





Article

# Use of high-plex data reveals novel insights into the tumour microenvironment of clear cell renal cell carcinoma

Raffaele De Filippis <sup>1,†</sup> , Georg Wölflein <sup>2,†</sup> , In Hwa Um <sup>1</sup>, Peter D Caie <sup>3</sup>, Sarah Warren <sup>4</sup>, Andrew White <sup>4</sup>, Elizabeth Suen <sup>1</sup>, Emily To <sup>1</sup>, Ognjen Arandjelović <sup>2,\*</sup>  and David J Harrison <sup>1,\*</sup> 

<sup>1</sup> School of Medicine, University of St Andrews, UK  
<sup>2</sup> School of Computer Science, University of St Andrews, UK  
<sup>3</sup> Indica Labs  
<sup>4</sup> NanoString Technologies  
\* Correspondence: [david.harrison@st-andrews.ac.uk](mailto:david.harrison@st-andrews.ac.uk) / [ognjen.arandjelovic@gmail.com](mailto:ognjen.arandjelovic@gmail.com)  
† First co-authors

**Simple Summary:** Cancer is a complex ensemble of morphological and molecular features whose role is still unclear. Moreover, their role may change in different areas of the same tumour. Artificial intelligence (AI) allows pathologists to go beyond human perception and bias, and may help better understand how these features influence disease progression. Furthermore, by capturing variation intrinsic to the tumour, AI may improve the accuracy of current prognostic tools, such as Leibovich Score (LS), in predicting patient outcome and response to therapy. For these reasons, we studied in clear cell renal cell carcinoma (ccRCC) tissue, in which molecular features and their coexpression in the same cell were quantified and mapped using AI-based image analysis software. We demonstrated a novel approach for investigating ccRCC and revealed new potential biomarkers of prognosis which may also be able to direct patients towards the most appropriate personalised therapy.

**Abstract:** Although Immune Checkpoint Inhibitors (ICIs) have significantly improved Clear Cell Renal Cell Carcinoma (ccRCC) prognosis, about one third of patients experience recurrence. Current prognostic algorithms like the Leibovich Score (LS) rely on morphological features manually assessed by pathologists, and are therefore subject to bias. Moreover, these tools do not consider the heterogeneous molecular milieu present in the Tumour Microenvironment (TME), which may have prognostic value. We systematically developed a semi-automated method to investigate 62 markers and their combinations in 150 primary ccRCCs using multiplex Immunofluorescence (mIF), NanoString GeoMx<sup>®</sup> Digital Spatial Profiling (DSP) and Artificial Intelligence (AI)-assisted image analysis in order to find novel prognostic signatures and investigate their spatial relationship. We found that coexpression of Cancer Stem Cell (CSC) and Epithelial-to-Mesenchymal Transition (EMT) markers such as OCT4 and ZEB1 are indicative of poor outcome. OCT4 and the immune markers CD8, CD34 and CD163 significantly stratified patients at intermediate LS. Furthermore, augmenting the LS with OCT4 and CD34 improved patient stratification by outcome. Our results support the hypothesis that combining molecular markers has prognostic value and can be integrated with morphological features to improve risk stratification and personalised therapy. To conclude, GeoMx<sup>®</sup> DSP and AI image analysis are complementary tools providing high multiplexing capability required to investigate the TME of ccRCC, while reducing observer bias.

**Keywords:** multiplex; immunofluorescence; nanostring; image analysis; pathology; kidney; spatial analysis

**Citation:** De Filippis, R.; Wölflein, G.; Um, I.H.; Caie, P.; Warren, S.; White, A.; Suen, E.; To, E.; Arandjelović, O.; Harrison, D.J. Use of high-plex data reveals novel insights into the tumour microenvironment of clear cell renal cell carcinoma. *Cancers* **2022**, *1*, 0. <https://doi.org/>

Received:  
Accepted:  
Published:

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Copyright:** © 2022 by the authors. Submitted to *Cancers* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Clear cell renal cell carcinoma (ccRCC) is the most frequent kidney cancer and the deadliest genitourinary disease. Over 10,000 new cases are diagnosed annually, with over 76,000 new RCC cases diagnosed in 2021 [1]. To date, the only curative method is

nephrectomy, although 30% of patients experience recurrence [2]. Even though tremendous progress has been made in systemic therapy with the advent of immune checkpoint inhibitors (ICIs), around one third of patients present with metastasis, and their 5-year survival drops to 12% [3,4].

The current clinical gold standard prognostic tools are TNM staging and International Society of Urological Pathology (ISUP) nuclear grading, obtained through Haematoxylin & Eosin (H&E) staining of resected tissue sections. The former considers the tumour size (pT), lymph node involvement (pN), and the presence of distant metastasis (pM); the latter takes into account nucleolar prominence and morphology, as well as the presence of rhabdoid and sarcomatoid features [5,6].

Several prognostic algorithms have been built to assign prognostic scores to patients and predict the risk of advanced disease. For instance, the widely used Leibovich score (LS) predicts the risk of recurrence after surgery by considering pT, pN, tumour size, nuclear grade, and the presence of necrosis (table 1). However, LS only considers tumour morphological features, which are subjectively assessed by pathologists, while there is a lack of general consensus about molecular features. For these reasons, current prognostic tools, such as LS, are subject to inter-observer bias and are unable to accurately predict patient outcome and response to therapy [6,7].

Other factors may influence tumour progression: ccRCC is characterised by a highly heterogeneous Tumour Microenvironment (TME), in which hypoxia, blood vessels, stromal, and immune cells play a crucial role for cancer progression. For example, variations in immune escape mechanisms, such as PD-1/PD-L1 interaction, which are at the base of some ICIs, are responsible for drug resistance [8]. Therefore, assessing molecular marker expression such as PD-1 and PD-L1 could predict the response to therapy, and distinguish patients who would benefit from treatment from those who would potentially experience adverse effects. After exposure to chronic inflammation, cytotoxic T cells express high levels of co-inhibitory receptors, including PD-1, LAG-3, TIM-3, and CTLA-4. This results in T cell loss of function, and tumour adaptation to the immune response, favouring disease progression. The main exhausted T cell molecular signature is represented by coexpression of CD8 along with PD-1 and TIM-3 and was associated to poor prognosis in ccRCC [9].

Moreover, recent evidence supports the hypothesis that high PD-1/PD-L1 expression and increased regulatory T cell infiltration promote epithelial-mesenchymal transition (EMT), a process in which epithelial cells acquire a mesenchymal phenotype [10]. These signals lead to cytoskeleton reorganisation and cell motility, which are associated with invasiveness and metastatic potential in cancer [11]. The hallmark of EMT is the upregulation of N-cadherin followed by the downregulation of E-cadherin [12]. Although the origin of EMT is still unclear, evidence suggests that the transforming growth factor  $\beta$  (TGF- $\beta$ ) induces EMT by activating transcription factors including vimentin, snail, and zinc finger E-box binding homeobox 1 (ZEB1), which suppress epithelial cell proliferation [13–15]. In particular, vimentin is involved in the cytoskeleton reorganisation [16], whereas ZEB1 represses E-cadherin expression and activates N-cadherin and vimentin expression [17].

Interestingly, it has recently been shown that cancer stem cells (CSCs) require a hybrid epithelial/mesenchymal phenotype in order to initiate cancer progression [18]. These de-differentiated tumour cells hold self-renewal potential and cell motility which contribute to tumour aggressiveness and poor prognosis [19].

However, no standard CSC markers exist for ccRCC due to the high heterogeneity of its TME, although some have shown a prognostic potential. For instance, CD44 is involved in the regulation of stem cell features via the Wnt/b-catenin signalling pathway, suggesting that this marker could contribute to EMT [20]. Moreover, CD44 was associated with poor 5-year survival, high nuclear grade and recurrence in ccRCC [21]. To note, in vitro studies using RCC cell lines observed that CD44 was upregulated in the presence of tumour-associated macrophages (TAMs) [22].

OCT4 has also been shown to be responsible for stemness, EMT, tumorigenesis, and cell survival mechanisms [23]. In ccRCC, OCT4a, either alone or in combination with

NANOG, was associated with stemness and poor prognosis [24]. CSCs have also been shown to bind to CTLA-4 receptor causing CTLA-4 upregulation. These evidences suggest that CSCs may also be involved in immune evasion and T cell exhaustion [25,26], showing how all the mechanisms described above are strictly connected to each other symbiotically modulating the TME.

Understanding the high ccRCC intra-tumour heterogeneity (ITH) is crucial for patient stratification in order to achieve personalised therapy. Therefore, a better investigation of ccRCC TME is necessary, and novel approaches, such as multiplex immunofluorescence (mIF) and spatial profiling *in situ* can play a crucial role. Artificial intelligence (AI) and machine learning (ML) applied to image analysis have revolutionised digital pathology, allowing more precise quantification, generating more data, and overcoming inter-observer variability [27].

We systematically developed a semi-automated method to assess 62 and 158 features in Tissue Microarray (TMA) and whole slide tissue sections, respectively. Subsequently, we assessed their relationship and their correlation with outcome using ML-based statistical tools (Figure 1). This approach may facilitate the discovery of novel molecular signatures able to predict outcome and in order to choose the best therapeutic strategy and avoid ineffective and potentially harmful drug administration.

**Table 1.** Leibovich score

<b>Tumour status</b>	<b>Score</b>
pT1a	0
pT1b	2
pT2	3
pT3a	4
pT3b	4
pT3c	4
pT4	4
<b>Lymph node involvement</b>	<b>Score</b>
pNx	0
pN0	0
pN1	0
pN2	0
<b>Tumour size</b>	<b>Score</b>
<10cm	0
≥10cm	1
<b>Nuclear grade</b>	<b>Score</b>
1	0
2	0
3	1
4	3
<b>Necrosis</b>	<b>Score</b>
Absent	0
Present	1
<b>Risk group</b>	<b>Score</b>
Low	0-2
Intermediate	3-5
High	≥6

## 2. Materials and Methods

### 2.1. Study population

Tissues samples resected from 150 patients who were diagnosed with ccRCC and treated by radical nephrectomy between the years 1983 and 2018 were extracted from the pathology archive in Edinburgh and verified by an experienced pathologist. Specimens

were collected prior to therapy administration: none of these patients had therapy before surgery, and none received immune blockade therapy.

The analysed cohort consisted of 60 females and 86 males: Gender of 4 patients was not available. The average age was 64 years and mean follow-up was 60 months. 106 patients (70%) were reported to be deceased, of which 75 (50%) experienced cancer-related death (CRD). 5-year survival rate was 66%. 46 patients (30%) experienced metastasis at the time of diagnosis and for them, overall survival (OS) rate was 46% and 5-year survival rate was 58%.

A total of 124 patients had LS data, although no necrosis data were available. The majority of patients presented stage 3 and intermediate LS risk, while nuclear grades 2 and 3 were the most common. Finally, a total of 18 patients (14.5%) presented lymph node involvement (pN=1) and 35 patients (28.2%) presented distant metastasis. More detailed information is available in supplementary table A.

From each primary sample, 3 tissue punches were randomly collected and stored in 11 tissue microarray (TMA) blocks. Clinical information was retrieved from the patients' medical records, including age, sex, cancer-specific survival, time after initial surgery, TNM stage and nuclear grade. The tissue samples, clinical data, and ethical approval for this retrospective study were issued by the Research Tissue Bank and the SAHSC Bio Resource on behalf of NHS Lothian (15/ES/0094).

## 2.2. Antibody optimisation

Antibodies were first optimised separately on optimisation TMAs, consisting of several different normal and cancer tissues (RCC, breast, gastric, colon, prostate, tonsil, spleen etc) and then on RCC TMAs from the study cohort. All antibodies were optimised first in brightfield immunoperoxidase and then for immunofluorescence (IF). For tumour cell segmentation in immunofluorescence (IF), a combination of CA9 and pan-Cadherin antibodies was applied. The list of antibodies used, along with the dilution used is shown in Table B1 in the appendix.

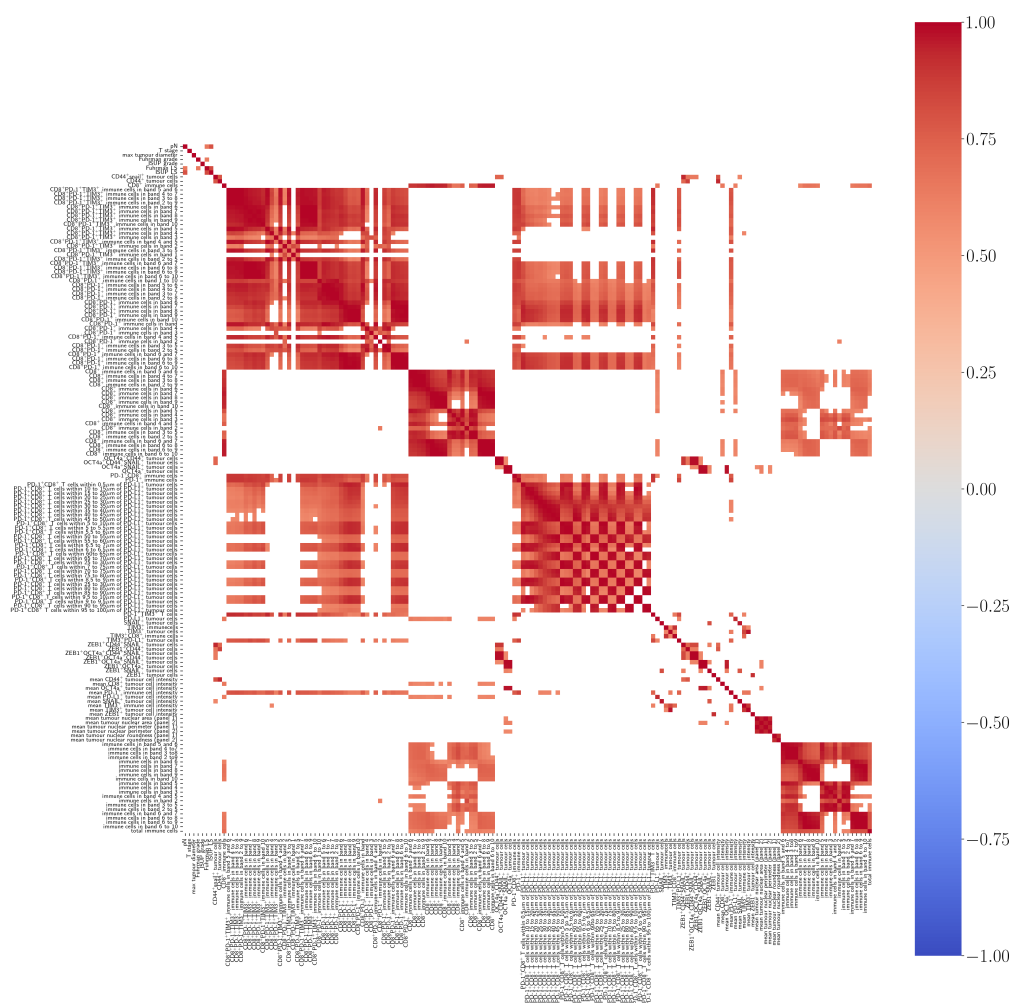
## 2.3. Multiplex immunofluorescence

Formalin-fixed, paraffin-embedded (FFPE) tissues were first de-waxed in xylene and then re-hydrated through descending concentrations of ethanol. In order to expose epitopes, heat-mediated antigen retrieval was performed in an electric pressure cooker using TRIS-EDTA buffer (pH 9). Subsequently, slides were treated with 3% hydrogen peroxide to block endogenous peroxidase activity, and with protein block (Agilent, X090930-2), in order to prevent non-specific staining. Slides were then incubated with primary antibodies for either one hour or 30 minutes at room temperature, or overnight at 4°C. After primary incubation, tissues were then incubated with secondary HRP-conjugated or biotin-conjugated antibodies for 30 minutes at room temperature. Slides were treated either with cyanine 3, cyanine 5, fluorescein, using Tyramide Signal Amplification (TSA) system, or Alexa Fluor 750 (ThermoFisher, S21384), diluted in amplification diluent (Perkin Elmer, FP1498) and antibody diluent (Agilent, S080983-2), respectively. Hoechst 33342 (ThermoFisher, H3570), diluted in PBS, was used for nuclear counterstaining. Finally, slides were dehydrated in ethanol, air-dried and mounted using Prolong anti-fade mounting medium (ThermoFisher, P36930). The full list of primary antibodies used is shown in table B1 of the appendix.

## 2.4. GeoMx<sup>®</sup> Digital spatial profiling

### 2.4.1. Sample preparation

Slides were prepared according to the digital spatial profiling (DSP) FFPE Protein Manual (MAN-10100-05). FFPE slides were baked for one hour at 60°C before being treated with CitroSolv (Fisher, 22-143-975) and descending concentrations of ethanol. Antigen retrieval in Citrate buffer (pH 6) was performed in a pressure cooker and slides were then washed in TBS-T, blocked with Buffer W (Iba 2-1003-100) and incubated overnight at 4°C with 62



**Figure 1.** Heatmap showing positive (red), negative (blue) or no (white) correlation among markers assessed in the Whole Slides (WSs) after unsupervised clustering. 156 features were assessed, including morphological features, molecular marker combination in tumour and immune cells, as well as spatial analysis data. Only correlations with  $p < 0.05$  after Chow-Denning correction are shown. No negative correlations were reported.

UV-cleavable Oligonucleotide-conjugated antibodies for DSP and Fluorescent-labelled antibodies for visualisation (Pan Cadherin – Alexa Fluor 488, Syto83 - Alexa Fluor 532, CD3e - Alexa Fluor 594, CD163 - Alexa Fluor 647). The Oligonucleotide-conjugated antibodies consisted of five panels (immune cell profiling, immune activation status, immune cell typing, tumour markers and drug targets), and are shown in table B2 of the appendix.

#### 2.4.2. DSP and nCounter<sup>®</sup> Readout

After incubation, the slides were loaded into the GeoMx<sup>®</sup> digital spatial profiler, where the tissues were scanned and digital images were generated for region of interest (ROI) selection. Each TMA core corresponded to a ROI. Subsequently, ROIs were illuminated with UV light and oligo tags were released into the aqueous layer just above the slide, which was aspirated through a micro capillary tube and stored in a single well of a 96-well plate. This process was repeated for each ROI. An average of 120 ROIs per slide were collected. nCounter Readout was processed according to the GeoMx DSP Readout Manual (MAN-10091-09), and as previously described [28,29].

#### 2.4.3. DSP Quality control and normalisation

Quality control was based on marker density and was automatically performed by the nSolver<sup>®</sup> software by NanoString Technology. Data normalisation was based on mouse IgG1 and IgG2 antibodies included in the antibody mix, which were considered as negative controls. Raw data were converted to a  $\log_2$  scale. For each ROI, corresponding to the TMA core, the mean of the IgG1 and IgG2 was used to calculate the signal-to-noise ratio (SNR) as

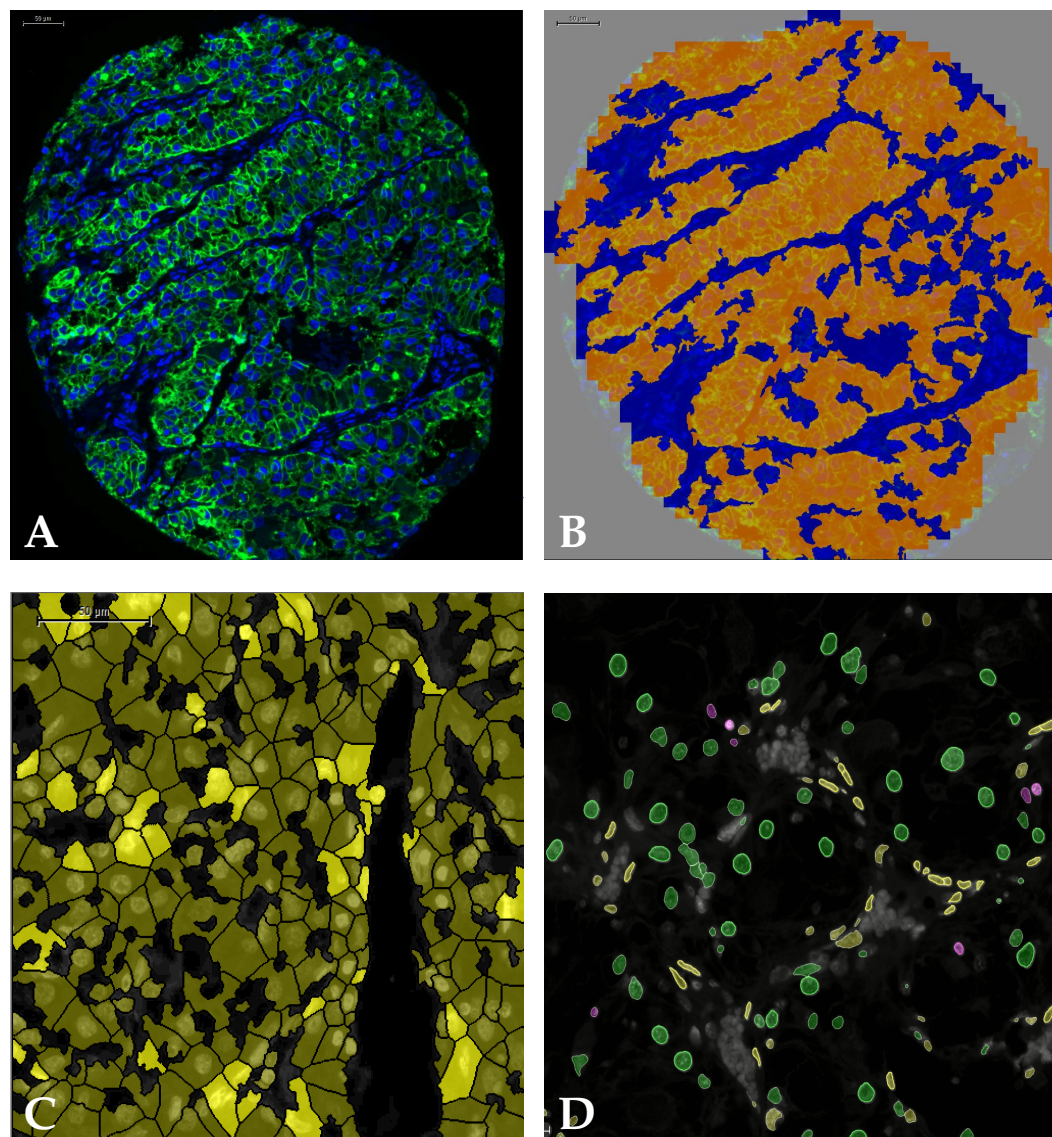
$$SNR = X - \frac{IgG_1 + IgG_2}{2} \quad (1)$$

where  $X$  is the count of a single marker in a single ROI. Since NanoString suggests that SNR should be  $> \log_2(3)$ , values greater than 1.5 were considered a positive signal. In order to keep the data intact, values below 1.5 were clamped to 1.5.

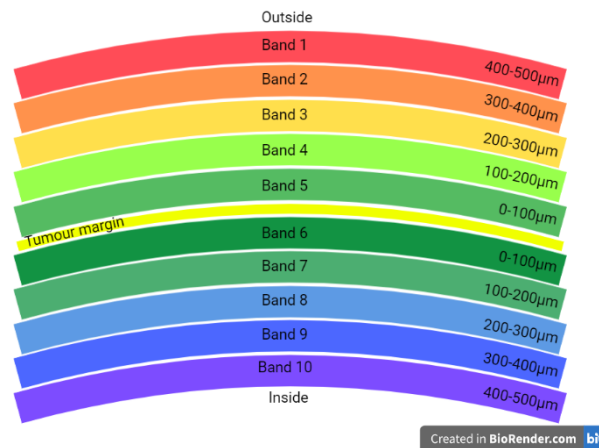
#### 2.5. Image analysis

TMA images were analysed using Definiens<sup>®</sup> Tissue Studio. Around 10% of images were used for training and several training rounds were performed to reach optimal results. For each TMA core, tumour-stroma segmentation was performed and only the tumour areas were considered since most cores consisted of over 90% tumour (Figures 2A and 2B). Nuclear segmentation was based on Hoechst channel. Subsequently, cells were reconstructed virtually. In the case of tumour cells, the expansion followed the tumour mask stain (Figure 2C). In the case of immune cells, the nuclear circumference was expanded by a specific radius. Whole slide images were analysed with Indica Labs Halo<sup>®</sup> AI software. Like with Tissue Studio, only tumour regions were considered after tumour-stroma segmentation, and nuclear segmentation was based on Hoechst intensity. After segmentation, Halo<sup>®</sup> AI was fed with hundreds of examples in order to distinguish three nuclear types: tumour, immune and stromal nuclei (Figure 2D). Cell simulation was only performed by expanding nuclear circumference since no tumour mask was used, and tumour cells were recognised from their nuclear morphology. Subsequently, cells were classified accordingly and single/multiple expressions were quantified.

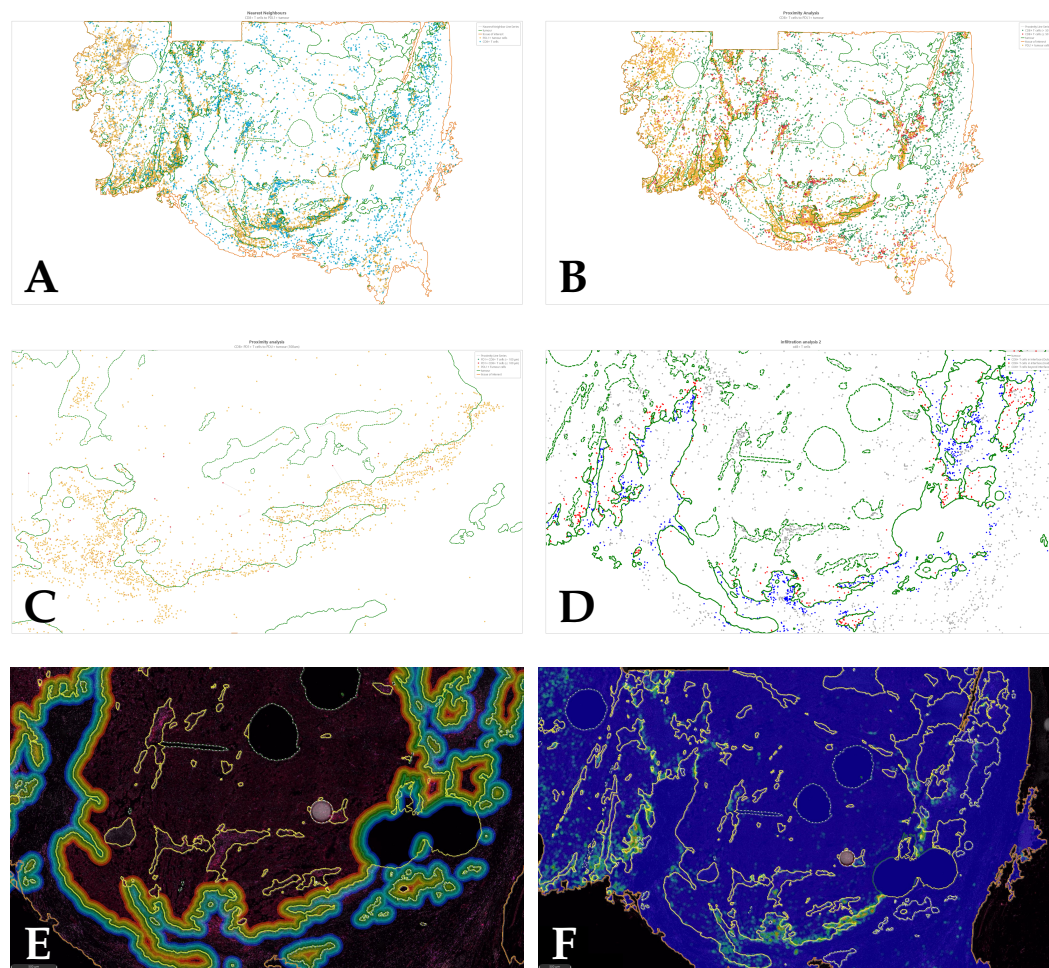
Spatial analysis was also performed for whole slides using dHalo<sup>®</sup> software. Four different analyses were performed: Infiltration analysis to determine the density of PD-1<sup>+</sup> T cells in 5 distance bands within and outside the tumour margin (Figure 3); nearest neighbour analysis to determine the number of PD-1<sup>+</sup> T cells nearby each PD-L1<sup>+</sup> tumour cell within a specific radius; proximity analysis to determine the distance of PD-1<sup>+</sup> T cells from PD-L1<sup>+</sup> tumour cells, and density heatmap analysis to visualise the density of PD-L1 tumour cell density across the tissue section (Figure 4).



**Figure 2.** (A) TMA core showing nuclear (blue) and tumour (green) stain; (B) same TMA core after tumour (orange) stroma (blue) segmentation in Definiens<sup>®</sup> Tissue Studio; (C) tumour cell simulation following tumour mask: positive tumour cells are shown in dark yellow, while tumour cells positive for PD-L1 are shown in bright yellow; (D) nuclear phenotyping only based on nuclear morphology using Halo<sup>®</sup> AI showing tumour nuclei (green), immune nuclei (purple) and fibroblast nuclei (yellow).



**Figure 3.** Schematic representation of distance bands assessed in infiltration analysis. The tumour margin is represented by the yellow line between band 5 and 6. Each band was 100 $\mu\text{m}$  in thickness. Marker densities were quantified in each band separately as well as in multiple adjacent bands (see figure 1).



**Figure 4.** (A) Nearest neighbour analysis quantified the number of PD-1<sup>+</sup> T cells nearby each PD-L1<sup>+</sup> tumour cell within a specific radius; (B) Proximity analysis quantified the distance of PD-1<sup>+</sup> T cells from PD-L1<sup>+</sup> tumour cells; (C) Proximity analysis, closer view; (D) Infiltration analysis quantified the density of PD-1<sup>+</sup> T cells in 5 distance bands within and outside the tumour margin; (E) Infiltration analysis, distance bands visualised on the real image; (F) Density heatmap showing PD-L1<sup>+</sup> tumour cells distributed mainly on the tumour margin.



## 2.6. Statistical analysis

Statistical analysis of 67 and 152 molecular features was performed in TMAs and WSs, respectively, using R Studio (version 1.1.423) and Python (version 3.8). Seven morphological features, including maximum diameter, pN, stage, Fuhrman grade, ISUP grade, classical LS, and ISUP LS were also investigated in both experiments. Unsupervised clustering according to Pearson's correlation was used to investigate relationships between markers. Cox's proportional hazard method was used to predict 5-year and OS, where CRD was considered as the event. OS and CRD data were collected from the Research Tissue Bank of the Royal Infirmary of Edinburgh. We conducted two sets of experiments: one where we took as input the continuous data, and another where each variable was binarised into *high* and *low* after determining the optimal cutpoint in relation with outcome, using the maximally selected rank statistics from the R package *maxstat*. Statistical significance was evaluated using the likelihood ratio test and log-rank test to obtain *p*-values for the continuous and categorical models, respectively. We exhaustively fit Cox models for every combination of one, two, and three markers, and applied Chow-Denning correction [30] to the *p*-values in order to account for multiple testing. Due to the vast number of models that needed to be fit (in the order of millions), we ran the computation in parallel on 20 Linux machines for approximately 12 hours using the Jug library [31] to parallelise the code.

## 3. Results (Tissue Microarray)

For each molecular marker, density was calculated by dividing the number of positive cells by the TMA core area. Subsequently, average density of replicates was calculated.

### 3.1. Pearson's correlation

Unsupervised clustering using Pearson's correlation found positive correlation among morphological, DSP-detected, and IF-detected markers, involved in immune response, immune evasion, T cell exhaustion, as well as CSCs and EMT (Figure 5).

Among DSP markers, positive correlation was found between CD163 and CD4, CD44, and CD68. CD3 was positively correlated with CD4, CD44, CD68, and CD8. Moreover, CD44 was positively correlated with CD68. CD8 was positively correlated with CD4 and CD68 (Figure 5).

Among IF markers,  $\beta$ -catenin was positively correlated with DSP Histone H3. PD1<sup>+</sup>CD8<sup>+</sup> cells were positively correlated with CD8<sup>+</sup>LAG-3<sup>+</sup> cells and with CD8<sup>+</sup>LAG-3<sup>+</sup>TIM-3<sup>+</sup> cells. Positive correlation was also found between ZEB1 and TIM-3, as well as CD8 and the following markers: PD-1, PD-L1<sup>+</sup> tumour cells, CD163, and LAG-3. Interestingly, positive correlation was also found between CD8<sup>+</sup>LAG-3<sup>+</sup> cells and CD8<sup>+</sup>PD-1<sup>+</sup> cells (Figure 5), indicating that PD-1 may contribute to an exhausted T cell phenotype.

### 3.2. Survival analysis

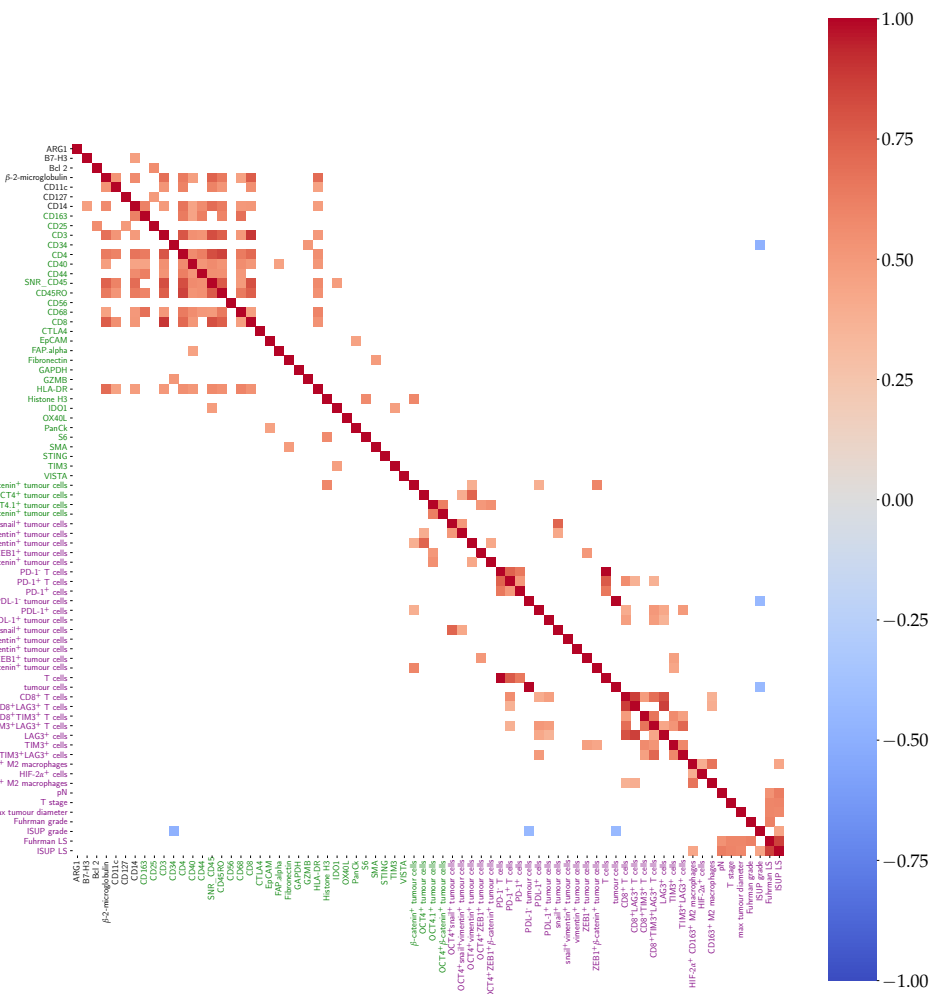
#### 3.2.1. Univariate Cox regression

Overall, morphological features showed the highest significance. In particular, the LS with the ISUP nuclear grade instead of the Fuhrman grade was the most significant feature, and was associated with poor OS, see Figure 6A) and 5-year survival, see Figure 6B). Similar results were obtained with pN, Fuhrman LS, and ISUP grade. Among molecular markers, OCT4 alone and in combination with ZEB1 was associated with poorer outcome. Moreover, OCT4<sup>+</sup>ZEB1<sup>+</sup> $\beta$ -catenin was associated with 5-year survival.

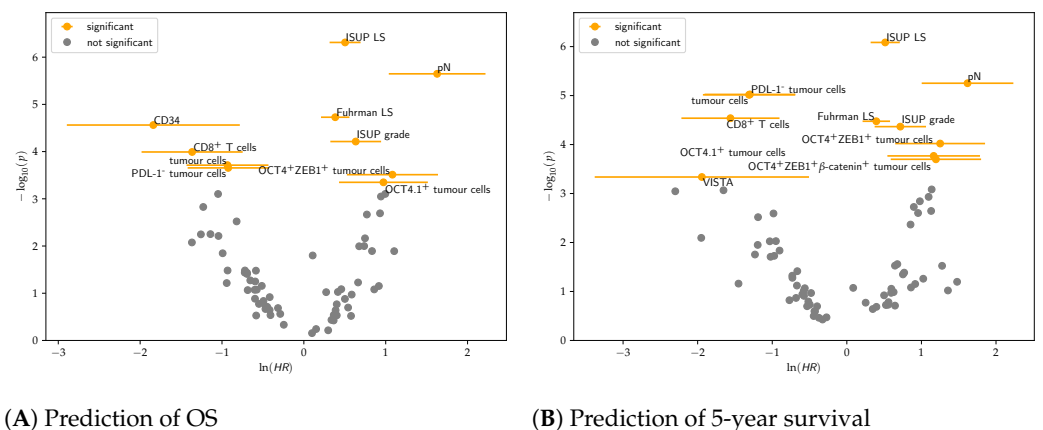
CD8, total tumour cell density, and PD-L1<sup>+</sup> tumour cell density were associated with longer OS and 5 year survival. CD34 density was associated with longer OS, while DSP-detected VISTA was associated with longer 5-year survival.

#### 3.2.2. Multivariate Cox regression

We exhaustively fit multivariate Cox regression models on every combination of two and three features, in order to find combinations of markers that exhibit statistically significant predictions of patient outcomes. For the TMA cohort with 74 features, we



**Figure 5.** Heatmap showing positive (red), negative (blue) or no (white) correlation among molecular markers assessed with DSP (black), IF (green), and morphological (purple) features, aggregating the TMA core replicates by mean per patient. Only correlations with  $p < 0.05$  after Chow-Denning correction are shown.



(A) Prediction of OS

(B) Prediction of 5-year survival

**Figure 6.** Volcano plots showing significant categorical univariate Cox regression features after Chow-Denning correction and their correlation with OS (a) and 5-year survival (b). Features with  $\log(HR) > 0$  are positively correlated with the event, whereas those with  $\log(HR) < 0$  are negatively correlated.

consequently fit  $\binom{74}{2} + \binom{74}{3} + \binom{74}{4} = 1,218,151$  Cox models. After Chow-Denning correction, we found numerous groups of markers that were predictive for OS or 5-year survival. As expected, most of these groups included features which individually were already statistically significant in the univariate model. However, some groups comprised entirely of markers that were not statistically significant when considering the markers individually in univariate Cox regression, but in combination achieved statistical significance.

In particular, coexpression of CD8 and LAG-3, representing an exhausted T cell phenotype, along with HLA DR, an isotype of the human leukocyte antigens, was identified as a prognostic combination for 5-year survival, suggesting that MHC II plays a crucial role in T cell exhaustion. OCT4 alone or coexpressed with vimentin, reached statistical significance for OS and 5-year survival, respectively, after pairing with Arginase 1 (Arg1), suggesting that CSC and EMT may be linked favouring cancer proliferation. OCT4 combination with HLA DR was also associated with shorter 5-year survival. OCT4<sup>+</sup>vimentin<sup>+</sup> cancer cells also stratified patients by OS when combined with PD-1 expression and snail. Unexpectedly, PD-1<sup>+</sup> cell expression was associated with longer OS when combined with snail<sup>+</sup>vimentin<sup>+</sup> tumour cells, therefore further investigation is needed. When OCT4<sup>+</sup> $\beta$ -catenin<sup>+</sup> cancer cell expression was combined with TIM-3, detected through DSP, it showed shorter OS and 5-year survival. Instead, OCT4<sup>+</sup>ZEB1<sup>+</sup> $\beta$ -catenin<sup>+</sup> tumour cell expression showed shorter OS when combined with snail<sup>+</sup>vimentin<sup>+</sup> tumour cell expression. This set of marker groups may provide insight into complex interactions between features that have an impact on predicting patient outcome. Table 2 lists all pairs of markers (i.e. combinations of two) in this category<sup>1</sup>.

As is evident in Figure 6A, Cox regression showed that OCT4<sup>+</sup>ZEB1<sup>+</sup> tumour cells were associated with CRD. However, tumour cells only expressing OCT4 were not found to predict the outcome with statistical significance and neither did the ones only expressing ZEB1. Furthermore, in bivariate Cox regression, OCT4<sup>+</sup> and ZEB1<sup>+</sup> *in combination* was not a statistically significant predictor of outcome. This means that merely knowing the densities of OCT4<sup>+</sup> cells and ZEB1<sup>+</sup> cells is insufficient; instead, knowledge of the density of cells coexpressing *both* markers is required to arrive at a prediction carrying statistical significance. This highlights the need for accurate and precise spatial analysis that can identify co-registered markers on the same cell. This result also suggests that coexpression of OCT4 and ZEB1 in tumour cells may not only be indicative of a dedifferentiation, but also that CSC and EMT may be dependent mechanisms.

### 3.3. Integrated LS

While the LS is able to stratify patients into low, intermediate and high risk groups, patients at intermediate risk still exhibit varied outcomes. In this section, we evaluate how a combination of the LS with molecular markers can result in better stratification.

As baseline, we assessed the stratification ability of the traditional LS on the TMA cohort. After scoring the patients, Kaplan-Meier analysis was able to accurately separate high and intermediate risk groups ( $p < 0.0001$ ) (Figure 7). Only six patients showed low risk score (0-3), and were therefore excluded from the analysis.

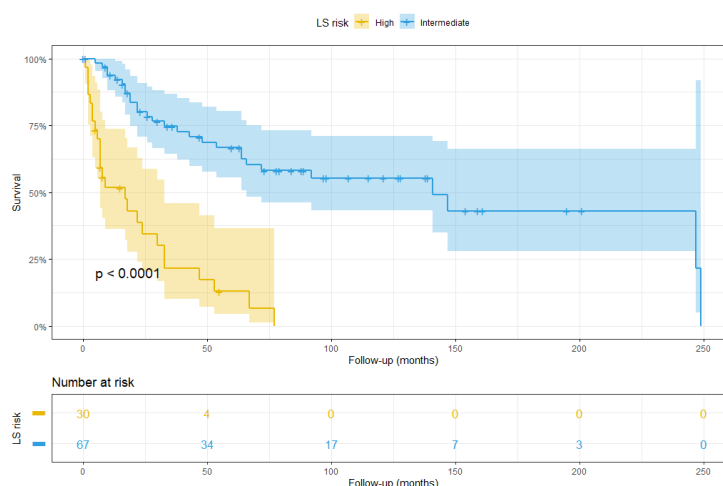
Since traditional LS only relies on morphological data, we tested the stratification ability of the molecular markers that resulted significant in survival analysis. In particular, we found that two specific molecular markers were able to further stratify the patients at intermediate risk according to ISUP LS: IF-detected OCT4<sup>+</sup> tumour cell density was used to stratify patients into intermediate-high and intermediate-low groups (Figure 8A). Eight patients were excluded from this analysis since their OCT4 data were not available. Similarly, CD34, detected through DSP, showed the same stratification ability (Figure 8B).

To further demonstrate the prognostic value of these markers, they were integrated in the LS algorithm. The patients were categorised in high and low risk according to

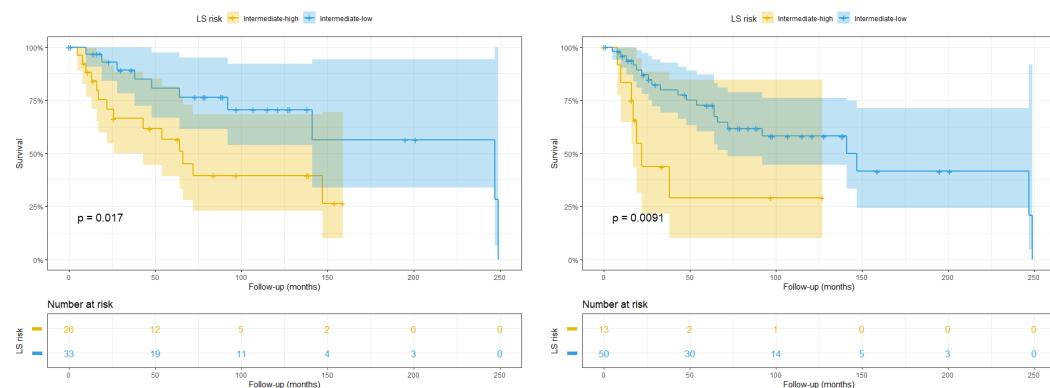
<sup>1</sup> We mention some statistically significant combinations of more than two markers, but do not provide the full list in this paper for the sake of brevity.

**Table 2.** Table of pairs of markers that were insignificant in univariate Cox regression individually ( $p > 0.05$ ), but when used together as input to a multivariate Cox model achieved  $p < 0.05$  after Chow-Denning correction. Markers annotated with  $\uparrow$  exhibit positive correlation with the event, whereas  $\downarrow$  denotes negative correlation.

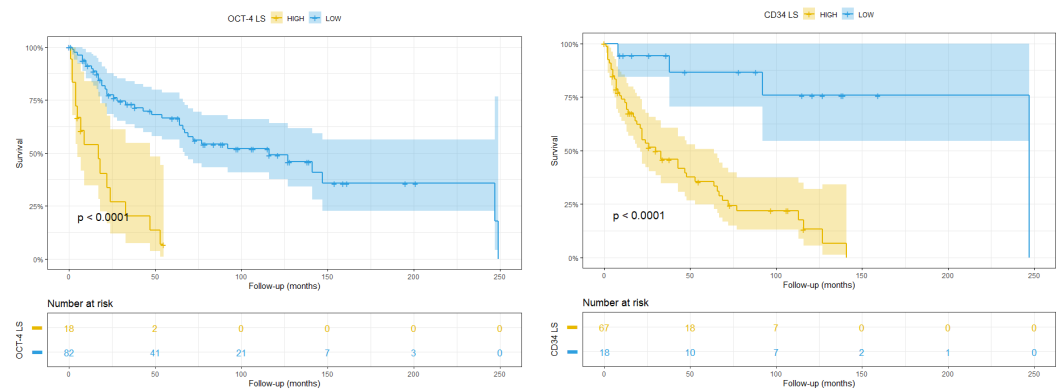
survival	markers	$p$ -value	
5yr	$\downarrow$ HLA-DR	$\downarrow$ CD8 <sup>+</sup> LAG3 <sup>+</sup> T cells	0.011
5yr	$\downarrow$ ARG1	$\uparrow$ OCT4 <sup>+</sup> tumour cells	0.016
5yr	$\downarrow$ TIM3	$\uparrow$ OCT4 <sup>+</sup> $\beta$ -catenin <sup>+</sup> tumour cells	0.020
5yr	$\downarrow$ ARG1	$\uparrow$ OCT4 <sup>+</sup> vimentin <sup>+</sup> tumour cells	0.027
5yr	$\uparrow$ OCT4 <sup>+</sup> vimentin <sup>+</sup> tumour cells	$\downarrow$ snail <sup>+</sup> vimentin <sup>+</sup> tumour cells	0.030
5yr	$\downarrow$ HLA-DR	$\uparrow$ OCT4 <sup>+</sup> vimentin <sup>+</sup> tumour cells	0.038
5yr	$\downarrow$ CD8 <sup>+</sup> TIM3 <sup>+</sup> T cells	$\uparrow$ TIM3 <sup>+</sup> cells	0.044
overall	$\uparrow$ OCT4 <sup>+</sup> vimentin <sup>+</sup> tumour cells	$\downarrow$ PD-1 <sup>+</sup> cells	0.011
overall	$\uparrow$ OCT4 <sup>+</sup> ZEB1 <sup>+</sup> $\beta$ -catenin <sup>+</sup> tumour cells	$\downarrow$ snail <sup>+</sup> vimentin <sup>+</sup> tumour cells	0.019
overall	$\downarrow$ TIM3	$\uparrow$ OCT4 <sup>+</sup> $\beta$ -catenin <sup>+</sup> tumour cells	0.022
overall	$\uparrow$ OCT4 <sup>+</sup> vimentin <sup>+</sup> tumour cells	$\downarrow$ snail <sup>+</sup> vimentin <sup>+</sup> tumour cells	0.022
overall	$\uparrow$ OCT4 <sup>+</sup> vimentin <sup>+</sup> tumour cells	$\downarrow$ snail <sup>+</sup> tumour cells	0.035
overall	$\downarrow$ PD-1 <sup>+</sup> cells	$\downarrow$ snail <sup>+</sup> vimentin <sup>+</sup> tumour cells	0.037



**Figure 7.** Kaplan-Meier plot showing patient at high (yellow) and intermediate (blue) risk according to ISUP LS.



**Figure 8.** Stratification of patients at intermediate ISUP LS, based on **(A)** OCT4 detected through multiplex IF and **(B)** CD34 detected through DSP.



(A) OCT4<sup>high</sup>

(B) CD34<sup>high</sup>

**Figure 9.** Kaplan-Meier plots showing stratification of patients with integrated LS using (A) IF OCT4<sup>high</sup> and (B) DSP CD34<sup>high</sup>.

marker expression and LS. Interestingly, OCT4<sup>high</sup> integrated in the ISUP LS was able to significantly stratify patients (Figure 9A). The same result was obtained when CD34<sup>high</sup> was integrated into the algorithm (Figure 9B).

#### 4. Results (Whole Slides)

##### 4.1. Pearson's correlation

Positive correlation was found between CD8<sup>+</sup> immune cells and PD-L1<sup>+</sup> tumour cells, between OCT4a<sup>+</sup> tumour cells and ZEB1<sup>+</sup>SNAIL<sup>+</sup> tumour cells. Interestingly, positive correlation was also found between OCT4a<sup>+</sup> tumour cells and mean tumour cell nuclear area, although no correlation was found with tumour grade. Positive correlation was found between PD1<sup>+</sup>CD8<sup>+</sup> T cells and PD-L1<sup>+</sup> tumour cells. TIM-3<sup>+</sup>PD-L1<sup>+</sup> tumour cells were positively correlated with PD-1<sup>+</sup>CD8<sup>+</sup> T cells.

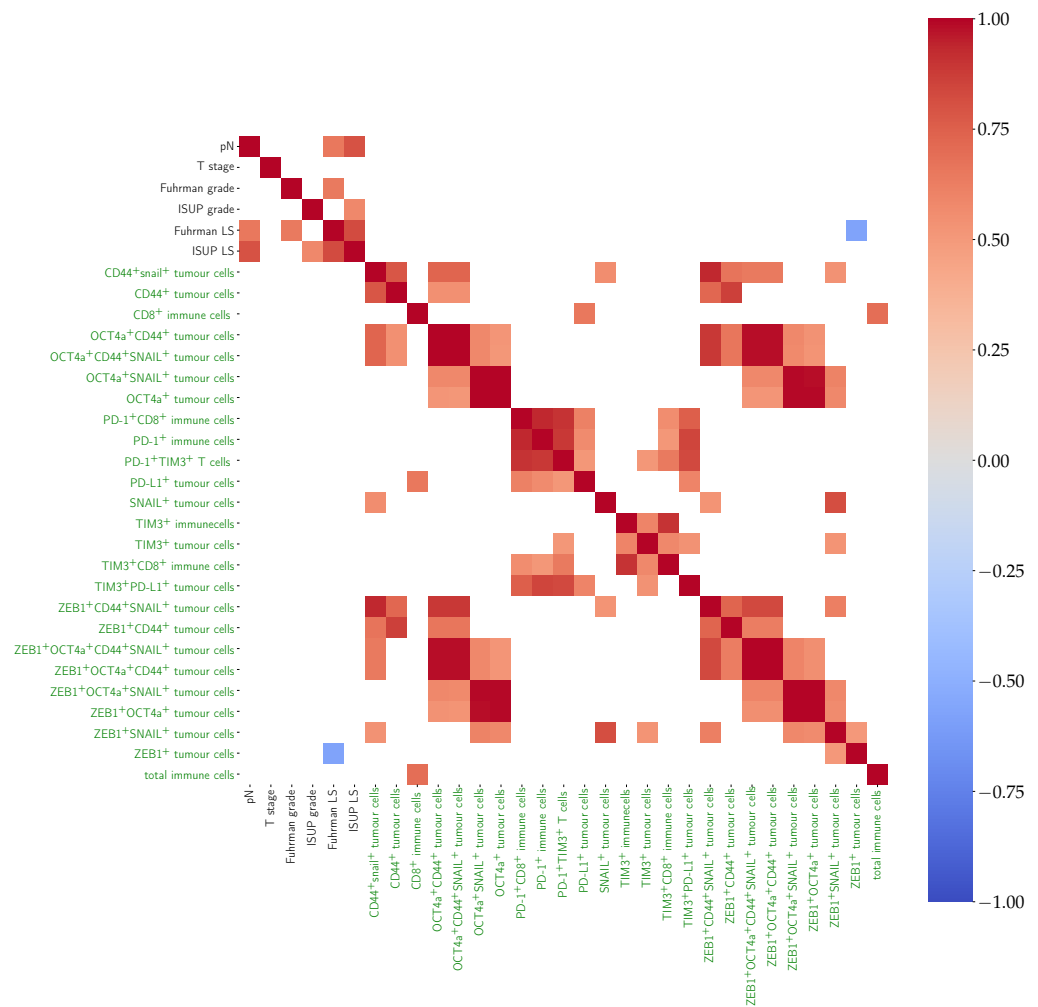
##### 4.2. Cox regression

Categorical univariate Cox regression showed that ZEB1<sup>+</sup>snail<sup>+</sup> tumour cells and snail<sup>+</sup>CD44<sup>+</sup> tumour cells were associated with longer OS and 5-year survival. Unexpectedly, snail<sup>+</sup>ZEB1<sup>+</sup>CD44<sup>+</sup> tumour cell density was associated with longer 5-year survival (Figure 11). Therefore, further analysis is needed in order to better reveal the role of these markers.

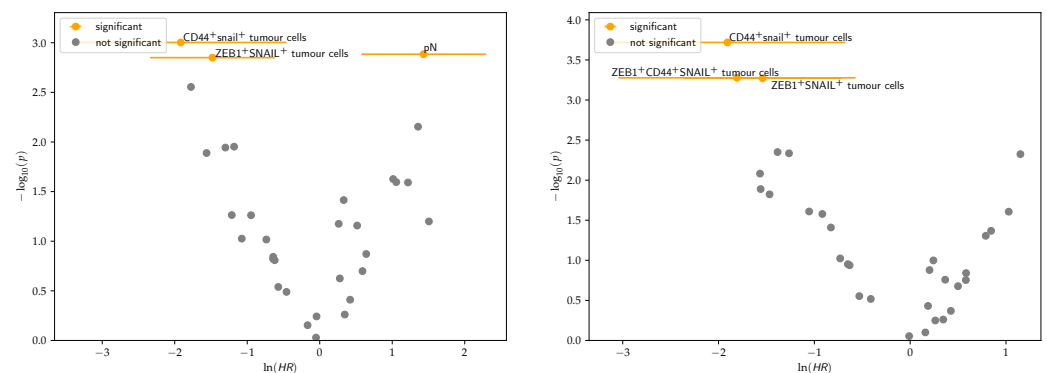
Multivariate analysis showed that combination of PD1<sup>+</sup> T cells and ZEB1<sup>+</sup>snail<sup>+</sup> predicted 5-year survival, whereas these two features did not reach statistical significance alone.

Survival analysis was performed to predict OS and 5-year survival using the classic ISUP and integrated LSs. The classic LS's stratification ability in the WS cohort is shown in Figure 12.

However, stratification of patients at intermediate risk did not show significant results due to the small sample size.



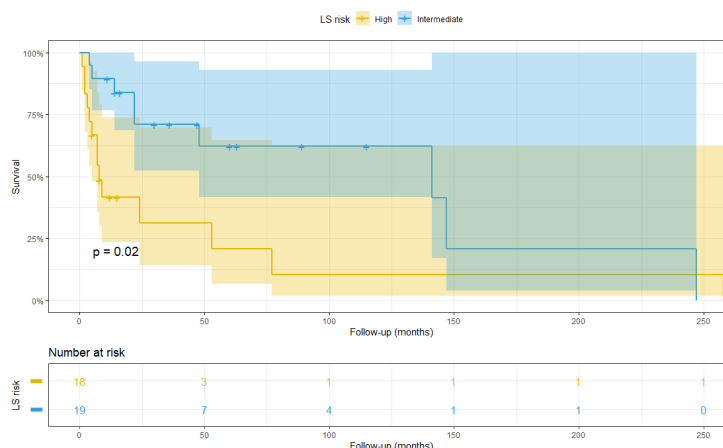
**Figure 10.** Heatmap showing positive (red), negative (blue) or no (white) correlation among morphological (black) and IF (green) markers in the WS cohort after unsupervised clustering. Features including OCT4a and ZEB1 showed the most correlation, while the only negative correlation was found between the Fuhrman LS and ZEB1<sup>+</sup> tumour cells.



(A) Prediction of OS

(B) Prediction of 5-year survival

**Figure 11.** Volcano plots showing significant categorical univariate Cox regression features after Chow-Denning correction and their correlation with (A) OS and (B) 5-year survival in the WS cohort. Features with  $\log(HR) > 0$  are positively correlated with the event, whereas those with  $\log(HR) < 0$  are negatively correlated.



**Figure 12.** Patients belonging to the whole slide cohort stratified at high (yellow) and low (blue) risk according to the ISUP LS.

## 5. Discussion

Although in recent years systems pathology has made great steps, it still lacks standardised methods with high-plex capability, which are necessary to investigate the complex tumour microenvironment. Nanostring GeoMx<sup>®</sup> DSP overcomes this problem by allowing a high-plex, spatially-resolved assessment of FFPE sections while retaining tissue integrity [32]. In our work, this high-throughput method allowed the identification of positive correlations between a large number of markers. In particular, the CSC marker CD44 was positively associated with immune markers, such as CD163, indicating a role of M2 macrophages in tumour dedifferentiation, or CD3 and CD4, which are indicative of an immune response. These findings support the hypothesis of a crosstalk between CSCs and TAMs, where TAMs exert a trophic function on stem cells while, in turn, CSCs drive M2 macrophage polarisation [33]. This mechanism may represent an innovative therapeutic target.

Combining GeoMx<sup>®</sup> DSP with mIF is a novel strategy to select specific regions of interest (ROIs) and obtain more precise information about the role that each marker plays in different areas of the tumour. In this study, GeoMx<sup>®</sup> DSP was performed using TMAs to allow assessment of a greater number of patients and acquire more robust data. In this setting, each ROI corresponded to a single TMA core. This allowed assessment of co-localisation and hence the relationship among markers. In multivariate analysis, OCT4 and ZEB1 co-expression was only significant for prognosis when the two molecules co-expressed in the same cell. When detected separately, the markers (that is, without co-localisation information) and used as two separate input features in a multivariate model did not reach statistical significance.

ZEB1 is a transcription factor that promotes epithelial-to-mesenchymal transition (EMT) by downregulating the epithelial marker E-cadherin, therefore facilitating cell motility and cancer dissemination [34]. In whole slides, ZEB1 correlated with longer overall survival (OS) when co-expressed in the same cell with snail, and longer 5-year survival when co-expressed with CD44 and snail, suggesting that ZEB1 is not the only determinant of outcome and that other factors may prevent EMT in these patients. By contrast, in TMAs, levels of ZEB1 were associated with shorter OS and 5-year survival when co-expressed with OCT4, consistent with the possibility that OCT4 and ZEB1 both play a role in the acquisition of either a CSC or EMT phenotype [35]. Further investigation is required.

The difference between whole slide and TMA may be because whole slide incorporates tumour core and invasive margin, whereas TMA cores were usually sampled from tumour core. ZEB1, snail, CD44, and OCT4 have been reported in the literature to be involved in tumour de-differentiation [36], therefore it is plausible that this mechanism occurs in the tumour core, which is more often hypoxic and hosts stem cell niches [37].

We have shown a pipeline of work that can be used to interrogate samples of cancer from patients using several analytical approaches. AI-based image analysis of both mIF and DSP produces an overwhelming amount of data which was possible to analyse only with machine learning-based statistical methods and required computational capability alongside substantial engineering efforts in terms of parallel computing to exhaustively assess all feature combinations. When analysing large volumes of data, the risk of overfitting is proportional to the number of features analysed, and further increases when the number of patients is not large enough [38]. We applied Chow-Denning correction [30] to account for multiple testing, which is slightly less conservative than Bonferroni correction [39]. However, further validation with an external cohort is required to confirm our findings.

It is noteworthy that samples were taken before therapy administration, and no therapy information was available in our metadata, so different therapeutic strategies may have been applied to the patients, therefore differently influencing outcome. This is important to consider also due to the large time span between the first and last sample collection, which may also be influenced by the evolution of operative techniques in sample collection. Another aspect to take into account is the deterioration of the FFPE tissue quality, since it has been reported that antigen quality may decay with time especially those expressed in the nucleus or on the cell surface [40]. However, we have conducted many studies using quantitative mIF and this has not appeared to be a problem. Moreover, mRNA extraction from similar old FFPE samples was previously validated [41], while no significant mRNA deterioration was found when FFPE slides at different storage time were compared after DSP analysis [42].

To conclude, both statistical analyses in whole slide and TMA samples found that morphological features on their own were stronger predictors than the molecular ones. An AI-based approach to better profile morphology of ccRCC TME could help standardise prognostic tools and overcome inter-observer bias. However, when molecular markers were integrated in the LS, the accuracy in stratifying patients at intermediate risk significantly improved. Therefore, molecular markers are crucial for identifying differences that are not noticeable by only looking at the tumour morphology. Moreover, adding prognostic immune markers to these scoring algorithms may add important information about ICI response, and may be crucial to apply the appropriate personalised therapy.

**Author Contributions:** Conceptualisation, R.D.F., G.W., D.J.H., P.D.C. and O.A.; methodology, R.D.F., I.H.U., A.W., G.W., E.S. and E.T.; software, R.D.F. and G.W.; validation, R.D.F. and G.W.; formal analysis, R.D.F. and G.W.; data curation, R.D.F. and G.W.; writing—original draft preparation, R.D.F. and G.W.; writing—review and editing, R.D.F., G.W., I.H.U., P.D.C., S.W., A.W., E.S., E.T., O.A. and D.J.H.; supervision, D.J.H., O.A. and S.W.; project administration, D.J.H. and O.A.; funding acquisition, D.J.H. and O.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by Medical Research Scotland (MRS), NHS Lothian, NanoString Technologies, and the Industrial Centre for AI Research in Digital Diagnostics (iCAIRD) which is funded by Innovate UK on behalf of UK Research and Innovation (UKRI) [project number: 104690].

**Institutional Review Board Statement:** The work was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of NHS Lothian NRS BioResource, REC-approved Research Tissue Bank (REC approval ref: 15/ES/0094, 23/04/2018).

**Data Availability Statement:** Data are currently in the process of being made available, pending ethical approval.

**Acknowledgments:** We would like to acknowledge Mrs Frances Rae and the SCOTRCC and NHS Lothian tissue banks for collecting and collating the TMA samples used in this study, Mr John O'Connor for preparing the whole slide samples, and East of Scotland Research Ethics Service for granting the ethical approval. Further acknowledgements go to Dr Jason Reeves for helping with the R software scripts.



**Table A.** Clinical data.

<b>Gender</b>	<b>Male</b> 86 (57%)	<b>Female</b> 60 (40%)	<b>NA</b> 4 (3%)		
<b>Age (years)</b>	<b>Mean</b> 64	<b>Min</b> 32	<b>Max</b> 94		
<b>Follow-up (months)</b>	<b>Mean</b> 60	<b>Min</b> 4	<b>Max</b> 252		
<b>Status at end of FU</b>	<b>Dead</b> 102 (68%)	<b>Alive</b> 40 (26.6%)	<b>NA</b> 8 (5.3%)		
<b>Cancer-related death</b>	<b>Yes</b> 75 (50%)	<b>No</b> 75 (50%)	<b>NA</b> -		
<b>Metastasis at diagnosis</b>	<b>Yes</b> 46 (30%)	<b>No</b> 105 (70%)	<b>NA</b> -		
<b>Leibovich Score risk</b>	<b>High</b> 17 (13.7%)	<b>Intermediate</b> 99 (79.8%)	<b>Low</b> 8 (6.5%)		
<b>Lymph-node involvement</b>	<b>Yes</b> 25 (16%)	<b>No</b> 125 (84%)	<b>NA</b> -		
<b>Stage</b>	<b>1</b> 8 (6.5%)	<b>2</b> 11 (8.9%)	<b>3</b> 100 (81.5%)	<b>4</b> 4 (3.2%)	<b>NA</b> -
<b>ISUP Grade</b>	<b>1</b> 19 (15.3%)	<b>2</b> 49 (39.5%)	<b>3</b> 30 (24.2%)	<b>4</b> 3 (2.4%)	<b>NA</b> -

**Conflicts of Interest:** Raffaele De Filippis was partly funded by Nanostring Technologies for his PhD. Sarah Warren and Andrew White are employed by NanoString Technologies. 416  
417

### Appendix A. Patient information 418

See Table A. 419

### Appendix B. Antibodies 420

The antibodies used in the experiments are listed in Tables B1 and B2. 421  
422

**Table B1.** Antibodies used in the IF experiments.

Antibody	Brand (Catalog No)	Species	Dilution	Target
CD163	Cell Marque (MRQ-26)	Mouse	1:3000	M2 TAMs
HIF-2 $\alpha$	Abcam (ab8265)	Mouse	1:400	Hypoxia
CD105	Abcam (Ab114052)	Mouse	1:1500	Blood vessels
CD8	Agilent (M7 103)	Mouse	1:200	Cytotoxic T cells
CA9	Novus Biological (NB100-417)	Rabbit	1:100	Tumour cells
Pan-Cadherin	Cell Signaling Technology (4068)	Rabbit	1:100	Tumour cells
$\beta$ -catenin	Cell Signaling Technology (8480S)	Rabbit	1:100	EMT
OCT4a	Cell Signaling Technology (C30A3)	Rabbit	1:800	CSLCs
SNAIL	Cell Signaling Technology (3879)	Rabbit	1:500	EMT
Vimentin	Cell Signaling Technology (5741)	Rabbit	1:200	EMT
ZEB1	Cell Signaling Technology (70512)	Rabbit	1:500	EMT
PD-1	Cell Signaling Technology (86163)	Rabbit	1:500	Immune escape
PD-L1	Cell Signaling Technology (13684)	Rabbit	1:500	Immune escape
TIM-3	Abcam (ab185703)	Rabbit	1:500	T cell exhaustion
LAG-3	Novus Biological (NBP1-97657)	Mouse	1:400	T cell exhaustion
CD44	Cell Signaling Technology (37259)	Rabbit	1:500	CSLCs
ZEB1	Proteintech (21544-1-AP)	Rabbit	1:500	EMT

**Table B2.** Antibodies used in the GeoMx<sup>®</sup> DSP experiment.

Antibody	Panel	Main expression site
4-1BB	Drug target	Cytotoxic T cells
ARG1	Tumour marker	Tumour cells
b-2-microglobulin	Immune cell profiling	Tumour cells
B7-H3	Drug target	Tumour cells
Bcl-2	Tumour marker	Tumour cells
CD11c	Immune cell profiling	Monocytes/macrophages
CD127	Immune activation status	Memory T cells
CD14	Immune cell typing	Monocytes, myeloid cells
CD163	Immune cell typing	M2 macrophages
CD20	Immune cell profiling	B cells
CD25	Immune activation status	T cells
CD3	Immune cell profiling	T cells
CD34	Immune cell typing	Hematopoietic cells
CD4	Immune cell profiling	T-helper cells
CD40	Immune activation status	B cells/APCs
CD45	Immune cell profiling	Macrophages
CD45RO	Immune cell typing	Memory T cells
CD56	Immune cell profiling	NK cells
CD68	Immune cell profiling	Macrophages
CD8	Immune cell profiling	T cells
CD80	Immune activation status	Myeloid cells
CTLA4	Immune cell profiling	T cells
EpCAM	Tumour marker	Tumour cells
FAP- $\alpha$	Immune cell typing	fibroblasts
Fibronectin	Immune cell profiling	Fibroblasts
GADPH	Positive control	Housekeeper
GITR	Drug target	T cells
GZMB	Immune cell profiling	Cytotoxic T cells
Histone H3	Positive control	Housekeeper
HLA-DR	Immune cell profiling	APCs
ICOS	Immune activation status	T cells
IDO1	Drug target	Myeloid cells
Ki-67	Tumour marker	Tumour cells
LAG-3	Drug target	T cells
OX40L	Drug target	Myeloid cells, T cells
PanCK	Tumour marker	Tumour cells
PD-1	Immune cell profiling	T cells
PD-L1	Immune cell profiling	APCs, tumour cells
PD-L2	Immune cell profiling	APCs, tumour cells
S6	Positive control	Housekeeper
SMA	Immune cell profiling	Fibroblasts
STING	Drug target	Immune cells
TGFB1	Tumour marker	Immune cells
TIM-3	Drug target	T cells
VISTA	Drug target	Myeloid cells, macrophages

## References

1. Siegel, R.L.; Miller, K.D.; Fuchs, H.E.; Jemal, A. Cancer statistics, 2021. *CA: a cancer journal for clinicians* **2021**, *71*, 7–33. 423
2. Eggener, S.E.; Yossepowitch, O.; Pettus, J.A.; Snyder, M.E.; Motzer, R.J.; Russo, P. Renal cell carcinoma recurrence after nephrectomy for localized disease: predicting survival from time of recurrence. *Journal of clinical oncology* **2006**, *24*, 3101–3106. 425
3. Padala, S.A.; Barsouk, A.; Thandra, K.C.; Saginala, K.; Mohammed, A.; Vakiti, A.; Rawla, P.; Barsouk, A. Epidemiology of renal cell carcinoma. *World journal of oncology* **2020**, *11*, 79. 427
4. Wei, W.; Lv, Y.; Gan, Z.; Zhang, Y.; Han, X.; Xu, Z. Identification of key genes involved in the metastasis of clear cell renal cell carcinoma. *Oncology letters* **2019**, *17*, 4321–4328. 428
5. Moch, H.; Cubilla, A.L.; Humphrey, P.A.; Reuter, V.E.; Ulbright, T.M. The 2016 WHO classification of tumours of the urinary system and male genital organs—part A: renal, penile, and testicular tumours. *European urology* **2016**, *70*, 93–105. 429
6. Warren, A.Y.; Harrison, D. WHO/ISUP classification, grading and pathological staging of renal cell carcinoma: standards and controversies. *World journal of urology* **2018**, *36*, 1913–1926. 430
7. Bektas, S.; Bahadir, B.; Kandemir, N.O.; Barut, F.; Gul, A.E.; Ozdamar, S.O. Intraobserver and interobserver variability of Fuhrman and modified Fuhrman grading systems for conventional renal cell carcinoma. *The Kaohsiung Journal of Medical Sciences* **2009**, *25*, 596–600. 431
8. Ren, D.; Hua, Y.; Yu, B.; Ye, X.; He, Z.; Li, C.; Wang, J.; Mo, Y.; Wei, X.; Chen, Y.; et al. Predictive biomarkers and mechanisms underlying resistance to PD1/PD-L1 blockade cancer immunotherapy. *Molecular cancer* **2020**, *19*, 1–19. 432
9. Giraldo, N.A.; Becht, E.; Vano, Y.; Petitprez, F.; Lacroix, L.; Validire, P.; Sanchez-Salas, R.; Ingels, A.; Oudard, S.; Moatti, A.; et al. Tumor-infiltrating and peripheral blood T-cell immunophenotypes predict early relapse in localized clear cell renal cell carcinoma. *Clinical Cancer Research* **2017**, *23*, 4416–4428. 433
10. Chen, X.; Wang, H.; et al. Prognostic value of vimentin is associated with immunosuppression in metastatic renal cell carcinoma. *Frontiers in oncology* **2020**, *10*, 1181. 434
11. Kalluri, R.; Weinberg, R.A.; et al. The basics of epithelial-mesenchymal transition. *The Journal of clinical investigation* **2009**, *119*, 1420–1428. 435
12. Dongre, A.; Weinberg, R.A. New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. *Nature reviews Molecular cell biology* **2019**, *20*, 69–84. 436
13. Andreiana, B.C.; Stepan, A.E.; Mărgăritescu, C.; Tăiesescu, O.; Osman, A.; Simionescu, C. Snail and E-cadherin immunoexpression in clear cell renal cell carcinoma. *Current Health Sciences Journal* **2019**, *45*, 185. 437
14. Wu, J.; Ru, N.; Zhang, Y.; Li, Y.; Wei, D.; Ren, Z.; Huang, X.; Chen, Z.; Bian, H. HAB18G/CD147 promotes epithelial–mesenchymal transition through TGF- $\beta$  signaling and is transcriptionally regulated by Slug. *Oncogene* **2011**, *30*, 4410–4427. 438
15. Zhang, P.; Sun, Y.; Ma, L. ZEB1: at the crossroads of epithelial-mesenchymal transition, metastasis and therapy resistance. *Cell cycle* **2015**, *14*, 481–487. 439
16. Liu, C.Y.; Lin, H.H.; Tang, M.J.; Wang, Y.K. Vimentin contributes to epithelial-mesenchymal transition cancer cell mechanics by mediating cytoskeletal organization and focal adhesion maturation. *Oncotarget* **2015**, *6*, 15966. 440
17. O'Brien, S.J.; Fiechter, C.; Burton, J.; Hallion, J.; Paas, M.; Patel, A.; Patel, A.; Rochet, A.; Scheurlen, K.; Gardner, S.; et al. Long non-coding RNA ZFAS1 is a major regulator of epithelial-mesenchymal transition through miR-200/ZEB1/E-cadherin, vimentin signaling in colon adenocarcinoma. *Cell death discovery* **2021**, *7*, 1–14. 441
18. Kröger, C.; Afeyan, A.; Mraz, J.; Eaton, E.N.; Reinhardt, F.; Khodor, Y.L.; Thiru, P.; Bieri, B.; Ye, X.; Burge, C.B.; et al. Acquisition of a hybrid E/M state is essential for tumorigenicity of basal breast cancer cells. *Proceedings of the National Academy of Sciences* **2019**, *116*, 7353–7362. 442
19. Myszczyszyn, A.; Czarnecka, A.M.; Matak, D.; Szymanski, L.; Lian, F.; Kornakiewicz, A.; Bartnik, E.; Kukwa, W.; Kieda, C.; Szczylik, C. The role of hypoxia and cancer stem cells in renal cell carcinoma pathogenesis. *Stem cell reviews and reports* **2015**, *11*, 919–943. 443
20. Zöller, M. CD44: can a cancer-initiating cell profit from an abundantly expressed molecule? *Nature Reviews Cancer* **2011**, *11*, 254–267. 444
21. Li, X.; Ma, X.; Chen, L.; Gu, L.; Zhang, Y.; Zhang, F.; Ouyang, Y.; Gao, Y.; Huang, Q.; Zhang, X. Prognostic value of CD44 expression in renal cell carcinoma: a systematic review and meta-analysis. *Scientific reports* **2015**, *5*, 1–8. 445
22. Ma, C.; Komohara, Y.; Ohnishi, K.; Shimoji, T.; Kuwahara, N.; Sakumura, Y.; Matsuiishi, K.; Fujiwara, Y.; Motoshima, T.; Takahashi, W.; et al. Infiltration of tumor-associated macrophages is involved in CD 44 expression in clear cell renal cell carcinoma. *Cancer Science* **2016**, *107*, 700–707. 446
23. Zhou, Y.; Chen, X.; Kang, B.; She, S.; Zhang, X.; Chen, C.; Li, W.; Chen, W.; Dan, S.; Pan, X.; et al. Endogenous authentic OCT4A proteins directly regulate FOS/AP-1 transcription in somatic cancer cells. *Cell death & disease* **2018**, *9*, 1–16. 447
24. Rasti, A.; Mehrzama, M.; Madjd, Z.; Abolhasani, M.; Zanjani, L.S.; Asgari, M. Co-expression of cancer stem cell markers OCT4 and NANOG predicts poor prognosis in renal cell carcinomas. *Scientific reports* **2018**, *8*, 1–11. 448
25. Waggoner, S.N.; Cornberg, M.; Selin, L.K.; Welsh, R.M. Natural killer cells act as rheostats modulating antiviral T cells. *Nature* **2012**, *481*, 394–398. 449
26. Comber, J.D.; Philip, R. MHC class I antigen presentation and implications for developing a new generation of therapeutic vaccines. *Therapeutic advances in vaccines* **2014**, *2*, 77–89. 450

27. Caie, P.D.; Dimitriou, N.; Arandjelović, O. Precision medicine in digital pathology via image analysis and machine learning. In *Artificial intelligence and deep learning in pathology*; Elsevier, 2021; pp. 149–173. 481
28. Geiss, G.K.; Bumgarner, R.E.; Birditt, B.; Dahl, T.; Dowidar, N.; Dunaway, D.L.; Fell, H.P.; Ferree, S.; George, R.D.; Grogan, T.; et al. 482  
Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature biotechnology* **2008**, *26*, 317–325. 483
29. Merritt, C.R.; Ong, G.T.; Church, S.E.; Barker, K.; Danaher, P.; Geiss, G.; Hoang, M.; Jung, J.; Liang, Y.; McKay-Fleisch, J.; et al. 484  
Multiplex digital spatial profiling of proteins and RNA in fixed tissue. *Nature biotechnology* **2020**, *38*, 586–599. 485
30. Chow, K.V.; Denning, K.C. A simple multiple variance ratio test. *Journal of Econometrics* **1993**, *58*, 385–401. 486
31. Coelho, L.P. Jug: Software for parallel reproducible computation in python. *Journal of Open Research Software* **2017**, *5*. 487
32. Beechem, J.M. High-plex spatially resolved RNA and protein detection using digital spatial profiling: a technology designed for 488  
immuno-oncology biomarker discovery and translational research. In *Biomarkers for Immunotherapy of Cancer*; Springer, 2020; pp. 489  
563–583. 490
33. Raggi, C.; Mousa, H.; Correnti, M.; Sica, A.; Invernizzi, P. Cancer stem cells and tumor-associated macrophages: a roadmap for 491  
multitargeting strategies. *Oncogene* **2016**, *35*, 671–682. 492
34. Harb, O.A.; Elfeky, M.A.; El Shafaay, B.S.; Taha, H.F.; Osman, G.; Harera, I.S.; Gertallah, L.M.; Abdelmonem, D.M.; Embaby, 493  
A. SPOP, ZEB-1 and E-cadherin expression in clear cell renal cell carcinoma (cc-RCC): Clinicopathological and prognostic 494  
significance. *Pathophysiology* **2018**, *25*, 335–345. 495
35. Shao, M.; Bi, T.; Ding, W.; Yu, C.; Jiang, C.; Yang, H.; Sun, X.; Yang, M. OCT4 potentiates radio-resistance and migration activity of 496  
rectal cancer cells by improving epithelial-mesenchymal transition in a ZEB1 dependent manner. *BioMed Research International* 497  
**2018**, *2018*. 498
36. Mani, S.A.; Guo, W.; Liao, M.J.; Eaton, E.N.; Ayyanan, A.; Zhou, A.Y.; Brooks, M.; Reinhard, F.; Zhang, C.C.; Shipitsin, M.; et al. 499  
The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **2008**, *133*, 704–715. 500
37. Najafi, M.; Farhood, B.; Mortezaee, K.; Kharazinejad, E.; Majidpoor, J.; Ahadi, R. Hypoxia in solid tumors: a key promoter of 501  
cancer stem cell (CSC) resistance. *Journal of cancer research and clinical oncology* **2020**, *146*, 19–31. 502
38. Xu, C.; Jackson, S.A. Machine learning and complex biological data, 2019. 503
39. Bonferroni, C. Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e* 504  
*Commerciabili di Firenze* **1936**, *8*, 3–62. 505
40. Grillo, F.; Bruzzone, M.; Pigozzi, S.; Prosapio, S.; Migliora, P.; Fiocca, R.; Mastracci, L. Immunohistochemistry on old archival 506  
paraffin blocks: is there an expiry date? *Journal of Clinical Pathology* **2017**, *70*, 988–993. 507
41. Bueno, R.; Hughes, E.; Wagner, S.; Gutin, A.S.; Lanchbury, J.S.; Zheng, Y.; Archer, M.A.; Gustafson, C.; Jones, J.T.; Rushton, K.; et al. 508  
Validation of a molecular and pathological model for five-year mortality risk in patients with early stage lung adenocarcinoma. 509  
*Journal of Thoracic Oncology* **2015**, *10*, 67–73. 510
42. Gupta, S.; Chen, T.; Destenaves, B. Quantitative RNA assessment and long-term stability in the FFPE tumor samples using Digital 511  
Spatial Profiler. *Immuno-Oncology and Technology* **2022**, *13*, 100069. 512  
513