Marc Escobosa Olmo

Spatial transcriptomics analysis of DUrvalumab and TREmelimumab versus chemotherapy as a NEOadjuvant approach to muscle-invasive urothelial bladder cancer (MIBC)

MASTER'S THESIS

supervised by Dr. Eduard Porta and Dr. Sergio Gómez

Master's Degree in Biomedical Data Science

Barcelona, 2024

Dr. Eduard Porta, certifies that the student Marc Escobosa Olmo has elaborated the work under his direction and he authorizes the presentation of this Master's Thesis for its evaluation.

Advisor(s) signature:

2. Abstract

The goal of this thesis was to investigate the distinctions between patients who positively respond to treatment and those who do not. Specifically, we aimed to understand the interactions at the border of the cancer and the tumour microenvironment (TME) and identify the mechanisms that either hinder or facilitate an effective response from the immune system and, consequently, the patient. Our approach involved utilising spatial transcriptomics and advanced computational data analysis to pinpoint the specific biomarkers that indicate treatment response and reveal the mechanisms that drive tumour suppression or progression.

We employed the Space Ranger tool for all the Visium data processing, the "ESTIMATE" R package for the cell type assignment, Semla for the sub-selection of the cancer border, and 16 cell-cell communication inference resources: Consensus, CellCall, CellChatDB, CellPhoneDB, Ramilowski2015, Baccin2019, LRdb, Kirouac2010, ICELLNET, iTALK, EMBRACE, HMPR, Guide2Pharma, ConnectomeDB2020, CellTalkDB, and OmniPath from which we did the selection of ligand-receptor pairs. The initial findings from MISTyR and LIANA showed little significance; however, using the linear model approach led to more meaningful results. Importantly, we discovered strong connections between gene pairs and treatment responses, revealing different gene expression patterns in each Arm. In particular, the standard arms had few upregulated genes in non-responders, whereas the DUTRENEO arm had fewer upregulated genes in complete responders, emphasising the treatment-specific effects on gene expression.

We are only grasping the surface of the problem; we have only begun to understand the development of MIBC and its response to treatment. Our goal is to offer valuable insights into potential biomarkers and targets for personalised medicine, aiming to advance oncology in the management of MIBC. For that purpose, we will delve deeper into the intrinsic mechanisms that this field encapsulates.

3. Table of Contents

Content

4. List of Abbreviations and Symbols

Abbreviation or Symbol	Description
MIBC	Muscle Invasive Bladder Cancer
DUTRENEO	DUrvalumab and TREmelimumab versus chemotherapy as a NEOadjuvant approach
L-R	Ligand - Receptor
TF	Trasncription Factor
FF	Freash Frozen
FPE	Formalin-fixed and Paraffin-Embedded
FxF	Fixed frozen
H&E	Hematoxylin and Eosin stain
TME	Tumour Micro Enviroment
LogFC	Log fold change
RC	Radial Cystectomy
pCR	pathological complete response rate (pCR)
TIS	Tumour Inflammation score
NSCLC	Non-Small Cell Lung Cancer
CCC	Cell-Cell Communication

5. Introduction

5.1 MIBC

Muscle-invasive urothelial bladder cancer (MIBC) happens when the tumour tissue from the bladder starts spreading into the muscle layer of the bladder wall. This event triggers into a more aggressive and potentially life-threatening cancer type. This event accounts for approximately 75% of all bladder cancer cases[1]. A standard diagnosis is based on the detection of clinical symptoms, imaging studies, cystoscopy, and histopathological examination of biopsy samples to confirm the diagnosis and determine the subtype of tumour[2]. Treatment approaches for MIBC can come in many forms and usually involve a collaborative effort among surgery, radiation, and

medical oncology. Always, the primary goal is to achieve a cure or durable remission while improving the patient's quality of life. Radical cystectomy (RC), the standard treatment, implies surgical removal of the bladder, surrounding lymph nodes, and sometimes adjacent organs like the prostate or uterus, often supplemented with chemotherapy and radiation therapy[3]. Lately, alternative approaches have emerged for patient's ineligibility for RC due to age or comorbidities. These methods involve using chemotherapy and radiation therapy to shrink the tumour before surgical removal. Additionally, immunotherapy, particularly checkpoint inhibitors, has become integral to MIBC treatment, either alone or in combination with traditional therapies to enhance outcomes[4]. While RC remains the gold standard in the treatment of MIBC, new strategies and treatments using immunotherapy techniques are increasingly improving treatment effectiveness and patient quality of life. In spite of all the advancements, there are still limitations and challenges with current treatments, which move the focus to neoadjuvant approaches.

5.2 Limitations of Current Treatments.

There are some clear limitations surrounding the current treatments. In regards of the latest advancements, chemotherapy or radiation therapy oftentimes does not result in satisfactory response rates for MIBC, resulting in a large portion of patients with no complete remission or facing disease progression[7]. Moreover, RC, the standard treatment, carries notable risks, especially among older patients or those with comorbidities, highlighting the need for less invasive alternatives. Every approach that implies surgical bladder removal involves a significant impact on the patient's quality of life, affecting urinary function and body image, emphasising the necessity for approaches preserving organ function[5,6]. On top of that resistance to chemotherapy is commonly found in MIBC patients, a fact that contributes to the treatment to unfavourable outcomes[8]. MIBC treatment protocols need to be standardised to improve consistency in outcomes and tailor treatments to individual patient needs.

5.3 Neoadjuvant Approaches

Here is where neoadjuvant therapy enters the stage. Based on the administration of therapeutic agents before a main treatment, neoadjuvant chemotherapy offers promising results by reducing tumour mass prior to surgery, which enhances the response, and potentially increases the likelihood of successful outcomes. Neoadjuvant approaches show a good synergy with immunotherapy. By combining both techniques, we can improve treatment efficacy, and patient prognosis. Neoadjuvant strategies mitigate the morbidity associated with RC, enabling less invasive surgical procedures and better post-operative recovery. On top of that, neoadjuvant approaches enable personalised medicine, tailoring treatments based on individual patient

characteristics and optimal outcomes. Reducing the need for radical surgeries and enhancing treatment responses, thanks to neoadjuvant approaches we can contribute to improved quality of life for MIBC patients[7,9].

5.4 DUTRENEO

The DUTRENEO phase II clinical trial[10] represents a significant advancement in the exploration of neoadjuvant approaches together with immunotherapy for managing muscle-invasive urothelial bladder cancer (MIBC) and an opportunity to research the intrinsic mechanisms separating the immune infiltration nature of tumours. Conducted as a randomised phase II trial, its primary objective was to assess the efficacy and safety of durvalumab and tremelimumab compared to chemotherapy in this setting. Durvalumab and tremelimumab, two drugs that belong to the monoclonal antibody group, have shown promising results for the treatment of various cancers, including muscle-invasive urothelial bladder cancer (MIBC). Durvalumab, an anti-programmed cell death ligand-1 (PD-L1) antibody, operates by obstructing the PD-L1 and PD-1 interaction, thereby reinforcing the immune system's response against cancer cells. On the other hand, tremelimumab, an anti-cytotoxic T lymphocyte-associated antigen 4 (CTLA-4) antibody, impedes the activity of regulatory T cells that dampen the immune response[11,12].

This trial targeted a subset of MIBC patients identified through a tumour pro-inflammatory IFN-gamma signature, quantified as the tumour inflammation score (TIS). The underlying hypothesis posited that the expression of certain genes within this signature could predict a favourable response to immunotherapy treatment.

Within the framework of the DUTRENEO trial, patients were stratified based on their TIS status. Those categorised with 'hot' tumours, characterised by scoring in the top two-thirds of the TIS distribution, were randomised to receive either durvalumab and tremelimumab or conventional chemotherapy. Conversely, patients with 'cold' tumours, representing the lower third of the TIS distribution, received standard chemotherapy alone.

The primary endpoint of the trial was the assessment of the complete pathological response rate (pCR) following neoadjuvant treatment. The results unveiled a notable achievement, with the combination of durvalumab and tremelimumab yielding a complete pathological response rate of 34.8%, coupled with a favourable safety profile. Intriguingly, however, the correlation between the tumour inflammation score (TIS) and treatment response within this group appeared less evident. [12]

The outcomes gleaned from the DUTRENEO trial underscore the burgeoning potential of neoadjuvant immunotherapy in reshaping the therapeutic landscape of MIBC, particularly for

patients harbouring 'hot' tumours. These findings advocate for continued exploration of immunotherapeutic interventions in the neoadjuvant setting, alongside endeavours to elucidate biomarkers capable of predicting treatment response.

Beyond the realm of MIBC, the synergistic effects of durvalumab and tremelimumab have been explored in diverse clinical settings. In a phase 3 trial targeting unresectable hepatocellular carcinoma, the combination exhibited encouraging clinical activity and safety, outperforming the standard treatment, sorafenib, in terms of overall survival. Similarly, in advanced non-small cell lung cancer (NSCLC), the combination showcased superior overall survival rates and a heightened rate of complete responses compared to durvalumab monotherapy.[13]

These collective findings underscore the favourable safety profile and therapeutic efficacy of the durvalumab and tremelimumab combination across various malignancies, including MIBC and NSCLC. However, further research from multiple points of view is indispensable to determining the optimal use of this combination and comprehending its shaded benefits and risks within distinct patient populations. That's where technologies such as spatial transcriptomics might bring an edge to the understanding of the treatments.

5.5 Spatial Transcriptomics

The Spatial transcriptomics approach has been increasing in popularity over the last years, this increase, is mainly due to the amount of value that provides researchers to analyse gene expression patterns within tissue samples in the spatial context of the cells. This approach to the analysis provides a powerful tool that will help to decipher the intricacies of tissue architecture and cellular interactions, adding an extra variable of complexity allowing for more detailed analysis.

Conventional transcriptome analysis techniques, such single-cell RNA sequencing (scRNA-seq), offer important insights into cellular resolution gene expression levels. These techniques, however, more often than not, need the dissociation of tissues into individual cells, which can blur our comprehension of how cells interact in their original tissue environments and disturb the spatial relationships between cells.

Spatial transcriptomics overcomes this limitation by focusing on the spatial landscape and gene expression distribution across the tissue samples. By capturing the spatial distribution of gene transcripts. Spatial transcriptomics provides a comprehensive view of how gene expression fluctuates across different regions of a tissue. This kind of data can reveal extremely important insights especially related to cell-to-cell interaction, tissue distribution, organisation, cell infiltration and disease mechanisms.

The spatial context provided by spatial transcriptomics can bring immeasurable value in fields especially in cancer research and immunology. This extra spatial context may help to understand the molecular mechanisms that cause tissue differentiation and organ formation, provide insight into the heterogeneity of tumours, microenvironment interactions and possible therapeutic targets in cancer research and even reveal the distribution and function of immune cells in tissues in the field of immunology. I believe that spatial transcriptomics has the power to enhance our understanding of complex biological systems by integrating spatial information with gene expression data[14].

5.6 Visium: Revolutionizing Spatial Transcriptomics

Leading the spatial transcriptomics technology and protocols there is Visium, a groundbreaking platform developed by 10x Genomics[15]. Using spatially resolved transcriptomics, Visium employs barcoded spots to capture and quantify gene expression at the single-cell level, which allows unravelling the spatial distribution and quantification of the genes present in the sample.

5.6.1 FFPE Tissue Samples

In spatial transcriptomics FFPE samples or formalin-fixed paraffin-embedded tissue samples are most commonly used, usually coming from biopsies or surgical resections, preserving cellular spatial organisation, and offering a window into gene expression within a tissue's natural context[16]. The Visium Spatial Gene Expression Solution by 10x Genomics[17] outstands in the field of spatial transcriptomics, providing a detailed protocol designed for FFPE tissue samples. This sophisticated approach reveals the complex spatial arrangement of gene expression in whole tissue sections, opening doors to groundbreaking discoveries in cellular behaviour. Allowing us to analyse gene expression while taking tissue structure into consideration.

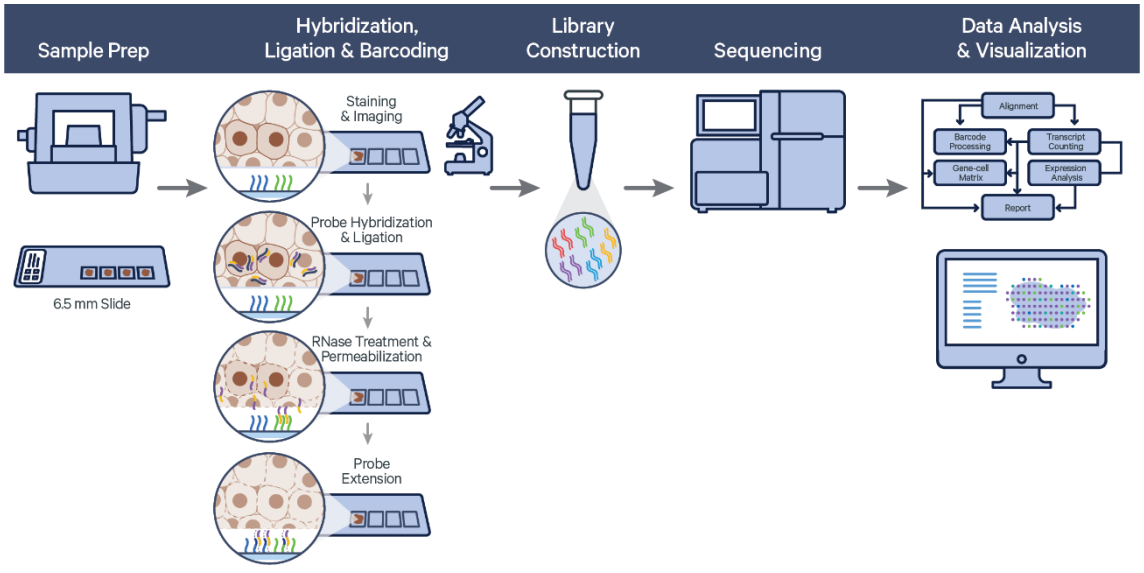

Figure 1: The Visium Spatial Gene Expression for FFPE workflow.

5.6.2 Workflow:

1. Tissue Preparation:

It start preparing the tissue sections, which in our case are formalin-fixed paraffin-embedded (FFPE). Initially, the Visium platform was developed for fresh frozen tissue sections, but a FFPE-compatible protocol has been developed to expand its applicability.

2. Probe Hybridization:

Gene-specific probes are used to capture mRNA molecules within the tissue sections. These probes are designed to target specific transcripts of interest.

3. In Situ Ligation:

In the case of FFPE samples, gene-specific probes are ligated together in situ within the tissue section using the target mRNA as a template. This process, known as RNA-templated ligation, improves target specificity.

4. Probe Release and Capture:

The ligated probes are then released from the tissue through permeabilization, diffused, and bound to barcoded capture oligonucleotides printed on a slide. This step allows for the spatial mapping of gene expression.

5. Barcode Addition and Amplification:

After capture, the probes are extended to incorporate a spatial barcode. Captured mRNA molecules undergo reverse transcription into complementary DNA (cDNA), followed by amplification to generate ample material for downstream sequencing.

6. Library Preparation:

Enzymatic fragmentation and ligation of sequencing adapters prepare amplified cDNA for high-throughput sequencing, ensuring robust data generation.

7. Sequencing:

Prepared libraries undergo sequencing on next-generation platforms, yielding millions of short reads corresponding to individual mRNA molecules.

8. Data Analysis:

10x Genomics' analysis pipelines process sequencing data, aligning reads, processing barcodes, and quantifying gene expression. Visualisation and exploration of spatial

transcriptomics data provides a general visualization of gene expression patterns and helps understanding the spatial context of the sample.

5.6.3 Spatial Resolution:

The Visium Spatial Gene Expression Slides are able to captures up to 4 samples at the same time in grids of 8 mm x 8 mm where the area where the barcoded spots are found measures 6.5mm x 6.5mm [Fig 2].

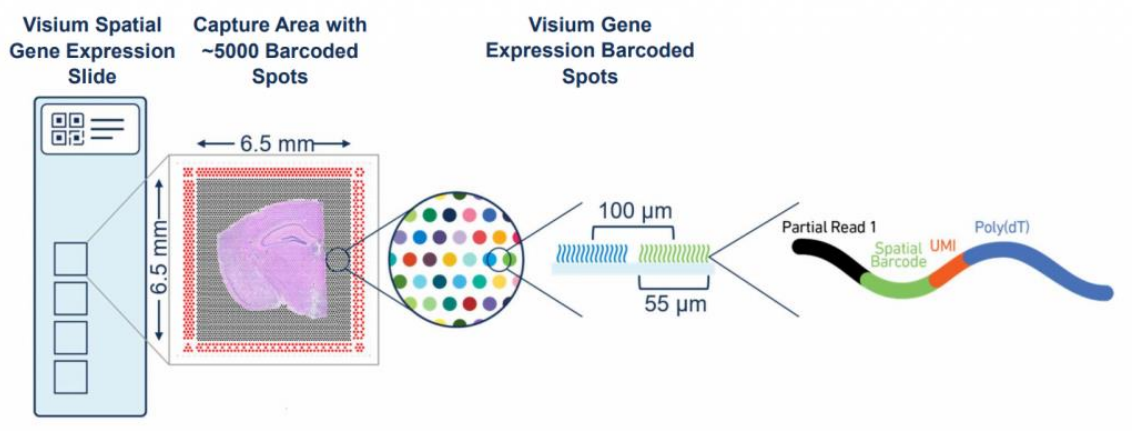

Figure 2: Visium Spatial Gene Expression Slide for FFPE (https://www.10xgenomics.com/).

5.6.4 Data Analysis:

Once the sequencing library is generated, we can perform downstream analysis to study gene expression patterns across the tissue section. Several computational methods and analysis tools can be used to deconvolve the expression profiles generated by the Visium technology and understand the relative cellular composition within the tissue.

5.6.5 Visium Key Properties:

Visium technology offers some clear advantages that made it the go to technology for the analysis of our samples;

- High Spatial Resolution: Visium offers a high spatial resolution and precise mapping of gene expression within tissue sections which facilitates the identification of cellular heterogeneity and distinct expression patterns.
- Multiplexed Analysis: The protocol supports analysis of thousands of genes within a single tissue section.
- Versatility and Scalability: Compatible with various tissue types, Visium's scalability allows for the analysis of multiple sections in parallel, facilitating large-scale spatial transcriptomics studies across diverse experimental models.

5.7 Available Data

The available dataset for our analysis, comprises tissue samples obtained from various segments of the DUTRENEO Visium FFPE slides, each uniquely identified with specific codes (DU1, DU2, DU3, etc.). These codes facilitate differentiation between the different samples utilised in the study. From the DUTRENEO clinical trial dataset we have access to 41 H&E DUTRENEO slides which will be used as input for the Visium workflow. The resulting data from the pipeline will be the playground from which all the analysis will be performed in this thesis. In total, we will have 41 samples from 33 patients.

COLD	COMPLETE	MISSING	NO	PARTIAL
DUTRENEO	0	0	0	0
STANDARD	14	0	6	0

Table 1: Number of samples available based on response and treatment for Cold tumours.

HOT	COMPLETE	MISSING	NO	PARTIAL
DUTRENEO	3	0	5	3
STANDARD	4	1	3	2

Table 2: Number of samples available based on response and treatment for Hold tumours.

5.8 Computational Tools and Packages

5.8.1 MistyR- Multiview Intercellular SpaTial modelling framework

MistyR, is a Multiview Intercellular Spatial Modelling Framework designed to analyse spatially resolved data without the need for cell type annotation. The way MISTy operates is by constructing a nonparametric and nonlinear model of the expression of each of the available markers as an intrinsic function. This tool enables a comprehensive exploration of marker interactions through the profiling of intra- and intercellular relationships. MISTy offers a versatile framework capable of handling a custom range, called “views”, depending on the scale they are referred as; Intraview, when the perspective focuses on capturing the interactions and regulatory mechanisms within a specific cell or location; Juxtaview, refers to the perspective that captures the interactions and relationships between a specific cell and its immediate neighbouring cells; Paraview, refers to the perspective that captures the interactions and relationships between a specific cell and its broader spatial context within a sample. MISTyR maintains agnosticism towards potential bias sources, building a nonparametric and nonlinear model for the expression of each marker. This model is established as a function of the simultaneous expression of all other markers (intrinsically) or considering the expression of other markers or features within the available views[18].

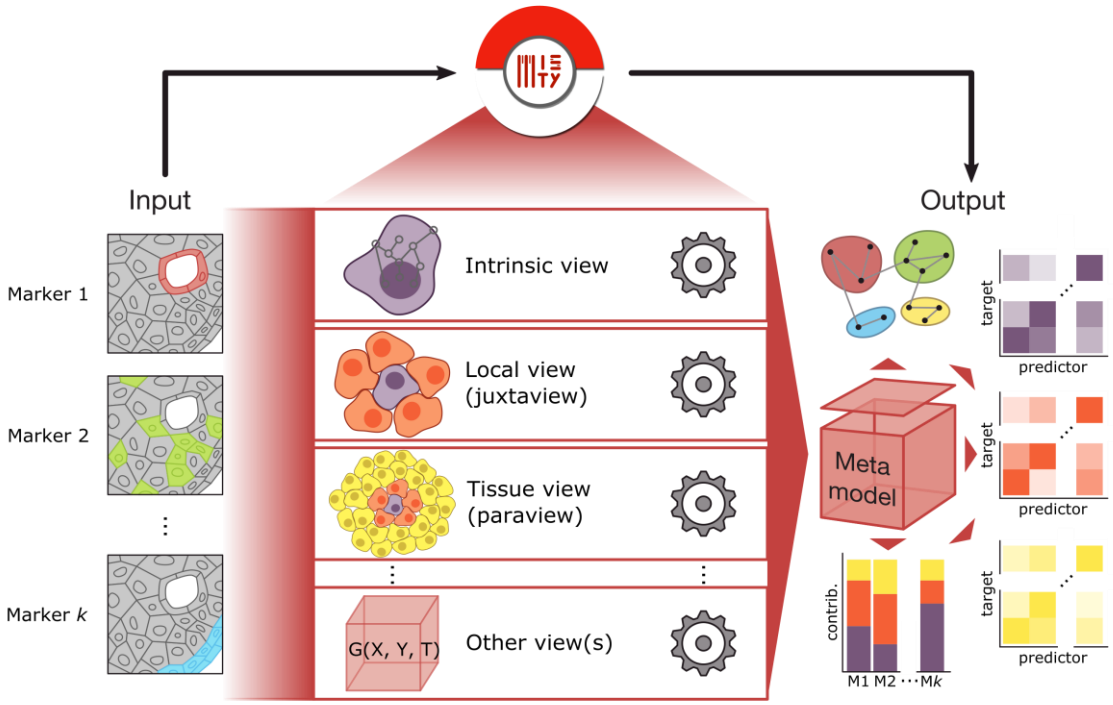

Figure 3: MISTy workflow

5.8.2 LIANA: a LIgand-receptor ANalysis framework

LIANA is the framework we decided to introduce to facilitate the comparison of cell-cell communication (CCC) inference resources and methods. It serves as an open-source interface that allows users to analyse and compare different resources and methods for predicting ligand-receptor interactions from single-cell RNA sequencing (scRNA-Seq) data.

The main purpose of LIANA is to provide a standardised platform for evaluating and prioritising cell-to-cell communication interactions. By using LIANA, we can easily assess the performance of various CCC inference resources and methods, understand the overlap and uniqueness of predicted interactions, and explore the influence of different resources and methods on the predicted intercellular interactions. This tool offers several key strengths that make it a valuable tool for analysing CCC.

LIANA provides a standardised platform for comparing and evaluating different cell-to-cell communication inference resources and methods. This allows us to assess the performance of various tools in a consistent manner. It also allows for the decoupling of cell-cell communication methods from their corresponding resources. This way we can explore the impact of different combinations of methods and resources on predicted interactions. Moreover, provides a

consensus ranking that ensures the finding of robust and reliable cell-to-cell communication interactions[19].

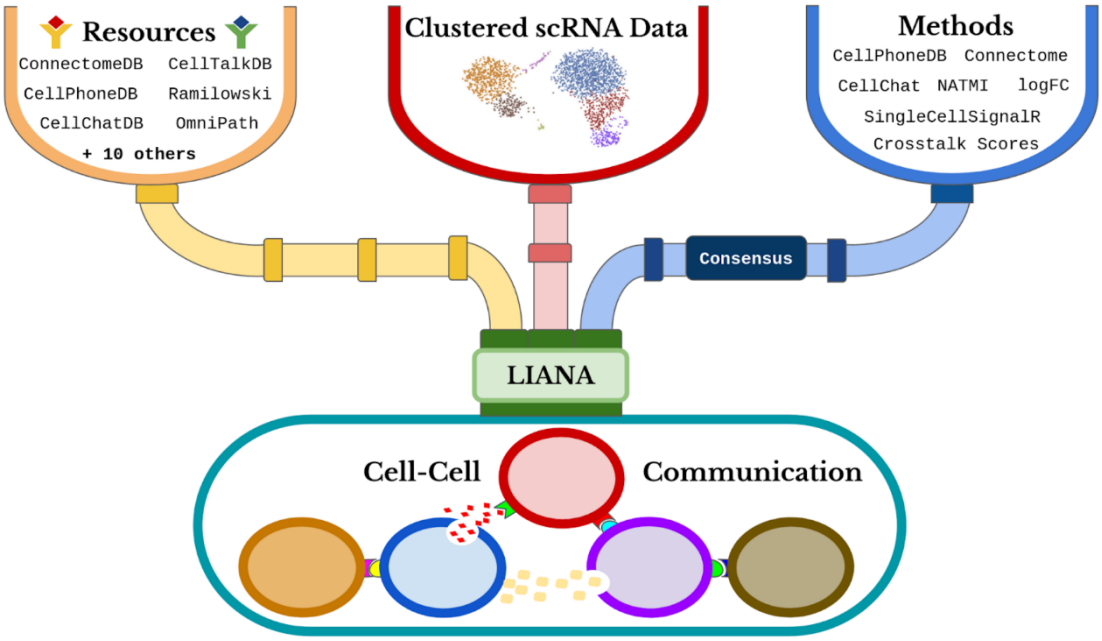

Figure 4: LIANA workflow

5.9 State-of-the-art

Spatial transcriptomics is a field that has been receiving constant updates in the form of new tools, pipelines, and technologies. This field finds itself at the intersection of technology, experimental biology, and computational analysis, offering a unique insight for the investigation of tissue environments. Recent technological innovations have enhanced the accessibility and resolution of spatial transcriptomics. Leading platforms such as Visium by 10X Genomics[15] and CosMx by Nanostring[20], which facilitate high-dimensional gene transcription analysis achieving even subcellular spatial resolution.

We find spatial transcriptomics increasingly being integrated with other omics techniques, such as spatial proteomics and spatial genomics. Fact that allows for a deeper understanding and analysis of the biological systems, encompassing gene expression, protein localization, and genomic features[14].

The application of spatial transcriptomics in disease research brings new points of view to the table, particularly in understanding how the spatial distribution of different cell types affects in the different disease mechanisms. In neuroscience, for instance, a study using spatial transcriptomics has evaluated gene expression patterns surrounding amyloid plaques in Alzheimer's disease models, adding an extra layer of information over processes such as inflammation and endocytosis[14,21]. In cancer research, it has been employed to characterise

transcriptional patterning within tumours, such as localizing and measuring the infiltration of both immune cells and cancer cells, helping to understand which factor lead to the progression of the disease.

5.9.1 Bioinformatic Analysis

Advancements in bioinformatic methods have come in many forms, almost every month there is a new analysis tool that offers deeper and more sophisticated insights over the raw spatial data. These tools and pipelines enable the processing of spatial transcriptomic data, allow to integrate the spatial context with single-cell RNA sequencing data, deduce the inference of cell-cell interactions, perform deconvolution to identify cell types, etc. Such analytical capabilities are crucial for deciphering the complex dynamics of cellular environments within tissues. The current state of bioinformatic analysis in spatial transcriptomics is marked by this rapid advancement, interdisciplinary collaborations, and an increasing interest in the field. These developments are far from slowing down which will continue to bring more and more advances in our understanding of complex biological systems and disease processes, using the seemingly unlimited potential of spatial transcriptomics in biomedical research.[21–24]

5.9.2 Analysis of the Tumour Microenvironment in MIBC.

Spatial transcriptomics among all the benefits and new insights that can brings to the research field it can specifically provide extremely valuable insights into the understanding of the interactions between cells and their microenvironment in tumour samples.[25]

The tumour microenvironment (TME) plays a critical role in the development and progression of any cancer, and is especially important in muscle-invasive bladder cancer (MIBC). The TME comprises a diverse array of cell types, including immune cells, stromal cells, and cancer cells, all of which interact with tumour cells to influence disease progression. In MIBC, spatial transcriptomics and single-cell analysis have been decisive in studying the TME and its impact on the disease. For instance, a study published in Genome Medicine utilised single-cell RNA sequencing to depict the tumour landscape of a chemo-resistant metastatic MIBC case[26]. The findings indicated that the TME contributed significantly to therapeutic failure and tumour evolution, underscoring the importance of monitoring the evolutionary trajectories of tumour cells and the surrounding TME using pre- and post-treatment samples.

Additionally, another study published in the Journal of Translational Cancer Research employed single-cell RNA sequencing and CITE-Seq analysis to characterise patient urine samples in an unbiased manner. This study discovered that urine from bladder cancer patients contains not

only tumour cells but also immune cells reflective of the TME. These findings suggest that non-invasive urine sampling could serve as a biomarker for diagnosing and monitoring MIBC.[27]

We can confidently say that spatial transcriptomics and single-cell analysis are really powerful tools for investigating the TME and its role in MIBC development and progression. Bringing important insights into the cell-to-cell interactions within the microenvironment.

5.9.3 Current gaps in the field

While significant lots progress has been made in the research on spatial transcriptomics, single-cell analysis, and the tumour microenvironment (TME) in muscle-invasive bladder cancer (MIBC), there are several gaps which remain unclear.

1. Ligand-Receptor Interactions in MIBC:

- Despite the evidence that we have regarding ligand-receptor interactions playing a crucial role in the development and progression of MIBC, comprehensive studies over specific pathways are lacking. Pathways and interactions that justify the aggressiveness of this kind of cancer or differentiate a responding from a non-responding patient to therapy. We need further research in order to find the specific mechanisms by which these interactions influence the disease.

2. Border Analysis of Tumour and TME:

- The interactions at the tumour-TME border are key for understanding tumour progression and response to therapy, as is there where the majority of interactions between the two tissues take place. There is a surprising lack of studies focusing on the border between the TME and the cancer tissue, talking the research in terms of immune penetration and cellular invasion. We believe that investigating these interactions at the tumour border can reveal crucial insights into how cancer cells communicate with their microenvironment and therefore be able to identify potential therapeutic targets.

3. Cell-to-cell communication in the TME:

- The TME is a complex environment where several interactions occur among cancer cells, immune cells, and stromal cells. We need more detailed insights that dissect the specific interactions that facilitate these cellular dialogues and their roles in MIBC progression.

6. Objectives

In this thesis we aim to address these gaps, providing a deeper understanding of the molecular interactions at the tumour-TME border, focusing on ligand-receptor dynamics and their implications for MIBC progression and treatment response.

The primary goal in this thesis is to delve deeply into the intricate interactions occurring between tumour cells and the tumour microenvironment (TME) in the area where these two collide, within muscle-invasive urothelial bladder cancer (MIBC). Focusing on addressing all of the gaps that I previously mentioned, investigating the specific network of signalling pathways and molecular dialogues taking place between tumour cells and the surrounding TME, focusing on ligand-receptor dynamics and their implications for MIBC progression and treatment response, taking a look at all the literature and databases for known ligand-receptor interactions. Basing our study in the DUTRENEO clinical trial data to unveil the spatial distribution of these crucial ligand-receptor interactions, transcription factors mediating cell-to-cell communication. With all this data gathered, we want to take a look at the differences in expression and the spatial distribution between the patients who fully responded to the treatment and those who did not, aiming to find enough statistical power to prove these differences. Some ways to see potential differences will be through the use of advanced computational analyses, including linear models and correlation analyses, also adding different tools available in the field such as MISTyR[18], and LIANA[19], tools that align with our research interests. In order to pinpoint spatial variations in gene expression profiles between hot and cold tumour regions, and identify molecular features associated with treatment response by analysing spatial transcriptomics data to find predictive biomarkers indicative of treatment efficacy.

7. Methodology

7.1 Space-Ranger

Space Ranger is a tool designed by and for 10x Genomics Visium data. It enables users like us to profile whole transcriptomes in formalin-fixed and paraffin-embedded (FFPE), fixed frozen (FxF), and fresh frozen (FF) tissues. Space Ranger provides multiple pipelines depending on the type of raw data that the user provides. In our case, we will use one specific pipeline named “Space-Ranger count” which is known to be the main pipeline and compatible with the majority of Visium data. In this pipeline we give as input a reference, the images of our samples, and the corresponding FASTQ files to generate feature-barcode matrices. Moreover, the pipeline offers a general preprocessing and analysis of our input data.

7.2 Count for Visium CytAssist, Spatial Gene and Protein Expression

The needed input consists of the species-specific pre-compiled transcriptome files, the species-specific probe set reference file in CSV format, sample sequencing files in FASTQ format, a single brightfield image with H&E staining in either TIFF or JPEG formats, the Visium slide serial number, and the Capture Area identifier on the Visium slide.

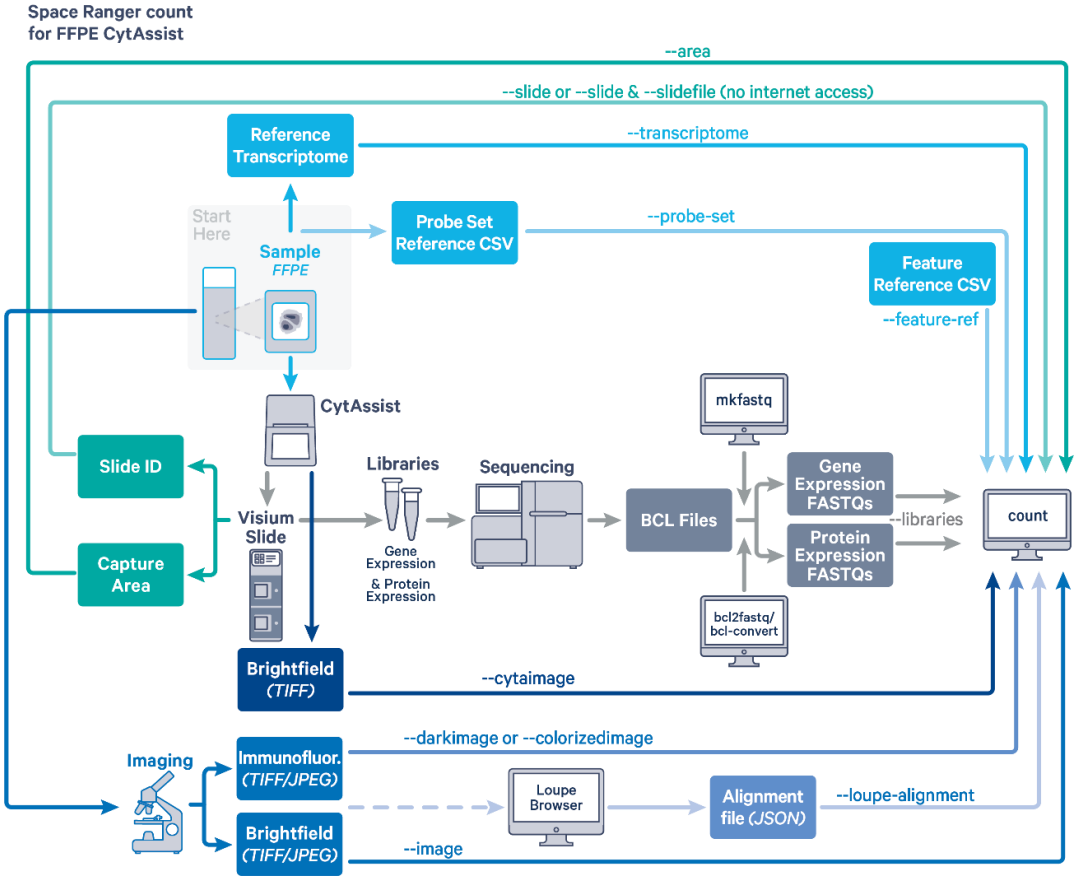

Figure 5: Space Ranger count for FFPE analysis pipeline

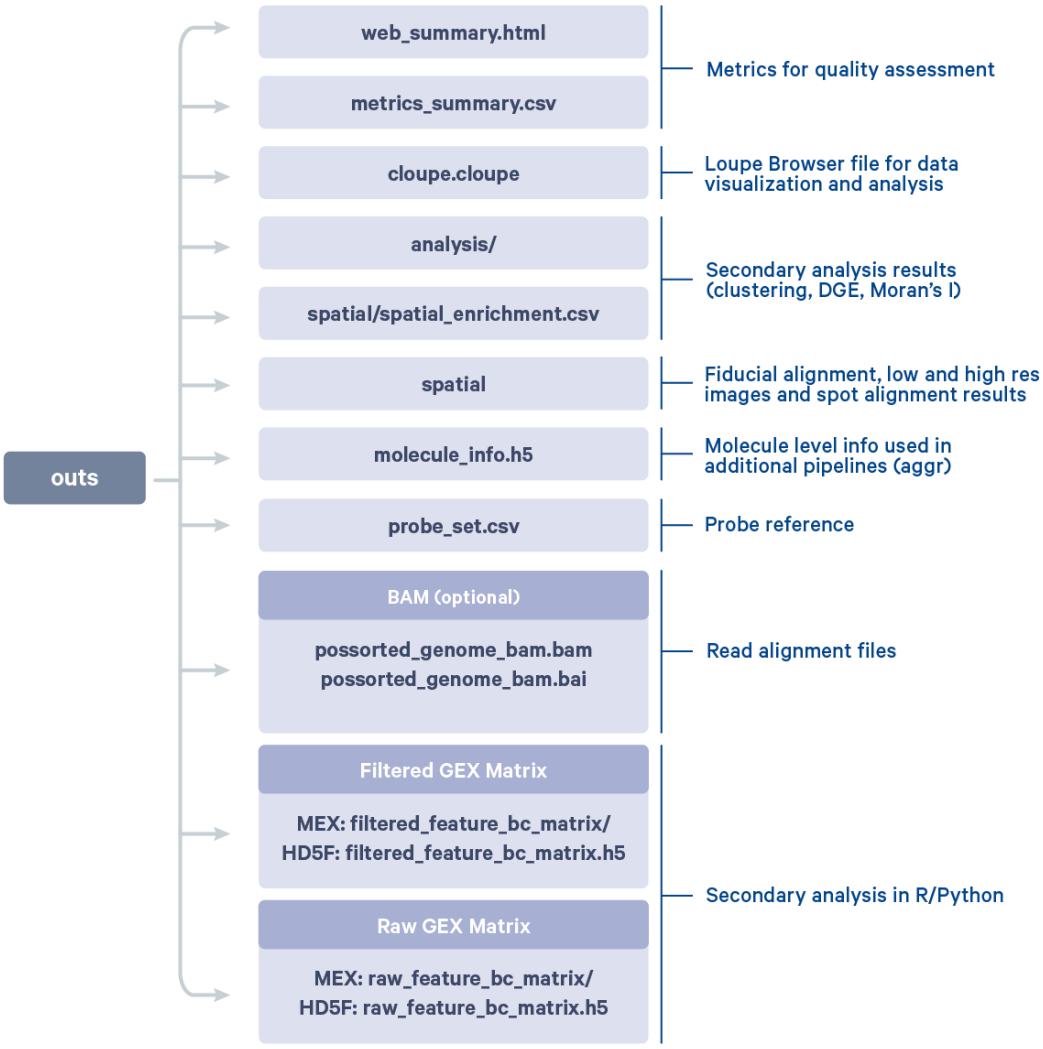

Figure 6: Output schema for the Space Ranger count FFPE analysis pipeline

7.3 Data preprocessing

All VISIUM data and raw fastq files of all 33 samples were processed with the Space Ranger pipeline (version 1.3.1, 10x Genomics) and mapped to the pre-built human reference genome (GRCh38).

The output of Space Ranger, was processed using the Semla R package[28]. Initially, genes with fewer than 5 UMIs were excluded from the analysis. Subsequently, mitochondrial and ribosomal genes, as well as non-protein-coding genes, were eliminated. Finally, spots were filtered out if the number of features was below a threshold value determined by the distribution in each sample. Each filtered sample's gene expression was then normalized individually using Seurat's SCTransform function for variance stabilising transformation.

7.4 Cell type assignment

The neoplastic cell purity within each (sub)spot in the VISIUM datasets was assessed using the "ESTIMATE" R package. This package has a successful track record of quantifying the presence of cancer, immune, or stromal cells in RNAseq profiles from bulk tumour samples. Given that each

(sub)spot in our dataset may contain multiple cells (usually between 1 and 10), we treated it as a mini-bulk RNAseq and applied the same deconvolution approach. To quantify, we first extracted the gene expression of each sub(spot) and converted it into GCT format. Subsequently, we used the "estimate" function to calculate the stroma, immune, and ESTIMATE scores. After adding ESTIMATE scores as metadata, they were utilized to form five clusters in each sample through the k-means algorithm. The cluster with the lowest ESTIMATE values indicates the highest purity of neoplastic cells, while the cluster with the highest ESTIMATE values indicates the lowest purity of neoplastic cells (TME). To determine if spots from cluster #3 could be categorized into either the cancer (clusters 1&2) or the TME (clusters 4&5) compartments, we conducted a differential gene expression analysis between cancer and TME compartments using the FindMarkers function from Seurat. Genes whose avg_log2FC >= 1 were considered upregulated in the cancer compartment, whereas those with avg_log2FC <= -1 were considered upregulated in the TME compartment. Next, we calculated the average expression of all genes from each list separately in each spot of cluster #3 (or "Buffer"), followed by taking the log2 of the division between the average expression of cancer genes and TME genes for each spot. Finally, if the Log_2FC value was >= 1 the spot was assigned to the cancer compartment, whereas if it was <= -1 it was assigned to the TME compartment. Spots who had Log_2FC values between -1 and 1 were kept as cluster #3 (or "Buffer").

$$Log_2FC = log_2\left(\frac{Row\ Means\ (Buff_{TME})}{Row\ Means\ \ (Buff_{cancer})}\right)$$

7.5 Patient Metadata

The Metadata for each patient included sample ID, type, treatment, and response, and it was saved into a csv file for easier manipulation and data retrieval at any moment. The main purpose of this metadata is to easily separate the samples based on the type of cancer (HOT, COLD) and treatment received (STANDARD, DUTRENEO). We decided that every analysis would be performed by grouping based on the "Arm". Each arm is the possible combination between the type and the treatment, resulting in 3 different "Arms"; Hot-Standard, Cold-Standard, and Hot-DUTRENEO.

7.6 Ligand-Receptor interactions Database

In order to obtain all of the ligand-receptor interactions, we decided to use LIANA, this tool integrates 16 cell-cell communication inference resources (Consensus[29], CellCall[30], CellChatDB[31], CellPhoneDB[32], Ramilowski2015[33], Baccin2019[34], LRdb[35], Kiroauc2010[36], ICELLNET[37], iTALK[38],

EMBRACE[39], HPMR[40], Guide2Pharma, ConnectomeDB2020[41], CellTalkDB[42], OmniPath[43]) giving us a comfortable number of possible interactions to explore. These Interactions were then filtered based on the following criteria; each of the genes in the pair should be present in at least 1% of the spots in every sample. This curated list of ligand-receptor pairs was then used as input for subsequent analyses, providing a reliable foundation for studying cell-to-cell communication.

7.7 Border Sub-setting

Using an R package called Semla[28], we were able to subset the samples, isolating only the spots corresponding to the outer border of the tumour and the adjacent areas. This subsampling strategy enabled us to focus on the interaction between the tumour microenvironment (TME) and the cancer cells at the border. At this critical interface, we expect to observe differences in cancer types and variations in cell-cell communication between responders and no responders. Specially in markers that regulate the immune infiltration and tumour proliferation.

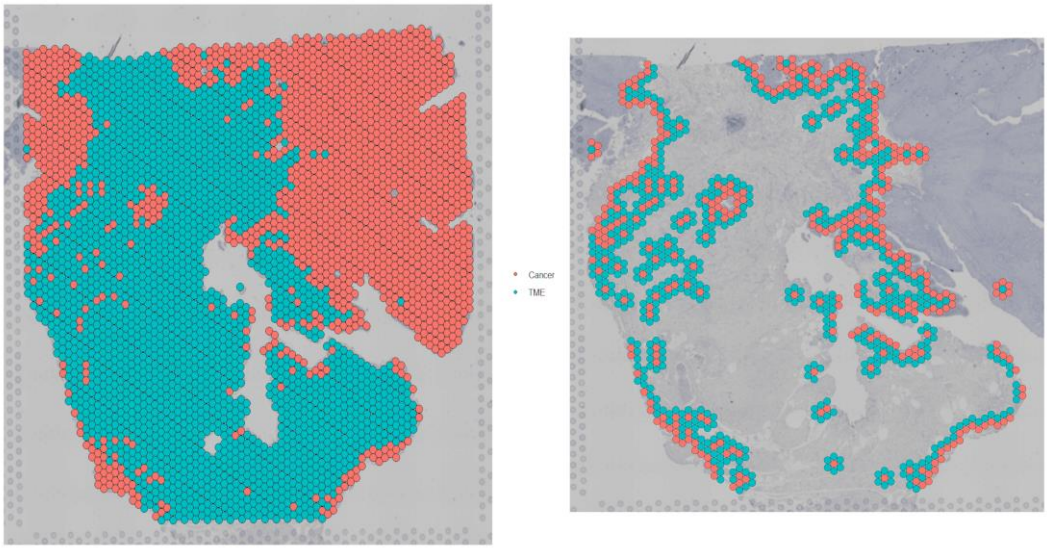

Figure 7: Before and after sub-selection of border region, blue dots represent the spots categorised as “TME” and red dots correspond to spots categorised as “cancer”

7.8 Arm Definition

To gain meaningful insights from the data, we divided it into three groups: Cold-Standard (Cold samples treated with Chemotherapy), Hot-Standard (Hot samples treated with Chemotherapy), and Hot-DUTRENEO (Hot samples treated with DUTRENEO). This division allowed us to focus on the differences between responders and non-responders and to understand the differences between the groups within each specific context.

7.9 MISTyR

The MISTy framework was employed to analyse our spatial data and try to infer cell-cell communication patterns within the samples. As stated before, we used all the genes collected and filtered from LIANA and fed them as input to MISTy. In order to obtain a wide variety of results to compare we calculated a set of different views: Intraview for local cell-cell communication, Juxtaview to find interactions between neighbouring spots at increasing distances going from 1 to 4 spots of radius, and Paraview. With all these views, we were able to gain a grasp of how the expression of the ligand-receptors was evolving as we went further into the tissue.

For each ligand-receptor pair, the mean magnitude value of the interaction importance was calculated across different distances and views for each sample group. This step was done to quantify each interaction's strength and relevance within different spatial contexts.

Gene pairs that exhibited a statistically significant difference in mean magnitude values between different Arms were identified. The difference in mean values for each pair was calculated to highlight interactions that were particularly distinct in specific Arms (Cold-Standard, Hot-Standard, and Hot-DUTRENEO).

To focus on the most significant interactions, a threshold was applied where values that were 2 standard deviations from the mean were considered significant. This thresholding aimed to ensure that only highly relevant and statistically significant interactions were selected for further analysis.

Interaction heatmaps were generated for each view, showing all groups together. These visualisations facilitated the interpretation of the spatial interaction patterns, in order to highlight key differences and similarities between groups.

7.10 LIANA

Initially, for each sample, we executed LIANA to generate a "LIANA object", facilitating the identification of the ligand-receptor interactions between the Cancer cluster and the Tumour Microenvironment (TME) cluster. Due to the high computational resources demanded by LIANA, it was necessary to use a high-performance computing cluster.

Subsequently, we proceeded to identify a unique set of genes by computing the average expression level in each sample. Genes expressed in at least 1% of the samples were retained for further analysis.

Then, for each classification category, we aggregated the interaction data obtained from LIANA results. To enhance clarity, we modified interaction symbols as necessary. Subsequently, we compiled data from all samples within each category. Further analysis involved calculating the mean magnitude for each interaction across all targets. To refine our findings, we filtered interactions based on their prevalence across the majority of samples within the category, as well as their mean values. Specifically, we retained interactions with mean values belonging to the 95th top percentile. Finally, to visualize and compare the interactions between the Cancer cluster and the TME cluster within each Arm, we generated plots using the `liana_dotplot` function.

7.11 Frontier Correlation

Taking as input the border spots of each sample individually, and then extracting the expression matrix that will be used to calculate the correlation scores between the ligand gene and the receptor gene, from this computation the p-value and correlation score will be stored. Then we will apply some filtering, taking only the significantly correlated pairs and only consider correlated those pairs that score higher than 0.4 or smaller than -0.4. Next, we further filtered the ligand-receptor pairs to those which mean correlation in the response group showed a difference of at least ± 0.2 with its counterpart.

Then we group them based on the arm to see if there was a significant difference in expression. When comparing the complete respondents against the non-respondents within each arm, despite finding some pairs that were statistically significant the p-values were too small. This was most likely due to the sample size. This resulted in finding little to no pairs that had the statistical power proven to be significantly more expressed despite clearly seeing a clear trend in the data. Therefore, we aimed for a different approach. To check for statistical differences between response groups we calculated the z-score for each ligand and receptor pair and stored the z-score for each spot-gene across each of the Arms. Then we fit a linear model to assess the relationship between each gene pair for treatment response.

$$Z\_score_{Receptor} = Z\_score_{Ligand} * Response + \epsilon$$

The linear model, expressed as `Z-Score receptor ~ Z-Score ligand*Response`, is a form of linear regression. Our models estimate coefficients for the intercept and each independent variable (Z-Score ligand, Response, and their interaction term). These coefficients represent the change in the dependent variable (Z-Score receptor) associated with a one-unit change in each independent variable, holding other variables constant.

The interaction term (`Z_Score_ligand*Response`) captures the combined effect of ligand expression and treatment response on receptor expression beyond their individual effects. If this term proved to be significant, it meant that the relationship between the ligand and the receptor was different between the response groups.

Using both the interaction term adjusted significance (p-value) and the difference between the mean correlation in the response groups, we built a volcano plot in order to clearly represent the most significant pairs among the findings.

8. Results

8.1 LIANA Results

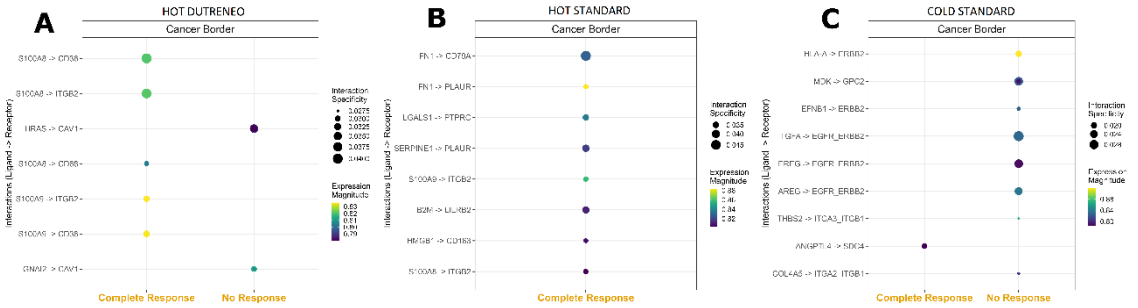

Figure 9: Dot plots illustrating ligand-receptor interactions at the cancer border for three conditions: (A) HOT DUTIRENEO, (B) HOT STANDARD, and (C) COLD STANDARD. Each plot displays interactions for samples with complete response (left) and no response (right). The y-axis lists the interactions, and the x-axis categorizes the responses. Dot sizes represent interaction specificity. Dot colours indicate expression magnitude, with a gradient from yellow (higher expression) to purple (lower expression).

8.2 MISTyR Resutls

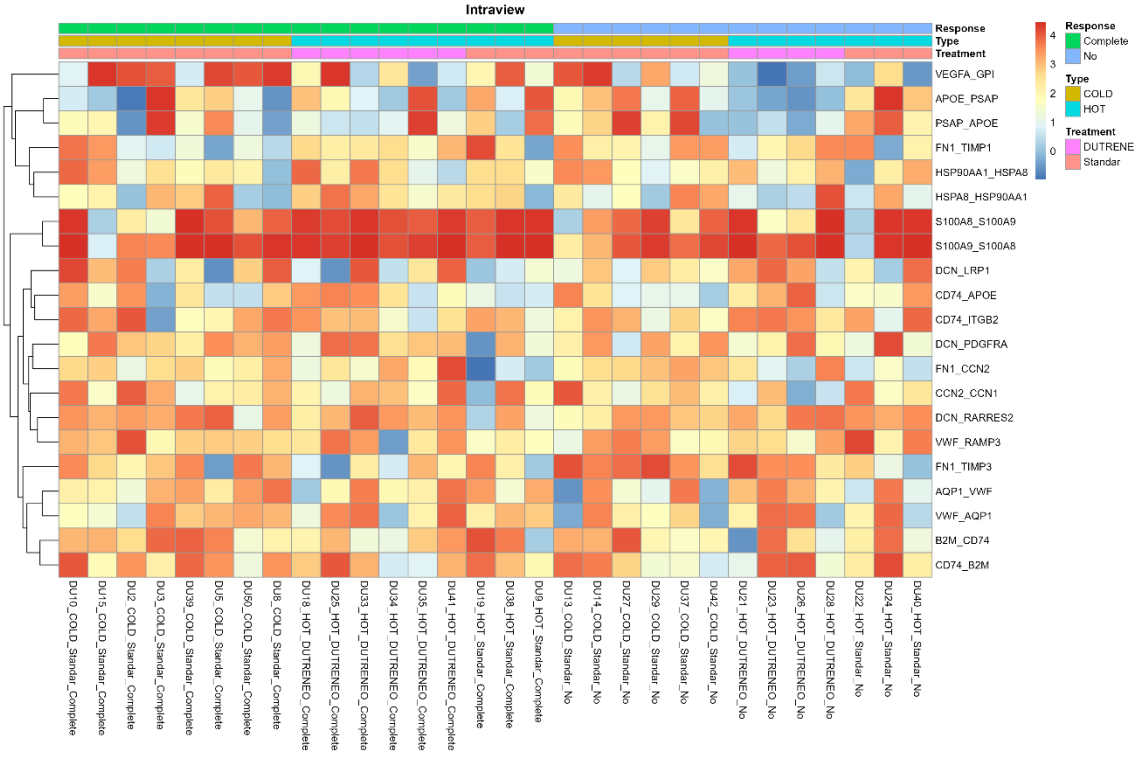

Figure 10: Heatmap showing hierarchical clustering of gene expression profiles across samples with different treatments and conditions in the Intraview. Columns represent samples, and rows represent gene pairs. Color gradients indicate expression levels (red = high, blue = low). Top colour bars show response type (green = Complete, blue = No), tumor type (orange = COLD, cyan = HOT), and treatment (magenta = DUTIRENEO, red = Standard). Dendrograms illustrate clustering of genes and samples based on expression similarities.

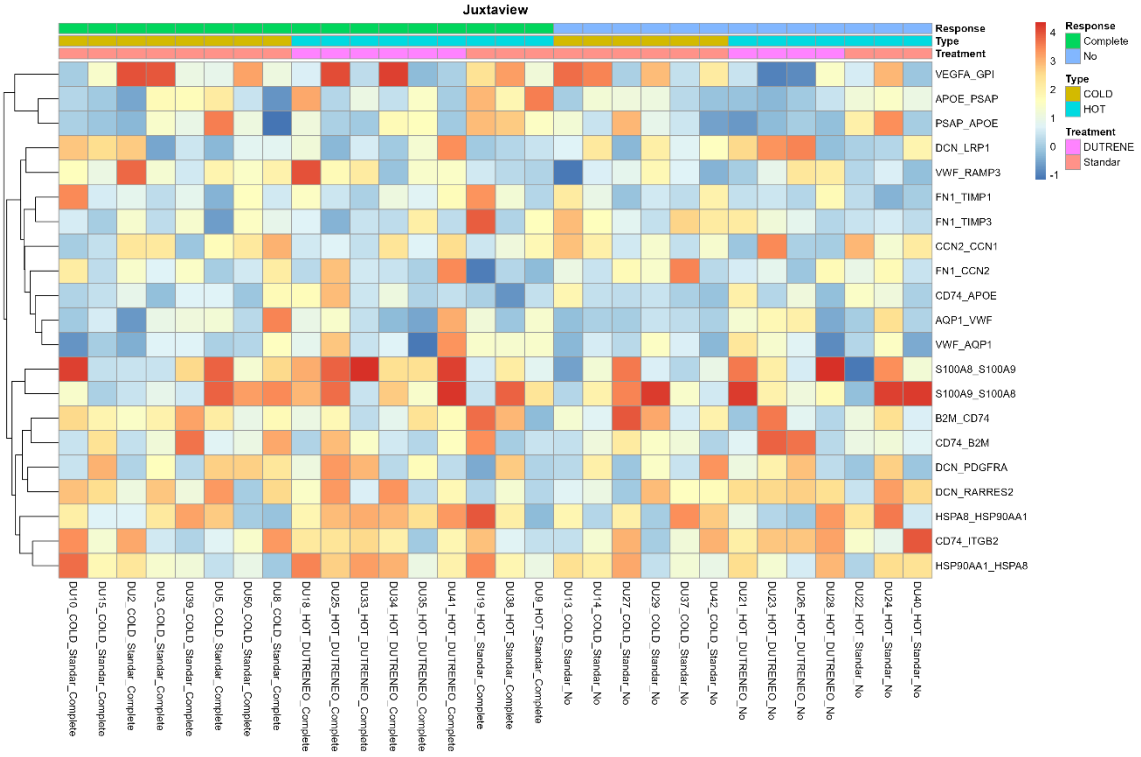

Figure 11: Heatmap showing hierarchical clustering of gene expression profiles across samples with different treatments and conditions in the Juxtaview with radius 2 spots. Columns represent samples, and rows represent gene pairs. Color gradients indicate expression levels (red = high, blue = low). Top colour bars show response type (green = Complete, blue = No), tumor type (orange = COLD, cyan = HOT), and treatment (magenta = DUTIRENEO, red = Standard). Dendrograms illustrate clustering of genes and samples based on expression similarities.

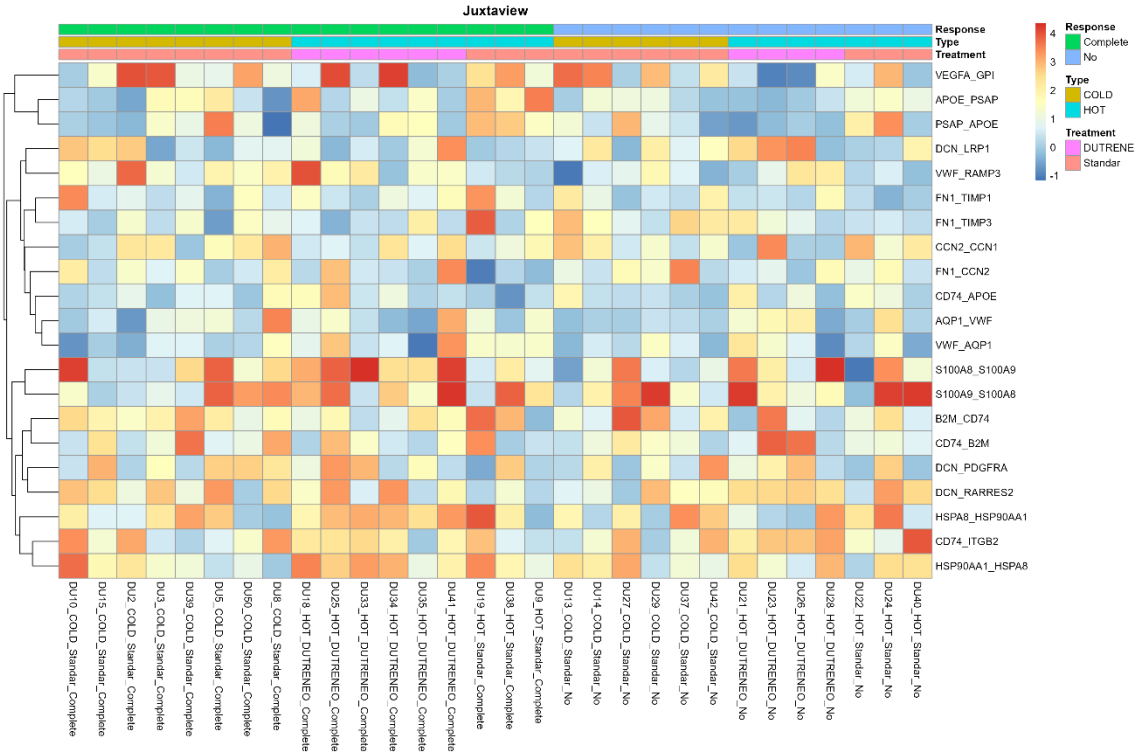

Figure 12: Heatmap showing hierarchical clustering of gene expression profiles across samples with different treatments and conditions in the Juxtaview with radius 3 spots. Columns represent samples, and rows represent gene pairs. Color gradients indicate expression levels (red = high, blue = low). Top color bars show response type (green = Complete, purple = No), temperature condition (cyan = COLD, orange = HOT), and treatment (dark blue = DUTIRENEC, yellow = Standard). Dendrograms illustrate clustering of genes and samples based on expression similarities.

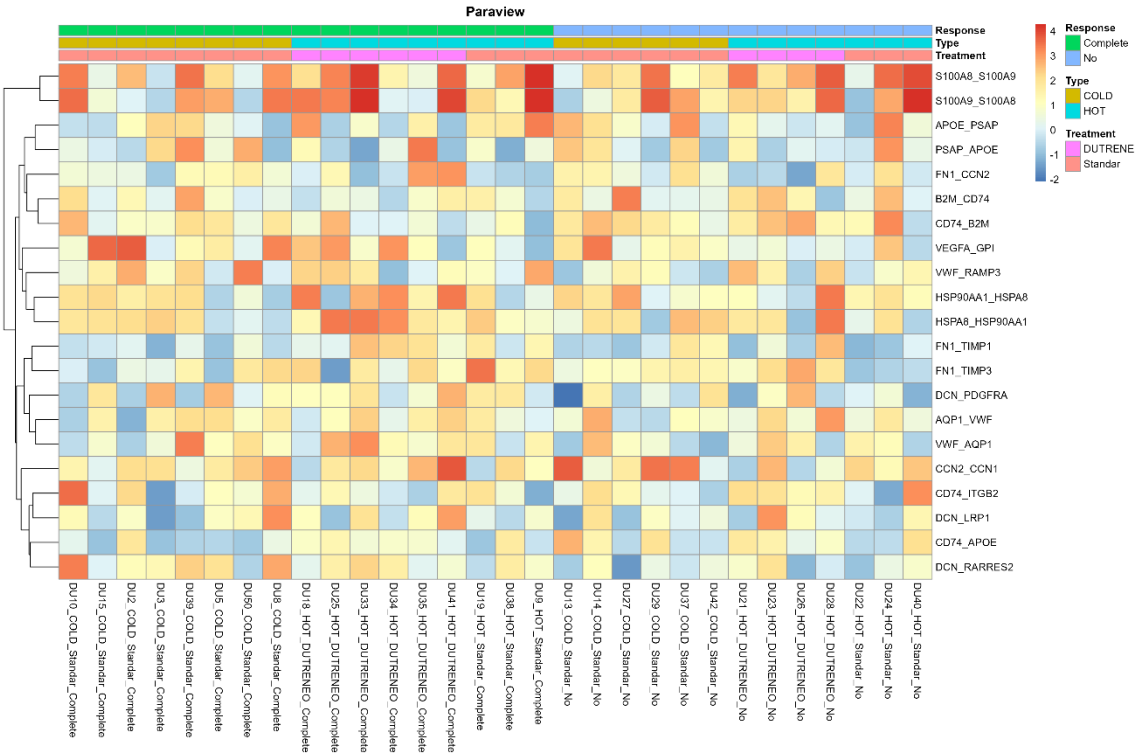

Figure 13: Heatmap showing hierarchical clustering of gene expression profiles across samples with different treatments and conditions in the Paraview. Columns represent samples, and rows represent gene pairs. Color gradients indicate expression levels (red = high, blue = low). Top color bars show response type (green = Complete, purple = No), temperature condition (cyan = COLD, orange = HOT), and treatment (dark blue = DUTIRENEC, yellow = Standard). Dendrograms illustrate clustering of genes and samples based on expression similarities.

8.3 Correlation and Linear Model Results

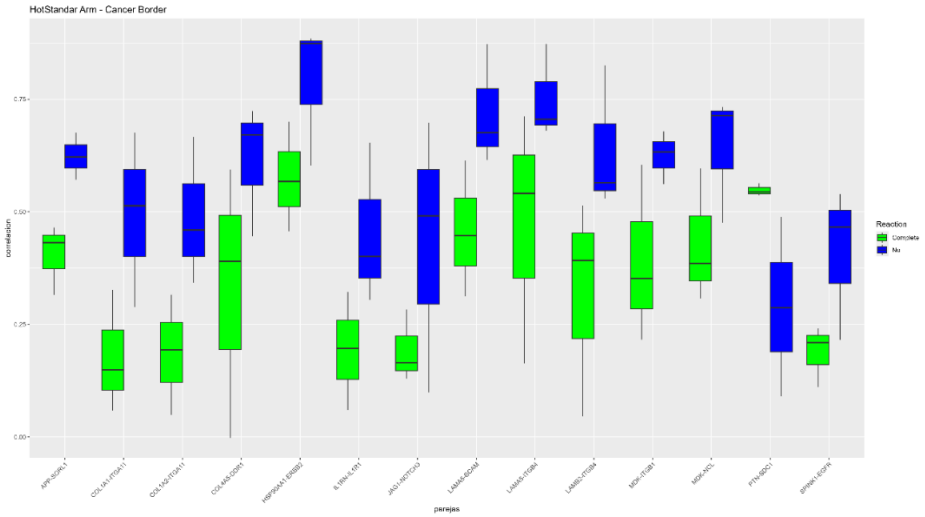

Figure 14: Box plot illustrating the correlation of ligand-receptor pairs at the cancer border in the Hot-Standard Arm. The x-axis represents different ligand-receptor pairs, and the y-axis shows the correlation values. The box plots compare the correlation between two groups: Complete Responders (green) and non-responders (blue).

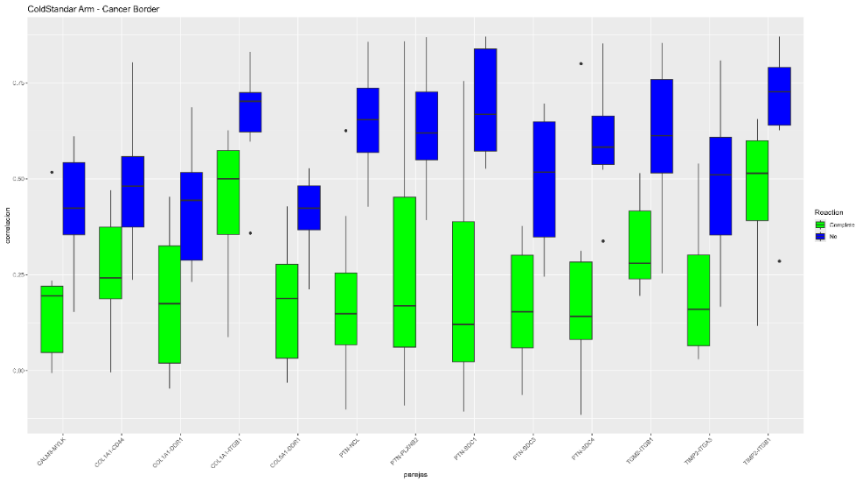

Figure 15: Box plot illustrating the correlation of ligand-receptor pairs at the cancer border in the Cold-Standard Arm. The x-axis represents different ligand-receptor pairs, and the y-axis shows the correlation values. The box plots compare the correlation between two groups: Complete Responders (green) and non-responders (blue).

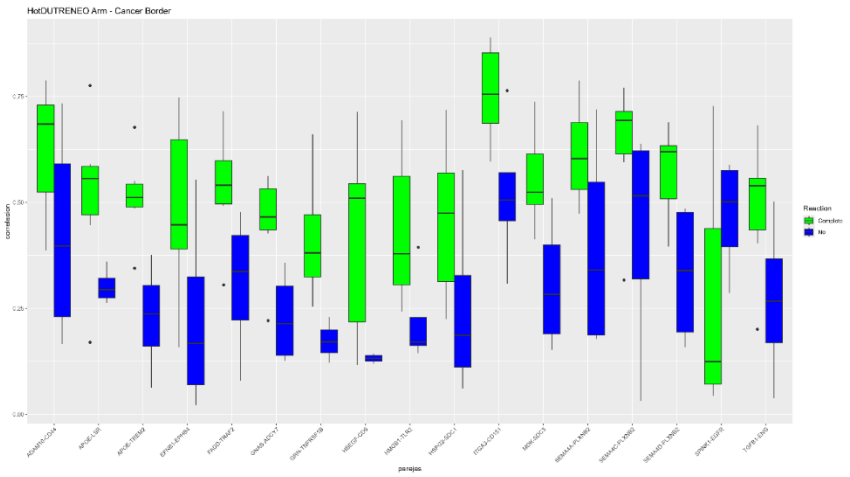

Figure 16: Box plot illustrating the correlation of ligand-receptor pairs at the cancer border in the Hot-DUTRENEO Arm.

The x-axis represents different ligand-receptor pairs, and the y-axis shows the correlation values. The box plots compare the correlation between two groups: Complete Responders (green) and non-responders (blue).

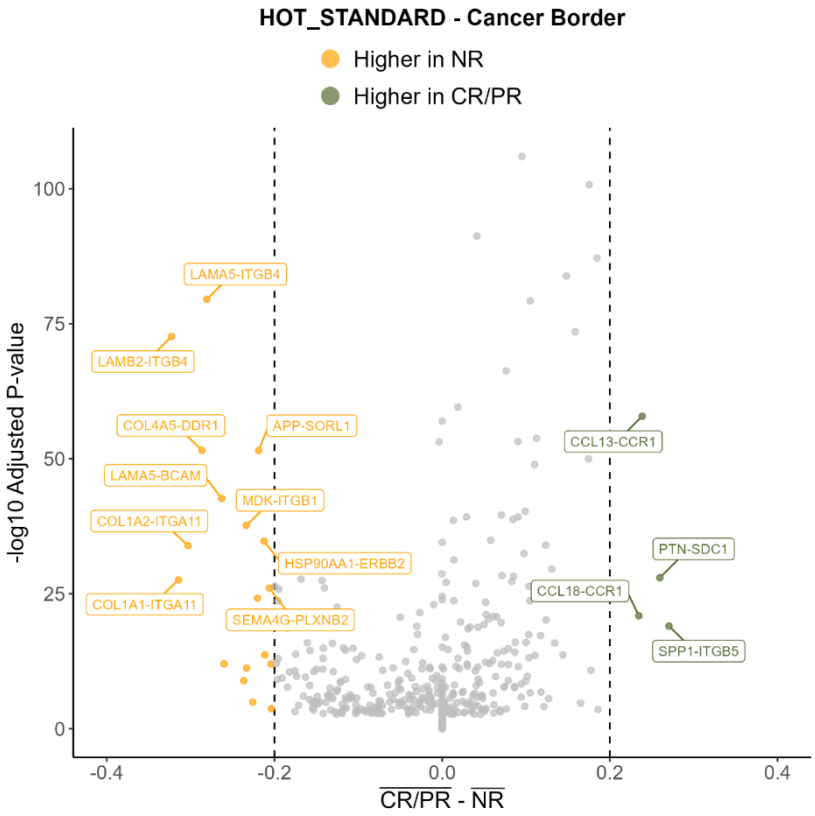

Figure 17: Volcano plot depicting differential gene expression at the cancer border in the Hot Standard arm. The x-axis represents the difference in mean co-expresion between Complete/Partial Responders (CR/PR) and Non-Responders (NR), while the y-axis shows the -log10 adjusted p-value, indicating the significance of the differential expression. Genes significantly higher in NR are highlighted in orange, and those higher in CR/PR are highlighted in green. Dashed vertical lines represent thresholds for difference in mean expression.

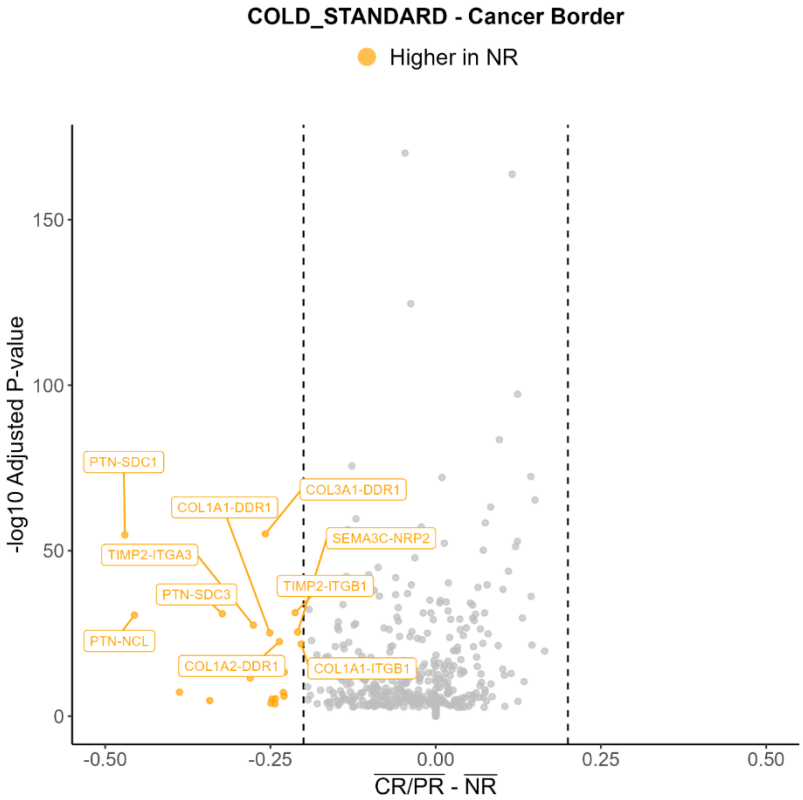

Figure 18: Volcano plot depicting differential gene expression at the cancer border in the Cold Standard arm. The x-axis represents the difference in mean co-expression between Complete/Partial Responders (CR/PR) and Non-Responders (NR), while the y-axis shows the -log10 adjusted p-value, indicating the significance of the differential expression. Genes significantly higher in NR are highlighted in orange, and those higher in CR/PR are highlighted in green. Dashed vertical lines represent thresholds for difference in mean expression.

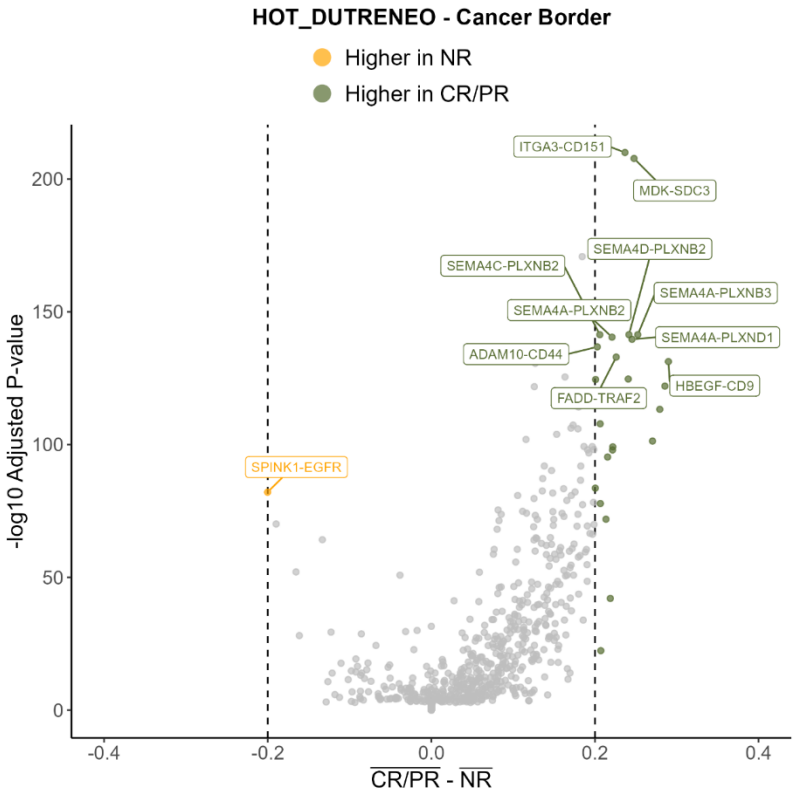

Figure 19: Volcano plot depicting differential gene expression at the cancer border in the Hot DUTRENEO arm. The x-axis represents the difference in mean co-expression between Complete/Partial Responders (CR/PR) and Non-

Responders (NR), while the y-axis shows the -log10 adjusted p-value, indicating the significance of the differential expression. Genes significantly higher in NR are highlighted in orange, and those higher in CR/PR are highlighted in green. Dashed vertical lines represent thresholds for difference in mean expression.

9. Discussion

In our exploration of MISTy and LIANA, which were expected to provide sophisticated insights, the outcomes fell short of our expectations [Figures 9-13]. While LIANA [Figure 9] revealed some distinctions between the "complete" and "no" respondent groups, a deeper examination unveiled an interaction specificity that did not surpass 0.05. This lack of confidence in the results raises concerns about their reliability. A similar issue arose with the MISTy results [Figures 10-13], where we observed a lack of discernible patterns or differentiation among the groups. While certain trends seemed to emerge, they were consistent across all samples rather than across conditions.

Overall, the results from both MISTy and LIANA did not yield any significant findings. We postulate that these disappointing results stem from the small sample size for each specific condition, which lacks sufficient statistical power for the tools to detect meaningful associations. This theory gains support from the observation that even when examining the correlation between gene pairs across conditions, no significance was found despite clear expression differences [Figures 14-16]. The sample size simply does not provide enough data, and while some pairs showed p-values smaller than 0.5, they lost significance upon correction. We have to keep in mind that those significance were made with samples sizes of 20 (14 Complete/Partial, 6 No) for Cold-Standard Arm, 11 (6 Complete/Partial, 5 No) for Hot-DUTRENEO Arm, and 9 (6 Complete/Partial, 3 No) for Hot-Standard [Table 1&2].

To address this limitation, we turned to a linear model approach, treating each point as a "sample." this alternative approach allowed us to have as much samples as spots in the Arm samples, significantly enhancing our statistical power and identified gene pairs worthy of further investigation. Consequently, we now have a selection of genes to explore with confidence in their functional relevance.

Notably, the Standard arms displayed few to no upregulated gene pairs in the “CR/PR” responders [Figures 17 & 18], while the DUTRENEO arm exhibited almost no genes in the “NR” responders [Figure 19]. This observation aligns well with the roles these genes play in cancer development and underscores the importance of considering treatment context in our analysis.

In summary, while our initial methodologies failed to produce significant results, the adoption of a linear model approach has provided us with a clearer direction for our investigation, shedding light on key genes and their implications in treatment response within distinct arms of our study.

10. Future Work

This thesis was pursuing some clear objectives, aiming to provide new insights into the molecular mechanisms underpinning tumour-TME interactions in MIBC, using spatial transcriptomics together with advanced computational analysis, specialised methodologies and tools. We believe that it is critical to comprehend the intricate interactions that occur between tumor cells and the tumor microenvironment (TME) in muscle-invasive urothelial bladder cancer (MIBC), as well as in any other cancer type, in order to advance research and enhance patient outcomes. We will continue with the investigations going further into the relation of all the ligand-receptor pairs found, looking at pathways that they are related, how do they interact between them, and even adding to the stage some known transcription factors or cell types. Given that our results ended up showing some unique signatures for each of the Arms with really strong statistical significance, we have something meaningful to keep researching in this direction. We are aware that we are barely scratching the surface of this colossal problem but we will continue to put our efforts in order to make a difference in the field.

11. Acknowledgements

All of this work would not have been possible without all of the people who supported and contributed to this project. First and foremost, I want to thank Eduard Porta for his guidance and mentorship throughout the entire project. His leadership and continuous support have been priceless. I am also immensely grateful to Daniela Grases and Elena Pérez for their expertise and dedication in managing the use of Visium technology and managing all of the samples. A very special thanks goes to Mustafa Sibai, with whom I am currently working to keep analysing all of the data that we have at our disposal. His work on the initial data preprocessing, cell type identification, and data integration. Lastly, I wanted to show my sincere appreciation to Paco Real for providing the raw data and serving as the connection with the clinical aspect.

12. Code Availability

All the code used in this thesis can be accessed in this GitHub repository: https://github.com/olmopolmo/TFM_source_code.git

13. References

1. Valderrama BP, González-del-Alba A, Morales-Barrera R, et al. SEOM-SOGUG clinical guideline for localized muscle invasive and advanced bladder cancer (2021). *Clin Transl Oncol*. 2022;24(4):613-624. doi:10.1007/s12094-022-02815-w

2. Aragon-Ching JB, Werntz RP, Zietman AL, Steinberg GD. Multidisciplinary Management of Muscle-Invasive Bladder Cancer: Current Challenges and Future Directions. *Am Soc Clin Oncol Educ Book*. 2018;(38):307-318. doi:10.1200/EDBK_201227

3. Malkowicz SB, van Poppel H, Mickisch G, et al. Muscle-invasive urothelial carcinoma of the bladder. *Urology*. 2007;69(1 Suppl):3-16. doi:10.1016/j.urology.2006.10.040

4. Muscle invasive bladder cancer staging. Accessed June 3, 2024. https://www.cancerresearchuk.org/about-cancer/bladder-cancer/types-stages-grades/muscle-invasive

5. Esteban-Villarrubia J, Torres-Jiménez J, Bueno-Bravo C, García-Mondaray R, Subiela JD, Gajate P. Current and Future Landscape of Perioperative Treatment for Muscle-Invasive Bladder Cancer. *Cancers*. 2023;15(3):566. doi:10.3390/cancers15030566

6. Patil G, Basu A. Emerging perioperative therapeutic approaches in muscle invasive bladder cancer. *Ther Adv Urol*. 2022;14:17562872221134389. doi:10.1177/17562872221134389

7. Ongoing Research Pushes for Neoadjuvant Alternatives to Chemotherapy in MIBC. OncLive. Published April 1, 2024. Accessed June 3, 2024. https://www.onclive.com/view/ongoing-research-pushes-for-neoadjuvant-alternatives-to-chemotherapy-in-mibc

8. EAU Guidelines on MIBC - DISEASE MANAGEMENT - Uroweb. Uroweb - European Association of Urology. Accessed June 3, 2024. https://uroweb.org/guidelines/muscle-invasive-and-metastatic-bladder-cancer/chapter/disease-management

9. Nair SS, Chakravarty D, Patel V, Bhardwaj N, Tewari AK. Genitourinary cancer neoadjuvant therapies: current and future approaches. *Trends Cancer*. 2023;9(12):1041-1057. doi:10.1016/j.trecan.2023.07.011

10. DUTRENEO Trial: A randomized phase II trial of DUrvalumab and TREmelimumab versus chemotherapy as a NEOadjuvant approach to muscle-invasive urothelial bladder cancer (MIBC) patients (pts) prospectively selected by an interferon (INF)-gamma immune signature. | Journal of Clinical Oncology. Accessed June 3, 2024. https://ascopubs.org/doi/10.1200/JCO.2020.38.15_suppl.5012

11. Fahmy O, Ahmed OAA, Khairul-Asri MG, et al. Adverse Events and Tolerability of Combined Durvalumab and Tremelimumab versus Durvalumab Alone in Solid Cancers: A Systematic Review and Meta-Analysis. *Biomedicines*. 2022;10(5):1101. doi:10.3390/biomedicines10051101

12. Capdevila J, Hernando J, Teule A, et al. Durvalumab plus tremelimumab for the treatment of advanced neuroendocrine neoplasms of gastroenteropancreatic and lung origin. *Nat Commun*. 2023;14(1):2973. doi:10.1038/s41467-023-38611-5

13. Yang B, Li Y, Deng J, Yang H, Sun X. Efficacy and safety of immune checkpoint inhibitors plus recombinant human endostatin therapy as second-line treatment in advanced non-

small-cell lung cancer with negative driver gene: a pilot study. *Front Oncol*. 2023;13. doi:10.3389/fonc.2023.1210267

14. An introduction to spatial transcriptomics for biomedical research | Genome Medicine | Full Text. Accessed June 3, 2024. https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-022-01075-1

15. Mastering Biology to Advance Human Health. 10x Genomics. Accessed June 3, 2024. https://www.10xgenomics.com/

16. 10x_LIT059_ProductSheet_VisiumSpatialGeneExpression_Letter_digital.pdf. Accessed June 3, 2024. https://pages.10xgenomics.com/rs/446-PBO-704/images/10x_LIT059_ProductSheet_VisiumSpatialGeneExpression_Letter_digital.pdf

17. 10x_LIT000128_Product_Sheet_Spatial_Biology_Without_Limits_Letter_digital.pdf. Accessed June 3, 2024. https://pages.10xgenomics.com/rs/446-PBO-704/images/10x_LIT000128_Product_Sheet_Spatial_Biology_Without_Limits_Letter_digital.pdf

18. Explainable multiview framework for dissecting spatial relationships from highly multiplexed data | Genome Biology | Full Text. Accessed June 3, 2024. https://genomebiology.biomedcentral.com/articles/10.1186/s13059-022-02663-5

19. Dimitrov D, Türei D, Garrido-Rodriguez M, et al. Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. *Nat Commun*. 2022;13(1):3224. doi:10.1038/s41467-022-30755-0

20. High-plex Multiomic Analysis in FFPE at Subcellular Level by Spatial Molecular Imaging | bioRxiv. Accessed June 3, 2024. https://www.biorxiv.org/content/10.1101/2021.11.03.467020v2

21. Chen WT, Lu A, Craessaerts K, et al. Spatial Transcriptomics and In Situ Sequencing to Study Alzheimer's Disease. *Cell*. 2020;182(4):976-991.e19. doi:10.1016/j.cell.2020.06.038

22. Lewis SM, Asselin-Labat ML, Nguyen Q, et al. Spatial omics and multiplexed imaging to explore cancer biology. *Nat Methods*. 2021;18(9):997-1012. doi:10.1038/s41592-021-01203-6

23. Asp M, Bergenstråhle J, Lundeberg J. Spatially Resolved Transcriptomes-Next Generation Tools for Tissue Exploration. *BioEssays News Rev Mol Cell Dev Biol*. 2020;42(10):e1900221. doi:10.1002/bies.201900221

24. Ståhl PL, Salmén F, Vickovic S, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*. 2016;353(6294):78-82. doi:10.1126/science.aaf2403

25. Chen Z, Chen D, Song Z, Lv Y, Qi D. Mapping the tumor microenvironment in bladder cancer and exploring the prognostic genes by single-cell RNA sequencing. *Front Oncol*. 2023;12. doi:10.3389/fonc.2022.1105026

26. Lee HW, Chung W, Lee HO, et al. Single-cell RNA sequencing reveals the tumor microenvironment and facilitates strategic choices to circumvent treatment failure in a

chemorefractory bladder cancer patient. *Genome Med*. 2020;12(1):47. doi:10.1186/s13073-020-00741-6

27. Tran M, Farkas A, Beaumont K, et al. 82 Single-cell RNA sequencing and CITE-Seq analysis of bladder cancer patient urine with matched tumor and peripheral blood suggests urine as a window into the tumor immune microenvironment. *J Immunother Cancer*. 2021;9(Suppl 2). doi:10.1136/jitc-2021-SITC2021.082

28. Larsson L, Franzén L, Ståhl PL, Lundeberg J. Semla: a versatile toolkit for spatially resolved transcriptomics analysis and visualization. *Bioinformatics*. 2023;39(10):btad626. doi:10.1093/bioinformatics/btad626

29. ConsensusPathDB. Accessed June 5, 2024. http://cpdb.molgen.mpg.de/

30. CellCall: integrating paired ligand–receptor and transcription factor activities for cell–cell communication | Nucleic Acids Research | Oxford Academic. Accessed June 5, 2024. https://academic.oup.com/nar/article/49/15/8520/6332819

31. Jin S, Guerrero-Juarez CF, Zhang L, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun*. 2021;12(1):1088. doi:10.1038/s41467-021-21246-9

32. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc*. 2020;15(4):1484-1506. doi:10.1038/s41596-020-0292-x

33. Ramilowski JA, Goldberg T, Harshbarger J, et al. A draft network of ligand-receptor-mediated multicellular signalling in human. *Nat Commun*. 2015;6:7866. doi:10.1038/ncomms8866

34. Baccin C, Al-Sabah J, Velten L, et al. Combined single-cell and spatial transcriptomics reveal the molecular, cellular and spatial bone marrow niche organization. *Nat Cell Biol*. 2020;22(1):38-48. doi:10.1038/s41556-019-0439-6

35. Cabello-Aguilar S, Alame M, Kon-Sun-Tack F, Fau C, Lacroix M, Colinge J. SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics. *Nucleic Acids Res*. 2020;48(10):e55. doi:10.1093/nar/gkaa183

36. Kirouac DC, Ito C, Csaszar E, et al. Dynamic interaction networks in a hierarchically organized tissue. *Mol Syst Biol*. 2010;6:417. doi:10.1038/msb.2010.71

37. Dissection of intercellular communication using the transcriptome-based framework ICELLNET | Nature Communications. Accessed June 5, 2024. https://www.nature.com/articles/s41467-021-21244-x

38. Wang Y, Wang R, Zhang S, et al. iTALK: an R Package to Characterize and Illustrate Intercellular Communication. Published online January 4, 2019:507871. doi:10.1101/507871

39. EMBRACE. Accessed June 5, 2024. https://www.embracestudy.dk/Public/Default.aspx?main=1&sub=0&embrace=embrace

40. Ben-Shlomo I, Yu Hsu S, Rauch R, Kowalski HW, Hsueh AJW. Signaling receptome: a genomic and evolutionary perspective of plasma membrane receptors involved in signal

transduction. *Sci STKE Signal Transduct Knowl Environ*. 2003;2003(187):RE9. doi:10.1126/stke.2003.187.re9

41. Raredon MSB, Yang J, Garritano J, et al. Computation and visualization of cell-cell signaling topologies in single-cell systems data using Connectome. *Sci Rep*. 2022;12(1):4187. doi:10.1038/s41598-022-07959-x

42. Shao X, Liao J, Li C, Lu X, Cheng J, Fan X. CellTalkDB: a manually curated database of ligand-receptor interactions in humans and mice. *Brief Bioinform*. 2021;22(4):bbaa269. doi:10.1093/bib/bbaa269

43. Lab SR. OmniPath :: Intra- & intercellular signaling knowledge. Accessed June 5, 2024. https://omnipathdb.org/