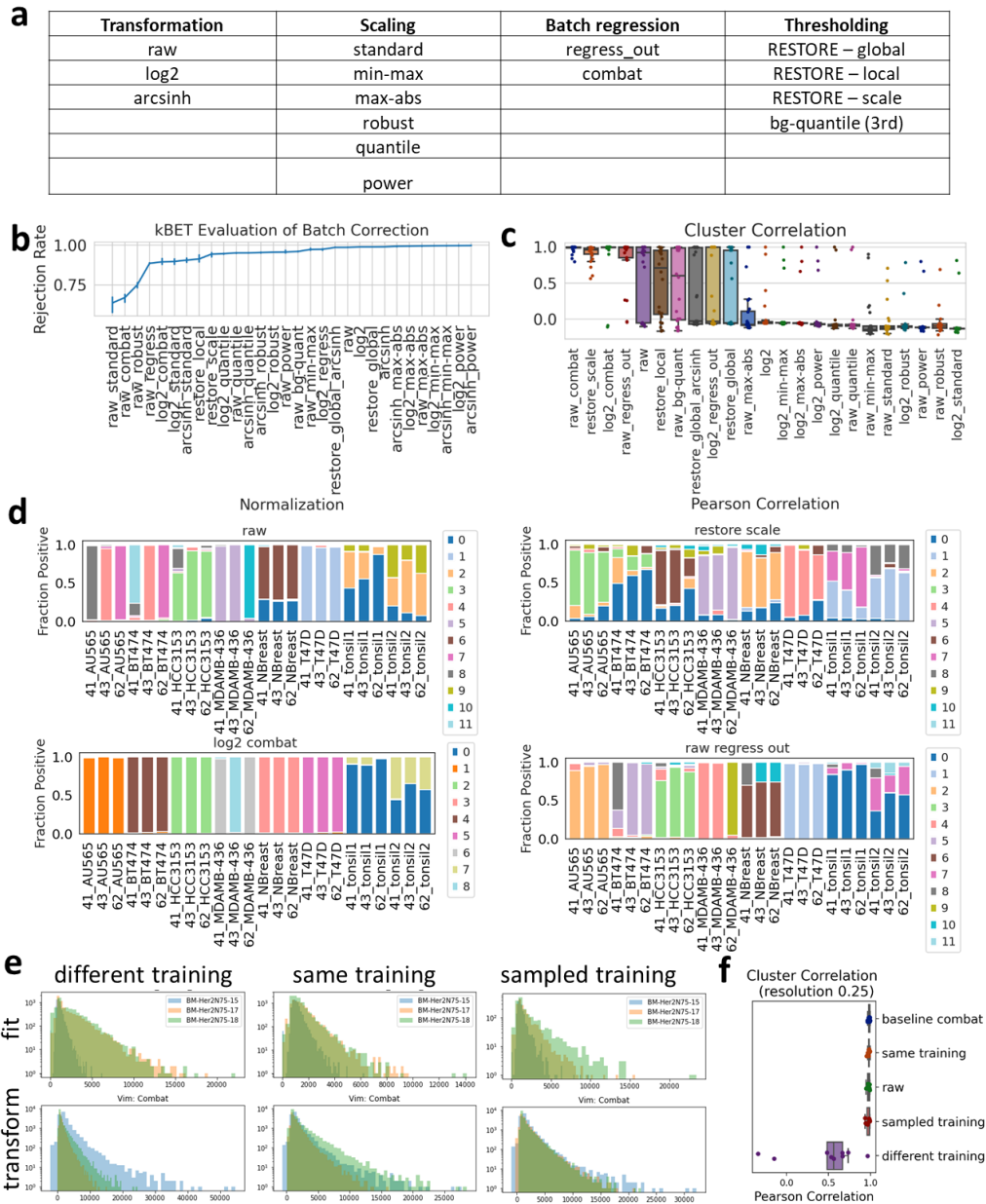
Supplementary Figure 13: Replicates of Three HER2+ Tumor TMA Cores. a. Tissue overview, marker and display range in x-axis label. b. Whole core area analyzed; yellow/blue = mask applied for signal-to-background (SBR) calculation. c. SBR, Percent Positive, Intensity, Background, and Tissue-specific SBR quantification, based on threshold in b. (Additional markers' thresholds and visualizations here: https://github.com/engjen/cyclIF_Validation)



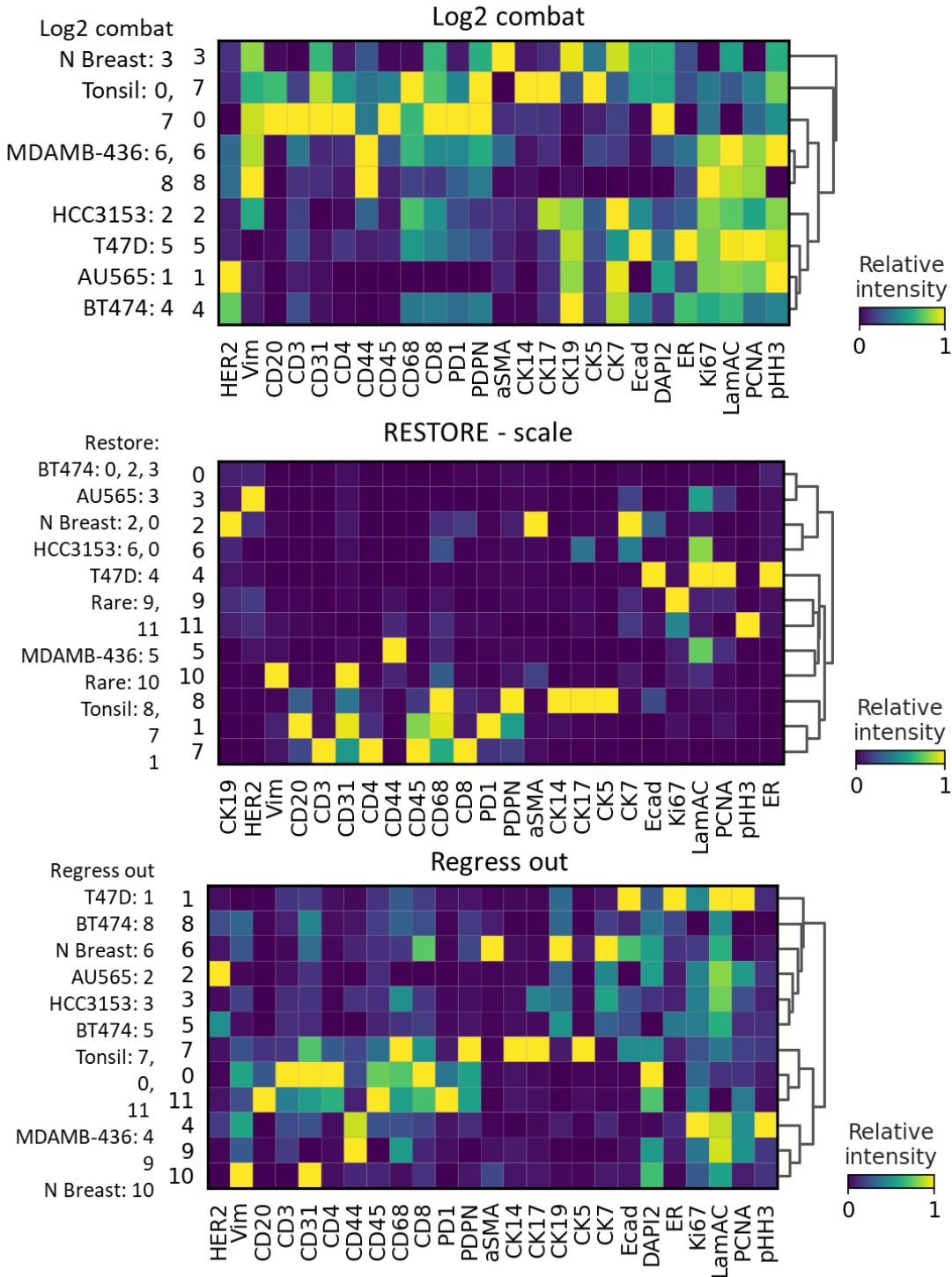
Supplementary Figure 14: Comparison of Manual Versus Automated SBR and Estimating Dynamic Range with different Quantiles. Dynamic range of each marker was estimated by the 5th and 95th quantiles of mean intensity. Manual threshold was set to identify positive cells based on staining pattern. a. Replicate cyclic experiments, estimated (“quantile”) SBR versus manual. b. We used different quantile ranges to estimate the relative SBR; top 0.05 – 0.999 and bottom, 0.05 to 0.999. The result was not different, indicating the relative quantiles between adjacent slides are stable. c-d. Histograms of single-cell intensity. Dashed line indicates manual threshold, and dotted lines are quantiles for standard versus cyclic (c) and replicate experiments (d). Manual thresholding SBR differs most from estimated SBR in markers that are negative (i.e. HER2 in normal breast, in c). e. Dynamic range of each marker in in three cores in three replicate TMAs. The dynamic range was increased to a greater degree in rare markers, e.g. aSMA and Ki67, when we increase the top quantile from 0.98 (top) to 0.999 (bottom).



Supplementary Figure 15: kBET evaluation of batch correction methods in HER2+ breast cancer tissue. a. Random samples of 1800 cells were drawn from adjacent TMA sections stained with the same cycIF panel. a. UMAP projection of raw or normalized mean intensity, colored by batch. b. kBET rejection rate (lower indicates successful batch correction) was calculated. Error bars S.E.M. (n=3). Several variations of RESTORE algorithm were tested (see methods), in addition to combat and regress out. c. Pearson correlation between replicates' cluster composition, n = 9 (3 cores x 3 replicates). d. Heatmaps of cluster mean intensity on tissue dataset. No autofluorescence (AF) subtraction was performed, and combat showed more influence of AF than RESTORE. e. Cluster composition of replicate TMA cores after combat normalization (left) and RESTORE normalization (right).



Supplementary Figure 16: kBET evaluation of batch correction methods. a. Factors tested for effect on normalization. b. Random samples of 2700 cells were drawn from adjacent TMA sections (from cell lines and normal tissue) stained with the same markers. kBET rejection rate (lower indicates successful batch correction) was calculated. Error bars S.E.M. (n=3). c. UMAP projection colored by batch, of representative batch correction strategies. c. After normalization, unsupervised Leiden clustering was applied (resolution = 0.25) and the Pearson correlation between replicates' cluster composition was calculated; n = 24 (8 cores x 3 replicates). d. Cluster composition of replicate TMA cores, left to right, top to bottom, raw data, RESTORE normalized, log2 transformed and combat normalized, and regress out normalized. The top two clusters per tissue were noted for evaluation of cluster identity. e. Evaluation of effect of training input into combat parameterization and normalization. Combat algorithm was parameterized with a different single tissue within each batch (left), the same single tissue within in batch (center), or a sample of cells from all three tissues in each batch (right). Histograms show intensity distributions used to fit the combat parameters (top) and the resulting normalized data (bottom). f. Under different combat parameterization strategies, unsupervised Leiden clustering was applied (resolution = 0.25) and the Pearson correlation between replicates' cluster composition was calculated.



Supplementary Figure 17: Normalization and cluster annotation. Unsupervised clustering of normalized data was performed with the Leiden algorithm. For each cell line and normal tissue the top two were noted for evaluation of cluster identity (annotated on left of heat maps). All showed reasonable results, e.g. AU565 should be HER2+, T47D ER+, MDAMB-436 CD44+, BT474 HER2+/ER+, HCC3153 triple negative, tonsil contains immune and basal cytokeratin cell types.

Tables

Supplementary Table 1: Double Application of Antibodies				
Round	AF488	AF555	AF647	AF750
R1	PCNA	CD8	PD1	CK19
R2	CK5	HER2	ER	CD45
R3	aSMA	CD68	CD4	Ecad
R4	Vim	AR	CD31	CD44
R5	CK7	CK14	Ki67	PgR
R6	pHH3	pS6RP	pERK	CD20
R7	CK17	EGFR	pAkt	GRNZB
R8	H3K27	PDPN	gH2AX	LamAC
R9	CK8	cPARP	pRB	FoxP3
R10	LamB1	H3K4	ColIV	ColI
R11	S100	AR*	LamB2	CD45*
R12	PCNA*	CD8*	PD1*	CK19*
R13	CK5*	HER2*	ER*	CD45**
R14	aSMA*	CD68*	CD4*	Ecad*
R15	Vim*	AR**	CD31*	CD44*
R16	CK7*	CK14*	Ki67*	PgR*
R17	pHH3*	pS6RP*	pERK*	CD20*
R18	CK17*	EGFR*	pAKT*	GRNZB*
R19	H3K27*	PDPN*	gH2AX*	ColI*
R20	CK8*	cPARP*	pRB**	LamAC*
R21	LamB1*	H3K4*	ColIV*	FoxP3*
R22	S100*	CD8**	LamB2*	CD3
* second application			** third application	

**Supplementary Table 2:
Antibody Order Optimization**

Marker	Original	Optimized
aSMA	3	3
CD4	3	7
CD8	1	7
CD20	6	8
CD31	4	10
CD44	4	3
CD45	2	6
CD68	3	6
CK5	2	4
CK7	5	6
CK8	9	8
CK14	5	1
CK17	7	9
CK19	1	1
ColI	8	9
ColIV	10	8
Ecad	3	2
ER	2	2
FoxP3	10	5
GRNZB	7	10
HER2	2	2
LamAC	11	4
LamB2	11	12
PCNA	1	2
PD1	1	6
PDPN	8	5
PgR	5	7
pS6RP	6	10

Supplementary Table 3: Control TMA

Scene	Cell Line/Tissue	Clinical Sub.	Neve et al.
10, 11	AU565	HER2 +	Luminal
8, 9	BT474	ER+/HER2+	Luminal
2	HCC1143	TN	Basal A
3	HCC3153	TN	Basal A
12, 13	MDAMB436	TN	Basal B
5, 6	T47D	ER+/PR+	Luminal
4	Nor. breast	na	na
1,7	Tonsil	na	na

Supplementary Table 4: Antibody Order in Replicate TMAs			
Marker	TMA-41	TMA-43	TMA-62
aSMA	5	3	4
CD20	7	8	6
CD3	6	11	4
CD31	6	10	6
CD4	1	7	7
CD44	3	3	3
CD45	2	6	5
CD68	5	6	6
CD8	1	7	7
CK14	3	1	1
CK17	6	9	5
CK19	1	1	1
CK5	7	4	2
CK7	4	6	6
DAPI2	2	2	2
Ecad	5	2	2
ER	2	2	2
HER2	2	2	2
Ki67	5	1	1
LamAC	4	4	10
PCNA	2	2	3
PD1	4	6	8
PDPN	6	5	5
pHH3	3	5	1
Vim	4	7	7