Quantifying response and dynamics to pre-operative treatments in urothelial cancer
Mapping the tumor microenvironment for better response predictions

Gil Jimenez, A.

DOI
10.4233/uuid:83ce79ab-6074-4990-833f-392a79851a4b

Publication date
2024

Document Version
Final published version

Citation (APA)
Gil Jimenez, A. (2024). *Quantifying response and dynamics to pre-operative treatments in urothelial cancer: Mapping the tumor microenvironment for better response predictions*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:83ce79ab-6074-4990-833f-392a79851a4b

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Quantifying response and dynamics to pre-operative treatments in urothelial cancer

Mapping the Tumor Microenvironment for better response predictions

Alberto Gil Jiménez

QUANTIFYING RESPONSE AND DYNAMICS TO PRE-OPERATIVE TREATMENTS IN UROTHELIAL CANCER

Mapping the Tumor Microenvironment for better response predictions

Alberto GIL JIMÉNEZ

This research was funded by the KWF Kankerbestrijding, Oncode Institute, and Bristol Myers Squibb.

Keywords: Urothelial cancer, Immunotherapy, Immune checkpoint inhibitors, Chemotherapy

Printed by: Gildeprint

Front: The Ballroom scene is an underground LGBTIQ+ subculture originating from Afro-American and Latino communities characterized by "balls" in which contestants "walk" (performing, dancing, walking, lip-syncing, modeling, etc) for categories to win trophies (more information can be found at https://www.vanvoguejam.com/ballroom-history or using the key term "Ball culture"). Old Way Vogue is a category within the Ballroom scene where dancers emulate the glamour and poses inspired by fashion magazines, martial arts and Egyptian hieroglyphics and is characterized by the formation of lines, precision and symmetry in the execution of the movement.

The front depicts a voguer emulating an Old Way pose emulating the shape, symmetry and lines of an antibody, which is an important component of this thesis. First, immunotherapies in the form of monoclonal antibodies are studied in chapters 3, 4, and 5. Second, a technique used to profile the tumor microenvironment in chapters 3, 4, and 5 known as multiplex immunofluorescence is based on fluorescent antibodies.

The drawn person is inspired by Willi Ninja, a famous voguer known for his appearance in the documentary "Paris is Burning", for exporting ballroom culture abroad, and for being an inspiration for Madonna's song and videoclip "Vogue".

Front by Oct Streitenberger.

The written content of this thesis has been improved using Grammarly and ChatGPT.

ISBN: 978-94-6496-264-2

An electronic version of this dissertation is available at https://repository.tudelft.nl/

QUANTIFYING RESPONSE AND DYNAMICS TO PRE-OPERATIVE TREATMENTS IN UROTHELIAL CANCER

Mapping the Tumor Microenvironment for better response predictions

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, Prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Tuesday 19 November 2024 at 12:30 o'clock

by

Alberto GIL JIMÉNEZ

Master of Science in Bioinformatics,
Vrije Universiteit Amsterdam and University of Amsterdam, Amsterdam, the Netherlands,
Born in Hospitalet de Llobregat, Spain

This dissertation has been approved by the promotor.

Composition of the doctoral committee:

Rector Magnificus,	Chairperson
Prof. dr. L. F. A. Wessels	Delft University of Technology and Netherlands Cancer Institute, promotor
Dr. M. S. van der Heijden	Netherlands Cancer Institute, copromotor
Dr. D. J. Vis	Netherlands Cancer Institute, copromotor

Independent members:

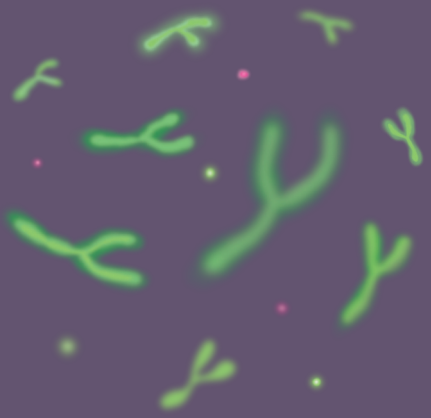
Prof. dr. M. J. T. Reinders	Delft University of Technology
Prof. dr. E. E. Voest	Netherlands Cancer Institute and University of Utrecht
Prof. dr. J. B. A. G. Haanen	Netherlands Cancer Institute and University of Leiden
Dr. M. Kok	Netherlands Cancer Institute

Reserved member:

Prof. dr. B. Rieger	Delft University of Technology

De acacias para arriba

TABLE OF CONTENTS



Glossary

1-NN	First nearest-neighbor
AI	Artificial Intelligence
AUC	Area under the curve
APC	Antigen-Presenting Cell
CDF	Cumulative distribution function
CD8	Cluster of differentiation 8
CIS	Carcinoma in situ
CNA	Copy Number Alteration
CTLA-4	Cytotoxic T-Lymphocyte-Associated Protein 4
DC	Dendritic cell
FDR	False Discovery Rate
HLA	Human Leukocyte Antigen
ICI	Immune Checkpoint Inhibitor
IF	Immunofluorescence
MHC	Major Histocompatibility Complex
MIBC	Muscle-Invasive Bladder Cancer
mIF	Multiplexed Immunofluorescence
MSI	Microsatellite Instability
mUC	Metastatic Urothelial Cancer
NAC	Neoadjuvant Chemotherapy
NGS	Next-Generation Sequencing
NK	Natural Killer
NMIBC	Non-Muscle-Invasive Bladder Cancer
OS	Overall Survival
PBMC	Peripheral Blood Mononuclear Cell
PD-1	Programmed Cell Death Protein 1
PD-L1	Programmed Death-Ligand 1
PDF	Probability Density Function
pCR	Pathological Complete Response
RC	Radical Cystectomy
RFS	Recurrence-Free Survival
TCGA	The Cancer Genome Atlas
TCR	T-Cell Receptor
TGF-beta	Transforming Growth Factor-beta
TMB	Tumor Mutational Burden
TME	Tumor Microenvironment
TLS	Tertiary Lymphoid Structure
TUR	Transurethral Resection
UC	Urothelial Cancer

1 INTRODUCTION

Alberto Gil-Jimenez

Cancer can be understood as a group of diseases characterized by an impairment of cellular regulatory processes at the genomic, transcriptomic, and protein levels, which results in aberrant cell proliferative behavior[1]. Understanding such altered biological processes is essential to develop anti-cancer treatments to control and stop the disease.

Essentially, each tumor is unique regarding its associated genomic aberrations and protein profiles. Nonetheless, tumors share characteristics and patterns, which, when linked to clinical outcomes, allow us to optimize treatments and understand treatment responses. This thesis aims to explore and identify novel shared patterns of urothelial cancer tumors collected from patient samples, link them to clinical outcomes (e.g., response to preoperative treatments), quantify shared changes in tumors by the effect of treatments, and find novel ways of modeling tumor characteristics, with the ultimate goal of expanding our current knowledge on urothelial cancer tumor biology that could serve to improve patient care.

1.1 UROTHELIAL CANCER AND THE CURRENT STATUS OF TREATMENT

Urothelial cancer (UC) is a disease characterized by cancer cells in the urinary tract. UC in the bladder is categorized as non-muscle-invasive urothelial cancer (NMIBC) when cancer cells are present at the inner layer of the bladder, as muscle-invasive bladder cancer (MIBC) once cancer cells invade the muscle layer of the bladder, and as metastatic UC (mUC) once the disease spreads to distant organs. For the scope of this thesis, the focus will be solely on non-metastatic disease, which will be referred to as UC for simplification purposes.

1.1.1 Prognosis and risk factors

The prognosis of UC is generally poor, even when the disease has not yet spread to distant organs, with a 5-year overall survival (OS) of around ~70% for localized MIBC, but only ~40% if cancer cells metastasized to the local lymph nodes[2]. The high recurrence rates observed in UC after standard-of-care treatments explain a great part of such poor survival rates. Together with the moderate decrease in mortality in UC patients over the past 10 years attributed to the lack of approval of novel UC treatments[2], there is an urgent need to improve the UC treatment landscape and to tailor treatments better to cure UC.

1.1.2 Hallmarks of muscle-invasive urothelial cancer

UC tumors are characterized by driver mutations altering the regulation of the cell cycle, chromatin, and kinase signaling pathways, as well as genomic instability resulting in a high mutational load[3,4]. Due to the disease heterogeneity at the molecular level, multiple expression

subtypes have been proposed to group tumors with shared expression profiles that associate with distinct genomic and clinical features[5].

1.1.3 Standard of care: radical cystectomy and neoadjuvant chemotherapy

The standard of care of UC is radical cystectomy (RC), a surgical procedure that removes the bladder and nearby lymph nodes. However, upon RC, a 5-year OS of 68% is usually achieved with variability depending on the baseline disease stage[6], suggesting that the treatment is imperfect to control the disease.

Before surgery, patients can be treated in a neoadjuvant setting. To date, neoadjuvant chemotherapy (NAC) is considered the standard-of-care[7] for treating UC, mainly using cisplatin-based combination chemotherapy. Despite decreasing the tumor burden at the surgical timepoint, increasing the likelihood of disease control, and reducing the likelihood of micro-metastasis, the NAC combined with surgery provides only a moderate increase in the 5-year OS (5%-10%) relative to surgery alone[8]. Because grade 3 or higher chemotherapy-related toxicities are observed in 10% of patients[9], at the cost of only a 5-10% 5-year OS increase relative to surgery alone[8], there is an urgent need to develop novel neoadjuvant treatment strategies for UC.

In the past two decades, the treatment landscape for non-metastatic UC has experienced no change since the approval of neoadjuvant chemotherapy. However, ongoing clinical trials involving immune checkpoint inhibitors (ICIs) hold promise as potential drug treatments, as they could potentially change the UC treatment landscape in the coming years[10], holding promise to increase patient survival.

1.2 TUMOR IMMUNOLOGY AND IMMUNE CHECKPOINT INHIBITORS

1.2.1 Tumor microenvironment

Tumors reside in a region that is called the tumor microenvironment (TME). The TME comprises immune cells, normal epithelial cells, fibroblasts, the extracellular matrix, and blood vessels[11]. These surrounding cells allow for interactions that influence the tumor behavior; for example facilitating tumor growth, progression, and metastatic dissemination, and modulating (anti-)tumorigenic functions mediated by the immune system[12].

Immune cells eliminate cancer through the cancer-immunity cycle[13]. Within this multi-step dynamic process, immune cells recognize cancer cells' molecular fingerprints in the form of antigens. At the immune priming phase, tumor antigens are released in the TME upon

tumor cell apoptosis. Such antigens are captured and processed by antigen-presenting cells, e.g., dendritic cells (DCs), which then traffic to the lymph nodes. Antigen-presenting cells present tumor antigens to T-cells via the major histocompatibility complex (MHC), inducing T-cell activation and clonal expansion. This step initiates the effector phase, in which T-cells recognize specific cancer antigens through their T-cell receptor (TCR). Then, cytotoxic T-cells infiltrate the tumor and recognize antigens presented by cancer cells via the MHC, eliminating cancer cells. New tumor antigens are released upon this last step, marking a new immunity cycle initiation.

Impairment of the cancer immunity cycle facilitates tumor progression[14]. Immune checkpoints are surface proteins expressed on cancer and immune cells that downregulate the immune system. An immune checkpoint expressed on cancer cells is the programmed death-ligand 1 (PD-L1), which allows for an inhibition of the cancer-immunity cycle effector phase upon interaction with its receptor programmed cell death protein 1 (PD-1), expressed among others by T-cells (**Figure 1**). Another checkpoint is cytotoxic T-lymphocyte-associated protein 4 (CTLA-4), which downregulates the immune priming phase upon interaction with its ligands CD80 (B7-1) or CD86 (B7-2) mostly expressed on APCs (**Figure 1A**).

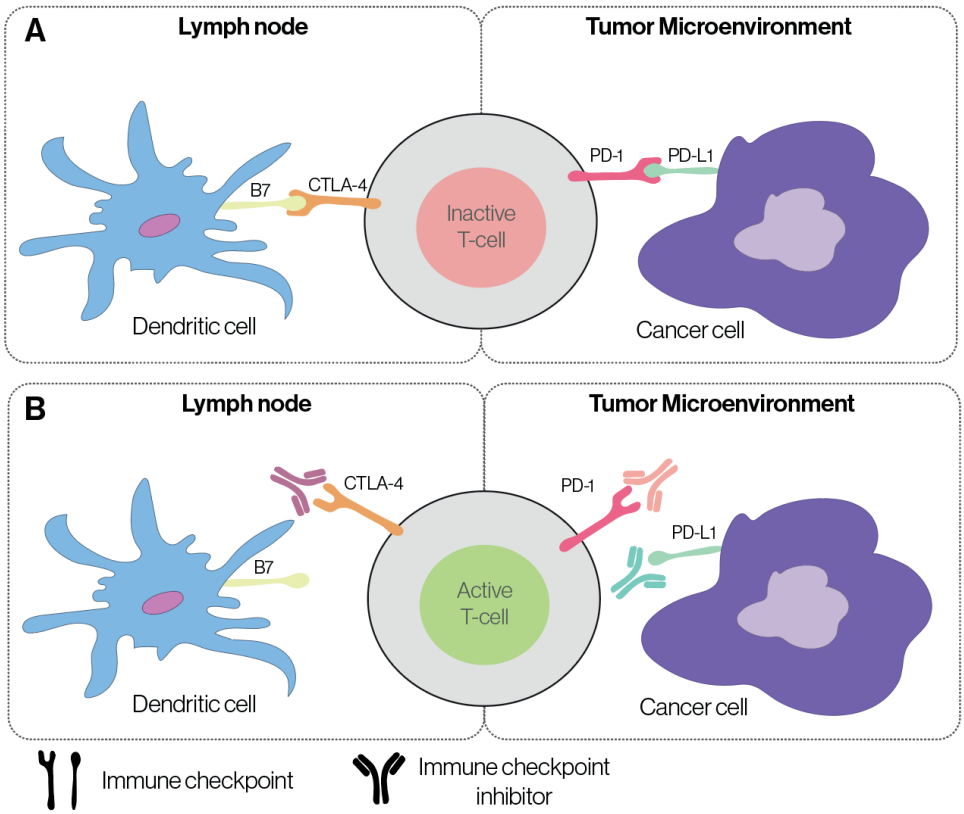


Figure 1. A. Schematic representation of a T-cell expressing the immune checkpoint CTLA-4 interacting with a Dendritic cell expressing the immune checkpoint B7 (left, occurring at the tumor-draining lymph nodes) and a T-cell expressing the immune checkpoint PD-L1 interacting with a cancer cell expressing the immune checkpoint PD-1 (right, occurring at the tumor microenvironment). Both interactions downregulate anti-cancer T-cell killing. **B**. Schematic representation of immune checkpoint inhibitors blocking the immune checkpoints CTLA-4, PD-L1, and PD-1.

1.2.2 Immune checkpoint inhibitors

Understanding regulatory cancer immune response mechanisms has been pivotal to developing novel anti-cancer therapeutic strategies in immunotherapies that modulate the immune system to promote anti-tumor T-cell activity. Immune checkpoint inhibitors (ICIs) are a group of antibody-based immunotherapies that aim to overcome the inhibitory signals posed by immune checkpoints to enhance anti-tumor T-cell activity. Like chemotherapy regimens, ICIs are a systemic therapy administered via intravenous injection.

In the last decade, ICI treatments targeting various immune checkpoints have been approved to treat multiple cancer types in the metastatic setting[15]. ICI treatments are currently being

investigated for suitability in the neoadjuvant (i.e., pre-operative) setting[16,17]. Some of the most widely used ICIs are monoclonal antibodies that block PD-1 (e.g., nivolumab), PD-L1 (e.g., pembrolizumab), CTLA-4 (e.g., ipilimumab), or a combination of them (**Figure 1B**). In UC treatment, ICIs have been approved for the metastatic setting[18]. To date, ongoing clinical research is exploring the efficacy and toxicity of ICIs in the neoadjuvant setting with promising results thus far regarding the long-term low rates of recurrence[19].

1.3 LEARNING FROM TUMOR MOLECULAR PROFILES

Within the context of UC, the primary sources of human tumor tissue material are from two surgical procedures: transurethral resection (TUR) and radical cystectomy (RC). Transurethral resection (TUR) is a semi-invasive procedure that allows sampling of a tumor primarily for diagnostic purposes, and yields sufficient tumor material for extensive profiling to determine associations with response to neoadjuvant treatments. Radical cystectomy (RC) material is typically used to investigate treatment dynamics and resistance mechanisms to neoadjuvant treatments[20]. Additionally, molecular profiling can be performed on samples from the primary tumor, locally metastasized lymph nodes, or at a systemic level through liquid biopsies obtained, for instance, from peripheral blood mononuclear cells (PBMCs)[21].

Interrogating the molecular characteristics of tumors is a powerful approach to enhance our understanding of tumor biology[22]. Together with the availability of human material before or on treatment, such data presents a unique opportunity to establish links between tumor characteristics and clinical outcomes, posing a valuable source of information to infer tumor and immunological characteristics.

Sequencing-based technologies have revolutionized the possibilities of molecular tumor profiling in a cost-effective and time-efficient manner due to the rise of next-generation sequencing (NGS) technologies. NGS allows for the profiling of genomic alterations (DNA sequencing), expression of RNA (RNA sequencing) as well as other molecular profiles such as methylation patterns (ChIP sequencing[23]).

Orthogonally, proteomics-based approaches allow the characterization of the tumor proteome and the TME[24]. For instance, multiplex immunofluorescence technologies simultaneously stain for multiple fluorescent antibodies on a single tissue slide[25]. In contrast to standard stainings with a single antibody, multiplex immunofluorescence enables staining for multiple markers of cancer and immune cells and resolves the TME spatially.

In general, bulk profiling techniques profile the entire tumor sample in contrast to single-cell profiling which sequences individual cells. A relatively new dimension to tumor profiling is the spatial resolution, which allows for a characterization of the spatial features of tumors and their surrounding TME[26]. Lastly, the panel of genes or proteins being profiled (e.g., whole-genome, whole-exome, targeted panel) determines the breadth of molecular characteristics that can be profiled from a tumor. Altogether, the resolution chosen depends on the research question's hypothesis and tissue and resource availability.

Signal processing steps are taken to obtain usable data amenable to hypothesis testing. Typically, DNA sequencing data requires the alignment of the DNA reads to a reference genome, calling variants (e.g., tumor mutations), and further filtering the data. These choices can impact, for instance, the list of mutations to be used in association studies. Similarly, for RNA sequencing data, an alignment, transcript quantification, and differential expression modeling framework must be carefully performed before downstream analysis. For multiplex immunofluorescence data, we first need to segment the signal on a per-cell basis, classify them (e.g., cancer cells, B-cells, T-cells), and segment the tissue (e.g., tumor vs. stroma). Finally, the data are usually summarized as normalized counts. Therefore, this implies that each data type must be accompanied by best practices established by the scientific community to ensure the validity of the measurements.

The rapid advancement and development of novel technologies to profile tumors implies that 1) best practices to analyze the data are not harmonized within the scientific community, 2) there is a need to develop novel algorithms to quantify the data concisely and reliably, and 3) benchmarking approaches are required to assess the validity of the analytical procedures. Once these challenges are overcome, novel opportunities arise to assess unexplored axes of tumor biology. For instance, the rapid development of technologies with a spatial resolution (e.g., spatial transcriptomics, spatial proteomics), accompanied by novel computational methodologies, provides valuable sources of information into tumor heterogeneity and spatial biology. Such information could be, for instance, used to discover novel biomarkers of response to anti-cancer treatments[27–29].

1.4 PREDICTIVE MARKERS OF RESPONSE TO NEOADJUVANT TREATMENTS IN UROTHELIAL CANCER

Predictive biomarkers are molecular traits that quantify the likelihood of a tumor's response to a particular treatment (**Figure 2**). In the current era of precision medicine, research aims to identify novel molecular tumor traits that could serve as biomarkers. Such biomarkers pose opportunities to tailor the right treatment to the right tumor. Therefore, predictive biomarkers allow for overcoming the current heterogeneous response rates observed across multiple treatments, avoiding patient over- and under-treatment and preventing unnecessary toxicities[30,31].

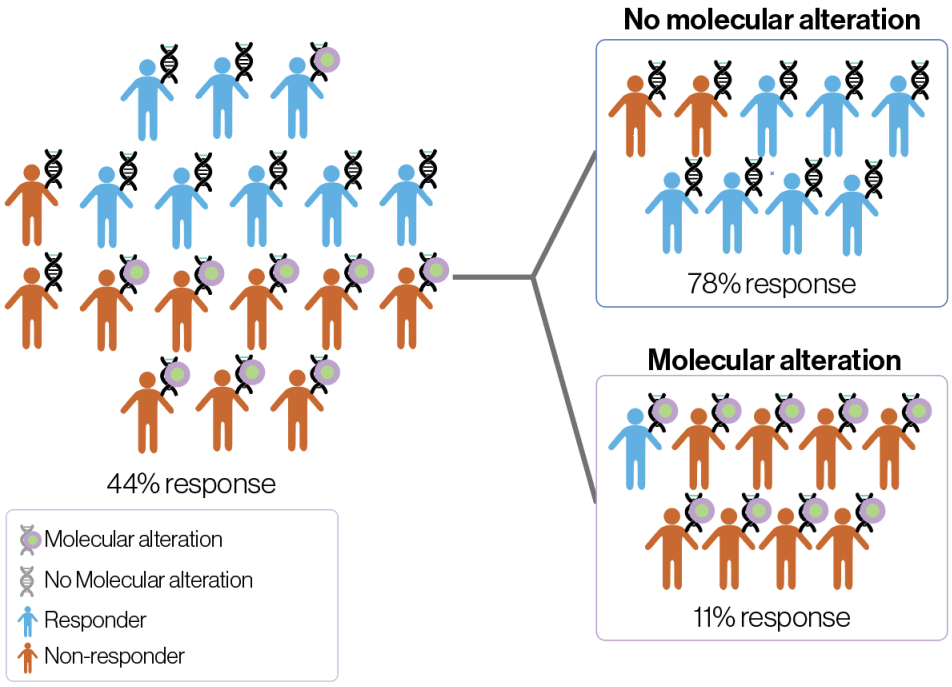


Figure 2. Schematic representation of a predictive biomarker: a molecular trait can be associated with a patient cohort's clinical outcome (e.g., response to treatment). Upon stratifying the patient cohort based on the presence or absence of a molecular alteration, the patients can be classified by having a higher or lower likelihood of responding to treatment.

1.4.1 Biomarkers of response to neoadjuvant immune checkpoint inhibitors

ICIs for the treatment of UC in the neoadjuvant setting (i.e., before surgery) are thus far not approved for clinical usage, in contrast to the metastatic setting in which several ICIs are approved for usage[15,32].

With more cancer types treated by ICI, we can now pool data from many studies to learn more about the mechanisms of response shared between the various types. Associations with treatment response have been measured based on a discovery-based or knowledge-based approach to the components of tumor-immune interactions. For instance, the cancer immunogram[33,34] serves as a fundamental framework to understand the baseline characteristics a tumor must satisfy to respond to ICIs. However, despite more data being available for research, the behavior of biomarkers and candidate associations of response still needs to be more heterogeneous between clinical trials. Existing ICI response biomarkers, based on pre-therapy molecular profiles, have displayed a varying performance in large validation cohorts (n>1000 patients), particularly in a pan-cancer setting[35,36] but also between UC trials[37]. Such variability can be attributed to heterogeneity between tumors and ICI treatment courses (e.g., monotherapy vs. combination therapy, or different immune checkpoints being targeted).

European drug agencies have restricted the treatment of adjuvant UC with nivolumab to patients having high expression of PD-L1 on their tumor cells[38]. Several other response biomarkers for ICIs have been proposed, such as the tumor mutational burden (TMB) across multiple cancer types[39]. A high mutational profile increases the foreignness of a tumor, increasing the chances of being recognized by the immune system as foreign[40]. Moreover, molecular expression profiles indicative of pre-existing $CD8^+$ T-cell immunity have been linked to an increased sensitivity of tumors to T-cell effectors[40]. Different gene expression signatures related to $CD8^+$ T-cell effectiveness, such as a general immune profile score (e.g., the tumor inflammation score signature) and interferon-gamma signaling[41] have shown associations with response to ICIs. Moreover, an impairment of the antigen presentation machinery has been linked to non-response[42]. On the other hand, expression profiles of a suppressive TME indicative of inhibitory tumor metabolism, such as TGF-beta signaling, have been linked to non-response to ICIs across multiple cancer types[43,44]. Lastly, the abundance and configuration of immune cells within the TME have also been proposed as candidate biomarkers, such as infiltration of $CD8^+$ T-cells, the configuration of immune phenotypes, and the presence of tertiary lymphoid structures (TLS)[45]. It is expected that a new avenue of molecular data paired with clinical data will be valuable to further develop robust candidate biomarkers for response to ICIs in UC.

1.4.2 Biomarkers of response to neoadjuvant chemotherapy

Although chemotherapy-based treatments have been approved to treat UC tumors for over 20 years[46], to date, there is not a single biomarker with a sufficiently high odds ratio to have clinical utility[8].

Several candidate biomarkers predict NAC response[8,46]. Among them, clinical response associated with mutations in the transcription-coupled nucleotide excision repair *ERCC2*[47,48],

receptor tyrosine kinase *ERBB2*[49], DNA repair genes (*ATM*/*RB1*/*FANCC*, *ERCC2*, *BRCA1*, *ERCC1*)[50,51], tumor suppressor *TP53*[52], and fibroblast growth receptor *FGFR3*[53]. In addition to genomic biomarkers, transcriptomic biomarkers have also been proposed as candidate response biomarkers to NAC, including aberrant expression profiles of p53[54], distinct molecular subtypes[55], and expression of DNA excision repair *ERCC1*[56], which all have been linked to an increased sensitivity to NAC. Altered methylation profiles have also been correlated with response to NAC, including the cell-free methylation profile at baseline and after one cycle of NAC[57].

Besides uni-modal biomarker strategies, multi-modal approaches that integrate multiple molecular data types have also been proposed to comprehensively understand tumors[58], such as a combination of high immune cell infiltration, high PD-1 protein expression, aberrations in chromosomal alterations, indels, *BRCA2* mutations and a non-basal/squamous subtype, which altogether are linked to an increased response to NAC.

Biomarkers are too often cohort-specific and fail to generalize, limiting their usage as response biomarkers and clinical utility[8,46]. Validation approaches in larger cohorts are still needed to generate more evidence to further develop response biomarkers of NAC.

1.5 CONTRIBUTION AND THESIS OBJECTIVE

A deeper molecular understanding of UC tumors, their associated TME, and their systemic components is essential to understanding the response and resistance mechanisms to treatments, providing a knowledge foundation to develop novel biomarkers for patient stratification in the precision medicine era, and serving as a basis to develop novel anti-cancer treatments.

This thesis aims to characterize the TME immune landscape of UC tumors, understand the response to neoadjuvant treatments in UC, and quantify neoadjuvant treatment dynamics upon ICIs. To do so, the contributions of each chapter (**Figure 3**) are as follows:

- **Chapter 2** assesses the validity of several proposed baseline genomic biomarkers to neoadjuvant cisplatin-based chemotherapy (NAC) in a large independent cohort.
- In **Chapter 3,** we first dissect the TME immune landscape of untreated UC tumors to understand the baseline immunological characteristics of UC tumors, quantify the immunological components and heterogeneity of tertiary lymphoid structures, and assess the differences between untreated and ICI-treated UC tumors to quantify treatment-associated effects, which altogether provide a basis to understand response and resistance to ICIs and to develop novel biomarkers and treatment schedules.

- In **Chapter 4,** we first show that a neoadjuvant combination of immune checkpoint inhibitors (ICIs) involving PD-1 and CTLA-4 blockade (ipilimumab + nivolumab) is feasible in a phase-I trial of high-grade UC tumors, explore biomarkers of response to ICIs, and quantify treatment dynamics components that impair response to ICIs.
- Lastly, in **Chapter 5**, we propose a novel framework to quantify spatial relationships within the TME and propose novel biomarkers of response to combination ICIs in UC, which are also validated in head and neck cancer tumors.

Altogether, our results provide a comprehensive foundation to deepen our knowledge of UC tumor biology and tumor immunology, which serve as a base to further understand the response to neoadjuvant treatments such as ICIs and chemotherapy and develop novel biomarkers of response.

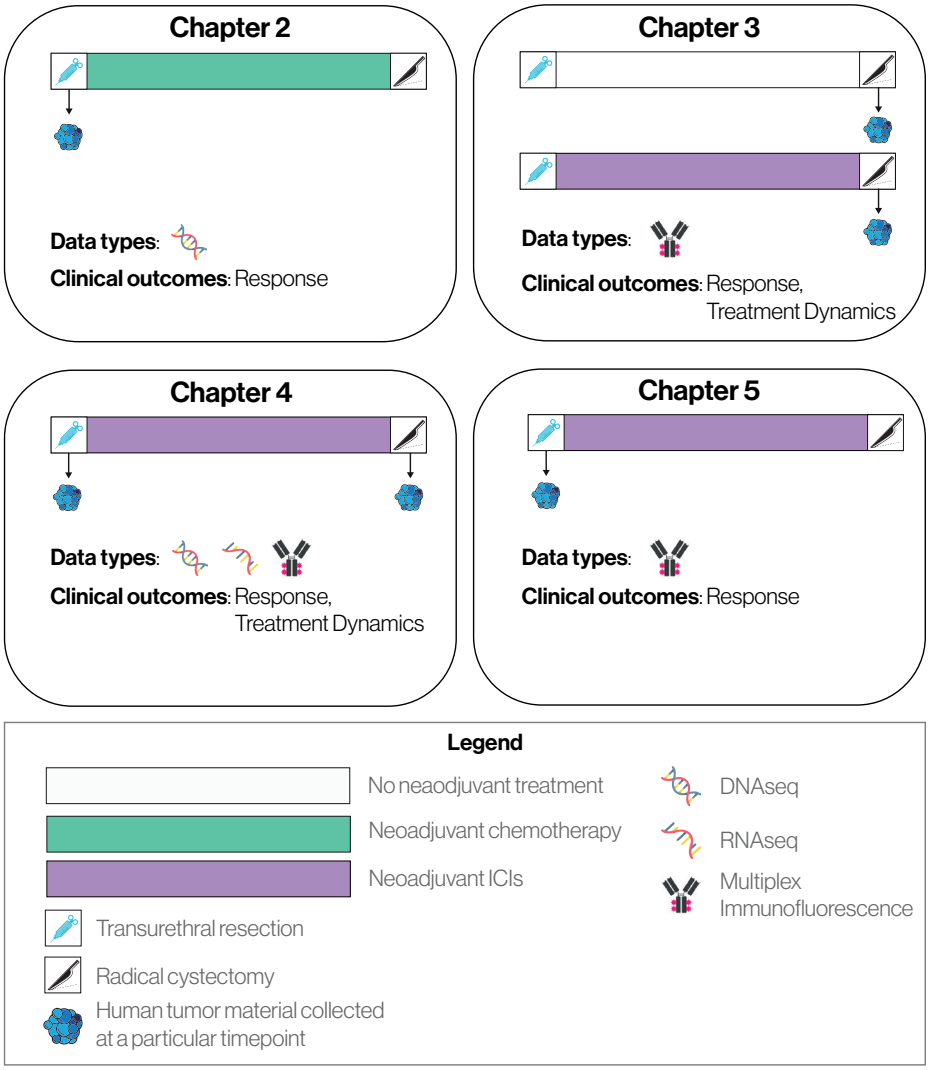


Figure 3. Graphical abstract of the content of the thesis. Each box represents a chapter. Each line denotes a treatment line for a patient cohort. Arrows indicate when the tumor material was collected (e.g., pre-treatment with transurethral resection, post-treatment with radical cystectomy). Coloring indicates the type of neoadjuvant treatment, and below is indicated the molecular and clinical data types collected in each chapter.

Abbreviations: ICIs: immune checkpoint inhibitors

REFERENCES

1. Tian, T., Olson, S., Whitacre, J. M. & Harding, A. The origins of cancer robustness and evolvability. *Integr. Biol.* **3**, 17–30 (2011).
2. Howlader, N. *et al.* SEER Cancer Statistics Review, 1975-2018, National Cancer Institute. Bethesda, MD. Preprint at (2021).
3. Robertson, A. G. *et al.* Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. *Cell* **171**, 540–556.e25 (2017).
4. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of urothelial bladder carcinoma. *Nature* **507**, 315–322 (2014).
5. Kamoun, A. *et al.* A Consensus Molecular Classification of Muscle-invasive Bladder Cancer. *Eur. Urol.* **77**, 420–433 (2020).
6. Stein, J. P. *et al.* Radical cystectomy in the treatment of invasive bladder cancer: long-term results in 1,054 patients. *J. Clin. Oncol.* **19**, 666–675 (2001).
7. Witjes, J. A. *et al.* European Association of Urology Guidelines on Muscle-invasive and Metastatic Bladder Cancer: Summary of the 2020 Guidelines. *Eur. Urol.* **79**, 82–104 (2021).
8. Buttigliero, C., Tucci, M., Vignani, F., Scagliotti, G. V. & Di Maio, M. Molecular biomarkers to predict response to neoadjuvant chemotherapy for bladder cancer. *Cancer Treat. Rev.* **54**, 1–9 (2017).
9. Choueiri, T. K. *et al.* Neoadjuvant dose-dense methotrexate, vinblastine, doxorubicin, and cisplatin with pegfilgrastim support in muscle-invasive urothelial cancer: pathologic, radiologic, and biomarker correlates. *J. Clin. Oncol.* **32**, 1889–1894 (2014).
10. Grivas, P. D., Melas, M. & Papavassiliou, A. G. The biological complexity of urothelial carcinoma: Insights into carcinogenesis, targets and biomarkers of response to therapeutic approaches. *Semin. Cancer Biol.* **35**, 125–132 (2015).
11. Anderson, N. M. & Simon, M. C. The tumor microenvironment. *Curr. Biol.* **30**, R921–R925 (2020).
12. de Visser, K. E. & Joyce, J. A. The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. *Cancer Cell* **41**, 374–403 (2023).
13. Chen, D. S. & Mellman, I. Oncology meets immunology: the cancer-immunity cycle. *Immunity* **39**, 1–10 (2013).
14. He, X. & Xu, C. Immune checkpoint signaling and cancer immunotherapy. *Cell Res.* **30**, 660–669 (2020).
15. Scott, E. C. *et al.* Trends in the approval of cancer therapies by the FDA in the twenty-first century. *Nat. Rev. Drug Discov.* (2023) doi:10.1038/s41573-023-00723-4.
16. Tang, Q. *et al.* PD-1/PD-L1 immune checkpoint inhibitors in neoadjuvant therapy for solid tumors (Review). *Int. J. Oncol.* **62**, (2023).
17. Wojtukiewicz, M. Z. *et al.* Inhibitors of immune checkpoints—PD-1, PD-L1, CTLA-4—new opportunities for cancer patients and a new challenge for internists and general practitioners. *Cancer Metastasis Rev.* **40**, 949–982 (2021).
18. Lopez-Beltran, A. *et al.* Immune Checkpoint Inhibitors for the Treatment of Bladder Cancer. *Cancers* **13**, (2021).
19. van Dorp, J. *et al.* High- or low-dose preoperative ipilimumab plus nivolumab in stage III urothelial cancer: the phase 1B NABUCCO trial. *Nat. Med.* **29**, 588–592 (2023).
20. Kim, L. H. C. & Patel, M. I. Transurethral resection of bladder tumour (TURBT). *Transl. Androl. Urol.* **9**, 3056–3072 (2020).
21. Crocetto, F. *et al.* Liquid biopsy in bladder cancer: State of the art and future perspectives. *Crit. Rev. Oncol. Hematol.* **170**, 103577 (2022).

22. Roelofsen, L. M., Kaptein, P. & Thommen, D. S. Multimodal predictors for precision immunotherapy. *Immunooncol Technol* **14**, 100071 (2022).
23. Tripathi, A. & Grivas, P. The utility of next generation sequencing in advanced urothelial carcinoma. *Eur Urol Focus* **6**, 41–44 (2020).
24. Sheng, W. *et al.* Multiplex Immunofluorescence: A Powerful Tool in Cancer Immunotherapy. *Int. J. Mol. Sci.* **24**, (2023).
25. Hatogai, K. *et al.* Multiplex immunofluorescence to assess the tumor microenvironment in bladder cancer. *Urologic Oncology: Seminars and Original Investigations* **38**, 907–908 (2020).
26. Li, X. & Wang, C.-Y. From bulk, single-cell to spatial RNA sequencing. *Int. J. Oral Sci.* **13**, 36 (2021).
27. Fu, T. *et al.* Spatial architecture of the immune microenvironment orchestrates tumor immunity and therapeutic response. *J. Hematol. Oncol.* **14**, 98 (2021).
28. Palla, G., Fischer, D. S., Regev, A. & Theis, F. J. Spatial components of molecular tissue biology. *Nat. Biotechnol.* (2022) doi:10.1038/s41587-021-01182-1.
29. Lewis, S. M. *et al.* Spatial omics and multiplexed imaging to explore cancer biology. *Nat. Methods* **18**, 997–1012 (2021).
30. Wang, D. *et al.* Adverse Effects and Toxicity of Immune Checkpoint Inhibitors For Patients With Urothelial Carcinoma. *Front. Pharmacol.* **12**, 710943 (2021).
31. Narayan, V. & Vaughn, D. Pharmacokinetic and toxicity considerations in the use of neoadjuvant chemotherapy for bladder cancer. *Expert Opin. Drug Metab. Toxicol.* **11**, 731–742 (2015).
32. Twomey, J. D. & Zhang, B. Cancer Immunotherapy Update: FDA-Approved Checkpoint Inhibitors and Companion Diagnostics. *AAPS J.* **23**, 39 (2021).
33. Blank, C. U., Haanen, J. B., Ribas, A. & Schumacher, T. N. The 'cancer immunogram'. *Science* **352**, 658–660 (2016).
34. van Dijk, N. *et al.* The Cancer Immunogram as a Framework for Personalized Immunotherapy in Urothelial Cancer. *Eur. Urol.* **75**, 435–444 (2019).
35. Litchfield, K. *et al.* Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. *Cell* **184**, 596–614.e14 (2021).
36. Chowell, D. *et al.* Improved prediction of immune checkpoint blockade efficacy across multiple cancer types. *Nat. Biotechnol.* **40**, 499–506 (2022).
37. Meeks, J. J. *et al.* Checkpoint Inhibitors in Urothelial Carcinoma—Future Directions and Biomarker Selection. *Eur. Urol.* (2023) doi:10.1016/j.eururo.2023.05.011.
38. ESMO. EMA recommends extension of indications for nivolumab. https://www.esmo.org/oncology-news/ema-recommends-extension-of-indications-for-nivolumab2.
39. Klempner, S. J. *et al.* Tumor Mutational Burden as a Predictive Biomarker for Response to Immune Checkpoint Inhibitors: A Review of Current Evidence. *Oncologist* **25**, e147–e159 (2020).
40. Cristescu, R. *et al.* Pan-tumor genomic biomarkers for PD-1 checkpoint blockade–based immunotherapy. *Science* **362**, eaar3593 (2018).
41. Ayers, M. *et al.* IFN-γ–related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Invest.* **127**, 2930–2940 (8 2017).
42. Thompson, J. C. *et al.* Gene signature of antigen processing and presentation machinery predicts response to checkpoint blockade in non-small cell lung cancer (NSCLC) and melanoma. *J Immunother Cancer* **8**, (2020).
43. Fan, Y. *et al.* TGF-β–Induced Upregulation of malat1 Promotes Bladder Cancer Metastasis by Associating with suz12. *Clin. Cancer Res.* **20**, 1531–1541 (2014).
44. Mariathasan, S. *et al.* TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* **554**, 544–548 (2018).

45. Helmink, B. A. *et al.* B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* **577**, 549–555 (2020).
46. Roviello, G. *et al.* Neoadjuvant Treatment in Muscle-Invasive Bladder Cancer: From the Beginning to the Latest Developments. *Front. Oncol.* **12**, 912699 (2022).
47. Van Allen, E. M. *et al.* Somatic ERCC2 mutations correlate with cisplatin sensitivity in muscle-invasive urothelial carcinoma. *Cancer Discov.* **4**, 1140–1153 (2014).
48. Liu, D. *et al.* Clinical Validation of Chemotherapy Response Biomarker ERCC2 in Muscle-Invasive Urothelial Bladder Carcinoma. *JAMA Oncol* **2**, 1094–1096 (2016).
49. Groenendijk, F. H. *et al.* ERBB2 Mutations Characterize a Subgroup of Muscle-invasive Bladder Cancers with Excellent Response to Neoadjuvant Chemotherapy. *Eur. Urol.* **69**, 384–388 (2016).
50. Plimack, E. R. *et al.* Defects in DNA Repair Genes Predict Response to Neoadjuvant Cisplatin-based Chemotherapy in Muscle-invasive Bladder Cancer. *Eur. Urol.* **68**, 959–967 (2015).
51. Miron, B. *et al.* Defects in DNA Repair Genes Confer Improved Long-term Survival after Cisplatin-based Neoadjuvant Chemotherapy for Muscle-invasive Bladder Cancer. *Eur Urol Oncol* **3**, 544–547 (2020).
52. Watanabe, J. *et al.* Clinical evaluation of p53 mutations in urothelial carcinoma by IHC and FASAY. *Urology* **63**, 989–993 (2004).
53. Teo, M. Y. *et al.* Fibroblast Growth Factor Receptor 3 Alteration Status is Associated with Differential Sensitivity to Platinum-based Chemotherapy in Locally Advanced and Metastatic Urothelial Carcinoma. *Eur. Urol.* **78**, 907–915 (2020).
54. Williams, P. D. *et al.* Concordant gene expression signatures predict clinical outcomes of cancer patients undergoing systemic therapy. *Cancer Res.* **69**, 8302–8309 (2009).
55. Sjödahl, G. *et al.* Different Responses to Neoadjuvant Chemotherapy in Urothelial Carcinoma Molecular Subtypes. *Eur. Urol.* **81**, 523–532 (2022).
56. Yang, Z. *et al.* Somatic FGFR3 Mutations Distinguish a Subgroup of Muscle-Invasive Bladder Cancers with Response to Neoadjuvant Chemotherapy. *EBioMedicine* **35**, 198–203 (2018).
57. Lu, Y.-T. *et al.* Cell-free DNA Methylation as a Predictive Biomarker of Response to Neoadjuvant Chemotherapy for Patients with Muscle-invasive Bladder Cancer in SWOG S1314. *Eur Urol Oncol* (2023) doi:10.1016/j.euo.2023.03.008.
58. Taber, A. *et al.* Molecular correlates of cisplatin-based chemotherapy response in muscle invasive bladder cancer by integrated multi-omics analysis. *Nat. Commun.* **11**, 4858 (2020).

2

Assessment of Predictive Genomic Blomarkers for Response to Cisplatin-based Neoadjuvant Chemotherapy in Bladder Cancer

Alberto Gil-Jimenez[a,b,#], Jeroen van Dorp[a,#], Alberto Contreras-Sanz[c], Kristan van der Vos[a], Daniel J. Vis[a,b], Linde Braaf[d], Annegien Broeks[d], Ron Kerkhoven[e], Kim E.M. van Kessel[f], María José Ribal[g], Antonio Alcaraz[g], Lodewyk F.A. Wessels[a,b,h], Roland Seiler[i,j], Jonathan L. Wright[k], Lourdes Mengual[g], Joost Boormans[f], Bas W.G. van Rhijn[l,m], Peter C. Black[c], Michiel S. van der Heijden[a,n]

Parts of this chapter have been published in European Urology (2022): https://www.sciencedirect.com/science/article/pii/S0302283822025386

[a]) Department of Molecular Carcinogenesis, Netherlands Cancer Institute, Amsterdam, the Netherlands
[b]) Oncode Institute, Utrecht, the Netherlands
[c]) The Vancouver Prostate Centre and Department of Urologic Sciences, University of British Columbia, Vancouver, British Columbia, Canada
[d]) Core Facility Molecular Pathology & Biobanking, Netherlands Cancer Institute, Amsterdam, the Netherlands
[e]) Core Facility Genomics, Netherlands Cancer Institute, Amsterdam, the Netherlands
[f]) Department of Urology, Erasmus MC Cancer Institute, University Medical Center, Rotterdam, the Netherlands
[g]) Laboratory and Department of Urology, Hospital Clínic de Barcelona, Institut d'Investigacions Biomèdiques August Pi I Sunyer (IDIBAPS), Universitat de Barcelona, Barcelona, Spain
[h]) Faculty of EEMCS, Delft University of Technology, Delft, the Netherlands
[i]) Department of BioMedical Research, University of Bern, Bern, Switzerland
[j]) Department of Urology, Hospital Center Biel, Biel, Switzerland
[k]) Department of Urology, University of Washington School of Medicine, Seattle, Washington, USA
[l]) Department of Surgical Oncology (Urology), Netherlands Cancer Institute, Amsterdam, the Netherlands
[m]) Department of Urology, Caritas St. Josef Medical Centre, University of Regensburg, Regensburg, Germany
[n]) Department of Medical Oncology, Netherlands Cancer Institute, Amsterdam, the Netherlands
[#]) These authors contributed equally

Keywords: Muscle-Invasive Bladder Cancer; Neoadjuvant Chemotherapy; DNA sequencing; Response prediction; Somatic mutations; Cisplatin-based Chemotherapy

ABSTRACT

Cisplatin-based neoadjuvant chemotherapy (NAC) followed by radical cystectomy is recommended for patients with muscle-invasive bladder cancer (MIBC). Somatic deleterious mutations in *ERCC2*, gain-of-function mutations in *ERBB2*, and alterations in *ATM*, *RB1*, and *FANCC* have been shown to correlate with pathological response to NAC in MIBC.

The objective of this study was to validate these genomic biomarkers in pre-treatment transurethral resection (TUR) material from an independent retrospective cohort of 165 MIBC patients who had subsequently undergone NAC and radical surgery. Patients with ypT0/Tis/Ta/T1N0 disease after surgery were defined as responders.

Somatic deleterious mutations in *ERCC2* were found in 9/68 (13%) evaluable responders and in 2/95 (2%) evaluable non-responders (p=0.009, FDR=0.03). No correlation was observed between response and alterations in *ERBB2* or in *ATM*, *RB1* or *FANCC* alone or in combination. In an exploratory analysis, no additional genomic alterations discriminated between responders and non-responders to NAC. No further associations were identified between the aforementioned biomarkers and complete pathological complete response (ypT0N0) after surgery.

In conclusion, we observed a positive association between deleterious mutations in *ERCC2* and pathological response to NAC, but not overall survival or recurrence-free survival. Other previously reported genomic biomarkers were not validated.

Patient Summary

It is currently unknown which patients will respond to chemotherapy before definitive surgery for bladder cancer. Previous studies described several gene mutations in bladder cancer that correlated with chemotherapy response. This study confirmed that patients with bladder cancer with a mutation in the *ERCC2* gene often respond to chemotherapy.

2.1 MAIN

Neoadjuvant cisplatin-based chemotherapy (NAC) followed by radical cystectomy is recommended for patients with muscle-invasive bladder cancer (MIBC)[1]. Pathological response after treatment with NAC is strongly associated with recurrence-free survival (RFS) and overall survival (OS)[2]. Currently, clinicians are unable to identify which patients will benefit from NAC. Genomic biomarkers have been described to correlate with response to NAC, including somatic deleterious mutations in *ERCC2,* gain-of-function mutations in *ERBB2*, and alterations in *ATM*, *RB1* and *FANCC*[3-7]. However, none of these biomarkers have been validated in larger independent cohorts and are consequently not used in clinical practice[1,8].

Here, we set out to to validate these genomic biomarkers in an independent multicenter retrospective cohort. Pre-treatment tissue derived from five centers was sequenced at the Netherlands Cancer Institute (NKI cohort, n=117) or Vancouver Prostate Centre (Vancouver cohort, n=48, **Supplementary Figure 1**). All patients were diagnosed with MIBC (cT2-4aN0M0 and/or cT1-4aN1-3M0) by transurethral resection (TUR) and were treated with at least two cycles of cisplatin-based NAC, followed by radical cystectomy. The primary endpoint of this study was pathological response, defined as ypT0/Tis/Ta/T1N0 after surgery[2,9]. Seventy of 165 patients (42%) were categorized as responders. Pathological complete response after surgery, defined as ypT0N0, was used as a secondary endpoint which was observed in 51 of 165 patients (31%).

Baseline age, gender, chemotherapy regimen, and number of cycles of chemotherapy did not differ between response groups, however cT-stage at baseline was higher in the non-responders (**Table 1**). Furthermore, baseline cT-stage and chemotherapy regimen differed between cohorts (**Supplementary Table 1**). Tumor DNA extracted from TUR samples obtained prior to NAC was sequenced using a targeted capture-based panel for the NKI cohort and whole exome sequencing for the Vancouver cohort. Somatic variants in *ERCC2*, *ERBB2*, *ATM*, *RB1* and *FANCC* were inferred from population databases (**Supplementary Methods**). Mutations were predicted to be functional (deleterious or gain-of-function) using the annotation databases OncoKB, ClinVar, SIFT, FATHMM, and PolyPhen-2 (**Supplementary Methods**). A high concordance between the observed and the TCGA mutation rates was observed (**Supplementary Table 2**).

Table 1. Baseline characteristics and response of 165 patients with muscle-invasive bladder cancer treated with neoadjuvant chemotherapy and radical cystectomy.

	NKI		Vancouver		
	Responders (ypT0/Tis/Ta/T1N0)	**Non-responders (≥ypT2N0)**	**Responders (ypT0/Tis/Ta/T1N0)**	**Non-responders (≥ypT2N0)**	**Significance Responders vs Non-responders (full cohort)[a]**
Number of patients	53	64	17	31	-
Median age in years (IQR)	71.0 (61.0, 75.0)	71.0 (61.0, 77.3)	61.2 (56.0, 66.0)	65.5 (58.3, 73.0)	0.2
Male sex (% of pts)	40 (76%)	38 (59%)	15 (88%)	24 (77%)	0.08
cT-stage (% of pts)					**0.04**
cT1	1 (2%)	0 (0%)	0 (0%)	0 (0%)	
cT2	27 (51%)	19 (30%)	2 (12%)	4 (13%)	
cT3	21 (40%)	26 (40%)	7 (41%)	20 (65%)	
cT4	4 (7%)	19 (30%)	8 (47%)	7 (22%)	
cN-stage (% of pts)					0.6
cN0	31 (59%)	40 (63%)	10 (59%)	12 (39%)	
cN+	22 (41%)	24 (37%)	7 (41%)	19 (61%)	
Chemotherapy regimen (% of pts)					0.16
Cis/Gem	40 (75%)	41 (64%)	14 (82%)	28 (90%)	
MVAC	11 (21%)	23 (36%)	3 (18%)	3 (10%)	
CMV	2 (4%)	0 (0%)	0 (0%)	0 (0%)	
Chemotherapy cycles received (% of pts)					0.8
2	2 (4%)	2 (3%)	0 (0%)	2 (6%)	
3	10 (19%)	13 (20%)	6 (35%)	7 (23%)	
4	39 (74%)	46 (72%)	10 (59%)	17 (55%)	

Table 1. *(Continued)*

	NKI		Vancouver		
	Responders (ypT0/Tis/ Ta/T1N0)	**Non-responders (≥ypT2N0)**	**Responders (ypT0/Tis/Ta/ T1N0)**	**Non-responders (≥ypT2N0)**	**Significance Responders vs Non-responders (full cohort)[a]**
>4	2 (4%)	3 (5%)	1 (6%)	5 (16%)	
Pathological response (% of pts)					
ypT0N0 (complete response)	40 (75%)	0 (0%)	13 (76%)	0 (0%)	
ypTis/Ta/T1N0	13 (25%)	0 (0%)	4 (24%)	0 (0%)	
≥ypT2N0 (non-response)	0 (0%)	64 (100%)	0 (0%)	31 (100%)	

[a] Fisher's exact test for binary predictors, t-test for numerical predictors. All the statistical tests were two-sided. No adjustments were made for multiple hypothesis testing. Significant associations are highlighted on bold. *Cis/Gem = cisplatin + gemcitabine; MVAC = methotrexate + vinblastine + docorubicin + cisplatin; CMV = cisplatin + methotrexate + vinblastine; IQR = interquartile range.*

After filtering, deleterious mutations in *ERCC2* were found in nine of 68 (13%) evaluable responders and in two of 95 (2%) evaluable non-responders (p=0.009, **Figure 1A**). We found relevant gain-of-function mutations in *ERBB2* in nine of 69 (13%) evaluable responders and five of 95 (5%) evaluable non-responders (p=0.09, **Figure 1A**). Twenty-seven of 70 (39%) responders had ≥1 alteration in *ATM*, *RB1* or *FANCC* compared to 25 of 95 (26%) non-responders (p=0.13, **Figure 1A**). Nine of eleven patients (82%) with a deleterious mutation in *ERCC2* had a pathological response after treatment with NAC, as opposed to 62 of 154 patients (40%) without any relevant mutations in *ERCC2* (**Supplementary Table 3**). After correction for multiple hypothesis testing (three hypotheses), mutations in *ERCC2* were significantly enriched in responders (FDR=0.03, **Figure 1B**). The association remained when adjusted for cT-stage in a multivariable logistic regression model (p_{ERCC2}=0.008, p_{cT2}=p_{cT3}=p_{cT4}>0.9), or when patients that received less than 3 cycles of NAC were excluded (**Supplementary Figure 2**). Baseline clinical difference between *ERCC2* mutated and wild-type patients were not identified (**Supplementary Table 3**).

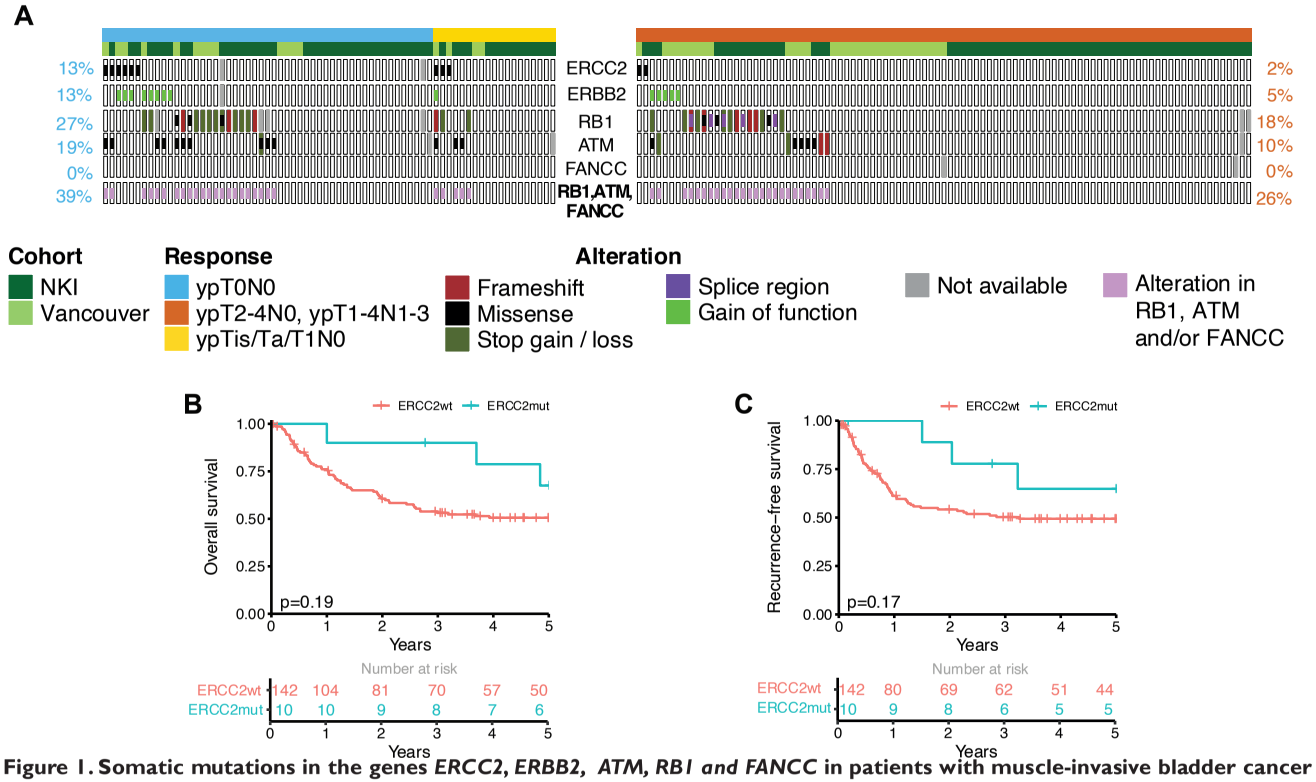


Figure 1. Somatic mutations in the genes *ERCC2*, *ERBB2*, *ATM*, *RB1* and *FANCC* in patients with muscle-invasive bladder cancer treated with neoadjuvant chemotherapy. A) Overview of relevant mutations for each patient **(Supplementary Methods - Variant calling and inference of somatic and functional variants).** Left panel shows patients with ypT0N0 (light blue, n=51) or ypTis/Ta/T1N0 (yellow, n=19) after neoadjuvant chemotherapy (responders), and right panel shows non-responders (orange, n=95). Percentages represent the number of patients with a relevant mutation relative to the total number of eligible patients for that specific gene for responders (left) and non-responders (right). Patients with an alteration in any one of *ATM*, *RB1* or *FANCC* are indicated in the last row. **B)** Five-year overall survival (OS) for patients with (blue) and without (red) mutations in *ERCC2*. The p-value indicates statistical significance by a log-rank test. **C)** 5-year recurrence-free survival (RFS) for patients with (blue) and without (red) mutations in *ERCC2*. The p-value inside the survival plot indicates statistical significance by a log-rank test. *NKI = patients from the NKI-cohort; wt = wild type; mut = mutant; Not available = gene-coverage below 20 reads.*

In contrast, alterations in *ERCC2*, *ERBB2*, or in any one of *ATM*, *RB1* or *FANCC* did not associate with a pathological complete response (ypT0N0) after correcting for multiple hypothesis testing (FDR_{ERCC2}=0.09, FDR_{ERBB2}=0.07, $FDR_{ATM/RB1/FANNC}$=0.07, **Supplementary Figure 3)**

The median duration of follow-up for patients using reverse censoring was 7.2 years. The 5-year OS rates for patients with and without mutations in *ERCC2* were 75% (95% confidence interval (CI): 50%-100%) and 52% (95% CI: 45%-62%), respectively (p=0.2, **Figure 1C**). The 5-year RFS rates were 65% (95% CI: 39%-100%) and 49% (95% CI: 42%-59%), respectively (p=0.2, **Figure 1D**). Thus, while the Kaplan Meier curves appear to separate according to *ERCC2* mutation status, we could not demonstrate a statistical difference for either OS or RFS, possibly due to the low frequency of *ERCC2* mutations.

Following earlier analyses by Plimack and colleagues[6], we assessed copy number alterations (CNA) for *ATM*, *RB1,* and *FANCC* by shallow whole genome sequencing for patients from the NKI cohort (n=117, **Supplementary Methods**). CNA on the Vancouver cohort could not be confidently assessed due to a lack of germline data. We found seven CNA in *ATM*, *RB1* and/or *FANCC* in all evaluable patients. Together with the previously described mutations, 22 of 53 (42%) responders had ≥1 alteration in *ATM*, *RB1* or *FANCC* versus 20 of 64 (31%) non-responders (p=0.052, **Supplementary Figure 4**).

In a further exploratory analysis, mutations frequently occurring in MIBC were assessed for their correlation with response to NAC (Supplementary Figure 5). This analysis included *FGFR3*, which was previously associated with negative outcome after chemotherapy (Supplementary Figure 6)[10]. No association with response was identified after correction for multiple hypothesis testing (Supplementary Figure 5).

There are several limitations to this study. The genomic data were derived using different sequencing technologies at different centers, leading to potential biases in the mutation frequency. Furthermore, we lacked germline data and somatic variants were filtered with the help of population databases to remove benign germline variants. As germline DNA is often unavailable, this approach is common practice and was also used in the original studies of *ERBB2* and *ATM*/*RB1*/*FANCC*[4, 6]. Multiple definitions of response have been used in previous studies, thus introducing heterogeneity between studies. Complete pathological response (ypT0N0) and pathological downstaging (ypT0/Tis/Ta/T1N0) are commonly used. Long-term clinical outcome is favorable in both groups, although patients with ypT0/TisN0 may have a modest survival benefit over patients with ypT0/Tis/Ta/T1N0[2, 9].

In summary, we attempted to validate mutations in *ERCC2*, *ERBB2*, *ATM, RB1* and *FANCC* as predictive markers of pathological response in a cohort of 165 patients treated with

NAC. We confirmed a positive association of deleterious mutations in *ERCC2* with pathological response (ypT0/Tis/Ta/T1N0), but not with complete response (ypT0N0), OS or RFS. Prospective evaluation of *ERCC2* mutations as a biomarker for response to NAC is needed to confirm our results.

Funding
This project was funded by NWO.

2.2 METHODS

2.2.1 Study population / Treatment

The full cohort consisted of 165 prospectively collected samples from five different centers. All patients had muscle-invasive bladder cancer (MIBC, cT2-4aN0M0 and/or cT1-4aN1-3M0) diagnosed by transurethral resection (TUR) and treated with at least two cycles of neoadjuvant chemotherapy (NAC) followed by radical cystectomy. NAC consisted of either cisplatin + gemcitabine (cis/gem), methotrexate + vinblastine + doxorubicin + cisplatin (MVAC) or cisplatin + methotrexate + vinblastine (CMV).

Patient cohorts were named after the center in which the sequencing was performed. The NKI cohort (n=117) included retrospectively collected MIBC samples from three centers: Amsterdam (The Netherlands), Rotterdam (The Netherlands), and Barcelona (Spain) and the Vancouver cohort (n=48) included MIBC samples from two institutions: Bern (Switzerland) and Seattle (Washington, USA) were compiled (**Supplementary Figure 1, Table 1**).

2.2.2 Targeted DNA sequencing (NKI cohort)

Formalin-fixed paraffin-embedded (FFPE) tumor blocks from TUR material were collected from the different hospitals and centrally reviewed by an experienced pathologist. Tumor area was marked for every tumor block and DNA was collected from subsequent FFPE slides (10 µm). Nanodrop 2000 (Thermofisher) was used to quantify the total amount of DNA, and Qubit® dsDNA HS Assay Kit (Invitrogen, cat no Q32851) was used to quantify the amount of double stranded DNA. Covaris shearing was used to fragment a maximum of 2000 ng of double stranded genomic DNA to get fragment sizes of 200-300bp. 2X Agencourt AMPure XP PCR Purification beads were used to purify samples following the manufacturer's instructions (Beckman Coulter, cat no A63881). The sheared DNA samples were qualified and quantified using a BioAnalyzer system (DNA7500 assay kit, Agilent Technologies cat no. 5067- 1506). Library preparation for Illumina sequencing was done using the KAPA HTP Prep Kit (KAPA Biosystems, KK8234) with an input of maximum 1 µg sheared DNA. Four PCR cycles were done during library amplification to obtain sufficient

yield for the exome capture. Libraries were cleaned up using 1X AMPure XP beads. The DNA libraries were analyzed on a Bioanalyzer system using the DNA7500 chips to determine the concentration. With 150 ng of each indexed sample three pools of eight samples were prepared, and 2 µl of IDT TS-mix universal blockers (IDT, Cat: 1075475) and 5 µl Human Cot-1 DNA (Invitrogen, Ref: 15279-011) was added to the pools. Then, a concentrator was used to dry the pool, and to rehydrate 8.5 µl of hybridization buffer, 3.4 µl Hybridization component A (SeqCap Hybridization and wash kit, Roche, Ref: 05634253001) and 1.1 µl nuclease-free water was added. The pool underwent incubation at RT for 10 minutes, and at 96 degrees Celsius for 10 minutes. The samples were hybridized with 4 µl of the custom 44 gene bladder cancer panel (which included muscle-invasive bladder cancer driver genes, clinically relevant genes, and frequently mutated genes) at 65°C for 24 hours. IDT protocol (Rapid protocol for DNA probe Hybridization and Target Capture using an Illumina TruSeq or Ion Torrent Library) was followed to capture and wash the pool, and amplified using 10 PCR cycles. The amplified pool was purified using AMPure® XP beads (Beckman Coulter). The purified pools were quantified on the Agilent Bioanalyzer 7500 system and one sequence pool was made by equimolar pooling. The sequence pool was diluted to a final concentration of 10 nM and subjected to sequencing on an Illlumina Nextseq 550 machine with a MID 150 cycle kit for a paired end 75bp run, following manufacturer's instructions.

Sequencing reads were aligned against the Human Reference Genome Ghr38 using Burrows-Wheeler aligner v0.7.17-r1188. Duplicated reads were marked and removed using MarkDuplicates v4.1.1.0, and base quality score recalibration was done using GATK ApplyBQSR v4.1.1.0. Indel realignment was not performed as per current GATK best practices (June 2021) it is not recommended when performing variant calling with Mutect2.

2.2.3 Exome DNA sequencing (Vancouver cohort)

To extract DNA from FFPE tumor samples, two FFPE cores per case were used to prepare hematoxylin and eosin-stained slides and a certified pathologist reviewed them for tumor content. Macro-dissection on tumor regions was done to enrich for tumor content. For paraffin removal, tissue re-hydration, tissue digestion and DNA extraction a M220 instrument (Covaris) and a truXTRAC FFPE DNA microTUBE Kit (Covaris) was used, and to quantify DNA the Qubit 2.0 fluorometer (Life Technologies) was used. The DNA quality was assessed by a multiplex-PCR assay (van Beers E.H. et al, British Journal of Cancer (2006) 94, 333-337), with usable DNA in samples with >400-bp PCR products.

To sequence the samples, 150-200 bp fragments were generated by fragmenting 1 µg of genomic DNA by hydrodynamic shearing (Covaris E210). DNA fragments were ligated to Illumina barcoded adapters after end repair, cleaned up. PCR amplification was used to enrich for adapter ligated fragments, and controlled for quality (Qubit and agarose gel). The library

was enriched by liquid phase hybridization using Agilent SureSelect XT Human All Exon v6 (Agilent Technologies; 60M bp target size) following manufacturer's recommendations, and amplified by PCR using indexing primers. Captured libraries were cleaned and controlled for quality (Qubit and agarose gel). To cluster the index-coded samples and the PE100 libraries sequencing a HiSeq4000 system (Illumina) was used.

Sequencing reads were aligned against the Human Reference Genome Ghr38 using Burrows-Wheeler aligner v0.7.17-r1188. Duplicated reads were marked and removed using MarkDuplicates v2.23.8, and base quality score recalibration was done using GATK ApplyBQSR v4.1.1.0.

2.2.4 Variant calling and inference of somatic and functional variants

GATK Mutect2 (v4.1.9.0) was used to call single-nucleotide variants (SNVs) and short insertions and deletions (indels) on tumor samples using the 'Tumor-only' mode. Variants were further filtered using GATK FilterMutectCalls (v4.1.9.0), and variants with an allele frequency below 5% or an alternate number or reads below 3 were filtered out.

Because the sequencing depth between samples and genes was variable, we annotated regions with a low sequencing coverage. First, we computed the sequencing depth using Samtools (v1.9) and estimated the average coverage per-gene. For each sample, we annotated each gene from the oncoplots (i.e. Figure 1A) as '*Not available*' when the average number of gene sequencing coverage was <20 reads. As an exception to this rule, when a relevant (non-germline and pathogenic) variant and with at least 3 alternate reads was detected for a low coverage gene, the variant was subjected to downstream analysis and the '*Not available*' annotation was removed.

For downstream analysis, samples showing low coverage (average <20 reads) in >40% of the studied genes were filtered out, which affected 4 patients from the NKI cohort, and 2 patients from the Vancouver cohort. Variants were processed in R 3.6.0 using the packages VariantAnnotation v1.24.5, ComplexHeatmap v1.17.1, tidyverse 1.2.1, and ggpubr 0.2.1.

To filter out germline variants, only variants annotated as COMMON != 1, CAF <0.05, TOPMED <0.05, GNOMAD dbNSFP_gnomAD_exomes_NFE_AF <0.05. and dbNSFP_gnomAD_exomes_AF <0.05 were retrieved. To further filter out potential germline variants, we retrieved the germline DNA sequencing data collected from NABUCCO cohort 1, and identified the germline variants FOXQ1 T60P, KMT2D T4629P (mutated in >70% of the patients), which were filtered out from our list.

To retrieve functional variants, we filtered out variants annotated as introns, non-coding, synonymous, downstream gene variant, 3' UTR variant, 5' UTR variant, t' UTR premature start codon gain, sequence feature. Then, we only retrieved variants being annotated with at least one of the following annotations: Pathogenic or Likely Pathogenic on CLNSIG, High Impact, Deleterious on SIFT or FATHMM, Damaging or Probably Damaging on HVAR, or as Oncogenic or Likely Oncogenic on OncoKB. For ERBB2, only variants annotated as Gain of Function by OncoKB were retrieved and reported in the manuscript.

2.2.5 Shallow whole-genome DNA sequencing for Copy number (NKI cohort)

For shallow genome sequencing, the protocol up to PCR amplification was analogous to the one indicated in 'DNA sequencing'. The uniquely indexed samples were mixed together by equimolar pooling. Different pools of samples were prepared, consisting of 5 batches of 12 (1 pool, 1 lane), and one batch of 34 samples (1 pool, run over 3 lanes). The pools were analyzed on the Agilent Technologies 2100 Bioanalyzer. The Pools were diluted to 10 nM, and measured on the qPCR. Then the pool was sequenced on an Illlumina HiSeq2500 machine, where each pool was in one lane of a single-end 65 base pairs (bp) run, following the manufacturer's instructions.

The low coverage whole genome samples were aligned to GRCh38.78 using the Burrows-Wheeler Aligner mem algorithm (bwa version 0.7.17). For every sample, and on segments of 20 kilobases (kb) on the genome, the mapping quality read counts were rated and tiled for 65 base pairs against a similar mapping of all known sequences for genome version GRCh38.78. A non-linear loess fit of mappabilities over 0.8 on autosomes was used to correct per 20kb for local GC effects. Then a scaling to the slope of a linear fit of the reference mappabilities after GC correction was done on the reference values. where the intercept was forced at the origin. We filtered out the ratios of corrected sample counts and the reference values left out bins with a mappability below 0.2 or overlapped with ENCODE blacklisted regions (ENCODE Project Consortium, Nature, 2012). The pipeline used in the count and log2ratio corrections is available at https://github.com/NKI-GCF/SeqCNV. For male patients, the log2 ratio was increased by one in genes from Chromosome X. Copy number ratios (CNR) at a gene level (CNR-gene) were estimated using a weighted average copy number ratio per gene. We defined deletions as $\log_2$(CNR-gene) <−0.7, and amplifications as $\log_2$(CNR-gene) >1.

2.2.6 TCGA cohort

Mutation data from Muscle-invasive bladder cancer patients from the TCGA cohort (n=412, Robertson, A. G., Kim, J., Al-Ahmadie, H., Bellmunt, J., Guo, G., Cherniack, A. D., …, Zwarthoff, E. C. (2017). Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. Cell, 171(3), 540-556.e25) was downloaded as provided on cBioportal on

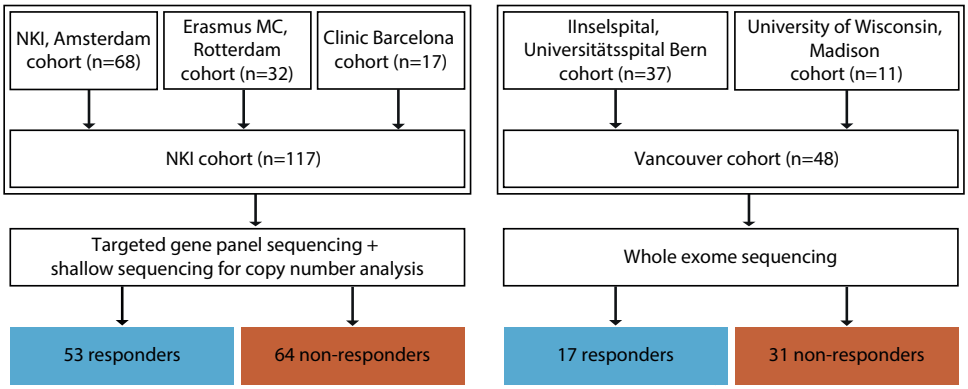
8th April 2022. Mutation data was aggregated by patient to compute mutation rates, and compared with mutation rates from our cohort by a two-sided Fisher's exact test.

2.2.7 Statistical analyses

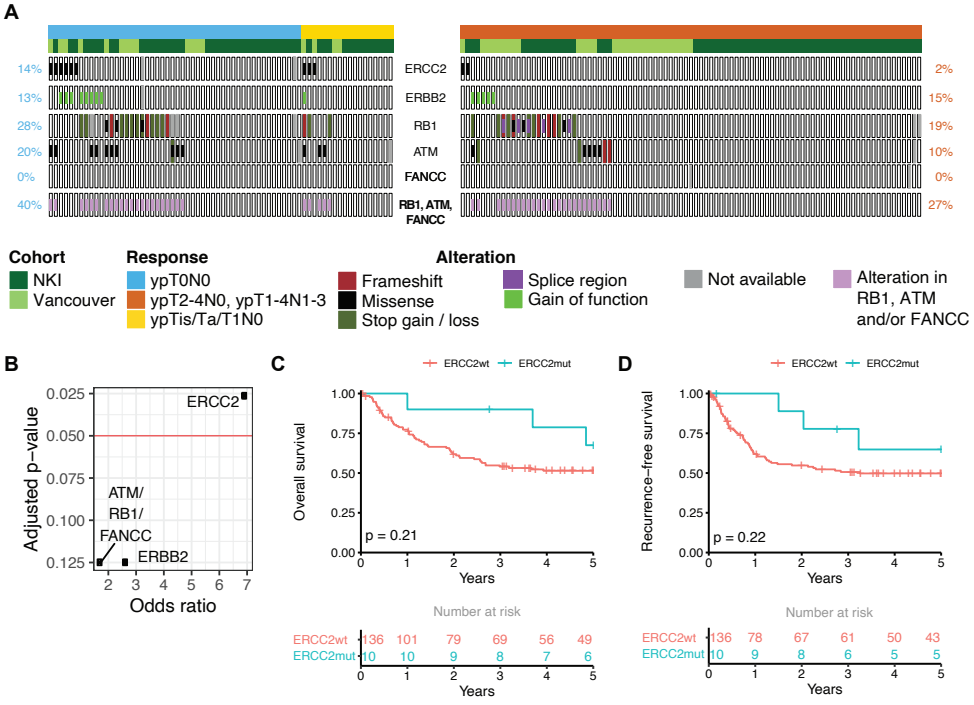
Associations between genomic mutations and clinical response were tested using a two-sided Fisher's exact test. We performed multiple hypothesis testing on our original set of 3 hypotheses (*ERCC2*, *ERBB2* and *ATM*, *RB1* or *FANCC*) using the Benjamini-Hochberg method. A 0.05 threshold for both the unadjusted p-values and the false discovery rates was used to define significance. Associations between baseline clinical characteristics and response (**Table 1**) were sought using a Fisher's exact test for binary predictors, and a t-test for numerical predictors. Unless otherwise stated, all the statistical tests were two-sided. The statical analysis was performed using R 3.6.0.

Survival analysis with time to event was performed with a Kaplan-Meier analysis, and statistical significance was tested with a log-rank test. Time to event was computed as the temporal window between the day of cystectomy and either time to death (overall survival, OS) or time to recurrence (recurrence-free survival RFS). For median follow-up estimation, reverse censoring was implemented using the reverse Kaplan-Meier method by reversing the event (death) and censoring labels. For OS and RFS analysis, censoring was implemented for patients lost to follow-up. Hazard ratio testing with a Cox proportional hazards regression model was not performed, due to the proportional hazard assumption being violated for a Cox model of the 5-year OS association with *ERCC2* mutations, in which a significant association between the Cox model residuals and time was identified (p=0.037). Analyses were performed using the R packages survival 2.44, prodlim version 2019.11.13, and survminer 0.4.6.

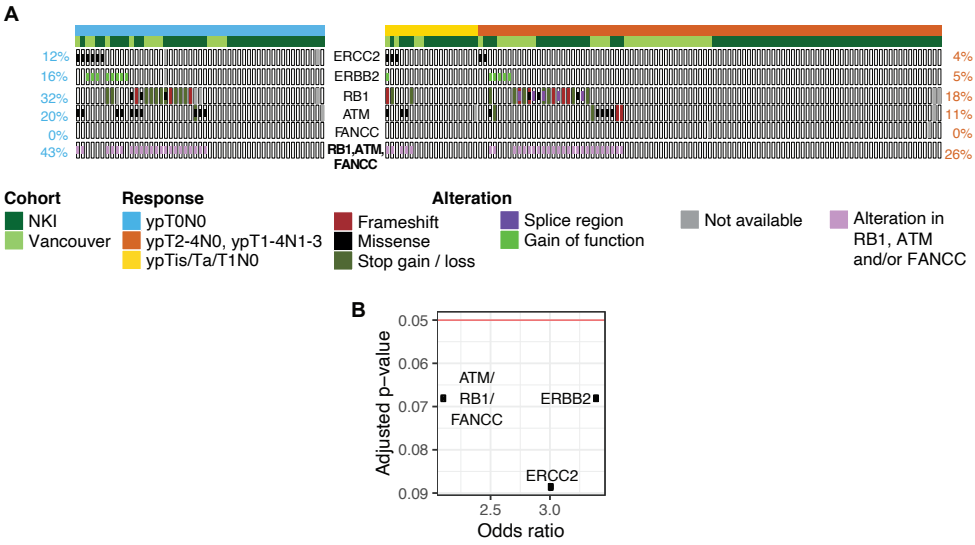
2.3 SUPPLEMENTARY FIGURES



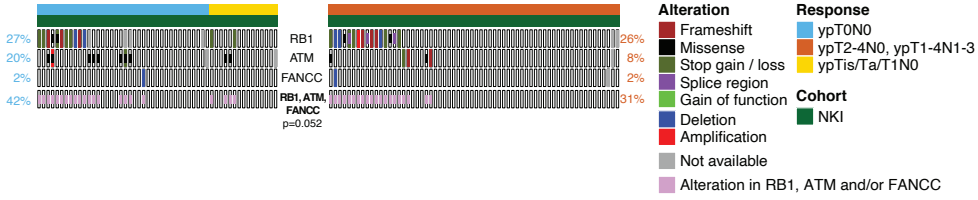
Supplementary Figure 1. Overview of patient cohorts from 5 different centers. The patients from Amsterdam (The Netherlands), Rotterdam (The Netherlands) and Barcelona (Spain) (n=117) were sequenced by the NKI using targeted DNA sequencing and shallow whole genome sequencing. Whole-exome sequencing was performed in Vancouver on tumor material from Bern (Switzerland) and Seattle (Washington) (n=48).



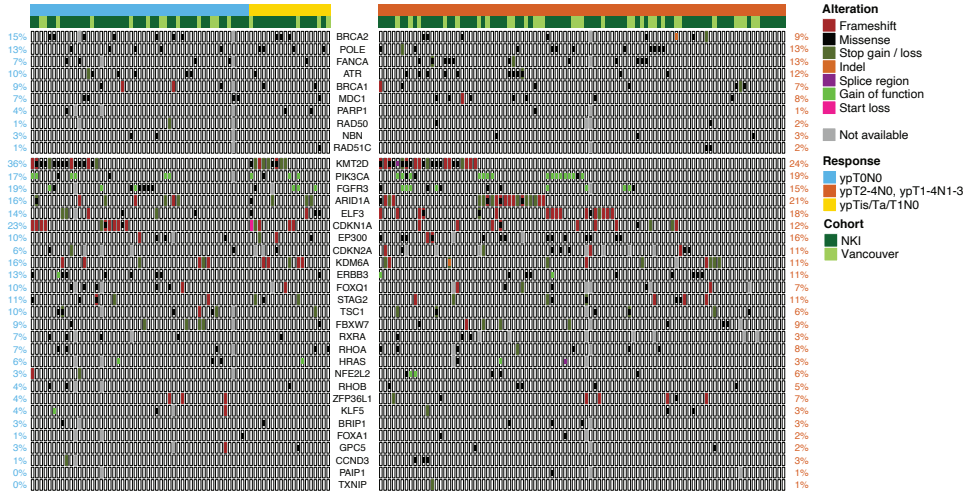
Supplementary Figure 2. Somatic mutations in the genes *ERCC2*, *ERBB2*, *ATM*, *RB1* and *FANCC* in patients with muscle-invasive bladder cancer treated with neoadjuvant chemotherapy with at least 3 NAC cycles. Six patients that received 2 cycles of chemotherapy were excluded from the analysis. **A)** Overview of relevant mutations for each individual patient. Left panel shows patients with pathological complete response (ypT0N0, light blue, n=50), and right panel shows patients with ypTis/Ta/T1N0 (yellow, n=18) and ≥ypT2N0 (orange, n=91). Percentages represent the number of patients per cohort with a relevant mutation relative to the total number of eligible patients for that specific gene for ypT0N0 (left) and >ypT0N0 (right) patients. Patients with an alteration in any one of *ATM*, *RB1* or *FANCC* are indicated in the last row. **B)** Adjusted p-value versus odds ratio for *ERCC2*, *ERBB2* and any one of *ATM*, *RB1* or *FANCC* between response groups. P-values were calculated by a two-sided Fisher's exact test, and adjusted for multiple hypothesis testing using the Benjamini-Hochberg procedure. *NKI = patients from the NKI-cohort; Not available = gene-coverage below 20 reads.*



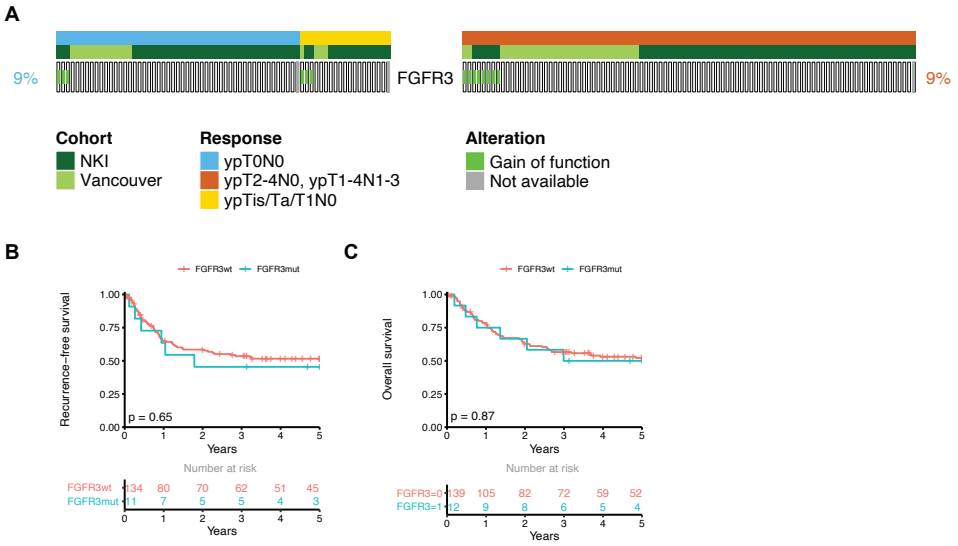
Supplementary Figure 3. Somatic mutations in the genes *ERCC2*, *ERBB2*, *ATM*, *RB1* and *FANCC* in patients with muscle-invasive bladder cancer treated with neoadjuvant chemotherapy related to complete pathological response. A) Overview of relevant mutations for each individual patient. Left panel shows patients with pathological complete response (ypT0N0, light blue, n=51), and right panel shows patients with ypTis/Ta/T1N0 (yellow, n=19) and ≥ypT2N0 (orange, n=95). Percentages represent the number of patients per cohort with a relevant mutation relative to the total number of eligible patients for that specific gene for ypT0N0 (left) and >ypT0N0 (right) patients. Patients with an alteration in any one of *ATM*, *RB1* or *FANCC* are indicated in the last row. **B)** Adjusted p-value versus odds ratio for *ERCC2*, *ERBB2* and any one of *ATM*, *RB1* or *FANCC* between response groups. P-values were calculated by a two-sided Fisher's exact test, and adjusted for multiple hypothesis testing using the Benjamini-Hochberg procedure. *NKI = patients from the NKI-cohort; Not available = gene-coverage below 20 reads.*



Supplementary Figure 4. Somatic mutations and copy number alterations in *ERCC2*, *ERBB2*, *ATM*, *RB1* and *FANCC* in patients from the NKI cohort. Overview of relevant mutations and copy number alterations for all patients for whom mutation data and copy number alterations were available (NKI cohort, n=117). Left panel shows responders (ypT0/Tis/Ta/T1N0), and the right panel shows non-responders. Percentages represent the number of patients with a relevant alteration relative to the number of eligible patients for that specific gene for responders (left) and non-responder (right) patients. Patients with an alterations in any one of *ATM*, *RB1* or *FANCC* are shown separately in the last row. *NKI = patients from the NKI-cohort; Not available = gene-coverage below 20 reads.*



Supplementary Figure 5. Somatic mutations in oncogenic driver genes and genes frequently-mutated in muscle-invasive bladder cancer. A) Overview of somatic mutations in driver genes and genes frequently altered in muscle-invasive bladder cancer in responders (ypT0/Tis/Ta/T1N0, left) and in non-responders (right). **B)** Adjusted p-value versus odds ratio for the alterations of each gene between response groups. P-values were calculated by a two-sided Fisher's exact test, and adjusted for multiple hypothesis testing using the Benjamini-Hochberg procedure. *NKI = patients from the NKI-cohort; Not available = gene-coverage below 20 reads.*



Supplementary Figure 6. Somatic mutations in the *FGFR3* gene in patients with muscle-invasive bladder cancer treated with neoadjuvant chemotherapy. A) Overview of relevant *FGFR3* mutations for each patient. Left panel shows patients with ypT0N0 (light blue, n=51) or ypTis/Ta/T1N0 (yellow, n=19) after neoadjuvant chemotherapy (responders), and right panel shows non-responders (orange, n=95). Percentages represent the number of patients with a relevant mutation relative to the total number of eligible patients for that specific gene for responders (left) and non-responders (right). **B)** Five-year overall survival (OS) for patients with (blue) and without (red) somatic mutations in *FGFR3*. The p-value indicates statistical significance by a log-rank test. **C)** 5-year recurrence-free survival (RFS) for patients with (blue) and without (red) somatic mutations in *FGFR3*. The p-value inside the survival plot indicates statistical significance by a log-rank test. *NKI = patients from the NKI-cohort; Not available = gene-coverage below 20 reads.*

2.4 SUPPLEMENTARY TABLES

Supplementary Table 1. Baseline characteristics and response of 165 patients with muscle-invasive bladder cancer treated with neoadjuvant chemotherapy and radical cystectomy and comparison between cohorts.

	NKI Cohort	Vancouver	Significance NKI vs Vancouver[a]
Number of patients	117	48	-
Median age in years (IQR)	71.0 (61.0, 75.0)	62.1 (57.3, 72.0)	**0.003**
Male sex (% of pts)	78 (67%)	39 (81%)	0.09
cT-stage (% of pts)			**0.002**
cT1	1 (1%)	0 (0%)	
cT2	46 (39%)	6 (13%)	
cT3	47 (40%)	27 (56%)	
cT4	23 (20%)	15 (31%)	
cN-stage (% of pts)			0.09
cN0	71 (61%)	22 (46%)	
cN+	46 (39%)	26 (54%)	
Chemotherapy regimen (% of pts)			**0.03**
Cis/Gem	81 (69%)	42 (88%)	
MVAC	34 (29%)	6 (13%)	
CMV	2 (2%)	0 (0%)	
Chemotherapy cycles received (% of pts)			0.15
2	4 (3%)	2 (4%)	
3	23 (20%)	13 (27%)	
4	85 (73%)	27 (56%)	
>4	5 (4%)	6 (13%)	

Supplementary Table 1. *(Continued)*

	NKI Cohort	Vancouver	Significance NKI vs Vancouver[a]
Pathological response (% of pts)			0.5
ypT0N0 (complete response)	38 (32%)	13 (27%)	
ypTis/Ta/T1N0	15 (13%)	4 (8%)	
≥ypT2N0 (non-response)	64 (55%)	31 (65%)	
Nodal response (% of pts)			0.7
ypN0	81 (69%)	32 (67%)	
≥ypN1	34 (29%)	16 (33%)	
ypN unavailable	2 (2%)	0 (0%)	

[a] Fisher's exact test for binary predictors, t-test for numerical predictors. All the statistical tests were two-sided. No adjustments were made for multiple hypothesis testing. Significant associations are highlighted on bold.

Cis/Gem = cisplatin + gemcitabine; MVAC = methotrexate + vinblastine + docorubicin + cisplatin; CMV = cisplatin + methotrexate + vinblastine; IQR = Interquartile range; pts = patients.

Supplementary Table 2. Comparison of mutation rates against TCGA Bladder Urothelial Carcinoma (n=412).

Gene	NKI/Vancouver mutation rate (%)	TCGA mutation rate (%)	Significance TCGA vs NKI/Vancouver[a]	Significance (false-discovery rate)[a]
KDM6A	13	33	0.0005	0.02
FANCA	10	4	0.01	0.16
CDKN1A	16	9	0.01	0.16
POLE	13	7	0.02	0.16
TXNIP	1	5	0.02	0.16
FOXQ1	8	3	0.02	0.17
MDC1	8	4	0.06	0.4

Supplementary Table 2. *(Continued)*

Gene	NKI/Vancouver mutation rate (%)	TCGA mutation rate (%)	Significance TCGA vs NKI/Vancouver[a]	Significance (false-discovery rate)[a]
ARID1A	19	30	0.15	0.6
ELF3	16	13	0.18	0.6
ATR	11	8	0.19	0.6
FANCC	0	1	0.19	0.6
PAIP1	1	3	0.19	0.6
RHOA	7	5	0.22	0.6
RB1	22	19	0.23	0.6
BRCA1	8	5	0.24	0.6
ERBB2	9	13	0.24	0.6
CDKN2A	9	7	0.28	0.7
CCND3	2	1	0.28	0.7
ERCC2	7	10	0.33	0.7
PIK3CA	18	26	0.36	0.8
KLF5	4	6	0.41	0.8
STAG2	11	15	0.41	0.8
FOXA1	2	3	0.42	0.8
FGFR3	17	15	0.44	0.8
GPC5	2	4	0.46	0.8
BRIP1	3	5	0.50	0.8
BRCA2	11	10	0.54	0.8
NFE2L2	5	7	0.56	0.9
RHOB	5	7	0.69	0.9
EP300	14	17	0.70	0.9

Supplementary Table 2. *(Continued)*

Gene	NKI/Vancouver mutation rate (%)	TCGA mutation rate (%)	Significance TCGA vs NKI/Vancouver[a]	Significance (false-discovery rate)[a]
RAD51C	2	1	0.72	0.9
FBXW7	9	8	0.73	0.9
ERBB3	12	11	0.77	0.9
RAD50	2	3	0.77	0.9
NBN	3	3	0.78	0.9
RXRA	5	6	0.84	1
KMT2D	29	34	0.92	1
ATM	14	15	1	1
HRAS	4	4	1	1
PARP1	2	3	1	1
TSC1	8	9	1	1
ZFP36L1	6	6	1	1

[a] Fisher's exact test for frequency of somatic mutations between TCGA and NKI/Vancouver cohorts. P-values were adjusted for multiple hypothesis testing by the Benjamini-Hochberg method.

Supplementary Table 3. Baseline characteristics and response rates of 163 patients with muscle-invasive bladder cancer treated with neoadjuvant chemotherapy and radical cystectomy and comparison between *ERCC2* mutation status.

	ERCC2$_{mut}$	*ERCC2*$_{wt}$	Significance *ERCC2*$_{mut}$ vs *ERCC2*$_{wt}$ [a]
Number of patients	11	152	
Median age in years (IQR)	57.0 (53.2, 74.0)	69.0 (60.9, 75.0)	0.12
Male sex (% of pts)	8 (72%)	107 (70%)	1
cT-stage (% of pts)			0.7
cT1	0 (0%)	1 (1%)	
cT2	2 (18%)	49 (32%)	
cT3	6 (55%)	67 (44%)	
cT4	3 (27%)	35 (23%)	
cN-stage (% of pts)			0.4
cN0	8 (73%)	84 (55%)	
cN+	3 (27%)	68 (45%)	
Chemotherapy regimen (% of pts)			0.8
Cis/Gem	9 (82%)	112 (74%)	
MVAC	2 (18%)	38 (25%)	
CMV	0 (0%)	2 (1%)	
Chemotherapy cycles received (% of pts)			0.5
2	0 (0%)	6 (4%)	
3	4 (36%)	32 (21%)	
4	6 (54%)	104 (68%)	
>4	1 (9%)	10 (7%)	
Pathological response (% of pts)			0.009
ypT0N0 (complete response)	6 (54%)	43 (28%)	
ypTis/Ta/T1N0	3 (27%)	16 (11%)	
≥ypT2N0 (non-response)	2 (18%)	93 (61%)	

[a] Fisher's exact test for binary predictors. t-test for numerical predictors. All the statistical tests were two-sided. No adjustments were made for multiple hypothesis testing. Significant associations are highlighted on bold.

Cis/Gem = cisplatin + gemcitabine; MVAC = methotrexate + vinblastine + docorubicin + cisplatin; CMV = cisplatin + methotrexate + vinblastine; mut = mutant; wt = wild-type; IQR = Interquartile range; pts = patients.

ACKNOWLEDGMENTS

We acknowledge the Genomics Core Facility (Roel Kluin), the Core Facility Molecular Pathology and Biobanking, and the Research High-Performance Computing facility, all at the Netherlands Cancer Institute; the sequencing facility at the Vancouver Prostate Center, and all the clinical teams involved in the 5-center patient cohort.

REFERENCES

1. Witjes, J.A., et al., European Association of Urology Guidelines on Muscle-invasive and Metastatic Bladder Cancer: Summary of the 2020 Guidelines. Eur Urol, 2021. **79**(1): p. 82-104.
2. Rosenblatt, R., et al., Pathologic downstaging is a surrogate marker for efficacy and increased survival following neoadjuvant chemotherapy and radical cystectomy for muscle-invasive urothelial bladder cancer. Eur Urol, 2012. **61**(6): p. 1229-38.
3. Van Allen, E.M., et al., Somatic ERCC2 mutations correlate with cisplatin sensitivity in muscle-invasive urothelial carcinoma. Cancer Discov, 2014. **4**(10): p. 1140-53.
4. Groenendijk, F.H., et al., ERBB2 Mutations Characterize a Subgroup of Muscle-invasive Bladder Cancers with Excellent Response to Neoadjuvant Chemotherapy. Eur Urol, 2016. **69**(3): p. 384-8.
5. Li, Q., et al., ERCC2 Helicase Domain Mutations Confer Nucleotide Excision Repair Deficiency and Drive Cisplatin Sensitivity in Muscle-Invasive Bladder Cancer. Clin Cancer Res, 2019. **25**(3): p. 977-988.
6. Plimack, E.R., et al., Defects in DNA Repair Genes Predict Response to Neoadjuvant Cisplatin-based Chemotherapy in Muscle-invasive Bladder Cancer. Eur Urol, 2015. **68**(6): p. 959-67.
7. Miron, B., et al., Defects in DNA Repair Genes Confer Improved Long-term Survival after Cisplatin-based Neoadjuvant Chemotherapy for Muscle-invasive Bladder Cancer. Eur Urol Oncol, 2020. **3**(4): p. 544-547.
8. Becker, R.E.N., et al., Clinical Restaging and Tumor Sequencing are Inaccurate Indicators of Response to Neoadjuvant Chemotherapy for Muscle-invasive Bladder Cancer. Eur Urol, 2021. **79**(3): p. 364-371.
9. Ravi, P., et al., Optimal pathological response after neoadjuvant chemotherapy for muscle-invasive bladder cancer: results from a global, multicentre collaboration. BJU Int, 2021. **128**(5): p. 607-614.
10. Teo, M.Y., et al., Fibroblast Growth Factor Receptor 3 Alteration Status is Associated with Differential Sensitivity to Platinum-based Chemotherapy in Locally Advanced and Metastatic Urothelial Carcinoma. Eur Urol, 2020. **78**(6): p. 907-915.

3

The Tumor Immune Landscape and Architecture of Tertiary Lymphoid Structures in Urothelial Cancer

Nick van Dijk[1,*], Alberto Gil-Jimenez[2,3,*], Karina Silina[4], Maurits L. van Montfoort[5], Sarah Einerhand[1], Lars Jonkman[6], Charlotte S. Voskuilen[1], Dennis Peters[7], Joyce Sanders[5], Yoni Lubeck[5], Annegien Broeks[7], Erik Hooijberg[5], Daniel J. Vis[2,3], Maries van den Broek[4], Lodewyk F. A. Wessels[2,3,8], Bas W. G. van Rhijn[9,10], Michiel S. van der Heijden[1,2,#].

Parts of this chapter have been published in Frontiers in Immunology (2021): https://www.frontiersin.org/journals/immunology/articles/10.3389/fimmu.2021.793964/full

* These authors contributed equally and share first authorship
# Corresponding author
[1] Department of Medical Oncology, The Netherlands Cancer Institute, Amsterdam, the Netherlands
[2] Department of Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam, the Netherlands
[3] Oncode Institute, Utrecht, the Netherlands
[4] Institute of Experimental Immunology, University of Zurich, Zurich, Switzerland
[5] Department of Pathology, The Netherlands Cancer Institute, Amsterdam, the Netherlands
[6] Department of Medical Oncology, Maastricht University Medical Center, Maastricht, the Netherlands
[7] Core Facility Molecular Pathology & Biobanking, The Netherlands Cancer Institute, Amsterdam, the Netherlands
[8] Faculty of EEMCS, Delft University of Technology, Delft, the Netherlands
[9] Department of Urology, The Netherlands Cancer Institute, Amsterdam, the Netherlands
[10] Department of Urology, Caritas St. Josef Medical Center, University of Regensburg, Regensburg, Germany

Keywords: Urothelial Cancer; Bladder Cancer; Immunotherapy; Tumor Microenvironment; Tertiary Lymphoid Structures; Multiplex Immunofluorescence

MANUSCRIPT CONTRIBUTION TO THE FIELD

Urothelial cancer (UC) is an aggressive disease that has a high chance of metastatic dissemination and death. Recent advances in immunotherapy, targeting inhibitory checkpoints, have changed the treatment paradigm in UC. Still, many patients do not respond to immunotherapy, and discriminative biomarkers for response are lacking. A detailed understanding of the UC immune landscape is needed to improve treatment. In this study, we comprehensively characterize the UC immune landscape with and without immunotherapy pretreatment, using computational analysis of immune cell stainings. We explore the distribution of immune cells in the main tumor compartments and assess markers of T-cell activation or exhaustion. Moreover, we provide an in-depth characterization of the Tertiary Lymphoid Structure (TLS) immune contexture to study associations with recurrence and immunotherapy response. In addition, we propose a hypothesis on the existence of a TLS subgroup located in superficial bladder tissue that may be unrelated to anti-tumor immunity and could cloud the interpretation of tumor-associated TLS studies. This study could form the foundation for further research into the UC tumor-immune landscape and strategies to improve immunotherapy response.

ABSTRACT

Candidate immune biomarkers have been proposed for predicting response to immunotherapy in urothelial cancer (UC). Yet, these biomarkers are imperfect and lack predictive power. A comprehensive overview of the tumor immune contexture, including Tertiary Lymphoid Structures (TLS), is needed to better understand the immunotherapy response in UC. We analyzed tumor sections by quantitative multiplex immunofluorescence to characterize immune cell subsets in various tumor compartments in tumors without pretreatment and tumors exposed to preoperative anti-PD1/CTLA-4 checkpoint inhibitors (NABUCCO trial). Pronounced immune cell presence was found in UC invasive margins compared to tumor and stroma regions. $CD8^+PD1^+$ T-cells were present in UC, particularly following immunotherapy. The cellular composition of TLS was assessed by multiplex immunofluorescence (CD3, CD8, FoxP3, CD68, CD20, PanCK, DAPI) to explore specific TLS clusters based on varying immune subset densities. Using a k-means clustering algorithm, we found five distinct cellular composition clusters. Tumors unresponsive to anti-PD-1/CTLA-4 immunotherapy showed enrichment of a $FoxP3^+$ T-cell-low TLS cluster after treatment. Additionally, cluster 5 (macrophage low) TLS were significantly higher after pre-operative immunotherapy, compared to untreated tumors. We also compared the immune cell composition and maturation stages between superficial (submucosal) and deeper TLS, revealing that superficial TLS had more pronounced T-helper cells and enrichment of early TLS than TLS located in deeper tissue. Furthermore, superficial TLS displayed a lower fraction of secondary follicle like TLS than deeper TLS. Taken together, our results provide a detailed quantitative overview of the tumor immune landscape in UC, which can provide a basis for further studies.

3.1 INTRODUCTION

Muscle-invasive urothelial cancer (UC) is an aggressive disease with limited treatment options that originates in the bladder and parts of the urinary tract. Although UC can be cured by resection of the bladder (cystectomy), recurrence rates are high and 5-year survival is only 60-70% for pT2N0 tumors, and even worse for high-risk patients having pT3-4aN0 (40-50%) or pTxN+ (10-35%) at cystectomy. Immune checkpoint inhibitors (ICIs) have changed the treatment paradigm in metastatic urothelial cancer. Currently, ICIs have been approved for the first-line and second-line treatment (1–5), and are being tested in the adjuvant and preoperative setting. In the PURE-01 trial[6] and ABACUS trial[7], preoperative pembrolizumab (anti-PD-1) and atezolumab (anti-PD-L1) were clinically tested in patients diagnosed with cT2-4N0 UC, respectively. These trials revealed promising pathological complete response (pCR) rates upon treatment with neo-adjuvant pembrolizumab and atezolizumab. However, pCR to ICI monotherapy was primarily found in patients having less extensive disease (cT2N0), whereas patients with more extensive disease (cT3-4N0) or loco-regional lymph node involvement (T2-4N+) showed only limited pCR to anti-PD1 or anti-PD-L1. Recent clinical studies testing combination strategies targeting PD-1/PD-L1 plus CTLA-4 in the metastatic setting found higher response rates than in trials testing anti-PD1 or anti-PD-L1 alone[8,9]. In the NABUCCO trial[10], preoperative ipilimumab plus nivolumab was tested in high-risk patients having locoregionally-advanced UC (cT3-4N0/cT2-4N1-3) without distant metastases. Histopathological examination showed that 58% of patients in NABUCCO had no remaining invasive disease (pT0 or CIS/pTa) after ipilimumab plus nivolumab[10]. A study testing preoperative tremelimumab plus durvalumab in cT2-4N0 UC observed a pCR in 37.5% (pT0 or CIS) of patients having surgery, whereas the pCR rate was 31.7% in all patients analyzed[8].

Associations between ICI response and candidate biomarkers, such as PD-L1 immunohistochemistry and tumor mutational burden (TMB), have been observed in metastatic UC. These biomarkers are currently imperfect and lack sufficient predictive power for clinical utility[11,12]. In addition, comparison of biomarker findings across trials is complicated by variability in biomarker assays (i.e. PD-L1 assessment) and heterogeneity in tumor tissue used to assess biomarkers. In the preoperative setting, the pCR rate to pembrolizumab in the PURE-01 study was high in TMB-high and PD-L1-high (PD-L1 >10%; tumor plus immune cells combined) tumors[6], whereas no significant associations were found for TMB-high and PD-L1-high (PD-L1 >5% of immune cells) subgroups in anti-PD-L1 treated patients in ABACUS[7]. Both studies found that baseline pre-existing $CD8^+$ T-cell immunity based on high CD8 presence and interferfon-γ signaling was associated with pCR to ICI monotherapy. Qualification of immune phenotypes by CD8 immunohistochemistry showed that "immune desert" tumors in ABACUS were unresponsive to ICI[7]. In sharp contrast, the clinical response

to combination ICI in NABUCCO was independent of baseline CD8+ T-cell density by multiplex immunofluoresence and inflammatory signatures such as interferon-gamma, tumor inflammation and T-cell effector signatures[10]. Similarly, baseline pre-existing CD8+ T-cell immunity did not differ between responders and non-responders to neo-adjuvant tremelimumab plus durvalumab[13], suggesting that the addition of anti-CTLA4 can induce responses in immunologically "cold" tumors.

Tertiary lymphoid structures (TLS) are ectopic lymph node formations that share functional features such as antigen presentation and B-cell activation with secondary lymphoid organs. TLS emerge upon chronic inflammatory stimuli in non-lymphoid tissues and can also be found in the tumor micro-environment. In an analysis of the presence of TLS, responders to tremelimumab plus durvalumab showed higher baseline TLS and B-cell abundance than non-pCR tumors. Intriguingly, baseline TLS and B-cell abundance did not differ between responders and non-responders in NABUCCO. However, both studies found that responders to combination ICI showed a higher TLS abundance in post-treatment tissue than non-responders[10,13]. Thus, conflicting results on baseline candidate biomarkers for immunotherapy response were found between comparable studies. The complex interplay between immune cells in the UC tumor-immune microenvironment and TLS is still poorly understood, hampering the discovery and development of novel cancer immunotherapy as well as predictive biomarkers for immunotherapy response, underscoring the urgent need to better characterize the tumor immune landscape in UC.

In this study, we employ quantitative multiplex immunofluorescence to assess the UC tumor-immune contexture in untreated and immunotherapy-treated tumors. We first provide a general overview of the UC tumor-immune microenvironment, followed by a more detailed assessment of the TLS immune composition in untreated and immunotherapy-treated tumors.

3.2 RESULTS

3.2.1 Untreated urothelial cancer demonstrates heterogeneous immune cell infiltration

To examine the UC immune context, we analyzed immune cell infiltration by multiplex immunofluorescence (IF) on whole-slide cystectomy tissue sections from untreated (n=32, **Table 1**) and ipilimumab (anti- CTLA-4) plus nivolumab (anti-PD1) treated (n=24, **Table 2**) UC patient cohorts (**Fig 1A)**. In the current study, cystectomy specimens obtained from NABUCCO are analyzed, while we previously[10] reported CD8+ and CD20+ immune cell presence in pretreatment biopsies. Additionally, we segmented tumor areas into various regions of interest. Our antibody panel allowed the quantitation of immune

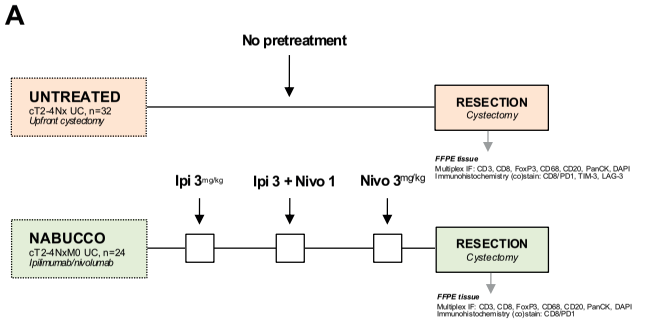
cells actively involved in anti-tumor immunity and response, such as B-cells ($CD20^+$), macrophages ($CD68^+$) and distinct $CD3^+$ T-cell populations. $CD3^+$ T-cell populations were further specified by expression of CD8 or FoxP3, resulting in CD8 T-cells ($CD3^+CD8^+$), FoxP3 T-cells ($CD3^+FoxP3^+$) and $CD4^+$ T-cells ($CD3^+CD8^-FoxP3^-$), a non-$CD8^+$/$FoxP3^+$ T-cell population which is likely to involve primarily CD4 T-cells. $CD3^+FoxP3^-CD8^-$ was thus used as an approximation of $CD4^+$ T-cells to make the manuscript easier to read. CD4 IF was not used in our multiplex panel given the expression of CD4 on other immune cells (including macrophages and dendritic cells) when using CD4 antibodies in our pilot studies. Immune cells were separately quantified for tumor and stroma areas within the central tumor and square grids were computed for spatial sampling to assesses heterogeneity of immune subsets within tumors (**Fig 1B, Supplementary Methods 1)**. We additionally quantified immune cell abundance in the tumor margin and TLS. The tumor margin was annotated from the outermost edge of the invasive tumor, with an extend of 250µm **(Supplementary Methods 1)**. To promote readability, immune cell labels and not markers are reported throughout the results.

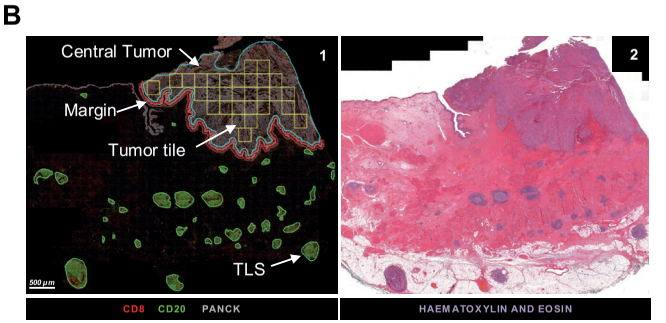
Table 1. Untreated cohort characteristics.

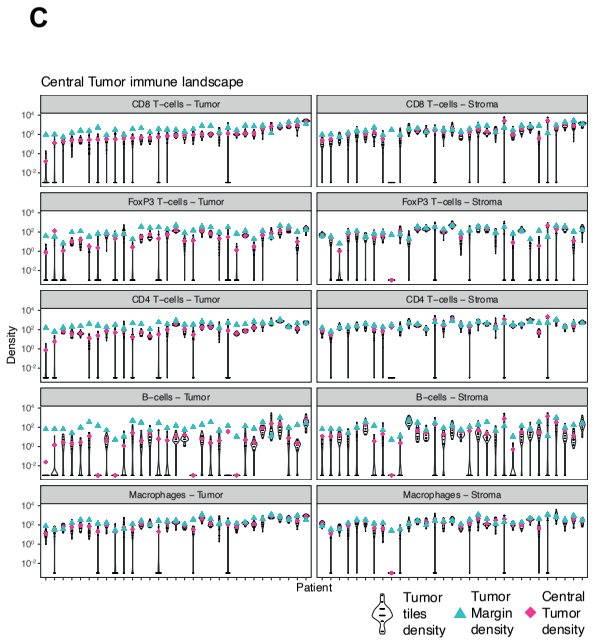
Baseline characteristics	Total (n=31)
Male sex, n (%)	24 (77%)
Median age – years [range]	64.79 [45.7, 78.7]
Pathological T stage, (%)	
pT1-4/pTis/pTaN0M0	20 (65%)
pT3-4N1-2M0	11 (35%)
Histology, (%)	
Urothelial Carcinoma	29 (94%)
Urothelial Carcinoma and Small cell carcinoma	1 (3%)
Urothelial Carcinoma and Squamous differentiation	1 (3%)
Adjuvant treatment, (%)	
No adjuvant treatment	25 (81%)
Adjuvant chemotherapy	2 (6%)
Adjuvant radiotherapy	3 (10%)
Adjuvant chemotherapy and adjuvant radiotherapy	1 (3%)

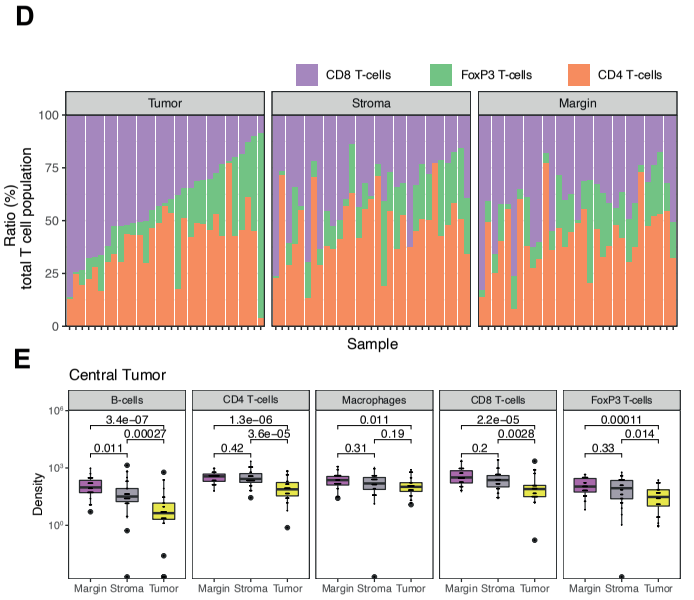
Table 2. Ipilimumab plus nivolumab treated cohort (NABUCCO Cohort 1) characteristics.

Study population characteristics	Total (n=24)
Male sex, n (%)	18 (75%)
Median age – years [range]	65 [50, 81]
Baseline clinical T stage, (%)	
cT3-4N0M0	14 (58%)
cT3-4N1	5 (21%)
cT2-3N2M0	5 (21%)
Post-treatment clinical stage, (%)	
ypT0/pTa/pTisN0M0/Mx	14 (58%)
ypT2-3N0M0	2 (8.5%)
ypT0-4N1M0	6 (25%)
ypT3N2-3M0	2 (8.5%)
Immunotherapy cycles, (%)	
2	6 (25%)
6	18 (75%)


A
No pretreatment
UNTREATED
cT2-4Nx UC, n=32
Upfront cystectomy
RESECTION
Cystectomy
FFPE tissue
Multiplex IF: CD3, CD8, FoxP3, CD68, CD20, PanCK, DAPI
Immunohistochemistry (co)stain: CD8/PD1, TIM-3, LAG-3
Ipi 3mg/kg
Ipi 3 + Nivo 1
Nivo 3mg/kg
NABUCCO
cT2-4NxM0 UC, n=24
Ipilimumab/nivolumab
RESECTION
Cystectomy
FFPE tissue
Multiplex IF: CD3, CD8, FoxP3, CD68, CD20, PanCK, DAPI
Immunohistochemistry (co)stain: CD8/PD1


B
Central Tumor
Margin
Tumor tile
TLS
1
2
500 µm
CD8 CD20 PANCK
HAEMATOXYLIN AND EOSIN


C
Central Tumor immune landscape
CD8 T−cells − Tumor
CD8 T−cells − Stroma
FoxP3 T−cells − Tumor
FoxP3 T−cells − Stroma
CD4 T−cells − Tumor
CD4 T−cells − Stroma
B−cells − Tumor
B−cells − Stroma
Macrophages − Tumor
Macrophages − Stroma
Density
Patient
Tumor tiles density
Tumor Margin density
Central Tumor density


D
CD8 T-cells
FoxP3 T-cells
CD4 T-cells
Tumor
Stroma
Margin
Ratio (%) total T cell population
Sample
E
Central Tumor
B-cells
CD4 T-cells
Macrophages
CD8 T-cells
FoxP3 T-cells
3.4e−07
0.00027
0.011
1.3e−06
3.6e−05
0.42
0.011
0.19
0.31
2.2e−05
0.0028
0.2
0.00011
0.014
0.33
Density
Margin Stroma Tumor

F

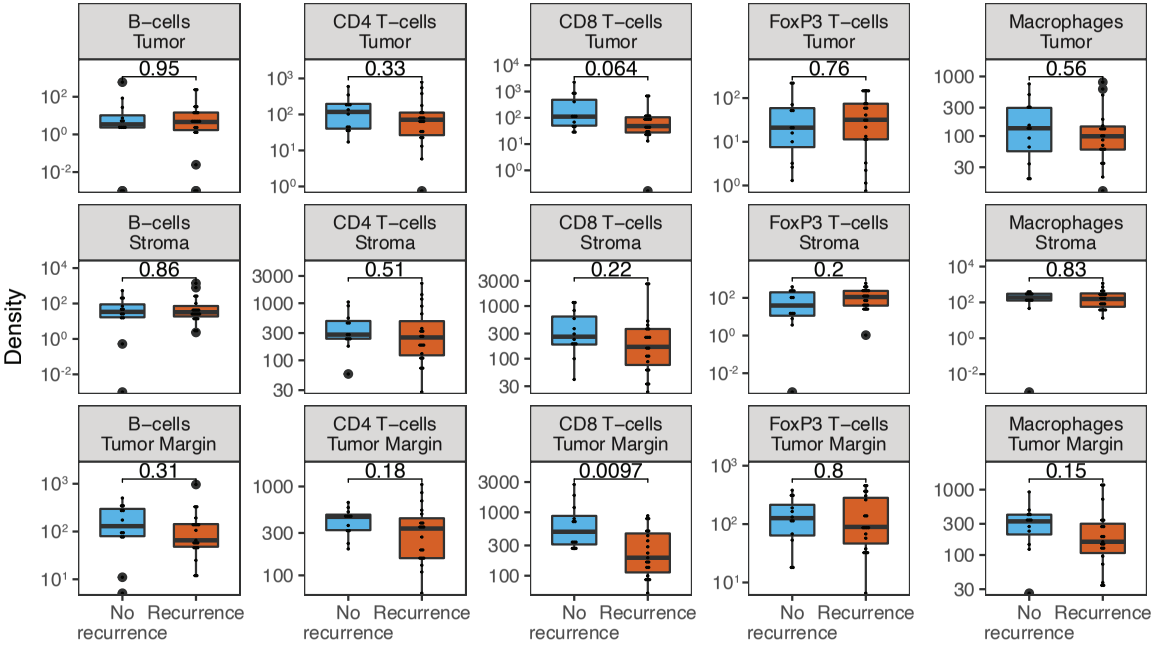


Figure 1. Untreated urothelial cancer demonstrates heterogeneous immune cell infiltration

A. Cohort interventions and timepoints of tissue collection for biomarker analysis. **B.** *1)* Example of annotated regions of interest in untreated urothelial cancer analyzed by multiplex immunofluorescence and HALO image analysis, involving annotations of central tumor (blue), central tumor tiles (yellow, n=30 tiles per slide), tumor margin (red, 250 micrometers diameter and tertiary lymphoid structures (green). Central tumor area can be distinguished in tumor and stroma area by employing and training a tissue classifier. *2)* Corresponding H&E slides were consulted to support annotation of regions of interest. **C.** Intratumoral and stroma immune subset densities per mm^2 within tumor tiles (violin plots, n=30 tiles per sample), whole tumor (pink**)** and tumor margin (cyan**)** for the untreated UC cohort (n=31). The median immune subset densities and distribution across tumor tiles were analysed by quantitative multiplex immunofluorescence. Samples were sorted by intratumoral CD8 T-cell central tumor density. **D.** Relative abundance of T-cell subsets in the total T-cell population in central tumor tissue classes and tumor margin by multiplex IF (n=31). Samples were sorted

by CD8 T-cell ratio. **E.** Immune subset densities per mm^2 for tumor tissue regions in the untreated UC cohort (n=31). **F.** Intratumoral, stroma and tumor margin immune subset densities per mm^2 for the combined untreated UC cohort (n=31) between recurrence (n=19) and non-recurrence (n=12) groups. The boxplots from the panels display the median and 25th and 75th percentiles, and the whiskers expand from the hinge to the largest value not exceeding the hinge 1.5×Interquantile range. Unless otherwise stated, a two-sided Mann-Whitney test was used for the comparison between distributions. The p-value is presented in-between boxplots. No adjustments were implemented for multiple comparisons.

Abbreviations: IF: Immuno-fluorescence, FFPE: Formalin-fixed paraffin-embedded tissue, Ipi: Ipilimumab, Nivo: Nivolumab, TLS: Tertiary lymphoid structure.

We first examined immune cell infiltration by multiplex IF for tumor and stroma areas to provide a comprehensive overview of the UC immune contexture and assess intratumor heterogeneity. We observed that the median density of immune subsets varied greatly across the untreated tumor cohort, particularly for B-cells, FoxP3 T-cells and CD8 T-cells **(Fig 1C)**. Variable intratumoral heterogeneity existed for specific immune cells upon a comparison of separate tiles in the computed square grid **(Fig 1C).** Next, we examined the relative abundance of T-cell subsets in the total T-cell population. We found that the fraction of CD4 T-cells was highly heterogeneous across tumors in the untreated cohort **(Supplementary Figure 1A)**. Further explorative analysis revealed that tumors having a low CD8 T-cell ratio demonstrated a higher proportion of FoxP3 T-cells in tumor **(Fig 1D, Supplementary Figure 1B)**. We then compared the immune cell density between central tumor regions and the tumor margin. A significantly higher presence of immune cells was found in tumor margins when compared to the tumor region ($p<0.02$ **Fig 1E**). In non-recurring tumors, the tumor margins displayed a significantly higher CD8 T-cell presence than in recurring tumors ($p=0.0097$**, Fig 1F)**, while immune cell presence in tumor and stroma did not inform clinical outcome in untreated tumors. In conclusion, the UC immune landscape is heterogeneous between tumors, and pronounced immune infiltration is found in the UC tumor margin[7,14].

3.2.2 Urothelial cancer immune phenotypes show distinct patterns of cytotoxic T-cell exclusion in the stroma and tumor margin

CD8 T-cell tumor infiltration patterns can be segregated into three immune phenotypes ("immune-inflamed", "immune-excluded" and "immune-desert") of pre-existing tumor-immunity[15]. Previous studies found that these distinct immune phenotypes harbor prognostic relevance[16] and predictive value[17,18] for an immunotherapy response, including in UC[7,14]. Currently, limited knowledge exists on the presence of distinct immune subsets beyond cytotoxic T-cells across CD8-based immune phenotypes in UC, while their presence may impact CD8 effector function and the extend of CD8 tumor-immunity. Using multiplex IF, immune phenotypes (**Fig 2A**) were classified based on CD8 T-cell density (**Supplementary Methods 1.2**) in the tumor and stroma compartment and the tumor margin in the untreated UC cohort. We first explored the distribution of tumor immune phenotypes in the untreated cohort and assessed possible correlations with prognosis for "inflamed",

"excluded" and "desert" tumors separately. In line with results in the ABACUS study[7], "immune-inflamed" (42%) tumors were most abundant in our cohort, whereas 32% and 26% of tumors exhibited the "excluded" and "desert" phenotype, respectively. The separate tumor immune phenotypes did not inform recurrence outcome in the untreated cohort (**Fig 2B**), although tumors qualified as "immune-desert" showed a high recurrence rate (87.5%, p=0.1). Next, we explored the immune composition in tumor subgroups qualified as "immune-inflamed", "immune-excluded" and "immune-desert" based on CD8-based immune phenotypes. Intratumoral immune cell densities were generally higher in "inflamed" tumors compared to "excluded" and "desert" tumors, as shown for the significantly higher macrophages compared to "desert" tumors (p=0.006. **Fig 2C**). In the stoma compartment, immune cell densities were lowest in "desert" tumors, as shown for the significantly lower CD4 T-cells when compared to "excluded" (p=0.027) and "inflamed" tumors (p=0.013) (**Fig 2D**). Interestingly, FoxP3 T-cells were an exception, as these cells were similar across immune phenotypes in absolute density and higher as a percentage of total T-cells in "desert" tumors, compared to "inflamed" tumors (p=0.037, Supplemental **Fig 2A**). Macrophage abundance in tumor margins of "inflamed" tumors was significantly higher than in "excluded" (p=0.049) and "desert" (p=0.005) tumors, (**Fig 2E**).

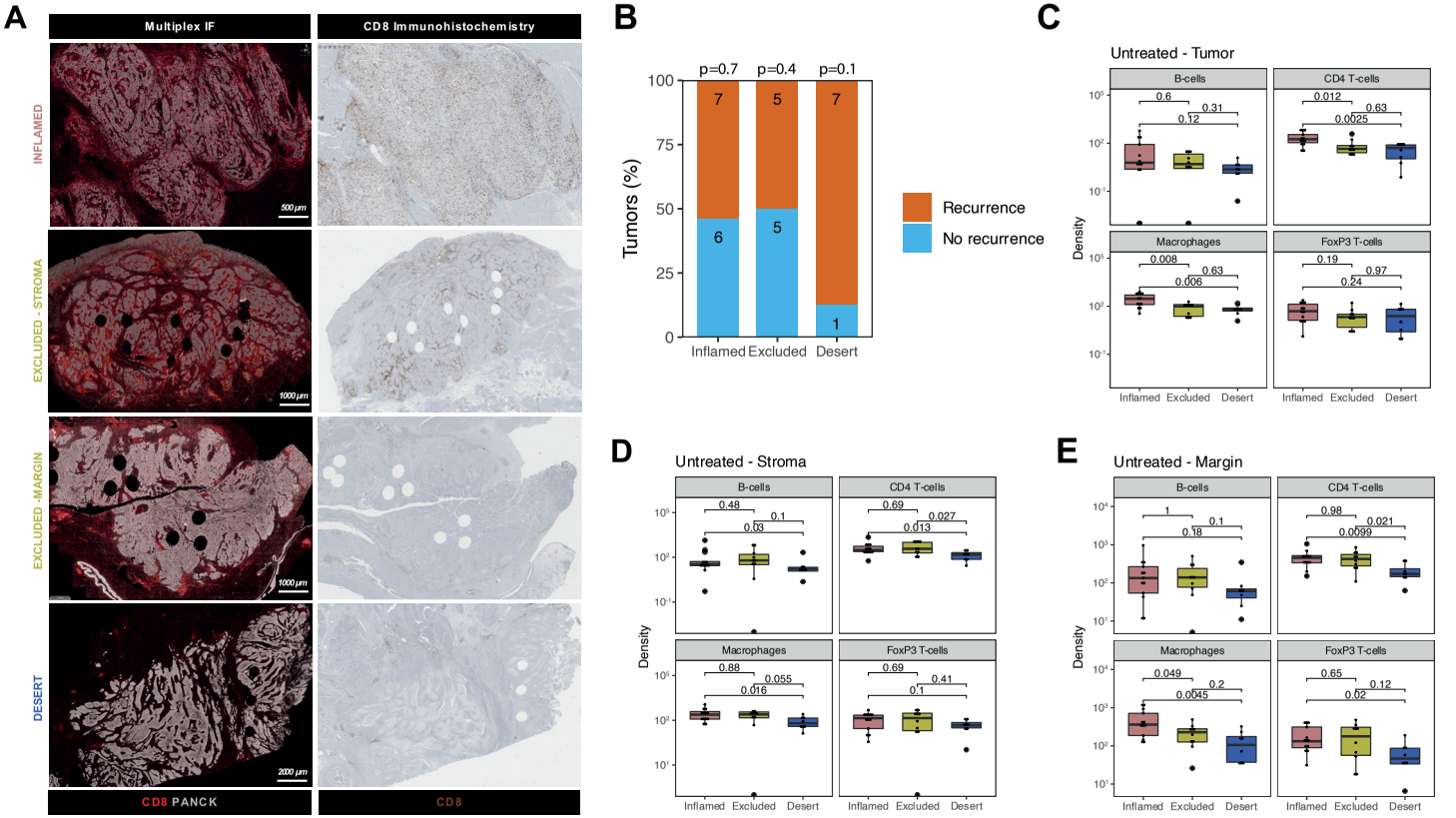


Figure 2. Urothelial cancer immune phenotypes display a varying abundance of immune cells and distinct patterns of cytotoxic T-cell exclusion in stroma and tumor margins

A. Examples of tumor immune phenotypes by multiplex immunofluorescence and CD8 immunohistochemistry in untreated urothelial cancer **B.** Proportion of patients having recurrence (n=19) and no recurrence (n=12) stratified by immune phenotype for the untreated UC cohort. The group size is indicated on each bar. A Fisher's

exact test was implemented on a 2x2 contingency table between recurrence and immune phenotype (i.e. Desert vs No Desert) for each phenotype. The p-value for each phenotype is indicated at the top of each bar. All statistical tests were two-sided. **C-E.** Comparison of immune subset densities per mm^2 in central tumor parenchyma (**C**), central tumor stroma (**D**) and tumor margin (**E**) between inflamed (n=13), excluded (n=10) and desert (n=8) tumors by quantitative multiplex IF.

The boxplots from the panels display the median and 25th and 75th percentiles, and the whiskers expand from the hinge to the largest value not exceeding the hinge 1.5×Interquantile range. Unless otherwise stated, a two-sided Mann-Whitney test was used for the comparison between distributions. The p-value is presented in-between boxplots. No adjustments were implemented for multiple comparisons.

Abbreviations: IF: immunofluorescence.

3.2.3 Markers of T-cell exhaustion in untreated and immunotherapy treated UC

Exhausted CD8 T-cells are characterized by impaired effector function and sustained expression of immune inhibitory checkpoints such as TIM3, LAG3 and PD1[19]. Immunotherapies targeting these checkpoints demonstrate promising therapeutic potential in several studies (20–26), presumably by reinvigorating exhausted T-cells. Given the implication of T-cell exhaustion as a target of immunotherapy, we employed immunohistochemistry in our untreated cohort to examine the expression of TIM3 and LAG3, as well as co-expression of CD8 and PD1. In untreated tumors, we observed considerable TIM-3 expression (example image in **Fig 3A**) on tumor-infiltrating lymphocytes (15% median positivity, range 5%-30%, **Supplementary Figure 3A**) in most central tumors, as well as in lymph nodal T-cell zones in rare cases having perivesical lymph nodes adjacent to the central tumor (**Supplementary Figure 3B**). In contrast to TIM-3, expression of LAG-3 was virtually non-existent in untreated tumors (**Supplementary Figure 3C**), as illustrated in **Supplementary Figure 3D**. Following CD8/PD1 co-staining, an algorithm was trained (**Supplementary methods 1.3**), based on a similar approach as in colorectal cancer[20], to assess $CD8^+PD1^+$ T-cells in tumor and stroma. $CD8^+PD1^+$ T-cells were clearly present in untreated UC, as shown in **Fig 3B**. Upon quantitation, we found that $CD8^+PD1^+$ T-cell abundance in tumor and stroma did not inform recurrence (**Fig 3C**). We then examined $CD8^+PD1^+$ T-cells in NABUCCO tumors having complete response (CR, qualified as pCR or CIS/pTa) and non-CR following ipilimumab plus nivolumab. $CD8^+PD1^+$ T-cells were enriched irrespective of response compared to untreated cystectomies, whereas $CD8^+PD1^+$ T-cells were highest in tumors achieving CR to immunotherapy (**Fig 3D**). Altogether, TIM-3 was highly expressed on lymphocytes and abundant $CD8^+PD1^+$ T-cells were found in cystectomies, particularly following immunotherapy, in both responders and non-responders.

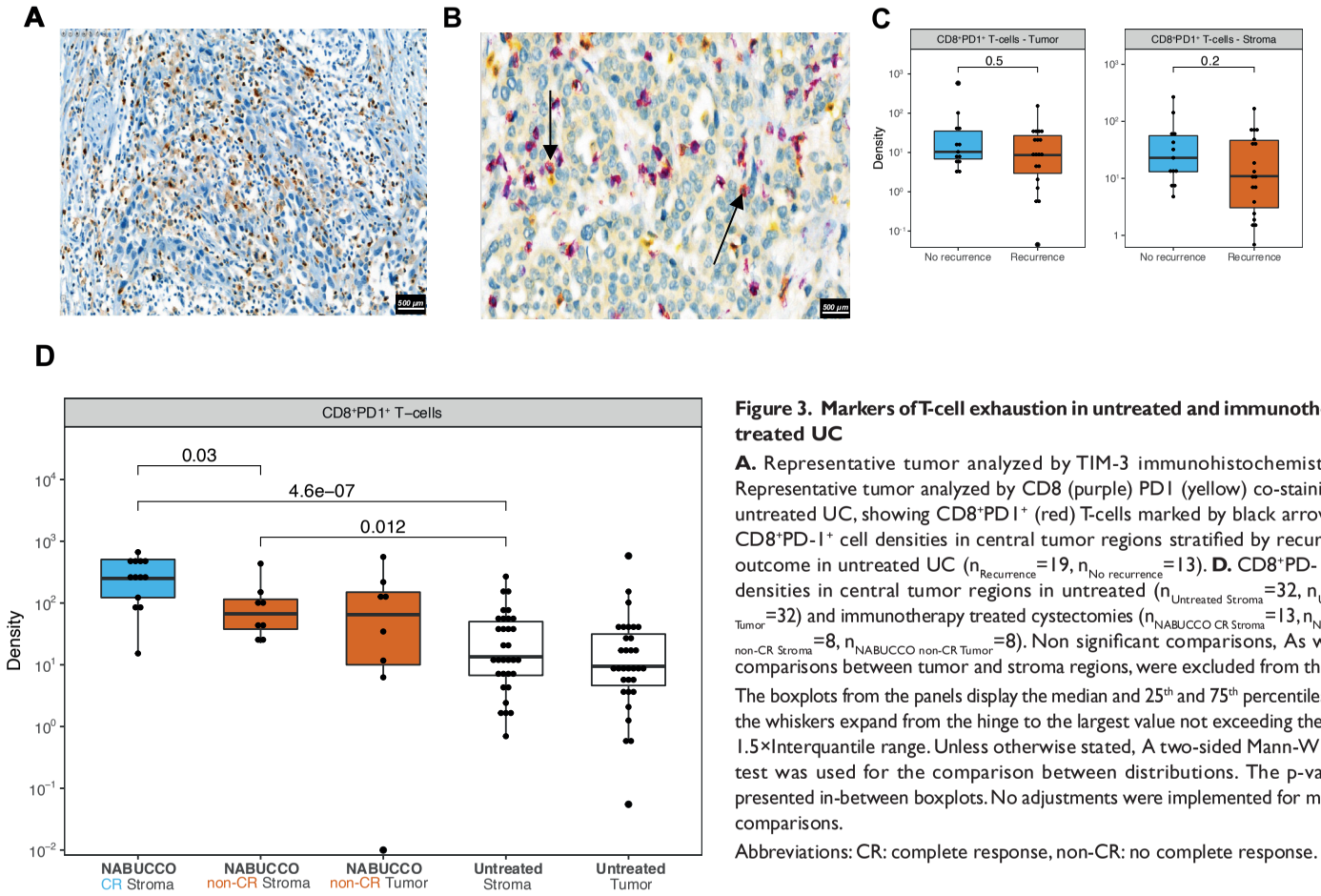


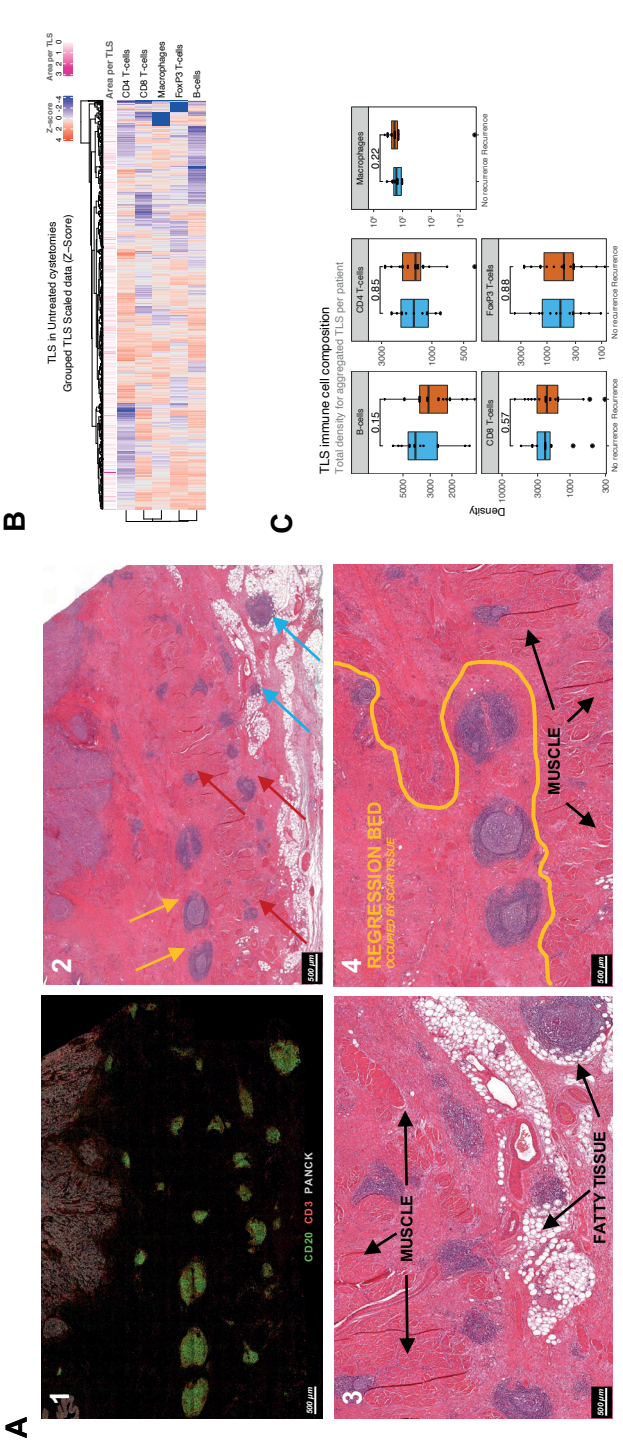
Figure 3. Markers of T-cell exhaustion in untreated and immunotherapy treated UC

A. Representative tumor analyzed by TIM-3 immunohistochemistry. **B.** Representative tumor analyzed by CD8 (purple) PD1 (yellow) co-stainings in untreated UC, showing $CD8^+PD1^+$ (red) T-cells marked by black arrows. **C.** $CD8^+PD\text{-}1^+$ cell densities in central tumor regions stratified by recurrence outcome in untreated UC ($n_{Recurrence}$=19, $n_{No\ recurrence}$=13). **D.** $CD8^+PD\text{-}1^+$ cell densities in central tumor regions in untreated ($n_{Untreated\ Stroma}$=32, $n_{Untreated\ Tumor}$=32) and immunotherapy treated cystectomies ($n_{NABUCCO\ CR\ Stroma}$=13, $n_{NABUCCO\ non\text{-}CR\ Stroma}$=8, $n_{NABUCCO\ non\text{-}CR\ Tumor}$=8). Non significant comparisons, As well as comparisons between tumor and stroma regions, were excluded from the plot. The boxplots from the panels display the median and 25th and 75th percentiles, And the whiskers expand from the hinge to the largest value not exceeding the hinge 1.5×Interquantile range. Unless otherwise stated, A two-sided Mann-Whitney test was used for the comparison between distributions. The p-value is presented in-between boxplots. No adjustments were implemented for multiple comparisons.

Abbreviations: CR: complete response, non-CR: no complete response.

3.2.4 Urothelial cancer TLS display distinct cellular composition clusters and checkpoint inhibitor-induced changes

In many cancers, the immune landscape exhibits highly organized B-cell-rich clusters related to TLS formation. The presence of TLS has been associated with favorable clinical outcomes in untreated and treated malignancies (13,27–29), whereas other studies found no correlation or immunosuppressive TLS function (30–33). We hypothesized that heterogeneity in TLS immune composition might impact anti-tumor-immunity and patient outcome in the untreated and treated setting. We employed multiplex IF to assess the cellular composition of TLS and associations with clinical outcome in our untreated cohort. TLS were automatically annotated by a trained algorithm and manually revised when needed. In total, 754 TLS aggregates were identified in untreated tumors mainly found around the muscularis propria regions, fatty tissue and fibroinflammatory regression beds (**Fig 4A**). TLS often co-localized with nerve bundles as confirmed on the corresponding H&E slide (**Supplementary Figure 4A**). Following TLS assessment by multiplex IF, most untreated tumors showed notable TLS presence, but no differences in TLS abundance were observed between recurrence groups (**Supplementary Figure 4B**). Upon quantitative analysis, TLS revealed a heterogeneous cellular immune composition, accompanied by strong variations in TLS size between TLS in untreated tumors (**Fig 4B**). No differences were found for immune subset density in aggregated TLS between recurrence groups (**Fig 4C**). As limited knowledge exists on TLS immune architecture and how immune composition impacts the clinical outcome, we grouped TLS based on immune cell density and their relative abundance in untreated tumors using a k-means clustering algorithm. We identified five distinct TLS clusters in untreated tumors (**Fig 4D**), characterized by varying abundance of immune cells (**Fig 4E**), whereas TLS cluster presence was balanced between immune phenotype subgroups (**Supplementary Figure 4C**). No differences were observed for TLS cluster abundance between outcome groups (**Fig 4F**) in untreated UC. Next, the relative abundance of TLS clusters was compared between untreated tumors and anti-PD-1/CTLA-4 treated tumors to examine how immunotherapy impacts these TLS clusters. In NABUCCO non-responders, cluster 1 (FoxP3 T-cell low) TLS were significantly enriched when compared to untreated tumors or NABUCCO responders (**Fig 4G**). Furthermore, cluster 5 (macrophage low) TLS were significantly higher in NABUCCO (non-CR or CR) tumors compared to untreated tumors (**Fig 4G**). These findings suggest that UC displays distinct TLS clusters that change in cellular composition upon immunotherapeutic treatment.


A
1
CD20 CD3 PANCK
500 µm
2
500 µm
3
MUSCLE
FATTY TISSUE
500 µm
4
REGRESSION BED
OCCUPIED BY DESMOPLASIA
MUSCLE
500 µm
B
TLS in Untreated cystectomies
Grouped TLS Scaled data (Z-Score)
Z-score
Area per TLS
Area per TLS
CD4 T-cells
CD8 T-cells
Macrophages
FoxP3 T-cells
B-cells
C
TLS immune cell composition
Total density for aggregated TLS per patient
Density
B-cells
0.15
CD4 T-cells
0.85
CD8 T-cells
0.57
FoxP3 T-cells
0.88
Macrophages
0.22
No recurrence
Recurrence

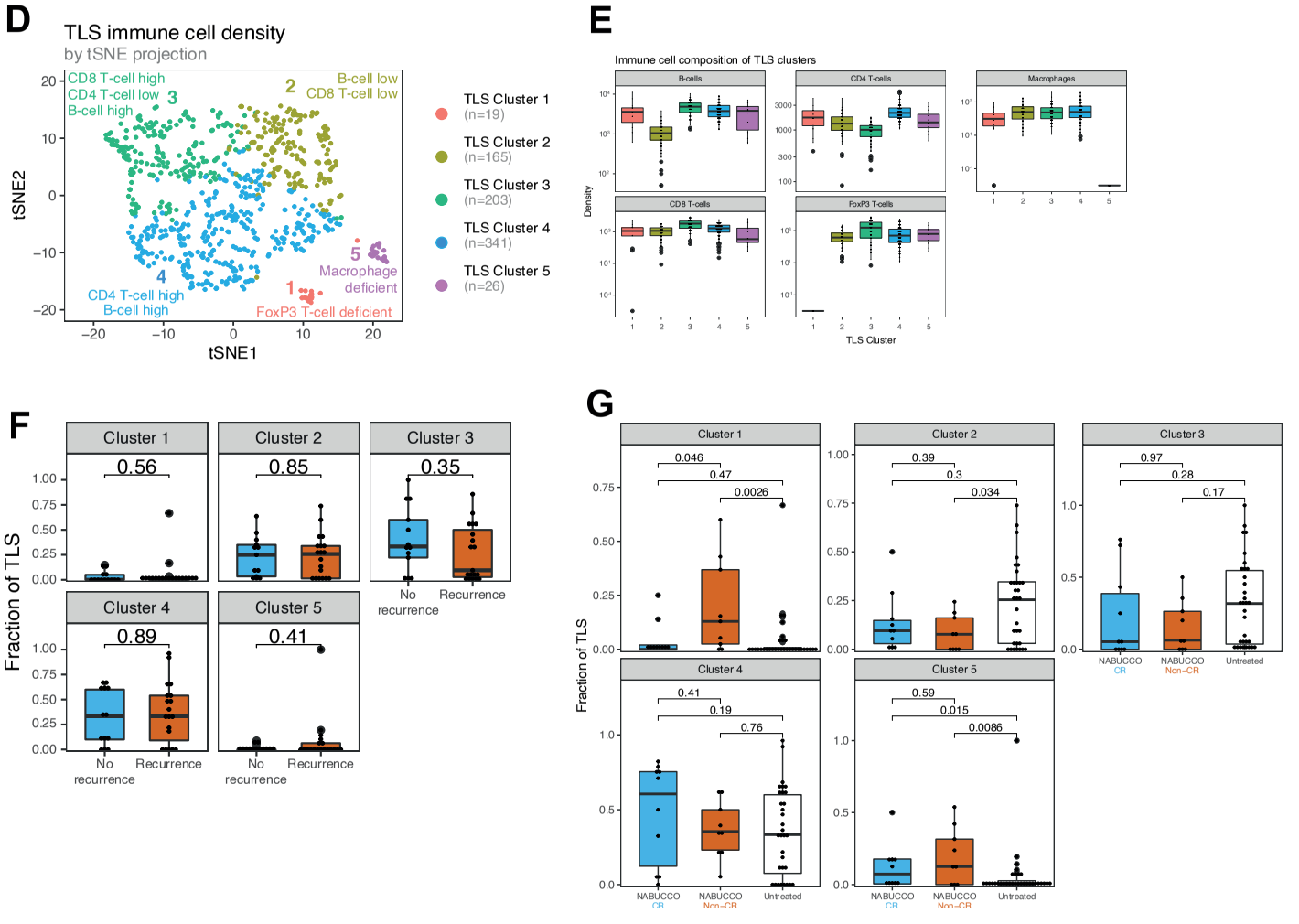

D
TLS immune cell density
by tSNE projection
CD8 T-cell high
CD4 T-cell low
B-cell high
B-cell low
CD8 T-cell low
CD4 T-cell high
B-cell high
Macrophage
deficient
FoxP3 T-cell deficient
tSNE1
tSNE2
TLS Cluster 1
(n=19)
TLS Cluster 2
(n=165)
TLS Cluster 3
(n=203)
TLS Cluster 4
(n=341)
TLS Cluster 5
(n=26)
E
Immune cell composition of TLS clusters
B-cells
CD4 T-cells
Macrophages
CD8 T-cells
FoxP3 T-cells
Density
TLS Cluster
F
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Fraction of TLS
No recurrence
Recurrence
G
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Fraction of TLS
NABUCCO CR
NABUCCO Non-CR
Untreated

Figure 4. Urothelial cancer displays distinct TLS clusters and differences in treatment effect on TLS composition between responders and non-responders.

A. 1) Multiplex immunofluorescence example showing substantial peritumoral TLS formation. 2) Corresponding haematoxylin and eosin stain, showing TLS formation in muscle (red arrow), fatty tissue (blue arrow) and fibroinflammatory regression bed (yellow). 3). Close-up image of A2, showing TLS formation around muscle, fatty tissue and in regression bed. 4) Regression bed TLS and depositions of scar tissue in areas previously harboring muscle suggest that pre-existing invasive tumor has been cleared and replaced by scar tissue, suggesting pre-existing antitumor immunity. **B.** Heatmap showing the variability of immune cell density in untreated UC TLS. Each column represents an individual TLS (n=754) from n=32 patients. Z-score high expression levels (red) and low expression levels (blue) and varying TLS size (pink) are indicated for each TLS. **C.** TLS immune subset densities per mm^2, stratified by recurrence outcome groups ($n_{Recurrence}$=19, $n_{No\ Recurrence}$=13). **D.** Clustering map upon computing a trained k-means model using 754 untreated TLS from 32 unique patients of the untreated cohort (Median 16.5 TLS per patient, Mean 24 TLS per patient, **Materials and Methods 3**). Each TLS type is assigned a colour label and an interpretation. **E.** Abundance of Immune subsets per mm^2 for each TLS cluster. TLS clusters are depicted in distinct colours. ($n_{Cluster\ 1}$=19, $n_{Cluster\ 2}$=165, $n_{Cluster\ 3}$=203, $n_{Cluster\ 4}$=341, $n_{Cluster\ 5}$=26) **F.** Comparisons of TLS relative area per cluster based on multiplex immunofluorescence between non-recurring tumors (n=13) and tumors having recurrence (n=19). **G.** Comparisons of post-treatment TLS cluster fractions between untreated tumors (n=32) and complete-responders (n=10) and non-responders (n=9) in NABUCCO.

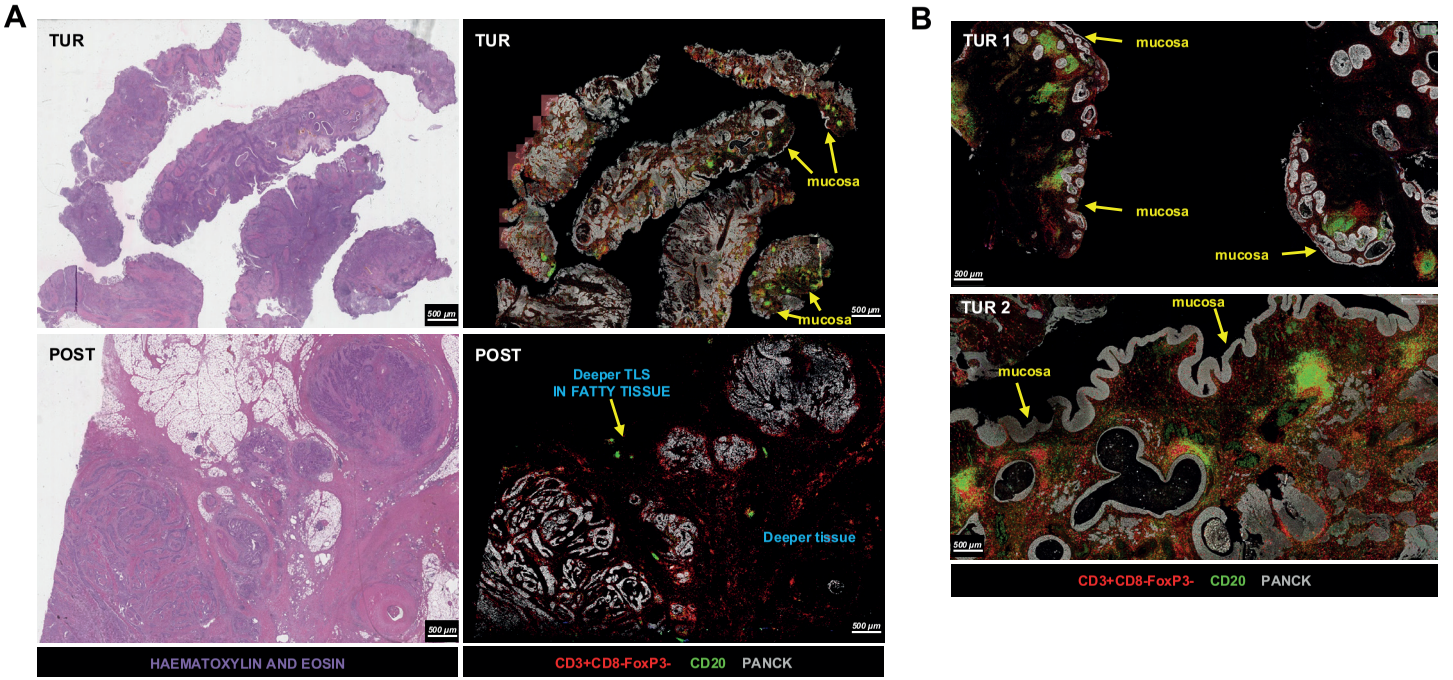
The boxplots from the panels display the median and 25th and 75th percentiles, and the whiskers expand from the hinge to the largest value not exceeding the hinge 1.5×Interquantile range. Unless otherwise stated, a two-sided Mann-Whitney test was used for the comparison between distributions. The p-value is presented in-between boxplots. No adjustments were implemented for multiple comparisons.

Abbreviations: CR: complete response, non-CR: no complete response, TLS: Tertiary lymphoid structures.

3.2.5 Discrepant TLS patterns and variable expression of CD4 T-cells between superficial and deeper TLS in urothelial cancer

Although pretreatment B-cell and TLS enrichment has been associated with favorable clinical outcomes and immunotherapy response, other studies reported no positive associations[10,13], suggesting that B-cells and TLS can have opposite roles. In NABUCCO, we previously found that immature TLS, B-cells, and genes associated with B-cell proliferation and plasma cells were enriched in pretreatment biopsies in non-CR tumors, compared to CR tumors[10]. Conversely, a study testing preoperative tremelimumab plus durvalumab in UC reported higher pretreatment TLS and B-cells in responders[13]. As other stimuli have been shown to induce TLS (31,34,35), we hypothesized that a subset of TLS may be unrelated to anti-tumor immunity, particularly in pretreatment tissue obtained by transurethral resection (TUR, debulking of a tumor from the luminal layer of the bladder). TUR biopsies primarily collect superficial tissue that is highly exposed to urinary toxins, microbial pathogens (especially in the presence of a bladder tumor) and inflammatory mediators (**Supplementary Figure 5A-B**). These TLS could cloud the tumor-associated TLS analysis, particularly in superficial parts of the tumor. To examine this, we explored whether TLS composition in superficial regions differed from TLS in deeper tissue regions. In line with quantitated results in our previous report[10], a high TLS presence was observed in NABUCCO pretreatment TUR, especially in non-CR tumors, while TLS abundance was limited in their corresponding post-treatment tissues (**Fig 5A**). TLS abundance in pretreatment TUR was particularly high in the urothelial submucosa (**Fig 5B**). TLS present in the urothelial submucosa (Superficial TLS)

were characterized by pronounced CD4 T-cell presence, whereas deeper TLS showed only limited CD4 T-cell contribution to the immune cell composition (**Fig 5B**). The predominant abundance of superficial TLS was also found in a subset of post-treatment specimens from NABUCCO (**Supplementary Figure 5C**) and untreated tumors (**Supplementary Figure 5D**), further supporting the existence of a distinct TLS population in superficial tissue. Next, we stratified superficial and deep TLS in untreated UC to compare TLS composition and the relative abundance of TLS clusters. In untreated tumors, superficial TLS showed a significantly higher CD4 T-cell presence ($p=0.012$, **Fig 5C**), which is in line with our visual observations. Next, we quantified TLS maturation stages for superficial and deep TLS using a 7-plex multiplex immunofluorescence panel on a separate, larger cohort (n=40, involving 20 patients from the original untreated cohort, **Supplementary Table 1**). Upon assigning TLS maturation, we found that superficial TLS displayed a higher fraction of early TLS and lower germinal center positive TLS when compared to deeper TLS ($p=0.001$ and $p=0.01$, respectively **Fig 5D**).


A
TUR
TUR
mucosa
mucosa
500 µm
POST
POST
Deeper TLS
IN FATTY TISSUE
Deeper tissue
HAEMATOXYLIN AND EOSIN
CD3+CD8-FoxP3- CD20 PANCK
B
TUR 1
mucosa
mucosa
mucosa
TUR 2
mucosa
mucosa
CD3+CD8-FoxP3- CD20 PANCK

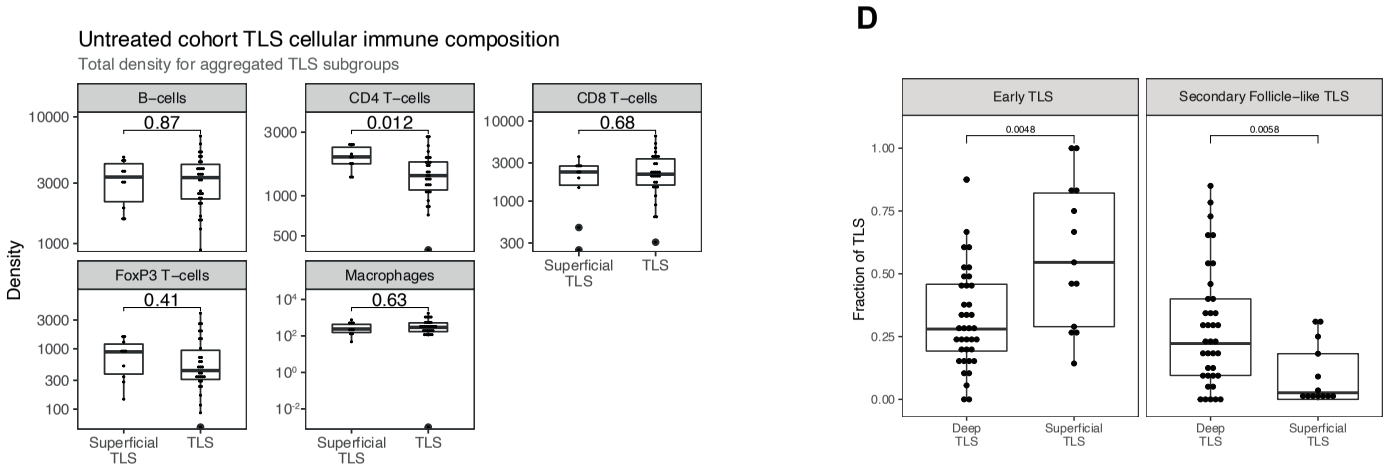


Figure 5. Discrepant TLS patterns and variable expression of T-helper cells between superficial and deeper TLS in urothelial cancer

A. Example of TLS abundance in baseline TUR and post-treatment cystectomy by multiplex immunofluorescence and haematoxylin and eosin stain in a non-responding patient in NABUCCO. Baseline TUR tissue shows a higher TLS presence than in the post-treatment specimen, particularly in submucosal regions. **B.** Two different TUR examples showing TLS that display pronounced CD4 T-cell presence. **C.** Comparison of TLS aggregated immune cell density (counts per mm^2) between superficial (n=10 patients) and deeper TLS (n=30 patients). **D.** TLS maturation states quantified by CD21-expressing Follicular Dendritic Cell networks and CD23+ Germinal center zones by multiples IF (Materials and Methods). Quantifications were done on 40 untreated UC cystectomies (18 patients from the original untreated cohort (**Fig 1A**) and 22 additional untreated UC cystectomies, **Supplementary Table 1**). Fraction of TLS maturation states are depicted for Deep TLS (n=37) and Superficial TLS (n=13), for Early TLS (germinal center negative) and Secondary Follicle-like TLS (germinal Center positive). 27 patients had only Deep TLS present, 3 patients had only Superficial TLS present, and 10 patients had both Superficial and Deep TLS present.

The boxplots from the panels display the median and 25th and 75th percentiles, and the whiskers expand from the hinge to the largest value not exceeding the hinge 1.5×Interquantile range. Unless otherwise stated, a two-sided Mann-Whitney test was used for the comparison between distributions. The p-value is presented in-between boxplots. No adjustments were implemented for multiple comparisons.

Abbreviations: TLS: Tertiary lymphoid structures, TUR: Transurethral Resection, UC: Urothelial Carcinoma

Altogether, our findings suggest that superficial TLS may be compositionally different from deeper TLS. These observations could impact the approach to immune biomarkers in UC and provides the rationale to dissect TLS populations further and study their precise role in anti-tumor immunity in the UC tumor-immune microenvironment.

3.3 DISCUSSION

The introduction of ICI changed the treatment landscape of UC. Despite recent successes, a substantial proportion of patients do not respond to immunotherapy[36,37]. As the biology driving anti-tumor immunity is still poorly understood, the characterization of the tumor immune contexture is critical to broaden our understanding of the immune landscape to ultimately improve immunotherapeutic treatment of UC patients[11].

The aim of our study was to characterize the immune landscape in tumor, stroma and TLS using computational analysis of multiplex IF. We started with a general overview of the UC immune landscape and observed substantial variation in immune subset presence across untreated tumors. Immune cells were more abundantly present in the tumor margin, compared to tumor and stroma. In previous UC immune biomarker studies, the tumor margin immune infiltrate was not specifically reported[6] or incorporated into the immune phenotype classification system[7,14]. In other cancer types such as colorectal cancer, breast cancer and melanoma, tumor margins have been extensively used for immune phenotype assessment[38]. In UC, T-cell exclusion by TGF-beta signaling has been proposed as a mechanism of resistance by excluding T-cells, emphasizing the importance of incorporating the tumor margin compartment in biomarker assessment in UC.

Tumor-specific T-cells can be re-activated through blocking immune inhibitory checkpoints (20–26). We observed high TIM-3 expression and abundant $CD8^+PD1^+$ T-cell presence in UC. $CD8^+PD1^+$ T-cells were enriched upon immunotherapy, and surprisingly, also in immunotherapy non-responders. These data suggest that, despite the immune system being able to mount an anti-cancer response upon checkpoint blockade, resistance mechanisms beyond the CTLA-4 and PD-1 checkpoints may limit cytotoxic T-cell effector function and tumor elimination in these cases. A further dissection of the tumor-immune landscape in non-responders is crucial to identify the resistance mechanism limiting the efficacy of checkpoint blockade.

In this study, we found that UC exhibits distinct TLS clusters with varying cellular composition. We observed that upon CTLA-4/PD-1 blockade, the fraction of TLS clusters 1 (FoxP3 T-cell low) was enriched in non-responding tumors when compared to untreated tumors and responding tumors. Tregs are generally believed to have immune-suppressive

functions, though limited data exist on the function of these cells within TLS. In a lung cancer mouse model, Treg presence in TLS was associated with a suppressed T-cell function[39]. Studies in colorectal cancer[40] and melanoma[41] found no correlation between Treg presence in TLS and patient survival. A possible reason for the enrichment of Treg-low TLS may be a direct therapeutic effect of anti-CTLA4, depleting Tregs in TLS. Despite Treg depletion, these tumors did not respond, suggesting that other causes for resistance might be present in these tumors[42,43].

Generally, TLS in the tumor-microenvironment are considered tumor-associated. Our findings suggest that superficial TLS may define a distinct TLS category in UC that may not be tumor-responsive. Superficial bladder tissue may exhibit immune features (e.g., TLS) unrelated to anti-tumor immunity, given the high exposure to urinary toxins or microbial pathogens, especially in the presence of a bladder tumor disrupting the mucosal barrier. We found that these superficial TLS had a higher density of CD4 T-cells. The proportion of secondary follicle-like TLS, which are required for the prognostic benefit of TLS in other cancer types[44,45], was significantly lower in superficial TLS compared to deep TLS. Given the similar characteristics, we hypothesize that superficial TLS may be related to Hunner-type interstitial cystitis, an idiopathic inflammatory disease characterized by submucosal lymphocytic pan-cystitis, lymphoid aggregates (Hunner lesions) with varying maturation stages[46] and expression of follicular T-helper cell markers[47]. In addition, a recent study showed that Hunner-type interstitial cystitis was associated with enrichment of B-cell receptor signaling genes and B-cell clonal expansion[48]. In line with these findings, we previously found that immature TLS, B-cells and genes associated with B-cell proliferation and plasma cells were enriched in baseline TUR tissue in non-CR tumors[10]. These discrepant findings in NABUCCO may be explained by the presence of tumor-unrelated TLS such as Hunner-type aggregates in the TUR samples. One can even speculate that high numbers of superficial TLS indicate prominent chronic inflammation with adverse effects on anti-tumor immunity, explaining the association with non-response. This hypothesis needs further testing. In biomarker assessments, the presence of submucosal TLS may possibly enrich B-cell and TLS levels independent of anti-tumor immunity, particularly in TUR (which removes superficial layers) and smaller biopsies. In non-UC patients, the prevalence of interstitial cystitis is 0.5% in the western world[49]. No data exists on interstitial cystitis in muscle-invasive bladder cancer, because of the prognostic impact of bladder cancer and overlapping locoregional symptoms.

The strengths of the current study are the comprehensive computational analysis and the automated nature of our assessments, enabling 1) in-depth analysis of the tumor bed, and 2) systematic assessment of tertiary lymphoid structure's immune architecture in untreated and ICI treated tumors. Combined, our study provides a unique overview of the UC immune landscape. Limitations include the limited sample size, which precluded robust assessment of

associations with outcome, and the number of immune markers profiled, which limited insight into the functional relevance of immune cells. Further limitations include the retrospective nature of our study and the risk of overinterpretation due to multiple testing.

In conclusion, our study provides a comprehensive overview of the tumor immune landscape and architecture of TLS in UC. We established distinct TLS clusters based on their cellular compositions. Compared to untreated tumors, TLS clusters showed a distinct immune cell composition in anti-CTLA-4/PD-1 ICI treated tumors. In addition, we identified a superficial TLS population, characterized by more pronounced CD4 T-cell expression than deeper TLS. The relevance of the superficial TLS population for antitumor immunity is currently unknown and warrants further investigation.

3.4 MATERIALS AND METHODS

3.4.1 Study cohort characteristics

Tumors were obtained from untreated patients and a prospective clinical trial testing the efficacy of preoperative ipilimumab (anti-CTLA-4) plus nivolumab (anti-PD-1) (NABUCCO: NCT03387761). In NABUCCO, a total of 24 patients with stage III resectable urothelial cancer (cT3-4aN0M0 and cT1-4aN1-3M0) were treated with preoperative ipilimumab 3 mg/kg (day 1), ipilimumab 3 + nivolumab 1 mg/kg (day 22), and nivolumab 3 mg/kg (day 43) followed by surgical resection. In the untreated cohort (n=31), patients had upfront cystectomy without prior systemic therapy following diagnosis of muscle-invasive carcinoma in pretreatment transurethral resection (TUR) specimen. Cystectomy specimens were preferred over TUR, given that TUR specimens provide a limited overview of the overarching tumor contexture, as shown in **Supplementary Figure 5**. The NABUCCO trial was approved by the institutional review board of the Netherlands Cancer Institute and was executed in accordance with the protocol and Good Clinical Practice Guidelines defined by the International Conference on Harmonization and the principles of the Declaration of Helsinki. Use of the cohort of untreated cystectomies was approved by the NKI-AVL institutional research board, following national regulations. Archival FFPE tumor tissue cystectomy specimens were used for immunohistochemistry and multiplex immunofluorescent analysis. Non-recurring patients and patients having recurrence were compared for explorative biomarker analysis. In NABUCCO, tumors with complete response (CR, defined as pCR, pTis or pTaN0) were compared to non-CR tumors for biomarker exploration. We included non-invasive disease in the CR definition, which is generally believed to be cured by surgery.

3.4.2 Multiplex immunofluorescence analysis and immunohistochemistry

Multiplex immunofluorescence of CD8/CD4 T-cells, B-cells, Macrophages, and B-cells

Analysis of immune cell subsets was performed by multiplex Immunofluorescence (IF) technology using an automated multiplex staining on a Discovery Ultra Stainer. Prior to multiplex staining, 3µm slides were cut on DAKO Flex IHC slides. Slides were then dried overnight and stored in +4°C. Before a run was started tissue slides were baked for 30 minutes at 70°C in an oven. Opan 7-color manual IHC kit (50 slides kit, Perkin Elmer, cat NEL81101KT) was used for staining. The protocol was initiated by heating the FFPE cuts for 28 minutes at 75°C, followed by dewaxing with Discovery Wash using the standard setting of 3 cycles of 8 minutes at 69°C. Cell Conditioning 1 (CC1,Ventana Medical Systems) was performed with Discovery CC1 buffer for 32 minutes at 95°C, after which Discovery Inhibitor was applied for 8 minutes to block endogenous peroxidase activity. Specific markers were detected consecutively on the same slide with the following antibodies, which included anti-CD3 (SP7, Cat RM-9107-S, ThermoScientific, 1/400 dilution 1 hour at RT), anti-CD8 (Clone C8/144B, Cat M7103, DAKO, 1/100 dilution 1 hour at RT), anti-CD68 (Clone KP1, M0814, Dako, 1/500 dilution, 1 hour at RT), anti-FoxP3 (clone 236A/47, Cat ab20034, Abcam, 1/50 dilution, 2 hours at RT), anti-CD20 (Clone L26, cat M0755, Dako, 1/500 dilution, 1 hour at RT) anti-PanCK (Clone AE1AE3, Cat MS-343P, Thermo Scientific, 1/100 dilution, 2 hours at RT).

Each staining cycle consisted of four steps: Primary Antibody incubation, Opal polymer HRP Ms+Rb secondary antibody incubated for 32 minutes at RT, OPAL dye incubation (OPAL520, OPAL540, OPAL570, OPAL620, OPAL650, OPAL690, 1/50 or 1/75 dilution as appropriate for 32 minutes at RT) and an antibody denaturation step using CC2 buffer for 20minutes at 95°C. Cycles were repeated for each new antibody to be stained. At the end of the protocol slides were incubated with DAPI (1/25 dilution in Reaction Buffer) for 12 minutes. After the run was finished slides were washed with demi water and mounted with Fluoromount-G (SouthernBiotech, cat 0100-01) mounting medium. After staining, imaging of the slides was done using the Vectra 3.0 automated imaging system (PerkinElmer). First, whole slide scans were made at 10x magnification. After selection of the region of interest, multispectral images were taken at 20x magnification. Library slides were created by staining a representative sample with each of the specific dyes. Using the InForm software version 2.4 and the library slides the multispectral images were unmixed into 8 channels: DAPI, OPAL520, OPAL540, OPAL570, OPAL620, OPAL650, OPAL690 and Auto Fluorescence and exported to a multilayered TIFF file. The multilayered TIFF's were fused with HALO software (Indica Labs, v2.3). Analysis was done using HALO (Indica Labs, v2.3) image analysis. Pragmatic definitions and delineation of tumor regions in a spatial context are described

in **Supplementary Methods 1.1**. Tumor and stroma regions were classified by HALO automated tissue segmentation. Quantitative assessment of central tumors was assessed in 31/32 patients, as one slide involved insufficient tumor material for appropriate assessment, but did involve notable TLS (**Supplementary Methods 1.4**).

Multiplex immunofluorescence of Tertiary Lymphoid Structures maturation states

TLS maturation was analyzed in tissue sections by 7-plex multiplex IF as previously described (Silina et al. 2018, Springer Protocols)[50]. Briefly, tissue sections were deparaffinized, rehydrated and retrieved all in one step using the Trilogy buffer (CellMarque) for 10 min at 110 °C in a pressure cooker. The following antibodies and dilutions were used for a 7-plex IF; CD21 (1:5000, clone 2G9 Leica), DC-LAMP (1:1000, clone 1010E1.01, Dendritics), CD23 (1:1000, clone SP3, Abcam), PNAd (1:5000, clone MECA-79, Biolegend), CD20 (1:5000, clone L26, Dako), CD3 (1:1000, clone SP7, ThermoScientific) and 200x magnified images were acquired by Vectra 3.0 multispectral microscope (PerkinElmer/Akoya). Area segregation was done by Inform tissue segmentation algorithm of the Inform software (Akoya).

TLS maturation stages were defined by the presence or absence of $CD21^+$ Follicular Dendritic cells (FDC) networks and $CD23^+$ Germinal Center (GC) cells in dense $CD20^+$ B-cell regions. Proportions of early TLS (no FDCs, no GC), primary follicle-like (PFL) TLS (has FDCs but no GC) and secondary follicle-like (SFL) TLS were determined as fractions out of all analyzed TLS for each patient.

Staining of TIM3, LAG3, and co-staining of CD8 and PD1

Stainings and co-stainings were performed by immunohistochemistry. Prior to the staining, 3µm sections were cut and dried overnight and subsequently transferred to Ventana Discovery Ultra autostainer. Briefly, paraffin sections were cut at 3 µm, heated at 75°C for 28 minutes, and deparaffinized in the instrument with EZ prep solution (Ventana Medical Systems). Heat-induced antigen retrieval was carried out using Cell Conditioning 1 (CC1, Ventana Medical Systems) for 64 minutes at 95°C. For the detection of TIM3, the clone D5D5R (Cell Signaling) was used (1/200 dilution, 1 hour, 370°C), and for the detection of LAG3, the clone 11E3 (1/50 dilution, 1 hour at 370°C, AbCam). The bound antibodies were detected using either Anti-Rabbit HQ (Ventana Medical Systems), 12 minutes at 37°C (TIM-3) or anti-mouse HQ (Ventana Medical Systems) for 12 minutes at 37°C (LAG-3) followed by Anti-HQ HRP (Ventana Medical Systems) for 12 minutes at 37°C and ChromoMap DAB Detection (Ventana Medical Systems). Slides were counterstained with Hematoxylin and Bluing Reagent (Ventana Medical Systems). For untreated tumors, the percentage of TIM-3 and LAG-3 expression on lymphocytes tumors was scored upon visual inspection of digital slides in Slidescore by a pathologist.

For the co-staining of PD-1 (yellow) and CD8 (purple), the protocol was adjusted. Detection of PD-1 was done using the antibody clone NAT105 (Ready-to-Use, 32 minutes at 37°C, Roche Diagnostics) in the first sequence. Visualization of the PD-1-bound antibody was done using anti-mouse NP (Ventana Medical Systems) for 12 minutes at 37°C, and subsequent anti-NP AP (Ventana Medical Systems) for 12 minutes at 37°C followed by the Discovery Yellow Detection Kit (Ventana Medical Systems). In the double-stain second sequence, CD8 was detected using the antibody clone C8/144B (Agilent, 1:200, 32 minutes at 37°C). CD8 was detected using anti-mouse HQ (Ventana Medical Systems) for 12 minutes at 37°C and subsequent anti-HQ horseradish peroxidase (Ventana Medical Systems) for 12 minutes at 37°C, followed by the Discovery Purple Detection Kit (Ventana Medical Systems). Slides were counterstained with Hematoxylin and Bluing Reagent (Ventana Medical Systems). All immunohistochemistry slides were uploaded to SlideScore for visual exploration.

3.4.3 TLS clustering approach

We employed an unsupervised learning strategy to identify TLS clusters with distinct immune cell composition. A k-Means algorithm was trained with the cellular densities (cells / mm2) of B-cells, CD4 T-cells, CD8 T-cells, FoxP3 T-cells, and macrophages in TLS using input from all TLS identified in the untreated cohort (n=754, **Fig 1A**, **Table 1**). Cellular densities per TLS (with a pseudo-count of 0.01 cells / mm2 to account for null densities) were transformed to a logarithmic scale and scaled by the standard deviation after subtracting the mean. The k-means clustering algorithm was trained by testing 1 to 10 centroids with a maximum of 300 iterations. An optimal number of k=5 clusters was selected based on a reduction or decrease of the total within-cluster sum of squares observed from k=5 to k=6 (**Supplementary Figure 6**), by visual exploration of the separation on a tSNE plot (**Fig 4D**), and by taking into account that only 5 features (distinct immune cell densities) were used to train the k-means algorithm.

To assign clusters to TLS identified in the treated NABUCCO cohort, cellular densities (with a pseudo-count of 0.01 cells / mm2 to account for null densities) were transformed to a logarithmic scale, followed by subtraction of means computed on the untreated, and scaling by the standard deviations computed on the untreated cohort. Then, we computed the distances between each TLS and each of the 5 centroids trained with the k-means clustering on the untreated cohort and predicted each TLS subtype by selecting the nearest centroid.

3.5 SUPPLEMENTARY METHODS

3.5.1 Image analysis of tumor and TLS regions upon multiplex immunofluorescence

Following multiplex immunofluorescent staining and VECTRA image acquisition (Akoya Biosciences, v3.0), image analysis was done in HALO (Indica Labs, v2.3). Using HALO, the tumor bulk was identified and the outermost edges were annotated manually. The tumor margin was defined as the region centered on the outermost border separating healthy tissue from malignant carcinoma. In HALO, the tumor margin was automatically computed from the outermost tumor edge, extending 125µm inside and beyond, covering 250µm centered on the border. Immune cell densities in tumor margins were assessed based on the tumor margin area, irrespective of tissue classifier. Based on observations, the tumor margin was arbitrarily selected to cover mostly immune cells just outside the tumor edge, without incorporating too much stroma that may dilute immune cell densities. The central tumor was defined as the area included within the outermost edges of tumor. Immune cells were quantified in central tumors for tumor and stroma, classified by the HALO random forest algorithm. automated tissue segmentation. This classifier was employed to discriminate stroma from Pan-Cytokeratin positive tumor regions. Square grids were computed in central tumors for spatial sampling to assess heterogeneity of immune subsets. In total, 30 square grid tiles were computed and normalized to fit the area within central tumor. Annotation of dense B-cell aggregates associated with tertiary lymphoid structures (TLS) formation was done manually, annotating the outer edge of dense B/T-cell presence that delineate TLS. For superficial TLS analysis, submucosal TLS were separated from TLS beyond the submucosa. The cellular density TLS immune subsets were analyzed based on the total number of cells in the total TLS area, by definition involving stroma. Tissue annotation and cell identification was performed together with a pathologist.

3.5.2 Quantification of immune phenotypes

Immune phenotypes in our urothelial cancer cohorts were assessed based on multiplex immunofluorescence analysis, resulting in 1) Immune-excluded (≥5 times higher CD8 in CT-stroma or in tumor margin compared to CT-tumor), 2) immune-desert (Below median CD8 79 T-cells per mm2] in CT-tumor, not excluded), or 3) Immune-inflamed (Above median CD8 in CT-tumor, not excluded). The median cut-off for central tumor CD8 T-cells that separated inflamed from desert tumors was supported by visual inspection, as notable CD8 infiltration was not found in tumors assigned as immune-desert. Tumor immune phenotype assignment was conducted by an experienced uro-pathologist.

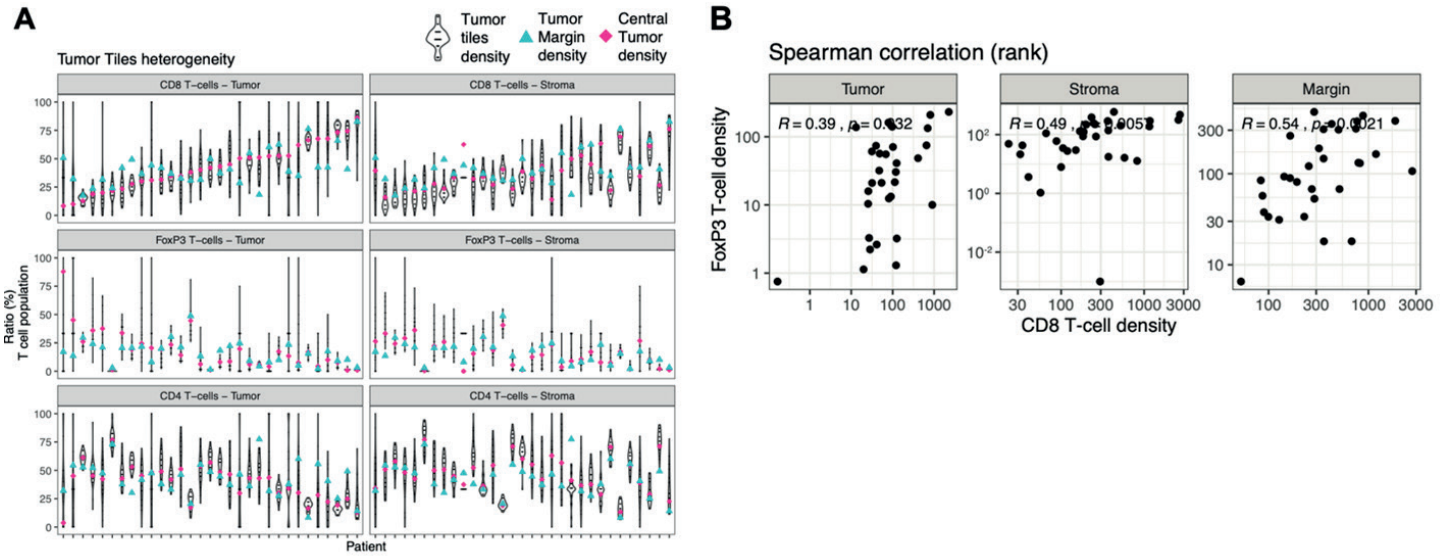
3.5.3 Analysis of CD8/P1 co-expression

Using HALO AI v3.2, we assessed CD8/PD1 positivity to assess CD8+PD1+ T-cell density in untreated and immunotherapy treated tumors. We annotated the central tumor based on the outermost border separating healthy tissue from invading tumor, as described in supplementary methods 1.1. For tissue segmentation, we annotated "tissue" and "tumor nests" within tissue to separate tissue and tumor form background on whole-slide sections. These annotations were used as a learning dataset to train a neural-network (Mininet) that is fast to run to identify tumor and stroma within central tumors. The distinct tissue classes were visualized as mark-up annotations and evaluated by a pathologist to facilitate an active learning cycle to optimize annotation accuracy. These mark-up annotations were then used as input to train a deep neural network (DenseNet), allowing more robust classifiers at higher resolutions. Cellular phenotypes were classified and analyzed using the HALO multiplex algorithm (v3.0.3). Segmentation of immune cells was done based on input from cellular parameters, including nuclei size, optical density and cellular roundness. CD8 and PD1 positivity in our cohorts was analyzed based on optimal marker color positivity thresholds established by a pathologist. The density of CD8+PD1+ T-cells was quantified in classified tissue regions separately.

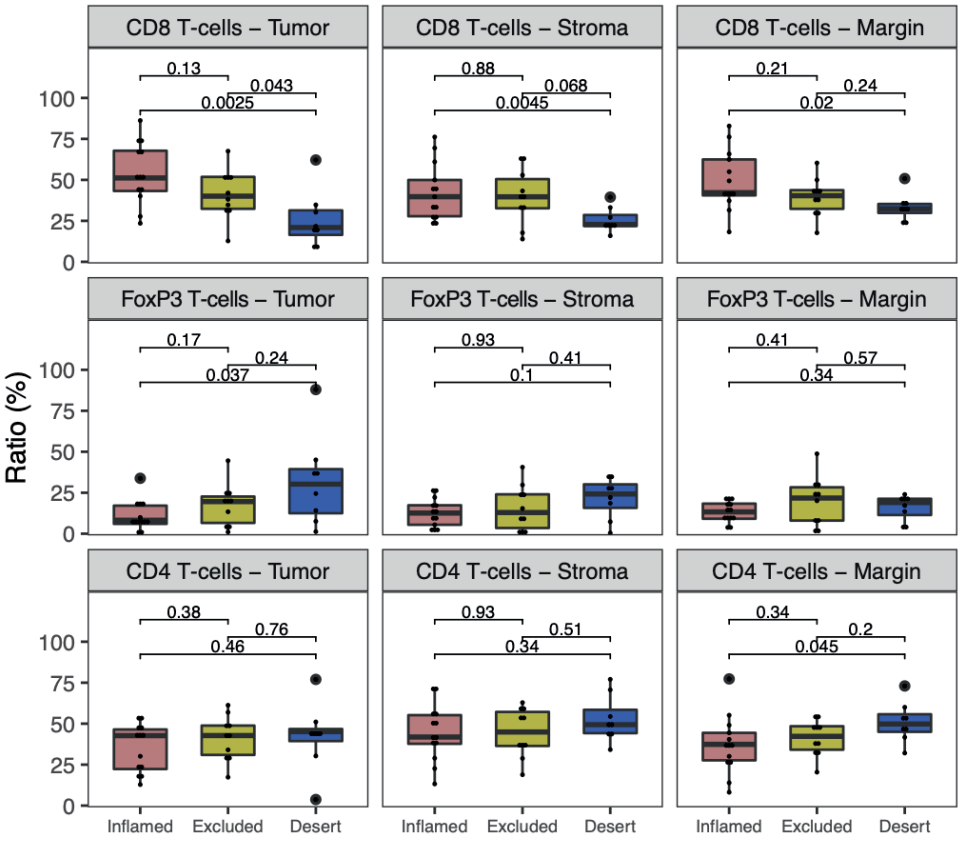
3.5.4 Quantification of central tumor and TLS on the untreated cohort

The quantification of central tumor for one patient (ypT4N0) from the untreated cohort (Table 1, Fig 1A) was not carried out due to insufficient tumor abundance present in the tissue slide from the selected tissue block. Thus, central tumor quantifications were carried out on 31 patients from the untreated cohort (Table 1). Nevertheless, because notable TLS (n=22) could be quantified in the slide, TLS quantifications on TLS for the patient were included in the TLS analysis. Thus, TLS quantifications were carried out in 32 patients from the untreated cohort (Table 1).

3.6 SUPPLEMENTARY FIGURES



Supplementary Figure 1. A. Ratio of T-cell subsets in the total T-cell population in central tumor tissue regions by multiplex immunofluorescence in untreated UC (n=31 patients) within tumor tiles (violin plot, n=30 tiles per sample), whole tumor (pink, n=31) and tumor margin (cyan, n=31). Samples were sorted by increasing intratumoral CD8 T-cell ratio (pink). **B.** The intratumoral CD8 T-cell density correlated with intratumoral ratio's of T-cell populations by multiplex immunofluorescence analysis (n=31). A Spearman correlation test was used to analyze these relations. Spearman's rank correlation coefficient and the correlation p-value are shown in the plot. The boxplots from the panels display the median and 25th and 75th percentiles, and the whiskers expand from the hinge to the largest value not exceeding the hinge 1.5×Interquantile range. Unless otherwise stated, a two-sided Mann-Whitney test was used for the comparison between distributions. The p-value is presented in-between boxplots. No adjustments were implemented for multiple comparisons.

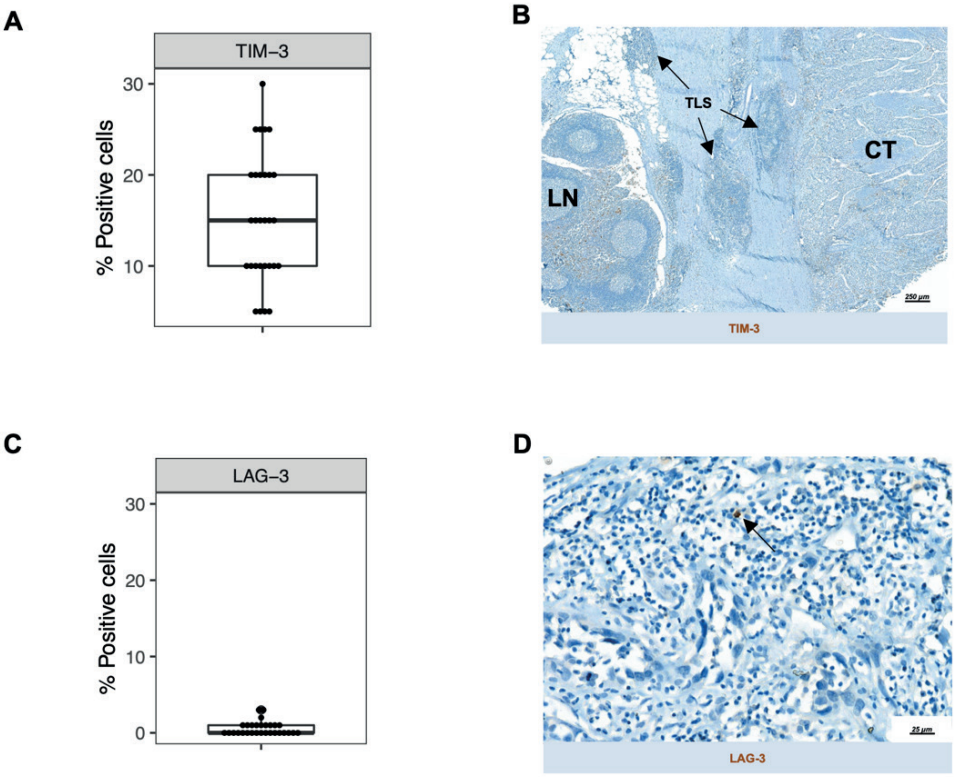


Supplementary Figure 2. A. Ratio of T-cell subsets in the total T-cell population in central tumor tissue classes and the tumor margin per immune phenotype between inflamed (n=13), excluded (n=10) and desert (n=8) tumors by multiplex immunofluorescence

The boxplots from the panels display the median and 25th and 75th percentiles, and the whiskers expand from the hinge to the largest value not exceeding the hinge 1.5×Interquantile range. Unless otherwise stated, a two-sided Mann-Whitney test was used for the comparison between distributions. The p-value is presented in-between boxplots. No adjustments were implemented for multiple comparisons.

Abbreviations: IF: immunofluorescence.

Supplementary Figure 3.

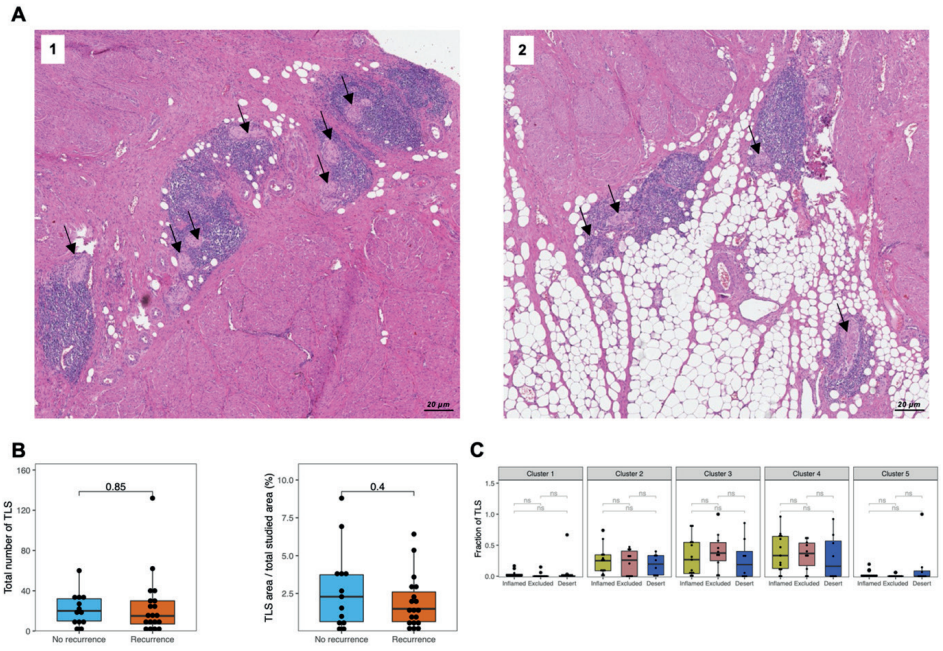


Supplementary Figure 3. A. Distribution of TIM-3 positivity by staining on the Untreated UC cohort (Table 1, n=29). Due to limited tissue size, assessment of TIM-3 on 3 tissue slides from the Untreated UC cohort was not carried out. **B.** Example of TIM-3 expression in a rare case displaying a perivesical lymph node (LN) in close proximity to the central tumor border. Prominent TIM-3 expression can be observed in the nodal T-cell zone, while tertiary lymphoid structures (TLS) marked by arrows show only limited expression. **C.** Distribution of LAG-3 positivity by staining on the Untreated UC cohort (Table 1, n=29). Due to limited tissue size, assessment of TIM-3 on 3 tissue slides from the Untreated UC cohort was not carried out. **D.** Example of LAG-3 expression in an untreated tumor. The arrow marks an isolated intratumoral LAG-3+ lymphocyte. In untreated UC, LAG-3 expression is virtually non-existent.

The boxplots from the panels display the median and 25th and 75th percentiles, and the whiskers expand from the hinge to the largest value not exceeding the hinge 1.5×Interquantile range.

Abbreviations: LN: lymph node, CT: Central tumor, TLS: tertiary lymphoid structures.

Supplementary Figure 4.

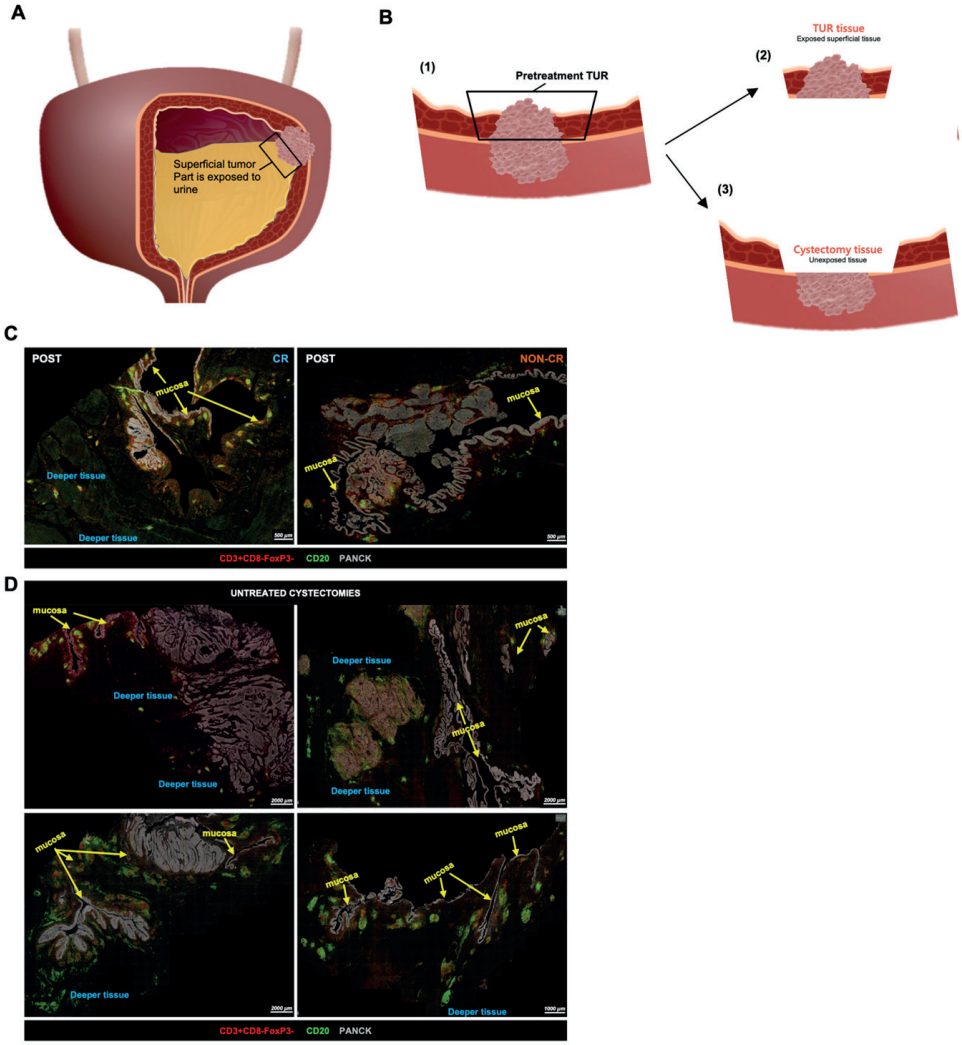


Supplementary Figure 4. A. Examples of tertiary lymphoid structures (TLS) co-localized with nerve bundles, as shown for TLS in muscularis propria regions[1] and TLS in fatty tissue[2]. **B.** Distribution of TLS abundance per patient in the untreated cohort. Left: Distribution of number of TLS between recurring (n=19) and non recurring tumors (n=13). **C.** Distribution of TLS clusters among immune phenotypes (n inflamed=13, excluded=10, desert=8).

The boxplots from the panels display the median and 25th and 75th percentiles, and the whiskers expand from the hinge to the largest value not exceeding the hinge 1.5×Interquantile range. Unless otherwise stated, a two-sided Mann-Whitney test was used for the comparison between distributions. The p-value is presented in-between boxplots. No adjustments were implemented for multiple comparisons.

Abbreviations: TLS: Tertiary lymphoid structures.

Supplementary Figure 5.

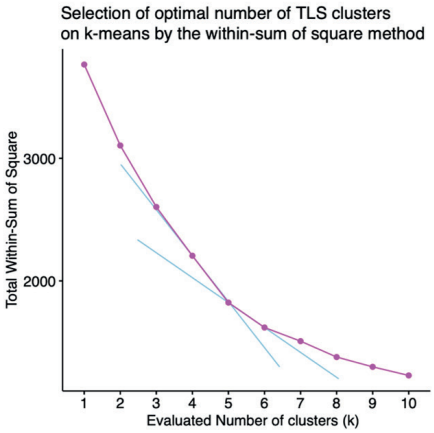


Supplementary Figure 5. A. Bladder tumor exposed to urine environment, particularly the superficial part that is resected through TUR **B.** Illustration showing 1) the tumor unexposed to TUR biopsy or therapy. Upon TUR resection (scraping tumor tissue), the tumor part exposed to urine is resected. The remaining tissue[3] does not involve superficial tissue, but deeper tissue regions involving tumor not chronically exposed to urine. **C.** Post-treatment example of TLS abundance in submucosal and deeper regions by multiplex immunofluorescence in a responding (left) and non-responding (right) tumor in NABUCCO. **D.** Examples displaying TLS abundance in submucosal and deeper regions and CD4 T-cells predominantly in superficial lymphoid aggregates in untreated cystectomies. The boxplots from the panels display the median and 25th and 75th percentiles, and the whiskers expand from the hinge to the largest value not exceeding the hinge 1.5×Interquantile range. Unless otherwise stated, a two-sided Mann-Whitney test was used for the comparison between distributions. The p-value is presented in-between boxplots. No adjustments were implemented for multiple comparisons.

Abbreviations: TUR: Transurethral resection, TLS: Tertiary Lymphoid Structure

Supplementary Figure 6.

A



Supplementary Figure 6. Selection of the optimal number of clusters for the TLS clustering approach (Materials and Methods). A k-means clustering algorithm was trained on the immune cell composition of B-cells, CD8 T-cells, CD4 T-cells, FoxP3 T-cells, and Macrophages of 754 unique TLS from the Untreated cohort (Fig 1A). The k-means algorithm was trained by changing the number of clusters from k=1 to k=10 (x axis). For each model, the within-cluster sum of squares was computed (y axis). As observed in the plot, the within-cluster sum of squares decreases rapidly from k=1 to k=5. When k is increased to 6, the slope of within-cluster sum of squares decreases. For illustration purposes, the lines of the within-cluster sum of squared interpolated from k=4 to k=5, and from k=5 to k=6, are shown as grey lines.

Abbreviations: TLS: Tertiary Lymphoid Structure.

3.7 SUPPLEMENTARY TABLES

Supplementary Table 1: Characteristics of an additional cohort of untreated tumors used for assessment of TLS maturation

Baseline characteristics	Total (n=40)
Male sex, n (%)	30 (75%)
Median age – years [range]	62.50 [39-82]
Pathological T stage, (%)	
pT1-4N0M0	25 (63%)
pT2-4N1-2M0	15 (38%)
Histology	
Urothelial Carcinoma (%)	40 (100%)
Adjuvant treatment	
No adjuvant treatment	34 (85%)
Adjuvant chemotherapy	6 (15%)
Adjuvant treatment	
No adjuvant treatment	31 (78%)
Adjuvant chemotherapy	5 (12%)
Adjuvant radiotherapy	3 (7%)
Adjuvant chemotherapy and radiotherapy (%)	1 (3%)

FUNDING

M.v.d.B. was financially supported by the Swiss National Science Foundation (CRSII5_177208 and 310030_175565), the University of Zurich Research Priority Program 'Translational Cancer Research', the Cancer Research Center Zurich and Worldwide Cancer Research (18–0629)

DATA AVAILABILITY STATEMENT

Multiplex Immunofluorescence raw data will be made available upon reasonable request for academic use by the corresponding author within the restrictions of the informed consent. A data access agreement will need to be signed with the Netherlands Cancer Institute, and reviewed by the institutional review board of the Netherlands Cancer Institute after approval.

ACKNOWLEDGMENTS

We would like to acknowledge the Core Facility Molecular Pathology and Biobanking (I. Hofland, S. Vonk, I. Seignette, W. Kieviet and J. Sanders) for support, and the Research High-Performance Computing facility, all at the Netherlands Cancer Institute.

REFERENCES

1. Rosenberg JE, Hoffman-Censits J, Powles T, van der Heijden MS, Balar AV, Necchi A, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. Lancet. 2016 May;387(10031):1909–20.
2. Bellmunt J, de Wit R, Vaughn DJ, Fradet Y, Lee JL, Fong L, et al. Pembrolizumab as Second-Line Therapy for Advanced Urothelial Carcinoma. N Engl J Med. 2017 Mar 16;376(11):1015–26.
3. Powles T, Durán I, van der Heijden MS, Loriot Y, Vogelzang NJ, De Giorgi U, et al. Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial. Lancet. 2017/12/23 ed. 2018 Feb 24;391(10122):748–57.
4. Sharma P, Retz M, Siefker-Radtke A, Baron A, Necchi A, Bedke J, et al. Nivolumab in metastatic urothelial carcinoma after platinum therapy (CheckMate 275): a multicentre, single-arm, phase 2 trial. The Lancet Oncology. 2017 Mar;18(3):312–22.
5. Massard C, Gordon MS, Sharma S, Rafii S, Wainberg ZA, Luke J, et al. Safety and Efficacy of Durvalumab (MEDI4736), an Anti–Programmed Cell Death Ligand-1 Immune Checkpoint Inhibitor, in Patients With Advanced Urothelial Bladder Cancer. 2016;34(26):3119–25.
6. Necchi A, Anichini A, Raggi D, Briganti A, Massa S, Lucianò R, et al. Pembrolizumab as Neoadjuvant Therapy Before Radical Cystectomy in Patients With Muscle-Invasive Urothelial Bladder Carcinoma (PURE-01): An Open-Label, Single-Arm, Phase II Study. 2018;36(34):3353–60.
7. Powles T, Kockx M, Rodriguez-Vida A, Duran I, Crabb SJ, Van Der Heijden MS, et al. Clinical efficacy and biomarker analysis of neoadjuvant atezolizumab in operable urothelial carcinoma in the ABACUS trial. Nat Med. 2019/11/07 ed. 2019 Nov;25(11):1706–14.
8. Sharma P, Siefker-Radtke A, de Braud F, Basso U, Calvo E, Bono P, et al. Nivolumab Alone and With Ipilimumab in Previously Treated Metastatic Urothelial Carcinoma: CheckMate 032 Nivolumab 1 mg/kg Plus Ipilimumab 3 mg/kg Expansion Cohort Results. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2019/05/18 ed. 2019 Jul 1;37(19):1608–16.
9. Powles T, van der Heijden MS, Castellano D, Galsky MD, Loriot Y, Petrylak DP, et al. Durvalumab alone and durvalumab plus tremelimumab versus chemotherapy in previously untreated patients with unresectable, locally advanced or metastatic urothelial carcinoma (DANUBE): a randomised, open-label, multicentre, phase 3 trial. The Lancet Oncology. 2020 Dec;21(12):1574–88.
10. van Dijk N, Gil-Jimenez A, Silina K, Hendricksen K, Smit LA, de Feijter JM, et al. Preoperative ipilimumab plus nivolumab in locoregionally advanced urothelial cancer: the NABUCCO trial. Nature Medicine [Internet]. 2020 Oct 12; Available from: https://doi.org/10.1038/s41591-020-1085-z
11. van Dijk N, Funt SA, Blank CU, Powles T, Rosenberg JE, van der Heijden MS. The Cancer Immunogram as a Framework for Personalized Immunotherapy in Urothelial Cancer. Eur Urol. 2018/10/03 ed. 2019 Mar;75(3):435–44.
12. McGrail DJ, Pilié PG, Rashid NU, Voorwerk L, Slagter M, Kok M, et al. High tumor mutation burden fails to predict immune checkpoint blockade response across all cancer types. Annals of Oncology. 2021 May;32(5):661–72.
13. Gao J, Navai N, Alhalabi O, Siefker-Radtke A, Campbell MT, Tidwell RS, et al. Neoadjuvant PD-L1 plus CTLA-4 blockade in patients with cisplatin-ineligible operable high-risk urothelial carcinoma. Nature Medicine [Internet]. 2020 Oct 12; Available from: https://doi.org/10.1038/s41591-020-1086-y
14. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGFbeta attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature. 2018/02/15 ed. 2018 Feb 22;554(7693):544–8.

15. Chen DS, Mellman I. Oncology Meets Immunology: The Cancer-Immunity Cycle. Immunity. 2013 Jul;39(1):1–10.
16. Echarti A, Hecht M, Büttner-Herold M, Haderlein M, Hartmann A, Fietkau R, et al. CD8+ and Regulatory T cells Differentiate Tumor Immune Phenotypes and Predict Survival in Locally Advanced Head and Neck Cancer. Cancers. 2019 Sep 19;11(9):1398.
17. Rizvi H, Sanchez-Vega F, La K, Chatila W, Jonsson P, Halpenny D, et al. Molecular Determinants of Response to Anti–Programmed Cell Death (PD)-1 and Anti–Programmed Death-Ligand 1 (PD-L1) Blockade in Patients With Non–Small-Cell Lung Cancer Profiled With Targeted Next-Generation Sequencing. Journal of Clinical Oncology. 2018;36(7):633–41.
18. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. Nature. 2017 Jan 18;541(7637):321–30.
19. Wherry EJ. T cell exhaustion. Nat Immunol. 2011 Jun;12(6):492–9.
20. Chalabi M, Fanchi LF, Dijkstra KK, Van den Berg JG, Aalbers AG, Sikorska K, et al. Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. Nature Medicine [Internet]. 2020 Apr 6; Available from: https://doi.org/10.1038/s41591-020-0805-8
21. Ngiow SF, von Scheidt B, Akiba H, Yagita H, Teng MW, Smyth MJ. Anti-TIM3 antibody promotes T cell IFN-gamma-mediated antitumor immunity and suppresses established tumors. Cancer Res. 2011 May 15;71(10):3540–51.
22. Kumagai S, Togashi Y, Kamada T, Sugiyama E, Nishinakamura H, Takeuchi Y, et al. The PD-1 expression balance between effector and regulatory T cells predicts the clinical efficacy of PD-1 blockade therapies. Nature Immunology [Internet]. 2020 Aug 31; Available from: https://doi.org/10.1038/s41590-020-0769-3
23. Das M, Zhu C, Kuchroo VK. Tim-3 and its role in regulating anti-tumor immunity. Immunol Rev. 2017 Mar;276(1):97–111.
24. Thommen DS, Koelzer VH, Herzig P, Roller A, Trefny M, Dimeloe S, et al. A transcriptionally and functionally distinct PD-1+ CD8+ T cell pool with predictive potential in non-small-cell lung cancer treated with PD-1 blockade. Nat Med. 2018 Jul;24(7):994–1004.
25. Rotte A, Jin JY, Lemaire V. Mechanistic overview of immune checkpoints to support the rational design of their combinations in cancer immunotherapy. Annals of Oncology. 2018 Jan;29(1):71–83.
26. Lipson EJ, Tawbi HA-H, Schadendorf D, Ascierto PA, Matamala L, Gutiérrez EC, et al. Relatlimab (RELA) plus nivolumab (NIVO) versus NIVO in first-line advanced melanoma: Primary phase III results from RELATIVITY-047 (CA224-047). JCO. 2021 May 20;39(15_suppl):9503–9503.
27. Siliņa K, Soltermann A, Attar FM, Casanova R, Uckeley ZM, Thut H, et al. Germinal Centers Determine the Prognostic Relevance of Tertiary Lymphoid Structures and Are Impaired by Corticosteroids in Lung Squamous Cell Carcinoma. 2018;78(5):1308–20.
28. Helmink BA, Reddy SM, Gao J, Zhang S, Basar R, Thakur R, et al. B cells and tertiary lymphoid structures promote immunotherapy response. Nature. 2020 Jan 1;577(7791):549–55.
29. Cabrita R, Lauss M, Sanna A, Donia M, Skaarup Larsen M, Mitra S, et al. Tertiary lymphoid structures improve immunotherapy and survival in melanoma. Nature. 2020 Jan 1;577(7791):561–5.
30. Goeppert B, Frauenschuh L, Zucknick M, Stenzinger A, Andrulis M, Klauschen F, et al. Prognostic impact of tumour-infiltrating immune cells on biliary tract cancer. Br J Cancer. 2013 Nov;109(10):2665–74.
31. Hill DG, Yu L, Gao H, Balic JJ, West A, Oshima H, et al. Hyperactive gp130/STAT3-driven gastric tumourigenesis promotes submucosal tertiary lymphoid structure development: gp130/STAT3 regulates TLSs in gastric cancer. Int J Cancer. 2018 Jul 1;143(1):167–78.
32. Figenschau SL, Fismen S, Fenton KA, Fenton C, Mortensen ES. Tertiary lymphoid structures are associated with higher tumor grade in primary operable breast cancer patients. BMC Cancer. 2015 Dec;15(1):101.

33. Lee HJ, Kim JY, Park IA, Song IH, Yu JH, Ahn J-H, et al. Prognostic Significance of Tumor-Infiltrating Lymphocytes and the Tertiary Lymphoid Structures in HER2-Positive Breast Cancer Treated With Adjuvant Trastuzumab. American Journal of Clinical Pathology. 2015 Aug 1;144(2):278–88.
34. Pipi E, Nayar S, Gardner DH, Colafrancesco S, Smith C, Barone F. Tertiary Lymphoid Structures: Autoimmunity Goes Local. Front Immunol. 2018 Sep 12;9:1952.
35. Rubio CA, Ásmundsson J, Silva P, Illies C, Hartman J, Kis L. Lymphoid aggregates in Crohn's colitis and mucosal immunity. Virchows Arch. 2013 Nov;463(5):637–42.
36. Szabados B, van Dijk N, Tang YZ, van der Heijden MS, Wimalasingham A, Gomez de Liano A, et al. Response Rate to Chemotherapy After Immune Checkpoint Inhibition in Metastatic Urothelial Cancer. Eur Urol. 2017/09/18 ed. 2018 Feb;73(2):149–52.
37. Gómez de Liaño Lista A, van Dijk N, de Velasco Oria de Rueda G, Necchi A, Lavaud P, Morales-Barrera R, et al. Clinical outcome after progressing to frontline and second-line Anti–PD-1/PD-L1 in advanced urothelial cancer. European Urology. 2020 Feb 1;77(2):269–76.
38. Kather JN, Suarez-Carmona M, Charoentong P, Weis C-A, Hirsch D, Bankhead P, et al. Topography of cancer-associated immune cells in human solid tumors. eLife. 2018 Sep 4;7:e36967.
39. Joshi NS, Akama-Garren EH, Lu Y, Lee D-Y, Chang GP, Li A, et al. Regulatory T Cells in Tumor-Associated Tertiary Lymphoid Structures Suppress Anti-tumor T Cell Responses. Immunity. 2015 Sep;43(3):579–90.
40. Yamaguchi K, Ito M, Ohmura H, Hanamura F, Nakano M, Tsuchihashi K, et al. Helper T cell-dominant tertiary lymphoid structures are associated with disease relapse of advanced colorectal cancer. OncoImmunology. 2020 Jan 1;9(1):1724763.
41. Lynch KT, Young SJ, Meneveau MO, Wages NA, Engelhard VH, Slingluff Jr CL, et al. Heterogeneity in tertiary lymphoid structure B-cells correlates with patient survival in metastatic melanoma. J Immunother Cancer. 2021 Jun;9(6):e002273.
42. Blank CU, Haanen JB, Ribas A, Schumacher TN. The "cancer immunogram." Science. 2016 May 6;352(6286):658–60.
43. van Dijk N, Funt SA, Blank CU, Powles T, Rosenberg JE, van der Heijden MS. The Cancer Immunogram as a Framework for Personalized Immunotherapy in Urothelial Cancer. European Urology. 2019 Mar;75(3):435–44.
44. Posch F, Silina K, Leibl S, Mündlein A, Moch H, Siebenhüner A, et al. Maturation of tertiary lymphoid structures and recurrence of stage II and III colorectal cancer. Oncoimmunology. 2017 Dec 18;7(2):e1378844–e1378844.
45. Silina K, Soltermann A, Attar FM, Casanova R, Uckeley ZM, Thut H, et al. Germinal Centers Determine the Prognostic Relevance of Tertiary Lymphoid Structures and Are Impaired by Corticosteroids in Lung Squamous Cell Carcinoma. Cancer research. 2017/12/28 ed. 2018 Mar 1;78(5):1308–20.
46. Akiyama Y, Luo Y, Hanno PM, Maeda D, Homma Y. Interstitial cystitis/bladder pain syndrome: The evolving landscape, animal models and future perspectives. Int J Urol. 2020 Jun;27(6):491–503.
47. Lu K, Wei S, Wang Z, Wu K, Jiang J, Yan Z, et al. Identification of novel biomarkers in Hunner's interstitial cystitis using the CIBERSORT, an algorithm based on machine learning. BMC Urol. 2021 Dec;21(1):109.
48. Akiyama Y, Maeda D, Katoh H, Morikawa T, Niimi A, Nomiya A, et al. Molecular Taxonomy of Interstitial Cystitis/Bladder Pain Syndrome Based on Whole Transcriptome Profiling by Next-Generation RNA Sequencing of Bladder Mucosal Biopsies. Journal of Urology. 2019 Aug;202(2):290–300.
49. Hanno PM, Erickson D, Moldwin R, Faraday MM. Diagnosis and Treatment of Interstitial Cystitis/Bladder Pain Syndrome: AUA Guideline Amendment. Journal of Urology. 2015 May;193(5):1545–53.
50. Siliņa K, Burkhardt C, Casanova R, Solterman A, van den Broek M. A Quantitative Pathology Approach to Analyze the Development of Human Cancer-Associated Tertiary Lymphoid Structures. In: Dieu-Nosjean M-C, editor. Tertiary Lymphoid Structures [Internet]. New York, NY: Springer New York; 2018 [cited 2021 Oct 7]. p. 71–86. (Methods in Molecular Biology; vol. 1845). Available from: http://link.springer.com/10.1007/978-1-4939-8709-2_5

4

Preoperative Ipilimumab Plus Nivolumab in Locoregionally Advanced Urothelial Cancer: The NABUCCO Trial

Nick van Dijk[1*], Alberto Gil-Jimenez[2,3*], Karina Silina[4], Kees Hendricksen[5], Laura Smit[6], Jeantine M. de Feijter[1], Maurits van Montfoort[6], Charlotte van Rooijen[6], Dennis Peters[7], Annegien Broeks[7], Henk van der Poel[4], Annemarie Bruining[8], Yoni Lubeck[2,6], Karolina Sikorska[9], Thierry Boellaard[8], Pia Kvistborg[10], Daniel J. Vis[2,3], Erik Hooijberg[6], Ton N. Schumacher[3,10], Maries van den Broek[4], Lodewyk F. Wessels[2,3], Christian U. Blank[1,10], Bas W. van Rhijn[5,11], Michiel S. van der Heijden[1,2,#]

Parts of this chapter have been published in Nature Medicine (2020): https://www.nature.com/articles/s41591-020-1085-z

[1]Department of Medical Oncology, Netherlands Cancer Institute, Amsterdam, the Netherlands.
[2]Department of Molecular Carcinogenesis, Netherlands Cancer Institute, Amsterdam, the Netherlands.
[3]Oncode Institute, Utrecht, the Netherlands. [4]Institute of Experimental Immunology, University of Zurich, Zurich, Switzerland,
[5]Department of Urology, Netherlands Cancer Institute, Amsterdam, the Netherlands.
[6]Department of Pathology, Netherlands Cancer Institute, Amsterdam, the Netherlands.
[7]Core Facility Molecular Pathology & Biobanking, Netherlands Cancer Institute, Amsterdam, the Netherlands.
[8]Department of Radiology, Netherlands Cancer Institute, Amsterdam, the Netherlands.
[9]Department of Biometrics, Netherlands Cancer Institute, Amsterdam, the Netherlands.
[10]Division of Molecular Oncology and Immunology, Netherlands Cancer Institute, Amsterdam, the Netherlands. [11]Department of Urology, Caritas St. Josef Medical Centre, University of Regensburg, Regensburg, Germany

*These authors contributed equally

[#] Corresponding author

ABSTRACT

Preoperative immunotherapy with anti-PD1 plus anti-CTLA4 antibodies has shown remarkable pathological responses in melanoma[1] and colorectal cancer[2]. In NABUCCO (ClinicalTrials.gov: NCT03387761), a single-arm feasibility trial, twenty-four stage III urothelial cancer patients received 2 doses of ipilimumab and 2 doses of nivolumab, followed by resection. The primary endpoint was feasibility to resect <12 weeks from treatment start. All patients were evaluable for the study endpoints and underwent resection; twenty-three (96%) within <12 weeks. Grade 3-4 immune-related adverse events occurred in 55% of patients; 41% when excluding clinically insignificant laboratory abnormalities. 46% of patients had a pathological complete response (pCR), meeting the secondary efficacy endpoint. Fourteen patients (58%) had no remaining invasive disease (pCR or pTisN0/pTaN0). In contrast to studies with anti-PD1/PD-L1 monotherapy, complete response to ipilimumab plus nivolumab was independent of baseline $CD8^+$ presence or T-effector signatures. Induction of TLS upon treatment was observed in responding patients. Our data indicate that combined CTLA-4 plus PD-1 blockade may provide an effective preoperative treatment strategy in locoregionally advanced UC, irrespective of pre-existing CD8+ T-cell activity.

4.1 MAIN

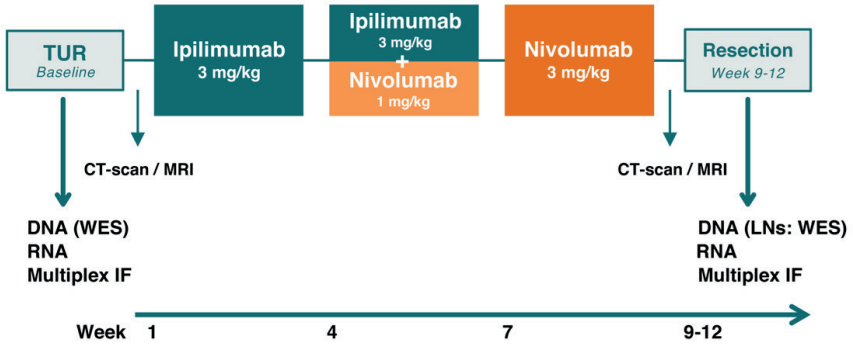
Recurrence rates after surgical resection of muscle-invasive urothelial cancer (UC) are high and many patients die of their disease[3]. Although neoadjuvant cisplatin-based chemotherapy shows impressive responses, including 22-40% pathological complete response (pCR)[4], the absolute survival benefit is only 5% at 5 years[5], accompanied by substantial toxicity. Given the high rate of distant recurrences after surgical resection of UC, there is a great need for more effective systemic treatment approaches.

Immune checkpoint inhibitors have shown durable responses in the metastatic setting[6-12]. Response to checkpoint inhibitors appears higher when treating in first-line rather than second-line metastatic disease and when treating patients with only lymph node metastases[13], providing the rationale for immunotherapeutic treatment in earlier disease settings. A study testing neo-adjuvant ipilimumab showed that this drug can safely be administered prior to cystectomy and does not preclude resection[14,15]. Recently, promising pCR rates were observed in trials testing preoperative anti-PD-1/PD-L1 in UC of the bladder[16-18]. However, complete responses were primarily observed in less advanced tumors and tumors with pre-existing $CD8^+$ T-cell immunity[16,18]. In the NABUCCO study, we investigated whether the addition of anti-CTLA-4 to PD-1 blockade is feasible as preoperative treatment in locoregionally advanced (stage III) UC. Furthermore, we studied whether this combination treatment could broaden efficacy to more advanced tumors and tumors with limited baseline $CD8^+$ T-cell immunity. Given the recently published work correlating B-cells and tertiary lymphoid structures (TLS) to immunotherapy response[19,20], we also assessed baseline B-cell presence and TLS dynamics.

In the NABUCCO study **(Fig. 1a)**, 24 stage III UC patients were enrolled **(Fig. 1b)** between February 2018 and February 2019 and treated with ipilimumab 3mg/kg (day 1), ipilimumab 3 mg/kg plus nivolumab 1 mg/kg (day 22) and nivolumab 3 mg/kg (day 43), followed by resection[21] (See **Supplementary Study protocol**). The median postoperative follow-up was 8.3 months (IQR 4.7 – 13.1 months). Fourteen (58%) patients had node-negative disease at baseline (cT3-4aN0) and 10 (42%) patients had baseline lymph node metastases (cT2-4aN1-3). Patients were cisplatin-ineligible or refused cisplatin-based chemotherapy. All 24 patients were evaluable for the primary endpoint, which was surgical resection within 12 weeks after initiation of preoperative therapy. All patients had surgical resection, despite inclusion of patients with bulky/extensive disease. Twenty-three (96%; 95% CI: 79-100%) patients underwent resection <12 weeks, thus meeting the study's primary endpoint. One patient had a delay of 4 weeks due to immune-related hemolysis. Grade 3-4 immune-related adverse events (irAEs) occurred in 55% of patients **(Supplementary Table 1)**; 41% when excluding clinically insignificant laboratory abnormalities (mainly lipase elevation). Eighteen

(75%) patients received all 3 treatment cycles. Six (25%) patients had only 2 cycles due to irAEs. There was no treatment-related mortality.

a



b

BASELINE CHARACTERISTICS	TOTAL *(N=24)*
Median age, years (range)	65 (50-81)
Male sex, N (%)	18 (75%)
Clinical T stage, (%)	
▪ cT3-4aN0M0	14 (58%)
▪ cT2-4aN1-3M0[1]	10 (42%)
Cisplatin eligibility status, N (%)	
▪ Cisplatin-ineligible	13 (54%)
▪ Refusal of cisplatin-based chemotherapy[2]	11 (46%)
WHO, N (%)	
▪ WHO 0	21 (88%)
▪ WHO 1	3 (12%)
Median tumor volume, cm^3 (range)	
▪ Primary tumor[3]	19.9 (0-153)
PD-L1 combined positive score (CPS)	
▪ Median PD-L1 CPS, %	32
▪ PD-L1 CPS > 10, N (%)	15 (63%)
▪ PD-L1 CPS < 10, N (%)	9 (37%)

[1] One patient had bladder + upper tract disease, all others had bladder only disease
[2] Cisplatin-eligible according to Galsky criteria. Patients had relative contraindications or refused cisplatin-based chemotherapy
[3] Assessed through multiparametric magnetic resonance

Figure 1. NABUCCO study design and baseline study population characteristics.
A. NABUCCO study treatment and time points of radiological assessment and sample collection. **B.** Baseline characteristics of patients enrolled in the study. Abbreviations: TUR: transurethral resection, MRI: magnetic Resonance Imaging, CT: computed tomography, TMB: tumor mutational burden, FFPE: Formalin-fixed paraffin-embedded tissue, LN: lymph node, IF: immunofluorescence, WES: whole-exome sequencing, RNA: ribonucleic acid, DNA: deoxyribonucleic acid

Following ipilimumab plus nivolumab, most resected primary tumors displayed extensive tumor regression at histopathological examination **(Fig. 2a).** Eleven (46%; 95% CI: 26–67%) patients had pCR (pT0N0), meeting the pre-specified secondary efficacy endpoint. Fourteen

(58%; 95% CI: 37–77%) patients had no residual invasive cancer (pCR or small CIS/pTa focus) after treatment. Two additional patients (8%) achieved a major pathological response (MPR), defined as <10% residual vital tumor and pN0[18], exhibiting extensive immune infiltration **(Fig. 2b)**. A 50% (95% CI: 29–71%) pCR rate was observed in patients without baseline lymph node metastases (cT3-4aN0), compared to 40% (95% CI: 12–73%) pCR in patients with clinically suspected node positive disease (cT2-4aN1-3). Interestingly, four patients achieved pCR/MPR in the primary tumor, while a resistant micrometastasis was observed in a concurrent lymph node. No specific genetic resistance mechanisms could be identified in discrepant mutations between the primary tumor and unresponsivene lymph nodes upon whole-exome sequencing (WES, see **Extended Data Fig. 1**). At the time of analysis, two patients (both non-pCR) had relapsed, one of these patients died of metastatic disease **(Fig. 2c)**.

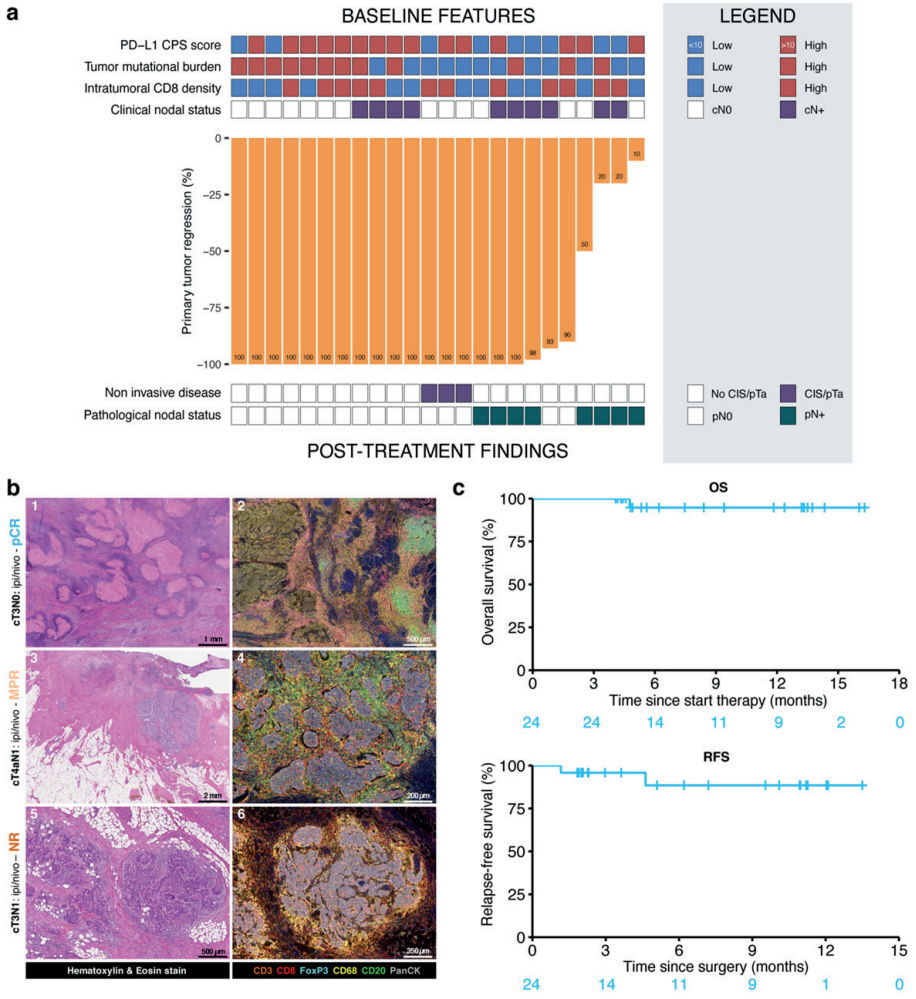


Figure 2. Pathological tumor regression and outcome to preoperative ipilimumab plus nivolumab. **A.** Percentage of pathological tumor regression per primary bladder tumor after preoperative ipilimumab+nivolumab, based on the percentage of residual viable tumor area, for the complete cohort. Baseline clinical and biomarker features are annotated for each patient in the top section. Median mutational burden cut offs; below median (Low <9.6) vs above median (High ≥9.6) mutations per megabase. Intratumoral median CD8 cell density assigned as below median (Low <145.6) vs above median (High ≥145.6) cells per mm2. Histopathological findings in the surgical resection specimen upon ipilimumab+nivolumab are listed in the lower panel **B1-2**. Pathological complete response (pCR) example, displaying large fields of necrosis surrounded by abundant CD8 T-cells and TLS in multiplex immunofluorescence images. **B3-4**. Major pathological response (MPR, < 10% residual viable tumor), characterized by reactive changes and small remaining tumor nests with substantial immune cell infiltration after treatment. **B5-6.** Non-response, displaying intact tumor fields and immune infiltration at the tumor margin. **C.** Kaplan Meyer analysis of recurrence-free survival (RFS) and overall survival (OS). At a median postoperative follow-up of 8.3 months, 2 (8%) patients have relapsed (both non-responders). One patient died 6 weeks after surgery due to

rapid progressive disease. Experiments and scorings related to the presented micrographs were conducted once.

We examined several published biomarkers for immunotherapy response in our cohort. For these biomarker analyses, we compared complete response (CR, defined as pCR or CIS/pTa) to non-CR. The CR rate in PD-L1 positive tumors (CPS>10) was 73% (95% CI: 45–92%), compared to 33% (95% CI: 7-70%) in PD-L1 negative tumors (p=0.15). Analysis of somatic mutations by WES (see **Extended Data Fig. 2**) of pretreatment tumor tissue and germline DNA showed a trend towards a higher tumor mutational burden (TMB) in tumors achieving CR to ipilimumab plus nivolumab compared to non-CR tumors (p=0.056, **Fig. 3a)**. We also assessed mutations in a set of DNA damage response (DDR) genes[22]. Alterations in DDR genes were more frequently observed in responding tumors than in non-CR tumors (p=0.03, **Fig. 3b)**. Recent data from studies in advanced and localized UC indicate that Transforming Growth Factor beta (TGF-β) signaling is associated with atezolizumab unresponsiveness[18,23]. In line with these data, we found a TGF-β expression signature to be associated with non-response **(Fig. 3c).** This result suggests that TGF-β mediated inhibition of the anti-cancer immune response cannot be overcome by the addition of anti-CTLA-4 to PD-1 blockade.

Previous studies in UC found that response to preoperative anti-PD1/PD-L1 monotherapy is primarily observed in tumors exhibiting pre-existing $CD8^+$ T-cell immunity[16,18]. This restriction could limit the general applicability of anti-PD1/PD-L1 monotherapy in the preoperative setting. We assessed whether the addition of anti-CTLA-4 enables response in tumors with limited pre-existing T-cell immunity in our cohort. Using quantitative multiplex immunofluorescence, we observed no correlation between baseline $CD8^+$ T-cell density and response to ipilimumab plus nivolumab **(Fig. 3d)**. The CR rate in both the lowest and highest $CD8^+$ cell density quartile was 67%. We further explored the correlation between pre-existing immunity and response by transcriptome analysis of baseline tumor tissue. There was no difference between CR and non-CR tumors in baseline interferon gamma (IFN-γ), tumor inflammation (TIS) and $CD8^+$ T-cell effector signatures **(Fig. 3e)**. Thus, our data suggest that the addition of anti-CTLA-4 to PD-1 blockade can induce pathological complete response in tumors irrespective of baseline $CD8^+$ T-cell immunity.

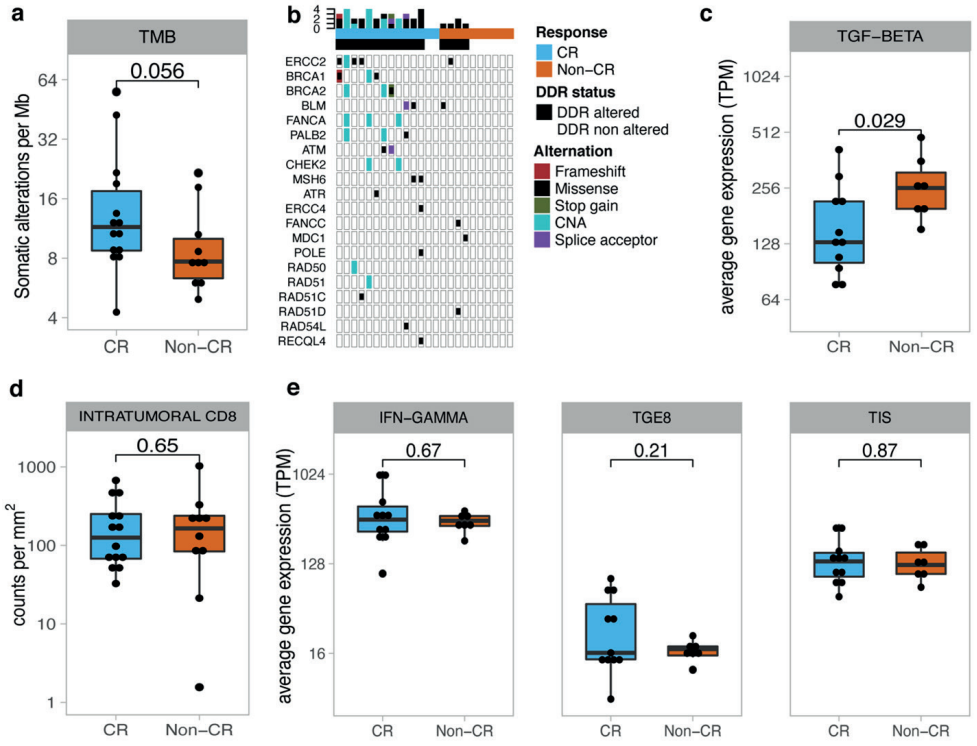


Figure 3. Pathological complete response to ipilimumab plus nivolumab is independent of baseline CD8+ T-cells and inflammatory signatures.

A. Association between response groups and tumor mutational burden (TMB). TMB is measured as the number of somatic, non-synonymous and exonic mutations, based on whole-exome sequencing (WES). The panel shows the TMB distribution in CR (n=14) and non-CR (n=10) tumors. **B.** Association between somatic alterations in DNA Damage Response and Repair (DDR) genes [17] and response (CR; n=14, non-CR; n=10). We excluded variants with a clinical significance labelled as Benign or Likely Benign (CLINVAR annotation), variants previously reported as population/germline SNPs (TOPMED and CAF annotations with alternate allele frequency >5%), and variants annotated as Tolerated with the SIFT predicted score. Patients were stratified by response (CR and non-CR) and DDR status (DDR altered and DDR non-altered) in a contingency table, and statistical significance was tested using a two-sided Fisher's exact test. **C.** Comparison of average TGF-β induced genes [18] at baseline between CR (n=11) and non-CR (n=7) tumors. **D.** Baseline comparison of CD8 density per mm2 between CR (n=14) and non-CR (n=10) tumors. Baseline intratumoral CD8 density per mm2 was analyzed by quantitative multiplex immunofluorescence. **E.** Comparison of average gene expression for the Interferon gamma signature (IFN, Ayers et at. 2017), eight-gene cytotoxic T cell transcriptional signature (tGE8, Mariathasan et al. 2018) and tumor inflammation (TIS, Damotte et al. 2019) signatures between CR (n=11) and non-CR (n=7) tumors. Boxplots in all panels represent the median and 25th and 75th percentiles. The whiskers expand from the hinge to largest value not exceeding 1.5× IQR from the hinge. Complete responders are marked in blue, while non-responders are displayed in orange. A two-sided t-test was used for comparisons of the gene expression features between CR and non-CR tumors, and a two sided Mann-Whitney test was used for comparison of the TMB between CR and non-CR tumors. The P-value is presented in-between boxplots. All statistical tests were two-sided. No adjustments were made for multiple comparisons. Abbreviations: CR: complete response, non-CR: non complete response, TMB: Tumor mutational burden, SNP: Single nucleotide polymorphism, CNA: Copy number aberration, DDR: DNA damage response and

repair, TGF-beta: Transforming growth factor beta, IFN-gamma: Interferon Gamma, tGE8: eight-gene cytotoxic T cell transcriptional signature, TIS: tumor inflammation signature

Baseline B-cell presence has been associated with immunotherapy response and prognosis in several malignancies[20,24,25], although other studies suggest immune-suppressive capacities of the intratumoral B-cell compartment[26,27]. Upon exploratory analysis of differentially expressed genes, we found that upregulation of B-cell related genes at baseline correlated with non-response **(Fig. 4a)**. Furthermore, these differentially expressed B-cell genes were found to positively correlate with $CD20^+$ B-cell counts in multiplex immunofluorescent analysis **(Fig. 4b)**. In non-responding tumors, stromal B-cells were more abundant than in responding tumors (p=0.043, **Fig. 4c)**, whereas the density of B-cells in the tumor compartment was numerically higher (p=0.07). The increased B-cell presence in non-CR tumors was observed irrespective of pre-existing $CD8^+$ T-cell immunity **(Extended Data Fig. 3)**. B-cells have been studied in conjunction with tertiary lymphoid structure (TLS) composition[20]. We used multiplex immunofluorescent staining and image analysis to visualize TLS composition and quantify TLS maturation stages **(Fig. 4d,e).** No correlation was observed between baseline TLS numbers and response **(Fig. 4f),** although immature TLS were higher in non-CR **(Extended Data Fig. 3)**. On-treatment analysis revealed enrichment in TLS (p<0.001, **Fig. 4g)** in CR tumors. Enrichment was observed across all TLS maturation stages **(Extended Data Fig. 4)**. Expression of TLS-related genes and a recently established signature[19] related to B-cell/TLS presence in melanoma increased upon treatment **(Extended Data Fig. 5)**, particularly in CR tumors. TLS development was dampened in patients having corticosteroids for immunotherapy toxicity management (p=0.014, **Fig. 4h,i)**, as observed in lung cancer[28]. TLS were further characterized by presence of $CD27^+$ B-cells **(Extended Data Fig. 6)**, T follicular helper cells, $BCL6^+CD4^-$ cells and expression of CXCL13 **(Extended Data Fig. 7)**. We observed no increase of $CD27^+$ B-cells and T follicular helper cells in TLS post-therapy. Early induction of these cells may have been missed due to the fact that samples were taken relatively late in the anti-cancer immunotherapy response, after tumor-clearance. To further study TLS dynamics, a bioinformatic algorithm was developed to annotate the TLS compartment in multiplex immunofluorescence data. Using this algorithm, we found that Tregs were reduced in TLS upon treatment with ipilimumab plus nivolumab **(Extended Data Fig. 8)**. These exploratory findings suggest that presence of B-cells and TLS at baseline does not predict for ipilimumab plus nivolumab response in UC, though TLS induction is observed in responding patients.

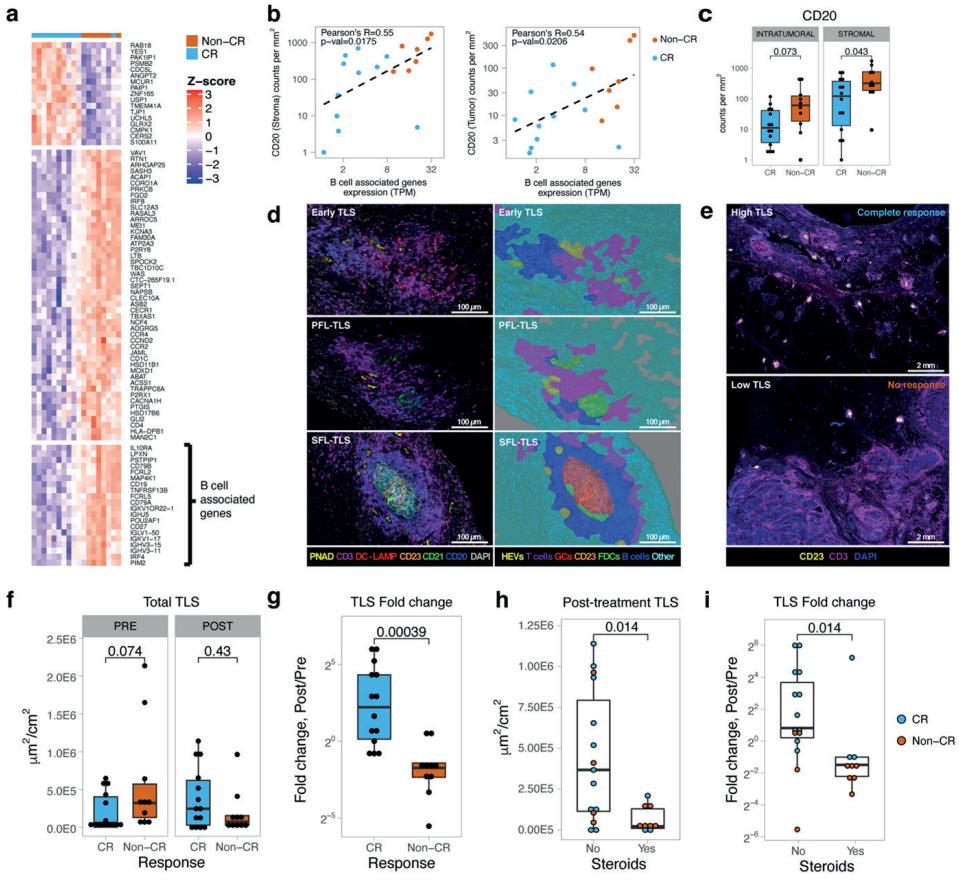


Figure 4. B-cell analysis and assessment of tertiary lymphoid structure dynamics upon preoperative ipilimumab plus nivolumab.

A. Hierarchical clustering of differentially expressed genes (FDR < 8%) at baseline between CR (blue; n=11) and non-CR (orange; n=7) tumors, showing up-regulation of B-cell-related genes in patients with non-CR. Gene expression levels are represented as counts per million (CPM), mean-centered and scaled (Z-scores). Higher expression levels (red) and lower expression levels (blue) are indicated per gene. For each gene, baseline differential expression between response groups was tested fitting a linear model to response using weighted least squares (limma R package). Contrasts that compared CR and non-CR tumor were fitted. For each gene the density-based False Discovery Rate (FDR) from the p-value distribution was computed to correct for multiple hypothesis testing. Only genes that scored an FDR < 8% are shown on the heatmap. **B.** The differentially expressed B-cell associated genes (derived from **A**) correlate with tumor and stroma CD20+ immune cell counts per mm^2 by multiplex immunofluorescence analysis. A two-sided Pearson's moment correlation test was used to analyze this correlation in CR (n=11) and non-CR (n=7). Pearson coefficient and correlation p-value are shown in the plot. **C** Baseline CD20+ immune cell densities per mm^2 in tumor and stroma by multiplex immunofluorescent analysis between CR (n=14) and non-CR (orange; n=10) tumors stratified by response categories for tumor and stroma. **D.** Tissue segmentation (PerkinElmer) was used to segregate TLS areas. Early-TLS display dense B cell clusters without follicular dendritic cells (FDCs) or germinal centers (GC). Primary follicle-like TLS (PFL-TLS) harbor FDC networks while lacking GC. Secondary follicle-like TLS (SFL-TLS) exhibit active GC reaction. **E.** Example images of the TLS landscape in a CR vs non-CR tumor after therapy **F.** TLS development was quantified as normalized TLS area (square microns

per tissue square centimeter). Pre-treatment and post-therapy samples were compared between CR (n=14) and non-CR (n=10). **G.** Immunotherapy-induced changes in normalized TLS area were assessed as fold change (post/pre) and compared between the CR (blue: n=14) and non-CR (orange: n=10) tumors. **H.** Normalized TLS area in post-therapy samples. between patients that received steroids (>20mg a day prior to surgery: n=9) and no steroids (n=15). A comparison was made using a two-sided Mann Whitney test. **I.** Immunotherapy-induced fold change in normalized TLS area were assesses as fold change (post/pre) in patients that received steroids (>20mg a day prior to surgery: n=9) and no steroids (n=15). Boxplots in all panels represent the median and 25th and 75th percentiles. The whiskers expand from the hinge to largest value not exceeding 1.5× IQR from the hinge. Complete responders are marked in blue, while non-responders are displayed in orange. All comparisons, except for the ones shown in panels A and B, were made using a two-sided Mann Whitney test. The P-value is presented in-between boxplots. All statistical tests were two-sided. No adjustments were made for multiple comparisons, except for the ones shown in panel A. Experiments and scorings related to the presented micrographs in D and E were conducted once. Abbreviations: CR: complete response, non-CR: non complete response, FDR: False discovery rate, TLS: Tertiary lymphoid structure.

We found that ipilimumab plus nivolumab is feasible and highly active as preoperative treatment in UC. Moreover, high pCR rates were observed in patients with locoregionally advanced disease. Previous studies testing preoperative anti-PD1/PD-L1 monotherapy in UC were mainly conducted in patients with less advanced tumors (cT2-3N0), showing encouraging pCR rates. Substantially lower pCR rates were observed with anti-PD-L1 in patients with more advanced disease[18]. Our results are in line with findings in locoregionally advanced melanoma, where the addition of anti-CTLA-4 to anti-PD-1 monotherapy before surgery broadened response and resulted in superior efficacy and outcome[1,29]. Furthermore, we found that response to preoperative ipilimumab plus nivolumab was independent of baseline $CD8^+$ T-cell presence and inflammatory signatures. This observation contrasts with findings from trials using anti-PD1/PD-L1 monotherapy, where pCR was primarily observed in tumors exhibiting pre-existing $CD8^+$ T-cell immunity[16,18]. A limitation of our study is the small sample size, which was powered to show feasibility of preoperative ipilimumab plus nivolumab and to determine an efficacy signal that would justify further clinical studies. Additionally, the treatment schedule can be further refined to find the optimal balance between toxicity and efficacy.

In conclusion, combined CTLA-4 plus PD-1 blockade may provide an effective preoperative treatment strategy in locoregionally advanced UC, regardless of pre-existing $CD8^+$ T-cell activity at baseline. The therapeutic efficacy in this cohort, with poor prognosis and high unmet clinical need, warrants further development through randomized trials.

4.2 METHODS

4.2.1 Study design and population

NABUCCO (ClinicalTrials.gov: NCT03387761) is an investigator-initiated, prospective, single-arm trial testing the feasibility of attenuated preoperative ipilimumab 3 mg/kg (day

1), ipilimumab 3 + nivolumab 1 mg/kg (day 22), and nivolumab 3 mg/kg (day 43) followed by cystectomy or nephro/urethrectomy with appropriate lymph node dissection[21] (see **Supplementary Study protocol**). A total of 24 patients with stage III resectable urothelial cancer (cT3-4aN0M0 and cT1-4aN1-3M0) according to AJCC guidelines were enrolled at the Netherlands Cancer Institute. Patients were ≥18 years old, cisplatin-ineligible or refused cisplatin-based chemotherapy, were anti-CTLA-4/PD-1/PD-L1 naïve, WHO performance score 0-1, and had diagnostic TUR blocks available. Key exclusion criteria were documented severe autoimmune disease, chronic infectious disease and use of systemic immunosuppressive medications. All patients provided written informed consent. Surgery was scheduled in weeks 9-11 from the first drug infusion. Baseline staging was based on diagnostic pathology and thorax/abdomen/pelvis computed-tomography (CT). CT-based response assessment occurred in week 8-10. Additional baseline and on-treatment MRI assessment was optional. Non-related and immune-related adverse events (irAEs) were graded and reported throughout the study. In case of serious irAEs after the second treatment cycle, the remaining dose of nivolumab was not administered to not interfere with the timing of surgery. The sponsor of the trial was the Netherlands Cancer Institute. Bristol-Myers Squibb provided the study drugs and funding for the trial. The trial protocol was designed and written by the authors (NvD, CUB, BvR and MSvdH) and is part of the Dutch uro-oncology study group (DUOS). A data safety monitoring board (DSMB) was established, consisting of Prof. T. Powles (medical oncologist, Barts Cancer Centre) and dr. J. Boormans (urologist, Erasmus Medical Centre). This study was approved by the institutional review board of the Netherlands Cancer Institute and was executed in accordance with the protocol and Good Clinical Practice Guidelines defined by the International Conference on Harmonization and the principles of the Declaration of Helsinki. After meeting the pre-specified primary and secondary endpoint in the current analysis (n=24), the study was expanded with new dose-exploratory cohorts (IRB approved September 2019).

4.2.2 Study endpoints and statistics

The primary endpoint of this trial was feasibility, testing whether preoperative ipilimumab and nivolumab is feasible within 12 weeks from first infusion and does not delay surgical resection, as this is an endpoint that is clinically meaningful for this population. We set the desired resection rate by 12 weeks at 90% of patients. The lower statistical boundary for futility was set at 60%. For 24 patients and one-sided alpha=0.025, the power of rejecting 60% resection rate, under the alternative of 90%, is 91%. Treatment efficacy (pCR) was selected as a secondary endpoint, and defined by the percentage of pathological complete response (complete absence of neoplastic cells, pT0N0) at surgical resection. For the efficacy endpoint we considered the disease burden of this cohort, which is higher than a typical neoadjuvant cohort, and cisplatin-ineligibility/refusal. We determined that a 40% pCR rate would be desirable, while 14% or lower would not warrant further investigation of treatment. Under

those assumptions, 24 evaluable patients provide 90% power to detect treatment efficacy with the one-sided alpha of 0.05. At least 7 responders are required. Histopathological assessment was performed on the resected primary tumor and LN specimens, and pathological staging was done according to international standards and protocols. In addition, percentage of residual vital tumor tissue was analyzed and calculated. Other secondary endpoints included translational endpoints and safety and are extensively described in the study protocol (see Study protocol in **Supplementary information)**. Unless specified otherwise, Wilcoxon signed ranked test was used for pairwise comparisons. R 3.4.4 was used for statistical analysis, together with the packages tidyverse 1.2.1, ggpubr 0.2.1.

4.2.3 PD-L1 assessment, multiplex immunofluorescence analyses and immunohistochemistry

Baseline PD-L1 immunohistochemistry was performed on FFPE sections by a PD-L1 IHC 22C3 1/40 dilution pharmDx qualitative immunohistochemical assay on a DAKO autostainer 48 system at Tergooi hospital laboratory. After staining, PD-L1 and Hematoxylin & Eosin slides were uploaded to SlideScore (www.slidescore.com) for manual scoring. An experienced uro-pathologist determined the combined positive score (CPS), PD-L1 positivity was qualified as CPS > 10[16]. Analysis of tumor immune cell infiltrates (anti-CD3 (1/400 dilution Clone P7, ThermoScientific), anti-CD8 (1/100 dilution Clone C8/144B, DAKO), anti-CD68 (1/500 dilution Clone KP1, Dako), anti-FoxP3 (1/50 dilution Clone 236A/47, Abcam), anti-CD20 (1/500 dilution Clone L26, Dako), anti-PanCK (1/100 dilution Clone AE1AE3, Thrermo Scientific) was performed by multiplex Immunofluorescence (mIF) technology on a ventana Discovery Ultra automated stainer using Perkin Elmer opal 7-color dyes,. In short, 3 µm FFPE sections were cut and heated at 75 °C for 28 minutes and subsequently deparaffinized in EZ prep solution, Ventana Medical Systems. Using Cell Conditioning 1 (CC1, Ventana Medical Systems), Heat-induced antigen retrieval was conducted at 95 °C for 32 min. Further analysis was done by VECTRA image acquisition (Akoya Biosciences, v3.0) and HALO (Indica Labs, v2.3) image analysis. Tumor and stroma regions were classified by HALO automated tissue segmentation. TLS composition was analyzed in tissue sections by 7-plex multiplex immunofluorescence[28] using the following antibodies and dilutions; CD21 (1:5000 2G9 Leica), DC-LAMP (1:1000 1010E1.01 Dendritics), CD23 (1:1000 SP3 Abcam), PNAd (1:5000 MECA-79 Biolegend), CD20 (1:5000 L26 Dako), CD3 (1:1000 SP7 ThermoScientific) and multispectral microscopy. Area segregation was done by Inform tissue segmentation algorithm. Additional information on antibodies used is provided in the Nature Research Reporting Summary.

For immunohistochemistry (including co-stainings), slides were cut at 3µm and dried overnight. Slides were transferred to the Ventana Discovery Ultra autostainer. Slides were then deparaffinised within the instrument and antigen retrieval was carried out at 95°C

for 64 minutes. For the double-staining of CD20 (yellow) followed by CD27 (purple), the CD20 was detected in the first sequence using clone L26 (1:800, 32 min at 37 °C, Agilent). CD20-bound antibody was visualized using Ready-to-Use anti-mouse NP (Ventana Medical systems) for 12 min at 37 °C followed by anti-NP AP (Ventana Medical systems) for 12 min at 37 °C, followed by the Discovery Yellow Detection Kit (Ventana Medical Systems). In the second sequence of the double-staining procedure, CD27 was detected using clone EPR8569 (Abcam, 1:3000 dilution, 32 min at 37 °C). CD27 was visualized using Ready-to-Use anti-mouse HQ (Ventana Medical systems) for 12 min at 37 °C followed by anti-HQ horseradish peroxidase (Ventana Medical systems) for 12 min at 37 °C, followed by the Discovery Purple Detection Kit (Ventana Medical Systems). Slides were counterstained with Hematoxylin and Bluing Reagent (Ventana Medical Systems). For the double-staining of CD4 (yellow) followed by Ready-to-Use BCL6 (purple), the CD4 was detected in the first sequence using clone SP35 (1:25, 32 min at 37 °C, Cell Marque). CD4-bound antibody was visualized using the Roche Diagnostics Ready-to-Use anti-mouse NP (Ventana Medical systems) for 12 min at 37 °C followed by anti-NP AP (Ventana Medical systems) for 12 min at 37 °C, followed by the Discovery Yellow Detection Kit (Ventana Medical Systems). In the second sequence of the double-staining procedure, BCL6 was detected using clone GI191E/A8 (Ready-to-use, Ventana Medical Systems, 32 min at 37 °C). BCL6 was visualized using the Roche Diagnostics Ready-to-Use anti-mouse HQ (Ventana Medical systems) for 12 min at 37 °C followed by the Roche Diagnostics Ready-to-Use anti-HQ horseradish peroxidase (Ventana Medical systems) for 12 min at 37 °C, followed by the Discovery Purple Detection Kit (Ventana Medical Systems). Slides were counterstained with Hematoxylin and Bluing Reagent (Ventana Medical Systems). CXCL13 was stained using the polyclonal goat R&D Systems antibody Clone AF801, in a 1:400 dilution. CXCL13 bound antibody was visualized using the Roche Diagnostics Ready-to-Use OmniMap anti-Goat HRP (Ventana Medical Systems) for 12 minutes at 37°C, followed by the Chromomap DAB detection kit. Slides were counterstained with Hematoxylin and Bluing Reagent (Ventana Medical Systems). All IHC slides were uploaded to SlideScore for manual scoring. Additional information on antibodies used is provided in the Nature Research Reporting Summary.

For biomarker analysis, tumors with complete response (CR, defined as pCR, pTis or pTaN0) were compared to non-CR tumors. We included non-invasive disease in the CR definition, as this is generally believed to be cured by surgery[30]. Figure panel colors are optimized for colorblind readers.

4.2.4 DNA sequencing

DNA and RNA were isolated in parallel from baseline and on treatment FFPE tumor material using the Qiagen AllPrep FFPE DNA/RNA isolation kit. Germline DNA was extracted from peripheral blood mononuclear cells using the QIAsymphony DSP DNA midi kit. Tumor

depositions in the lymph nodes were isolated by using laser microdissection. DNA sequencing was carried out following the Human IDT Exome Target Enrichment Protocol. Covaris shearing was used to fragment the DNA to 200-300 base pairs. KAPA HTP DNA Library Kit was used for library preparation. IDT Human Exome capture V1.0 was used for exome enrichment. Libraries were sequenced with 100 base pairs paired-end reads using the Hiseq 2500 (Illumina) with a high output mode PE 100. Burrows-Wheeler aligner 37 was used to align the raw reads to the human reference genome GRCh38. Duplicated reads were marked using MarkDuplicates (http://broadinstitute.github.io/picard), and lastly GATK BaseRecalibrator was used to recalibrate the quality scores. Somatic single-nucleotide variants (SNVs), short insertions and deletions (indels) were called for tumor samples matched with germline samples using Strelka 2.9.2[31]. Only variants with a high reliability score (HighEVS) and allele frequency > 5% were maintained, and annotated using SnpEff v4.3t (build 2017-11-24 10:18). Tumor mutational burden (TMB) was assessed as the number of somatic variants annotated as non-synonymous, non-intronic or non-intergenic. For biomarker discovery, we retrieved variants classified as indels, missense, frameshift, stop codon gain or loss, splice acceptors or donor, transcription ablation, exon loss, and structural interaction variant. Copy number calling was done using CNVkit v0.9.6[32] and copy number ratios were estimated against a pooled normal reference. Gene copy number ratios (cnr) were calculated with a weighted average copy number ratio per gene. Deep deletions were defined as $\log_2(cnr) < -0.7$, shallow deletions as $-0.7 \leq \log_2(cnr) < -0.5$ and Amplifications as $\log_2(cnr) > 1$. All genomic data was analyzed using the R packages VariantAnnotation v1.24.5 and ComplexHeatmap v1.17.1. For associations between somatic alterations in DNA Damage Response and Repair (DDR) genes[22] and response, we excluded variants with a clinical significance labelled as Benign or Likely Benign, variants previously reported as population/germline SNPs (TOPMED and CAF annotations with alternate allele frequency >5%), and variants annotated as Tolerated with the SIFT predicted score. We explored alterations in DNA mismatch repair machinery (*MLH1, MSH2, MSH6, PMS1, PMS2*), nucleotide excision repair (*ERCC2, ERCC3, ERCC4, ERCC5*), homologous recombination (*BRCA1, MRE11A, NBN, RAD50, RAD51, RAD51B, RAD51D, RAD52, RAD54L*), Fanconi anemia (*BRCA2, BRIP, FANCA, FANCC, PALB2, RAD51C, BLM*), checkpoint (*ATM, ATR, CHEK1, CHEK2, MDC1*) and others (*POLE, MUTYH, PARP1, RECQL4*). LN and primary tumors with discrepant response were sequenced using 3 different pools to reach >150x exomic coverage. Somatic SNVs and short indels for each tumor – germline and LN - germline sample pairs were called using Strelka. For each patient, we retrieved all variants that scored at least in one of the tumor or lymph nodes variant calling a high reliability score (HighEVS), recomputed the allele frequencies from the bam files, and filtered by allele frequency > 5%.

4.2.5 RNA sequencing

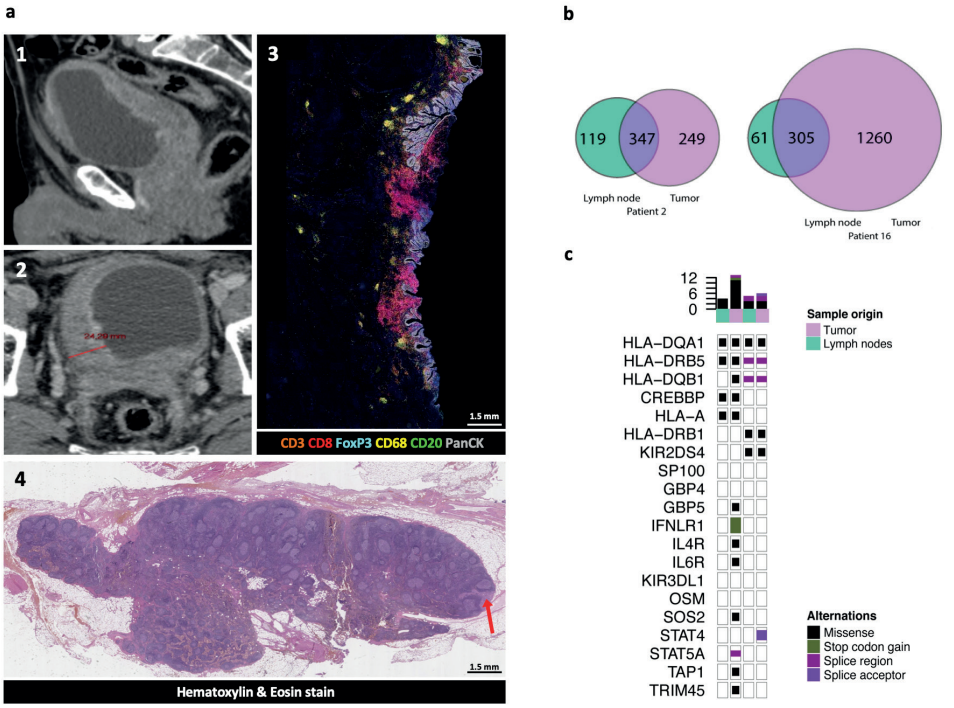
The TruSeq Stranded mRNA Sample Prep Kit from Illumina was used to generate strand-specific libraries according to the manufacturer's instructions. Twelve cycles of PCR were done to amplify the 3' end-adenylated and adapter-ligated RNA. Quality and quantity of the total RNA from FFPR was assessed by the 2100 Bioanalyzer using a 7500 Nano chip (Agilent, Santa Clara, CA). The percentage of RNA fragments > 200nt fragment distribution values (DV200) were determined using the region analysis method according to the manufacturer's instructions manual (Illumina, technical-note-470-2014-001). The libraries were subsequently diluted and pooled equimolar in a multiplex sequencing pool. Storage was done at -20 degrees Celsius. The pooled libraries were enriched for target regions using the probe Coding Exome Oligos set (CEX, 45MB) according to the manufacturer instruction (Illumina, #1000000039582v01). Briefly, cDNA libraries and magnetic bead bound capture probes were combined followed by hybridization using a denaturation step of 95°C for 10 minutes and an incubation step from 94°C to 58 °C having a ramp of 18 cycles with 1 minute incubation and 2°C per cycle. The enriched fraction was subjected to two stringency washes, an elution step and a second round of enrichment followed by a cleanup using AMPure XP beads (Beckman, A63881) and PCR amplification of 10 cycles. The target enriched pools were analyzed on a 2100 Bioanalyzer using a 7500 chip (Agilent, Santa Clara, CA), diluted and subsequently pooled equimolar into a multi-plex sequencing pool. A HiSeq 2500 System was used to sequence libraries with 65 base pair single-end reads in Illumina high output mode using V4 chemistry (Illumina Inc., San Diego). HISAT2 aligner was used to align the raw reads to the human reference genome GRCh38 and to quantify gene expression using default parameters. For quality control, samples with a low cDNA yield, low sequencing coverage (reads < 9000000) and fraction of aligned reads to human reference genome < 85% were filtered out (6 baseline samples). Genes with low gene expression variability (null counts in >80% of the samples and/or read counts mean < 1) were filtered out. EdgeR v3.20.9[33] was used to calculate the normalization factors In order to account for variable library sizes. Only genes with maximum counts per million (CPM) > 1, and with a CPM >1 in at least 3 samples, were considered for biomarker analysis (~7000 genes). Differential expression (DE) was tested using limma v3.34.9[34]. First, we used the voom function to account for different weights from the mean-variance relationships. Then, we built linear models to test DE between different baseline/on-treatment response groups. We corrected for multiple hypothesis testing by modelling the density-based False Discovery Rate (FDR) from the p-value distribution using fdrtool[35]. Genes that scored an FDR < 8% were considered as significant.

4.2.6 Bioinformatic algorithm to detect TLS in Multiplex immunofluorescence

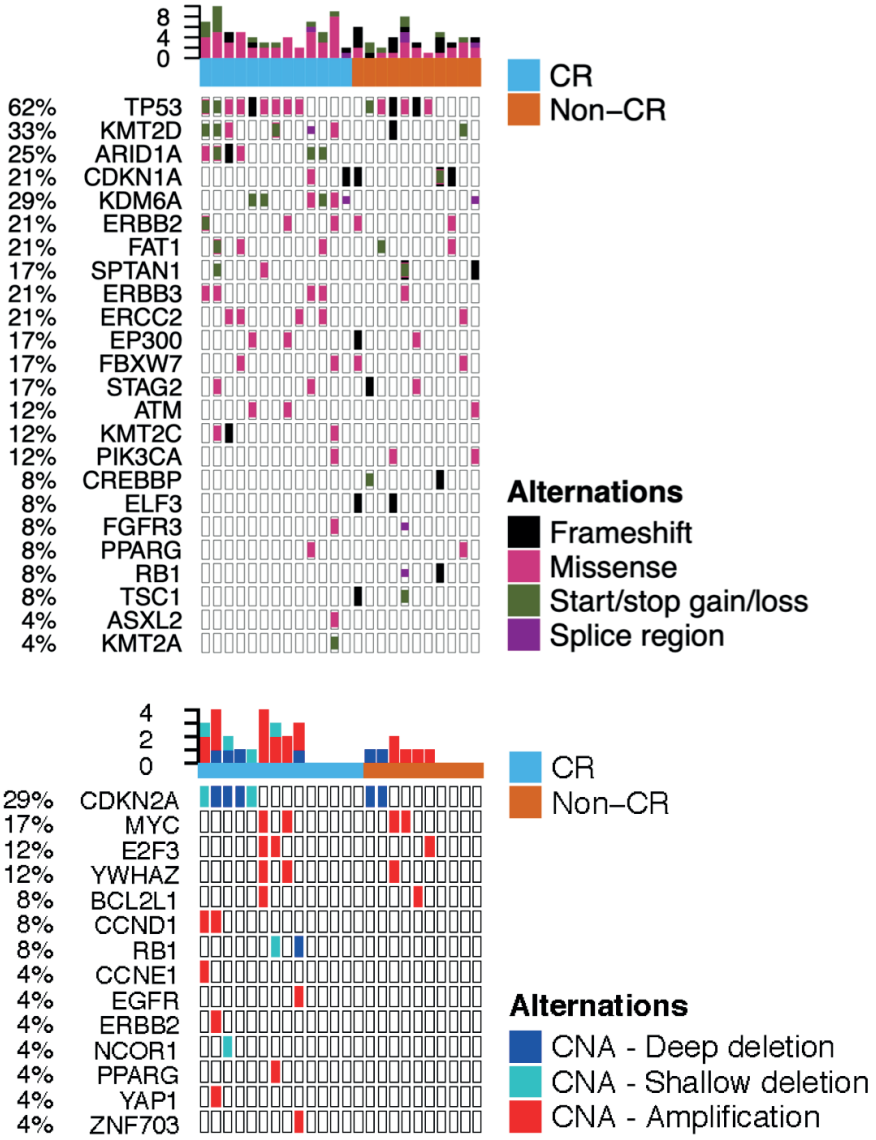
Tertiary lymphoid structure (TLS) clusters in multiplex immunofluorescence (CD3, CD8, FoxP3, CD20, CD68) data were detected by using DBSCAN v1.1-4. The cell phenotypes $CD20^{+}$, $CD3^{+}CD8^{-}FoxP3^{-}$ and $CD3^{+}FoxP3^{+}CD8^{-}$ were used to identify putative TLS

clusters in stromal regions. These same three phenotypes were quantified residing either inside TLS detected regions or in stromal/tumoral regions. Mann-Whitney U tests were used to test for difference in TLS composition between pre/post treatment samples of phenotypes frequencies $CD20^{+}$, $CD3^{+}CD8^{-}FoxP3^{-}$ and $CD3^{+}CD8^{-}FoxP3^{+}$. The difference in TLS T-reg($CD3^{+}CD8^{-}FoxP3^{+}$)/Naïve T-cell ($CD3^{+}CD8^{-}FoxP3^{-}$) ratio was explored with a Mann-Whitney U-test between pre/post samples of CR and Non-CR. The $CD3^{+}CD8^{-}FoxP3^{-}$ and CD20 location ratios (in TLS/in tumor+stroma) was tested for difference in ratio between pre/post samples in CR and NCR with a Mann-Whitney U test.

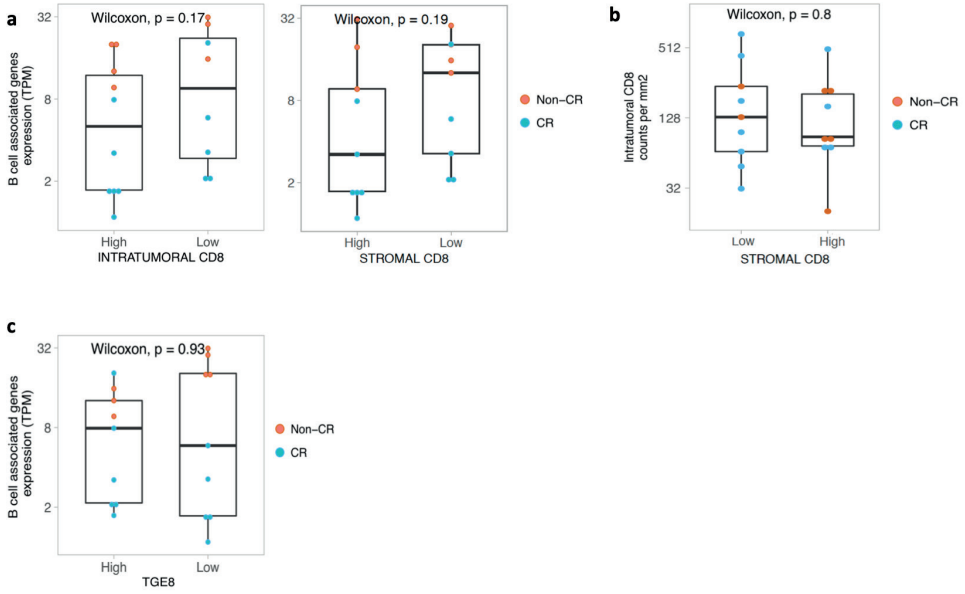
4.3 EXTENDED DATA



Extended Data Fig. I Assessment of response discrepancies between the responding primary bladder tumor and unresponsive local lymph node micrometastases upon ipilimumab plus nivolumab.
a1-2. Imaging shows a cT4aN1 bladder tumor (enlarged lymph node not shown). Pathological assessment revealed a complete response in the bladder, as shown by multiplex immunofluorescent analysis **(a-3)** and a non-responding lymph node micrometastasis, annotated by a red arrow **(a-4).** Experiments and scorings related to the presented micrographs were conducted once. **b,** Whole-exome sequencing was employed to assess whether lymph node micrometastases are genetically distinct from the primary tumor and could explain unresponsiveness (n=2 patients). The figure shows overlap between somatic variants identified in the baseline tumor (purple) and the post-treatment lymph node metastasis (green). **c,** Genomic alterations in genes related to interferon gamma signaling, JAK/STAT signaling or antigen presentation machinery for baseline tumor samples (purple, n=2) and matching lymph nodes metastases (green, n=2). Sample labels and genetic alteration type are displayed in figure legends. No specific genetic cause for resistance could be identified in the discrepant mutations.

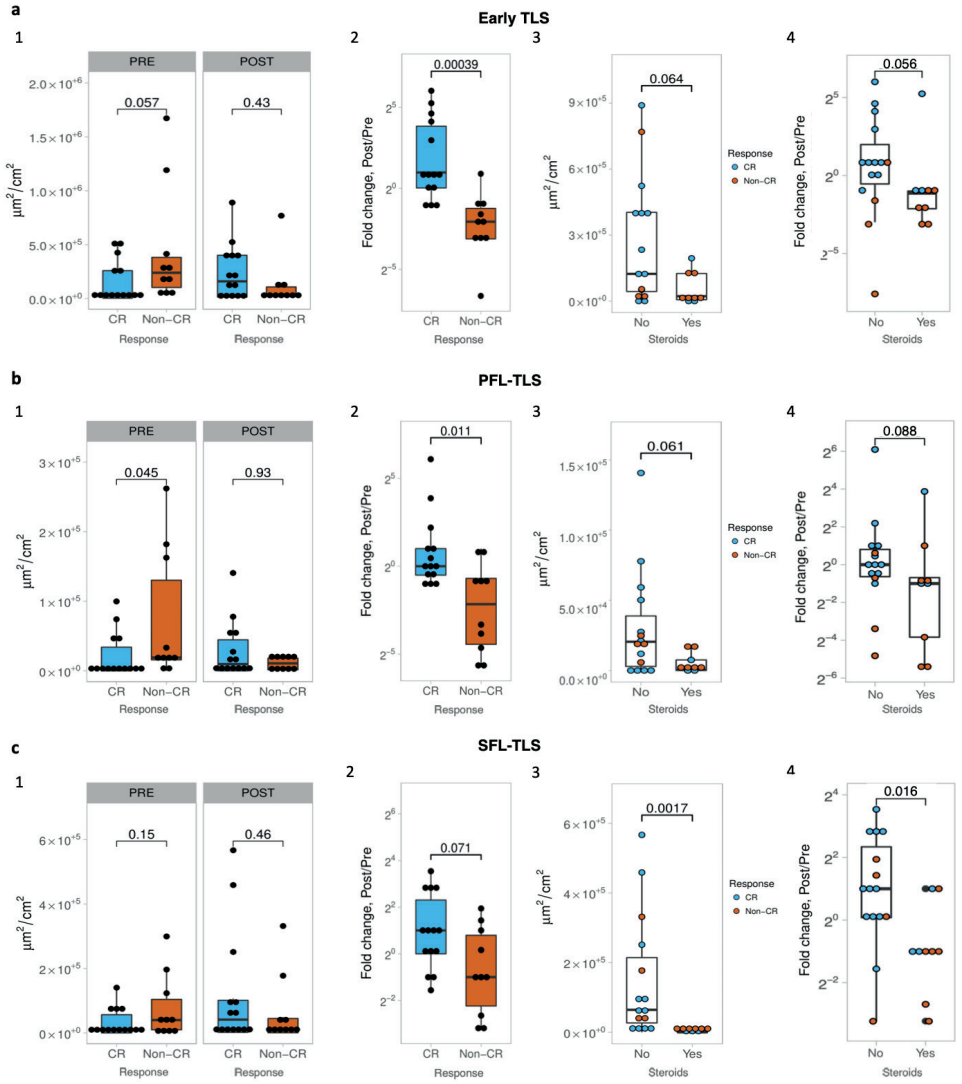


Extended Data Fig. 2 Analysis of baseline genomic alterations by whole-exome sequencing.
Whole-exome sequencing of pretreatment tumor tissue and germline DNA was employed to identify somatic mutations in baseline tumors (n=24). An oncoprint figure of genetic alterations in significantly mutated bladder cancer genes (by TCGA, Robertson et al, Cell 2018) is shown. Alterations are clustered by response categories, including 14 CR and 10 non-CR tumors. Sample labels and genetic alteration type are displayed in the figure. Abbreviations: CR: complete response, non-CR: non complete response, CNA: copy number alterations.



Extended Data Fig. 3 Assessment of interdependence between baseline B cell presence and CD8 T cell infiltration.

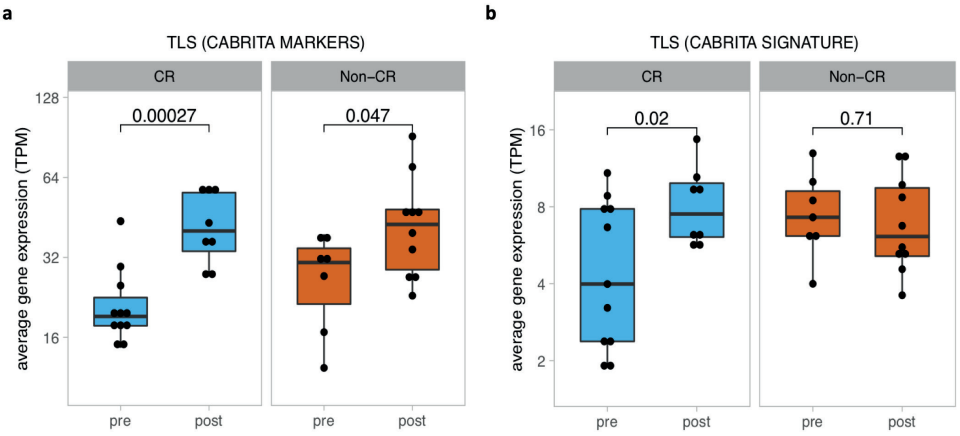
a, Average expression of B-cell-related differentially expressed genes at baseline stratified by intratumoral and stromal CD8+ higher and lower than median groups on multiplex immunofluorescence (mIF). Gene expression levels are represented as transcripts per million (TPM), mean-centered and scaled (Z-scores). **b,** Intratumoral CD8 density per mm2 by multiplex immunofluorescence, stratified by stromal CD20 higher and lower than median groups on multiplex immunofluorescence (mIF). **c,** Average expression of B-cell-associated genes at baseline stratified by average TGE8 signature groups (higher or lower than median) showing similar expression of B-cell-related genes in the TGE8 signature groups. Gene expression levels are represented as transcripts per million (TPM), mean-centered and scaled (Z-scores). A Wilcoxon signed tank test was used for all comparisons. The P-value is presented in-between boxplots. All statistical tests were two-sided. Analyses include CR (blue; n=11) and non-CR (orange; n=7) tumors. All boxplots display the median and 25th and 75th percentiles. The whiskers expand from the hinge to largest value not exceeding 1.5× IQR from the hinge. No adjustments were made for multiple comparisons. Abbreviations: CR: complete response, non-CR: non complete response.



Extended Data Fig. 4 Dynamics of tertiary lymphoid structure spectrum upon immunotherapy for response and steroid groups.

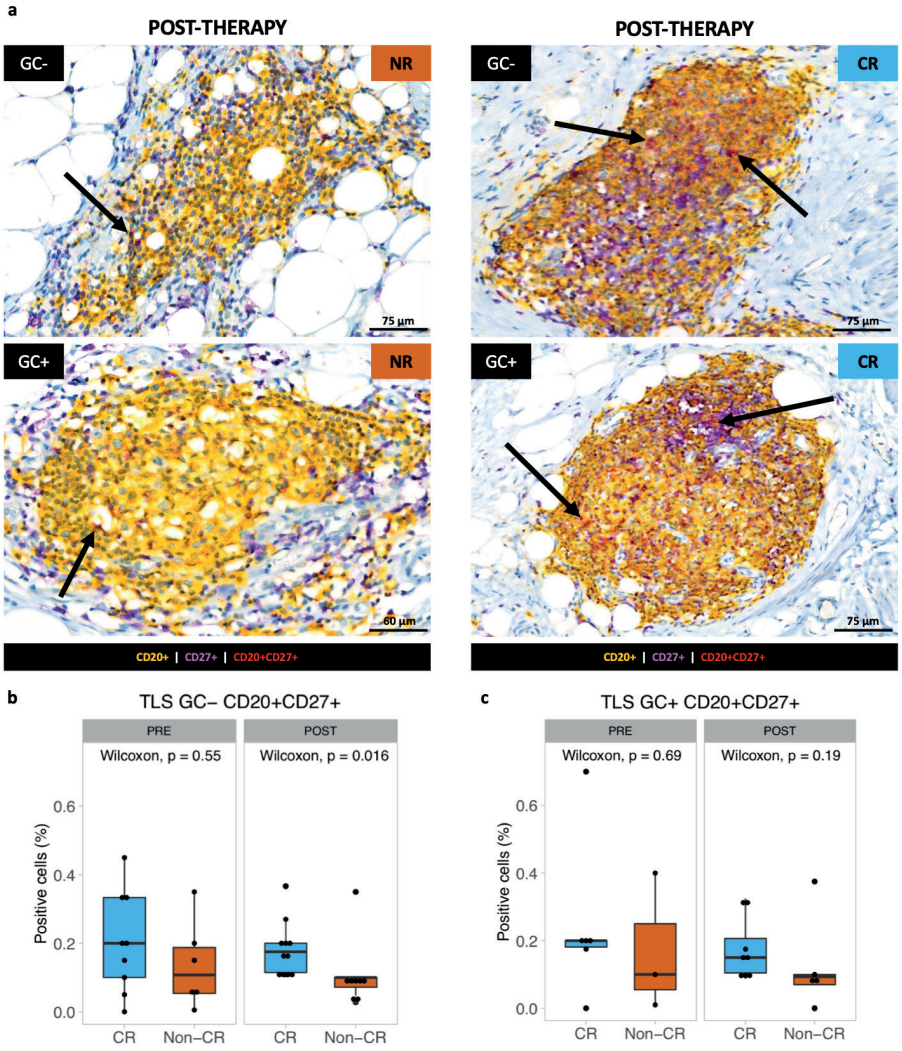
Upon multiplex immunofluorescent staining and segregation of tertiary lymphoid structure (TLS) areas, **a,** Early-TLS, **b,** Primary follicle-like TLS and **c,** Secondary follicle-like TLS were quantified as normalized TLS area (square microns per tissue square centimeter). For each TLS maturation stage, four different analysis were performed; 1) Comparison of normalized TLS areas in baseline and post-therapy samples between response groups by Mann Whitney test, 2) Normalized TLS area assessed as fold change (post/pre) upon treatment between response groups, 3) Normalized TLS area comparison in post-therapy samples between patients receiving steroids and no steroids, 4) Normalized TLS area assessed as fold change (post/pre) upon treatment in patients that received steroids and no steroids. Unless otherwise noted, all boxplots display the median and 25th and 75th percentiles. The whiskers

expand from the hinge to largest value not exceeding 1.5× IQR from the hinge. A Mann Whitney test was used for comparisons between response and steroid groups. The P-value is presented in-between boxplots. All statistical tests were two-sided. Analyses have 14 CR and 10 non-CR, or 9 steroids (>20mg a day prior to surgery) and 15 no steroid patients. No adjustments were made for multiple comparisons. Abbreviations: TLS: tertiary lymphoid structures, CR: complete response, non-CR: no complete response.

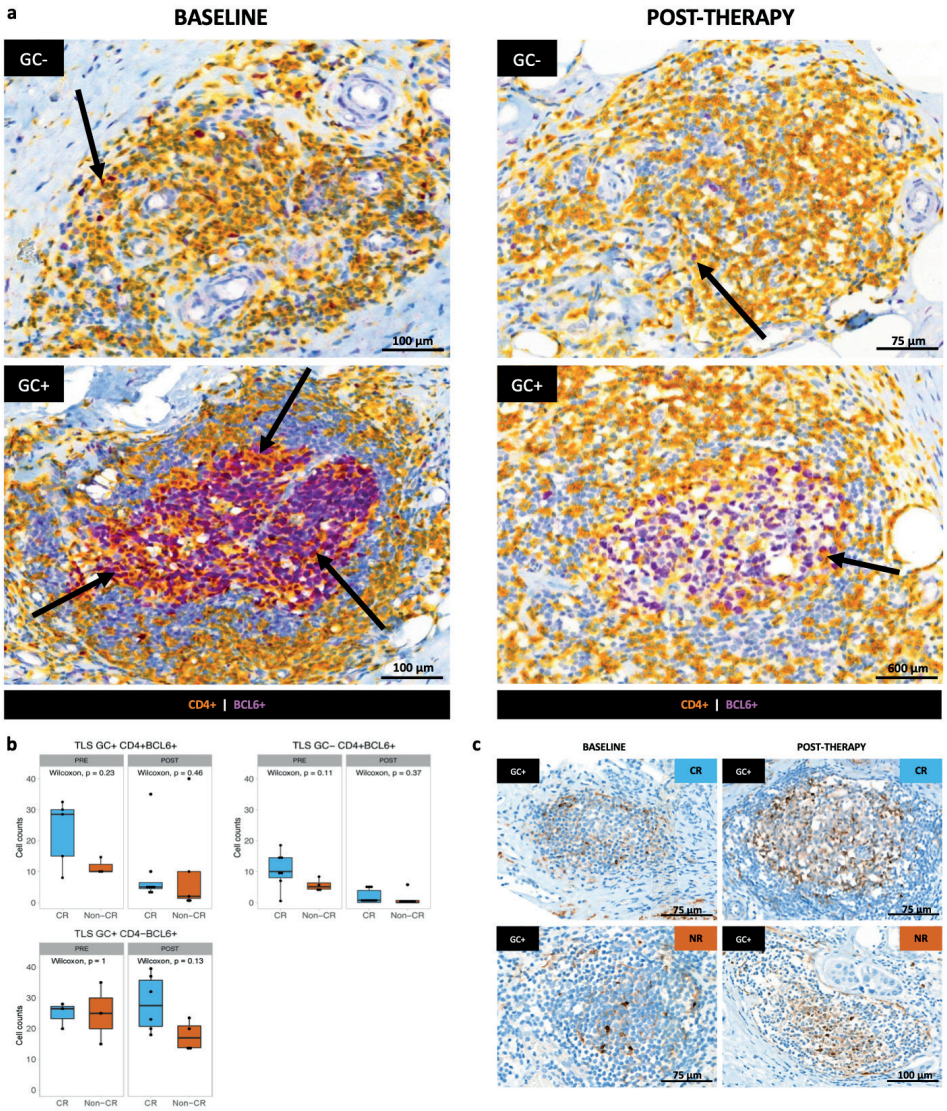


Extended Data Fig. 5 Exploratory assessment of TLS markers and a TLS signature upon immunotherapy for response groups.

a, Average gene expression for TLS-related genes (CCL19, CCL21, CXCL13, CCR7, SELL, LAMP3, CXCR4, CD86, BCL6), as published by Cabrita et al. 19. Pre- and post-treatment samples were compared for CR (n=11 pre-treatment, n=8 post-treatment) and non-CR (n=7 pre-treatment, n=10 post-treatment) tumors using a two-sided t-test. Baseline expression was not significantly different between CR and non-CR (p=0.28). **b,** TLS gene signature (Cabrita et al., Nature 2020) derived from genes specifically upregulated in CD8+CD20+ metastasized melanoma tumors (CD79B, CD1D, SKAP1, CETP, EIF1AY, RBP5, PTGDS). LAT and CCR6 genes were lowly expressed and thus removed from the analysis. Gene signatures were compared between baseline and post- treatment samples for CR and non-CR tumors. Baseline expression was not significantly different between CR and non-CR (p=0.05). Boxplots in all panels represent the median and 25th and 75th percentiles. The whiskers expand from the hinge to largest value not exceeding 1.5× IQR from the hinge. CR samples are marked in blue, while non-CR samples are displayed in orange. A t-test test was used for comparisons between CR and non-CR. The p-value is presented in-between boxplots. All statistical tests were two-sided. All analyses involved 11 CR and 7 non-CR tumors in pre-treatment samples, and 10 CR and 8 non-CR tumors in post-treatment samples. Abbreviations: TLS: tertiary lymphoid structures, CR: complete response, Non-CR: non complete response.



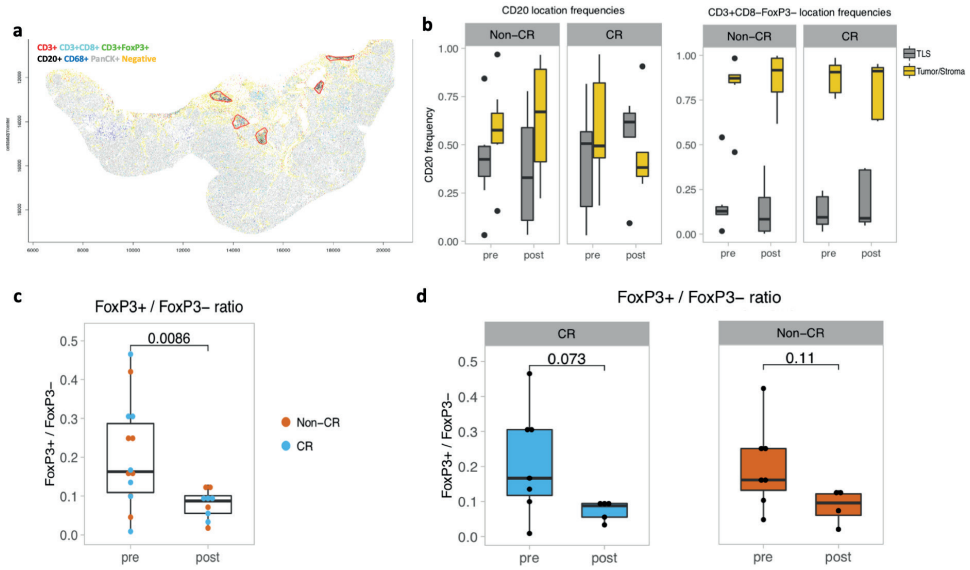
Extended Data Fig. 6 Assessment of CD27+ B-cells in TLS upon immunotherapy for response groups. **a,** Example images of CD20 (yellow) and CD27 (purple) IHC co-staining, revealing CD27+ B-cells (CD20+CD27+) in red, as indicated by the black arrow. Experiments and scorings related to the presented micrographs were conducted once. **b/c.** Baseline and post-treatment comparison of the mean percentage of CD20+CD27+ cells in germinal center (GC) negative (b) and GC+ **(c)** TLS per patient between CR (n=9 GC- pre-treatment, n=6 GC+ pre-treatment, n=11 GC- post-treatment, n=8 GC+ post-treatment) and non-CR (n=6 GC- pre-treatment, n=3 GC+ pre-treatment, n=8 GC- post-treatment, n=5 GC+ post-treatment) tumors. The percentage CD20+CD27+ cells in the CD20+ population in TLS was estimated by a pathologist (L.S.). Boxplots represent the median and 25th and 75th percentiles. The whiskers expand from the hinge to largest value not exceeding 1.5× IQR from the hinge. Complete responders are marked in blue, while non-responders are displayed in orange. A Wilcoxon signed tank test was used to compare the percentage of CD20+CD27+ between CR and non-CR tumors. The P-value is presented in-between boxplots. All statistical tests were two-sided. No adjustments were made for multiple comparisons. Abbreviations: CR: complete response, NR: no response, GC: germinal center.



Extended Data Fig. 7 Assessment of CD4+ T-cells and CD4+BCL6+ follicular T helper cells in tumor and TLS regions upon immunotherapy.

a, Example images of CD4 (yellow) and BCL6 (purple) IHC co-stainings, showing CD4+BCL6+ follicular T helper cells pre- and post-treatment, characterized by a purple nucleus and deep orange cytoplasmatic staining. **b,** Mean absolute CD4+BCL6+ cell counts in co-stains for mature and immature TLS and CD4−BCL6+ cell counts for mature TLS only; for CR tumors (n=7 GC- pre-treatment, n=5 GC+ pre-treatment, n=10 GC- post-treatment, n=8 GC+ post-treatment) and non-CR tumors (n=4 GC- pre-treatment, n=3 GC+ pre-treatment, n=8 GC- post-treatment, n=5 GC+ post-treatment). Co-stainings were assessed and scored (number of cells per TLS) by a pathologist. **c,** Representative example images of CXCL13 (brown) in TLS for CR and non-CR tumors

by immunohistochemistry in pre- and post-treatment specimens. CXCL13 positivity is clearly present in TLS, emphasizing that TLS are characterized by CXCL13 expression. No post-treatment differences were observed between response groups (Quantified data not shown). Unless otherwise noted, all boxplots display the median and 25th and 75th percentiles. The whiskers expand from the hinge to largest value not exceeding 1.5× IQR from the hinge. Complete responders are marked in blue, while non-responders are displayed in orange. A Wilcoxon signed tank test was used to compare the CD4+BCL6+ counts between CR and non-CR tumors. The P-value is presented in-between boxplots. All statistical tests were two-sided. No adjustments were made for multiple comparisons. Experiments and scorings related to the presented micrographs in A and C were conducted once. Abbreviations: CR: complete response, non-CR: non complete response, GC: germinal center.



Extended Data Fig. 8 Exploratory assessment of cellular distribution within TLS regions using a bioinformatic algorithm to analyse multiplex immunofluorescence slides.

a, Data map displaying immune cell subsets by digital multiplex immunofluorescent analysis in baseline tissue in a spatial context. A bioinformatic algorithm (methods) was developed to identify and segment TLS-like structures, as annotated in red in the example map. Tumor is depicted in gray (panCK) **b,** Distribution of CD3+CD8− and CD20+ cells over the analysed area, between algorithm assigned tumor bed (orange) and TLS (grey) pre- and post-therapy, in CR (n=7 pre, n=5 post) and non-CR (n=7 pre, n=4 post) patients. **c,** The ratio of FOXP3+/FOXP3- cells was calculated within the TLS CD3+CD8- compartment upon algorithmic TLS segmentation and quantitation of immune cell subsets in multiplex immunofluorescence images (Methods). Pre-treatment and post-therapy samples were compared in the complete cohort (C; n=14 pre, 9 post) p=0.0086) or between CR (n=7 pre, 5 post; p=0.073) and non-CR (n=7 pre, 4 post; p=0.11) d, by Mann Whitney test. Unless otherwise noted, all boxplots display the median and 25th and 75th percentiles. The whiskers expand from the hinge to largest value not exceeding 1.5× IQR from the hinge. Complete responders are marked in blue, while non-responders are displayed in orange. A Wilcoxon signed tank test was used to compare the CD4+BCL6+ counts between CR and non-CR tumors. The P-value is presented in-between boxplots. All statistical tests were two-sided. No adjustments were made for multiple comparisons. Abbreviations: CR: complete response, non-CR: non complete response, GC: germinal center.

4.4 SUPPLEMENTARY TABLES

Supplementary Table 1. Immune-related adverse events registered up to data cut-off

-Immune-related adverse events	All grade, N (%)	Grade 3/4, N (%)
Any IRAE	**24 (100%)**	**13 (55%)**
Lipase increased	8 (33%)	6 (25%)
Fatigue	6 (25%)	1 (4%)
Alanine aminotransferase increased	5 (21%)	3 (12%)
Diarrhea	5 (21%)	3 (12%)
Aspartate aminotransferase increased	4 (17%)	1 (4%)
Skin rash	5 (21%)	0 (0%)
Alkaline phosphatase increased	3 (12%)	1 (4%)
Colitis	3 (12%)	2 (8%)
Gamma-glutamyltransferase increased	3 (12%)	2 (8%)
Headache	3 (12%)	0 (0%)
Hyperthyreoidism	3 (12%)	0 (0%)
Serum amylase increased	3 (12%)	0 (0%)
Adrenal insufficiency	2 (8%)	0 (0%)
Hypothyreoidism	2 (8%)	0 (0%)
Toxicoderma, DRESS syndrome	1 (4%)	1 (4%)
Blood bilirubin increased	1 (4%)	1 (4%)
Peripheral motor neuropathy	1 (4%)	1 (4%)
Hemolysis	1 (4%)	1 (4%)
Hyperglycemia	1 (4%)	1 (4%)

4.5 SUPPLEMENTARY INFORMATION

Because of printing limitations, the NABUCCO study protocol is only available online: https://www.nature.com/articles/s41591-020-1085-z#Sec12

DATA AVAILABILITY STATEMENT

DNA and RNA sequencing data has been deposited in the European Genome-phenome Archive under the accession code EGAS00001004521 and will be made available on reasonable request for academic use and within the limitations of the provided informed consent by the corresponding author upon acceptance. Every request will be reviewed by the institutional

review board of the NKI; the researcher will need to sign a data access agreement with the NKI after approval. Sequencing data corresponds with Figure 3 and Figure 4. Multiplex immunofluorecence raw quantification data corresponding to Figure 3d and 4 can be made available on reasonable academic request within the limitations of informed consent by the corresponding author upon acceptance.

ACKNOWLEDGEMENTS

We would like to acknowledge the clinical trial data managers (C. Hagenaars and J. Kant), clinical trial nurses (A. Lechner, E. van der Laan and S. van der Kolk), the Genomics Core Facility and the Core Facility Molecular Pathology and Biobanking (I. Hofland, S. Cornelissen, L. Braaf and J. Sanders) for support, all at the Netherlands Cancer Institute. We also thank B. Stegenga at Bristol-Myers Squibb. MvdB was financially supported by the Swiss National Science Foundation (CRSII5_177208 and 310030_175565), the University of Zurich Research Priority Program "Translational Cancer Research", the Cancer Research Center Zurich, and Worldwide Cancer Research (18-0629).

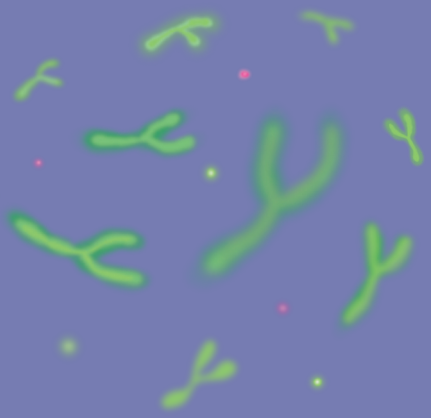
REFERENCES

1. Rozeman, E.A., *et al.* Identification of the optimal combination dosing schedule of neoadjuvant ipilimumab plus nivolumab in macroscopic stage III melanoma (OpACIN-neo): a multicentre, phase 2, randomised, controlled trial. *The Lancet. Oncology* (2019).
2. Chalabi, M., *et al.* Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. *Nature medicine* (2020).
3. Stein, J.P., *et al.* Radical cystectomy in the treatment of invasive bladder cancer: long-term results in 1,054 patients. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **19**, 666-675 (2001).
4. Zargar, H., *et al.* Multicenter Assessment of Neoadjuvant Chemotherapy for Muscle-invasive Bladder Cancer. *European urology* (2014).
5. Advanced Bladder Cancer (ABC) meta-analysis group. Neoadjuvant chemotherapy in invasive bladder cancer: update of a systematic review and meta-analysis of individual patient data. *European urology* **48**, 202-205 (2005).
6. Bellmunt, J., *et al.* Pembrolizumab as Second-Line Therapy for Advanced Urothelial Carcinoma. *N Engl J Med* **376**, 1015-1026 (2017).
7. Balar, A.V., *et al.* Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: a single-arm, multicentre, phase 2 trial. *Lancet* **389**, 67-76 (2017).
8. Sharma, P., *et al.* Nivolumab in metastatic urothelial carcinoma after platinum therapy (CheckMate 275): a multicentre, single-arm, phase 2 trial. *The Lancet. Oncology* **18**, 312-322 (2017).
9. Rosenberg, J.E., *et al.* Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *Lancet* **387**, 1909-1920 (2016).
10. Massard, C., *et al.* Safety and Efficacy of Durvalumab (MEDI4736), an Anti–Programmed Cell Death Ligand-1 Immune Checkpoint Inhibitor, in Patients With Advanced Urothelial Bladder Cancer. **34**, 3119-3125 (2016).
11. Powles, T., *et al.* Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial. *Lancet* **391**, 748-757 (2018).
12. Sharma, P., *et al.* Nivolumab Alone and With Ipilimumab in Previously Treated Metastatic Urothelial Carcinoma: CheckMate 032 Nivolumab 1 mg/kg Plus Ipilimumab 3 mg/kg Expansion Cohort Results. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **37**, 1608-1616 (2019).
13. Balar, A.V., *et al.* First-line pembrolizumab in cisplatin-ineligible patients with locally advanced and unresectable or metastatic urothelial cancer (KEYNOTE-052): a multicentre, single-arm, phase 2 study. *Lancet Oncol* (2017).
14. Carthon, B.C., *et al.* Preoperative CTLA-4 blockade: tolerability and immune monitoring in the setting of a presurgical clinical trial. *Clinical cancer research : an official journal of the American Association for Cancer Research* **16**, 2861-2871 (2010).
15. Liakou, C.I., *et al.* CTLA-4 blockade increases IFNgamma-producing CD4+ICOShi cells to shift the ratio of effector to regulatory T cells in cancer patients. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 14987-14992 (2008).
16. Necchi, A., *et al.* Pembrolizumab as Neoadjuvant Therapy Before Radical Cystectomy in Patients With Muscle-Invasive Urothelial Bladder Carcinoma (PURE-01): An Open-Label, Single-Arm, Phase II Study. **36**, 3353-3360 (2018).
17. Necchi, A., *et al.* Updated Results of PURE-01 with Preliminary Activity of Neoadjuvant Pembrolizumab in Patients with Muscle-invasive Bladder Carcinoma with Variant Histologies. *European urology* (2019).

18. Powles, T., *et al.* Clinical efficacy and biomarker analysis of neoadjuvant atezolizumab in operable urothelial carcinoma in the ABACUS trial. *Nature medicine* **25**, 1706-1714 (2019).
19. Cabrita, R., *et al.* Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature* **577**, 561-565 (2020).
20. Helmink, B.A., *et al.* B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* **577**, 549-555 (2020).
21. Meerveld-Eggink, A., *et al.* Short-term CTLA-4 blockade directly followed by PD-1 blockade in advanced melanoma patients: a single-center experience. *Annals of oncology : official journal of the European Society for Medical Oncology* **28**, 862-867 (2017).
22. Teo, M.Y., *et al.* Alterations in DNA Damage Response and Repair Genes as Potential Marker of Clinical Benefit From PD-1/PD-L1 Blockade in Advanced Urothelial Cancers. **36**, 1685-1694 (2018).
23. Mariathasan, S., *et al.* TGFbeta attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* **554**, 544-548 (2018).
24. Hollern, D.P., *et al.* B Cells and T Follicular Helper Cells Mediate Response to Checkpoint Inhibitors in High Mutation Burden Mouse Models of Breast Cancer. *Cell* **179**, 1191-1206.e1121 (2019).
25. Cabrita, R., *et al.* Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature* **577**, 561-565 (2020).
26. Largeot, A., Pagano, G., Gonder, S., Moussay, E. & Paggetti, J. The B-side of Cancer Immunity: The Underrated Tune. *Cells* **8**, 449 (2019).
27. Griss, J., *et al.* B cells sustain inflammation and predict response to immune checkpoint blockade in human melanoma. *Nature Communications* **10**, 4186 (2019).
28. Silina, K., *et al.* Germinal Centers Determine the Prognostic Relevance of Tertiary Lymphoid Structures and Are Impaired by Corticosteroids in Lung Squamous Cell Carcinoma. *Cancer Res* **78**, 1308-1320 (2018).
29. Amaria, R.N., *et al.* Neoadjuvant immune checkpoint blockade in high-risk resectable melanoma. *Nat Med* **24**, 1649-1654 (2018).

METHODS-ONLY REFERENCES

30. Zargar, H., et al. Multicenter assessment of neoadjuvant chemotherapy for muscle-invasive bladder cancer. European urology 67, 241-249 (2015).
31. Kim, S., et al. Strelka2: fast and accurate calling of germline and somatic variants. Nature Methods 15, 591-594 (2018).
32. Talevich, E., Shain, A.H., Botton, T. & Bastian, B.C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLoS computational biology 12, e1004873 (2016).
33. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139-140 (2009).
34. Ritchie, M.E., et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research 43, e47 (2015).
35. Strimmer, K. fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. Bioinformatics 24, 1461-1462 (2008).34. Strimmer, K. fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. *Bioinformatics* **24**, 1461-1462 (2008).



5

Spatial Relationships in The Urothelial and Head and Neck Tumor Microenvironment Predict Response to Combination Immune Checkpoint Inhibitors

Alberto Gil-Jimenez[1,2], Nick van Dijk[3], Joris L. Vos[4,5], Yoni Lubeck[6], Maurits L. van Montfoort[6], Dennis Peters[7], Erik Hooijberg[6], Annegien Broeks[7], Charlotte L. Zuur [4,5,8], Bas W. G. van Rhijn[9,10], Daniel J. Vis[1,2,*], Michiel S. van der Heijden[1,3,*], Lodewyk F. A. Wessels[1,2,11,*]

Parts of this chapter have been published in Nature Communications (2024), https://www.nature.com/articles/s41467-024-46450-1

[1]Department of Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam, the Netherlands
[2]Oncode Institute, Utrecht, the Netherlands
[3]Department of Medical Oncology, The Netherlands Cancer Institute, Amsterdam, the Netherlands
[4]Department of Head and Neck Surgery and Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[5]Division of Tumor Biology & Immunology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[6]Department of Pathology, The Netherlands Cancer Institute, Amsterdam, the Netherlands
[7]Core Facility Molecular Pathology & Biobanking, Netherlands Cancer Institute, Amsterdam, the Netherlands
[8]Department of Otorhinolaryngology Head and Neck Surgery, Leiden University Medical Center, Leiden, The Netherlands
[9]Department of Urology, The Netherlands Cancer Institute, Amsterdam, Netherlands
[10]Department of Urology, Caritas St. Josef Medical Centre, University of Regensburg, Regensburg, Germany
[11]Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, the Netherlands
*: equal contribution

Keywords: Immunotherapy; Urothelial cancer; Head and neck cancer; Spatial analysis; Multiplex immunofluorescence; Tumor microenvironment; Tumor immunology; Immune cells.

ABSTRACT

Immune checkpoint inhibitors (ICI) can achieve remarkable responses in urothelial cancer (UC), which may depend on tumor microenvironment (TME) characteristics. However, the relationship between the TME, usually characterized by immune cell density, and response to ICI is unclear.

Here, we quantify the TME immune cell densities and spatial relationships (SRs) of 24 baseline UC samples, obtained before pre-operative combination ICI treatment, using multiplex immunofluorescence. We describe SRs by approximating the first nearest-neighbor distance distribution with a Weibull distribution and evaluate the association between TME metrics and ipilimumab+nivolumab response.

Immune cell density does not discriminate between response groups. However, the Weibull SR metrics of $CD8^+$ T cells or macrophages to their closest cancer cell positively associate with response. $CD8^+$ T cells close to B cells are characteristic of non-response. We validate our SR response associations in a combination ICI cohort of head and neck tumors.

Our data confirm that SRs, in contrast to density metrics, are strong biomarkers of response to pre-operative combination ICIs.

5.1 INTRODUCTION

Immune checkpoint inhibitors (ICI) block inhibitory signals between immune and neoplastic cells that can result in cancer cell killing[1]. Inhibitors targeting PD-1 and PD-L1 have shown durable responses in a subset of urothelial cancer (UC) patients[2,3]. Still, most tumors do not respond to treatment[1,4], and ICI causes grade≥3 immune-related adverse events in 10% of UC patients[5]. Therefore, it is crucial to identify biomarkers that aid the stratification of responding patients so that alternative lines of treatment can be considered for non-responding patients and prevent unnecessary toxicities.

The surrounding of a tumor, known as the tumor microenvironment (TME), contains immune cells, normal epithelial cells, and fibroblasts that continuously interact[6]. Components of the TME indicative of pre-existing immunity have shown associations with response to ICI, such as $CD8^+$ T cell infiltration and transcription factors related to T cell activity[7–11]. However, biomarkers do not behave consistently in UC trials. For instance, the baseline presence of $CD8^+$ T cells correlates with ICI monotherapy response in the pre-operative (anti-PD-L1)[7] or metastatic setting[12]. Still, in pre-operative ICI combination therapy (anti-PD(L)-1 + anti-CTLA-4), the response is independent of $CD8^+$ T cell density[9,10]. The lack of robust response biomarkers highlights the need to dissect tumor-immune interactions in more detail[13].

A technology enabling a TME characterization at single-cell resolution is multiplex immunofluorescence (mIF), which spatially profiles a tissue slide using multiple antibodies simultaneously[14,15]. MIF-derived metrics were found to predict anti-PD-1 and anti-PD-L1 response across different tumor types[16], highlighting its ability to quantify crucial immune components that determine ICI response. Typically, mIF data are summarized as immune cell densities, informing about immune cell counts, and typically topologically assessed in different compartments, i.e., tumor and stroma[17].

By definition, immune cell density and abundance metrics ignore the immune interactions relevant to an anti-tumor response[18]. Ignoring these interactions is suboptimal, as many immune interactions require proximity. For instance, a T cell receptor and antigen interaction require physical binding, which requires the cells involved to be in close proximity to each other. In contrast, an immunosuppressive TME will have few immune cells close to cancer cells due to their inability to infiltrate a tumor[19]. Such distance or adjacency patterns between cells at the TME can be measured through their spatial relationship (SR), allowing for a mathematical description amenable to downstream analysis[20]. Notably, associations between SRs at the TME and prognosis[18,21,22] and response to monotherapy ICIs have been reported across different cancer types[23,24,25]. The SRs allow for quantitative exploration of the TME,

providing a basis for improving our understanding of tumor immunology and scrutinizing new associations with ICI response. SRs in the UC's TME are poorly understood and, to our knowledge, largely unexplored in pre-operative combination ICI treatments, and this study aims to address that.

Several analytical frameworks aim to measure the SRs of the TME, such as cell-cell interactions and tissue modules, using spatially-resolved protein-derived data across multiple data types[19]. Methodologies are predominantly topologically based (graph-, networks-, and cell-counting-based methods) or distance-based[20]. Distance-based methods, such as the first-nearest-neighbor (1-NN) distribution, allow modeling proximity patterns within the TME using spatially-resolved data in a simple yet informative manner. Because distances following a 1-NN distribution are asymmetrical, approaches that estimate the 1-NN distribution mean[25] provide inadequate data summaries. A common approach to model 1-NN distributions is through the cumulative distribution function (CDF), known as the G-function. Nevertheless, the downstream analyses require an additional summary by estimating the area under the curve (AUC) at a predefined threshold[26,27]. Currently, there is a lack of spatial methodologies that describe the distance distribution without using a threshold and that model variation between individuals.

In this study, we spatially profile cancer cells, T cells, macrophages, and B cells using mIF and present a methodology to quantify the TME SRs using 24 pre-operative baseline tumor resection UC samples from the NABUCCO trial[10]. In NABUCCO, pre-operative combination ICI with ipilimumab and nivolumab is administered in UC. We fit a Weibull distribution to the 1-NN distance distribution between pairwise cell relationships to extract a two-parameter describing the distribution. These spatial descriptions outperform immune cell densities when quantifying the differences in immune cell SRs between response groups to ICI. To demonstrate generalizability of our findings, we confirm the baseline associations between SRs and response in an independent cohort of 25 mostly HPV-negative head and neck squamous carcinoma (HNSCC) patients receiving pre-operative ipilimumab and nivolumab treatment[28].

5.2 RESULTS

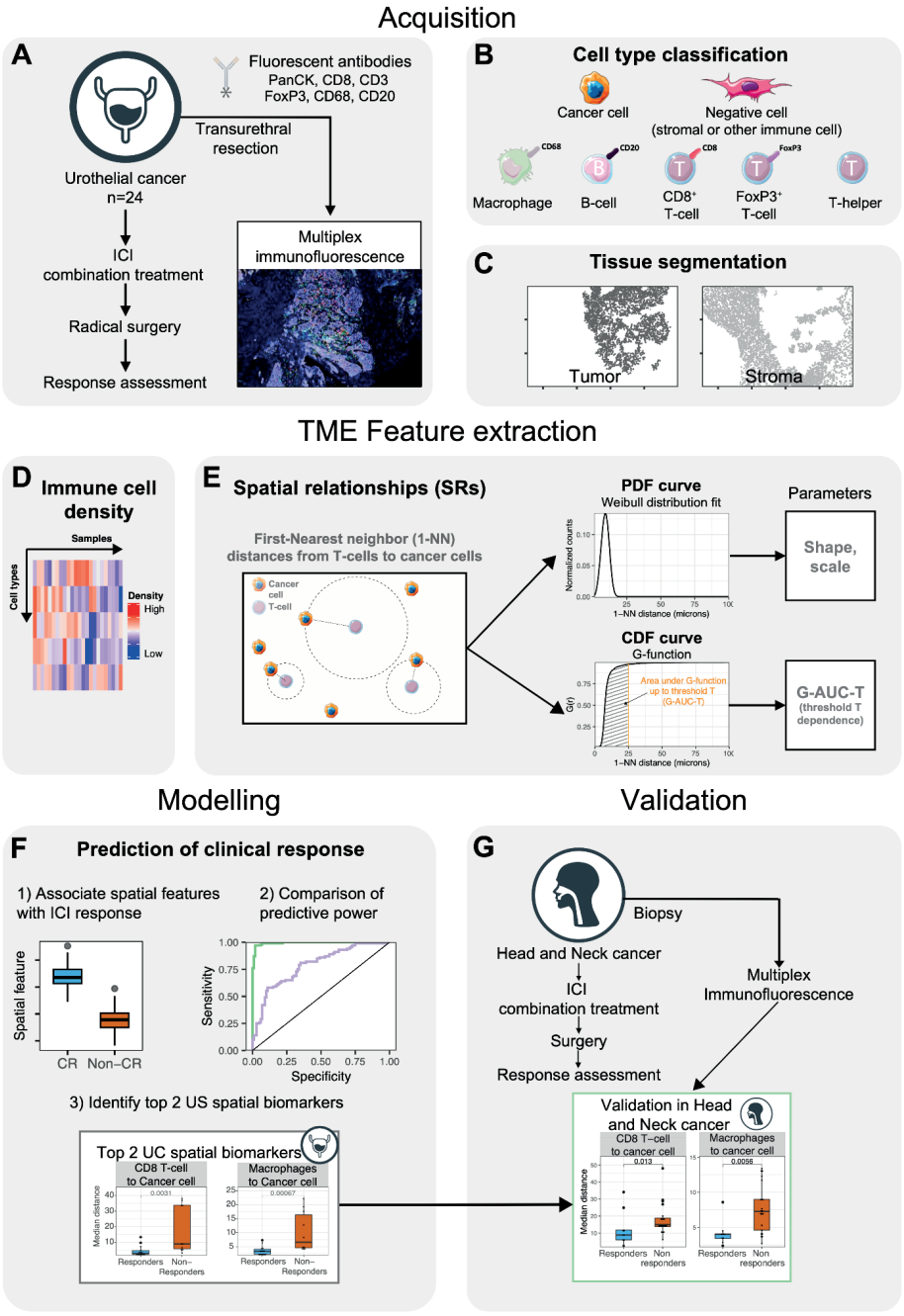

Acquisition
A
Fluorescent antibodies
PanCK, CD8, CD3
FoxP3, CD68, CD20
Transurethral resection
Urothelial cancer n=24
ICI combination treatment
Radical surgery
Response assessment
Multiplex immunofluorescence
B
Cell type classification
Cancer cell
Negative cell (stromal or other immune cell)
Macrophage
B-cell
CD8+ T-cell
FoxP3+ T-cell
T-helper
C
Tissue segmentation
Tumor
Stroma
TME Feature extraction
D
Immune cell density
Samples
Cell types
Density
High
Low
E
Spatial relationships (SRs)
First-Nearest neighbor (1-NN) distances from T-cells to cancer cells
Cancer cell
T-cell
PDF curve
Weibull distribution fit
Normalized counts
1–NN distance (microns)
Parameters
Shape, scale
CDF curve
G-function
G(r)
Area under G-function up to threshold T (G-AUC-T)
G-AUC-T
(threshold T dependence)
Modelling
Validation
F
Prediction of clinical response
1) Associate spatial features with ICI response
2) Comparison of predictive power
Spatial feature
CR
Non–CR
Sensitivity
Specificity
3) Identify top 2 US spatial biomarkers
Top 2 UC spatial biomarkers
CD8 T-cell to Cancer cell
Macrophages to Cancer cell
Median distance
Responders
Non–Responders
G
Biopsy
Head and Neck cancer
ICI combination treatment
Surgery
Response assessment
Multiplex Immunofluorescence
Validation in Head and Neck cancer
CD8 T–cell to cancer cell
Macrophages to cancer cell
Median distance
Responders
Non responders

Figure 1. Profiling of immune cell density and spatial relationships of the urothelial cancer tumor micro-environment by multiplex immunofluorescence.

(A) Biopsy samples from 24 patients from the NABUCCO trial were profiled using mIF.

(B) Cell type classification by comparing antibody marker positivity.

(C) Tissue segmentation into tumor and stroma regions by comparing the local densities of cancer cell marker positive and negative cells.

(D) Immune cell density in the tumor and stroma compartments was calculated in each tissue compartment (tumor and stroma).

(E) SRs were summarized using the 1-NN statistic studied from a reference cell type to a target cell type. The resulting 1-NN distances vector was studied using 2 approaches: modeling a Weibull distribution to the Probabilistic Density Function (PDF) (top), and using the cumulative distribution function (CDF) using the G-function

(F) Association of SR parameters with ICI response and comparison of the discriminative power between SR and density TME parameters.

(G) Validation of associations between SR parameters and response identified in UC in an independent cohort of HNSCC tumors.

Icons from panel A, B, F and G were adapted from bioIcons (cancerous-cell-1, lymphocytes-4, macrophage, t-lymphocyte, b-lymphocyte, fibroblast-1 licensed under CC-BY 3.0 Unported by Servier), flaticon.com (bladder icon, https://www.flaticon.com/free-icon/bladder_1453578; head neck icon, https://www.flaticon.com/free-icon/injection_4418017).

Abbreviations*: TME*: tumor micro-environment; *SR*: spatial relationship; *mIF*: multiplex immunofluorescence; *ICI*: immune checkpoint inhibitors; *1-NN*: first nearest neighbor; *PDF*: probabilistic density function; *CDF*: cumulative density function; *G-AUC-T*: G-function evaluated at a threshold T; *T*: threshold; *UC*: urothelial cancer; *HNSCC*: head and neck squamous cell carcinoma.

5.2.1 Multiplex immunofluorescence and modeling of immune cell densities and spatial relationships of urothelial and head and neck cancer samples

We collected multiplex immunofluorescence (mIF) data from baseline formalin-fixed, paraffin-embedded (FFPE) stage-III urothelial cancer (UC) samples (n=24), of patients recruited in the NABUCCO trial (Figure 1A, Table 1). Patients received pre-operative combination ICI, consisting of two or three cycles of ipilimumab (anti-CTLA-4) and nivolumab (anti-PD-1)[10]. We determined the position and identity of cells using mIF and identified B cells, T cells ($CD8^+$ T cells, $FoxP3^+$ T cells, and T-helper cells), macrophages, and cancer cells (Figure 1B). Negative cells scored negative for all the antibodies ($CD8^-$, $CD3^-$, $FoxP3^-$, $CD20^-$, $CD68^-$, $PanCK^-$); this group contains all stromal cells and immune cells not covered by our antibody panel. Next, by comparing the local density of cancer and negative cells, we virtually segmented the tissue into the tumor and stroma compartments (Figure 1C, Supplementary Figure 1) and quantified immune cell density in both compartments (Figure 1D, Supplementary Data Table 1).

Table 1. Clinical trial, treatment and sample characteristics for the cohorts used in this study.

	NABUCCO *NCT03387761*	**IMCISION** *NCT03003637*
Sample size	24	25*
Cancer type	Urothelial cancer	Head and neck cancer
Tissue source	Bladder, n=24 (100%)	Oral cavity, n=24 (96%) Oropharynx, n=1 (4%)
Tissue sampling	Transurethral resection (primary tumor, FFPE)	Primary tumor biopsy (FFPE)
Treatment type	Pre-operative (neoadjuvant)	Pre-operative (neoadjuvant)
Treatment dosage per cycle	1: ipilimumab 2: ipilimumab + nivolumab 3: nivolumab	1: ipilimumab + nivolumab 2: nivolumab
Tumor type	Primary, n=24 (100%)	Primary, n=20 (80%) Recurrence, n=5 (20%)
HPV-positivity (%)	n/a	23 (92%)
Response definition	Pathological response assessment (complete pathological response or residual disease)	Pathological response assessment combined with comparison of tumor cells decrease from baseline vs. on-treatment sample
Response (%)	14 (58%)	9 (36%)
Clinical Stage	cT3-4aN0M0 n=14 (58%) cT2-4aN1-3M0 n=10 (42%)	cT2N0 n=5 (20%) cT3-4aN0M0 n=10 (40%) cT2-4aN1-3M0 n=10 (40%)
Sex (%)	Male n=18 (75%)	Male n=20 (80%)

*only samples from IMCISION Arm B (combination ICIs) have been used in this manuscript

We quantified the pairwise SRs between all cell types in the TME using the first nearest-neighbor (1-NN) distance statistic (Supplementary Data Table 2). In brief, the statistic is measured between a reference cell type (cell from) and a target cell type (cell to). The distances between each reference cell type and their closest target cell type yielded a 1-NN distance vector (Figure 1E). Then, we fitted a Weibull distribution function using a non-linear mixed effect model to summarize the 1-NN distance distribution. The model has two parameters: shape and scale, which uniquely describe the properties of the 1-NN distance distribution (Figure 1E, top) in a threshold-independent manner. We estimated the Weibull parameters (scale and shape) for all 49 pairwise relationships between cell types for all samples using the data from the whole tissue slide. Additionally, we evaluated the threshold-dependent G-function derived from the cumulative distribution function (CDF) of the 1-NN distance distribution and broadly documented in the spatial statistics literature[29]. The G-functions were summarized by computing the area under the curve (AUC) at different thresholds T, which we refer to as G-AUC-T (which included T = 25 [Figure 1E, bottom], T = 50- and T = 100-micron) across all pairwise cell type SRs and samples.

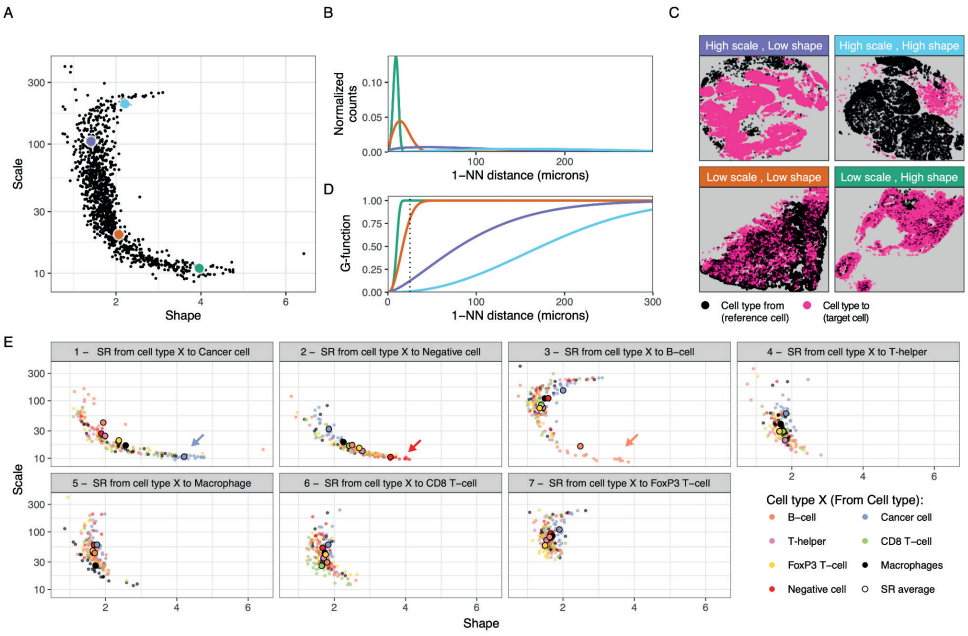


Figure 2. Exploration of pairwise cell type SRs in the TME using the 1-NN distance statistic.

(A) Scale vs shape SR parameters fitted on the 1-NN distribution for the 24 samples and the 7x7 cell type combinations. Representative examples are highlighted in green, orange, purple and cyan, with their associated 1-NN distance distributions (B), point patterns (C) and G-functions (D).

(E) Scatter plot of the scale-shape parameter space by target cell type (cell type to in the SR). For instance, the first facet represents SRs studied from any reference cell type to cancer cells. There, the coloring denotes the reference cell type (cell type from); i.e. the orange dots represent spatial relations studied from B cells to cancer cells. Cohort averages for their associated SR parameters are highlighted as big dots for each cell type-cell type combination.

Abbreviations: *SR*: spatial relationship; *TME*: Tumor microenvironment; *1-NN*: First nearest neighbor.

We then compared the spatial (Figure 1E) and density (Figure 1D) parameters with response to ICIs and compared their predictive power (Figure 1F). Lastly, we validated the associations between TME parameters and response in an independent cohort of mostly HPV-negative head and neck squamous cell carcinoma (HNSCC, n=25) using baseline primary tumor samples of patients recruited in the IMCISION trial that received pre-operative ipilimumab and nivolumab combination treatment (Figure 1G).

5.2.2 Exploration of spatial relationships across the urothelial cancer tumor micro-environment

We quantified the SRs for all pairwise relationships of immune, cancer, and negative cells by estimating the Weibull parameters (shape, scale) characterizing the 1-NN distance distributions (Figure 1E). Next, we explored the shape-scale parameter space across patients (Figure 2A).

To illustrate what the Weibull parameters (shape, scale) represent, we use four combinations of the SR metrics that characterize distinct instances of cellular spatial distributions (Figure 2A, colored dots). For the green dot in Figure 2A, we observe that the 1-NN distance distribution is characterized by relatively low distances (Figure 2B, green curve), which originate from the black cells being close to a pink cell (Figure 2C, "Low scale, High shape"). The low scale/high shape value signals densely packed pink cells, meaning there is always a pink cell close to a black cell. Another type of SR characterized by relatively short 1-NN distances but with a higher variance in 1-NN distances, is illustrated in Figures 2A-B (orange dot and curve). It represents an SR where one or both cell types are arranged in overlapping, densely packed clusters, such as the pink and black cells in Figure 2C (*"Low scale, Low shape"*). Cases with an even larger spread in distances (Figures 2A-B, purple dot and curve) from the black to the nearest pink cells, are described by high scale and low shape parameter values (Figure 2C, *"High scale, Low shape"*). Lastly, the cyan dot (Figure 2A) represents a 1-NN distribution shifted towards higher 1-NN distances (Figure 2B, cyan curve), characteristic of a repulsion pattern, i.e., where both cell types are clustered in relatively large clusters (Figure 2C, *"High scale, High shape"*). Therefore, increasing the scale parameter results in an increase of the distance distribution spread (i.e., the distribution width, Supplementary Figure 2A-B), while the shape parameter is related to the distinct forms of the distribution behavior (Supplementary Figure 2C-F, Supplementary Figure 3).

Furthermore, the associated threshold-dependent G-function results showed corresponding differences between the four scenarios, which is expected, as the G-function is the cumulative distribution of the 1-NN distances (Figure 2D, equation 4). However, in contrast to the Weibull approach, a threshold value (T) is required to generate the summary metric G-AUC-T.

We then dissected the SRs by reference and target cell type to explore patterns of SRs across the TME (Figure 2E). First, we investigated self-self relationships, which are relationships between cells of the same type. The self-self relationship for tumor cells falls in the "Low scale, High shape" scenario, with shorter distances between cells as tumor cells are typically densely packed in the tumor regions (Figure 2E-1, blue dots and arrow). Similar behavior was observed for negative cells, indicating a tight packing of negative cells (Figure 2E-2, red dots and arrow). Then, we explored the self-self relationships of immune cells. We observed B cells clustering for all patients (Figure 2E-3, orange dots and arrow). We did not observe the same clustering behavior for self-self SRs of other immune cell types (e.g., green dots in Figure 2E-6, which shows a behavior more akin to the *"Low scale, Low shape"* scenario). Subsequently, we assessed the SRs between different cell types. We observed a high variation in the Weibull parameters across samples and pairwise cell type combinations (Figure 2E) and a dependence on the SR perspective, i.e., whether a given cell type is the cell type from or

the cell type to, which can be attributed to the asymmetric property of the 1-NN statistic, as illustrated in Supplementary Figure 4.

In short, we created a framework to quantify, interpret, and study SRs in the TME using the Weibull parameters extracted from the 1-NN distance distributions. The framework allows exploring distinct cellular organization patterns and quantifying specific SRs (e.g., T cell to B cell vs. B cell to T cell) amenable for downstream analyses aimed at furthering our understanding of the TME and its relationship to response to ICIs.

5.2.3 Comparison between spatial relationship metrics derived from the first nearest neighbor distance statistic

A common approach to extracting parameters from the 1-NN distance distribution statistic is through the G-function, which represents the cumulative density function (CDF) of the 1-NN distribution and requires a particular threshold to summarize the data for downstream analyses. We compared our Weibull parameters with the G-function summary. We evaluated the G-function using its area under the curve (AUC) up to a 25-micron distance, which we defined as the "G-AUC-25". We chose this threshold after inspecting the distance region in which the G-function showed the highest variability. Because we observed variability in G-functions across pairwise cell-type relationships, other thresholds were evaluated and denoted as "G-AUC-T", in which T denotes the threshold in microns. We observed a non-linear relationship between the shape, scale and the G-AUC-25 (Figure 3A), G-AUC-12 and G-AUC-50 (Supplementary Figures 5A-B).

Upon summarizing the G-function, and for a given SR's G-AUC-25 value, we observed that the associated shape and scale parameters can show substantial variation (Figure 3A; e.g., dots coloured in red mapping to a wide range of shape parameter values). Figure 3A highlights a pair of SRs with large differences between their shape or scale parameters, which can be visually confirmed by their 1-NN curves (Figures 3B, D, F). In contrast to the Weibull parameters, the corresponding G-AUC-25 values do not differ substantially between the pair members (Figures 3C, 3E, and 3G, showing the G-function curves corresponding to the pairs in Figures 3B, 3D, and 3F, respectively), which can hinder interpretations on the associated SR. Specifically, comparing Figures 3B and 3C, we observe that while the Weibull parameters are quite different (scale=14 and shape=6.4 for B-light, scale=14 and shape=3.0 for B-dark) the G-AUC-T values are quite similar (G-AUC-25=11.5 for B-light and G-AUC-25=11.8 for B-dark). While in the other examples the differences in the G-AUC-25 values were small, we found that higher values of the summary threshold could better capture the difference between the SR pairs in Figure 3E (Weibull density in 3D) and Figure 3G (Weibull density in 3F).

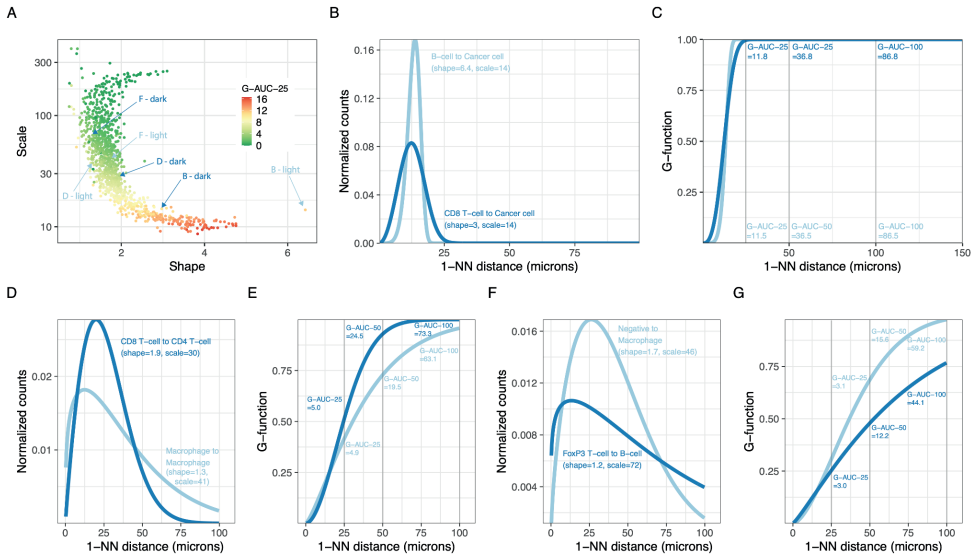


Figure 3. Comparison of spatial relationship parameters derived from the 1-NN distance statistic.
(A) Scale/shape parameter space associated with the G-AUC-25 (coloring, G-function AUC evaluated at 25 microns).
(B, D, F) 1-NN PDF distance distribution of pairs of examples with a substantial difference in the shape parameters. The plot reports each curve's shape, scale, and SR. Each SR is annotated in panel (A).
(C, E, G) Associated G-functions (CDF) of the examples shown in panel A of pairs of examples without a substantial difference in G-AUC-25. G-AUCs evaluated at different thresholds T are reported in the figure.
Abbreviations: *AUC*: Area under the curve; *G-AUC-T*: G-function AUC evaluated at a threshold T; *T*: threshold; *1-NN*: first nearest neighbor; *SR*: spatial relationship.

To further illustrate differences between SRs captured by the shape or scale parameters but not by the G-AUC-25 parameter, we compared G-AUC-Ts for different pairwise cell-cell relationships (Supplementary Figure 5C). Here, for different samples but the same SRs (e.g., Macrophages to B cell), the magnitude of the increase in the G-AUC-T value when altering the evaluation threshold T, depending on the studied pairwise cell-cell relationship, the G-function's shape (rapidly or slowly reaching the maximum value), and the sample. In some SRs, the G-AUC's increase was linear (e.g., Cancer cell to cancer cell) because the G-function saturated at low thresholds (Supplementary Figure 5D). Still, in others (e.g., Macrophages to B cell), the increase was not always linear because the G-function reaches saturation at higher thresholds (Supplementary Figure 5D). Therefore, when using a G-function statistic, such as the G-AUC-T, the SR quantification critically depends on the threshold used.

In short, our data shows that the G-function threshold introduces variance in the downstream metric G-AUC-T. Furthermore, optimizing the threshold to maximize the effect size of SR-biomarkers for treatment response creates a risk of overfitting. Therefore, different

results using the same SR data can be obtained when varying the threshold, which can hinder downstream interpretation.

5.2.4 Spatial relationships associated with immune checkpoint blockade response

Multiplex immunofluorescence data are usually summarized as cell type fractions or immune cell density. We quantified the density of T cells, B cells, and macrophages in both the tumor and stromal compartments. However, we found no significant differences between response groups (Figure 4A), indicating no differences between response groups in immune cell abundances in either the tumor and stromal compartments. In addition, immune cells spatially distribute following configurations of immune phenotypes[30], being Excluded (high immune cell abundance in the stroma), Inflamed (high immune cell abundance in the tumor), and Desert (low immune cell abundance in the tumor and stroma). We quantified exclusion ratios (ratio between stromal and intratumoral immune cell density) and used them as a proxy of immune phenotypes for each immune cell. However, again we found no significant associations with treatment response (Figure 4B), suggesting similar immune cell configurations in the response groups.

Motivated by the invariance between immune cell abundances and density ratios in the TME between response groups, we investigated whether spatial relationships derived from the TME were predictive of response to ICI combination treatment. We first investigated whether the SRs of all pairwise cell types, characterized by the Weibull parameters (shape and scale), were associated with clinical response (Figure 1F). After correction for multiple hypothesis testing, we identified nine SRs that were associated with clinical response (FDR<0.10) for the Weibull parameters shape and scale (Figures 4C, D). The association between G-AUC-T and response using a rank-based statistic strongly depended on the selected value of the threshold, with the fold change decreasing with increasing values of the threshold (T) (Supplementary Figure 6). When selecting a low threshold value (T = 25 microns), we found no significant associations between SRs quantified by G-AUC-25 and response (Supplementary Figure 6A). Upon increasing the threshold value (T = 50 microns), we found three associations between SRs quantified by G-AUC-50 and response (Supplementary Figure 6B), of which two were also identified using the Weibull parameters and one SR (FoxP3$^+$ T cell to negative cell) was trending but not significant (FDR_{scale}=0.21, FDR_{shape}=0.11) using the Weibull parameters (Supplementary Figures 6, 7).

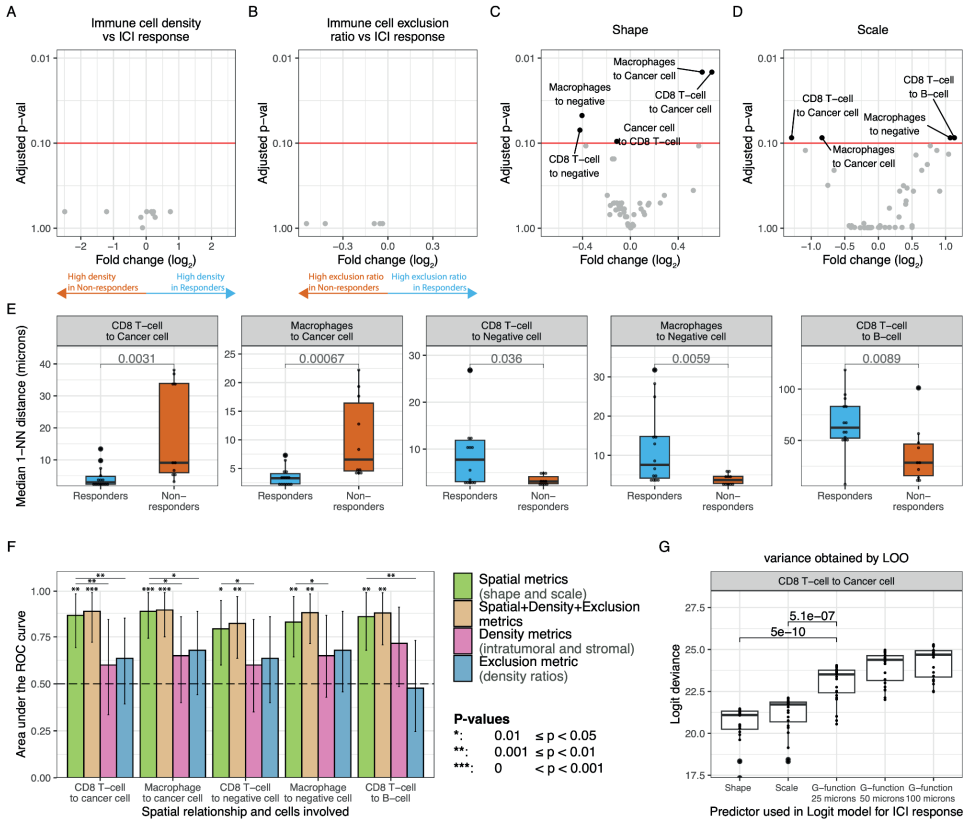


Figure 4. Association of spatial relationships with response to pre-operative ipilimumab+nivolumab in urothelial cancer.

(A) Volcano plot showing the fold change on the intratumoral and stromal immune cell densities between response groups (x-axis) and statistical significance by t-test adjusted by multiple hypothesis testing (y-axis).

(B) Volcano plot showing the fold change on the exclusion ratio (ratio between stromal and intratumoral immune cell density) between response groups (x-axis) and statistical significance by t-test adjusted by multiple hypothesis testing (y-axis).

(C) Volcano plot showing the fold change on the shape parameter between response groups (x-axis) and statistical significance by t-test adjusted by multiple hypothesis testing (y-axis).

(D) Volcano plot showing the fold change on the scale parameter between response groups (x-axis) and statistical significance by t-test adjusted by multiple hypothesis testing (y-axis).

(E) Median first nearest-neighbor distance distribution per response group as calculated by the associated shape and scale for the SRs that are significantly associated with response to ICI treatment (n=14 independent responders and n=10 independent non-responders) and statistical significance by a Mann-Whitney test. No adjustments were made for multiple comparisons.

(F) ROC curve AUCs of the discriminative power of distinct TME parameters when predicting ICI response on n=14 independent responders and n=10 independent non-responders: all spatial parameters (green), all spatial parameters, density and exclusion ratio metrics (gold), all density metrics (pink), and exclusion ratio between stroma and tumor density (blue). Confidence intervals (error bar of each barplot) denote the 95% confidence interval as estimated by bootstrapping the 24 samples 500 times as implemented in pROC. The lines denote whether a statistical significance on the associated AUC was achieved. Significance symbols above bar plots denote whether the AUC is significantly

different from AUC=0.5 as assessed by a two-sided Mann-Whitney test. Significance symbols between bar plots denote whether the ROC-AUC-spatial (green) is significantly greater than the ROC-AUC-spatial-density-exclusion (gold) or ROC-AUC-density (pink) or ROC-AUC-exclusion (blue). Statistical significance between ROC curves was assessed by re-calculating AUCs by bootstrapping each ROC plot 500 times, and significance was assessed by a one-sided t-test as implemented in pROC. Exact p-values are reported in the Source Data file. No adjustments were made to correct for multiple comparisons.

(G) Logistic regression deviance of a univariate logistic regression model predicting ICI response (n=14 independent responders and n=10 independent non-responders) using as a predictor the shape, the scale, or the G-function evaluated at different thresholds (G-AUC-T). Variability on the AIC was evaluated by leave-one-out cross-validation on the 24 samples and significance was tested by a student's t-test. No adjustments were made for multiple comparisons.

The box plots in each panel show the middle 50% of the data, with the box itself representing the median and the interquartile range (IQR) between the 25th and 75th percentiles. The whiskers extend from the box to the furthest data points within 1.5 times the IQR from the median. Unless otherwise stated, all statistical tests were two-sided. Significance symbols: *: p<0.05, **: p<0.01, ***: p<0.001.

Abbreviations: *ICI*: Immune checkpoint inhibitor; 1-NN: first nearest neighbor; *ROC*: Receiver operating characteristic; *AUC*: area under the curve; *TME*: tumor microenvironment; *G-AUC-T*: G-function evaluated at a threshold T; *T*: threshold; *logit*: logistic regression; LOO: leave-one-out cross validation.

To guide interpretation, we computed, for each SR significantly associated with clinical response to combination immunotherapy and each patient, the median 1-NN distances (Figure 4E) and compared them using a rank-based statistic. Median distances exhibit a non-linear relationship with the shape and scale parameters (as indicated in equation 6) but provide enhanced clarity in the interpretation of our findings. In responding tumors, the distances from either $CD8^+$ T cells or macrophages to the closest cancer cells were smaller than in non-responders (median 1-NN distance $CD8^+$ T cell to cancer cell, responders=4±3μm, non-responders=18±15μm; Macrophage to cancer cell, responders=4±2μm, non-responders=10.2±7μm). Conversely, responding tumors had the largest 1-NN distances for the SR from $CD8^+$ T cells or macrophages to the closest negative cell (median 1-NN distance $CD8^+$ T cell to negative cell, responders=9±8μm, non-responders=3±1μm; Macrophage to negative cell, median 1-NN distance responders=12±10μm, non-responders=4±1μm). Despite the clear differences in the associated median 1-NN distances, the G-function approach did not identify the associations of the SRs with response at a low threshold (T = 25 microns, Supplementary Figure 3A) nor the associations of the SRs involving $CD8^+$ T cells and response at a higher threshold (T = 50 microns, Supplementary Figure 3B, FDR=0.14 and FDR=0.20 for $CD8^+$ T cell to cancer cell and $CD8^+$ T cell to negative cell, respectively). Furthermore, non-responding tumors were characterized by small distances from $CD8^+$ T cells to B cells (median 1-NN distance $CD8^+$ T cell to B cell, responders=66±27μm, non-responders=36±27μm) when compared to responding tumors. We identified an association between the SR from cancer cells to $CD8^+$ T cells and response with a small fold change for the shape parameter ($|FC_{shape}|$=0.11, FDR_{shape}=0.09), pointing to a difference in distribution that was not detected in terms of median 1-NN distances (Supplementary Figures 7A-E, p=0.98), suggesting that this may well be a false positive. Lastly, the single SR biomarker identified by the G-function approach at a 50 microns but not at a 25 microns threshold (FoxP3 T cell to negative cell, $FDR_{G\text{-}AUC\text{-}25}$=0.12, $FDR_{G\text{-}AUC\text{-}50}$=0.04, Supplementary Figure 6 and Supplementary Figure 7J) that was not identified

by the Weibull approach was trending (Supplementary Figure 7I, FDR_{shape}=0.11) and showed relative differences in the associated median 1-NN distances (Supplementary Figure 7F).

In contrast to the SRs, the immune cell density and exclusion ratios were not associated with response. We confirmed, using simulated data, that density affects SRs between rare (e.g., immune to immune cells) but not between abundant and rare cell types (e.g., cancer to immune cells) (Supplementary Note 1). To further confirm independence between density and SR metrics in the predictive setting, we compared the predictive power for clinical response of the SR Weibull parameters (shape and scale) and their associated relevant density and exclusion metrics. The comparisons were made for each SR that was significantly associated with treatment response. For example, the SR from $CD8^+$ T cells to cancer cells was associated with response (FDR_{shape}= 0.01, FDR_{scale}=0.09). We compared its predictive power with the $CD8^+$ T cell density (intratumoral and stromal) and exclusion ratio of $CD8^+$ T cells. For this comparison, we employed a logistic regression model and the resulting area under the ROC curve (AUROC). The AUROCs for the five SR associations (depicted in Figure 4E) and the associated density and exclusion metrics are shown in Figure 4F. No density or exclusion metric reached significance as all 95% CI of the associated AUROCs included AUROC=0.5. In contrast, all the SR Weibull parameters reached significance with AUROC values around 0.8 and 95% CI that do not include AUROC=0.5 (Table 2, Supplementary Figure 8), highlighting the superior predictive power of the SR metrics. Lastly, we tested whether adding density and exclusion ratio as covariates to the SR-based model improved the performance, but this was not the case (Figure 4F, Table 2).

Table 2. Predictive power of logistic regression models as measured by the ROC-AUC for logistic regression models trained using only spatial metrics (shape, scale), density metrics (intratumoral and stromal density), exclusion ratio metrics (ratio between stromal and intratumoral density), and all metrics (all the metrics above). P-value against spatial ROC-AUC denotes the difference between each model and the model trained spatial metrics from the associated SR by bootstrapping 500 times (testing whether the AUC is greater). Confidence intervals of ROCs were estimated by bootstrapping samples 500 times. Error bars denote the 95% confidence interval as estimated by bootstrapping the 24 samples 500 times as implemented in pROC. Significance was estimated by evaluating whether the ROC-AUC is significantly greater than the ROC-AUC-spatial (ROC plot built with spatial parameters) by re-calculating AUCs by bootstrapping each ROC plot 500 times, and significance was assessed by a one-sided t-test as implemented in pROC. No adjustments were made to correct for multiple hypothesis testing.

Spatial relationship	Features included in model	AUC (95% CI)	P-value against Spatial ROC-AUC
$CD8^+$ T cell to cancer cell	Spatial (shape and scale)	0.86 (0.70, 0.98)	-
$CD8^+$ T cell to cancer cell	Density (Intratumoral $CD8^+$ T cell density and Stromal $CD8^+$ T cell density)	0.6 (0.34 - 0.85)	0.005
$CD8^+$ T cell to cancer cell	Exclusion ratio (ratio between Stromal $CD8^+$ T cell density and Intratumoral $CD8^+$ T cell density)	0.64 (0.40 - 0.86)	0.02
$CD8^+$ T cell to cancer cell	All TME features (spatial + density + exclusion ratio)	0.89 (0.74 - 0.99)	0.79
$CD8^+$ T cell to negative cell	Spatial (shape and scale)	0.79 (0.59 - 0.95)	-
$CD8^+$ T cell to negative cell	Density (Intratumoral $CD8^+$ T cell density and Stromal $CD8^+$ T cell density)	0.60 (0.34 - 0.84)	0.03
$CD8^+$ T cell to negative cell	Exclusion ratio (ratio between Stromal $CD8^+$ T cell density and Intratumoral $CD8^+$ T cell density)	0.64 (0.39 - 0.86)	0.08
$CD8^+$ T cell to negative cell	All TME features (spatial + density + exclusion ratio)	0.82 (0.63 - 0.96)	0.73
Macrophage to cancer cell	Spatial (shape and scale)	0.89 (0.73 - 0.99)	-
Macrophage to cancer cell	Density (Intratumoral Macrophage density and Stromal Macrophage density)	0.65 (0.41 - 0.86)	0.02

Table 2. *(Continued)*

Spatial relationship	Features included in model	AUC (95% CI)	P-value against Spatial ROC-AUC
Macrophage to cancer cell	Exclusion ratio (ratio between Stromal Macrophage density and Intratumoral Macrophage density)	0.68 (0.46 - 0.88)	0.03
Macrophage to cancer cell	All TME features (spatial + density + exclusion ratio)	0.89 (0.74 - 0.99)	0.76
Macrophage to negative cell	Spatial (shape and scale)	0.83 (0.64 - 0.96)	-
Macrophage to negative cell	Density (Intratumoral Macrophage density and Stromal Macrophage density)	0.65 (0.41 - 0.86)	0.04
Macrophage to negative cell	Exclusion ratio (ratio between Stromal Macrophage density and Intratumoral Macrophage density)	0.68 (0.45 - 0.89)	0.08
Macrophage to negative cell	All TME features (spatial + density + exclusion ratio)	0.87 (0.71 - 0.99)	0.92
$CD8^+$ T cell to B cell	Spatial (shape and scale)	0.86 (0.68 - 0.98)	-
$CD8^+$ T cell to B cell	Density (Intratumoral B cell density and Stromal B cell density) **	0.72 (0.48 - 0.91)	0.16
$CD8^+$ T cell to B cell	Exclusion ratio (ratio between Stromal B cell density and Intratumoral B cell density)	0.48 (0.23 - 0.72)	0.008
$CD8^+$ T cell to B cell	All TME features (spatial + density + exclusion ratio)	0.88 (0.70 - 0.99)	0.58

Abbreviations: ROC: receiver operating characteristic; AUC: area under the curve; SR: spatial relationship; CI: confidence interval.

**: CD8+T cell density was not included in the model as it was already included in the density models of the SRs CD8+T cell to Cancer cell and CD8+T cell to negative cell

Next, we compared different SR metrics derived from the 1-NN distance distribution (shape, scale, G-AUC-T at different T) for their ability to describe clinical response. To do so, we compared the Weibull parameters to the G-AUC-T metrics in the predictive setting. Specifically, we compared the logistic regression deviance, in which lower values indicate better model fits. We did so for the six SRs that were found to be significantly associated with response. The SR from $CD8^+$ T cells to cancer cells model showed that the shape or the scale parameters scored significantly better than the G-AUC-T trained models for T = 25, 50, and 100 (Figure 4G). For the remaining SRs significantly associated with clinical response (Figures 4C-D), we observed that the models trained using Weibull parameters (shape and scale) outperformed the models trained using G-function parameters, except for the SR from macrophages to negative cells and from CD8 T cells to B cells where G-AUC-25 performed similarly as in the Weibull approach (Supplementary Figure 9).

In summary, we observed that mIF-derived spatial relationships in the TME hold superior predictive power for clinical response compared to immune cell density or immune phenotypes. Our results convincingly demonstrate that the Weibull parameters (shape and scale) are superior to the G-function metrics (G-AUC-T) in predicting clinical response to combination checkpoint therapy.

5.2.5 Validation of spatial relationships biomarkers of ICI response in a cohort of head and neck cancer

We tested whether our spatial biomarkers also predicted response in other cancer types. We used a cohort of head and neck squamous cell carcinoma (HNSCC) patients from the IMCISION trial[28] to validate our findings. A subset of 25 IMCISION patients was treated with pre-operative ipilimumab+nivolumab (similar to NABUCCO), and successfully provided tumor sample profiling with the same mIF antibody panel as the UC cohort (Figure 1G, Table 1).

We first compared the SR parameter space in HNSCC (Supplementary Data Table 3) with that of the UC cohort. We observe the same C-shape distribution in the shape-scale space as we observed in UC (Figure 5A). Second, we found a high concordance between NABUCCO and IMCISION for the shape and scale population averages across pairwise SRs between all cell types (Figure 5B, Supplementary Figures 10A-D). These results suggest that the distance between cell types follows a characteristic pattern preserved across these two cancer types. For instance, similar behavior of B cell to B cell 1-NN distances compatible with the "Low scale, High shape" behavior was observed in HNSCC.

Next, we evaluated whether the SR biomarkers of ICI response identified in UC were also predictive of reaching a major pathological response upon combination ICI in the HNSCC cohort. The validation was assessed for the strongest biomarkers (FDR<0.04) to maximize

the likelihood of validation, which involved the SRs *CD8$^+$ T cells to cancer cells* and *Macrophages to cancer cells* from Figures 4C-D. Both SRs showed the same direction of association with response in HNSCC (Figure 5B) and, importantly, showed a statistically significant association with response after multiple testing correction: CD8$^+$ T cells to cancer cells (FDR_{shape}=0.045) and macrophages to cancer cells (FDR_{shape}=0.0076, FDR_{scale}=0.00094) (Supplementary Figures 11A-B), which matches with the spatial proximity behavior (lower 1-NN distances) identified in the responding UC tumors (Figure 4E). We confirmed the earlier established superiority of the SR metrics over density metrics by showing that, in the HNSCC cohort, immune cell density was not associated with response (Supplementary Figure 11C), except for stromal macrophage and CD8$^+$ T cell densities.

In conclusion, the TME spatial biomarkers for pathological response to ICI combination treatment in UC validated in an HNSCC cohort, suggesting that the SRs between CD8$^+$ T cells and macrophages to cancer cells could be an important context-independent biomarker for clinical response to ipilimumab+nivolumab.

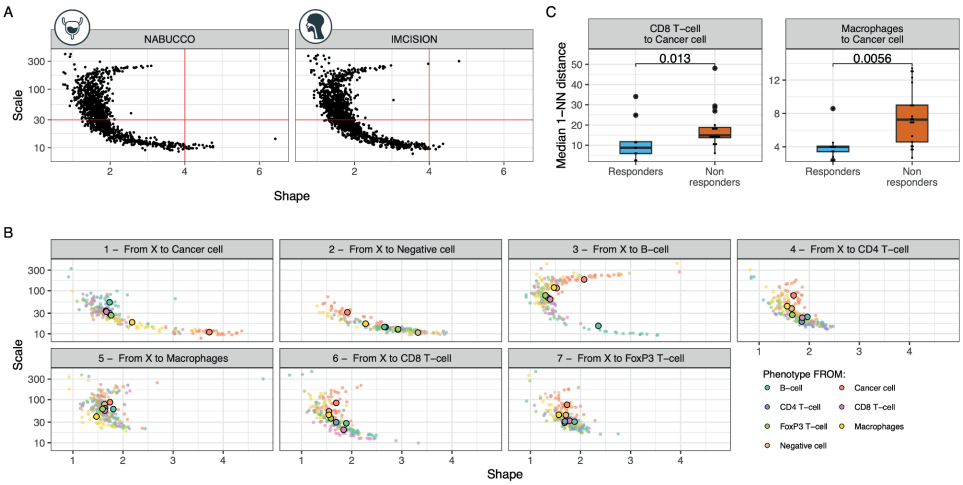


Figure 5. Validation of spatial relationship biomarkers of ICI response in an independent cohort of pre-operative ipilimumab+nivolumab in head and neck cancer.

(A) Scale vs. shape SR parameters fitted on the 1-NN distribution for the 25 samples and the 7x7 cell type combinations obtained in the head and neck cancer data (right, IMCISION) and UC data (left, NABUCCO).

(B) Scatter plot of the scale-shape parameter space by neighbor cell type (cell to) obtained in the IMCISION trial data. Cohort averages for their associated spatial parameters are highlighted as big dots for each cell type-cell type combination.

(C) Median first nearest-neighbor distance distribution in head and neck cancer samples per response group (n=9 independent responders and n=16 independent non-responders) as calculated by the associated shape and scale for the SRs that significantly (FDR<0.04) associated with response in UC and statistical significance by a Mann-Whitney test. No adjustments were made for multiple comparisons (adjustments for multiple hypothesis testing were done on the shape and scale parameters space in Supplementary Figure 11).

The box plots in each panel show the middle 50% of the data, with the box itself representing the median and the interquartile range (IQR) between the 25th and 75th percentiles. The whiskers extend from the box to the furthest data points within 1.5 times the IQR from the median. Unless otherwise stated, all statistical tests were two-sided.

Icons from panel A, B, F and G were adapted from bioIcons (cancerous-cell-1, lymphocytes-4, macrophage, t-lymphocyte, b-lymphocyte, fibroblast-1 licensed under CC-BY 3.0 Unported by Servier), flaticon.com (bladder icon, https://www.flaticon.com/free-icon/bladder_1453578; head neck icon, https://www.flaticon.com/free-icon/injection_4418017).

Abbreviations: *SR:* spatial relationship; *ICI:* immune checkpoint inhibitors; *1-NN:* first nearest neighbor; *UC:* urothelial cancer.

5.3 DISCUSSION

Advances in ICI have resulted in pembrolizumab (anti-PD1) becoming the second-line standard of care for advanced UC[3], and avelumab (anti-PD-L1) as the standard of care for maintenance after chemotherapy treatment[31]. Results from pre-operative clinical trials show that patients can have a pathological complete response to only two or three cycles of immunotherapeutic treatment[9,10,32,33]. These promising clinical results need biomarkers that

stratify individual patients and improve our understanding of the immunological background of (non-)response. In this study, we provided a comprehensive quantitative exploration of the, thus far, poorly characterized SRs in the UC TME. We show the potential for clinical utility in predicting the response to pre-operative combination ICIs and provide a quantitative basis for follow-up research.

We found an association between the proximity of the SR from $CD8^+$ T cells to cancer cells and response in UC and confirmed that this association also holds in HNSCC. In contrast, no differences between response groups in $CD8^+$ T cell density were found, revealing that abundance alone is likely insufficient to explain treatment response. Tumors with an immune excluded phenotype exhibit an enrichment of $CD8^+$ T cells at the stroma due to mechanisms preventing T cells from reaching the tumor. Quantifying immune phenotypes is not trivial as distinct patterns of exclusion and topography exist[17,34]. In our study, we used exclusion ratios and densities as a proxy to estimate immune phenotypes, but we found no difference between response groups. In contrast, our SR parameters served as a distance metric that objectively quantifies proximity differences between $CD8^+$ T cells and cancer cells between response groups. Therefore, while abundances or ratios between abundances of $CD8^+$ T cells was insufficient to explain response to combination ICIs, a more complex quantification of their relative spatial distribution in relation to cancer cells was a more informative way to describe their behavior within a tumor. Our observations suggest that therapeutic strategies that enhance $CD8^+$ T cell migration closer to cancer cells may overcome resistance to ICI. These results align with the observation that immunosuppressive mechanisms, such as TGF-beta signaling, are associated in UC[10,12] with a $CD8^+$ T cell excluded phenotype and resistance to ICI[35]. Similar results have been reported in the ICI context for melanoma, in which responding tumors to different ICI treatments were characterized using a 1-NN statistic by proximity between proliferating antigen-experienced $CD8^+$ T cells ($CD45RO^+Ki67^+$) and their closest cancer cell[23]. Moreover, in gynecological and non-small cell lung cancer, the SR between tumor-infiltrating lymphocytes (TIL) and non-TILs (e.g., cancer cells) demonstrated its utility for clinical outcome prediction in an ICI cohort[24], which is compatible with our observations in UC and HNSCC.

We found that the proximity of macrophages to cancer cells was positively associated with response in the UC and HNSCC cohorts. Interpretation of this candidate biomarker warrants further investigation due to the plasticity and potential pro- or anti-tumorigenic behavior of macrophages[36], which results in macrophage subtype heterogeneity not covered by our mIF antibody panel. Literature in the ICI context suggests that macrophages can express PD-L1 and PD-1[37] but can also prevent T cells from reaching cancer cells[38]. Data from pancreatic cancer suggest that anti-tumorigenic macrophages (M1-macrophages) are closer to cancer cells than pro-tumorigenic macrophages (M2 macrophages)[39], which

indicates that our proximity signal between macrophages and cancer cells in responding tumors may originate from an M1-type macrophage lineage. In locally advanced esophageal squamous cell carcinoma treated with chemoradiotherapy and SHR-1210 (anti-PD-1 ICI), a prognostic signal using the 1-NN statistic median reported PD-L1^{+} tumor cells closer to PD-L1^{-} macrophages associated with a better OS after treatment[25]. Lastly, non-responding tumors were associated with close proximity between B cells surrounding CD8^{+} T cells, which is in line with the high baseline expression of genes involved in B cell signaling we found in non-responding UC tumors in NABUCCO[10].

We compared the spatial and density metrics' predictive power to corroborate the SR metrics' importance. Our results show a superior predictive power for SR metrics and enhance the limited view that count-derived data, such as density or exclusion ratios, provide of the TME. Furthermore, we compared the SR quantifications on our Weibull parameters with the conventional G-function. While both are based on the 1-NN distance statistic, we showed that the G-function dependence on a distance threshold (T) reduced its utility for group comparisons because the associated G-function's range of values, variance, and predictive power was threshold dependent. Besides, due to the heterogeneity in the G-function evaluations across cell-cell pairwise relationships, there is no unique optimal threshold that maximizes differences between clinical groups of interest for all SRs. Therefore, optimization methodologies for the threshold of choice depend on the SR and cohort, potentially leading to under- or over-fitting and generalization issues. Earlier work on the G-function metric usage for pancreatic cancer grade prediction reported that a single threshold evaluation cannot model all the inherent signals from the data[26]. A higher predictive power could only be achieved by discretizing the G-function at multiple thresholds, which limits its interpretability and utility because of an increased number of summary parameters[26]. On the other hand, our Weibull parameters (shape and scale) allowed for an invariant summary of the SRs without any threshold, which was achieved by, instead of having an empirical summary or discretizing it, modeling all its inherent structure using a curve-fitting approach. Furthermore, the mixed model methodology allowed us to smooth the data and model the parameter variance at a cohort level, making it more suitable for group comparisons when correlating them with clinical phenotypes of interest because of the reduced leverage of outlier samples[40].

The validation of our TME candidate biomarkers for pre-operative combination ICIs response identified in UC indicates biologically relevant SR differences consistent across cancer types. Crucially, despite HNSCC being a different organ and morphologically distinct tumor type, showing variability within their biopsy locations (oral cavity, oropharynx), using our proposed spatial approach, we observed similar average SR distances in both tumor types. A combination of pathological complete response and near-complete response defined response for exploratory analyses in IMCISION (HNSCC). However, in NABUCCO (UC), treatment

response was defined as a pathological complete downstaging at the time of surgery. The response rate in IMCISION was lower than in NABUCCO (36% vs. 58%, p=0.04), thus decreasing the statistical power to quantify differences between response groups. Despite these differences and the relatively small sample sizes (24 and 25 for NABUCCO and IMCISION, respectively), the translatability of our findings on the associations between the SR parameters identified in our UC cohort and treatment response in the HNSCC cohort is promising.

Limitations of our study include the number of antibody markers profiled in mIF data, which restricted the types of cells we could detect. Transurethral resections provide a superficial spatial sampling of the whole TME architecture, therefore allowing for a limited profiling of the tumor margin, which is known to contain a higher abundance of immune cells in UC[34] compared to intratumoral tissue. Nevertheless, the literature suggests that transurethral resection (TUR) material in UC is representative of the whole UC tumor spatial heterogeneity in ~58-73% of cases at an immune cell density level[41], but their associated SRs remain yet unexplored. Limitations to our methodological framework include quantifying SRs by studying only the first nearest neighbor and not beyond. While considering higher-order neighbors could facilitate exploring more distant spatial patterns, the trade-off involves a compromise in the interpretability of the data. Network or graph-based approaches would allow for a broader spatial representation of the TME. However, these topology-based methods usually ignore distances and require more complex SR representations. Furthermore, combinations of samples and pairwise cell type SRs involving noisy distance distributions, such as SRs derived from lowly-populated cells (e.g., FoxP3$^+$ T cells in a subset of UC samples), are excluded from the analysis only when convergence is not reached in the mixed model fitting. However, only 2% of our SRs (24 out of 1176 SRs) were rejected for this reason. While this might have consequences in associations with clinical outcomes of interest (e.g., clinical response), such rare cell types SRs lack robustness. Lastly, our sample sizes are relatively small, and our results warrant further validation in independent and larger cohorts.

In short, our study provides a systematic framework to quantify SRs. It demonstrates that SRs provide a complementary summary of the TME outperforming count-derived metrics, such as density, for identifying biomarkers with a clinical utility. Our results reveal proximity between CD8$^+$ T cells to cancer cells and macrophages to cancer cells as candidate biomarkers for response to pre-operative combination ICIs, which have been thus far unexplored and provide a complementary view of the TME that warrants further investigation.

5.4 METHODS

5.4.1 Study oversight

The studies from this manuscript received approval from the institutional review board of the Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital. The execution of these studies strictly adhered to the protocols and Good Clinical Practice Guidelines outlined by the International Conference on Harmonization, along with the principles established in the 1964 Declaration of Helsinki. Approval for the trial protocols and any subsequent amendments was obtained from the Medical Research Ethics Committee of the Netherlands Cancer Institute—Antoni van Leeuwenhoek Hospital (MREC AVL, https://english.ccmo.nl/mrecs/accredited-mrecs/mrec-netherlands-cancer-institute-the-antoni-van-leeuwenhoek-hospital). Before enrolling in the clinical trials, all participating patients provided written informed consent to partake in the studies.

5.4.2 Urothelial cancer study population and treatment (NABUCCO trial)

Twenty-four pre-treatment Urothelial Cancer (UC) samples from the NABUCCO trial (NCT03387761, Cohort 1) were used for analyses. In the trial patients underwent a combination treatment (2 or 3 cycles) of ipilimumab (anti-CTLA-4) and nivolumab (anti-PD-1) prior to surgical resection. The trial cohort consisted of high-grade stage III muscle-invasive urothelial cancer (cT3-4aN0M0 or cT1-4aN1-3M0). Details of the trial are reported[10].

Response to treatment was evaluated by pathological response assessment on radical surgery. Tumors with a complete pathological response (ypT0N0) or residual disease (<=ypT1N0) were classified as responders (n=14), and tumors with a >=ypT2N0 were classified as non-responders (n=10).

5.4.3 Head and Neck squamous cell carcinoma population and treatment (IMCISION trial)

Thirty-one head and neck squamous cell carcinoma (HNSCC) tumor samples of multiple subsites (oral cavity n=27, oropharynx, n=4) were obtained from the IMCISION trial (NCT03003637). Patients underwent either two cycles of nivolumab (Arm A, n=6) or a combination treatment of 2 cycles of ipilimumab (anti-CTLA-4) and nivolumab (anti-PD-1) (Arm B, n=25) prior to surgical resection. The trial cohort consisted of HNSCC tumors with a histological grade T2–T4N0–N3b and metastatic grade M0 primary or recurrent of mostly HPV-negative head and neck squamous cell carcinoma (HPV negative, n=23; HPV positive, n=2). Details of the trial can be found elsewhere[28]. Only samples from Arm B (n=25) were analyzed in this manuscript.

Response to treatment was evaluated by pathological response assessment on surgery and by comparison of tumor cells decrease from baseline to on-treatment samples[28]. Tumors with <=10% tumor cell percentage (TCP) at surgery and a decrease of 90-100% in tumor cells from baseline to on-treatment were classified as major pathological responders (MPR, n=9); tumors with <=50% TCP at surgery and a decrease of 50-89% in tumor cells from baseline to on-treatment were classified as partial pathological responders (PPR, n=1); else tumors were classified as no pathological responders (NPR, n=15). Patients with an MPR were classified as Responders, and patients with a PPR or NPR were classified as Non-Responders.

5.4.4 Multiplex Immunofluorescence

Multiplex immunofluorescence (mIF) was performed on pre-operative baseline formalin-fixed paraffin-embedded (FFPE) tumor resections and assessed on an immune panel (DAPI, PanCK, CD8, CD3, FoxP3, CD20, CD68) as previously described for UC[10] (NABUCCO) and HNSCC[28] (IMCISION). The experimental protocol and data processing is reported elsewhere[10,28].

Antibodies used for the NABUCCO trial dataset were CD3 (1/400 dilution, Clone P7, Cat RM-9107-S, ThermoScientific), CD8 (1/100 dilution, Clone C8/144B, Cat M7103, DAKO), CD68 (1/500 dilution, Clone KP1, M0814, Dako), FoxP3 (1/50 dilution, Clone 236A/47, Cat ab20034, Abcam), CD20 (1/500 dilution, Clone L26, cat M0755, Dako), PanCK (1/100 dilution, Clone AE1AE3, Cat MS-343P, Thermo Scientific). Antibodies used for the IMCISION trial dataset were CD3 (clone SP7, ThermoScientific, CatalogNo: RM-9107-S, LotNo: 9107S1805A), CD8 (clone C8/144B, DAKO / Agilent, CatalogNo: M7103, LotNo: 20048132), CD68 (clone KP1, DAKO / Agilent, CatalogNo: M0814, LotNo: 20040389), FoxP3 (clone 236A/47, DAKO / Agilent, CatalogNo: ab20034, LotNo: GR3220121-1), CD20 (clone L26, DAKO / Agilent, CatalogNo: M0755, LotNo: 20038880), PanCK (clone AE1AE3, Thermoscientific, CatalogNo: MS-343P, LotNo: 343P1205H).

Upon mIF profiling, cells were segmented by marker positivity, and classified as Cancer cells ($PanCK^+$), CD8 T cells ($CD3^+CD8^+$), FoxP3 T cells ($CD3^+FoxP3^+$), CD4 T cells ($CD3^+CD8^-FoxP3^-$), Macrophages ($CD68^+$) and B cells ($CD20^+$).

Multiplex immunofluorescence in IMCISION was assessed as in NABUCCO and the experimental protocol is published in the original manuscript[28]. To ensure consistency in the mIF spatial data between the HNSCC and UC cohorts, including similar tumor purity, for each sample from IMCISION we aligned analysis methods by discarding stromal tissue residing beyond 150 microns distance from the tumor tissue by filtering out all cells classified as belonging to the 'Stroma compartment' (by segmentation) if their closest cancer cell lay beyond 150 microns by using the nncross method from *spatstat*.

The data subjected to downstream analysis represented the position of each cell in the tissue (x- and y-coordinates of the nuclei) and its corresponding cell type.

5.4.5 Segmentation of tumor and stroma compartments

To segment the tumor and stroma regions from each tissue, we first split each individual tissue island from each sample biopsy using dbscan v1.1-6 (density-based spatial clustering of applications with noise) by setting the size of the epsilon neighborhood to 300 and the minimum number of points in the epsilon neighborhood to 50 (Supplementary Figure 1A). Each tissue island was named foci.

To segment the tumor and stroma compartments for each focus, we first computed the kernel density estimation (KDE) of the point pattern defined by cancer cells (KDE_{tumor}), and by negative cells ($KDE_{negative}$). The KDE was estimated as implemented by density in the stats v3.6.3 package. The smoothing bandwidth for the KDE was optimized using likelihood cross-validation as implemented in *bw.ppl* in spatstat v1.64-1. Then, the KDEs were normalized to their maximum value ($KDE_{tumor} = KDE_{tumor} / \max(KDE_{tumor})$) to allow comparison between the KDE of the tumor and negative cells. To segment the tissue, for each position populated by a cell, we compared both KDEs, and classified them as "Tumor" when $KDE_{tumor} > KDE_{negative}$ and as "Stroma" otherwise (Supplementary Figure 1B-C).

5.4.6 Calculation of tumor and stroma compartment areas

To compute the covered area by each segmented tissue compartment ("tumor" or "stroma"), we first computed the kernel density estimation (KDE) of the point pattern defined by the cancer cells (KDE_{tumor}) and normalized by maximum KDE intensity. Then, to compute the area of the tumor compartment, we filtered out all the KDE pixels with a normalized intensity < 0.1. We selected this threshold based on visual exploration for all cells. We then estimated the tumor compartment area as the aggregated area of all non-filtered pixels from KDE_{tumor} (thus with intensity >= 0.1).

To compute the total tissue area, we also computed the $KDE_{negative}$ and normalized it by the maximum value. We then summed the KDE_{tumor} and $KDE_{negative}$, and filtered out pixels with a normalized intensity < 0.1. We then estimated the Total tissue area as the aggregated area of non-filtered pixels from ($KDE_{tumor} + KDE_{negative}$) (intensity >= 0.1).

To estimate the area of the "stroma" compartment, we subtracted the "tumor area" from the "total tissue" area (Supplementary Figure 1D). This process was performed for each foci.

5.4.7 Spatial analysis: quantification of the first nearest neighbor (1-NN) distance distribution

The spatial relationships between all cells within the tumor microenvironment were studied using the first-nearest neighbor (1-NN) statistic as implemented in *spatstats*. In brief, the approach is studied from a reference cell type to a neighbor cell type. For each reference cell type ("cell type from"), the distance to the closest neighbor cell type ("cell type to") was measured using *nndist* (Figure 1E). Then, we constructed a histogram from the vector of 1-NN distances. We smoothed the distribution by sliding a 5-micron window across the 1-NN histogram and iteratively counting the frequency of the 1-NN distances for each micron. We normalized the distribution to achieve a unit area under the curve (AUC) by dividing for the numerical AUC. SRs were quantified using the data from the whole tissue slide (i.e., not making a distinction beween tumor and stroma compartments).

5.4.8 Spatial analysis: fitting of Weibull distribution to the 1-NN distances vector

To summarize the 1-NN distance distribution, we fitted a Weibull distribution to the empirical probability density function (PDF), which is a 2-parameter distribution based on (the positive parameters) shape and scale, defined as:

$$(1)\ f\left(x, b, a\right) = \frac{a}{b}\left(\frac{x}{b}\right)^{a-1} e^{-(x/b)^{a}}$$

Here, b denotes the scale, and a denotes the shape.

We implemented a methodology based on a functional data analysis approach to fit the Weibull distribution. First, to have an initial estimate of the distribution parameters for each patient (n=24) and cell type-cell type combination (n=49), we employed maximum likelihood estimation (MLE) using fitdist as implemented on fitdistrplus v1.1.3 package to have an initial estimate of the scale and shape parameters. Then, for each pairwise cell type relationship (cell type from vs. cell to), we implemented a non-linear mixed effect model (nlme v3.1-144) to fit a Weibull distribution on all patient samples, having the shape / scale intercept as fixed effects (fixed = a+b~1) and allowing a random effect for the scale/shape on each sample (random = list{sample=pdDiag(a+b~1)}) by modeling the correlation structure of the random effects a with diagonal positive-definite matrix. The nlme model was implemented by maximizing the restricted log-likelihood (method='REML'), with the parameters set to their default values and the control values set as:

- 1000 maximum iterations for the optimization algorithm (maxIter = 1000).
- 200 maximum iterations for the optimization step which is inside the nlme optimization process (msMaxIter=200).
- 1e-1 tolerance for the PNLS step convergence criterion (pnlsTol=1e-1).

- 1e-6 tolerance for the convergence criterion in nlme algorithm (tolerance=1e-6).
- Nonlinear minimization optimizer (opt="nlm").

To filter out low-quality distributions that lead to non-convergence of the models, we sequentially filtered out samples based on the number of cells (n) from the reference cell type (nFROM) or neighbor cell type (nTO). First, we fitted a model with the data for the 24 samples. If convergence was not achieved, we filtered out all samples with fewer than 20 cells (nFROM < 20 or nTO < 20). If the model still did not converge, we repeated filtering samples with fewer than 50, 70, or 100 cells. The approach allowed us to model as much data as possible unless the goodness of fit was compromised. For 24 spatial relationships for which a data fit could not be carried out (2% of the total combinations of data points from 24 samples and 7x7 cell type pairwise relationships) in our cohort.

Because the positivity constraint for the Weibull distribution parameters (b>0, a>0) could not be optimally implemented using a constrained non-linear mixed effect model, we re-parameterized the shape and scale parameters to force the values to be positive. In brief, we re-parameterized the scale and shape employing new parameters *A* and *B*, which were unconstrained:

(2) $a = \frac{10}{1 + e^{-A}}$, with A: unconstrained and aϵ(0, 10)

(3) $b = \frac{500}{1 + e^{-B}}$, with B: unconstrained and b $\epsilon^{0,500}$

Here, *a* is the shape and *b* is the scale.

Unless otherwise stated, the SR parameters reported in the manuscript correspond to the ones calculated on the first-nearest neighbor distributions using the spatial distribution of the whole tissue slide (thus, not using data stratified by tumor or stroma).

Sources of the variability of the SR parameters were also quantified, as reported in Extended Methods.

5.4.9 Spatial analysis: Computation of the G-function

Alternatively to fitting a distribution to the 1-NN distance distribution, we computed the G-function, which is defined as the cumulative distribution function (CDF) of the first-nearest neighbor distance distribution:

(4) $G - function\,(r) = probability(1 - NN\ distance \leq r)$

The G-function was computed as implemented by *gest* in the spatstat package. Because our SRs were studied between different cell types, we used the multitype nearest-neighbor function $G_{ij}(r)$ as implemented in Gcross, which was calculated from the first nearest neighbor distances from a cell of type i to the nearest point of type j. Then, to summarize the G-function, we computed the Area Under the Curve (AUC) of the G-function, named G-AUC-T, at different thresholds T's, which included 25, 50, and 100 microns (i.e., G-AUC-25, G-AUC-50, and G-AUC-100, respectively).

5.4.10 Spatial analysis: Weibull-derived G-function

Using the properties of the Weibull distribution, we analytically constructed the Cumulative Distribution Function (CDF) using the shape and the scale parameters:

$$(5)\ Analytical\ CDF\left(r\right) = 1 - exp\left(-(r/b)^{a}\right)$$

Here, *b* and *a* denote the scale and shape, respectively, and *r* denotes the first nearest neighbor distance.

The Analytical CDF is analogous to the Analytical G-function. The AUC of the Analytical CDF at was evaluated at different distance thresholds T's, and referred to in the manuscript as Weibull-G-AUC-T.

5.4.11 Comparison of discriminative power between spatial and density-related parameters

We compared the predictive power between ICI response groups of the spatial-related parameters (shape, scale), density-based parameters (intratumoral and stromal immune cell density) and exclusion-based parameters (ratio between stromal and intratumoral immune cell densities). First, for each set of parameters (e.g., spatial-related parameters), we trained a logistic regression model using glm to predict response. Then, a ROC curve was built using the logistic regression's fitted values (probabilities) using pROC v1.17.0.1. For each ROC-AUC, we tested whether the AUC was significantly different from AUC=0.5 using a two-sided Wilcoxon signed-rank test between cases and controls. We used a bootstrapped approach (n=500) to estimate the confidence intervals of the ROC-AUCs and to test whether the ROC-AUCs from the spatial parameters were significantly greater than the ROC-AUCs from either the density or the exclusion ratio parameters.

5.4.12 Comparison of discriminative power between Weibull-derived or G-function-derived parameters

Logistic regression deviance of Weibull parameters (shape, scale) and G-function parameters (G-AUC-T evaluated at different Ts) was evaluated by training univariate logistic regression (LR) models using data from each feature as a predictor, and clinical response labels as the dependent variable. Logistic regression deviance (logit deviance) was assessed, in which lower

values denote better model fits. The uncertainty of the AIC was estimated by performing a leave-one-out (LOO) variant of the analysis. A two-sided student's t-test tested statistical significance.

5.4.13 Spatial biomarker validation in HNSCC

We validated our top SR biomarkers associated with clinical response in our UC cohort (NABUCCO). First, we selected the top two biomarkers identified using our pipeline in UC (FDR < 0.04 in either the shape or scale parameter, which yielded the SRs *$CD8^+$ T cells to cancer cells* and Macrophages to cancer cells*)*. Second, we evaluated the shape and scale parameters for the biomarkers mentioned above between clinical response groups in the external HNSCC (IMCISION) cohort using a two-sided t-test and adjusted for multiple hypothesis testing using the Benjamini-Hochberg method. Then, we evaluated the median 1-NN distances using the analytical derivation from the shape and scale parameters:

$$(6)\ Median\ 1-NN\ distance\left(shape, scale\right) = scale * (ln2)^{1/shape}$$

5.4.14 Statistical analysis

Unless otherwise stated, a two-sided student's t-test was used for group comparisons. We modeled density and count data in a logarithmic space. For spatial data, the shape of the Weibull parameters (*a*) was modeled on a non-logarithmic scale, and the scale (*b*) was studied on a logarithmic scale. Multiple hypothesis testing corrections were done using the Benjamini-Hochberg method. Unless otherwise noted, statistical significance was defined as p<0.05 and False Discovery Rate (FDR) < 0.10 (10%), and all statistical tests were two-sided. All statistical analyses were performed in R 3.6.3. The following packages were used in this study:

- Spatstat 1.64 [29]
- Dplyr 1.0.4
- Fitdistrplus 1.1.3
- Patchwork 1.1.1
- Survival 1.3.24
- ComplexHeatmap 2.2
- Circlize 0.4.12
- Glmnet 4.1.1
- RColorBrewer 1.1.2
- Nlme 3.1.144
- Spastat 1.7.0
- Ggpubr 0.4.0
- Ggrepel 0.9.1
- Plyr 1.8.6

- Tidyverse 1.3.0
- Ggplot 2.3.3.3
- Tibble 3.0.6
- ggrastr version 1.0.1
- pROC 1.17.0.1

5.5 SUPPLEMENTARY METHODS

5.5.1 Spatial analysis: a simulation study to quantify spatial relationship parameters sources of variation

Using the segmentation between the tumor and stroma compartments and the tissue architecture, we simulated different states of immune cell infiltration or cell location to quantify sources of variation in the spatial patterns. We used our original samples' point patterns and the delineated tumor and stroma compartments in all the simulations. We altered the abundance and location of three cell types: immune, cancer, and stroma.

First, we simulated immune cell density at different values. Here, all the cells belonging to the tumor compartment were labeled as cancer cells, and the cells belonging to the stroma compartment were labeled as stroma cells. Then, we randomly re-labeled cells a fraction of cancer cells (at a 1%, 2%, 5%, 7%, and 10% fraction) as immune cells, in which the SR of the immune cells resembled a homogeneous spatial distribution. Analogously, the same simulation study was carried out by randomly re-labeling stroma cells or altogether re-labeling cancer and stroma cells. These perturbations led to three different simulation studies (immune cells only in tumor, only in stroma, or in tumor and stroma), in which the SR parameters from/to cancer, immune and stromal cells (3x3=9 SRs) were estimated. For each simulation, the immune cell density (either intratumoral or stromal) was associated with the distinct SR Weibull parameters (shape and scale) obtained for the 9 SRs being studied.

In a second simulation study, we compared the SR parameters between the previous three different simulations that matched distinct immune phenotypes, being *Inflamed* (simulated immune cells in Tumor), *Excluded* (simulated immune cells in Stroma), and *Mixed* (simulated immune cells altogether in Tumor and Stroma). This study aggregated and compared simulations between groups, e.g., *Inflamed* vs *Excluded (*i.e., *Inflamed* combines simulations with an inflamed phenotype at multiple densities), as we aimed to quantify the global effect of the perturbation.

Lastly, we simulated differences in the local immune cell arrangement by perturbing the clustering of immune cells. In *Clustered* simulations, we allocated the immune cells next

to each other by setting an anchor point in which immune cells were present at different abundances (1%, 2%, 5%, and 10% fraction of cells being immune cells). Using this simulation (*Clustered*), we compared the SR parameters with the previous simulation, in which immune cells were placed along the tissue following a *Homogeneous* spatial distribution.

5.6 SUPPLEMENTARY NOTES

5.6.1 Supplementary Note 1: Quantifying sources of variation of the spatial relationship parameters

To investigate whether SR metrics are susceptible to changes in immune cell density, we performed a simulation study based on the real patient mIF data (Supplementary Methods, *Spatial analysis: a simulation study to quantify spatial relationship parameters sources of variation*). We preserved the spatial positions of cells (but not their identity) and the tumor and stroma delineation. We then placed immune cells at an increasing density at the preserved cell positions for the cohort. Supplementary Figure 12A shows a representative sample where the immune cells are only placed in the cancer compartment, i.e., replacing a cancer cell with an immune cell. We varied the percentage of immune cells in the tumor compartment from 1 to 10%, where the percentage is a fraction of the total number of cells on a slide. This range is representative of the variation in immune cell fraction we observed in our cohort (2-1583 immune cells/micron^2).

We started by simulating multiple intratumoral immune cell densities, which included *Low* (1% of immune cells), *Medium* (4% of immune cells), or *High* (7% of immune cells) densities (Supplementary Figure 12A). We then estimated all the pairwise SR parameters between cancer, immune and negative cells. As expected, an increase in immune cell density decreased the 1-NN distances between immune cells (Supplementary Figure 12B, and Supplementary Figure 13A, right panel). The associated medians for each 1-NN spatial distribution shown in Supplementary Figure 12B confirmed these findings (Supplementary Figure 13A). In contrast, changes in immune cell density did not affect the 1-NN curves nor the metrics quantifying the SR from immune to cancer cells (Supplementary Figure 13A). Our findings showed little variation between the different patient samples (Supplementary Figures 14A-B). In summary, we found that density can affect the SR metrics, but that is highly dependent on the cell types involved and the directionality of the relationship.

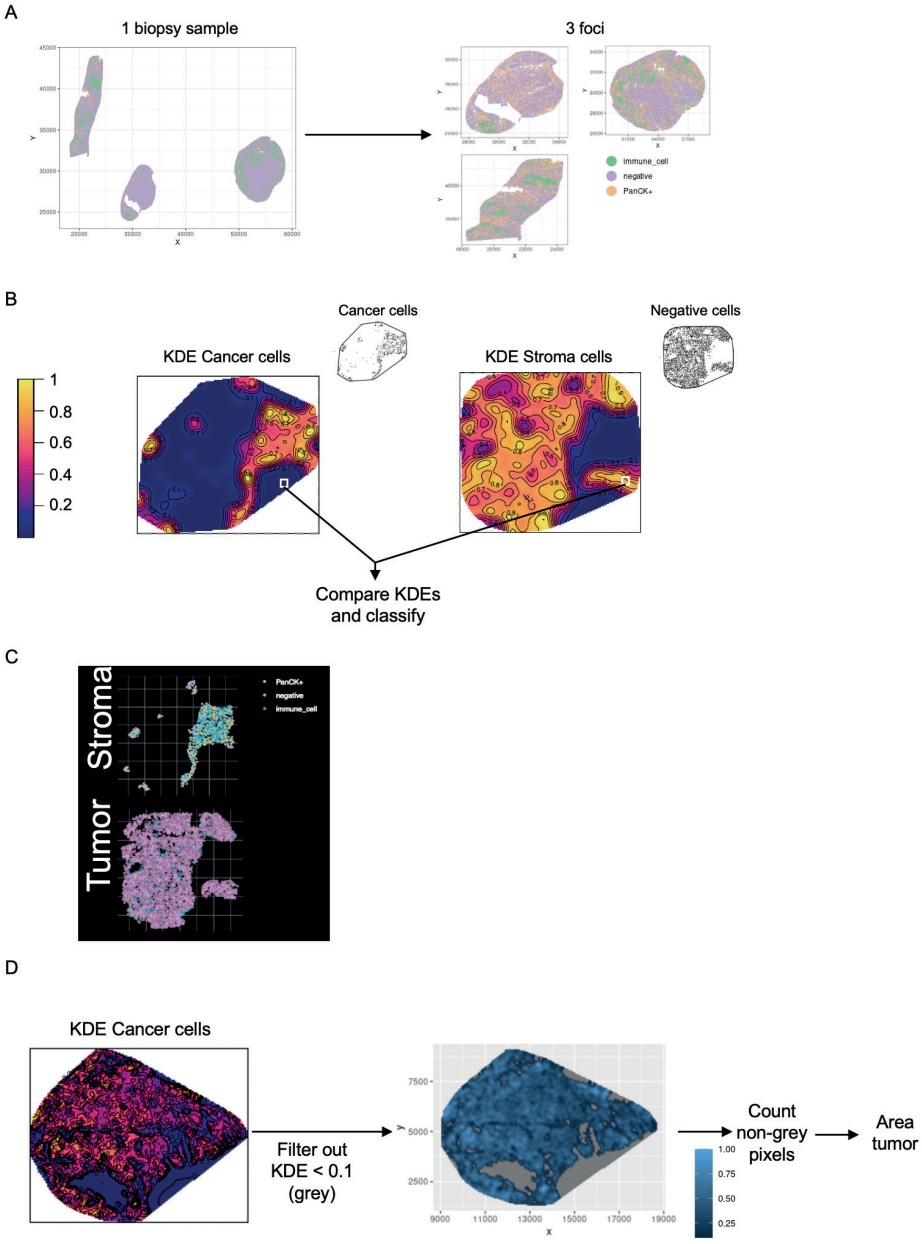
We can conceive the following immune cells spatially distributions giving rise to the following configurations of immune phenotypes[30]. First, being Excluded (higher immune cell abundance in the stroma), Inflamed (high immune cell abundance in the tumor), and Desert (low immune cell abundance in the tumor and stroma). In order to quantify the effect of immune

phenotypes on the SR metrics, we simulated the associated extreme scenarios of the immune phenotypes, and named them as *Excluded* (immune cells only in the stroma and not in the tumor), *Mixed* (immune cells both in the stroma and tumor), and *Inflamed* (immune cells only in the tumor), all at distinct immune cell densities (Supplementary Figure 12C). As expected, the different immune phenotypes have a distinct effect on the SR of immune to cancer cells. In contrast, the SR immune to immune cells remains unaffected, similar to the observation when the immune cell density was varied (Supplementary Figures 12D, 13B). Notably, these effects were remarkably stable across the complete cohort (small variance in Supplementary Figure 14A-B), indicating that the specific positions of cells and the arrangement of tumor and stroma in a particular tumor do not have a strong effect on the simulated SRs. Cases from a particular configuration (e.g., *Inflamed*) consisted of aggregated simulations of distinct immune cell densities compatible with such configuration to assess the true effect between immune phenotypes in our comparisons (e.g., *Inflamed* vs. *Excluded*). Again, little sample variability within the whole cohort was observed between the associations of immune phenotype perturbation and SR parameters (Supplementary Figure 14C-D).

Next, we compared the SR metrics for homogeneously and heterogeneously distributed immune cells. Examples of such instances include T-cells homogeneously infiltrating a tumor or B-cells forming immune cell clusters resulting in non-homogeneous distributions of B-cells. We simulated these two scenarios (Supplementary Figure 12E) and compared the associated SR metrics. We found that immune cell clustering can affect the SR metrics from immune to immune cells and from immune to cancer cells (Supplementary Figures 12F, 13C, 14E-F).

To summarize our findings, we aggregated simulations for all the cell type pairwise relationships and samples (Supplementary Figure 14G). Changes in the SR parameters upon the posed perturbations highly depend on which reference and target cell types were studied (Supplementary Figure 15). The SRs between abundant cell types (i.e., cancer and negative cells) were not altered upon such perturbations. In contrast, the SRs between rare cell types (i.e., immune cells) were affected by the density or local clustering perturbations and SRs from abundant to immune cells were affected by the density perturbations. Moreover, immune phenotypes and clustering perturbations affected SRs from immune to cancer cells. In conclusion, we identified multiple factors that affect the magnitude of SR parameters, which depending on the cell type context being studied, need to be considered in downstream analysis.

5.7 SUPPLEMENTARY FIGURES



Supplementary Figure 1. Segmentation of tumor and stroma compartments and tissue area assessment.

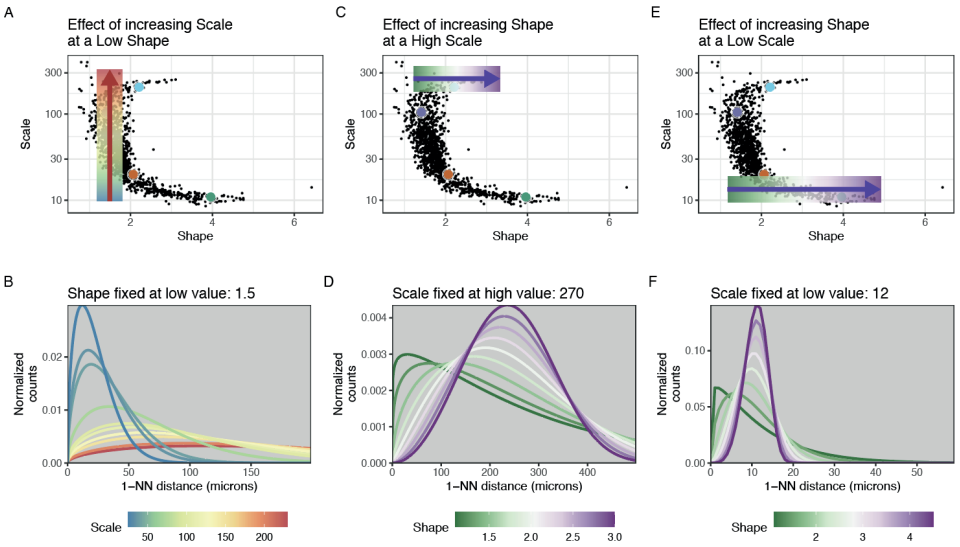
(A) Visual representation of a biopsy tissue slide (left) that contains multiple tissue islands. Each tissue island was identified using the dbscan algorithm and named foci (right).

(B) Tumor and stroma regions segmentation by comparison of the Kernel density estimation (KDE) computed using only cancer cells (left) or stroma cells (right). Prior to segmentation, KDEs were normalized by the maximum value.

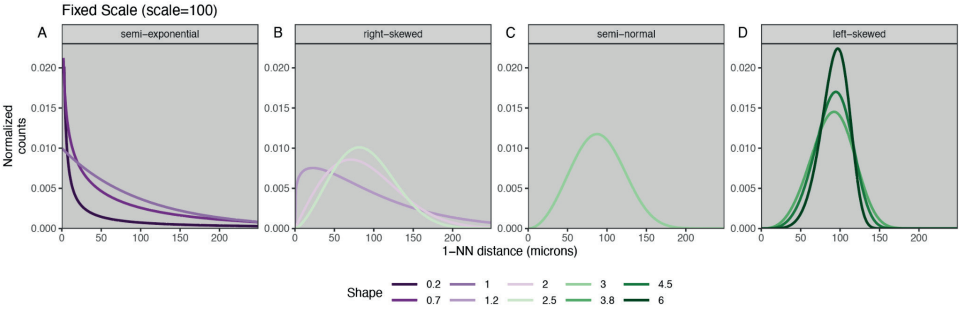
(C) Example of tumor and stroma classifications. Cells in the top and bottom panels are classified as being in the stroma or the tumor compartment, respectively. Coloring denotes the cell type classification by HALO.

(D) Visual representation of the area calculation of the tumor region. First, a KDE using tumor cells was computed and normalized by the maximum value. Then, all the pixels with a KDE below 0.1 were filtered out (drawn as gray in the middle panel). Lastly, all the remaining pixels were counted, and used for the area estimation.

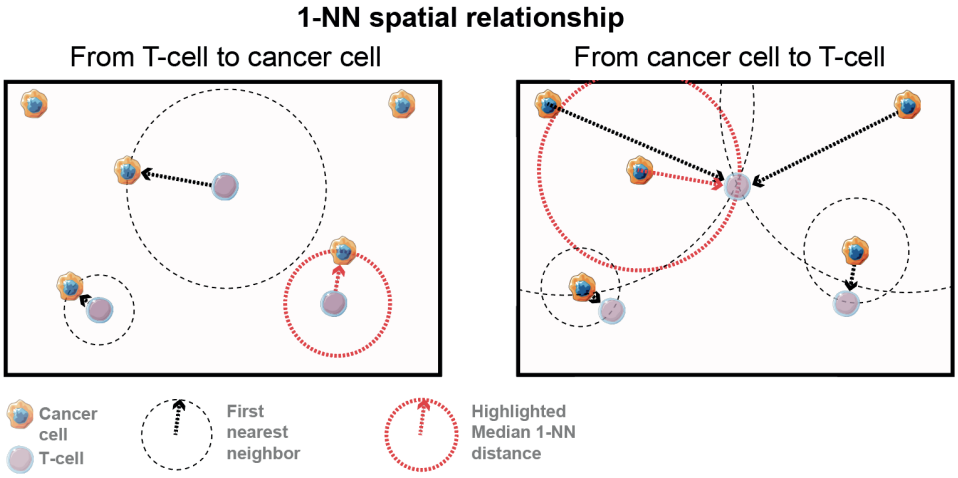
Abbreviations: *KDE*: kernel density estimation.



Supplementary Figure 2. Effect of the modeled shape and scale parameters on the Weibull distribution.
A-B) Differences at the Weibull distribution when increasing the scale parameter at a fixed shape.
C-D) Differences at the Weibull distribution when increasing the shape parameter at a fixed scale (high scale).
E-F) Differences at the Weibull distribution when increasing the shape parameter at a fixed scale (low scale).



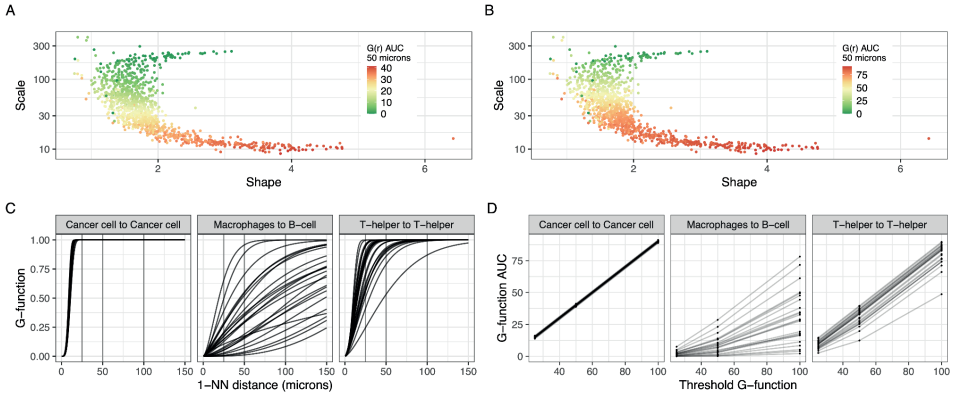
Supplementary Figure 3. Representation of a Weibull distribution when altering the shape parameter at a fixed scale. Each facet groups Weibull distributions grouped by a similar behavior: (A) Represents Weibull distributions resembling a semi-exponential distribution, (B) represents Weibull distributions resembling a right-skewed distribution, (C) represents Weibull distribution resembling a semi-normal distribution and (D) represents Weibull distributions resembling a left-skewed distribution.



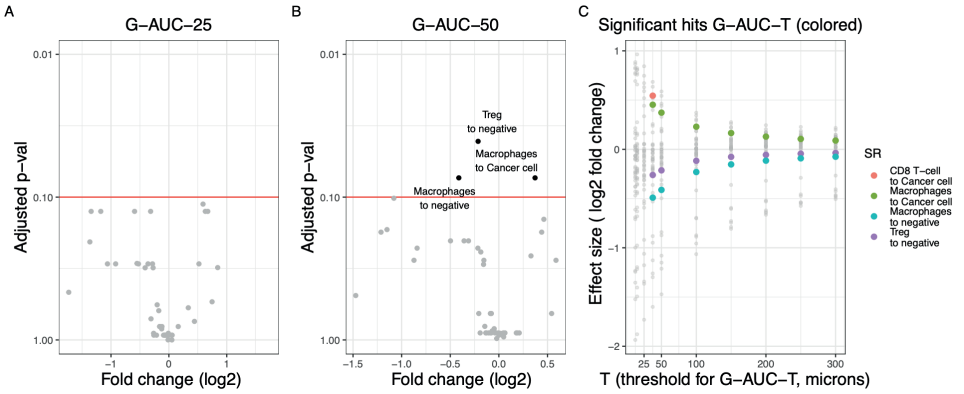
Supplementary Figure 4. SRs were summarized using the 1-NN statistic studied from a reference cell type to a target cell type. The 1-NN statistic is not symmetrical; panels A and B illustrate the statistic when T-cells (left panel) or cancer cells (right panel) are selected as a reference cell type, respectively. Only the 1-NN distances are drawn in each panel. Distances in red indicate the median 1-NN distance for that particular SR.

Icons from panel were adapted from bioIcons (cancerous-cell-1, lymphocytes-4, t-lymphocyte licensed under CC-BY 3.0 Unported by Servier).

Abbreviations: *SR*: spatial relationship; 1-NN: first nearest-neihbor.



Supplementary Figure 5. Comparison of spatial parameters derived from the 1-NN distance statistic.
(A) Scale vs. shape parameter space associated with the G-function evaluated at 12.5 microns (coloring, G-AUC-12).
(B) Scale vs. shape parameter space associated with the G-function evaluated at 50 microns (coloring, G-AUC-50).
(C) G-functions for the B-cell to B-cell, $CD8^+$ T-cell to Cancer cell, and Macrophages to B-cell SRs. Lines join samples. Vertical lines denote the thresholds 25, 50 and 100 microns.
(D) G-AUC-T evaluated at 25, 50 and 100 micron threshold for the examples shown in panel (C). Lines join SRs from the same samples.
Abbreviations: *G-AUC-T*: G-function evaluated at a threshold T; *T*: threshold; *AUC*: Area under the curve, *1-NN*: first nearest-neighbor.



Supplementary Figure 6. Association of spatial relationships derived from the G-function with response to pre-operative ipilimumab+nivolumab in urothelial cancer.

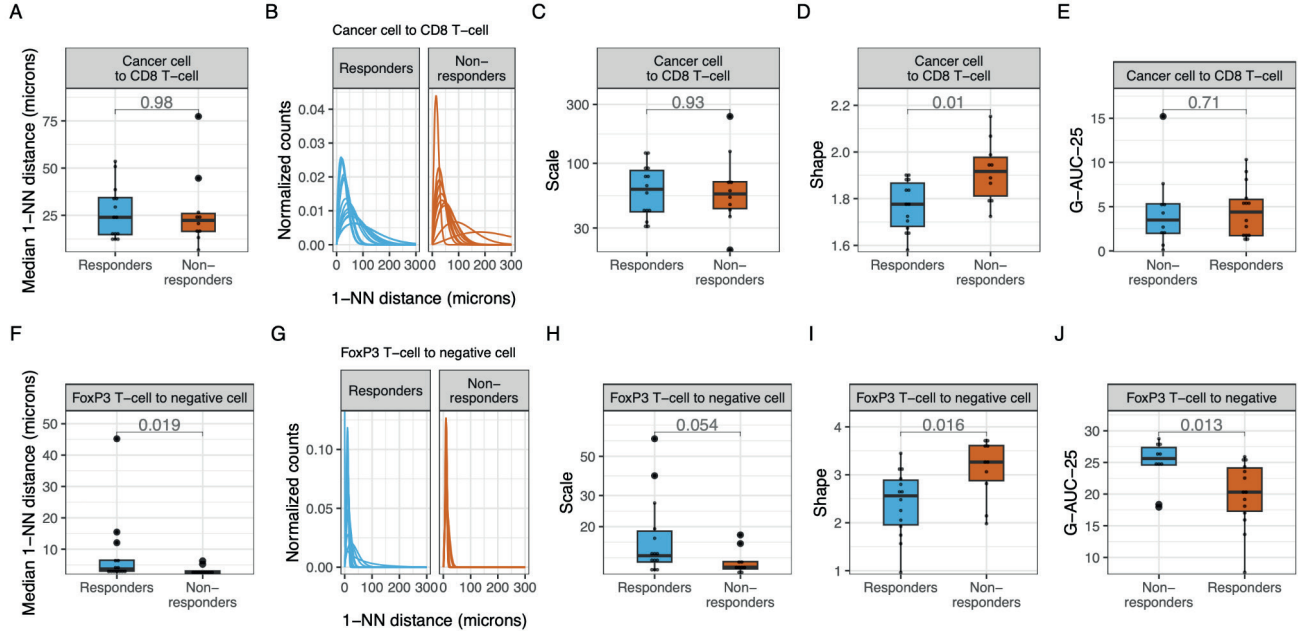
(A) Volcano plot showing the fold change on the G-function parameter evaluated at 25 microns (G-AUC-25) between response groups (x-axis, $\log_2$ fold change, n=14 independent responders and n=10 independent non-responders) and statistical significance by Wilcoxon test adjusted by multiple hypothesis testing (y-axis).

(B) Volcano plot showing the fold change on the G-function parameter evaluated at 50 microns (G-AUC-50) between response groups (x-axis, $\log_2$ fold change, n=14 independent responders and n=10 independent non-responders) and statistical significance by Wilcoxon test adjusted by multiple hypothesis testing (y-axis).

(C) Fold change (y-axis, $\log_2$ scale) on the G-function parameter evaluated at different thresholds T (x-axis) between response groups. Fold changes that are significantly associated with response as evaluated with a Wilcoxon test adjusted by multiple hypothesis testing are depicted in color. Non-significant fold changes are depicted in grey.

Unless otherwise stated, all statistical tests were two-sided.

Abbreviations: *G-AUC-T*: G-function evaluated at a threshold T; *T*: threshold; *AUC*: Area under the curve; *p-val*: p-value; *SR*: spatial relationship.

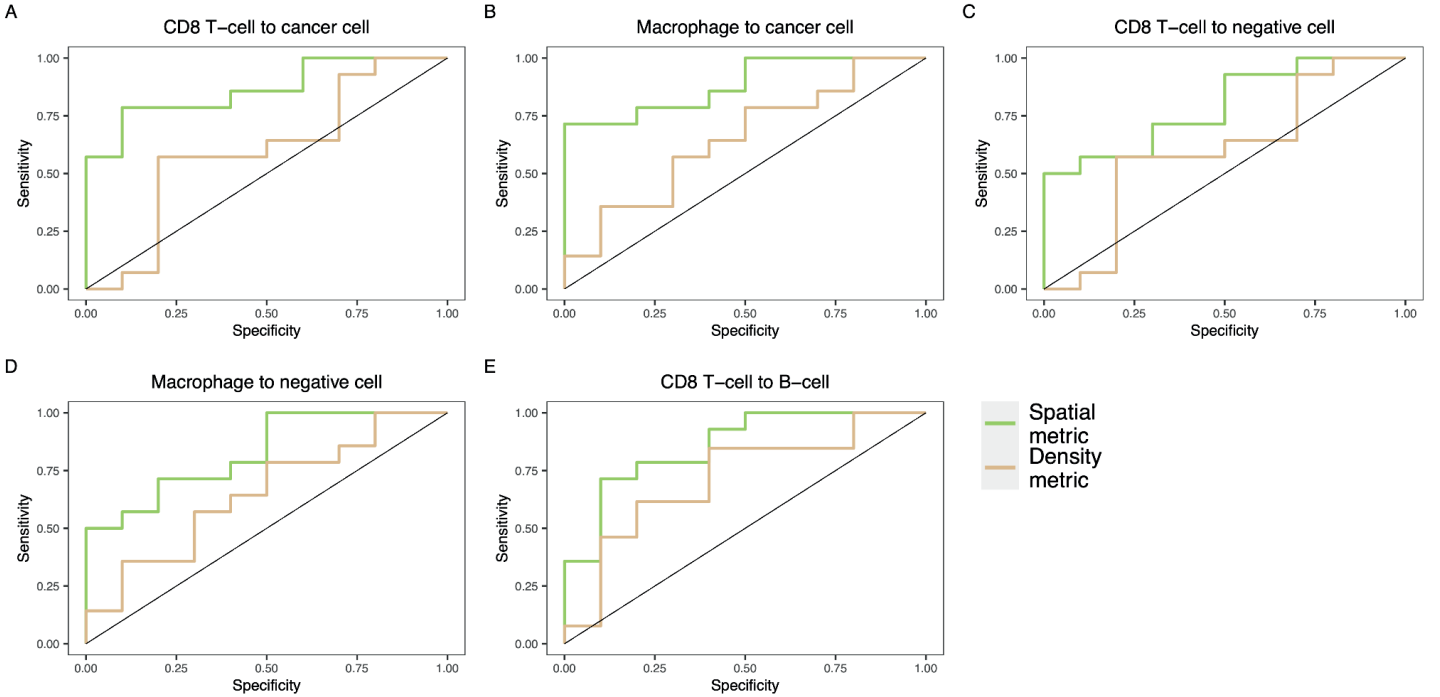


Supplementary Figure 7. Association of spatial relationships with response to pre-operative ipilimumab+nivolumab in urothelial cancer in hits with a weak association.

(A) Median 1-NN distances, (B) associated 1-NN curves, (C) scale parameters and (D) shape parameters and (E) G-AUC-25 parameters stratified by response groups (n=14 independent responders and n=10 independent non-responders) for the SR from cancer cell to CD8$^+$ T-cells (top row) and (F) Median 1-NN distances, (G) associated 1-NN curves, (H) scale parameters and (I) shape parameter and (J) G-AUC-25 parameters stratified by response groups for the SR from FoxP3 T-cells to negative cells (bottom row). Statistical significance was assessed by a Mann-Whitney test (panels A, E, F, J) and a t-test (C, D, H, I).

The box plots in each panel show the middle 50% of the data, with the box itself representing the median and the interquartile range (IQR) between the 25th and 75th percentiles. The whiskers extend from the box to the furthest data points within 1.5 times the IQR from the median. Unless otherwise stated, all statistical tests were two-sided and no adjustments for multiple hypotheses were made. Adjusted p-values: FDR_{scale}(cancer cell to CD8$^+$ T-cell)=0.98, FDR_{shape}(cancer cell to CD8$^+$ T-cell)=0.095, $FDR_{G\text{-}AUC\text{-}25}$(cancer cell to CD8$^+$ T-cell)=0.93, FDR_{scale}(FoxP3 T-cell to negative cell)=0.21, FDR_{shape}(FoxP3 T-cell to negative cell)=0.11, $FDR_{G\text{-}AUC\text{-}25}$(FoxP3 T-cell to negative cell)=0.13, $FDR_{G\text{-}AUC\text{-}25}$(FoxP3 T-cell to negative cell)=0.04.

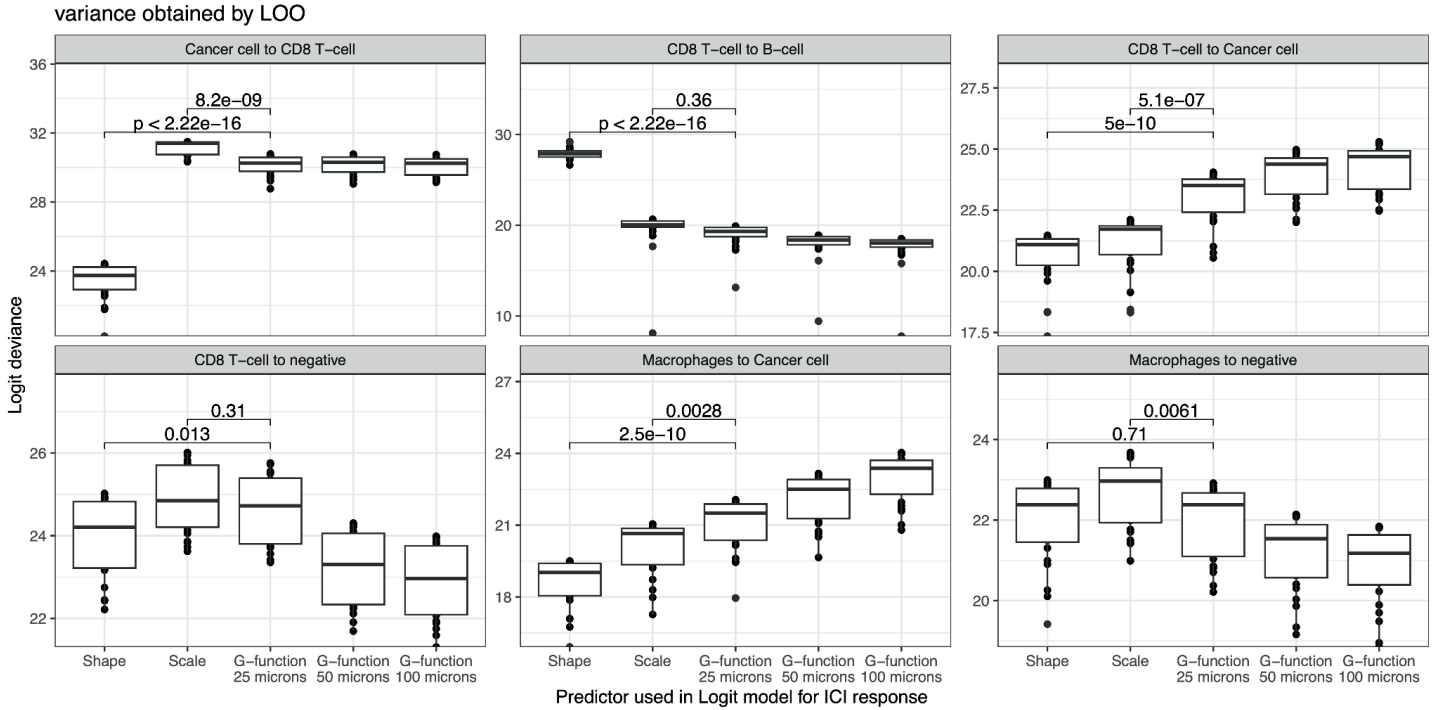
Abbreviations: *1-NN*: first nearest neighbor; *SR*: spatial relationship; *G-AUC-25*: AUC of the G-function evaluated at 25 microns; *AUC*: area under the curve..



Supplementary Figure 8. Comparison of the predictive power of associations with response to pre-operative ipilimumab+nivolumab in urothelial cancer between spatial relationships and density parameters.

ROC curves of the discriminative power of the spatial and density-related parameters involving significant SRs that are associated with response (n=14 independent responders and n=10 independent non-responders), involving (A) CD8 T-cells to cancer cells, (B) Macrophages to cancer cells, (C) CD8 T-cells to negative cells, (D) Macrophages to negative cells, and (E) CD8 T-cells to B-cells. ROC-plots from 'Spatial metrics' (green) were built upon the predictions of a logistic regression model trained on the shape and scale parameters of the associated SR. ROC-plots from 'Density metrics' (orange) were built upon the predictions of a logistic regression model trained on the intratumoral density, stromal density, and exclusion ratio (ratio between stromal and intratumoral density) for the immune cells associated in the SR.

Abbreviations: *SR*: spatial relationship; *ROC*: receiver operating characteristic.

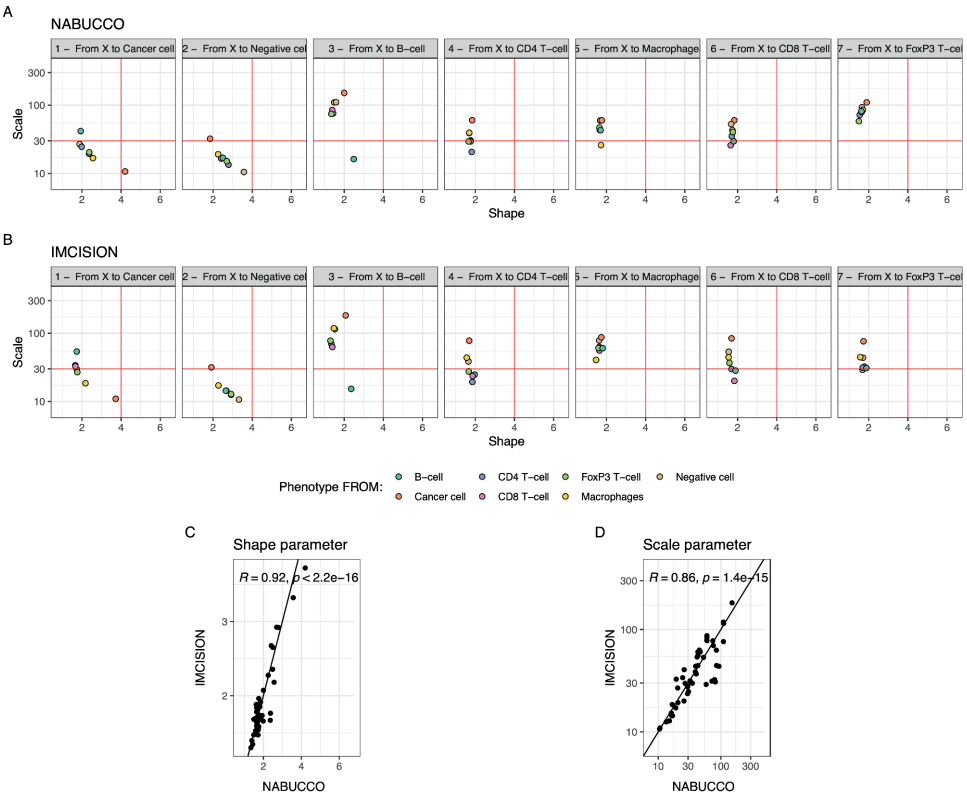


Supplementary Figure 9. Comparison of predictive models of response to pre-operative ipilimumab+nivolumab in urothelial cancer trained using spatial parameters derived from the 1-NN distance statistic.

Logistic regression deviance of a univariate logistic regression model predicting ICI response (n=14 independent responders and n=10 independent non-responders) using as a predictor the shape, the scale, or the G-function evaluated at different thresholds (G-AUC-T). Variability on the AIC was evaluated by leave-one-out cross-validation. Each panel denotes models trained using SRs studied for different cell-cell pairwise relationships.

The box plots in each panel show the middle 50% of the data, with the box itself representing the median and the interquartile range (IQR) between the 25th and 75th percentiles. The whiskers extend from the box to the furthest data points within 1.5 times the IQR from the median. Unless otherwise stated, all statistical tests were two-sided. Statistical significance was assessed by a two-sided t-test.

Abbreviations: *SR*: spatial relationship; *ICI*: Immune checkpoint inhibitors; *G-AUC-T*: G-function evaluated at a threshold T; *T*: threshold; logit: logistic regression.



Supplementary Figure 10. Comparison of spatial relationship parameters derived from pre-operative urothelial cancer (NABUCCO) and head and neck cancer (IMCISION) samples.

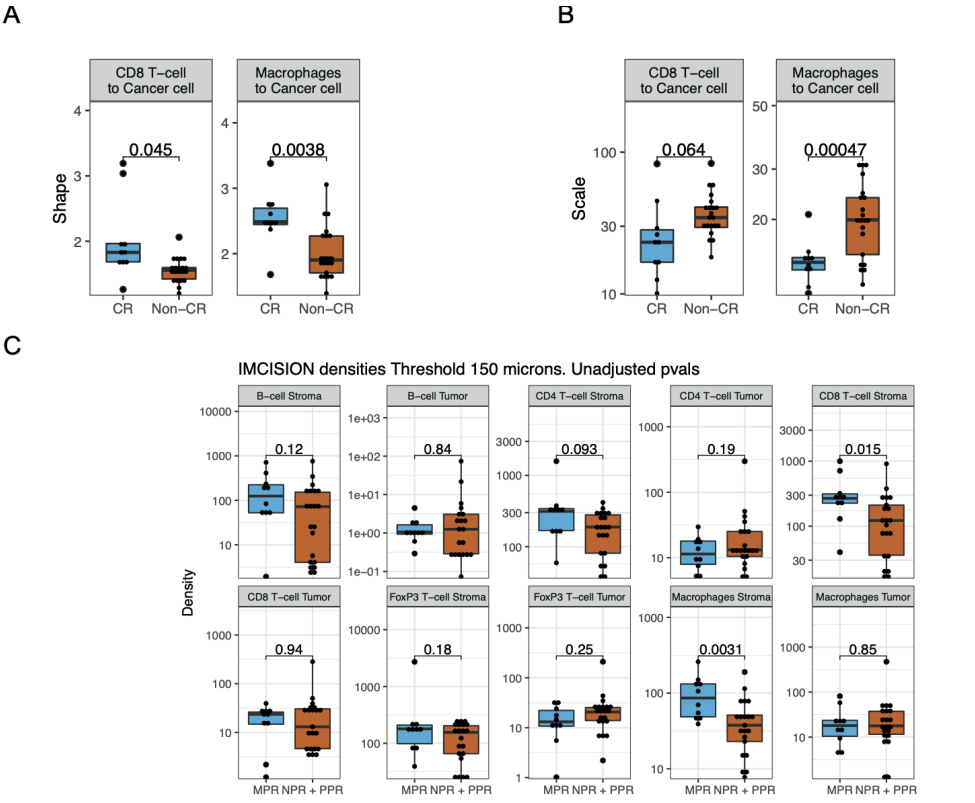
(A, B) Cohort averages for each SR's cell type pairwise relationships shape and scale parameters in UC (top, NABUCCO, A) and HNSCC (bottom, IMCISION, B). Each facet represents SRs to a specific target cell type. For instance, the first facet represents SRs studied from any reference cell type to cancer cells, and the color indicates the cell type from which the SR was studied (reference cell type).

(C) Correlation between shape parameters quantified in NABUCCO (x-axis) and IMCISION (y-axis). Pearson's coefficient and correlation p-value are shown in the plot.

(D) Correlation between scale parameters quantified in NABUCCO (x-axis) and IMCISION (y-axis). Pearson's coefficient and correlation p-value are shown in the plot.

All statistical tests were two-sided. No adjustments were made to correct for multiple comparisons.

Abbreviations: *SR:* spatial relationship; *UC:* urothelial cancer; *HNSCC:* head and neck squamous cell carcinoma.



Supplementary Figure 11. Validation of SR biomarkers of ICI response in head and neck cancer cohort (IMCISION trial).

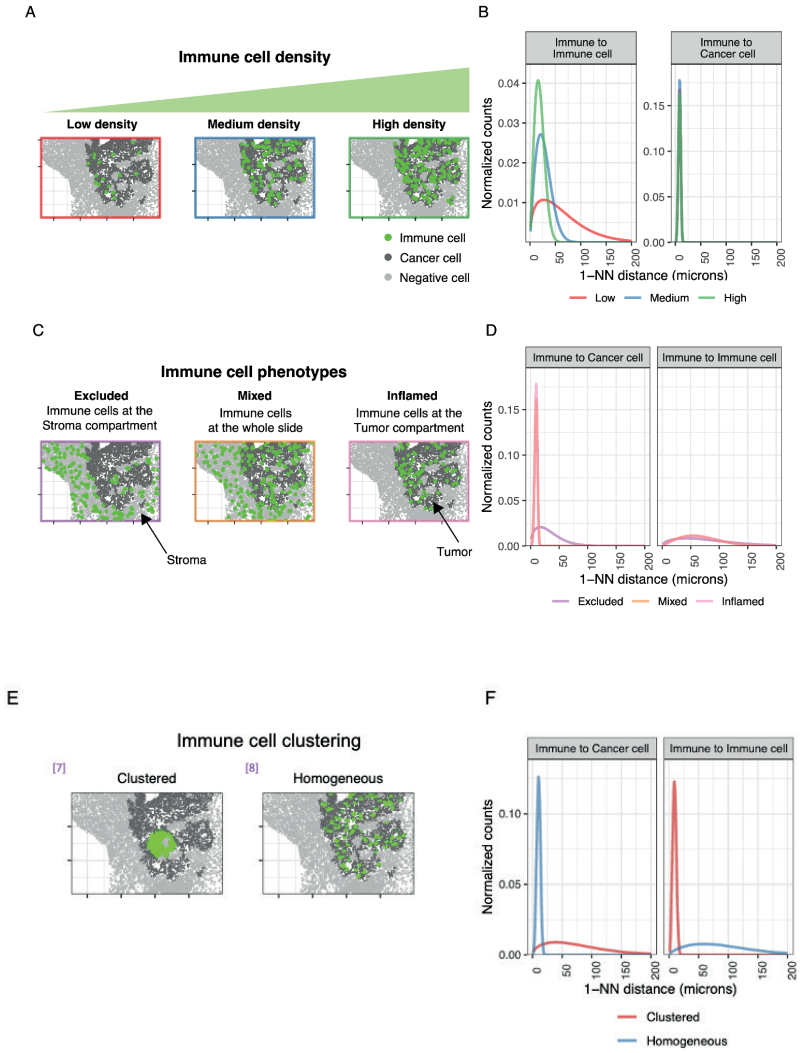
(A) Shape parameters for the SRs from $CD8^+$ T-cells to Cancer cells (left) and from Macrophages to cancer cells (right) between ICI response groups. Adjusted p-values: $FDR^{CD8\ T\text{-}cell\ to\ Cancer\ cell}$=0.045, $FDR^{Macrophages\ to\ Cancer\ cell}$=0.0076.

(B) Scale parameters for the SRs from $CD8^+$ T-cells to Cancer cells (left) and from Macrophages to cancer cells (right) between ICI response groups. A two-sided t-test was used for comparisons between response groups. Adjusted p-values: $FDR^{CD8\ T\text{-}cell\ to\ Cancer\ cell}$=0.064, $FDR^{Macrophages\ to\ Cancer\ cell}$=0.00094.

(C) Immune cell densities between response groups in IMCISION.

The box plots in each panel show the middle 50% of the data, with the box itself representing the median and the interquartile range (IQR) between the 25th and 75th percentiles. The whiskers extend from the box to the furthest data points within 1.5 times the IQR from the median. Unless otherwise stated, all statistical tests were two-sided. A two-sided t-test was used for comparisons between response groups.

Abbreviations: SR: spatial relationship; *ICI*: immune checkpoint inhibitors.



Supplementary Figure 12. Simulation study to quantify sources of variation of the spatial relationship parameters.

Representative data for a sample: (A) Immune cell density was simulated at different values.

(B) 1-NN curves vs immune cell density for three representative SR parameters.

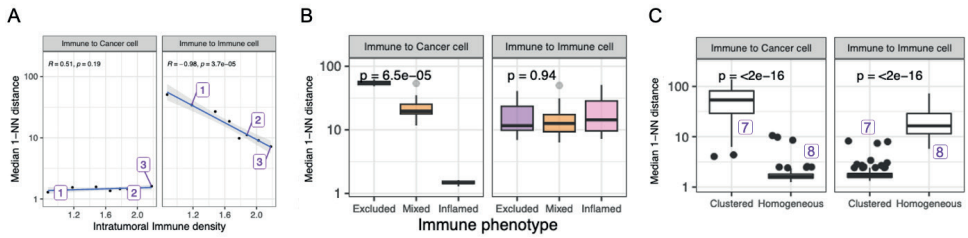
(C) Immune cells were simulated to be present only in the stroma region (*Excluded*, left), both at the stroma and the tumor region (*Mixed*, middle), and only in the tumor region (*Inflamed*, right).

(D) 1-NN curves vs immune cell phenotypes for three representative SR parameters.

(E) Immune cell clustering was simulated with values *Clustered* (left) and *Homogeneous* (right) spatial pattern.

(F) First-nearest neighbor distance curves for the clustered and homogeneous examples from panel C for the SRs studied from Immune to cancer cells and from Immune to immune cells.

Abbreviations: *SR*: spatial relationship; *ICI*: immune checkpoint inhibitors; *1-NN*: first nearest-neighbor.



Supplementary Figure 13. Associations between simulation study perturbations and spatial relationship parameters for a representative sample.

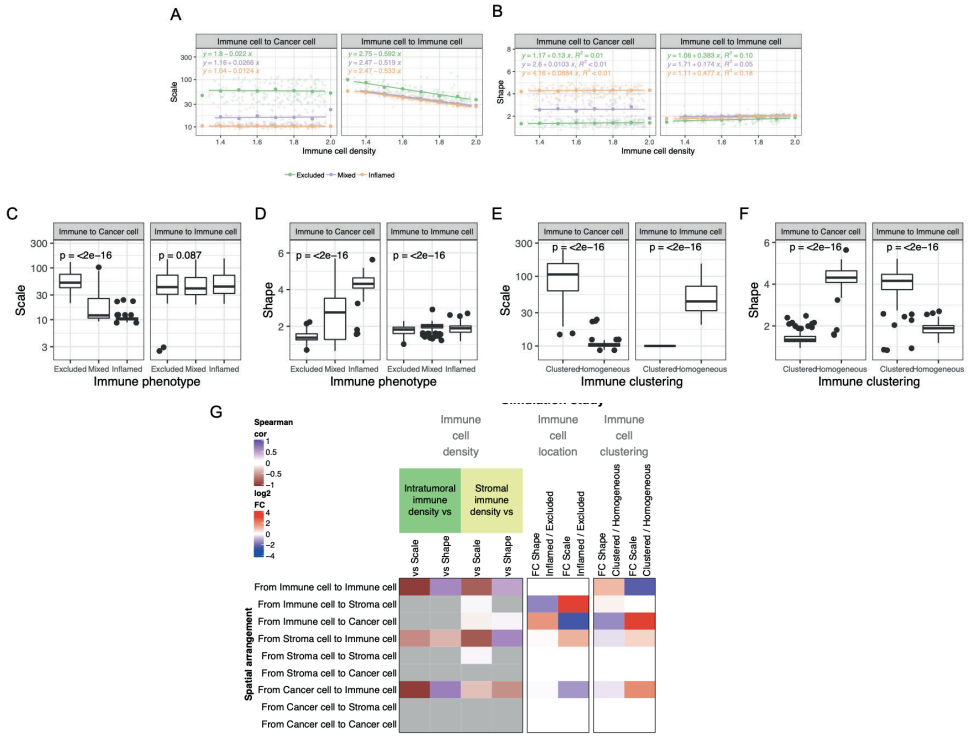
(A) Median 1-NN distances vs. simulated intratumoral immune cell density for the SRs from Immune to Cancer cells and from Immune to Immune cell. Two-sided Pearson's moment correlation test was used to test for the association. The Pearson's coefficient and correlation p-value are represented in the plot. 1, 2, and 3 annotations match with Low, Medium and High densities from Supplementary Figure 9A.

(B) Median 1-NN distances vs simulated immune phenotypes for the SRs from Immune to Cancer cells and from Immune to Immune cell. Differences between groups were tested by a Kruskan-Wallis test.

(C) Median 1-NN distances vs simulated immune cell clustering for the SRs from Immune to Cancer cells and from Immune to Immune cell. A two-sided t-test tested differences between groups.

The box plots in each panel show the middle 50% of the data, with the box itself representing the median and the interquartile range (IQR) between the 25th and 75th percentiles. The whiskers extend from the box to the furthest data points within 1.5 times the IQR from the median. All statistical tests were two-sided. Unless otherwise stated, no adjustments for multiple hypothesis testing were made.

Abbreviations: 1-NN: first-nearest neighbor; SR: spatial relationship.



Supplementary Figure 14. Associations between simulation study perturbations and spatial relationship parameters for the whole simulated cohort.

(A) Scatter plot of the scale parameter for two representative SRs for simulations of immune cell at different density values (x-axis) and at distinct tissue compartments, including in the Tumor (Inflamed, orange), Stroma (Excluded, green) and in the Tumor and Stroma (Mixed, purple). A linear model was fitted using the data simulated at each tissue compartment (e.g. Inflamed), and the slope of the fit and significance is highlighted in the figure. Dots represent all simulations' average scale parameters at a discretized density.

(B) Scatter plot of the shape parameter for two representatives SRs for simulations of immune cell at different density values (x-axis) and at distinct tissue compartments, including in the Tumor (Inflamed, orange), Stroma (Excluded, green) and in the Tumor and Stroma (Mixed, purple). A linear model was fitted using the data simulated at each tissue compartment (e.g. Inflamed), and the slope of the fit and significance is highlighted in the figure. Dots represent all simulations' average shape parameters at a discretized density.

(C) Scale parameter distribution for simulations of immune cells present at different tissue compartments (Inflamed: in Tumor; Excluded: in Stroma; Mixed: in Tumor and Stroma). Each boxplot (e.g. Tumor) contains data for different samples and simulations at distinct densities. Statistical association was assessed by an ANOVA test.

(D) Shape parameter distribution for simulations of immune cells present at different tissue compartments (Inflamed: in Tumor; Excluded: in Stroma; Mixed: in Tumor and Stroma). Each boxplot (e.g. Tumor) contains data for different samples and simulations at distinct densities. An ANOVA test assessed statistical association.

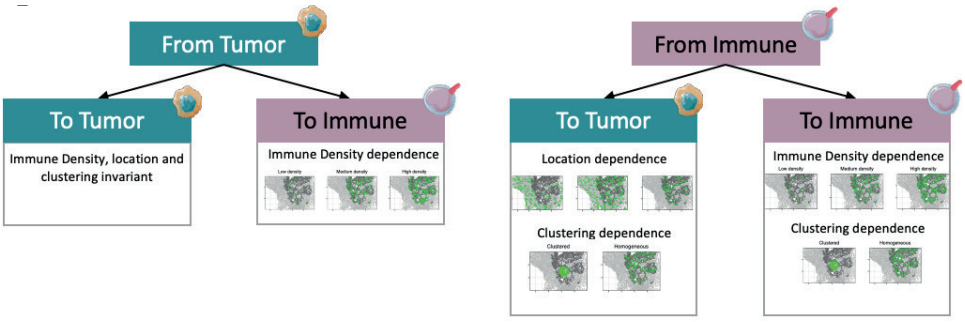
(E) Scale parameter distribution for simulations of immune cells following a homogeneous spatial distribution, or a clustered spatial distribution. Each boxplot (e.g. clustered) contains data for different samples and simulations at distinct densities. A t-test assessed statistical association.

(F) Shape parameter distribution for simulations of immune cells following a homogeneous spatial distribution, or a clustered spatial distribution. A t-test assessed statistical association.

(G) First 4 columns: Spearman correlation between the spatial parameters (shape or scale) and intratumoral (green columns) or stromal immune cell density (yellow) for each SR. Fifth and sixth columns: fold change on differences on the shape or the scale parameters between Inflamed and Excluded simulations (non-significant fold changes are coloured in white). Statistical significance was assessed by a t-test. Seventh and eighth columns: fold change on differences on the shape and scale parameters between immune cell clustered or homogeneous simulations by a t-test. Non-significant correlations or associations are labeled in gray and white.

The box plots in each panel show the middle 50% of the data, with the box itself representing the median and the interquartile range (IQR) between the 25th and 75th percentiles. The whiskers extend from the box to the furthest data points within 1.5 times the IQR from the median. All statistical tests were two-sided. Unless otherwise stated, no adjustments for multiple hypothesis testing were made.

Abbreviations: *UC*: urothelial cancer; *SR:* Spatial relationship.



Supplementary Figure 15. Interpretations on the associations between simulation study perturbations and spatial relationship parameters for the whole simulated cohort.

Summary of associations between simulated perturbations and effect on SR parameters by reference and target cell type by aggregating data for the whole cohort (Supplementary Figure 21H). Here, the interpretation of when *Cancer cells* are replaced by *Negative cells* are analogous (abundant cell types). Immune cells are referenced in text text as rare cell types. Icons from panel adapted from bioIcons (cancerous-cell-1, lymphocytes-4, t-lymphocyte licensed under CC-BY 3.0 Unported by Servier).

Abbreviations: *SR:* spatial relationship.

5.8 SUPPLEMENTARY TABLES AND SUPPLEMENTARY DATA

Due to printing limitations the Supplementary Tables, Supplementary Data and Source Data files are provided online: https://www.nature.com/articles/s41467-024-46450-1

DATA AVAILABILITY

The multiplex immunofluorescence data consisting on the spatial coordinates and immune cell types linked with the clinical data used for this manuscript are available under restricted access for data privacy regulations governing clinical and molecular data collected from human samples and will be made available upon reasonable academic and scientifically request and

within the limitations of the provided informed consent by the corresponding author upon reasonable request. The institutional review board of the Netherlands Cancer Institute will review every request. After approval, the researcher will need to sign the Netherlands Cancer Institute data access agreement.

Multiplex immunofluorescence derived data (spatial parameters and densities) is made available as Supplementary Data 1 and Source data files to reproduce the findings from the manuscript.

CODE AVAILABILITY

Code to reproduce the main findings will be made available in a Github repository (https://github.com/tropicalberto/nabucco_spatial_manuscript).

ACKNOWLEDGMENTS

We thank the facilities and staff at the Netherlands Cancer Institute (NKI) that contributed to this project, including the Core Facility Molecular Pathology and Biobanking (I. Hofland, S. Vonk, and W. Kievit), the Pathology department (I. M. Seignette, C. van Rooijen, M. Almekinders, L. Smit, and J. Sanders), the Urology department (K. Hendricksen, H. van der Poel) and the Research High-Performance Computing facility. We thank all patients who participated in the NABUCCO and IMCISION clinical trials and the study teams involved. This research at the NKI was supported by an institutional grant from the Dutch Cancer Society (KWF) and the Dutch Ministry of Health, Welfare, and Sport.

REFERENCES

1. Wei, S. C., Duffy, C. R. & Allison, J. P. Fundamental Mechanisms of Immune Checkpoint Blockade Therapy. *Cancer Discovery* vol. 8 1069–1086 Preprint at https://doi.org/10.1158/2159-8290.cd-18-0367 (2018).
2. Powles, T. *et al.* Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial. *Lancet* **391**, 748–757 (2018).
3. Bellmunt, J. *et al.* Pembrolizumab as Second-Line Therapy for Advanced Urothelial Carcinoma. *N. Engl. J. Med.* **376**, 1015–1026 (2017).
4. Sharma, P., Hu-Lieskovan, S., Wargo, J. A. & Ribas, A. Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. *Cell* **168**, 707–723 (2017).
5. Wu, Z. *et al.* Adverse Events of Immune Checkpoint Inhibitors Therapy for Urologic Cancer Patients in Clinical Trials: A Collaborative Systematic Review and Meta-analysis. *Eur. Urol.* **81**, 414–425 (2022).
6. Blank, C. U., Haanen, J. B., Ribas, A. & Schumacher, T. N. The 'cancer immunogram'. *Science* **352**, 658–660 (2016).
7. Powles, T. *et al.* Clinical efficacy and biomarker analysis of neoadjuvant atezolizumab in operable urothelial carcinoma in the ABACUS trial. *Nature Medicine* vol. 25 1706–1714 Preprint at https://doi.org/10.1038/s41591-019-0628-7 (2019).
8. Necchi, A. *et al.* Pembrolizumab as neoadjuvant therapy before radical cystectomy in patients with muscle-invasive urothelial bladder carcinoma (PURE-01): an open-label, single-arm, phase II study. *J. Clin. Oncol.* **36**, (2018).
9. Gao, J. *et al.* Neoadjuvant PD-L1 plus CTLA-4 blockade in patients with cisplatin-ineligible operable high-risk urothelial carcinoma. *Nat. Med.* **26**, 1845–1851 (2020).
10. van Dijk, N. *et al.* Preoperative ipilimumab plus nivolumab in locoregionally advanced urothelial cancer: the NABUCCO trial. *Nat. Med.* **26**, 1839–1844 (2020).
11. van Dijk, N. *et al.* The Cancer Immunogram as a Framework for Personalized Immunotherapy in Urothelial Cancer. *Eur. Urol.* **75**, 435–444 (2019).
12. Mariathasan, S. *et al.* TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* **554**, 544–548 (2018).
13. Havel, J. J., Chowell, D. & Chan, T. A. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. *Nat. Rev. Cancer* **19**, 133–150 (2019).
14. Hoyt, C. C. Multiplex Immunofluorescence and Multispectral Imaging: Forming the Basis of a Clinical Test Platform for Immuno-Oncology. *Front Mol Biosci* **8**, 674747 (2021).
15. van Dam, S., Baars, M. J. D. & Vercoulen, Y. Multiplex Tissue Imaging: Spatial Revelations in the Tumor Microenvironment. *Cancers* vol. 14 3170 Preprint at https://doi.org/10.3390/cancers14133170 (2022).
16. Lu, S. *et al.* Comparison of Biomarker Modalities for Predicting Response to PD-1/PD-L1 Checkpoint Blockade: A Systematic Review and Meta-analysis. *JAMA Oncol* **5**, 1195–1204 (2019).
17. Kather, J. N. *et al.* Topography of cancer-associated immune cells in human solid tumors. *Elife* **7**, (2018).
18. Feng, Z. *et al.* Multiparametric immune profiling in HPV- oral squamous cell cancer. *JCI Insight* **2**, (2017).
19. Palla, G., Fischer, D. S., Regev, A. & Theis, F. J. Spatial components of molecular tissue biology. *Nat. Biotechnol.* (2022) doi:10.1038/s41587-021-01182-1.
20. Parra, E. R. Methods to Determine and Analyze the Cellular Spatial Distribution Extracted From Multiplex Immunofluorescence Data to Understand the Tumor Microenvironment. *Front Mol Biosci* **8**, 668340 (2021).
21. Huang, Y.-K. *et al.* Macrophage spatial heterogeneity in gastric cancer defined by multiplex immunohistochemistry. *Nat. Commun.* **10**, 3928 (2019).

22. Zhang, W. *et al.* Identification of cell types in multiplexed in situ images by combining protein expression and spatial information using CELESTA. *Nat. Methods* **19**, 759–769 (2022).
23. Moldoveanu, D. *et al.* Spatially mapping the immune landscape of melanoma using imaging mass cytometry. *Sci Immunol* **7**, eabi5072 (2022).
24. Wang, X. *et al.* Spatial interplay patterns of cancer nuclei and tumor-infiltrating lymphocytes (TILs) predict clinical benefit for immune checkpoint inhibitors. *Sci Adv* **8**, eabn3966 (2022).
25. Ma, X. *et al.* Spatial Distribution and Predictive Significance of Dendritic Cells and Macrophages in Esophageal Cancer Treated With Combined Chemoradiotherapy and PD-1 Blockade. *Front. Immunol.* **12**, 786429 (2021).
26. Barua, S. *et al.* A Functional Spatial Analysis Platform for Discovery of Immunological Interactions Predictive of Low-Grade to High-Grade Transition of Pancreatic Intraductal Papillary Mucinous Neoplasms. *Cancer Informatics* vol. 17 11769351878288 Preprint at https://doi.org/10.1177/11769351878288O (2018).
27. Barua, S. *et al.* Spatial interaction of tumor cells and regulatory T cells correlates with survival in non-small cell lung cancer. *Lung Cancer* **117**, 73–79 (2018).
28. Vos, J. L. *et al.* Neoadjuvant immunotherapy with nivolumab and ipilimumab induces major pathological responses in patients with head and neck squamous cell carcinoma. *Nat. Commun.* **12**, 7348 (2021).
29. Baddeley, A., Rubak, E. & Turner, R. *Spatial Point Patterns: Methodology and Applications with R*. (CRC Press, 2015).
30. Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer-immune set point. *Nature* **541**, 321–330 (2017).
31. Powles, T. *et al.* Avelumab Maintenance Therapy for Advanced or Metastatic Urothelial Carcinoma. *N. Engl. J. Med.* **383**, 1218–1230 (2020).
32. Basile, G. *et al.* Neoadjuvant pembrolizumab and radical cystectomy in patients with muscle-invasive urothelial bladder cancer: 3-year median follow-up update of PURE-01 trial. *Clin. Cancer Res.* (2022) doi:10.1158/1078-0432.CCR-22-2158.
33. van Dorp, J. *et al.* High- or low-dose preoperative ipilimumab plus nivolumab in stage III urothelial cancer: the phase 1B NABUCCO trial. *Nat. Med.* (2023) doi:10.1038/s41591-022-02199-y.
34. van Dijk, N. *et al.* The Tumor Immune Landscape and Architecture of Tertiary Lymphoid Structures in Urothelial Cancer. *Frontiers in Immunology* vol. 12 Preprint at https://doi.org/10.3389/fimmu.2021.793964 (2021).
35. Oelkrug, C. & Ramage, J. M. Enhancement of T cell recruitment and infiltration into tumours. *Clin. Exp. Immunol.* **178**, 1–8 (2014).
36. Shapouri-Moghaddam, A. *et al.* Macrophage plasticity, polarization, and function in health and disease. *J. Cell. Physiol.* **233**, 6425–6440 (2018).
37. Noguchi, T. *et al.* Temporally Distinct PD-L1 Expression by Tumor and Host Cells Contributes to Immune EscapeThe Role of Tumor versus Host PD-L1 in Tumor Immune Escape. *Cancer immunology research* **5**, 106–117 (2017).
38. Peranzoni, E. *et al.* Macrophages impede CD8 T cells from reaching tumor cells and limit the efficacy of anti–PD-1 treatment. *Proceedings of the National Academy of Sciences* **115**, E4041–E4050 (2018).
39. Väyrynen, Zhang, Yuan & Väyrynen. Composition, Spatial Characteristics, and Prognostic Significance of Myeloid Cell Infiltration in Pancreatic CancerMyeloid Cells in Pancreatic Cancer. *Clinical Cancer Research* **27**, 1069–1081 (2021).
40. Bates, D. Fitting and evaluating mixed models using lme4.
41. van Wilpe, S. *et al.* Spatial and Temporal Heterogeneity of Tumor-Infiltrating Lymphocytes in Advanced Urothelial Cancer. *Front. Immunol.* **12**, 802877 (2021).

Discussion

6

Alberto Gil-Jimenez

For decades, urothelial cancer (UC) of the bladder has been treated with either surgery (radical cystectomy; RC) with or without neoadjuvant chemotherapy (NAC) prior to the procedure[1]. However, the advent of clinical trial data involving the usage of immune checkpoint inhibitors (ICIs) in the preoperative setting and the long-term low progression rates observed hold promise to change UC's current standard of care treatment landscape.

It is essential to understand UC tumors, their surrounding tumor microenvironment (TME), and their interplay with clinical response to standard and novel treatments to have a real chance of refining precision medicine strategies in the clinical management of UC and develop novel treatment strategies that overcome the current resistance mechanisms to anti-cancer drugs.

The work presented in this thesis advanced our understanding of UC, its surrounding TME before or upon treatment, and the associated clinical responses to neoadjuvant treatments involving either NAC or ICIs. To do so, the contributions to this thesis are as follows:

- In **Chapter 2**, we assessed several genomic biomarkers of response to NAC and validated them in a large independent cohort;
- In **Chapter 3,** we characterized the UC TME in untreated surgical specimens and treated (post-treatment samples) surgical specimens upon neoadjuvant ICIs;
- In **Chapter 4**, we explored current and novel biomarkers of response to neoadjuvant ICIs in UC tumors as well as treatment dynamics induced by ICIs at the UC TME;
- In **Chapter 5**, we proposed a novel framework to model spatial relationships (SRs) at the TME using multiplex immunofluorescence data of pre-treated tumors. We explored the SRs in the context of response to neoadjuvant ICIs to identify novel biomarkers of ICIs.

In this discussion, the general findings from this thesis will be synthesized and linked to the current knowledge in the scientific literature. Furthermore, the strengths and limitations of this research will be highlighted, the potential implications of the results for the clinical management of UC tumors will be discussed, and the potential future research efforts to improve how we understand and treat UC tumors will be emphasized.

6.1 GENOMIC BIOMARKERS OF NEOADJUVANT CHEMOTHERAPY

Genomic alterations can modulate sensitivity to anti-cancer treatments. Precision medicine has benefitted from this characteristic of tumors, in which molecular targeted therapies target tumor-specific druggable alterations[2] and have significantly improved cancer care in the past decade. In locally advanced and metastatic UC, tumors with somatic mutations in the *FGFR2/3*

gene can be treated with an *FGFR* inhibitor[3], resulting in an objective anti-tumor response in approximately 40% of pre-treated patients. However, in the context of non-metastatic UC, no targeted treatments are clinically approved[3].

Association studies showed that genomic alterations can sensitize tumors to non-targeted therapies, such as chemotherapy. In this case, the molecular alteration (e.g., a somatic mutation) enhances treatment response and can, therefore, be considered as a response biomarker. In **Chapter 2**, we set out to validate previously proposed genomic biomarkers of response to neoadjuvant chemotherapy (NAC) in a large independent cohort of 165 pre-treatment samples from tumors acquired prior to cisplatin-based NAC. We first set out to validate published associations between NAC response and somatic mutations [4–7] and also searched for novel biomarkers of NAC response based on sequencing data derived from 44 bladder cancer-related genes.

Among all the tested hypotheses, we could only confirm a positive association between *ERCC2* mutations and pathological downstaging to NAC. Although mutations in *ERCC2* conferred a prolonged 5-year overall survival (OS) and progression-free survival (PFS), the association was not statistically significant. Our results align with previous publications showing associations between *ERCC2* mutations and response and overall survival to NAC in different independent cohorts[8–10].

Enhancing our understanding of the mechanism of action of cisplatin-based chemotherapies will help generate more appropriate hypotheses and, therefore, refine biomarkers of response. Although we understand how cisplatin-based chemotherapies interfere with the damaged DNA leading to cancer cell death, there are still unknowns regarding the associated mechanisms, such as cisplatin uptake from the general circulation into tumor cells[11]. Nevertheless, a mechanism of action has been proposed for cisplatin-based chemotherapies in tumors with alterations in *ERCC2* to explain the enhanced drug sensitivity. More specifically, Li et al. functionally assessed *ERCC2* mutations in UC cell lines. They demonstrated that most *ERCC2* mutations in the helicase domain turn off nucleotide excision repair activity and fail to rescue genomic instability and UV sensitivity, which enhances treatment sensitivity[8]. These results support our conclusion that alterations in *ERCC2* make tumors more sensitive to cisplatin chemotherapies.

The wealth of data involving associations between *ERCC2* mutations and response to NAC leads to a natural question: are we ready to implement this finding in clinical practice? First, the mutation rate of *ERCC2* in UC tumors is generally low, ranging from 9-16%[12], depending on the patient population and database. Second, the expected response rate for the general UC patient population upon NAC is 40%[13]. Therefore, its utility as a response biomarker

may only be relevant to a small (5%-10%) subset of patients, and alternative mechanisms of response to NAC exist, underscoring the necessity for additional biomarkers to benefit the broader UC patient population comprehensively. Our results from **Chapter 2** align with such estimations, in which somatic *ERCC2* mutations conferred a low sensitivity (11%) but a high specificity (97%) for response prediction. Therefore, the biomarker effectively identifies a subset of responding patients but must be coupled with orthogonal biomarkers for effective patient stratification.

6.2 BIOMARKERS OF NEOADJUVANT IMMUNE CHECKPOINT INHIBITORS

The wealth of clinical trial data involving ICIs in the UC neoadjuvant setting has shown that the treatment is effective and leads to prolonged responses[14]. Still, the majority of patients do not respond to ICIs. Tumors with certain types of molecular characteristics, such as intratumoral $CD8^+$ T-cell infiltration, can be more likely to respond to ICIs[15–17]. It is, therefore, essential to identify the shared characteristics of responding tumors to ICIs to develop novel response biomarkers that could aid patient stratification. However, ICIs biomarkers exhibit varied behavior between cohorts[17], and there are still unknowns regarding the associated resistance mechanisms[18]. For instance, recently, it was discovered that the efficacy of anti-CTLA4 ICIs requires T-regulatory cell depletion[19]. This has led to a lack of consensus regarding biomarkers of response to neoadjuvant ICIs in UC for blockade of PD-1, PD-L1, and CTLA-4, or a combination of them across and within cancer types.

To further map the landscape of response biomarkers and treatment dynamics of ICIs in UC, in **Chapter 3,** we assessed the effect in the UC TME induced by combination ICIs, and in **Chapters 4** and **5**, we evaluated the behavior of known and novel biomarkers of response to ICIs using samples from the NABUCCO cohort, which tested a combination treatment of anti-PD1 plus anti-CTLA4.

6.2.1 Genomic biomarkers

The tumor mutational burden (TMB) has been proposed to be a surrogate for tumor immunogenicity and foreignness[20–22]. In **Chapter 4,** we showed that tumors responding to ICIs have a numerically higher TMB but not a significant difference when compared to non-responders. Recent literature has shown that the TMB fails to predict response to ICIs across all cancers. However, it does correlate with response in tumors where neoantigen load positively correlates with immune cell infiltration, such as in UC[21,22]. However such associations were derived for monotherapy ICIs (for anti-PD-L1 or anti-PD-1) in the context of metastatic UC, thus deviating from our non-metastatic cohort described in **Chapters**

3-5 that received combination therapy. Effective immune responses also require multiple antigen-presentation machinery and T-cell recognition mechanisms even when a tumor has a mutation. The TMB metric often ignores such mechanisms[23]. Instead, more reliable estimates of the tumor immunogenicity are obtained by estimating the neoantigen load and their clonality, performing a neo-epitope analysis, HLA typing, and T-cell receptor sequencing[17]. Combining such metrics allows us to estimate whether an antigen can be presented by tumor cells and recognized by T-cells[24–26].

Other genomic biomarkers for ICIs response exist[17,27], such as microsatellite instability (MSI)[28], which induces more neoantigens. However, this rarely occurs in UC tumors[29] and was not present in any tumors from the NABUCCO cohort. Cancer immune evasion via a genomic impairment of the antigen presentation machinery[30] (e.g., deleterious point mutations, loss of heterozygosity, loss of beta2-microglobulin) and loss of function at the *JAK/STAT* pathway can also explain non-responsiveness to ICIs[31]. Such alterations are rare in UC tumors and the diversity of the HLA locus is large[32]. Due to the small cohorts from this thesis, we could not study this. Recently, it was discovered that loss of the Y chromosome, which commonly occurs in bladder cancer, can enhance sensitivity to anti-PD1 ICI[33]. Whether other genomic biomarkers for ICIs exist is an ongoing research focus[17], and novel insights are anticipated soon.

6.2.2 Modulators of the immune response: expression of immune checkpoints

PD-L1 expression in tumors associated with response to ICIs[34]. However, in **Chapter 4,** we showed that PD-L1 scoring was not significantly associated with response to combination ICIs. Other clinical trials in metastatic UC have shown that PD-L1 positivity is associated with response to anti-PD-L1[35] and anti-PD-1[36]. However, comparative trials failed to show that PD-L1 positivity confers a prolonged overall survival in anti-PD-(L)1 treatment when compared to chemotherapy[37–39]. Moreover, in the neoadjuvant setting, anti-PD-L1 treatment in UC has shown no significant association with PD-L1 scoring[40]. Still, a recent randomized phase 3 trial of adjuvant anti-PD-1 vs placebo did show a clinical benefit for ICI treatment in the PD-L1 high population[41]. These contradictory findings suggest that despite PD-L1 scoring being the only EMA-approved biomarker to treat UC with anti-PD-1 (adjuvant nivolumab; pembrolizumab or atezolizumab in cisplatin-ineligible first-line metastatic UC)[42], it is unclear whether the PD-L1 expression biomarker could be translated to stratify patients in the neoadjuvant setting.

The role of other factors related to immune regulation and exhaustion in ICI response are also being investigated. Pre-clinical data derived from human samples have suggested that double-positive T-cells expressing CD8 and PD-1 ($CD8^+PD\text{-}1^+$) enhance ICI response in non-small cell lung cancer[43,44]. In **Chapter 3**, we characterized the abundance of $CD8^+PD\text{-}1^+$ cells upon ICIs and in untreated UC tumors to investigate their presence and dynamic

behavior upon neoadjuvant ICIs. Our findings suggest that $CD8^+PD\text{-}1^+$ cell presence does not improve prognosis in untreated tumors but that neoadjuvant ICIs induce $CD8^+PD\text{-}1^+$ cell expansion, with responding tumors having higher levels of $CD8^+PD\text{-}1^+$ cells after treatment. In **Chapter 3**, we also assessed the presence of other immune exhaustion checkpoints in the untreated UC TME and found that TIM-3, but not LAG-3, is highly expressed. These results warrant further investigation and a combined analysis of the exhaustion profile of the UC TME (involving PD-L1, PD-1, and TIM-3) could generate new hypotheses regarding response and mechanisms of action to ICIs.

6.2.3 T-cell immunity

In **Chapter 4,** we showed that, in contrast to UC clinical trials with neoadjuvant monotherapy ICIs, response to combination ICIs is independent of pre-existing $CD8^+$ T-cell immunity at the proteomic (e.g., intratumoral $CD8^+$ T-cell density) and transcriptomic (e.g., $CD8^+$ T-cell effector signature and interferon-gamma signatures) levels. Our data is in line with a neoadjuvant combination ICI trial in UC (*NCT02812420*[45]) that also showed a lack of association between response and baseline $CD8^+$ T-cell immunity, which suggests that other mechanisms of action beyond CD8 T-cell immunity exist for combination ICIs such as PD-1 plus CTLA-4 blockade.

T-cells exhibit diverse spatial distributions across the TME stromal and intratumoral compartments, presenting distinct immune phenotypes with prognostic and predictive value in response to ICIs[46]. Our data from **Chapter 3** suggested that untreated tumors with a desert phenotype are more prone to recurrence. However, in **Chapters 4** and **5,** we showed that response to ICIs was not associated with immune phenotypes. This is in contrast to what has been observed for monotherapy ICIs, in which response is associated with inflamed tumors[40]. Our data suggest that adding CTLA-4 to the anti-PD-1 ICI regimen, results in a more effective immune response in cold tumors.

Recent literature suggested that many tumor-infiltrating T-cells are bystanders, i.e., unrelated to anti-cancer immunity[47]. The bystander component is typically ignored in T-cell assessments within the TME and such T-cells are all assumed to be tumor-reactive. Furthermore, there is no standardized method to distinguish bystander from tumor-reactive T-cells in bulk measurements. Therefore, performing a thorough examination of tumor-reactive T-cell immunity is crucial. This differentiation will be beneficial when considered for the study of ICI biomarkers and therapy resistance.

6.2.4 Presence of immunosuppressive cytokines

Immune resistance mechanisms can impair response to ICIs. For instance, TGF-beta signaling represses the antitumor functions of immune cells at the TME and is associated

with non-response to ICIs in metastatic UC[27,35,48]. In **Chapter 4,** we corroborated that TGF-beta signaling negatively correlated with ICI response. Such negative correlation can be explained by a candidate mechanism that induces CD8 T-cell and natural killer (NK) cell exclusion[49]. This data emphasizes the importance of studying the TME stroma in the immune response context, focusing on cancer-associated fibroblasts (CAFs), which play an essential role due to their interactions with cancer and immune cells[50].

6.2.5 Tertiary lymphoid structures and B-cell immunity

Tertiary lymphoid structures (TLS) are ectopic lymphoid structures appearing at sites of chronic inflammation, such as tumors[51]. Recent data demonstrated that TLS presence confers anti-tumor activity and is beneficial for a response to ICIs[52–54]. Upon this discovery, a wealth of data associating the presence, abundance, and molecular architecture of TLS has been investigated in the context of ICIs.

There is no standard approach to assess TLS in tumor tissue to date. In **Chapter 4,** we assessed TLS using a customized multiplex immunofluorescence panel that allowed for TLS abundance and maturation state quantification. Alternatively, in **Chapter 3,** we assessed TLS by an immune cell multiplex panel not explicitly designed to profile TLS. However, that allowed quantification of TLS abundance but not the associated maturation states. Such an approach showed that TLS immune cell abundance is heterogeneous showing differences between TLSs from untreated vs ICI-treated tumors. Our data suggest that ICIs modulate TLS characteristics with a different magnitude between response groups, which could aid in understanding ICI resistance mechanisms.

Moreover, in **Chapter 3** we showed that TLS maturation states vary depending on their proximity to the central tumor. Earlier maturation states and differential immune cell abundance were found in TLSs close to the central tumor area compared to those close to the deeper/mucosal part of a tumor. Such an observation suggests an impact on TLS assessment as a candidate biomarker. While pre-treatment clinical trials are mostly based on TURs (lowly enriched in TLSs close to the central tumor), the nature of the tissue obtained from TURs does not allow for deeper tissue layer profiling as the surgical procedure is based on the scraping of the superficial layer of a tumor at the bladder lining. The superficial layer of a bladder tumor can be exposed to the presence of infections, urinary toxins, and inflammatory mediators, which can induce immune responses and the presence of bystander TLSs that can be detected in collected tumor tissue[55,56]. Nevertheless, because the unique characteristics of anti-tumor TLSs are yet unknown, current TLS assessment methodologies do not allow distinguish them from bystander TLSs.

B-cells are heterogeneous immune cells regarding their functional programs in the TME. They have both anti-tumorigenic and pro-tumorigenic roles[57] and are highly abundant within TLSs. In a whole-transcriptome unbiased differential expression from **Chapter 4**, we found that B-cell signaling negatively correlated with response. Our results contradict the findings from a comparable trial of combination ICIs, in which a positive correlation between B-cell signaling and response to ICIs was found[45]. Moreover, B-cell presence has been associated with both good and poor prognosis in UC[58,59]. Altogether, the role of B-cells in the context of ICIs is still in its infancy and understudied in the context of ICIs. The contradictory findings from the literature and our research warrant further research to decipher their interplay with ICIs[60].

6.2.6 Spatial relationships at the tumor microenvironment

One of the thesis objectives was to identify novel biomarkers of clinical response of neoadjuvant ICIs in UC beyond the currently existing biomarkers. Motivated by the lack of association between response and immune cell infiltration both at the transcriptomic level (e.g., inflammation or T-cell effector gene signatures) and the immune cell level (e.g., $CD8^+$ T-cell density) identified in **Chapter 4**, we aimed to explore novel ways to profile the TME and thereby gain biological insights. To do so, we explored spatial relationships (SRs) in the TME, which have already been proposed as a possible avenue to uncover novel biological insights into immune cell behavior[61–63]. Moreover, it also allowed us to further explore our preliminary conclusion that response to neoadjuvant ipilimumab and nivolumab combination is independent of $CD8^+$ T-cell immunity.

In **Chapter 5**, we proposed a statistical framework to quantify spatial relationships (SRs) in the TME. We re-assessed the pre-treatment multiplex immunofluorescence data from **Chapter 4** to determine alternative factors that determine response to ICIs. Our antibody panel allowed us to derive SRs between T-cell populations ($CD8^+$ T-cells, $FoxP3^+$ T-cells, T-helpers), B-cells, macrophages, cancer cells, and other cells not covered by the panel (negative cells). Our findings indicate that SRs hold a higher predictive power when compared to the standard way of analyzing multiplex immunofluorescence data using immune cell density. Moreover, response to ICIs is more likely upon proximity 1) from $CD8^+$ T-cells to cancer cells, and 2) from macrophages to cancer cells.

Our findings underscore the importance of quantifying SRs within the TME, potentially yielding novel biomarkers of response to ICIs with clinical utility. However, due to the spatial profiling technologies' novelty, there is no consensus regarding SR quantification within the TME[61]. Diverse methodologies exist to quantify SRs (e.g., distance-based, network-based) and statistically model the data (e.g., median derivation, probabilistic distribution fitting). To illustrate this point, in **Chapter 5,** we evaluated the importance of the modeling approach

for SR quantification. We showed that our proposed statistical approach was superior to using existing metrics (e.g., G-functions) when assessing a predictive signal. We concluded that the modeling approach is important and that guidelines and best practices still need to be established within the community to assess robust biomarkers[61].

Importantly, we confirmed the robustness of our SR biomarkers from **Chapter 5** by validating them in an independent cohort of head and neck cancer tumors, suggesting that our candidate biomarkers may be a context-independent feature determining response to combination ICIs. Integrating our existing spatial data and orthogonal data modalities will be required to understand further the mechanistic programs activated upon proximity between $CD8^+$ T-cells and cancer cells and between macrophages and cancer cells. The lack of availability of single-cell and spatial RNA sequencing or spatial proteomics (e.g., cyTOF) in our cohorts did not allow us to make these connections. Other data types that allow for the profiling of interacting cells, such as PICseq, or cell-cell communication profiles derived from single-cell RNA sequencing, pose an important base from which to further navigate the interactions activated by spatial proximity biomarkers that determine response to ICIs. Lastly, additional and larger datasets involving multiplex immunofluorescence in UC before neoadjuvant ICI - which do not exist - are still needed to validate our findings and refine our modeling strategy.

6.3 TOWARDS A ROBUST BIOMARKER

In the domain of predictive biomarkers for treatment response, the conventional approach involves correlating the molecular and clinical data from prospective or retrospective cohorts. Our approaches from **Chapters 2-5** have evaluated the current and novel biomarkers of response to neoadjuvant treatments in UC. Our results suggest that, to date, complexity exists in the UC response biomarker domain due to the heterogeneity that we observed between clinical trials or retrospective cohorts in terms of biomarkers for response. Here, we will discuss the potential limitations of our approaches, all in the context of what is required to develop a robust biomarker:

- **Assessment in large cohorts**: biomarker associations must be (re-)assessed in large cohorts. In UC, biomarkers of response to neoadjuvant ICIs are typically identified in early-phase clinical trials, implying that the cohort sizes are relatively small (~10-50 patients). Larger cohorts (>>100 samples) are still required to assess whether biomarker associations are robust.
- **Validation in biomarker-guided randomized clinical trials**: biomarker associations must be validated in randomized, prospective clinical trials, ideally following a biomarker-based approach. This approach is usually expensive and only carried out

for robust biomarkers. For instance, a clinical trial involving ICIs in melanoma stratified patients based on a baseline signature of interferon-gamma signaling[64,65], allowing them to prospectively stratify patients to monotherapy or combination therapy ICIs and confirm the predictive power of the signature.

- **Robustness, reproducibility, and independence**: biomarkers must be robust, reproducible, and observer-independent. Some molecular traits are derived from techniques prone to batch effects and sampling bias, affecting the heterogeneity of the signal. For instance, technical variation in RNA-sequencing data induces large batch effects[66], posing a challenge when models are evaluated in external cohorts. An alignment of technical pipelines, a consensus of best practices, and an improvement of batch effect correction algorithms are necessary.
- **Statistical rigor for evaluation of the predictive power**: a robust statistical approach should be followed when evaluating the predictive power of biomarkers and underscore its potential cohort biases. Double loop, nested cross-validation is usually required to avoid overfitting and to obtain an unbiased performance estimate[67].
- **Validation in external cohorts:** biomarker associations must be validated in external cohorts (test sets) to prevent overfitting and assess their generalizability to a broader patient population.
- **Cost-effectiveness and feasibility**: Biomarkers must be cost-effective and easy to measure[68]. For instance, some novel biomarkers are derived from single-cell experiments. The associated time and costs make it impractical to be implemented on a large scale - at least, in the current stage. In that case, an alternative approach should be derived from the primary biomarker discovery.

Following these statements/requirements, we will navigate the limitations of our cohorts assessed in our chapters.

6.3.1 Limitations from Chapter 2

In **Chapter 2**, we validated biomarkers in a large UC retrospective cohort. Despite the relatively large cohort size (n=165), the low frequency of events (only 12 [7%] *ERCC2* mutations observed in the cohort) did not allow a cross-validation approach. However, given that we used our cohort as a validation cohort independent from the current data published in the literature, our association suggests robustness.

Further efforts to validate the *ERCC2* mutation biomarker (ideally prospectively) to confirm its predictive power, are required. To our knowledge, only one prospective clinical trial has assessed the association between response and distinct genomic biomarkers (involving somatic mutations in *ERCC2*, *FANCC*, *RB1*, and *ATM*) as a secondary endpoint[69–71]. Although the data suggested ERCC2 mutations were more predominant in responding tumors, the

association was not significant. Nevertheless, the trial design involved a combination of NAC with ICIs (nivolumab) and excluded patients with baseline nodal disease (N+), differing from the treatment course from **Chapter 2** and the nodal disease rate, which accounted for 44% of patients, making such results not directly comparable with our findings. Other prospective trials have assessed alternative biomarkers and responses to chemotherapy. For instance, a trial assessed the predictive power of a gene expression signature (COXEN) and response to NAC, but it could only confirm the prognostic power of the signature[72]. Preliminary secondary use of the data reported a positive association between NAC response and a mutation in any of *ATM*, *RB1*, *FANCC*, or *ERCC2*, but the translational results have not yet been published[73]. Further prospective trials will be required to assess the true predictive power of *ERCC2* mutations or alternative molecular traits as candidate biomarkers.

Our approach to distinguish driver from passenger mutations and somatic mutations from germline variants in the cohort employed a customized bioinformatic pipeline, combining previous knowledge on germline variants reported on population databases and functional annotations from experimentally validated clinical databases and bioinformatic predictions on variant pathogenicity. An alignment of such approaches between research centers, which is currently lacking, will be beneficial to guarantee the reproducibility of the biomarker and its robustness.

Lastly, and as discussed in the discussion section from **Chapter 2**, to date, there is heterogeneity in response assessment upon neoadjuvant chemotherapy in UC, which hinders the interpretations when comparing results from different studies.

6.3.2 Limitations from Chapters 3-5

In **Chapters 4** and **5**, we used a modestly sized (n=24) prospective cohort to evaluate associations between molecular tumor traits and response to neoadjuvant ICIs. Data integration from external cohorts or additional patients from the trial is still needed to investigate our preliminary observations further and achieve more statistical power. Moreover, the small cohort size limited the possibility of employing a cross-validation approach. Therefore, our results must be validated in independent cohorts.

Our results from **Chapter 4** suggest a divergence of response biomarkers to ICIs when compared to similar cohorts. A possible explanation for such differences is that while our neoadjuvant cohort from **Chapters 3-5** consists of a combination treatment of ICIs in UC (PD-1 plus CTLA-4 blockade), most of the published clinical trials involve either neoadjuvant monotherapy ICI (e.g., only PD-1 blockade) or were performed in the metastatic UC setting. Moreover, the population characteristics from clinical trials differed, which could induce differences between trials. While our cohorts from **Chapters 3-5** consisted of stage

III UC patients, other clinical trials, such as ABACUS[40] and PURE-01[74], recruited patients with a lower tumor stage. Another potential explanation for such differences is the small cohort size from the current studies. Data from larger, later-phase clinical trials (e.g., phase II and beyond) will be beneficial to assess the robustness of the current candidate biomarkers.

Our data also suggested that tumor tissue sampling poses a limitation for biomarker assessment in the context of response prediction. In **Chapter 3**, we used surgical radical cystectomy (RC) material to quantify the UC TME. Typically, biopsy/transurethral resection (TUR) material is used to estimate biomarkers of response to neoadjuvant treatments (e.g., neoadjuvant ICIs), in which we assume that biopsy material is representative of the whole tumor. Our data from **Chapter 3** suggested that although biopsy material can be representative of surgical material in most cases, there is a high intratumoral immune cell heterogeneity in the TME. Moreover, we showed that tumor margins, in comparison with the tumor or stromal compartment, generally contain a higher immune cell density. TUR material gives an incomplete picture of the tumor bed and the surrounding TME and therefore does not allow for extensive tumor margin profiling. These results warrant caution when interpreting response biomarkers in the UC neoadjuvant setting. The literature suggests that TUR material provides a representative sample of RC material with a 60%-70% likelihood[75], which aligns with our interpretations. However, this study did not compare TLS assessment within the UC TME. This, coupled with our observation that differences in maturation states between superficial and deep TLSs exist, warrants further investigation.

Chapters 3 and **4** assessed the TME dynamics upon treatment with neoadjuvant ICIs by comparing pre- vs post-treatment samples. This comparison posed a limitation in responding tumors, which, upon a good response to ICIs, did not allow the collection of post-treatment tumor tissue and impeded assessing such associations. However, quantifying post-treatment stromal tissue still allowed us to uncover novel insights regarding response to ICIs, such as responding tumors having a high abundance of $CD8^+PD1^+$ T-cells.

Furthermore, because the TMB is a continuous metric, currently, there is no consensus on the ideal cutoff for patient stratification. Clinical trial data usually stratifies the cohort by median TMB, which makes the data between clinical trials not directly comparable. Moreover, different filtering methodologies can be applied to account for the non-synonymous mutations for TMB quantification, affecting the metric value. Calibration is required between assays and bioinformatic pipelines[76]. To date, there is no consensus, which contributes to the candidate biomarker's complexity[77].

Digital pathology allows for the quantification of tumor slide stainings and has reproducibly revolutionized tumor material profiling[78]. However, some algorithms still require refinement

to be reproducible for translational medicine. For instance, in **Chapter 3,** we assessed CD8 and PD-1 double-positive cells by integrating two independent stainings using an artificial intelligence algorithm. Our algorithm required human refinement since the employed AI approach is still in its infancy and was hence prone to errors, which limits the reproducibility and robustness of the assessments if the data is used in the future for biomarker discovery. However, to date, there is no alternative strategy to quantify double-positive cells due to the lack of suitable antibodies for CD8 and PD-1 that could allow for simultaneous staining of a slide in a multiplexed fashion.

On the other hand, PD-L1 expression shows a dynamic behavior across tumors, as PD-L1 is expressed both on neoplastic and tumor-infiltrating immune cells[34], and its expression profile shows spatial heterogeneity[79] and can change over time[80]. Moreover, multiple immunohistochemical assays exist to assess PD-L1 expression. Consequently, further standardization for PD-L1 expression assessment across research centers is required to enhance its utility as a candidate biomarker of ICI response[18,34]. To date, it is still being determined what the most suitable scoring approach is to assess PD-L1.

This thesis has not covered other data modalities that allow further profiling of tumors, the surrounding TME, and the systemic effects of ICIs. These methods include T-cell receptor and B-cell receptor sequencing[81], single-cell sequencing data, microbiome profiling, metabolic profiling, and techniques for quantifying tumor heterogeneity such as spatial proteomics or transcriptomics[82]. These orthogonal data modalities are expected to improve our understanding of ICI response and benefit the discovery of biomarkers.

6.4 FUTURE DIRECTIONS IN IMMUNO-ONCOLOGY FOR UROTHELIAL CANCER

Approximately half of muscle-invasive UC patients are not eligible for standard-of-care cisplatin regimens due to age and comorbidities[83]. The advancements in immuno-oncological treatments are promising because high response rates across cohorts are observed. Moreover, ICIs offer alternative treatment options for UC patients who are not eligible for standard-of-care chemotherapy treatments, and the associated treatment-related toxicity profiles, accounting for 6%-41% of the patients, are deemed acceptable[83].

The NABUCCO trial described in **Chapters 4** and **5** has been expanded with 30 additional patients to confirm treatment efficacy and optimize treatment dosage[84]. Importantly, it will allow us to validate our spatial predictors of response from **Chapter 5** as well as molecular associations of response from **Chapter 4**.

The associations identified in this thesis, in combination with existing literature, have allowed for the generation of novel hypotheses to improve treatments. One finding, the negative correlation between ICI response and TGF-beta signaling, has initiated research into combining ICIs with TGF-beta blockade, hypothesizing a synergistic effect in anti-cancer treatment[85]. However, clinical data for such combination treatment is currently lacking, and early-phase clinical trials involving UC are still in the recruiting phase[49]. Our thesis revealed an important correlation between the response to ICI treatment and the proximity of $CD8^+$ T-cells to cancer cells. This suggests that optimizing ICI treatments could be achieved by inducing proximity between these cells through therapeutic interventions. One potential approach could be to use bispecific antibodies that target cancer cells expressing immune checkpoints like PD-L1 and $CD8^+$ T-cell markers (such as CD8). Although bispecific antibodies have effectively targeted both TGF-beta and PD-L1[86] in the ICI context, combining drug regimens or increasing the number of targets can increase toxicity. Therefore, further (pre-)clinical data collection is necessary to optimize such treatments.

Lastly, while this thesis explored the molecular characteristics of tumors undergoing neoadjuvant chemotherapy or ICIs, it is worth noting that chemotherapy regimens, such as cisplatin, can modulate the TME[87]. For instance, chemotherapy can trigger antigen release, enhance immunogenicity, and promote T-cell infiltration[88]. Consequently, the combination of chemotherapy and ICI regimens is being investigated in clinical trials. While clinical data for such combination treatments in UC, especially in the neoadjuvant setting, remains either unavailable or incomplete[3], clinical trials of chemotherapy plus immunotherapy in the metastatic UC setting have reported inconclusive results regarding their synergistic effects[89–93].

In addition to the inconclusive clinical trial data, the research effort devoted to exploring biomarkers for response and resistance to treatments combining ICIs and chemotherapy regimens has been limited. Therefore, collecting additional data, and combining it with existing response biomarker data from a combination of chemotherapy and ICI regimens, could improve future disease management and patient stratification.

6.5 GENERAL CONCLUSIONS AND FUTURE PERSPECTIVES

In this thesis we have used genomic, transcriptomic, and proteomic profiling of UC tumors in the untreated and pre-operative settings collected patients. We explored the cancer-intrinsic factors to predict and understand better the surrounding TME.

Our study revealed that certain data types contain valuable information amenable to alternative modeling approaches. For instance, we explored spatial modeling on spatially-resolved proteomic data beyond the standard modeling strategies to gain insight into the complexity and heterogeneity of human cancer datasets. Such an approach allowed us to propose novel candidate biomarkers of response to pre-operative combination ICIs, underscoring the need to explore existing datasets beyond standard analytical methodologies.

The work to fully comprehend the mechanisms underlying neoadjuvant combination ICIs in UC remains ongoing. Despite the significant progress in improving our understanding of response to combination ICIs, the differences observed between our findings and those reported in the literature indicate that our understanding of treatment response and mechanism is yet limited. Larger cohorts will be necessary to quantify more robust associations, validate the current preliminary findings, and better understand patient heterogeneity regarding response and resistance to treatment. Incorporation of additional data types that profile the TME, especially with a single-cell and a spatial resolution, will be required to model further the differences observed between patient populations and ideally to exploit such heterogeneity for biomarker purposes. Moreover, the study of the interaction between cancer, adaptive immunity, and response to immune checkpoint inhibition must be done beyond T-cell-related immunity, as other immune system components can play a role in determining ICI treatment response, such as shown in this thesis with macrophages. Lastly, future studies integrating the current knowledge on response to standard-of-care treatments such as neoadjuvant chemotherapy with the novel knowledge on response to ICIs still need to be met to propose novel ways of patient stratification. Ultimately, these insights will contribute to the refinement and optimization of the treatment management of UC tumors.

Four key areas for exciting new developments in the near future are identified based on the work presented in this thesis: 1) clinical trials further exploring the efficacy of neoadjuvant (monotherapy or combination) ICIs in UC, 2) experimental and computational approaches to derive biomarkers of response, 3) a better understanding of treatment dynamics to aid overcoming treatment resistance in standard-of-care treatments and ICIs and 4) a change in the neoadjuvant treatment landscape of UC by the incorporation of ICIs.

REFERENCES

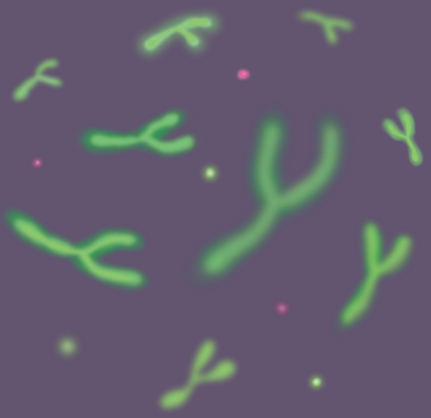
1. Grivas, P. D., Melas, M. & Papavassiliou, A. G. The biological complexity of urothelial carcinoma: Insights into carcinogenesis, targets and biomarkers of response to therapeutic approaches. *Semin. Cancer Biol.* **35**, 125–132 (2015).
2. Min, H.-Y. & Lee, H.-Y. Molecular targeted therapy for anticancer treatment. *Exp. Mol. Med.* **54**, 1670–1694 (2022).
3. Patil, G. & Basu, A. Emerging perioperative therapeutic approaches in muscle invasive bladder cancer. *Ther. Adv. Urol.* **14**, 17562872221134389 (2022).
4. Van Allen, E. M. *et al.* Somatic ERCC2 mutations correlate with cisplatin sensitivity in muscle-invasive urothelial carcinoma. *Cancer Discov.* **4**, 1140–1153 (2014).
5. Groenendijk, F. H. *et al.* ERBB2 Mutations Characterize a Subgroup of Muscle-invasive Bladder Cancers with Excellent Response to Neoadjuvant Chemotherapy. *Eur. Urol.* **69**, 384–388 (2016).
6. Teo, M. Y. *et al.* Fibroblast Growth Factor Receptor 3 Alteration Status is Associated with Differential Sensitivity to Platinum-based Chemotherapy in Locally Advanced and Metastatic Urothelial Carcinoma. *Eur. Urol.* **78**, 907–915 (2020).
7. Plimack, E. R. *et al.* Defects in DNA Repair Genes Predict Response to Neoadjuvant Cisplatin-based Chemotherapy in Muscle-invasive Bladder Cancer. *Eur. Urol.* **68**, 959–967 (2015).
8. Li, Q. *et al.* ERCC2 Helicase Domain Mutations Confer Nucleotide Excision Repair Deficiency and Drive Cisplatin Sensitivity in Muscle-Invasive Bladder Cancer. *Clin. Cancer Res.* **25**, 977–988 (2019).
9. Iyer, G. *et al.* Multicenter Prospective Phase II Trial of Neoadjuvant Dose-Dense Gemcitabine Plus Cisplatin in Patients With Muscle-Invasive Bladder Cancer. *J. Clin. Oncol.* **36**, 1949–1956 (2018).
10. Liu, D. *et al.* Clinical Validation of Chemotherapy Response Biomarker ERCC2 in Muscle-Invasive Urothelial Bladder Carcinoma. *JAMA Oncol* **2**, 1094–1096 (2016).
11. Makovec, T. Cisplatin and beyond: molecular mechanisms of action and drug resistance development in cancer chemotherapy. *Radiol. Oncol.* **53**, 148–158 (2019).
12. Choi, W. & McConkey, D. J. ERCC2 Mutation: The Marker for Chemosensitivity in Primary and Secondary Muscle-invasive Bladder Cancers. *European urology* vol. 75 240–241 (2019).
13. Zargar, H. *et al.* Multicenter assessment of neoadjuvant chemotherapy for muscle-invasive bladder cancer. *Eur. Urol.* **67**, 241–249 (2015).
14. Ghatalia, P., Kaur, J. & Sonpavde, G. Muscle invasive bladder cancer: where is the field headed? *Expert Opin. Biol. Ther.* 1–15 (2023).
15. Li, H., van der Merwe, P. A. & Sivakumar, S. Biomarkers of response to PD-1 pathway blockade. *Br. J. Cancer* **126**, 1663–1675 (2022).
16. Maleki Vareki, S., Garrigós, C. & Duran, I. Biomarkers of response to PD-1/PD-L1 inhibition. *Crit. Rev. Oncol. Hematol.* **116**, 116–124 (2017).
17. Litchfield, K. *et al.* Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. *Cell* **184**, 596–614.e14 (2021).
18. Barone, B. *et al.* Immune Checkpoint Inhibitors as a Neoadjuvant/Adjuvant Treatment of Muscle-Invasive Bladder Cancer: A Systematic Review. *Cancers* **14**, (2022).
19. Lax, B. M. *et al.* Both intratumoral regulatory T cell depletion and CTLA-4 antagonism are required for maximum efficacy of anti-CTLA-4 antibodies. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2300895120 (2023).
20. van Dijk, N. *et al.* The Cancer Immunogram as a Framework for Personalized Immunotherapy in Urothelial Cancer. *Eur. Urol.* **75**, 435–444 (2019).

21. McGrail, D. J. *et al.* High tumor mutation burden fails to predict immune checkpoint blockade response across all cancer types. *Ann. Oncol.* **32**, 661–672 (2021).
22. McGrail, D. J. *et al.* Validation of cancer-type-dependent benefit from immune checkpoint blockade in TMB-H tumors identified by the FoundationOne CDx assay. *Ann. Oncol.* **33**, 1204–1206 (2022).
23. Kallingal, A., Olszewski, M., Maciejewska, N., Brankiewicz, W. & Baginski, M. Cancer immune escape: the role of antigen presentation machinery. *J. Cancer Res. Clin. Oncol.* **149**, 8131–8141 (2023).
24. Snyder, A. *et al.* Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* **371**, 2189–2199 (2014).
25. Zou, X.-L. *et al.* Prognostic Value of Neoantigen Load in Immune Checkpoint Inhibitor Therapy for Cancer. *Front. Immunol.* **12**, 689076 (2021).
26. Gaud, G., Lesourne, R. & Love, P. E. Regulatory mechanisms in T cell receptor signalling. *Nat. Rev. Immunol.* **18**, 485–497 (2018).
27. Bagaev, A. *et al.* Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* **39**, 845–865.e7 (2021).
28. Sahin, I. H. *et al.* Immune checkpoint inhibitors for the treatment of MSI-H/MMR-D colorectal cancer and a perspective on resistance mechanisms. *Br. J. Cancer* **121**, 809–818 (2019).
29. Robertson, A. G. *et al.* Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. *Cell* **171**, 540–556.e25 (2017).
30. Sade-Feldman, M. *et al.* Resistance to checkpoint blockade therapy through inactivation of antigen presentation. *Nat. Commun.* **8**, 1136 (2017).
31. Dhatchinamoorthy, K., Colbert, J. D. & Rock, K. L. Cancer Immune Evasion Through Loss of MHC Class I Antigen Presentation. *Front. Immunol.* **12**, 636568 (2021).
32. Martínez-Jiménez, F. *et al.* Genetic immune escape landscape in primary and metastatic cancer. *Nat. Genet.* **55**, 820–831 (2023).
33. Abdel-Hafiz, H. A. *et al.* Y chromosome loss in cancer drives growth by evasion of adaptive immunity. *Nature* **619**, 624–631 (2023).
34. Powles, T., Walker, J., Andrew Williams, J. & Bellmunt, J. The evolving role of PD-L1 testing in patients with metastatic urothelial carcinoma. *Cancer Treat. Rev.* **82**, 101925 (2020).
35. Mariathasan, S. *et al.* TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* **554**, 544–548 (2018).
36. Balar, A. V. *et al.* First-line pembrolizumab in cisplatin-ineligible patients with locally advanced and unresectable or metastatic urothelial cancer (KEYNOTE-052): a multicentre, single-arm, phase 2 study. *Lancet Oncol.* **18**, 1483–1492 (2017).
37. Powles, T. *et al.* Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial. *Lancet* **391**, 748–757 (2018).
38. Powles, T. *et al.* Durvalumab alone and durvalumab plus tremelimumab versus chemotherapy in previously untreated patients with unresectable, locally advanced or metastatic urothelial carcinoma (DANUBE): a randomised, open-label, multicentre, phase 3 trial. *Lancet Oncol.* **21**, 1574–1588 (2020).
39. Fradet, Y. *et al.* Randomized phase III KEYNOTE-045 trial of pembrolizumab versus paclitaxel, docetaxel, or vinflunine in recurrent advanced urothelial cancer: results of >2 years of follow-up. *Ann. Oncol.* **30**, 970–976 (2019).
40. Powles, T. *et al.* Clinical efficacy and biomarker analysis of neoadjuvant atezolizumab in operable urothelial carcinoma in the ABACUS trial. *Nature Medicine* vol. 25 1706–1714 Preprint at https://doi.org/10.1038/s41591-019-0628-7 (2019).

41. Bajorin, D. F. *et al.* Adjuvant Nivolumab versus Placebo in Muscle-Invasive Urothelial Carcinoma. *N. Engl. J. Med.* **384**, 2102–2114 (2021).
42. ESMO. EMA recommends extension of indications for nivolumab. https://www.esmo.org/oncology-news/ema-recommends-extension-of-indications-for-nivolumab2.
43. Thommen, D. S. *et al.* A transcriptionally and functionally distinct PD-1+ CD8+ T cell pool with predictive potential in non-small-cell lung cancer treated with PD-1 blockade. *Nat. Med.* **24**, 994–1004 (2018).
44. Caushi, J. X. *et al.* Transcriptional programs of neoantigen-specific TIL in anti-PD-1-treated lung cancers. *Nature* **596**, 126–132 (2021).
45. Gao, J. *et al.* Neoadjuvant PD-L1 plus CTLA-4 blockade in patients with cisplatin-ineligible operable high-risk urothelial carcinoma. *Nat. Med.* **26**, 1845–1851 (2020).
46. Kather, J. N. *et al.* Topography of cancer-associated immune cells in human solid tumors. *Elife* **7**, (2018).
47. Meier, S. L., Satpathy, A. T. & Wells, D. K. Bystander T cells in cancer immunology and therapy. *Nat Cancer* **3**, 143–155 (2022).
48. Jenkins, R. W., Barbie, D. A. & Flaherty, K. T. Mechanisms of resistance to immune checkpoint inhibitors. *Br. J. Cancer* **118**, 9–16 (2018).
49. Benjamin, D. J. & Lyou, Y. Advances in Immunotherapy and the TGF-β Resistance Pathway in Metastatic Bladder Cancer. *Cancers* **13**, (2021).
50. Pei, L. *et al.* Roles of cancer-associated fibroblasts (CAFs) in anti- PD-1/PD-L1 immunotherapy for solid cancers. *Mol. Cancer* **22**, 29 (2023).
51. Schumacher, T. N. & Thommen, D. S. Tertiary lymphoid structures in cancer. *Science* **375**, eabf9419 (2022).
52. Sautès-Fridman, C., Petitprez, F., Calderaro, J. & Fridman, W. H. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat. Rev. Cancer* **19**, 307–325 (2019).
53. Cabrita, R. *et al.* Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature* **577**, 561–565 (2020).
54. Petitprez, F. *et al.* B cells are associated with survival and immunotherapy response in sarcoma. *Nature* **577**, 556–560 (2020).
55. Tang, H., Zhu, M., Qiao, J. & Fu, Y.-X. Lymphotoxin signalling in tertiary lymphoid structures and immunotherapy. *Cell. Mol. Immunol.* **14**, 809–818 (2017).
56. Domblides, C. *et al.* Tumor-Associated Tertiary Lymphoid Structures: From Basic and Clinical Knowledge to Therapeutic Manipulation. *Front. Immunol.* **12**, 698604 (2021).
57. Laumont, C. M. & Nelson, B. H. B cells in the tumor microenvironment: Multi-faceted organizers, regulators, and effectors of anti-tumor immunity. *Cancer Cell* **41**, 466–489 (2023).
58. Zirakzadeh, A. A. *et al.* Tumour-associated B cells in urothelial urinary bladder cancer. *Scand. J. Immunol.* **91**, e12830 (2020).
59. Fridman, W. H. *et al.* B cells and tertiary lymphoid structures as determinants of tumour immune contexture and clinical outcome. *Nat. Rev. Clin. Oncol.* **19**, 441–457 (2022).
60. Helmink, B. A. *et al.* B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* **577**, 549–555 (2020).
61. Palla, G., Fischer, D. S., Regev, A. & Theis, F. J. Spatial components of molecular tissue biology. *Nat. Biotechnol.* (2022) doi:10.1038/s41587-021-01182-1.
62. Bressan, D., Battistoni, G. & Hannon, G. J. The dawn of spatial omics. *Science* **381**, eabq4964 (2023).
63. Wang, X. Q. *et al.* Spatial predictors of immunotherapy response in triple-negative breast cancer. *Nature* **621**, 868–876 (2023).

64. Reijers, I. L. M. *et al.* Personalized combination of neoadjuvant domatinostat, nivolumab and ipilimumab in macroscopic stage III melanoma patients stratified according to the interferon-gamma signature: The DONIMI study. *J. Clin. Orthod.* **38**, TPS10087–TPS10087 (2020).
65. Reijers, I. L. M. *et al.* IFN-γ signature enables selection of neoadjuvant treatment in patients with stage III melanoma. *J. Exp. Med.* **220**, (2023).
66. Goh, W. W. B., Yong, C. H. & Wong, L. Are batch effects still relevant in the age of big data? *Trends Biotechnol.* **40**, 1029–1040 (2022).
67. Diaz-Uriarte, R. *et al.* Ten quick tips for biomarker discovery and validation analyses using machine learning. *PLoS Comput. Biol.* **18**, e1010357 (2022).
68. Burdett, N. & Desai, J. New biomarkers for checkpoint inhibitor therapy. *ESMO Open* **5**, e000597 (2020).
69. Galsky, M. D. *et al.* Co-primary endpoint analysis of HCRN GU 16-257: Phase 2 trial of gemcitabine, cisplatin, plus nivolumab with selective bladder sparing in patients with muscle-invasive bladder cancer (MIBC). *J. Clin. Orthod.* **41**, 447–447 (2023).
70. Galsky, M. D. *et al.* Phase 2 trial of gemcitabine, cisplatin, plus nivolumab with selective bladder sparing in patients with muscle- invasive bladder cancer (MIBC): HCRN GU 16-257. *J. Clin. Orthod.* **39**, 4503–4503 (2021).
71. Galsky, M. D. *et al.* Gemcitabine and cisplatin plus nivolumab as organ-sparing treatment for muscle-invasive bladder cancer: a phase 2 trial. *Nat. Med.* **29**, 2825–2834 (2023).
72. Flaig, T. W. *et al.* Long-term Outcomes from a Phase 2 Study of Neoadjuvant Chemotherapy for Muscle-invasive Bladder Cancer (SWOG S1314; NCT02177695). *Eur. Urol.* (2023) doi:10.1016/j.eururo.2023.06.014.
73. Plimack, E. R. *et al.* S1314 correlative analysis of ATM, RB1, ERCC2, and FANCC mutations and pathologic complete response (pT0) at cystectomy after neoadjuvant chemotherapy (NAC) in patients with muscle invasive bladder cancer (MIBC): Implications for bladder preservation. *J. Clin. Orthod.* **40**, 4581–4581 (2022).
74. Necchi, A. *et al.* Updated Results of PURE-01 with Preliminary Activity of Neoadjuvant Pembrolizumab in Patients with Muscle-invasive Bladder Carcinoma with Variant Histologies. *Eur. Urol.* **77**, 439–446 (2020).
75. van Wilpe, S. *et al.* Spatial and Temporal Heterogeneity of Tumor-Infiltrating Lymphocytes in Advanced Urothelial Cancer. *Front. Immunol.* **12**, 802877 (2021).
76. Vega, D. M. *et al.* Aligning tumor mutational burden (TMB) quantification across diagnostic platforms: phase II of the Friends of Cancer Research TMB Harmonization Project. *Ann. Oncol.* **32**, 1626–1636 (2021).
77. Anagnostou, V., Bardelli, A., Chan, T. A. & Turajlic, S. The status of tumor mutational burden and immunotherapy. *Nat Cancer* **3**, 652–656 (2022).
78. Baxi, V., Edwards, R., Montalto, M. & Saha, S. Digital pathology and artificial intelligence in translational medicine and clinical practice. *Mod. Pathol.* **35**, 23–32 (2022).
79. Ben Dori, S., Aizic, A., Sabo, E. & Hershkovitz, D. Spatial heterogeneity of PD-L1 expression and the risk for misclassification of PD-L1 immunohistochemistry in non-small cell lung cancer. *Lung Cancer* **147**, 91–98 (2020).
80. Grossman, J. E., Vasudevan, D., Joyce, C. E. & Hildago, M. Is PD-L1 a consistent biomarker for anti-PD-1 therapy? The model of balstilimab in a virally-driven tumor. *Oncogene* **40**, 1393–1395 (2021).
81. Schina, A. *et al.* Intratumoral T-cell and B-cell receptor architecture associates with distinct immune tumor microenvironment features and clinical outcomes of anti-PD-1/L1 immunotherapy. *J Immunother Cancer* **11**, (2023).
82. Kashyap, A. *et al.* Quantification of tumor heterogeneity: from data acquisition to metric generation. *Trends Biotechnol.* **40**, 647–676 (2022).
83. Nair, S. S., Chakravarty, D., Patel, V., Bhardwaj, N. & Tewari, A. K. Genitourinary cancer neoadjuvant therapies: current and future approaches. *Trends Cancer Res.* (2023) doi:10.1016/j.trecan.2023.07.011.
84. van Dorp, J. *et al.* High- or low-dose preoperative ipilimumab plus nivolumab in stage III urothelial cancer: the phase 1B NABUCCO trial. *Nat. Med.* **29**, 588–592 (2023).

85. Castiglioni, A. *et al.* Combined PD-L1/TGFβ blockade allows expansion and differentiation of stem cell-like CD8 T cells in immune excluded tumors. *Nat. Commun.* **14**, 4703 (2023).
86. Li, T. *et al.* Bispecific antibody targeting TGF-β and PD-L1 for synergistic cancer immunotherapy. *Front. Immunol.* **14**, 1196970 (2023).
87. Vennin, C. *et al.* Taxanes trigger cancer cell killing in vivo by inducing non-canonical T cell cytotoxicity. *Cancer Cell* **41**, 1170–1185.e12 (2023).
88. Li, J.-Y., Chen, Y.-P., Li, Y.-Q., Liu, N. & Ma, J. Chemotherapeutic and targeted agents can modulate the tumor microenvironment and increase the efficacy of immune checkpoint blockades. *Mol. Cancer* **20**, 27 (2021).
89. Galsky, M. D. *et al.* Atezolizumab with or without chemotherapy in metastatic urothelial cancer (IMvigor130): a multicentre, randomised, placebo-controlled phase 3 trial. *Lancet* **395**, 1547–1557 (2020).
90. Cathomas, R. *et al.* Perioperative Chemoimmunotherapy With Durvalumab for Muscle-Invasive Urothelial Carcinoma: Primary Analysis of the Single-Arm Phase II Trial SAKK 06/17. *J. Clin. Oncol.* JCO2300363 (2023).
91. van der Heijden, M. S. *et al.* Nivolumab plus Gemcitabine–Cisplatin in Advanced Urothelial Carcinoma. *N. Engl. J. Med.* (2023) doi:10.1056/NEJMoa2309863.
92. Powles, T. *et al.* Pembrolizumab alone or combined with chemotherapy versus chemotherapy as first-line therapy for advanced urothelial carcinoma (KEYNOTE-361): a randomised, open-label, phase 3 trial. *Lancet Oncol.* **22**, 931–945 (2021).
93. van der Heijden, M. S. *et al.* LBA7 Nivolumab plus gemcitabine-cisplatin versus gemcitabine-cisplatin alone for previously untreated unresectable or metastatic urothelial carcinoma: Results from the phase III CheckMate 901 trial. *Ann. Oncol.* **34**, S1341 (2023).



Summary

Urothelial cancer (UC) is a malignancy originating in the bladder with a challenging prognosis, even in its early disease stages. Currently, the standard of care to cure the disease involves surgical removal of the tumor through radical cystectomy (RC), which is often complemented by pre-operative chemotherapy (NAC, neoadjuvant chemotherapy). In the past two decades, the disease treatment landscape has remained unchanged.

Immune checkpoint inhibitors (ICIs) are a novel treatment type that blocks specific proteins, known as immune checkpoints, that downregulate an immune response. Upon immune checkpoint blockade, immune cells can effectively recognize cancer cells and induce an immune response against cancer cells. In the UC context, early-phase clinical trials have demonstrated both the feasibility and efficacy of pre-operative ICIs and have shown long-lasting clinical responses. Therefore, ICIs hold promise for changing UC clinical management in the future. However, not all patients respond to ICIs, and substantial rates of treatment-related toxicities are observed, highlighting the need for identifying biomarkers that can aid patient stratification.

The studies presented in this thesis focus on elucidating the role of the UC tumor microenvironment (TME) in determining responses to pre-operative treatments, such as chemotherapy and ICIs. Through a comprehensive multi-omics approach, we quantified intrinsic and extrinsic characteristics of UC tumors collected from human samples. We associated them with clinical outcomes such as treatment response and treatment dynamics.

In **Chapter 1**, we introduce the current status of the treatment of UC and the immune system's role in cancer. We then highlight the (limited) current knowledge landscape of molecular biomarkers of response to pre-operative treatments in UC, such as NAC and ICIs. Lastly, we highlight the contribution and objective of this thesis.

In **Chapter 2**, we performed targeted DNA sequencing to validate candidate biomarkers of response to pre-operative chemotherapy in a large multi-center retrospective cohort of 165 patients. Among the three hypotheses tested, we could only confirm a positive association between pathological response and deleterious alterations in *ERCC2*. While this correlation held for progression-free survival, it did not translate to overall survival, underscoring the necessity for further evaluation in prospective studies.

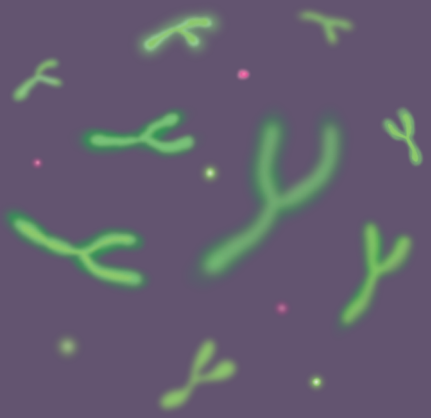
In **Chapter 3**, we employed multiplex immunofluorescence and protein stainings on UC RC specimens from untreated (n=31) and tumors treated with pre-operative combination ICIs (n=24) to characterize the UC TME. Our multiplexed immune panel enabled us to characterize diverse immune cell populations (e.g., T-cells, B-cells, macrophages) and tertiary lymphoid structures (TLS), the role of which in cancer is still being elucidated. We observed that upon ICIs the invasive margin of tumors showed a higher abundance of $CD8^{+}PD1^{+}$

T-cells. Moreover, the immune cell proportions within TLSs differed between treatment-naive and ICI-treated tumors, and their maturation states varied depending on their proximity to the tumor bed, which suggests that not all TLSs present in tumor tissue may be directly related to anti-tumor activity. Our data provided a foundation for further research in biomarker discovery within the context of pre-operative ICI treatments for UC.

In **Chapter 4**, we first presented the clinical data from NABUCCO, a phase-I clinical trial testing in 24 patients the feasibility of resecting a tumor upon a combination of pre-operative combination ICIs blocking the immune checkpoints CTLA-4 and PD-1 (ipilimumab and nivolumab, respectively). The trial achieved the primary endpoint, and durable responses were observed in more than fifty percent of patients. We then conducted a translational analysis to quantify tumor baseline response characteristics to pre-operative combination ICIs and treatment-associated dynamics. Making use of pre- and post-treatment samples, we profiled whole-exome DNA sequencing, bulk RNA sequencing, and protein abundance (multiplex immunofluorescence and stainings), allowing us to conclude that response was independent of $CD8^+$ T-cell related immunity and that in responding patients, the treatment induced TLS.

In **Chapter 5**, we quantified spatial relationships (SRs) in the TME of UC tumors and associated them with response to combination ICIs. First, we proposed a methodology to quantify SRs based on a curve fitting to the first nearest-neighbor spatial statistic. Such an approach allowed us to model spatially-resolved data, such as multiplex immunofluorescence quantitatively. We found that SRs are significantly associated with response to pre-operative combination ICIs and that SRs outperform immune cell abundance metrics - the standard way of quantifying multiplex immunofluorescence data - when predicting treatment response. Responding tumors were characterized by the proximity of $CD8^+$ T-cells to cancer cells and macrophages to cancer cells. Importantly, we validated our main findings in an independent cohort of head and neck cancer tumors (n=25) with a comparable combination ICI treatment, suggesting that our candidate biomarkers could be a context-independent feature determining response to combination ICI treatment.

The work from this thesis contributed to validate genomic biomarkers of response to pre-operative chemotherapy in UC, suggesting combination CTLA-4 plus PD-1 blockade as an effective treatment to cure UC, to advance our understanding of response to ICIs in UC, and to propose a novel way to characterize the TME employing spatial relationships. Altogether our findings expanded our current knowledge of UC tumor biology that could serve to improve patient care in the future.



Samenvatting

Urotheelcelcarcinoom (UC) is een maligniteit die ontstaat in de blaas met een slechte prognose, zelfs in de vroege stadia van de ziekte. De standaardbehandeling is een chirurgische verwijdering van de tumor door middel van een operatie bekend als radicale cystectomie (RC), vaak gevolgd door pre-operatieve chemotherapie (NAC, neoadjuvante chemotherapie). In de afgelopen twee decennia is het behandelingslandschap van de ziekte onveranderd gebleven.

Checkpointremmers zijn een vorm van immuuntherapie die specifieke eiwitten, bekend als immuuncheckpoints, die een immuunrespons downreguleren. Het blokkeren van deze checkpoints haalt de rem op de immuunresponse tegen kanker cellen weg. In de context van UC hebben vroege klinische trials zowel de haalbaarheid als de werkzaamheid van pre-operatieve checkpointremmers aangetoond. Daarbij werden er bij patiënten langdurige responsen op behandeling gezien. Om die redenen zijn checkpointremmers een veelbelovende behandelmethode die de behandeling van UC in de toekomst kan veranderen. Niettemin reageren niet alle patiënten op deze therapie en worden aanzienlijke percentages therapie gerelateerde toxiciteit waargenomen. Dit benadrukt dat er behoefte is aan het identificeren van biomarkers die kunnen helpen bij de stratificatie van patiënten die baat hebben bij therapie met checkpointinhibitors.

Het onderzoek dat in dit proefschrift wordt gepresenteerd, richt zich op het verduidelijken van de rol van de tumor micro-omgeving (TMO) van UC bij het bepalen van reacties op pre-operatieve behandelingen, zoals chemotherapie en checkpointremmers. Via een uitgebreide multi-omics-benadering hebben we intrinsieke en extrinsieke biomarkers verzameld uit UC humane weefsel en deze gekwantificeerd. We hebben deze gekoppeld aan klinische uitkomsten zoals behandelingsrespons en behandeling dynamiek.

In **Hoofdstuk 1** introduceren we zowel de huidige status van de behandeling van UC als de rol van het immuunsysteem bij kanker. We belichten vervolgens de (beperkte) huidige wetenschappelijke literatuur van moleculaire biomarkers voor respons op pre-operatieve behandelingen bij UC, zoals NAC en ICIs. Tot slot benadrukken we de bijdrage en het doel van dit proefschrift.

In **Hoofdstuk 2** hebben we targeted DNA-sequencing uitgevoerd om biomarkers voor respons op pre-operatieve chemotherapie te valideren in een grote retrospectieve multicentercohort van 165 patiënten. Van de drie geteste hypothesen konden we alleen een positieve associatie bevestigen tussen pathologische respons en *ERCC2* mutaties. Hoewel deze correlatie gold voor progressievrije overleving, vertaalde deze zich niet naar algehele overleving. Daarvoor is verdere evaluatie in prospectieve studies noodzakelijk.

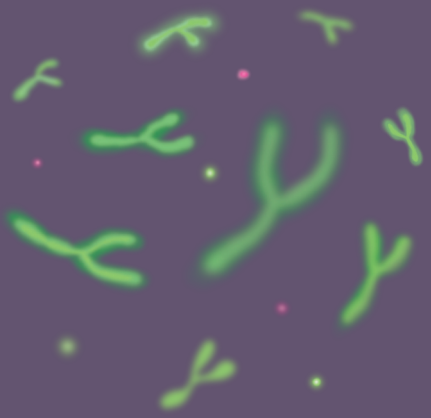
In **Hoofdstuk 3** hebben we multiplex immunofluorescentie en eiwitkleuring toegepast op UC-blaasweefsel van onbehandelde tumoren (n=31) en tumoren behandeld met pre-operatieve combinatie checkpoint inhibitors (n=24) om de UC TMO te karakteriseren. Ons gemultiplexte antilichaampanel stelde ons in staat diverse immuuncelpopulaties (bijv., T-cellen, B-cellen, macrofagen, etc.) en tertiaire lymfoïde structuren (TLS) - waarvan de rol in kanker nog steeds wordt opgehelderd - te karakteriseren. We hebben waargenomen dat checkpointinhibitors de invasieve marge van tumoren een hogere hoeveelheid aan $CD8^+PD1^+$ T-cellen vertoonden. Bovendien verschilden de proporties immuuncellen binnen TLS's tussen tumoren die waren behandeld met of zonder checkpointinhibitors, en hun rijpingsfasen varieerden afhankelijk van hun nabijheid tot het tumorgebied. Dit suggereert dat niet alle TLS's in tumorweefsel direct gerelateerd kunnen zijn aan antitumoractiviteit. Onze resultaten legden de basis voor verder onderzoek naar biomarkers binnen de context van pre-operatieve ICI-behandelingen voor UC.

In **Hoofdstuk 4** hebben we eerst de klinische gegevens gepresenteerd van NABUCCO. NABUCCO is een fase-I klinische trial die de haalbaarheid testte van het verwijderen van een tumor na een combinatie van pre-operatieve checkpointremmers ipilimumab en nivolumab bij 24 patiënten. De primaire uitkomstmaat - de haalbaarheid van de behandeling - van de studie werd behaald en duurzame respons op behandeling bij meer dan vijftig procent van de patiënten werd waargenomen. Vervolgens voerden we een analyse uit om tumorkarakteristieken van respons op pre-operatieve combinatie checkpointinhibitors en behandelingsgerelateerde dynamiek te kwantificeren. In tumorweefsel vooraf en achteraf aan behandeling, hebben we hele exoom DNA-sequencing, 'bulk' RNA-sequencing en eiwitexpressiepatronen geprofileerd. Hierdoor konden we concluderen dat respons onafhankelijk was van CD8+ T-cel-gerelateerde immuniteit en dat bij patiënten met respons op therapie, de behandeling vorming van TLS induceerde.

In **Hoofdstuk 5** hebben we spatial relaties (SRs) gekwantificeerd in de TMO van UC-tumoren en deze geassocieerd met respons op een combinatie van checkpointremmers. Eerst hebben we een methode opgesteld om SR's te kwantificeren op basis van een 'curve fitting' op de 'first-nearest neighbor' statistiek. Deze benadering stelde ons in staat om ruimtelijke gegevens kwantitatief te modelleren, zoals multiplex immunofluorescentie data. We ontdekten dat SR's significant geassocieerd zijn met de respons op pre-operatieve combinatie ICIs en dat SR's beter presteren dan kwantitatieve gegevens van immuuncellen. Dit laatste is de standaard manier om gegevens over multiplex immunofluorescentie te kwantificeren - bij het voorspellen van de behandelingsrespons. Tumoren die respons op therapie vertoonden, werden gekenmerkt door de nabijheid van $CD8^+$ T-cellen tot kankercellen en macrofagen tot kankercellen. Onze belangrijkste bevindingen hebben we bevestigd in een onafhankelijke cohort van tumoren van hoofd- en halskanker (n=25) met een vergelijkbare behandeling met

een checkpointinhibitorcombinatie. Dit suggereert dat de biomarkers die we hadden gevonden, tumoronafhankelijk zijn en respons op combinatiebehandeling met checkpointinhibitors voorspellen.

Het werk in dit proefschrift heeft bijgedragen aan het valideren van genomische biomarkers voor de respons op pre-operatieve chemotherapie bij UC. Daarnaast heeft dit proefschrift bijgedragen aan het voorstellen van een combinatie van pre-operatieve CTLA-4 en PD-1 blokkade als een effectieve behandeling tegen UC, om ons begrip van de respons op checkpointinhibitors als therapeutische optie bij UC te verbeteren en een nieuwe manier voor te stellen om de TMO te karakteriseren door gebruik te maken van SRs. Samenvattend hebben onze bevindingen onze huidige kennis van de biologie van UC-tumoren uitgebreid, wat kan bijdragen aan een verbeterde patiëntenzorg in de toekomst.



Resumen

El carcinoma urotelial (CU) es una malignidad que se origina por la presencia de células cancerosas en la vejiga que causan un pronóstico desfavorable, incluso en las primeras etapas de la enfermedad. Actualmente, el tratamiento más común para curar la enfermedad consiste en la extirpación quirúrgica del tumor mediante cistectomía radical (CR), que a menudo se complementa con quimioterapia preoperatoria (NAC, quimioterapia neoadyuvante). En las últimas dos décadas, el tratamiento del CU ha permanecido sin cambios.

La terapia con inhibidores de puntos de control inmunitario, también conocida como terapia de immune checkpoint inhibitors (ICIs), es un novedoso tratamiento que bloquea proteínas específicas, conocidas como puntos de control inmunológico o checkpoints immunológicos, los cuáles regulan desfavorablemente la respuesta inmune. Tras el bloqueo de los checkpoint immunológicos mediante ICIs, las células inmunitarias pueden reconocer más eficazmente las células cancerosas e inducir una respuesta inmunitaria contra las células cancerosas. En el contexto del CU, los ensayos clínicos de fase temprana han demostrado tanto la viabilidad como la eficacia de los ICIs en fase preoperatoria, y, además, han demostrado respuestas clínicas duraderas. Por lo tanto, los ICIs son una terapia prometedora con posibilidades de ser incorporados en el manejo clínico del CU en el futuro. Sin embargo, no todos los pacientes responden a los ICIs y además los ICIs causan unas proporciones sustanciales de toxicidades. Por lo tanto, existe una necesidad de identificar biomarcadores (biomarkers) que ayuden a identificar qué paciente tiene más probabilidad de responder satisfactoriamente al tratamiento.

Los estudios presentados en esta tesis doctoral se centran en dilucidar el rol del microambiente tumoral (TME, tumor microenvironment) del CU en la inducción de respuestas clínicas a tratamientos preoperatorios, como la quimioterapia y los ICIs. A través de un enfoque multiómico, en esta tesis hemos cuantificado las características intrínsecas y extrínsecas de los tumores de CU recolectados de muestras humanas. Estas características tumorales han sido asociadas con variables clínicas como las respuestas clínicas y las dinámicas (cambios) asociadas a tratamientos preoperatorios.

En el **Capítulo 1**, presentamos el panorama actual del tratamiento del CU y el rol del sistema inmunológico en el cáncer. Además, destacamos el cuadro general del (limitado) conocimiento actual de los biomarcadores moleculares asociados con respuestas clínicas a tratamientos preoperatorios en el CU, como la quimioterapia y los ICIs. Por último, destacamos la contribución y objetivo de esta tesis.

En el **Capítulo 2**, realizamos una secuenciación dirigida (targeted) de ADN para validar biomarcadores asociados con respuestas clínicas a la quimioterapia preoperatoria en una cohorte retrospectiva multicéntrica de 165 pacientes. Entre las tres hipótesis testeadas,

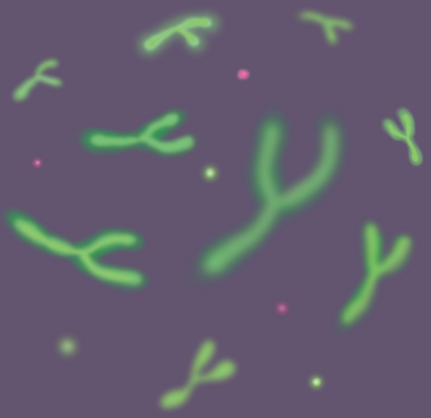
solo pudimos confirmar una asociación positiva entre la respuesta patológica y mutaciones en *ERCC2*. Aunque esta correlación fue también significante con la supervivencia libre de progresión (la probabilidad temporal de que un tumor progrese), no hubo correlación entre las mutaciones en *ERCC2* y la supervivencia general (la probabilidad temporal de que un paciente sobreviva), lo que subraya la necesidad de realizar futuras evaluaciones en estudios prospectivos.

En el **Capítulo 3**, empleamos inmunofluorescencia múltiple (multiplex immunofluorescence) y tinciones de proteínas en muestras quirúrgicas mediante RC de tumores de CU no tratados (n = 31) y tratados con ICIs preoperatorios (n = 24) para caracterizar el TME del CU. Nuestro panel inmunológico multiplexado nos permitió caracterizar diversas células inmunes (por ejemplo, células-T, células-B, macrófagos,…) y estructuras linfoides terciarias (TLS), cuyo papel en el cáncer aún se está dilucidando. Las muestras tratadas con ICIs se caracterizaron por una una mayor abundancia de células-T $CD8^{+}PD1^{+}$. Además, las proporciones de células inmunitarias dentro de las estructuras TLS difirieron entre los tumores sin tratamiento previo y los tratados con ICI, y los estados de maduración de las TLS variaron dependiendo de su proximidad al margen tumoral. Estos resultados sugieren que no todos las estructuras TLS presentes en el tejido tumoral están directamente relacionadas con actividad anti-tumoral. Nuestros datos proporcionan una base para futuras investigaciones relacionadas con el descubrimiento de biomarcadores en el contexto de los tratamientos preoperatorios con ICIs para el CU.

En el **Capítulo 4**, presentamos los datos clínicos de NABUCCO. NABUCCO es un ensayo clínico de fase temprana que investiga en 24 pacientes la viabilidad de resectar tumores de CU después de un tratamiento con una combinación de ICIs en fase preoperatoria que bloquean los checkpoints immunológicos CTLA-4 y PD-1 (con los fármacos ipilimumab y nivolumab, respectivamente). El criterio de endpoint primario del ensayo clínico fue alcanzado, y se observaron respuestas clínicas duraderas en más de la mitad de los pacientes. Posteriormente, realizamos un análisis para investigar qué características tumorales determinan una respuesta favorable al tratamiento con ICIs y qué características son alteradas por efecto del tratamiento. Usando muestras tumorales recolectadas previa y posteriormente al tratamiento con ICIs, los tumores fueron cuantificados mediante secuenciación del exoma completo del ADN (whole-exome DNA sequencing), mediante secuenciación del ARN y mediante tinciones y inmunofluorescencia múltiple (multiplex immunofluorescence) para cuantificar proteínas. Nuestros resultados nos permitieron concluir que la respuesta clínica a los ICIs es independiente de la inmunidad relacionada con las células-T $CD8^{+}$ y que el tratamiento con ICIs induce una mayor presencia de estructuras TLSs en pacientes que respondieron al tratamiento.

En el **Capítulo 5**, cuantificamos las relaciones espaciales (spatial relationships, SR) en el TME de los tumores de CU y asociamos las SRs con la respuesta clínica a la combinación de ICIs en fase pre-operativa. Primero, proponemos una metodología para cuantificar las SR basada en un ajuste de curvas a la estadística espacial del vecino más próximo (first nearest-neighbor). Este método nos permitió modelar cuantitativamente datos resueltos espacialmente, como por ejemplo la inmunofluorescencia múltiplex. Mediante la cuantificación de SRs observamos que las SRs se asocian significativamente con la respuesta clínica a la combinación de ICIs, y que las SRs tienen un poder predictivo superior comparado a las métricas de abundancia de células inmunitarias (densidad celular; la manera estándar de cuantificar los datos de inmunofluorescencia múltiple) en modelos de predicción de la respuesta clínica al tratamiento con ICIs. Los tumores que respondieron a ICIs se caracterizaron por la proximidad de las células-T $CD8^+$ a las células cancerosas y de los macrófagos a las células cancerosas. Nuestros hallazgos fueron validados en una cohorte independiente de tumores cancerosos de cabeza y cuello (n = 25) con un tratamiento con ICIs comparable al del cohorte de CU, lo que sugiere que los biomarcadores propuestos en este capítulo pueden ser una característica independiente del contexto tumoral que determina una respuesta favorable al tratamiento con ICIs.

El trabajo de esta tesis ha contribuído a validar biomarcadores genómicos de respuesta a la quimioterapia preoperatoria en la CU, a sugerir la combinación de bloqueo de CTLA-4 más PD-1 como un tratamiento eficaz para curar el CU, a avanzar nuestra comprensión de la respuestas clínicias a los ICIs en el CU, y a proponer una metodología novedosa para caracterizar el TME empleando relaciones espaciales. En conjunto, nuestros hallazgos han ampliado el conocimiento actual sobre la biología tumoral del CU que podría ser empleado para mejorar la atención al paciente en el futuro.



Resum

El càncer urotelial (CU) és una malignitat originada per la presència de cèl·lules canceroses a la bufeta que causen un pronòstic desfavorable, fins i tot en les primeres fases de la malaltia. Actualment, la forma estàndard per curar la malaltia implica l'extirpació quirúrgica del tumor mitjançant cistectomia radical (RC), que sovint es complementa amb quimioteràpia preoperatòria (NAC, quimioteràpia neoadjuvant). En les últimes dues dècades, el panorama del tractament de la malaltia s'ha mantingut inalterat.

Els inhibidors del punt de control immunitari (també coneguts com immune checkpoint inhibitors, o ICIs) són un nou tipus de tractament que bloqueja proteïnes específiques, conegudes com a punts de control immunitari, que regulen la resposta immune. Mitjançant el bloqueig dels punt de control immunitari amb ICIs, les cèl·lules immunitàries poden reconèixer més eficaçment les cèl·lules canceroses i induir una resposta immune contra les cèl·lules canceroses. En el context del CU, els assaigs clínics de fase inicial han demostrat tant la viabilitat com l'eficàcia dels ICI en fase preoperatòria, i també han demostrat respostes clíniques de llarga durada. Per tant, els ICI prometen canviar la gestió clínica del CU en un futur. Tanmateix, no tots els pacients mostren una resposta clínica a els ICI, i a més una quantitat substancial de pacients mostra toxicitats relacionades amb el tractament, cosa que posa de manifest la necessitat d'identificar biomarcadors que puguin ajudar a la identificació del pacients que tenen més probabilitat de respondre al tractament.

Els estudis presentats en aquesta tesi se centren en dilucidar el rol del microambient tumoral del CU (tumor microenvironment, TME) en la inducció de respostes clíniques als tractaments preoperatoris, tals com la quimioteràpia i els ICIs. Mitjançant un enfocament multiòmic, hem quantificat les característiques intrínseques i extrínseques dels tumors de CU recollits de mostres humanes. Les característiques tumorals han estat associades amb variables clíniques, tals i com la resposta clínica al tractament i la dinàmica associada amb el tractament.

Al **capítol 1**, presentem tant l'estat actual del tractament del CU com el paper del sistema immunitari en el càncer. A més, destaquem el panorama (limitat) actual del coneixement de biomarcadors moleculars de resposta clínica a tractaments preoperatoris pel CU com la quimioteràpia i els ICIs. Finalment, destaquem la contribució i l'objectiu d'aquesta tesi.

Al **capítol 2**, realitzem una seqüenciació dirigida d'ADN (targeted sequencing) amb l'objectiu de validar biomarcadors de resposta clínica a la quimioteràpia preoperatòria en un cohort retrospectiu multicèntric de 165 pacients. Entre les tres hipòtesis testejades, només vam poder confirmar una associació positiva entre la resposta patològica al tractament i les mutacions al gen *ERCC2*. Tot i que aquesta correlació es va mantenir significant amb la supervivència sense progressió (progression-free survival), la correlació no es va mantenir

significant amb la supervivència global (overall survival), cosa que subratlla la necessitat d'una futura avaluació addicional en estudis prospectius.

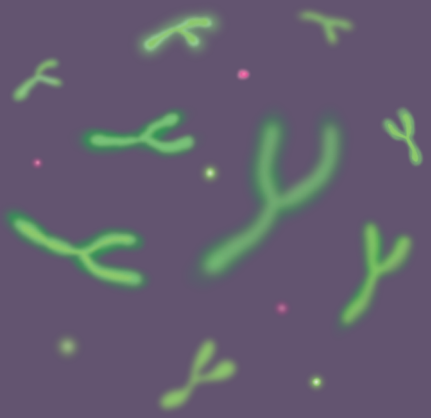
Al **capítol 3**, vam utilitzar immunofluorescència múltiple i tincions de proteïnes en mostres de RC de tumors de CU no tractats (n = 31) i de tumors de CU tractats amb una combinació d'ICIs preoperatoris (n = 24) per caracteritzar el TME del CU. El nostre panell immunitari multiplexat ens va permetre caracteritzar diverses poblacions de cèl·lules immunitàries (per exemple, cèl·lules-T, cèl·lules-B, macròfags, ...) i estructures limfoides terciàries (TLS), el paper de les quals en el càncer encara s'està dilucidant. El marge invasiu dels tumors tractacts amb ICIs van mostrar una major abundància de cèl·lules-T $CD8^+PD1^+$. A més, les proporcions de cèl·lules immunitàries dins de les TLS van diferir entre els tumors no tractacts i els tumors tractats amb ICIs, i els estats de maduració de les TLS van variar en funció de la seva proximitat al llit tumoral, cosa que suggereix que no totes les TLS presents al teixit tumoral estan directament relacionades amb activitat anti-tumoral. Les nostres dades proporcionen una base per a futures investigacions sobre el descobriment de biomarcadors en el context dels tractaments preoperatoris amb ICIs pel CU.

Al **capítol 4**, primer presentem les dades clíniques de NABUCCO, un assaig clínic de fase I que testejava la viabilitat d'extirpar un tumor posteriorment al tractament amb una combinació d'ICIs preoperatoris que bloquegen els punts de control immunitari CTLA-4 i PD-1 (ipilimumab i nivolumab, respectivament) en 24 pacients. El criteri d'endpoint primari de l'assaig clínic va ser assolit, i es van observar respostes clíniques duradores en més de la meitat dels pacients. A continuació, vam realitzar una anàlisi translacional per quantificar les característiques del tumors que determinen resposta clínica al tractament amb ICIs i que son alterats pels efectes dinàmics associats al tractament. Fent ús de mostres prèvies i posteriors al tractament, vam quantificar els tumors amb seqüenciació de l'exoma sencer de l'ADN de l'exoma sencer (whole-exome DNA sequencing), amb seqüenciació de l'ARN, i amb immunofluorescència múltiple i tincions per perfilar l'abundància de proteïnes. Els nostres resultats ens van permetre concloure que la resposta clínica al tractament amb ICIs és independent de la immunitat relacionada amb les cèl·lules-T $CD8^+$, i que en pacients que responen al tractament amb ICIs el tractament indueix la presència de TLS.

Al **capítol 5**, vam quantificar les relacions espacials (spatial relationships, SRs) al TME dels tumors de CU i vam associar les SRs amb la resposta clínica al tractament pre-operatori amb combinació d'ICIs. En primer lloc, vam proposar una metodologia per quantificar els SR basats en un ajust de corbes a l'estadística espacial del veí més proper (first-nearest neighbor). Aquest enfocament ens va permetre modelar quantitativament dades resoltes espacialment, com ara la immunofluorescència múltiple. Els SR s'associen significativament amb la resposta clínica a les combinacions d'ICI preoperatòries, i els SR van superar les

mètriques d'abundància de cèl·lules immunitàries (la forma estàndard de quantificar les dades d'immunofluorescència múltiple), quan es prediuen la resposta clínica al tractament amb ICIs. Els tumors que responen a ICIs es van caracteritzar per la proximitat de les cèl·lules-T $CD8^+$ a les cèl·lules canceroses i els macròfags a les cèl·lules canceroses. És important destacar les nostres principals troballes es van validar en una cohort independent de tumors de càncer de cap i coll (n = 25) amb un tractament similar de combinació d'ICIs, cosa que suggereix que els nostres biomarcadors podrien ser una característica independent del context tumoral que determina la resposta al tractament amb combinació d'ICIs.

El treball d'aquesta tesi ha contribuït a validar biomarcadors genòmics de resposta a la quimioteràpia preoperatòria al CU, ha suggerit la combinació de bloqueig de CTLA-4 més PD-1 com a tractament eficaç per curar el CU, ha avançat la nostra comprensió de la resposta clínica als tractaments pre-operatius amb ICIs al CU, i ha proposat una nova manera de caracteritzar el TME emprant relacions espacials (SRs). En conjunt, les nostres dades han ampliat el nostre coneixement actual de la biologia del tumor del CU que podria servir per millorar l'atenció al pacient en un futur.



List of publications

MANUSCRIPTS

Published

van Dijk, N.*, **Gil-Jimenez, A.***, Silina, K., Hendricksen, K., Smit, L. A., de Feijter, J. M., van Montfoort, M. L., van Rooijen, C., Peters, D., Broeks, A., van der Poel, H. G., Bruining, A., Lubeck, Y., Sikorska, K., Boellaard, T. N., Kvistborg, P., Vis, D. J., Hooijberg, E., Schumacher, T. N., van den Broek, M., Wessels, L. F. A., Blank, C. U., van Rhijn, B. W., van der Heijden, M. S. Preoperative ipilimumab plus nivolumab in locoregionally advanced urothelial cancer: the NABUCCO trial. ***Nature Medicine**, 26, 1839–1844* (2020)

van Dijk, N.*, **Gil-Jimenez, A.***, Silina, K., Montfoort, M. L. van, Einerhand, S., Jonkman, L., Voskuilen, C. S., Peters, D., Sanders, J., Lubeck, Y., Broeks, A., Hooijberg, E., Vis, D. J., van den Broek, M., Wessels, L. F. A., Rhijn, B. W. G., van der Heijden, M. S. The Tumor Immune Landscape and Architecture of Tertiary Lymphoid Structures in Urothelial Cancer. ***Frontiers in Immunology***, *12*, 5450 (2021)

Gil-Jimenez, A.*, van Dorp, J.*, Contreras-Sanz, A., van der Vos, K., Vis, D. J., Braaf, L., Broeks, A., Kerkhoven, R., van Kessel, K. E. M., Ribal, M. J., Alcaraz, A., Wessels, L. F. A., Seiler, R., Wright, J. L., Mengual, L., Boormans, J., van Rhijn, B. W. G., Black, P. C., van der Heijden, M. S. Assessment of Predictive Genomic Biomarkers for Response to Cisplatin-based Neoadjuvant Chemotherapy in Bladder Cancer. ***European Urology*,** 83(4), 313-317 (2022)

de Vries, H. M., Rafael, T. S., **Gil-Jimenez, A.**, de Feijter, J. M., Bekers, E., van der Laan, E., Lopez-Yurda, M., Hooijberg, E., Broeks, A., Peters, D., Seignette, I. M., Pos, F. J., Horenblas, S., van Rhijn, B. W. G., Jordanova, E. S., Brouwer, O. R., Schaake, E., van der Heijden, M. S. Atezolizumab With or Without Radiotherapy for Advanced Squamous Cell Carcinoma of the Penis (The PERICLES Study): A Phase II Trial. ***Journal of Clinical Oncology*** , 41(31), 4872-4880 (2023)

Gil-Jimenez, A., van Dijk, N., Vos, J. L., Lubeck, Y., van Montfoort, M. L., Peters, D., Hooijberg, E., Broeks, A., Zuur, C. L., van Rhijn, B. W. G., Vis, D. J., van der Heijden, M. S., Wessels, L. F. A. Spatial relationships in the urothelial and head and neck tumor microenvironment predict response to combination immune checkpoint inhibitors. *Nature Communications*, 15(1), 1-15 (2024).

Accepted

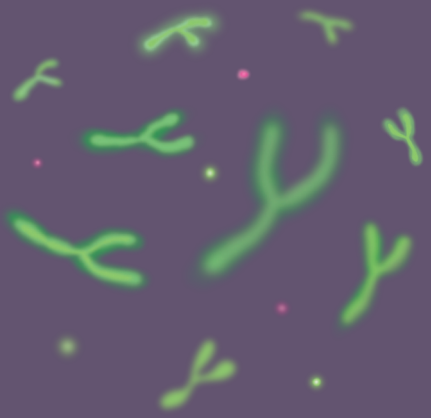
Rijnders, M., Nakauma-González, J. A., Robbrecht, D. G. J., **Gil-Jimenez, A.**, Balcioglu, H. E., Oostvogels, A. A. M., Aarts, M. J. B., Boormans, J. L., Hamberg, P., van der Heijden, M. S., Szabados, B. E., van Leenders, G. J. L. H., Mehra, N., Voortman, J., Westgeest, H. M., de Wit, R., van der Veldt, A. A. M., Debets, R., Lolkema, M. P. T Cell-to-Stroma Enrichment (TSE) score: a gene expression metric that predicts response to immune checkpoint inhibitors in patients with urothelial cancer. ***BioRxiv***, (2023). Accepted for publication in ***Nature Communications***

PUBLISHED CONFERENCE ABSTRACTS

Dijk, N. van, **Gil-Jimenez, A.**, Silina, K., Hendricksen, K., Smit, L., Feijter, J. de, van Montfoort, M. L., Broeks, A., Lubeck, Y., Sikorska, K., Boellaard, T. N., Kvistborg, P., Vis, D. J., Hooijberg, E., Schumacher, T., van den Broek, M., Wessels, L. F. A., Blank, C. U., van Rhijn, B. W. G., Heijden, M. S. van der. Biomarker analysis and updated clinical follow-up of preoperative ipilimumab (ipi) plus nivolumab (nivo) in stage III urothelial cancer (NABUCCO). ***Journal of Clinical Oncology***, 38(15_suppl), 5020–5020 (2020).

van Dorp, J., Pipinikas, C., van Dijk, N., Jones, G., **Gil-Jimenez, A.**, Marsico, G., van Montfoort, M. L., Hackinger, S., Braaf, L., McLay, K., van den Broek, D., van Rhijn, B. W., Rosenfeld, N., van der Heijden, M. S. Abstract 1273: Predicting pathological response after ipilimumab plus nivolumab in stage III urothelial cancer by liquid-biopsy assessment of plasma and urine ctDNA using the RaDaR assay. ***Cancer Research***, 82(12_Supplement), 1273–1273 (2022).

Gil-Jimenez, A., van Dijk, N., Lubeck, Y., van Montfoort, M. L., Peters, D., Hooijberg, E., Broeks, A., Vos, J. L., Zuur, C. L., van Rhijn, B., Vis, D. J., van der Heijden, M. S., Wessels, L. F. Abstract 5786: Spatial relationships in the tumor microenvironment predict response to immune checkpoint inhibitors in urothelial and head and neck cancer. ***Cancer Research***, 83(7_Supplement), 5786–5786 (2023).



PhD Portfolio

Courses - Discipline skills	
2024	Clinical internship (Antoni van Leeuwenhoek – NKI)
2023	Machine Learning with Python (Amsterdam Data Academy)
2022	Using MOFA for integration of omics data (VIB - Flemish Institute for Biotechnology)
2022	Histopathology of Human tumors (OOA)
2022	Advanced Immunology (UMC,VUMc)
2022	Nextflow for reproducible and automated data analysis (VIB - Flemish Institute for Biotechnology)
2020	Single-cell analysis (BioSB)
2020	Basic Oncology (OOA,VU)
2019	Bioinformatics Summer school (UCLouvain)

Courses - Transferable skills	
2023	How to be a good mentor (Mentor{y}teller)
2023	Designing scientific posters and covers for theses with Adobe InDesign (TU Delft)
2023	Career development: job interview preparation (TU Delft)
2023	Graphic design, posters, and infographics (CENL)
2023	Design your thesis with InDesign (OOA)
2023	Career development: marketing tools to position yourself in the job market (TU Delft)
2023	Career development: preparing for your next career step in academia (TU Delft)
2023	Career development: looking for work in the Netherlands (TU Delft)
2022	Career development: exploring a research career outside academia (TU Delft)
2022	Language certification NT2 Staatsexamen - Netherlands als tweede taal (Dutch B2) (Dienst Uitvoering Onderwijs)
2022	Project management for PhD candidates (TU Delft)
2021	Creative tools for scientific writing (TU Delft)
2021	Teamwork, leadership, and group dynamics (TU Delft)
2020	Time Management (TU Delft)
2020	Analytic storytelling (TU Delft)
2020	How to write research papers (OOA)
2020	Language course: Dutch B2 speaking (Klestmaatjes)
2020 - 2021	Language course: Dutch B2 (Toptaal)
2020	Language course: Dutch B1 pre-intermediate (VU University Amsterdam)
2019	PhD startup modules B & C (TU Delft)

Courses - Research skills	
2021	Ethics and Integrity in Science (OOA)
2021	Basic problem-solving and decision-making for researchers (TU Delft)
2020	Speedreading and Mindmapping (TU Delft)

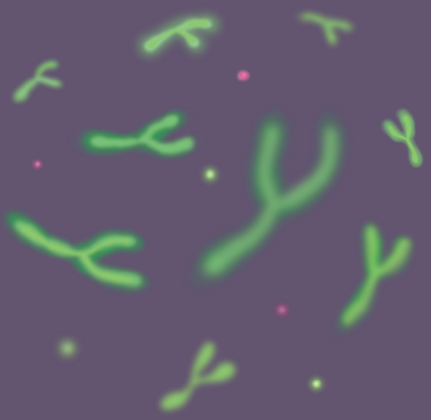
Funding obtained	
2022	Oncode technology access pilot - Spatial Transcriptomics (10k), to perform Nanostring Digital Spatial Profiling. Applicants: Lodewyk F. A. Wessels, Michiel S. van der Heijden (Oncode Institute)

Supervision	
2019, 2020	MSc student, Yoni Lubeck. Co-supervised with Dr. Daniel J. Vis, Prof. Lodewyk F. A. Wessels, Dr. Michiel van der Heijden (NKI)

Organized seminars, workshops and divulgation	
2023	Event: We are all CENL: Science is for everyone - an event on Diversity, Equity, and Inclusion for science. Co-organized with Leticia G. Leon and Alba Muniesa, CENL - The Association for Spanish Scientists in the Netherlands (Utrecht, the Netherlands)
2023	Seminar: Sex and gender in Biomedical research, by Sabine Oertelt-Prigione, NKI (Amsterdam, the Netherlands)
2023	Divulgation workshop on Bioinformatics and DNA for young students (5-10 years old), co-organized with Lety G. Leon, CENLab-CENL - The Association for Spanish Scientists in the Netherlands (Utrecht, the Netherlands)
2022	Divulgation blogs on the intersection between Science and (LGBTQ+) Diversity, Equity and Inclusion. Co-written with Lety G. Leon and Alba Muniesa, CENL - The Association for Spanish Scientists in the Netherlands (Utrecht, the Netherlands)
2022	Divulgation workshop on Bioinformatics and DNA for young students (5-16 years old), co-organized with Lety G. Leon, Cervantes Institute scientific fair, CENL - The Association for Spanish Scientists in the Netherlands (Utrecht, the Netherlands)
2021	Seminar: Computational analysis of cancer genomes, Nuria Lopez-Bigas, NKI (Amsterdam, the Netherlands)

Invited speaker	
2023	Computational Methods for Spatial Data Analysis, Oral presentation, BioSB HotTopics Symposium (Leiden University Medical Center (LUMC), the Netherlands)
2023	OncoCENL, Oral presentation, CENL - The Association for Spanish Scientists in the Netherlands (Cervantes Institute, Utrecht, the Netherlands)
2022	Round table on equity, equality, and diversity in the immunological working environment, panelist, yEFIS Equity and Diversity working group (online)

Attended conferences	
2023	American Association for Cancer Research (AACR) annual meeting, *Oral presentation* (Orlando, USA)
2022	European Conference on Computational Biology (ECCB), *Poster presentation* (Sitges, Spain)
2022	European Student Council Symposium from ISCB (ESCS), *Oral presentation* and Poster presentation (Sitges, Spain)
2022	International PhD student Cancer Conference (IPSCC), *Oral presentation* (Heidelberg, Germany)
2021	Bioinformatics & Systems Biology Conference (BioSB), *Oral presentation* (online, the Netherlands)
2021	Oncode annual meeting, *Poster presentation* (online, the Netherlands)
2019	OOA retreat, *Poster presentation* (Renesse, the Netherlands)



Curriculum Vitae

Alberto GIL JIMÉNEZ (he/him*)

20-03-1993	Born in Hospitalet de Llobregat, Spain

Education	
2016 – 2018	**MSc in Bioinformatics and Systems Biology,** *Cum* laude Vrije Universiteit Amsterdam and University of Amsterdam (Amsterdam, The Netherlands)
2011 – 2016	**BSc in Physics** Autonomous University of Barcelona (Barcelona, Spain)
2011 – 2016	**BSc in Chemistry** Autonomous University of Barcelona (Barcelona, Spain)
2009 – 2011	**High School, Health Sciences track** Salesians de Sarrià (Barcelona, Spain)
2005 – 2009	**High School** Academia Cultura (Hospitalet de Llobregat, Spain)

Professional Experience	
2019 – 2023	**PhD, Computational Cancer Biology**: study of neoadjuvant treatments for urothelial cancer Supervised by Prof. Lodewyk F. A. Wessels, Dr. Michiel S. van der Heijden, Dr. Daniel J. Vis Netherlands Cancer Institute (Amsterdam, the Netherlands) and Delft University of Technology (Delft, the Netherlands)
2018	**MSc internship, Computational Cancer Biology:** study of the role of germline variants associated with breast cancer survival Supervised by Prof. Marjanka Schmidt, Dr. Sander Canisius, Dr. Maria Escala-Garcia and Dr. Sanne Aben Netherlands Cancer Institute (Amsterdam, the Netherlands)

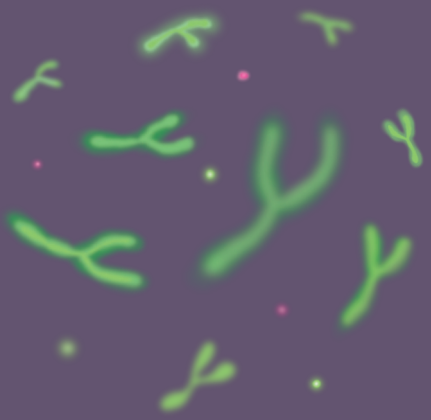
Professional Experience	
2018	**Teacher Assistant,** ***Machine Learning*** Supervised by Dr. Peter Bloem Vrije Universiteit Amsterdam (Amsterdam, the Netherlands)
2017 - 2018	**MSc internship, Data Science**: Interpretable machine learning to understand hospital readmission rates Supervised by Dr. Thomas Kelder and Dr. Sanne Aben EdgeLeap (Utrecht, the Netherlands)
2018 – 2019	**Bioinformatician**: data services on observational data and large-scale analytics The Hyve (Utrecht, the Netherlands)
2017	**Teacher Assistant,** ***Fundamentals of Bioinformatics*** **and** ***Programming in Python*** Supervised by Dr. Anton Feenstra and Dr. Sanne Aben Vrije Universiteit Amsterdam (Amsterdam, the Netherlands)
2015 – 2016	**BSc internship**, **Computational Chemistry**: modelling of the fluorescent properties of a protein Supervised by Prof. Ricard Gelabert and Dr. Jordi Faraudo Autonomous University of Barcelona (Bellaterra, Spain) and CSIC-The Institute of Materials Science of Barcelona (Bellaterra, Spain)
2015	**Fellow Researcher, Chemometrics**: study of the infrared spectra of proteins Supervisors: Prof. Jens Biegert and Dr. Idoia Martí i Aluja ICFO - The Institute of Photonic Sciences (Castelldefels, Spain)
2012 – 2016	**Private teacher**: physics, chemistry, maths, English Freelance (Barcelona, Spain)

Extracurricular and sideline activities	
2023 - present	**Coordinator - Diversity, Gender and Equity committee** CENL (The Association of Spanish Scientists in the Netherlands)
2023 - present	**Participation - Communication committee** CENL (The Association of Spanish Scientists in the Netherlands)
2020 - 2023	**Participation - Diversity, Equity and Inclusion committee** Netherlands Cancer Institute (Amsterdam, the Netherlands)
2018, 2019, 2022	**Volunteer** International Queer and Migrant film festival and Bingham cup
2017 - 2018	**Volunteer**, travel committee ASVGay student association
2014	**Hostel Assistant** Scottish Youth Hostels Association (Edinburgh, Scotland, UK)

Extracurricular and sideline activities	
2013	**Front desk receptionist** Lagan Backpacker's Hostel (Belfast, Northern Ireland, UK)

* Sharing pronouns is a little effort and makes it inclusive for everyone. By openly stating my pronouns, I try to contribute to a more respectful and supportive environment where individuals of diverse gender identities feel seen and acknowledged. It fosters a culture of understanding and helps break down assumptions, creating a space where everyone's identity is valued.

If you wish to learn more, please read our blog on how to be a good LGBTQ+ ally: https://www.cenetherlands.nl/en/lgbtq-allies-101-gender-identity-and-how-to-be-a-good-ally/



Acknowledgements

What an adventure these past five years have been, far from the path I originally imagined. I'm very grateful that this time has not only helped me grow as a scientist and professional but also on a personal level.

My gratitude goes to everyone who contributed to making this journey a reality, helping me experience change and growth. Summarizing it in a few words feels almost impossible, so this section will be an attempt to do so.

First, I want to thank the team of supervisors who made the work on this thesis possible and supported me over these years. At times, I felt like I was in a boat in the middle of the ocean, close to the destination yet desperately in need of direction. Your guidance has been crucial for reaching the right direction, and your influence on this book has been profound. I value your mentorship and passion for science, which helped me navigate the PhD years and brought me so many opportunities to learn and grow. **Michiel**, before starting this journey, I never imagined developing such an interest in the clinical side of research. A few years later I'm glad the labels "clinical" and "immunology" are my favorites ones when choosing a project. I'm very grateful for the opportunity you brought me when I was accepted to join the group, even though little I knew by that time about immunotherapy and immunology in general. I want to give you a massive thank you for your guidance on clinical and translational research, supervising so many projects over the years, helping set up so many collaborations, introducing me to the immuno-oncology field, for all your patience and for providing me access to so much invaluable data. **Lodewyk,** your guidance has been so important to making it through the PhD, thank you for helping me (and sometimes pushing me) to see the 'big picture' from each project, for all your help to keep in balance the countless collaborations and for supporting me throughout the PhD. Our discussions and your critical perspective have been very useful over the years, to move forward our projects and to keep up the motivation, I appreciate all your help. Even when the data was not telling us what it was supposed to tell us, you showed me how staying calm and cheerful at times was the right attitude to move forward or help looking at the data differently. **Daniel**, after meeting with you every week for the past five years, I can truly see the change—thank you! Thank you very much for your guidance and help, your help in each weekly meeting and for all the (statistical, computational, scientific, daily life) knowledge and advice that helped me progress and save so much time over the years. I also want to thank you for your patience, you've always had a tip on how to fix unexpected issues, how to be a "data detective" and be alert on unexpected things (like batch effects) that could mess up the analyses, or how to fix my unwanted attempts to crash the RHPC. I feel very grateful for your constant support and motivation and readiness to help.

I am very grateful to the **OOA committee** that accompanied me during this journey: **Marleen**, **Wilbert**, and **Joana**. Even though we only met yearly, your input has been very valuable and I saw every time we met as an opportunity to reflect on what has changed over the years. It made me realize how time flies and how to overcome frustrations while celebrating successes. Thank you so much for your time and independent help with the content of the thesis.

Moreover, I would like to extend my sincere thanks to the PhD reading committee: **Emile**, **John**, **Marleen**, **Marcel**, and the **rector magnificus**. Thank you for taking the time to assess and evaluate the content of my thesis and for your insights during the thesis defense.

The next shoutout goes to my dear paranymphs. **Sofía**, I still remember when we first met at the elevator when I was a naïve master student, and a few months later we eventually went out together, despite barely knowing each other. The connection and energy were intense and you've always been such a good ally and supporter: from overcoming frustrating moments over the PhD to overcoming all the problems and things from life and sharing our successes from the PhD (like sharing vitamin D at PR ;)) and helping me stay motivated. I'm glad we've shared so much time together in and out of the NKI and look forward to continuing our friendship. **Marcos**, who would have imagined that after meeting for the first time while being neighbors camping at the BBK, years later we would be sharing so many moments in Amsterdam. First, thank you so much for being my paranymph! Second, thank you for all the time we have spent over the years, I feel these years in Amsterdam have been full of adventures, joy, happiness, tears, successes, gossip, and overall growth. I'm glad we've been able to align our energies so well and helped each other so much over the years.

The research from this thesis has been greatly influenced by the interactions within the "**bladder cancer clinical team**", which helped me value how important interdisciplinary teams are for translational projects. The clinical interactions within our group have been so important in driving forward our translational projects and also into helping me understand what happens at the clinic side. **Nick**, it's been a pleasure collaborating on the first steps of my PhD and over the years within the NABUCCO and immune landscape projects. Before starting the PhD, little I knew about multiplex immunofluorescence, immune cells and how tumor slides work but you had enough patience to guide me through them and share all your knowledge in your translational projects. I'm grateful for all our collaborations and discussions, and I wish you the best in your future endeavors. **Jeroen**, we started working together with small projects and then ended up working together on the big chemo cohort for which I'm thankful for how you helped us with the survival data and to interpret the data. Thank you for always being there to answer questions about the patients and for the fun vibes to the group. **Chantal**, although the data from our collaborations did not

make it to this thesis, I enjoyed and I am looking forward to seeing about the upcoming new outcomes from the NABUCCO project. Thank you so much for the critical input in our collaborations, your patience during our discussions, and the positive mindset you brought to the group. **Maksim**, after a few years of being the only bioinformatician in the bladder group I was very happy to see a new bioinformatician joining us. Thank you so much for sharing all your knowledge with me, testing some of my scripts in your projects, for bringing so many new perspectives into our discussions on the data and overall advice on how to apply bioinformatics. **Hielke**, although we have been working on the PERICLES trial, my contribution did not start until you were already gone. Nevertheless, thank you for still being there to answer our questions even when you were already gone. Your joy is missed in the group ;) **Tysha**, our collaboration started by the end of my PhD adventure, but your motivation and questions were super useful to keep the translational work from PERICLES ongoing. Thanks for sharing so much knowledge during our collaboration and our discussion sessions on the project. I'm really looking forward to the final outcome! **Kristan**, we briefly overlapped at my naive times of Phd-year-0 but I appreciate we could work with your sequencing data from the chemo project that made it to this thesis. **Sander**, our overlap was very short, you brought a different perspective to the group with all your lab data, good luck on your post-phd adventures. **Charlotte**, thank you so much for assessing the slides from NABUCCO! I would also like to thank **Sarah**, even though we have briefly collaborated in the group our discussions have always been useful. Good luck on your last steps of your PhD. **Jan-Jaap**, our overlap was short during the end of my PhD but I'm happy to see how motivated you are. Good luck with your PhD endeavors. **Hamza**, our time together was very short but it has been intense - in a positive way! I am very impressed with your motivation to work on the bioinformatics world almost from scratch and I hope you make it through all the translational projects from the group. I wish everyone from the team the best in your PhD journey.

A massive thanks goes to everyone from the **Computational Cancer Biology group** who has been always willing to help and contributed to a nice environment in which to work. Growing in the field would not have been possible without critical input. Over these five years there have been countless interactions and help: thank you so much for your tips and critical input! **Marie**, we met when we both decided to start our adventures in bioinformatics and I'm glad that we "survived together" sharing master projects, teacher assistantships, and being office roomies at the NKI and Edgeleap. You've always been willing to help - I wish you all the best with your family and ongoing PhD adventures. **Tesa**, thank you for all the tips on how to build patience when analyzing DNA sequencing data and for all the fun vibes you brought to the office. I still miss all the scientific and life discussions we had there! **Marlous**, you might not remember but, when I was looking for an internship at the NKI, you helped build the bridge toward the NKI. My sincere thanks for that, and also

for the nice times we had sharing office and all your tips during my first steps in analyzing translational data, all the best in your teaching adventures! **Sander**, we started our adventures together when you supervised my master thesis and you showed me how fun it can be to analyze gene expression data. Believe it or not our project helped me grow an interest in the field and motivated me to come back to the NKI to pursue a PhD - thank you for your invaluable help and support! **Soufiane**, I still remember how lost I was during my first weeks as a PhD and you were one of the first people to give me a warm welcome to the group. Not only that but you've always provided feedback and help with projects. I had a great time in and outside the NKI. All the best in the UK! **Jin**, I wish you the best with your new adventure in your hometown, thank you so much for your help in "fishing" the important bits from the biological data. **Kathy**, you've always been so friendly and positive. Thank you for being so patient in explaining to us the experimental side from the lab, for all the help at the DEI committee, and for all your advice for daily life. I value your perseverance a lot ;) **Silvana**, it's been a while since I'm not at B7 but, I still remember the sound of your laugh resonating through the corridors. Thank you for bringing such good vibes to the department, always being so happy around us, and helping keep the good vibes. **Maarten**, it was so much fun sharing office during the PhD. You also helped me so much with the translational and immunological projects and my first attempts with R, thank you so much for your good vibes and help. **Joris**, thank you so much for the time spent and all the immunology knowledge that you brought to the group and all your feedback, I wish you the best. **Tom**, thank you for all the fun adventures during the PhD, we overlapped for most of the time and I value a lot your help and fun times in and out of the NKI. To the cheerful **Tim**, you've always brought good vibes to the group, I admire your positivity and also your motivation into bringing the more mathematical side of research into bioinformatics. **Alex**, thank you so much for all your help and advice with cancer immunology, coding, bioinformatics, and for the time spent outside the NKI. You have a vibrant mindset ;) **Bram**, it's been so nice to see your transition toward becoming a senior postdoc in the group. Thank you for all the advice and help in the critical discussions, and for showing us an example of how motivation into research pays off. **Tycho,** thank you so much for all the advice and help you gave me on several projects in this thesis, from DNAseq, RNAseq, to multiplexed data, I am very thankful for that. **Mysha**, you always had such good questions to twist the way we thought about a project, we miss your positive mindset to the group. **Kat**, thank you not only for the critical input that you brought into our discussions and for helping me think beyond the data. Your energy spent into improving the overall way of working at the group is much appreciated. **Lindsay**, I'm so glad you always brought so much fun to the group! I'm also thankful for all the knowledge you brought to us from our "downstairs neighbors" which has been so important for better understanding the immunological side of our research. **Duco**, it's been so much fun sharing office with you, thank you not only for our joint discussions but also for all the time spent at the borrels and outings! **Pedro**, it was nice to see you return to the

group to pursue your PhD and it's been fun to be together at the office. Thanks for all your advice on the translational projects that we've had, you brought so much knowledge to the group. **Roebi**, we never worked together but you've always brought so much fun energy to the group, I wish you all the best with your research and life. **Olga**, we started our PhDs almost at the same time and we've shared so many moments together. I wish you the best wrapping up your projects and I'm very grateful for all the help and advice you've given me to navigate the immuno-oncology field. **Ewald**, it was nice to see you coming back to the group, I wish you all the best with your new research line. **Michael**, the critical thinking that you brought into our group discussions, slack discussions and Friday borrel discussions was invaluable! All the best in the next step in your scientific career. **Evert**, we overlapped little during my PhD but I'm very grateful for your invitation to the BioSB spatial session. **Joe**, we were at the same office but covid times happened so in the end we barely had chances to sit together at the same time. Despite that, I appreciated the few moments we did have together and your critical discussions during the group meetings. **Guizela**, I value so much all the time you spent explaining to us your mathematical concepts with such easy to understand diagrams, and your input at the group discussions. **Franziska**, I wish you continued success in your research and all the best in your career. **Vanessa,** you were one of the first ones to bring AI into the group. Thank you for being so patient while explaining this side of the research to us, and more importantly, thank you for helping me keep my plants alive while I was out of the country :). **Paolo**, it's a pity that we shortly overlapped by the end of the PhD. You brought so many fun vibes to the office, I wish you the best on your PhD :) **Petros,** we met a long time ago around the NKI and it's a pity we only shared office and department in my last months. Thank you so much for all the bioinformatic and life advice you've given me over the years, I look forward to keeping in touch. **Joleen** we overlapped (at least, at the CCB) at the end of the PhD and your input has always been super useful, thank you so much for the bioinformatic and immuno-oncological advice and all the best with your new adventures. **Matt**, we briefly overlapped as I was about to depart the group, I wish you all the best in the group. **Niek**, your effort in applying the spatial approach into a newer and bigger dataset was invaluable. Thanks so much for such effort. To the master students **Lisa**, **Miguel**, **Juan**, **Huub**, **Niek**, **Lucía**, **Aaron, Bram,** you brought such good vibes to the group, it's been super nice to see so many fresh and bright minds bringing new ideas to the group.

This thesis wouldn't have been possible without the incredible support and collaboration from the NABUCCO team (for which, part of the data made it to three of the chapters from this thesis) and the **Urothelial Cancer team**, with a special mention to all the collaborators of the NABUCCO study. **Bas**, your role as a surgeon has been pivotal to the work on the trial, thank you for being one of the main characters in the NABUCCO manuscript and all your help in the neoadjuvant chemotherapy project, your clinical insights have been invaluable. Also

a big shoutout to all the **NABUCCO team** that made possible the success of the trial: **Jeantine**, **Maurits**, **Thierry**, and all the team. **Christian**, thank you for providing so much critical input in the development of NABUCCO project, your contributions added depth and richness to our research efforts. I would also like to thank all the **patients and their relatives**, from the clinical trials as well as from the retrospective data that has been part of this dissertation, for consenting the use of their data in our research. A big thanks also goes to the PERICLES team: **Hielke, Tysha,** and all the clinical team involved. I joined the "ship" later in the project but you were always there to help in our discussions.

A big thanks also comes to the neoadjuvant chemotherapy team which helped in gathering a large cohort of patients across continents (and across my two "homes": Barcelona and the Netherlands) as well as interpreting such a large collection of sequencing data: to the Rotterdam team (**Joost**, **Kim**, **Roland**, **Jonathan**), to the Barcelona Hospital Clinic team (**María José**, **Antonio**, **Lourdes**), and to the Vancouver team (**Peter Black, Alberto Contreras**).

Over the years and despite all possible changes, the only thing that has remained permanent has been being at the **B7 Department**. The coffee machine that virtually divides the wet lab side with the dry lab side has been so important over these years: not only to keep the focus with the coffee but to interact between us, to share successes and failures, and to keep a good atmosphere at the department. I would like to thank everyone who is part of it for the countless interactions, shared time and fun moments at our group outings. First and foremost, to the cornerstone of the department, B7 wouldn't be the same without you: **Patty**, you've always been willing to help to keep all the science functioning, the department running smoothly, the group outings being so fun, and my Dutch and my life in the Netherlands in check. Thank you so much for your invaluable knowledge and for keeping the "wheel" turning. I'd like to thank the invaluable knowledge and scientific output that our PIs bring to the department: **René** and **Roderick**, over the years we've seen so many successes coming from your side to the department which makes me proud to belong to it. **Antonio**, thank you for bringing all the joy and fun vibes to the department, as well as for bringing the Spanish vibes to the department together with **Isa** and **Begoña**. **Fleur**, I feel we started talking more during the end of our PhDs and I value these conversations a lot. It's been fun to share adventures during our film acting, thanks for being such an inspiration in graphic design, science and life! Your advice is greatly appreciated. **Robin**, I am impressed how quickly you managed to begin and wrap up your project with so much joy (starting after me and finishing before me), which inspired me a lot. Again, your illustrations were very impressive. **Rodrigo**, if I remember correctly you were the first one to show me a tumor slide. Even though we didn't overlap for a long time I'm inspired by your career and I'm grateful we're still in touch. Thank you also for the countless coffee breaks, the fun moments at our outings, and all the fun chats

we've had at the borrels: **Giulia**, **Sara**, **Ziva**, **Matheus**, **Cor**, **Johan**, **Mani**, **Liqin**, **Marielle**, **Katrin**, **João, Astrid**, **Marielle**, **Katrien**, **Gabriele**, **Mengnuo**, **Haojie**, **Arnout, Chysa, Gabriele, Johan, Marieke.** Also, **Mohammed**, you've been so many times at B7 that I consider you as part of our B7 family, gracias amigo por la buena onda!

Part of this thesis wouldn't have been possible without the fruitful collaborations that happened over the years. **Lotje** and **Joris Vos**: by the time we were preparing our spatial manuscript you kindly helped us improve the significance of our findings with your data from IMCISION which you kindly shared with us and provided invaluable input to the manuscript. I appreciate your help to validate our findings and help us publish our manuscript.

Moreover, the content from NABUCCO and the immune landscape significantly improved with the experts on Tertiary Lymphoid Structures. **Karina**, I appreciate your passion for what you do, and all the experimental data as well as input that you provided us to the manuscript and to upcoming projects. I wish you all the best in your research group and your grants. **Maries**, thank you so much for all the immunological input provided that significantly improved the manuscripts.

I would also like to thank the funders and financial support that made this work possible: **BMS**, **KWF**, and **Oncode Institute**.

To the **Genomics Core Facility team**, more than half of the content of this thesis wouldn't have been possible without your work. **Ron**, **Roderick**, **Roel**, **Arno**, **Wim**, and **all the members**: thank you so much not only for quantifying such important data, but also for always being there to answer our questions about how to interpret the analysis and outcomes.

The second cornerstone of the thesis comes from the invaluable work from the **Core Facility Molecular Pathology** and **Molecular Biobanking (CFMPB)**, who helped provide a wealth of tissue and molecular data that has been crucial for the development of the thesis. Thank you **Annegien**, **Dennis, Iris**, **Linde**, **Sten**, **Maartje and the rest of the team** for your invaluable work into handling and quantifying so many markers and so many samples that have been of use for the content of the thesis. A big thanks goes to the **Pathology department**, **Maurits**, **Erik**, **Joyce** and the rest team. Thank you so much for all your work in analyzing biological samples. **Yoni**, it was a great pleasure to work together with you on a dual side: on the one hand I would like to thank you for all your help providing us data from your facility, and on the other hand I would like to thank for being a master student that I co-supervised during the PhD; your work proved valuable input and a basis to write a chapter from this thesis.

Last but not least, the third component from this thesis comes from all the computational resources provided by the **Research High Performance Computing (RHPC) team**. **Torben**, **Daniel**, and now **Ameer**, thank you for providing us and supporting the tools and resources that help us carry our work - without you, doing our work would not be possible.

I'd also like to acknowledge the support from the Biostatistics team and their seminars. **Renée**, thank you for your inspiration and career advice, and the feedback you provided me into our spatial approach.

Last but not least, a massive thanks for all the facilities from the NKI: the **Library** for all the support in how to handle the publications, repositories, databases and data management, all the **support staff,** our colleagues from **Tiffany's**, and the **front-workers**: your work is invaluable. Moreover, I'd like to thank the core of the NKI: **Henri**, without your work most of the activities of the institute wouldn't be possible, thank you so much for all the hard work, and for facilitating the whole research to the institute and all the support to the committees.

Something I truly enjoyed during the PhD was being part of the **NKI Diversity and Inclusion committee**. Thanks to all the motivation given by **Migaisa**, **Leila**, and **Joana**, your help and leadership at the diversity committee has been truly a motivation and an inspiration. Moreover, thanks to the best team that made this possible, I'm proud of what we achieved together: **Danielle**, **Lety**, **Stefan**, **Chao**, **Ronak**, **Margarida**, **Olga**, **Ana Moisés.** Also to the new members that brought new vibrant energy into the committee: **Beatrice**, **Allard**, **Jitske**, **Minh**, **Floor**, **Jael**, **Vanessa**, thank you for helping us revive the committee!

My career path hasn't been a straight one: coming from a background in Physics and Chemistry it took time and a set of discoveries ending up in the Cancer Research world. For this I'd like to thank **Maria** and **Sander**, for welcoming me at C2 for doing my internship. Your combo helped me raise interest in cancer research and continue later in time into my PhD adventure. I'd also like to thank you **Marjanka**, for letting me join the group and for your feedback on the project. It didn't make it into a manuscript but it was a great learning experience! Also a big thanks to all the members from the C2 department from back on the days: **Hayra**, **Danielle**, **Anna**, **Marcelo**, **Myriam**, **Renske**, and all the **C2 department**. Lastly, thank you **Julia Houthuizen** for sharing your knowledge into CAFs.

Next I'd like to acknowledge all the fun moments that the NKI brought at borrels, conferences, retreats and unofficial outings, to all the celebrated successes and all the support on

times when things didn't go as planned. To the "**lo que unió el OOA retreat**" team: **Carmen**, **Ino**, **Ana Krotengerg**, **Rebeca**, I find it funny that 5 years later we still talk about the adventures we had back in the days. Thanks for all the memorable borrels and stories. To the unofficial committee from the NKI, the "**Spanish committee**": thank you for all the laughs and making me feel at home while being abroad: **Sofía**, **Celia**, **Daniela**, **Ana Krotenberg**, **Rebeca**, **Julia-C2**, **Fernando**, **Inés**, **Ángela**, **Xabier**, **Alberto**, **Isa**, **Fernando**, **Davinia**, **Noemí, Lali**, **Mercedes**, **Sebastiaan**, **Juan Simon, Miguel.** To all the countless fun interactions at the Friday borrel and the post-borrels: **Felipe**, **Ana Moisés**, **Simon**, **Lisa**, **Itske**, **Steven**, **Catrin**, **Mattia**, **Kelly**, **Shannon**, **Ferhat**, **Christina**, **Sebastian**, **Annoek**, **Dorine**, **Dimitrios**, **Maria**, **Ronak**, **Ziva**, **Mathias**, **Nils**, **Sra**, **Ronak**, **Francesca**, thank you for keeping always the good vibes on Fridays. **Trieu**, **Alex**, I had so much fun at the beginning of my PhD sharing the social endeavors in Tiffany's as well as outside the NKI on our shared travel, I wish you all the best in your careers. To the Summer party committee (**Tom**, **Christ**, **Jacobien**): it was so much deciding on a short notice (yes, we didn't know each other by that time) and fun to take part in the preparation of such a party (even though it was during my first months at the NKI). **Monique**, you've always been around for a chat and always with such good energy, thank you for all the life advice. **Daniil**, **Will**, **Stefan**, it's a pity we met in my last months at the NKI but I'm looking forward to more fun adventures! **Shirley**, **Manon**, our love for immunology brought us together and I'm glad for all the fun times we've had afterwards.

A significant encounter that happened during this journey was finding CENL **(the Association for Spanish Scientists in the Netherlands)**. By no means this led to a growth in skills and it made me realize how important (and how useful!) is to work on our soft skills. For this adventure I'd like to first thank **Lety** with which we've explained to kids how cool bioinformatics are, but also we've raised awareness on LGBTQ+ rights into the commission. Thank you also for all the fun times we had around Amsterdam and all your life advice that you have given me :) . **Marina**, thank you for showing your enthusiasm in such a divulgation project – it helped me so much getting out of my comfort zone and realizing how to explain in a cool way some 'bits' from this thesis. **Rodrigo** and **Sofía**, thank you so much for your countless work into the association, for bringing all your motivation to everyone and all for providing **Lety** and me a platform (together with **Aitor**, our blog master) to set up the Diversity, Gender and Equity committee. I'd also like to thank the members of the committee: **Alba** and **Cris**, thank you so much for what we have done so far with our divulgation blogs and our kick-off event 'We are ALL CENL' for which I'm super proud, gathering so many nice people in a place to discuss such important topics has been a dream that came true thanks to all the help from everyone. Moreover, **Andrea** (and now **Sofía)**, thank you so much for coordinating the communication committee and giving us some inspiration when preparing the RRSS post for the association: it's been a

nice opportunity in which to practice some illustrations for which I learnt a lot. Thanks to the rest of the team from the countless interactions. I'd also like to thank all the members from the **Oncology commission** for all the support. And to all the members for the fun after activities that we do.

I would also like to thank all the support from the teachers that taught me all about Bioinformatics back in the Master-naïve days when I barely knew anything about it. **Sanne**, **Anton**, **Jaap**, thank you so much for your help during the master and for helping me set up my first contacts in the field by then.

I'd also like to thank the teachers from the Universitat Autonoma de Barcelona, but especially the ones that raised my interest in the world of proteins and biology. **Ricard,** I always felt 'intimidated' by your work on quantum chemistry and your help during the bachelor thesis was so valuable – 10 years later I still keep a good memory about those times. Perhaps I didn't continue into the 'quantum' world but your guidance on the protein world was how I discovered what Bioinformatics was: thanks a lot for that. Also to **Jordi**, **Idoia**, thank you for bringing so much passion into your work.

Time flies and it's been around 8 years since I started building a new life in Amsterdam. I don't have enough words to describe what this time has meant to me in terms of personal growth and change and I'm super grateful to all the people that have been around me, thank you so much for your support. First, to **los chicos del Prik**. **Marcos**, you've already been acknowledged before ;) To the wisest from the group, **Rodri**, thank you so much for repeating to me so many times "you have to finish the thesis"; well, now it's finally done! I'm so glad for all the moments we've shared together in the city, all your help and support over the years and all the yummy food you've shared with us. **Fran**, we already miss you so much since you left Amsterdam :(You've always brought so many fun and crazy moments to the group, so many references (referencias!), so much life advice, so many laughs, and so many arrossets, thank you so much for being yourself. *Buenas noches* to **las Pamelas**: **Natsumi**, I really appreciate how we after crossing paths in my first bartender attempt became friends. Who would have thought that a year later we would be all in Paris to cheer you up at the runway. Thank you so much for all the crazy moments we've had together, for always being so positive, and for being a role model. **Jafet**, it has been a pleasure meeting you this last year. Your PR vibes brought so many fun moments. I wish you all the best with your theatrical adventures and look forward to finally meeting you again. To the **maruja limón** group: **Libs**, my dear, you really have seen me grow (since I was a naive, 16 year old). You have such a unique personality and style, thank you for all the warmth that you and **Thijs** gave me on the first days in the Netherlands, and for all our musical and cultural outings we've had since then. **Luís,** I've always been so attracted by your fun mood. I'm

glad we crossed paths here in the city and have shared so many moments together. Thank you for your attempts to read my scientific papers at a party and for all the fun moments :). I'd like to also thank the countless encounters in the city, sharing laughs and moments with **Pablo**, **Javi**, **Lara**, **Flor, Sol, Jess** and many more! Thank you for bringing the sunshine on the gray days.

Thank you to all the best classmates I could have ever imagined, with whom I've shared not only passion but also trips and lots of friends. **Elena**, over these years we've shared so many moments, so many (bus) trips, so many laughs, and even a roof, thank you so much for your countless support here in the city. **Vicky**, I'm so happy that despite the andalusian vibes didn't last long in Amsterdam we've been able to meet so many times in Barcelona. I'm glad to see how happy you are with **Peña** and look forward to keeping in touch over the years. And last, **Radi**, our favorite Bulgarian, I miss you and your fun vibes a lot around Amsterdam, you showed us how we bioinformaticians can be very cool.

To the Uilenstede crowd: we all started lost and confused in a foreign country and made it through it. I still keep in my heart all the times we spent at unit 11 back in the naive days. To the wise **Julia**, the first one to begin with an "adult life" while still keeping all the joy, fun and laughs about it, thank you for always being so funny! **Javier**, you've always been so unique and mysterious at the same time. I'm so thankful for all the travels, outings (and missed attempts for outings when the energy was low), all the knowledge you brought to us and the moments that we spent together. To the *malaka* **Yannis**, it's been fun to keep connecting over the years, to see how our goals have changed, and see that we're still around the best city in the world :p I wish you all the best with your PhD! To **Janne**, it's been so nice seeing you coming back from time to time to Amsterdam, thank you for bringing your unique vibes to the group. And to my favorite Italians: **Sara** and **Fede**, sometimes so similar and sometimes so different. It's been a while since we've seen each other and I miss the time we spent together.

I'd also like to thank the countless people doing cute things in Amsterdam. This city has been a vibrant place where I feel at home and not judged. Without the great inspiration that I've seen during the day and night it wouldn't have been possible. Thank you to the **International queer and migrant film festival** team, it has been a cool opportunity to help out and expand my network. Also **Yao**, I'm so glad we met – I wish you the best with your African adventures. To the teachers that have motivated me: countless thanks to **The Living room** - Dutch ballroom scene team – I feel my boundaries are less boundaries thanks to your work and motivation.

Jamie, you helped me out a long time during the 'thesis time'. I appreciate the time we spent together, and I wish you the best in the UK – you helped me grow and you'll always occupy the best memories I could think of.

To raise the importance of online connections: back in the days (>14 years ago!) before smartphones existed, I had the opportunity to meet a bunch of people online in a forum and gladly we still keep in touch with each other. To all the **Skinianos**, thank you for being there, you are the most unique group of people I've ever met, with such a cool mixture of personalities and interests. **Merka**, **Oct**, **Libs**, **Xandra**, **Ester**, **Nusky**, **Bea**, **Juls, Coco**, **Tarik**, **Ceci**. Also, **Octavio**, thank you so much for all your work designing the cover from this book <3.

By the end of this thesis I had several encounters on the "other side of the world" which helped me grow and reflect on the future opportunities. I'd like to thank all the connections that have helped me grow in such a short time and I would like to especially thank **Sebas** for the time spent together.

Then, back to my roots. Barcelona has been a vibrant city to grow, and I'm glad to still keep along the years of connections that I value so much. **Xandra**, fuiste una de mis primeras aliadas y hemos compartido tantos momentos, tantas risas, tantos conciertos (y tantos gritos), tantas locuras, gracias por estar siempre ahí y por tantos momentos compartidos y por seguir siendo amis a pesar de estar tan lejos. Al dúo de **Zaida** y **Rubén**, siempre habéis estado ahí con vuestro humor, hemos también compartido muchos conciertos, mucha inspiración cultural, muchas risas, comidas, muchísimas gracias por siempre estar tan cerca (junto a Leia :p). Me gustaría también agradecer a toda la gente con quien compartí tantas salidas, conciertos y festivales: **Pacheco**, **Álex**, **Montse**, **Bitta**, **Fanny**, os deseo lo mejor :).

A mis amigos del cole, los del barrio, con los que hemos compartido tantos momentos, y crecido juntos con nuestras similitudes y diferencias. **Raquel**, siempre has tenido una personalidad tan única, una energía tan vibrante, casi inagotable :p gracias por aguantar tantos años juntos. **Omar**, han sido muchos momentos, muchas salidas juntos, gracias por mantener siempre el equilibrio entre lo calmado y lo loco. **Cristian**, siempre tan presente en el barrio, muchas gracias por todo el tiempo compartido y me alegro mucho tu proyecto con **Aida**, os deseo lo mejor. **Silvia**, ya no estamos ninguno de los dos en Barcelona y cada vez es más difícil vernos por la distancia, pero me alegra que aún así mantengamos la amistad durante los años. Gracias por traer siempre tantos momentos de alegría junto a **Irene** :).

També a les meves amigues que em van acollir al BBK, les de "**maldita ****** la nuestra"**, que tot i que sen's fa tan difícil veurens des de que vaig marxar, guardo un molt bon record

de tot el que hem vivit junts. Moltes gràcies **Elena**, per tots els dies (i bueno, millor dit, nits) junts, per tants concerts passat junts, per tantes xerrades i rises. I a la **Laura**, sempre amb tanta energía, amb tanta passió, tan única, la que encara m'intriga com s'ho fa per fer tantes coses ahora

I al millor grup de l'ICFO, **Mariona**, **Maria Bea**, **Luís**, **Ximo**, per tot els moments de diversió que hem passat.

Ahora toca agradecer el apoyo de mi familia durante todos estos años. En especial, me gustaría agradecer a mi **Mama** y mi **Papa** por todo el apoyo incondicional y amor durante estos años, me siento privilegiado por toda la ayuda que me habéis dado especialmente durante mi mudanza tan lejos de casa, gracias por confiar en mis decisiones. Me gustaría agradecer también a toda mi familia por ser quienes son, por estar siempre ahí para los buenos y malos momentos, por tal mezcla de personalidades, y por todas las historias y momentos vividos, gracias a mi querido **Abuelo** que siempre seré su cazurrete, a mis tías **Merche, Esther, Ana, Ángeles** que siempre me han dado tantos consejos, a mi tío **Chal** que siempre ha sido tan inspirador, a mis primos **César, Adrià,** y **Pau,** al **Kiko,** a mis queridas **Yaya** y **Tieta** que siempre han estado ahí para todo**,** a la **Mireia, Nuri, Juan,** a los que por desgracia ya no están con nosotros, **Abuela**, **Yayo**, y a toda la familia extendida, que son muchísimos. Por último me gustaría agradecer a la "familia postiza", los que siempre han estado durante los años, gracias **Susana** por estar ahí desde que era un niño, **Gely**, **Miguel**, **Ana**, **Albert**, **Judith**, **Edu**, **Andrea**, **Aida**.

Trying to remember every person that has made a significant contribution to my growth is challenging. I apologize in advance if someone is missing.