White Paper

Development and Verification of the HTG EdgeSeq™ Reveal Immune, Stroma and TME Signatures

Introduction

The Tumor Microenvironment (TME) is a complex and dynamic environment, which plays a role in the body's response to growth and metastasis of tumors[1,2]. The TME is composed of malignant, proliferating tumor cells in addition to numerous distinct non-cancerous cell types. These additional cell types include a variety of innate, stromal, and immune cells all of which interact with the tumor to varying degrees. Mounting evidence suggests that understanding how the body responds to a malignant tumor, in the form of an immune response, is a known indicator in predicting patient outcome to therapy[3]. Interestingly, patient response to tumor inflammation varies by tumor indication. In some indications, an increased immune response is associated with improved progression-free survival (PFS) while in other indications, no association is observed between immune characteristics and PFS[4]. This highlights the need to accurately classify the inflammation status of a tumor.

Historically, tumor immune response has been characterized using immunohistochemistry (IHC) to determine the abundance of T-cells. With the advent of targeted therapies, and immunotherapeutic agents, our understanding of tumor immune response has grown. This has led to an increased evaluation of the immune response, which encompasses additional cell types in the tumor microenvironment. IHC has limited utility in the characterization of the TME, as this can include analysis of various immune and stromal cell types, and often a single IHC stained slide is required to analyze each cell type. Gene Expression Profiling (GEP), however, offers a powerful new tool that utilizes the power of Next-Generation Sequencing (NGS) to identify gene expression from hundreds to thousands of genes from significantly less tissue.

The purpose of this white paper is to describe the development and verification of three new HTG EdgeSeq Reveal RUO Oncology Signatures, built upon the HTG EdgeSeq RUO Precision Immuno-Oncology Panel (PIP). The three signatures are the HTG EdgeSeq Reveal Immune Signature (Immune), the HTG EdgeSeq Reveal Stroma Signature (Stroma) and the HTG EdgeSeq Reveal Tumor Microenvironment (TME) signature. The data presented here, demonstrate that the HTG EdgeSeq Immune, Stroma, and TME scores offer an accurate and robust method for determining the inflammation and stroma status of tumor FFPE samples.

Methods

Workflow Overview

Signature development and primary verification utilized data from 1,073 colorectal, gastric, and ovarian cancer formalin-fixed paraffin embedded (FFPE) tissue samples. First, a reference signature was generated using the xCell algorithm[5] to establish reference immune, stroma and TME scores for each sample. The xCell signature is a peer-reviewed GEP-based method for identifying the relative abundance of 64 individual immune and stromal cell types. The algorithm uses the abundance of several of these cell types to calculate an Immune, Stroma and TME score. HTG EdgeSeq PIP data were then generated for the same set of samples and randomly split into training and test sets. The HTG EdgeSeq Immune, Stroma and TME signatures were developed using the training set (805 samples) and performance was verified to the xCell reference using the test set (268 samples). Lastly, an orthogonal comparison was made between the HTG EdgeSeq Reveal Immune Signature and five IHC stains, which were applied to lung cancer FFPE tissue samples. The use of these additional samples not only establishes correlation with IHC, it also shows that the signatures perform well and can be successfully applied across indications.

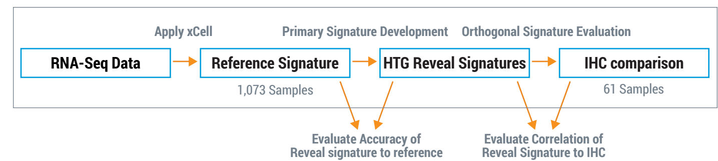


Figure 1. *Overall signature development workflow included primary signature development with orthogonal method comparison to ensure the accuracy of each signature.*

Samples

For signature development and primary verification, a total of 1,200 FFPE samples were tested on both the HTG EdgeSeq PIP assay and Illumina TruSeq RNA Access (RNA-Seq). Data generation for the HTG EdgeSeq PIP assay was performed at the HTG VERI/O commercial laboratory in Tucson, AZ and RNA-Seq data were generated at Q2 Solutions, Morrisville, NC. FFPE tissue samples collected from patients with colorectal (CRC), gastric (GC), or ovarian cancer (OVC) were commercially procured. After removal of samples due to quality control failures, the effective sample size was 1,073. **Table 1** summarizes key demographic variables of the samples used in the study.

Table 1. *Summary statistics of key demographic variables by indication.*

	Colorectal N=402	Gastric N=398	Ovarian N=400
Sex, Female (%)	181 (45.0)	156 (39.2)	400 (100.0)
Age, Mean (SD)	65.3 (12.0)	60.5 (11.5)	59.7 (12.2)
20<Age<40, n (%)	16 (4.0)	7 (1.8)	19 (4.8)
40≤Age<60, n (%)	87 (21.6)	183 (46.0)	182 (45.5)
60≤Age<80, n (%)	259 (64.4)	189 (47.5)	179 (44.8)
80≤Age<100, n (%)	40 (10.0)	18 (4.5)	20 (5.0)
Missing, n (%)	0 (0)	1 (0.3)	0 (0)

Sixty-one non-small cell lung cancer (NSCLC) FFPE samples, including 40 adenocarcinomas and 21 squamous cell carcinomas were included in the orthogonal signature verification. These samples were run on the HTG EdgeSeq Precision Immuno-Oncology Panel, passed the quality control specifications for the assay, and were stained for CD3, CD4, CD8, CD68 and PD-1 via IHC.

Signature Development and Verification

Reference Signatures

Reference immune, stroma and TME signatures were generated using RNA-Seq data analyzed with a well-documented gene signature-based method. The xCell algorithm uses an equal-weighted average of 10 immune cell types, including B-cells, CD4 and CD8-positive T-cells, dendritic cells, eosinophils, macrophages, monocytes, mast cells, neutrophils and natural killer cells to produce the Immune score and an average of three cell types including adipocytes, endothelial cells and fibroblasts to produce the Stroma score. The TME score is the summation of immune and stroma scores.

Development and Primary Verification

The development of the HTG EdgeSeq Oncology Signature algorithm was conducted in three steps: 1) separation of samples into the training and test sets; 2) HTG EdgeSeq signature predictive model fitting on the training set; and 3) HTG EdgeSeq signature predictive model verification on the test set. The Oncology Signature prediction models were trained for each signature as a linear Least Absolute Shrinkage and Selection Operator (LASSO) model. The LASSO model is a regularized regression method that incorporates variable selection through a shrinkage step of the parameter estimation, wherein the coefficients (or weights) for “unimportant” predictors preferentially are shrunk to zero. The optimal model was developed after comparison of over 1000 bootstrap models applied to the linear LASSO model to select informative probes and estimate coefficients. The trained linear prediction models were evaluated on the testing sets by assessing Lin’s concordance correlation coefficients (Lc) between the predicted and reference signature scores.

The predictive accuracy, i.e., the agreement measured by Lc between the predicted and reference immune scores is shown in **Figure 2** for the training and testing sets separately. High Lc values for all three signatures indicated the LASSO model had good agreement in signature scores for both training and testing sets. In addition to Lc, the estimated intercept and slope from Passing-Bablok (PB) regression are also shown in **Figure 2**. For all training and testing sets, the estimated PB intercepts were close to 0 and PB slopes were close to 1, suggesting there were no constant or proportional biases between the predicted and reference immune scores. Together these data show a high degree of agreement between the reference scores generated using the xCell algorithm and the HTG EdgeSeq Reveal Oncology Signature model.

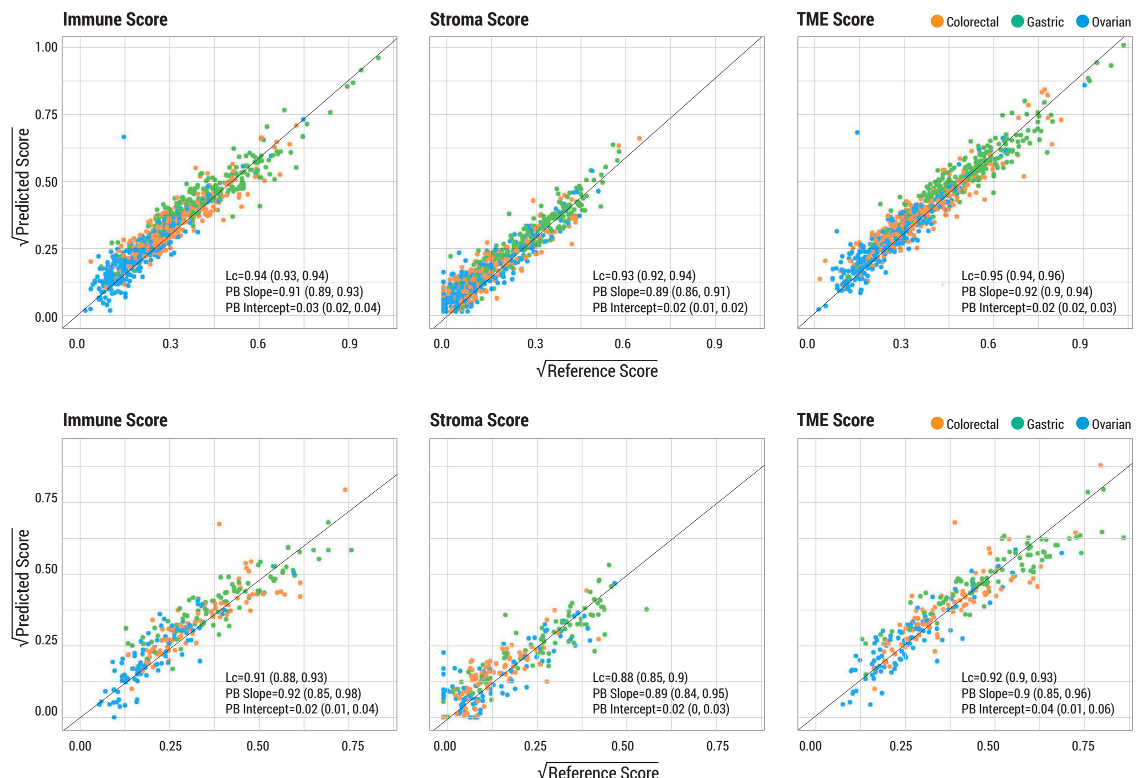


Figure 2. Agreement between Predicted and Reference Immune Score for LASSO Models. *Data for the (A) training and (B) test data comparison to the corresponding reference xCell signature for all three signatures. The Y axis shows the square root of the score predicted by the HTG EdgeSeq Oncology Signature and the x-axis shows the square root of the value of the reference xCell score for Immune, Stroma and TME.*

The reference scores ranged from 0 to 1.07 for the immune score, 0 to 0.45 for the stroma score, and 0 to 1.16 for the TME score. The reference scores generated from the test data were compared to available data from The Cancer Genome Atlas (TCGA), which contains data from more than 9,000 samples covering 33 cancer types, to evaluate if the dynamic range of the evaluated sample set adequately covers the range of scores expected in a variety of cancer specimens (**Figure 3**). The HTG EdgeSeq data, obtained with the 1,073 FFPE samples used to develop and test the Reference Signatures (see **Figure 3**), were in-line with the data obtained from TGCA showing that the samples used to develop the HTG EdgeSeq Reveal Signatures adequately cover the range of scores seen in the TCGA data set.

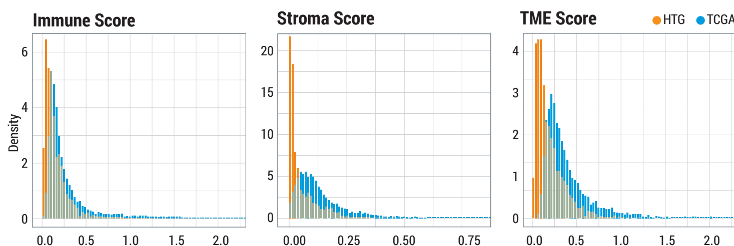


Figure 3. *Comparison of the overall distribution of the immune, stroma, and TME scores generated using data from TCGA RNA-Seq data (blue) and HTG EdgeSeq PIP gene expression data (orange).*

Orthogonal Method Comparison

To ensure that the three signatures are reflective of the tumor biology, an orthogonal method comparison was performed. Among other methodologies, IHC has long been considered the 'gold standard' for determining tumor inflammation. Therefore, the orthogonal verification method evaluated the output of the HTG EdgeSeq Reveal Immune Signature score against five well-known biomarkers. It is important to evaluate these two methods as they both aim to measure tumor inflammation; however, perfect correlations are not expected due to the differences in technology and what each platform measures (RNA or protein). HTG EdgeSeq PIP is measuring the RNA expression in multiple cell types to generate the Immune score, whereas IHC is only measuring abundance of a single immune cell marker. IHC was performed for CD3, CD4, CD8, CD68, and PD-1 antibodies on 61 non-small cell lung carcinoma FFPE samples (see **Table 2**). The number of cells staining positive for each of the immune cell markers in the FFPE samples was counted across five areas that were selected and screened by a pathologist to confirm the presence of viable tumor cells, excluding necrosis. The average number of positive immune cells for each of the markers across the five areas was tabulated and labeled as the tumor area immune cell count (TAIC).

Table 2. *Antibody information for IHC Comparison*

Biomarker	Antibody	Marker for
CD3	Dako, Clone#: F7.2.38	T-cells
CD4	Novocastra, Clone #: 4B12	Helper T-cells
CD8	Thermo Scientific, Clone # C8/144B	Cytotoxic T-cells
CD68	Dako, Clone # PG-M1	Macrophages and dendritic cells
PD-1	Abcam, Clone # EPR4877	Activated T-cells

Serial sections of the same FFPE samples that were evaluated for IHC staining were also run on the HTG EdgeSeq PIP. The HTG EdgeSeq Reveal Immune scores were then computed for the 61 samples that passed QC metrics for both methods. Only the Immune scores were measured as all the IHC biomarkers tested here are specific to immune cells. **Figure 4** (top panel) visualizes the correlation between the immune signature scores and IHC measures of CD3, CD4, CD8, CD68, and PD-1 along with the Spearman rank correlation coefficient. The immune signature scores had moderate to good correlation with markers targeting T-cells, such as CD3, CD4, CD8, and PD-1 (Spearman correlations are 0.7, 0.62, 0.71, and 0.52, respectively). The Reveal Immune score had a relatively lower correlation with CD68, a marker for macrophages and dendritic cells. To better understand the difference in correlation to CD68, the distribution of the IHC tumor area immune cell count was plotted (**Figure 4**, bottom panel). The range of TAIC densities for all the immune cells ranged from 0 to nearly 5000 positive immune cells / mm^2. The CD3, CD4, and CD8 positive T-cells displayed the widest range of counts across the samples and were highly expressed, while the CD68 positive cells were the least represented. The narrow range of IHC expression for CD68 likely contributed to the lower correlation. Additionally, the Reveal Immune Signature was designed to look at the whole tumor immune environment, with markers for ten different immune cell types, and thus is not specific for each independent marker. These data demonstrate that as the density of these immune cell markers, measured by IHC increases, so does the immune score, indicating that it is reflective of the inherent tumor biology/inflammation. In addition to establishing correlation to IHC, these data also show that the signatures perform well in lung tumor samples and can be successfully applied across indications.

Taken together, the data from the primary and orthogonal verification studies show that the HTG EdgeSeq Reveal Immune, Stroma and TME signatures exhibit a high degree of accuracy to the xCell reference scores and are reflective of inherent tumor biology/ inflammation.)

Signature Robustness

The intent of the HTG EdgeSeq Reveal Oncology Signatures is to generate a score based on the relative abundance of immune and stromal cell types present in the tissue sample. Therefore, it is critical that the measurements, or scores, generated for the Signatures are robust to fluctuations in both sample input and macrodissection. That is, the signature must generate similar scores for a sample whether it is run at the recommended sample input of 6.0 mm^2 / well or whether it is run at half or double that concentration. Similarly, a sample that is macrodissected to include the tumor only, tumor plus invasive margin, or the entire FFPE tissue should generate a similar score.

To ensure the Signatures described here are robust to variations in sample input, a range of sample inputs were evaluated for precision and compared to the recommended sample input of 6.0 mm^2 / well. This assessment was made using FFPE tissue samples from melanoma (Mel), non-small cell lung cancer (NSCLC), and squamous cell carcinoma of the head and neck (SCCHN). The signature scores for each input were compared to the scores at the standard sample input for HTG EdgeSeq PIP (6.0 mm2 / well), which is also the input at which the signatures were developed. The difference between the standard sample input compared to the average scores for each additional sample input (1.5, 3, 12 and 24 mm^2 / well) is shown in **Figure 5**. These data demonstrate that there are no significant differences in the average Immune, Stroma or TME scores across all sample inputs compared to the recommended sample input and suggests that the signatures are robust to variability in sample input between 1.5 and 24.0 mm^2 per well. In addition, these data also add to the number of indications tested, showing that the signatures can be successfully applied across indications.

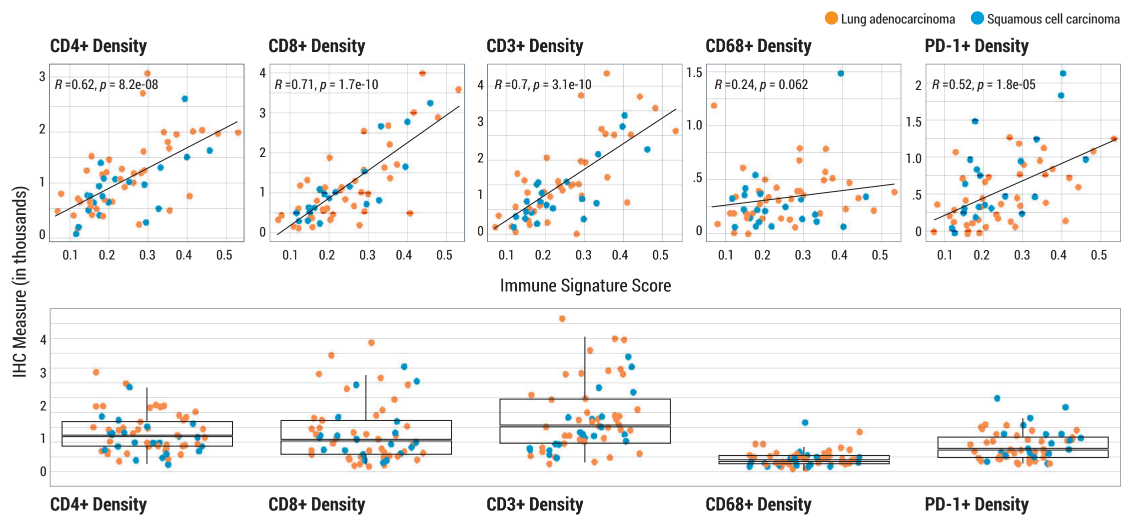


Figure 4. Correlation between Immune Signature and IHC Cell Counts. The top plots depict the correlation between the immune signature score (x-axis), and the tumor immune cell counts of CD3, CD4, CD8, CD68, and PD-1 (y-axis). The gray lines represent the linear fitted line. The Spearman correlation coefficient along with P-value are included in the left upper corner. The bottom panel illustrates the distribution of the IHC tumor area immune cell count (y-axis) for the different immune cell markers (x-axis). Each dot represents one of 61 NSCLC samples, which are color-coded by histological subtype (orange: Lung adenocarcinoma and blue: Lung squamous cell carcinoma).

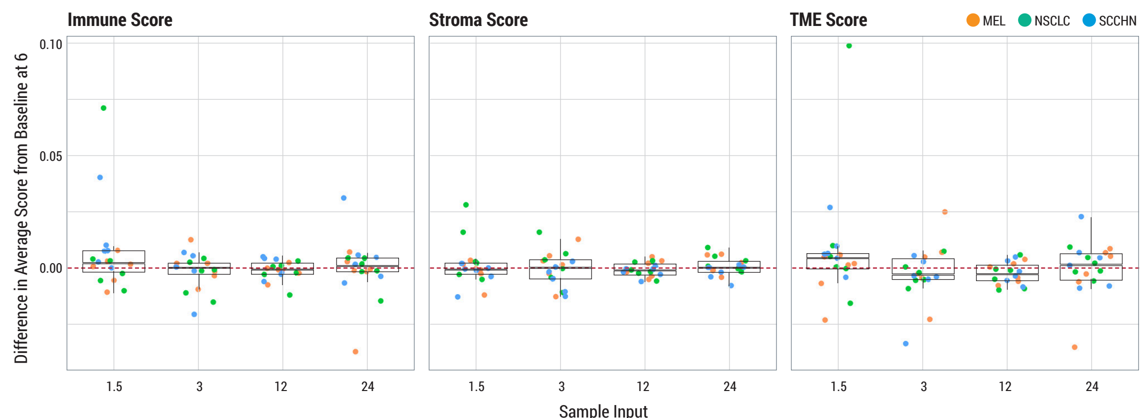


Figure 5. Accuracy of the HTG EdgeSeq Reveal RUO Oncology Signatures in relation to sample input. The x-axis lists the sample inputs tested (1.5–24 mm^2 / well); the y-axis illustrates how different the scores are from the average scores generated at 6.0 mm^2 / well. Sample indications are color coded, orange = melanoma, green = non-small cell lung cancer, and blue = squamous cell carcinoma of the head and neck.

Extensive literature analysis suggests that immune cells are known to stack up on the border, or invasive margin, of a tumor. Therefore, it is critical to understand robustness of the signatures to variations in sample macrodissection. To evaluate variability in macrodissection, three conditions were tested: tumor only, tumor plus invasive margin or tumor plus invasive margin plus adjacent tissue. First, samples were macrodissected to enrich for just the malignant cells of the tumor (gray in **Figure 6**; "macrodissected"), the standard recommendation for HTG EdgeSeq PIP sample processing. Second, the slide was macrodissected to include the invasive margin tissue along with the tumor (gray + orange in **Figure 6**; "invasive margin"). Third, all tissue on the slide was utilized, including the tumor, invasive margin, and any adjacent tissue present (gray + orange + blue in **Figure 6**; "non-macrodissected").

The performance of the Signatures was assessed by calculating Lc and the Passing-Bablok intercept and slope for each macrodissection approach and comparing them to one another. A total of 10 FFPE samples, including melanoma, ovarian, prostate, gastric, breast, colorectal and non-small cell lung cancer were tested for each of the three conditions and **Figure 7** summarizes the results. The Lc values ranged between 0.90 and 0.99 for the Immune, Stroma, and TME scores across the different macrodissection approaches. The PB intercepts were close to 0, and PB slopes were close to 1. While the non-macrodissected samples generated slightly greater stroma scores, the difference was not statistically significant. This is likely due to the increase of stromal cells inherent in the non-macrodissected samples. Collectively, these findings demonstrate that the Signature scores are robust to variations in sample macrodissection.

Taken together, the data presented here shows that the HTG EdgeSeq Reveal Immune, Stroma and TME scores are robust to variations in sample input and macrodissection.

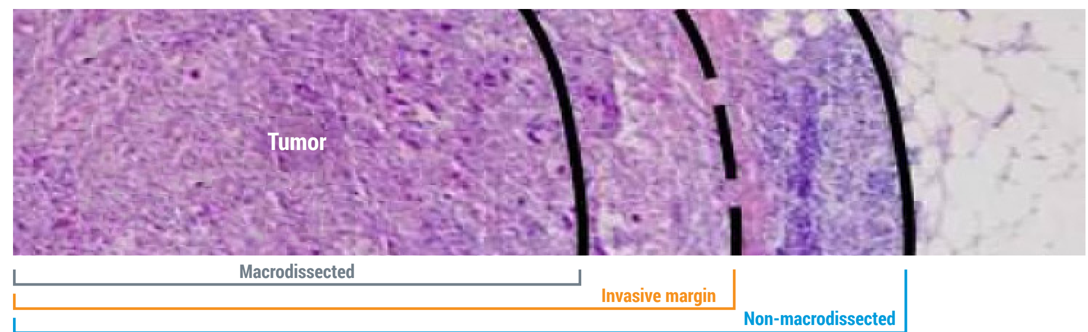


Figure 6. Three sample inputs for evaluation of Signature robustness to macrodissection. *Three conditions, tumor only, tumor plus invasive margin and tumor plus invasive margin and adjacent tissue were tested. Gray is tumor; orange is invasive margin and blue is adjacent tissue.*

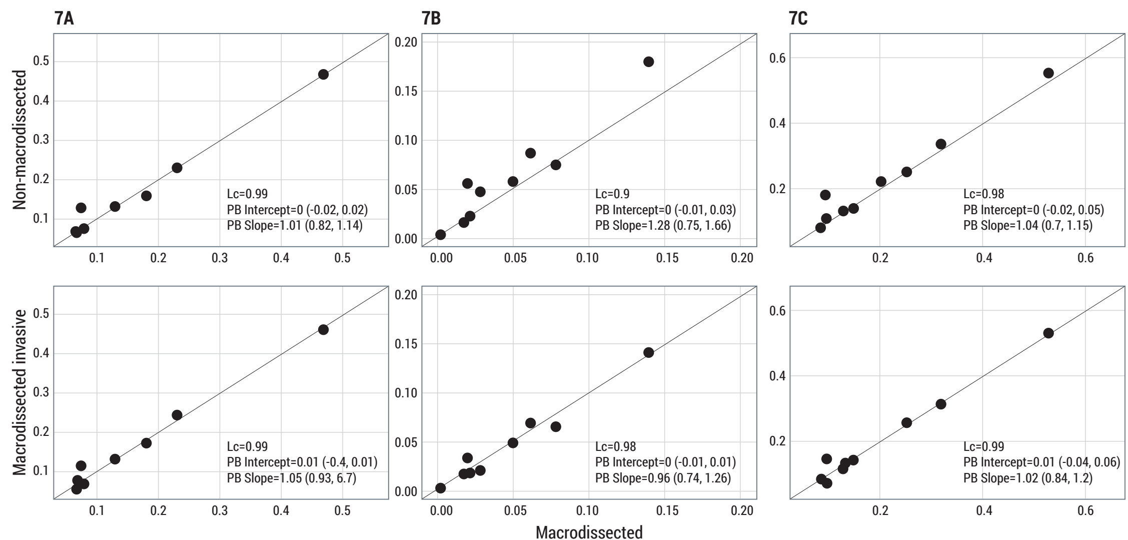


Figure 7. Impact to the HTG EdgeSeq Reveal RUO Oncology Signatures in relation to macrodissection. *Tissue samples were macrodissected with tumor tissue only (macrodissected), tumor tissue plus invasive margin (macrodissected invasive) or not macrodissected (non-macrodissected). Axes are marked as either macrodissected, invasive margin or non-macrodissected as defined in **Figure 6**. Each dot represents one of 10 different FFPE samples. LC and PB slope and intercept values (including their respective 95% confidence intervals) are shown for (A) Immune, (B) Stroma and (C) TME scores.*

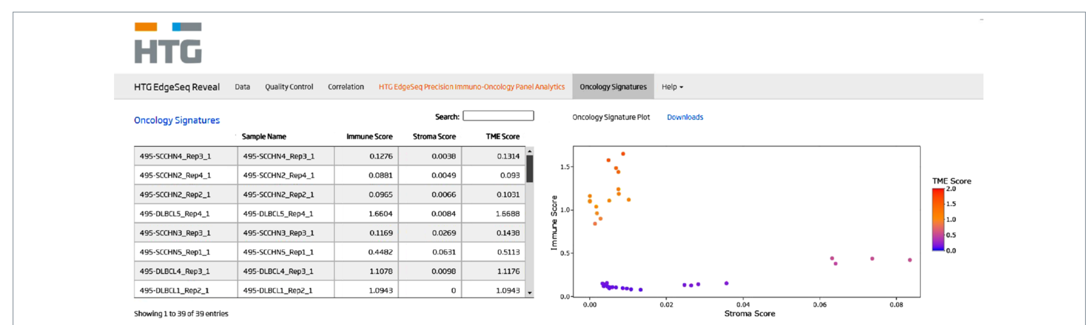
Signature Implementation and Data Interpretation

The Immune, Stroma and TME Signatures are applied to HTG EdgeSeq PIP data using the HTG EdgeSeq Reveal software. The HTG EdgeSeq Reveal software is a fully integrated web-based data analysis software suite that can analyze data quality, generate publication quality figures, and now apply the HTG EdgeSeq Reveal Oncology Signatures to generate Immune, Stroma, and TME scores. Any HTG EdgeSeq PIP dataset that is loaded into the Reveal software will prompt a new tab to appear, titled 'Oncology Signatures' (**Figure 8**). This applies to new data generated after release of the new Signature as well as to all previously generated HTG EdgeSeq PIP data. Clicking on this tab will generate an interactive table and chart to help visualize the Immune, Stroma and TME scores. Data can also be exported in Excel format to facilitate further data analysis by researchers.

Summary

This white paper compared the performance of the HTG EdgeSeq Reveal Immune, Stroma and TME Signatures to the xCell signatures, an established, peer-reviewed gene signature method based on the presence of 64 individual cell types commonly found in the TME. Development and verification of the HTG EdgeSeq Reveal Signatures used approximately 1,100 FFPE tumor tissue samples, run on both RNA-Seq and HTG EdgeSeq platforms. Primary Signature verification showed good accuracy, as measured by Lc, in predicting signature scores (0.88 - 0.92). In addition, Passing-Bablok (PB) regression did not indicate the presence of constant or proportional biases between the predicted and reference immune scores. Orthogonal verification evaluated the correlation between the immune signature scores and IHC measures of CD3, CD4, CD8, CD68, and PD-1 measured as the Spearman rank correlation coefficient. The immune signature scores had moderate to good correlation with markers targeting T-cells, such as CD3, CD4, CD8, and PD-1 (Spearman correlation were 0.7, 0.62, 0.71, and 0.52, respectively). The Reveal Immune score had a relatively lower correlation with CD68, a marker for macrophages and dendritic cells, likely caused by the low abundance and narrow range of IHC expression. Robustness studies showed that there were no significant differences in the average Immune, Stroma or TME scores across sample inputs between 1.5 and 24.0 mm^2 / well, and there were no significant differences when different macrodissection techniques were employed. The use of numerous cancer types, nine in total, shows that the signatures performed well across different tumor indications and can be successfully applied to indications other than gastric, ovarian and CRC.

The data presented here show that the HTG EdgeSeq Reveal Immune, Stroma and TME Signatures are accurate and robust, allowing researchers to accurately determine the inflammation and stroma status of tumor FFPE samples.



	Sample Name	Immune Score	Stroma Score	TME Score
495-SCCHN4_Rep3_1	495-SCCHN4_Rep3_1	0.1276	0.0038	0.1314
495-SCCHN2_Rep4_1	495-SCCHN2_Rep4_1	0.0881	0.0049	0.093
495-SCCHN2_Rep2_1	495-SCCHN2_Rep2_1	0.0965	0.0066	0.1031
495-DLBCL5_Rep4_1	495-DLBCL5_Rep4_1	1.6604	0.0084	1.6688
495-SCCHN3_Rep3_1	495-SCCHN3_Rep3_1	0.1169	0.0269	0.1438
495-SCCHN5_Rep1_1	495-SCCHN5_Rep1_1	0.4482	0.0631	0.5113
495-DLBCL4_Rep3_1	495-DLBCL4_Rep3_1	1.1078	0.0098	1.1176
495-DLBCL1_Rep2_1	495-DLBCL1_Rep2_1	1.0943	0	1.0943

Figure 8: Example visual output from the HTG EdgeSeq Reveal Immune, Stroma and TME Signatures. *The HTG EdgeSeq Reveal software generates scores for all three Signatures and displays them in an interactive table and chart to help visualize the scores using HTG EdgeSeq PIP data.*

References

1: Galon J, Costes A, Sanchez-Cabo F, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science. 2006;313(5795):1960-1964.

2: Hanahan D, Coussens LM. Accessories to the crime: functions of cells recruited to the tumor microenvironment. Cancer Cell. 2012;21(3):309-322.

3: Roma-Rodrigues C, Mendes R, Baptista PV, Fernandes AR. Targeting Tumor Microenvironment for Cancer Therapy. Int J Mol Sci. 2019;20(4):840.

4: Öjlert ÅK, Halvorsen AR, Nebdal D, et al. The immune microenvironment in non-small cell lung cancer is predictive of prognosis after surgery. Mol Oncol. 2019;13(5):1166-1179.

5: Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. 2017;18(1):220.

For Research Use Only. Not for use in diagnostic procedures.
All trademarks are the property of HTG Molecular and its subsidiaries unless otherwise specified.