



MASTERARBEIT | MASTER'S THESIS

Titel | Title

ProteoPicking: bridging t-CyclIF with precision tissue
microdissection for spatial proteomics exploration of tumors
and their microenvironments

verfasst von | submitted by
Lina Maltrovsky B.Sc.

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree
programme code as it appears on the
student record sheet:

UA 066 863

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student
record sheet:

Masterstudium Biologische Chemie

Betreut von | Supervisor:

Univ.-Prof. Dr. Christian Friedrich Wilhelm Becker

Acknowledgment

I would like to start by thanking my supervisor, Univ.-Prof. Dr. Christian Friedrich Wilhelm Becker, for taking me on as a master's student and for giving me the opportunity to carry out my thesis abroad under your guidance. Your support throughout the organizational and writing process made things much smoother, and I appreciated how easy it was to communicate and work together.

I am also very thankful to my second supervisor, MD PhD Anniina Färkkilä, for offering me the chance to work in your lab and for all the support you provided during my time in Finland. I am so grateful for your guidance throughout the research and organizational aspects of my thesis. It is thanks to you that I was able to carry out this project in Helsinki, Finland, and the experience of living and working there for nearly eight months was absolutely amazing. The country, the weather, the culture, and the people made my final year of my Master's program something truly special, and none of it would have been possible without the chance and trust you gave me.

Speaking of the people I met in Finland, I would like to take this opportunity to thank all my lab colleagues from the Färkkilä Lab. From day one, you made me feel welcome, and I truly couldn't have asked for a better group of people to work with. The daily conversations, coffee breaks, and even UniCafe lunches were something I looked forward to every day. What I especially appreciated was that there was always someone to turn to, whether I had a scientific question or simply needed a friendly ear. You weren't just colleagues, but became real friends, and I feel incredibly lucky to have worked alongside such amazing people.

A special thanks goes to my co-supervisor, MSc Ada Junquera Mencia, for being a constant source of help and guidance throughout the whole process. Your expertise and input made a real difference, and I learned so much from you. I am truly grateful for the time we spent working together, and I'm happy to have had the opportunity to get to know you both in and out of the lab.

Lastly, I want to thank my family and friends, both old and new ones. You've been there for me every step of the way, and even when we were miles apart, I always felt your support and love. Whether it was during walks in the Finnish winter, long phone calls, shared writing sessions, technical support, or proofreading when you had no idea what the subject matter was, you were always there for me. Your encouragement and motivation kept me going.

Zusammenfassung

Bei hochgradig serösem Ovarialkarzinom handelt es sich um den häufigsten und tödlichsten Subtyp von Eierstockkrebs. Es zeichnet sich durch eine hohe Heterogenität, genomische Instabilität und beträchtliche Rückfallraten aus. Besonders in fortgeschrittenen Stadien sind herkömmliche Behandlungen langfristig oft wenig effektiv, was den dringenden Bedarf an neuen, gezielteren Therapieansätzen unterstreicht. Ein grundlegendes Verständnis der Mechanismen, die der Tumorentwicklung zugrunde liegen, ist dafür unerlässlich, insbesondere im Hinblick auf das Tumormikroumfeld und den Zellzyklus, die beide entscheidenden Einfluss auf Tumorprogression und Therapieerfolg haben. Multiplex-Bildgebungstechniken und räumlich aufgelöste Omics-Methoden haben sich hierfür als wertvolle Werkzeuge erwiesen. Trotz bedeutender Fortschritte bestehen jedoch weiterhin Herausforderungen in Bezug auf Skalierbarkeit, Sensitivität und Komplexität der Arbeitsabläufe, was den Bedarf an methodischer Weiterentwicklung verdeutlicht.

In dieser Masterarbeit wird mit ProteoPicking eine neue räumlich aufgelöste Proteomics-Methode entwickelt. Sie kombiniert zyklische Immunfluoreszenzfärbung mit massenspektrometrischer Proteomics-Analyse und ermöglicht die Untersuchung von Gewebeproben hinsichtlich ihrer strukturellen, funktionellen und molekularen Merkmale. Im Rahmen dieses Projekts wurde die Kompatibilität beider Methoden gezeigt. Darauf aufbauend wurde ein mehrstufiger experimenteller Arbeitsablauf zur gezielten Isolierung und Analyse spezifischer Geweberegionen entwickelt und durch die Untersuchung verschiedener Einflussfaktoren wie Gewebemenge, Regionstyp und Probenpufferwahl optimiert. Der Workflow umfasst zyklische Immunfluoreszenzfärbung, Probenahme ausgewählter Geweberegionen durch Mikrodissektion mit Mikronadel und massenspektrometrische Analyse der isolierten Zellproben. Zusätzlich wurde eine robuste Datenanalyse-Pipeline etabliert, die Bildverarbeitung, Zellphänotypisierung, Analyse der zellulären Probenzusammensetzung, Proteinexpressionsanalyse sowie Korrelationsanalyse zwischen Proteinexpression und Zellzusammensetzung vereint. Dies ermöglicht die Verknüpfung mikroskopischer Bilddaten mit zellulären Proteomics-Profilen. Insgesamt adressiert diese Arbeit zentrale technische Herausforderungen und schafft eine Grundlage für die Weiterentwicklung räumlich aufgelöster Proteomics-Methoden.

Abstract

High-grade serous ovarian carcinoma represents the most common and lethal subtype of ovarian cancer, characterized by its significant heterogeneity, genomic instability, and high recurrence rates. Standard treatments are often limited in their long-term effectiveness, particularly at advanced stages, stressing the need for novel, improved therapies. For this, a thorough understanding of the biology behind cancer progression is necessary, including key factors like the tumor microenvironment and cell cycle progression. Over the past years, highly multiplexed imaging techniques and spatially resolved omics methods have become valuable tools for studying complex, heterogeneous tissues, like tumors and their environments. But despite significant advancements, spatially resolved methodologies still face limitations in scalability, sensitivity, and workflow complexity, highlighting the ongoing need for further development and optimization in this field.

This thesis establishes ProteoPicking, a novel spatially resolved proteomics method that integrates tissue-based cyclic immunofluorescence imaging with mass spectrometry-based proteomics analysis. The approach enables the investigation of tissue samples with respect to their structural, functional, and molecular characteristics. As part of this project, the compatibility of both techniques was demonstrated. Building on this, a multi-step experimental workflow was developed for the targeted isolation and analysis of specific tissue regions. It comprises cyclic staining, region selection based on imaging results, microdissection of regions of interest using a microneedle, and subsequent mass spectrometry analysis of the collected cells. The process was optimized by evaluating key factors influencing protein yield, including sample size, tissue type, and lysis buffer used. In addition, a robust data analysis pipeline was established, combining image processing, cell phenotyping, compositional analysis of tissue samples, differential protein expression analysis, and correlation analysis between protein expression levels and sample compositions. This integrative approach allows for the combined evaluation of microscopic imaging data and spatially resolved proteome profiles. Altogether, this work addresses central technical challenges and lays the foundation for the further development of spatially resolved proteomics methods.

Contents

Acknowledgment	I
Zusammenfassung	II
Abstract	III
List of Abbreviations	VI
1 Introduction	1
1.1 Ovarian cancer	1
1.1.1 High-grade serous ovarian carcinoma	1
1.1.2 Clinical treatment	3
1.2 Tumor microenvironment	6
1.2.1 Key cell populations within the TME	7
1.2.2 TME in ovarian cancer	10
1.3 Cell cycle	11
1.3.1 Out-of-cycle states	13
1.3.2 Cell cycle and tumorigenesis	14
1.3.3 Cell cycle staging	16
1.4 Spatial resolution methodologies	17
1.4.1 Highly multiplexed imaging techniques	18
1.4.2 Proteomics	22
2 Aim of the Project	26
3 Materials & Methods	27
3.1 Materials	27
3.1.1 Reagents & instruments	27
3.1.2 Patients & specimens	29
3.2 Methods	29
3.2.1 Sample sectioning	30
3.2.2 t-CycIF staining	31
3.2.3 Image acquisition	34
3.2.4 Cell picking	36
3.2.5 Image processing, phenotyping, analysis and visualization . . .	39
3.2.6 Proteomics analysis	43
3.3 Code availability	44

4	Results	45
4.1	Pilot experiment	45
4.1.1	Experimental design	45
4.1.2	IF and t-CycIF staining enables the characterization of cell types and tissue regions	48
4.1.3	IF and t-CycIF staining patterns guide the selection of ROIs .	48
4.1.4	Proteomics results demonstrate compatibility of t-CycIF with MS-based proteomics analysis	50
4.2	Optimization experiment	53
4.2.1	Experimental design	53
4.2.2	t-CycIF staining provides tissue and cellular characterization, enabling ROI selection	56
4.2.3	Image segmentation was achieved by combining two image enhancement methods with the Mesmer nuclear segmentation model	57
4.2.4	Quantification and phenotyping enable single-cell characterization of cells within ROIs	62
4.2.5	Compositional analysis facilitates accurate sample categorization	64
4.2.6	Generated proteomics data of picked samples enables development and application of an integrated analysis pipeline combining protein expression and cellular composition	68
5	Discussion	77
	List of Figures	86
	List of Tables	89
	References	90

List of Abbreviations

ASHLAR alignment by simultaneous harmonization of layer/adjacency registration.

BaSiC background and shading correction.

CAF cancer-associated fibroblast.

CDK cyclin-dependent kinases.

CLAHE contrast limited adaptive histogram equalization.

CODEX co-detection by indexing.

CTL CD8+ T cytotoxic cell, also cytotoxic T lymphocyte.

CTLA-4 cytotoxic T-lymphocyte-associated protein 4.

D dormant.

DC dendritic cell.

DEA differential expression analysis.

ECM extracellular matrix.

FFPE formalin-fixed paraffin-embedded.

FIGO international federation of gynecology and obstetrics.

G0 quiescent phase.

G1 gap 1 phase.

G2 gap 2 phase.

H&E hematoxylin and eosin.

HGSC high-grade serous ovarian carcinoma.

HR homologous recombination.

HRD homologous recombination deficiency.

HRP homologous recombination proficiency.

HUH Helsinki University Hospital.

ICI immune checkpoint inhibitor.

IDS interval debulking surgery.

IF immunofluorescence.

IHC immunohistochemistry.

IMC imaging mass cytometry.

LC-MS/MS liquid chromatography-tandem mass spectrometry.

M phase mitosis phase.

MALDI matrix-assisted laser desorption/ionization.
MDSC myeloid-derived suppressor cell.
Mesmer mesoscopic segmentation of cellular and nuclear regions.
MIBI multiplex ion beam imaging.
MS mass spectrometry.

N non-dormant.
NACT neoadjuvant chemotherapy.
NHEJ non-homologous end joining.
NK cell natural killer cell.

OC ovarian cancer.
OIS oncogene-induced senescence.
ORR objective response rate.
OS overall survival.

P proliferating.
PARP poly(ADP-ribose)polymerase.
PBS phosphate-buffered saline.
PD-1 programmed cell death protein-1.
PD-L1/2 programmed death ligand 1/2.
PDS primary debulking surgery.
PFS progression free survival.

Rb retinoblastoma.
ROI region of interest.

S phase synthesis phase.

t-CycIF tissue-based cyclic immunofluorescence.
TAM tumor-associated macrophage.
TBS tris-buffered saline.
Th cell CD4+ T helper cell.
TIL tumor-infiltrating lymphocyte.
TME tumor microenvironment.
Treg regulatory T cell.
TSI tumor-stroma interface.

1 Introduction

1.1 Ovarian cancer

With more than 300.000 new cases and over 200.000 deaths in 2022, ovarian cancer (OC) is the 18th most common cancer worldwide and the 8th most common one for women (World Cancer Research Fund International, 2024).

OC is a histologically heterogeneous disease (Kurman & Shih, 2016; Scully et al., 1998). It can be classified into subtypes with different cells of origin, molecular compositions, risk factors, clinical features, and treatment approaches. Approximately 90 % of emerging OC is identified to arise from epithelial cells, germ cell tumors account for another 5 % of cases, and the remaining 5 % are sex cord-stromal tumors (Doubeni et al., 2016; Matulonis et al., 2016). Epithelial OC can further be divided into five histological subtypes, namely high-grade serous ovarian carcinoma (HGSC), low-grade serous ovarian carcinoma, mucinous carcinoma, clear cell carcinoma, and endometrioid carcinoma (Lheureux et al., 2019). The latter four are type I tumors according to the dualistic model of epithelial ovarian carcinogenesis. They are low-grade, slowly proliferating neoplasms that are mostly genetically stable and differ in their distinct molecular changes from type II tumors. Type II tumors are genetically unstable, aggressive, high-grade neoplasms. HGSC makes up the majority of this type (Shih & Kurman, 2004).

1.1.1 High-grade serous ovarian carcinoma

HGSC is the most common and lethal subtype of OC, accounting for 75 % of all epithelial OC cases and 70 % of all OC related deaths (Lheureux et al., 2019; Peres et al., 2019). Most cases originate from the epithelium of the distal fallopian tube and not, as long assumed, from the ovarian surface epithelium (Labidi-Galy et al., 2017). It grows in a heterogeneous way, typically as a solid mass of cells with slit-like fenestrations and frequent necrotic areas. Its structure is similar to the surface epithelium of the fallopian tube with glandular, cribriform, or papillary architecture. The cells usually display a high grade of nuclear atypia and have a high mitotic index, which indicates high proliferation (Kurman et al., 2014).

HGSC is diagnosed among all ages, but the most frequently affected are women aged 55 to 74. The 5-year relative survival rate amounts to 50 %. However, it must be taken into account that this rate greatly differs between different stages of HGSC. There are four stages of OC based on the progress of the disease and defined by the international federation of gynecology and obstetrics (FIGO). In early stages, the tumor is limited to the ovaries (FIGO stage I) or involves one or both ovaries and

has spread within the pelvic region or originates as primary peritoneal cancer (FIGO stage II). More advanced stages are defined by an additional spread of the tumor beyond the pelvis to the peritoneum and/or the retroperitoneal lymph nodes (FIGO stage III) or even a spread to more distant sites in the body (FIGO stage IV) (Prat, on Gynecologic Oncology et al., 2014). While over 90 % of women with cancer in early stages survive more than five years after their diagnoses, the survival rate for distant stage OC, FIGO stage IV, is only around 30 %. Moreover, the majority of OC cases is diagnosed at advanced stages, FIGO stages III and IV (National Cancer Institute, Surveillance, Epidemiology, and End Results Program, 2024). This is explained by the fact that there is no effective screening strategy for the early detection of OC and early stages of this disease are often asymptomatic. If there are any symptoms they are usually non-specific for this type of cancer. Possible symptoms include back pain, fatigue, abdominal pain, abdominal bloating, constipation, or urinary symptoms (Goff et al., 2004; Matulonis et al., 2016).

HGSC is a type II epithelial ovarian carcinoma. It displays a high degree of genetic instability with two characterizing aspects - whole genome duplications and high-frequency somatic copy number alterations. Additionally, genetic mutations in *TP53* typically occur. Beside this, there are only a few other recurrent genetic mutations in HGSC. A study conducted by the Cancer Genome Atlas Research Network in 2011 (Cancer Genome Atlas Research Network and others, 2011) showed that in more than 96 % of the analyzed samples a *TP53* mutation was present. Further recurrent mutations, even though less common than *TP53*, were germline and somatic *BRCA1* and *BRCA2* mutations in 22 % of the samples and mutations in *RB1*, *NF1*, *FAT3*, *CSMD3*, *GABRA6*, or *CDK12* in less than 10 %. Moreover, the study revealed that in more than half of the analyzed HGSC samples the homologous recombination (HR) pathway of DNA repair was defective. This can be associated with germline and somatic *BRCA* mutations. HR repair is a key mechanism for HGSC to fix DNA double-strand breaks (Cancer Genome Atlas Research Network and others, 2011). It uses homologous DNA sequences as a template to accurately recreate the former sequence and fix the double-strand break. Such breaks can derive from exogenous anticancer treatments like topoisomerase inhibition, platinum-based drugs, or ionizing radiation and if they are not correctly repaired they can lead to mutations, cell cycle arrest, or apoptosis. Besides HR, there are other pathways in cells that can be alternatively used to repair DNA double-strand breaks. One example would be non-homologous end joining (NHEJ). NHEJ is the direct ligation of blunt, complementary, and non-complementary ends after certain enzymatic modifications without a homo-

logous sequence as a template. During this process small scale mutations at the repair site can occur, which is why NHEJ is more error-prone than HR (Pfeiffer et al., 2000). Hence, a HR deficiency, which hinders cells from repairing DNA damage induced by cancer therapy, strongly influences the treatment sensitivity of HGSC. Other commonly altered pathways are RB1 (67 %), RAS/PI3K (45 %), FOXM1 (87 %), and NOTCH (22 %)(Cancer Genome Atlas Research Network and others, 2011).

1.1.2 Clinical treatment

For the primary treatment of HGSC after diagnosis there are two different approaches - primary debulking surgery (PDS) with adjuvant chemotherapy or neoadjuvant chemotherapy (NACT) followed by interval debulking surgery (IDS) (National Comprehensive Cancer Network, 2024). The treatment strategy is determined by the stage of cancer, the resectability of the tumor, and patient variables such as the patient's age or the general health condition (Fagotti et al., 2020; Van Meurs et al., 2013; I. Vergote et al., 2018).

If the patient's tumor is deemed resectable, that means complete cytoreduction can be achieved, and the patient is in suitable condition for surgery, primary debulking surgery (PDS) is the preferred approach (I. B. Vergote et al., 2016). Afterwards, adjuvant chemotherapy is recommended for all stages of HGSC (National Comprehensive Cancer Network, 2024). In practice, six cycles of platinum-based adjuvant therapy are administered. The preferred regimen is intravenous administration of a combination of carboplatin with paclitaxel every three weeks. In cases with high risk of progression (FIGO stage III patients with R2 resection and inoperable FIGO stage III and IV patients) bevacizumab is sometimes added to the therapy with carboplatin and paclitaxel and afterwards used for maintenance therapy (Burger et al., 2011; National Comprehensive Cancer Network, 2024; Perren et al., 2011).

When patients are too ill for surgery or when their cancer burden is too extensive for a macroscopic complete resection, typically in advanced disease stages (FIGO stages III and IV), NACT and IDS are recommended instead of PDS. By this, the cancer burden is reduced before surgery is performed (Gill et al., 2017). A combination of carboplatin and paclitaxel is administered for three cycles before IDS is performed. Afterwards, six cycles of post-operative chemotherapy are recommended (National Comprehensive Cancer Network, 2024).

Both surgical approaches aim for complete cytoreduction by removing the complete disseminated tumor mass from the peritoneal cavity of the patient. Successful resection occurs when after surgery there is no macroscopic residual of the diseased tissue left and is classified as complete resection (R0 resection). However, in practice often

some tissue parts remain and the resection is classified according to the postoperative amount of tumor. When the diameter of the residual cancer is less than 1 cm, it is an optimal resection (R1 resection). Otherwise, it counts as a suboptimal resection (R2 resection)(Matulonis et al., 2016). The achieved level of resection after PDS or IDS is the most important prognostic factor concerning patient survival (Du Bois et al., 2009).

The majority of patients initially has a favorable response to platinum- and taxane-based chemotherapies, as shown in several clinical studies (McGuire et al., 1996; Ozols et al., 2003; Piccart et al., 2000). However, recurrence rates remain high, affecting about 75 % of patients, with 23 % experiencing relapse within six months following the completion of primary therapy (Du Bois et al., 2009). This demonstrates how treatment remains challenging with standard approaches being ineffective in most cases in the long term, particularly at advanced stages. Consequently, there is an urgent need for investigating novel, improved therapy strategies for HGSC.

In recent years, alternative treatment approaches such as targeted therapy and immunotherapy have been and continue to be increasingly investigated (Wilson et al., 2024).

One promising group of drugs are so called PARP inhibitors, small-molecular inhibitors that target poly(ADP-Ribose)polymerase (PARP). PARP is especially vital for cancer cells with defective DNA repair mechanisms, for example HR deficiencies in HGSC patients due to mutations in *BRCA1/2* or other genes associated with HR deficiency like several *RAD* or *FA* genes. Those cells heavily rely on PARP mediated base excision repair of single-strand breaks to deal with DNA damage. Hence, inhibition of PARP inactivates DNA repair, resulting in synthetic lethality and anti-tumor activity (Ashworth et al., 2011; Cancer Genome Atlas Research Network and others, 2011; Eustermann et al., 2015; Farmer et al., 2005). Based on this mechanism the first PARP inhibitor finding application in clinics was olaparib. It is used as a treatment for patients with recurrent cancer who have initially responded to prior platinum-based chemotherapy and for newly diagnosed patients with advanced-stage HR-deficient tumors who have experienced tumor reduction or clearance following platinum-based chemotherapy (European Medicines Agency, 2025). The effect of olaparib differs in patients as demonstrated in a phase II clinical trial with a diverse patient cohort (Gelmon et al., 2011). Of the 63 enrolled ovarian cancer patients especially patients with *BRCA* mutations exhibited impressive partial responses, but also a proportion of patients with *BRCA* wildtype displayed significant responses to the

treatment with olaparib. A second factor that influences the response to olaparib is the platinum resistance status of patients. The results of the trial demonstrated that while the response rate in platinum sensitive patients was around 50% for *BRCA* wildtype patients and 60% for *BRCA* mutation patients, it was significantly lower in platinum resistant patients with around 5% respectively 30%. Other examples of PARP inhibitors that are used for treating HGSC are rucaparib and niraparib (Mirza et al., 2016; Syed, 2017).

Another routinely used targeted therapy uses anti-angiogenic agents to interfere with the development of new blood vessels within the tumor. This process, called angiogenesis, is essential for tumor growth, prevents hypoxia and necrosis, and promotes invasion, metastasis, and ascites. By inhibition of angiogenesis tumor growth is restrained (Hanahan & Weinberg, 2000). One example for such an agent is the humanized monoclonal antibody bevacizumab. It targets the cytokine VEGF-A, an important growth factor responsible for the recruitment of blood vessels during angiogenesis (Burger, 2011). Bevacizumab, as already mentioned above, sometimes is added to the standard carboplatin-paclitaxel regimen and used as maintenance therapy to increase progression free survival (PFS) as shown in the ICON7 and GOG218 clinical trials (Burger et al., 2011; Oza et al., 2015). Apart from bevacizumab, there are also other anti-angiogenic agents such as pazopanib, nintedanib, or cediranib (Du Bois et al., 2013; Ledermann et al., 2013).

Also several immunotherapeutic approaches are actively investigated (Wilson et al., 2024). HGSC is an immunogenic tumor, that means it can cause an immune response, and there are implications that patients could be responsive to immunotherapy. The patients immune system is able to identify the tumor, for example through the recognition of autologous tumor-associated antigens by T cells (L. Zhang et al., 2003). Despite this, previous clinical trials on various immunotherapeutic agents have largely demonstrated limited response rates (Lheureux et al., 2019).

One investigated type of immunotherapies are immune checkpoint inhibitors (ICIs). ICIs are a class of drugs designed to enhance the immune system's ability to target cancer cells by blocking specific inhibitory pathways known as checkpoints. Those checkpoints regulate immune responses, prevent excessive immune activation, and protect against autoimmunity. However, tumor cells can exploit these mechanisms to evade immune detection. By inhibiting immune checkpoints, the immune system activity against cancer cells can be restored. Two examples for immune checkpoints are the T cell inhibitory receptors cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) and programmed cell death protein-1 (PD-1) receptor. CTLA-4 downregu-

lates the activation of T cells, while PD-1, upon binding to its ligands PD-L1/2 on tumor cells, suppresses immune activity, allowing tumors to evade detection (Pardoll, 2012). ICIs, including ipilimumab (National Cancer Institute (NCI), 2020), which targets CTLA-4 as well as nivolumab (Hamanishi et al., 2015) and pembrolizumab (Matulonis et al., 2019), both targeting PD-1, have been evaluated in clinical trials for advanced recurrent ovarian cancer. The outcomes of these studies indicate modest response rates. Nivolumab achieved an objective response rate (ORR) of 15 %, while pembrolizumab demonstrated a lower ORR of 8 %. Similarly, the trial investigating ipilimumab reported an ORR of approximately 10 %. These results suggest that while checkpoint inhibitors show some clinical activity, their efficacy as therapy in recurrent ovarian cancer remains limited.

Overall, the challenges in effectively treating HGSC highlight the critical need for further research in this field. Advancing our understanding of HGSC’s mechanisms and molecular characteristics is essential for identifying new biomarkers as potential therapeutic targets, and developing new, more effective therapies capable of addressing advanced disease, therapy resistance, and tumor recurrence.

1.2 Tumor microenvironment

The tumor microenvironment (TME) is a highly heterogeneous, complex, and dynamic ecosystem, consisting of a wide variety of tumor, stroma, and immune cells, along with various regulatory signaling factors such as chemokines, cytokines, and growth factors. It evolves together with the tumor, contributing to tumor heterogeneity by fostering the development of diverse cancer cell clones. The TME plays a crucial role in tumor survival, progression, metastasis, and immune evasion, either promoting or suppressing the immune response through various cellular interactions (Hanahan & Weinberg, 2011; Jia et al., 2022). Additionally, it can significantly influence a patient’s response to treatment (Wu & Dai, 2017). Due to its heavy influence on tumor development, the TME is an interesting and important target in the investigation of novel, effective treatment strategies.

Within the TME, stroma cells such as cancer-associated fibroblasts (CAFs), angiogenic vascular cells, endothelial cells, and pericytes are integral for tumor progression. They do not only provide structural support, but also contribute to the promotion of several hallmarks of cancer (Hanahan & Weinberg, 2011). This is accomplished through their secretion of signaling factors, extracellular matrix (ECM) components, and growth factors, which influence the TME. In doing so, they help fa-

cilitate immune evasion, disrupt growth suppression, enable invasion and metastasis, and alter the metabolic pathways of tumor cells (Hanahan & Coussens, 2012).

The immune cells in the TME are composed of cells of the innate immune system, like dendritic cells (DCs), natural killer cells (NK cells), macrophages, and myeloid-derived suppressor cells (MDSCs) as well as the adaptive immune system, including CD4+ T cells, CD8+ T cells, regulatory T cells (Tregs), and B cells. By displaying their different cellular functions and interacting with each other and other parts of the TME they either contribute to a pro-tumor activity by supporting tumor growth and immune evasion, or an anti-tumor immune response (Giraldo et al., 2019).

DCs capture and present tumor-associated antigens to T cells, triggering an immune response. NK cells are essential for identifying and directly killing tumor cells through their cytotoxic activity, recognizing cells that display stress markers or altered self-antigens. CD4+ T helper cells activate other immune cells, including CD8+ cytotoxic T cells, which directly target and kill tumor cells, or B cells, which contribute to the immune defense by producing antibodies that bind to tumor antigens, aiding in the recognition and elimination of cancer cells.

In contrast, Tregs play a pivotal role in maintaining immune tolerance by suppressing the activity of other immune cells, thus allowing tumors to evade detection and destruction. MDSCs inhibit T cell activity and promote an immunosuppressive environment beneficial to tumor growth. Lastly, M2 macrophages also promote tumor progression by supporting tissue remodeling, angiogenesis, and immune suppression through the secretion of anti-inflammatory cytokines (Giraldo et al., 2019; Hanahan & Weinberg, 2011).

Given the central role these diverse cell populations play in shaping the TME, it is essential to understand their individual functions and interactions to develop more targeted and effective therapies for tumor treatment.

1.2.1 Key cell populations within the TME

Cancer-associated fibroblasts

CAFs are fibroblasts within the stroma region of the TME. They primarily differentiate from mesenchymal-derived cells, but can also transform from other cell types. They produce various proteins, including most components of the extracellular matrix (ECM) like collagens, fibronectin, and laminin as well as ECM regulating proteins (Kalluri & Zeisberg, 2006). CAFs support tumor progression via several mechanisms,

for example by expressing chemokine CXCL14, which plays an important role in promoting tumor growth, and the expression of the migration and invasion enhancing fibroblast activation protein α (Sjöberg et al., 2016; W. Yang et al., 2013). Moreover, they support the infiltration of Tregs into the TME, which leads to immune inhibition, and release angiogenesis promoting molecules (Givel et al., 2018).

Various studies suggest that CAFs in OC can serve as biomarkers for poor prognosis (Givel et al., 2018; W. Yang et al., 2013; Zhao et al., 2017). One study conducted in HGSC showed that a high accumulation of a specific CAF subtype, CAF-S1, is a hallmark of a dismal prognosis (Givel et al., 2018). Furthermore, CAFs contribute to increased resistance to platinum-based treatments and expedite tumor recurrence by expressing fibroblast activation protein α . The overexpression of this protein in the tumor stroma is one of the hallmarks of platinum resistance and predicts shorter recurrence in patients, as shown in a study which analyzed over 300 epithelial OC sample (Mhaweche-Fauceglia et al., 2015). All in all, CAFs play a critical role in ovarian cancer progression, immune evasion, and treatment resistance with their presence and subtypes serving as key prognostic indicators of poor outcomes.

Tumor-associated macrophages

Macrophages are innate immune cells involved in inflammation and tumorigenesis. In the TME, they are referred to as tumor-associated macrophages (TAMs), which can either be tissue-resident macrophages or infiltrating macrophages derived from monocytes originating in the bone marrow. In general, there are two main phenotypes of TAMs: M1 and M2 macrophages. However, these phenotypes represent the extremes of a continuous spectrum, with TAM subpopulations displaying varying degrees of M1 or M2 characteristics, or even sharing features of both. M1-like macrophages develop from monocytes stimulated with interferon-gamma $\text{IFN-}\gamma$, granulocyte-macrophage-colony-stimulating factor, or bacterial lipopolysaccharide. They possess anti-tumor functions, including cytotoxicity, tumor-suppression, and immune-stimulation. In contrast, M2-like macrophages differentiate from monocytes stimulated with tumor expressed cytokines, like IL-4, IL-10, or IL-13. They are pro-tumor macrophages (Galdiero et al., 2013; Grivennikov et al., 2010; Sica et al., 2008). Through remodeling and regulation of the ECM by degrading ECM components, M2 macrophages facilitate tumor migration, invasion, and metastasis (Finkernagel et al., 2016). Moreover, they contribute to immune evasion and angiogenesis via the expression of various immunosuppressive and pro-angiogenic cytokines as well as VEGF-A, an important angiogenic factor (Sica et al., 2008).

In OC, TAMs of mixed-polarization are present (Reinartz et al., 2014). As tumor and TME evolve, changes in TAM infiltration and polarization occur. Studies have shown that higher TAM and M2 macrophage density correlates with more advanced cancer stages and poorer prognosis. Additionally, M1/M2 ratio has been identified as a prognostic factor for PFS and overall survival(OS) in epithelial OC, with higher ratios associated with extended survival (M. Zhang et al., 2014). In conclusion, TAMs and particularly their polarization states play a crucial role in shaping the TME of OC, influencing tumor progression, immune evasion, and patient prognosis.

Tumor-infiltrating lymphocytes

Tumor-infiltrating lymphocytes (TILs) are a major component of the TME and comprise of B, T, and innate lymphoid cells. Among these, T lymphocytes are the main cell type acting in adaptive cancer immunity.

T cells develop in the thymus to naive T cells. Only upon encountering antigens, the naive cells are activated and start differentiating, proliferating, and maturing. Depending on the cytokine environment, they differentiate into various subtypes with different functions within the TME. Mature T lymphocytes can be categorized into three main types: CD4+ T helper cells (Th cells), CD8+ T cytotoxic cells (also cytotoxic T lymphocytes (CTLs)), and Tregs.

Th cells regulate immune responses by releasing cytokines that can either promote or inhibit inflammation. Based on this, they are classified into pro-inflammatory Th1 cells, which help CTLs in their anti-tumor activity, or anti-inflammatory Th2 cells, which promote tumor progression. CTLs exhibit tumor-suppressive activity. Like Th1 cells, they secrete inflammatory cytokines, but in addition they express cell lytic molecules like perforin and granzymes to destroy infected or malignant cells after recognizing them. Tregs are cells whose primary function is to prevent autoimmune diseases by inhibiting self-reactive and auto-reactive cells. But within the TME they act as immunosuppressors, promoting tumorigenesis through mechanisms such as the secretion of immunosuppressive molecules, cytotoxicity of effector cells, metabolic disruption, and modulation of dendritic cell maturation (Chen & Mellman, 2013; Y. Yang et al., 2020).

In OC, the distribution and localization of TILs play a crucial role in determining clinical outcomes. A study analyzing over 180 tissue samples from advanced OC stages found that patients with T cell-infiltrated tumors had a higher 5-year OS rate compared to those with no T cell infiltration. This study also suggested a significant association between intratumoral T cells and delayed relapse (L. Zhang et al., 2003).

Several studies support these findings, showing that patients with CTL-infiltrated tumors experienced extended PFS and OS ([hamanishi2007programd](#); Clarke et al., 2009; Sato et al., 2005). In contrast, high Treg infiltration was associated with poor clinical outcomes, as demonstrated by Curiel et al. (Curiel et al., 2004). Cancer immunity phenotypes are typically categorized into three major types based on the extent of TIL presence: immune-desert, immune-excluded, and inflamed tumors (Chen & Mellman, 2013). While immune-deserted tumors lack TILs, and immune-excluded tumors only feature immune cells at the stroma-epithelial interface but not within the tumor, inflamed tumors display significant TIL infiltration. These immune profiles have also been observed in HGSC, where distinct immune compositions have been identified (Kandalaft et al., 2022).

All in all, TILs, particularly their distribution and polarization within the TME, play a crucial role in shaping ovarian cancer progression and patient prognosis, with immune profiles offering valuable insights into disease outcomes.

1.2.2 TME in ovarian cancer

In line with the unique heterogeneity of HGSC, the TME of this cancer type is also highly diverse and intrinsically heterogeneous (Jiménez-Sánchez et al., 2020). Moreover, it can be dynamically modulated by patient treatment such as chemotherapy (Khairallah et al., 2018). While commonly used chemotherapy strategies induce response-specific changes within the TME, the precise mechanisms driving these alterations are still not completely understood. Although chemotherapies do not induce immunogenic cell death, they still promote inflammation and anti-tumor responses through various pro-inflammatory processes. These include an increase in MHC I expression on antigen-presenting cells, depletion of Tregs, polarization of TAMs toward the M1 phenotype, and enhanced tumor permeability, which facilitates the infiltration of TILs (Böhm et al., 2016).

A study conducted by Jiménez-Sánchez et al. investigated the TME in metastatic HGSC before and after NACT. Their findings revealed the coexistence of immune-suppressive, immune cell-excluded regions alongside immuno-activated, inflammatory regions within the same tumor site of treatment-naïve HGSC. Furthermore, the study demonstrated that NACT led to increased immune cell infiltration, enhancing the tumor’s immunogenicity, and indicating that chemotherapy locally activates the immune system (Jiménez-Sánchez et al., 2020).

Interestingly, the post-NACT immune landscape and its balance between pro-tumor and anti-tumor immune cell populations have been shown to correlate with patient outcomes. Specifically, the presence of stromal TILs offers the most predictive insights

(Leary et al., 2021; Mesnage et al., 2017). These findings underscore the critical role that immune cell composition plays in predicting treatment response and survival.

While significant progress has been made, our understanding of the interplay between treatment and immune cell populations remains far from complete. Continued research is essential to fully unravel the complex dynamics of the TME and its impact on tumor behavior and treatment response.

1.3 Cell cycle

The cell cycle is a complex and tightly regulated series of events that result in cell growth, the duplication of genetic information, and the subsequent distribution of the duplicated genome between two daughter cells. This process is crucial for proliferation, maintaining genomic integrity, and ensuring accurate cell division (Israels & Israels, 2000; Williams & Stoeber, 2012).

Generally, it can be divided into four distinct phases: gap 1 phase (G1), synthesis phase (S phase), gap 2 phase (G2), and mitosis phase (M phase). During G1, the cell grows in size and prepares for the following phase by expressing genes that are necessary for DNA replication. In S phase, the cell duplicates its DNA. In G2, the cell continues to grow and synthesize components essential for cell division. At this stage, the cell also performs a final check to ensure that DNA replication is complete and error-free. Finally, M phase involves the separation of the replicated genome into two equal parts, followed by cytokinesis, which divides the cytoplasm and completes the formation of two daughter cells (Israels & Israels, 2000; Ligasová et al., 2023).

The progression through these phases must be tightly regulated to ensure proper cell division. Key regulators of the cell cycle are cyclin-dependent kinases (CDKs), cyclins, and CDK inhibitors. These molecules work together to control cell cycle progression (Ligasová et al., 2023; Williams & Stoeber, 2012). CDKs and cyclins form complexes, in which CDKs serve as the catalytic unit, while cyclins act as the regulatory unit. Upon activation through phosphorylation, these complexes themselves phosphorylate specific target proteins, driving the cell cycle forward. Different cyclin-CDK pairs are active at distinct stages of the cell cycle, as the expression levels of cyclins vary during different phases - cyclin levels increase or decrease to regulate the timing of cell cycle transitions (Israels & Israels, 2000). Additionally, CDK activity is controlled by CDK inhibitors, which can bind to and inhibit CDK-cyclin complexes, ensuring the cell cycle progresses only under appropriate conditions (Besson et al., 2008).

Alongside the regulation provided by cyclin-CDK complexes, the cell cycle is further controlled by cell cycle checkpoints. These checkpoints serve as regulatory mechanisms that ensure the cycle halts if DNA is damaged or if cells fail to meet necessary conditions for division. Through this, it is guaranteed that only healthy, fully prepared cells proceed through each phase, preventing the transmission of mutations to daughter cells (Ligasová et al., 2023; Williams & Stoeber, 2012).

One of the most important checkpoints in the cell cycle is the restriction point in G1, which is regulated by the retinoblastoma (Rb) protein. Rb functions as a gatekeeper, controlling the progression from G1 to S phase. In its unphosphorylated form, Rb tightly binds to the E2F transcription factor, inhibiting its activity and preventing the transcription of genes required for DNA replication. If conditions are favorable, such as in the absence of genomic damage, Rb is phosphorylated by cyclin D-CDK4/6 complexes, causing it to dissociate from E2F. This release allows E2F to activate the transcription of genes necessary for DNA synthesis, facilitating progression past the restriction point into S phase. In the presence of genomic damage, however, Rb remains unphosphorylated, preventing the activation of E2F and thereby leading to a delay or abandonment of the cell cycle to allow for DNA repair or apoptosis. Notably, Rb remains phosphorylated throughout S, G2, and M phase, and is only dephosphorylated upon completion of mitosis, ensuring that the cell cycle does not proceed prematurely (Henley & Dick, 2012; Israels & Israels, 2000).

Another critical checkpoint in the cell cycle is regulated by the tumor suppressor protein p53. It is a transcription factor that continuously monitors the cellular genome integrity. In response to DNA damage, it can initiate cell cycle arrest, facilitate DNA repair, or in cases of severe damage promote apoptosis. Thereby, it helps to maintain genomic stability. When DNA damage is detected, p53 binds to sequence-specific DNA sites. This induces an increased p53 protein synthesis and a subsequent activation of p53 via phosphorylation. Activated p53 controls the cell cycle by upregulating the expression of p21, a CDK inhibitor that specifically targets CDK 2, 4, and 6. Inhibition of these prevents the phosphorylation of Rb and subsequently halts the cell cycle progression. The cell remains in G1, giving the cell time to repair the damaged DNA. If the damage is irreparable, p53 directs the cell towards cell death by upregulating pro-apoptotic proteins (Engeland, 2022; Israels & Israels, 2000).

As outlined, cyclin-CDK complexes and cell cycle checkpoints are essential in regulating the cell cycle, maintaining genomic integrity, and ensuring accurate cell division. Disruptions in these regulatory mechanisms can lead to uncontrolled cell prolifera-

tion, which is a hallmark of cancer (Williams & Stoeber, 2012). This underlines the importance of understanding these processes, as they not only play a central role in normal cell development but also in tumorigenesis.

1.3.1 Out-of-cycle states

In addition to active cell cycle phases, there are also so-called out-of-cycle states. Those are distinct non-dividing, non-proliferative stages that play important roles for the cell in both physiology and disease. There are quiescence, senescence, and dormancy, which represent different modes of cell cycle arrest characterized by varying triggers, reversibility, and biological functions. While all involve stopping proliferation, the molecular mechanisms and implications of each state differ.

Quiescence (G0 phase) is a non-proliferative state, in which cells reversibly exit the cell cycle in response to specific environmental conditions such as nutrient depletion, growth factor deprivation, or increased cell density during early G1. This reversible arrest allows cells to conserve resources and maintain tissue homeostasis under unfavorable conditions. Upon cellular exposure to mitogenic and growth signaling, G0 can be easily reversed and cells enter the cell cycle at G1, resuming proliferation (Coller et al., 2006; Israels & Israels, 2000). Quiescent cells typically exhibit low metabolic activity and suppressed expression of proliferation markers but retain the capacity for rapid activation. In the context of cancer, quiescence may enable tumor cells to survive therapies targeting dividing cells, contributing to disease persistence and relapse (Valcourt et al., 2012).

In contrast, cellular senescence is a distinct, non-proliferative, out-of-cycle state that is characterized by the irreversible arrest of the cell cycle. Unlike quiescence, senescence represents a permanent interruption of cell division. Senescent cells cannot re-enter the cell cycle, even in the presence of mitogenic stimuli. It occurs in diploid cells as an end of proliferation and acts as a protective mechanism in response to various cellular stresses such as DNA damage, telomere shortening, oncogene activation, or oxidative stress. Senescence ensures that cells with significant damage or potential to become harmful do not continue to divide, thus safeguarding the organism from compromised cell function (Calcinotto et al., 2019; Williams & Stoeber, 2012).

Lastly, dormancy describes a prolonged, reversible, non-proliferative state of cancer cells that allows them to survive undetected for years after removal of the primary tumor (Endo & Inoue, 2019). In general, dormancy can be classified into two forms: tumor mass dormancy, where proliferation and cell death are balanced by factors like limited blood supply and immune surveillance, and cellular dormancy, characterized

by individual cancer cells entering a quiescent-like state with cell cycle arrest and stress tolerance. Dormant cells show resistance to therapy and reduced metabolism, enabling them to evade treatment and later cause relapse. Although dormancy shares some features with senescence, it is distinct in being reversible and metabolically subdued, whereas senescence is typically irreversible and metabolically active (Min & Lee, 2023).

Understanding these distinct out-of-cycle states is critical for comprehending how cells regulate proliferation under normal and pathological conditions. While quiescence allows cells to temporarily pause proliferation in response to environmental cues, senescence represents a permanent arrest triggered by stress signals to prevent propagation of damaged DNA. Dormancy occupies an intermediate role, particularly relevant in cancer, where reversible arrest facilitates long-term survival of disseminated tumor cells. Together, these states reveal complex cellular strategies to balance growth, repair, and survival, and underline their importance in research.

1.3.2 Cell cycle and tumorigenesis

As discussed, the cell cycle is a tightly regulated process, facilitating cell division and growth while ensuring genomic stability. However, in cancer this controlled system becomes disrupted, leading to dysregulated proliferation, genomic instability, and tumorigenesis (Glaviano et al., 2025).

Cell cycle dysregulation is a multilayered process involving genetic mutations, aberrant cyclin-CDK complexes, and loss of critical checkpoints. Oncogene mutations typically drive the cell cycle forward and therefore increase cell division rates either by directly enhancing the activity of cyclin-CDK complexes or activating downstream pathways that subsequently stimulate cyclin and CDK expression. For instance, mutations in *MYC* or *CCND1*, the gene that encodes cyclin D1, lead to the overexpression of cyclins, especially cyclin D1, and CDKs. Similarly, mutations in various *RAS* genes activate cyclin and CDK activating signaling pathways such as MAPK/ERK and PI3K/AKT, which in turn stimulate cyclins and CDKs. The overexpression and increased activity of cyclin-CDK complexes accelerate progression through the cell cycle, allowing cells to bypass critical checkpoints, ultimately driving uncontrolled proliferation. Mutations in tumor suppressor genes further compromise these checkpoints, removing the critical brakes on cell division, and impairing processes like cell cycle arrest and apoptosis. Two examples of key tumor suppressors involved in cell cycle regulation that are often mutated in various cancers are p53 and Rb. As discussed earlier, p53 and Rb are essential checkpoints that halt cell cycle progres-

sion when DNA damage occurs, making them both essential for maintaining genomic integrity. Mutations in *RB1*, the encoding gene of Rb, can lead to uncontrolled S-phase entry. Mutations in p53's encoding gene *TP53*, a mutation often encountered in HGSC, can result in the loss of cell cycle arrest and apoptotic response. Both allow cells with damaged DNA to proliferate unchecked, contributing to tumorigenesis (Glaviano et al., 2025; Madaminov et al., 2021).

While mutations in oncogenes typically drive the cell cycle forward, leading to uncontrolled proliferation, in some cases, these same mutations can instead trigger a protective mechanism known as oncogene-induced senescence (OIS). This process is a natural protection against tumorigenesis. Instead of promoting cell cycle progression, it induces a permanent cell cycle arrest. Thus, senescence is a physiological tumor-suppressing mechanism, inhibiting the transformation from benign tumor to malignancy. For example, *RAS* mutations without further mutations in oncogenes or tumor suppressor genes can trigger OIS. Alone this mutation is not sufficient to drive the progression into cancer, but instead leads to permanent cell cycle arrest.

Next to senescence, quiescence and dormancy also play significant roles in tumor development. Quiescence allows cancer cells to reversibly exit the cell cycle, enabling survival under adverse conditions and resistance to therapies targeting dividing cells. Dormancy, often seen in disseminated tumor cells, permits long-term survival in a non-proliferative state, contributing to tumor relapse years after initial treatment. Both states enable cancer cells to evade treatment and may serve as reservoirs for tumor recurrence (Endo & Inoue, 2019; Valcourt et al., 2012).

In recent years, cell cycle-targeting therapies have emerged as promising treatments for cancers driven by cell cycle dysregulation. These therapies typically aim to restore normal cell cycle control by inhibiting key regulators such CDKs, which are often overactive in cancer cells. Additionally, some therapies aim to induce senescence in tumor cells, effectively halting their proliferation. These therapies exploit the natural tumor-suppressive role of senescence, triggering a permanent cell cycle arrest in cancer cells, thus reducing tumor growth. However, effectively managing this as well as other out-of-cycle states remains a challenge (Glaviano et al., 2025).

Overall, the cell cycle and its control represent a critical aspect of tumor heterogeneity and cancer development, influencing both cell proliferation and cell death. Understanding the mechanisms underlying cell cycle dysregulation and out-of-cycle states is key to improving cancer treatments.

1.3.3 Cell cycle staging

The kinetics of the tumor cell cycle impact prognosis and influence how a tumor responds to treatments targeting specific cell cycle phases. There are several methods to determine the cell cycle stage a cell is in, each with its advantages depending on the research question (Williams & Stoeber, 2012).

One frequently used method measures the amount of DNA within the cells. While cells in the earlier cycle stages, G0 and G1, contain only one genome set, in stages after the S phase, G2 and M, the amount of DNA in cells doubles. The DNA content during S phase is somewhere in between. To measure the DNA content, DNA is stained by fluorescent dyes that stoichiometrically bind to the DNA, for example Hoechst 33258, Hoechst 33342, propidium iodide, or 4,6-diamidino-2-phenylindole. The fluorescent signal is analyzed and the percentages of cells in G0/G1, S, and G2/M phase can be calculated using several mathematical models (Ligasová et al., 2023).

Another widely used approach, many cell cycle analysis methods are based on, is immunofluorescence (IF) staining. It relies on detecting specific cell cycle proteins with antibodies conjugated with fluorescent dyes and thus enables the visualization and qualitative and quantitative analysis of cell cycle progression and cell proliferation (Bruhn et al., 2014; Williams & Stoeber, 2012). An example for a method based on IF staining is Ki67 staining. Ki67 is a nuclear protein that is expressed during all active phases of the cell cycle (G1, S, G2, and M) but is absent in non-dividing cells (G0, dormant, and senescent cells). Through this characteristic, it is widely used to assess cell proliferation and the fraction of proliferating cells in a population (Ligasová et al., 2023). Another method using this approach aims to identify cells in S phase. It is based on culturing cells in media containing nucleoside analogs like bromodeoxyuridine. Cells in S phase incorporate the artificial building block into the newly synthesized DNA and can be detected by staining cells with antibodies that specifically bind to the added modified nucleoside (Bruhn et al., 2014; Ligasová et al., 2023).

Beyond those two examples, many other markers can be used to study and determine the different cell cycle stages. Several examples are listed in Table 1.

Table 1: Examples of commonly used cell cycle markers.

Cell cycle marker	Cell cycle phase	Reference
Ki67	G1, S, G2, M (proliferation marker)	Ligasová et al., 2023
Cyclin D1	G1	Israels and Israels, 2000
pRb	G1/S transition	Israels and Israels, 2000
Bromodeoxyuridine	S phase	Ligasová et al., 2023
Cyclin A	S/G2 transition	Israels and Israels, 2000
Cyclin B	G2	Bruhn et al., 2014
MPM2	M phase	Tapia et al., 2006
pH3	M phase	Tapia et al., 2006
p21	cell cycle arrest	Calcinotto et al., 2019

1.4 Spatial resolution methodologies

Spatial resolution refers to the ability to distinguish and map features within a tissue based on their location and structure. It is essential in biological research for understanding the organization of cells, tissues, and molecular interactions. In tumor research, spatial resolution is crucial for examining how cancer cells interact with the TME. These interactions are key to understanding cancer progression and therapy resistance, making spatial resolution critical for developing targeted treatments.

For spatial analysis, tissue samples must be prepared to preserve their biological structure on a molecular and cellular level. Formalin fixation followed by paraffin embedding is the most common method used in clinical settings, producing stable formalin-fixed paraffin-embedded (FFPE) blocks that can be stored long-term and used for various staining techniques. Another method, cryosectioning, rapidly freezes tissues, preserving molecular structures. Resulting blocks from both methods are sliced with a microtome into thin tissue sections, before they can be analyzed using various techniques (Bodenmiller, 2016).

Traditional staining methods such as hematoxylin and eosin (H&E), immunohistochemistry (IHC), and IF are widely used to visualize tissue structures. H&E staining is a standard practice in pathology for diagnosis and prognostic evaluation. It provides a general overview of tissue architecture, but lacks molecular specificity. IHC uses antibodies to visualize specific proteins with a chromogenic signal, typically detecting

one or two markers at a time. IF, using fluorescence-labeled antibodies, allows detection of up to five markers with high resolution, but is limited by photo bleaching and fluorophore overlap, restricting the number of markers visible at once.

While valuable, these methods have limitations when studying the complex TME. H&E offers minimal molecular information, and both IHC and IF are constrained by their low multiplex capacity, making it difficult to analyze multiple tissue components simultaneously (Fischer et al., 2008; Tan et al., 2020).

In contrast to the limited amount of molecular information H&E, IHC, and IF can provide, omics approaches such as genomics, transcriptomics, and proteomics offer the ability to generate vast amounts of molecular data, even at a single-cell level. Techniques like whole genome sequencing, RNA-seq, and mass spectrometry (MS)-based proteomics enable comprehensive profiling, thereby providing deep insights into molecular landscapes. However, these methods lack spatial context, which is critical for understanding how cells interact within their environment (Hasin et al., 2017).

The low multiplex capacity of traditional staining methods as well as the lack of spatial resolution in omics technology highlight the need for novel methods that integrate molecular data with tissue architecture, offering a more complete picture of the structural and functional landscape of the heterogeneous TME. This has led to the development of advanced imaging technologies, that enable highly multiplexed spatial analysis, and novel spatially resolved omics methods. While these emerging approaches, which are discussed in the following chapters, offer exciting potential for simultaneously analyzing multiple biomarkers and providing spatial context, they still have some limitations (Merritt et al., 2020). These challenges present an ongoing opportunity for the development of new technologies and methodologies that could overcome the limitations of existing approaches and enable a more complete and scalable integration of molecular and spatial data.

1.4.1 Highly multiplexed imaging techniques

Highly multiplexed imaging techniques enable the simultaneous detection of multiple biomarkers within a single tissue sample, allowing for a comprehensive spatial analysis of molecular composition and cellular interactions. Unlike traditional methods such as IHC and IF, these advanced techniques are not limited to a few markers and can measure a large number of proteins, nucleic acids, or metabolites in the same sample, making them essential for studying complex heterogeneous environments like the TME.

The existing techniques can be divided into two categories: mass cytometry/ spectro-

metry-based and fluorescence-based imaging. Mass cytometry/spectrometry-based methods provide higher sensitivity in comparison to fluorescence-based methods, but are limited in their smaller coverage area and therefore low throughput, and more complex, time-intensive analysis. In contrast, fluorescence-based methods are not as detailed, but are less dependent on specialized reagents, expensive equipment, and allow a higher throughput, thereby making the analysis of larger tissue areas possible (Giesen et al., 2014; Lin et al., 2018).

Mass cytometry/spectrometry-based imaging

Mass cytometry/spectrometry-based imaging combines mass cytometry or MS with spatial tissue analysis, allowing for simultaneous detection of multiple biomarkers in tissue sections with high precision at subcellular resolution. The general principle involves using antibodies conjugated to heavy metal isotopes or other molecules to label the tissue, which is then ablated. The released ions are analyzed to map their abundance across the tissue, revealing detailed molecular patterns.

One widely used method is imaging mass cytometry (IMC). It uses metal-tagged antibodies and a laser that ablates specific tissue regions. Resulting ions are analyzed by time-of-flight mass cytometry, allowing the detection of over 40 protein markers simultaneously with high-resolution, high-plex imaging (up to 1 μm). However, it is limited by the size of the region of interest (ROI) and low throughput due to long analysis time per sample (Giesen et al., 2014; Milosevic, 2023).

Building on similar principles, multiplex ion beam imaging (MIBI) uses secondary ion mass cytometry. Here, an ion beam hits the tissue to release secondary ions, which are analyzed to map the distribution of metal-labeled antibodies. It provides excellent spatial resolution (down to 200 nm) for single-cell analysis and can detect up to seven markers simultaneously, with the potential to expand detection by iterating up to 10 times over tissue layers. MIBI offers better single-cell resolution than IMC but is also limited by small ROIs and the requirement for specialized equipment (Angelo et al., 2014; Baharlou et al., 2019).

In contrast, matrix-assisted laser desorption/ionization (MALDI)-imaging MS takes a label-free approach. Using laser ablation, it ionizes molecules like proteins, lipids, and metabolites for analysis via MS. It offers spatial resolution between 10-50 μm and can detect up to 10-20 markers simultaneously. While it provides detailed molecular information over a wide variety of molecules, MALDI-imaging MS is limited in multiplex capacity and requires careful sample preparation (Zaima et al., 2010).

Taken together, mass cytometry/MS-based imaging methods provide high-plex and quantitative analysis, making them ideal for studying complex tissues such as the TME. These techniques can analyze a broad range of biomarkers in parallel with sub-cellular resolution. Nonetheless, they are often time-consuming, require specialized instrumentation with complex data analysis, and have relatively low throughput. Additionally, while MALDI-imaging MS offers broad coverage, its multiplex capacity is lower than that of IMC or MIBI, and the methods are generally restricted to smaller tissue sections.

Fluorescence-based imaging

Fluorescence-based highly multiplexed imaging techniques use antibodies conjugated to fluorescent molecules to detect extra- and intracellular proteins in tissue sections. These methods allow the detection of multiple markers within a single tissue sample by applying multiple cycles of fluorescence staining and visualization, followed by fluorophore bleaching. This enables the simultaneous identification of numerous biomarkers, offering detailed phenotypic and neighborhood information.

Tissue-based cyclic immunofluorescence (t-CycIF) is one such technique, enabling the detection of up to 60 markers within a single tissue section. With t-CycIF, even large tissue areas can be explored with single-cell resolution. It uses sequential cycles of staining with DNA dye and 3-5 fluorescence-labeled antibodies (depending on the number of available fluorescence channels of the imaging device), imaging, and fluorophore inactivation through oxidation in hydrogen peroxide under light exposure. The resulting low-plex images are then overlaid to visualize the tissue in a high-plexed way (up to 60-plex images). t-CycIF stands out for not requiring specialized reagents, its compatibility with commercially available antibodies, and its customizable antibody panels, making it adaptable for various projects. It also works well with FFPE samples and allows for H&E staining afterwards. While the resolution depends on the imaging system, super-resolution imaging (<200 nm (Schermelleh et al., 2019)) is possible. Albeit a key limitation is tissue degradation after several cycles. Consequently, the truly possible number of biomarkers is restricted by the number of cycles the tissue integrity allows for (Lin et al., 2018).

Another method, co-detection by indexing (CODEX), is based on DNA-barcoded antibodies that are conjugated to target proteins. Fluorescence-labeled reporter sequences bind to the DNA barcodes, enabling the visualization of multiple markers in

a tissue section. After imaging, the reporter sequences are chemically stripped, allowing for the application of new markers and fluorescence labels in subsequent cycles. This iterative process enables the detection of up to 60 markers at high resolution (up to 200 nm), with three markers per cycle, while minimizing batch effects typical for other techniques. However, it is limited by costly reagents, the availability of non-cross-reacting and barcoded antibodies, and small tissue areas (4 mm $\times$ 4 mm). The system is also sensitive to photo bleaching (Black et al., 2021).

A further approach, Opal multiplex IHC assay, uses tyramine signal amplification to visualize proteins. In this method, target-specific primary antibodies bind to the proteins, and secondary antibodies conjugated with horseradish peroxidase bind to the primary antibodies. The secondary antibodies react with tyrosine linked to a fluorophore and lead to the formation of covalent bonds between fluorophore and tissue in the vicinity of the target antigen. Afterwards, heat-induced antigen retrieval removes the antibodies to enable the next round of antibody staining. After all marker signal amplification rounds, the tissue is stained with a nuclear dye, before the staining is visualized. Through the use of secondary antibodies and signal amplification, Opal enables the detection of low-abundant proteins. Moreover, there is no antibody cross reactivity, achievable resolution is relatively high, up to 250 nm, and it provides a higher signal-to-noise ratio and photostability compared to other fluorescence techniques. Disadvantages of this method include the limitation to 8-10 markers per slide due to fluorophore overlap and tissue decay after multiple rounds of staining (Harms et al., 2023; Opal Multiplex, 2022).

All in all, fluorescence-based imaging methods offer a powerful way to visualize multiple markers in a single tissue section with single-cell resolution. They are becoming more widely used with a growing popularity due to the abundant availability of reagents and antibodies as well as the ability to examine larger tissue areas, in contrast to the smaller ROIs typical in mass cytometry. However, they also encounter challenges such as background fluorescence (particularly in FFPE tissue blocks), signal overlap between markers, and the potential for signal leakage caused by overlapping emission and absorption spectra.

The following table, Table 2, contains a summary of the key characteristics of the discussed mass cytometry/spectrometry-based and fluorescence-based imaging methods.

Table 2: Key characteristics of different highly multiplexed imaging techniques.

Method	Detection method	Resolution [nm]	simultaneous Biomarker Nr	Cyclic process
IMC	Mass cytometry	1000	> 40	No
MIBI	Mass cytometry	200	7	Yes 10 cycles
MALDI-imaging MS	Mass spectrometry	10000-50000	10-20	No
t-CycIF	Antibody-fluorescence conjugates	< 200	3-5	Yes > 15 cycles
CODEX	Antibody-DNA barcode conjugates	200	3	Yes 20 cycles
Opal	Tyramine signal amplification	250	8-10	No

1.4.2 Proteomics

Proteomics is the investigation of structure, function, composition, activity, and interactions of proteins (Al-Amrani et al., 2021). Since protein expression changes throughout the cell cycle and adjusts to environmental conditions, proteomics enables an actual insight into current cellular events and is an important foundation for the characterization of cellular systems. Because of this, it is widely used in various research areas, especially medical fields like oncology for finding new diagnostic biomarkers and therapy targets (Aslam et al., 2016; Vitzthum et al., 2005).

In general, there are three major forms of proteomics:

Expression proteomics focuses on the qualitative and quantitative analysis of proteins differently expressed within two cellular states, such as diseased and healthy cells. With it disease specific biomarkers and proteins involved into the disease mechanism can be identified. Usually, techniques like 2-dimensional gel electrophoresis and MS are used. A second approach concentrates on the protein structure. Structural proteomics utilizes methods like X-ray crystallography and nuclear magnetic resonance spectroscopy to investigate 3-dimensional complex structures. Lastly, there is the field of functional proteomics. It aims to unravel the biological role of proteins within the cell, especially by focusing on protein-protein interactions (Al-Amrani et al., 2021).

For investigating the cellular proteome, a series of different methodologies can be utilized. Purification of protein samples is often done using various chromatographic methods. Complex protein mixtures can be separated using gel based approaches

such as 2-dimensional gel electrophoresis. While for the analysis of individual proteins methods like western blotting, enzyme-linked immunosorbent assays, or protein microarrays/chips are commonly applied, the latter one also as high-throughput method, whole proteome characterizations are usually performed with high-sensitivity, high-throughput methods like MS. And, as already mentioned, X-ray crystallography and nuclear magnetic resonance spectroscopy, also high-throughput, are used for determining the 3-dimensional structure of proteins. Next to these listed techniques, there are many more that are employed in proteomics analysis, but listing all would exceed the scope of this work (Aslam et al., 2016).

MS-based proteomics analysis

Over the last decade, MS has become the main method for qualitative and quantitative proteomics analysis. With its high sensitivity, resolution, accuracy, and possibility to perform it on a broad range of scales, from identification of single proteins to many thousand proteins, it is ideal for studying a wide variety of protein samples in great detail (Rozanova et al., 2021; Shuken, 2023). The method is based on ionizing peptides and proteins, separating them based on their mass-to-charge ratio m/z , and ultimately detecting the ions via measuring time, m/z , and signal intensity. Commonly, electrospray ionization, MALDI, or surface enhanced laser desorption/ionization are used for the ionization of the analytes (Aslam et al., 2016; Rozanova et al., 2021).

The used MS approach usually depends on the research question. There are two main strategies: untargeted and targeted quantification. For global proteomics, untargeted quantification is ideal for analyzing large numbers of proteins across samples, providing a broad proteomics overview. Targeted quantification, however, focuses on specific proteins or modifications, offering higher sensitivity for applications like biomarker validation.

MS-based quantification can be label-based, where stable isotopes enable precise protein comparisons across samples, or label-free, which relies on natural peptide intensities but demands high experimental reproducibility. Furthermore, quantification can be relative, comparing protein levels between samples, or absolute, measuring the exact concentration of a protein within a sample (Rozanova et al., 2021).

Moreover, there are two different concepts of sample preparation - the experimental strategy can be either bottom-up or top-down. The bottom-up method involves digesting proteins into smaller peptides for high-throughput analysis, which is effective for complex samples but loses the relationship between peptides and their parent proteins, complicating interpretation. In contrast, the top-down approach analyzes

whole proteins, preserving their structural context and enabling the identification of protein modifications, but it is less widely used due to its complexity and lower throughput (Al-Amrani et al., 2021; Shuken, 2023).

From these broad MS approaches, a variety of specific techniques have emerged, enabling a very focused, problem-specific exploration of protein functions, modifications, and interactions in various biological contexts. One prominent example is shotgun proteomics, which is commonly used for untargeted analysis. In shotgun proteomics, proteins are first digested into peptides, which are then analyzed using data-dependent or data-independent acquisition. In the former, a subset of the most abundant precursor ions is selected for fragmentation and analysis, whereas in the latter, all precursor ions within a predefined mass window are fragmented simultaneously, offering more reproducibility and less bias in comparison (Rozanova et al., 2021).

Spatially resolved proteomics

Spatially resolved proteomics is an advancing field that integrates proteomics analysis with spatial resolution tissue analysis, providing insights into functional roles of proteins in situ (Lundberg & Borner, 2019). Understanding complex cellular reactions, like gene expression and signaling cascades, within their spatial context is essential, as tissue usually has a high level of cellular heterogeneity and cellular behavior is influenced by its microenvironment. While bulk-proteomics offers valuable insights into tissue-wide mechanisms, it averages out the molecular heterogeneity present in complex tissues, such as tumors and their TME, thereby obscuring important intracellular variation. In contrast, highly multiplexed imaging methods, like those discussed before, provide this needed spatial context while enabling the simultaneous detection of multiple protein markers. They can therefore also be seen as a type of spatial proteomics. However, these imaging techniques often face challenges related to low throughput and limited multiplexing capacity, restricting their ability to provide truly comprehensive spatial proteomics information. To address these limitations of traditional proteomics approaches and imaging technologies, several methods have been developed to achieve spatially resolved proteomics (Taylor et al., 2021).

Firstly, there is laser capture microdissection. It is a method combining targeted tissue isolation by cutting tissue with a laser with MS-based proteomics analysis of the isolated tissue region. As a result, it provides spatial and proteomics information with high precision and high resolution around 100-200 μm (Zhu et al., 2018). On the downside, laser capture microdissection is limited by a low protein yield due to small

sample sizes, a high amount of manual steps, making it work-intensive and prone to mistakes, and compromised sensitivity due to small tissue sizes (Gutstein & Morris, 2007; Taylor et al., 2021).

Another method, nano-droplet processing in one pot for trace samples, is especially useful for small samples. It enables highly sensitive analysis on a single-cell level by isolating proteins in nano-droplets that are then processed and analyzed typically by liquid chromatography-tandem MS (LC-MS/MS). Through the droplets sample loss is minimized, the method's accuracy is increased, and also low-abundant proteins can be detected. However, it also has limitations: only small sample sizes are feasible, and the approach requires specialized equipment, which makes it less accessible (Kelly et al., 2019).

A third approach is called ProteomEx. Its general principle is based on the integration of tissue expansion with MS-based proteomics to achieve high-resolution spatially resolved proteomics. Expansion is achieved by embedding the tissue in a hydrogel. Through this growth in size, protein distribution within the tissue can be investigated more easily and in more detail. The expanded tissue is then microdissected and analyzed by MS, typically LC-MS/MS, preserving protein localization while enabling the identification of thousands of proteins. With this method, a resolution up to 160 μm can be achieved, and it is also feasible for small tissue regions without losing spatial information. Additionally, necessary reagents are commercially available and the method is cost effectively, robust, and easy to implement. Key drawbacks of ProteomEx are a possible tissue distortion due to the expansion, lower protein yields, and, as a consequence, weaker protein signals that may lead to less detected proteins (Li et al., 2022).

All in all, spatially resolved proteomics approaches provide the possibility to analyze proteins within their cellular and tissue context. Thereby, they enable investigation of complex tissues, such as the TME, with high spatial precision and sensitivity, creating functional protein landscapes. However, despite their promising nature, current spatial proteomics methods face limitations in scalability, sensitivity, and workflow complexity. While some achieve relatively high spatial resolution, none meet the precision of spatially resolved imaging techniques. Additionally, they often require complex protocols, costly equipment, and specialized reagents, limiting accessibility. Therefore, there remains a clear need for a spatial proteomics method with high spatial precision, simple protocols, and affordable materials. Addressing this gap will improve our ability to map the proteomics architecture of tumors, paving the way for better-targeted therapies.

2 Aim of the Project

This project aims to establish the basis for a new spatially resolved proteomics method called ProteoPicking. It will integrate t-CycIF staining with MS-based proteomics for the targeted analysis of specific ROIs. By interlinking multiplexed imaging analysis, which enables detailed compositional and phenotypic characterization of ROIs, with proteomics profiling of the isolated cells, ProteoPicking aims to provide a holistic view of the structural, functional, and molecular landscape of tissue samples.

The project is guided by the following objectives:

1. Evaluation of the feasibility of ProteoPicking and compatibility of t-CycIF staining with MS-based proteomics analysis.
2. Establishment and optimization of an experimental workflow comprising t-CycIF staining, tissue sample collection of specific ROIs, and subsequent MS analysis.
3. Exploration of key factors influencing the number of detected proteins within samples, like the amount of tissue per sample, the targeted tissue region, or used sample lysis buffer.
4. Development of a data analysis pipeline that integrates image analysis data with proteomics analysis data, enabling comparative and spatially informed analysis of the structural, functional, and molecular landscape of different tissue samples.

By systematically addressing these objectives, this project aims to overcome key technical and analytical challenges in developing the method, laying the groundwork for its effective application. With continued refinement, the successful implementation of ProteoPicking has the potential to provide an enhanced spatially resolved proteomics platform for detailed molecular profiling of complex heterogeneous tissues.

3 Materials & Methods

3.1 Materials

3.1.1 Reagents & instruments

Table 3: Used reagents and instruments.

Reagent/Instrument	Catalog Nr.	Manufacturer
Antibodies	see Table 4	
Antigen retrieval Chamber High Voltage	DCARC0003	Biocare Medical
Citric acid monohydrate, ACS reagent, $\geq 99.0\%$	C1909-2.5 kg	Sigma-Aldrich
CyteFinder II	-	RareCyte
CytePicker Retrieval Module	-	RareCyte
Epredia Menzel Microscope Coverslips, 24x60 mm	BB02400600 A113FST0	Thermo Fisher Scientific
Ethanol, 96 %	13210132	Berner Medical
Hoechst 33342	H3570	Invitrogen
Hydrogen peroxide (H_2O_2), 30 % in water	BP2633500	Thermo Fisher Scientific
Glycerol, for molecular biology, $\geq 99.0\%$	G5516-500ML	Sigma-Aldrich
Kartell Polystyrene Microscope Slide Box	10586191 0027800	Thermo Fisher Scientific
LED Ceiling Luminaire	9290026655	Philips
LED Drive, Constant Current LED Power Supply	EHC012S290 DNTP01	Philips
Microm 355, motorized rotary microtome	905480	Thermo Fisher Scientific
Microtome blade MX35, ultra low profile	3053835	Thermo Fisher Scientific
Odyssey blocking buffer, 50x	927-40150	Licor Biosciences
PCR tubes, 0.2 mL	FG-021	Nippon Genetics

Reagent/Instrument	Catalog Nr.	Manufacturer
Phosphatase inhibitor tablets, PhosSTOP	04906845001	Sigma-Aldrich
Protease inhibitor tablets, cOmplete, Mini, EDTA-free	04693159001	Sigma-Aldrich
Sodium hydroxyde (NaOH), pellets Emplura	1064621000	Sigma-Aldrich
SuperFrost Plus Adhesion Microscope Slides (75x25 mm)	631-9483	Thermo Fisher Scientific
2,2,2-Trifluoroethanol, BioUltra, for molecular biology, $\geq 99.0\%$	96924-250 mL-F	Sigma-Aldrich
Tris-buffered saline (TBS, 10x), pH 7.4	J60764.K7	Thermo Fisher Scientific
Tris-HCl, 1 M soln., pH 8.0, Molecular Biology Grade, Ultrapure	J22638.K2	Thermo Fisher Scientific
Urea, 99.5 %, for analysis	AC140750050	Thermo Fisher Scientific
Xylene, ACS, 99 %	10741624	Acros Organics

Table 4: Used antibodies for t-CyclIF.

Antibody	Clone	Catalog Nr.	Company
CD45	2D1	FAB1430P	RnD systems
CyclinD1	EPR2241	ab203448	Abcam
Iba1	EPR6136(2)	ab195031	Abcam
Ki67	D3B5	11882S	CST
MPM2	-	16-220	Millipore-Merck
PAX8	EPR18715	ab215953	Abcam
pH3	D2C8	3465S	CST
pRb	D20B12	8957S	CST
p21	12D1	8587S	CST
Vimentin	D21H3	69227S	CST

3.1.2 Patients & specimens

All tissue samples were obtained from the ONCOSYS-OVA prospective study collection or the Helsinki University Hospital (HUU). The ethics board of the Women’s Clinic, HUU, and Helsinki Biobank have granted approval for the use of all samples and associated clinical data in this study. Patient identifiers and personal information are securely managed within the HUU server, utilizing the HUU Datalake infrastructure provided by Helsinki Biobank. Only authorized personnel have access to clinical data. To ensure patient confidentiality in publicly available results, each individual is assigned a randomized publication ID, reducing the risk of identification. All used samples in this project are derived from the omentum of HGSC patients after NACT and interval debulking surgery. Table 5 contains an overview of all patient samples used in this project.

Table 5: Overview of key characteristics of used patients samples. HRP - homologous recombination proficiency, HRD - homologous recombination deficiency.

Patient	Treatment	Tissue	HR status
C129	NACT + IDS	Omentum	HRP
C331	NACT + IDS	Omentum	HRD
C369	NACT + IDS	Omentum	HRD
C758	NACT + IDS	Omentum	HRD
C426	NACT + IDS	Omentum	HRD
C078	NACT + IDS	Omentum	HRD
C791	NACT + IDS	Omentum	HRD

3.2 Methods

For this Master’s thesis, two experiments - a pilot experiment and an optimization experiment - were performed. Below, the experimental procedure of both and any differences between pilot and optimization experiment are described. The changes in the protocol were necessary since the analysis of the proteomics results of the pilot experiment showed that the samples were contaminated with keratin and a not further identified type of polymer, which interfered to a certain degree with the analysis.

Figure 1 depicts the workflow of the method.

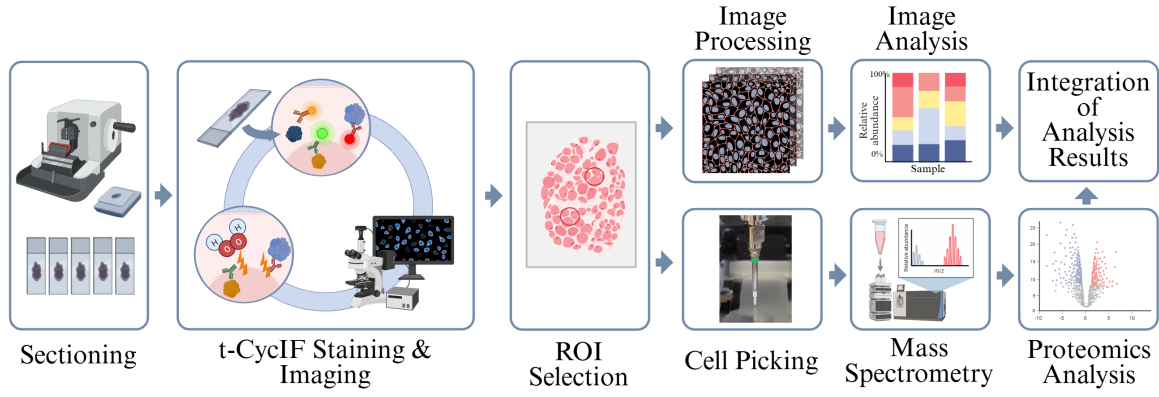


Figure 1: Schematic overview of the main steps of ProteoPicking. Figure created in BioRender, 2025.

3.2.1 Sample sectioning

Tissue samples were processed into FFPE blocks by the tissue preparation and histochemistry unit of the Anatomy Department of the University of Helsinki. Of those FFPE blocks, tissue sections were cut with a Microm HM 355 motorized rotary microtome. For the pilot experiment, 8 μm thick slides were sectioned, while the sections for the optimization experiment were 10 μm thick. The sliced sections were transferred into a waterbath set to 40 $^{\circ}\text{C}$ for approximately 1 min to completely unfold. Only then the sections were scooped up with SuperFrost Plus Adhesion Microscope Slides, dried as thoroughly as possible, placed onto a tray, baked at 56 $^{\circ}\text{C}$ for 2 h, and dried overnight at room temperature. Afterwards the samples were stored at 4 $^{\circ}\text{C}$ until further usage.

Throughout the optimization experiment, several practices were implemented to avoid contamination. Gloves were used throughout the entire experiment. The microtome, the tools for transferring the sections, the work surface, and the storage box of the microscope slides were cleaned thoroughly beforehand with ethanol. The water of the waterbath was replaced with fresh water directly from the Milli-Q filtering system, instead of using pre-filled Milli-Q water from reservoir containers, and the waterbath was cleaned before refilling. A new package of microscope slides was opened to prevent contamination from earlier experiments, and the slides were stored vertically in a plastic box with a lid, which was only removed during the baking step to allow the remaining water to evaporate from the slides. In comparison, during the pilot experiments the slides were transported and stored on wooden trays without proper cover.

3.2.2 t-CycIF staining

The t-CycIF protocol (Lin et al., 2018) was adapted to be compatible with proteomics MS assays. Within the standard protocol, phosphate-buffered saline (PBS) is used as a buffer to maintain pH and osmolarity to preserve tissue morphology and protein integrity (Visikol, 2024). But since we perform MS-based proteomics analysis following the t-CycIF process and the phosphate content of PBS would interfere with this analysis, it was replaced with tris-buffered saline (TBS), which provides comparable buffering properties (Cappiello et al., 1997).

Deparaffinization, rehydration, antigen retrieval, and pre-staining

First, slides were baked at 56 °C for 30 min. Then, they were deparaffinized and rehydrated by placing them first three times in xylene for 5 min each followed by immersing them for 3 min each in a series of ethanol solution with gradually decreasing ethanol concentration (100 %, 90 %, 70 %, 50 % and 30 % in Milli-Q). Finally, the slides were washed twice for 1 min in 1x TBS (50 mM tris-HCl, pH 7.6; 150 mM NaCl). In a next step, samples were submerged in 225 mL 10 mM citric acid buffer (10 mM citric acid monohydrate, pH 6.0) and heated up to 95 °C for 15 min in an Antigen retrieval Chamber High Voltage. This is done to fully retrieve the antigen molecules, which can be slightly modified and cross-linked after the formalin fixation. Afterwards, the slide holder with slides and filled with the hot citric acid buffer was taken out of the cooker to cool down for 20 min at room temperature and then slides were washed three times for 3 min each in 1x TBS. To minimize unspecific antibody binding of the tissue, 200-300 μ L of filtered Odyssey blocking buffer were pipetted onto the tissue samples, which were then incubated at room temperature for 30 min in a moist Kartell Polystyrene Microscope Slide Box incubation chamber to prevent evaporation and dehydration of the tissue. After incubation, the Odyssey blocking buffer was removed by vertically tapping the slides on tissue paper. The samples were then washed three times for 3 min in 1x TBS.

For the optimization experiment, again protocol changes were implemented to avoid any kind of contamination. All used equipment, like slide holders, slide chambers, antigen retrieval chamber, incubation chamber, and tweezers, as well as work surfaces were cleaned thoroughly before the experiment. Fresh xylene was used and the ethanol chain was freshly prepared, in comparison to the pilot experiment where pre-prepared solutions were reused. Reagents and buffers were prepared in glass Schott flasks and the water was acquired directly from the Milli-Q filtering system, instead of

using pre-filled Milli-Q water from reservoir containers. Moreover, the use of plastic containers was avoided as much as possible. Plastic slide holders and chambers were replaced with glass slide holders and chambers, except for the antigen retrieval chamber, which was only available in a plastic version.

Antibody staining

Before the actual staining was performed, antibody panels for both experiments were planned with three cycles for the pilot project (Table 6) respectively two cycles for the optimization experiment (Table 7). The antibody panels used in this project had already been validated and established in previous experiments, making further validation not necessary. The antibodies were selected to detect tumor, stroma, and immune cells as well as to identify different phases of the cell cycle.

For each cycle of staining, an antibody mastermix was prepared containing all antibodies for one cycle and a nuclear dye, diluted accordingly in filtered Odyssey blocking buffer. Hoechst 33342 was used for nuclear staining in each cycle. Adding it to each staining mastermix is especially important for subsequent image processing and analysis steps (see chapter *Image acquisition, processing, analysis, and visualization*). 100 μ L of the antibody mixture were pipetted per sample. The slides were then incubated in a moist incubation chamber at 4 °C overnight. Since only dye-conjugated antibodies were chosen for the antibody panel no secondary antibodies were necessary on the next day and the slides were directly washed three times in 1x TBS for 3 min each. Afterwards, 80-100 μ L of 70 % glycerol in Milli-Q were pipetted onto the tissue samples and the slides were covered with Eprelia Menzel Microscope Coverslips for the following imaging. After the imaging, coverslips were removed again by placing the slides vertically in 1x TBS. After a short while the coverslips detached on their own and the coverless slides were either prepared for the next staining cycle (fluorophore inactivation) or stored in 1x TBS at 4 °C for further steps.

For the pilot project, two of the three used tissue samples were only stained with one cycle of antibodies (cycle 3) to have two samples that did not undergo the aggressive bleaching step, which is performed in between each staining cycle (see chapter *Fluorophore inactivation (bleaching)*). This approach allows the comparison of proteomics results between sections with and without fluorophore inactivation, assessing its potential impact on MS-based proteomics analysis outcomes.

Also during this phase of the optimization experiment, special precautions were taken. Materials and work surfaces were again cleaned before usage. A new package of cov-

erslips was opened to prevent contamination from earlier experiments. And for the washing and decoverslipping steps, glass slide holders and chambers were used instead of plastic ones.

Table 6: Antibody panel for pilot experiment.

Cycle	Antibody	Target	Location	Fluorescence label [nm]	Dilution
1	Hoechst1	DNA	Nuclear	395	10000
1	Iba1	Macrophages	Cytoplasmic	488	600
1	CyclinD1	G1 phase	Nuclear & cytoplasmic	555	200
1	p21	Cell arrest	Nuclear	647	200
2	Hoechst2	DNA	Nuclear	395	10000
2	Ki67	Proliferation	Nuclear	488	300
2	CD45	Immune cells	Cytoplasmic	555	100
2	MPM2	late G2 phase	Nuclear	647	200
3	Hoechst3	DNA	Nuclear	395	10000
3	pH3	M phase	Nuclear	488	300
3	pRb	S phase	Nuclear	555	100
3	PAX8	Tumor cells	Nuclear	647	300
3	Vimentin	Stroma	Cytoplasmic	750	100

Table 7: Antibody panel for optimization experiment.

Cycle	Antibody	Target	Location	Fluorescence label [nm]	Dilution
1	Hoechst1	DNA	Nuclear	395	10000
1	Iba1	Macrophages	Cytoplasmic	488	600
1	CD45	Immune cells	Cytoplasmic	555	100
1	PAX8	Tumor cells	Nuclear	647	300
2	Hoechst2	DNA	Nuclear	395	10000
2	Ki67	Proliferation	Nuclear	488	300
2	pRb	S phase	Nuclear	555	100
2	p21	Cell arrest	Nuclear	647	200
2	Vimentin	Stroma	Cytoplasmic	750	100

Fluorophore inactivation (bleaching)

The slides need to undergo fluorophore inactivation between each staining cycle to guarantee that no fluorescent signal from the previous cycle remains. For this, the slides were placed in a transparent plastic container with 50 mL freshly prepared bleaching solution (40 mL 1x TBS, 7.5 mL 30 % H_2O_2 , 2.5 mL 480 mM NaOH). 50 mL are sufficient for 7 to 8 slides. The completely submerged slides were exposed to LED lights from an upper and lower light panel for 90 min at room temperature. After the first staining cycle, the bleaching solution was replaced two times, once after 30 min and again after 60 min. For every following bleaching step, the solution was exchanged only once after 45 min. Afterwards, the slides were washed four times in 1x TBS for 3 min each.

During this step of the protocol, for both experiments the slides were submerged in bleaching solution in plastic dishes because there was no glass replacement that could have been used for the optimization experiment. The slide holders and chambers for the washing step were again replaced with glass versions for the second experiment.

3.2.3 Image acquisition

The images of the tissue samples were acquired using a CyteFinder II imaging instrument. It is a high-throughput automated microscope device for digital imaging that enables multiplexed, high-resolution tissue imaging of whole slide sections. Within the scanner, there is a slide holder with space for two slides and multiple microscope lenses to scan the slides at different magnifications (4x, 10x, 20x, and 40x). With its different channels, up to four antibodies can be scanned simultaneously in addition to a nuclear dye. The scanner is connected to various software applications to control and perform the imaging. The RareCyte Instrument Controller helps to operate the microscope. The RareCyte CyteHub application is a control center for scanning assays and image analysis. With the RareCyte CyteFinder program, the scan process can be set up and monitored, and the RareCyte Imager software allows real-time imaging and picking control (see chapter *Cell picking*).

Before the scanning, covered slides were placed upside down in the scanner since the microscope lenses are inverted and positioned beneath the slide holder. Then, a scanning assay was created within the RareCyte CyteHub application. In it, all necessary parameters for the scanning process were defined, see Table 8.

Table 8: Overview of used parameters during the imaging process.

Parameter	Setting	Notes
Sample type	Tissue	Specifies the type of sample
Slide type	Custom ROI	Scanned area can be manually adjusted
Illumination type	Fluorescence	Specifies type of illumination for the imaging
Lens	10x	Only magnification compatible with the RareCyte Imager picking application
Binning	1	No binning applied for highest resolution
Scan overlap	Automatic	Corresponds to a 20 μm overlap in x and y direction
Focus range [μm]	270	Specifies range around the nominal z , in which the scanner is still able to focus accurately
Focus steps	Automatic	Scanner focuses automatically
Nominal z	Individual value for each slide	Specifies the exact z position for each slide for an optimal focus of the image
Registration channel	1:BV421	Specifies the channel used for assembling the tiles, here the nuclear dye channel is used
Blend panel overlap	Yes	Ensures that overlap of adjacent scanning tiles are blended smoothly
Focus channel	BV421	Nuclear dye channel is used for focusing during the imaging process
Channels	BV421, A488, Sytox, A647 & A750	For detection of fluorescent labels at wavelengths 395 nm, 488 nm, 555 nm, 647 nm and 750 nm
Biomarker	Used markers of the respective cycle	Specifies marker name and location
Exposure [s]	Individual value for each marker	Suitable exposure times are determined using the RareCyte Imager software, usually 100-200 ms

After the assay was created, the slide scanning was started using the RareCyte CyteFinder program. The first time a slide was scanned, a pre-scan was performed to identify the tissue area. This enables manual adjustment of the scanning region and selection of focus points necessary for accurate overlay of images from multiple cycles of one section. The chosen scanning region and focus points were stored in the RareCyte CyteHub program and were applied to every future scan of the slide.

The actual scanning process was performed automatically. For this, the scanning area was divided into a grid of micro-regions, so-called “tiles”, by the system. The tiles had overlapping regions with their adjacent neighbor tiles to enable correct assembling of the tiles into a complete image during the image processing (see chapter *Image processing*). The size and number of tiles depended on the chosen magnification and were automatically calculated by the system. Tiles were scanned sequentially, with

the next tile being scanned only after all channels of the prior tile were completed. After the whole selected area was scanned, the raw data of each tile was stored in a `rcpnl` file for further processing. Moreover, for each scanning cycle an `ome.tiff` file was created from the RareCyte CyteFinder II program, containing the completely assembled image with all channels of the respective cycle. These files were later used for the ROI selection in the cell picking process (see chapter *Cell picking*).

In Figure 2, an overview of the image acquisition process is depicted.

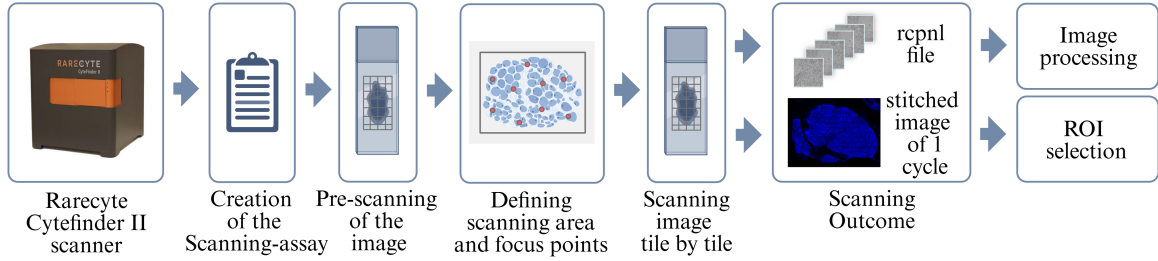


Figure 2: Overview of the image acquisition workflow for multiplexed tissue imaging - tissue sections are imaged using a Rarecyte CyteFinder II scanner, generating multi-channel, tiled image stacks (`rcpnl` files) with associated metadata and individual stitched images of each cycle. Figure created in BioRender, 2025.

3.2.4 Cell picking

The cell picking step is a critical connecting step in this workflow that interlinks t-CycIF-based spatial analysis with MS-based proteomics analysis. The protocol for the targeted microdissection of selected ROIs was adapted from Pick-seq, a spatially resolved transcriptomics approach established by RareCyte (Ericson et al., 2019; Podymnugin et al., 2019).

ROIs selection

For selecting the ROIs on the slide, the `ome.tiff` files of the individual cycles were opened in the tissue result section of the RareCyte CyteHub software. The desired cells were identified and marked with circles with a 40 μm diameter, which corresponds to the diameter of the micro needle that was used for the picking. The selected ROIs were stored as a `rcpickmap` file.

Picking preparations

The necessary needles were loaded into their holding spots along with three 0.2 mL PCR tubes that contained each 200 μ L 1x TBS for needle rinsing, retrieval of the cells from the slide, and waste discarding. Additional tubes were added for the collection of the picked cells. Depending on the experiment, they either contained 20 μ L of urea buffer (50 mM tris-HCl, 6 M urea, pH 7.5 to 8, protease and phosphatase inhibitor tablets) or 20 μ L TFE lysis buffer (50 % (v/v) 2,2,2-trifluoroethanol, 300 mM tris-HCl, pH 8.0, protease and phosphatase inhibitor tablets). When everything was prepared, the retrieval module was placed into the right slide holder position in the CyteFinder II device.

Next, the RareCyte Imager software was opened, and the calibration of the needle was performed. This step ensures precise control over the needle by aligning the actual position (x, y) of it with the position displayed in the RareCyte Imager and by assuring that the tip of the needle actually touches the slide (z) during the picking process. For the calibration, an empty slide with fingerprints on it was placed into the left slide holder position. The fingerprints helped with focusing the image. Then a needle was selected, loaded, and flushed three times with the rinsing solution. Afterwards, the needle was positioned above the slide for either an automatic calibration or a manual one, where the parameters of the needle position (x, y, z) and slide position (z) were adjusted.

In a last step, the created `rcpickmap` file, which contained the selected ROIs, was imported into the RareCyte Imager software with the Imager Main Control together with the `metadata` file of the slide.

Picking & collection

For controlling the picking and collection process, the RareCyte Imager software was used. First, the picking parameters were set, see Table 9. Then the slide the cells were to be picked from was put into the left slide holder position facing upwards, and its image was focused using the Hoechst channel. Subsequently, the picking process was performed. It comprised of a series of steps, see Figure 3. The needle was flushed two to three times in the rinsing solution, then it was filled with retrieval solution before the needle was moved to the previously selected ROIs. There, the cells were checked again for the respective biomarkers with the various channels and then the cells were picked. After the picking, the cells were deposited in the respective collection tube. As soon as the cells were collected, the next cell picking cycle was started until all

selected ROIs of one slide were picked. Afterwards, either a new slide was put into the slide holder and the rinsing, retrieval, and waste solution together with the collection tubes were exchanged, or the picking process was finished by rinsing the needle three times in Milli-Q water and then dropping it into the needle discard tray.

After all ROIs from one slide were picked, the remaining tissue was imaged again. For this, as described in chapter *Antibody staining*, a coverslip was positioned above the tissue after pipetting 80-100 μL of 70 % glycerol in Milli-Q onto the sample. Subsequently, the image was acquired following the workflow described in chapter *Image acquisition*. Since the tissue was scanned before during the t-CycIF process, the already specified scanning area and focus points were used. Following the imaging, the coverslip was removed as before, by placing the slides vertically in 1x TBS.

After the imaging, the remaining tissue was scraped into a tube containing 50 μL of TFE buffer (pilot experiment) or urea buffer (optimization experiment) using a scalpel. Lastly, all picked cell samples as well as the whole tissue samples were snap frozen with liquid nitrogen and stored in a -80°C freezer for further experiments.

Table 9: Overview of used parameters during the picking process.

Parameter	Setting	Notes
Pick standby	150 μm	Starting and ending z position of the needle during cell retrieval
Wet standby	0 μm	Distance between needle tip and tissue surface from where retrieval solution is dispensed onto tissue
Wet volume	120 nL	Amount of retrieval solution dispensed onto tissue during retrieval
Overdrive	8 μm	Distance the needle pushes down after touching the surface of the tissue, helps to isolate the targeted cells
Stage wiggles	2	Number of times the needle moves in a wiggle pattern while touching the tissue
Wiggle radius	15 μm	Radius of the wiggle motion
Focus in	BV421	Channel the system focuses on to determine the slide z position
Backoff	0 μm	Distance between needle tip and tissue surface during cell retrieval
Draw volume	120 nL	Amount of retrieval solution aspirated during cell retrieval
Speed	SFP	Speed at which the plunger of the needle dispenses retrieval solution during cell retrieval and deposit
Fill buffer	0 nL	Amount of retrieval solution the needle is filled with before the cell picking in addition to the Dispense volume
Wet wait	0 s	Time between dispensing the wet volume on the tissue and drawing sample and retrieval solution back into the needle
Distance	500 μm	z position of needle tip during sample deposit
Dispense	210 nL	Amount of retrieval solution containing the sample that is deposited after picking
Step	0 nL	Dispense volume is segmented into multiple dispenses equal to the specified volume, here sample is deposited within one dispense step

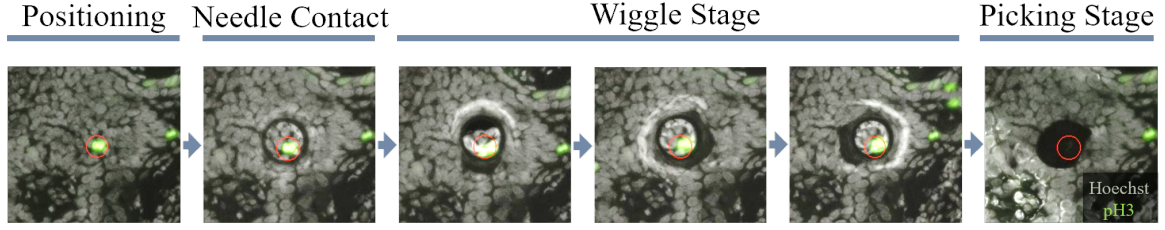


Figure 3: Visualization of the four picking steps - positioning of the needle, lowering the needle to dissect the tissue, detaching the tissue via wiggling, and collecting the tissue.

3.2.5 Image processing, phenotyping, analysis and visualization

In the pilot experiment, `rcpn1` files from imaging were simply combined into whole images through stitching and registration (see below) to enable tissue visualization. The optimization experiment, however, involved comprehensive processing of the generated `rcpn1` files, following the entire workflow described below - including all processing steps, phenotyping, and image analysis.

An overview of image processing, phenotyping, and visualization is shown in Figure 4.

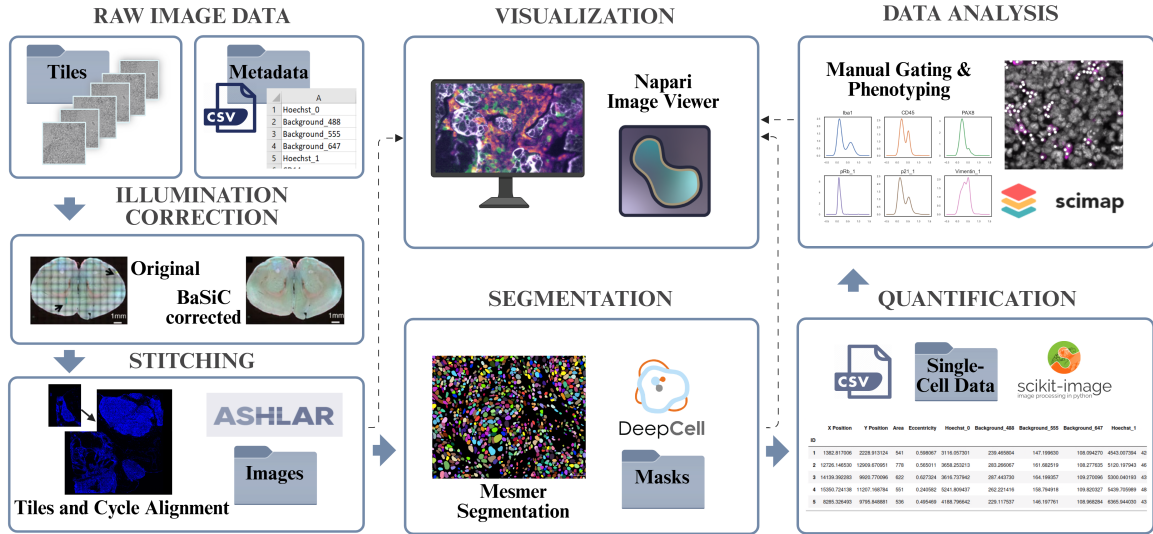


Figure 4: Overview of the image processing (including illumination correction, stitching and registration, segmentation, and quantification), phenotyping, and visualization. Figure created in BioRender, 2025.

Image processing

The first step of the image processing workflow is illumination correction. Uneven illumination, particularly darker edges in tiles, can compromise quantitative analysis and impact subsequent processing. To address this, the BaSiC program (background and shading correction, Peng et al., 2017) is used to correct `rcpn1` files generated

by the RareCyte CyteFinder application. BaSiC models and compensates for illumination bias (the non-uniform brightness from the light source) and shading bias caused by optical effects like vignetting. This correction results in uniform image illumination, ensuring consistent fluorescence intensity measurements and reliable downstream analysis.

Next, a complete multiplex image is created by assembling all tiles, channels, and imaging cycles, including the tiles from the post-picking image, from a single slide. For this, the open-source Python package ASHLAR (alignment by simultaneous harmonization of layer/adjacency registration, Muhlich et al., 2022) was used to process the illumination-corrected `rcpn1` files. ASHLAR’s workflow consists of two main steps: stitching and registration. In those steps, it relies on the nuclear staining through Hoechst, as nuclei provide high-contrast, and well-defined features with clear boundaries. During stitching, adjacent tiles within the same imaging cycle are aligned. During registration, images of different channels from different cycles covering the same tissue regions are aligned and overlaid. The output is an `ome.tiff` file - a multichannel, pyramidal stacked image containing multiple resolution layers. This format enables faster visualization and efficient processing by loading only the necessary portions and resolutions of the image into memory as needed.

Following stitching and registration, rectangular regions containing all selected ROIs of one sample were extracted from the whole slide images. The coordinates defining these regions in the original whole slide images were determined using post-picking image channels, showing the remaining tissue holes, and the ‘new shape layer’ feature of the open-source multidimensional image viewer Napari (Napari Contributors, 2019). Cropped images were then generated by cutting out the areas corresponding to these coordinates. All subsequent processing steps, phenotyping, and analysis were performed on these cropped images rather than the entire slides.

After extracting the ROI areas from the whole slide images, gamma correction and unsharp masking were applied to enhance nuclear visibility by increasing background intensity and sharpening cell edges. The used parameters for this step, $\gamma = 2/3$, $\sigma = 1.0$, and strength = 2.5, were determined beforehand through empirical testing. Using the two contrast and sharpness enhancement methods was necessary to improve the performance of the subsequent processing step - image segmentation.

Segmentation is the process of detecting and outlining individual cells in an image by assigning each pixel to either a specific cell or the background. For this step, the nuclear segmentation model Mesmer (mesoscopic segmentation of cellular and nuclear regions, Greenwald et al., 2022) was used. It is a set of pre-trained deep learning models designed for high-quality biological image analysis. Using a convolutional

neural network Mesmer generates per-pixel probability values for nuclear regions, cell boundaries, and background. A distance transform and post-processing algorithm, similar to the watershed method, refines these probability maps to accurately segment individual nuclei, even in densely packed regions. Mesmer operates on two input channels - either one nuclear and one cytoplasmic, or two nuclear channels. In this project, the Hoechst channels from cycles one and two were used. The resulting cell masks were saved as `ome.tiff` files.

As the final processing step, single-cell information was extracted by quantifying the segmented images using the scikit-image Python package (Van der Walt et al., 2014). It uses the assembled images and generated mask files to produce a data frame with multiple parameters for spatial analysis for each image. Each data frame includes key features such as CellID, x and y coordinates, eccentricity, area, and median marker intensity values for each object across all channels.

Phenotyping

Following image processing, the next step is cell phenotyping, in which each cell is classified based on the expression levels of specific biomarkers. This is carried out using Scimap, a scalable toolkit for analyzing spatial molecular data (Nirmal & Sorger, 2024). The package operates on the quantified single-cell data frames generated in the last image processing step, using marker intensity values along with the x and y coordinates of each cell.

A key step in this workflow is gating, which involves defining intensity thresholds for key phenotype markers to separate expressing from non-expressing cells. For the images of the optimization experiment, nearly all markers from both cycles were gated, except for the two Hoechst channels. Gating thresholds were manually defined using the Scimap gate finder function, which overlays marker expression data on the original tissue image in the image viewer Napari. This approach allowed for the visual exploration of intensity ranges and the determination of cutoff values to classify cells as positive or negative for each relevant marker. Based on this binary classification, each cell is then assigned a phenotype using a logic table (see Table 10). This table defines which combinations of positive and negative markers correspond to specific cell types. The resulting phenotype annotations were added to the single-cell data frame for downstream analysis.

Table 10: Logic table for phenotyping.

		Iba1	CD45	PAX8	Ki67	pRb1	p21	Vimentin
all	Tumor			anypos				neg
all	Non-Tumor			neg				pos
Tumor	T - Ki67+p21-				allpos		allneg	
Tumor	T - Ki67- p21-				allneg		allneg	
Tumor	T - p21+					anypos	anypos	allpos
Tumor	T - Ki67- pRb- p21+				allneg	allneg	allpos	
Non-Tumor	Immune	anypos	anypos					
Non-Tumor	Stroma	allneg	allneg					
Stroma	S - p21-				allneg		allneg	
Stroma	S - p21+					anypos	anypos	allpos
Stroma	S - Ki67- pRb- p21+				allneg	allneg	allpos	

Image analysis

For image analysis, a combined data frame was created containing single-cell data only from the cells retrieved during the picking process, along with metadata specifying the patient sample and the ROI from which the cells originated. This was achieved by cropping the single-cell data frames generated during image processing and phenotyping to retain only cells within the selected ROIs, adding the relevant metadata, and then merging the resulting data frames. The cropping was based on the x - and y -coordinates of the cells and the coordinates of the ROIs. The coordinates of the ROIs within the cropped images were determined using post-picking image channels, showing the remaining tissue holes, and the ‘new shape layer’ feature of the open-source multidimensional image viewer Napari.

The combined data frame served as the basis for quantitative analysis. Compositional analysis of the picked samples was performed to determine the total number of cells in each sample, the proportions of tumor, immune, and stroma cells within samples from tumor, tumor-stroma interface (TSI), and stroma regions as well as the relative abundances of defined tumor phenotypes within tumor cells in tumor and TSI region samples and stroma phenotypes within stroma cells in TSI and stroma region samples. Altogether, this provided an actual insight into sample composition.

Image visualization

The open-source multi-dimensional image viewer Napari (Napari Contributors, 2019) was used for visualization throughout the workflow described above. It facilitated the visualization and combination of different channels while allowing adjustments to contrast, intensity, opacity, and color to enhance image interpretation. Additionally, Napari enabled the overlay of segmentation masks onto images, providing a clear depiction of cell boundaries and aiding in quality control of the cell segmentation results. Its interactive capabilities were particularly useful for visually inspecting and validating each processing step, the phenotyping results, and any analysis outputs, thereby ensuring the accuracy of the whole workflow.

3.2.6 Proteomics analysis

Mass spectrometry analysis

MS analysis of the picked cell samples and subsequent proteomics data processing were performed by the Turku Proteomics facility, a unit of the Turku Bioscience Centre, University of Turku. The samples were processed using a bottom-up approach, following the facility's established protocols. The used method was data-independent acquisition shotgun proteomics via LC-MS/MS. The resulting proteomics data was processed using the Spectronaut software (Biognosys, version 19.2.240905.62635) employing the direct data-independent acquisition workflow for protein identification. Label-free quantification was performed using the MaxLFQ algorithm, with normalization applied according to Callister et al., 2006.

The output included a report containing the number of identified proteins of each sample and metadata specifying the analyzed samples as well as a data frame containing all identified proteins, the number of distinct peptides detected per protein in each sample, and the normalized signal intensities for each protein across all samples. Both were generated with the Spectronaut software.

Proteomics data analysis

For assessing the proteomics data of both experiments, the software Freestyle was used. With it the chromatograms of the MS runs were visualized and their quality assessed. This step allowed for manual inspection and ensured the reliability of the MS data prior to further analysis. Following this, Spectronaut reports of both experiments were used to evaluate and compare the number of identified proteins in between

samples with different cell types and picking conditions.

While the analysis of the pilot experiment was limited to those two steps, for the optimization experiment a comprehensive analysis of the proteomics data was performed. Initially, samples with fewer than 500 identified proteins were excluded due to low protein yield. The remaining data was then explored using clustering heatmaps, which grouped samples based on similar protein detection patterns. Subsequently, differential expression analysis (DEA) was conducted to identify proteins with statistically significant expression. For this analysis, samples were grouped by tissue region of origin (tumor, TSI, or stroma) and by cell type (proliferating or dormant). Differential expression was assessed by comparing proliferation and dormancy pairs within each tissue region. Proteins with a p -value < 0.05 and a $\log_2$ fold change $> \pm 2$ were considered differentially expressed.

Finally, DEA results were integrated with image analysis results by correlating differentially expressed proteins with the actual proportions of cells exhibiting a proliferative/dormant phenotype within samples. This integrated approach enabled a comprehensive understanding of the relationship between protein expression profiles and cellular phenotypes of samples.

3.3 Code availability

Code and scripts used in this project are available at the ProteoPicking repository on the Färkkilä lab GitHub, University of Helsinki (Färkkilä lab, [2025](#)). Parts of the image processing pipeline, in particular illumination correction, stitching and registration, and quantification as well as the script for image visualization in Napari were conducted using parts of the already established image processing workflow from the ImageProcessing repository on the Färkkilä lab GitHub (Färkkiä lab, [2024](#)). The code for the segmentation with Mesmer is adapted from scripts available at DeepCellApplications repository on the Van Valen lab GitHub (Van Valen Lab, [2023](#)).

4 Results

4.1 Pilot experiment

4.1.1 Experimental design

The pilot experiment was designed to assess the feasibility of ProteoPicking and explore key factors influencing the number of detected proteins in tissue samples.

In total, 25 samples from three sections and two patients were analyzed. One section was cut from the FFPE-block of patient C129 and two sections were cut from the FFPE-block of patient C331. To determine whether t-CycIF staining, particularly the aggressive fluorophore bleaching step, affects protein abundance within the tissue, one section from each patient was subjected to standard IF staining (equivalent to one cycle of t-CycIF staining, but without fluorophore inactivation), while the second C331 section underwent three full cycles of t-CycIF staining. To investigate a possible difference in the protein abundance between tumor and stroma regions, ROIs of both tissue regions were selected from the two C331 tissue sections. Tumor regions were defined as PAX8+ and further classified based on cell cycle markers, pRB+ for S phase or pH3+ for M phase. This distinction between tumor cells in different cell cycle phases was made to make the ROI selection more targeted and assess how complicated the determination of suitable regions is. Stroma regions were defined as Vimentin+ and PAX8-. Of the section from patient C129, only tumor region samples were analyzed. Since the exact protein content per pick was unknown, it was tested how many picks were necessary for a reliable protein detection. Samples were collected with 1, 5, 10, and 20 picks from tumor and stroma regions. The prepared tissue sections were 8 μ m thick, whereas standard t-CycIF is typically performed on 4 μ m to 5 μ m sections. The increased thickness meant that each pick contained more tissue, reducing the number of picks required to obtain sufficient material for MS analysis. To optimize sample preparation, two different lysis buffers were compared: urea and TFE. While all initial samples were prepared in urea, additional 10-pick tumor samples from both C331 sections were prepared in TFE to compare protein detection efficiency between the two buffers. To establish a reference for the number of detectable proteins and to ensure that the IF- and t-CycIF-stained sections contained any detectable proteins, whole-section lysates were prepared of each tissue section as positive controls.

A depiction of the experimental design is shown in Figure 5A. Table 11 contains a summary of key characteristics and picking conditions of each analyzed tissue sample.

Table 11: Sample specifications of the pilot experiment.

Sample Nr	Patient	Staining	Region	CellState	Nr of Picks	Buffer
01	C331	IF	T	M phase	1	Urea
02	C331	IF	T	M phase	5	Urea
03	C331	IF	T	M phase	10	Urea
04	C331	IF	T	M phase	20	Urea
05	C331	IF	T	M phase	10	TFE
06	C331	IF	S	M phase	1	Urea
07	C331	IF	S	M phase	5	Urea
08	C331	IF	S	M phase	10	Urea
09	C331	IF	S	M phase	20	Urea
10	C331	IF	Scrape	Scrape	Scrape	TFE
11	C331	t-CycIF	T	M phase	1	Urea
12	C331	t-CycIF	T	M phase	5	Urea
13	C331	t-CycIF	T	M phase	10	Urea
14	C331	t-CycIF	T	M phase	20	Urea
15	C331	t-CycIF	T	M phase	10	TFE
16	C331	t-CycIF	S	M phase	1	Urea
17	C331	t-CycIF	S	M phase	5	Urea
18	C331	t-CycIF	S	M phase	10	Urea
19	C331	t-CycIF	S	M phase	20	Urea
20	C331	t-CycIF	Scrape	Scrape	Scrape	TFE
21	C129	IF	T	S phase	1	Urea
22	C129	IF	T	S phase	5	Urea
23	C129	IF	T	S phase	10	Urea
24	C129	IF	T	S phase	20	Urea
25	C129	IF	Scrape	Scrape	Scrape	TFE

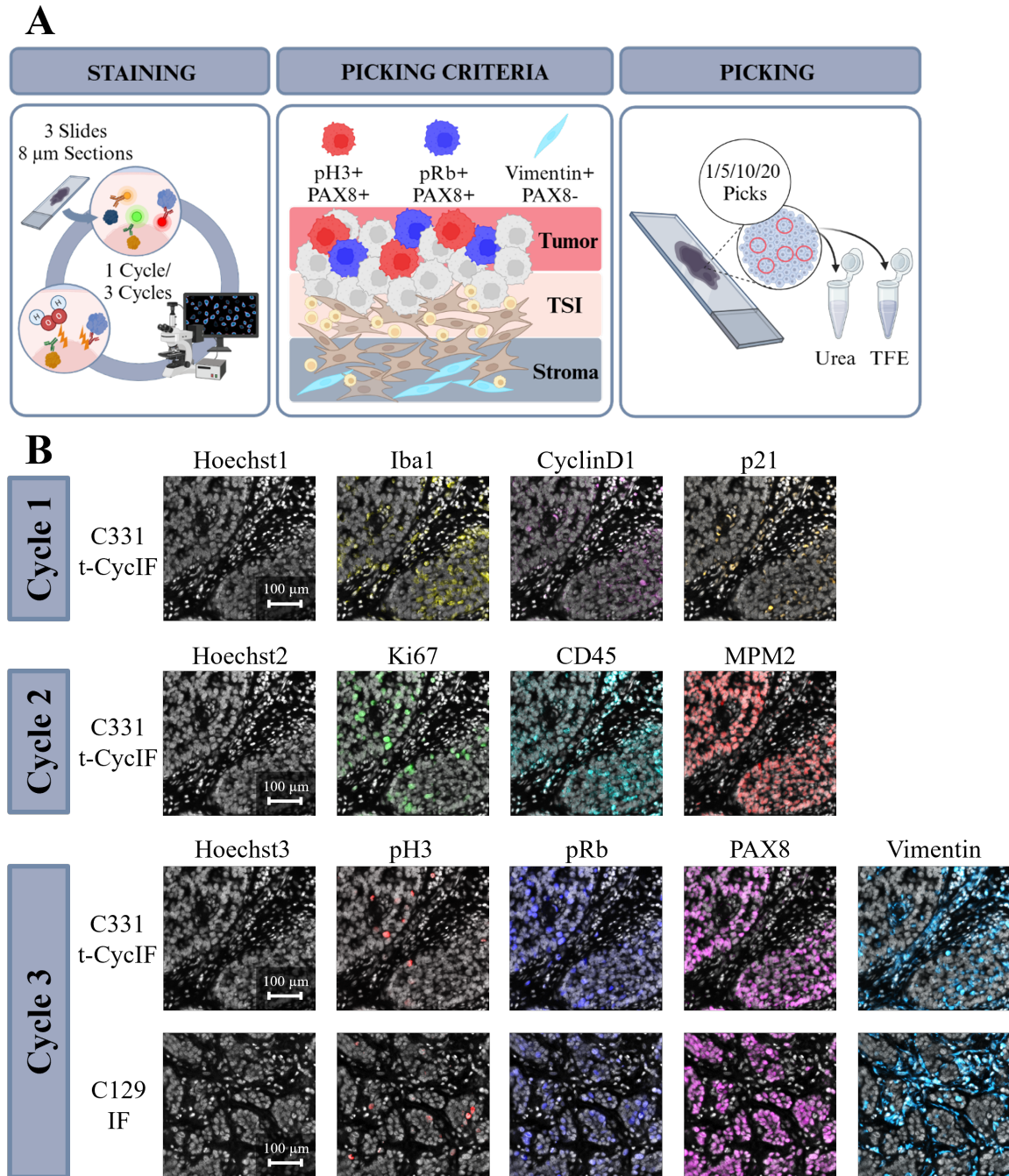


Figure 5: Overview of the experimental design of the pilot project and staining patterns of all cycles. **A** Schematic representation of the experimental setup of the pilot project. Figure created in BioRender, 2025. **B** IF and t-CycIF staining patterns of the markers used during all staining cycles. Each image shows Hoechst staining either alone or merged with the respective markers.

4.1.2 IF and t-CycIF staining enables the characterization of cell types and tissue regions

One of the three sample sections underwent three cycles of t-CycIF staining, while the other two sections were stained using standard IF staining, which corresponds to one cycle of t-CycIF staining without a fluorophore inactivation step. For IF as well as t-CycIF the protocol was followed as described in the method section. The antibodies for each cycle and the respective used concentrations are shown in Table 6. The IF samples were stained with antibodies of the third cycle only. For the staining, an antibody panel was chosen that was already validated in earlier projects of the lab. It consisted of the following biomarkers: the tumor marker PAX, the stroma marker Vimentin, the immune markers CD45 and Iba1, and cell cycle markers Ki67, Cyclin D1, pRb, MPM2, pH3, and p21.

As described earlier, the stained tissues were thicker than those usually used in t-CycIF experiments. Nevertheless, nearly all used markers resulted in clear staining patterns with sufficient signal-to-noise ratios, as shown in Figure 5B. Only the staining of CD45 of the t-CycIF sample displays a noisier background than the other markers. However, it was still possible to distinguish positive from negative cells, especially in combination with Iba1, the second immune cell marker used in this project. Both t-CycIF and IF staining results provided clear characterization of tumor and stroma areas and enabled a distinction between mitotic tumor cells and tumor cells in S phase.

Based on nuclear staining with Hoechst, images of cycles one to three of the t-CycIF sample section in addition to the image acquired after the picking process were accurately stitched with no overlapping or missing areas. This additional fourth cycle enabled the exact tracking of the picked tissue regions. There were no tissue shifts, detachments, or deformations in the images that would interfere with ROI selection or other following steps.

4.1.3 IF and t-CycIF staining patterns guide the selection of ROIs

ROIs for IF and t-CycIF samples were selected based on visual assessment of the staining patterns of the sections. Specific positive and negative marker combinations were used to identify cells and regions of interest within the images in preparation for the picking, as shown in Figure 6A. Moreover, the final staining cycle was used during the picking process to identify pre-selected regions and validate their expression profile in real time. For this, cycle composition and order were designed in a way that for the picking relevant markers were used in the last staining cycle.

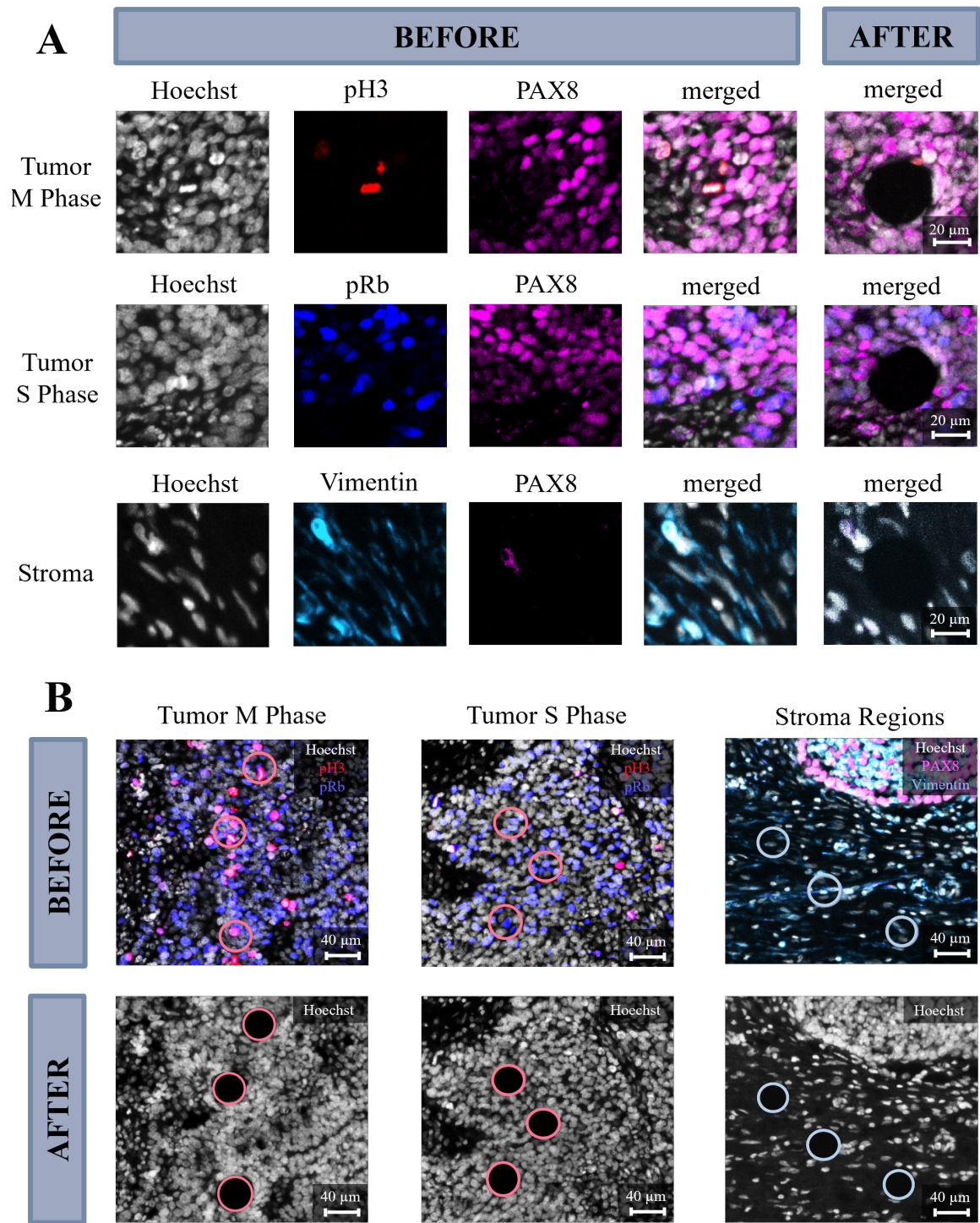


Figure 6: Picking criteria for ROI selection and examples for picked areas. **A** The ROI selection of tumor M and S phase samples as well as stroma samples is based on specific combinations of positive and negative markers. After the picking a hole remains in the tissue. **B** Examples of picked tumor M phase, S phase, and stroma regions before and after the picking process.

First, tumor and stroma areas were distinguished through PAX8 and Vimentin staining patterns as well as Hoechst staining, which revealed tissue architecture and structure. PAX8+ cells were identified as tumor, while PAX8– and Vimentin+ cells were categorized as stroma. Markers pH3 and pRb were used to further characterize tumor cells as either mitotic tumor cells or tumor cells in S phase. Regions with tumor cells only in S phase were abundant and therefore easy to determine. In contrast, mitotic regions were not as frequent, which made it partly difficult to select suitable regions. Often mitotic cells were surrounded by S phase regions. In the end, tumor ROIs were selected that contained either only mitotic tumor cells, mitotic tumor cells and as few S phase tumor cells as possible. For stroma ROIs, stroma regions were selected that were surrounded only by other stroma cells and not directly next to tumor regions. Examples for selected ROIs pre- and post-picking are shown in Figure 6B.

4.1.4 Proteomics results demonstrate compatibility of t-CycIF with MS-based proteomics analysis

Initial evaluation of the MS-derived proteomics data was performed using Freestyle, a software tool for visualizing chromatograms and corresponding spectra. Inspection of the chromatograms confirmed successful detection of peptide-derived signals across all runs, indicating that protein digestion and MS acquisition were successful. Peptide identification was further supported by spectral peaks with charges of +2, which are characteristic for tryptic peptides generated during sample preparation (Steen & Mann, 2004). However, in addition to expected peptide signals, several non-peptide impurities were consistently detected in all samples. These included prominent keratin and polymer contaminants, both of which can interfere with protein detection and quantification. These impurities were identified in collaboration with the Turku Proteomics Facility and were primarily characterized by a charge state of +1 in the mass spectra. Keratin-related peaks appeared at earlier retention times, while polymer-associated signals were detected at later stages of the chromatographic runs. Representative chromatograms and mass spectra from one tumor and one stroma sample are shown in Figure 7A, with highlighted regions indicating areas of contamination.

Following chromatographic assessment, the number of identified proteins per sample was extracted from the generated Spectronaut analysis reports. Figure 7B, C, and D depict comparisons of the number of identified proteins within different samples. Despite variability across all samples, some trends were observed.

Firstly, the number of identified proteins was comparable between samples stained

with IF and those stained with t-CycIF, indicating that the additional staining and bleaching steps in t-CycIF did not markedly impair protein detection. Moreover, a general trend was observed that showed that samples with a higher number of picked regions contained more identifiable proteins. Samples with only one pick generally yielded low protein counts and showed considerable variability. With 5 or 10 picks per sample, the number of identified proteins increased but remained variable between samples. For 20 pick samples, protein detection tended to be more consistent and frequently achieved a high protein yield relative to the rest of the dataset. This highlights that larger pick numbers generally improved reliability and proteome coverage. However, several 5 and 10 pick samples also reached protein counts that were sufficiently high to provide meaningful proteomics profiles, making them suitable for downstream analysis despite their smaller input size.

Then, tumor and stroma samples with equal numbers of picks per sample were compared. In general, tumor-derived samples tended to yield higher numbers of identified proteins than stroma-derived ones. The only exception was observed in the one-pick samples, where stroma-derived samples showed higher protein counts than their tumor-derived counterparts. However, one-pick samples also displayed considerable variability overall, and this observation may reflect an outlier rather than a consistent trend.

To further evaluate protocol performance, a comparison was conducted between 10 pick tumor samples processed with urea buffer versus those processed with TFE. Overall, the number of proteins identified was comparable between the two buffer types. However, the IF sample processed with TFE resulted in a slightly higher number of detected proteins than the one processed with urea buffer. In contrast, the t-CycIF sample processed with urea yielded a substantially higher number of proteins than its TFE-processed counterpart.

Lastly, the number of identified proteins of the post-picking scrape samples was compared with the protein yield of all small region samples. As expected, these whole-tissue lysates yielded significantly higher protein counts, ranging between 6,000 and 7,000 identified proteins. This difference reflects the substantially greater amount of input material, which increases the likelihood of detecting low-abundance proteins that may fall below the detection threshold in smaller samples. All in all, this comparison illustrates the expected difference in depth of proteome coverage when analyzing small, spatially restricted regions versus entire tissue sections.

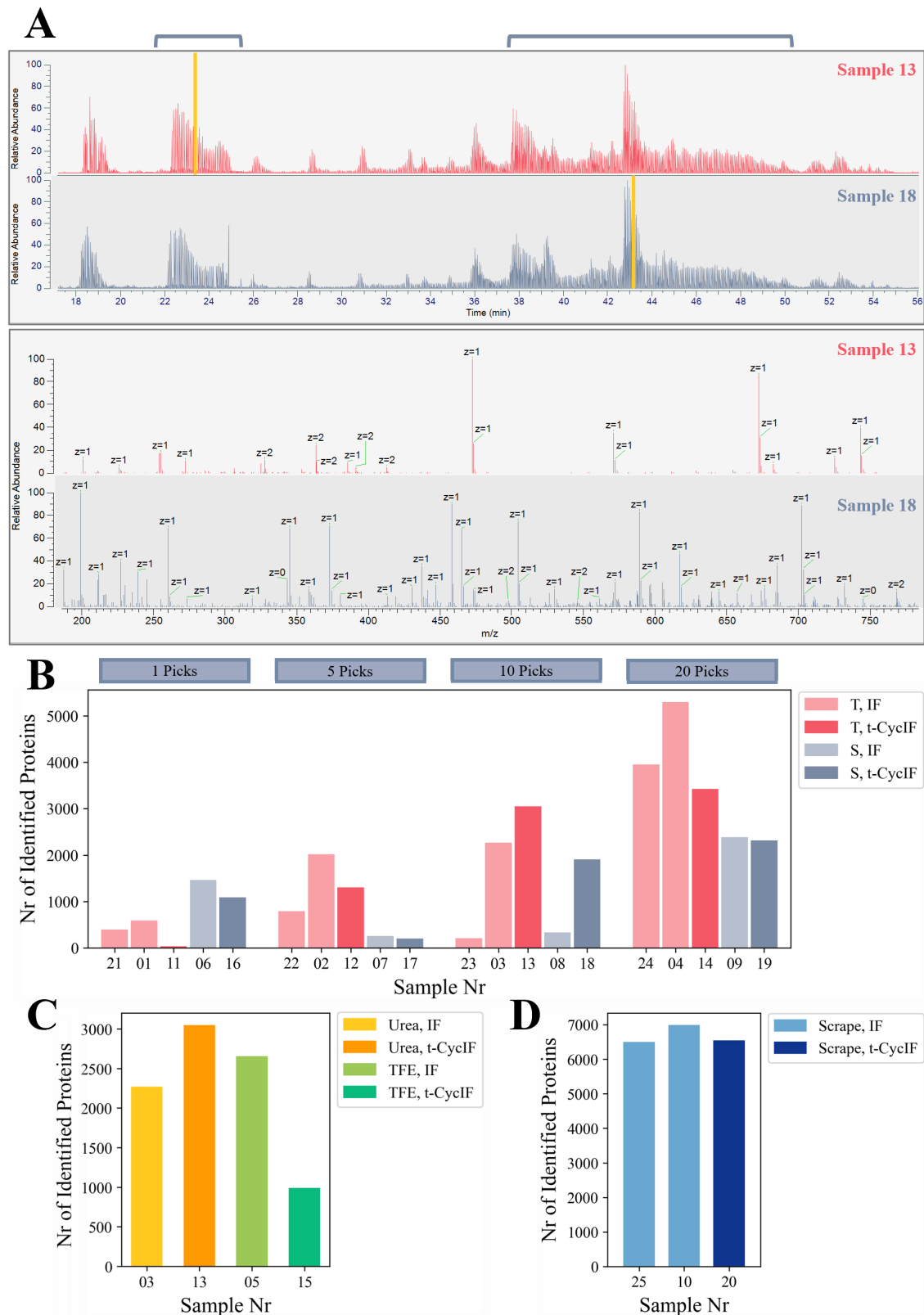


Figure 7: Proteomics data analysis results of the pilot project. **A** Representative chromatograms and spectra illustrating peptide signals and contaminants. **B** Comparison of the number of identified proteins in tumor and stroma samples with different numbers of picks using urea buffer, with either IF- or t-CycIF-staining **C** Comparison of protein yields in 10 pick samples processed with either urea or TFE buffers, with either IF- or t-CycIF-staining. **D** Protein identification count in whole-section lysate samples as a reference.

4.2 Optimization experiment

4.2.1 Experimental design

The optimization experiment was designed with the goal to improve the experimental protocol, in particular with regard to minimizing any kind of contamination. Moreover, it was constructed that based on the generated data, an analysis pipeline could be developed for analyzing image and proteomics data of various samples in a comparative way.

To avoid contaminants as far as possible, especially from polymeric materials and exogenous keratin sources like skin flakes, hair, and dust, several precautionary measures were implemented in the workflow. Glass laboratory ware was consistently used wherever possible. Prior to each experimental step, surfaces, instruments, and tools were thoroughly cleaned. Only newly prepared or unopened reagents and solutions were utilized, with all solutions and buffers stored in glass bottles to ensure no contamination from plastic particles potentially migrating into the solutions. Finally, consumables such as microscope slides and coverslips were taken from newly opened packages to further limit contamination risks. To monitor possible contaminants introduced by the microscope slide alone and contamination caused by the t-CycIF staining, two picking samples were taken from empty microscope slides as negative controls. One of those slides underwent the whole t-CycIF staining process, sample 0, while the other was freshly taken out of the package before the picking was performed, sample 00.

In addition to the two negative controls, 42 samples from six different patients, seven samples per patient, were analyzed. To improve comparability despite the inherent heterogeneity of HGSC, patient samples were selected based on uniform treatment, NACT with IDS, and consistent HR status, HRD. Additionally, tissue composition of the patient samples was checked with previous image analysis data to ensure that enough suitable picking regions matching the objective of this project were present. Based on the results of the pilot project, it was decided to cut 10 μm sections and combine 7 picks per sample in urea buffer, which approximately corresponds to the tissue amount of 10 picks from 8 μm sections. Before the sample picking, each section underwent two cycles of t-CycIF staining with the tumor marker PAX8, immune markers CD45 and Iba1, the stroma marker Vimentin, and the cell cycle markers Ki67, pRb, and p21. These biomarkers were selected to enable the distinction between tumor, TSI, and stroma tissue areas as well as the differentiation between different cell types, namely tumor cells, PAX+, stroma cells, PAX8– and Vimentin+, and immune cells, PAX8–, Vimentin+, and CD45+ or Iba1+, and contrasting cellular states like pro-

liferation, Ki67+ and p21−, non-dormancy, Ki67− and p21−, and dormancy, Ki67−, pRb−, and p21+. Based on the staining patterns, the targeted collection of proliferating and dormant tumor cells from tumor and TSI regions and non-dormant and dormant stroma cells from stroma regions was planned to generate the necessary data for the intended comparative analysis pipeline of contrasting cellular states (proliferative vs. dormant, respectively non-dormant vs. dormant). Additionally, again whole-section lysate samples were prepared from each tissue section post-picking to function as positive controls and provide a reference for protein detection levels. A depiction of the experimental design is shown in Figure 8A. Table 12 contains a summary of picking conditions and key characteristics of each analyzed tissue sample.

Table 12: Sample specifications of the optimization experiment.

Sample Nr	Patient	Staining	Region	CellState	Nr of Picks	Buffer
0	-	t-CyclIF	-	-	7	Urea
00	-	-	-	-	7	Urea
01	C331	t-CyclIF	T	P	7	Urea
02	C331	t-CyclIF	T	D	7	Urea
03	C331	t-CyclIF	TSI	P	7	Urea
04	C331	t-CyclIF	TSI	D	7	Urea
05	C331	t-CyclIF	S	N	7	Urea
06	C331	t-CyclIF	S	D	7	Urea
07	C331	t-CyclIF	Scrape	Scrape	Scrape	Urea
08	C791	t-CyclIF	T	P	7	Urea
09	C791	t-CyclIF	T	D	7	Urea
10	C791	t-CyclIF	TSI	P	7	Urea
11	C791	t-CyclIF	TSI	D	7	Urea
12	C791	t-CyclIF	S	N	7	Urea
13	C791	t-CyclIF	S	D	7	Urea
14	C791	t-CyclIF	Scrape	Scrape	Scrape	Urea
15	C369	t-CyclIF	T	P	7	Urea
16	C369	t-CyclIF	T	D	7	Urea

Sample Nr	Patient	Staining	Region	CellState	Nr of Picks	Buffer
17	C369	t-CyclIF	TSI	P	7	Urea
18	C369	t-CyclIF	TSI	D	7	Urea
19	C369	t-CyclIF	S	N	7	Urea
20	C369	t-CyclIF	S	D	7	Urea
21	C369	t-CyclIF	Scrape	Scrape	Scrape	Urea
22	C758	t-CyclIF	T	P	7	Urea
23	C758	t-CyclIF	T	D	7	Urea
24	C758	t-CyclIF	TSI	P	7	Urea
25	C758	t-CyclIF	TSI	D	7	Urea
26	C758	t-CyclIF	S	N	7	Urea
27	C758	t-CyclIF	S	D	7	Urea
28	C758	t-CyclIF	Scrape	Scrape	Scrape	Urea
29	C426	t-CyclIF	T	P	7	Urea
30	C426	t-CyclIF	T	D	7	Urea
31	C426	t-CyclIF	TSI	D	7	Urea
32	C426	t-CyclIF	TSI	P	7	Urea
33	C426	t-CyclIF	S	N	7	Urea
34	C426	t-CyclIF	S	D	7	Urea
35	C426	t-CyclIF	Scrape	Scrape	Scrape	Urea
36	C078	t-CyclIF	T	P	7	Urea
37	C078	t-CyclIF	T	D	7	Urea
38	C078	t-CyclIF	TSI	P	7	Urea
39	C078	t-CyclIF	TSI	D	7	Urea
40	C078	t-CyclIF	S	N	7	Urea
41	C078	t-CyclIF	S	D	7	Urea
42	C078	t-CyclIF	Scrape	Scrape	Scrape	Urea

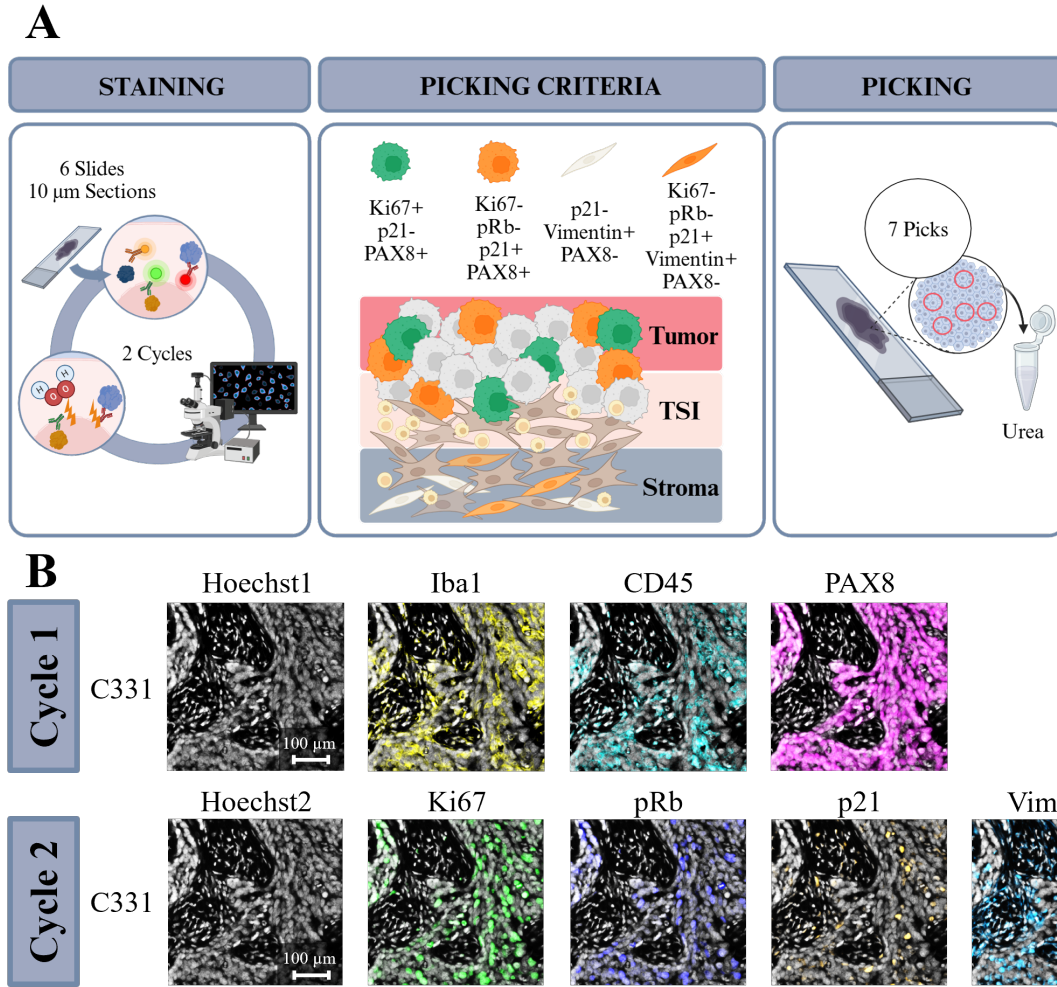


Figure 8: Overview of the experimental design of the optimization project and staining patterns of all cycles. **A** Schematic representation of the experimental setup of the pilot project. Figure created in BioRender, 2025. **B** t-CycIF staining patterns of the markers used during all staining cycles. Each image shows Hoechst staining either alone or merged with the respective markers.

4.2.2 t-CycIF staining provides tissue and cellular characterization, enabling ROI selection

During this experiment, all tissue sections were stained with two cycles of t-CycIF staining. For this, 10 μ m sections were used and a slightly adjusted t-CycIF protocol was followed, incorporating the described precautionary steps to minimize contamination. Again, an antibody panel was used that had already been validated during earlier projects of the lab. It comprised tumor, stroma, immune, and cell cycle markers.

Even with thicker tissue sections than during the pilot project and minor protocol changes, all used markers displayed a clear staining pattern with little background

noise, as shown in Figure 8B. Based on the nuclear staining channels, the stitching and registration of both cycles and the post-picking images were successful. The resulting multichannel images were accurately overlaid, showing no overlaps or missing areas. ROIs were identified using the images of the stained tissue. During the picking, the second staining cycle provided guidance for identifying pre-selected ROIs and validating expected marker combinations. The positive and negative marker combinations used for identifying cells and regions of interest are shown in Figure 9 together with examples of selected areas for tumor, TSI, and stroma samples. Based on tissue structure, revealed by nuclear staining and staining with PAX8 and Vimentin, tumor, TSI, and stroma areas were determined. ROIs for tumor samples were selected to be predominantly surrounded by tumor cells. In contrast, TSI ROIs were chosen at the border of tumor and stroma regions, encompassing a mixture of both cell types. Finally, stroma ROIs were confined to areas exclusively surrounded by stroma cells, avoiding proximity to tumor regions. In general, ROIs for one sample were selected from the same tissue region. The specific cell selection was made based on absence or presence of proliferation marker Ki67, cell cycle arrest marker p21, and pRb in tumor and stroma cells. Noteworthy, the selection of tumor and TSI ROIs was based on staining patterns of tumor cells in these areas, while stroma cell expression patterns guided the stroma ROI selection.

During ROI selection, particular care was taken with two patient samples, C369 and C758, which showed minor tissue detachment at the borders. Although these areas were limited and did not impact the overall analysis, ROIs were deliberately placed away from these detached regions.

4.2.3 Image segmentation was achieved by combining two image enhancement methods with the Mesmer nuclear segmentation model

Next step in the image processing pipeline, after registration and stitching, is segmentation. The laboratory’s standard segmentation pipeline typically utilizes the Stardist model (Schmidt et al., 2018) for nuclear segmentation. However, its performance on this dataset was suboptimal. Stardist failed to segment a substantial portion of cells accurately. This was likely due to the higher cell density and overlap in the images, caused by the use of thicker tissue sections compared to those commonly used in standard t-CycIF experiments.

To address these challenges, two additional segmentation models, Cellpose (Stringer et al., 2021) and Mesmer (Greenwald et al., 2022), were evaluated. Both models were specifically developed for biomedical image segmentation and were selected based on their performance in a recent benchmarking study comparing state-of-the-art compu-

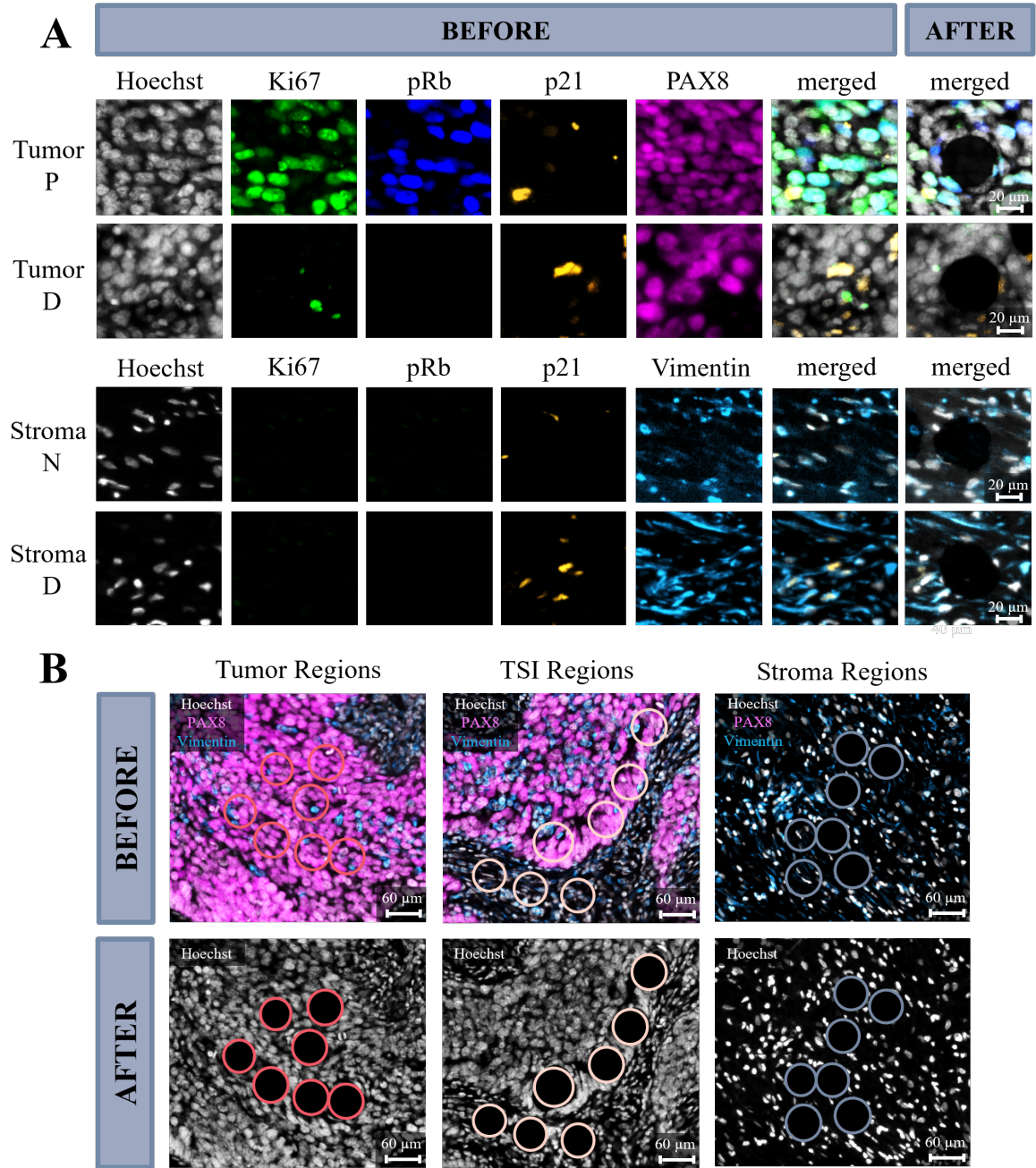


Figure 9: Picking criteria for ROI selection and examples for picked areas. **A** The ROI selection of tumor proliferative (P) and dormant (D) phase samples as well as stroma non-dormant (N) and dormant (D) phase samples is based on specific positive and negative marker combinations. After the picking, a hole remains in the tissue. **B** Examples of picked samples from tumor, TSI, and stroma regions before and after the picking process.

tational segmentation methods (Wang et al., 2024). In this study, Cellpose demonstrated strong overall performance across multiple evaluation metrics, while Mesmer ranked lower in most categories but showed particular strength in scalability. In contrast, Stardist, the laboratory’s standard model, exhibited a moderate performance, placing near the middle of the ranked methods. Due to considerations of time efficiency, the comparison was conducted on a representative image crop rather than whole-slide images. As illustrated in Figure 10A, Mesmer outperformed both Stardist and Cellpose, successfully segmenting the majority of cells visible in the sample. Notably, Cellpose performed the poorest, primarily segmenting well-separated cells but struggling with densely clustered cell regions - an outcome that contrasts with its strong performance reported in the benchmarking study. Based on these results, Mesmer was selected as the primary segmentation tool for subsequent analysis.

However, despite the improved segmentation performance, Mesmer’s segmentation remained imperfect, in particular for cells exhibiting low signal intensity, indistinct boundaries, or located in less focused image planes. To further enhance segmentation accuracy, image processing techniques aimed at improving intensity and contrast were implemented prior to segmentation. Four methods were evaluated: contrast limited adaptive histogram equalization (CLAHE), gamma correction, unsharp masking, and Richardson-Lucy deconvolution.

CLAHE enhances local contrast in an image by improving the visibility of details within small regions. It works by dividing the image into non-overlapping tiles, the size of which is defined by the kernel size parameter, where smaller kernels create more tiles and allow for more localized contrast adjustments. Within each tile, a pixel intensity histogram is computed and redistributed into a fixed number of bins, which determines the resolution of intensity levels used. To prevent excessive contrast amplification and noise, a clip limit restricts the maximum height of histogram bins - any counts above this limit are clipped and then evenly redistributed across all bins (Zuiderveld et al., 1994).

In contrast, gamma correction is a method that modifies image brightness through a nonlinear adjustment of pixel values by raising them to the power of γ . For $\gamma < 1$, darker pixels are brightened, enhancing contrast and revealing details in shadowed regions. Conversely, for $\gamma > 1$, the image darkens, compressing highlights and increasing contrast in bright areas.

Unsharp masking enhances image sharpness by subtracting a blurred (unsharp) version of the image from the original, which increases the contrast of edges and fine details. The blurring is performed using a Gaussian filter, where the parameter σ

defines the standard deviation of the Gaussian function and thus controls the spread of the blur. A larger σ results in a smoother, more widely spread blur. How much of the blurred image is subtracted from the original is determined with the strength parameter. This parameter acts as a scaling factor that controls the intensity of the sharpening effect: higher strength values increase the contribution of the blurred image subtraction, leading to more pronounced edge enhancement, while lower values produce a subtler sharpening effect (Gonzalez & Woods, 2018).

Lastly, Richardson-Lucy deconvolution is a method that aims to reduce image blur and sharpen details. It relies on a point spread function, which models how the image was blurred during imaging. Here, a Gaussian function with standard deviation σ was used to represent the blur level. The process is repeated for a set number of iterations, where more iterations lead to a clearer, sharper image but also require more processing time (Lucy, 1974; Richardson, 1972).

Parameter values for each image processing method were determined through empirical testing. This involved the visual assessment of Mesmer segmentation results on images processed with the different enhancement methods and various parameter combinations to ultimately identify those yielding the most accurate segmentation appearance. The optimized parameter settings were as follows: CLAHE with kernel size = 8×8 , bin number = 200, and clip limit = 0.03, gamma correction with $\gamma = 2/3$, unsharp masking with $\sigma = 1.0$ and strength = 2.5, and Richardson-Lucy deconvolution with $\sigma = 1.0$ and 30 iterations.

The final Mesmer segmentation results using these optimized parameters are presented in Figure 10B. The different enhancement approaches showed only marginal differences in performance. While CLAHE combined with Mesmer performed similarly to Mesmer alone, likely because Mesmer’s internal workflow already incorporates a CLAHE step, gamma correction, unsharp masking, and Richardson-Lucy deconvolution achieved slightly better segmentation accuracy than Mesmer alone. Ultimately, computational efficiency alongside accuracy guided the choice of enhancement method. Although Richardson-Lucy improved image sharpness with additional iterations, the increased computational time did not proportionally translate into improved segmentation. In contrast, gamma correction and unsharp masking provided a balance of enhanced segmentation performance and faster processing times. Therefore, a combination of both was adopted, leveraging their complementary effects - gamma correction brightens darker regions to improve intensity, while unsharp masking enhances image sharpness. This combination yielded a satisfactory segmentation accuracy, as demonstrated in Figure 10C.

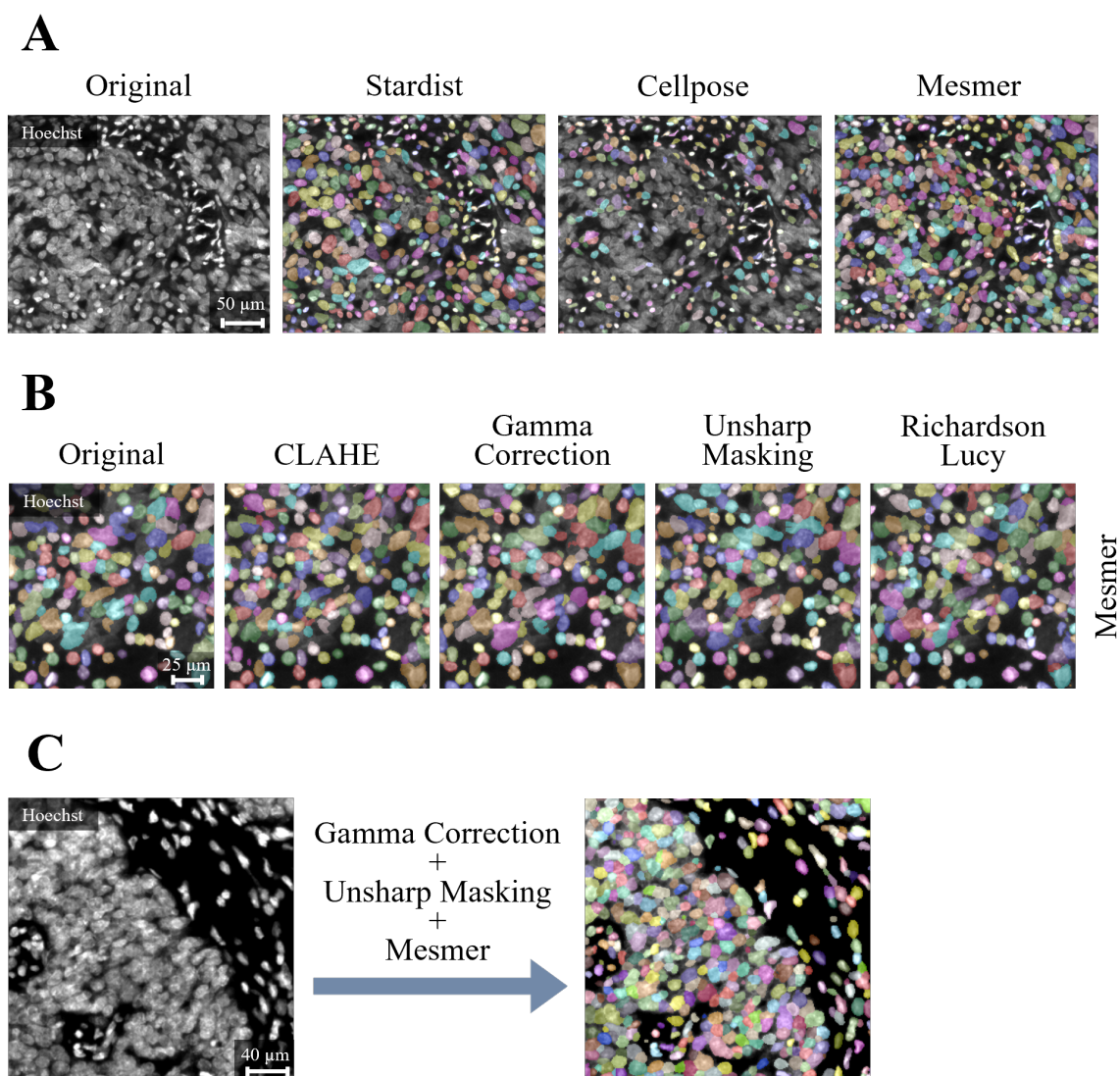


Figure 10: Overview of different segmentation model performances and image enhancement techniques. **A** Comparison of segmentation performance on a representative image crop showing the nuclear staining channel alongside overlays of masks generated by Stardist, Cellpose, and Mesmer. **B** Segmentation performance of Mesmer on the original nuclear staining image compared to its performance on images enhanced using CLAHE (kernel size = 8×8 , bin number = 200, clip limit = 0.03), gamma correction ($\gamma = 2/3$), unsharp masking ($\sigma = 1.0$, strength = 2.5), and Richardson-Lucy deconvolution ($\sigma = 1.0$, 30 iterations). Comparison is made by overlaying nuclear staining with segmentation masks. **C** Segmentation outcome using a combined approach of gamma correction and unsharp masking with the nuclear segmentation model Mesmer.

Following the segmentation optimization performed on image crops, efforts were made to adapt the workflow for whole-slide images. The Mesmer model, adapted from the Van Valen Lab's DeepCell Applications repository GitHub, was unable to process large tissue areas directly. Therefore, whole-slide images were tiled prior to segment-

ation, with each tile being segmented separately and subsequently being reassembled. A workflow incorporating these steps was developed and operated as intended; however, three major issues arose:

First, the computational runtime required for segmenting a single whole-slide image was very long, exceeding more than three days, rendering this approach impractical for processing multiple samples within a reasonable time frame. Second, the resulting cell masks lacked unique identifiers - individual cell IDs, which typically distinguish cells from one another, were partly assigned to multiple cells. This overlap would have posed significant challenges for downstream processing and analysis. Lastly, although segmentation of the tissue and regions surrounding it generally achieved satisfactory accuracy, the image borders, black background areas completely devoid of cells, exhibited extensive over-segmentation, generating numerous near pixel-sized masks, which would have complicated downstream analysis.

Due to these limitations, whole-slide image segmentation was deemed unsuccessful. As a practical alternative, rectangular regions encompassing all selected ROIs for each sample were extracted from the whole-slide images. A total of six small crops were extracted from each tissue section, two each from tumor, TSI, and stroma regions, with one representing a proliferating/non-dormant area and one a dormant area, resulting in 36 crops overall. All subsequent processing and analysis steps were conducted using these small crops of the whole slide images rather than the entire tissue section images.

4.2.4 Quantification and phenotyping enable single-cell characterization of cells within ROIs

To obtain detailed single-cell information, small crops containing selected ROIs were quantified and phenotyped. Cells located outside of ROIs were excluded from further analysis to specifically focus on those captured during the tissue picking process. Cell type annotations were based on seven markers (see Figure 11A) and phenotyping was performed using the Scimap toolkit (see chapter *Image processing, phenotyping, analysis, and visualization*). The accuracy of this classification was thoroughly validated through visual inspection with a customized Napari application, which enabled overlaying cell annotation masks onto the t-CycIF image crops. By visualizing multiple markers simultaneously, different cell types were clearly distinguished. The assigned phenotypes corresponded well with expected marker profiles, as demonstrated in Figure 11B.

The tumor marker PAX8 served as the primary marker to distinguish tumor cells from non-tumor populations. Both stroma and immune cells were identified based on

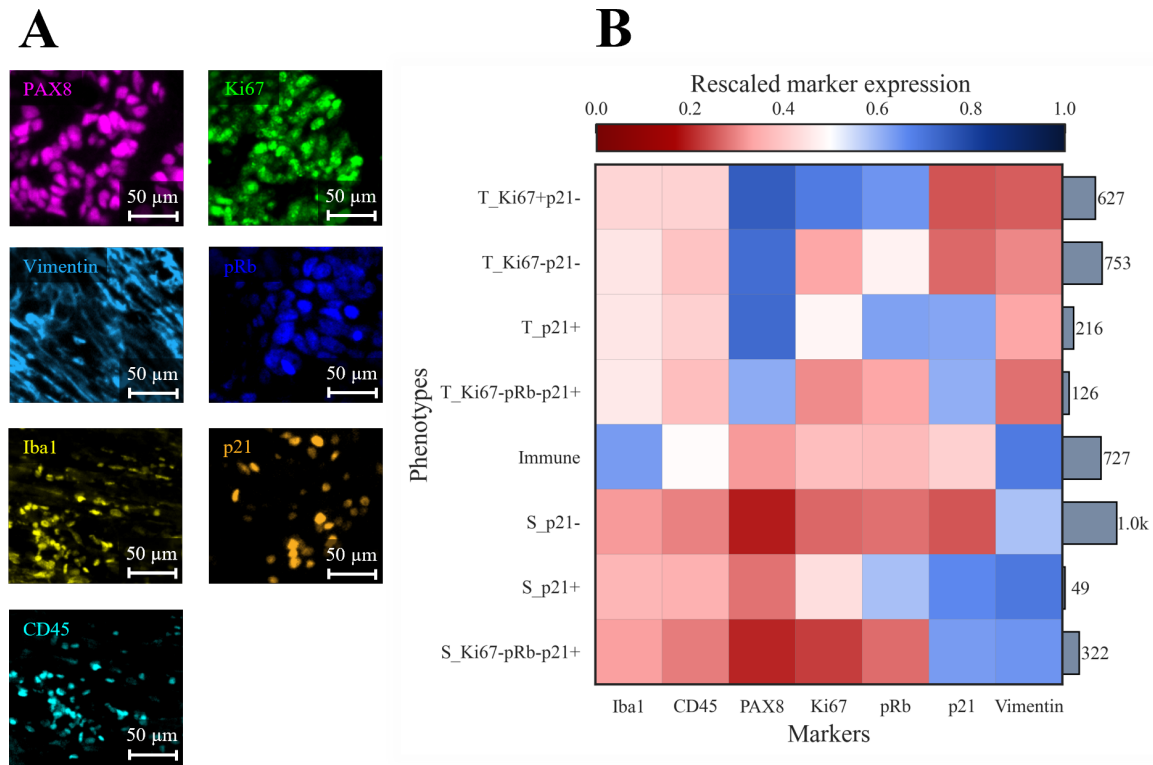


Figure 11: Overview of the cell type annotation marker panel and marker expression patterns of the various phenotypes. **A** Phenotyping was based on seven selected biomarkers, each exhibiting distinct expression patterns. **B** Heatmap illustrating the different marker expression profiles of the distinct phenotypes within the ROIs as well as the total abundance of the different cell types within all picked samples. For better comparability, log transformed and rescaled marker expression values were used (log transformation and rescaling was performed using Scimap functions). Expression values above 0.5 are counted as positive for the specific marker.

Vimentin expression, and further distinguished by the immune cell markers Iba1 and CD45: immune cells were positive for one or both, while stroma cells were negative for both.

Tumor and stroma cells were further characterized based on their cell cycle marker profiles, with a particular focus on proliferating, non-dormant, and dormant cells. Tumor cells expressing the proliferation marker Ki67 and lacking the cell cycle arrest marker p21 were classified as proliferating. Cells negative for both Ki67 and p21 were categorized as an intermediate phenotype, neither clearly proliferating nor dormant. Arrested tumor cells were defined in two subgroups: one more inclusive group required only p21 positivity, while a stricter group required p21 positivity combined with negativity for both Ki67 and pRb. In the stromal compartment, Ki67 expression was rare. As a result, the analysis focused less on proliferating stroma cells and instead emphasized the identification of non-dormant cells, those lacking p21 expres-

sion. Dormant stroma cells were again split into two phenotypes: one defined by p21 positivity alone, and a stricter variant requiring additional negativity for Ki67 and pRb. As can be seen in the heatmap in Figure 11B, tumor and stroma phenotypes defined solely by p21 positivity, also showed a moderate rescaled mean expression of pRb, while the rescaled Ki67 expression was, on average below 0.5 and therefore regarded as negative, supporting their classification as arrested but not proliferative. In total, the 3847 cells located within ROIs were identified and phenotyped in the described way. Of these, 1,722 (44.8 %) were classified as tumor cells, 727 (18.9 %) as immune cells, and 1,398 (36.3 %) as stroma cells. Their exact distribution between the different phenotypes is shown in Figure 11B.

In conclusion, the applied phenotyping approach enabled accurate and spatially resolved identification of diverse cell types and functional states within the picked tissue regions, laying the foundation for detailed downstream analysis.

4.2.5 Compositional analysis facilitates accurate sample categorization

Building on the phenotyping results, a compositional analysis was performed to evaluate the cellular makeup of each individual sample. This allowed for an overview of how successfully different cell types were targeted across picked regions and how their relative abundances varied by tissue context and cell state.

Samples were categorized based on both the tissue region they were picked from, tumor, TSI, and stroma as well as the cellular state that was targeted, proliferating or dormant for tumor and TSI, and non-dormant or dormant for stroma. This yielded six distinct groups of samples, organized into three main categories based on the region of origin of the picked cells. Each was comprised of two subgroups based on cellular state, enabling structured comparison across biological conditions.

As shown in Figure 12A, the distribution of the three major cell types, tumor, stroma, and immune cells, varied noticeably across sample groups. Tumor samples were generally enriched in tumor cells, with the majority of samples containing more than 50 %. Nonetheless, all tumor samples also included a notable fraction of stroma and immune cells. As intended during ROI selection, TSI samples consistently contained a mixture of tumor, stroma, and immune cells. The relative abundance of these populations varied substantially across samples: tumor cell content ranged from approximately 20 % to nearly 70 %, stroma cells from around 15 % to over 50 %, and immune cells from below 10 % to roughly 50 %. Compared to tumor and stroma samples the cellular composition of TSI samples exhibited the greatest compositional heterogeneity. Stroma samples demonstrated the highest specificity. All samples

contained a predominant proportion of stroma cells, while tumor cells were largely absent, typically making up less than 5 % of cells when present at all. Notably, subtle trends were observed within the phenotypic subgroups. Tumor samples classified as dormant tended to contain a lower proportion of tumor cells than their proliferating counterparts. A similar pattern was seen in stroma samples, where dormant samples included fewer stroma cells on average than non-dormant ones. Additionally, tumor samples more frequently exhibited stroma contamination, whereas stroma samples were rarely contaminated by tumor cells. Immune cells were detected across most samples, often in varying proportions, despite not being specifically targeted during ROI selection.

Taken together, these findings suggest that ROI selection was generally effective in capturing the intended tissue regions. When considering overall cell identity rather than specific phenotypic states, the resulting samples reflected the expected enrichment of tumor, stroma, or mixed cell types according to their region of origin.

A complementary analysis examined the number of cells per sample, grouped by tissue region, see Figure 12B. Tumor samples exhibited the highest cell counts per sample and the greatest variability. Stroma samples, by contrast, showed consistently lower cell counts with minimal variation across samples. TSI samples were more similar to tumor samples in both cell count and variability, although they generally contained slightly fewer cells and exhibited a more constrained distribution.

Beyond the overall cell type composition and analysis of the number of picked cells, a more detailed analysis was performed to investigate the distribution of phenotypic states within each targeted cell population across the samples. For tumor and TSI samples, the analysis focused on tumor cells only, while for stroma samples it was restricted to stroma cells. Additionally, for exploratory purposes, stroma cell phenotypes within TSI samples were also assessed, even though they were not specifically targeted during ROI selection (figures 12C and D). To simplify interpretation, both dormant phenotypes, p21 positive alone and the stricter variant defined by additional Ki67 and pRb negativity, are considered together in the following analysis.

In five of the six tumor samples initially categorized as proliferating, samples 1, 8, 15, 22, and 36, the majority of tumor cells exhibited a proliferative phenotype, each exceeding 50 %. Dormant cells in these samples typically accounted for 5 % to 15 %, with one outlier reaching approximately 30 %. The remaining proportion of tumor cells in all samples consisted of cells negative for both Ki67 and p21, whose cell state could not be further specified based on the available biomarkers. Only sample 29,

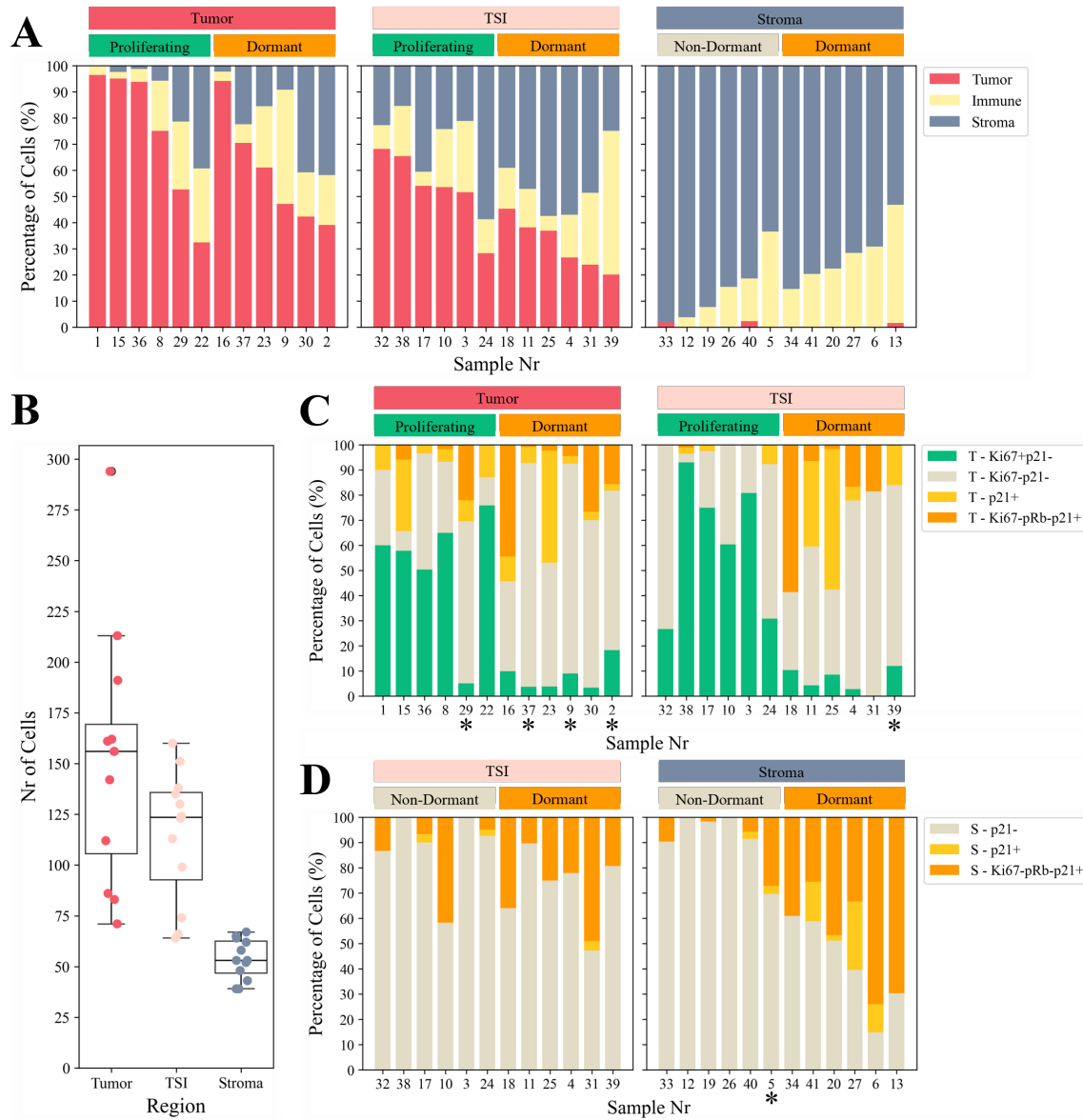


Figure 12: Overview of cell type and phenotype composition across all analyzed samples. **A** Cell type composition across all samples grouped by tissue region and targeted cell state. **B** Boxplots displaying the distribution of total cell counts within all picked samples grouped by region of origin. **C** Distribution of tumor cell phenotypes within tumor cells across tumor and TSI samples. **D** Distribution of stroma cell phenotypes within stroma cells across stroma and TSI samples. Asterisks in **C** and **D** mark samples that were relabeled based on compositional analysis.

originally classified as proliferating, displayed a markedly low proportion of proliferating cells around 5 % and a higher dormant fraction, approximately 30 %. Based on this phenotypic distribution, it was reclassified into the dormant group. Among the six samples initially labeled as dormant, three, samples 16, 23, and 30, showed clear enrichment of the dormant phenotype, comprising 30 % to 55 % of tu-

mor cells. Proliferating tumor cells in these samples did not exceed 10 %, and the remainder were Ki67–p21– tumor cells without further specification. The other three samples, 2, 9, and 37, exhibited less than 20 % dormant cells and comparable or higher levels of proliferating cells. They lacked a clear enrichment of either phenotype and were therefore reassigned to a non-specific sample group.

Notably, the proportion of proliferating tumor cells in proliferating samples was, on average, higher than the proportion of dormant cells in the dormancy group. This imbalance should be taken into account in subsequent analysis steps.

All six TSI samples initially categorized as proliferating, samples 3, 10, 17, 24, 32, and 38, were confirmed to be dominated by proliferating tumor cells. Four samples exhibited proliferative fractions exceeding 60 %, while the remaining two showed a more moderate enrichment of around 30 %. Dormant tumor cells, when present, generally accounted for less than 10 %. As in the tumor samples, the remaining tumor cells were Ki67–p21– and could not be phenotypically specified.

Among the six TSI samples categorized as dormant, five showed an enrichment of the dormant phenotype. Of these, sample 11, 18, and 25 showed a clear enrichment of dormant tumor cells, with dormant fractions ranging from approximately 40 % to 60 % and only 5 % to 10 % proliferating cells. In contrast, the other two samples, 4 and 31, exhibited lower absolute proportions of dormant cells, around 20 % to 25 %. Nevertheless, they still showed a dominant dormant phenotype due to the near-complete absence of proliferating tumor cells in the samples. Only one sample, sample 39, showed similar proportions of proliferating and dormant tumor cells, approximately 15 % each, and was therefore reclassified as non-specific.

Although stromal cells in TSI samples were not specifically targeted, the analysis revealed that several of the dormant-labeled samples, 4, 11, 18, 25, 31, and 39, contained moderate to substantial fractions of dormant stromal cells, ranging from 10 % to nearly 50 %. In contrast, the proliferating-labeled TSI samples differed more, with some samples containing no dormant cells at all, and other samples containing up to 40 % of dormant cells.

The analysis of stromal cell phenotypes showed that five of the six samples initially categorized as non-dormant, samples 12, 19, 26, 33, and 40, contained indeed mostly non-dormant stroma cells, with dormant cells comprising less than 10 % of the stromal population. One sample, 5, however, exhibited a dormant stroma cell fraction of approximately 30 %, a level more comparable to the dormant group than the other non-dormant samples, though still noticeably lower. As it did not clearly align with either category, it was reclassified as non-specific.

Conversely, all six stroma samples originally labeled as dormant showed clear enrich-

ment of dormant stroma cells, with proportions exceeding 40 % and reaching over 80 % in some cases. Notably, the stricter dormancy phenotype, characterized by Ki67 and pRb negativity and p21 positivity, accounted for at least 50 % of dormant stroma cells in each of these samples, and in several cases even made up the entire dormant stroma cell population.

Overall, the phenotypic composition analysis confirmed successful enrichment of the intended cell states in the majority of samples, establishing a solid basis for subsequent sample analysis. Moreover, it provided a basis for relabeling the few samples with a composition that did not match their original label.

4.2.6 Generated proteomics data of picked samples enables development and application of an integrated analysis pipeline combining protein expression and cellular composition

Building on the validated sample composition, proteomics analysis was performed to evaluate the molecular profiles of the enriched cellular states.

First, similar to the pilot experiment, a chromatographic evaluation of the MS runs was conducted using Freestyle. Compared to the pilot experiment, a markedly lower number of overall signals was observed in the chromatograms across all runs of all picked samples, see Figure 13A. While a certain reduction was expected and favorable, indicating successful implementation of improved sample preparation protocols for reducing impurities, it also reflects a general decrease in peptide-derived signals across all sample types. This was consistent in all types of picked samples, proliferating, non-dormant, and dormant samples from tumor, TSI, and stroma regions.

Subsequent analysis using Spectronaut confirmed the limited protein detection. Across all picked samples, the number of identified proteins ranged from fewer than 200 up to 1400-1500 proteins, with a substantial proportion of samples falling below 1000 identified proteins, as shown in Figure 13B. In this experiment, 7 pick samples from 10 μ m tissue sections were analyzed, which is comparable to the 10 pick samples from 8 μ m sections in the pilot study. Although the pilot dataset was small, with only seven 10-pick samples in total, and showed considerable variability, most samples in that study had protein yields between 1000-3000 proteins, with two outliers below this range. Therefore, based on those results, a higher protein yield would have been anticipated in the current samples.

As with the pilot project, positive samples in the form of full-section scrapes were taken for this experiment. These controls served to confirm the overall presence and abundance of proteins in the tissue, providing a reference point for interpreting the protein yields obtained from smaller, spatially restricted regions. The full-section

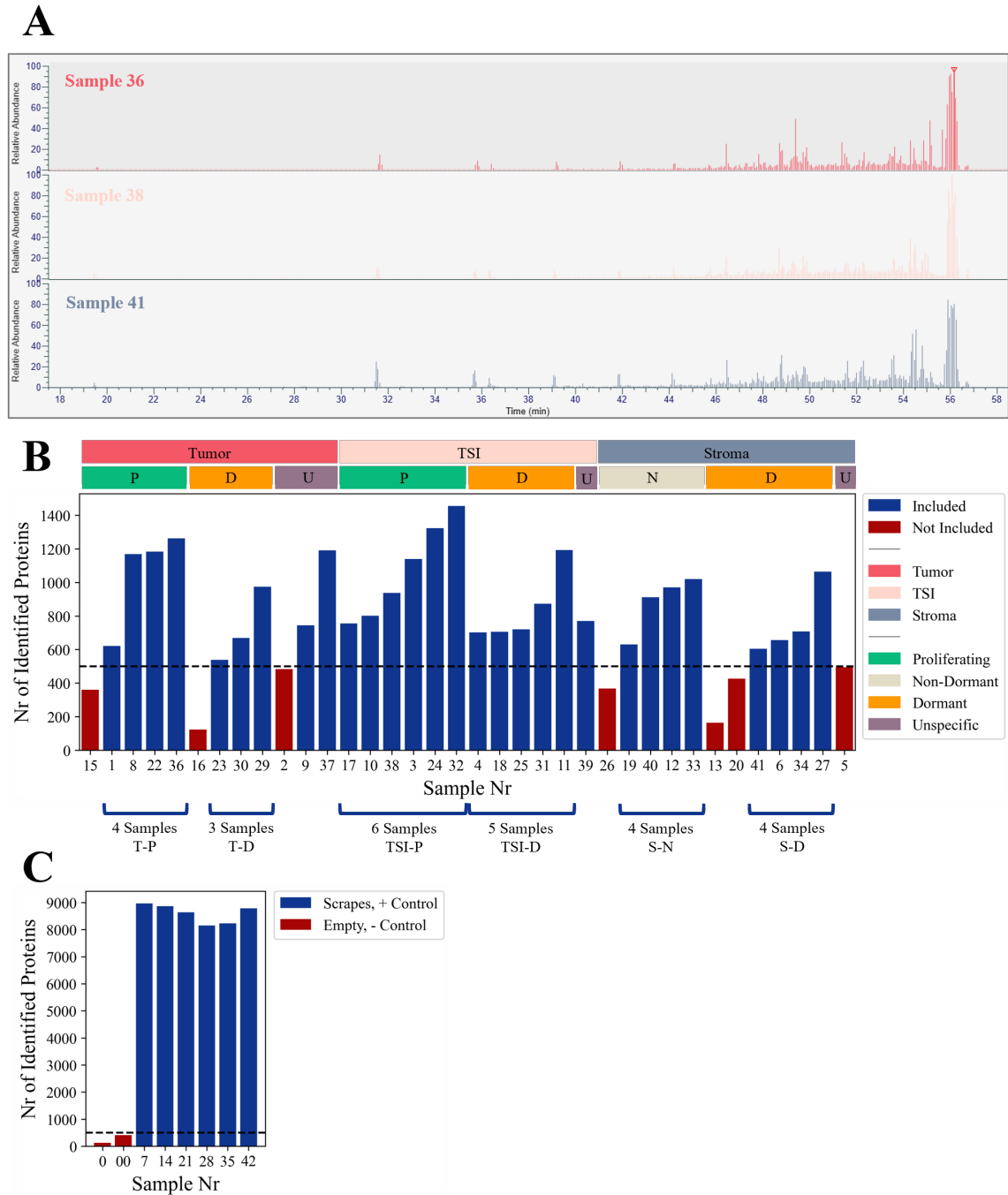


Figure 13: Overview of proteomics data quality and yielded protein detection numbers. **A** Representative chromatograms from three samples, a proliferating tumor sample, a proliferating TSI sample, and a dormant stroma sample, illustrating the similar protein detection across sample types. **B** Number of identified proteins per sample, with samples excluded from downstream analysis highlighted in red due to protein counts below 500 and purple due to unspecific cellular state. The horizontal dashed line indicates the 500 protein threshold for sample inclusion. **C** Protein yield in positive and negative control samples.

scrapes exhibited robust protein detection, with 8000 to 9000 proteins identified per

sample, which is even higher than the pilot project results. Negative control samples were also analyzed to assess background protein identification. The negative control picked from an empty slide that underwent the full t-CycIF staining process, sample 0, yielded approximately 150 identified proteins, which is expected due to unavoidable background contamination. However, the negative control obtained from a microscope slide taken directly from the package, sample 00, showed a higher protein yield, with close to 500 identified proteins. This level of detection is comparable to some of the low-yielding picked samples. Both negative and positive control protein yields are depicted in Figure 13C.

For subsequent analysis steps, samples that had been relabeled as unspecific based on the compositional analysis were excluded. Additionally, samples with fewer than 500 identified proteins were removed from the dataset due to their low protein yield, which was comparable to the negative controls and therefore considered unreliable. Following these inclusion and exclusion criteria, each sample group retained between 3 and 6 samples. These criteria and the final sample distribution are summarized in Figure 13B, where samples with red bars indicate those excluded for low protein counts (below the 500-protein threshold), and samples labeled as “U” (unspecific, purple) were excluded due to lack of clear enrichment for proliferating/non-dormant or dormant tumor or stroma cell states. Despite the exclusion of low-quality and unspecific samples, the following analysis should be interpreted with caution, as their statistical robustness and significance may be limited due to the relatively low sample number and the lower-than-expected protein yields. Nevertheless, the analysis was carried out to establish and demonstrate a pipeline for future proteomics investigations using similarly processed, spatially resolved samples.

To explore the protein detection patterns across samples, a cluster heatmap was generated based on the binary presence or absence of proteins, as illustrated in Figure 14A. Detected and identified proteins are shown in blue, while proteins not detected appear in white. Both samples and proteins were clustered based solely on shared detection profiles. In this context, non-detected proteins were not considered informative for clustering, as absence of detection does not necessarily indicate a true absence of the protein in the sample. It may also reflect levels below the MS detection threshold.

The heatmap reveals several distinct patterns in protein detection across the analyzed samples. A core set of proteins was consistently identified in nearly all samples. In contrast, a subset of proteins was sporadically detected across samples without a consistent distribution.

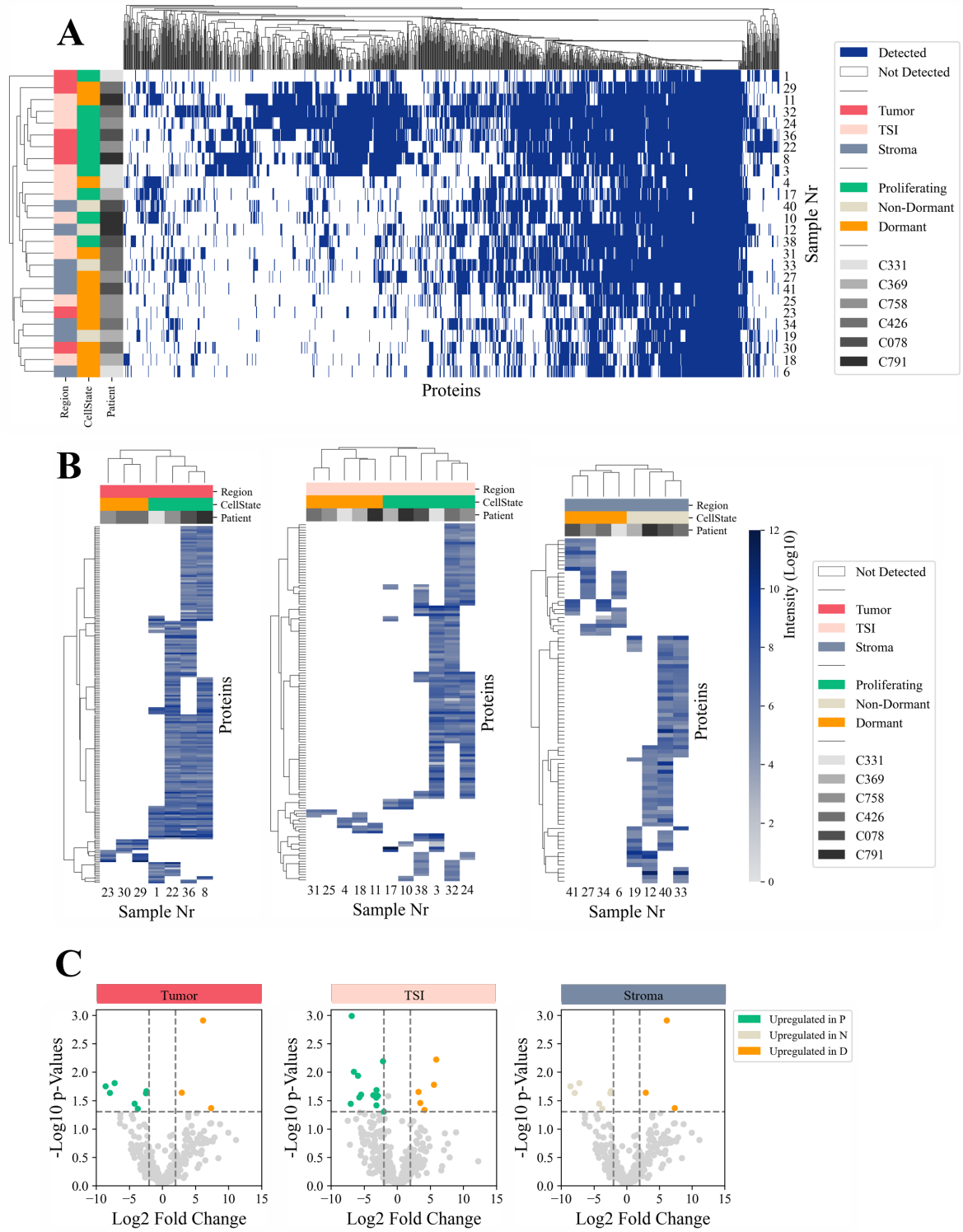


Figure 14: Proteomics data exploration and DEA. **A** Cluster heatmap showing binary detection profiles of proteins across samples. Samples and proteins are clustered based on shared detection patterns. **B** Visualization of proteins excluded from the DEA. Heatmaps contain only proteins detected in more than one sample, exclusively in one sample group. Samples and proteins are clustered based on shared detection patterns. **C** Volcano plots depicting results of DEA for each tissue region. Proteins with $|\log_2 \text{fold change}| > 2$ and $p\text{-value} < 0.05$ ($-\log_{10} p\text{-value} > 1.3$) are highlighted as significantly differentially expressed.

Notably, a considerable number of proteins was detected predominantly in proliferating samples but was largely absent in dormant ones. Dormant samples also generally showed a slightly reduced number of detected proteins in comparison to the proliferation samples.

Clustering of the samples based on their detection profiles showed partial grouping according to both tissue type and proliferation status. There was no clustering of samples based on the patient sample from which they were derived. Proliferating and dormant samples as well as tumor/TSI and stroma-derived samples, tended to form distinct clusters, though some outliers were observed. Some dormant tumor or TSI samples clustered more closely with dormant stroma samples, and some non-dormant stroma samples grouped with tumor/TSI samples.

Following the clustering analysis, a DEA was conducted to compare protein abundance between sample categories. For this, samples were grouped first by tissue region (tumor, TSI, and stroma), and within each region further subdivided by cellular state (proliferating/non-dormant vs. dormant).

This analysis compares protein intensities between two sample categories using unpaired two-sided t -tests to assess statistical differences in expression levels. Due to the requirements of this test, only proteins detected in both groups and in more than one sample per group were included. This is essential, as statistical comparisons such as the t -test require numerical input from each group to calculate differences. In MS data, the absence of detection does not represent a true zero value and therefore cannot be substituted into the analysis. Additionally, proteins detected in only a single sample within one group may reflect random or technical variation rather than a consistent biological signal. Including such proteins would risk introducing noise rather than capturing meaningful differences. These filtering criteria ensure that the analysis is based on consistent, quantifiable data rather than isolated or potentially artifactual detections.

As a result of these criteria, many proteins were excluded from the DEA, especially those detected exclusively in one group. This is a notable limitation, particularly given the earlier heatmap results (Figure 14A), which suggested that certain proteins are detected almost exclusively in proliferating or dormant samples. These group-specific proteins, although excluded from DEA, were visualized separately in heatmaps in Figure 14B, where $\log_{10}$ -transformed intensities are shown for proteins detected in only one group in more than one sample. While they were not formally tested for differential expression due to statistical constraints, they represent a set of biologically interesting proteins that demand further investigation beyond this thesis's scope.

DEA was performed separately for each tissue region group, comparing the defined sample categories. The results are summarized in Figure 14C in volcano plots, where the x axis represents the $\log_2$ fold change between groups, and the y axis shows the $-\log_{10}$ of the p -values from the unpaired two-sided t -tests. Proteins with an absolute $\log_2$ fold change greater than 2 and a $-\log_{10} p$ -value greater than 1.3, corresponding to a p -value below 0.05, were considered significantly differentially expressed and classified as upregulated in either proliferating or dormant tumor cells, or in non-dormant or dormant stroma cells.

To address the issue of false positives due to multiple hypothesis testing, the false discovery rate correction was applied using the Benjamini-Hochberg method. However, all adjusted q -values exceeded 0.05, indicating no proteins remained statistically significant after correction. However, given the small sample sizes and the exploratory nature of this study, primarily focused on establishing a functional analysis pipeline, further analysis steps were based on unadjusted p -values. While this approach does not satisfy stringent statistical criteria, it can be accepted since the main goal is method development rather than definitive biomarker discovery.

In addition to the volcano plots, results of the DEA were visualized in heatmaps in Figure 15. They show the expression patterns of proteins identified as differentially expressed for each tissue region. To facilitate comparison both within and between proteins, measured intensities were min-max scaled, normalizing each protein's expression range from 0 to 1. This scaling overcomes the limitation of measured intensities, which are only comparable across samples for a single protein but not between different proteins due to varying abundance and detection sensitivity. The heatmaps were clustered based on scaled intensities, grouping samples primarily according to their cellular state, proliferating/non-dormant or dormant, and segregating proteins into two main clusters representing those upregulated in one cell state versus the other.

To evaluate the relationship between protein expression and cellular composition, Pearson correlation coefficients were calculated between $\log_{10}$ -transformed intensities of proteins identified as significantly upregulated in the DEA and the proportions of corresponding cell states within each sample. For instance, proteins elevated in proliferating tumor cells were analyzed in relation to the fraction of proliferating cells within the tumor and TSI samples. Importantly, cell state proportions were calculated relative to their specific cell compartments rather than total cell counts, as illustrated in Figure 12C and D. As illustrated in Figure 16A, several of the proteins exhibited statistically significant correlation ($p < 0.05$) between protein intensity and cellular proportion. Notably, proteins upregulated in samples from tumor regions

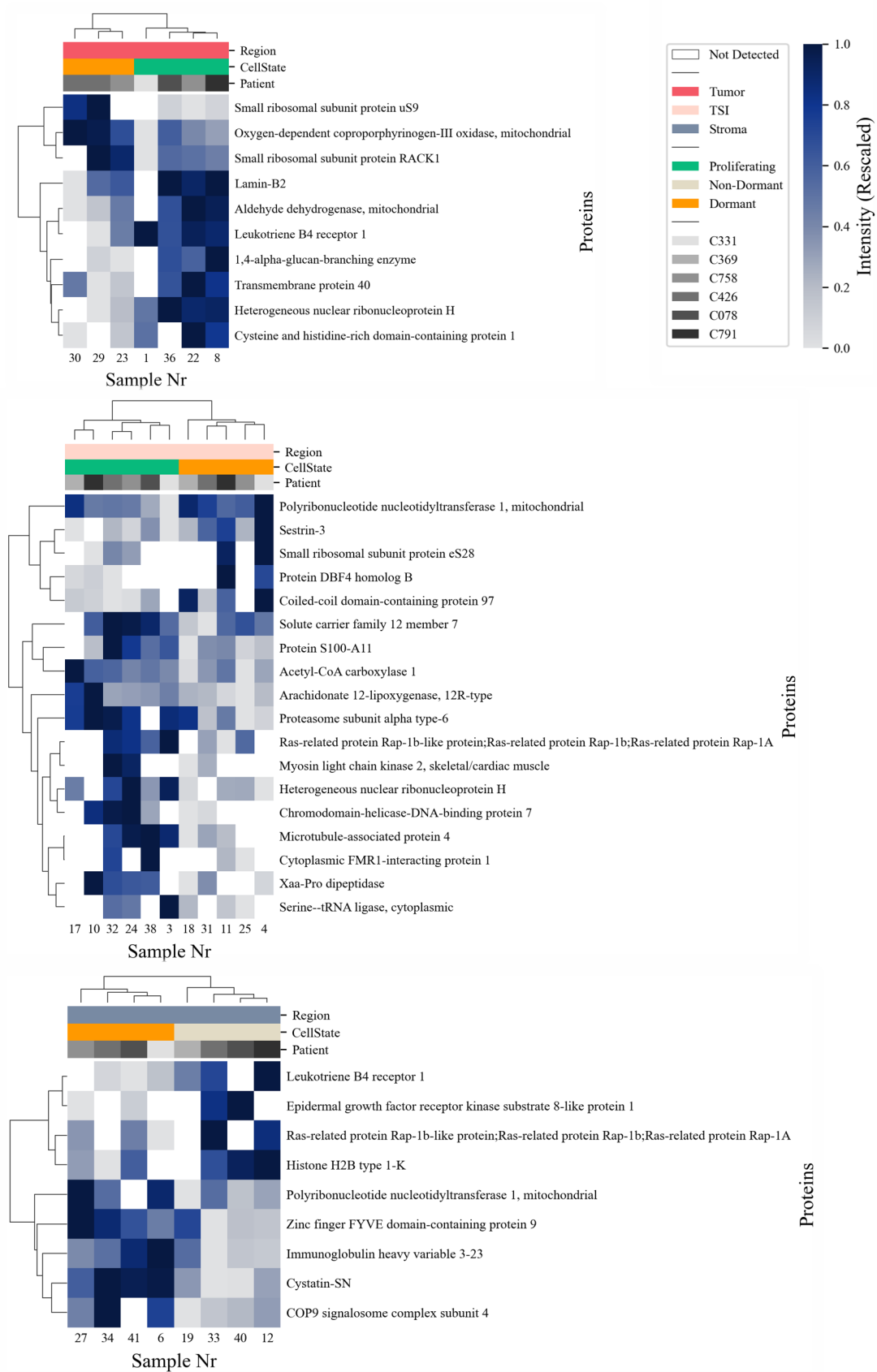


Figure 15: Heatmaps of differentially expressed proteins in tumor, TSI, and stroma regions. Intensities are min-max scaled and clustering was based on rescaled protein intensities.

showed strong correlations, with Pearson correlation coefficients consistently above 0.8. In stroma samples, upregulated proteins demonstrated robust correlations with coefficients above 0.7. Most proteins upregulated in TSI samples also had coefficients above 0.7, except for two with values between 0.6 and 0.7, all in all indicating moderate to strong relationships. Figure 16B depicts representative scatter plots for each tissue region and cell state, showing significant positive correlations between protein intensity and cellular proportion.

Overall, the analysis presented here establishes and applies a comprehensive proteomics pipeline that integrates spatially resolved protein expression data with cellular composition measurements. This pipeline encompasses chromatographic quality assessment, protein identification and quantification using Spectronaut, visualization of expression patterns, DEA, and integration with spatial cellular composition via correlation analysis.

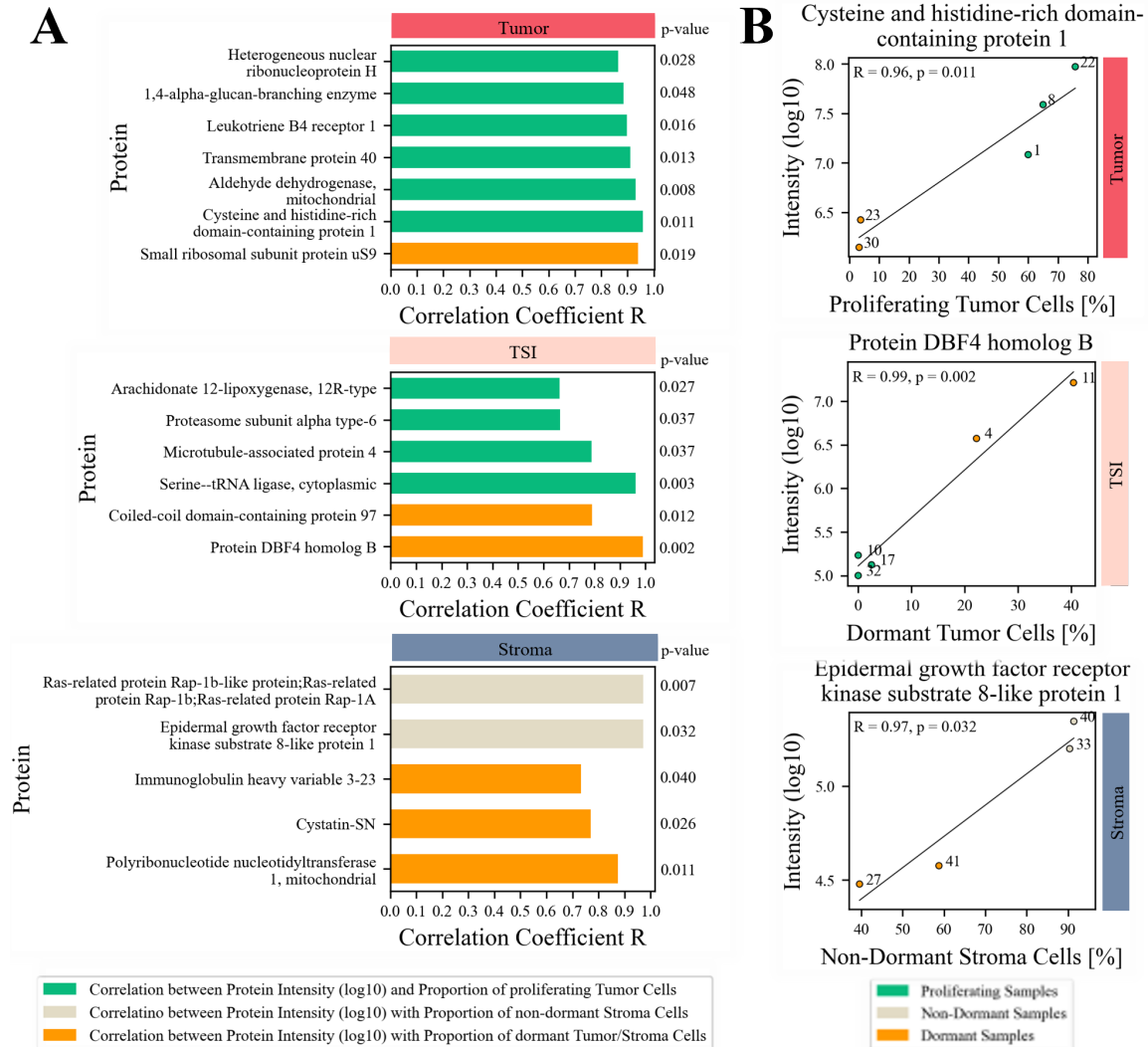


Figure 16: Correlation of protein expression with cellular composition across tissue regions. **A** Bar plots showing proteins with statistically significant Pearson correlations ($p < 0.05$) between $\log_{10}$ -transformed protein intensities and the proportion of corresponding cell states within tumor, TSI, and stroma samples. Correlation coefficients R are showcased for each protein, with $R > 0.8$ indicating a strong positive linear correlation and $0.8 > R > 0.4$ indicating a moderate positive linear correlation. **B** Scatter plots for representative proteins for each tissue region and cell state illustrating the positive correlations between protein intensities and cellular proportion.

5 Discussion

This study demonstrates the successful integration of t-CycIF staining with MS-based proteomics and establishes a functioning experimental workflow and a robust data analysis pipeline for spatially resolved tissue analysis. The combined approach enables detailed cellular and tissue characterization while providing proteomics insights from precisely selected ROIs.

t-CycIF staining & subsequent ROI selection

A key step in the project is the t-CycIF staining. In particular, it was important to demonstrate that the staining protocol works effectively with thicker tissue sections, up to 10 μm , than those typically used in t-CycIF experiments. Despite the increased thickness, nearly all antibodies showed clear staining patterns with good signal-to-noise ratios, indicating that the protocol is robust under these conditions. However, re-validating some antibody markers originally optimized for thinner sections on thicker tissues could further improve staining quality and ensure reliable interpretation.

Based on the successful t-CycIF staining, the ROI selection and the subsequent picking were effectively guided. The staining facilitated the identification of key tissue regions, including tumor, TSI, and stroma as well as the further characterization of various cells. A notable advantage of the approach was the inclusion of biomarkers relevant for ROI selection in the final staining cycle, which permitted real-time validation and adjustment during the picking process of the pre-selected ROIs. This strategy enhanced the accuracy of targeting specific regions and cell types, improving the overall reliability of spatially resolved sampling. However, identifying suitable ROIs was still challenging, especially when the target cell populations were scarce or unevenly distributed within the tissue. In such cases, locating regions with sufficient representation of the desired biomarkers required increased effort and careful examination. To address this limitation, it is recommended to utilize any available prior staining data or histological information from the tissue samples before starting with the experiment. Pre-assessing tissue characteristics allows for informed selection of samples that are more likely to contain regions exhibiting the phenotypes of interest, thereby increasing the efficiency and success rate of ROI targeting. It is also important to keep in mind that some cell types or marker-defined populations are naturally rare or unevenly distributed. This variability should be considered when designing research questions and planning future experiments using ProteoPicking, as it affects how easily these cells can be isolated and how the obtained data can be interpreted.

Following t-CycIF staining, ROI selection, and sample picking, the next essential step was the establishment of a robust image processing pipeline.

As mentioned before, illumination correction, stitching and registration of the images were successful. The pre-established BaSiC and ASHLAR workflows from the Färkkilä GitHub ImageProcessing repository work not only for the usually used 4 μm to 5 μm tissue section images, but also for the thicker ones used for this project. Thus, they were implemented into the image processing pipeline of ProteoPicking without further problems.

The next critical component of the processing step was the segmentation of the images, which needed to accurately delineate individual cells within especially densely clustered regions due to the higher thickness of the imaged tissue sections. Among the segmentation models tested, Mesmer demonstrated superior performance in handling the images. Its effectiveness was further enhanced by applying image pre-processing steps, in particular gamma correction and unsharp masking, which improved image intensity and contrast, enabling more accurate nuclear segmentation. Based on these results, gamma correction, unsharp masking, and Mesmer were combined and successfully integrated into the processing pipeline to segment small image crops containing the selected ROIs.

Despite all advancements, the current segmentation workflow exhibits limitations and considerable room for refinement. Future efforts should focus on a systematic, quantitative evaluation of various image enhancement techniques and combinations of them to objectively assess their individual and combined impact on segmentation accuracy. In addition, extensive objective and quantified parameter optimization is needed to further improve model performance. Based on evaluation and parameter optimization, the most effective and computationally efficient techniques should be implemented to improve the current performance of the method. Moreover, as a measure to address the current limitation to small image crops, adapting the workflow for whole-slide image segmentation would be a sensible next step. At present, the developed workflow for whole-slide image segmentation is computationally intensive, has a long running time, and poses the problem of over-segmentation of the image background. Solving these problems and optimizing the workflow, for example by evaluating optimal tile size, number, and overlap, would significantly improve the scalability and facilitate broader application of the pipeline in future projects. Additionally, it would allow for whole-slide quantification and phenotyping, enabling more comprehensive

analysis of the picked ROIs, especially with regard to their large-scale spatial context.

With individual cells segmented, the next step involved quantification and phenotyping to characterize the cellular makeup of the picked tissue regions in detail. Phenotyping, although time-consuming due to the need for manual verification, was crucial to accurately identify cell types and functional states based on marker expression. Beyond its role in analysis, phenotyping also served as a form of quality control, providing an overview of the total cell counts and phenotypic distributions across all samples. Thereby, it offers early indications of how successfully targeted phenotypes were captured as well as the extent of non-target cell inclusion. During the optimization project, the phenotyping results already served as a first representation of the picking outcome. Phenotyping results showed that nearly 20 % of all isolated cells were immune cells, even though in the picking process they were not targeted and rather tried to avoid being picked. Moreover, it revealed that while the majority of picked dormant stroma cells belonged to the targeted stricter dormant phenotype, Ki67–pRb–p21+, two thirds of the dormant tumor cells were part of the less strict phenotype. These results could be due to various reasons. The total phenotype distribution could be influenced by individual samples, whose ROI selection and picking were not sufficiently targeted, despite prior control based on the captured tissue images and real-time monitoring during the picking process. Another reason for the displayed distribution could be the inherent composition and structure of the used HGSC samples with their given immune infiltration and possible differences in dormant tumor and stroma cells. Lastly, it is important to consider that during the picking process not only the targeted cells alone are isolated, but also all cells within the 40 μm needle diameter. If ROIs are not evaluated thoroughly beforehand, this can lead to the isolation of non-targeted cells in addition to the targeted ones, contributing to sample heterogeneity. However, to make more accurate assumptions about the distribution of the various phenotypes, the cell composition of each individual sample needs to be considered, as was done during the compositional analysis.

Image analysis & sample composition

While phenotyping provides a global overview, detailed, sample-wise compositional analysis was necessary to determine whether the composition of individual ROIs truly reflected the intended targets and to break down the global cellular distribution in greater depth.

The analysis showed that tumor and stroma cells were distributed as expected based

on ROI locations. That means that tumor samples mainly consisted of tumor cells, stroma samples mainly of stroma cells, and TSI samples were a mix of tumor and stroma cells. This confirms the overall success of the picking strategy. However, despite efforts to avoid immune cell-rich areas, immune cells were detected in all samples. This supports the assumption that the significant amount of immune cells in the global sample composition could be due to the inherent character of the analyzed tissues. In general, the TME is often highly infiltrated by immune cells, which could make their complete exclusion from samples taken from tumor tissues very difficult or even nearly impossible. Comparing the composition of main cell types in the samples derived from the different tissue regions, it became evident that stroma samples were well targeted, with very little tumor contamination. In contrast, tumor samples consistently included, next to the unavoidable immune cells, stroma cells, which may be attributed to the complex, slit-like architecture characteristic for HGSC, making clean separation of tumor cells particularly challenging. For TSI samples, a mixed cellular composition was expected due to their location at the tumor-stroma interface, and this was consistently reflected in the observed data.

Examining the distribution of phenotypes within tumor and stroma populations revealed further trends. In proliferating, respectively non-dormant samples, the proportion of proliferating or non-dormant cells was generally higher than the proportion of dormant cells in the corresponding dormant samples. This likely reflects practical challenges. Regions enriched in proliferating or non-dormant cells were more abundant and spatially clustered, facilitating easier targeted picking. In contrast, dormant cells appeared less clustered and harder to isolate, requiring less narrow selection criteria that may have reduced sample purity. Although this observation hints at possible biological differences in the spatial organization of proliferating/non-dormant and dormant cells, further investigation, such as spatial neighborhood analysis, would be necessary to support such conclusions, especially given the limited sample size.

When comparing the two dormancy phenotypes defined in this study, stroma samples generally exhibited a higher proportion of the stricter dormancy phenotype compared to tumor and TSI samples. Whether this difference reflects true biological variation or sampling constraints remains unclear. Similarly, the co-occurrence of dormant tumor and stroma cells in several TSI regions raises the possibility of spatial associations between cell states. However, this observation remains speculative and would require further targeted investigation.

All in all, this underlines the importance of spatial validation of the sample com-

position in this workflow. Reviewing imaging data from all staining cycles can confirm whether picked regions truly reflect their intended content. Thereby, it can be determined whether samples can retain their original pre-picking labels or require reassignment or exclusion, as done during the optimization experiment, ultimately enhancing the accuracy and reliability of the dataset.

Proteomics data exploration

Following spatial validation of sample composition, the quality and reliability of proteomics data from picked tissue regions were evaluated to identify feasibility as well as strengths and limitations of the current workflow, and to guide future improvements.

The pilot proteomics experiment demonstrated that t-CycIF staining is compatible with MS-based proteomics, yielding meaningful protein identification even from small, spatially defined samples. Although some contamination from keratin and polymers was detected, slightly impacting protein detection, the overall success of protein digestion and MS acquisition was confirmed through consistent peptide signals and charge states characteristic for tryptic peptides. Importantly, t-CycIF-stained samples showed comparable protein yields to those stained with IF, indicating that additional staining and bleaching steps did not markedly reduce proteome quality, which further solidifies the feasibility of ProteoPicking.

A positive correlation between the number of picked regions per sample and the depth of proteome coverage was observed. Samples with fewer than ten picks displayed high variability and lower protein counts, whereas those with ten or twenty picks generally achieved greater consistency and higher protein yields. This suggests that increasing the number of picks can enhance data robustness, though several lower pick samples still provided sufficient protein yields. Additionally, tumor-derived samples tended to yield higher protein counts than stroma samples, potentially reflecting biological differences such as higher metabolic activity and protein abundance in tumor cells, or a difference in the extraction efficiency between the tissue types.

Buffer choice also influenced protein yield. Although limited by sample size, urea buffer performed slightly better than TFE, especially in t-CycIF-stained samples. This fact, together with urea's lower hazard profile, makes urea buffer preferable as a lysis buffer for future experiments. Finally, whole-tissue lysates yielded substantially higher protein counts compared to spatially restricted samples, underscoring the intrinsic limitation of proteome depth when analyzing small regions. This reflects the reduced total protein quantity in small picks rather than methodological failure.

In the subsequent optimization experiment, improvements in sample preparation successfully reduced contaminating impurities, as indicated by chromatograms showing fewer non-peptide signals. But in addition, an unexpected overall reduction in peptide-derived signals was observed. Therefore, it is not completely clear if the reduction in signals derived from contaminants was caused by implemented protocol changes or just by the general lower detection rate.

The observed reduction in protein yield compared to the pilot experiment remains incompletely explained but likely arises from one or multiple of the differences between the two projects. One difference was the methodological adjustments between the two experiments. They were generally minor and aimed to improve sample quality and minimize contamination. Based on the conducted protocol changes, it is unlikely that they are responsible, at least not solely, for the substantial decrease in protein identification count in the picked samples. Another difference was technical problems with the CyteFinder II during the picking process of the optimization experiment. Even though a workaround was found, some issues during the picking or sample depositing may have persisted and could have caused the lower protein yield. Lastly, a further critical difference was the extended storage duration of samples prior to MS analysis. In the pilot study, samples were stored at -80°C for approximately two to three weeks before processing. In contrast, the samples of the optimization experiment had a storage period exceeding two months. Although frozen storage should theoretically preserve protein integrity, the combination of very low sample volumes and protein concentrations raises concerns about potential degradation, buffer evaporation, or adhesion of proteins to tube walls over prolonged periods. Such effects, while not definitively demonstrated here, could contribute to diminished protein recovery and detection sensitivity. Importantly, full-section tissue lysate controls processed alongside the small picks exhibited robust protein yields, indicating that the general workflow, t-CycIF staining, extraction, and MS protocols remained effective. This suggests that the reduced protein yield in picked samples is more likely caused by factors specific to sample handling, storage, or picking efficiency, rather than fundamental limitations of the proteomics workflow. Taken together, these considerations highlight the need for further systematic investigation into sample handling, storage conditions, and instrument performance to optimize the analysis.

In addition to low protein yields of the picked samples, analysis of negative controls revealed additional challenges. The negative control picked from the slide taken directly from the package showed an unexpectedly high protein count. Notably, this sample was collected last, after all other picks had been completed, raising suspicion

of potential sample carry-over during the picking process. This carry-over may result from residual material adhering to the picking needle between samples. To mitigate this risk in future experiments, implementing additional needle washing steps or increasing the frequency of needle replacements could be effective strategies.

Altogether, these findings confirm the feasibility of integrating t-CycIF-guided spatial sampling with MS-based proteomics, while highlighting several critical areas for improvement. Variability in protein detection among samples with low pick numbers reinforces the importance of improving picking consistency and increasing replicate numbers to ensure statistical robustness and enable exclusion of poor-quality data. Finally, advances in MS instrumentation sensitivity and optimized sample processing protocols hold promise for enabling proteomics on smaller tissue regions, thereby improving spatial resolution in the future.

Integrated proteomics analysis pipeline

Next, based on the proteomics data generated during the optimization experiment and despite the generally low protein yields, a functional analysis pipeline was successfully established. This pipeline provides a solid foundation for extracting meaningful information from spatially resolved proteomics datasets. Nevertheless, there is still room for complementary analysis like pathway enrichment analysis or network modeling, to deepen biological insights in future studies, which would have exceeded the scope of this study.

Using this pipeline, the proteomics profiles of picked samples were explored in regard of their region of origin and their cell state. Due to the limited sample number, variable protein yields, and conservative quality control criteria applied, the statistical power and significance of findings are limited. Thus, the results presented here are primarily hypothesis-generating and intended to demonstrate the utility of the pipeline rather than serve as definitive biological conclusions.

Cluster Heatmap Analysis revealed a core proteome consistently detected across nearly all samples, suggesting a stable baseline of abundant proteins in all investigated cell types. In contrast, a subset of proteins appeared sporadically across all samples and groups, indicating they were neither consistently detected nor specific to any region or cell state, likely reflecting low-abundance targets or technical variability. Dormant samples generally exhibited fewer identified proteins than proliferating samples, which may indicate reduced proteome complexity or lower protein

expression associated with dormancy. Clustering patterns also suggested that protein expression profiles were influenced by both cell state and tissue origin, with some non-dormant stroma samples clustering alongside tumor/TSI samples and some dormant tumor/TSI samples clustering alongside dormant stroma samples. This potentially reflects molecular similarities between activated stroma and tumor regions, and dormant tumor and stroma regions.

The performed DEA faced limitations due to strict filtering and the resulting low number of samples per comparison group. Moreover, all proteins uniquely detected in one group needed to be excluded from statistical testing due to the inherent type of the performed analysis, though these may represent biologically relevant biomarkers specific for one cell state. There thorough exploration and analysis would have exceeded the scope of this thesis, therefore warranting further investigation.

After performing the DEA, results were assessed. No proteins reached statistical significance after multiple testing correction, a likely consequence of the small sample size. Nevertheless, determining upregulated proteins by relying on unadjusted p -values combined with fold-change thresholds remains a widely accepted approach in exploratory proteomics studies. This strategy provides a practical means to identify candidate proteins for pipeline validation and future hypothesis generation, while acknowledging the need for cautious interpretation.

For further exploration of differentially expressed proteins, heatmaps normalized via min-max scaling were generated to highlight relative expression differences. These visualizations reinforced distinct proteomics signatures associated with proliferating versus dormant states. The group-specific clusters likely encompass proteins involved in cell cycle regulation, metabolism, or microenvironment interactions relevant to cellular proliferation and dormancy. An extensive investigation into the function and background of all these proteins would be interesting, but it is not conducted here as it exceeds the scope of this study and would be too extensive to address within the available framework. Given the exploratory nature and limited sample size, these heatmaps serve as qualitative validation tools rather than definitive evidence of protein regulation. Follow-up studies using targeted proteomics or complementary functional assays are needed to confirm these observations.

Integration of proteomics data with compositional analysis via correlation assessments revealed significant positive correlations between protein expression levels and the relative abundance of corresponding cell states. These correlations not only suggest that certain proteins may serve as markers or functional effectors of specific cellular

phenotypes within tumor and stroma compartments but also provide orthogonal validation of sample characterization. It is important to note that correlation does not imply causation, and some observed relationships may reflect co-variation with other factors. And again, the limited sample size further reduces the robustness of these findings, implying the necessity of a cautious interpretation. Nonetheless, these proteins represent promising candidates for future targeted validation, functional studies, and biomarker development.

These results demonstrate that ProteoPicking, the developed spatially resolved proteomics approach combining t-CycIF staining with MS-based proteomics analysis, effectively captures biologically relevant molecular signatures reflective of cellular composition and state. This work lays important groundwork for more comprehensive and statistically powered studies aimed at unraveling functional proteomics heterogeneity within complex tissue microenvironments. Nevertheless, while the method is already applicable and yields valuable insights, it still requires further refinement and optimization.

In conclusion, the spatially resolved proteomics analysis method developed and established in this thesis highlights both the promise of the approach and the continuing need for methodological advancements in this field.

List of Figures

1	Schematic overview of the main steps of ProteoPicking. Figure created in BioRender, 2025.	30
2	Overview of the image acquisition workflow for multiplexed tissue imaging - tissue sections are imaged using a Rarecyte CyteFinder II scanner, generating multi-channel, tiled image stacks (rcpn1 files) with associated metadata and individual stitched images of each cycle. Figure created in BioRender, 2025.	36
3	Visualization of the four picking steps - positioning of the needle, lowering the needle to dissect the tissue, detaching the tissue via wiggling, and collecting the tissue.	39
4	Overview of the image processing (including illumination correction, stitching and registration, segmentation, and quantification), phenotyping, and visualization. Figure created in BioRender, 2025.	39
5	Overview of the experimental design of the pilot project and staining patterns of all cycles. A Schematic representation of the experimental setup of the pilot project. Figure created in BioRender, 2025. B IF and t-CycIF staining patterns of the markers used during all staining cycles. Each image shows Hoechst staining either alone or merged with the respective markers.	47
6	Picking criteria for ROI selection and examples for picked areas. A The ROI selection of tumor M and S phase samples as well as stroma samples is based on specific combinations of positive and negative markers. After the picking a hole remains in the tissue. B Examples of picked tumor M phase, S phase, and stroma regions before and after the picking process.	49
7	Proteomics data analysis results of the pilot project. A Representative chromatograms and spectra illustrating peptide signals and contaminants. B Comparison of the number of identified proteins in tumor and stroma samples with different numbers of picks using urea buffer, with either IF- or t-CycIF-staining C Comparison of protein yields in 10 pick samples processed with either urea or TFE buffers, with either IF- or t-CycIF-staining. D Protein identification count in whole-section lysate samples as a reference.	52

8	Overview of the experimental design of the optimization project and staining patterns of all cycles. A Schematic representation of the experimental setup of the pilot project. Figure created in BioRender, 2025. B t-CycIF staining patterns of the markers used during all staining cycles. Each image shows Hoechst staining either alone or merged with the respective markers.	56
9	Picking criteria for ROI selection and examples for picked areas. A The ROI selection of tumor proliferative (P) and dormant (D) phase samples as well as stroma non-dormant (N) and dormant (D) phase samples is based on specific positive and negative marker combinations. After the picking, a hole remains in the tissue. B Examples of picked samples from tumor, TSI, and stroma regions before and after the picking process.	58
10	Overview of different segmentation model performances and image enhancement techniques. A Comparison of segmentation performance on a representative image crop showing the nuclear staining channel alongside overlays of masks generated by Stardist, Cellpose, and Mesmer. B Segmentation performance of Mesmer on the original nuclear staining image compared to its performance on images enhanced using CLAHE (kernel size = 8×8 , bin number = 200, clip limit = 0.03), gamma correction ($\gamma = 2/3$), unsharp masking ($\sigma = 1.0$, strength = 2.5), and Richardson-Lucy deconvolution ($\sigma = 1.0$, 30 iterations). Comparison is made by overlaying nuclear staining with segmentation masks. C Segmentation outcome using a combined approach of gamma correction and unsharp masking with the nuclear segmentation model Mesmer. .	61
11	Overview of the cell type annotation marker panel and marker expression patterns of the various phenotypes. A Phenotyping was based on seven selected biomarkers, each exhibiting distinct expression patterns. B Heatmap illustrating the different marker expression profiles of the distinct phenotypes within the ROIs as well as the total abundance of the different cell types within all picked samples. For better comparability, log transformed and rescaled marker expression values were used (log transformation and rescaling was performed using Scimap functions). Expression values above 0.5 are counted as positive for the specific marker.	63

12	Overview of cell type and phenotype composition across all analyzed samples. A Cell type composition across all samples grouped by tissue region and targeted cell state. B Boxplots displaying the distribution of total cell counts within all picked samples grouped by region of origin. C Distribution of tumor cell phenotypes within tumor cells across tumor and TSI samples. D Distribution of stroma cell phenotypes within stroma cells across stroma and TSI samples. Asterisks in C and D mark samples that were relabeled based on compositional analysis.	66
13	Overview of proteomics data quality and yielded protein detection numbers. A Representative chromatograms from three samples, a proliferating tumor sample, a proliferating TSI sample, and a dormant stroma sample, illustrating the similar protein detection across sample types. B Number of identified proteins per sample, with samples excluded from downstream analysis highlighted in red due to protein counts below 500 and purple due to unspecific cellular state. The horizontal dashed line indicates the 500 protein threshold for sample inclusion. C Protein yield in positive and negative control samples. .	69
14	Proteomics data exploration and DEA. A Cluster heatmap showing binary detection profiles of proteins across samples. Samples and proteins are clustered based on shared detection patterns. B Visualization of proteins excluded from the DEA. Heatmaps contain only proteins detected in more than one sample, exclusively in one sample group. Samples and proteins are clustered based on shared detection patterns. C Volcano plots depicting results of DEA for each tissue region. Proteins with $ \log_2 \text{ fold change} > 2$ and $p\text{-value} < 0.05$ ($-\log_{10} p\text{-value} > 1.3$) are highlighted as significantly differentially expressed.	71
15	Heatmaps of differentially expressed proteins in tumor, TSI, and stroma regions. Intensities are min-max scaled and clustering was based on rescaled protein intensities.	74

16	Correlation of protein expression with cellular composition across tissue regions. A Bar plots showing proteins with statistically significant Pearson correlations ($p < 0.05$) between $\log_{10}$ -transformed protein intensities and the proportion of corresponding cell states within tumor, TSI, and stroma samples. Correlation coefficients R are showcased for each protein, with $R > 0.8$ indicating a strong positive linear correlation and $0.8 > R > 0.4$ indicating a moderate positive linear correlation. B Scatter plots for representative proteins for each tissue region and cell state illustrating the positive correlations between protein intensities and cellular proportion.	76
----	---	----

List of Tables

1	Examples of commonly used cell cycle markers.	17
2	Key characteristics of different highly multiplexed imaging techniques.	22
3	Used reagents and instruments.	27
4	Used antibodies for t-CyclIF.	28
5	Overview of key characteristics of used patients samples. HRP - homologous recombination proficiency, HRD - homologous recombination deficiency.	29
6	Antibody panel for pilot experiment.	33
7	Antibody panel for optimization experiment.	33
8	Overview of used parameters during the imaging process.	35
9	Overview of used parameters during the picking process.	38
10	Logic table for phenotyping.	42
11	Sample specifications of the pilot experiment.	46
12	Sample specifications of the optimization experiment.	54

References

- Al-Amrani, S., Al-Jabri, Z., Al-Zaabi, A., Alshekaili, J., & Al-Khabori, M. (2021). Proteomics: Concepts and applications in human medicine. *World journal of biological chemistry*, 12(5), 57.
- Angelo, M., Bendall, S. C., Finck, R., Hale, M. B., Hitzman, C., Borowsky, A. D., Levenson, R. M., Lowe, J. B., Liu, S. D., Zhao, S., et al. (2014). Multiplexed ion beam imaging of human breast tumors. *Nature medicine*, 20(4), 436–442.
- Ashworth, A., Lord, C. J., & Reis-Filho, J. S. (2011). Genetic interactions in cancer progression and treatment. *Cell*, 145(1), 30–38.
- Aslam, B., Basit, M., Nisar, M. A., Khurshid, M., & Rasool, M. H. (2016). Proteomics: Technologies and their applications. *Journal of chromatographic science*, 1–15.
- Baharlou, H., Canete, N. P., Cunningham, A. L., Harman, A. N., & Patrick, E. (2019). Mass cytometry imaging for the study of human diseases—applications and data analysis strategies. *Frontiers in immunology*, 10, 2657.
- Besson, A., Dowdy, S. F., & Roberts, J. M. (2008). Cdk inhibitors: Cell cycle regulators and beyond. *Developmental cell*, 14(2), 159–169.
- BioRender. (2025). Biorender scientific illustration tool [Accessed: 2025-05-28]. <https://biorender.com>
- Black, S., Phillips, D., Hickey, J. W., Kennedy-Darling, J., Venkataraman, V. G., Samusik, N., Goltsev, Y., Schürch, C. M., & Nolan, G. P. (2021). Codex multiplexed tissue imaging with dna-conjugated antibodies. *Nature protocols*, 16(8), 3802–3835.
- Bodenmiller, B. (2016). Multiplexed epitope-based tissue imaging for discovery and healthcare applications. *Cell systems*, 2(4), 225–238.
- Böhm, S., Montfort, A., Pearce, O. M., Topping, J., Chakravarty, P., Everitt, G. L., Clear, A., McDermott, J. R., Ennis, D., Dowe, T., et al. (2016). Neoadjuvant chemotherapy modulates the immune microenvironment in metastases of tubo-ovarian high-grade serous carcinoma. *Clinical Cancer Research*, 22(12), 3025–3036.
- Bruhn, C., Kroll, T., & Wang, Z.-Q. (2014). Systematic characterization of cell cycle phase-dependent protein dynamics and pathway activities by high-content microscopy-assisted cell cycle phenotyping. *Genomics, Proteomics & Bioinformatics*, 12(6), 255–265.
- Burger, R. A. (2011). Overview of anti-angiogenic agents in development for ovarian cancer. *Gynecologic oncology*, 121(1), 230–238.
- Burger, R. A., Brady, M. F., Bookman, M. A., Fleming, G. F., Monk, B. J., Huang, H., Mannel, R. S., Homesley, H. D., Fowler, J., Greer, B. E., et al. (2011).

- Incorporation of bevacizumab in the primary treatment of ovarian cancer. *New England Journal of Medicine*, 365(26), 2473–2483.
- Calcinotto, A., Kohli, J., Zagato, E., Pellegrini, L., Demaria, M., & Alimonti, A. (2019). Cellular senescence: Aging, cancer, and injury. *Physiological reviews*, 99(2), 1047–1078.
- Callister, S. J., Barry, R. C., Adkins, J. N., Johnson, E. T., Qian, W.-j., Webb-Robertson, B.-J. M., Smith, R. D., & Lipton, M. S. (2006). Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. *Journal of proteome research*, 5(2), 277–286.
- Cancer Genome Atlas Research Network and others. (2011). Integrated genomic analyses of ovarian carcinoma. *Nature*, 474(7353), 609.
- Cappiello, A., Famiglini, G., Rossi, L., & Magnani, M. (1997). Use of nonvolatile buffers in liquid chromatography/mass spectrometry: Advantages of capillary-scale particle beam interfacing. *Analytical chemistry*, 69(24), 5136–5141.
- Chen, D. S., & Mellman, I. (2013). Oncology meets immunology: The cancer-immunity cycle. *immunity*, 39(1), 1–10.
- Clarke, B., Tinker, A. V., Lee, C.-H., Subramanian, S., Van De Rijn, M., Turbin, D., Kalloger, S., Han, G., Ceballos, K., Cadungog, M. G., et al. (2009). Intraepithelial t cells and prognosis in ovarian carcinoma: Novel associations with stage, tumor type, and brca1 loss. *Modern Pathology*, 22(3), 393–402.
- Coller, H. A., Sang, L., & Roberts, J. M. (2006). A new description of cellular quiescence. *PLoS biology*, 4(3), e83.
- Curiel, T. J., Coukos, G., Zou, L., Alvarez, X., Cheng, P., Mottram, P., Evdemon-Hogan, M., Conejo-Garcia, J. R., Zhang, L., Burow, M., et al. (2004). Specific recruitment of regulatory t cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. *Nature medicine*, 10(9), 942–949.
- Doubeni, C. A., Doubeni, A. R., & Myers, A. E. (2016). Diagnosis and management of ovarian cancer. *American family physician*, 93(11), 937–944.
- Du Bois, A., Floquet, A., Kim, J. W., Rau, J., Del Campo, J. M., Friedlander, M., Pignata, S., Fujiwara, K., Vergote, I., Colombo, N., et al. (2013). Randomized, double-blind, phase iii trial of pazopanib versus placebo in women who have not progressed after first-line chemotherapy for advanced epithelial ovarian, fallopian tube, or primary peritoneal cancer (aeoc): Results of an international intergroup trial (ago-ovar16).
- Du Bois, A., Reuss, A., Pujade-Lauraine, E., Harter, P., Ray-Coquard, I., & Pfisterer, J. (2009). Role of surgical outcome as prognostic factor in advanced epithelial ovarian cancer: A combined exploratory analysis of 3 prospectively randomized

- phase 3 multicenter trials: By the arbeitgemeinschaft gynaekologische onkologie studiengruppe ovarialkarzinom (ago-ovar) and the groupe d'investigateurs nationaux pour les etudes des cancers de l'ovaire (gineco). *Cancer*, 115(6), 1234–1244.
- Endo, H., & Inoue, M. (2019). Dormancy in cancer. *Cancer science*, 110(2), 474–480.
- Engeland, K. (2022). Cell cycle regulation: P53-p21-rb signaling. *Cell Death & Differentiation*, 29(5), 946–960.
- Ericson, N. G., Podyminogin, R., Chow, J., Nirmal, A. J., Lin, J.-R., Chen, Y.-A., Maliga, Z., Sorger, P., Teplitz, K., Ramirez, A. B., Kaldjian, E. P., & George, T. (2019). Pick-seq: Spatial tissue analysis via multi-parameter imaging and rna sequencing of micro-regions. Retrieved May 15, 2025, from https://rarecyte.com/pdf/PickSeq_Poster_SITC_2019.pdf
- European Medicines Agency. (2025). Lynparza, olaparib. Retrieved January 9, 2025, from <https://www.ema.europa.eu/en/medicines/human/EPAR/lynparza>
- Eustermann, S., Wu, W.-F., Langelier, M.-F., Yang, J.-C., Easton, L. E., Riccio, A. A., Pascal, J. M., & Neuhaus, D. (2015). Structural basis of detection and signaling of dna single-strand breaks by human parp-1. *Molecular cell*, 60(5), 742–754.
- Fagotti, A., Ferrandina, M. G., Vizzielli, G., Pasciuto, T., Fanfani, F., Gallotta, V., Margariti, P. A., Chiantera, V., Costantini, B., Alletti, S. G., et al. (2020). Randomized trial of primary debulking surgery versus neoadjuvant chemotherapy for advanced epithelial ovarian cancer (scorpion-nct01461850). *International Journal of Gynecological Cancer*, 30(11), 1657–1664.
- Färkkiä lab. (2024). Image processing. Retrieved February 10, 2025, from https://github.com/farkkilab/image_processing/tree/main
- Färkkilä lab. (2025). Proteopicking. Retrieved May 28, 2025, from https://github.com/farkkilab/ProteoPick_LinasThesis
- Farmer, H., McCabe, N., Lord, C. J., Tutt, A. N., Johnson, D. A., Richardson, T. B., Santarosa, M., Dillon, K. J., Hickson, I., Knights, C., et al. (2005). Targeting the dna repair defect in brca mutant cells as a therapeutic strategy. *Nature*, 434(7035), 917–921.
- Finkernagel, F., Reinartz, S., Lieber, S., Adhikary, T., Wortmann, A., Hoffmann, N., Bieringer, T., Nist, A., Stiewe, T., Jansen, J. M., et al. (2016). The transcriptional signature of human ovarian carcinoma macrophages is associated with extracellular matrix reorganization. *Oncotarget*, 7(46), 75339.

- Fischer, A. H., Jacobson, K. A., Rose, J., & Zeller, R. (2008). Hematoxylin and eosin staining of tissue and cell sections. *Cold spring harbor protocols, 2008*(5), pdb-prot4986.
- Galdiero, M. R., Garlanda, C., Jaillon, S., Marone, G., & Mantovani, A. (2013). Tumor associated macrophages and neutrophils in tumor progression. *Journal of cellular physiology, 228*(7), 1404–1412.
- Gelmon, K. A., Tischkowitz, M., Mackay, H., Swenerton, K., Robidoux, A., Tonkin, K., Hirte, H., Huntsman, D., Clemons, M., Gilks, B., et al. (2011). Olaparib in patients with recurrent high-grade serous or poorly differentiated ovarian carcinoma or triple-negative breast cancer: A phase 2, multicentre, open-label, non-randomised study. *The lancet oncology, 12*(9), 852–861.
- Giesen, C., Wang, H. A., Schapiro, D., Zivanovic, N., Jacobs, A., Hattendorf, B., Schüffler, P. J., Grolimund, D., Buhmann, J. M., Brandt, S., et al. (2014). Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nature methods, 11*(4), 417–422.
- Gill, S. E., McGree, M. E., Weaver, A. L., Cliby, W. A., & Langstraat, C. L. (2017). Optimizing the treatment of ovarian cancer: Neoadjuvant chemotherapy and interval debulking versus primary debulking surgery for epithelial ovarian cancers likely to have suboptimal resection. *Gynecologic oncology, 144*(2), 266–273.
- Giraldo, N. A., Sanchez-Salas, R., Peske, J. D., Vano, Y., Becht, E., Petitprez, F., Validire, P., Ingels, A., Cathelineau, X., Fridman, W. H., et al. (2019). The clinical role of the tme in solid cancer. *British journal of cancer, 120*(1), 45–53.
- Givel, A.-M., Kieffer, Y., Scholer-Dahirel, A., Sirven, P., Cardon, M., Pelon, F., Mägagna, I., Gentric, G., Costa, A., Bonneau, C., et al. (2018). Mir200-regulated *excl12β* promotes fibroblast heterogeneity and immunosuppression in ovarian cancers. *Nature communications, 9*(1), 1056.
- Glaviano, A., Singh, S. K., Lee, E. H. C., Okina, E., Lam, H. Y., Carbone, D., Reddy, E. P., O'Connor, M. J., Koff, A., Singh, G., et al. (2025). Cell cycle dysregulation in cancer. *Pharmacological Reviews, 77*(2), 100030.
- Goff, B. A., Mandel, L. S., Melancon, C. H., & Muntz, H. G. (2004). Frequency of symptoms of ovarian cancer in women presenting to primary care clinics. *Jama, 291*(22), 2705–2712.
- Gonzalez, R. C., & Woods, R. E. (2018). *Digital image processing* (4th) [See Chapter 3: Intensity Transformations and Spatial Filtering]. Pearson.
- Greenwald, N. F., Miller, G., Moen, E., Kong, A., Kagel, A., Fullaway, C. C., McIntosh, B. J., Leow, K. H., Schwartz, M. S., Yang, S., Gee, J. C., Brown, M., &

- Van Valen, D. (2022). Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nature Biotechnology*, 40, 555–565. <https://doi.org/10.1038/s41587-021-01197-w>
- Grivennikov, S. I., Greten, F. R., & Karin, M. (2010). Immunity, inflammation, and cancer. *Cell*, 140(6), 883–899.
- Gutstein, H. B., & Morris, J. S. (2007). Laser capture sampling and analytical issues in proteomics. *Expert review of proteomics*, 4(5), 627–637.
- Hamanishi, J., Mandai, M., Ikeda, T., Minami, M., Kawaguchi, A., Murayama, T., Kanai, M., Mori, Y., Matsumoto, S., Chikuma, S., et al. (2015). Safety and antitumor activity of anti-pd-1 antibody, nivolumab, in patients with platinum-resistant ovarian cancer. *Journal of Clinical Oncology*, 33(34), 4015–4022.
- Hanahan, D., & Coussens, L. M. (2012). Accessories to the crime: Functions of cells recruited to the tumor microenvironment. *Cancer cell*, 21(3), 309–322.
- Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. *cell*, 100(1), 57–70.
- Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *cell*, 144(5), 646–674.
- Harms, P. W., Frankel, T. L., Moutafi, M., Rao, A., Rimm, D. L., Taube, J. M., Thomas, D., Chan, M. P., & Pantanowitz, L. (2023). Multiplex immunohistochemistry and immunofluorescence: A practical update for pathologists. *Modern Pathology*, 36(7), 100197.
- Hasin, Y., Seldin, M., & Lusis, A. (2017). Multi-omics approaches to disease. *Genome biology*, 18, 1–15.
- Henley, S. A., & Dick, F. A. (2012). The retinoblastoma family of proteins and their regulatory functions in the mammalian cell division cycle. *Cell division*, 7, 1–14.
- Israels, E., & Israels, L. (2000). The cell cycle. *The oncologist*, 5(6), 510–513.
- Jia, Q., Wang, A., Yuan, Y., Zhu, B., & Long, H. (2022). Heterogeneity of the tumor immune microenvironment and its clinical relevance. *Experimental hematology & oncology*, 11(1), 24.
- Jiménez-Sánchez, A., Cybulska, P., Mager, K. L., Koplev, S., Cast, O., Couturier, D.-L., Memon, D., Selenica, P., Nikolovski, I., Mazaheri, Y., et al. (2020). Unraveling tumor-immune heterogeneity in advanced ovarian cancer uncovers immunogenic effect of chemotherapy. *Nature Genetics*, 52(6), 582–593.
- Kalluri, R., & Zeisberg, M. (2006). Fibroblasts in cancer. *Nature reviews cancer*, 6(5), 392–401.

- Kandalajt, L. E., Dangaj Laniti, D., & Coukos, G. (2022). Immunobiology of high-grade serous ovarian cancer: Lessons for clinical translation. *Nature Reviews Cancer*, 22(11), 640–656.
- Kelly, R., Zhu, Y., Liang, Y., Cong, Y., Piehowski, P., Dou, M., Zhao, R., Qian, W.-J., Burnum-Johnson, K., & Ansong, C. (2019). Single cell proteome mapping of tissue heterogeneity using microfluidic nanodroplet sample processing and ultrasensitive lc-ms. *Journal of Biomolecular Techniques: JBT*, 30(Suppl), S61.
- Khairallah, A. S., Genestie, C., Auguste, A., & Leary, A. (2018). Impact of neoadjuvant chemotherapy on the immune microenvironment in advanced epithelial ovarian cancer: Prognostic and therapeutic implications. *International Journal of Cancer*, 143(1), 8–15.
- Kurman, R. J., Carcangiu, M. L., Herrington, C. S., & Young, R. H. (2014). *WHO classification of tumours of female reproductive organs*. International Agency for Research on Cancer.
- Kurman, R. J., & Shih, I.-M. (2016). The dualistic model of ovarian carcinogenesis: Revisited, revised, and expanded. *The American journal of pathology*, 186(4), 733–747.
- Labidi-Galy, S. I., Papp, E., Hallberg, D., Niknafs, N., Adleff, V., Noe, M., Bhattacharya, R., Novak, M., Jones, S., Phallen, J., et al. (2017). High grade serous ovarian carcinomas originate in the fallopian tube. *Nature communications*, 8(1), 1093.
- Leary, A., Genestie, C., Blanc-Durand, F., Gouy, S., Dunant, A., Maulard, A., Drusch, F., Cheaib, B., Michels, J., Bentivegna, E., et al. (2021). Neoadjuvant chemotherapy alters the balance of effector to suppressor immune cells in advanced ovarian cancer. *Cancer Immunology, Immunotherapy*, 70, 519–531.
- Ledermann, J., Perren, T., Raja, F., et al. (2013). Randomised double-blind phase iii trial of cediranib (azd 2171) in relapsed platinum sensitive ovarian cancer: Results of the icon6 trial [abstract e17-7020]. *Presented at: ECCO-ESMO-ESTRO European Cancer Congress*.
- Lheureux, S., Gourley, C., Vergote, I., & Oza, A. M. (2019). Epithelial ovarian cancer. *The Lancet*, 393(10177), 1240–1253.
- Li, L., Sun, C., Sun, Y., Dong, Z., Wu, R., Sun, X., Zhang, H., Jiang, W., Zhou, Y., Cen, X., et al. (2022). Spatially resolved proteomics via tissue expansion. *Nature Communications*, 13(1), 7242.
- Ligasová, A., Frydrych, I., & Koberna, K. (2023). Basic methods of cell cycle analysis. *International Journal of Molecular Sciences*, 24(4), 3674.

- Lin, J.-R., Izar, B., Wang, S., Yapp, C., Mei, S., Shah, P. M., Santagata, S., & Sorger, P. K. (2018). Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-cycif and conventional optical microscopes. *elife*, 7, e31657.
- Lucy, L. B. (1974). An iterative technique for the rectification of observed distributions. *The Astronomical Journal*, 79, 745.
- Lundberg, E., & Borner, G. H. (2019). Spatial proteomics: A powerful discovery tool for cell biology. *Nature Reviews Molecular Cell Biology*, 20(5), 285–302.
- Madaminov, A., Khasanov, A., Khatamov, S., Abdurakhmonov, O., Amonov, A., Shukurov, Z., Khudayorov, M., Bekmirzaev, R., & Nishonboev, L. (2021). Cancer-dysregulation of the cell cycle and transduction of cascade signals. *Romanian Journal of Rhinology*, 11(43).
- Matulonis, U., Shapira-Frommer, R., Santin, A., Lisysanskaya, A., Pignata, S., Vergote, I., Raspagliesi, F., Sonke, G., Birrer, M., Provencher, D., et al. (2019). Antitumor activity and safety of pembrolizumab in patients with advanced recurrent ovarian cancer: Results from the phase ii keynote-100 study. *Annals of Oncology*, 30(7), 1080–1087.
- Matulonis, U., Sood, A. K., Fallowfield, L., Howitt, B. E., Sehouli, J., & Karlan, B. Y. (2016). Ovarian cancer. *Nature reviews Disease primers*, 2(1), 1–22.
- McGuire, W. P., Hoskins, W. J., Brady, M. F., Kucera, P. R., Partridge, E. E., Look, K. Y., Clarke-Pearson, D. L., & Davidson, M. (1996). Cyclophosphamide and cisplatin compared with paclitaxel and cisplatin in patients with stage iii and stage iv ovarian cancer. *New England Journal of Medicine*, 334(1), 1–6.
- Merritt, C. R., Ong, G. T., Church, S. E., Barker, K., Danaher, P., Geiss, G., Hoang, M., Jung, J., Liang, Y., McKay-Fleisch, J., et al. (2020). Multiplex digital spatial profiling of proteins and rna in fixed tissue. *Nature biotechnology*, 38(5), 586–599.
- Mesnager, S., Auguste, A., Genestie, C., Dunant, A., Pain, E., Drusch, F., Gouy, S., Morice, P., Bentivegna, E., Lhomme, C., et al. (2017). Neoadjuvant chemotherapy (nact) increases immune infiltration and programmed death-ligand 1 (pd-l1) expression in epithelial ovarian cancer (eoc). *Annals of oncology*, 28(3), 651–657.
- Mhawech-Fauceglia, P., Yan, L., Sharifian, M., Ren, X., Liu, S., Kim, G., Gayther, S. A., Pejovic, T., & Lawrenson, K. (2015). Stromal expression of fibroblast activation protein alpha (fap) predicts platinum resistance and shorter recurrence in patients with epithelial ovarian cancer. *Cancer Microenvironment*, 8, 23–31.

- Milosevic, V. (2023). Different approaches to imaging mass cytometry data analysis. *Bioinformatics Advances*, 3(1), vbad046.
- Min, H.-Y., & Lee, H.-Y. (2023). Cellular dormancy in cancer: Mechanisms and potential targeting strategies. *Cancer Research and Treatment: Official Journal of Korean Cancer Association*, 55(3), 720–736.
- Mirza, M. R., Monk, B. J., Herrstedt, J., Oza, A. M., Mahner, S., Redondo, A., Fabbro, M., Ledermann, J. A., Lorusso, D., Vergote, I., et al. (2016). Niraparib maintenance therapy in platinum-sensitive, recurrent ovarian cancer. *New England Journal of Medicine*, 375(22), 2154–2164.
- Muhlich, J. L., Chen, Y.-A., Yapp, C., Russell, D., Santagata, S., & Sorger, P. K. (2022). Stitching and registering highly multiplexed whole-slide images of tissues and tumors using ashlar. *Bioinformatics*, 38(19), 4613–4621.
- Napari Contributors. (2019). Napari: A multi-dimensional image viewer for python [Accessed: 2025-02-11]. <https://doi.org/10.5281/zenodo.3555620>
- National Cancer Institute (NCI). (2020). Phase ii study of ipilimumab monotherapy in recurrent platinum-sensitive ovarian cancer. Retrieved March 11, 2025, from <https://www.clinicaltrials.gov/study/NCT01611558?tab=table>
- National Cancer Institute, Surveillance, Epidemiology, and End Results Program. (2024). Cancer stat facts: Ovarian cancer. Retrieved October 22, 2024, from <https://seer.cancer.gov/statfacts/html/ovary.html>
- National Comprehensive Cancer Network. (2024). Clinical practice guidelines in oncology (nccn guidelines): Ovarian cancer including fallopian tube cancer and primary peritoneal cancer. Retrieved December 11, 2024, from https://www.nccn.org/professionals/physician_gls/pdf/ovarian.pdf
- Nirmal, A. J., & Sorger, P. K. (2024). Scimap: A python toolkit for integrated spatial analysis of multiplexed imaging data. *Journal of open source software*, 9(97).
- Opal Multiplex, I. (2022). Guide and image acquisition information. Retrieved May 12, 2025, from https://www.akoyabio.com/wp-content/uploads/2020/04/Opal-Reagents_Brochure_Opal-Assay-Development-Guide.pdf
- Oza, A. M., Cook, A. D., Pfisterer, J., Embleton, A., Ledermann, J. A., Pujade-Lauraine, E., Kristensen, G., Carey, M. S., Beale, P., Cervantes, A., et al. (2015). Standard chemotherapy with or without bevacizumab for women with newly diagnosed ovarian cancer (icon7): Overall survival results of a phase 3 randomised trial. *The Lancet Oncology*, 16(8), 928–936.
- Ozols, R. F., Bundy, B. N., Greer, B. E., Fowler, J. M., Clarke-Pearson, D., Burger, R. A., Mannel, R. S., DeGeest, K., Hartenbach, E. M., & Baergen, R. (2003). Phase iii trial of carboplatin and paclitaxel compared with cisplatin and paclit-

- axel in patients with optimally resected stage iii ovarian cancer: A gynecologic oncology group study. *Journal of clinical oncology*, 21(17), 3194–3200.
- Pardoll, D. M. (2012). The blockade of immune checkpoints in cancer immunotherapy. *Nature reviews cancer*, 12(4), 252–264.
- Peng, T., Thorn, K., Schroeder, T., Wang, L., Theis, F. J., Marr, C., & Navab, N. (2017). A basic tool for background and shading correction of optical microscopy images. *Nature communications*, 8(1), 14836.
- Peres, L. C., Cushing-Haugen, K. L., Köbel, M., Harris, H. R., Berchuck, A., Rossing, M. A., Schildkraut, J. M., & Doherty, J. A. (2019). Invasive epithelial ovarian cancer survival by histotype and disease stage. *JNCI: Journal of the National Cancer Institute*, 111(1), 60–68.
- Perren, T. J., Swart, A. M., Pfisterer, J., Ledermann, J. A., Pujade-Lauraine, E., Kristensen, G., Carey, M. S., Beale, P., Cervantes, A., Kurzeder, C., et al. (2011). A phase 3 trial of bevacizumab in ovarian cancer. *New England Journal of Medicine*, 365(26), 2484–2496.
- Pfeiffer, P., Goedecke, W., & Obe, G. (2000). Mechanisms of dna double-strand break repair and their potential to induce chromosomal aberrations. *Mutagenesis*, 15(4), 289–302.
- Piccart, M. J., Bertelsen, K., James, K., Cassidy, J., Mangioni, C., Simonsen, E., Stuart, G., Kaye, S., Vergote, I., Blom, R., et al. (2000). Randomized intergroup trial of cisplatin–paclitaxel versus cisplatin–cyclophosphamide in women with advanced epithelial ovarian cancer: Three-year results. *Journal of the National Cancer Institute*, 92(9), 699–708.
- Podyminogin, R., Ericson, N., Chow, J., Lin, J.-R., Chen, Y.-A., Maliga, Z., Teplitz, K., Tsai, T. P.-C., Quintanilla, A., Sorger, P., et al. (2019). Pick-seq (r): A novel technology to retrieve tissue micro-regions for rna sequencing. 121, 1–2.
- Prat, J., on Gynecologic Oncology, F. C., et al. (2014). Staging classification for cancer of the ovary, fallopian tube, and peritoneum. *International Journal of Gynecology & Obstetrics*, 124(1), 1–5.
- Reinartz, S., Schumann, T., Finkernagel, F., Wortmann, A., Jansen, J. M., Meissner, W., Krause, M., Schwörer, A.-M., Wagner, U., Müller-Brüsselbach, S., et al. (2014). Mixed-polarization phenotype of ascites-associated macrophages in human ovarian carcinoma: Correlation of cd163 expression, cytokine levels and early relapse. *International journal of cancer*, 134(1), 32–42.
- Richardson, W. H. (1972). Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America*, 62(1), 55–59.

- Rozanova, S., Barkovits, K., Nikolov, M., Schmidt, C., Urlaub, H., & Marcus, K. (2021). Quantitative mass spectrometry-based proteomics: An overview. *Quantitative methods in proteomics*, 85–116.
- Sato, E., Olson, S. H., Ahn, J., Bundy, B., Nishikawa, H., Qian, F., Jungbluth, A. A., Frosina, D., Gnjjatic, S., Ambrosone, C., et al. (2005). Intraepithelial cd8+ tumor-infiltrating lymphocytes and a high cd8+/regulatory t cell ratio are associated with favorable prognosis in ovarian cancer. *Proceedings of the National Academy of Sciences*, 102(51), 18538–18543.
- Schermelleh, L., Ferrand, A., Huser, T., Eggeling, C., Sauer, M., Biehlmaier, O., & Drummen, G. P. (2019). Super-resolution microscopy demystified. *Nature cell biology*, 21(1), 72–84.
- Schmidt, U., Weigert, M., Broaddus, C., & Myers, G. (2018). Cell detection with star-convex polygons. *Medical image computing and computer assisted intervention—MICCAI 2018: 21st international conference, Granada, Spain, September 16–20, 2018, proceedings, part II 11*, 265–273.
- Scully, R. E., Young, R. H., & Clement, P. B. (1998). *Tumors of the ovary, maldeveloped gonads, fallopian tube, and broad ligament*. American Registry of Pathology.
- Shih, I.-M., & Kurman, R. J. (2004). Ovarian tumorigenesis: A proposed model based on morphological and molecular genetic analysis. *The American journal of pathology*, 164(5), 1511–1518.
- Shuken, S. R. (2023). An introduction to mass spectrometry-based proteomics. *Journal of proteome research*, 22(7), 2151–2171.
- Sica, A., Allavena, P., & Mantovani, A. (2008). Cancer related inflammation: The macrophage connection. *Cancer letters*, 267(2), 204–215.
- Sjöberg, E., Augsten, M., Bergh, J., Jiström, K., & Östman, A. (2016). Expression of the chemokine cxcl14 in the tumour stroma is an independent marker of survival in breast cancer. *British journal of cancer*, 114(10), 1117–1124.
- Steen, H., & Mann, M. (2004). The abc’s (and xyz’s) of peptide sequencing. *Nature reviews Molecular cell biology*, 5(9), 699–711.
- Stringer, C., Wang, T., Michaelos, M., & Pachitariu, M. (2021). Cellpose: A generalist algorithm for cellular segmentation. *Nature methods*, 18(1), 100–106.
- Syed, Y. Y. (2017). Rucaparib: First global approval. *Drugs*, 77, 585–592.
- Tan, W. C. C., Nerurkar, S. N., Cai, H. Y., Ng, H. H. M., Wu, D., Wee, Y. T. F., Lim, J. C. T., Yeong, J., & Lim, T. K. H. (2020). Overview of multiplex immunohistochemistry/immunofluorescence techniques in the era of cancer immunotherapy. *Cancer Communications*, 40(4), 135–153.

- Tapia, C., Kutzner, H., Mentzel, T., Savic, S., Baumhoer, D., & Glatz, K. (2006). Two mitosis-specific antibodies, mpm-2 and phospho-histone h3 (ser28), allow rapid and precise determination of mitotic activity. *The American journal of surgical pathology*, 30(1), 83–89.
- Taylor, M. J., Lukowski, J. K., & Anderton, C. R. (2021). Spatially resolved mass spectrometry at the single cell: Recent innovations in proteomics and metabolomics. *Journal of the American Society for Mass Spectrometry*, 32(4), 872–894.
- Valcourt, J. R., Lemons, J. M., Haley, E. M., Kojima, M., Demuren, O. O., & Coller, H. A. (2012). Staying alive: Metabolic adaptations to quiescence. *Cell cycle*, 11(9), 1680–1696.
- Van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., & Yu, T. (2014). Scikit-image: Image processing in python. *PeerJ*, 2, e453.
- Van Meurs, H. S., Tajik, P., Hof, M. H., Vergote, I., Kenter, G. G., Mol, B. W. J., Buist, M. R., & Bossuyt, P. M. (2013). Which patients benefit most from primary surgery or neoadjuvant chemotherapy in stage iiic or iv ovarian cancer? an exploratory analysis of the european organisation for research and treatment of cancer 55971 randomised trial. *European journal of cancer*, 49(15), 3191–3201.
- Van Valen Lab. (2023). Deepcell applications. Retrieved February 11, 2025, from <https://github.com/vanvalenlab/deepcell-applications>
- Vergote, I., Coens, C., Nankivell, M., Kristensen, G. B., Parmar, M. K., Ehlen, T., Jayson, G. C., Johnson, N., Swart, A. M., Verheijen, R., et al. (2018). Neoadjuvant chemotherapy versus debulking surgery in advanced tubo-ovarian cancers: Pooled analysis of individual patient data from the eortc 55971 and chorus trials. *The Lancet Oncology*, 19(12), 1680–1687.
- Vergote, I. B., Van Nieuwenhuysen, E., & Vanderstichele, A. (2016). How to select neoadjuvant chemotherapy or primary debulking surgery in patients with stage iiic or iv ovarian carcinoma.
- Visikol. (2024). Tbs vs pbs in immunohistochemistry: Navigating the differences. Retrieved February 10, 2025, from <https://visikol.com/blog/2024/01/04/tbs-vs-pbs-in-immunohistochemistry-navigating-t>
- Vitzthum, F., Behrens, F., Anderson, N. L., & Shaw, J. H. (2005). Proteomics: From basic research to diagnostic application. a review of requirements & needs. *Journal of proteome research*, 4(4), 1086–1097.

- Wang, Y., Zhao, J., Xu, H., Han, C., Tao, Z., Zhou, D., Geng, T., Liu, D., & Ji, Z. (2024). A systematic evaluation of computational methods for cell segmentation. *Briefings in Bioinformatics*, 25(5), bbae407.
- Williams, G. H., & Stoeber, K. (2012). The cell cycle and cancer. *The Journal of pathology*, 226(2), 352–364.
- Wilson, E. M., Eskander, R. N., & Binder, P. S. (2024). Recent therapeutic advances in gynecologic oncology: A review. *Cancers*, 16(4), 770.
- World Cancer Research Fund International. (2024). Ovarian cancer statistics. Retrieved October 22, 2024, from <https://www.wcrf.org/cancer-trends/ovarian-cancer-statistics/>
- Wu, T., & Dai, Y. (2017). Tumor microenvironment and therapeutic response. *Cancer letters*, 387, 61–68.
- Yang, W., Han, W., Ye, S., Liu, D., Wu, J., Liu, H., Li, C., & Chen, H. (2013). Fibroblast activation protein- α promotes ovarian cancer cell proliferation and invasion via extracellular and intracellular signaling mechanisms. *Experimental and molecular pathology*, 95(1), 105–110.
- Yang, Y., Yang, Y., Yang, J., Zhao, X., & Wei, X. (2020). Tumor microenvironment in ovarian cancer: Function and therapeutic strategy. *Frontiers in cell and developmental biology*, 8, 758.
- Zaima, N., Hayasaka, T., Goto-Inoue, N., & Setou, M. (2010). Matrix-assisted laser desorption/ionization imaging mass spectrometry. *International journal of molecular sciences*, 11(12), 5040–5055.
- Zhang, L., Conejo-Garcia, J. R., Katsaros, D., Gimotty, P. A., Massobrio, M., Regnani, G., Makrigiannakis, A., Gray, H., Schlienger, K., Liebman, M. N., et al. (2003). Intratumoral t cells, recurrence, and survival in epithelial ovarian cancer. *New England journal of medicine*, 348(3), 203–213.
- Zhang, M., He, Y., Sun, X., Li, Q., Wang, W., Zhao, A., & Di, W. (2014). A high m1/m2 ratio of tumor-associated macrophages is associated with extended survival in ovarian cancer patients. *Journal of ovarian research*, 7, 1–16.
- Zhao, L., Ji, G., Le, X., Wang, C., Xu, L., Feng, M., Zhang, Y., Yang, H., Xuan, Y., Yang, Y., et al. (2017). Long noncoding rna linc00092 acts in cancer-associated fibroblasts to drive glycolysis and progression of ovarian cancer. *Cancer research*, 77(6), 1369–1382.
- Zhu, Y., Dou, M., Piehowski, P. D., Liang, Y., Wang, F., Chu, R. K., Chrisler, W. B., Smith, J. N., Schwarz, K. C., Shen, Y., et al. (2018). Spatially resolved proteome mapping of laser capture microdissected tissue with automated sample transfer to nanodroplets. *Molecular & Cellular Proteomics*, 17(9), 1864–1874.

Zuiderveld, K. J., et al. (1994). Contrast limited adaptive histogram equalization.
Graphics gems, 4(1), 474–485.