



Evaluation of the EdgeSeq Precision Immuno-Oncology Panel for Gene Expression Profiling From Clinical Formalin-Fixed Paraffin-Embedded Tumor Specimens

Yang Shi, Xiaopeng Ma, Wei Shen, Tengfei Liu, Liang Liang, Silu Liu, Zhirong Shen, Yun Zhang* and Pei Zhang*

BeiGene (Beijing) Co., Ltd., Beijing, China

OPEN ACCESS

Edited by:

Fangqing Zhao,
Beijing Institutes of Life Science (CAS),
China

Reviewed by:

Bo Wang,
Sun Yat-sen Memorial Hospital, China
Alexander Deutsch,
Medical University of Graz, Austria

*Correspondence:

Yun Zhang
yun.zhang@beigene.com
Pei Zhang
pei.zhang@beigene.com

Specialty section:

This article was submitted to
Cancer Cell Biology,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 18 March 2022

Accepted: 28 April 2022

Published: 27 May 2022

Citation:

Shi Y, Ma X, Shen W, Liu T, Liang L,
Liu S, Shen Z, Zhang Y and Zhang P
(2022) Evaluation of the EdgeSeq
Precision Immuno-Oncology Panel for
Gene Expression Profiling From Clinical
Formalin-Fixed Paraffin-Embedded
Tumor Specimens.
Front. Cell Dev. Biol. 10:899353.
doi: 10.3389/fcell.2022.899353

Characterizing the tumor microenvironment (TME) of archived clinical tissues requires reliable gene expression profiling (GEP) of formalin-fixed paraffin-embedded (FFPE) samples. The EdgeSeq Precision Immuno-oncology Panel (PIP) is a targeted GEP assay designed for TME characterization but lacks widespread technical validation on a large cohort of clinical samples. Here, we evaluated its performance by exploring its concordance with multiple orthogonal platforms using 1,220 FFPE samples across various cancer types. Quantitative comparisons with RNA-seq and NanoString showed strong correlations at the sample level (median $\rho = 0.73$ and 0.81) and moderate correlations at the single-gene level (median $\rho = 0.49$ and 0.57). Gene signature analysis revealed high concordance with RNA-seq on widely used signatures for TME characterization and immune checkpoint inhibitor (ICI) efficacy prediction, though some genes in these signatures are not targeted by EdgeSeq PIP. From a histopathological viewpoint, the tumor/immune abundances derived from hematoxylin and eosin (H & E) staining were well recapitulated by the transcriptomic profiles assessed by EdgeSeq PIP. Furthermore, the mRNA level of PD-L1 assessed by EdgeSeq PIP was moderately correlated with the PD-L1 score ($\rho = 0.65$) estimated by immunohistochemistry (IHC); the mRNA level of CD8A aligned well ($\rho = 0.55$) with the IHC-derived abundance of CD8⁺ T cells. Overall, our results showed that EdgeSeq PIP generated well-correlated data with independent approaches at mRNA, protein, and histological levels, thus providing strong technical support for further using EdgeSeq PIP in biomarker studies and companion diagnostic (CDx) development.

Keywords: gene expression profile, tumor microenvironment, immunotherapy, platform evaluation, edgeseq PIP

INTRODUCTION

Over the past years, immune checkpoint inhibitor (ICI)-based therapies have transformed the treatment landscape of cancer (Sanmamed and Chen, 2018). ICIs such as humanized monoclonal antibodies against cytotoxic T lymphocyte antigen 4 (CTLA4), programmed cell death protein 1 (PD-1) and programmed death ligand 1 (PD-L1) have demonstrated impressive efficacy and have

been approved as first-line or second-line therapies for an ever-growing list of malignancies (Vaddepally et al., 2020). However, only a small fraction of patients benefits from ICI-based therapy, and there are urgent needs to identify the mechanisms driving response or resistance and develop new biomarkers to guide personalized therapy (Havel et al., 2019). Unraveling the landscape of immune cell subpopulations in the tumor microenvironment (TME) and investigating their interactions with tumor and stromal cells is a critical step in this process (Chen and Mellman, 2017).

TME characterization heavily relies on precise and comprehensive gene expression profiling (GEP). Among multiple GEP methods, RNA-seq has been established as the gold standard and is widely used in cancer research (Wang et al., 2009), such as The Cancer Genome Atlas (TCGA). However, in clinical practice, RNA-seq (and derived methods, such as RNA Exome) has several limitations. On the one hand, it requires large amounts of high-quality samples, which is usually not feasible in clinical trials in which most samples are formalin-fixed and paraffin-embedded (FFPE) and exhibit moderate-to-severe degradation (Evers et al., 2011). This requirement also contradicts the scarcity of tumor tissues in most clinical trials, especially when other typical biomarker assays, such as tumor mutation burden (TMB) and immunohistochemistry (IHC) assays are in competition (Aisner et al., 2016). On the other hand, RNA-seq is untargeted and thus not cost-effective for the development of companion diagnosis assay, as only a small proportion of genes are of strong interest (Thorsson et al., 2018).

Given these limitations of RNA-seq and the increasing demands for profiling immuno-oncology-related genes in clinical trials, targeted assays such as EdgeSeq PIP (HTG Molecular, Tucson, AZ), PanCancer IO 360 (NanoString, Seattle, WA) and Oncomine IRRA (Thermo Fisher Scientific, Waltham, MA), have emerged in recent years. These assays significantly reduced requirements for sample quality and quantity, thus mitigating the challenges of sample acquisition and making them clinically feasible. EdgeSeq PIP, which focuses on ~1,300 key genes related to immuno-oncology, has the highest coverage among all these assays. By utilizing quantitative nuclease protection chemistry, EdgeSeq PIP quantifies RNAs *via* an extraction-free approach, which eliminates the risk of extraction bias induced by the removal of short or fragmented RNAs (Martel et al., 2002; Qi et al., 2016). The data generated from FFPE samples has good concordance with that from fresh-frozen samples. Moreover, the extraction-free approach circumvents the loss of RNA from sample during extraction; thus, less tissue input is required to generate an equivalent amount of RNA. Last, EdgeSeq PIP has the unique advantage of utilizing samples previously subjected to hematoxylin and eosin (H & E) or IHC staining, which further expands its clinical utility when samples are extremely limited (Qi et al., 2019). Although successfully used in several clinical trials (Wang et al., 2018; Desai et al., 2020; Garg et al., 2020; Martin-Broto et al., 2020; Song et al., 2021), the reliability of EdgeSeq PIP in real clinical settings and its concordance with other platforms have not been well studied. Previous studies either compared only EdgeSeq PIP with RNA-

seq at the single-gene level or were limited by a small number of samples to draw solid conclusions (Anguiano et al., 2020; Ran et al., 2020). In this study, using 1,220 FFPE samples across several cancer types from clinical trials, we performed comprehensive comparisons of EdgeSeq PIP and multiple platforms and confirmed its fidelity at the RNA, protein, and histological levels.

MATERIALS AND METHODS

Patient Cohort

Baseline FFPE samples were collected from the following seven clinical studies on tislelizumab monotherapy or tislelizumab combined with chemotherapy or anti-PD-L1 therapy: A317-001 (NCT02407990), A317-102 (NCT04068519), 900-101 (NCT03379259), A317-204 (NCT04004221), A317-205 (NCT03469557), A317-206 (NCT03432598), RATIONALE 307 (NCT03594747), and RATIONALE 309 (NCT03924986). The major cancer types in each study were summarized in **Supplementary Table S1**.

Sample Preparation, Library Construction and Sequencing

All archived FFPE blocks were prepared as previously described. After confirming the presence of malignant cells by histological H & E staining, the samples were processed *via* standardized procedures for biomarker investigation. In general, a tumor content of at least 20% and sufficient tumor area (>20 mm² for RNA-seq and >2.5 mm² for EdgeSeq PIP) were required for sample inclusion.

For EdgeSeq PIP, sample processing, library construction and sequencing were performed in accordance with OP-00034, OP-00035, OP-00079 (HTG EdgeSeq instrument method). Briefly, tissues were scraped and lysed using lysis buffer from HTG Molecular Diagnostics. Next, Proteinase K was added to digest proteins and remove potential contaminations. Gene-specific nuclease protection probes were then added to the lysed samples to form the probe-target RNA heteroduplexes, after which S1 nuclease was added to degrade non-hybridized molecules. Then samples were individually barcoded using a 16-cycle PCR, purified using Agencourt AMPure XP beads and loaded into Illumina MiSeq (Illumina, San Diego, CA) for 50 bp single-end sequencing.

For RNA-seq, RNA was extracted using the AllPrep DNA/RNA FFPE Kit (QIAGEN, Hilden, Germany). The amount of RNA was quantitated by the fluorescence method using Qubit RNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA), and the quality was assessed by Agilent 2,100 Bioanalyzer System (Agilent Technologies, Santa Clara, CA). Only samples with RNA >40 ng and DV200 >20% were included for downstream steps. rRNA depletion, cDNA synthesis and NGS library preparation were performed using the TruSeq RNA Exome (Illumina, San Diego, CA). The libraries were then loaded into HiSeq X Ten instrument (Illumina, San Diego, CA) for 150 bp paired-end sequencing.

For NanoString, all processing steps were performed according to the manufacturer's instructions. In brief, RNA was extracted using RNeasy Mini Kit (QIAGEN, Hilden, Germany). The amount of RNA was quantitated by the fluorescence method using Qubit RNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA), and the quality was assessed by Agilent 2,100 Bioanalyzer System (Agilent Technologies, Santa Clara, CA). After hybridization with the PanCancer IO 360 Panel, sample analysis was performed on a nCounter Digital Analyzer (NanoString Technologies, Seattle, WA).

EdgeSeq Precision Immuno-Oncology Panel Data Processing

Demultiplexed FASTQ files from the Illumina MiSeq were parsed by the EdgeSeq parser (HTG Molecular Diagnostics, Inc.). Three post-sequencing quality control metrics were derived and used for filtering samples: 1) >15% percentage of reads allocated to the positive process control probe; 2) >1.5 million reads; and 3) a relative standard deviation (RSD) of reads allocated to each probe >0.10. After removing samples not meeting these QC requirements, the raw read counts for each sample were transformed to the log₂ counts per million (CPM) scale.

RNA-Seq Data Processing

Adapters in raw reads were detected and trimmed by Trimmomatic (0.36) (Bolger et al., 2014). After trimming, reads shorter than 50 bp were removed. These reads were then mapped to the human genome (GRCh38) using STAR (2.7.10a) (Dobin et al., 2013). Gene expression was quantified using the RSEM workflow (1.3.3) with default parameters (Li and Dewey, 2011). Only samples with >1.5 million reads and >80% reads confidently mapped to the transcriptome were retained. Then, transcripts per million (TPM) values were log₂-transformed for downstream analysis.

Differential Expression and Gene Set Enrichment Analysis

Differentially expressed genes were identified following the limma-voom workflow (3.50) (Law et al., 2014; Ritchie et al., 2015). Gene set variation analysis (GSVA), single-sample gene set enrichment analysis (ssGSEA) and Z scores for each signature were calculated using the R package GSVA (1.42) by switching the "method" parameter (Hanzelmann et al., 2013). In addition to the signatures collected from specific studies, gene sets from the MSigDB-C2-Canonical Pathway (KEGG, BioCarta, PID, Reactome, WikiPathways) were also included in the signature analysis (Liberzon et al., 2011).

Programmed Death Ligand 1 Immunohistochemistry

PD-L1 expression was assessed by the VENTANA PD-L1 (SP263) IHC assay (Ventana Medical Systems, Oro Valley, AZ, United States). The level of PD-L1 was then scored by the percentage of PD-L1 membranous staining on tumor cells (TC).

CD8 Immunofluorescence

CD8 immunofluorescence analysis of FFPE samples was performed with CD8A antibody (SP57, Ventana 790-4460) in a College of American Pathologist (CAP)-controlled area within the Oncology and Immunology Unit of WuXi AppTec using the IF 6-colorWJJ-CD30 protocol on a Leica BOND Rx platform. Whole-slide images were acquired by Leica Aperio VERSA 8. Z1. Image analysis was performed using the HALO software package (Indica Labs, United States).

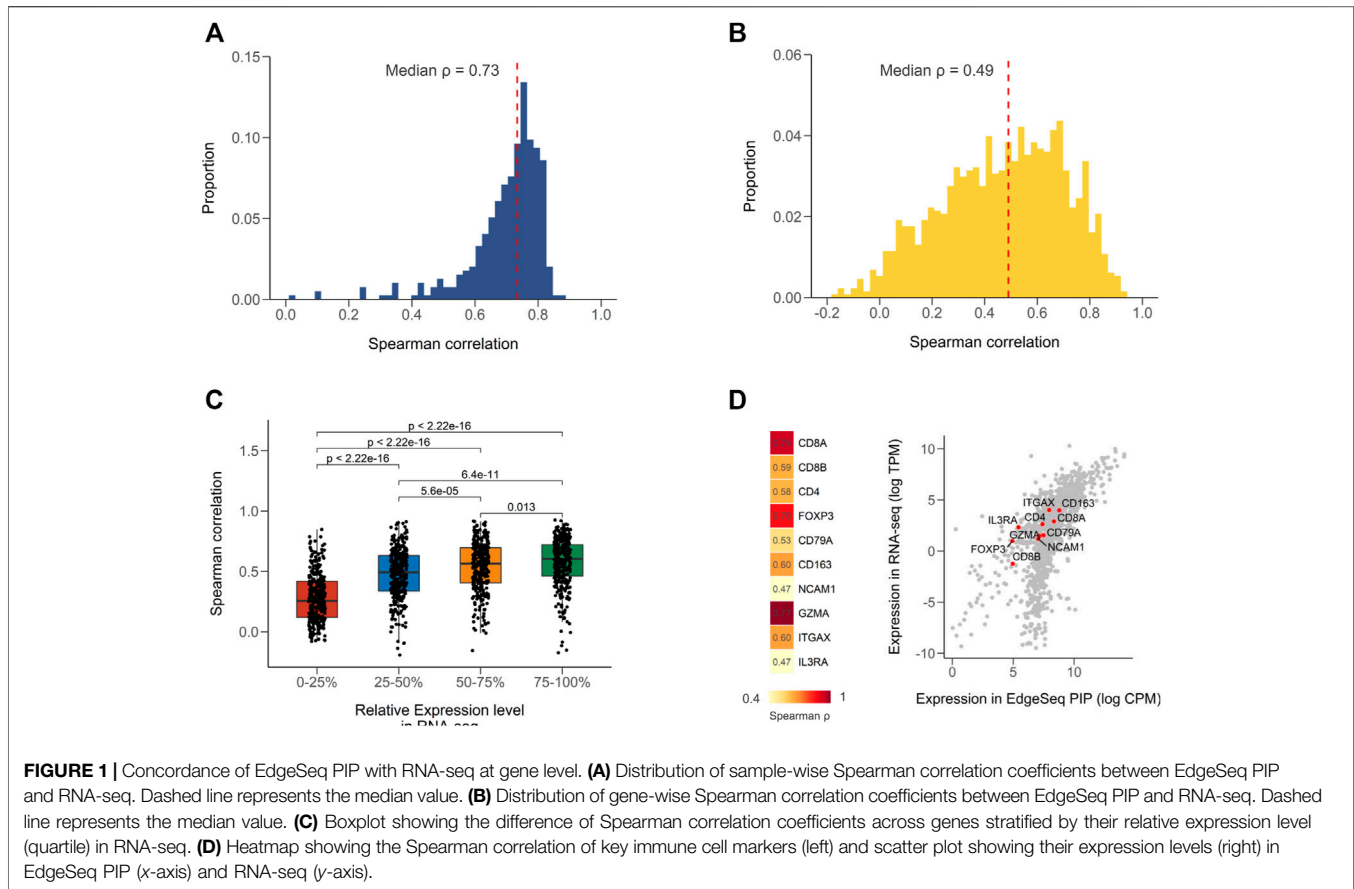
Statistical Analysis

Correlation between two continuous variables was assessed by Spearman correlation. The interpretation of correlation was defined as: negligible (0–0.09), weak (0.10–0.39), moderate (0.4–0.69), strong (0.70–0.89) and very strong (0.90–1.00) (Schober et al., 2018). Differences in medians or continuous variables between two groups were assessed by non-parametric Wilcoxon rank-sum tests. The alpha level for all comparisons was 0.05 unless indicated otherwise. All statistical analyses and visualizations were performed with R (v.4.0.2).

RESULTS

High Concordance Between EdgeSeq Precision Immuno-Oncology Panel and Other GEP Platforms at Gene Level

We first calculated Spearman correlation coefficients for all common samples between EdgeSeq PIP and RNA-seq (**Supplementary Table S2**). The strength of correlation was defined according to a widely used guiding rule for correlation interpretation in medical research (Materials and Methods) (Schober et al., 2018). Of the 395 samples, most had strong correlation (median $\rho = 0.73$) (**Figure 1A**). When assessing the concordance of the two platforms using a gene-by-gene approach, we found a moderate correlation (median $\rho = 0.49$) (**Figure 1B** and **Supplementary Table S3**), which was comparable with the result from a previous study comparing RNA-seq with NanoString (Kwong et al., 2018). We presumed that the weak correlation of some genes was partially attributed to their low expression levels. By stratifying genes according to their relative expression in RNA-seq, we found that the correlation coefficients significantly decreased as the expression level decreased (**Figure 1C**). The genes with expression levels in the lowest 25% only had a median correlation of 0.26 (**Figure 1C**). In addition, the dynamic range of gene expression measured by the median absolute deviation (MAD) also associated with the cross-platform correlation (**Supplementary Figure S1A; Supplementary Table S3**). This might indicate that correlation analysis was not suitable for numbers with low variances. After excluding genes with low expression (<1 TPM) and dynamic range (MAD <0.98, 25% quartile), the median correlation improved from 0.49 to 0.60 (**Supplementary Figure S1B**). Furthermore, considering the intended use of EdgeSeq PIP for profiling immune statuses in tumors, we evaluated the expression of several key immune markers and found all genes had moderate



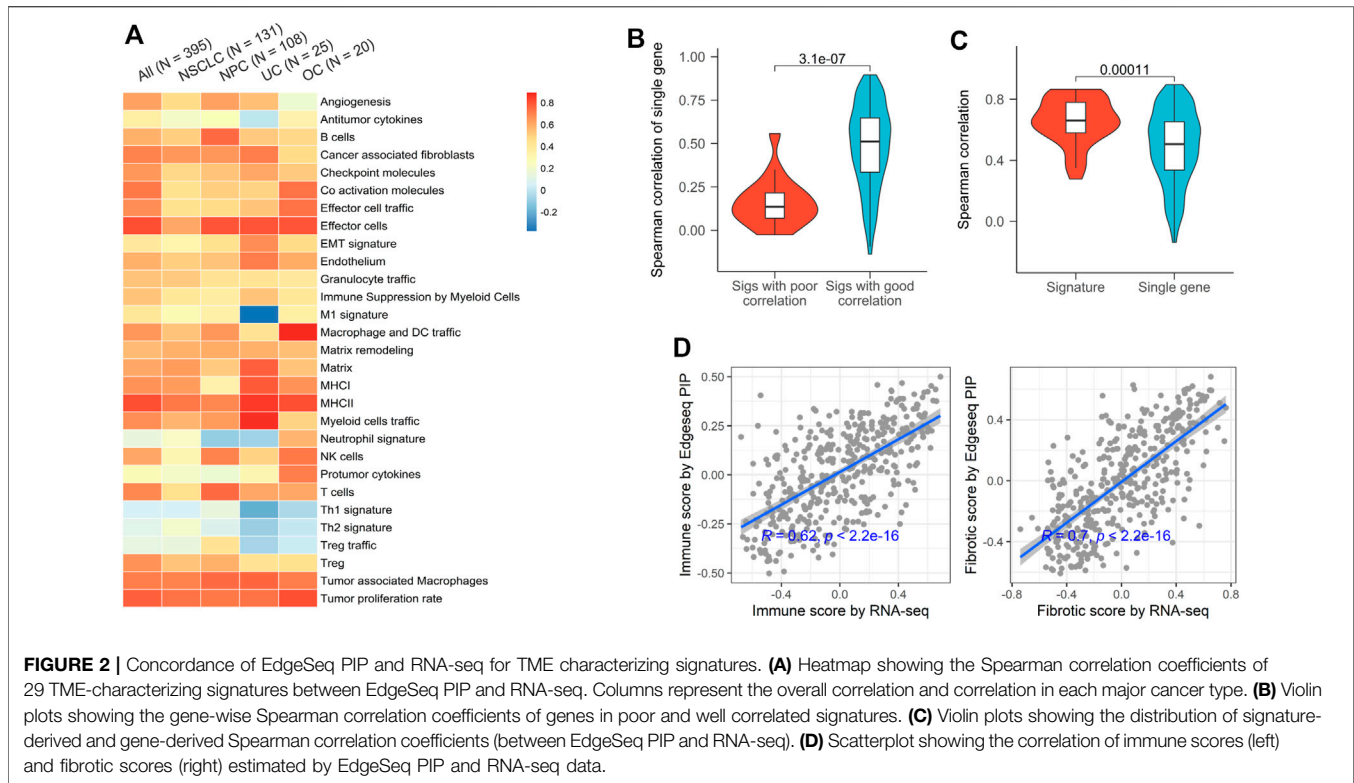
to strong correlations (ρ range from 0.47 to 0.77) (Figure 1D and Supplementary Figure S1C).

We also performed comparisons of EdgeSeq PIP and NanoString in another 17 samples and found strong correlations (median $\rho = 0.81$) using the sample-by-sample approach and moderate correlations (median $\rho = 0.57$) using the gene-by-gene approach (Supplementary Figures S2A,B, Supplementary Table S4). Similar to the results in RNA-seq, the immune markers assessed by EdgeSeq PIP also showed a high level of concordance with NanoString (6/9 genes had strong correlations with $\rho > 0.7$) (Supplementary Figure S2C). Taken together, our findings showed that EdgeSeq PIP had a high degree of agreement with RNA-seq and NanoString at the single-gene level.

High Concordance Between EdgeSeq Precision Immuno-Oncology Panel and RNA-Seq for Characteristic Tumor Microenvironment Gene Signatures

Gene signature, a set of genes with similar expression patterns or biological functions, have been widely used in TME profiling (Liberzon et al., 2011). Compared to single gene, signature-based approach has many significant advantages, including dimension reduction and greater biological interpretability. Here, we used 29 signatures that successfully classified the TMEs of tumors from

TCGA and investigated the applicability of EdgeSeq PIP for TME subtyping (Bagaev et al., 2021) (Supplementary Table S5). First, we assessed the concordance of gene set variation analysis (GSVA) scores derived from EdgeSeq PIP and RNA-seq (Materials and Methods). Most signatures had moderate positive correlations (median $\rho = 0.61$), which persisted even after samples were divided into smaller groups according to cancer types (Figure 2A). Nevertheless, five signatures: neutrophil, Th1, Th2, Treg, and Treg traffic had negligible correlations (median $\rho = 0.01$) between the two platforms (Figure 2A). This result may have been due to the incomplete coverage of gene list in certain signatures by EdgeSeq PIP or to the involvement of genes with low correlations between the two platforms. We first explored whether the percentage of missed genes in each signature influenced the correlation. Surprisingly, the signatures with poor correlation did not have significantly higher rates of missing genes than the others (Supplementary Figure S3B). In contrast, most genes in these five signatures with worse correlations had lower RNA-seq expression levels and thus weaker correlations with EdgeSeq PIP at the single-gene level (Figure 2B and Supplementary Figure S3C). This result suggested that not the incompleteness of genes in the signature but rather the correlation at single-gene level drove the concordance of signatures. After removing these weakly correlated signatures, the remaining signatures had significantly higher correlations as well as lower variations



than individual genes (Figure 2C), which indicated that signature-based analysis had the property of noise reduction.

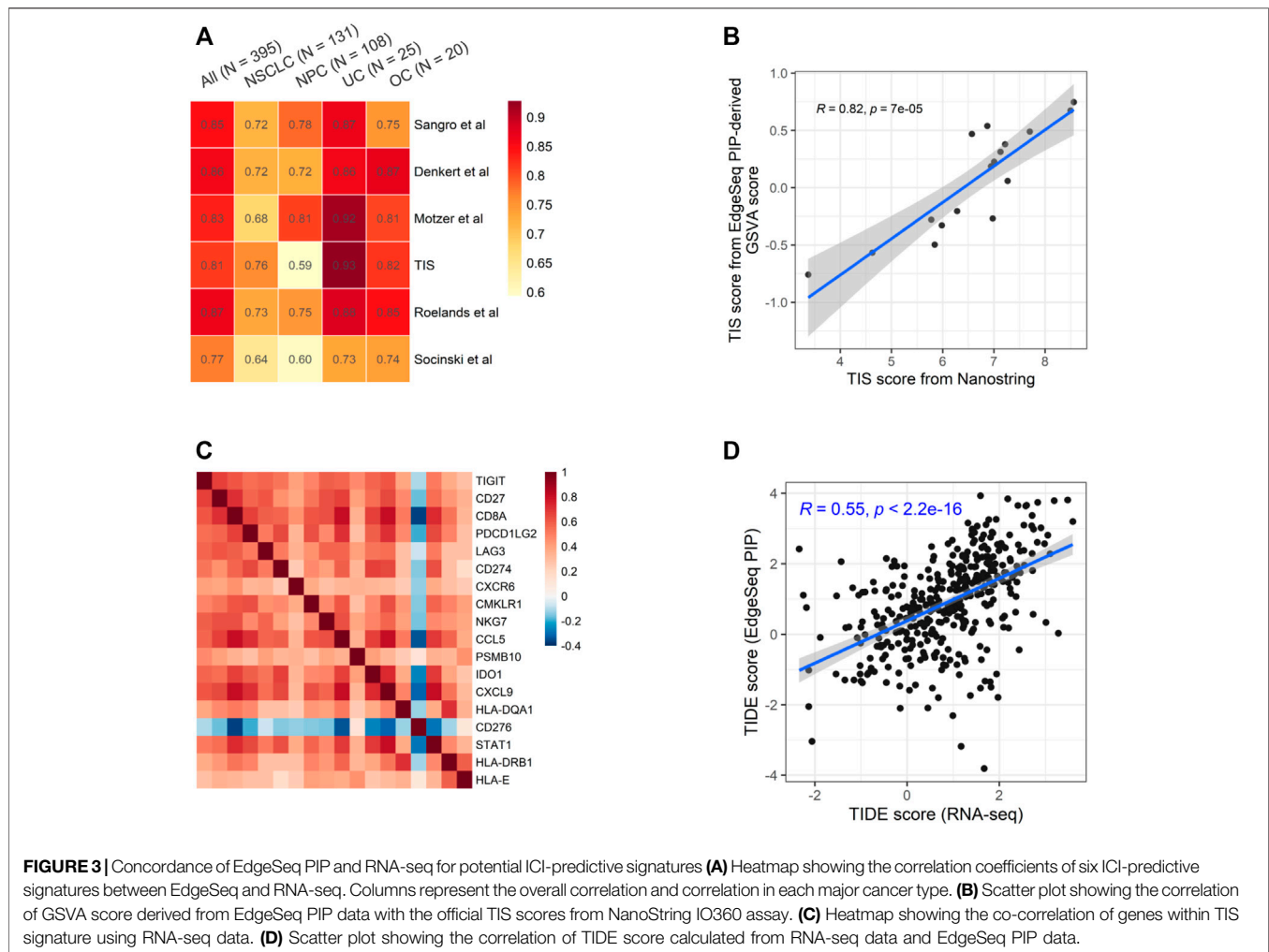
In the previous study, the TME of TCGA tumors were classified into 4 subgroups and annotated according to their enrichment of immune or fibroblast components (Bagaev et al., 2021). By averaging the scores of corresponding signatures (Supplementary Table S5), we derived “immune enrichment” and “fibroblast enrichment” score for each sample and found the scores generated by EdgeSeq PIP and RNA-seq correlated well ($\rho = 0.62$ and 0.70 separately) (Figure 2D). Taken together, the data from EdgeSeq PIP generated reliable signature scores and TME subtype results.

High Concordance Between EdgeSeq Precision Immuno-Oncology Panel and Other Gene Expression Profiling Platforms Regarding Potential Immune Checkpoint Inhibitor-Predictive Gene Signatures

We chose six well-known ICI-predictive signatures and studied the fidelity of calculating them using EdgeSeq PIP data (Supplementary Table S6) (Denkert et al., 2015; Cristescu et al., 2018; Socinski et al., 2018; Motzer et al., 2020; Roelands et al., 2020; Sangro et al., 2020). By calculating the correlation between signature scores derived from EdgeSeq PIP and RNA-seq, we found a high degree of concordance for all signatures (median $\rho = 0.84$) (Figure 3A and Supplementary Figure S4A). As expected, most genes in these signatures were well-correlated between the two platforms (median $\rho = 0.57$) (Supplementary Figure S4B).

Among these signatures, tumor inflammation signature (TIS) is originally derived from the NanoString platform using predefined algorithm and weight (Cristescu et al., 2018). Using the 17 samples with both EdgeSeq PIP and NanoString data, we applied GSVA on EdgeSeq PIP data to get “TIS-PIP score” and then compared it with the “TIS-NanoString score” derived from its original algorithm. We found a moderate correlation ($\rho = 0.61$) of TIS score despite of different platforms and algorithms (Figure 3B). In addition to GSVA, we also tested other methods for EdgeSeq PIP, such as ssGSEA and Z-score (Materials and Methods) and found consistent results (Supplementary Figures S5A,B). We reasoned that in addition to the cross-platform concordance at the single-gene level (Supplementary Figure S5C), the strong intra-signature correlation of genes could also account for the high correlation coefficient. Indeed, we found that except for *CD276*, all other 17 genes were moderately correlated with each other (median $\rho = 0.46$) (Figure 3C and Supplementary Figure S5D).

Additionally, we assessed tumor immune dysfunction and exclusion (TIDE), a computational framework integrating multiple signatures of tumor evasion that successfully predicts the responses of melanoma and lung cancer patients to ICIs (Jiang et al., 2018). We found that the TIDE scores generated by EdgeSeq PIP had moderate correlations with those generated by RNA-seq ($\rho = 0.55$) (Figure 3D). Overall, our results suggested that data from EdgeSeq PIP can be reliably used for most well-known ICI efficacy prediction algorithms.



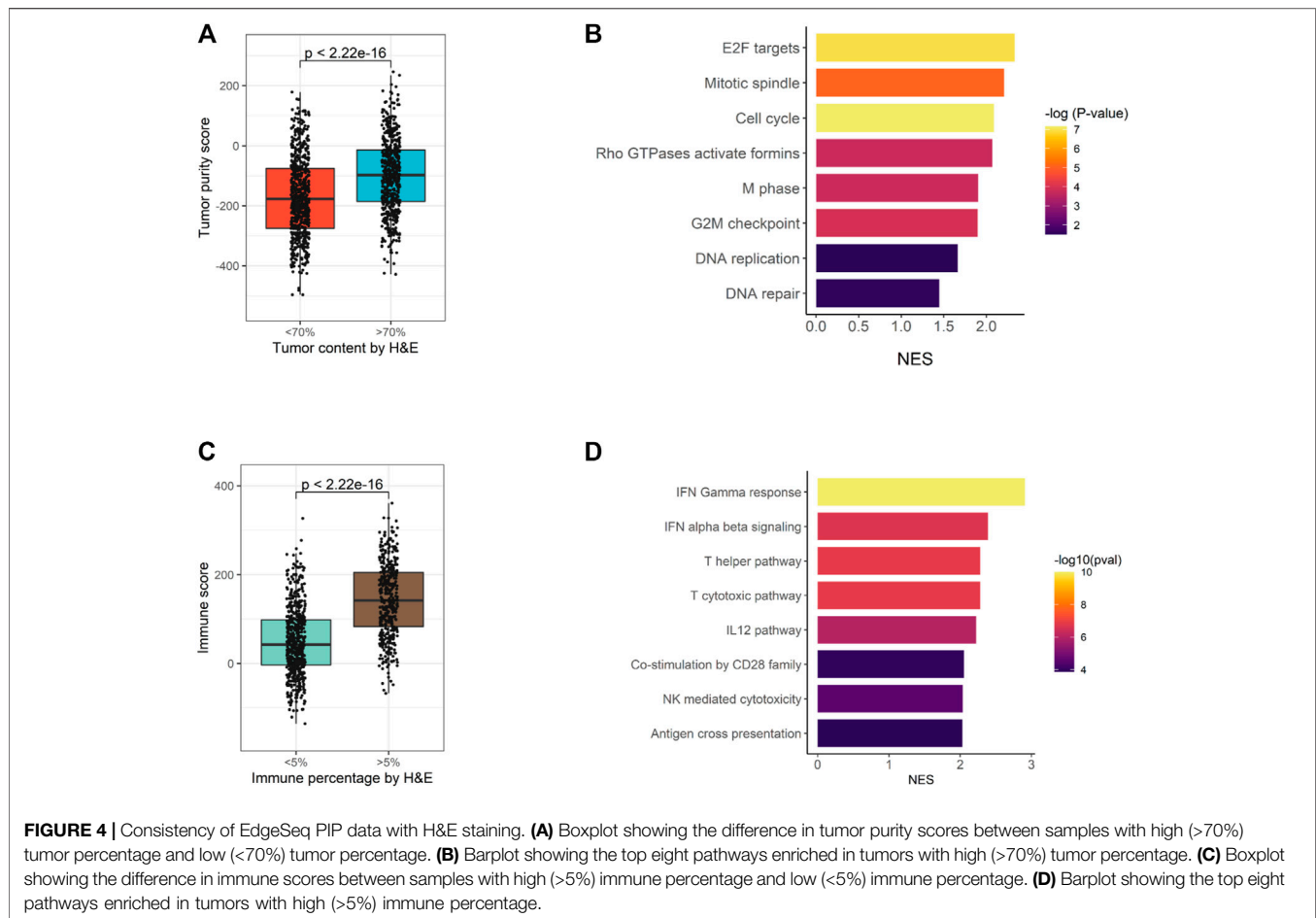
High Consistency Between EdgeSeq Precision Immuno-Oncology Panel and H & E Staining Results

For 1,174 samples assessed by EdgeSeq PIP, H & E staining was also performed to evaluate the percentages of tumor cells and immune cells. Here, we assessed whether the H&E staining results could be recapitulated by EdgeSeq PIP at mRNA level. First, we used ESTIMATE (Yoshihara et al., 2013) to calculate the tumor purity score for each sample based on EdgeSeq PIP data. We found consistent result with H & E staining: the samples with a high tumor percentage (>70%, assessed by H & E) had significant higher tumor purity scores than those with low tumor percentage (<70%) (Figure 4A). In addition, we identified the differentially expressed genes between these two groups and performed gene set enrichment analysis (GSEA) (Materials and Methods) (Supplementary Table S7). Our results showed that cell cycle-related pathways were enriched among the genes upregulated in samples with >70% tumor percentage (Figure 4B and Supplementary Figures S6A,B). This finding aligned with the H & E result as most tumor cells had an extremely high proliferation rate.

Similarly, we studied the concordance of EdgeSeq PIP with the percentage of infiltrated immune cells estimated by H & E. On the one hand, tumors with >5% immune cell percentage (by H & E) had significant higher ImmunoScores (calculated by ESTIMATE) than others (Figure 4C). On the other hand, immune-related pathways and representative immune marker genes such as *CD8A*, *STAT1* and *GZMK* were up-regulated in tumors with >5% immune percentage (Supplementary Figures S6C,D, Supplementary Table S8). Overall, this evidence suggested that the abundance of tumor or immune cells assessed by histological methods was well reflected by EdgeSeq PIP.

High Consistency Between EdgeSeq Precision Immuno-Oncology Panel and Immunohistochemistry

PD-L1 IHC score has been widely used for predicting responses to ICI-based therapies. A previous study showed that the mRNA level of PD-L1 assessed RNA-seq had good concordance with its protein level assessed by IHC (Conroy et al., 2019). Therefore, the correlation with PD-L1 IHC could be used to evaluate the fidelity of NGS-based approaches. In 369 samples with PD-L1 IHC data



(Supplementary Figure S8), we found that the percentage of PD-L1 positive tumor cells had a moderate correlation ($\rho = 0.65$) with the mRNA level of PD-L1 (*CD274*) assessed by EdgeSeq PIP (Figure 5A), which was comparable to its association with RNA-seq ($\rho = 0.69$). It suggested that the remaining lack of correlation was mainly due to the discrepancy between mRNA and protein, rather than technical issues specific to platform. In addition, we calculated the correlation of PD-L1 IHC with every gene and found that *CD274* had the highest correlation coefficient among all genes (Figure 5B), which further suggested the fidelity of EdgeSeq PIP data.

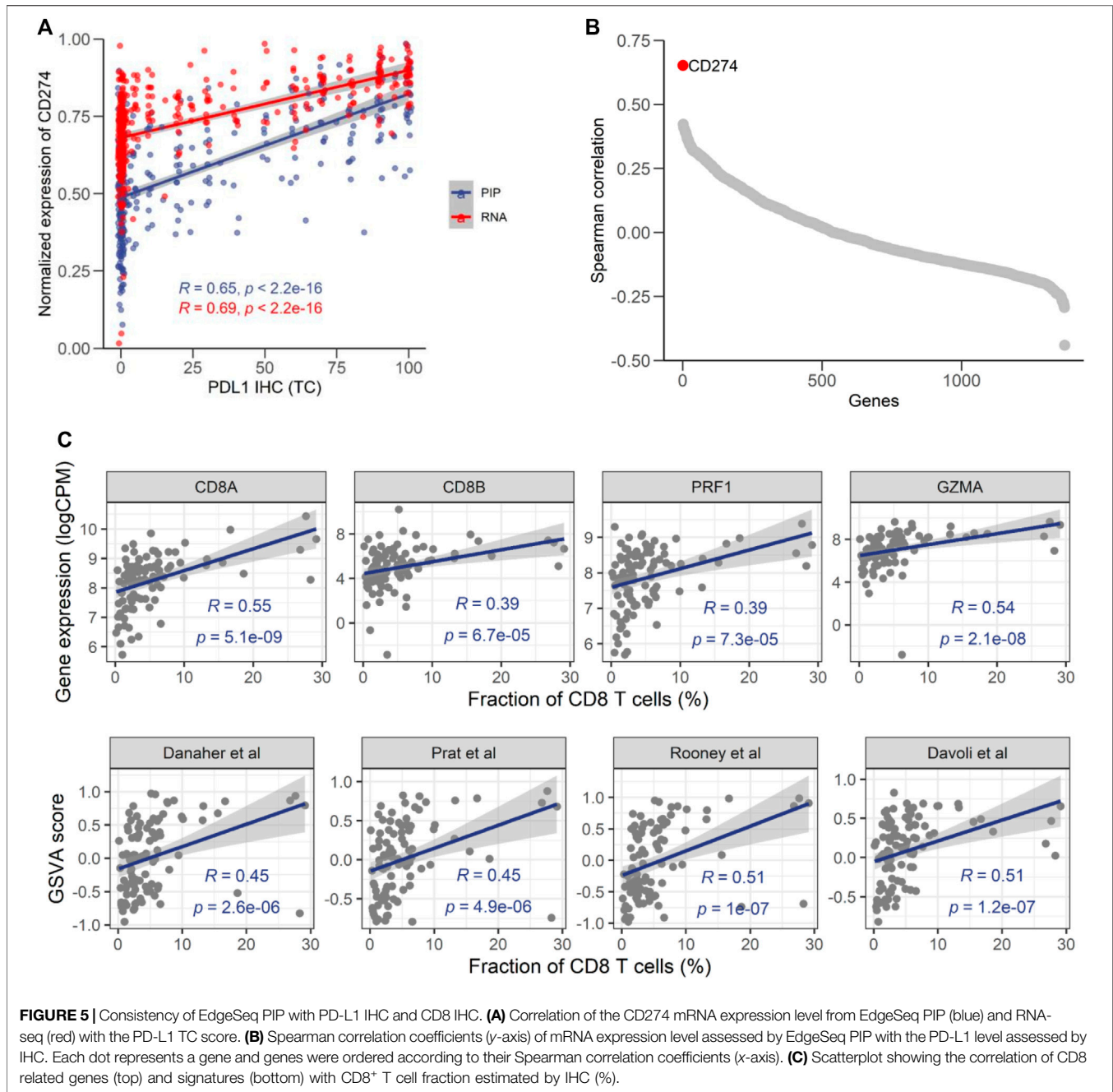
To further investigate the concordance of EdgeSeq PIP with IHC, we assessed another important immuno-oncology marker, the fraction of CD8⁺ T cells, by CD8A IHC in another 98 tumors (Supplementary Figure S9). We investigated the correlation of CD8 abundance estimated by IHC with the expression level of 4 traditional CD8⁺ T cell markers (*CD8A*, *CD8B*, *PRF1*, *GZMA*) from EdgeSeq PIP and the EdgeSeq PIP-derived GSVA scores of 4 well-known CD8⁺ T signatures (Supplementary Table S9) (Rooney et al., 2015; Danaher et al., 2017; Davoli et al., 2017; Prat et al., 2017). All estimates for CD8⁺ T cells had significant positive correlations with the IHC results (Figure 5C). Interestingly, *CD8A* outperformed the other genes ($\rho = 0.55$) and even the gene signatures. This result suggested that at least for data

generated by EdgeSeq PIP, the expression level of *CD8A* is sufficient for estimating the fraction of CD8⁺ T cells in tumors. Overall, our results indicated that the data generated by EdgeSeq PIP were comparable to IHC data and could be used as a complementary method for evaluating the protein levels of PD-L1 and the CD8⁺ T cell fraction.

DISCUSSION

Gene expression profiling plays an important role in immuno-oncology related biomarker studies, as it helps to elucidate the TME landscape, and many predictive biomarkers have been derived for ICI-based therapies. As clinical specimens in certain cancers are obtained from core needle biopsies of small size and suboptimal quality, tissue availability restricts the clinical application of traditional GEP methods (e.g., RNA-seq). EdgeSeq PIP, an assay focused on well-studied genes related to immuno-oncology, uses an extraction-free technology to minimize sample requirements. This method mitigates the challenge for sample acquisition and has been widely used in clinical trials.

In this study, we provided evidence that EdgeSeq PIP is a robust assay with a high degree of concordance with multiple platforms, including RNA-seq, tumor/immune cell fraction



estimation by H & E staining, PD-L1 scoring and CD8⁺ T cell abundance estimation by IHC. By analyzing data from 395 clinical FFPE samples, we found that EdgeSeq PIP showed good correlation with RNA-seq at the single-gene level, especially after excluding genes with low expression levels or dynamic ranges. Importantly, we demonstrated that EdgeSeq PIP, with incomplete coverage of genes in some signatures, still generated signature scores that were well correlated with those of RNA-seq. This result indicated that EdgeSeq PIP data can be reliably utilized for signature-based analyses of TME subtypes and ICI efficacy prediction. In addition, we found that EdgeSeq PIP aligned well with the overall percentage of tumor/immune

cells determined by H & E staining. Furthermore, our results showed that the RNA levels determined by EdgeSeq PIP aligned well with the protein levels determined by the IHC assessment of two important immunotherapy biomarkers, the fraction of PD-L1-positive tumor cells and CD8⁺ T cells.

As the method of choice for transcriptome probing, RNA-seq has always been used as a reference standard to benchmark other GEP assays. In our results, EdgeSeq PIP correlated well with RNA-seq except for a small proportion of genes with low expression levels. This result was within expectations, as true signals for genes with low expression are hard to distinguish from noise, which would diminish the correlation (Kwong

et al., 2018). In addition, loss of degraded transcripts during RNA extraction, especially for low-expressing genes in archived FFPE samples, could further exacerbate the discordancy between the two platforms. There was evidence that EdgeSeq PIP, using its extraction-free technique, offer superior sensitivity on these “discordant” genes (Ran et al., 2020). Nevertheless, such genes should be used with caution and should even excluded from some analyses involving data from both RNA-seq and EdgeSeq PIP.

In this study, we found that the gene signature score derived from EdgeSeq PIP was strongly correlated with that from RNA-seq, although some genes in given signatures were not covered by EdgeSeq PIP probes. We reasoned that such result was attributed to the coexpression of genes within the signature (e.g., TIS). Such coexpression led to gene “redundancy” and made the signature score calculation robustly susceptible to gene dropout and different algorithms. Though some poorly correlated genes resulted in a low correlation among certain signatures, after their exclusion, most signatures defined based on RNA-seq data could be directly transferred and used on EdgeSeq PIP data. However, we must admit that the signatures used in our study were limited to those related to immuno-oncology, which predominantly consisted of surface markers sharing similar expression patterns. For other signatures, such as cellular signaling pathways, the effect of “gene dropout” needs to be further investigated.

PD-L1 IHC has been used to predict responses to ICI-based therapy. A previous study revealed that the expression of PD-L1 measured by RNA-seq had a strong association with that measured by IHC (Conroy et al., 2019). Here, we found that the mRNA level of PD-L1 generated by EdgeSeq PIP had a high correlation with the PD-L1 IHC level as well. Although a biological gap exists between mRNAs and proteins, EdgeSeq PIP at least provides another layer of PD-L1 quantification that can be exploited for the prediction of ICI efficacy. In addition, EdgeSeq PIP quantifies PD-L1 transcripts without the need for subjective scoring and cell type discrimination by pathologists, which might introduce substantial discrepancy across studies (Hirsch et al., 2017).

In summary, for the first time, we comprehensively benchmarked EdgeSeq PIP with multiple platforms using large-scale clinical FFPE samples, and the results are reflective of its actual performance in clinical practice. Our results showed that EdgeSeq PIP generated data comparable to those generated by classical methods. Considering its low sample requirement and cost efficiency, the technology would get an ever-increasing application for biomarker studies in clinical trials.

REFERENCES

Aisner, D. L., Rumery, M. D., Merrick, D. T., Kondo, K. L., Nijmeh, H., Linderman, D. J., et al. (2016). Do More with Less: Tips and Techniques for Maximizing Small Biopsy and Cytology Specimens for Molecular and Ancillary Testing: The

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the **Supplementary Material**, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by please refer to the clinical trials mentioned in our manuscript. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YS, XM, and PZ conceived the study. YS and XM performed the bioinformatics analysis. WS, SL, and TL analyzed the immunohistochemistry data. YS, XM, WS, and PZ wrote the article. All authors discussed the results and commented on the manuscript.

FUNDING

This study was funded by BeiGene. Funding for open access charge: BeiGene.

ACKNOWLEDGMENTS

We thank Jayesh Desai (Peter MacCallum Cancer Centre, University of Melbourne), Jun Guo (Harbin Medical University Cancer Hospital), Lin Shen (Peking University Cancer Hospital and Institute), Jie Wang (Cancer Institute and Hospital, Chinese academy of Medical Sciences), Yilong Wu (Guangdong Lung Cancer Institute), Jianmin Xu (The fifth Medical Center, Chinese PLA General Hospital), Dingwei Ye (Fudan University) and Li Zhang (Sun Yat-sen University) for leading the clinical trials included in this study. We also thank Mo Liu (BeiGene) for his kind suggestions on this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2022.899353/full#supplementary-material>

University of Colorado Experience. *Arch. Pathol. Lab. Med.* 140, 1206–1220. doi:10.5858/arpa.2016-0156-ra

Anguiano, E., Desai, K., Hayati, S., Szabo, P. M., Pant, S., Wang, L., et al. (2020). Abstract 1997: Comparison of Gene Expression Profiling Platforms: Translatability of Tumor Inflammation Gene Signatures. *Cancer Res.* 80, 1997–1997. doi:10.1158/1538-7445.am2020-1997

- Bagaev, A., Kotlov, N., Nomic, K., Svekolkin, V., Gafurov, A., Isaeva, O., et al. (2021). Conserved Pan-Cancer Microenvironment Subtypes Predict Response to Immunotherapy. *Cancer Cell* 39, 845–865. doi:10.1016/j.ccell.2021.04.014
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170
- Chen, D. S., and Mellman, I. (2017). Elements of Cancer Immunity and the Cancer-Immune Set Point. *Nature* 541, 321–330. doi:10.1038/nature21349
- Conroy, J. M., Pabla, S., Nesline, M. K., Glenn, S. T., Papanicolaou-Sengos, A., Burgher, B., et al. (2019). Next Generation Sequencing of PD-L1 for Predicting Response to Immune Checkpoint Inhibitors. *J. Immunother. Cancer* 7, 18. doi:10.1186/s40425-018-0489-5
- Cristescu, R., Mogg, R., Ayers, M., Albright, A., Murphy, E., Yearley, J., et al. (2018). Pan-tumor Genomic Biomarkers for PD-1 Checkpoint Blockade-Based Immunotherapy. *Science* 362. doi:10.1126/science.aar3593
- Danaher, P., Warren, S., Dennis, L., D'Amico, L., White, A., Disis, M. L., et al. (2017). Gene Expression Markers of Tumor Infiltrating Leukocytes. *J. Immunother. cancer* 5, 18. doi:10.1186/s40425-017-0215-8
- Davoli, T., Uno, H., Wooten, E. C., and Elledge, S. J. (2017). Tumor Aneuploidy Correlates with Markers of Immune Evasion and with Reduced Response to Immunotherapy. *Science* 355. doi:10.1126/science.aaf8399
- Denkert, C., von Minckwitz, G., Brase, J. C., Sinn, B. V., Gade, S., Kronenwett, R., et al. (2015). Tumor-infiltrating Lymphocytes and Response to Neoadjuvant Chemotherapy with or without Carboplatin in Human Epidermal Growth Factor Receptor 2-positive and Triple-Negative Primary Breast Cancers. *Jco* 33, 983–991. doi:10.1200/jco.2014.58.1967
- Desai, J., Zhou, Q., Deva, S., Zhao, J., Wang, J., Tan, W., et al. (2020). 76 Potential Mechanisms of Resistance Identified through Analysis of Multiple Biomarkers in Immune Hot Non-responders with Non-small Cell Lung Cancer (NSCLC) Treated with Tislelizumab. *J. Immunother. Cancer* 8, A47–A47.
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* 29, 15–21. doi:10.1093/bioinformatics/bts635
- Evers, D. L., Fowler, C. B., Cunningham, B. R., Mason, J. T., and O'Leary, T. J. (2011). The Effect of Formaldehyde Fixation on RNA. *J. Mol. Diagnostics* 13, 282–288. doi:10.1016/j.jmoldx.2011.01.010
- Garg, S. K., Welsh, E. A., Fang, B., Hernandez, Y. I., Rose, T., Gray, J., et al. (2020). Multi-Omics and Informatics Analysis of FFPE Tissues Derived from Melanoma Patients with Long/Short Responses to Anti-PD1 Therapy Reveals Pathways of Response. *Cancers (Basel)* 12. doi:10.3390/cancers12123515
- Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSEA: Gene Set Variation Analysis for Microarray and RNA-Seq Data. *BMC Bioinforma.* 14, 7. doi:10.1186/1471-2105-14-7
- Havel, J. J., Chowell, D., and Chan, T. A. (2019). The Evolving Landscape of Biomarkers for Checkpoint Inhibitor Immunotherapy. *Nat. Rev. Cancer* 19, 133–150. doi:10.1038/s41568-019-0116-x
- Hirsch, F. R., McElhinny, A., Stanforth, D., Ranger-Moore, J., Jansson, M., Kulangara, K., et al. (2017). PD-L1 Immunohistochemistry Assays for Lung Cancer: Results from Phase 1 of the Blueprint PD-L1 IHC Assay Comparison Project. *J. Thorac. Oncol.* 12, 208–222. doi:10.1016/j.jtho.2016.11.2228
- Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T Cell Dysfunction and Exclusion Predict Cancer Immunotherapy Response. *Nat. Med.* 24, 1550–1558. doi:10.1038/s41591-018-0136-1
- Kwong, L. N., De Macedo, M. P., Haydu, L., Joon, A. Y., Tetzlaff, M. T., Calderone, T. L., et al. (2018). Biological Validation of RNA Sequencing Data from Formalin-Fixed Paraffin-Embedded Primary Melanomas. *JCO Precis. Oncol.* 2018. doi:10.1200/po.17.00259
- Law, C. W., Chen, Y., Shi, W., and Smyth, G. K. (2014). Voom: Precision Weights Unlock Linear Model Analysis Tools for RNA-Seq Read Counts. *Genome Biol.* 15, R29. doi:10.1186/gb-2014-15-2-r29
- Li, B., and Dewey, C. N. (2011). RSEM: Accurate Transcript Quantification from RNA-Seq Data with or without a Reference Genome. *BMC Bioinforma.* 12, 323. doi:10.1186/1471-2105-12-323
- Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdottir, H., Tamayo, P., and Mesirov, J. P. (2011). Molecular Signatures Database (MSigDB) 3.0. *Bioinformatics* 27, 1739–1740. doi:10.1093/bioinformatics/btr260
- Martel, R. R., Botros, I. W., Rounseville, M. P., Hinton, J. P., Staples, R. R., Morales, D. A., et al. (2002). Multiplexed Screening Assay for mRNA Combining Nuclease Protection with Luminescent Array Detection. *ASSAY Drug Dev. Technol.* 1, 61–71. doi:10.1089/154065802761001310
- Martin-Broto, J., Cruz, J., Penel, N., Le Cesne, A., Hindi, N., Luna, P., et al. (2020). Pazopanib for Treatment of Typical Solitary Fibrous Tumours: a Multicentre, Single-Arm, Phase 2 Trial. *Lancet Oncol.* 21, 456–466. doi:10.1016/s1470-2045(19)30826-5
- Motzer, R. J., Robbins, P. B., Powles, T., Albiges, L., Haanen, J. B., Larkin, J., et al. (2020). Avelumab Plus Axitinib versus Sunitinib in Advanced Renal Cell Carcinoma: Biomarker Analysis of the Phase 3 JAVELIN Renal 101 Trial. *Nat. Med.* 26, 1733–1741. doi:10.1038/s41591-020-1044-8
- Prat, A., Navarro, A., Paré, L., Reguart, N., Galván, P., Pascual, T., et al. (2017). Immune-Related Gene Expression Profiling after PD-1 Blockade in Non-small Cell Lung Carcinoma, Head and Neck Squamous Cell Carcinoma, and Melanoma. *Cancer Res.* 77, 3540–3550. doi:10.1158/0008-5472.can-16-3556
- Qi, Z., Wang, L., Desai, K., Cogswell, J., Stern, M., Lawson, B., et al. (2019). Reliable Gene Expression Profiling from Small and Hematoxylin and Eosin-Stained Clinical Formalin-Fixed, Paraffin-Embedded Specimens Using the HTG EdgeSeq Platform. *J. Mol. Diagnostics* 21, 796–807. doi:10.1016/j.jmoldx.2019.04.011
- Qi, Z., Wang, L., He, A., Ma-Edmonds, M., and Cogswell, J. (2016). Evaluation and Selection of a Non-PCR Based Technology for Improved Gene Expression Profiling from Clinical Formalin-Fixed, Paraffin-Embedded Samples. *Bioanalysis* 8, 2305–2316. doi:10.4155/bio-2016-0186
- Ran, D., Moharil, J., Lu, J., Gustafson, H., Culm-Merdek, K., Strand-Tibbitts, K., et al. (2020). Platform Comparison of HTG EdgeSeq and RNA-Seq for Gene Expression Profiling of Tumor Tissue Specimens. *Jco* 38, 3566–3566. doi:10.1200/jco.2020.38.15_suppl.3566
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007
- Roelands, J., Hendrickx, W., Zoppoli, G., Mall, R., Saad, M., Halliwill, K., et al. (2020). Oncogenic States Dictate the Prognostic and Predictive Connotations of Intratumoral Immune Response. *J. Immunother. Cancer* 8. doi:10.1136/jitc-2020-000617
- Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G., and Hacohen, N. (2015). Molecular and Genetic Properties of Tumors Associated with Local Immune Cytolytic Activity. *Cell* 160, 48–61. doi:10.1016/j.cell.2014.12.033
- Sangro, B., Melero, I., Wadhawan, S., Finn, R. S., Abou-Alfa, G. K., Cheng, A.-L., et al. (2020). Association of Inflammatory Biomarkers with Clinical Outcomes in Nivolumab-Treated Patients with Advanced Hepatocellular Carcinoma. *J. Hepatology* 73, 1460–1469. doi:10.1016/j.jhep.2020.07.026
- Sanmamed, M. F., and Chen, L. (2018). A Paradigm Shift in Cancer Immunotherapy: From Enhancement to Normalization. *Cell* 175, 313–326. doi:10.1016/j.cell.2018.09.035
- Schober, P., Boer, C., and Schwarte, L. A. (2018). Correlation Coefficients. *Anesth. Analgesia* 126, 1763–1768. doi:10.1213/ane.0000000000002864
- Socinski, M. A., Jotte, R. M., Cappuzzo, F., Orlandi, F., Stroyakovskiy, D., Nogami, N., et al. (2018). Atezolizumab for First-Line Treatment of Metastatic Nonsquamous NSCLC. *N. Engl. J. Med.* 378, 2288–2301. doi:10.1056/nejmoa1716948
- Song, Y., Gao, Q., Zhang, H., Fan, L., Zhou, J., Zou, D., et al. (2021). Tislelizumab for Relapsed/refractory Classical Hodgkin Lymphoma: 3-year Follow-Up and Correlative Biomarker Analysis. *Clin. Cancer Res.* 28, 1147–1156. doi:10.1158/1078-0432.CCR-21-2023
- Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The Immune Landscape of Cancer. *Immunity* 48, 812–e14. doi:10.1016/j.immuni.2018.03.023
- Vaddepally, R. K., Kharel, P., Pandey, R., Garje, R., and Chandra, A. B. (2020). Review of Indications of FDA-Approved Immune Checkpoint Inhibitors Per NCCN Guidelines with the Level of Evidence. *Cancers (Basel)* 12. doi:10.3390/cancers12030738

- Wang, L., Saci, A., Szabo, P. M., Chasalow, S. D., Castillo-Martin, M., Domingo-Domenech, J., et al. (2018). EMT- and Stroma-Related Gene Expression and Resistance to PD-1 Blockade in Urothelial Cancer. *Nat. Commun.* 9, 3503. doi:10.1038/s41467-018-05992-x
- Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-seq: a Revolutionary Tool for Transcriptomics. *Nat. Rev. Genet.* 10, 57–63. doi:10.1038/nrg2484
- Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-García, W., et al. (2013). Inferring Tumour Purity and Stromal and Immune Cell Admixture from Expression Data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612

Conflict of Interest: YS, XM, WS, TL, LL, SL, ZS, YZ, and PZ were employed by the company BeiGene (Beijing) Co., Ltd.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Shi, Ma, Shen, Liu, Liang, Liu, Shen, Zhang and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.