Multiple Omics Profiling of the PDAC Tumor Immune Microenvironment
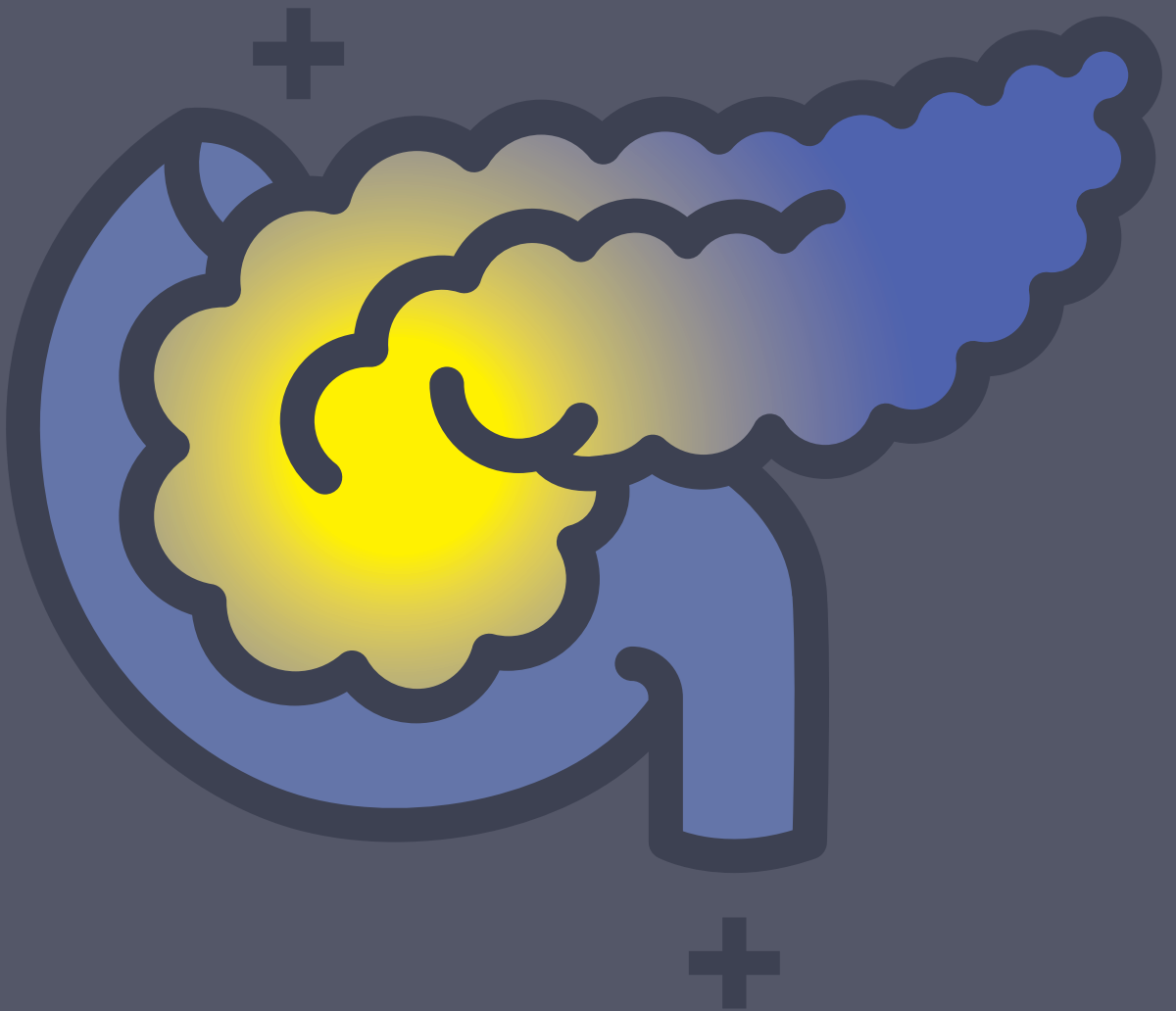
Thesis by
Willem de Koning

Multiple Omics Profiling of the PDAC Tumor Immune Microenvironment

Willem de Koning

Multiple Omics Profiling of the PDAC Tumor Immune Microenvironment

Meerdere omics-profilering van de PDAC tumor immuun micro-omgeving

Thesis
to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus
Prof. Dr. A.L. Bredenoord
And in accordance with the decision of the Doctorate Board.

The public defence shall be held on
Tuesday 20 June 2023, at 15.30hrs
by
Willem de Koning
Born in Woerden, The Netherlands

SUPPORT CASPER

Erasmus University Rotterdam

Doctoral Committee

Promotor:	Prof. Dr. C. H. J. van Eijck
Other members:	Prof. Dr. J.W. Wilmink
	Prof. Dr. F.J. van Kemenade
	Dr. M. Summer-Kutmon
Copromotors:	Dr. A.P. Stubbs
	Dr. D.A.M. Mustafa

Table of Contents

Chapter 1
Introduction and outline of this thesis

PANCREATIC DUCTAL ADENOCARCINOMA

Pancreatic cancer arises when cells in the pancreas, a glandular organ behind the stomach, begin to multiply out of control and form a mass. Pancreatic cancer is an aggressive cancer with an overall survival rate of 11%, the lowest survival rate of all cancers [1], where pancreatic ductal adenocarcinoma (PDAC) is the most common type of pancreatic cancer [2]. PDAC develops within the exocrine compartment of the pancreas. Signs and symptoms of PDAC may include yellow skin, abdominal or back pain, unexplained weight loss, light-colored stools, dark urine, and loss of appetite [3]. Usually, no symptoms are seen in the early stages, and symptoms that are specific enough to suggest pancreatic cancer typically do not develop until the disease has reached an advanced stage and has often spread to other parts of the body [4]. The number of patients with PDAC continues to increase and is expected to become the second leading cause of cancer-related deaths by the year 2040 [5]. Screening the general population has not been found to be effective. The 20% of the patients that are diagnosed with localized PDAC are still eligible for surgical resection with curative intent. For the other 80%, the tumor has either advanced locally (LAPC) or has metastasized to other parts of the body. These patients receive chemotherapy consisting of fluorouracil, leucovorin, irinotecan, and oxaliplatin (FOLFIRINOX) sometimes in combination with radiotherapy. However, not all patients respond to FOLFIRINOX, and the tumor will continue growing or develop metastasis. The reason for the lack of response to treatment in PDAC patients can partially be blamed on the heterogeneity of the tumor [6]. PDAC is known to have a tumor environment (TME) that stimulates tumor growth, invasion, and metastasis [7]. The TME is a complex system made up of many different cell types, the secreted products of these cell types, such as cytokines and chemokines, and other non-cellular extracellular matrix components [8].

THE TUMOR IMMUNE MICROENVIRONMENT OF PDAC

The TME of PDAC is characterized by a dense stromal architecture, regulating processes such as tumor growth, drug responsiveness, and metastasis [9]. This physically dense structure prevents the infiltration of immune cells and traditional therapies, like FOLFIRINOX. Furthermore, the TME consists of pancreatic stellate cells (PSCs), cancer-associated fibroblasts (CAFs), immune cells, extracellular matrix (ECM), and other immune-suppressive molecules. PSCs are a unique type of cells that exist in normal pancreatic tissue [10]. These PSCs differentiate into CAFs, which play a significant role in the TME of PDAC by producing ECM and their interaction with tumor and other stromal cells. PDAC has a unique fibrotic TME composed of ECM proteins produced by CAFs, also known as the desmoplastic stroma. This ECM encloses various cell types, including different immune cells such as lymphocytes, macrophages, mast cells, and myeloid-derived suppressor cells (MDSCs), along with endothelial and neuronal cells[11]. Tumor-infiltrating lymphocytes (TILs) are often observed in resected cancer tissue and are believed to participate in the host immune response against cancer [12]. The heterogeneity of the tumor-infiltrating immune cells within the TME plays a key role in the tumor's response to therapy [13]. A heterogeneous tumor immune microenvironment (TIME) within PDAC has access to many different tumor-promoting aspects and can thus be responsible for the poor survival rate of PDAC [14]. In the past, many omics studies (proteomics, transcriptomics, etc.) have been utilized to study the TIME of PDAC.

ANTIBODY-BASED PROFILING: A RAPID AND SIMPLE METHOD TO DETERMINE THE COMPOSITION OF THE PDAC TIME

Antibody-based profiling can target proteins that are expressed by specific immune cell types. Furthermore, it can detect proteins that are secreted by a cell, such as cytokines or chemokines, which are involved in cell signaling and transcriptional regulation. Thus, Antibody-based profiling can give us an insight into the composition and functionality of cells in the PDAC TIME. The currently most used methods are immune histochemistry (IHC), flow cytometry, and mass spectrometry.

Immunohistochemistry (IHC): Since the first production of reagents against typhus and cholera using a red stain in 1934, the techniques used for immunohistochemistry have been improved immensely [15]. Nowadays, in the clinical diagnostic field pathologist can classify tumors based on their expression of lineage-specific markers, oncogenic somatic mutations, and epigenetic modifications [16,17]. IHC provides a rapid and simple method to determine the origin of neoplastic tissue or investigate the behavior or progression of a given neoplasm [18]. In short, IHC is performed on formalin-fixed paraffin-embedded (FFPE) tumor slides. Antibodies are used in conjunction with a coloring dye to visualize and detect the antigen of interest in the tissue section. Usually, manual scoring is performed to evaluate the location and intensity of the markers [19]. Standardization by automated IHC machines has improved reproducibility and enabled high throughput. Instead of manual scoring sophisticated algorithms are implemented for automatic quantification. Both the standardization of the machines and implementation of the algorithms reduce biases [20]. Further, multiplex IHC provides the possibility to stain more proteins. This way a more complete overview of the TIME can be achieved. Multiplex IHC enables the researcher to study functionally active subsets of immune cells or co-localization of antigens on tumors or immune cells.

Flow cytometry: This technique delivers rapid identification and quantification of multiple cell types in a heterogeneous sample [21]. Since its invention in the 1960s, flow cytometry has been widely applied to characterize and quantify immune cells [22]. Flow cytometry is a technology that utilizes lasers to produce light signals that are converted into electronic signals, which can be analyzed with a computer. Based on the fluorescent or light-scattering characteristics, cell populations can be identified and quantified [23]. It allows for the simultaneous characterization of mixed populations of cells from blood and bone marrow as well as solid tissues that can be dissociated into single cells such as lymph nodes, spleen, mucosal tissues, solid tumors, etc. In addition to the identification and quantification of populations of cells, a major utility of flow cytometry is sorting cells into uniform populations to be used for further downstream analysis [23]. The automation systems for sample preparation and sample acquisition provide a more reproducible platform and increase the throughput of processed samples. These developments

have enabled high-resolution quantification of cell types, analysis of cell surface and intracellular molecules, immunophenotyping analysis, functional characteristics of different cell populations, and drug discovery for a variety of diseases [24,25].

Mass spectrometry (MS): This is a method to detect and quantify changes in the proteome of clinical samples [26]. Since the foundation of mass spectroscopy in 1898, mass spectrometry has been used to identify chemical substances. MS traces ions through an ion source, analyses the ion beam according to their mass-to-charge ratio, and uses detectors that are capable of measuring/recording the currents of the beams. This principle makes it possible to characterize T cells, the tumor-infiltrating lymphocytes (TILs) as well as biofluids [27]. MS is emerging as a tool for rapid diagnostics.

Spatial proteomics: Over the last few years, spatial proteomics has been developed to study the spatial architecture of tissues. Spatial proteomics provides a technology for the visualization of proteins in their native cellular environment without the need for cell lysis or the physical separation of compartments or organelles before proteomic analysis [28]. This allows for studying the spatial distribution of the immune cells in the TIME. Currently, there are several spatial profiling approaches. The approaches are based on antibodies linked to photocleavable oligonucleotide tags (GeoMx® Protein Assays), use an agonistic approach by detection of isotope-labeled antibodies with mass spectrometry (PhenoImager™), or use a combination of DNA-conjugated antibodies and multicyclic addition of complementary fluorescently labeled DNA probes (PhenoCycler™-Fusion). Using spatial proteomics to study the TIME of PDAC could enhance our understanding IHC of the spatial composition and interactions between cells in a high-throughput manner.

TRANSCRIPTOMICS: FROM BULK TO A SINGLE-CELL DESCRIPTION OF THE PDAC TIME

Transcriptomics technologies are used to study ribonucleic acids (RNA), which have diverse cellular and biological functions. The measurement of RNAs gives information on how genes are regulated and gives a perspective on the molecular activity in cells. Thus, transcriptomics can give us an insight into the activity of cells in the PDAC TIME. Bulk transcriptomics and targeted gene transcriptomics are the two major approaches for transcriptome studies.

Bulk transcriptomics: The PDAC TIME can be described from bulk transcriptomics data. Bulk omics studies are important to describe tumor pathogenesis and for the identification of different phenotypes. The main downside of using bulk omics data is that this data represents an average expression of the tumor, ignoring possible intratumor heterogeneity [29]. Bulk transcriptomics gives us a general understanding of what is going on in PDAC. Tools have been developed to identify the abundance of immune cell populations within the TIME. In heterogeneous tissue samples, transcriptomic measurements average signals originating from the individual cells. The deconvolution of these signals can yield estimates of cell population proportions in samples [30]. CIBERSORT is a method for enumerating cell compositions from bulk transcriptomics [31]. By using cell type expression signatures CIBERSORT allows you to measure intra-sample proportions of immune cells from bulk RNA admixtures without physical isolation. MCP-counter calculates an abundance estimate per cell population. This allows you to perform inter-sample comparisons [32]. The development of long-read RNA sequencing enabled the sequencing of full-length cDNA transcripts without clonal amplification or transcript assembly [33]. Therefore, amplification and read mapping steps introduce fewer false positive splice sites and can detect transcript isoforms. Furthermore, the introduction of single-cell sequencing has made it possible to profile the transcriptome at single-cell resolution [34]. Although single-cell transcriptomics introduced a method to study the intra-sample differences, it cannot capture spatial information [35].

Targeted gene expression: Using targeted gene expression, genes can be measured based on probes or primers. Conventional PCR can be used to detect the number of PCR products generated. In reverse transcription PCR (RT-PCR), RNA is first reverse transcribed into cDNA and then amplified. This allows us to measure the amount of RNA. Later real-time quantitative RT-qPCR was developed, which counts how many reaction cyclers are used to reach an amplification slope. It consists of four steps, (1) reverse transcription of the mRNA template, (2) primer pair annellation to the target cDNA template and a DNA polymerase extends from the ligated primer, (3) thermal cycler varies the temperature settings to perform the three stages of the PCR: denaturation, annealing, and extension, (4) a standard dissociation curve can be studied to determine the initial amount of RNA (cDNA) transcripts [36]. After PCR an array technology, microarray, was developed in the late 1990s. Microarrays have been used for the quantification of

RNA (and DNA) using fluorescence. Fluorescent-labeled RNA hybridizes on the surface of the slide and the absolute intensity of the hybridization signal is measured [37]. Recently the nCounter® technology was developed by NanoString Technologies, which is a variation on the microarrays. It employs two ~50 base probes per mRNA that hybridize in solution. The mRNA is immobilized by the capture probe, while the reporter probe carries the signal. This allows for the detection of the color codes by the Digital Analyzer. The Digital Analyzer detects the color-coded probes and captures the signal intensities as gene expression mRNA levels. Reverse transcription of the RNA to cDNA is not needed, as the probes anneal directly to the RNA [38]. As earlier described CIBERSORT and MCP-counter can be used to enumerate cell compositions from bulk transcriptomics and are therefore also applicable to microarray data [32,39]. To estimate the abundance of immune cell populations in nCounter data, marker genes are used [40,41].

Spatial transcriptomics: A new development in transcriptomics is the addition of spatial information. Spatial transcriptomics, like spatial proteomics, uses technologies to visualize the cellular environment [42]. Techniques such as GeoMx[43] and CosMx by NanoString or Visium by 10X Genomics were developed. These three techniques stain slides with morphological markers. GeoMx creates high-quality images with single-cell visibility. Regions of interest are selected and illuminated. These regions are collected in wells and sequenced. CosMx takes this a step further by generating detailed cell segmentations and can perform sequencing on single-cell level. Visium permeabilizes the tissue to release RNA which binds to adjacent capture probes which can be sequenced. Using spatial transcriptomics to study the TIME of PDAC could further enhance our understanding of the spatial composition and interactions between cells in a high-throughput manner.

In this thesis, we investigated the use of multiple omics techniques in tissue samples to determine the heterogeneous PDAC TIME. Furthermore, with these techniques, we analyzed peripheral blood samples of PDAC patients to identify alterations in the immune cell profile during chemotherapy.

OUTLINE OF THIS THESIS

After an introduction in **Chapter one,** where we describe, among other things, different multiple omics techniques, in **chapter two** we describe a method to identify different subtypes of immune cells specifically in PDAC tissue samples with the use of marker genes. In **chapter three**, blood samples collected at baseline and after the first FOLFIRINOX cycle of 68 patients were used to develop an immune gene signature from the targeted gene expression measurements to predict response after a single cycle of FOLFIRINOX. **Chapter four** presents a study that investigates the safety of adding IMM-101 to SBRT and the immuno-modulatory effects of the combination treatment in the peripheral blood of locally advanced pancreatic cancer PDAC patients. Targeted gene-expression profiling and multicolor flow cytometry were performed for longitudinal immune monitoring of the peripheral blood. **Chapter five** describes the identification of the peripheral immune cell profile alterations after one cycle of FOLFIRINOX in PDAC patients using the gold standard, flow cytometry, and further explore these results with targeted gene expression profiling. In **chapter six** we discuss the pros and cons of using RNA-sequencing for fresh frozen samples and targeted gene expression immune profiles for formalin-fixed, paraffin-embedded samples to characterize the tumor immune microenvironment. The results presented in this thesis and a future perspective are discussed in **chapter seven**. A summary of the work is given in **chapter eight**, and detailed information about the author can be found in **chapter nine**. In **chapter ten** everybody who was involved in the process of creating this thesis was thanked.

REFERENCES

1. Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2022. *CA Cancer J Clin* **72**, 7–33 (2022).

2. Kleeff, J. *et al.* Pancreatic cancer. *Nature Reviews Disease Primers 2016 2:1* **2**, 1–22 (2016).

3. PDQ Adult Treatment Editorial Board. Pancreatic Cancer Treatment (Adult) (PDQ®): Patient Version. *PDQ Cancer Information Summaries* (2002).

4. Longo, D. L., Ryan, D. P., Hong, T. S. & Bardeesy, N. Pancreatic Adenocarcinoma. *https://doi.org/10.1056/NEJMra1404198* **371**, 1039–1049 (2014).

5. Rahib, L., Wehner, M. R., Matrisian, L. M. & Nead, K. T. Estimated Projection of US Cancer Incidence and Death to 2040. *JAMA Netw Open* **4**, e214708–e214708 (2021).

6. Samuel, N. & Hudson, T. J. The molecular and cellular heterogeneity of pancreatic ductal adenocarcinoma. *Nat Rev Gastroenterol Hepatol* **9**, 77–87 (2011).

7. Murakami, T. *et al.* Role of the tumor microenvironment in pancreatic cancer. *Ann Gastroenterol Surg* **3**, 130 (2019).

8. Schiavoni, G., Gabriele, L. & Mattei, F. The tumor microenvironment: A pitch for multiple players. *Front Oncol* **3 APR**, 90 (2013).

9. Neesse, A. *et al.* Stromal biology and therapy in pancreatic cancer. *Gut* **60**, 861–868 (2011).

10. Erkan, M. *et al.* StellaTUM: current consensus and discussion on pancreatic stellate cell research. *Gut* **61**, 172–178 (2012).

11. Pothula, S. P., Pirola, R. C., Wilson, J. S. & Apte, M. v. Pancreatic stellate cells: Aiding and abetting pancreatic cancer progression. *Pancreatology* **20**, 409–418 (2020).

12. Lohneis, P. *et al.* Cytotoxic tumour-infiltrating T lymphocytes influence outcome in resected pancreatic ductal adenocarcinoma. *Eur J Cancer* **83**, 290–301 (2017).

13. Bindea, G. *et al.* Spatiotemporal Dynamics of Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer. *Immunity* **39**, 782–795 (2013).

14. Grünwald, B. T. *et al.* Spatially confined sub-tumor microenvironments in pancreatic cancer. *Cell* **184**, 5577-5592.e18 (2021).

15. Marrack, J. Nature of Antibodies. *Nature 1934 133:3356* **133**, 292–293 (1934).

16. Huang, T. *et al.* Detection of histone H3 K27M mutation and post-translational modifications in pediatric diffuse midline glioma via tissue immunohistochemistry informs diagnosis and clinical outcomes. *Oncotarget* **9**, 37112–37124 (2018).

17. Gajjar, A., Pfister, S. M., Taylor, M. D. & Gilbertson, R. J. Molecular Insights into Pediatric Brain Tumors Have the Potential to Transform Therapy. *Clinical Cancer Research* **20**, 5630–5640 (2014).

18. Painter, J. T., Clayton, N. P. & Herbert, R. A. Useful Immunohistochemical Markers of Tumor Differentiation. *Toxicol Pathol* **38**, 131 (2010).

19. Kim, S. W., Roh, J. & Park, C. S. Immunohistochemistry for Pathologists: Protocols, Pitfalls, and Tips. *J Pathol Transl Med* **50**, 411–418 (2016).

20. Rozowsky, J. S. *et al.* A Toolkit for Profiling the Immune Landscape of Pediatric Central Nervous System Malignancies. *Front Immunol* **13**, 1448 (2022).

21. Young, Y. K., Bolt, A. M., Ahn, R. & Mann, K. K. Analyzing the tumor microenvironment by flow cytometry. *Methods in Molecular Biology* **1458**, 95–110 (2016).

22. Fulwyler, M. J. Electronic Separation of Biological Cells by Volume. *Science (1979)* **150**, 910–911 (1965).

23. McKinnon, K. M. Flow Cytometry: An Overview. *Curr Protoc Immunol* **120**, 5.1.1-5.1.11 (2018).

24. Maiques, O., Georgouli, M. & Sanz-Moreno, V. Recent advances in tissue imaging for cancer research. *F1000Research 2019 8:1980* **8**, 1980 (2019).

25. Ding, M. & Baker, D. Recent advances in high-throughput flow cytometry for drug discovery. *Expert Opin Drug Discov* **16**, 303–317 (2021).

26. Macklin, A., Khan, S. & Kislinger, T. Recent advances in mass spectrometry based clinical proteomics: applications to cancer research. *Clinical Proteomics 2020 17:1* **17**, 1–25 (2020).

27. Urbiola-Salvador, V., Miroszewska, D., Jabłońska, A., Qureshi, T. & Chen, Z. Proteomics approaches to characterize the immune responses in cancer. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1869**, 119266 (2022).

28. Lundberg, E. & Borner, G. H. H. Spatial proteomics: a powerful discovery tool for cell biology. *Nature Reviews Molecular Cell Biology 2018 20:5* **20**, 285–302 (2019).

29. Li, Y., Ma, L., Wu, D. & Chen, G. Advances in bulk and single-cell multi-omics approaches for systems biology and precision medicine. *Brief Bioinform* **22**, 1–18 (2021).

30. Shen-Orr, S. S. & Gaujoux, R. Computational deconvolution: extracting cell type-specific information from heterogeneous samples. *Curr Opin Immunol* **25**, 571–578 (2013).

31. Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods 2015 12:5* **12**, 453–457 (2015).

32. Becht, E. *et al.* Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol* **17**, 1–20 (2016).

33. Schadt, E. E., Turner, S. & Kasarskis, A. A window into third-generation sequencing. *Hum Mol Genet* **19**, (2010).

34. Tang, F. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* **6**, 377–382 (2009).

35. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science (1979)* **353**, 78–82 (2016).

36. Gadkar, V. J. & Filion, M. New developments in quantitative real-time polymerase chain reaction technology. *Curr Issues Mol Biol* **16**, 1–6 (2014).

37. de Mello-Coelho, V. & Hess, K. L. A conceptual and practical overview of cDNA microarray technology: implications for basic and clinical sciences. *Brazilian Journal of Medical and Biological Research* **38**, 1543–1552 (2005).

38. Tsang, H. F. *et al.* NanoString, a novel digital color-coded barcode technology: current and future applications in molecular diagnostics. *https://doi.org/10.1080/14737159.2017.1268533* **17**, 95–103 (2016).

39. Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M. & Alizadeh, A. A. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol* **1711**, 243 (2018).

40. Danaher, P. *et al.* Gene expression markers of Tumor Infiltrating Leukocytes. *J Immunother Cancer* **5**, 18 (2017).

41. de Koning, W. *et al.* Identification, Validation, and Utilization of Immune Cells in Pancreatic Ductal Adenocarcinoma Based on Marker Genes. *Front Immunol* **12**, 1449 (2021).

42. Marx, V. Method of the Year: spatially resolved transcriptomics. *Nature Methods 2021 18:1* **18**, 9–14 (2021).

43. Merritt, C. R. *et al.* Multiplex digital spatial profiling of proteins and RNA in fixed tissue. *Nat Biotechnol* **38**, 586–599 (2020).

Chapter 2
Identification, validation, and utilization of immune cells in pancreatic ductal adenocarcinoma based on marker genes

W. de Koning[1,2], D. Latifi[3], Y. Li[1,2], C.H.J. van Eijck[3], A.P. Stubbs[1], D.A. Mustafa[2]

[1]Unit of Clinical Bioinformatics, Department of Pathology, Erasmus University Medical Centre, Rotterdam, The Netherlands

[2]Tumor Immuno-Pathology Laboratory, Department of Pathology, Erasmus University Medical Centre, Rotterdam, The Netherlands

[3]Department of Surgery, Erasmus University Medical Centre, Rotterdam, The Netherlands

ABSTRACT

The immune response affects tumor biological behavior and progression. The specific immune characteristics of pancreatic ductal adenocarcinoma (PDAC) can determine the metastatic abilities of cancerous cells and the survival of patients. Therefore, it is important to characterize the specific immune landscape in PDAC tissue samples, and the effect of various types of therapy on that immune composition. Previously, a set of marker genes was identified to assess the immune cell composition in different types of cancer tissue samples. However, gene expression and subtypes of immune cells may vary across different types of cancers. The aim of this study was to provide a method to identify immune cells specifically in PDAC tissue samples. The method is based on defining a specific set of marker genes expressed by various immune cells in PDAC samples. A total of 90 marker genes were selected and tested for immune cell type-specific definition in PDAC; including 43 previously used, and 47 newly selected marker genes. The immune cell-type specificity was checked mathematically by calculating the "pairwise similarity" for all candidate genes using the PDAC RNA-sequenced dataset available at The Cancer Genome Atlas. A set of 55 marker genes that identify 22 different immune cell types for PDAC was created. To validate the method and the set of marker genes, an independent mRNA expression dataset of 24 samples of PDAC patients who received various types of (neo)adjuvant treatments was used. The results showed that by applying our method we were able to identify PDAC-specific marker genes to characterize immune cell infiltration in tissue samples. The method we described enabled identifying different subtypes of immune cells that were affected by various types of therapy in PDAC patients. In addition, our method can be easily adapted and applied to identify the specific immune landscape in various types of tissue samples.

Keywords: pancreatic ductal adenocarcinoma, marker genes, immune cells, immune microenvironment, mRNA expression

INTRODUCTION

Pancreatic cancer is one of the deadliest diseases with a 5-year survival rate of 9% [1]. The most prevalent neoplastic disease of the pancreas is pancreatic ductal adenocarcinoma (PDAC) [2]. Failure of treatment is partially due to the high heterogeneity of the disease [3]. The interaction between cancer and immune cells, known as the immune microenvironment (TME), leads to diverse mechanisms of immune evasion [4]. The abundance and composition of tumor-infiltrating lymphocytes (TILs) are fundamental to tumor immunogenicity [5,6]. The variety of TILs and their interaction with pancreatic cancer cells influence tumor progression [7]. During the early stages of tumor development, immune cells such as natural killer (NK) and CD8+ T cells facilitate the destruction of immunogenic cancer cells [8]. As the tumor evolves, different immune cells infiltrate and have an impact on the tumor's fate. For instance, high infiltration of CD4+ T cells correlates with a good prognosis [9], while high infiltration levels of regulatory T cells (Tregs) correlate with a poor prognosis [10]. In addition, TME and TILs influence the survival of PDAC patients. The high levels of CD8/Tregs ratio correlate with longer survival of the patients [11]. Taken together, the accurate determination of the immune infiltration in PDAC tissue samples is important because it provides valuable information regarding how the host immune response interacts with cancer cells. This information can be used in guiding the immunomodulatory approaches to treat PDAC patients.

The gold standard to identify and quantify immune cells in blood samples is Flow cytometry. Immune cells in the blood samples do not need enzymatic disassociation and they can be detected relatively easily after binding to antibodies. However, immune cells in fixed tissue samples, like Fresh-Frozen (FF) or Formalin-Fixed, Paraffin-Embedded (FFPE) samples, are more difficult to quantify by flow cytometry. The methods and enzymes used to dissociate cells in tissue severely harm membranous antigens making it more challenging to bind to the antibodies. The preferred method to use for tissue samples is immunohistochemistry (IHC) which showed to be clinically useful [12]. However, relatively a lot of tissue sections are needed to measure only a few immune markers. The recent development of this technique enabled multiplexing measurements of various antibodies using one section sample [13]. Nevertheless, the number of immune cells that can be identified using IHC-based techniques is still limited and dependent on the availability and accuracy of the antibodies. Alternatively, gene expression profiling is a promising and clinically applicable method for measuring the diversity of TILs in FFPE samples. Various techniques can be used to measure the gene expression profiles of tissue samples. Most of the techniques are based on using enzymatic reactions to synthesize cDNA and amplify it, prior to measuring the expression of the genes or sequence of the fragments of RNA. However, the targeted gene expression measurements using nCounter® technology (NanoString) enable counting the copies of RNA fragments of tissue samples directly without any enzymatic reactions or amplification steps. It facilitates detecting low abundance targets, down to 0.1–0.5 fM RNA targets, with high sensitivity and high reproducibility ($R^2 > 0.98$) [14,15]). These features enable determining the immune cell

repertoires in FFPE samples. Moreover, many mRNA expression profile databases are available online and can be re-analyzed either to identify the immune cells in a specific cancer type or to validate the findings of a specific analysis. However, an accurate method to identify the TILs based on gene expression per cancer subtype is needed.

To accurately estimate the abundance of the various immune cell populations within the TME, a set of marker genes is needed for each cell type. Previously, a set of immune-specific marker genes were identified to determine cell type across various types of cancer [16,17]. However, gene expression levels are highly affected by the type of tumor. In addition, the marker genes used to identify immune cells may differ in various types of cancer. The aim of this study was to identify a set of marker genes that can be used to characterize the immune landscape in PDAC tissue samples. To that aim, we selected a set of candidate genes (PDAC-cMG), then checked their accuracy to identify immune cells in PDAC samples. Genes that passed our definition criteria were chosen to create the set of marker genes to identify immune cells in PDAC tissue samples (i.e., PDAC-MGICs). To demonstrate the utility of PDAC-MGICs, we applied them to evaluate the effect of therapy on the immune cell infiltration between PDAC patient groups that have been treated with a combination of surgery and neoadjuvant therapy.

MATERIALS AND METHODS

Selecting PDAC candidate marker genes (PDAC-cMG)

A set of PDAC-specific candidate marker genes (PDAC-cMG) were selected based on genes that were previously used to identify immune cells across different types of cancer (n=43) [16]. The PDAC-cMG gene list was enriched by genes that were found to identify immune cells in the literature (n=47) (**Table 1, Column 2**). A total of 90 candidate genes were included in the PDAC-cMG, representing 23 immune cell types.

Table 1 Summary of the candidate gene set and the selected marker genes used to identify immune cell types in PDAC tissue samples.

Column 1	Column 2	Column 3	Column 4
Cell type	**Candidate marker genes (PDAC-cMG)**	**Selected marker genes (PDAC-MGICs)**	**Default marker genes used in the nSolver® Advanced Analysis**
B cells	**BLK**[16] **BLNK**[16] **CCR9**[16] **CD19**[16] **CD22**[18] **CD24**[19] **CR2**[16] **HLA-DOB**[16] **HLA-DQA1**[16] **MEF2C**[16] **MS4A1**[16]	BLK, CD19, CD22, CR2, MS4A1	BLK, CD19, MS4A1, TNFRSF17, FCRL2*, KIAA0125*, PNOC*, SPIB*, TCL1A*
Plasma B cells	**CD27**[20] **CD38**[21] **SLAMF7**[22] **TNFRSF17**[22]	*CD27 CD38 SLAMF7 TNFRSF17*	
Regulatory B cells	**CD1D**[23] **CD5**[23] **IL10**[23]	*CD1D CD5*	
Cytotoxic cells	**GZMA**[16] **GZMB**[16] **GZMH**[16] **KLRB1**[16] **KLRD1**[16] **KLRK1**[16] **PRF1**[16] **CTSW**[16] **GNLY**[16]	GZMA, GZMB, GZMH, KLRB1, KLRD1, KLRK1, PRF1	GZMA, GZMB, GZMH, KLRB1, KLRD1, KLRK1, PRF1, CTSW, GNLY, NKG7*
Dendritic cells	**CCL13**[16] **CD1A**[24] **CD1C**[24] **CD209**[16] **HSD11B1**[16]	*CD1A CD1C*	CCL13, CD209, HSD11B1
Conventional Dendritic cells 1	**BTLA**[24] **XCR1**[24] **DPP4**[24] **THBD**[24]	*BTLA XCR1*	
Conventional Dendritic cells 2	**CD2**[24] **ITGAM**[24] **ITGAX**[24]	*ITGAM ITGAX*	
Plasmacytoid Dendritic cells	**CLEC4C**[24] **IL3RA**[25] **NRP1**[24]		
Macrophages	**CD68**[16] **FCGR2A**[26]	CD68, FCGR2A	CD163, CD68, CD84, MS4A4A*
Antigen-presenting cells	**CCR7**[27] **CD80**[28] **CD86**[26]	*CD80 CD86*	
M2 Macrophages	**CD163**[29] **CD36**[27] **MRC1**[27]	*CD163 MRC1*	
Mast cells	**C2**[30] **CMA1**[16] **CTSG**[16] **FCER1A**[30] **MS4A2**[16] **PLAU**[30] **TPSAB1**[16]	MS4A2, TPSAB1	MS4A2, TPSAB1, CPA3*, HDC*, TPSB2*

Column 1	Column 2	Column 3	Column 4
Monocytes	**CD14**[27] **CD33**[30] **TLR2**[31]	*CD14 CD33 TLR2*	
Natural Killer cells	**NCR1**[16] **XCL2**[16]	NCR1	NCR1, XCL1*, XCL2
Natural Killer CD56+ dim cells	**IL21R**[16] **KIR3DL1**[16]	KIR3DL1	IL21R, KIR3DL1, KIR2DL3, KIR3DL2
Neutrophils	**CSF3R**[16] **FCGR3A**[30] **S100A12**[16]	CSF3R, FCGR3A	CEACAM3*, CSF3R, FCAR*, FCGR3B*, FPR1*, S100A12, SIGLEC5*
T cells	**CD3D**[16] **CD3E**[16] **CD3G**[16] **CD6**[16] **SH2D1A**[16]	CD3D, CD3E, CD3G, CD6, SH2D1A	CD3D, CD3E, CD3G, CD6, SH2D1A, TRAT1*
CD4+ T cells	**CD4**[32] **SELL**[25]	*CD4 SELL*	
CD8+ T cells	**CD8A**[16] **CD8B**[16]	CD8A, CD8B	CD8A, CD8B
Exhausted CD8+ T cells	**CD244**[16] **HAVCR2**[33] **LAG3**[16] **PDCD1**[33] **TIGIT**[33]	LAG3, PDCD1, TIGIT	CD244, EOMES, LAG3, PTGER4
Helper 1 T cells	**TBX21**[16] **ALCAM**[25] **CD70**[16]	TBX21	TBX21
Regulatory T cells	**FOXP3**[16] **CD274**[34] **IDO1**[35] **IL2RA**[30] **TNFRSF18**[30]	FOXP3, IL2RA	FOXP3
CD45+	PTPRC[16]	PTPRC	PTPRC

Downloading data from The Cancer Genome Atlas (TCGA)

The gene expression profiling data of pancreatic adenocarcinoma (PAAD) from the TCGA database (Level 3 RSEM-normalized, Illumina RNA-seq, Version2) was downloaded [36]. The TCGA PAAD dataset is filtered for patients with PDAC primary tumors that received no treatment prior to surgery (n=147). The expression data were log2-transformed prior to pairwise similarity calculations.

Calculating the pairwise similarity

The pairwise similarity statistic between all pairs in the PDAC-cMG per cell type was calculated using an adaptation of Pearson's correlation metric. The adapted Pearson's correlation metric was proven to be better than the simple Pearson correlation [16]:

$$similarity(x, y) = \frac{\sum (x - \bar{x})(y - \bar{y})}{\frac{(n-1)}{2}(var(x) + var(y))}$$

The log2-transformed vectors of the gene expression of two genes are denoted by x and y, where the sample means are denoted by and . The sample variance is indicated by var (x) and var (y). This adaptation takes the slope of the correlation into account; hence two ideal cell type marker genes would have a similarity of 1. All calculations were completed using R version 4.0.3.

Identifying immune cells in PDAC data

The pairwise similarity for all 90 candidate genes was calculated, and genes with high pairwise similarity (≥ 0.6) were selected to be included per cell type in the PDAC-MGICs (**Table 1, Column 3**) [16]. Each immune cell type was represented by at least two unique genes included in the PDAC-MGICs [37–40]. The specificity of the selected gene markers was confirmed by creating heatmaps showing the pairwise similarity of all selected marker genes per immune cell type.

Validating the immune cell marker genes using published PDAC profile data

Previously published data (Farren et al., 2020, GEO accession: GSE129492) from 6 PDAC patients who received no systemic therapy prior to surgery (i.e. Surgery Only) were used to validate that PDAC-MGICs are robust and valid immune marker genes in other PDAC cohorts. The database was created by measuring the PanCancer Immune profiles panel. It contained gene expression profiles of 730 immune-related genes and 40 housekeeping genes measured by using the nCounter® platform of NanoString technology (Platform GPL19965). The expression level of the 55 genes of the PDAC-MGICs set was checked and confirmed to be higher than the detection threshold in at least 50% of the samples. The gene expression was normalized and log2-transformed using nSolver® (version 4.0) and the Advanced Analysis module (version 2.0) of NanoString technology (NanoString, Seattle, WA, USA). The mean pairwise similarities for the PDAC-MGICs were calculated following the same method that was described earlier.

Concordance of the new PDAC-MGICs

We validated the marker genes' concordance by calculating *p*-values for the cell type gene sets as implemented in the nSolver® Advanced Analysis module (version 2.0). The null hypothesis that a given gene set exhibits no greater cell type-specific-like behavior than a randomly selected gene set of similar size was tested. Therefore, the concordance was calculated for each cell type (i.e., a metric of a gene set's adherence to the assumption of cell type-specific and consistent expression):

$$\text{concordance}(X) = \frac{1}{trace\big(Cov(X)\big)} \left(p^{-\frac{1}{2}}, ..., p^{-\frac{1}{2}}\right) Cov(X) \left(p^{-\frac{1}{2}}, ..., p^{-\frac{1}{2}}\right)^T$$

The matrix of log2-transformed expression values of the gene set for a specific cell type is denoted by X, and p is the number of genes. The concordance function returns 1 if all genes are perfectly correlated with a slope of 1 and degrade to 0 as this pattern weakens. This concordance is compared to the concordance of 1000 random gene sets of size p, denoted by X0'. The *p*-value equals the proportion of concordance (X0') values greater than concordance(X0), where concordance(X0) is the concordance of the selected

marker genes. The concordance of the gene markers was compared to the default gene markers in the nSolver® Advanced Analysis module.

Validation using PDAC samples affected by various types of neoadjuvant therapy

The performance of PDAC-MGICs for samples affected by treatments is tested for 18 samples of patients that were subjected to different neoadjuvant therapy prior to surgery: 6 patients received FOLFIRINOX chemotherapy, 6 patients received FOLFIRINOX + stereotactic body radiotherapy (SBRT), and 6 patients received FOLFIRINOX + conventional radiotherapy (XRT) [41]. The samples were matched based on lymphovascular invasion and perineural invasion. The database was created by measuring the PanCancer Immune profiles panel. It contained gene expression profiles of 730 immune-related genes and 40 housekeeping genes measured by using the nCounter® platform of NanoString technology (GEO accession: GSE129492). Gene expression profiles were normalized and log2-transformed using nSolver® and the Advanced Analysis module, and the pairwise similarity and concordance were calculated as described earlier.

To identify immune cells within the nSolver® advanced analysis module, genes that are annotated to define immune cells within the probe annotation file (provided by NanoString) were changed. After that, the modified probe annotation file was uploaded to a new analysis file, and the average of all genes confirmed to identify a specific immune cell was calculated resulting in a score of a cell type. The scores of cells were compared between the groups of interest, and the significance was calculated using a t-test between the groups.

Utilization of PDAC-MGICs

The clinical utility is demonstrated by uploading our defined marker genes PDAC-GMICs in the nSolver® Advanced Analysis module to identify immune cells in the "cell type profiling" section. The PDAC-GMIC set was used to assess the composition of the immune microenvironment for all patients in GSE129492 [41]. The abundance of the immune cell types is represented by a cell score which is the average log2-transformed expression value of their corresponding marker genes. To correct for the total tumor-infiltrating immune cells per patient, the abundance was calculated relatively to the CD45+ cells. The relative abundance of a cell type in a group of patients is the average log2-scale expression of the marker genes divided by the average log2-scale expression of CD45+. To demonstrate the impact of changing the definition of cells, the relative cell abundances based on the default marker genes of the nSolver® Advanced Analysis module were compared to the relative cell abundance based on PDAC-MGICs.

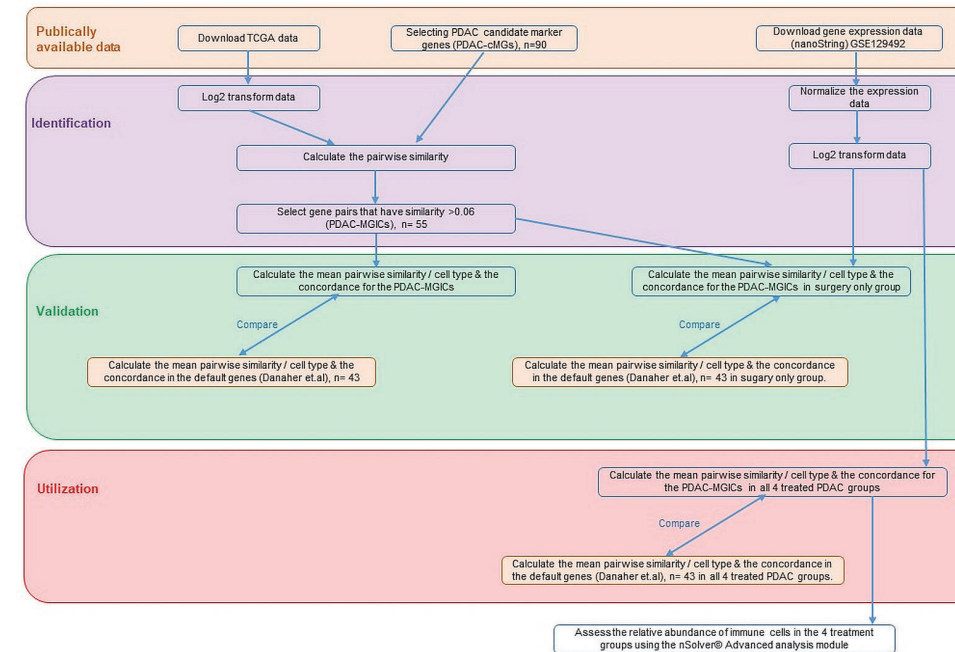An overview of our method is presented in (**Figure 1**).

Figure 1 An overview of the method used to determine the definition of immune cells in PDAC tissue samples.

RESULTS

PDAC-MGICs enabled identifying 22 immune cell types

The PDAC-cMGs with a pairwise similarity ≥ 0.6 were selected and summarized in (**Table 1, Column 3**). For example, 11 genes were chosen as candidate genes to define B cells. Calculating the pairwise similarity of all 11 genes showed that 5 genes had a high pairwise similarity (≥ 0.6), while the other 6 genes had a low pairwise similarity (< 0.6). Therefore, B cells were defined using the 5 genes with high pairwise similarity (**Figure 2**). Following the same method, a total of 55 genes were selected as marker genes to identify 22 immune cells in PDAC (**Supplementary file 1**).
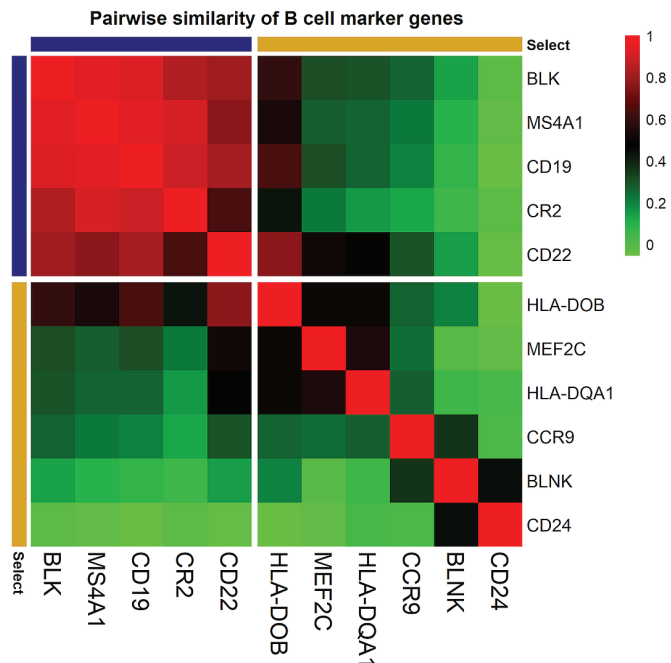

Figure 2 Correlation plot of the pairwise similarity of candidate marker genes tested to identify B cells. The pairwise similarity varies between the 11 selected genes. Five genes (blue) showed a high pairwise similarity (≥ 0.6). These genes were selected to identify B cell infiltration in PDAC tissue samples. Six genes (yellow) showed a low pairwise similarity (< 0.6). these genes were not used to identify B cell infiltration in PDAC tissue samples. The red color in the correlation plot presented the highest correlation score between the genes ($R^2 = 1$); the green color presented the lowest correlation score ($R^2 = 0$).

Enriching the PDAC-MGICs list with additional genes from the literature enabled identifying 8 additional immune cells that were not included in the default setting of the nSolver® Advanced Analysis module. These cells are plasma B cells, regulatory B cells (Bregs), 2 types of conventional dendritic cells (cDC), antigen-presenting cells (APCs), M2 macrophages, monocytes, and CD4+ T cells. In addition, the enrichment of the PDAC-MGICs increased the accuracy to identify Tregs cells, B cells, and macrophages,

because more genes were used to identify these cells as compared to the default settings. To validate the specificity of the selected markers, pairwise similarities were calculated across all marker genes. The results are shown as a correlation plot for all the marker genes (**Figure 3**). The highest correlation was achieved between marker genes that were used to identify a specific immune cell. However, a relatively high correlation was also seen across other types of cells. For example, the 5 genes that identified B cells showed the highest correlation for B cells. But they also showed a lower but still good correlation in identifying T cells. This highlights the importance of trusting marker genes that have the highest pairwise similarity to identify a specific cell type. Marker genes that were used to identify cDC2 showed a relatively high correlation in Monocytes and M2 macrophages, highlighting that the definition of cDC2 is challenging using gene expressions of the PanCancer Immune profile panel in PDAC tissue samples, and can be improved by selecting additional marker genes.
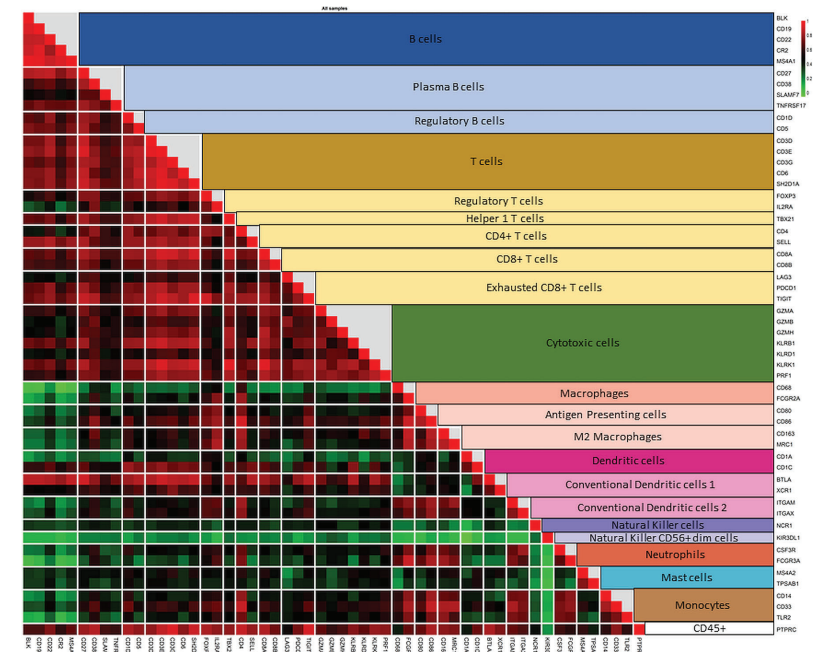

Figure 3 Correlation plot of the pairwise similarity of all 55 marker genes selected to identify the immune repertoire in the PDAC tissue sample. The pairwise similarity plot shows a high correlation between marker genes that identify a specific immune family and the subtype of that family. The highest correlation is shown between the marker genes that identify a specific type of immune cell. In addition, a relatively high correlation is shown between the subtypes of immune cells of the same family (B cells and various subtypes of B cells; T cells and various subtypes of T cells). The correlation between T cells and cytotoxic cells is lower than the other subtypes of T cells because cytotoxic cells include both T and NK cells. The correlation plot also shows a high pairwise similarity and a high specificity of marker genes that identify macrophages and their subtypes in PDAC tissue samples. However, the various types of dendritic cells (DCs)are more difficult to identify. Genes used to identify DCs show a good correlation with T cells and macrophages, highlighting the need to use other marker genes (not measured by the PanCancer Immune profile panel) to increase the accuracy of identifying DCs in PDAC tissue samples.

Validation of PDAC marker genes

To validate the accuracy of the PDAC-MGICs, the mean pairwise similarities between the corresponding marker genes were calculated in the TCGA PAAD dataset (**Table 2, Column 2**) and the Surgery Only dataset (**Table 2, Column 4**). These similarities were compared to those calculated between the cell types defined by the default gene markers in nSolver® Advanced Analysis (**Table 2, Columns 3 & 5**). Using the PDAC-MGICs resulted in an improved pairwise similarity (≥ 0.6) in both datasets for B cells, cytotoxic cells, DCs, neutrophils, and T cells. From the eight newly defined immune cells, five have a mean pairwise similarity ≥ 0.6 in both datasets. The exceptions were Bregs and the 2 types of cDCs. Furthermore, the concordance per cell type of the PDAC-MGICs in the Surgery Only was calculated (**Table 2, Column 6**) and was compared to the default gene markers (**Table 2, Column 7**). The *p*-value for concordance improved for all PDAC-MGICs

compared to the default markers in nSolver® except for macrophages, CD8+ T cells, and exhausted CD8+ T cells.

Validation of PDAC-MGICs in PDAC samples subjected to neoadjuvant therapy

The usability of the PDAC-MGICs was checked by calculating the pairwise similarity (**Table 2, Column 8**) and concordance (**Table 2, Column 10**) in 18 samples of patients that received neoadjuvant therapy prior to surgery and compared to the default gene markers (**Table 2, Column 9, 11**). Similar results to the Surgery Only group were achieved. An improvement of the pairwise similarity and concordance was shown for all PDAC-MGICs except for neutrophils that were not robustly identified in the default settings.

Cell type	Column 2 PDAC-MGICs mean pairwise similarity in the TCGA PAAD dataset	Column 3 Default marker genes mean pairwise similarity in the TCGA PAAD dataset	Column 4 PDAC-MGICs mean pairwise similarity in GSE129492 Surgery Only	Column 5 Default marker genes mean pairwise similarity in GSE129492 Surgery Only*	Column 6 PDAC-MGICs concordance in GSE129492 Surgery Only	Column 7 Default marker genes concordance in GSE129492 Surgery Only*	Column 8 PDAC-MGICs mean pairwise similarity in GSE129492 Neoadjuvant	Column 9 Default marker genes mean pairwise similarity in GSE129492 Neoadjuvant*	Column 10 PDAC-MGICs concordance in GSE129492 Neoadjuvant	Column 11 Default marker genes concordance in GSE129492 Neoadjuvant*
B cells	0.84	0.71	0.92	0.87	0.00	0.00	0.70	0.59	0.00	0.01
Plasma B cells	0.71		0.90		0.00		0.71		0.00	
Regulatory B cells	0.72		0.44		0.24		0.60		0.06	
Cytotoxic cells	0.7	0.64	0.59	0.54	0.00	0.01	0.43	0.46	0.01	0.01
Dendritic cells	0.7	0.48	0.80	0.19	0.04	0.43	0.37	0.08	0.21	0.58
Conventional Dendritic cells 1	0.68		0.53		0.18		0.71		0.02	
Conventional Dendritic cells 2	0.75		0.48		0.21		0.37		0.20	
Macrophages	0.55	0.64	0.49	0.55	0.23	0.07	0.68	0.59	0.03	0.01
Antigen-presenting cells	0.8		0.85		0.02		0.58		0.08	
M2 Macrophages	0.84		0.67		0.09		0.76		0.01	
Mast cells	0.73	0.76	0.67	0.67	0.09	0.10	0.59	0.59	0.08	0.06
Monocytes	0.71		0.62		0.04		0.48		0.06	
Natural Killer cells		0.42		0.25		1.00		0.45		0.18
Natural Killer CD56+ dim cells		0.26		0.37		0.70		0.28		0.13
Neutrophils	0.67	0.52	0.68	0.46	0.09	0.22	0.24	0.43	0.35	0.16
T cells	0.88	0.87	0.82	0.82	0.00	0.00	0.51	0.51	0.01	0.01
CD4+ T cells	0.61		0.66		0.10		0.48		0.14	
CD8+ T cells	0.86	0.86	0.07	0.07	0.58	0.546	0.73	0.73	0.02	0.02
Exhausted CD8+ T cells	0.68	0.43	0.54	0.59	0.11	0.074	0.21	0.06	0.35	0.56
Regulatory T cells	0.73		0.81		0.04		0.70		0.04	

Utilization of PDAC marker genes

The composition of the immune microenvironment for all samples published previously [41] was assessed by the nSolver® Advanced Analysis module (NanoString). The relative abundance of the immune cell types is compared between PDAC-MGICs and the default marker genes in the Surgery Only samples (**Figure 4**). Defining immune cells based on the PDAC-MGICs showed a significant effect in the relative scores of macrophages, neutrophils, natural killer cells (NK), and Tregs compared to the default settings (**Figure 4**). The effect of neoadjuvant therapy on the relative scores of cells was tested and shown in (**Figure 5**). Defining immune cells using PDAC-GMICs revealed that FOLFIRINOX + SBRT had the biggest effect on immune cells compared to the Surgery Only group. The results indicate an elevation of the cells as an effect of various types of neoadjuvant treatments, except for Bregs. The results are coherent with the previously reported results [41], but more immune subtypes were identified.
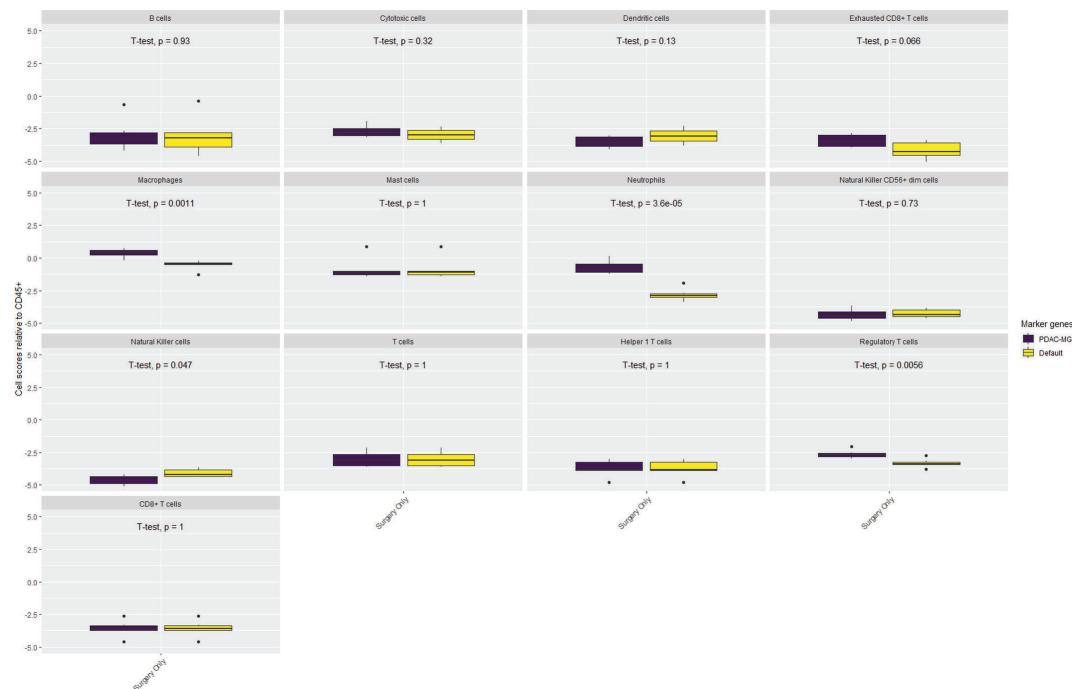

Figure 4 The impact of using PDAC-MGICs to identify immune cells in PDAC tissue samples. Comparing the relative immune scores using mRNA expression data of 6 tissue samples of patients who were subjected to surgery before receiving any treatment (Surgery Only). Immune cells were identified using the PDAC-MGICs set (purple) or the default marker genes in the nSolver® Advanced Analysis module of NanoString technology (yellow). All cell types were relative to the total infiltration of CD45+ expression. Identifying immune cells based on the PDAC-MGICs shows a significant variation (p-value < 0.05) in Macrophages, Neutrophils, Natural Killer cells, and Tregs cells.
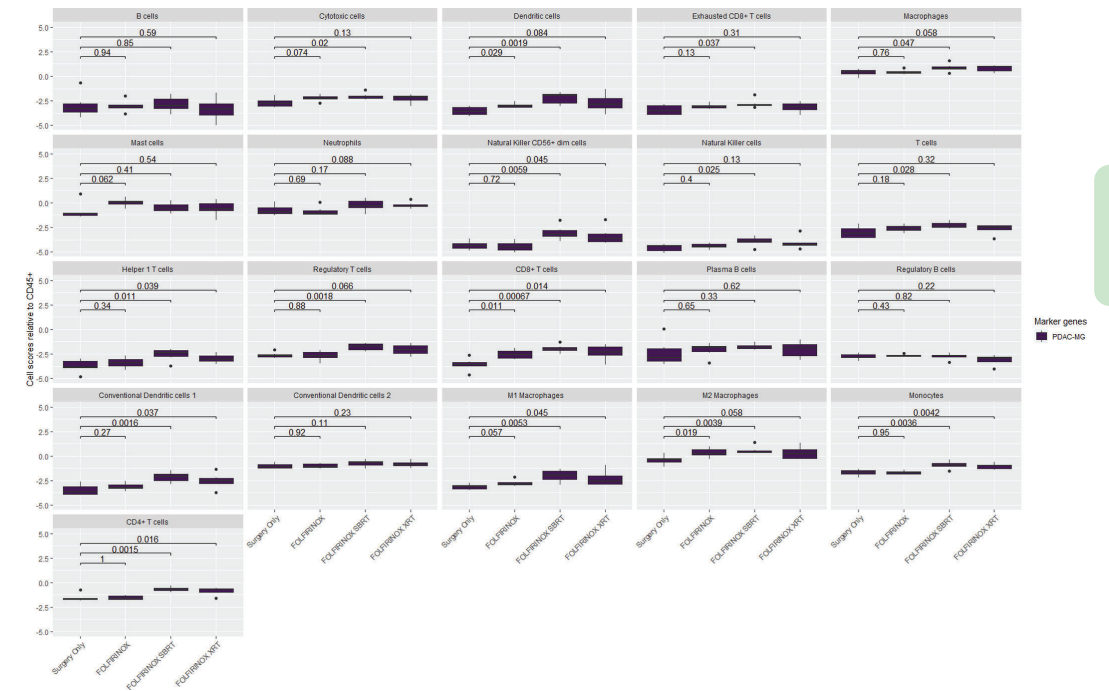

Figure 5 The relative immune abundance in PDAC tissue samples that received neoadjuvant therapy compared to treatment naïve samples using PDAC-MGICs. Comparing the relative immune scores using mRNA expression data of 18 PDAC tissue samples of patients who receive three types of neoadjuvant therapy compared to patients who were subjected to surgery before receiving any treatment (Surgery Only). Immune cells were identified using the PDAC-GMICs set and were presented relative to the total infiltration of CD45+ expression. The treatment effect of FOLFIRINOX + SBRT treated samples is most apparent. The p-values are the result of two-sided t-tests between Surgery Only and the other treatment groups individually. Surgery only (purple); neoadjuvant FOLFIRINOX (blue), FOLFIRINOX + SBRT (green), FOLFIRINOX + XRT (yellow).

DISCUSSION

We have identified and validated specific marker genes to define immune cells in PDAC tissue samples (PDAC-MGICs). The PDAC-MGCI are more PDAC-specific than the marker genes used to define immune cells across various types of tumor tissue samples (PanCancer marker genes). In addition, PDAC-MGICs enabled identifying eight additional immune cells (**Table 1**). To the best of our knowledge, our method is the only PDAC-specific method that enables identifying 22 immune cells from 730 genes only. Moreover, it is the only method to describe the effect of (neo)adjuvant therapy in all 22 immune cells of PDAC tissue samples.

The method we provided is adapted from the previously published method based on the mathematical calculation of the pairwise similarities between the marker genes [16]. Our method is based on using genes that are expressed above the background threshold. It differs in the number of genes used to identify immune cell types. We identified cells based on using at least two unique marker genes for each cell type. In addition, to increase the accuracy of cell definition, we chose a higher cut-off for the pairwise similarity (≥ 0.6) [42] compared to > 0.2 that was used in the previous publication. By increasing the threshold of the pairwise similarity, some important genes may not be used to identify an immune cell type. However, the accuracy of the identified immune cells will increase, which will be reflected in the time and money that will be spent on validating immune cells. The threshold can be adjusted to different levels in each experiment. We used ≥ 0.6 to identify immune cells with a high level of accuracy that will minimize the amount of future validation. The set of genes used to identify immune cells has been reported to be expressed by a specific type of immune cells and showed a similar pattern of expression in the PDAC database which increases the cumulative evidence to be included as a marker gene in the PDAC-MGIC. The identification of immune cells infiltrating the tumor is very important to understand the underlying mechanisms of tumor immunogenicity [5,6]. While the previously described PanCancer marker genes [16] can give a comprehensive understanding of the relative immune cells' abundance in various types of tumor tissue samples, it is highly important to check the pairwise similarity in a given database to ensure the accurate definition of immune cells. This importance becomes clear by checking DCs. The previously reported marker genes for DCs are shown to be insufficient for PDAC samples in contrast to pan-cancer samples (**Table 1**). Incorporating the PDAC-MGIC in nSolver® Advanced Analysis software enabled discovering the effect of neoadjuvant therapy on the immune profiling of PDAC tissue samples. Our method showed the same effect of neoadjuvant therapy in PDAC samples as was reported before [41]. However, it highlighted more clearly that the addition of a radiotherapy regimen to FOLFIRINOX induces more profound changes in gene expression than FOLFIRINOX alone. This was reflected in the relative scores of B cells, exhausted CD8+ T cells, macrophages, and neutrophils. The same types of cells had similar scores comparing the Surgery Only group to the FOLFIRINOX group, (**Supp Figure S1-S5**). Taken together, the results indicate that the addition of radiotherapy is necessary to stimulate immune cell infiltration in PDAC patients.

It should be noted that our method should be used to describe the relative scores of immune cells in two or more groups of samples, but it does not support estimating the absolute number of immune cells. Using a gene expression-based method to identify immune cells does not allow distinguishing between the number and the activity of cells. In addition, the definition of immune cells based on using one marker gene only like NK cells, NK CD56+ dim cells, and Helper 1 T cells, or cells that showed pairwise similarity < 0.6, remains not very accurate. However, in this study, we showed that the pairwise similarity is consistent between different databases (**Table 2**). Few exceptions were found, for example in CD8+ T cells, which highlights the huge effect of neoadjuvant therapy on the expression of genes that identify CD8+ T cells. CD8+ T cells were identified by using very specific and accurate genes: CD8A and CD8B genes. Therefore, the results reflect the effect of neoadjuvant therapy on the relative scores of CD8 cells. A recent publication described the immune landscape by estimating 22 different immune cells in PDAC samples using CIBERSORTx [43]. The immune estimation was correlated to the molecular subtypes and the survival of the patients. Interestingly, the number of estimated immune cells was the same as we identified in our method. However, the immune subtypes do not completely overlap (**Supp Table S1**). In the study of Liu et al. [43], immune cells were computed by using the LM22 gene signature containing 547 genes as a reference. Opposite to our method, genes are not mutually exclusive. Although an assumption is made by using mutually exclusive genes, our method can be used to estimate the relative abundance of 22 immune cells using 55 genes only. Furthermore, all marker genes described in our method are specifically measured in PDAC tissue samples, contrary to the LM22 gene signature reference. In addition, our method can be applied using gene expression data generated from samples that were preserved differently like FF and FFPE tissue samples or blood samples. Identifying PDAC-specific immune cells using the PDAC-MGICs enables revealing the effect of any type of therapy in various clinical settings and clinical trials. Moreover, applying the method to data generated from blood samples supports monitoring the progression of patients, and may be informative to direct therapeutic decisions.

Our method is easily tailored and applicable to identify specific immune cells in any type of tissue sample. Nevertheless, we highlight the importance of selecting and testing the marker genes critically for each tissue type. It has been shown that marker genes can be specified for each type of tissue sample [17]. The candidate marker gene list can be checked and narrowed down to a more specific marker gene list by calculating the pairwise similarity between all pairs of genes to ensure the accurate identification of immune cells in any type of tissue sample. Once the marker genes for each immune cell have been identified and checked, the expression of the genes for each immune cell type can be compared between the groups of interest. This method can be applied using any RNA databases (sequencing or gene expressions). The use of single-cell sequencing has shown

that cells of the same type can have different gene expressions present [44]. Furthermore, the assumption that the gene markers are exclusively expressed by one specific cell type is in many cases hard to prove. Therefore, we believe that the described method can accurately estimate the relative score of immune cells based on their marker genes definition.

CONCLUSION

We provided a method to identify specific immune cells in PDAC tissue samples based on using mRNA expression of marker genes (PDAC-MGICs). In addition, we validated and utilized the PDAC-MGICs to delineate the effect of various (neo)adjuvant treatments on the immune landscape in PDAC tissue samples. The PDAC-MGICs set reflects the immune microenvironment of the PDAC tumor tissue sample, however, it can be easily tailored and applicable to identify specific immune cells in any type of tissue sample.

REFERENCES

1. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2019. *CA Cancer J Clin* **69**, 7–34 (2019).

2. Kleeff, J. *et al.* Pancreatic cancer. *Nat Rev Dis Primers* **2**, 16022 (2016).

3. Bai, X. *et al.* Characteristics of Tumor Infiltrating Lymphocyte and Circulating Lymphocyte Repertoires in Pancreatic Cancer by the Sequencing of T Cell Receptors. *Sci Rep* **5**, 13664 (2015).

4. Karamitopoulou, E. Tumour microenvironment of pancreatic cancer: immune landscape is dictated by molecular and histopathological features. *Br J Cancer* **121**, 5–14 (2019).

5. Gooden, M. J. M., de Bock, G. H., Leffers, N., Daemen, T. & Nijman, H. W. The prognostic influence of tumour-infiltrating lymphocytes in cancer: a systematic review with meta-analysis. *Br J Cancer* **105**, 93–103 (2011).

6. Boon, T., Cerottini, J.-C., Van den Eynde, B., van der Bruggen, P. & Van Pel, A. Tumor Antigens Recognized by T Lymphocytes. *Annu Rev Immunol* **12**, 337–365 (1994).

7. Li, K.-Y. *et al.* Pancreatic ductal adenocarcinoma immune microenvironment and immunotherapy prospects. *Chronic Dis Transl Med* **6**, 6–17 (2020).

8. Teng, M. W. L. *et al.* IL-12 and IL-23 cytokines: from discovery to targeted therapies for immune-mediated inflammatory diseases. *Nat Med* **21**, 719–729 (2015).

9. Fukunaga, A. *et al.* CD8+ Tumor-Infiltrating Lymphocytes Together with CD4+ Tumor-Infiltrating Lymphocytes and Dendritic Cells Improve the Prognosis of Patients with Pancreatic Adenocarcinoma. *Pancreas* **28**, e26–e31 (2004).

10. Hiraoka, N., Onozato, K., Kosuge, T. & Hirohashi, S. Prevalence of FOXP3+ Regulatory T Cells Increases During the Progression of Pancreatic Ductal Adenocarcinoma and Its Premalignant Lesions. *Clinical Cancer Research* **12**, 5423–5434 (2006).

11. Sideras, K. *et al.* Prognostic value of intra-tumoral CD8 + /FoxP3 + lymphocyte ratio in patients with resected colorectal cancer liver metastasis. *J Surg Oncol* **118**, 68–76 (2018).

12. Galon, J. *et al.* Cancer classification using the Immunoscore: a worldwide task force. *J Transl Med* **10**, 205 (2012).

13. Yanagita, E., Imagawa, N., Ohbayashi, C. & Itoh, T. Rapid Multiplex Immunohistochemistry Using the 4-antibody Cocktail YANA-4 in Differentiating Primary Adenocarcinoma From Squamous Cell Carcinoma of the Lung. *Applied Immunohistochemistry & Molecular Morphology* **19**, 509–513 (2011).

14. Geiss, G. K. *et al.* Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol* **26**, 317–325 (2008).

15. Veldman-Jones, M. H. *et al.* Evaluating Robustness and Sensitivity of the NanoString Technologies nCounter Platform to Enable Multiplexed Gene Expression Analysis of Clinical Samples. *Cancer Res* **75**, 2587–2593 (2015).

16. Danaher, P. *et al.* Gene expression markers of Tumor Infiltrating Leukocytes. *J Immunother Cancer* **5**, 1–15 (2017).

17. Hausser, J. *et al.* Tumor diversity and the trade-off between universal cancer tasks. *Nat Commun* **10**, 5423 (2019).

18. Clark, E. A. & Giltiay, N. V. CD22: A Regulator of Innate and Adaptive B Cell Responses and Autoimmunity. *Front Immunol* **9**, (2018).

19. Yakimchuk, K. Development and specific markers of T and B lymphocytes. *Materials and Methods* **6**, (2016).

20. Jung, J., Choe, J., Li, L. & Choi, Y. S. Regulation of CD27 expression in the course of germinal center B cell differentiation: the pivotal role of IL-10. *Eur J Immunol* **30**, 2437–2443 (2000).

21. Hammarlund, E. *et al.* Plasma cell survival in the absence of B cell memory. *Nat Commun* **8**, 1781 (2017).

22. Frigyesi, I. *et al.* Robust isolation of malignant plasma cells in multiple myeloma. *Blood* **123**, 1336–1340 (2014).

23. Mauri, C. & Menon, M. Human regulatory B cells in health and disease: therapeutic potential. *Journal of Clinical Investigation* **127**, 772–779 (2017).

24. Collin, M. & Bigley, V. Human dendritic cell subsets: an update. *Immunology* **154**, 3–20 (2018).

25. Charoentong, P. *et al.* Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep* **18**, 248–262 (2017).

26. Xue, J. *et al.* Transcriptome-Based Network Analysis Reveals a Spectrum Model of Human Macrophage Activation. *Immunity* **40**, 274–288 (2014).

27. Martinez, F. O., Gordon, S., Locati, M. & Mantovani, A. Transcriptional Profiling of the Human Monocyte-to-Macrophage Differentiation and Polarization: New Molecules and Patterns of Gene Expression. *The Journal of Immunology* **177**, 7303–7311 (2006).

28. Raggi, F. *et al.* Regulation of Human Macrophage M1–M2 Polarization Balance by Hypoxia and the Triggering Receptor Expressed on Myeloid Cells-1. *Front Immunol* **8**, (2017).

29. Hu, J. M. *et al.* CD163 as a marker of M2 macrophage, contribute to predict aggressiveness and prognosis of Kazakh esophageal squamous cell carcinoma. *Oncotarget* **8**, 21526–21538 (2017).

30. Franzén, O., Gan, L.-M. & Björkegren, J. L. M. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database* **2019**, (2019).

31. Schimunek, L. *et al.* Early decreased TLR2 expression on monocytes is associated with their reduced phagocytic activity and impaired maturation in a porcine polytrauma model. *PLoS One* **12**, e0187404 (2017).

32. Golubovskaya, V. & Wu, L. Different Subsets of T Cells, Memory, Effector Functions, and CAR-T Immunotherapy. *Cancers (Basel)* **8**, 36 (2016).

33. Wherry, E. J. & Kurachi, M. Molecular and cellular insights into T cell exhaustion. *Nat Rev Immunol* **15**, 486–499 (2015).

34. Moradi, B. *et al.* CD4+CD25+/highCD127low/- regulatory T cells are enriched in rheumatoid arthritis and osteoarthritis joints—analysis of frequency and phenotype in synovial membrane, synovial fluid and peripheral blood. *Arthritis Res Ther* **16**, R97 (2014).

35. Spranger, S. *et al.* Up-Regulation of PD-L1, IDO, and Tregs in the Melanoma Tumor Microenvironment Is Driven by CD8+ T Cells. *Sci Transl Med* **5**, 200ra116-200ra116 (2013).

36. Raphael, B. J. *et al.* Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell* **32**, 185-203.e13 (2017).

37. Parmigiani, G., Garrett, E. S., Anbazhagan, R. & Gabrielson, E. A statistical framework for expression-based molecular classification in cancer. *J R Stat Soc Series B Stat Methodol* **64**, 717–736 (2002).

38. Lee, H. K. Coexpression Analysis of Human Genes Across Many Microarray Data Sets. *Genome Res* **14**, 1085–1094 (2004).

39. Cheng, W.-Y., Yang, T.-H. O. & Anastassiou, D. Biomolecular Events in Cancer Revealed by Attractor Metagenes. *PLoS Comput Biol* **9**, e1002920 (2013).

40. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).

41. Farren, M. R. *et al.* Immunologic alterations in the pancreatic cancer microenvironment of patients treated with neoadjuvant chemotherapy and radiotherapy. *JCI Insight* **5**, (2020).

42. Miao, Y. *et al.* ImmuCellAI: A Unique Method for Comprehensive T Cell Subsets Abundance Prediction and its Application in Cancer Immunotherapy. *Advanced Science* **7**, 1902880 (2020).

43. Liu, R., Liao, Y.-Z., Zhang, W. & Zhou, H.-H. Relevance of Immune Infiltration and Clinical Outcomes in Pancreatic Ductal Adenocarcinoma Subtypes. *Front Oncol* **10**, (2021).

44. Szabo, P. A. *et al.* Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat Commun* **10**, 4706 (2019).

Chapter 3

A multigene circulating biomarker to predict the lack of FOLFIRINOX response after a single cycle in patients with pancreatic ductal adenocarcinoma

C.W.F. van Eijck[1,2], W. de Koning[2,3], F. van der Sijde[1,2], M. Moskie[1], B. Groot Koerkamp[1], M.Y.V. Homs[4], S.H. van der Burg[5], C.H.J. van Eijck[1,2], D.A.M. Mustafa[2].

[1]Department of Surgery, Erasmus University Medical Center, Rotterdam, The Netherlands

[2]Tumor Immuno-Pathology Unit, Department of Pathology, Erasmus University Medical Center, Rotterdam, The Netherlands

[3]Clinical Bioinformatics Unit, Department of Pathology, Erasmus University Medical Center, Rotterdam, The Netherlands

[4]Department of Medical Oncology, Erasmus University Medical Center, Rotterdam, The Netherlands

[5]Department of Medical Oncology, Oncode Institute, Leiden University Medical Center, Leiden, The Netherlands

ABSTRACT

Introduction: 5-fluorouracil, folinic acid, irinotecan and oxaliplatin (FOLFIRINOX) is promising in treating patients with pancreatic ductal adenocarcinoma. However, many patients and physicians are reluctant to start FOLFIRINOX due to its high toxicity and limited clinical response rates. In this study, we investigated the effect of a single FOLFIRINOX cycle, in combination with a granulocyte colony-stimulating factor, on the blood immune transcriptome of patients with pancreatic ductal adenocarcinoma. We aimed to identify an early circulating biomarker to predict the lack of FOLFIRINOX response.

Methods: Blood samples of 68 patients from all disease stages, who received at least four FOLFIRINOX cycles, were collected at baseline and after the first cycle. The response to treatment was radiologically evaluated following the Response Evaluation Criteria in Solid Tumours criteria 1.1. Targeted immune-gene expression profiling (GEP) was performed using NanoString technologies. To predict the lack of FOLFIRINOX response, we developed a FOLFIRINOX delta GEP (FFX-ΔGEP) score.

Results: A single FOLFIRINOX cycle significantly altered 395 genes, correlating to 30 significant alterations in relative immune cell abundances and pathway activities. The eight-gene (*BID, FOXP3, KIR3DL1, MAF, PDGFRB, RRAD, SIGLEC1* and *TGFB2*) FFX-ΔGEP score predicted the lack of FOLFIRINOX response with a leave-one-out cross-validated area under the curve (95% confidence interval) of 0.87 (0.60–0.98), thereby outperforming the predictiveness of absolute and proportional Δ carbohydrate antigen19-9 values.

Conclusions: A single FOLFIRINOX cycle, combined with granulocyte colony-stimulating factor, alters the peripheral immune transcriptome indisputably. Our novel FFX-ΔGEP is, to our knowledge, the first multigene early circulating biomarker that predicts the lack of FOLFIRINOX response after one cycle. Validation in a larger independent patient cohort is crucial before clinical implementation.

Keywords: Pancreatic ductal adenocarcinoma (PDAC), FOLFIRINOX chemotherapy, Lack of response, Blood immune transcriptome, FFX-ΔGEP score, Precision medicine

INTRODUCTION

Pancreatic ductal adenocarcinoma (PDAC) is one of the most lethal and aggressive solid malignancies with a poor prognosis [1,2]. The 5-year overall survival (OS) rate for all stages of PDAC is only 9% [3]. The poor prognosis is, among other things, related to the lack of distinctive symptoms, the lack of reliable biomarkers for early diagnosis, progressive metastatic spread, and the complex tumor (immune) microenvironment (TME) [4]. Surgical resection with or without chemotherapy is the only curative treatment for early-stage PDAC, but only 20% of the tumors are resectable at diagnosis, and more than 50% of patients present with metastatic disease [5–7].

The combined chemotherapeutic regimen of 5-fluorouracil, folinic acid, irinotecan, and oxaliplatin (FOLFIRINOX) is considered one of the most effective adjuvant chemotherapy and first-line treatment for patients with locally advanced (LAPC) and metastatic pancreatic cancer [8]. Multiple studies have demonstrated that FOLFIRINOX treatment is associated with prolonged OS compared to gemcitabine treatment in all stages of the disease [9–11]. A meta-analysis combining 11 studies reported improved OS in LAPC (24.2 months vs. 6-13 months) [9]; a multicenter, randomized, phase 2-3 trial reported improved OS in metastatic patients (11.1 months vs. 6.8 months) [10]; and a multicenter, randomized, phase 3 trial reported the most prolonged OS in patients with stage I-II or borderline resectable patients [11]. In addition, neoadjuvant FOLFIRINOX followed by surgical resection showed favorable outcomes in resectable pancreatic cancer patients [12]. Despite the generally improved FOLFIRINOX response rates, 25% of PDAC patients experience disease progression during treatment [10,13].

FOLFIRINOX treatment has also been associated with a higher incidence of toxicity-related events compared to gemcitabine treatment [10,14]. To prevent FOLFIRINOX-induced neutropenia, patients are frequently treated with a prophylactic granulocyte colony-stimulating factor (G-CSF) which stimulates granulocyte production in the bone marrow [15–18]. Currently, treatment response is evaluated through computed tomography (CT) imaging, but not until after four cycles of FOLFIRINOX. Exposure to ineffective but toxic treatment reduces patients' quality of life, carries unnecessary costs, and withholds patients from a potentially effective treatment. Hence, it is desirable to identify a biomarker that predicts the lack of response to FOLFIRINOX at an early stage. Carbohydrate Antigen 19-9 (CA19-9) is the only FDA-approved biomarker for routine management of PDAC [19] but has only been shown to predict FOLFIRINOX response after multiple cycles [20].

Several studies have demonstrated that oxaliplatin, 5-FU, and irinotecan can enhance tumor antigen presentation in poor immunogenic cancer types such as PDAC [21,22]. The increase in HLA-I and programmed death-ligand 1 (PD-L1) could synthesize the tumor for immune checkpoint (IC)-based immunotherapy and stimulates CD8[+] cytotoxic T lymphocyte activation [23]. Particularly oxaliplatin has been shown to induce immunogenic

cancer cell death and modulate the immune response, resulting in increased antigenicity and enhanced adaptive immune responses [24–26]. Additionally, oxaliplatin has been shown to elicit a systemic immune response against the tumor [27]. However, the peripheral immune effects of FOLFIRINOX have not been studied.

We hypothesized that the immune-modulating effects of FOLFIRINOX may be detectable in the peripheral blood after a single cycle of treatment. To test this hypothesis, we conducted targeted immune-gene expression profiling on the blood of PDAC patients. The aim of this study was to investigate the impact of a single FOLFIRINOX cycle on the peripheral immune transcriptome and to identify an early circulating biomarker predictive of the lack of FOLFIRINOX response in PDAC patients.

MATERIALS AND METHODS

Patient population

A total of 80 PDAC patients were included in this study. Patients were hospitalized at the Erasmus University Medical Centre Rotterdam between February 2018 and February 2021. Twenty-three patients with (borderline) resectable PDAC participated in the randomized clinical trial PREOPANC-2 (Dutch trial register NL7094), and 57 patients with locally advanced or metastasized PDAC participated in the prospective cohort study iKnowIT (Dutch trial register NL7522). Exclusion criteria were < 18 years of age, previous treatment with FOLFIRINOX, or co-treatment with another chemotherapeutic

Clinical procedure

Following histological confirmation of the primary tumor or metastases, patients were treated with FOLFIRINOX chemotherapy. All patients were prophylactically treated with the long-acting G-CSF lipegfilgrastim (Lonquex®; Teva Ltd, Petach Tikva, Israel), 24 hours after each cycle, to reduce FOLFIRINOX-induced neutropenia [18,28]. Two whole blood samples from each patient were collected: at baseline (immediately before the first cycle) and 14 days after the first but just before the second FOLFIRINOX cycle. As part of the standard clinical routine, serum CA19-9 concentrations were determined at the same time points using an enzyme-linked immunosorbent assay (ELISA). A patient's response to FOLFIRINOX was assessed based on a CT scan made after four cycles, following the Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 criteria (Figure 1) [29].
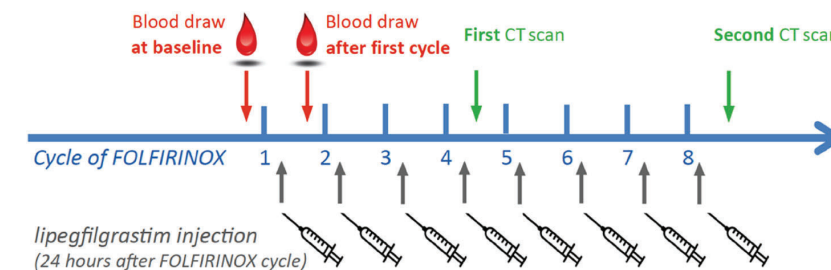

Figure 1 Schematic description of clinical procedure. Cycles of FOLFIRINOX chemotherapy (blue), lipegfilgrastim injection 24 hours after each cycle (grey), blood draw time points (red), and patient evaluation time points using CT scan (green).

Clinicopathological groups

To compare immune profiles, patients were grouped based on their clinicopathological characteristics. Disease stages at baseline included resectable, locally advanced pancreatic cancer (LAPC), and metastatic patients. Baseline CA19-9 values included patients with low (35-150 µmol/L) and high (> 1500 µmol/L) values. Patients who showed stable disease, partial response, or complete radiological response were defined as "disease

control". Patients showing disease progress were defined as "progressive disease". All patients that were radiologically evaluated received at least four cycles of FOLFIRINOX.

Whole blood sample collection and RNA isolation

Whole blood samples were collected in Tempus tubes (Applied Biosystems, Foster City, CA, USA) and stored at -80°C. Tempus tubes contain an RNA stabilizing reagent, which preserves the RNA quality and enables measuring gene expression profiles without isolating the peripheral blood mononuclear cells [30]. Total RNA was extracted from blood in Tempus tubes using the Tempus Spin RNA Isolation Kit of Thermo Fisher Scientific (Waltham, MA, USA) following the manufacturer's instructions. RNA quality control was done using the Agilent 2100 BioAnalyzer (Santa Clara, CA, USA). Samples with RNA concentrations less than 35 mg/mL were excluded. Corrected RNA concentrations were calculated based on the percentage of fragments of 300-4000 nucleotides to correct RNA degradation.

Targeted multiplex gene expression

Targeted gene expression profiling was performed using the nCounter® FLEX system and PanCancer Immune profiling panel, which includes 40 housekeeping genes and 730 immune-related genes [31]. A total of 200 ng RNA per sample in a maximum of 7 μL was used for hybridization, which was performed at 65°C for 17 hours using the SimpliAmp Thermal Cycler (Applied Biosystems). Gene expression was counted by scanning 490 Fields of View (FOV).

Data processing and analysis

Data quality control, normalization, and analysis were performed using the nSolver™ software (version 4.0) and the Advanced Analysis module (version 2.0) of NanoString Technology Inc. [32]. A patient's gene expression profile was included if all positive and negative control genes were within the expected values and if binding density values ranged between 0.5 and 3.0. Raw gene counts were normalized based on the most stable 34 housekeeping genes, identified by the geNorm algorithm [33], and all normalized data were $\log_2$ transformed. Genes were included when they were higher than the limit of detection of 4.384 $\log_2$, calculated as the average of all eight negative control genes multiplied by two, in > 80% of the gene expression profiles. Differentially expressed genes (DEGs) were identified using simplified negative binomial models, a mixture of negative binomial models, or log-linear models based on the convergence of each gene. Genes with a P-value < 0.05 after correction for multiple testing with the Benjamin-Hochberg (BH) method were considered DEGs.

Immune cell type analysis with the NanoString nSolver module

The peripheral abundance of various immune cell types was quantified using the nSolver Advanced Analysis module, which assigns relative immune cell type scores to each sample [34]. Marker genes, that identify specific immune cell types, were selected based on the pairwise similarities method tailored specifically for PDAC [35]. Marker genes were accepted to define an immune cell type when pairwise similarity was sufficient ($R^2 \geq 0.6$). Accordingly, the relative abundance of immune cells was calculated between the tested groups (Table S2).

Pathway analysis with the NanoString nSolver module and the Cytoscape plug-in ClueGO

Genes were clustered into predefined pathways using the nSolver Advanced Analysis module to examine immune-associated pathway alterations. We calculated the square root of the average squared t-statistic of all genes in the corresponding pathway [34], resulting in a pathway score for each sample. In addition, to explore the potential role of unique DEGs in disease control and progressive disease patients, we performed functional enrichment analysis using the Cytoscape plug-in ClueGO [36]. DEGs were included in the ClueGO analysis if they met two criteria: (1) a $\log_2$ fold-of-change (FOC) > |0.5| after a single FOLFIRINOX cycle and (2) a $\log_2$ FOC > |0.5| difference between disease control and progressive disease patients.

Statistical analysis

Statistical testing and data visualization were performed with R Statistical Software (v.4.1.2) [37]. Data were tested for normality with Shapiro-Wilk tests. We used paired or unpaired two-sided student t-tests for parametrical data and paired Wilcoxon tests or unpaired Mann-Whitney U tests for non-parametrical data. All tests were corrected with the BH correction for multiple testing. We used the R packages ggplot2 [38] and EnhancedVolcano [39] for data visualization.

The FOLFIRINOX delta gene expression profiling (FFX-ΔGEP) score

A gene signature representing an early predictive circulating biomarker of the lack of FOLFIRINOX response was identified (the FFX-ΔGEP score). Briefly, $\log_2$ normalized gene expression counts of baseline samples were subtracted from the $\log_2$ normalized gene expression counts of samples after a single FOLFIRINOX cycle, resulting in Δ expression counts for each gene. Genes that showed statistically significant differences (BH.P < 0.05) in Δ expression count between disease control and progressive disease patients were identified as candidate genes for the FFX-ΔGEP score. Patients were randomly split into training and test sets (75%/25%). To find the combination of candidate genes predicting the lack of FOLFIRINOX response most accurately, the least absolute shrinkage

and selection operator (LASSO) multivariate regression analysis was conducted on the training set with leave-one-out cross-validation. Weights (regression coefficient) were assigned to the candidate genes to improve model robustness and avoid overfitting. Genes weighted with a regression coefficient of 0 were excluded from the FFX-ΔGEP score. The fitted model was used in the corresponding test set to predict the lack of FOLFIRINOX response. The overall predictive performance was assessed by receiver operating characteristic (ROC) analysis depicting the area under the curve (AUC) value with a 95% confidence interval (CI).

The absolute (μmol/L) and proportional (%) change in CA19-9 was calculated to compare the predictive performance to the FFX-ΔGEP score. CA19-9 values of baseline samples were subtracted from those after a single FOLFIRINOX cycle to obtain absolute Δ CA19-9 values. The proportional Δ CA19-9 values were calculated by dividing the absolute Δ CA19-9 values by their baseline Δ CA19-9 values. ROC analysis was performed for both absolute and proportional Δ CA19-9 values, and the AUC value was compared to the AUC value of the FFX-ΔGEP score.

RESULTS

Samples and patient characteristics

Blood samples from 80 patients were collected at baseline and after the first cycle (Figure 1). RNA isolation was performed for all 160 blood samples. However, eight were excluded due to poor RNA concentration (< 35 mg/mL) and four were excluded due to poor binding density (< 0.5 or > 3.0). After removing corresponding pairs, 68 patients (136 samples) were included in the data analysis. The response to four FOLFIRINOX cycles was assessed by CT scan evaluation in 58 out of 68 patients, which resulted in 48 disease control and 10 progressive disease patients. No CT scan was performed in ten patients due to toxicity or early progression during FOLFIRINOX treatment. The overall survival [95% CI] for the disease control and the progressive disease patients was 40.2 months [32.7 − 46.5] and 13.8 months [11.2 − 15.5]. All clinicopathological characteristics are summarized in Table 1.

Table 1 Clinicopathological characteristics of patients in the study.

	All patients	Treatment response	
	(n = 68)	Disease control (n = 48)	Progressive disease (n = 10)
Age (y), mean (range)	65 (47 − 81)	65 (49 − 78)	60 (47 − 69)
Sex, no (%)			
Male	35 (51%)	25 (52%)	5 (50%)
Female	33 (49%)	23 (48%)	5 (50%)
Alcohol, no (%)			
Yes	35 (51%)	22 (46%)	3 (30%)
No	33 (49%)	26 (54%)	7 (70%)
Smoking, no (%)			
Yes	40 (59%)	28 (58%)	6 (60%)
No	28 (41%)	20 (42%)	4 (40%)
Diabetes Mellitus (DM), no (%)			
Yes	14 (11%)	11 (33%)	2 (20%)
No	54 (79%)	37 (77%)	8 (80%)
Disease stage, no (%)			
Resectable disease	21 (31%)	17 (35%)	2 (20%)
LAPC	28 (41%)	19 (40%)	5 (50%)
Metastatic disease	19 (28%)	12 (25%)	3 (30%)
Baseline CA19-9 (U/mL), no (%)			
Mean (± SD)	2919 (± 10893)	1352 (± 4182)	10503 (± 25859)

	All patients	Treatment response	
	(n = 68)	Disease control (n = 48)	Progressive disease (n = 10)
No expression (< 35)	13 (19%)	9 (19%)	3 (30%)
Low expression (35-150)	15 (22%)	13 (27%)	0 (0%)
Moderate expression (150-1500)	25 (37%)	19 (39.5%)	3 (30%)
High expression (> 1500)	15 (22%)	7 (14.5%)	4 (40%)
CA19-9 difference after a single cycle, compared to baseline			
Mean (U/mL) (± SD)	225.5 (± 1883)	128.1 (± 1777)	242.5 (± 2042)
Mean (%) (range)	15 (-62 − 304)	20% (-53 − 304)	5% (-62 − 49)
Baseline clinical parameters, mean (± SD)			
CEA (µg/L)	19.09 (± 49)	11.86 (± 24)	32.70 (± 69)
Bilirubin (µmol/L)	13 (± 12)	13 (± 8.0)	20 (± 23)
CRP (mg/L)	16 (± 24)	17 (± 25)	17 (± 24)
SII	1182 (± 1151)	1116 (± 1067)	1189 (± 713)
NLR	4.0 (± 2.8)	4.0 (± 3.0)	3.8 (± 1.8)
Total cycles of FOLFIRINOX, mean (± SD)	7.0 (± 3.0)	8.0 (± 2.0)	4.0 (± 2.0)
Median OS (months), mean (± SD)	11.7 (± 6.6)	13.0 (± 5.2)	8.27 (± 7.3)

Overall survival (OS) is defined as the difference in months (±SD) between the date of the first FOLFIRINOX cycle and the date of death. Abbreviations CT: Computed Tomography; LAPC: Locally Advanced Pancreatic Cancer; CA19-9: Carbohydrate Antigen 19-9; CEA: Carcinoembryonic Antigen; CRP: C-Reactive Protein; SII: Systemic Immune-inflammation Index; NLR: Neutrophil-to-Lymphocyte Ratio; SD: Standard Deviation; OS: Overall Survival.

PDAC patients with different disease stages or different *baseline CA19-19 values show comparable* immune profiles

Immune profiles based on the three disease stages (resectable, LAPC, metastatic) and based on the two baseline CA19-9 values (low and high) were compared at baseline and after a single FOLFIRINOX cycle. Baseline immune profiles revealed eight DEGs between the three disease stages and no DEGs between low and high baseline CA19-9 values (Figure S1). The pathway activity in baseline samples was not altered in any of the comparisons (BH.P > 0.05). Two immune cell types were relatively different between the three disease stages (BH.P < 0.05). Resectable patients showed relatively lower NK cells compared to LAPC and metastatic patients, and relatively lower conventional dendritic cells type 2 (cDC2s) compared to metastatic patients (BH.P < 0.05). No immune cell types were relatively different between the two baseline CA19-9 values (BH.P > 0.05; Figure S1).

A single FOLFIRINOX cycle altered the peripheral immune transcriptome of PDAC patients

Data analysis revealed 395 DEGs (BH.P < 0.01) in samples after a single FOLFIRINOX cycle compared to baseline samples (Figure 2A). Filtering the DEGs based on a $\log_2$ FOC ≥ |1.0| revealed 36 upregulated genes and three downregulated genes after a single FOLFIRINOX cycle (Figure 2B).
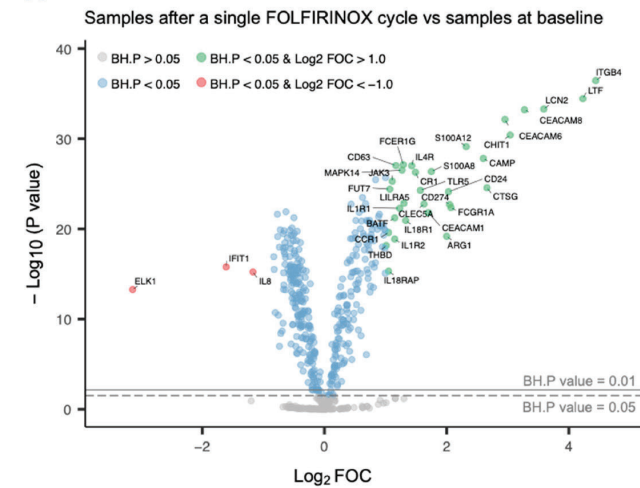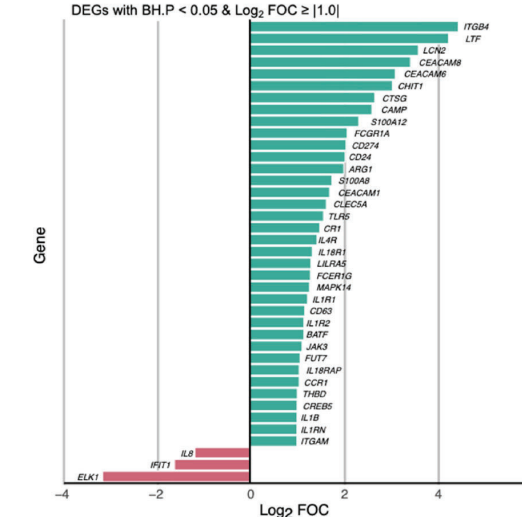

Figure 2 Identified DEGs after a single FOLFIRINOX cycle. A: Volcano plot of the identified DEGs using the paired analysis embedded in the Advanced Analysis module. Each dot is a gene, all DEGs genes (blue), upregulated DEGs with FOC < -1.0 in baseline samples (red), upregulated DEGs with FOC > 1.0 in samples after a single FOLFIRINOX cycle (green), and non-significant genes (grey). B: Waterfall plot displaying DEGs of BH.P < 0.05 & $\log_2$ FOC ≥ |1.0|. Upregulated DEGs with FOC < -1.0 in baseline samples (red), and upregulated DEGs with FOC > 1.0 in samples after a single FOLFIRINOX cycle (green). Abbreviations: BH.P: Benjamin-Hochberg P value; FOC: Fold Of Change.

Pathway analysis revealed alterations among all immune-associated pathways (BH.P < 0.001; Figure 3A-3C). Pathway-specific genes with $\log_2$ FOC ≥ |1.0| were considered key pathway drivers (Table S1). The pathways of adhesion, chemokines, cytokines, interleukins, macrophage function, pathogen defense, toll-like receptor (TLR), and tumor necrosis factor (TNF) superfamily were enhanced after a single FOLFIRINOX cycle while the immune-associated pathways of antigen processing, B cell, NK cell, and T cell functions, and cytotoxicity were diminished. Immune cell type analysis revealed alterations among all relative peripheral immune cell type abundances (BH.P < 0.05). The total immune cells (PTPRC, CD45$^+$), cDC2, monocytes, NK cells, and neutrophils increased while the B cells, cytotoxic cells, NK CD56^{dim} cells, total T cells, T regulatory (Treg) cells, and CD8$^+$ T cells decreased after a single FOLFIRINOX cycle (Figure 3D).

A single FOLFIRINOX cycle altered the expression of IC regulatory genes

The expression of the IC inhibitory genes *PDCD1* (PD-1), *CD274* (PD-L1), and *PDCD1LG2* (PD-L2) were upregulated after a single FOLFIRINOX cycle compared to baseline with an average $\log_2$ FOC of 1.1 (BH.P < 0.001). In contrast, the IC inhibitory genes *BTLA*, *CTLA4* and its ligand *CD86* (B7-2), *HAVCR2* (TIM-3), and *TIGIT* were statistically significantly downregulated (BH.P < 0.001; Figure 3E). The IC inhibitory gene *LAG3* was not altered (Figure S4).
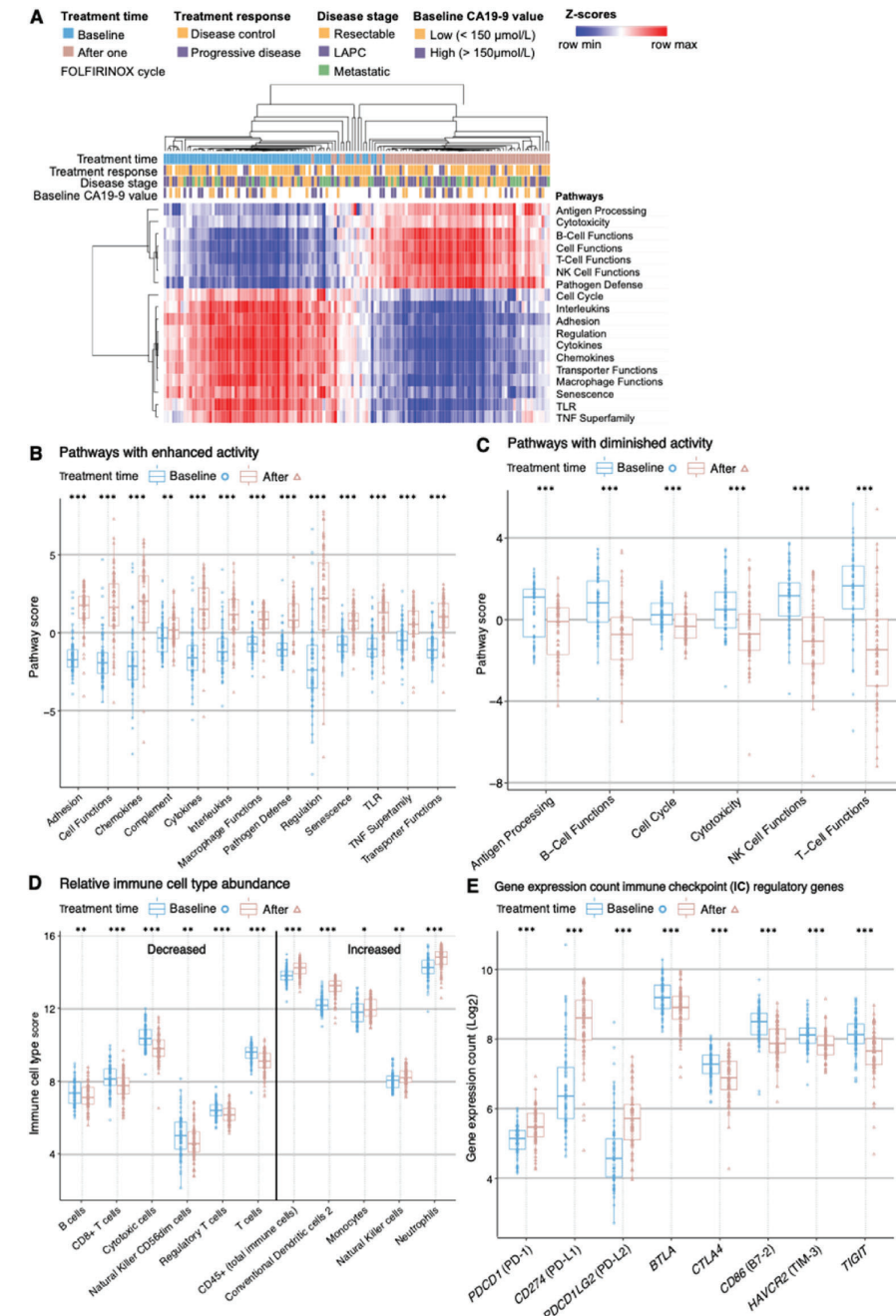

Figure 3 Immune profile alterations after a single FOLFIRINOX cycle. Samples at baseline (blue) and after one FOLFIRINOX cycle (brown). A: Heatmap of pathway scores showing sample clustering based on time of collection. B and C: Boxplots of pathways with enhanced (B) and diminished (C) activity. D: Boxplots of relative immune cell type abundance. E: Boxplots of IC regulatory genes. Statistical significance: *BH.P < 0.05, ***BH.P < 0.001. Abbreviations: NK: Natural killer; TLR: Toll-like receptor; TNF: Tumor Necrosis Factor; BH.P: Benjamin-Hochberg P value; IC: Immune Checkpoint; PD: Programmed Cell Death.

Subtle differences in the immune transcriptome of disease control and progressive disease patients after a single FOLFIRINOX cycle

Immune profiles of the disease control and progressive disease patients were compared at baseline and after a single FOLFIRINOX cycle. Baseline immune profiles revealed no differences in pathway activities between the two groups. However, a relatively high abundance of the total immune cells and Treg cells were observed in progressive disease patients (Figure 4A). Immune profiles after a single FOLFIRINOX cycle revealed 400 DEGs in disease control and 256 DEGs in progressive disease patients (Figure S3), which were used in the ClueGo analysis based on the criteria in the materials and method section. Two key genes involved in the negative regulation of type-I interferon-mediated (IFN-I) signaling pathway were downregulated in disease control but not in progressive disease patients (BH.P < 0.01; Figure 4C). The change in IC regulatory gene expression, pathway activity, and immune cell type abundance was comparable in both groups (Figure S4), with one immune cell type exception. Driven by its solitary marker *KIR3DL1*, the relative abundance of NK CD56^{dim} cells was decreased in disease control but increased in progressive disease patients (BH.P < 0.05; Figure 4B).
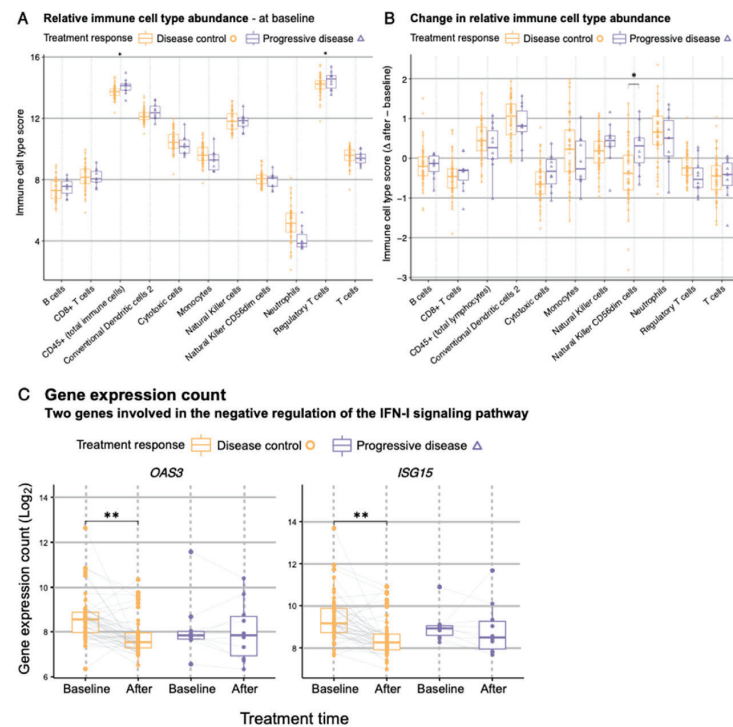

Figure 4 Differences in immune profiles between disease control (n = 48, yellow) and progressive disease (n = 10, purple) patients. A: Boxplots of samples at baseline displaying relative immune cell type abundances. B: Boxplots of the change in relative immune cell type abundances after a single FOLFIRINOX cycle. C: Boxplots of OAS3 and ISG15 expression counts (Log$_2$), involved in the negative regulation of the IFN-I signaling pathway, in baseline and after a single FOLFIRINOX cycle samples. Statistical significance: *BH.P < 0.05. Abbreviations: BH.P: Benjamin-Hochberg P value.

An eight-gene FFX-ΔGEP score predicted the lack of response after a single FOLFIRINOX cycle

To identify an early circulating biomarker that predicts the lack of FOLFIRINOX response, we developed an FFX-ΔGEP score. The Δ gene expression count, which results from subtracting the log$_2$ normalized gene expression counts of baseline samples from samples after a single FOLFIRINOX cycle, revealed fourteen candidate genes that differed significantly between disease control and progressive disease patients (BH.P < 0.05; Table 2).

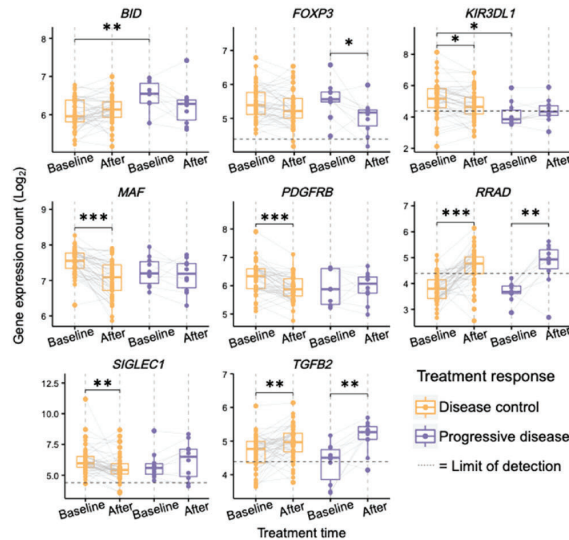*Table 2 The fourteen candidate genes selected for the FFX-ΔGEP score*

Gene	Disease control mean (± SD)		Progressive disease Mean (± SD)		BH.P value	Weights
BID	0.030	(± 0.48)	-0.237	(± 0.42)	0.030	-1.63
FOXP3	0.012	(± 0.67)	-0.518	(± 0.43)	0.012	-0.10
KIR3DL1	0.018	(± 0.59)	0.255	(± 0.78)	0.018	0.26
KLRC1	0.035	(± 0.56)	0.075	(± 0.71)	0.035	0
KLRD1	0.044	(± 0.46)	-0.293	(± 0.53)	0.044	0
KLRG1	0.041	(± 0.46)	-0.204	(± 0.47)	0.041	0
MAF	0.023	(± 0.44)	-0.111	(± 0.47)	0.023	0.54
NFATC2	0.024	(± 0.43)	-0.136	(± 0.37)	0.024	0
PDGFRB	0.038	(± 0.60)	0.011	(± 0.49)	0.038	0.31
PLAU	0.043	(± 0.99)	0.611	(± 0.71)	0.043	0
REL	0.028	(± 0.32)	0.278	(± 0.37)	0.028	0
RRAD	0.008	(± 0.46)	1.412	(± 0.70)	0.008	0.97
SIGLEC1	0.020	(± 0.84)	0.403	(± 1.24)	0.020	0.21
TGFB2	0.020	(± 0.64)	0.790	(± 0.60)	0.020	0.81

The mean values of the Δ gene expression counts (± SD) per response group. The BH.P value between disease control and progressive disease patients. The assigned weights are calculated using LASSO multivariate regression analysis. Abbreviations: SD: Standard Deviation; BH.P: Benjamin-Hochberg P value; LASSO: Least Absolute Shrinkage and Selection Operator.

LASSO multivariate regression analysis, which constructed the most optimal combination of candidate genes by assigning a regression coefficient (weight) to all candidate genes, was conducted. Six candidate genes were assigned a weight of zero and the FFX-ΔGEP score was composed of the remaining eight genes (Figure 5):

$$FOLFIRINOX\ delta\ gene\ expression\ profiling\ (FFX\ \Delta GEP)\ score$$
$$= (0.26 * KIR3DL1) + (0.54 * MAF) + (0.31 * PDGFRB) + (0.97 * RRAD)$$
$$+ (0.21 * SIGLEC1) + (0.81 * TGFB2) - (1.63 * BID) - (0.10 * FOXP3)$$

The eight-gene FFX-ΔGEP score ranged from 3.82 to -1.76 among all patients, and the performance to predict lack of FOLFIRINOX response after a single cycle was assessed by ROC analysis (Figure 6A). The leave-one-out cross-validated AUC [95% CI] was 0.87 [0.60 − 0.98], indicating that the FFX-ΔGEP score could distinguish between disease control and progressive disease patients. The predictive performance of the currently used absolute and proportional Δ CA19-9 values [95% CI] were 0.70 [0.27 − 1.0] and 0.52 [0.24 − 0.80]. Importantly, the FFX-ΔGEP score outperformed Δ CA19-9 values with less overlap in the designation of disease control and progressive disease patients (Figures 6B-6D).
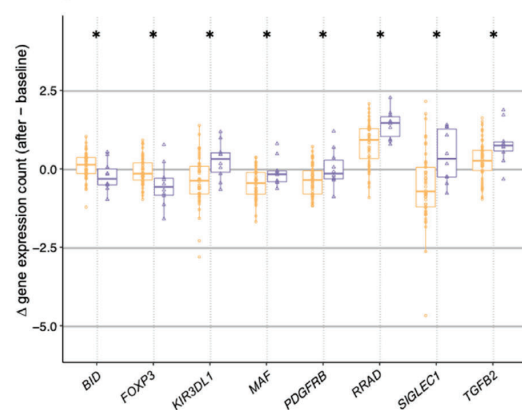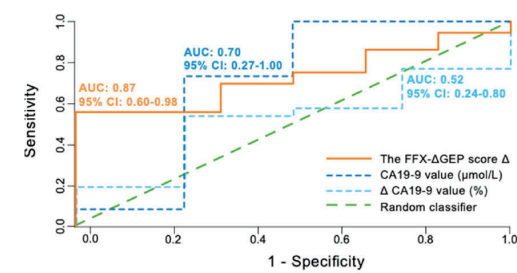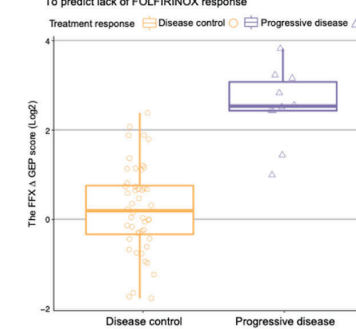
Figure 5 The eight genes that compose the FFX-ΔGEP score. All boxplots compare disease control (n = 48, yellow) and progressive disease patients (n = 10, purple). A: Boxplots of gene expression counts (Log₂) in baseline and after a single FOLFIRINOX cycle samples. B: Boxplots of the change (Δ) in gene expression count after a single FOLFIRINOX cycle and baseline samples. Statistical significance: *BH.P < 0.05, **BH.P < 0.01, ***BH.P < 0.001. Abbreviations: BH.P: Benjamin-Hochberg P value. FFX-ΔGEP: FOLFIRINOX-delta Gene Expression Profiling.
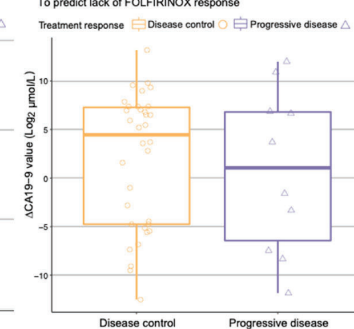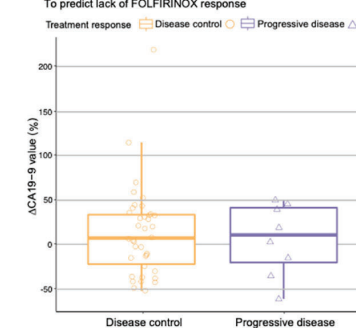
Figure 6 The FFX-ΔGEP score. A: The ROC of the FFX-ΔGEP score (orange), absolute (μmol/L) log₂ ΔCA19-9 score (dark blue), proportional (%) ΔCA19-9 score (light blue), and the random classifier (green). B-D: Boxplots of the FFX-ΔGEP score (B), the absolute (μmol/L) log₂ ΔCA19-9 score (C), and the proportional (%) ΔCA19-9 score (D) for disease control (n = 48, yellow) and progressive disease (n = 10, purple) patients. Abbreviations AUC: Area Under the Curve; CI: Confidence Interval; ROC: Receiver Operating Characteristic; FFX-ΔGEP: FOLFIRINOX-delta Gene Expression Profiling; CA19-9: Carbohydrate Antigen 19-9.

DISCUSSION

In this study, we used paired blood samples of 68 PDAC patients to investigate the effect of a single FOLFIRINOX cycle on the immune profile. We aimed to identify an early circulating biomarker to predict the lack of response to FOLFIRINOX. We revealed an eight-gene FFX-ΔGEP score that predicted the lack of FOLFIRINOX response only after the first cycle, independent of disease stage or change in CA19-9. This novel multigene FFX-ΔGEP score is, to our knowledge, the first gene expression-based early circulating biomarker predicting the lack of FOLFIRINOX response in PDAC patients from all disease stages.

The FFX-ΔGEP score is composed of eight immune-related genes. Four of these genes (*FOXP3*, *KIR3DL1*, *MAF*, and *SIGLEC1*) are associated with immune cell types [40–43], while the other four (*BID*, *PDGFRB*, *RRAD*, and *TGFB2*) are associated with the tumor or chemotherapeutic efficacy [44–47]. *FOXP3*, a marker for Tregs, is associated with poor PDAC prognosis [40,48] but their peripheral abundance could be reduced by neoadjuvant FOLFIRINOX [49,50]. *KIR3DL1* inhibits NK cell activity [41] and is associated with PDAC progression [51]. In this study, *KIR3DL1* expression decreased indeed only in disease control patients. *MAF* is a transcription factor that can impair CD8+ T cell function [42] and is often highly expressed in M2 macrophages [52], which are associated with poor prognosis in PDAC [53]. Correspondingly, our results showed a decrease in MAF expression in disease control patients only. *SIGLEC1* is a protein that mediates phagocytosis and endocytosis [54] and is expressed in the blood by activated CD14+ monocytes in reaction to IFN-I [43,55]. In contrast, our results showed downregulation of *SIGLEC1* but increased activation of the IFN-I pathway in disease control patients. *BID* encodes pro-apoptotic proteins [44] and their deregulated expression is associated with apoptotic resistance in PDAC [56]. Accordingly, our results showed downregulation in *BID* expression in progressive disease patients only. *PDGFRB* is associated with poor disease-free survival, cancer cell invasion, and metastasis in PDAC [45,46]. These pro-tumoral effects are in line with our results showing no change in *PDGFRB* expression in progressive disease but downregulation in disease control patients. In gastric and colorectal cancer, 5-FU and oxaliplatin, two chemotherapeutic agents of FOLFIRINOX, displayed increased efficacy when combined with *RRAD* inhibition [57]. Transforming growth factor- 2 (*TGFB2*) plays a complex role in PDAC and can both promote and inhibit tumor growth [58,59].

This study has some limitations. Firstly, our FFX-ΔGEP score needs to be validated in a larger cohort of patients because the current sample was not sufficient to accurately calculate a cut-off value for response or lack of response to FOLFIRINOX treatment. Secondly, the patients in this study received FOLFIRINOX in combination with G-CSF, but it would have been ideal to also include a patient group who received FOLFIRINOX only. However, at least in the Netherlands, the combination of G-CSF and FOLFIRINOX is standard practice due to the high risk of neutropenia. Therefore, we have accepted this omission and did not evaluate the performance of the FFX-ΔGEP score in patients who did not receive G-CSF. Thirdly, we did not evaluate the specificity of the FFX-ΔGEP score

for FOLFIRINOX by comparing its predictive ability in a control cohort of patients treated with another chemotherapy regimen, such as Nab-Paclitaxel-Gemcitabine. While this would have been ideal, the availability of patients treated with these regimens is limited due to the superior effectiveness of FOLFIRINOX. However, we plan to conduct a multi-center clinical trial to validate our results. Lastly, treatment response was evaluated using CT scans, but some patients experience prolonged OS without showing an imaging response. Therefore, it is important to determine whether the FFX-ΔGEP score can predict FOLFIRINOX-induced prolonged OS.

To our knowledge, this study is the first to describe the effect of a single FOLFIRINOX cycle, in combination with G-CSF, on the immune transcriptome of PDAC patients. We discovered that this treatment significantly changed the expression of 395 immune-related genes, even after two weeks of recovery. Our results showed that the relative peripheral abundance of total immune cells (CD45+), B cells, cDC2, cytotoxic cells, monocytes, NK CD56dim cells, and all T cell subsets (total, CD8+, and Treg) were reduced while the relative neutrophil abundance was increased after a single cycle of treatment. The increase in granulocyte-derived cells could be due to G-CSF which can affect the relative abundance of the other immune cells. In line with our results, previous studies described a rapid recovery of total lymphocytes, cDCs, and monocytes after two weeks of chemotherapy [47,60] and increased cDC2s after G-CSF treatment [61].

Importantly, the immune transcriptome in patients with different disease stages or different baseline CA19-9 values was similar. This suggests that the progression of PDAC does not stimulate the systemic immune response. Additionally, we could not predict the lack of FOLFIRINOX response using baseline samples only. This highlights the challenges we phase in applying precision medicine protocols or stratifying PDAC patients to receive their most suitable treatment. Based on our results, at least one cycle of FOLFIRINOX is needed to predict the lack of response in PDAC patients.

CONCLUSIONS

In this study, we developed a novel multigene FFX-ΔGEP score using targeted immune-gene expression profiling and found that it could predict the lack of FOLFIRINOX response in pancreatic cancer patients after only one cycle. In our cohort, the FFX-ΔGEP score predicted the lack of FOLFIRINOX response with more accuracy than the absolute or proportional change in CA19-9 levels. In addition, we were the first to describe the pronounced effect of a single FOLFIRINOX cycle on the immune transcriptome in the blood of PDAC patients from all disease stages. Further research is needed to validate our results in a larger cohort of patients, preferably including those who did not receive C-GSF treatment or were treated with another type of chemotherapeutic regimen.

REFERENCES

1. Sarantis, P., Koustas, E., Papadimitropoulou, A., Papavassiliou, A. G. & Karamouzis, M. v. Pancreatic ductal adenocarcinoma: Treatment hurdles, tumor microenvironment and immunotherapy. *World J Gastrointest Oncol* **12**, 173–181 (2020).

2. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **71**, 209–249 (2021).

3. Rangarajan, K., Pucher, P. H., Armstrong, T., Bateman, A. & Hamady, Z. Z. R. Systemic neoadjuvant chemotherapy in modern pancreatic cancer treatment: A systematic review and meta-analysis. *Ann R Coll Surg Engl* **101**, 453–462 (2019).

4. Kikuyama, M. *et al.* Early Diagnosis to Improve the Poor Prognosis of Pancreatic Cancer. *Cancers 2018, Vol. 10, Page 48* **10**, 48 (2018).

5. Adamska, A., Domenichini, A. & Falasca, M. Pancreatic Ductal Adenocarcinoma: Current and Evolving Therapies. *Int J Mol Sci* **18**, (2017).

6. Fan, J. Q. *et al.* Current advances and outlooks in immunotherapy for pancreatic ductal adenocarcinoma. *Mol Cancer* **19**, (2020).

7. CL, W. *et al.* Recent progress in pancreatic cancer. *CA Cancer J Clin* **63**, 228–230 (2013).

8. van der Sijde, F. *et al.* Treatment Response and Conditional Survival in Advanced Pancreatic Cancer Patients Treated with FOLFIRINOX: A Multicenter Cohort Study. *J Oncol* **2022**, (2022).

9. Suker, M. *et al.* FOLFIRINOX for locally advanced pancreatic cancer: a systematic review and patient-level meta-analysis. *Lancet Oncol* **17**, 801–810 (2016).

10. Conroy, T. *et al.* FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. *N Engl J Med* **364**, 1817–1825 (2011).

11. Conroy, T. *et al.* FOLFIRINOX or Gemcitabine as Adjuvant Therapy for Pancreatic Cancer. *N Engl J Med* **379**, 2395–2406 (2018).

12. Janssen, Q. P. *et al.* Neoadjuvant FOLFIRINOX in Patients With Borderline Resectable Pancreatic Cancer: A Systematic Review and Patient-Level Meta-Analysis. *J Natl Cancer Inst* **111**, 782–794 (2019).

13. Perri, G. *et al.* Response and Survival Associated With First-line FOLFIRINOX vs Gemcitabine and nab-Paclitaxel Chemotherapy for Localized Pancreatic Ductal Adenocarcinoma. *JAMA Surg* **155**, 832–839 (2020).

14. Thibodeau, S. & Voutsadakis, I. A. FOLFIRINOX Chemotherapy in Metastatic Pancreatic Cancer: A Systematic Review and Meta-Analysis of Retrospective and Phase II Studies. *J Clin Med* **7**, (2018).

15. Richards, M. K., Liu, F., Iwasaki, H., Akashi, K. & Link, D. C. Pivotal role of granulocyte colony-stimulating factor in the development of progenitors in the common myeloid pathway. *Blood* **102**, 3562–3568 (2003).

16. Steger, G. *et al.* Use of lipegfilgrastim for the prophylaxis of chemotherapy-induced neutropenia: Pan-European non-interventional study. *Annals of Oncology* **29**, viii607–viii608 (2018).

17. Terazawa, T. *et al.* Efficacy of Prophylactic G-CSF in Patients Receiving FOLFIRINOX: A Preliminary Retrospective Study. *Intern Med* **54**, 2969–2973 (2015).

18. Timmer-Bonte, J. N. H. *et al.* Lipegfilgrastim for prophylaxis of chemotherapy-induced neutropenia in Dutch patients. *Neth J Med* **78**, 270–276 (2020).

19. Winter, J. M., Yeo, C. J. & Brody, J. R. Diagnostic, prognostic, and predictive biomarkers in pancreatic cancer. *J Surg Oncol* **107**, 15–22 (2013).

20. van der Sijde, F. *et al.* Circulating Biomarkers for Prediction of Objective Response to Chemotherapy in Pancreatic Cancer Patients. *Cancers (Basel)* **11**, (2019).

21. Principe, D. R. *et al.* Long-Term Gemcitabine Treatment Reshapes the Pancreatic Tumor Microenvironment and Sensitizes Murine Carcinoma to Combination Immunotherapy. *Cancer Res* **80**, 3101–3115 (2020).

22. Zhou, C. *et al.* Camrelizumab plus carboplatin and pemetrexed versus chemotherapy alone in chemotherapy-naive patients with advanced non-squamous non-small-cell lung cancer (CameL): a randomised, open-label, multicentre, phase 3 trial. *Lancet Respir Med* **9**, 305–314 (2021).

23. Park, S. J. *et al.* Cisplatin and oxaliplatin induce similar immunogenic changes in preclinical models of head and neck cancer. *Oral Oncol* **95**, 127–135 (2019).

24. Bezu, L. *et al.* Combinatorial strategies for the induction of immunogenic cell death. *Front Immunol* **6**, (2015).

25. Tesniere, A. *et al.* Immunogenic death of colon cancer cells treated with oxaliplatin. *Oncogene* **29**, 482–491 (2010).

26. Liu, W. M., Fowler, D. W., Smith, P. & Dalgleish, A. G. Pre-treatment with chemotherapy can enhance the antigenicity and immunogenicity of tumours by promoting adaptive immune responses. *Br J Cancer* **102**, 115–123 (2010).

27. Stojanovska, V. *et al.* Oxaliplatin Treatment Alters Systemic Immune Responses. *Biomed Res Int* **2019**, (2019).

28. Lyman, G. H., Yau, L., Nakov, R. & Krendyukov, A. Overall survival and risk of second malignancies with cancer chemotherapy and G-CSF support. *Ann Oncol* **29**, 1903–1910 (2018).

29. Eisenhauer, E. A. *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* **45**, 228–247 (2009).

30. van der Sijde, F. *et al.* RNA from stabilized whole blood enables more comprehensive immune gene expression profiling compared to RNA from peripheral blood mononuclear cells. *PLoS One* **15**, (2020).

31. Cesano, A. nCounter(®) PanCancer Immune Profiling Panel (NanoString Technologies, Inc., Seattle, WA). *J Immunother Cancer* **3**, (2015).

32. Geiss, G. K. *et al.* Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol* **26**, 317–325 (2008).

33. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**, (2002).

34. NanoString Technologies®. nCounter Advanced Analysis 2.0 User Manual. Preprint at (2018).

35. de Koning, W. *et al.* Identification, Validation, and Utilization of Immune Cells in Pancreatic Ductal Adenocarcinoma Based on Marker Genes. *Front Immunol* **12**, 1449 (2021).

36. Bindea, G. *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093 (2009).

37. R Core Team. R: A Language and Environment for Statistical Computing. Preprint at https://www.R-project.org/ (2021).

38. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2016).

39. Blighe, K., Rana, S. & Lewis, M. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. Preprint at https://github.com/kevinblighe/EnhancedVolcano (2021).

40. Hu, L., Zhu, M., Shen, Y., Zhong, Z. & Wu, B. The prognostic value of intratumoral and peritumoral tumor-infiltrating FoxP3+Treg cells in of pancreatic adenocarcinoma: a meta-analysis. *World J Surg Oncol* **19**, (2021).

41. Peng, Y. P. *et al.* Comprehensive analysis of the percentage of surface receptors and cytotoxic granules positive natural killer cells in patients with pancreatic cancer, gastric cancer, and colorectal cancer. *J Transl Med* **11**, (2013).

42. Speiser, D. E., Ho, P. C. & Verdeil, G. Regulatory circuits of T cell function in cancer. *Nat Rev Immunol* **16**, 599–611 (2016).

43. Oliveira, J. J. *et al.* The plasma biomarker soluble SIGLEC-1 is associated with the type I interferon transcriptional signature, ethnic background and renal disease in systemic lupus erythematosus. *Arthritis Res Ther* **20**, (2018).

44. Adams, J. M. & Cory, S. Bcl-2-regulated apoptosis: mechanism and therapeutic potential. *Curr Opin Immunol* **19**, 488–496 (2007).

45. Singh, P. K. *et al.* Platelet-derived growth factor receptor beta-mediated phosphorylation of MUC1 enhances invasiveness in pancreatic adenocarcinoma cells. *Cancer Res* **67**, 5201–5210 (2007).

46. Weissmueller, S. *et al.* Mutant p53 drives pancreatic cancer metastasis through cell-autonomous PDGF receptor signaling. *Cell* **157**, 382–394 (2014).

47. Bang, S. *et al.* Differences in immune cells engaged in cell-mediated immunity after chemotherapy for far advanced pancreatic cancer. *Pancreas* **32**, 29–36 (2006).

48. Ino, Y. *et al.* Immune cell infiltration as an indicator of the immune microenvironment of pancreatic cancer. *Br J Cancer* **108**, 914–923 (2013).

49. Peng, H. *et al.* Neoadjuvant FOLFIRINOX Therapy Is Associated with Increased Effector T Cells and Reduced Suppressor Cells in Patients with Pancreatic Cancer. *Clin Cancer Res* **27**, 6761–6771 (2021).

50. Michelakos, T. *et al.* Tumor Microenvironment Immune Response in Pancreatic Ductal Adenocarcinoma Patients Treated With Neoadjuvant Therapy. *J Natl Cancer Inst* **113**, 182–191 (2021).

51. Lee, H. S. *et al.* Peripheral natural killer cell activity is associated with poor clinical outcomes in pancreatic ductal adenocarcinoma. *J Gastroenterol Hepatol* **36**, 516–522 (2021).

52. Kang, K. *et al.* Interferon- Represses M2 Gene Expression in Human Macrophages by Disassembling Enhancers Bound by the Transcription Factor MAF. *Immunity* **47**, 235-250.e4 (2017).

53. Yang, S., Liu, Q. & Liao, Q. Tumor-Associated Macrophages in Pancreatic Ductal Adenocarcinoma: Origin, Polarization, Function, and Reprogramming. *Front Cell Dev Biol* **8**, (2021).

54. O'Neill, A. S. G., van den Berg, T. K. & Mullen, G. E. D. Sialoadhesin - a macrophage-restricted marker of immunoregulation and inflammation. *Immunology* **138**, 198–207 (2013).

55. Xiong, Y. S. *et al.* Increased expression of Siglec-1 on peripheral blood monocytes and its role in mononuclear cell reactivity to autoantigen in rheumatoid arthritis. *Rheumatology (Oxford)* **53**, 250–259 (2014).

56. Campani, D. *et al.* Bcl-2 expression in pancreas development and pancreatic cancer progression. *J Pathol* **194**, 444–450 (2001).

57. Kim, H. K. *et al.* RRAD expression in gastric and colorectal cancer with peritoneal carcinomatosis. *Sci Rep* **9**, (2019).

58. Murphy, J. E. *et al.* Total Neoadjuvant Therapy With FOLFIRINOX in Combination With Losartan Followed by Chemoradiotherapy for Locally Advanced Pancreatic Cancer: A Phase 2 Clinical Trial. *JAMA Oncol* **5**, 1020–1027 (2019).

59. Chen, J. *et al.* Targeting transforming growth factor- signaling for enhanced cancer chemotherapy. *Theranostics* **11**, 1345–1363 (2021).

60. Markowicz, S. *et al.* Recovery of dendritic cell counts and function in peripheral blood of cancer patients after chemotherapy. *Cytokines Cell Mol Ther* **7**, 15–24 (2002).

61. Bonanno, G. *et al.* Effects of pegylated G-CSF on immune cell number and function in patients with gynecological malignancies. *J Transl Med* **8**, (2010).

Chapter 4

Immunomodulatory Effects of Stereotactic Body Radiotherapy and Vaccination with Heat-Killed Mycobacterium Obuense (IMM-101) in Patients with Locally Advanced Pancreatic Cancer

F.R. van 't Land*[1], S.P. Lau*[1], W. de Koning[2], L. Klaase[3], M. Vink[3], A. van Krimpen[3], J. Dumas[4], D. Vadgama[4], J.J. Nuyttens[5], D.A.M. Mustafa[4], R. Stadhouders[3,6], M. Willemsen[3], A.P. Stubbs[2], J.G. Aerts[3,7], C.H.J. van Eijck[1]

[1]Department of Surgery, Erasmus University Medical Center, 's-Gravendijkwal 230, 3015 CE Rotterdam, The Netherlands

[2]Department of Pathology and Clinical Bioinformatics, Erasmus University Medical Center, 's-Gravendijkwal 230, 3015 CE Rotterdam, The Netherlands

[3]Department of Pulmonary Medicine, Erasmus University Medical Center, 's-Gravendijkwal 230, 3015 CE Rotterdam, The Netherlands

[4]Department of Pathology, the Tumor Immuno-Pathology Laboratory, Erasmus University Medical Center, 's-Gravendijkwal 230, 3015 CE Rotterdam, The Netherlands

[5]Department of Radiation Oncology, Erasmus University Medical Center, 's-Gravendijkwal 230, 3015 CE Rotterdam, The Netherlands

[6]Department of Cell Biology, Erasmus University Medical Center, 's-Gravendijkwal 230, 3015 CE Rotterdam, The Netherlands

[7]Erasmus MC Cancer Institute, Erasmus University Medical Center, 's-Gravendijkwal 230, 3015 CE Rotterdam, The Netherlands

*These authors contributed equally to this work

ABSTRACT

Background: Patients with locally advanced pancreatic cancer (LAPC) are treated with chemotherapy. In selected cases, stereotactic body radiotherapy (SBRT) can be added to the regimen. We hypothesized that adding an adjuvant containing a heat-killed mycobacterium (IMM-101) to SBRT may lead to beneficial immuno-modulatory effects, thereby improving survival. This study aims to investigate the safety of adding IMM-101 to SBRT and to investigate the immuno-modulatory effects of the combination treatment in the peripheral blood of LAPC patients.

Methods: LAPC patients were treated with SBRT (40 Gy) and six intradermal vaccinations of one milligram IMM-101. The primary endpoint was an observed toxicity rate of grade 4 or higher. Targeted gene-expression profiling and multicolor flow cytometry were performed for longitudinal immune-monitoring of the peripheral blood.

Results: Twenty patients received study treatment. No treatment-related adverse events of grade 4 or higher occurred. SBRT/IMM-101 treatment induced a transient decrease in different lymphocyte subsets and an increase in CD14+CD16-CD11b+HLA-DR^{low} myeloid-derived suppressor cells. Importantly, treatment significantly increased activated ICOS+, HLA-DR+ and Ki67+PD1+ T and NK cell frequencies. This was not accompanied by increased levels of most inhibitory markers, such as TIM-3 and LAG-3.

Conclusions: Combination therapy with SBRT and a heat-killed mycobacterium vaccine was safe and had an immune-stimulatory effect.

Keywords: pancreatic ductal adenocarcinoma (PDAC); locally advanced pancreatic cancer (LAPC); stereotactic body radiotherapy (SBRT); mycobacterium vaccines; cancer immunotherapy; immuno-oncology

INTRODUCTION

Pancreatic ductal adenocarcinoma (PDAC) is a notoriously lethal malignancy with a five-year survival rate of less than 5% [1]. About thirty-five percent of patients present with locally advanced pancreatic cancer (LAPC) [2]. LAPC is treated with induction chemotherapy, preferably with the multi-agent FOLFIRINOX regimen in young and fit patients [3]. Next to FOLFIRINOX, gemcitabine combined with nab-paclitaxel is another adequate first-line treatment option, which is often better tolerated than FOLFIRINOX [3]. Stereotactic body radiotherapy (SBRT) can be added to the treatment regimen if there are no signs of disease progression after the chemotherapy [4–6].

Radiation therapy is the cornerstone of treatment for many cancer types, with fifty percent of cancer patients being treated with some form of radiotherapy throughout their illness [7]. Traditionally, radiation therapy has been utilized for its direct cytotoxic properties, inducing tumor cell apoptosis [8]. However, besides the direct cytotoxic effect, there is emerging evidence that radiation, particularly SBRT, has potential immuno-modulatory effects. Upregulation of immunogenic cell surface markers such as ICAM-1, MHC-1 and Fas on tumor cells has been described following radiotherapy [9–13]. Cancer cells may escape immune surveillance trough the downregulation of MHC-1 molecules [14]. The upregulation of MHC-1 molecules by radiation therapy may revert this escape mechanism. Additionally, irradiation can induce an upregulation of FAS molecules on tumor cells, thereby improving the cytotoxic efficacy of T cells [12]. Moreover, radiotherapy has been demonstrated to be able to induce immunogenic cell death [15], thereby reinforcing the cancer-immunity cycle [16,17]. In our previous LAPC-1 trial, LAPC patients were treated with FOLFIRINOX followed by SBRT [5]. The SBRT treatment was found to be safe, and the median overall survival (OS) in patients who received SBRT after FOLFIRINOX was 17 months (95% CI 14–21). As PDAC is considered an immunological cold tumor, the anti-tumor immune response in LAPC patients treated with SBRT monotherapy after systemic chemotherapy is probably not optimal. Adding an adjuvant to SBRT could improve the immunological conditions for an effective immune response. In this first-in-human trial, the addition of a vaccine containing a heat-killed mycobacterium obuense (IMM-101), to SBRT was investigated. IMM-101 has been demonstrated to induce the activation and maturation of dendritic cells in vitro [18]. Moreover, in a pancreatic cancer murine model, IMM-101 demonstrated to be able to produce protective CD8+ T cell responses [19]. A previous randomized controlled trial in patients with advanced pancreatic cancer investigated the value of adding IMM-101 to gemcitabine treatment [20]. The addition of IMM-101 to gemcitabine was associated with an improvement in OS from 4.4 to 7.0 months (95% CI 0.33–0.87, p = 0.01) in a pre-defined metastatic subgroup [20]. Next to this, an interesting case report presented a case of a patient with metastasized pancreatic cancer who underwent a synchronous resection of the primary tumor and liver metastases, after multimodality treatment with chemotherapy, IMM-101 and chemoradiation. This patient was free of disease four years after diagnosis [21]. Additionally, promising outcomes have been reported in melanoma patients treated with IMM-101 as well [22,23]. We hypothesize

that IMM-101 vaccinations can enhance a host's innate immune response, improving the immuno-modulatory effects and in situ vaccination efficacy of SBRT.

In this study, we present the results of the immuno-monitoring of the peripheral blood in patients with locally advanced pancreatic cancer treated with SBRT and IMM-101, as well as their clinical outcome.

TREATMENT SCHEME AND METHODS

Study Design and Participants

The LAPC-2 trial was a single-center, single-arm, non-randomized, open-label, phase I/II trial treating biopsy proven LAPC patients with SBRT and IMM-101, after prior treatment with at least 4 cycles of FOLFRINOX. LAPC was defined according to the guidelines of the Dutch Pancreatic Cancer Group as >90 contact with the superior mesenteric artery, the celiac axis and/or any hepatic artery and/or >270 contact with the superior mesenteric vein or the portal vein and/or occlusion of these veins [24]. Main inclusion criteria were (1) age > 18 years and < 75 years, (2) WHO performance status of 0 or 1, (3) normal renal and liver function, (4) largest tumor size <7 cm x 7 cm x 7 cm, and (5) no evidence of metastatic disease. Main exclusion criteria were (1) prior radiotherapy, chemotherapy other than FOLFIRINOX or pancreatic resection, (2) current or previous treatment with immunotherapeutic drugs, and (3) use of corticosteroids. The study was approved by the Central Committee on Research involving Human Subjects (NL68762.078.19) as defined by the Medical Research Involving Human Subjects Act. Procedures followed were in accordance with the ethical standards of these committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. The trial is registered with the Netherlands Trial Register, NL7578. Written informed consent was obtained from each subject. All detailed inclusion and exclusion criteria are listed in Supplementary Table S1.

SBRT and IMM-101 Vaccination

The tumors were irradiated with the Cyberknife (Accuray Incorporated, Sunnyvale, CA, USA). To accurately guide the radiation, the gastroenterologist placed three radiopaque markers in or near the tumor (within 3cm of the tumor). Patients received a total of 40 Gray (Gy) of SBRT in five fractions on consecutive days. Radiation started at week 2, just after patients received the second vaccination of IMM-101. Immodulon Therapeutics Ltd. (Uxbridge, UK) produced and shipped pre-labelled IMM-101 vials to the pharmacy of the Erasmus MC University Medical Center. IMM-101 was injected intradermally over the deltoid muscle by the standard Mantoux intradermal injection technique. One mL was injected, which contained one milligram of IMM-101. IMM-101 was administered six times: i.e., on week 0, week 2, week 4, week 8, week 10 and week 12. Figure 1 illustrates the treatment schedule. At week 0, week 2, week 4, week 8 and week 14 blood draws were performed for immune-monitoring;, i.e., before planned study drug administration or SBRT treatment. One red 10 mL clot activator tube from BD Vacutainer®, one 3 mL Tempus™ RNA stabilisator tube and two 10 mL EDTA tubes from BD Vacutainer® were collected. The blood was processed within six hours after collection. Plasma, serum and peripheral blood mono-nuclear cells (PBMCs) were isolated and cryopreserved.
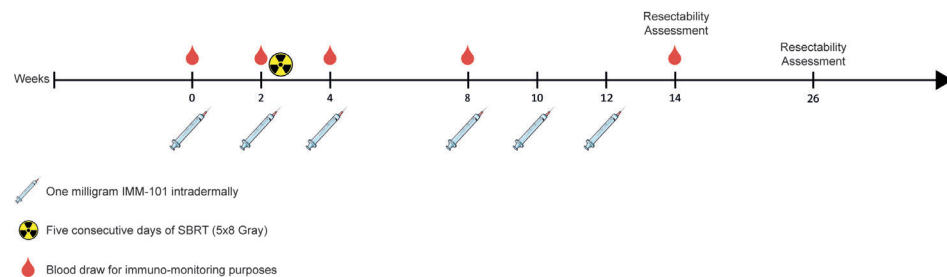
Figure 1 Schematic treatment schedule. After discontinuation of FOLFIRINOX treatment, patients were included in the trial. Patients received three bi-weekly intradermal vaccinations of IMM-101 at weeks 0, 2 and 4. At week 2, after the second vaccination, stereotactic body radio-therapy (SBRT) treatment started. Patients received 5 × 8 Gy of SBRT. They received three more vaccinations at weeks 8, 10 and 12. At week 14, the first resectability assessments was per-formed. Some patients were offered an explorative laparotomy with possible resection.

Follow up and Resectability Assessments

At week 14, resectability was assessed based on CT scans, biochemical response and the patients' clinical situation. An explorative laparotomy was performed in fit patients with a possibly resectable tumor and a >50% decrease in CA 19.9. In case of local and distant tumor progression, the patient was referred to the medical oncologist. The decision for an explorative laparotomy was made by a multidisciplinary tumor board consisting of at least a radiologist specialized in abdominal radiology, an experienced pancreas surgeon and a medical oncologist. After completion of IMM-101 treatment, routine follow-up was started until the time of death or 5 years after completion of SBRT. Follow-up visits included regular CT scans and tumor-marker assessments.

Objectives and Endpoints

The primary objective of the phase I study was to determine the safety of adding IMM-101 to SBRT. The endpoint for this objective was an observed toxicity rate of grade 4 or higher related to the study treatment. Toxicities were scored according to CTCAE criteria version 5.0 [25]. The secondary objective was to investigate the immuno-modulatory effects of the combination treatment in the peripheral blood. Endpoints for this were the changes in the circulating immune cell compartment on RNA and protein level.

Targeted Gene-Expression Profiling

RNA was isolated from Tempus blood tubes using Tempus TM Spin RNA Isolation Reagent Kit (Thermo Fisher Scientific, Breda, The Netherlands). Isolated RNA was puri-fied using RNeasy® MinElute® Cleanup Kit (Qiagen, Leiden, The Netherlands). The RNA quantity and quality were measured using the Agilent 2100 BioAnalyzer (Santa Clara, CA, USA). The RNA concentration was corrected to include the fragments ≥300 bp. For each sample, 200 ng of RNA was hybridized with probes of the PanCancer Immune profiling panel (730 innate and adaptive immune related genes and 40 housekeeping genes) for 17

h at 65 °C, following the manufacturing procedure (NanoString Technologies Inc., Seattle, WA, USA). The nCounter® FLEX platform was used to wash the extra probes, and genes were counted by scanning 490 Fields-of-view (FOV). The raw data of gene counts were uploaded to the nSolver™ Data Analysis software (version 4.0, NanoString, Seattle, WA, USA). The gene counts were normalized using the Advanced Analysis module (version 2.0) of nSolver™.

Flow Cytometry Immuno-Monitoring

For the enumeration of immune subsets, whole blood was freshly stained for flow cytometry. In addition, longitudinal immuno-monitoring was performed on liquid nitro-gen stored PBMCs. Cell surface staining was carried out after blocking Fc receptors by incubating cells with fluorescently conjugated mAbs directed against CD4 (SK3), CD11b (ICRF44), CD14 (M5E2), CD19 (HIB19), CD20 (2H7) CD56 (NCAM16.2), CD86 (FUN-1), HLA-DR (G46-6), ICOS (DX29) and ICOS-L (2D3/B7-H2) (all BD Biosciences, Erebodegem, Belgium); CD8 (SK1), CD11c (BV605), CD15 (HI98), CCR7 (G043H7), LAG-3 (11C3C65), PD-1 (EH12.2H7), TIM-3 (F38-2E2) (all BioLegend, Amsterdam, The Netherlands); and CD3 (UCHT1), CD33 (WM-53), CD45RA (MEM-56), CTLA-4 (14D3), FOXP3 (236A/E7), Ki-67 (20Raj1) (all Thermo Fisher Scientific). Intracellular transcription factor staining was performed using the FoxP3 Staining Buffer Set (Thermo Fisher Scientific). Cells were in addition stained for viability using fixable LIVE/DEAD aqua cell stain (Thermo Fisher Scien-tific). Data were acquired using the Symphony flow cytometer (BD Biosciences) and analyzed with FlowJo v10.7. Cell subsets are gated as previously described [26,27].

Statistical Analysis—Sample Size Calculation

The primary objective of the phase I trial was to determine the safety of adding IMM-101 to SBRT. In our previous LAPC-1 trial, the grade 4 toxicity rate of SBRT was 10% [5]. With a sample size of 20 for the phase I trial, we were able to estimate a toxicity rate of 10% within a 95% confidence interval of [1.2–31.7%] using the binomial exact method. This means that a maximum of 6/20 (30%) patients were allowed to have grade 4 toxicity or higher for the treatment to be regarded as safe and before proceeding to the phase II trial.

Statistical Analysis—Data Analysis and Visualisation

Baseline patient characteristics are summarized using the median and interquartile range for continuous variables and using counts and percentages for categorical varia-bles. PFS and OS were calculated from start date of FOLFIRINOX chemotherapy to the first documented event. Survival estimates were calculated using Kaplan–Meier method. Flow cytometry data were normalized for baseline. Paired Wilcoxon signed-ranks tests were used to test for significance between baseline measurements and other time points. Figures were made using GraphPad Prism software v8.0. Gene-expression data were

corrected for multiple testing using the Benjamini–Hochberg procedure. In all cases, a p-value of 0.05 and below was considered significant (*), $p < 0.01$(**) and $p < 0.001$ (***) as highly significant. The heat map was generated using the average log2 normalized gene expression of the significant differentially expressed genes per week. The heat map was visualized using the web-based tool Morpheus [28]. The Spearman correlations were calculated using the PFS or OS and the absolute difference between baseline and week 4 (after IMM101/SBRT) of activated cell frequencies. The volcano plots and correlations were visualized in R (version 4.1.1).

RESULTS

Patient and Treatment Characteristics

A total of 21 patients were included in the phase I, LAPC-2 trial, between October 2019 and June 2020. The first included patient (IMM001) had a liver metastasis, which was found during endoscopic ultrasound that was performed to place the radio-opaque markers for the SBRT. This patient was, therefore, excluded. Eventually, 20 patients received study treatment. Patients were treated with a median of 8 (8–9) cycles of FOLFI-RINOX before inclusion in the trial. The median time between FOLFIRINOX and the first IMM-101 vaccination was 6.4 (5.2–7.8) weeks. The median age was 63 (60–68) years and 11 (55%) were male. Their median body mass index was 24 (21–28) kg/m2. All patients received the total dose of 40 Gy of SBRT. Nineteen patients received the six planned vaccinations with IMM-101 and one patient received only three vaccinations due to disease progression. Immune analyses of the PBMCs were performed in 19/20 patients due to the absence of sufficient PBMCs in patient IMM016. Gene expression analyses were performed in 19/20 patients because we were not able to isolate RNA from IMM017. Detailed patient and treatment characteristics are shown in Table 1.

*Table 1 Patient and treatment characteristics. Statistics: Continuous variables are shown as medians with interquartile range and categorical variables are shown as counts with percentages. Abbreviations: BMI = body mass index, ECOG = Eastern Cooperative Oncology Group, CA 19.9 = carbohydrate antigen 19.9, CEA = carcinoembryonic antigen, SII = Systemic-Immune-Inflammation index, NLR = neutrophil to lymphocyte ratio, PLR = platelet to lymphocyte ratio, SBRT = stereotactic body radiotherapy, N = neutrophils, P = platelets, L = lymphocytes. * ECOG performance status 0 = Fully active, able to carry on all pre-disease performance without restriction. ECOG performance status 1 = Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light housework, office work.*

Patient Characteristics	N = 20 (IQR) or [%]
Age, years	63 (60–68)
Male sex	11 [55]
BMI, kg/m²	24 (21–28)
ECOG performance status *	
0	4 [20]
1	16 [80]
CA 19.9 at inclusion, kU/L	101 (43–137)
CEA at inclusion, µg/L	4.4 (3.5–5.8)
Leukocyte count at inclusion, ×10⁹/L	6.7 (4.7–9.9)
Platelet count at inclusion, ×10⁹/L	195 (133–232)
Neutrophil count at inclusion, ×10⁹/L	3.6 (2.7–7.2)
Lymphocyte count at inclusion, ×10⁹/L	1.4 (1.2–1.8)
SII, (N x P) / L	624 (311–889)

Patient Characteristics	N = 20 (IQR) or [%]
NLR	3.1 (2.3–5.0)
PLR	147 (87 – 171)
Treatment characteristics	
Biliary stenting at diagnosis	9 [45]
Diagnostic laparoscopy at diagnosis	6 [30]
FOLFIRINOX treatment	20 [100]
FOLFIRINOX, cycles	8 (8–9)
Interval stop FOLFIRINOX and start IMM-101, weeks	6.4 (5.2–7.8)
40 Gray of SBRT	20 [100]
IMM-101	20 [100]
Six vaccinations	19 [95]
Three vaccinations	1 [5]
Resection	4 [20]

Safety and Clinical Outcome

In 6/20 patients, we observed eleven grade 3 adverse events, of which three were considered to be possibly related to SBRT. None were related to IMM-101. Toxicity of grade 4 or higher was not observed. All patients experienced mild injection-site reactions, ranging from erythema to skin abscesses, with none resulting in systemic symptoms. Table 2 shows all grade 3 or higher toxicities. At present, (i.e., May 2022), 18/20 (90%) patients have experienced progression (local or distant) of disease and 17/20 (85%) patients have died. In all patients, the median PFS was 11.7 months (95% CI: 10.2–13.3) and the median OS was 17.8 months (95% CI: 11.3–24.4). The median PFS and median OS of the unresected patients (n = 16) was 11.2 (95% CI: 8.0–14.4) and 17.8 (95% CI: 12.0–23.6) months, respectively. Four (20%) patients underwent a resection of the tumor. In one patient, a small, solitary liver metastasis was found during explorative laparotomy and the primary tumor and the metastasis were both resected. This patient was free of disease 15 months after the operation. Another patient experienced local recurrence of disease four months after the resection. This was treated with systemic chemotherapy. In the absence of disease progression, a re-resection was performed 12 months after the initial resections. This patient was free of disease 8 months after the re-resection. Two patients died from complications from the operation.

Table 2 Grade 3 or higher adverse events. Toxicities were scored according to Common Terminology Criteria for Adverse Events (CTCAE) version 5.0 [25]. The treating physicians judged the possibility of a relation to the study treatment. Adverse events not related to SBRT or IMM-101 were considered to be related to pancreatic ductal adenocarcinoma. Abbreviations: SBRT = stereotactic body radiotherapy.

Subject	Adverse Event Term	Grade	Relation to SBRT	Relation to IMM-101
IMM003	Gastro-intestinal haemorrhage	3	Possibly	Unrelated
IMM006	Gastro-intestinal haemorrhage	3	Possibly	Unrelated
IMM007	Gastro-intestinal haemorrhage	3	Unrelated	Unrelated
IMM007	Gastro-intestinal haemorrhage	3	Possibly	Unrelated
IMM007	Stent disfunction	3	Unrelated	Unrelated
IMM007	Cholangitis	3	Unrelated	Unrelated
IMM007	Stent disfunction	3	Unrelated	Unrelated
IMM008	Cholestasis	3	Unrelated	Unrelated
IMM008	Cholangiosepsis	3	Unrelated	Unrelated
IMM009	Vertigo	3	Unrelated	Unrelated
IMM014	Duodenal obstruction	3	Unrelated	Unrelated

Downregulation of Genes Related to Lymphocyte Subsets and Immune inhibition after IMM-101/SBRT

Targeted gene expression profiling was performed to investigate the effect of IMM-101 and SBRT on the immune cells. Apart from increased expression of three genes (i.e., LTF, CAMP and LCN2) at baseline, no significant differences were observed between baseline (week 0) and after one vaccination IMM-101 (week 2) (Supplementary Figure S1). However, in week 4, after SBRT combined with IMM-101, profound changes were observed in immune-related gene expression (Figure 2A,B). Various genes related to lymphocyte subsets were downregulated (i.e., CD8a, MS4A1, CD22, CD79A, KLR family genes). Furthermore, genes related to lymphocyte inhibition/exhaustion (i.e., BTLA, TBX21, KLRC1) were also downregulated after IMM101/SBRT treatment. These results indicate changes in the circulating lymphoid compartment of LAPC patients specifically after combined IMM-101/SBRT treatment.
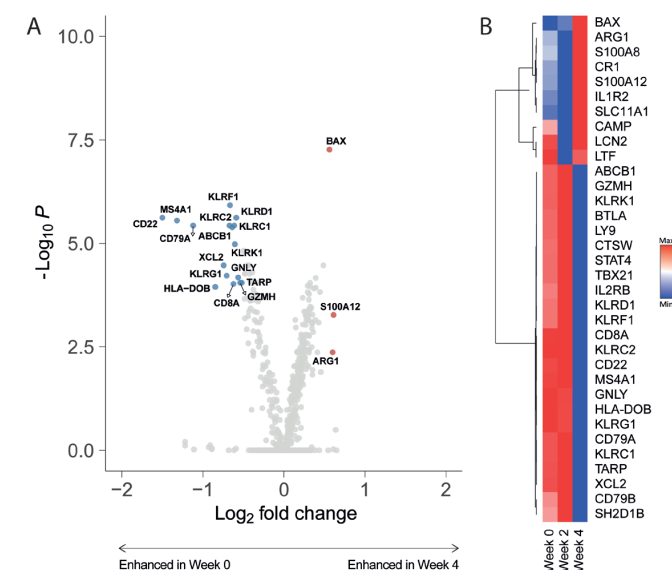

Figure 2 SBRT/IMM-101 induced gene expression. (a) Volcano plot demonstrating genes upregulated at baseline versus week 4. Highlighted genes underwent a log2fold change <−0.5 or >0.5 and p-value < 0.05. (b) Heat map of significantly differentially expressed genes between week 0, week 2 and week 4.

Reduced Peripheral Lymphocyte Numbers following IMM-101/SBRT

We additionally assessed various immune subsets in the peripheral blood using flow cytometry. No significant changes in immune subsets were observed two weeks after the first vaccination with IMM101. The addition of SBRT transiently reduced CD4+ and CD8+ T cells, CD19+ B-lymphocytes and CD56+ NK cells (Figure 3). SBRT did not curtail the myeloid compartment (i.e., CD15+CD16- eosinophils, CD15+CD16+ neutrophils, CD14+CD16- monocytes, CD14-CD16-CD11c+ dendritic cells). Additionally, the number of CD14+CD16-CD11b+HLA-DR^{low} MDSCs increased after combining SBRT and IMM-101 (Supplementary Figure S2). Lymphocyte cell numbers recovered at week 8, within 6 weeks after SBRT.
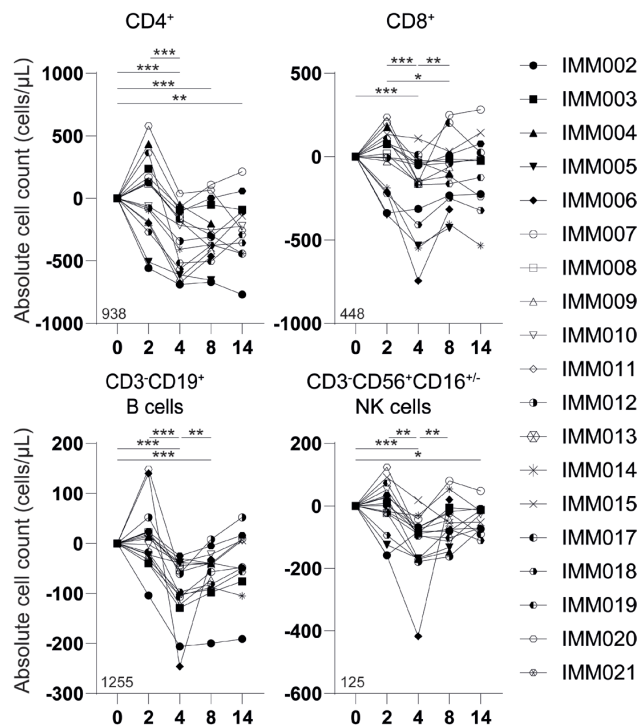

*Figure 3 SBRT/IMM-101 induced transient lymphodepletion. Number of CD4+, CD8+, CD3−CD19+ and CD3−CD56+CD16+/−peripheral blood lymphocytes per μL blood. N = 19. Data were normalized for baseline (week 0) and paired per patient. Percentage in the bottom left corner is the average frequency at baseline. Significance was determined using the paired Wilcoxon signed-rank test. * p < 0.05, ** p < 0.01, *** p < 0.0001.*

IMM-101/SBRT Increased Proportions of Activated Lymphocytes

In-depth longitudinal immune monitoring was performed to further describe the phenotypic characteristics of immune cells following study therapy. We did not find changes in activation or inhibitory marker expression on CD4+ regulatory T cells or CD4+ T helper cells or cytotoxic CD8+ T cells after one vaccination with IMM-101 in week 2. In contrast,

the addition of SBRT significantly increased the frequencies of activated CD4+ and CD8+ T cells and CD56+ NK cells in week 4 as indicated by the markers ICOS, HLA-DR as well as the combined increase in Ki67 and PD-1 levels. Notably, this increase was not observed for the inhibitory markers PD-1, TIM-3 and LAG-3, although we did observe significantly upregulated CTLA-4 levels on the CD4+ Non-Tregs after combination therapy. Furthermore, the increase in activated CD4+ and CD8+ T cell frequencies was mainly driven by the memory compartment (i.e., CCR7+CD45RA- central memory and CCR7-CD45RA- effector memory) [Not shown]. One vaccination of IMM-101 did significantly increase activated CD86+CD19+ B cell frequencies in week 2. The addition of SBRT further activated these CD19+ B cells demonstrated by increased Ki67+PD-1+ and CD86+ frequencies. Lastly, IMM-101/SBRT transiently induced higher frequencies of CD11c+ dendritic cells, HLADR+ CD14+ macrophages and HLA-DR-CD14-CD15- DN-MDSCs. Data are shown in detail in Figures 4 and S3.
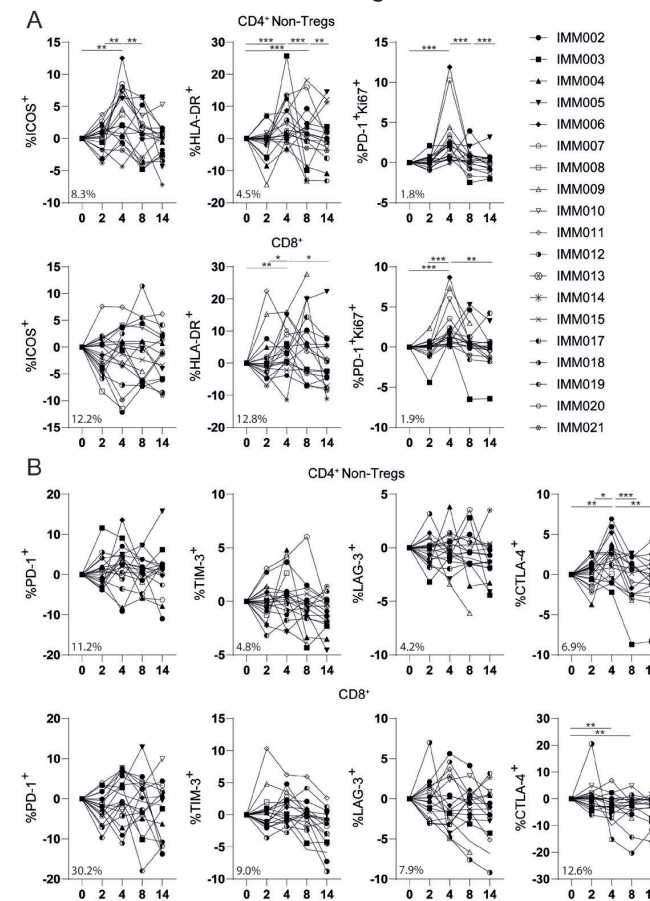
*Figure 4 SBRT/IMM-101 induced T-cell activation. (a) Percentage of ICOS+, HLA-DR+, PD-1+/Ki67+ subsets of CD4+ Non-Tregs and CD8+ cells. (b) Percentage of PD-1+, TIM-3+, LAG-3, CTLA-4+ subsets of CD4+ Non-Tregs and CD8+ cells. N = 19. Data were normalized for baseline (week 0) and paired per patient. Percentage in the bottom left corner is the average frequency at baseline. Significance was determined using the paired Wilcoxon signed-rank test. * p < 0.05, ** p < 0.01, *** p < 0.0001.*

Treatment-Induced Increase in Activated Lymphocytes Is Correlated with Survival

To explore if treatment-induced effects could be translated to clinical outcome, we analyzed if absolute differences in immune cell status between treatment-naïve (week 0) and study treatment samples (week 4) were correlated with survival. Patients who underwent a resection (n = 4) were excluded from this analysis, since a resection possibly influences PFS and OS outcomes. Another patient (IMM016) was excluded from the analysis due to an absence of sufficient PBMCs. Therefore, eventually 15 patients were included in the analysis. We found that increased levels of CD28+ effector memory (CCR7−CD45RA+) cytotoxic T cells correlated with improved PFS and OS (Figure 5).
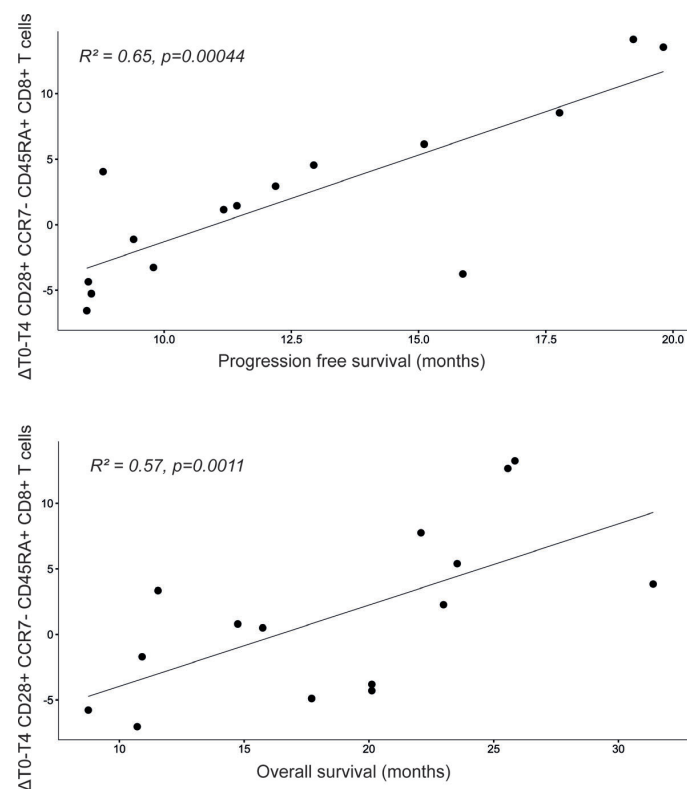

Figure 5 Treatment-induced T-cell activation correlated with improved progression-free survival. Spearman correlation plots demonstrating a positive correlation between IMM101/SBRT-induced absolute difference of CD28+ CCR7− CD54RA+ cytotoxic T cells and progression-free survival and overall survival. N = 15.

DISCUSSION

In this first-in-human trial, we firstly assessed the safety of IMM101/SBRT treatment, in patients with LAPC after prior treatment with FOLFIRINOX chemotherapy.

All patients experienced injection site reactions, which were uncomfortable for some patients. Eleven grade 3 toxicities were observed, of which three were possibly related to SBRT treatment. No grade 4 or higher toxicities were reported and none of the observed toxicities were considered to be related to IMM-101. This treatment approach demonstrated to be safe, and the trial proceeded to the phase II trial.

Secondly, we investigated the immunomodulatory effects of IMM-101/SBRT treatment in the peripheral blood. Two weeks after the first vaccination with IMM-101, no explicit changes on gene expression and protein level in the immune system of LAPC patients could be demonstrated. After treatment with IMM-101 with SBRT, we observed a downregulation of genes related to lymphocyte subsets, and this lymphodepletion was confirmed by flow cytometry. Interestingly, IMM-101/SBRT treatment did induce a rise in the number of MDSCs. Radiotherapy-induced MDSC expansion in patients with PDAC has previously been described [29]. It is also likely that SBRT and not IMM-101 induced the lymphodepletion, seeing that, in a previous study, external beam radiotherapy caused systemic immune-cell depletion [30]. Except MDSCs, cell numbers of other cell subsets within the myeloid compartment did not significantly increase. The latter may be explained by the fact that the radio-resistance of suppressive myeloid cells is stronger than that of lymphocytes [31].

Our combined gene expression and flow cytometry analyses demonstrated therapy-induced activation of T cell and NK cell subsets, with no increase in most inhibitory markers (i.e., PD-1, TIM-3 and LAG-3). Interestingly, therapy-induced activation of T cells occurred mainly in the memory compartment, which may be beneficial for seeding the tumor with antigen-specific T cells to mount successful anti-tumor responses. In agreement with this notion, improved PFS and OS were correlated with increased levels of activated effector memory cytotoxic T cells. In pre-clinical models, ablative doses of radiotherapy have been associated with improved intratumoral CD8+ T cell infiltration due to increased antigenicity of malignant cells, or by promoting immuno-stimulatory signals to recruit and activate antigen-presenting cells [32,33].

We found limited significant changes 2 weeks after the first vaccination with IMM-101. Still, the CD86+ expression on B cells increased. Adding SBRT further augmented the B cell activation, as demonstrated by the increase in Ki67+PD-1+ and CD86+ frequencies. A higher B cell activation may be beneficial, as B cell activation has been associated with positive responses to cancer-immunotherapy [34,35].

SBRT may hypothetically improve the anti-tumor efficacy of IMM-101 through anti-gen release upon tumor destruction, inducing in situ vaccination. IMM-101 could concurrently provide enhanced innate immunity to engage robust T cell responses. Unfortunately, the current study design did not allow for us to investigate this mechanism. Next to this, the common limitations of phase I/II trials, such as a small sample size and the lack of a control group, also applied to this study. However, the sample size was adequate to prove the safety of the combination treatment. Moreover, despite the low number of patients, a clear trend in immunological changes could be observed in most patients, which strengthens the hypothesis that treatment-induced immune modulation existed. Due to the lack of a control group, the observed changes could theoretically be better explained by time than by a cause–effect phenomenon caused by the treatment. However, certain factors argue against this. Firstly, between week 0 and 2, no significant changes occurred. In contrast, between week 2 and week 4, drastic changes were observed in the peripheral immunity. This occurred after the second vaccination and the SBRT treatment. The lack of changes in the first two weeks, compared to the extensive changes that occurred between week two and four, combined with the timing of treatment, argue against the hypothesis that the immunological changes were mostly impacted by time. Secondly, the observed immunological changes after SBRT/IMM-101 treatment tended to restore mostly to base-line after time progressed. If time and, thus, disease progression was the main factor ex-plaining the changes in the immune system, one would expect these changes to persist as time progressed. Another limitation of this study is that our analysis was only focused on peripheral immunity. A local assessment of the immune composition would have improved understanding of the study–treatment effect, as SBRT acts directly on the tumor. Nonetheless, the upregulation of immune checkpoints on circulating T cells, including CTLA-4, endorse the addition of immune-checkpoint blocking antibodies in future studies. Moreover, combining checkpoint-blocking antibodies with radiotherapy alone, or possibly with IMM-101, has shown promising results in pre-clinical models [36,37]. In addition to combination with immune-checkpoint-blocking antibodies, intratumoral administration of IMM-101 could improve its clinical efficacy. The most-used mycobacterium vaccine is the live-attenuated Mycobacterium Bovis Bacillus Calmette-Guérin (BCG) vaccine [38]. This tuberculosis vaccine was demonstrated to be able to induce potent anti-tumor immunity and adjuvant intravesical BCG instillations after a transurethral resection of bladder cancer, and was proved to be effective in preventing bladder cancer recurrence [39–43]. The administration of the vaccine at the disease site might be important to its efficacy.

CONCLUSIONS

In this open-label, single-center, phase I study, the safety and immunomodulatory effects of intradermal IMM-101 with SBRT were investigated in patients with LAPC. We observed transient lymphodepletion and enhanced T cell activation in the peripheral blood. Increased levels of activated T cells after treatment correlated with improved PFS and OS. Future studies are needed to provide mechanistic insights into how these observations are linked to clinical efficacy. The intratumoral administration of IMM-101 and combinations with other immunotherapeutic agents focusing on adaptive responses (e.g., immune checkpoint blockade, adoptive cell transfer therapy) may lead to improved efficacy for this group of patients with limited treatment options.

REFERENCES

1. Latenstein, A. E. J. *et al.* Nationwide trends in incidence, treatment and survival of pancreatic ductal adenocarcinoma. *Eur J Cancer* **125**, 83–93 (2020).

2. Mizrahi, J. D., Surana, R., Valle, J. W. & Shroff, R. T. Pancreatic cancer. *The Lancet* **395**, 2008–2020 (2020).

3. Tempero, M. A. *et al.* Pancreatic adenocarcinoma, Version 2.2021. *JNCCN Journal of the National Comprehensive Cancer Network* **19**, 439–457 (2021).

4. Herman, J. M. *et al.* Phase 2 multi-institutional trial evaluating gemcitabine and stereotactic body radiotherapy for patients with locally advanced unresectable pancreatic adenocarcinoma. *Cancer* **121**, 1128–1137 (2015).

5. Suker, M. *et al.* Efficacy and feasibility of stereotactic radiotherapy after folfirinox in patients with locally advanced pancreatic cancer (LAPC-1 trial). *EClinicalMedicine* **17**, (2019).

6. Quan, K. *et al.* Results of a prospective phase 2 clinical trial of induction gemcitabine/capecitabine followed by stereotactic ablative radiation therapy in borderline resectable or locally advanced pancreatic adenocarcinoma. *Pract Radiat Oncol* **8**, 95–106 (2018).

7. Delaney, G., Jacob, S., Featherstone, C. & Barton, M. The role of radiotherapy in cancer treatment: Estimating optimal utilization from a review of evidence-based clinical guidelines. *Cancer* **104**, 1129–1137 (2005).

8. Sia, J., Szmyd, R., Hau, E. & Gee, H. E. Molecular Mechanisms of Radiation-Induced Cancer Cell Death: A Primer. *Front Cell Dev Biol* **8**, (2020).

9. Chen, Y., Gao, M., Huang, Z., Yu, J. & Meng, X. SBRT combined with PD-1/PD-L1 inhibitors in NSCLC treatment: A focus on the mechanisms, advances, and future challenges. *J Hematol Oncol* **13**, (2020).

10. Sharabi, A. B., Lim, M., DeWeese, T. L. & Drake, C. G. Radiation and checkpoint blockade immunotherapy: Radiosensitisation and potential mechanisms of synergy. *Lancet Oncol* **16**, e498–e509 (2015).

11. Reits, E. A. *et al.* Radiation modulates the peptide repertoire, enhances MHC class I expression, and induces successful antitumor immunotherapy. *Journal of Experimental Medicine* **203**, 1259–1271 (2006).

12. Chakraborty, M. *et al.* Irradiation of tumor cells up-regulates Fas and enhances CTL lytic activity and CTL adoptive immunotherapy. *J Immunol* **170**, 6338–6347 (2003).

13. Sharabi, A. B. *et al.* Stereotactic radiation therapy augments antigen-specific PD-1-mediated antitumor immune responses via cross-presentation of tumor antigen. *Cancer Immunol Res* **3**, 345–355 (2015).

14. Cornel, A. M., Mimpen, I. L. & Nierkens, S. MHC class I downregulation in cancer: Underlying mechanisms and potential targets for cancer immunotherapy. *Cancers (Basel)* **12**, 1–33 (2020).

15. Golden, E. B. & Apetoh, L. Radiotherapy and Immunogenic Cell Death. *Semin Radiat Oncol* **25**, 11–17 (2015).

16. Ahmed, A. & Tait, S. W. G. Targeting immunogenic cell death in cancer. *Mol Oncol* **14**, 2994–3006 (2020).

17. Kroemer, G., Galluzzi, L., Kepp, O. & Zitvogel, L. Immunogenic cell death in cancer therapy. *Annu Rev Immunol* **31**, 51–72 (2013).

18. Bazzi, S. *et al.* Immunomodulatory effects of heat-killed Mycobacterium obuense on human blood dendritic cells. *Innate Immun* **23**, 592–605 (2017).

19. Elia, A., Lincoln, L., Brunet, L. & Hagemann, T. Treatment with IMM-101 induces protective CD8+ T cell responses in clinically relevant models of pancreatic cancer. *J Immunother Cancer* **1**, P215 (2013).

20. Dalgleish, A. G. *et al.* Randomised, open-label, phase II study of gemcitabine with and without IMM-101 for advanced pancreatic cancer. *British Journal of Cancer 2016 115:7* **115**, 789–796 (2016).

21. Costa Neves, M., Giakoustidis, A., Stamp, G., Gaya, A. & Mudan, S. Extended Survival after Complete Pathological Response in Metastatic Pancreatic Ductal Adenocarcinoma Following Induction Chemotherapy, Chemoradiotherapy, and a Novel Immunotherapy Agent, IMM-101. *Cureus* (2015) doi:10.7759/CUREUS.435.

22. Stebbing, J. *et al.* An intra-patient placebo-controlled phase I trial to evaluate the safety and tolerability of intradermal IMM-101 in melanoma. *Annals of Oncology* **23**, 1314–1319 (2012).

23. Dalgleish, A. G., Mudan, S. & Fusi, A. Enhanced effect of checkpoint inhibitors when given after or together with IMM-101: Significant responses in four advanced melanoma patients with no additional major toxicity. *J Transl Med* **16**, (2018).

24. Dutch Pancreatic Cancer Group. CT staging for adenocarcinoma of the pancreatic head and uncinate process (DPCG, 2012). https://dpcg.nl/wp-content/uploads/2020/04/Criteria_resectabiliteit.pdf (2018).

25. U.S. Department of Health and Human Services. Common terminology criteria for adverse events (CTCAE) version 5.0. *National Institutes of Health, National Cancer Institute* **5**, (2017).

26. Lau, S. P. *et al.* Autologous dendritic cells pulsed with allogeneic tumour cell lysate induce tumour-reactive T-cell responses in patients with pancreatic cancer: A phase I study. *Eur J Cancer* **169**, 20–31 (2022).

27. Kunert, A. *et al.* CD45RA+CCR7- CD8 T cells lacking co-stimulatory receptors demonstrate enhanced frequency in peripheral blood of NSCLC patients responding to nivolumab. *J Immunother Cancer* **7**, (2019).

28. Morpheus: Versatile matrix visualization and analysis software. https://software.broadinstitute.org/morpheus/.

29. Oweida, A. J. *et al.* Response to radiotherapy in pancreatic ductal adenocarcinoma is enhanced by inhibition of myeloid-derived suppressor cells using STAT3 anti-sense oligonucleotide. *Cancer Immunology, Immunotherapy* **70**, 989–1000 (2021).

30. van Meir, H. *et al.* Impact of (chemo)radiotherapy on immune cell composition and function in cervical cancer patients. *Oncoimmunology* **6**, (2017).

31. Barker, H. E., Paget, J. T. E., Khan, A. A. & Harrington, K. J. The tumour microenvironment after radiotherapy: Mechanisms of resistance and recurrence. *Nat Rev Cancer* **15**, 409–425 (2015).

32. Lee, Y. *et al.* Therapeutic effects of ablative radiation on local tumor require CD8 + T cells: Changing strategies for cancer treatment. *Blood* **114**, 589–595 (2009).

33. Verbrugge, I. *et al.* Radiotherapy increases the permissiveness of established mammary tumors to rejection by immunomodulatory antibodies. *Cancer Res* **72**, 3163–3174 (2012).

34. Helmink, B. A. *et al.* B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* **577**, 549–555 (2020).

35. Petitprez, F. *et al.* B cells are associated with survival and immunotherapy response in sarcoma. *Nature* **577**, 556–560 (2020).

36. Wei, J. *et al.* Sequence of PD-1 relative to local tumor irradiation determines the induction of abscopal antitumor immune responses. *Sci Immunol* **6**, (2021).

37. Crooks, J. *et al.* The effects of combination treatment of IMM-101, a heat-killed whole cell preparation of Mycobacterium obuense (NCTC 13365) with checkpoint inhibitors in pre-clinical models. in *Cell* vol. 10 20 (2016).

38. Mukherjee, N., Julián, E., Torrelles, J. B. & Svatek, R. S. Effects of Mycobacterium bovis Calmette et Guérin (BCG) in oncotherapy: Bladder cancer and beyond. *Vaccine* **39**, 7332–7340 (2021).

39. Malmström, P. U. *et al.* An Individual Patient Data Meta-Analysis of the Long-Term Outcome of Randomised Studies Comparing Intravesical Mitomycin C versus Bacillus Calmette-Guérin for Non-Muscle-Invasive Bladder Cancer. *Eur Urol* **56**, 247–256 (2009).

40. Shelley, M. D. *et al.* A systematic review of intravesical bacillus Calmette-Guérin plus transurethral resection vs transurethral resection alone in Ta and T1 bladder cancer. *BJU Int* **88**, 209–216 (2001).

41. Han, R. F. & Pan, J. G. Can intravesical bacillus Calmette-Guérin reduce recurrence in patients with superficial bladder cancer? A meta-analysis of randomized trials. *Urology* **67**, 1216–1223 (2006).

42. Böhle, A., Jocham, D. & Bock, P. R. Intravesical bacillus Calmette-Guerin versus mitomycin C for superficial bladder cancer: A formal meta-analysis of comparative studies on recurrence and toxicity. *Journal of Urology* **169**, 90–95 (2003).

43. Shelley, M. D. *et al.* Intravesical bacillus Calmette-Guérin is superior to mitomycin C in reducing tumour recurrence in high-risk superficial bladder cancer: A meta-analysis of randomized trials. *BJU Int* **93**, 485–490 (2004).

Chapter 5

Analyzing flow cytometry or targeted gene expression data influences clinical discoveries; profiling blood samples of pancreatic ductal adenocarcinoma patients

W. de Koning[1,2], C.W.F. van Eijck[1,3], F. van der Sijde[1,3], G.J. Strijk[1,3], A.A.M. Oostvogels[4], R. Debets[4], C.H.J. van Eijck[3], D.A.M. Mustafa[1]

[1] Tumor Immuno-Pathology Laboratory, Department of Pathology & Clinical Bioinformatics, Erasmus University Medical Centre, Rotterdam, Netherlands

[2] Clinical Bioinformatics Unit, Department of Pathology & Clinical Bioinformatics, Erasmus University Medical Centre, Rotterdam, Netherlands

[3] Department of Surgery, Erasmus University Medical Center, Rotterdam, The Netherlands.

[4] Department of Medical Oncology, Laboratory of Tumor Immunology, Erasmus University Medical Centre, Rotterdam, Netherlands

ABSTRACT

Introduction: Monitoring the therapeutic response of pancreatic ductal adenocarcinoma (PDAC) patients is crucial to determine treatment strategies. Several studies have examined the effectiveness of FOLFIRINOX as a first-line treatment in patients with locally advanced pancreatic cancer, but little attention has been paid to the immunologic alterations in peripheral blood caused by this chemotherapy regimen. Furthermore, the influence of the measurement type (e.g., flow cytometry, and targeted gene expression) on the clinical discoveries is unknown. Therefore, we aimed to scrutinize the influence of using flow cytometry or targeted immune-gene expression to study the immunological changes in blood samples of PDAC patients who were treated with a single-cycle FOLFIRINOX combined with lipegfilgrastim (FFX-Lipeg).

Material & Methods: Whole blood samples from 44 PDAC patients were collected at two time points; before the first FOLFIRINOX cycle and 14 days after the first cycle. EDTA blood tubes were used for multiplex flow cytometry analyses to quantify 18 immune cell populations and for complete blood count tests as standard clinical routine. The flow cytometry data were analyzed with FlowJo software. In addition, Tempus blood tubes were used to isolate RNA and measure 1,230 immune-related genes using the NanoString Technology®. Data quality control, normalization, and analysis were performed using the nSolver™ software and the Advanced Analysis module.

Results: FFX-Lipeg treatment increased the number of neutrophils and monocytes as shown by flow cytometry and complete blood count in concordance with elevated gene expression measured by targeted gene expression profiling analysis. Interestingly, flow cytometry analysis showed an increase in the number of B and T cells after treatment, while targeted gene expression analysis showed a decrease in B and T cell-specific gene expression.

Conclusion: Targeted gene expression complements the flow cytometry analysis to provide a comprehensive understanding of the effect of FFX-Lipeg. Flow cytometry and targeted gene expression showed an increase in the neutrophils and monocytes after FFX-Lipeg. The number of lymphocytes is increased after treatment, nevertheless, their cell-specific gene expression levels are downregulated. This highlights that different techniques influence clinical discoveries. Therefore, it is important to carefully select the measurement technique to study the effect of treatment.

Keywords: Pancreatic ductal adenocarcinoma (PDAC), FOLFIRINOX, Flow cytometry, Targeted gene expression, Peripheral immune cell profile

INTRODUCTION

Pancreatic ductal adenocarcinoma (PDAC) is a malignancy that develops from the epithelial cells that line pancreatic ducts and is one of the most lethal cancer types [1]. The complex tumor (immune) microenvironment, the ecosystem that surrounds the tumor cells, is composed of many cell types and the extracellular matrix. It includes immune cells, blood vessels, cancer-associated fibroblasts, and other cells that constantly interact and influence each other. The first-line treatment for patients with locally advanced and metastatic PDAC is a chemotherapeutic regimen of either a combination of gemcitabine with nab-paclitaxel or 5-fluorouracil, folinic acid, irinotecan, and oxaliplatin (FOLFIRINOX) [2]. FOLFIRINOX is considered the optimal adjuvant treatment in resected patients with good performance status and is currently being investigated as neoadjuvant therapy in several randomized control clinical trials [3]. However, FOLFIRINOX is associated with increased toxicity, specifically neutropenia [3–8]. Prophylactic treatment with a granulocyte-colony stimulating (G-CSF), such as lipegfilgrastim, could prevent FOLFIRINOX-induced neutropenia and is therefore considered standard therapy by most medical oncologists [9–11].

FOLFIRINOX has been shown to alter the intra-tumoral immune cell profile of PDAC patients. Increased effector T cells and reduced suppressor cells were reported in the pancreatic tumor after neoadjuvant FOLFIRINOX treatment [12]. Furthermore, FOLFIRINOX enhanced tumor antigen presentation, potentially synthesizing the pancreatic tumor for treatment with immune checkpoint inhibitors [13,14]. Nevertheless, the effect of FOLFIRINOX-lipegfilgrastim (FFX-Lipeg) on the patient's peripheral immune profile remains unclear. More knowledge on the immunological changes caused by FOLFIRINOX could pave the way to improved immunotherapeutic approaches and are therefore of clinical interest.

Flow cytometry is one of the most used techniques to determine the immune cell composition of the peripheral blood. This laser-based technique is used to detect and analyze the chemical and physical characteristics of cells or particles in fluids [15]. Flow cytometry allows for the precise quantification of immune populations from peripheral blood. It has the advantage of a more detailed, single-cell analysis. Subpopulations can be distinguished although common protein markers are present, and cells can be assigned to subpopulations based on negative protein markers. However, flow cytometry does not allow for robust quantification of subpopulations based on intracellular characteristics [16]. Furthermore, flow cytometry requires viable cells that require costly and time-consuming processing and storage. Longitudinal sample collection needs cryopreservation, which has been shown to alter the expression of important markers for immune subsets [17–19].

Another emerging technique to study different cellular processes, such as immune responses and cell types that are present in the peripheral blood, is RNA-based

transcriptome analysis [20]. Targeted gene expression profiling provides a more comprehensive immune-related data set compared to flow cytometry. The composition of the immune cell subpopulations with RNA-based transcriptome analysis is inferred from the generated bulk gene expression dataset and can be defined based on intracellular characteristics [21,22]. Furthermore, gene expression analysis is possible even when far fewer viable cells are available, as it only requires 25-100 ng of total RNA.

Recently, we identified the FOLFIRINOX delta Gene Expression Profiling (FFX-ΔGEP) score that predicts the lack of FOLFIRINOX response in PDAC patients after only one cycle of FFX-Lipeg treatment [23]. In this study, we aimed to scrutinize the influence of using flow cytometry or targeted immune-gene expression to study the immunological changes in blood samples of PDAC patients who were treated with a single cycle of FFX-Lipeg.

MATERIAL AND METHODS

Patient cohort and blood collection

The 44 PDAC patients included in this study were hospitalized at the Erasmus University Medical Centre Rotterdam between February 2018 and October 2020. Fourteen of those patients had (borderline) resectable PDAC and participated in the randomized clinical trial PREOPANC-2 (Dutch trial register NL7094). Thirty patients participated in the prospective cohort study iKnowIT (Dutch trial register NL7522) of which 19 had locally advanced and 11 had metastasized PDAC. Exclusion criteria were < 18 years of age, previous treatment with FOLFIRINOX, or co-treatment with another chemotherapeutic. The medical ethics committee of the Erasmus University Medical Centre Rotterdam approved both studies (MEC-2018-087 and MEC-2018-004), and patient samples were only used when written informed consent was provided. Following histological confirmation of the primary tumor or metastases, patients were treated with FOLFIRINOX chemotherapy. In addition, all patients were prophylactically treated with the long-acting G-CSF lipegfilgrastim (Lonquex®; Teva Ltd, Petach Tikva, Israel), 24 hours after each cycle, to reduce FOLFIRINOX-induced neutropenia [9,24]. Two types of whole blood samples (2 EDTA and 1 Tempus) were collected from the 44 PDAC patients at two time points: at baseline (on the same day, but before the first cycle) and 14 days after the first cycle, but before the second FOLFIRINOX cycle (Figure 1).

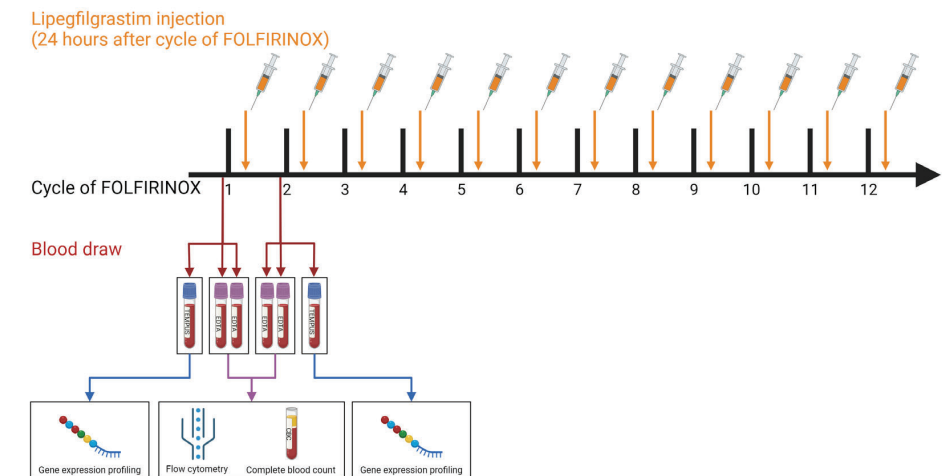

Figure 1: Schematic overview of the sample collection and measurements. The cycles of FOLFIRINOX chemotherapy (black), lipegfilgrastim injections (orange), and the blood draw time points (red). The blood is collected in EDTA and Tempus tubes. After the blood draw, flow cytometry, and complete blood count are performed on the EDTA tubes. The Tempus tubes are used for gene expression profiling.

Flow cytometry and complete blood count

The whole blood samples for flow cytometry and complete blood count (CBC) were collected in EDTA tubes. As part of the standard clinical routine, CBC tests were performed to evaluate lymphocyte, neutrophil, and thrombocyte counts. To quantify 18 immune cell populations, whole blood was stained and analyzed by multiplex flow cytometry as described previously [25]. In short, granulocytes, monocyte, lymphocytes, and T cell populations were gated separately in a scatter plot of $CD45^+$ staining versus side scatter. Immune cell populations were further defined using markers for eosinophils ($CD15^+ CD16^-$), mature neutrophils ($CD15^{high} CD16^{high}$), immature neutrophils ($CD15^+ CD16^+$), classical monocytes ($CD14^+ CD16^-$), intermediate monocytes ($CD14^+ CD16^+$), non-classical monocytes ($CD14^- CD16^+$), dendritic cells ($CD14^- CD16^- CD11c^+$), myeloid-derived suppressor cells ($CD14^+ CD16^- CD11b^+ HLA-DR^{low}$), B cells ($CD3^- CD19^+$), natural killer cells ($CD3^- CD56^+ CD16^{+/-}$), T cells ($CD3^+$), T cells ($CD3^+ TCR^+$), CD4+ T cells ($CD3^+ TCR^- CD4^+$) and CD8+ T cells ($CD3^+ TCR^- CD8^+$). Flow-Count Fluorospheres (Beckman Coulter, Brea, CA, USA) were used to determine the absolute cell counts. Data were analyzed with FlowJo software (Tree Star, San Carlos, CA, USA).

Targeted immune-gene expression profiling

The whole blood samples for targeted gene expression analysis were collected in Tempus tubes (Applied Biosystems, Foster City, CA, USA) and stored at -80°C. Tempus tubes contain an RNA stabilizing reagent, which preserves the RNA quality and enables measuring gene expression profiles without isolating the peripheral blood mononuclear cells [26]. Total RNA was extracted from blood in Tempus tubes using the Tempus Spin RNA Isolation Kit of Thermo Fisher Scientific (Waltham, MA, USA) following the manufacturer's instructions. RNA quality control was done using the Agilent 2100 BioAnalyzer (Santa Clara, CA, USA). Samples with RNA concentrations less than 35 mg/mL were excluded. Corrected RNA concentrations were calculated based on the percentage of fragments of 300-4000 nucleotides to correct RNA degradation. The nCounter® PanCancer Immune Profiling (IP) Panel and nCounter® Myeloid Innate Immunity (MII) Panel, consisting each of 730 genes and 40 housekeeping genes, were used for NanoString targeted gene expression analysis of the 88 whole blood samples. The IP panel targeted different immune-related pathways as well as immune and adaptive immune cell-related genes. The MII panel targeted genes involved in the innate immune response of myeloid-derived cells. A total of 200 ng RNA per sample in a maximum of 7 μL was hybridized with the two panels for 17 hours at 65⁰C, following the manufacturing procedure (NanoString Technologies Inc., Seattle, WA, USA). The nCounter® FLEX platform was used to wash the unbound probes, and genes were counted by scanning 490 Fields-of-view (FOV). Data quality control, normalization, and analysis were performed using the nSolver™ software (version 4.0) and the Advanced Analysis module (version 2.0) of NanoString Technology Inc. [27]. Raw gene counts were normalized based on the most stable 19 housekeeping genes, identified by the geNorm algorithm [28], scaling was performed based on

the 265 overlapping genes between the two panels, and all normalized data were log_2 transformed. Genes were included when they were higher than the limit of detection of 4.35 log_2, calculated as the average count of the negative controls plus two standard deviations in > 50% of the gene expression profiles. Differentially expressed genes (DEGs) were identified using simplified negative binomial models, a mixture of negative binomial models, or log-linear models based on the convergence of each gene. Genes with a P-value < 0.05 after correction for multiple testing with the Benjamin-Hochberg (BH) method were considered DEGs. To characterize the immune cell abundance, candidate cell-specific gene markers were identified as described previously [22]. In short, gene markers were selected by calculating the pairwise similarity between all pairs of the 61 candidate marker genes that were above the detection. The gene pairs with a pairwise similarity above 0.6 were selected to describe the immune cells. Each immune cell type needed at least two unique genes. The abundance of the immune cell types is the average expression value of their corresponding marker genes corrected for the total CD45+ infiltration. Pathway scores were extracted from the Advanced analysis of the nSolver software. Pathway enrichment analyses were performed with the differentially expressed genes (|Log_2 fold change| > 0.5, P.BH < 0.05) using Metascape [29] and ClueGo [30].

Statistical analysis

Statistical testing and data visualization were performed with R Statistical Software (v.4.1.2) [31]. We used paired two-sided student t-tests as a parametric test. We used the R packages ggplot2 [32] and EnhancedVolcano [33] for data visualization. Heatmaps were generated using the Log_2-normalized count data of significantly differentially expressed (|Log_2 Fold change| > 0.5, P.BH < 0.05). Genes that were determined to be outliers using Tukey's rule were removed [34]. The heatmap was visualized using the web-based tool Morpheus by Broad Institute (RRID: SCR_017386).

RESULTS

Patients' characteristics

A total of 44 patients were included who received one FFX-Lipeg cycle. Three whole blood samples (2 EDTA, 1 Tempus) were collected at baseline (on the same day before the first cycle) and 14 days after the first cycle but before the second cycle. The mean overall survival [95% CI] was 9 [11 – 14] months, calculated as the months between the first FFX-Lipeg cycle and the date of death. All clinicopathological characteristics are summarized in Supplementary Table 1.

FFX-Lipeg therapy results in enhanced frequencies of granulocytes and monocytes in the blood

Flow cytometry analyses showed that one cycle of FFX-Lipeg significantly increased 17 out of 18 immune cell types. The most pronounced increase was observed in the number of granulocytes and monocytes (Figure 2). The only cell type that was not significantly altered were the CD16+ NK cells (Supp Fig. 1). The CBC measurements showed a significant increase in lymphocytes and neutrophils but a significant decrease in thrombocytes after treatment (Supp Fig. 2).
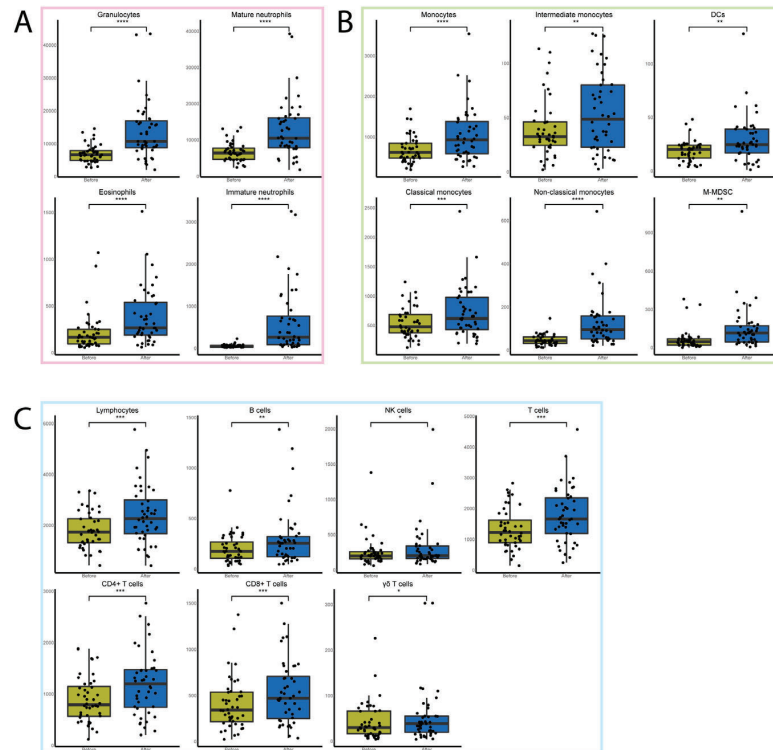

Figure 2: The effect of one cycle FFX-Lipeg on the immune cells measured by flow cytometry. (A) The number of granulocytes (pink) including the subtypes that were significantly increased after treatment (blue) in comparison with before treatment (yellow). (B) The number of monocytes (green) including the subtypes that were significantly increased after treatment. **(C)** The number of lymphocytes (light blue) including the subtypes that were significantly increased after treatment.

Targeted immune expression profiling showed that the total infiltration of CD45+ cells significantly increased after FFX-Lipeg treatment. The definition of infiltrated immune cells was based on 42 genes that showed a pairwise similarity higher than 0.6 and defined thirteen different immune cell types. The relative peripheral abundance of neutrophils, monocytes, and mast cells significantly increased after treatment. In contrast to the flow cytometry measurements, the relative peripheral abundance of T cells and B cells significantly decreased after treatment (Figure 3). The relative abundance of Plasma B cells was not significantly altered after treatment (Supp Fig. 3).
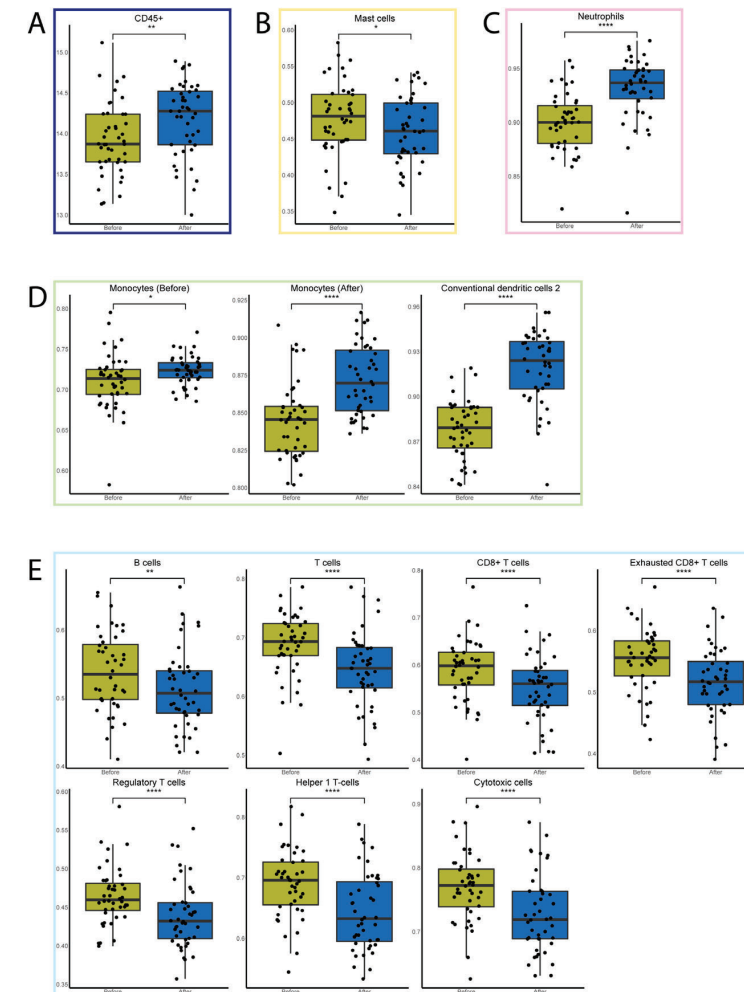
Figure 3: The effect of one cycle FFX-Lipeg on the immune cells measured by targeted gene expression. (A) The total immune cell (dark blue) abundance was significantly increased after treatment (blue) in comparison with before treatment (yellow). (B) The abundance of mast cells (yellow) was significantly increased after treatment. (C) The abundance of granulocytes' (pink) subtype neutrophils was significantly increased after treatment. (D) The abundance of monocytes (green) including the subtypes was significantly increased after treatment. (E) The abundance of lymphocytes (light blue) including the subtypes was significantly decreased after treatment.

FFX-Lipeg results in an increased number of lymphocytes, yet pathways associated with lymphocyte functions are downregulated

Flow cytometry results showed that T, B, and NK cells and their subtypes were increased after FFX-Lipeg treatment. On the contrary, pathway analysis using the nSolver software revealed decreased scores of natural killer, B cell, and T cell functions. On the other hand, both data from flow cytometry and targeted gene expression revealed an increase in the myeloid compartment after FFX-Lipeg therapy. In concordance, pathway analysis resulted in an increased score of the differentiation and maintenance of myeloid cells after FFX-Lipeg treatment. (Figure 4).
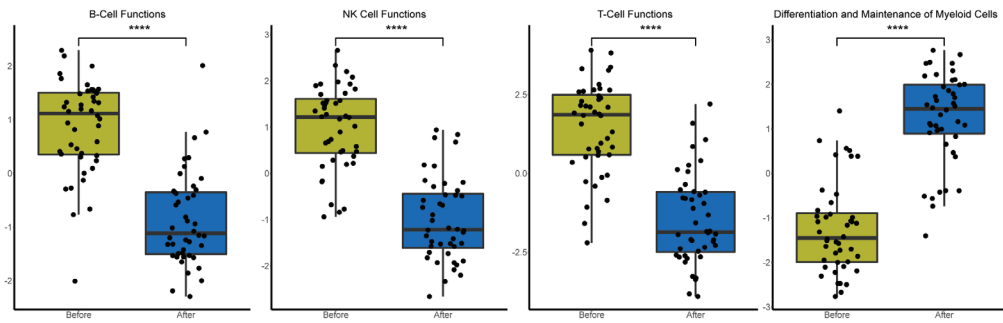
Figure 4: Differential expression of predefined pathway genes after one cycle of FFX-Lipeg. A significant upregulation of leukocyte functions and the differentiation and maintenance of myeloid cells was measured. On the contrary natural killer, B cell, and T cell functions were significantly downregulated.

Metascape enrichment analysis using the 39 downregulated (P.BH ≤ 0.05 and log2 fold of change ≤ -0.5) genes showed that the FFX-Lipeg treatment negatively affected the MHC class II antigen presentation (Supp Fig. 4A). The 170 upregulated (P.BH ≤ 0.05 and $\log_2$ fold of change ≥ 0.5) genes showed an enriched neutrophil degranulation, IRAK4 deficiency (TLR5), leukotriene metabolic pathway, IL-4 signaling pathway, and activation of matrix metalloproteinases (Supp Fig. 4B). ClueGo functional analysis using the 39 downregulated (P.BH ≤ 0.05 and log2 fold of change ≤ -0.5) showed a negative regulation of T cell-mediated immunity and the innate immune response (Supp Fig. 5A). On the contrary, myeloid, and neutrophil cell-related functions were stimulated and genes related to lymphocyte proliferation were enriched in the 170 upregulated (P.BH ≤ 0.05 and $\log_2$ fold of change ≥ 0.5) genes (Supp Fig. 5B).

In order to scrutinize the differences in the results between flow cytometry and gene expression, we performed a correlation analysis of proteins targeted by flow cytometry with their corresponding marker genes in the targeted immune-gene expression profile for the two time points separately. Lymphocyte subtypes showed a higher correlation before FFX-Lipeg treatment in comparison to the correlation after treatment (Supp Fig. 6).

FFX-Lipeg induces a distinct gene expression profile in the peripheral blood

A distinct genetic profile was observed between baseline and after a single FFX-Lipeg cycle (Figure 5A). Out of the 870 genes that were above the detection limit, 209 were differentially expressed (|Log$_2$ Fold change| > 0.5, P.BH < 0.05), of which 170 genes were upregulated and 39 genes were downregulated after FFX-Lipeg treatment (Figure 5B).
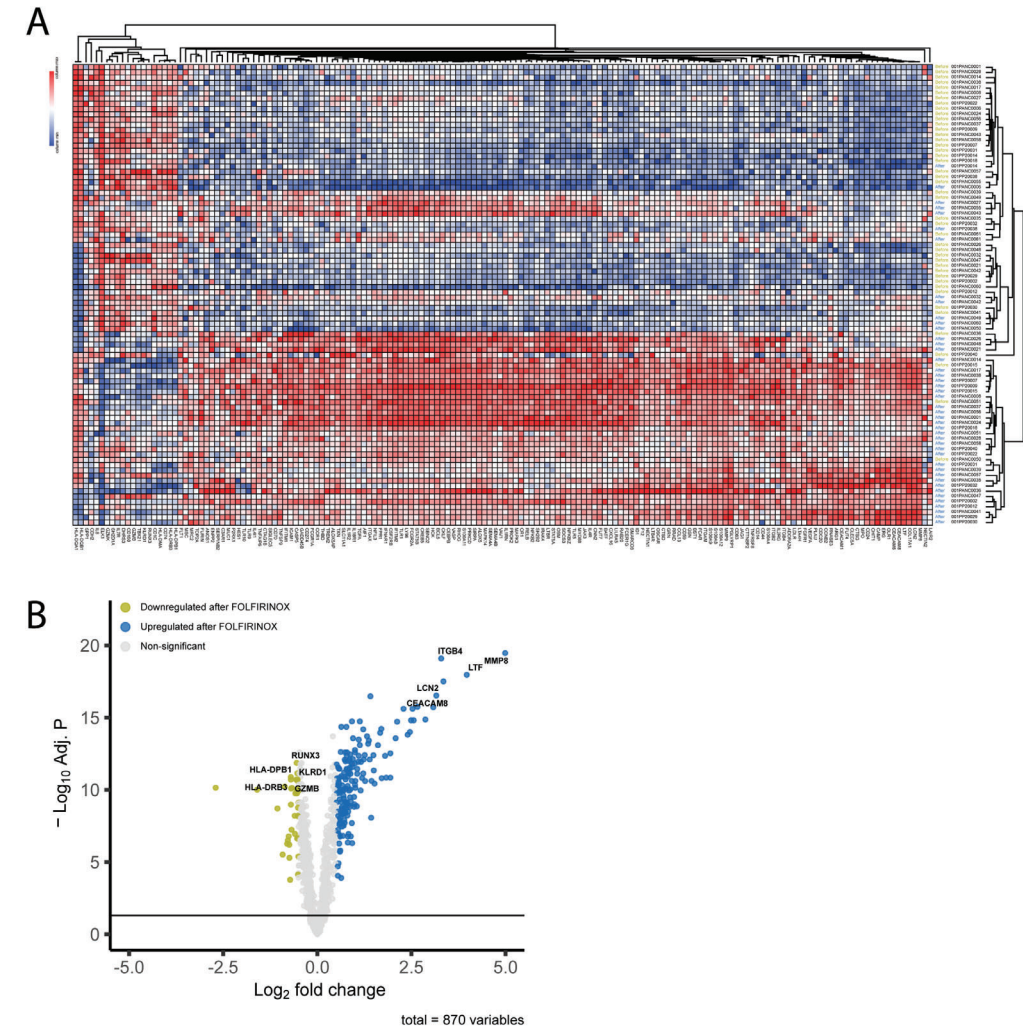
A

B

Figure 5: The effect of one cycle FFX-Lipeg on the gene expression profile. (A) The clustering of the genes showed an almost distinct genetic profile after one cycle. (B) A volcano plot highlighting the genes that got significantly altered (243 genes upregulated and 213 genes downregulated) after one cycle.

The most upregulated differentially expressed genes include Matrix metalloproteinase-8 (MMP8, P.BH = 3.38E-20, FOC = 4.99), Lactotransferrin (LTF, P.BH = 1.09E-18, FOC = 3.97) and Carcinoembryonic Antigen-Related Cell Adhesion Molecule 8 (CEACAM8, P.BH = 1.58E-19, FOC = 3.16), which are mainly expressed by neutrophils. MMP8 is involved in the breakdown of extracellular matrix and cleaves numerous substrates, including collagens and cytokines [35–37]. LTF is an iron-binding protein, uniquely abundant in polymorphonuclear neutrophils [38]. CEACAM8 is a surface glycoprotein that plays a role in heterophilic cell adhesion within activated neutrophils together with other carcinoembryonic antigen-related cell adhesion molecules, such as CEACAM6 [39]. Furthermore, Integrin Subunit Beta 4 (ITGB4, P.BH = 8.05E-20, FOC = 3.29) and Lipocalin 2 (LCN2, P.BH = 3.12E-18, FOC = 3.35) were also found upregulated. ITGB4 promotes cell migration and invasion in pancreatic cancer, however, the exact role in these processes remains unclear [40]. LCN2 is an iron-trafficking protein involved in multiple processes such as apoptosis, innate immunity, and renal development [41].

On the other hand, (cytotoxic) T-cell and natural killer cell-specific genes, RUNX Family Transcription Factor 3 (RUNX3, P.BH = 1.37E-12, FOC = -0.549), Granzyme B (GZMB, P.BH = 2.43E-11, FOC = -0.527), and Killer Cell Lectin Like Receptor D1 (KLRD1, P.BH = 1.96E-11, FOC = -0.525) were downregulated consistently in all samples after treatment. Furthermore, HLA class genes like Major Histocompatibility Complex, Class II, DP Beta 1 (HLA-DPB1, P.BH = 1.30E-11, FOC = -0.704), and Major Histocompatibility Complex, Class II, DR Beta 3 (HLA-DRB3, P.BH = 1.96E-11, FOC = -0.704) were also downregulated.

DISCUSSION

In this study, we used paired blood samples of 44 FFX-Lipeg treated PDAC patients to scrutinize the influence of using flow cytometry or targeted immune-gene expression to study the immunological changes. Flow cytometry and targeted gene expression profiling revealed a similar effect caused by a single cycle of FFX-Lipeg regarding granulocytes and monocytes. However, the measurement technique affects the observed changes regarding lymphocytes.

FFX-Lipeg treatment increased the number of neutrophils and monocytes as shown by flow cytometry and complete blood count in concordance with elevated gene expression measured by targeted gene expression profiling analysis. Contrarily, previously reported results showed a reduction of granulocytes and monocytes due to FOLFIRINOX treatment [12]. Nevertheless, studies on the effect of modified FFX-Lipeg show an increase in granulocytes, caused by lipegfilgrastim. The addition of lipegfilgrastim has been shown to decrease the incidence of neutropenic events and prolonged the progression-free survival of the patients [42,43].

Interestingly, flow cytometry analysis showed an increase in the number of B and T cells after treatment, while targeted gene expression analysis showed a decrease in B and T cell-specific gene expression. Both granulocytes and monocytes are potent suppressors of T cell functions and inhibit anti-tumor immune responses [44,45]. This could explain that even though the number of lymphocytes showed an increase after treatment in the flow cytometry data and complete blood count measurements, there is a decrease in cell-specific gene expression based on the gene expression analysis.

The correlation analysis highlights that the effect of FFX-Lipeg therapy influences the number of cells differently than the cell-specific gene expression. This could indicate that regardless of the number of lymphocytes, the function of those immune cells cannot be fulfilled. The pathway analyses further highlighted this downregulation of the functions of the lymphocytes. Based on the gene expression analysis, it seems that lymphocytes have a lower expression of functional genes after FFX-Lipeg treatment.

In this study, we measured 1,230 immune-related genes by targeted gene expression profile and 18 immune cell types by flow cytometry. The added value of targeted gene expression analysis is shown by the discovery that a single FFX-Lipeg cycle changed the expression of 209 immune-related genes significantly and caused a distinct genetic profile between the samples before and after treatment. Targeted gene expression profiles that specifically measure part of the immune-related genes enabled us to identify an FFX-ΔGEP score to predict the lack of treatment response after a single FFX-Lipeg cycle, as described [23]. As far as we know, measuring with flow cytometry did not lead to a similar discovery.

Besides the biological explanation, several technical factors might influence the discrepancy between the two measurement techniques. The higher detection of lymphocytes measured by flow cytometry could be due to non-specific bindings of lower-quality antibodies. Whereas the lower detection of immune cell-specific gene expression could be caused by the amount of RNA present in the samples, as well as mRNA stability. Furthermore, the comparison was performed on heterogenous samples and different types of blood samples (EDTA/Tempus).

CONCLUSION

Flow cytometry could be used for precise quantification of immune cell populations, whereas gene expression analysis gave a broader understanding of the immune expression activity of those cells. This highlights that measuring the number of cells in the blood does not reflect the immune functionality of these cells. To study the effect of treatment, different techniques must be used to get a more complete overview. This study revealed that the measurement technique influences clinical discoveries.

REFERENCES

1. Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2022. *CA Cancer J Clin* **72**, 7–33 (2022).

2. van der Sijde, F. *et al.* Treatment Response and Conditional Survival in Advanced Pancreatic Cancer Patients Treated with FOLFIRINOX: A Multicenter Cohort Study. *J Oncol* **2022**, (2022).

3. Conroy, T. *et al.* FOLFIRINOX or Gemcitabine as Adjuvant Therapy for Pancreatic Cancer. *N Engl J Med* **379**, 2395–2406 (2018).

4. Thibodeau, S. & Voutsadakis, I. A. FOLFIRINOX Chemotherapy in Metastatic Pancreatic Cancer: A Systematic Review and Meta-Analysis of Retrospective and Phase II Studies. *J Clin Med* **7**, (2018).

5. Suker, M. *et al.* FOLFIRINOX for locally advanced pancreatic cancer: a systematic review and patient-level meta-analysis. *Lancet Oncol* **17**, 801–810 (2016).

6. Chun, J. W. *et al.* A real-world analysis of nanoliposomal-irinotecan with 5-fluorouracil and folinic acid as third- or later-line therapy in patients with metastatic pancreatic adenocarcinoma. *Ther Adv Med Oncol* **14**, 175883592211195 (2022).

7. Irisawa, A. *et al.* Incidence of and risk factors for severe neutropenia during treatment with the modified FOLFIRINOX therapy in patients with advanced pancreatic cancer. *Sci Rep* **12**, 15574 (2022).

8. Conroy, T. *et al.* FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. *N Engl J Med* **364**, 1817–1825 (2011).

9. Timmer-Bonte, J. N. H. *et al.* Lipegfilgrastim for prophylaxis of chemotherapy-induced neutropenia in Dutch patients. *Neth J Med* **78**, 270–276 (2020).

10. Terazawa, T. *et al.* Efficacy of Prophylactic G-CSF in Patients Receiving FOLFIRINOX: A Preliminary Retrospective Study. *Intern Med* **54**, 2969–2973 (2015).

11. Steger, G. *et al.* Use of lipegfilgrastim for the prophylaxis of chemotherapy-induced neutropenia: Pan-European non-interventional study. *Annals of Oncology* **29**, viii607–viii608 (2018).

12. Peng, H. *et al.* Neoadjuvant FOLFIRINOX Therapy Is Associated with Increased Effector T Cells and Reduced Suppressor Cells in Patients with Pancreatic Cancer. *Clinical Cancer Research* **27**, 6761–6771 (2021).

13. Zhou, C. *et al.* Camrelizumab plus carboplatin and pemetrexed versus chemotherapy alone in chemotherapy-naive patients with advanced non-squamous non-small-cell lung cancer (CameL): a randomised, open-label, multicentre, phase 3 trial. *Lancet Respir Med* **9**, 305–314 (2021).

14. Principe, D. R. *et al.* Long-Term Gemcitabine Treatment Reshapes the Pancreatic Tumor Microenvironment and Sensitizes Murine Carcinoma to Combination Immunotherapy. *Cancer Res* **80**, 3101–3115 (2020).

15. Maecker, H. T., McCoy, J. P. & Nussenblatt, R. Standardizing immunophenotyping for the Human Immunology Project. *Nature Reviews Immunology 2012 12:3* **12**, 191–200 (2012).

16. McKinnon, K. M. Flow Cytometry: An Overview. *Curr Protoc Immunol* **120**, (2018).

17. Costantini, A. *et al.* Effects of cryopreservation on lymphocyte immunophenotype and function. *J Immunol Methods* **278**, 145–155 (2003).

18. Campbell, D. E. *et al.* Cryopreservation decreases receptor PD-1 and ligand PD-L1 coinhibitory expression on peripheral blood mononuclear cell-derived T cells and monocytes. *Clin Vaccine Immunol* **16**, 1648–1653 (2009).

19. Weinberg, A. *et al.* Optimization and limitations of use of cryopreserved peripheral blood mononuclear cells for functional and phenotypic T-cell characterization. *Clin Vaccine Immunol* **16**, 1176–1186 (2009).

20. Bagaev, A. *et al.* Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* **39**, 845-865.e7 (2021).

21. Danaher, P. *et al.* Gene expression markers of Tumor Infiltrating Leukocytes. *J Immunother Cancer* **5**, 18 (2017).

22. de Koning, W. *et al.* Identification, Validation, and Utilization of Immune Cells in Pancreatic Ductal Adenocarcinoma Based on Marker Genes. *Front Immunol* **12**, 1449 (2021).

23. van Eijck, C. W. F. *et al.* A multigene circulating biomarker to predict the lack of FOLFIRINOX response after a single cycle in patients with pancreatic ductal adenocarcinoma. *Eur J Cancer* **181**, 119–134 (2023).

24. Lyman, G. H., Yau, L., Nakov, R. & Krendyukov, A. Overall survival and risk of second malignancies with cancer chemotherapy and G-CSF support. *Ann Oncol* **29**, 1903–1910 (2018).

25. Kunert, A. *et al.* CD45RA+CCR7– CD8 T cells lacking co-stimulatory receptors demonstrate enhanced frequency in peripheral blood of NSCLC patients responding to nivolumab. *J Immunother Cancer* **7**, 149 (2019).

26. van der Sijde, F. *et al.* RNA from stabilized whole blood enables more comprehensive immune gene expression profiling compared to RNA from peripheral blood mononuclear cells. *PLoS One* **15**, (2020).

27. Geiss, G. K. *et al.* Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature Biotechnology 2008 26:3* **26**, 317–325 (2008).

28. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**, 1–12 (2002).

29. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications 2019 10:1* **10**, 1–10 (2019).

30. Bindea, G. *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091 (2009).

31. R Core Team. R: A Language and Environment for Statistical Computing. Preprint at https://www.R-project.org/ (2022).

32. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2016).

33. Blighe, K., Rana, S. & Lewis, M. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. Preprint at https://github.com/kevinblighe/EnhancedVolcano (2021).

34. Tukey, J. W. Exploratory data analysis. Addison-Wesley, Reading, MA. *Exploratory data analysis. Addison-Wesley, Reading, MA.* (1977).

35. Sirniö, P. *et al.* High-serum MMP-8 levels are associated with decreased survival and systemic inflammation in colorectal cancer. *Br J Cancer* **119**, 213–219 (2018).

36. Kruzel, M. L., Zimecki, M. & Actor, J. K. Lactoferrin in a Context of Inflammation-Induced Pathology. *Front Immunol* **8**, (2017).

37. Ribon, M., Mussard, J., Semerano, L., Singer, B. B. & Decker, P. Extracellular Chromatin Triggers Release of Soluble CEACAM8 Upon Activation of Neutrophils. *Front Immunol* **10**, (2019).

38. Zhao, X. *et al.* Beneficial Role of Neutrophils Through Function of Lactoferrin After Intracerebral Hemorrhage. *Stroke* **49**, 1241–1247 (2018).

39. Yamanaka, T., Kuroki, M., Kinugasa, T., Matsuo, Y. & Matsuoka, Y. Preparation and Characterization of Two Human Carcinoembryonic Antigen Family Proteins of Neutrophils, CD66b and c, in Silkworm Larvae. *Protein Expr Purif* **7**, 438–446 (1996).

40. Meng, X. *et al.* Integrin beta 4 (ITGB4) and its tyrosine-1510 phosphorylation promote pancreatic tumorigenesis and regulate the MEK1-ERK1/2 signaling pathway. *Bosn J Basic Med Sci* (2019) doi:10.17305/bjbms.2019.4255.

41. Yang, J. *et al.* An Iron Delivery Pathway Mediated by a Lipocalin. *Mol Cell* **10**, 1045–1056 (2002).

42. Yamao, K. *et al.* Clinical Safety and Efficacy of Secondary Prophylactic Pegylated G-CSF in Advanced Pancreatic Cancer Patients Treated with mFOLFIRINOX: A Single-center Retrospective Study. *Internal Medicine* **58**, 1993 (2019).

43. Carvalho de Brito, A. B., Riechelmann, R. P. & Fonseca de Jesus, V. H. Impact of Granulocyte Colony-Stimulating Factor (G-CSF) on the Outcomes of Patients With Metastatic Pancreatic Adenocarcinoma (MPA) During First-Line Treatment With FOLFIRINOX: A Single-Center Retrospective Analysis. *Cancer Control* **30**, (2023).

44. Bronte, V. *et al.* Recommendations for myeloid-derived suppressor cell nomenclature and characterization standards. *Nature Communications 2016 7:1* **7**, 1–10 (2016).

45. Nywening, T. M. *et al.* Targeting tumour-associated macrophages with CCR2 inhibition in combination with FOLFIRINOX in patients with borderline resectable and locally advanced pancreatic cancer: a single-centre, open-label, dose-finding, non-randomised, phase 1b trial. *Lancet Oncol* **17**, 651–662 (2016).

Chapter 6

Characterizing the tumor immune microenvironment of ependymomas using targeted gene expression profiles and RNA sequencing

W. de Koning[1,2], F.F. Feenstra[3], F.G.J. Calkoen[3], J. van der Lugt[3], L.A. Kester[3], D.A.M. Mustafa[1]

[1]Tumor Immuno-Pathology Laboratory, Department of Pathology & Clinical Bioinformatics, Erasmus University Medical Centre, Rotterdam, Netherlands

[2]Clinical Bioinformatics Unit, Department of Pathology & Clinical Bioinformatics, Erasmus University Medical Centre, Rotterdam, Netherlands

[3]Princess Máxima Center for Pediatric Oncology, Utrecht, Netherlands

ABSTRACT

Background: Defining the tumor immune microenvironment (TIME) of patients using transcriptome analysis is gaining more popularity. Here, we examined and discussed the pros and cons of using RNA sequencing for Fresh Frozen (FF) samples and targeted gene expression immune profiles (NanoString) for Formalin-Fixed, Paraffin-Embedded (FFPE) samples to characterize the TIME of ependymoma samples.

Results: Our results showed a stable expression of the 40 housekeeping genes throughout all samples. The Pearson correlation of the endogenous genes was high. To define the TIME, we first checked the expression of the *PTPRC* gene, known as *CD45*, and found it was above the detection limit in all samples by both techniques. T cells were identified consistently using the two types of data. In addition, both techniques showed that the immune landscape was heterogeneous in the 6 ependymoma samples used for this study.

Conclusions: The low abundant genes were detected in higher quantities using the NanoString technique, even when FFPE samples were used. RNA sequencing is better suited for biomarker discovery, fusion gene detection, and getting a broader overview of the TIME. The technique that was used to measure the samples had a considerable effect on the type of immune cells that were identified. The limited number of tumor-infiltrating immune cells compared to the high density of tumor cells in ependymoma can limit the sensitivity of RNA expression techniques regards the identification of the infiltrating immune cells.

Keywords: Tumor immune microenvironment, Targeted gene expression immune profiles, RNA sequencing, NanoString, Transcriptome analysis, Ependymoma

BACKGROUND

In the past years, the interest in molecular targeted therapy is rising. However, it has primarily focused on genomics. Transcriptome analysis is a high-accuracy strategy to define the tumor immune microenvironment (TIME) of patients based on RNA [1]. The TIME is known to be related to cancer progression and therapeutic outcomes [1]. Transcriptome analysis is a useful technique to study different cellular processes, such as immune responses and cell types that are present in the TIME [1].

Over the last decades, RNA sequencing is the most used transcriptomic analysis to understand genomic functions [2,3]. Fresh frozen (FF) samples are used to sequence the whole transcriptome via amplification. RNA sequencing can be used to find biomarkers since it does not require specific probes [4]. Profiling tumors with RNA sequencing can provide insights regarding classification and progression. The four main steps are [1] mRNA transcript fragmentation, followed by random primer binding, [2] cDNA synthesis via reverse transcription of the mRNA, [3] tagging the ends with a phosphate group and poly(A) tail and [4] ligation of adapters which enables PCR amplification and sequencing. Although this method is the most used, it has several disadvantages. High-quality RNA is needed for the amplification step of RNA sequencing, and due to the low signal/noise ratio, some transcripts are difficult to detect [4]. Furthermore, analyzing these RNA sequence data could be expensive, and it needs technical and bioinformatical skills [5].

In contrast to RNA sequencing, NanoString gene expression analysis is targeted and measures a selected set of genes. In addition, NanoString targeted gene expression analysis does not require amplification, which reduces workflow errors and increases reproducibility [6,7]. Formalin-fixed paraffin-embedded (FFPE) samples are used to perform targeted gene expression analysis [6]. This method is used in diagnostics, and it can measure up to 800 genes and even low counts [5,6]. NanoString's nCounter is a hybridization-based method, instead of amplification-based RNA sequencing, to detect RNA transcripts [7]. Different gene panels can be ordered via NanoString, for instance, the PanCancer Immune Profiling Panel [8]. NanoString targeted gene expression analysis is known to be a robust method with minimal background. This method can identify genes regardless of low-quality RNA or less-than-ideal FFPE preparation before gene expression analysis [5]. In addition, only a little amount of RNA (approximately 25ng) is needed for this method [5]. The workflow is more user-friendly than the workflow of RNA sequencing since it does not include library preparation[4]. Furthermore, NanoString provides software to analyze your RNA data, which is called nSolver™. This nSolver™ software can identify cellular processes and cell types based on RNA expression levels [9].

Ependymomas account for 8-10% of pediatric brain tumors, and the standard therapy consists of surgery (as radical as possible) and radiation therapy [10–12]. However, this treatment remains unchanged for the past two decades [10–12]. Although 50-70% of the tumors are successfully treated with surgery and radiation therapy, no standard care therapy is

available for recurrent or persistent ependymomas resulting in a dismal prognosis for these patients [13]. In addition, recurrence occurs in almost 50% of ependymoma patients [10]. Therefore, there is a need for alternative treatments for patients with ependymoma. Nine different subgroups of ependymomas can be defined based on the location of the tumor, genetics, and epigenetic DNA methylation [10,14]. The focus of this study is on posterior fossa group A (PF-A) ependymomas, which occur in children (aged 0-18 years old). PF-A ependymomas are located in the cerebellum and arise from regional radial glial-like cells [15]. PF-A ependymomas are known to be aggressive due to low mutation burden and high activation of several pathways, such as proliferation and angiogenesis, leading to a poor prognosis [13]. Another aggressive ependymoma subgroup is located in supertentorium (ST). In this study, both subgroups were included for analysis.

Knowledge of the immune system has great importance for treating cancer patients with immune therapies. In recent years, the Food and Drugs Administration (FDA) approved immune checkpoint inhibitors for solid tumors in adults and chimeric antigen receptor (CAR) T-cell therapy for children with leukemia [11]. However, immune therapies are challenging in brain tumors since the brain is protected by the blood-brain barrier (BBB), which is known to limit the infiltration of therapies [16]. To date, little is known about the TIME of ependymomas [17,18]. The TIME of pediatric CNS tumors tends to be immune suppressive [19,20], and there are indications that the TIME of ependymomas might also be immune suppressive or 'cold' TIME, indicating a lack of tumor-infiltrating T-cells [21].

Previous studies have demonstrated the high correlation between RNA sequencing and the NanoString gene expression analysis [5,22,23]. However, none of these studies focus on specific cancer tissue, nor did the studies investigate the TIME. This study aims to highlight the (dis)concordance in the identification of the TIME in a cold tumor, like ependymomas, using gene expression data generated by measuring bulk RNA with the two most common techniques: targeted gene expression and RNA sequencing.

MATERIAL AND METHODS

Sample collection and processing

Six formalin-fixed paraffin-embedded (FFPE) and fresh frozen (FF) ependymoma tumor samples (n=2 ST, n=4 PF-A ependymomas) from the same tumor were collected at the Princess Maxima Center (Utrecht, The Netherlands). Each sample is a primary tumor before either radiation therapy or chemotherapy and was collected between 2019 and 2020. The BioBank committee of the Princess Maxima Center approved the application. In addition, this study is in line with the declaration of Helsinki.

NanoString Immune profiling

Six FFPE ependymoma tumor samples were collected. Before RNA isolation the samples were sectioned with HM 340E Electronic Rotary Microtome. The whole sample was used to include the TIME. First, samples were put at -15°C. Sections of 10μm were cut to finally have 100-200μm of each sample. These sections were put in a 42°C water batch and afterward added to a slide. The slides were dried overnight before deparaffinization was performed. The nCounter® PanCancer Immune Profiling Panel was used for NanoString targeted gene expression analysis of six FFPE ependymoma tumor samples. This panel consists of 730 genes and 40 housekeeping genes targeting both innate and adaptive immune cells, and different pathways such as checkpoint signaling and antigen processing, which are considered an important part of the TIME (10). The total RNA was isolated from tumor tissue using the RNeasy® FFPE isolation kit (Qiagen, Leiden, The Netherlands). The RNA quantity and quality were measured using the Agilent 2100 BioAnalyzer (Santa Clara, CA, USA). RNA concentration was corrected to include fragments ≥ 300 bp. For each sample, 300 ng of RNA was hybridized with the PanCancer Immune Profiling probes for 17 hours at 67⁰C, following the manufacturing procedure (NanoString Technologies Inc., Seattle, WA, USA). The nCounter® FLEX platform was used to wash the extra probes, and genes were counted by scanning 490 Fields-of-view (FOV).

Data analysis of NanoString Immune profiling

The raw data of gene counts were uploaded to the nSolver™ Data Analysis software (version 4.0, NanoString, Seattle, WA, USA). Genes that had an expression level below the average count of the negative controls plus two standard deviations are considered undetected. The gene counts were normalized using the most stable housekeeping genes using the Advanced Analysis module (version 2.0) of nSolver™.

RNA sequencing

Six fresh frozen (FF) ependymoma tumor samples were collected. For all samples, 300 ng input material was collected and processed using RNA sequencing transcriptome

analysis at the Princess Maxima Center. The RNA sequencing library was prepared with the Roche KAPA hyperprep kit, including the amplification step. The library was sequenced with Illumina Novaseq 6000 using 2x150bp sequencing. The paired-end sequencing reads were aligned to the human reference genome (GRCh38.p12, hg38) using STAR [24] and annotated with transcript annotation (Gencode Release 31). Transcript quantification was performed using Subread featureCounts [25]. The counts were gene length trimmed mean of M-values (geTMM) normalized afterward [26].

Housekeeping genes

The PanCancer Immune Profiling Panel includes 40 housekeeping genes. These housekeeping genes were tested in the RNA sequencing and NanoString results for their stability using the geNorm algorithm [27]. The algorithm identified a minimum number of genes required to calculate a normalization factor as a geometric mean, which was used in the NanoString normalization. The most stable housekeeping genes from both techniques were compared by looking at the variation in the gene expression between samples.

Cell identification

To characterize the immune cell infiltration, gene markers were identified [28]. The gene markers were selected by calculating the pairwise similarity between all pairs of candidate marker genes that were above the detection limit in at least 50% of the samples. The gene pairs with a pairwise similarity above 0.6 were selected to describe the immune cells. Each immune cell type needed at least two unique genes. The abundance of the immune cell types is the average expression value of their corresponding marker genes.

The identification of tumor-infiltrating leukocytes (TILs) with the RNA sequencing data was performed using the CIBERSORT method [29] and the MCP-counter [30]. CIBERSORT combines support vector regression with prior knowledge from expression profiles from purified leukocyte subsets to estimate the immune composition. The validated leukocyte gene signature matrix (LM22) including 547 marker genes was used to quantify 22 human hematopoietic subsets. The absolute mode was run together with 1000 permutations without quantile normalization as recommended by the developer [29]. MCP-counter is a method that allows absolute abundance calculations of eight immune and two stromal cell populations with the use of transcriptomics markers [30].

Immunohistochemistry was performed at the UMC Utrecht Pathology department on 5-μm FFPE tumor tissue sections using a Ventana Immunostainer.

Statistical analysis

The official gene symbols approved by the HGNC were used to identify matches between the NanoString and RNA sequence datasets. The Pearson correlations of the samples between the two platforms were computed with R software, version 4.1.1 [31]. To detect the genes that differ the most from the expected correlation a linear model for each sample was created and the residuals were reported.

RESULTS

Patients' characteristics

A total of six patients were included that were diagnosed with ependymoma. The FFPE samples and Fresh Frozen samples of the 6 patients were collected and RNA was successfully isolated. The patients were between 0 and 13 years old (mean age is 4 years). All clinicopathological characteristics are summarized in Table 1.

*Table 1 Clinicopathological characteristics at the time of diagnosis from the six patients included. * Sample clusters as RELA but a ZFTA or RELA fusion was not detected.*

Sample	Age	Sex	Diagnosis	Tumor location	Resection location	Molecular subgroup	Fusion status
EPN-01	13	F	WHO III	ST	Supratentorial right frontal		RELA fusion*
EPN-02	2	F	WHO III	ST	Supratentorial left frontal		RELA fusion
EPN-03	1	M	WHO III	PF	Fossa cranii posterior	Group 1	PF-A
EPN-04	2	M	WHO III	PF	Fossa cranii posterior	Group 2	PF-A
EPN-05	2	M	WHO III	PF	Fossa cranii posterior	Group 2	PF-A
EPN-06	3	M	WHO III	PF	4th ventricle	Group 2	PF-A

Characterization of ependymoma using NanoString

Of the 770 genes included in the nCounter® PanCancer Immune Profiling Panel of NanoString Technology, 752 genes were detected in at least one sample. The total number of genes detected per sample was comparable between the six samples. However, the number of the detected genes was lower in three samples, namely EPN-04, EPN-05, and EPN-06 (Supplementary Figure 1A).

Characterization of ependymoma using RNA sequencing

In total 49,136 unique RNA molecules were detected in at least one sample by RNA sequencing out of the 58,804 features described in the transcript annotation. This included 19,051 protein-coding sequences (Table 2). The protein-coding sequences are the only features that can be compared to the 770 genes included in the NanoString panel.

Table 2 The number of unique RNA molecules detected in at least one sample per RNA molecule category by RNA-sequencing.

Category	Frequency
Protein coding	19051
Long non-coding RNA	14696
Processed pseudogenes	10219
Others	2860
Small nuclear RNA	1238
Micro RNA	1072

The total number of features detected per sample was comparable between the six samples. However, the number of detected features was the least in sample EPN-05 (Supplementary Figure 1B).

The expression stability of the housekeeping genes

The 40 assigned housekeeping genes in NanoString measurements were checked in the NanoString data. The twenty-four most stable housekeeping genes based on the expression ratio of the genes between all samples were selected by the geNorm algorithm [27] in the advanced analysis module of NanoString data (Figure 1A). Applying the same algorithm to select the most stable housekeeping genes in the RNA sequence data resulted in a selection of 26 genes (Figure 1B). Twenty housekeeping genes overlapped in the two selections.
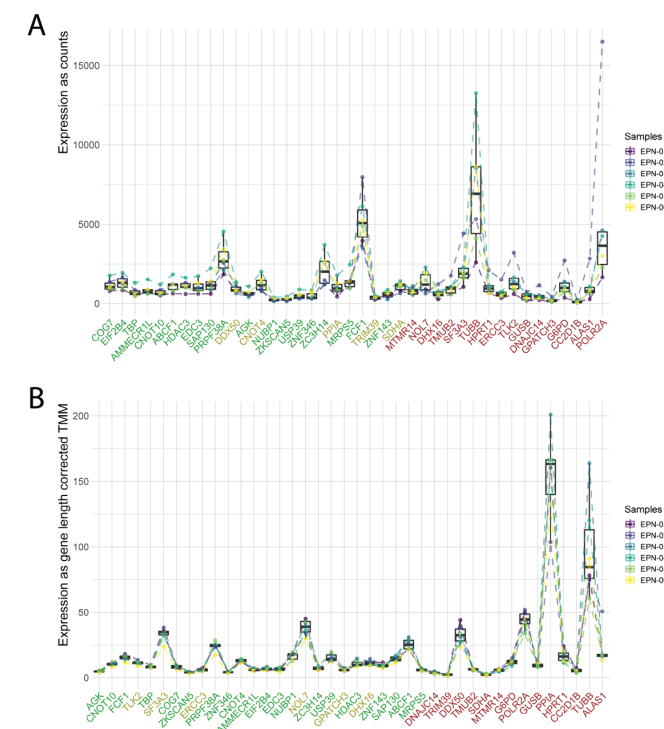

Figure 1 (A) The trend line of the forty housekeeping genes in NanoString samples. (B) The trend line of the forty housekeeping genes in RNA sequence samples.
Legend: *Green labels on the x-axis highlight the genes that were selected in both techniques, yellow is unique for the technique, and red means it was not selected for both techniques.*

The detection level of the nCounter® PanCancer Immune Profiling Panel

The 730 genes included in the nCounter® PanCancer Immune Profiling Panel were examined in both datasets. From the 730 genes, 722 were identified in the RNA sequencing data. Interestingly, 7 of the 8 genes that were not identified in the RNA sequence data were above the detection limit in NanoString data. However, one gene (*KIR3DS1*) was not detected in NanoString nor identified in RNA sequencing (Table 3).

Table 3 The NanoString gene expression values of the genes that were not identified with RNA sequencing.

Gene Symbol	Gene symbol NanoString	Alternative gene symbol	Average	Min.	Max.	Detected in # samples
BAGE	BAGE	CT2.1, BAGE1	52.7	undetected	101.3	3
HLA-DRB3	HLA-DRB3	HLA-DR3B	11030.93	1020.6	28529.6	6
HLA-DRB4	HLA-DRB4	DR4, DR-4, DRB4, HLA-DR4B	779.0	undetected	1399.1	3
KIR3DS1	KIR_Activating_ Subgroup_1	nkat10	undetected	undetected	undetected	0
KIR2DS1	KIR_Activating_ Subgroup_2	CD158H, EB6ActI, EB6ActII	51.3	undetected	51.3	1
LTBR	LTBR	D12S370, TNFCR, TNFR-RP, TNFR2-RP, TNF-R-III, TNFRSF3	317.2	106.7	564.8	6
MCAM	MCAM	MUC18, CD146, MelCAM, METCAM, HEMCAM,	431.1	103.2	941.8	6
TARP	TARP	CD3G, TCRG, TCRGC1, TCRGC2	110.7	47.1	301.3	5

Comparability of NanoString and RNA sequence expression

To determine the comparability between the two techniques, the Pearson correlation between genes that were above the detection limit was performed, and the correlation coefficient (R^2) was calculated. EPN-01 is shown as an example in Figure 2.
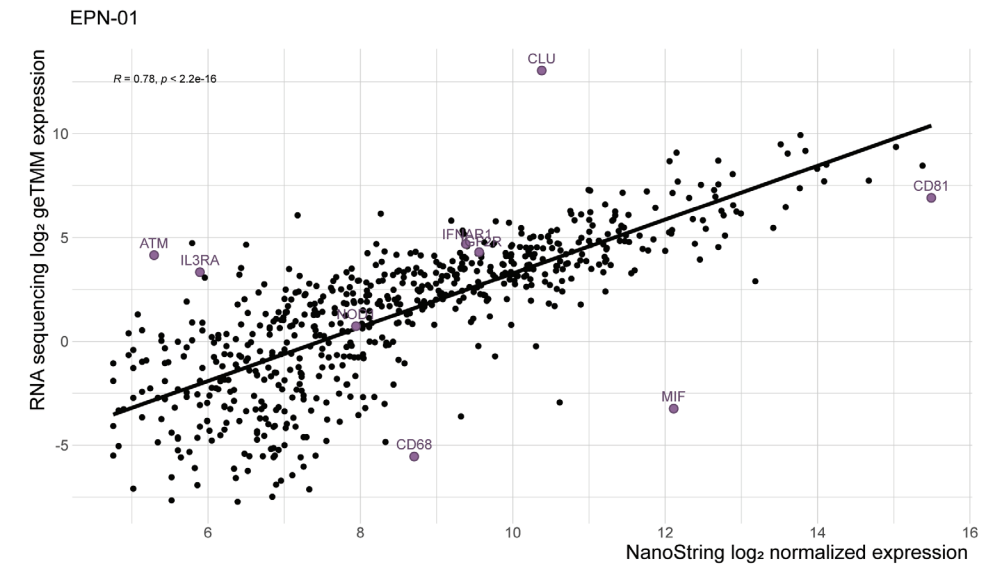
Figure 2 Pearson correlation of the genes that are above the detection limit and overlapping between both techniques in sample EPN-01. The three genes that are most divergent negatively (MIF, CD68, and CD81) and positively (CLU, ATM, and IL3RA) are highlighted next to the three genes that are the least divergent (IGF2R, NOD1, IFNAR1).

Interestingly, some genes like Clusterin (*CLU*), ATM Serine/Threonine Kinase (*ATM*), and Interleukin 3 Receptor Subunit Alpha (*IL3RA*) were detected at a higher level with the RNA sequence technique in all samples. On the other hand, Macrophage Migration Inhibitory Factor (MIF), Cluster of Differentiation 81 (*CD81*), Cluster of Differentiation 68 (*CD68*), and Cluster of Differentiation 81 (*CD81*) were detected at a higher level with the NanoString technique in all samples. The same results were confirmed by calculating the residuals, i.e., genes that fail to correlate, of the linear models per sample (Supplementary Figure 2, Supplementary Table 1). The average Pearson correlation of the genes detected by NanoString, and RNA sequencing was 0.82 (min. 0.78, max. 0.85), with an average R^2 of 0.67 (Supplementary Figure 3-7).

Dividing the 722 endogenous genes into bins based on their median expression from low to high showed an increased gene-specific inter-sample correlation for the highly expressed genes in both techniques. The average Spearman correlation was 0.3 across all 722 genes. In the NanoString data, 268 genes showed low expression (0-50 counts), 120 genes showed intermediate expression (50-200 counts) and 334 genes showed high expression (> 200). The Spearman correlation for the low-expressed genes (average R = 0.13) is significantly lower compared to the intermediate (average R = 0.34, p-value < 0.0001) and high-expressed genes (average R = 0.40, p-value < 0.0001). In the RNA sequence data, 261 genes showed low expression (0-1 count), 120 genes showed intermediate expression (1-3 counts) and 334 genes showed high expression (> 3 counts). The Spearman correlation for the low expressed genes (average R = 0.19, p-value < 0.0001)

and intermediate expressed genes (average R = 0.27, p-value < 0.001) are significantly lower compared to the high expressed genes (average R = 0.39), Supplementary Figure 8.

The detection limit of nCounter in comparison to RNA sequencing

From the 722 identified genes (out of 730 genes) in both platforms, 464 genes were expressed above the detection limit and therefore measured by both methods in all samples. The gene *TLR9* was detected in four samples, but the expression levels were below the detection limit in the other two samples for both methods. The other 257 genes were in at least one sample detected by one method but below the detection limit in the other of which four genes were only detected in all samples with RNA sequencing. There was not any gene only detected in all samples by NanoString (Supplementary Figure 9).

Comparison of the sensitivity of NanoString versus RNA sequencing

To examine the sensitivity of both techniques, genes that were above the detection limit for only one technique were investigated. Sixty genes were detected by NanoString technology but were under the detection limit by RNA sequencing in at least one sample. Remarkably, Macrophage Migration Inhibitory Factor (*MIF*) was detected with an expression value higher than 4,000 counts in EPN-06 by the NanoString technique (Figure 3A). On the contrary, 223 genes were only detected by RNA sequencing in at least one sample. The expression values from these genes vary from zero up to 28 (Figure 3B). The expression of *BLK*, *CCL28*, *CMA1*, and *FUT7* was not detected in any sample with NanoString but detected in all samples by RNA sequencing with expression values ranging from zero up to six (Figure 3C).
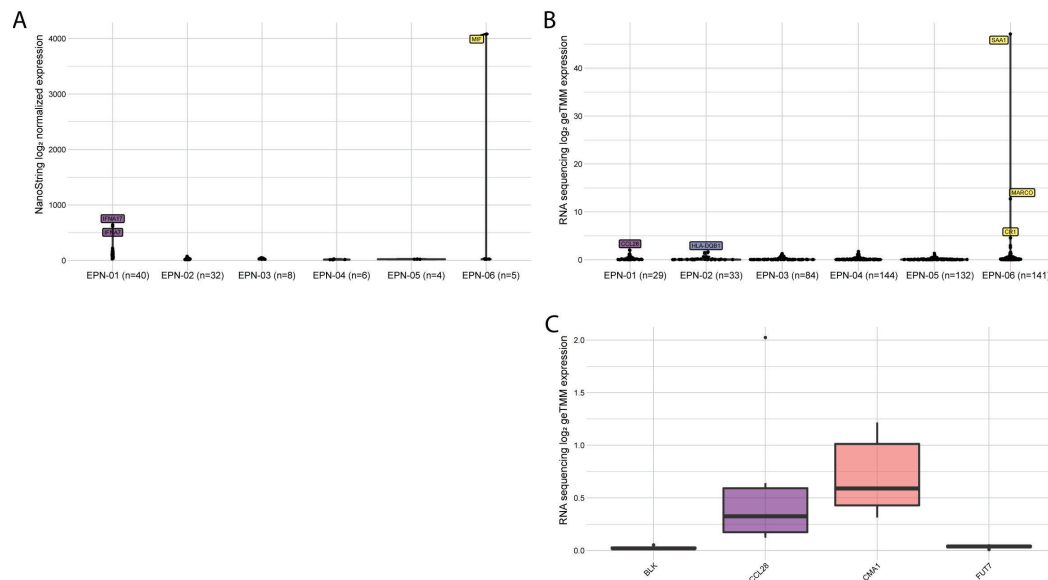
The identification of immune cells

The protein Tyrosine Phosphatase Receptor Type C (*PTPRC*) gene, also known as *CD45* was above the detection limit in all samples by NanoString and RNA sequencing, which reflects leukocytes infiltrated into the ependymoma tumor (Supplementary Table 2). To define the type of immune cells that infiltrated the ependymoma samples measured with NanoString, the pairwise similarities for the 55 candidate marker genes were calculated. Genes that showed acceptable pairwise similarities (>0.6) were selected as marker genes (Supplementary Document 1). In total, 25 genes were selected to identify 11 different immune cell types in ependymoma. These marker genes were then used to calculate the relative abundance per sample (Figure 4A). The same method was used to identify immune cells in the RNA sequencing data. The marker gene-based method was investigated to identify immune cell types using RNA sequencing data. However, the pairwise similarities of most cells were below 0.6. (Supplementary Document 2). Therefore, the marker gene-based method was only applicable to a small set of immune cell types in the RNA sequencing data (Figure 4B). The expression level of the selected marker genes was between zero and 2.1 B cells, plasma B cells, cytotoxic cells, and T cells were identified using the two types of data. Nevertheless, both techniques show a different pattern of immune infiltration. The second method, CIBERSORT, was only applicable to RNA sequence data, as it expects an expression panel of 547 genes to be measured. This mixture-based method resulted in significant estimations for EPN-03 and EPN-06 (Figure 4C). It showed infiltration of naïve b-cells, M1 macrophages, and neutrophils in EPN-03, whereas infiltration of M2 macrophages, monocytes, and resting NK cells was shown in EPN-06. A third method, MCP-counter, was used to determine the abundance of the immune cells based on the RNA sequence data. Similarly, MCP-counter expects a large set of 111 genes to be measured. Therefore, it was only applicable to RNA sequence data (Figure 4D). MPC-counter showed the highest immune infiltration in EPN-06, but similarly to the gene maker method, a high infiltration of T cells was found in EPN-01. Immunohistochemistry showed infiltration of CD3+ cells in EPN-01, but CD3+ cells were not found in EPN-05 and EPN-06.
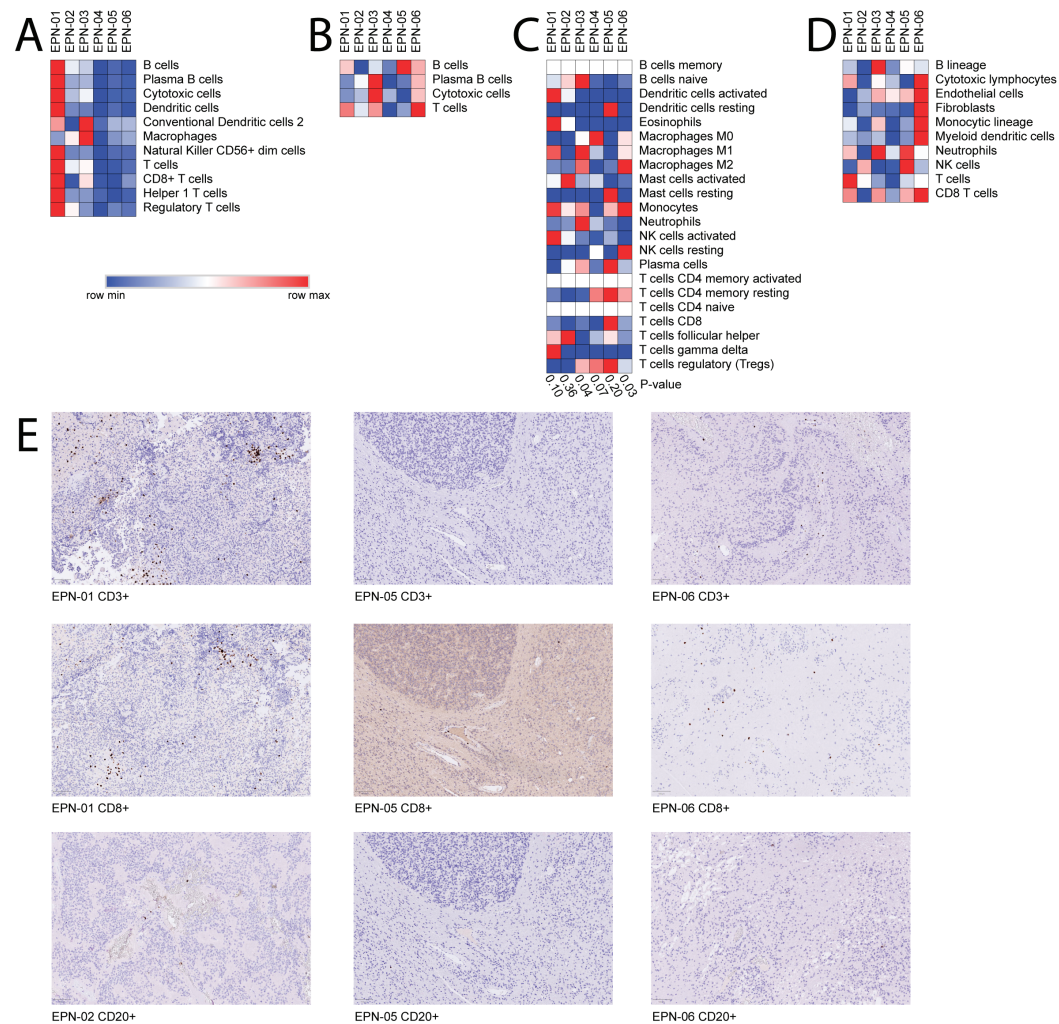
Figure 4 (A) The marker gene-based immune cell type abundances per sample with NanoString data (B) The marker gene-based immune cell type abundances per sample with RNA sequence data. (C) The CIBERSORT mixture-based immune cell abundances per sample in the RNA sequence data. (D) The MCP-counter mixture-based immune cell abundances per sample in the RNA sequence data. (E) The immunohistochemistry staining of CD3+ (EPN-01, EPN-05, and EPN-06), CD8+ (EPN-01, EPN-05, and EPN-06), and CD20+ (EPN-02, EPN-05, and EPN-06) cells within the samples.

DISCUSSION

The characterization of the ependymoma's TIME can be done with different techniques, each having its advantages and disadvantages. In that regard, the nCounter (PanCancer Immune Profiling Panel) of NanoString is compared to RNA sequencing to get a better understanding of these advantages and disadvantages. NanoString uses probes to measure up to 800 genes, whereas RNA sequencing is used to profile transcriptome-wide, including approximately 26,000 genes. RNA sequencing can be used to identify novel and rare transcripts like noncoding RNA and fusion RNA (Table 2). This makes RNA sequencing a suited tool for biomarker detection and measuring a broad scope of samples. The 40 housekeeping genes studied from both techniques showed a stable expression throughout all samples. The geNorm algorithm selected a similar number of housekeeping genes that were stable (Figure 1). This shows that both methods are comparable regarding these housekeeping genes, which is to be expected. Eight out of the 730 endogenous genes of the PanCancer Immune Profiling Panel were not available in the RNA sequencing data (Table 3). This is most likely due to the complexity of the downstream analysis with RNA sequencing. The measurement of mRNA with RNA sequencing has FASTQ files as output, after which the downstream analysis is started including quality control, mapping, and normalization. Each of these steps can introduce errors and therefore experienced bioinformatics knowledge is needed. The complexity of the downstream analysis of RNA sequencing can cause important information to be lost. Especially the quality control and mapping steps are tedious work. NanoString uses pre-defined panels with identifiers for each barcode detected, therefore the processing of the measurements is less error-prone.

The significant Pearson correlations of the genes detected by both techniques (avg. Pearson correlation of 0.71) show, like the housekeeping genes, that comparable results can be found for most of the genes (Figure 2). Nevertheless, there are discrepancies found between the two techniques. These discrepancies could be caused by the different types of samples used for each technique since FFPE samples were used for nCounter NanoString, whereas FF samples were used for RNA sequencing. Due to the frozen state of the FF samples, the RNA molecules are better preserved than in FFPE samples. The FF samples are recommended for RNA sequencing as high-quality RNA in a higher abundance is needed. On the other hand, PCR amplification can cause bias by amplifying specific genes. It is shown that the higher expressed genes have a more concordant expression between the two methods. The differences in the low abundant genes can be due to the high sensitivity of the nCounter from NanoString by using barcodes of 100bp, which makes it possible to measure low-abundant and lower-quality RNA more accurately even though FFPE samples were used [32]. In addition, RNA can only be measured with RNA sequencing after the RNA molecules undergo cDNA synthesis via reverse transcription of the mRNA and PCR amplification. In case the cDNA synthesis does not work properly the genes are less likely to be detected by RNA sequencing. The gene *ATM* (146,036 bases) was higher detected with RNA sequencing whereas *CD68* (2,621 bases)

was higher detected with NanoString. This could be explained by the amplification bias in RNA sequencing. During PCR amplification it is more likely to amplify a gene that has more reads present [33]. Nevertheless, this does not explain the different detection levels for *CLU* (17,784 bases) and *CD81* (21,221 bases). The 464 genes that were above the detection level in all samples confirm the comparability between the two techniques. The 257 genes that were detected by one method, but below the detection limit in the other in at least one sample, show the discrepancies that are caused by the technique or sample type. Although the RNA is expected to be more degraded in the FFPE samples sixty genes were detected by NanoString technology but were under the detection limit by RNA sequencing in at least one sample with expression values higher than 4,000 counts (Figure 3A). That these genes are not detected with RNA sequencing could be due to the instability of the RNA molecule, as RNA sequencing needs stable reads over the whole exon region to detect the gene. The 223 genes that were only detected by RNA sequencing in at least one sample had an expression value ranging from zero up to 28 (Figure 3B). *BLK*, *CCL28*, *CMA1*, and *FUT7* were not detected in any sample with NanoString but detected in all samples by RNA sequencing with expression values ranging from zero up to six (Figure 3C). These genes are most likely not detected by NanoString due to the degradation of the RNA molecule in FFPE.

The detection of *CD45* indicates that leukocytes infiltrated into the ependymoma tumor. Even though microglia are the most dominant immune cells in the brain (80%), other immune cell types have been identified in the brain including B cells, dendritic cells, macrophages, monocytes, myeloid cells, natural killer (NK) cells, and T cells [10,34–36]. Nevertheless, the limited number of tumor-infiltrating immune cells compared to the high density of tumor cells can limit the sensitivity of bulk techniques [37]. The estimation of immune cell types in the samples by NanoString and RNA sequencing shows mostly incomparable results (Figure 4). The marker gene method shows that the highest immune infiltration is in EPN-01 (Figure A), whereas the same method shows a higher immune infiltration in EPN-06 based on the RNA sequencing data. The low expression values in the RNA sequence data lead to low pairwise similarities and are therefore expected not to be accurate and are not suitable for most cell types using this method. The marker genes method cannot detect the absolute abundance of immune cells but can only be used to describe the relative abundance between samples. The panel-based method CIBERSORT is only applicable to RNA sequencing results as it uses a large panel of genes that are not included in the standard panels of NanoString. CIBERSORT results in the absolute cell abundance which can be used to examine the relative abundance between samples. The RNA sequencing results include more cell types than estimated with the NanoString panel. Therefore, a fair comparison regards the immune cell types cannot be made. MPC-counter is more comparable to the gene-marker method used for NanoString. In the RNA sequencing data, both the marker gene method and MPC-counter describe the highest immune infiltration in EPN-06. On the other hand, based on the NanoString data analyzed with the marker gene method and the RNA sequencing data analyzed with

the MPC-counter method show the highest infiltration of T cells in EPN-01. This has been confirmed with immunohistochemistry (Figure 4E).

The number of samples is limited in this study. Nevertheless, the aim of this study was to highlight the (dis)concordance in the identification of the TIME in a cold tumor, like ependymomas, using gene expression data generated by the two most common techniques. Therefore, the variation in the included cohort of samples can be considered positive rather than negative because it covers a wider range of immune-cold tissue samples. The identification of immune cells is sample-based. Therefore, the number of samples included to compare the two techniques has a limited effect on the presented results.

CONCLUSION

In conclusion, both methods have their advantages and disadvantages. RNA sequencing is better suited for biomarker discovery, getting a broad overview of the samples regarding the TIME, detecting fusion genes, and mutation detection, although the low abundance genes might be missed out. The NanoString technique is easier to understand due to the simplified pre-processing steps both in the laboratory and dry lab. Furthermore, NanoString allows you to detect low-abundant genes in the panel of interest. The technique that was used to measure the samples had a considerable effect on the type of immune cells that were identified.

REFERENCES

1. Bagaev, A. *et al.* Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* **39**, 845-865.e7 (2021).

2. Stark, R., Grzelak, M. & Hadfield, J. RNA sequencing: the teenage years. *Nat Rev Genet* **20**, 631–656 (2019).

3. Wang, B., Kumar, V., Olson, A. & Ware, D. Reviving the Transcriptome Studies: An Insight Into the Emergence of Single-Molecule Transcriptome Sequencing. *Front Genet* **10**, 384 (2019).

4. Bourré, L. Genome-Wide RNAseq and Array-Based NanoString Transcriptomic Technologies: Which To Use and When? 1 https://blog.crownbio.com/genome-wide-rnaseq-and-array-based-nanostring-transcriptomic-technologies (2020).

5. Speranza, E. *et al.* Comparison of Transcriptomic Platforms for Analysis of Whole Blood from Ebola-Infected Cynomolgus Macaques. *Sci Rep* **7**, 1–15 (2017).

6. M'Boutchou, M.-N. & van Kempen, L. C. Analysis of the Tumor Microenvironment Transcriptome via NanoString mRNA and miRNA Expression Profiling. *Methods Mol Biol* **1458**, 291–310 (2016).

7. Omolo, B. *et al.* Adaptation of a RAS pathway activation signature from FF to FFPE tissues in colorectal cancer. *BMC Med Genomics* **9**, 1–10 (2016).

8. NanoString. nCounter® PanCancer Immune Profiling Panel. 1 https://www.nanostring.com/products/ncounter-assays-panels/oncology/pancancer-immune-profiling/ (2021).

9. NanoString. Solver™ Data Analysis Support. https://www.nanostring.com/products/analysis-solutions/ncounter-analysis-solutions/nsolver-data-analysis-support/ (2021).

10. Hoffman, L. M. *et al.* Molecular sub-group-specific immunophenotypic changes are associated with outcome in recurrent posterior fossa ependymoma. *Acta Neuropathol* **127**, 731–745 (2014).

11. Witt, D. A. *et al.* Specific expression of PD-L1 in RELA-fusion supratentorial ependymoma: Implications for PD-1-targeted therapy. *Pediatr Blood Cancer* **65**, 1–9 (2018).

12. Mack, S. C. *et al.* Therapeutic targeting of ependymoma as informed by oncogenic enhancer profiling. *Nature* **553**, 101–105 (2018).

13. Reni, M., Gatta, G., Mazza, E. & Vecht, C. Ependymoma. *Crit Rev Oncol Hematol* **63**, 81–89 (2007).

14. Pajtler, K. W. *et al.* Molecular Classification of Ependymal Tumors across All CNS Compartments, Histopathological Grades, and Age Groups. *Cancer Cell* **27**, 728–743 (2015).

15. Vladoiu, M. C. *et al.* Childhood cerebellar tumours mirror conserved fetal transcriptional programs. *Nature* **572**, 67–73 (2019).

16. Sayour, E. J. & Mitchell, D. A. Immunotherapy for Pediatric Brain Tumors. *Brain Sci* **7**, 137 (2017).

17. Griesinger, A. M. *et al.* Characterization of Distinct Immunophenotypes across Pediatric Brain Tumor Types. *The Journal of Immunology* **191**, 4880–4888 (2013).

18. Nam, S. J. *et al.* Tumor-infiltrating immune cell subpopulations and programmed death ligand 1 (PD-L1) expression associated with clinicopathological and prognostic parameters in ependymoma. *Cancer Immunol Immunother* **68**, 305–318 (2019).

19. Wang, S. S., Bandopadhayay, P. & Jenkins, M. R. Towards Immunotherapy for Pediatric Brain Tumors. *Trends Immunol* **40**, 748–761 (2019).

20. Vermeulen, J. F. *et al.* Prognostic relevance of tumor-infiltrating lymphocytes and immune checkpoints in pediatric medulloblastoma. *Oncoimmunology* **7**, (2018).

21. Duan, Q., Zhang, H., Zheng, J. & Zhang, L. Turning Cold into Hot: Firing up the Tumor Microenvironment. *Trends Cancer* **6**, 605–618 (2020).

22. Zhang, W. *et al.* A large-scale comparative study of isoform expressions measured on four platforms. *BMC Genomics* **21**, 1–14 (2020).

23. Bondar, G. *et al.* Comparing NGS and NanoString platforms in peripheral blood mononuclear cell transcriptome profiling for advanced heart failure biomarker development. *J Biol Methods* **7**, e123 (2020).

24. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

25. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

26. Smid, M. *et al.* Gene length corrected trimmed mean of M-values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons. *BMC Bioinformatics* **19**, 1–13 (2018).

27. Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**, RESEARCH0034 (2002).

28. de Koning, W. *et al.* Identification, Validation, and Utilization of Immune Cells in Pancreatic Ductal Adenocarcinoma Based on Marker Genes. *Front Immunol* **12**, (2021).

29. Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* **12**, 453–457 (2015).

30. Becht, E. *et al.* Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol* **17**, 218 (2016).

31. R Core Team. R: A Language and Environment for Statistical Computing. Preprint at https://www.r-project.org/ (2020).

32. Veldman-Jones, M. H. *et al.* Evaluating Robustness and Sensitivity of the NanoString Technologies nCounter Platform to Enable Multiplexed Gene Expression Analysis of Clinical Samples. *Cancer Res* **75**, 2587–2593 (2015).

33. Gao, L., Fang, Z., Zhang, K., Zhi, D. & Cui, X. Length bias correction for RNA-seq data in gene set analyses. *Bioinformatics* **27**, 662 (2011).

34. Korin, B. *et al.* High-dimensional, single-cell characterization of the brain's immune compartment. *Nat Neurosci* **20**, 1300–1309 (2017).

35. Donson, A. M. *et al.* Immune Gene and Cell Enrichment Is Associated with a Good Prognosis in Ependymoma. *The Journal of Immunology* **183**, 7428–7440 (2009).

36. Gillen, A. E. *et al.* Single-Cell RNA Sequencing of Childhood Ependymoma Reveals Neoplastic Cell Subpopulations That Impact Molecular Classification and Etiology. *Cell Rep* **32**, 108023 (2020).

37. Rozowsky, J. S. *et al.* A Toolkit for Profiling the Immune Landscape of Pediatric Central Nervous System Malignancies. *Front Immunol* **13**, (2022).

6

6

Chapter 7

General discussion and future perspective

GENERAL DISCUSSION

Nowadays multiple omics techniques are applied to various sample types to determine the heterogeneity of the TIME [1-5]. Bioinformatics can improve the identification and quantification of differences between samples.

Immune cell identification

In this thesis, a method to identify and validate specific marker genes to define immune cells in any tissue type applied to PDAC tissue samples was shown. These markers were more PDAC-specific than the marker genes used to define immune cells across various types of tumor tissue samples (PanCancer marker genes) and enabled identifying eight additional immune cells. This method can be applied using gene expression data generated from samples that were preserved differently like FF and FFPE tissue samples or blood samples. It highlights the importance of selecting and testing the marker genes critically for each tissue type (Chapter 2). This method has been implemented in several clinical studies [6-9]. In the second study, paired blood samples of 68 PDAC patients were used to investigate the effect of a single FOLFIRINOX cycle on the immune profile. An application of the previously described method was used to identify the immune cells. Furthermore, the first gene expression-based circulating biomarker predicting the lack of FOLFIRINOX response in PDAC patients from all disease stages was identified. The biomarker entails an eight-gene FFX-ΔGEP score that predicted the lack of FOLFIRINOX response only after the first cycle, independent of disease stage or change in CA19-9. This highlights the value of targeted gene expression analysis on blood samples (Chapter 3). Furthermore, an application of the same targeted gene expression profiling in combination with flow cytometry was used to investigate the immune-modulatory effects of IMM-101/SBRT treatment in the peripheral blood. The targeted gene expression showed a downregulation of genes related to lymphocyte subsets induced by the treatment. This was confirmed by the flow cytometry data, which showed a transient decrease in different lymphocyte subsets and an increase in CD14+CD16−CD11b+HLA−DRlow myeloid-derived suppressor cells. Importantly, treatment significantly increased activated ICOS+, HLA-DR+, and Ki67+PD1+ T and NK cell frequencies (Chapter 4).

Comparing immune cell profiling measuring methods

As shown in the previous studies multiple techniques can be used to identify immune cells. Nevertheless, to get a better understanding of the advantages and disadvantages a comparison study between targeted gene expression, flow cytometry, and RNA bulk sequencing was necessary. The comparison between targeted gene expression and flow cytometry was performed on two types of whole blood samples that were collected from 44 PDAC patients before the first FOLFIRINOX- cycle and 14 days after the first cycle. EDTA blood samples were used for multiplex flow cytometry analyses and complete blood counts, whereas Tempus blood samples were used for targeted gene expression

analysis. Although similar results were found for granulocytes and monocytes, contradicting results were found for lymphocytes. This shows that different conclusions might be drawn based on the technique used. Flow cytometry and complete blood count measurements reflect the absolute number of immune cells, whereas gene expression analysis measures the total mRNA expression (Chapter 5). The characterization of the TIME can also be achieved with bulk RNA sequencing. Six FFPE ependymoma tumor samples measured with targeted gene expression were compared with six FF ependymoma tumor samples measured with bulk RNA sequencing from the same tumor. The low abundant genes were detected in higher quantities using the NanoString technique, even when FFPE samples were used. RNA sequencing is better suited for biomarker discovery, fusion gene detection, and getting a broader overview of the TIME. The technique that was used to measure the samples had a considerable effect on the type of immune cells that were identified. The limited number of tumor-infiltrating immune cells compared to the high density of tumor cells in ependymoma can limit the sensitivity of RNA expression techniques regards the identification of the infiltrating immune cells (Chapter 6). The comparison between RNA sequencing and targeted gene expression was performed in ependymoma, but similar conclusions can be drawn in PDAC samples.

The studies included in this thesis highlight the importance of choosing the right technique based on a well-defined hypothesis. Furthermore, it is of great importance to address the limitations and strengths of each technique in the discussion of each publication.

THE IMPACT ON THE PATIENTS

Improvements in the understanding of the single omics could lead to improved strategies to treat PDAC patients [10]. The findings in this thesis do not directly influence the treatment strategy. Nevertheless, the strengths and weaknesses of each of the measuring techniques can now be considered when a new study is set up. PDAC is considered to be a "cold" tumor, which means that the low quantity of immune cells that is present will be harder to detect. Choosing targeted gene expression in combination with the improved cell definitions described in this thesis have allowed us to detect that B cells were found to be significantly increased in the TIME of long-term survivors [11]. This emphasizes that studies using the correct measuring technique in combination with accurate data analysis led to understanding the importance of B cells and B cell–based therapy for future personalized immunotherapy in PDAC patients.

FUTURE PERSPECTIVE

Decades of clinical and pre-clinical research have achieved more insight, new treatments, and a better understanding of PDAC. Nevertheless, the 5-year survival of the patients remains low. Further improvements in the research could be made by studying a more complete view of the patient, combining multiple measurements in a multi-omics study, including genomics and metabolomics. Multi-omics studies rely on large datasets and therefore multi-omics studies could benefit from the big steps that are being made in the artificial intelligence (AI) methods. Implementation of the findable, accessible, interoperable, and reusable (FAIR) principles could facilitate the collection of large datasets usable for multi-omics studies. To achieve further progress in the treatment of PDAC many collaborations are set up and further collaboration is needed between the patients, doctors, and researchers.

Multi-omics: A spectrum of continuous and highly dynamic interactions with different immune and stromal cells

Each type of omics study gives us different insights into the highly heterogeneous PDAC TIME as shown in this thesis. Combining these omics studies will create a more complete understanding of the structure of the PDAC TIME and how it functions [12–14]. The combination of these omics studies is called multi-omics. Besides combining omics studies from research papers, more and more multi-omics technologies are being set up to immediately study the PDAC TIME from multiple directions. However, there are still limitations and challenges to full-scale multi-omics studies [15]. Firstly, when combining two or more omics studies, the data needs to be harmonized. The data might be scaled or normalized differently. Furthermore, the dimensions of multi-omics datasets quickly become very large, leading to the need for dimensionality reduction. A lot of research has been committed to the PDAC TIME. A common conclusion from all these studies is that the PDAC TIME is very heterogeneous both between different tumors and within a single tumor. Moreover, studies have already been committed to incorporating multi-omics methods to study the PDAC TIME from a broader perspective. Multi-omics research is a field with a lot of potential that can create deeper insights into the mechanisms underlying the biological processes, molecular functions, interactions, and cell fates in PDAC TIMEs. Further extensive multi-omics studies need to be performed to gain a more complete understanding of the heterogeneity in the PDAC TIME, and the implications of this heterogeneity. A solution however lies in the use of bioinformatics methods to correctly analyze this highly dimensional data. More advanced bioinformatics tools, including the vastly developing AI tools, need to be developed to obtain useful information from large multi-omics studies.

Artificial intelligence: A step towards precision medicine and prevention

Advances in artificial intelligence give a completely new way to analyze multi-omics data. Quantification and characterization of the PDAC TIME data including genomics, proteomics, transcriptomics, epigenomics, and metabolomics are getting closer to reality, where many aspects play a role. Machine learning (ML) can be used to integrate the different omics and together with the clinical information it can be used to create predictive models that identify risk before the condition is apparent [16]. This is especially important with a disease like PDAC, which does not show clinical symptoms in the early stages. After the disease gets apparent, we nowadays use evidence-based medicine data to treat patients, whereas, in future treatment, regimens might be suggested based on algorithms that use the patient's phenotype including omics data to individualize treatment. Nevertheless, single omics will continue to be used as a diagnostic measure, as it is practical and cost-effective. Artificial intelligence is dependent on the input data provided by the user. A model can only be as accurate as the metadata that is provided with the omics data and the understanding of the single omics. In this thesis, an effort has been made to get a better understanding of these single omics by making different comparisons and describing the effect of the technique on the results found. The metadata that could further enhance the multi-omics studies includes clinical information, technical information, lab processes, and bioinformatics tool information. The FAIR principles that emphasize machine-actionability are therefore created.

FAIR(R) data & tools: A framework to assure accurate data and maintained tools

FAIR data principles describe that the data should be findable, accessible, interoperable, and reusable [17]. These principles are getting more important, as an enormous amount of data is being created. Findability should make it easier for researchers to find data that fits research questions. Obviously, after finding the right data, it should be accessible. A broadly applicable language should be used to accurately compare (meta) data, meaning it is interoperable. Furthermore, the data should be of a reusable quality and annotated with relevant attributes. A term that is missing from this equation is reproducibility. Academic articles describe results based on processed data that could be of use to other researchers. Nevertheless, many of those articles are not transparent enough about the process that they use in the wet and dry lab. Beyond the creation of FAIR data, a similar principle should be created for bioinformatics tools to ensure a longer lifespan. Far too often bioinformatics tools become obsolete, because of several reasons. The main reason is unfortunately that tools are being created in an academic setting, where the developer focuses on the publication only. Writing and maintaining high-quality code takes time but is often neglected after publication. This leads to bioinformatics tools that are not sustainable and difficult to use. Unfortunately, reproducing the results from this thesis will not be easy without the help of somebody involved with the original research. This is mainly due to the pressure on publishing articles and creating valuable results instead of creating sustainable tools.

Collaborations: The Achilles' heel of translational research

The most important people in translational human disease research are the patients. In case the patients give consent to participate in research, the clinicians must have a complete record of the patient and use this information to enroll the patient in a research study. This research study needs to have a clear hypothesis and aim. Based on the aim of the study the clinicians must discuss the technique they will use to measure the samples with the bioinformaticians and the time points of measurements and sample size with the statisticians. The lab technicians will perform the processing of the samples in the lab, after which the bioinformatician will process the data. Only when all these steps are discussed in a team and understood by the person performing the final analysis the research will result in useful information for the patients.

CONCLUSION

In this thesis, multiple omics techniques were used to determine the heterogeneous PDAC TIME in tissue samples. Furthermore, we used these techniques to analyze peripheral blood samples of PDAC patients to identify alterations in the immune cell profile during chemotherapy. Lastly, comparisons between omics techniques were made. Studies aiming to improve the analysis of single omics, but also the implementation of multi-omics analysis has to continue and could improve the understanding of the PDAC TIME even further.

REFERENCES

1. Collisson, E. A. *et al.* Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nature Medicine 2011 17:4* **17**, 500–503 (2011).

2. Moffitt, R. A. *et al.* Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nature Genetics 2015 47:10* **47**, 1168–1178 (2015).

3. Bailey, P. *et al.* Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature 2016 531:7592* **531**, 47–52 (2016).

4. Qian, Z. R. *et al.* Association of Alterations in Main Driver Genes With Outcomes of Patients With Resected Pancreatic Ductal Adenocarcinoma. *JAMA Oncol* **4**, e173420–e173420 (2018).

5. Huber, M. *et al.* The Immune Microenvironment in Pancreatic Cancer. *International Journal of Molecular Sciences 2020, Vol. 21, Page 7307* **21**, 7307 (2020).

6. de Geus, V. *et al.* Identifying Molecular Changes in Early Cervical Cancer Samples of Patients That Developed Metastasis. *Front Oncol* **11**, 5706 (2022).

7. Broekhuizen, M. *et al.* The Placental Innate Immune System Is Altered in Early-Onset Preeclampsia, but Not in Late-Onset Preeclampsia. *Front Immunol* **12**, 5386 (2021).

8. Aziz, H. M. *et al.* Spatial genomics reveals a high number and specific location of B cells in the pancreatic ductal adenocarcinoma microenvironment of long-term survivors. *Front Immunol* **13**, 7072 (2023).

9. van Eijck, C. W. F. *et al.* A multigene circulating biomarker to predict the lack of FOLFIRINOX response after a single cycle in patients with pancreatic ductal adenocarcinoma. *Eur J Cancer* **181**, 119–134 (2023).

10. Karamitopoulou, E. Tumour microenvironment of pancreatic cancer: immune landscape is dictated by molecular and histopathological features. *British Journal of Cancer 2019 121:1* **121**, 5–14 (2019).

11. Aziz, H. M. *et al.* Spatial genomics reveals a high number and specific location of B cells in the pancreatic ductal adenocarcinoma microenvironment of long-term survivors. *Front Immunol* **13**, 7072 (2023).

12. Fan, X. *et al.* Integrated single-cell multiomics analysis reveals novel candidate markers for prognosis in human pancreatic ductal adenocarcinoma. *Cell Discovery 2022 8:1* **8**, 1–16 (2022).

13. Montagne, J. M., Jaffee, E. M. & Fertig, E. J. Multiomics Empowers Predictive Pancreatic Cancer Immunotherapy. *J Immunol* **210**, 859–868 (2023).

14. Fraunhoffer, N. A. *et al.* Multi-omics data integration and modeling unravels new mechanisms for pancreatic cancer and improves prognostic prediction. *npj Precision Oncology 2022 6:1* **6**, 1–16 (2022).

15. Tarazona, S., Arzalluz-Luque, A. & Conesa, A. Undisclosed, unmet and neglected challenges in multi-omics studies. *Nature Computational Science 2021 1:6* **1**, 395–402 (2021).

16. Kaissis, G. *et al.* A machine learning algorithm predicts molecular subtypes in pancreatic ductal adenocarcinoma with differential response to gemcitabine-based versus FOLFIRINOX chemotherapy. *PLoS One* **14**, e0218642 (2019).

17. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data 2016 3:1* **3**, 1–9 (2016).

Chapter 8
Summary

SUMMARY

In this thesis, the methodology and application of determining the heterogeneous PDAC TIME with the use of multiple omics techniques are discussed. In **chapter two** we presented a method to identify different subtypes of immune cells specifically in PDAC tissue samples with the use of marker genes. Ninety immune cell-type specific marker genes were checked by calculating the pairwise similarity using the PDAC RNA-sequenced dataset available at The Cancer Genome Atlas. A set of 55 marker genes that identify 22 different immune cell types for PDAC were identified and validated with an independent mRNA expression dataset of 24 samples of PDAC patients who received various types of (neo)adjuvant treatments. In **chapter three**, blood samples collected at baseline and after the first FOLFIRINOX cycle of 68 patients from all disease stages of PDAC measured with the PanCancer Immune Profiling panel were used to develop an immune gene signature to predict response after a single cycle of FOLFIRINOX. **Chapter four** describes a study that investigates the safety of adding IMM-101 to SBRT and the immuno-modulatory effects of the combination treatment in the peripheral blood of locally advanced pancreatic cancer PDAC patients. Targeted gene-expression profiling and multicolor flow cytometry were performed for longitudinal immune monitoring of the peripheral blood. SBRT/IMM-101 treatment induced a transient decrease in different lymphocyte subsets and an increase in CD14+CD16−CD11b+HLA−DRlow myeloid-derived suppressor cells. Importantly, treatment significantly increased activated ICOS+, HLA-DR+, and Ki67+PD1+ T and NK cell frequencies. Combination therapy with SBRT and a heat-killed Mycobacterium vaccine was safe and had an immune-stimulatory effect. **Chapter five** showed the effect of one-cycle FOLFIRINOX in combination with lipegfilgrastim on the peripheral immune cell profile of pancreatic ductal adenocarcinoma (PDAC) patients using flow cytometry and was further explored by targeted immune-gene expression profiling. In **chapter six** we discussed the pros and cons of using RNA-sequencing for fresh frozen samples and targeted gene expression immune profiles for formalin-fixed, paraffin-embedded samples to characterize the tumor immune microenvironment.

NEDERLANDSE SAMENVATTING

In dit proefschrift worden de methodologie en de toepassing van het bepalen van de heterogene tumor micro-omgeving van ductaal adenocarcinoom in de avleesklier (PDAC) met behulp van meerdere omics technieken besproken. In **hoofdstuk twee** presenteerden wij een methode om verschillende subtypes van immuuncellen te identificeren met behulp van specifieke markergenen in PDAC-weefselmonsters. Negentig immuuncel specifieke markergenen werden gecontroleerd door de paarsgewijze overeenkomst te berekenen met behulp van de PDAC RNA-sequencedataset die beschikbaar is in The Cancer Genome Atlas (TCGA). Een set van 55 markergenen die 22 verschillende typen immuuncellen voor PDAC identificeren, werd geïdentificeerd en gevalideerd met een onafhankelijke mRNA-expressie dataset van 24 monsters van PDAC-patiënten die verschillende typen (neo)adjuvante behandelingen kregen. In **hoofdstuk drie** werden met het PanCancer Immune Profiling panel gemeten bloedmonsters, verzameld bij baseline en na de eerste FOLFIRINOX-cyclus van 68 patiënten uit alle ziektestadia van PDAC, gebruikt om een immuun-gen-handtekening te ontwikkelen die de respons na één cyclus FOLFIRINOX voorspelt. **Hoofdstuk vier** beschrijft een studie waarin de veiligheid van toevoeging van IMM-101 aan SBRT en de immuunmodulerende effecten van de combinatiebehandeling in het perifere bloed van lokaal gevorderde PDAC-patiënten met alvleesklierkanker worden onderzocht. Gerichte genexpressieprofilering en multicolor flowcytometrie werden uitgevoerd voor longitudinale immuunmonitoring van het perifere bloed. Behandeling met SBRT/IMM-101 veroorzaakte een afname van verschillende lymfocytensubsets en een toename van CD14+CD16-CD11b+HLA-DRlow myeloid-derived suppressor cells. Belangrijk is dat de behandeling de frequenties van geactiveerde ICOS+, HLA-DR+ en Ki67+PD1+ T- en NK-cellen aanzienlijk verhoogde. Combinatietherapie met SBRT en een hittedood Mycobacterium vaccin was veilig en had een immuunstimulerend effect. **Hoofdstuk vijf** toonde het effect van één-cyclus FOLFIRINOX in combinatie met lipegfilgrastim op het perifere immuuncelprofiel van PDAC-patiënten met behulp van flowcytometrie en werd verder onderzocht door gerichte immuun-gen expressie profilering. In **hoofdstuk zes** bespraken wij de voor- en nadelen van het gebruik van RNA-sequencing voor vers ingevroren monsters en gerichte genexpressie-immuunprofielen voor formeel gefixeerde, paraffine-ingebedde monsters om de tumorimmuunmicro-omgeving te karakteriseren.

Chapter 9

About the author

BIOGRAPHY

Willem de Koning was born in Woerden, the Netherlands on the 23rd of January 1995. After growing up in Montfoort, he graduated HAVO at Minkema College, Woerden, in 2012. Following his interest in biology and computers, he chose to study for a Bachelor of Science in Bioinformatics at Hogeschool Leiden. He obtained the degree in June 2016, after which he continued for a Master of Science degree in Bioinformatics at Wageningen University & Research till 2019. For his master's internship, he became a member of the group of Dr. Andrew Stubbs in 2018, where he decided to continue his postgraduate education under the supervision of Prof. Dr. Casper van Eijck, Dr. Andrew Stubbs, and Dr. Dana Mustafa. Results of the work performed in the 4 years Ph.D. are presented in this thesis.

LIST OF PUBLICATIONS

Tumor lineage-specific immune response in brain metastatic disease: opportunities for targeted immunotherapy regimen? Najjary, S., Kros, J. M., **de Koning, W.**, Vadgama, D., Lila, K., Wolf, J., & Mustafa, D. A. M. (2023). *Acta Neuropathologica Communications*, *11*(1), 64. https://doi.org/10.1186/S40478-023-01542-9

Spatial genomics reveals a high number and specific location of B cells in the pancreatic ductal adenocarcinoma microenvironment of long-term survivors. Aziz, H. M., Saida, L., **de Koning, W.**, Stubbs, A. P., Li, Y., Sideras, K., Palacios, E., Feliu, J., Mendiola, M., van Eijck, C. H. J., & Mustafa, D. A. M. (2023). *Frontiers in Immunology*, *13*. https://www.frontiersin.org/articles/10.3389/fimmu.2022.995715

Characterizing the tumor immune microenvironment of ependymomas using targeted gene expression profiles and RNA sequencing. **de Koning, W.**, Feenstra, F. F., Calkoen, F. G. J., van der Lugt, J., Kester, L. A., & Mustafa, D. A. M. (2023). *Cancer Immunology, Immunotherapy : CII*. https://doi.org/10.1007/S00262-023-03450-2

A multigene circulating biomarker to predict the lack of FOLFIRINOX response after a single cycle in patients with pancreatic ductal adenocarcinoma. van Eijck, C. W. F., **de Koning, W.**, van der Sijde, F., Moskie, M., Groot Koerkamp, B., Homs, M. Y. V, van der Burg, S. H., van Eijck, C. H. J., & Mustafa, D. A. M. (2023). *European Journal of Cancer*, *181*, 119–134. https://doi.org/https://doi.org/10.1016/j.ejca.2022.12.024

Increase of mast cells in COVID-19 pneumonia may contribute to pulmonary fibrosis and thrombosis. Wismans, L. V, Lopuhaä, B., **de Koning, W.**, Moeniralam, H., van Oosterhout, M., Ambarus, C., Hofman, F. N., Kuiken, T., Endeman, H., Mustafa, D. A. M., & von der Thüsen, J. H. (2022). *Histopathology*, *n/a*(n/a). https://doi.org/https://doi.org/10.1111/his.14838

Feasibility and Potential of Transcriptomic Analysis Using the NanoString nCounter Technology to Aid the Classification of Rejection in Kidney Transplant Biopsies. Varol, H., Ernst, A., Cristoferi, I., Arns, W., Baan, C. C., van Baardwijk, M., van den Bosch, T., Eckhoff, J., Harth, A., Hesselink, D. A., van Kemenade, F. J., **de Koning, W.**, Kurschat, C., Minnee, R. C., Mustafa, D. A. M., Reinders, M. E. J., Shahzad-Arshad, S. P., Snijders, M. L. H., Stippel, D., Stubbs, A. P., von der Thüsen, J., Wirths, K., Becker, J. U., & Clahsen-van Groningen, M. C. (2022). *Transplantation*. https://doi.org/10.1097/TP.0000000000004372

The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. Afgan, E., Nekrutenko, A., Grüning, B. A., Blankenberg, D., Goecks, J., Schatz, M. C., Ostrovsky, A. E., Mahmoud, A., Lonie, A. J., Syme, A., Fouilloux, A., Bretaudeau, A., Nekrutenko, A., Kumar, A., Eschenlauer, A. C., DeSanto, A. D., Guerler, A., Serrano-Solano, B., Batut, B., Grüning, B. A., Langhorst, B. W., Carr, B., Raubenolt, B. A.,

Hyde, C. J., Bromhead, C. J., Barnett, C. B., Royaux, C., Gallardo, C., Blankenberg, D., Fornika, D. J., Baker, D., Bouvier, D., Clements, D., de Lima Morais, D. A., Tabernero, D. L., Lariviere, D., Nasr, E., Afgan, E., Zambelli, F., Heyl, F., Psomopoulos, F., Coppens, F., Price, G. R., Cuccuru, G., Corguillé, G. Le, Von Kuster, G., Akbulut, G. G., Rasche, H., Hotz, H.-R., Eguinoa, I., Makunin, I., Ranawaka, I. J., Taylor, J. P., Joshi, J., Hillman-Jackson, J., Goecks, J., Chilton, J. M., Kamali, K., Suderman, K., Poterlowicz, K., Yvan, L. B., Lopez-Delisle, L., Sargent, L., Bassetti, M. E., Tangaro, M. A., van den Beek, M., Čech, M., Bernt, M., Fahrner, M., Tekman, M., Föll, M. C., Schatz, M. C., Crusoe, M. R., Roncoroni, M., Kucher, N., Coraor, N., Stoler, N., Rhodes, N., Soranzo, N., Pinter, N., Goonasekera, N. A., Moreno, P. A., Videm, P., Melanie, P., Mandreoli, P., Jagtap, P. D., Gu, Q., Weber, R. J. M., Lazarus, R., Vorderman, R. H. P., Hiltemann, S., Golitsynskiy, S., Garg, S., Bray, S. A., Gladman, S. L., Leo, S., Mehta, S. P., Griffin, T. J., Jalili, V., Yves, V., Wen, V., Nagampalli, V. K., Bacon, W. A., **de Koning, W.**, Maier, W., & Briggs, P. J. (2022). *Nucleic Acids Research*, *50*(W1), W345–W351. https://doi.org/10.1093/nar/gkac247

Identifying Molecular Changes in Early Cervical Cancer Samples of Patients That Developed Metastasis. de Geus, V., Ewing-Graham, P. C., **de Koning, W.**, de Koning, M. N. C., van den Bosch, T. P. P., Nigg, A. L., van Eijck, C. H. J., Jozwiak, M., van Beekhuizen, H. J., & Mustafa, D. A. M. (2022). *Frontiers in Oncology*, *11*. https://www.frontiersin.org/articles/10.3389/fonc.2021.715077

Autologous dendritic cells pulsed with allogeneic tumour cell lysate induce tumour-reactive T-cell responses in patients with pancreatic cancer: A phase I study. Lau, S. P., Klaase, L., Vink, M., Dumas, J., Bezemer, K., van Krimpen, A., van der Breggen, R., Wismans, L. V, Doukas, M., **de Koning, W.**, Stubbs, A. P., Mustafa, D. A. M., Vroman, H., Stadhouders, R., Nunes, J. B., Stingl, C., de Miranda, N. F. C. C., Luider, T. M., van der Burg, S. H., Aerts, J. G., & van Eijck, C. H. J. (2022). *European Journal of Cancer*, *169*, 20–31. https://doi.org/https://doi.org/10.1016/j.ejca.2022.03.015

Immunomodulatory Effects of Stereotactic Body Radiotherapy and Vaccination with Heat-Killed Mycobacterium Obuense (IMM-101) in Patients with Locally Advanced Pancreatic Cancer. van 't Land, F. R., Lau, S. P., **de Koning, W.**, Klaase, L., Vink, M., van Krimpen, A., Dumas, J., Vadgama, D., Nuyttens, J. J., Mustafa, D. A. M., Stadhouders, R., Willemsen, M., Stubbs, A. P., Aerts, J. G., & van Eijck, C. H. J. (2022). *Cancers*, *14*(21). https://doi.org/10.3390/cancers14215299

Identification, Validation, and Utilization of Immune Cells in Pancreatic Ductal Adenocarcinoma Based on Marker Genes. **de Koning, W.**, Latifi, D., Li, Y., van Eijck, C. H. J., Stubbs, A. P., & Mustafa, D. A. M. (2021). *Frontiers in Immunology*, *12*. https://www.frontiersin.org/articles/10.3389/fimmu.2021.649061

The Role of the Respiratory Microbiome and Viral Presence in Lower Respiratory Tract Infection Severity in the First Five Years of Life. Hoefnagels, I., van de Maat, J., van

Kampen, J. J. A., van Rossum, A., Obihara, C., Tramper-Stranders, G. A., Heikema, A. P., **de Koning, W.**, van Wermerskerken, A.-M., Horst-Kreft, D., Driessen, G. J. A., Punt, J., Smit, F. J., Stubbs, A., Noordzij, J. G., Hays, J. P., & Oostenbrink, R. (2021). *Microorganisms*, *9*(7). https://doi.org/10.3390/microorganisms9071446

Rintatolimod Induces Antiviral Activities in Human Pancreatic Cancer Cells: Opening for an Anti-COVID-19 Opportunity in Cancer Patients? Mustafa, D. A. M., Saida, L., Latifi, D., Wismans, L. V, de **Koning, W.**, Zeneyedpour, L., Luider, T. M., van den Hoogen, B., & van Eijck, C. H. J. (2021). *Cancers*, *13*(12). https://doi.org/10.3390/cancers13122896

NanoGalaxy: Nanopore long-read sequencing data analysis in Galaxy. **de Koning, W.**, Miladi, M., Hiltemann, S., Heikema, A., Hays, J. P., Flemming, S., van den Beek, M., Mustafa, D. A., Backofen, R., Grüning, B., & Stubbs, A. P. (2020). *GigaScience*, *9*(10), giaa105. https://doi.org/10.1093/gigascience/giaa105

RNA from stabilized whole blood enables more comprehensive immune gene expression profiling compared to RNA from peripheral blood mononuclear cells. van der Sijde, F., Li, Y., Schraauwen, R., **de Koning, W.**, van Eijck, C. H. J., & Mustafa, D. A. M. (2020). *PLOS ONE*, *15*(6), e0235413-. https://doi.org/10.1371/journal.pone.0235413

Lessons learnt from the introduction of nanopore sequencing. Heikema, A., **de Koning, W.**, Li, Y., Stubbs, A., & Hays, J. P. (2020). *Clinical Microbiology and Infection*, *26*(10), 1286–1288. https://doi.org/https://doi.org/10.1016/j.cmi.2020.05.035

Comparison of Illumina versus Nanopore 16S rRNA Gene Sequencing of the Human Nasal Microbiota. Heikema, A. P., Horst-Kreft, D., Boers, S. A., Jansen, R., Hiltemann, S. D., **de Koning, W.**, Kraaij, R., de Ridder, M. A. J., van Houten, C. B., Bont, L. J., Stubbs, A. P., & Hays, J. P. (2020). *Genes*, *11*(9). https://doi.org/10.3390/genes11091105

PHD PORTFOLIO

Name PhD candidate: Willem de Koning
Erasmus MC department: Pathology & Clinical Bioinformatics
PhD period: April 2019- April 2023
Promotors: Prof. Dr. C. H. J. van Eijck
     Dr. Andrew P. Stubbs
     Dr. Dana A. M. Mustafa

	Year	ECTS
Courses, seminars, and workshops		
Erasmus MC - ImmuneAID goes Big Data	2019	0,60
Erasmus MC - The NGS in DNA Diagnostics Course	2019	1,00
BioSB - Machine Learning for Bioinformatics & Systems Biology	2019	3,00
Erasmus MC - Survival Analysis	2019	0,60
Erasmus MC - Biomedical English Writing	2020	2,00
Erasmus MC - Basic and Translational Oncology	2020	1,80
Erasmus MC - Molecular Medicine	2021	0,70
Erasmus MC - Digital PhD day 2021	2021	0,20
Erasmus MC - Scientific Integrity	2021	0,30
Erasmus MC - Photoshop and Illustrator CC 2021 Workshop	2021	0,30
National and international conference attendance and presentations		
Galaxy Community Conference, Freiburg, Germany; Poster, Training and Small talk	2019	1,70
BioHackathon Europe, Online; Programming	2020	1,40
"How FAIR are you" Hackathon, Online; Attendance	2021	0,30
Pancreasdag, Online; Attendance	2021	0,20
X-omics festival "the future of multi-omics research is now!", Nijmegen, The Netherlands; Attendance	2022	0,30
The Dutch Bioinformatics & Systems Biology conference (BioSB), Lunteren, The Netherlands; Poster presentation	2022	0,75

	Year	ECTS
Teaching and supervision		
Erasmus MC - Galaxy Training Course	2019	0,75
Erasmus MC - RNA-seq for beginners	2019	0,60
Erasmus MC - Gene expression data analysis using R: How to make sense out of your RNA-Seq/microarray data	2019	2,00
Avans Hogeschool Breda - Galaxy Training Workshop	2020	0,80
Erasmus MC - Gene expression data analysis using R: How to make sense out of your RNA-Seq/microarray data	2020	2,00
Supervision of Dennis Dollée, B. Sc. internship and report (9 months)	2021	15,00
GTN Smörgåsbord: A Global Galaxy Course	2021	0,35
Erasmus MC - Gene expression data analysis using R: How to make sense out of your RNA-Seq/microarray data	2021	2,00
GTN Smörgåsbord 2: A Global Galaxy Course	2022	0,30
BioSB - RNA-seq course	2022	0,85
Supervision of Esther Hoogerwerf, M. Sc. internship and review	2022	1,00
MGC course - Next Generation Sequencing Data Analysis	2022	0,30
Total sum of ECTS		**41,10**

CURRICULUM VITAE

Education

April 2019 – April 2023	Ph.D. at the Department of Pathology & Clinical Bioinformatics at the Erasmus Medical Center Rotterdam, the Netherlands
September 2016 – January 2019	Master of Science degree in Bioinformatics at Wageningen University & Research, the Netherlands
September 2012 – June 2016	Bachelor of Science degree in Bioinformatics at Hogeschool Leiden, the Netherlands
June 2012	Graduation from high school at Minkema College, Woerden, the Netherlands

Working Experience

Ph.D. project (Four years)	*Multiple Omics Profiling of the PDAC Tumor Immune Microenvironment* Promoters: Prof. Dr. Casper van Eijck, Ph.D.; Dr. Andrew Stubbs, Ph.D.; Dr. Dana Mustafa, PhD
Master internship (Four months)	*NanoGalaxy - A Galaxy toolkit for the detection of Species, plasmid, and antibiotic resistance from Nanopore NGS* Supervisor: Dr. Andrew Stubbs, Ph.D. (Erasmus Medical Center Rotterdam) Results: Antimicrobial resistance (AMR) has become one of the biggest threats to global health, food security, and development. An AMR specialized nanopore sequence analysis toolkit, NanoGalaxy, is incorporated in Galaxy. It contains 13 tools that cover quality control, de novo assembly, species, and assembly detection, AMR detection, and reporting of the analysis. This results in a user-friendly environment, with "end-to-end" workflows.
Master thesis project (Six months)	*Improving microbiome analysis* Supervisor: Dr. Guido Hooiveld, PhD (Wageningen University & Research) Results: To determine the impact of different diets on the composition of the microbiome, 16s rRNA is analyzed. Until now 16s rRNA is only used to find the bacteria present in the fecal samples, but a more in-depth study of the functionality of those bacteria is desired. Therefore, a specialized pipeline for analysis of the microbiome data was created.
Bachelor internship (Nine months)	*Agent-based modeling of maturation of B cells in germinal centers* Supervisor: Dr. Gooitzen Zwanenburg, Ph.D. (University of Amsterdam) Germinal centers are located within the secondary lymphoid organs and are the places where high-affinity antibodies are generated. This occurs by proliferation, apoptosis by selection, and migration of the B cell. These processes are not well understood. In this study, a germinal center is modeled with an agent-based model (made in Compucell3D and NetLogo) to gain insight into the events of the germinal center. The germinal center is modeled with experimentally determined normalized antibody affinities for neutralizing antibodies against the stem epitope of H1N1 influenza hemagglutinin and morphology parameters from the literature.
Bachelor external project (Five months)	*Interactions between endothelial cells and pericytes during blood vessel formation* Supervisor: Dr. Lisanne Rens, PhD (Centrum Wiskunde & Informatica) Results: The interaction between pericytes and endothelial cells during angiogenesis is researched. Therefore, 2D agent-based models are created with the cellular Potts model. These 2D models are based on existing models, but a further elaboration on the influence of the different parameters is tested. These models are used to see the influence of both cell types on the forming of blood vessels and which cell initiate this forming.
Bachelor external project (Five months)	*Analysis of metagenomic contig-binners* Supervisor: Dr. Lex Overmars, Ph.D. (University of Amsterdam) Results: Binning is making clusters of contigs possibly originating from the same organism. We compared two supervised (MetaWatt and VizBin) and two unsupervised binning methods (CONCOCT and MaxBin). Synthetic data were used to test precision and recall. Biological data was used to determine the effectiveness and reliability of the binning methods by mapping it to a bacteria database. Based on the gathered information, advice is given for future use.
Part-time Job (Five years)	Service desk at Gamma Utrecht-de Meern, the Netherlands

Skills

Languages	Dutch (native), English (fluent)
Computer languages	Python, R, Java, Bash, MATLAB, NetLogo
Omics	Genomics, Transcriptomics