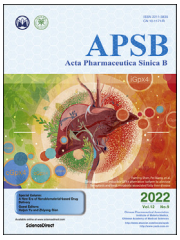
Acta Pharmaceutica Sinica B 2022;12(9):3594–3601

Chinese Pharmaceutical Association
Institute of Materia Medica, Chinese Academy of Medical Sciences

Acta Pharmaceutica Sinica B

www.elsevier.com/locate/apsb
www.sciencedirect.com



TOOLS

An integrative approach of digital image analysis and transcriptome profiling to explore potential predictive biomarkers for TGFβ blockade therapy

Robert Pomponio[a], Qi Tang[b], Anthony Mei[b], Anne Caron[c], Bema Coulibaly[c,#], Joachim Theilhaber[d,#], Maximilian Rogers-Grazado[d,#], Michele Sanicola-Nadel[e,#], Souad Naimi[c,#], Reza Olfati-Saber[b], Cecile Combeau[a], Jack Pollard[d], Tun Tun Lin[e,#], Rui Wang[a,*]

[a]*Sanofi Translational Medicine Early Development Cambridge, MA 02142, USA*
[b]*Sanofi Artificial Intelligence and Deep Analytics Data and Data Sciences, Cambridge, MA 02142, USA*
[c]*Sanofi Translational Science, Vitry sur Seine 94403, France*
[d]*Sanofi Precision Oncology Cambridge, MA 02142, USA*
[e]*Sanofi Oncology Early Development Cambridge, MA 02142, USA*

Received 14 December 2021; received in revised form 15 January 2022; accepted 3 March 2022

KEY WORDS

Digital pathology;
Artificial intelligence;
TGFβ;
Predictive biomarker;
Tumor topography;
T cell infiltration;
Transcriptomic profiling;
Machine learning

Abstract Increasing evidence suggests that the presence and spatial localization and distribution pattern of tumor infiltrating lymphocytes (TILs) is associate with response to immunotherapies. Recent studies have identified TGFβ activity and signaling as a determinant of T cell exclusion in the tumor microenvironment and poor response to PD-1/PD-L1 blockade. Here we coupled the artificial intelligence (AI)-powered digital image analysis and gene expression profiling as an integrative approach to quantify distribution of TILs and characterize the associated TGFβ pathway activity. Analysis of T cell spatial distribution in the solid tumor biopsies revealed substantial differences in the distribution patterns. The digital image analysis approach achieves 74% concordance with the pathologist assessment for tumor-immune phenotypes. The transcriptomic profiling suggests that the TIL score was negatively correlated with TGFβ pathway activation, together with elevated TGFβ signaling activity observed in excluded and

*Corresponding author.
E-mail address: rui.wang@sanofi.com (Rui Wang).
[#]was an employee of Sanofi at the time the studies were performed.
Peer review under responsibility of Chinese Pharmaceutical Association and Institute of Materia Medica, Chinese Academy of Medical Sciences.

https://doi.org/10.1016/j.apsb.2022.03.013

desert tumor phenotypes. The present results demonstrate that the automated digital pathology algorithm for quantitative analysis of CD8 immunohistochemistry image can successfully assign the tumor into one of three infiltration phenotypes: immune desert, immune excluded or immune inflamed. The association between "cold" tumor-immune phenotypes and TGFβ signature further demonstrates their potential as predictive biomarkers to identify appropriate patients that may benefit from TGFβ blockade.

1. Introduction

Recent studies suggest that the spatial context of tumor infiltrating lymphocytes is an important prognostic determinant of tumor progression and response[1–3]. The distribution pattern of tumor infiltrating lymphocytes in the tumor microenvironment (TME) can be classified as three T cell infiltration phenotypes: immune-inflamed, immune-excluded and immune-desert, based on the spatial localization of immune cells with respect to the tumor and stromal compartments[4]. Immune-inflamed tumors are usually associated with increased level of infiltrating antigen-specific $CD8^+$ T cells and better response to immunotherapies[5]. The immune-desert has been associated with general lack of immune response due to minimal infiltration of TILs[6,7]. Finally, the immune-excluded tumors display a phenotype in which $CD8^+$ T cells are restricted to surrounding tumor stroma compartment but not able to infiltrate into tumor parenchyma[4]. Thus, overcoming the restriction barriers to enable re-distribution of effector T cells from stroma to close contact with cancer cells might serve as a strategy to improve efficacy of immune checkpoint blockade.

Recent studies have identified TGFβ activity and signaling in the TME as a determinant of cytotoxic CD8 T cell exclusion and poor response to PD-1/PD-L1 blockade[8]. Analysis of transcriptome sequencing data suggested significantly enhanced TGFβ signaling pathway in biopsies from patients not responding to PD-L1 antagonist atezolizumab in metastatic urothelial cancer study. Further back translational research in syngeneic mouse tumor models which recapitulated the immune-excluded TME demonstrated that dual targeting of PD-L1 and TGFβ reverses immune exclusion by inducing $CD8^+$ T cell infiltration into tumors, while treatment of either therapy alone is not effective[9]. The results shed light on leveraging TGFβ inhibition to enhance antitumor response of immune checkpoint inhibitors by changing the immune landscape in TME.

Over the years the clinical development of TGFβ inhibitors have been hampered by lack of efficient predictive biomarkers to identify patients who are likely to benefit from the treatment[10]. The recent study indicating the role of TGFβ signaling pathway as key driver of immune-exclusion provided additional insight for combining TGFβ blockade with checkpoint inhibitors in biomarker-driven clinical trials[11,12]. We applied an integrative approach using image analysis together with gene expression profiling to quantify distribution of TILs and characterize the associated TGFβ pathway activity. The goal is to use this approach to identify tumors with immune-excluded phenotype and high levels of TGFβ pathway activity, as predictors of response to TGFβ inhibition to facilitate conversion of "cold" (immune excluded and desert) tumors to "hot" in the future clinical studies.

However, determining a patient's immune phenotype from their digital imaging slides can be a subjective process based on the individual pathologist's analysis—discrepancies between pathologists (inter-rater) and discrepancies between analyses from the same pathologist but at different times (intra-rater) are somewhat common challenges with these datasets[13,14]. To overcome these challenges, we developed a quantitative measure of immune phenotypes (TILS/NTILS) and applied it to our image processing algorithm-based model in an automated AI-driven fashion to reduce inter-rater and intra-rater variance in immune phenotype classification.

2. Materials and methods

2.1. Archival formalin-fixed, paraffin-embedded (FFPE) tumor blocks

110 individual tumors mounted in FFPE comprising colorectal cancer (CRC, $n = 30$), bladder cancer (BC, $n = 20$), ovarian cancer (OC, $n = 20$), head and neck squamous cell carcinoma (HNSCC, $n = 20$), and gastric cancer (GC, $n = 20$) were selected based on a pathologist review of available blocks for cases with a balance of tumor and stroma represented and low necrosis. Selected blocks were purchased from a commercial vendor (Asterand) who obtained informed consent for collection and research use from each donor and the respective IRB of the participating institutions. Sectioning and staining of FFPE blocks for IHC analysis were performed by Indivumed GmbH (Hamburg, Germany). An additional set of 5-μm sections were collected as FFPE curls for RNA isolation and RNAseq analysis. Of those initial 110 blocks, 105 passed QC after staining and were used for the subsequent analyses. Additional details of the tumor tissue have been summarized in Table 1.

2.2. Multiplex immunohistochemistry (IHC) assay and digital pathology

A multiplex fluorescent IHC assay was developed by Indivumed GmbH for PanCK (polyclonal)/CD3 (clone F7.2.38)/CD8 (clone SP16), and 4′,6-diamidino-2-phenylindole (DAPI). Image analysis was optimized to detect localization of $CD3^+$ and $CD8^+$ lymphocytes on the procured tumor slides. The multiplex immunofluorescent slide with a hematoxylin counterstain were scanned at 20 × magnification for the image analysis. Pathologist annotations were used to determine tumor and stroma regions in each sample based on the distribution of the T-cells and PanCK staining. The spatial localization of lymphocyte subsets as defined by the $CD8^+$ staining was used to assign each case to one of the three infiltration phenotypes (immune desert, immune excluded or immune inflamed). In addition, the digital imaging analysis platform for

Table 1 Summary of procured tumor biopsies from five indications.

Case	CRC	HNSCC	Ovarian	Bladder	Gastric
Female cases	20	4	20	8	9
(mean age years at time of biopsy [range])	(64.25 [50−85])	(68 [58−82])	(63.8 [43−78])	(69 [37−82])	(65.4 [43−80])
Male cases	10	15	Not applicable	12	11
(mean age years at time of biopsy [range])	(64 [45−84])	(64.3 [56−88])		(62 [46−81])	(61.2 [27−77])
AJCC/UI tumor stage group (% of cases)	IIIB (70%)	III (60%)	IIIC (80%)	G2 (20%)	G1 (15%)
	IV (20%)	IVA (20%)	IV (20%)	G3 (80%)	G2 (15%)
	IVA (7%)	IVC (20%)			G3 (70%)
	IVB (3%)				

quantitative assessment of immune phenotypes has been developed and applied to 103 tumor samples.

2.3. *Development and refinement of the artificial intelligence digital pathology algorithm*

The full set of the $CD8^+$ fluorescent images from the 103 tumor samples were first split into a training and an independent test dataset based on different tumor types to better understand the generalizability of the immune phenotyping model. Image data from colorectal cancer, gastric, and bladder cancer indications were input as the training set ($n = 68$), while HNSCC and OC data were used as the test data set ($n = 35$) (Fig. 3A).

Multiple digital image morphology operations were combined to preprocess and normalize each image. Several methods were considered: Otsu's thresholding[15], fixed thresholding, median and variance based thresholding. For each channel a median and standard deviation-based thresholding to separate signal from noise was used: threshold = (median of positive pixel value within tissue level tumor mask) + 1.5 (standard deviation of positive pixel value within tissue level tumor mask) and each channel had its own global threshold which was applied to all tiles. The size of the tile was chosen to be able to cover several tumor nests but not too large such that it could not be read into the computer due to memory constraints. The chosen size for this project was 2500 × 2500. Four types of stains were used to stain the four channels of the immunofluorescence images. The pan-CK stains epithelial cells, majority of which are tumor cells in the images since they are from tumor biopsy samples. DAPI stains cell nuclei, but sometimes there are staining artifacts.

The probes for CD3 and CD8 positive lymphocytes, can also stain red blood cells so these two stains and the cell nuclei channel were used to exclude false CD3, CD8 staining. In the phenotype classification, only three channels were used: CD8, pan-CK and DAPI. Top-hat image morphology operation was used to correct out-of-focus region and staining artifacts[16]. A kernel size of 75 was used for the top-hat operation and was chosen to be larger than the size of largest cells but smaller than the smallest artifact or region out-of-focus such that it removes the artifacts but not the cells. The opening image morphology operation was used with a kernel that is just right size, bigger than largest size of the small artifacts but smaller than the smallest cells. A kernel size of 20 was used for the epithelial channel while a size of five was used for the rest of channels. The closing image morphology operation was used with a kernel that is just right size. A kernel size of 35 was used for the epithelial channel while a size of 10 was used for the nuclei channel. The watershed image morphology operation was used to segment $CD8^+$ lymphocytes.

2.4. *TGFβ pathway activation gene signature*

TGFβ-stimulated and anti-TGFβ-treated MDA-MB-231 human breast cancer cell lines were cultured *in vitro* and used to derive a gene signature for TGFβ pathway activation. This resulted in a 159 gene expression signature that was validated *in vivo* by RNA profiling of the MDA-MD-231 xenograft model[17]. *In silico* validation was accomplished through comparison to other TGFβ signature scores[18,19]. The full gene signature and the gene list has been disclosed previously[20].

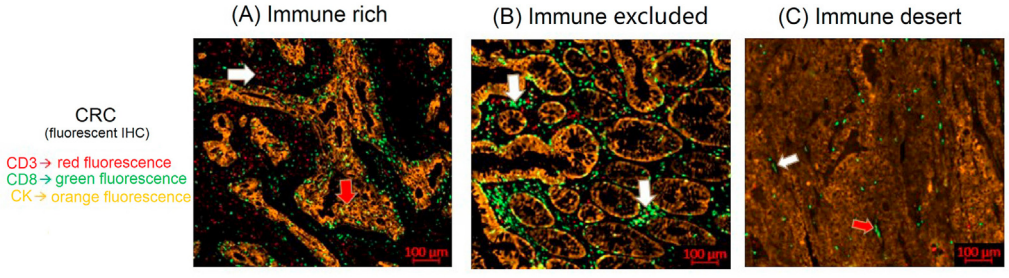


Figure 1 Characterization of the spatial distribution of tumor infiltrating T lymphocytes using multiplex immunohistochemistry and immune fluorescence assay. Representative multiplex immunofluorescent image of colorectal cancer illustrating localization of $CD3^+$ and $CD8^+$ T cells. Panel A: Immune rich phenotype. T lymphocytes are numerous both in stroma (white arrow) and between tumor cells (red arrow). Panel B: Immune excluded pattern. T lymphocytes are restricted in the stroma around tumor cells (white arrows). Panel C: Immune desert pattern. T lymphocytes are scarce both in stroma (white arrow) and between tumor cells (red arrow).

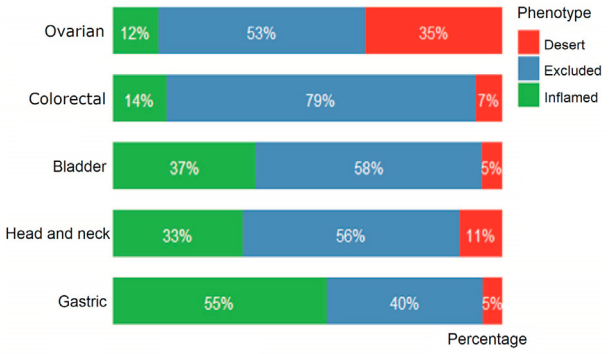


Figure 2 Distribution and frequency by tumor type for each immune cell phenotype for the cases in this study.

2.5. *Gene expression profiling*

Total RNAs of individual CRC samples ($n = 29$) were isolated from 5-µm thick FFPE sections of tumor blocks. Gene expression profiling was then performed using next-generation sequencing, briefly as follows: whole-transcriptome libraries were generated from the isolated RNA using a single-stranded protocol (NuGen), sequencing was performed using the Illumina HiSeq2000 platform, sequencing reads were then informatically mapped to transcripts, and relative abundances computed as transcripts per million (TPMs). Median sequencing depth was 14.7 million, with median mapping rate to the genome of 79%. Of mapped reads, on median 87% mapped to exonic sequences.

A regulated Kolmogorov–Smirnov analysis was used to generate enrichment scores for the TGFβ pathway activation signature according to a previously disclosed approach[20]. The profiles arising from each study were independently quantile-normalized, $\log_2$ transformed, and then Z transformed (standardized) on a gene-by-gene basis, before being used to generate the enrichment scores. The final enrichment scores were expressed in terms of "$\log_2$C" scores, with sign equal to the inferred relative activation state of the pathway ("on" = 1, "off" = −1) and magnitude equal to $\log_2$ of the largest of the left or right leading-edge slopes of the regulated sample distribution relative to the global gene population rank distribution.

2.6. *Statistical analysis*

A trend test[21] was used to compute global significance of differences between cancer immune phenotype groups in the analysis of the TGFβ pathway activation score. A two-sided *t*-test against 0 was conducted to determine whether there is an increasing trend of TGFβ pathway activation score from inflamed phenotype to excluded and desert phenotype.

3. Results

3.1. *Characterization and spatial distribution of tumor-infiltrating T cells and profile the immune topographies across five different solid tumors by multiplex immunohistochemistry*

Of the 110 cases analyzed, 105 yielded staining and images of sufficient quality to support further analysis. Five (1 CRC, one HNSCC, one OC and 2 GC) cases failed due to quality issues mostly related to the sample not adhering to the slide during processing and files from two other cases were corrupted during data transfer and were excluded from the study. Multiplex fluorescent images from 103 cases from the five tumor types were collected. A set of CRC tumors from individual cases provided representative examples of each immune cell infiltration phenotype and are depicted in Fig. 1. The immune rich or inflamed phenotype is illustrated by CD3 and CD8 positive cell staining throughout the sample, indicating an immune infiltration across the boundaries of the tumor and adjacent stroma. Given $CD8^+$ T cells are one of the primary tumor infiltrating lymphocytes that play a key role in anti-tumor responses and CD8 T cells are considered as a useful biomarker for prediction of prognosis and response to immunotherapy[22,23], as we move forward with the digital pathology platform establishment, CD8 staining was

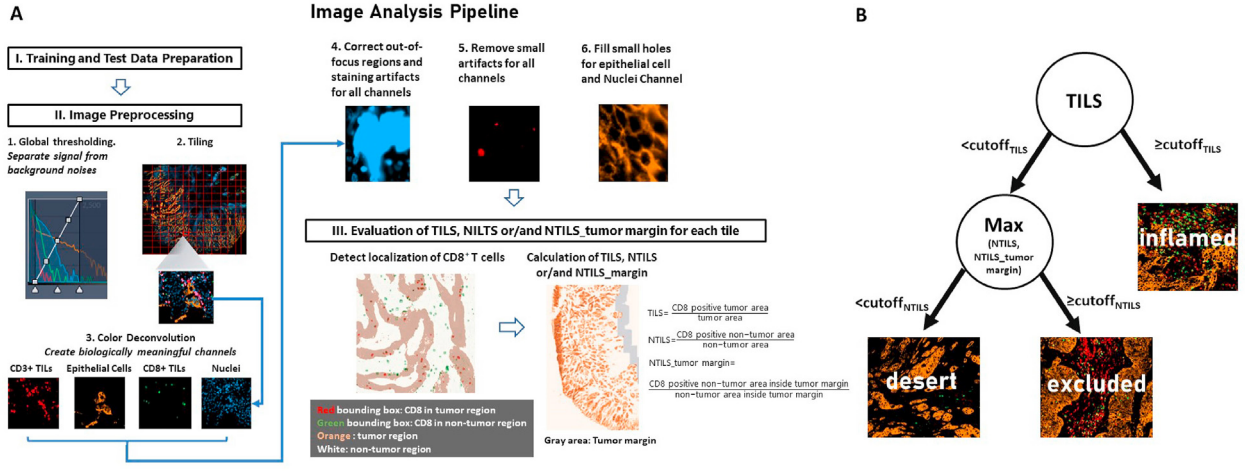


Figure 3 Digital image analysis pipeline for the quantitative evaluation of tumor immune cell phenotypes using $CD8^+$ immunofluorescent images. (A) Overview of workflow for digital image analysis pipeline. (B) Phenotype classification rule applied to sample images in the determination of immune cell phenotype.

prioritized as the leading marker to distinguish the cancer-immune cell distribution within TME.

The resulting multiplex images were reviewed and annotated by one Board Certified Medical Pathologist (BC) and one PhD in Molecular Histopathology (AC) which resulted in a "manual" classification of each case into one of the three immune cell phenotypes. The distribution of each phenotype by tumor type studied is depicted in Fig. 2. The profile of T cell infiltration phenotypes across the five different tumor types (Bladder, CRC, HNSCC, Gastric Cancer, and Ovarian Cancer) showed a distinct distribution pattern of CD 8^+ T lymphocytes in these biopsies ($n = 103$). Of interest is that the excluded phenotype exists as the majority phenotype in all but the gastric cases (40%) and is predominant in the colorectal cases in our study (79%). This observation is supported by similar findings in CRC by image analysis whereby a majority of cases were also found to have the lowest amount of tumor infiltrating lymphocytes and highest tumor-stroma ratio coinciding with the CMS-4 (mesenchymal) phenotype[24]. Overall, the gastric tumors had the highest percentage of the immune inflamed phenotype whereby ovarian and colorectal the lowest (12% and 14% respectively) (Fig. 2).

3.2. *Development of digital image analysis platform for quantitative evaluation of tumor immune phenotypes*

Due to heterogeneities of immunofluorescence images, an image preprocess procedure was conducted to remove staining artifacts, such as out-of-focus regions, false positive background staining and false positive red blood cell staining. Also included as part of the preprocessing process, tumor regions were annotated by a pathologist, which were utilized later to rule out PanCK stained normal epithelial cells when PanCK staining was used to segment out tumor regions. The annotation also helped the determination of the tumor margin for which distribution of $CD8^+$ lymphocytes were evaluated. After preprocessing, each image went through into a tiling process and only tiles contain enough nuclei were included for the rest of the analysis. To localize $CD8^+$ lymphocytes, a whole slide specific threshold was determined based on the mean and variance of the pixel values of the CD8 staining channel to segment out CD8 positive lymphocytes. Once segmentation was completed with watershed, a bounding box was generated for each segmented cell to locate each cell (summarized in Fig. 3A), which can be useful for both visual validation and the determination of the size of each cell to rule out large size staining artifacts that were removed during the preprocessing process. Similar thresholding and segmentation procedures were carried out for the rest of the channels of an image. Continuous scores were obtained to quantify spatial distribution of $CD8^+$ tumor infiltrating lymphocytes (TILs) for up to three compartments of each tile of an image, tumor compartment, stroma compartment and tumor margin compartment. If the area of a tile covered by the tumor compartment is greater than a prespecified threshold (say 5%), the percentage of tumor area covered by $CD8^+$ TILs will be evaluated and denoted as tumor infiltrating lymphocyte score (TILS) for that tile. If the area of a tile covered by the stroma compartment is greater than a prespecified threshold (say 0.1%), the percentage of stroma area covered by $CD8^+$ TILs will be evaluated and denoted as non-tumor-infiltrating-lymphocyte score (NTILS). If the area of a tile covered by tumor margin compartment is greater than a prespecified percentage (say 10%), the percentage of tumor margin area covered by $CD8^+$ TILs will be evaluated and denoted by NTILS_margin. For the calculation of all these scores, $CD8^+$ area was calculated as an aggregated sum of each segmented $CD8^+$ lymphocytes inside the compartment associated with each score. The scores were aggregated by taking the median of the scores of each tile to obtain the corresponding patient level TILS, NTILS and NTILS_margin scores. Based on these scores, a slide can be classified into one of the three immune phenotypes based on a tree decision rule outlined in Fig. 3B. The current method was found to be applicable to the five cancer types tested in the present study with decent performance. Sensitivity analysis on the impact of tile size to the phenotype classification has also been performed which confirms the robustness of current approach regarding to the choice of tile sizes as indicated in Table 2. Comparisons of digital to manual annotation were performed on a training set of tumors to refine thresholds and hyperparameters and applied to a testing set of tumors to derive specificity and sensitivity of the digital image analysis method. The refined thresholds for the immune phenotype classification rule in Fig. 3B are $\text{cutoff}_{\text{TILS}} = 0.6\%$, which can be interpreted as about 212 $CD8^+$ TILs per mm^2 of tumor cells, and $\text{cutoff}_{\text{NTILS}} = 0.3\%$, which can be interpreted as about 106 $CD8^+$ TILs per mm^2 of tumor cells, assuming the average diameter of a $CD8^+$ TIL is 6 μm.

3.3. *Performance of the AI digital pathology algorithm*

Hyperparameters for the digital pathology algorithm was trained and finetuned on the training set (Fig. 4A) and its performance was evaluated in the independent test set to derive specificity and sensitivity of the method (Fig. 4B). Our approach is different from the standard deep learning type of training process, where the training dataset and testing dataset use the samples from the same type of disease. In contrast, we used CRC and gastric cancer images for the training set while the images from HNSCC, ovarian and bladder cancers were used as the testing dataset. The purpose of creating training set and testing set with different cancer types was to examine the generalizability of the digital pathology algorithm to cancer types for which it was not trained on. The Receiver Operating Characteristic (ROC) curves, a commonly used metrics for evaluating performance of a classifier in machine learning, for predicting inflamed *vs*. non-inflamed phenotypes were visualized together with the ROC curves for predicting excluded *vs*. desert phenotypes in Fig. 4. Both ROC curves have more than 85% area under the curve (AUC) and demonstrated the applicability of the algorithm to support accurate characterization across the tumor types studied (Fig. 4C).

3.4. *Association between TGFβ pathway activation signature score and the infiltration phenotypes*

The RNAseq output of RNA derived from the 29 archival CRC cases was analyzed using the 159-gene expression signature representing TGFβ pathway activation[20] to generate a gene set enrichment score using a regulated Kolmogorov−Smirnov analysis for each case. Published studies have demonstrated TGFβ upregulation as a key mechanism of tumor immune evasion by contributing to a tumor microenvironment that is characterized by an immune cell exclusionary zone[25,26]. Previous analysis by our group confirmed that TGFβ pathway activation is significantly higher in a patient population that was unresponsive to anti-PD-L1 or anti-PD-1 checkpoint therapies[20] contributing to a resistance mechanism in these tumors. In this current study when we compared to the number of TILs in the archival CRC tumor

Table 2 Sensitivity analysis on the impact of tile size to the cancer immune phenotype classifications.

Tile size (# of pixels)	AUC for inflamed *vs.* non-inflamed (95% C.I.)	AUC for excluded *vs.* desert (95% C.I.)
1500	89.6% (72.9%, 100%)	86.8% (72.3%, 100%)
2500	86.1% (63.7%, 100%)	89.5% (77.4%, 100%)
3500	87.5% (66.8%, 100%)	88.2% (74.9%, 100%)

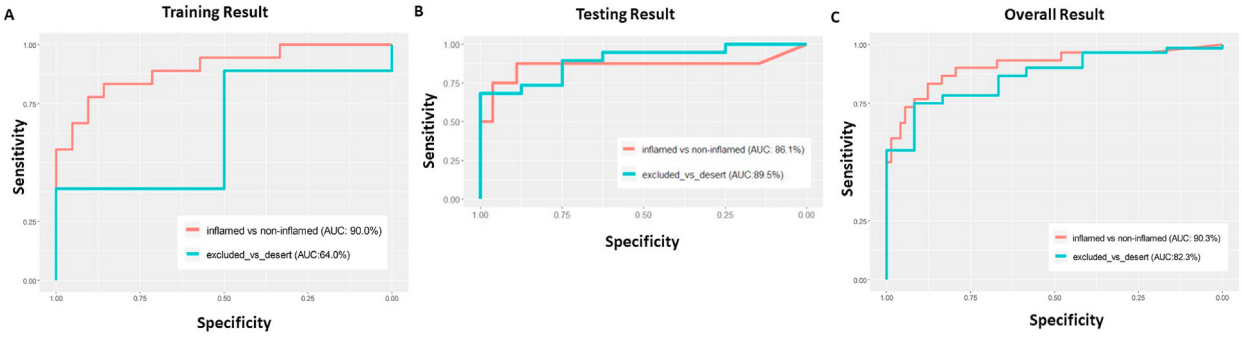


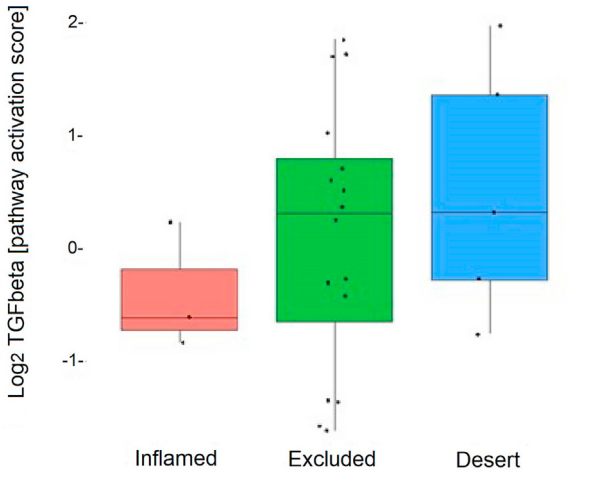
Figure 4 ROC curves depicting performance of the approach demonstrates the improvement in AUC from the (A) training data set to the (B) testing data set in correctly determining immune cell phenotype and the (C) overall generalizability to the five different tumor types studied. *X*-axis and *Y*-axis depict specificity and sensitivity, respectively.

sample to their respective TGFβ pathway activation signature score, there is a clear and inverse relationship between the TGFβ pathway activation score and number of TILS in the sample, especially in the tumors exhibiting an immune excluded or desert phenotype. This relationship was further extended to the level of TGFβ in the sample as a predictor of the immune phenotype exhibited by the sample. On average, higher TGF-β pathway activation score is observed in tumors with an excluded or desert phenotype. A trend test was conducted to test the increasing trend of TGFβ across three phenotypes, and two-sided P-value $= 0.32$ (Fig. 5).

4. Discussion

Assessing immune cell infiltration is an important component of the "Cancer Immunogram" for patient stratification in future immunotherapy trials[27]. However, robustly classifying tumor-immune phenotypes on a tumor and case by case basis with histological metrics has been challenging. Herein we have proposed the incorporation of a quantitative measure of immune phenotypes (TILS/NTILS) that can be automatically determined by our digital image process-based model to reduce inter-rater and intra-rater variance in immune phenotype classification. Traditional methods focus on discriminating the immune infiltration phenotypes based on the location of TILs in tumor or stroma areas, which can distinguish inflamed *vs.* non-inflamed phenotypes successfully but might be challenging to correctly differentiate between tumors with an excluded *vs.* desert phenotype.

In addition, as an important immune cell in tumor infiltrated lymphocytes, $CD8^+$ T cell has been regarded as a predictive biomarker and a key player in characterizing cancer-immune phenotypes[23,28]. However, with increasing studies pointing out that the heterogeneous distribution of CD8 T cells within a single tumor and across different tumors could complicate quantification of CD8 expression. Moreover, CD8 TIL is typically measured by immunohistochemistry for expression quantification purpose, thus, inter-observer variation between examining pathologists is another factor that contributing to the inconsistency of TIL quantification in clinical practice[22,29,30]. Taken together, intra- and inter-tumoral heterogeneity could distort the TIL quantification using CD8 IHC on single FFPE biopsy slide. Consequently, numerous studies have investigated novel cutting-edge technology, such as digital pathology analysis to quantify spatial heterogeneity of immune markers (CD8, PD-L1, etc.)[31], to pave the road for comprehensively studying tumor heterogeneity across multiple studies. Herein, our methodology goes beyond the tumor and non-tumor compartments and assesses the TILs in the tumor invasive-margin boundary which is defined as the region centered on the border between the tumor nest and host tissue if there is any. This allows for the measurement of TILs in the tumor margins which can indicate whether the tumor core is surrounded by TILs and thus specify the phenotype is excluded or desert based on their presence or absence, respectively. Additionally we developed a quantitative, continuous and automated analysis algorithm that is similar to the automated quantitative analysis technology that allows the detection of biomarker expression within specific compartment area to produce quantitative and reproducible score for each field of view[32]. This is consistent with the way that pathologists evaluate the tumor-immune phenotypes, and thus more accurate than using a simple cell count since tumor cells are known to have morphologic heterogeneity even within an individual tumor[33]. It is also worth noting that the present algorithm has high level of generalizability without the excessive overfitting need[34] which has the potential to be developed as an independent, pan-cancer approach: the algorithm was developed using a training dataset that included CRC, gastric and bladder cancer and then was applied to the test dataset that included HNSCC and ovarian cancer, with the results suggesting good accuracy in both the training and the independent test set. Last but not least, this AI type of digital pathology algorithm has an advantage over the more popular and widespread methods such as neural network as there isn't a need for upfront intensive annotation by a pathologist[35].



Phenotype	Inflamed	Excluded	Desert
N	3	16	5
Median	-0.612	0.304	0.319

Figure 5 Integrative approach suggests association between TGFβ pathway activity and T cell infiltration phenotypes. Higher TGFβ pathway activation score tends to be observed in tumors with an excluded or desert phenotype.

To further evaluate the association of TGFβ pathway in tumor-immune microenvironment, we performed transcriptomic analysis in the CRC biopsy samples and then correlated the TGFβ pathway activity with TIL level as well as the tumor-immune phenotypes. Our results suggest that 1) the levels of $CD8^+$ T cell infiltration negatively correlate with TGFβ pathway activation and 2) tumors presenting excluded or desert phenotypes have higher level of TGFβ pathway activity than the inflamed phenotype; which are all consistent with the previous reports on the pivotal role of TGFβ in immune suppression and cancer progression[12,24,36]. Additionally, the present study confirms the link between TGFβ-signaling and immunosuppressive tumor phenotypes, which provides supportive information on using them as enrichment biomarkers to select patients, with higher TGFβ activation or excluded/desert cancer-immune phenotypes, who may benefit from TGFβ blockade therapy[37–39].

Nevertheless, some limitations to the present study must be acknowledged. First, the machine-learning model for tumor-immune phenotype classification was established using a small number of archived tumor FFPE blocks involving five tumor types and would benefit from a further validation in a prospective clinical study in greater numbers of cases. Second, the cut-off value of the TGFβ activation gene signature and the correlation to the phenotypes established in the present study would benefit being applied in a larger cohort study allowing further optimization. Last, it remains a big challenge to reveal cancer-immune phenotypes due to complex spatial distribution of infiltrated $CD8^+$ T cells and the coexistence of various phenotypic and genotypic profiles caused by inter- and intra-tumoral heterogeneity. Thus, it is worth mentioning that further exploration is ongoing to validate the current integrative platform of digital pathology and transcriptomics in a large-scale analysis using dataset including procured and clinical tumor biopsies, which, we believe, may pave the road for implementing this approach to clinical study.

5. Conclusions

In this study, we established an image morphology operation-based automated digital pathology analysis approach that can characterize and classify the cancer immune phenotypes in an artificial intelligence fashion. In addition, the gene expression analysis revealed the association between TGFβ activation signature and immune phenotypes further demonstrates their potential as predictive biomarkers to identify appropriate patients that may benefit from TGFβ blockade.

Acknowledgments

We thank Jeff Lee, Jeff Weber and Brigitte Demers for their review and suggestions.

Author contributions

Rui Wang, Robert Pomponio, Michele Sanicola-Nadel, Souad Naimi, Reza Olfati-Saber, Cecile Combeau, Jack Pollard and Tun Tun Lin participated in the research design. Qi Tang, Anthony Mei, Anne Caron, Bema Coulibaly, Joachim Theilhaber and Maximilian Rogers-Grazado contribute to conducting experiments or performing the data analysis. Rui Wang, Robert Pomponio and Qi Tang wrote the manuscript.

Conflicts of interest

Authors are current or former employees of Sanofi (USA). This work was funded by Sanofi. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. Saltz J, Gupta R, Hou L, Kurc T, Singh P, Nguyen V, et al. Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Rep* 2018;**23**:181−193.e7.
2. Hiss S, Eckstein M, Segschneider P, Mantsopoulos K, Iro H, Hartmann A, et al. Tumour-infiltrating lymphocytes (TILs) and PD-L1 expression correlate with lymph node metastasis, high-grade transformation and shorter metastasis-free survival in patients with acinic cell carcinoma (AciCC) of the salivary glands. *Cancers* 2021;**13**:1−14.
3. Moore MR, Friesner ID, Rizk EM, Fullerton BT, Mondal M, Trager MH, et al. Automated digital TIL analysis (ADTA) adds prognostic value to standard assessment of depth and ulceration in primary melanoma. *Sci Rep* 2021;**11**:2809.
4. Hegde PS, Karanikas V, Evers S. The where, the when, and the how of immune monitoring for cancer immunotherapies in the era of checkpoint inhibition. *Clin Cancer Res* 2016;**22**:1865−74.
5. Klein S, Mauch C, Brinker K, Noh KW, Knez S, Büttner R, et al. Tumor infiltrating lymphocyte clusters are associated with response to immune checkpoint inhibition in BRAF V600 E/K mutated malignant melanomas. *Sci Rep* 2021;**11**:1834.
6. Echarti A, Hecht M, Büttner-Herold M, Haderlein M, Hartmann A, Fietkau R, et al. $CD8^+$ and regulatory T cells differentiate tumor immune phenotypes and predict survival in locally advanced head and neck cancer. *Cancers* 2019;**11**:1398.

7. Mlynska A, Vaišnorė R, Rafanavičius V, Jocys S, Janeiko J, Petrauskytė M, et al. A gene signature for immune subtyping of desert, excluded, and inflamed ovarian tumors. *Am J Reprod Immunol* 2020; **84**:e13244.
8. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* 2018;**554**:544−8.
9. Perren TJ, Swart AM, Pfisterer J, Ledermann JA, Pujade-Lauraine E, Kristensen G, et al. A phase 3 trial of bevacizumab in ovarian cancer. *N Engl J Med* 2011;**365**:2484−96.
10. Ganesh K, Massagué J. TGF-β inhibition and immunotherapy: checkmate. *Immunity* 2018;**48**:626−8.
11. Gordian E, Welsh EA, Gimbrone N, Siegel EM, Shibata D, Creelan BC, et al. Transforming growth factor β-induced epithelial-to-mesenchymal signature predicts metastasis-free survival in non-small cell lung cancer. *Oncotarget* 2019;**10**:810−24.
12. Desbois M, Udyavar AR, Ryner L, Kozlowski C, Guan Y, Dürrbaum M, et al. Integrated digital pathology and transcriptome analysis identifies molecular mediators of T-cell exclusion in ovarian cancer. *Nat Commun* 2020;**11**:5583.
13. Kather JN, Suarez-Carmona M, Charoentong P, Weis CA, Hirsch D, Bankhead P, et al. Topography of cancer-associated immune cells in human solid tumors. *Elife* 2018;**7**:e36967.
14. Lawrie CH, Ballabio E, Soilleux E, Sington J, Hatton CSR, Dirnhofer S, et al. Inter- and intra-observational variability in immunohistochemistry: a multicentre analysis of diffuse large B-cell lymphoma staining. *Histopathology* 2012;**61**:18−25.
15. Otsu N. A threshold selection method from gray-level histograms. *IEEE Transact Syst Man Cybernet* 1979;**9**:62−6.
16. Gonzalez RC, Woods RE. In: *Digital image processing*. 3rd ed. Upper Saddle River: Prentice Hall; 2008. p. 954.
17. Ganapathy V, Ge R, Grazioli A, Xie W, Banach-Petrosky W, Kang Y, et al. Targeting the transforming growth factor-β pathway inhibits human basal-like breast cancer metastasis. *Mol Cancer* 2010;**9**:122.
18. Padua D, Zhang XHF, Wang Q, Nadal C, Gerald WL, Gomis RR, et al. TGFβ primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell* 2008;**133**:66−77.
19. Kang Y, Siegel PM, Shu W, Drobnjak M, Kakonen SM, Cordón-Cardo C, et al. A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell* 2003;**3**:537−49.
20. Greco R, Qu H, Qu H, Theilhaber J, Shapiro G, Gregory R, et al. Pan-TGFβ inhibition by SAR439459 relieves immunosuppression and improves antitumor efficacy of PD-1 blockade. *OncoImmunology* 2020;**9**:1811605.
21. Bauer P. *Multiple comparisons—theory and methods*. London: Chapman & Hall; 1996xiv+277.
22. Obeid JM, Hu Y, Erdag G, Leick KM, Slingluff CL Jr. The heterogeneity of tumor-infiltrating CD8$^+$ T cells in metastatic melanoma distorts their quantification: how to manage heterogeneity?. *Melanoma Res* 2017;**27**:211−7.
23. Li F, Li C, Cai X, Xie Z, Zhou L, Cheng B, et al. The association between CD8$^+$ tumor-infiltrating lymphocytes and the clinical outcome of cancer immunotherapy: a systematic review and meta-analysis. *EClinicalMedicine* 2021;**41**:101134.
24. Yoo SY, Park HE, Kim JH, Wen X, Jeong S, Cho NY, et al. Whole-slide image analysis reveals quantitative landscape of tumor-immune microenvironment in colorectal cancers. *Clin Cancer Res* 2020;**26**: 870−81.
25. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The immune landscape of cancer. *Immunity* 2018;**48**:812−30. e14.
26. Newsted D, Banerjee S, Watt K, Nersesian S, Truesdell P, Blazer LL, et al. Blockade of TGF-β signaling with novel synthetic antibodies limits immune exclusion and improves chemotherapy response in metastatic ovarian cancer models. *OncoImmunology* 2019;**8**: e1539613.
27. Pagès F, Mlecnik B, Marliot F, Bindea G, Ou FS, Bifulco C, et al. International validation of the consensus immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet* 2018;**391**:2128−39.
28. Sun C, Zhang L, Zhang W, Liu Y, Chen B, Zhao S, et al. Expression of PD-1 and PD-L1 on tumor-infiltrating lymphocytes predicts prognosis in patients with small-cell lung cancer. *OncoTargets Ther* 2020;**13**: 6475−83.
29. Nguyen PHD, Ma S, Phua CZJ, Kaya NA, Lai HLH, Lim CJ, et al. Intratumoural immune heterogeneity as a hallmark of tumour evolution and progression in hepatocellular carcinoma. *Nat Commun* 2021; **12**:227.
30. Obeid JM, Wages NA, Hu Y, Deacon DH, Slingluff CL Jr. Heterogeneity of CD8$^+$ tumor-infiltrating lymphocytes in non-small-cell lung cancer: impact on patient prognostic assessments and comparison of quantification by different sampling strategies. *Cancer Immunol Immunother* 2017;**66**:33−43.
31. Mi H, Gong C, Sulam J, Fertig EJ, Szalay AS, Jaffee EM, et al. Digital pathology analysis quantifies spatial heterogeneity of CD3, CD4, CD8, CD20, and FoxP3 immune markers in triple-negative breast cancer. *Front Physiol* 2020;**11**:583333.
32. McCabe A, Dolled-Filhart M, Camp RL, Rimm DL. Automated quantitative analysis (AQUA) of *in situ* protein expression, antibody concentration, and prognosis. *J Natl Cancer Inst* 2005;**97**:1808−15.
33. Chung GG, Zerkowski MP, Ghosh S, Camp RL, Rimm DL. Quantitative analysis of estrogen receptor heterogeneity in breast cancer. *Lab Invest* 2007;**87**:662−9.
34. Camp RL, Chung GG, Rimm DL. Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nat Med* 2002;**8**:1323−7.
35. Abdolhoseini M, Kluge MG, Walker FR, Johnson SJ. Segmentation of heavily clustered nuclei from histopathological images. *Sci Rep* 2019; **9**:4551.
36. Chen YP, Lv JW, Mao YP, Li XM, Li JY, Wang YQ, et al. Unraveling tumour microenvironment heterogeneity in nasopharyngeal carcinoma identifies biologically distinct immune subtypes predicting prognosis and immunotherapy responses. *Mol Cancer* 2021;**20**:14.
37. Pai SI, Cesano A, Marincola FM. The paradox of cancer immune exclusion: immune oncology next frontier. *Cancer Treat Res* 2020; **180**:173−95.
38. Groeneveldt C, van Hall T, van der Burg SH, ten Dijke P, van Montfoort N. Immunotherapeutic potential of TGF-β inhibition and oncolytic viruses. *Trends Immunol* 2020;**41**:406−20.
39. Majidpoor J, Mortezaee K. The efficacy of PD-1/PD-L1 blockade in cold cancers and future perspectives. *Clin Immunol* 2021;**226**: 108707.