The Journal of Molecular Diagnostics, Vol. 21, No. 5, September 2019

the Journal of Molecular Diagnostics
jmd.amjpathol.org

Reliable Gene Expression Profiling from Small and Hematoxylin and Eosin—Stained Clinical Formalin-Fixed, Paraffin-Embedded Specimens Using the HTG EdgeSeq Platform

Zhenhao Qi,* Lisu Wang,* Keyur Desai,* John Cogswell,* Mark Stern,† Byron Lawson,† Sid P. Kerkar,* and Patrik Vitazka*

From Translational Medicine, Bristol-Myers Squibb, Princeton, New Jersey; and Pharmaceutical Services & Companion Diagnostics,† HTG Molecular Diagnostics, Inc., Tucson, Arizona*

Accepted for publication April 16, 2019.

Address correspondence to Lisu Wang, M.D., Translational Medicine, Bristol-Myers Squibb, 311 Pennington-Rocky Hill Rd., 3B-2.01, Pennington, NJ 08534. E-mail: lisu.wang@bms.com.

Clinical biomarker studies are often hindered by the scarcity or suboptimal quality of biological specimens. EdgeSeq, a transcriptomics analysis platform, combines quantitative nuclease protection assay technology with next-generation sequencing, using small amounts of starting material and delivering reproducible gene expression profiles from challenging material, such as formalin-fixed, paraffin-embedded (FFPE) tissue. To evaluate EdgeSeq for analysis of archives of stained FFPE tissue, EdgeSeq was performed on unstained, hematoxylin and eosin (H&E)—stained, and immunohistochemistry-stained slides from patients with small-cell and non—small-cell lung cancer. Pairwise comparisons of gene expression profiles from stained and unstained slides showed higher Pearson correlation coefficients with H&E staining (0.86 to 0.97) than with immunohistochemistry staining (0.21 to 0.56). A 25-gene interferon-γ signature score from unstained slides showed a Pearson correlation coefficient of 0.92 with H&E-stained slides and a significant Spearman correlation ($P = 0.0025$) with immune scores. To test gene expression profiling in small samples, FFPE sample equivalents were examined from 5.0 to 0.08 mm^2 of a section (5 μm thick); sample equivalents $\geq$0.31 mm^2 showed alignment rates >69% and pairwise Pearson correlation coefficients $\geq$0.87. EdgeSeq can, thus, be used to profile small and H&E-stained FFPE tumor specimens to obtain biomarker data from limited tissue in oncology clinical trials and enable research into tumor microenvironment and immune cell engagement with tumors at the locoregional level. *(J Mol Diagn 2019, 21: 796—807; https://doi.org/10.1016/j.jmoldx.2019.04.011)*

Genomic profiling of patient samples by next-generation sequencing is extremely valuable, revolutionizing our knowledge of the molecular basis of cancer and allowing the identification of patients who may benefit from personalized therapy.[1] In particular, genomic profiling is currently being used to stratify populations of patients with cancer of the lung and bronchus, the leading cause of death by cancer in the United States.[2] Histopathologic or molecular subtyping of lung cancers guides clinicians in the selection of appropriate treatments because there are now several therapies with varying efficacy on different histologic subtypes.[3,4] Assessment of molecular biomarkers also has the potential to predict the responses of patients with lung cancer to specific therapies.[5—9]

Tissue availability can be a major limiting factor in the study of cancer therapy. In patients with lung cancer, for example, tissue acquisition can be difficult, because 70% of

Supported by a collaborative project between Bristol-Myers Squibb and HTG Molecular Diagnostics, Inc.

Disclosures: L.W. and K.D. are employees of and stockholders in Bristol-Myers Squibb; J.C., Z.Q., S.P.K., and P.V. are former employees of Bristol-Myers Squibb; Z.Q. and J.C. are currently employed by and own shares in Daiichi Sankyo, Inc.; Z.Q. owns shares in Bristol-Myers Squibb; M.S. and B.L. are employees of HTG Molecular Diagnostics, Inc.; B.L. is also involved in patent US8741564B2 for improvements over proprietary sequencing.

Current address of Z.Q. and J.C., Daiichi Sankyo, Inc., Basking Ridge, NJ; of P.V., Roche Sequencing Solutions, Inc., Pleasanton, CA; of S.P.K., Boehringer Ingelheim Corporation, Ridgefield, CT.

https://doi.org/10.1016/j.jmoldx.2019.04.011

lung cancers are unresectable.[4] Discomfort and risk of complications lead to an unwillingness from patients and physicians to obtain multiple biopsies.[10,11] Furthermore, many patients are diagnosed in the advanced or metastatic stages of the disease. Surgical procedures are avoided in these patients, and there is a requirement to obtain biological samples using the least invasive method possible.[12] Lung cancer samples are, therefore, commonly obtained from core needle, small transbronchial or small endobronchial biopsies, or cytologic samples.[3,4,12]

Several established biomarkers can predict increased response to targeted therapies in patients with lung cancer, including oncogenic rearrangements in anaplastic lymphoma kinase (*ALK*)[13] and point mutations in epidermal growth factor receptor (*EGFR*).[14,15] Expression of programmed death ligand 1 (PD-L1) protein can be associated with greater efficacy of anti−programmed death-1 (PD-1) or anti−PD-L1 immunotherapy in non−small-cell lung cancer (NSCLC),[16−18] but may not be predictive for all lung cancer subtypes or immunotherapies.[19]

Recent studies have supported the diagnosis of lung cancer subtypes from molecular signatures in mRNA or genomic DNA.[20−23] Low expression or absence of *ERCC1* has been linked to longer overall survival in patients with locally advanced NSCLC who received neoadjuvant cisplatin-based chemotherapy.[24,25] In addition to PD-L1 expression, interferon-γ (IFN-γ) mRNA levels may predict increased survival in patients with NSCLC treated with anti−PD-L1 therapy.[26]

New biomarkers are being discovered and developed at a rapid rate. Tissue-based screening for established and newly developed biomarkers has led to competition for patient tissue samples, leaving limited samples for research.[27] In lung cancer trials, there are typically multiple biomarker assays, such as immunohistochemistry (IHC), gene expression profiling, and analysis of tumor mutational burden, to compete for those limited cancer tissue samples.[28] There is, therefore, an unmet need for reliable biomarker assays using limited cancer tissues or those of suboptimal quality. The scarcity of biological specimens, the suboptimal quality of collected material, and the need for use of a broad spectrum of testing modalities make clinical diagnostic and biomarker studies challenging.[4,29,30] These studies would benefit from using archives of formalin-fixed, paraffin-embedded (FFPE) samples, accelerating the generation of reliable biomarker data. In addition, these samples may not need renewed patient consent for use.[31,32] Previous studies have shown that archived FFPE NSCLC samples are reliable for assessment of biomarkers, such as PD-L1 expression.[33,34]

The EdgeSeq system (HTG Molecular Diagnostics, Inc., Tucson, AZ), a novel genomics analysis platform, combines quantitative nuclease protection assay technology[35] with next-generation sequencing. Libraries are generated using samples such as FFPE tissues, negating the requirement for nucleic acid extraction and potentially mitigating the challenges of sample acquisition. This method can, therefore, reduce sample input requirements, as libraries typically generated from 5 to 10 mm^2 of sections (5 μm thick) can deliver high-quality gene expression profiling data.

EdgeSeq has been used previously to classify FFPE lung cancer samples on the basis of mRNA and miRNA signatures,[20,36] and may provide an opportunity to tap into archives where tissue samples are small or of suboptimal quality. In this study, the potential of EdgeSeq to exploit archives of tissue that had been previously used for pathologic diagnostic methods, such as hematoxylin and eosin (H&E) staining or IHC, was tested. The use of this technology for gene expression profiling of smaller lung cancer samples by exploring the lower limit of FFPE sample requirement for this assay was also validated.

Materials and Methods

Tissue Samples

Anonymized FFPE tissue blocks from patients with NSCLC ($n = 45$) were obtained from the Oregon Health & Science University Knight Cancer Institute (Portland, OR) or TriStar Technology Group LLC (Rockville, MD). Small-cell lung cancer (SCLC) samples ($n = 4$) were purchased from Conversant Biologics, Inc. (Huntsville, AL) or Proteogenex (Inglewood, CA). FFPE muscle tissue was obtained from Proteogenex. RNA samples from fresh-frozen colorectal adenocarcinoma samples ($n = 9$) and matched FFPE blocks ($n = 9$) were obtained from Asterand Bioscience (Detroit, MI). Patient and sample characteristics can be found in Supplemental Tables S1, S2, S3, S4, and S5. Serial sections (5 μm thick) were used for H&E and PD-L1 IHC staining along with unstained slides. IHC included deparaffinization, antigen retrieval, peroxidase blocking, anti−PD-L1 labeling, and visualization using the PD-L1 IHC 28-8 pharmDx assay (Dako, an Agilent Technologies, Inc., company, Santa Clara, CA; catalog number SK005).[37,38] Serial sections were also used for sample size titration experiments.

EdgeSeq Analysis

The EdgeSeq assay (HTG Molecular Diagnostics, Inc.) is designed to ensure fast and reliable transcriptomic analysis with optimized hands-on time and a turnaround time of as little as 36 hours. The workflow included extraction-free sample processing and quantitative nuclease protection assay using the EdgeSeq processor (HTG Molecular Diagnostics, Inc.), sequencing the resultant library, and data processing, including parsing, quality control, and normalization.

Sample Preparation

H&E- or IHC-stained sections were submerged in xylene for approximately 2 days at room temperature to remove the coverslip. The area of FFPE tissue was measured for subsequent input normalization and prepared using the Sample Prep Kit (HTG Molecular Diagnostics, Inc.). Tissues were scraped into a microfuge tube and lysed with lysis buffer

(HTG Molecular Diagnostics, Inc.; 28 μL per 12.5-mm^2 tissue). Tissue lysates were incubated for 10 minutes at 95°C. Proteinase K was added to a final concentration of 1 mg/mL, and samples were incubated for 2 hours at 50°C. Each sample (28 μL) was added to a well of a 96-well plate.

For the comparison of RNA from fresh-frozen and FFPE tissue, DNA and RNA were coextracted from FFPE colorectal adenocarcinoma slides by Reniguard Life Sciences (Exton, PA) using an AllPrep DNA/RNA FFPE kit (Qiagen, Germantown, MD; catalog number 80234).

Library Generation and Sequencing

The nuclease protection step was performed using the EdgeSeq processor. The EdgeSeq Oncology Biomarker Panel (OBP) assay (HTG Molecular Diagnostics, Inc.; catalog number 916-002-008) contained 2560 gene-specific nuclease protection probes, flanked by sequences to enable barcoding. The OBP probe set allowed analysis of tumor gene expression in 24 key oncology groups and pathways, including cell cycle, DNA repair, angiogenesis, and immuno-oncology, as well as signaling pathways, such as epidermal growth factor (EGF)/platelet-derived growth factor, EGFR/human epidermal growth factor receptor 2 (HER2), Janus kinase/signal transducer and activator of transcription (JAK/STAT), hedgehog (HH), mitogen-activated protein kinase (MAPK), and wingless int-1 (WNT).[39] This panel also included key drug targets, such as EGFR, mitogen-activated protein kinase kinase (MEK), mesenchymal-epithelial transition factor (MET), programmed death-1 (PD-1), PD-L1, and mammalian target of rapamycin (MTOR) (full list available from HTG Molecular Diagnostics, Inc.).

Nuclease protection probes were added to the lysed samples, and probe/target RNA heteroduplexes were formed. S1 nuclease was then added to degrade non-hybridized probes and nontargeted RNA, followed by heat inactivation (85°C, 20 minutes, pH 9). Samples were then individually barcoded using a 16-cycle PCR to add adapters and molecular barcodes. Barcoded samples were individually purified using Agencourt AMPure XP beads (Beckman Coulter, Indianapolis, IN; catalog number A63880) and quantitated using a KAPA Library Quantification kit (Kapa Biosciences, Wilmington, MA; catalog number KK4873). The pooled library was sequenced on an Illumina MiSeq using 50 cycles of a V3 150-cycle kit (Illumina, San Diego, CA; catalog number MS-102-3001) with two index reads. PhiX (Illumina; catalog number FC-110-3001) was spiked into the library at 5% as an Illumina sequencing control.

Cleanup Optimization by SPRI

Agencourt AMPure XP is a solid-phase reversible immobilization (SPRI) paramagnetic bead technology used for library purification that specifically binds double-stranded DNA. The length of fragments that SPRI binds can be controlled by varying the bead amount that is added to the reaction. To optimize the bead/DNA ratio, a crude lysate from 5 mm^2 of a section (5 μm thick) of FFPE SCLC tissue was diluted twofold in lysis buffer to achieve input equivalents of 5, 2.5, 1.25, and 0.625 mm^2, and duplicate wells were run on the EdgeSeq processor. The same molecular barcode was used for all samples during the PCR step. Five PCRs were performed for each group. These were pooled after PCR, split into four subgroups, and purified using

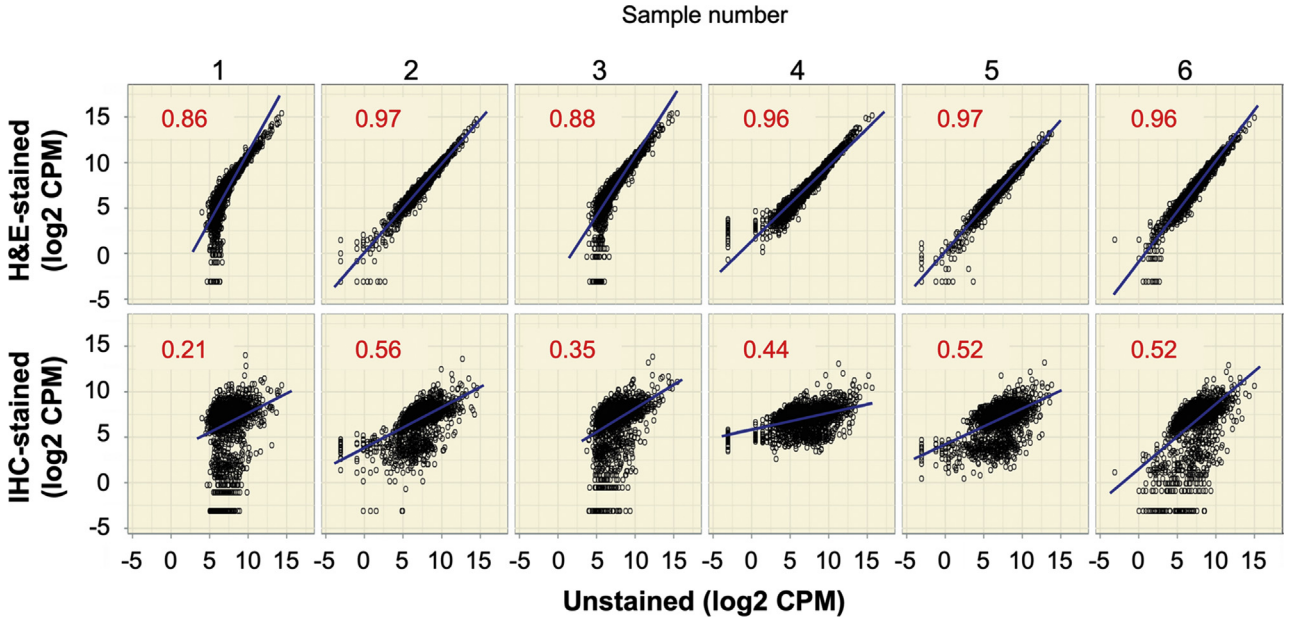

Figure 1 EdgeSeq Oncology Biomarker Panel analysis on unstained, hematoxylin and eosin (H&E)—stained, and immunohistochemistry (IHC)-stained formalin-fixed, paraffin-embedded non—small-cell lung cancer (samples 1 to 3) and small-cell lung cancer (samples 4 to 6) sections. Scatterplots show pairwise comparisons of gene expression data from six different unstained tissue samples with H&E- or IHC-stained sections. Numbers in red are Pearson correlation coefficients. **Blue lines** are from analysis by linear regression. CPM, counts per million.

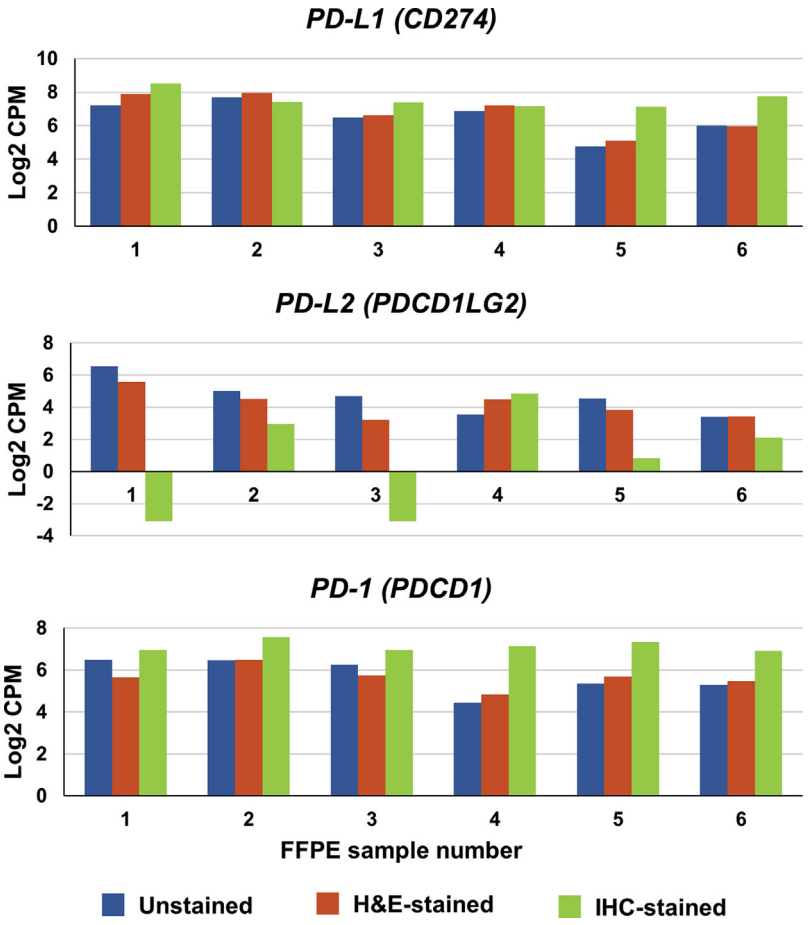

Figure 2 Expression of programmed death ligand 1 (PD-L1), programmed death ligand 2 (PD-L2), and programmed death-1 (PD-1) in libraries derived from six different unstained, hematoxylin and eosin (H&E)—stained, and immunohistochemistry (IHC)-stained formalin-fixed, paraffin-embedded (FFPE) non—small-cell lung cancer samples. CPM, counts per million.

AMPure XP (Beckman Coulter; catalog number A63880) and SPRIselect (Beckman Coulter; catalog number B23318) beads at four different bead/DNA volume ratios of 2.5, 2.0, 1.5, and 1.0. The purified libraries were analyzed by Kapa and Agilent Bioanalyzer with a High Sensitivity DNA Kit (Agilent Technologies, Inc.; catalog number 5067-4626).

EdgeSeq Sample Input Testing

A twofold serial dilution curve was performed on FFPE SCLC crude lysate in lysis buffer to the input equivalents from 5 to 0.08 mm^2 of a section (5 μm thick). Each sample was individually barcoded and purified by the SPRIselect beads with a bead/DNA ratio of 1.5.

EdgeSeq OBP Assay Data Analysis

Demultiplexed FASTQ files from the Illumina MiSeq were parsed by the EdgeSeq parser (HTG Molecular Diagnostics, Inc.) and aligned to the probe FASTQ files to collate the data. For each sample, the parser reported counts on 2568 probes (2559 unique genes plus 4 positive and 5 negative control probes). A statistical process quality control step filtered out those samples that failed on the basis of the negative control samples. In the analyses reported herein, 6 unstained, H&E-stained, and IHC-stained triplets, 42 unstained and H&E-stained pairs, and 7 SCLC dilutions were separately normalized. R software version 3.3.1 (R Foundation for Statistical Computing, Vienna, Austria; *https://cran.r-project.org/bin/windows/base/old/3.3.1*, last accessed January 24, 2019) was used for the normalization and analysis. The trimmed mean of M-values method implemented in the edgeR package was used for normalization, and the normalized counts were transformed to the log2 counts per million (CPM) scale.[40] Normalized gene expression values can be found in Supplemental Tables S6, S7, and S8. The R packages ggplot2 version 2.2.1[41] and GGally version 1.4.0 (R Foundation for Statistical Computing, *https://cran.r-project.org/web/packages/GGally/GGally.pdf*, last accessed January 24, 2019) were used to generate the scatter plots. A Pearson sample correlation was used as the measure of similarity, and linear regression as well as LOESS smoothing were used to assess the trends in the data.

Immune Scoring of NSCLC H&E Slides

Standard FFPE sections from 42 NSCLC specimens (including adenocarcinoma and squamous cell carcinoma) were stained with H&E, coverslipped, and digitally scanned at 40× objective with an Aperio digital scanner (Leica Biosystems, Wetzlar, Germany). Digital images were evaluated using a high-resolution Dell monitor using the Aperio

ImageScope (Leica Biosystems) digital slide viewer. The degree of immune cell infiltration within the tumor region was evaluated by a pathologist (S.P.K.) visually scanning the whole slide image and identifying the immune cell population (lymphocytes, macrophages, dendritic cells, plasma cells, neutrophils, and eosinophils) within tumors on the digitally scanned H&E-stained section. A score of 0 to 3 was assigned manually by a pathologist (S.P.K.) to record the percentage of infiltrating immune cells within the tumor region (0 indicates no immune cells present; 1, <1% of cells present; 2, 1% to 5% of cells present; and 3, >5% of cells present). Immune scoring assessments can be found in Supplemental Table S9.

IFN-γ Gene Expression Signature Scoring

The 25-gene list for the IFN-γ signature was obtained from Sharma et al.[42] Two gene expression matrices of size 25 × 42 containing normalized log2 CPM values for unstained and H&E-stained samples were prepared first. For each of the matrices, the log2 CPM values were transformed to their corresponding z-scores by standardizing each gene to a mean of 0 and an SD of 1. Next, for each of the samples, the z-scores were averaged across 25 genes, which resulted in an IFN-γ signature score for each unstained and H&E-stained sample. Unstained versus H&E-stained scores were plotted, and R^2 was computed by linear regression. The 42 samples were then reordered on the basis of their unstained IFN-γ signature score, and the z-score heat maps were prepared using a heat map function from the Nonnegative Matrix Factorization R package.[43] Finally, the box plots of IFN-γ signature score versus the pathologist-derived immune score, stratified by the unstained versus H&E-stained sample type, were plotted using the ggplot2 R package.[41] A Spearman rank-sum correlation test was used to assess the strength of association between the IFN-γ score and the immune score.

Results

EdgeSeq Analysis on Unstained, H&E-Stained, and IHC-Stained FFPE Slides

EdgeSeq OBP analysis was performed on matched unstained, H&E-stained, and IHC-stained slides derived from three different NSCLC and three different SCLC FFPE samples. Pairwise comparisons of gene expression profiles were made between unstained and stained (H&E or IHC) sections across the entire panel of 2560 genes (Figure 1). Pearson correlation coefficients ranging from 0.86 to 0.97 were observed between unstained and H&E-stained sections. Sequence reads from IHC-stained slides were slightly attenuated, but met the minimum requirements for further analysis and showed little evidence of strand, sequence, or gene length bias. Comparisons between unstained and IHC-stained sections showed Pearson correlation coefficients of 0.21 to 0.56, with increased correlations observed among moderately and highly expressed genes, compared with genes with low expression. Exploratory work performed in parallel with colorectal cancer samples compared the results of EdgeSeq analysis on matched fresh-frozen and FFPE tissue. Pearson correlation coefficients (0.79 to 0.90) between fresh-frozen and FFPE tissue were similar to those seen with unstained and H&E-stained sample comparisons (Supplemental Figure S1).

In a more focused gene-wise comparison, expression levels of the NSCLC biomarker genes PD-L1 (*CD274*), programmed death ligand 2 (PD-L2; *PDCD1LG2*), and the immunotherapy target programmed death-1 (PD-1; *PDCD1*) were compared for the three types of sections. Expression patterns for these genes were similar between the unstained

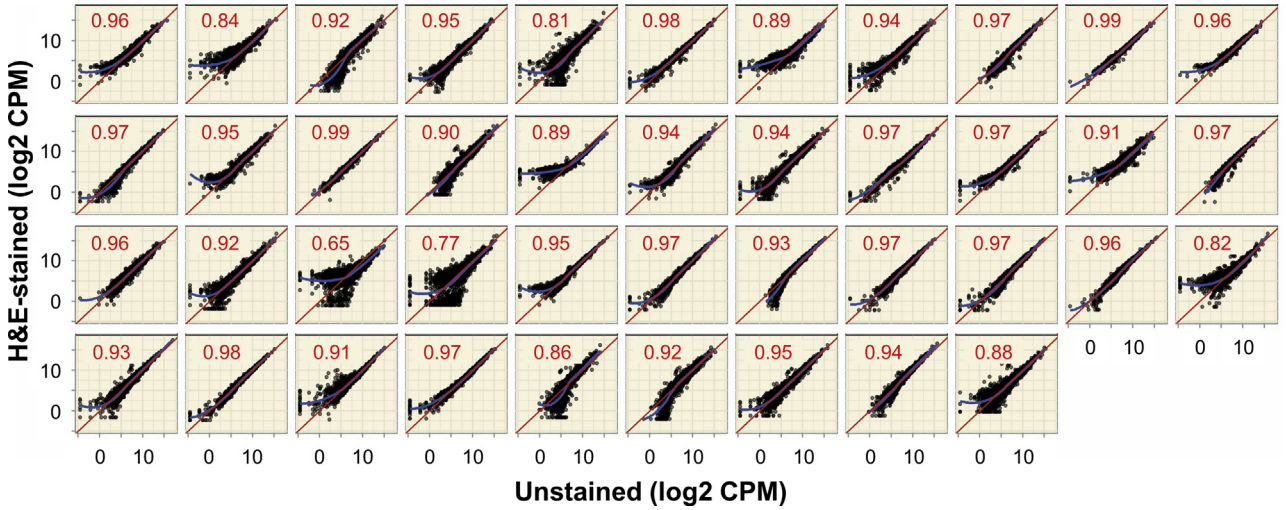

Figure 3 EdgeSeq Oncology Biomarker Panel analysis on 42 pairs of unstained and hematoxylin and eosin (H&E)—stained formalin-fixed, paraffin-embedded non—small-cell lung cancer samples. Numbers in red are Pearson correlation coefficients. **Red lines** show the ideal linear behavior for a 1:1 relationship between stained and unstained samples; **blue lines**, LOESS curves. CPM, counts per million.

FFPE and matched H&E slides across all six samples, but different in matched IHC slides (Figure 2). For programmed death ligand 2 gene expression, for example, the maximum difference (log2 CPM) between unstained and H&E-stained samples was 1.5, whereas it was 9.6 between unstained and IHC-stained samples.

After the pilot study comparing six tumor samples, the comparison between unstained and H&E-stained sections was extended to a further set of 42 FFPE NSCLC samples. Pairwise comparison of unstained and H&E-stained samples confirmed high correlation in gene expression data (Figure 3). Of the 48 comparisons made between unstained and H&E-stained sections (Figures 1 and 3), 90% (43 of 48 pairs) showed Pearson correlation coefficients $\geq$0.85.

IFN-γ Signature Analysis Using EdgeSeq Gene Expression on Unstained and H&E-Stained FFPE Slides and Their Correlation with Immune Score

The reproducibility of targeted gene expression data was tested by comparing expression levels in the 25-gene IFN-γ signature that was previously shown to be associated with response to immunotherapy.[42] Correlations between unstained and H&E-stained samples in *CXCL9* and *CXCL10* expression were 0.88 and 0.85, respectively, across the 42 NSCLC samples. Genes with expression levels below or close to the limit of detection showed weaker correlations (Supplemental Figure S2).

Heat maps of these 25 genes were generated from the normalized gene expression profiles of 42 paired unstained and H&E-stained FFPE NSCLC samples (Figure 4A). Samples were segregated into high, medium, or low IFN-γ score tertiles, on the basis of the average z-score for expression across 25 genes. The 14 unstained samples with high IFN-γ scores were also scored as high in the matched H&E-stained samples. Of the 14 samples, 9 were concordant for IFN-γ scores in both the medium and low groups.

Numerical IFN-γ signature scores were assigned to each of the unstained and H&E-stained samples. Pairwise comparisons of scores from the unstained and H&E-stained samples showed a Pearson correlation coefficient of 0.92 (Figure 4B), suggesting that H&E-stained IFN-γ

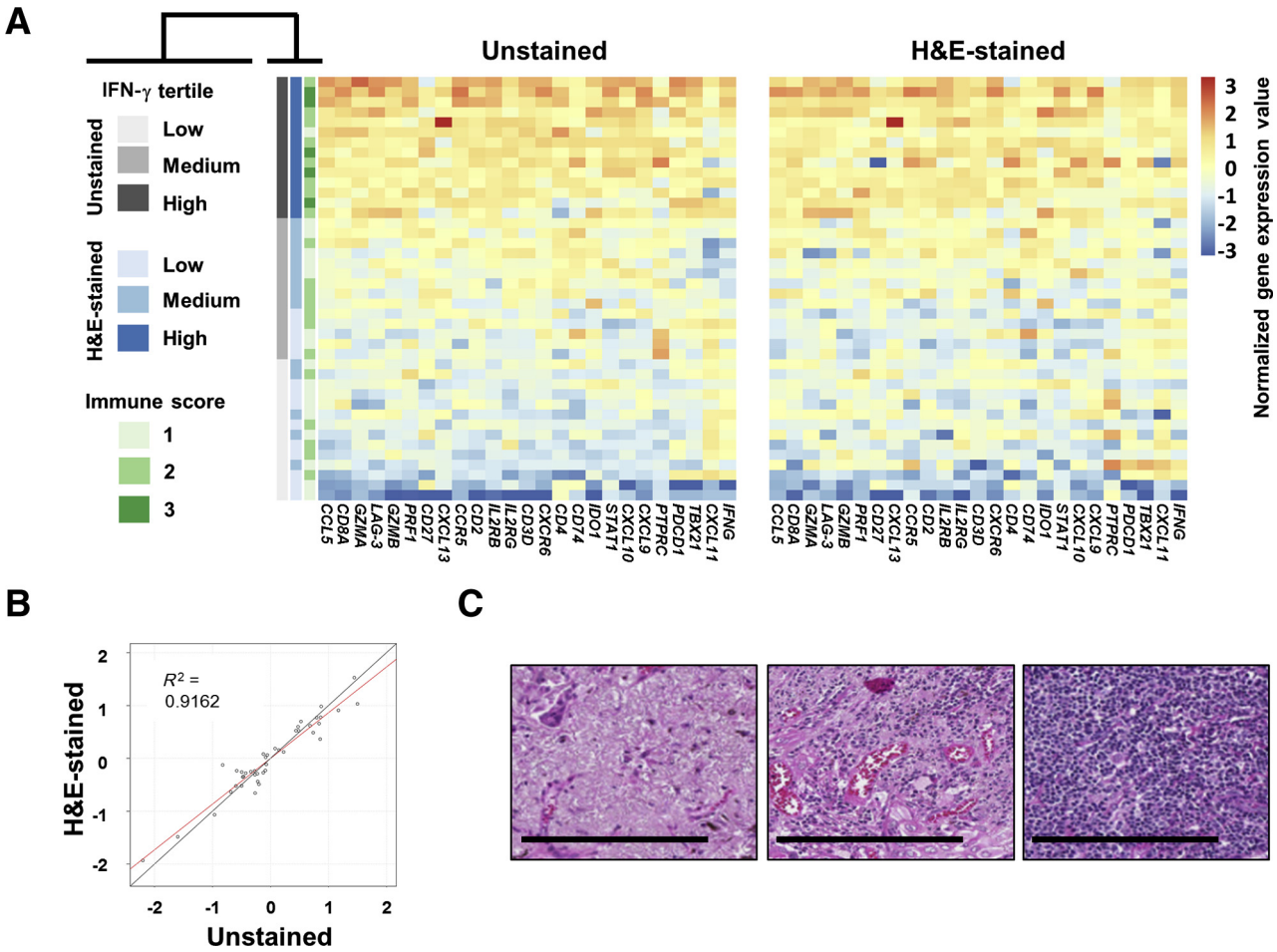

Figure 4 **A:** Heat maps of normalized gene expression for unstained versus hematoxylin and eosin (H&E)–stained slides; there are 42 samples and 25 genes; samples are ordered on the basis of the interferon-γ (IFN-γ) score. **B:** Unstained versus H&E-stained IFN-γ score. **Black line** shows the ideal linear behavior for a 1:1 relationship between stained and unstained samples; **red line**, linear regression analysis for this data set. **C:** Representative H&E images for samples with immune score of 1 (**left panel**), 2 (**middle panel**), or 3 (**right panel**). Scale bars = 200 μm (**C**).

Table 1 Distribution of IFN-γ Score versus a Pathologist-Derived Immune Score

Immune score	Median IFN-γ score (lower to upper quartile)	
	Unstained	H&E stained
1	−0.27 (−0.07 to −0.49)	−0.24 (0.02 to −0.34)
2	−0.10 (0.58 to −0.29)	−0.08 (0.56 to −0.45)
3	0.80 (1.17−0.68)	0.77 (0.91−0.62)
Correlation estimate (*P* value)	0.4847 (0.0011)	0.4552 (0.0025)

The median (lower quartile to upper quartile), correlation estimate, and *P* value from Spearman correlation test of association between IFN-γ score and immune score are shown. Additional information on immune scoring is available in Supplemental Table S9.

H&E, hematoxylin and eosin; IFN-γ, interferon-γ.

signature scores can be used as a proxy for the unstained IFN-γ gene signature score in assessing gene signatures from tumor samples. The 42 H&E-stained FFPE NSCLC samples were then subjected to histologic evaluation by a pathologist (S.P.K.) to derive immune scores of 1, 2, or 3 for each sample (<1%, 1% to 5%, or >5% infiltrating immune cells, respectively) (Figure 4C). The Spearman correlation test showed a statistically significant correlation ($P = 0.0025$) between the immune scores and IFN-γ signature scores in the H&E-stained samples (Table 1), similar to the *P* values observed ($P = 0.0011$) for comparisons between the immune scores on the H&E-stained samples and the IFN-γ signature scores generated from unstained samples. Thus, our IFN-γ signature analysis suggests that gene expression analysis from H&E-stained sections can be used in place of that from unstained samples to identify and assess potential correlations with complex phenotypes, such as immune infiltrates.

EdgeSeq Analysis on Smaller FFPE Tissue Input

After library preparation and cleanup, the yield of DNA produced from FFPE samples was analyzed by the Agilent Bioanalyzer. Species containing quantitative nuclease protection assay probe sequences flanked by adapter primers are approximately 187 bp. However, libraries can be contaminated with adapter dimers, producing a peak of DNA at approximately 70 bp, corresponding to adapter dimers with no intervening quantitative nuclease protection assay probe (Figure 5A), resulting in low aligned read counts. The SPRI cleanup procedure was optimized on DNA produced from a twofold serial dilution from 5.0 to

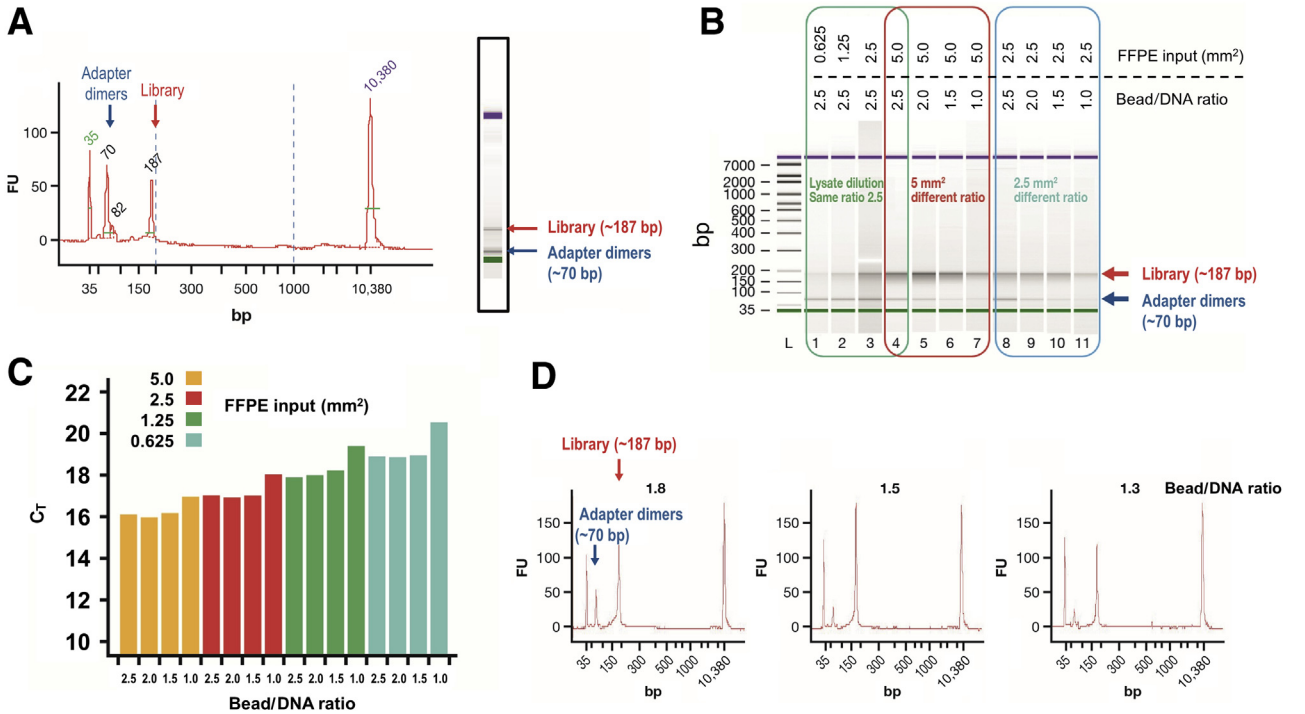

Figure 5 **A:** EdgeSeq library adapter dimer contamination detection. Agilent Bioanalyzer trace shows the yield of library (approximately 187 bp) and adapter (approximately 70 bp) dimers from the amplification procedure. Purple and green bands correspond to 10,380- and 35-bp standards, respectively, used to calibrate the trace. **B:** Bioanalyzer electrophoresis runs showing removal of adapter dimers from a library produced from 5.0 mm^2 (**red boxed area**) or 2.5 mm^2 (**blue boxed area**) formalin-fixed, paraffin-embedded (FFPE) small-cell lung cancer (SCLC) sample equivalent with solid-phase reversible immobilization (SPRI) bead/DNA ratios of 2.5, 2.0, 1.5, and 1.0. **Green boxed area** indicates the effect of sample dilution from 5.0 to 0.625 mm^2 FFPE sample equivalents. Purple and green bands correspond to 10,380- and 35-bp standards, respectively. **C:** Quantification of library yield by Kapa quantitative PCR, following production from FFPE sample inputs from 5.0 to 0.625 mm^2 and SPRI cleanup with bead/DNA ratios of 1.0:1 to 2.5:1. **D:** Further optimization of the SPRI cleanup procedure from a library produced from 5.0 mm^2 FFPE SCLC sample using bead/DNA ratios of 1.8, 1.5, and 1.3. FU, fluorescence unit.

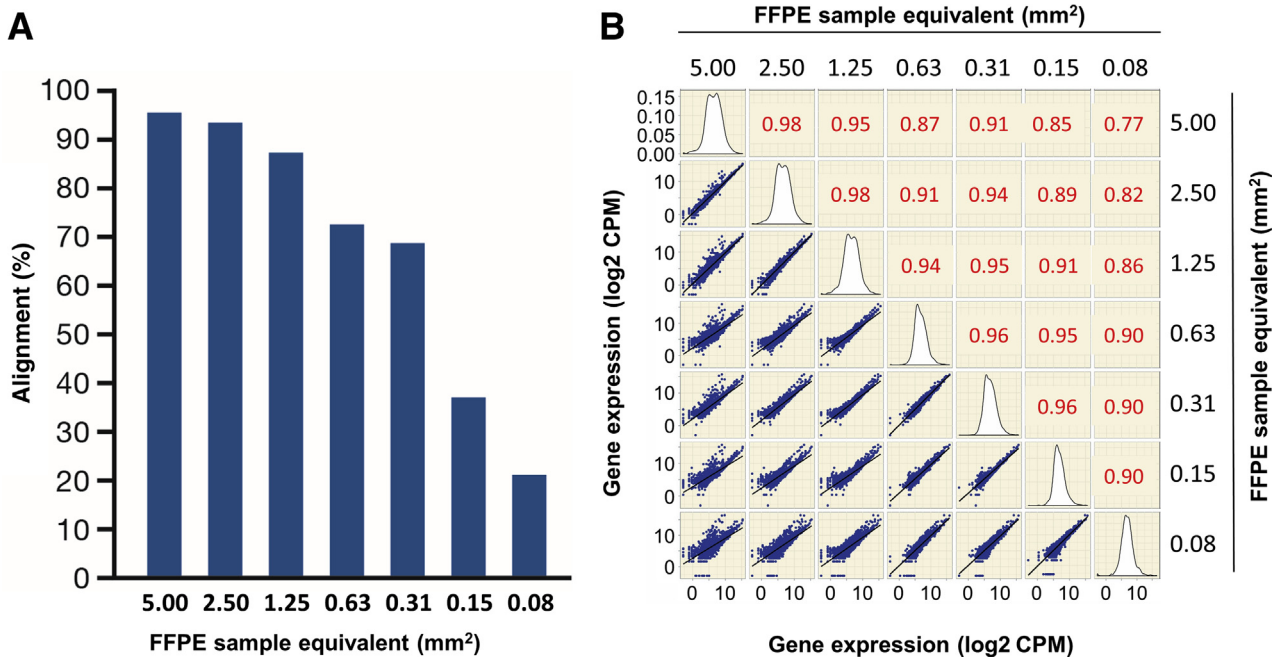

Figure 6 Testing the effect of decreasing sample input from 5.0 to 0.08 mm^2 tissue equivalents on next-generation sequencing (NGS) of libraries derived from formalin-fixed, paraffin-embedded (FFPE) small-cell lung cancer. **A:** Aligned rates of NGS data. **B:** Pairwise comparisons of gene expression data. Lower-left section shows scatterplots of the pairwise comparisons. **Black lines** are from linear regression analysis. Histograms in white denote the density of gene expression for each sample input. Upper-right section shows Pearson correlation coefficients in red. CPM, counts per million.

0.625 mm^2 FFPE SCLC sample equivalents, by adding SPRI beads in a bead/DNA ratio of 2.5:1, 2.0:1, 1.5:1, or 1.0:1 for each dilution. Adapter dimer removal was optimal at bead/DNA ratios of 1.5:1 and 1.0:1. However, using a bead/DNA ratio of 1.0:1 resulted in library loss (Figure 5B). Analysis of the products showed that decreasing sample input led to decreased library yield, as shown by a change in C_T from 16 to 19 (a low C_T corresponds to a high library concentration) (Figure 5C). Although similar C_T values were obtained at bead/DNA ratios of 2.5:1, 2.0:1, and 1.5:1, a ratio of 1.0:1 resulted in approximately a twofold to fourfold loss in library yield (Figure 5C). Thus, a bead/DNA ratio of 1.5:1 was observed to be optimal for library yield and elimination of adapter dimers.

Adapter removal by SPRI was further tested at bead/DNA ratios of 1.8:1 and 1.3:1. Figure 5D shows that a bead/DNA ratio of 1.8:1 resulted in poor removal of adapter dimers, and a bead/DNA ratio of 1.3:1 resulted in loss of library yield. A bead/DNA ratio of 1.5:1 was found to be optimal in reducing adapter dimer contamination while the library concentration remained unchanged.

The effect of decreasing sample input size on next-generation sequencing data quality was initially tested on libraries generated from FFPE muscle tissue ranging from 20 to 1 mm^2 in size. EdgeSeq analysis was possible on FFPE samples from 1 to 20 mm^2 with Pearson correlation coefficients ranging from 0.89 to 0.99 (Supplemental Figure S3). To examine the effect of decreasing sample input size even further, libraries generated from diluted SCLC samples equivalent to 5.0, 2.5, 1.25, 0.625, 0.31, and 0.08 mm^2 were sequenced. Next-generation sequencing data showed a good alignment rate (>69%) from 5.0 to 0.31 mm^2 FFPE sample equivalents, whereas libraries generated from 0.15 to 0.08 mm^2 showed alignment rates of <40% (Figure 6A). Pairwise comparison of gene expression levels showed a Pearson correlation coefficient >0.85 in samples equivalent to 5.0 to 0.31 mm^2, but lower correlations in sample input less than the equivalent of 0.31 mm^2 (Figure 6B).

Discussion

Tissue accessibility remains a limiting factor in biomarker analysis across many clinical trials.[27] In early immuno-oncology trials of patients with lung or head and neck cancer, FFPE samples from core needle biopsies were the predominant type of clinical specimen available. Remaining clinical samples are extremely valuable for multiple biomarker screens to study the tumor and its microenvironment, including IHC, gene expression, and assays for genomic signatures such as tumor mutational burden.[4,44–47] After multiple assays are performed, a single H&E-stained slide may be the only remaining sample from the patient.

Formalin fixation with paraffin embedding is the most common method of tissue preparation used in the clinic because it preserves tissue morphology and aids histopathologic subtyping.[48] There are now vast banks of FFPE

samples in worldwide tissue banks and pathology laboratories.[49] However, gene expression profiling from FFPE tissue is challenging because the fixation process can lead to RNA modification and/or degradation, which can inhibit downstream processes, such as RNA extraction, PCR, sequencing, or microarray analysis.[48,50,51]

Technologies that overcome the challenges associated with molecular subtyping of FFPE samples enable the exploitation of limited and archived tissue. The EdgeSeq system has been previously used to identify molecular signatures to aid subtype stratification of FFPE samples from several different tumors, including urothelial carcinoma,[42,52] small-cell bladder cancer,[53] NSCLC,[20,36,54] colorectal cancer,[55] and diffuse large B-cell lymphoma.[56] It is likely that this technology could be applied to biopsies from other types of tumors. EdgeSeq has also been used to generate data for biomarkers associated with response to immunotherapy and other drug treatments.[42,52,57]

In this study, we have demonstrated the feasibility of using EdgeSeq to generate reliable gene expression data from limited FFPE NSCLC and SCLC samples, even after staining with H&E. This introduces the possibility that EdgeSeq can be used for gene expression profiling of archival H&E-stained tissue, enabling investigators to fill a critical gap in the generation of biomarker data caused by limited supply of unstained samples. Our results show that EdgeSeq analysis delivers comparable data from fresh-frozen, unstained FFPE, or H&E-stained tumor samples. In contrast, a limitation to EdgeSeq technology is that it is not feasible for samples that have been used for IHC. A likely explanation for this is that steps in the IHC procedure, such as antigen retrieval, which can involve the slide being heated at boiling or subboiling temperatures in differing unmasking solutions, or hydrogen peroxide treatment to block endogenous peroxidase activity, can cause further deterioration of already fragmented RNA to a degree no longer suitable for EdgeSeq analysis.

Our analysis suggests that gene expression profiles from H&E-stained slides can be used as a meaningful proxy for profiles from unstained slides to assess gene expression signatures that may predict complex phenotypes, such as immune cell infiltration. For unstained versus H&E-stained comparisons, a Pearson correlation of $\geq$0.85 was observed in 43 of the 48 samples (90%) across the whole OBP, and a correlation R^2 value of 0.92 was observed for the probes used to calculate the IFN-γ score. Two of the samples in our analysis produced notably poor correlations across the OBP. Furthermore, some probes in the IFN-γ set showed consistently weaker correlations between unstained and H&E-stained sections at the patient level. It is possible that biological factors, such as tumor heterogeneity or cellularity, plus preanalytic factors, such as sample processing and/or storage, could contribute to variability in results.[51,58] It would be useful to build a statistical predictor of this correlation so that an additional quality control step could be applied to remove the weakly correlating samples or probe sets. Strong correlations were observed in those probes showing >8 log2 CPM. If necessary, data quality could be enhanced by applying a rigorous filter at this threshold on the basis of optimal lower limit of quantification.

Sequencing data quality was improved by minimizing adapter dimer contamination from the library. Quality control methods, such as the Bioanalyzer analysis, are important to identify the level of adapter dimer contamination before sequencing. The experiments described herein confirm that a bead/DNA ratio for SPRI of 1.5:1 is optimal for minimizing adapter dimer contamination in libraries produced from FFPE samples across a wide input range, while having minimal effects on library yield. The calculation of FFPE tissue area is semiquantitative and affected by factors such as tissue type, cell type and density, tumor content and uniformity, necrosis, and/or the degree of nucleic acid degradation. Also, preanalytical handling processes, such as scraping to transfer the tissue into the tube, can be challenging. The development of more-sensitive methods is needed for accurate quantification and control of preanalytical variables.

Gene expression profiling is routinely performed on tumor samples using other analysis methods, such as microarrays or multiplexed quantitative PCR. However, because these methods involve RNA extraction, substantially more starting material is often required (up to 10 mg of tissue or multiple sections).[49,59–61] It has been shown herein that the EdgeSeq platform (HTG Molecular Diagnostics, Inc.) can generate reliable expression data from as little as 1 mm^2 FFPE tissue and crude FFPE tissue lysates equivalent to surface areas as low as 0.31 mm^2 of a section (5 μm thick). This is approximately half the size of a small tissue microarray core,[62] and hundreds or even thousands of times smaller than that used for analysis by alternative methods.[49,59–61] Thus, the requirement for tissue with an equivalent surface area of at least 0.31 mm^2 does not represent a significant limitation of this method. However, smaller samples may be prone to increased variation, resulting from tissue or tumor heterogeneity.[63]

Together, these data show that scarce, small, and previously H&E-stained material could be used for gene expression analysis and finer regional follow-up of histopathologic features. Similarly, a single FFPE section could be macrodissected into different regions of interest (eg, normal versus tumor tissue, specific immune cell infiltrated regions, or tumor-stroma interface regions), followed by further gene expression analysis by EdgeSeq. The small sample input requirement for EdgeSeq analysis also means that this method could be conducted on tissue microarray sections and thus allow rapid gene expression analysis on hundreds to thousands of tumor specimens represented on tissue microarrays.

In summary, the EdgeSeq system can be considered as the platform of choice for the quick, reliable, and quantitative analysis of biological specimens of limited quantity or suboptimal quality. This technology enables research into gene expression patterns, tumor microenvironment, mechanisms of response and resistance, and immune cell engagement with tumors at the locoregional level.

Acknowledgments

We thank Qiuyan Wu (Bristol-Myers Squibb) for support and preparation of unstained, hematoxylin and eosin−stained, and immunohistochemistry-stained formalin-fixed, paraffin-embedded slides; and Bristol-Myers Squibb for funding the medical writing support and editorial assistance provided by Stuart Rulten, Ph.D., and Jay Rathi, M.A., of Spark Medica Inc., in accordance with Good Publication Practice guidelines.

Z.Q., J.C., and P.V. conceived and directed the project; M.S. and B.L. performed EdgeSeq analysis; L.W. optimized the protocol; K.D. analyzed gene expression; S.P.K. performed immune scoring.

Supplemental Data

Supplemental material for this article can be found at *http://doi.org/10.1016/j.jmoldx.2019.04.011*.

References

1. Coco S, Truini A, Vanni I, Dal Bello MG, Alama A, Rijavec E, Genova C, Barletta G, Sini C, Burrafato G, Biello F, Boccardo F, Grossi F: Next generation sequencing in non-small cell lung cancer: new avenues toward the personalized medicine. Curr Drug Targets 2015, 16:47−59
2. Siegel RL, Miller KD, Jemal A: Cancer statistics, 2017. CA Cancer J Clin 2017, 67:7−30
3. Mukhopadhyay S: Utility of small biopsies for diagnosis of lung nodules: doing more with less. Mod Pathol 2012, 25 Suppl 1: S43−S57
4. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger K, Yatabe Y, Ishikawa Y, Wistuba I, Flieder DB, Franklin W, Gazdar A, Hasleton PS, Henderson DW, Kerr KM, Petersen I, Roggli V, Thunnissen E, Tsao M: Diagnosis of lung cancer in small biopsies and cytology: implications of the 2011 International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society classification. Arch Pathol Lab Med 2013, 137: 668−684
5. Korpanty GJ, Graham DM, Vincent MD, Leighl NB: Biomarkers that currently affect clinical practice in lung cancer: EGFR, ALK, MET, ROS-1, and KRAS. Front Oncol 2014, 4:204
6. Gridelli C, Ardizzoni A, Barberis M, Cappuzzo F, Casaluce F, Danesi R, Troncone G, De Marinis F: Predictive biomarkers of immunotherapy for non-small cell lung cancer: results from an Experts Panel Meeting of the Italian Association of Thoracic Oncology. Transl Lung Cancer Res 2017, 6:373−386
7. Masucci GV, Cesano A, Eggermont A, Fox BA, Wang E, Marincola FM, Ciliberto G, Dobbin K, Puzanov I, Taube J, Wargo J, Butterfield LH, Villabona L, Thurin M, Postow MA, Sondel PM, Demaria S, Agarwala S, Ascierto PA: The need for a network to establish and validate predictive biomarkers in cancer immunotherapy. J Transl Med 2017, 15:223
8. Hellmann MD, Callahan MK, Awad MM, Calvo E, Ascierto PA, Atmaca A, Rizvi NA, Hirsch FR, Selvaggi G, Szustakowski JD, Sasson A, Golhar R, Vitazka P, Chang H, Geese WJ, Antonia SJ: Tumor mutational burden and efficacy of nivolumab monotherapy and in combination with ipilimumab in small-cell lung cancer. Cancer Cell 2018, 33:853−861
9. Hellmann MD, Ciuleanu TE, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, Minenza E, Linardou H, Burgers S, Salman P, Borghaei H, Ramalingam SS, Brahmer J, Reck M, O'Byrne KJ, Geese WJ, Green G, Chang H, Szustakowski J, Bhagavatheeswaran P, Healey D, Fu Y, Nathan F, Paz-Ares L: Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. N Engl J Med 2018, 378:2093−2104
10. Roy UB, Mantel S, Jacobson M, Ferris A: Willingness for multiple biopsies to improve quality of lung cancer care: understanding the patient perspective. J Thorac Oncol 2017, 12:S306−S307
11. Wiener RS, Schwartz LM, Woloshin S, Welch HG: Population-based risk for complications after transthoracic needle lung biopsy of a pulmonary nodule: an analysis of discharge records. Ann Intern Med 2011, 155:137−144
12. Ofiara LM, Navasakulpong A, Ezer N, Gonzalez AV: The importance of a satisfactory biopsy for the diagnosis of lung cancer in the era of personalized treatment. Curr Oncol 2012, 19:S16−S23
13. Solomon BJ, Mok T, Kim DW, Wu YL, Nakagawa K, Mekhail T, Felip E, Cappuzzo F, Paolini J, Usari T, Iyer S, Reisman A, Wilner KD, Tursi J, Blackhall F; Profile Investigators: First-line crizotinib versus chemotherapy in ALK-positive lung cancer. N Engl J Med 2014, 371:2167−2177
14. Rosell R, Carcereny E, Gervais R, Vergnenegre A, Massuti B, Felip E, et al: Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. Lancet Oncol 2012, 13:239−246
15. Sequist LV, Yang JC, Yamamoto N, O'Byrne K, Hirsh V, Mok T, Geater SL, Orlov S, Tsai CM, Boyer M, Su WC, Bennouna J, Kato T, Gorbunova V, Lee KH, Shah R, Massey D, Zazulina V, Shahidi M, Schuler M: Phase III study of afatinib or cisplatin plus pemetrexed in patients with metastatic lung adenocarcinoma with EGFR mutations. J Clin Oncol 2013, 31:3327−3334
16. Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready NE, Chow LQ, Vokes EE, Felip E, Holgado E, Barlesi F, Kohlhaufl M, Arrieta O, Burgio MA, Fayette J, Lena H, Poddubskaya E, Gerber DE, Gettinger SN, Rudin CM, Rizvi N, Crino L, Blumenschein GR Jr, Antonia SJ, Dorange C, Harbison CT, Graf Finckenstein F, Brahmer JR: Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. N Engl J Med 2015, 373:1627−1639
17. Reck M, Rodriguez-Abreu D, Robinson AG, Hui R, Csoszi T, Fulop A, Gottfried M, Peled N, Tafreshi A, Cuffe S, O'Brien M, Rao S, Hotta K, Leiby MA, Lubiniecki GM, Shentu Y, Rangwala R, Brahmer JR; Keynote Investigators: Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. N Engl J Med 2016, 375:1823−1833
18. Wang W, Sun J, Li F, Li R, Gu Y, Liu C, Yang P, Zhu M, Chen L, Tian W, Zhou H, Mao Y, Zhang L, Jiang J, Wu C, Hua D, Chen W, Lu B, Ju J, Zhang X: A frequent somatic mutation in CD274 3'-UTR leads to protein over-expression in gastric cancer by disrupting miR-570 binding. Hum Mutat 2012, 33:480−484
19. Brahmer J, Reckamp KL, Baas P, Crino L, Eberhardt WE, Poddubskaya E, Antonia S, Pluzanski A, Vokes EE, Holgado E, Waterhouse D, Ready N, Gainor J, Aren Frontera O, Havel L, Steins M, Garassino MC, Aerts JG, Domine M, Paz-Ares L, Reck M, Baudelet C, Harbison CT, Lestini B, Spigel DR: Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. N Engl J Med 2015, 373:123−135
20. Girard L, Rodriguez-Canales J, Behrens C, Thompson DM, Botros IW, Tang H, Xie Y, Rekhtman N, Travis WD, Wistuba II, Minna JD, Gazdar AF: An expression signature as an aid to the histologic classification of non−small cell lung cancer. Clin Cancer Res 2016, 22:4880−4889
21. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, et al: Somatic mutations affect key pathways in lung adenocarcinoma. Nature 2008, 455:1069−1075
22. Kim ES, Herbst RS, Wistuba II, Lee JJ, Blumenschein GR Jr, Tsao A, Stewart DJ, Hicks ME, Erasmus J Jr, Gupta S, Alden CM, Liu S,

Tang X, Khuri FR, Tran HT, Johnson BE, Heymach JV, Mao L, Fossella F, Kies MS, Papadimitrakopoulou V, Davis SE, Lippman SM, Hong WK: The BATTLE trial: personalizing therapy for lung cancer. Cancer Discov 2011, 1:44−53

23. Papadimitrakopoulou V, Lee JJ, Wistuba II, Tsao AS, Fossella FV, Kalhor N, Gupta S, Byers LA, Izzo JG, Gettinger SN, Goldberg SB, Tang X, Miller VA, Skoulidis F, Gibbons DL, Shen L, Wei C, Diao L, Peng SA, Wang J, Tam AL, Coombes KR, Koo JS, Mauro DJ, Rubin EH, Heymach JV, Hong WK, Herbst RS: The BATTLE-2 study: a biomarker-integrated targeted therapy study in previously treated patients with advanced non−small-cell lung cancer. J Clin Oncol 2016, 34:3638−3647
24. Li J, Li ZN, Yu LC, Bao QL, Wu JR, Shi SB, Li XQ: Association of expression of MRP1, BCRP, LRP and ERCC1 with outcome of patients with locally advanced non-small cell lung cancer who received neoadjuvant chemotherapy. Lung Cancer 2010, 69:116−122
25. Olaussen KA, Dunant A, Fouret P, Brambilla E, Andre F, Haddad V, Taranchon E, Filipits M, Pirker R, Popper HH, Stahel R, Sabatier L, Pignon JP, Tursz T, Le Chevalier T, Soria JC; Ialt Bio Investigators: DNA repair by ERCC1 in non-small-cell lung cancer and cisplatin-based adjuvant chemotherapy. N Engl J Med 2006, 355:983−991
26. Higgs BW, Morehouse C, Streicher K, Rebelatto MC, Steele K, Jin X, Pilataxi F, Brohawn PZ, Blake-Haskins JA, Gupta AK, Ranade K: Relationship of baseline tumoral IFNγ mRNA and PD-L1 protein expression to overall survival in durvalumab-treated NSCLC patients. J Clin Oncol 2016, 34, Abstract 3036
27. Hassanein M, Callison JC, Callaway-Lane C, Aldrich MC, Grogan EL, Massion PP: The state of molecular biomarkers for the early detection of lung cancer. Cancer Prev Res (Phila) 2012, 5:992−1006
28. Aisner DL, Rumery MD, Merrick DT, Kondo KL, Nijmeh H, Linderman DJ, Doebele RC, Thomas N, Chesnut PC, Varella-Garcia M, Franklin WA, Camidge DR: Do more with less: tips and techniques for maximizing small biopsy and cytology specimens for molecular and ancillary testing: the University of Colorado experience. Arch Pathol Lab Med 2016, 140:1206−1220
29. Goossens N, Nakagawa S, Sun X, Hoshida Y: Cancer biomarker discovery and validation. Transl Cancer Res 2015, 4:256−269
30. Waddell N, Cocciardi S, Johnson J, Healey S, Marsh A, Riley J, da Silva L, Vargas AC, Reid L; kConFab Investigators, Simpson PT, Lakhani SR, Chenevix-Trench G: Gene expression profiling of formalin-fixed, paraffin-embedded familial breast tumours using the whole genome-DASL assay. J Pathol 2010, 221:452−461
31. Cambon-Thomsen A: The social and ethical issues of post-genomic human biobanks. Nat Rev Genet 2004, 5:866−873
32. Olson EM, Lin NU, Krop IE, Winer EP: The ethical use of mandatory research biopsies. Nat Rev Clin Oncol 2011, 8:620−625
33. Herbst RS, Baas P, Perez-Gracia JL, Felip E, Kim DW, Han JY, Molina J, Kim JH, Arvis CD, Ahn MJ, Majem M, Fidler MJ, Surmont V, De Castro G Jr, Garrido M, Shentu Y, Dolled-Filhart M, Im E, Garon EB: P2.41 (also presented as PD1.06): pembrolizumab vs docetaxel for previously treated NSCLC (KEYNOTE-010): archival vs new tumor samples for PD-L1 assessment: track: immunotherapy. J Thorac Oncol 2016, 11:S242−S243
34. Midha A, Sharpe A, Scott M, Walker J, Shi K, Ballas M, Garassino M, Rizvi M: PD-L1 expression in advanced NSCLC: primary lesions versus metastatic sites and impact of sample age. J Clin Oncol 2016, 34, Abstract 3025
35. Qi Z, Wang L, He A, Ma-Edmonds M, Cogswell J: Evaluation and selection of a non-PCR based technology for improved gene expression profiling from clinical formalin-fixed, paraffin-embedded samples. Bioanalysis 2016, 8:2305−2316
36. Harpole D, Bueno R, Richards W, Beer D, Ballman K, Tsao M, Shepard F, Merrick D, Van Bokhoven A, Franklin W, Govindan R, Watson M, Gandara DR, Chen G, Chen Z, Chirieac L, Chui H, Genova C, Joshi M-B, Kowalewski A, Onaitis M, Rivard C, Sporn T, Hirsch FR: Creation and early validation of prognostic miRNA signatures for squamous cell lung carcinoma by the SPECS Lung Consortium. J Thorac Oncol 2017, 12:S612−S613
37. U.S. FDA: PD-L1 IHC 28-8 pharmDx (SK005): 50 tests for use with Autostainer Link 48. Silver Spring, MD: U.S. Food & Drug Administration, 2015. Available at https://www.accessdata.fda.gov/cdrh_docs/pdf15/P150025c.pdf (accessed June 14, 2019)
38. Phillips T, Simmons P, Inzunza HD, Cogswell J, Novotny J Jr, Taylor C, Zhang X: Development of an automated PD-L1 immunohistochemistry (IHC) assay for non-small cell lung cancer. Appl Immunohistochem Mol Morphol 2015, 23:541−549
39. Reinholz MM, Thompson DM, Botros I, Rounseville M, Roche PC: Abstract 1383: NGS-based measurement of gene expression of 2560 oncology-related biomarkers in formalin-fixed, paraffin-embedded (FFPE) tissues. Cancer Res 2016, 76:1383
40. Robinson MD, McCarthy DJ, Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010, 26:139−140
41. Wickham H: ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer-Verlag, 2016
42. Sharma P, Retz M, Siefker-Radtke A, Baron A, Necchi A, Bedke J, Plimack ER, Vaena D, Grimm MO, Bracarda S, Arranz JA, Pal S, Ohyama C, Saci A, Qu X, Lambert A, Krishnan S, Azrilevich A, Galsky MD: Nivolumab in metastatic urothelial carcinoma after platinum therapy (CheckMate 275): a multicentre, single-arm, phase 2 trial. Lancet Oncol 2017, 18:312−322
43. Gaujoux R, Seoighe C: A flexible R package for nonnegative matrix factorization. BMC Bioinformatics 2010, 11:367
44. Leidner R, Kang H, Haddad R, Segal NH, Wirth LJ, Ferris RL, Hodi FS, Sanborn RE, Gajewski TF, Sharfman W, McDonald D, Srivastava S, Gu X, Phillips P, Passey C, Seiwert T: Preliminary efficacy from a phase I/II study of the natural killer cell−targeted antibody lirilumab in combination with nivolumab in squamous cell carcinoma of the head and neck. J Immunother Cancer 2016, 4:91, Abstract O95
45. Dietel M, Bubendorf L, Dingemans AM, Dooms C, Elmberger G, Garcia RC, Kerr KM, Lim E, Lopez-Rios F, Thunnissen E, Van Schil PE, von Laffert M: Diagnostic procedures for non-small-cell lung cancer (NSCLC): recommendations of the European Expert Group. Thorax 2016, 71:177−184
46. Kayser G, Csanadi A, Otto C, Plones T, Bittermann N, Rawluk J, Passlick B, Werner M: Simultaneous multi-antibody staining in non-small cell lung cancer strengthens diagnostic accuracy especially in small tissue samples. PLoS One 2013, 8:e56333
47. Carbone DP, Reck M, Paz-Ares L, Creelan B, Horn L, Steins M, Felip E, van den Heuvel MM, Ciuleanu T-E, Badin F, Ready N, Hiltermann TJN, Nair S, Juergens R, Peters S, Minenza E, Wrangle JM, Rodriguez-Abreu D, Borghaei H, Blumenschein GR Jr, Villaruz LC, Havel L, Krejci J, Corral Jaime J, Chang H, Geese WJ, Bhagavatheeswaran P, Chen AC, Socinski MA: First-line nivolumab in stage IV or recurrent non−small-cell lung cancer. N Engl J Med 2017, 376:2415−2426
48. Guo Y, Wu J, Zhao S, Ye F, Su Y, Clark T, Sheng Q, Lehmann B, Shu XO, Cai Q: RNA sequencing of formalin-fixed, paraffin-embedded specimens for gene expression quantification and data mining. Int J Genomics 2016, 2016:9837310
49. Reis PP, Waldron L, Goswami RS, Xu W, Xuan Y, Perez-Ordonez B, Gullane P, Irish J, Jurisica I, Kamel-Reid S: mRNA transcript quantification in archival samples using multiplexed, color-coded probes. BMC Biotechnol 2011, 11:46
50. von Ahlfen S, Missel A, Bendrat K, Schlumpberger M: Determinants of RNA quality from FFPE samples. PLoS One 2007, 2:e1261
51. Srinivasan M, Sedmak D, Jewell S: Effect of fixatives and tissue processing on the content and integrity of nucleic acids. Am J Pathol 2002, 161:1961−1971
52. Galsky MD, Saci A, Szabo P, Wang L, Zhu J, Azrilevich A, Fischer BS, Necchi A, Siefker-Radtke AO, Sharma P: Fibroblast

growth factor receptor 3 (FGFR3), peroxisome proliferator-activated receptor gamma (PPARg), and outcomes with nivolumab (nivo) in metastatic urothelial cancer (UC). J Clin Oncol 2018, 36 6_suppl:511
53. Koshkin VS, Dhawan A, Hu M, Reynolds J, Elson P, McKenney J, Saunders LR, Isse K, Rashid S, Balyimez A, Ornstein MC, Gilligan TD, Lee B, Scott JG, Rini BI, Garcia JA, Grivas P, Mian OY: Correlation between gene expression and prognostic biomarkers in small cell bladder cancer (SCBC). J Clin Oncol 2018, 36 15_suppl:4546
54. Tang X, Lu W, Zhang J, Berhens C, Parra ER, Wineman J, Zhang J, Gibbons DL, Koeppel M, Kerns B, Stern M, Sepesi B, Lee JJ, Wistuba II: Abstract 4054: gene expression difference (GED) revealed immune function gene down-regulation as tumor-associated inflammatory cell (TAIC) infiltration in microenvironment in non-small cell lung cancer (NSCLC). Cancer Res 2018, 78 Suppl:4054
55. Sunakawa Y, Wang E, Roberts C, Yang D, Liu Q, Thompson D, Botros I, Moran M, Astrow SH, Hsiang J, Zhang W, Stintzing S, Tsuji A, Takahashi T, Denda T, Takeuchi M, Fujii M, Nakajima T, Ichikawa W, Lenz H-J: Association of gene signature to identify molecular subtypes with clinical outcomes of 1st-line cetuximab (cet) treatment for metastatic colorectal cancer (mCRC). J Clin Oncol 2016, 34 15_suppl:3592
56. Reinholz MM, Thompson D, Botros I, Sportmann P, Liu Q, Roche P, Xu-Monette Z, Young KH, LaFleur B: Next generation sequencing for DLBCL classification. J Clin Oncol 2016, 34 15_suppl:11559
57. Sunakawa Y, Yang D, LaFleur B, Luecke J, Thompson D, Moran M, Astrow SH, Hsiang J, Zhang W, Tsuji A, Takahashi T, Tanioka H, Negoro Y, Takagane A, Tani S, Yamaguchi T, Eto T, Fujii M, Ichikawa W, Lenz H-J: Association of immune-related genes to neutrophil-lymphocyte ratio (NLR) with survival of cetuximab treatment for metastatic colorectal cancer (mCRC): JACCRO CC-05/06AR. J Clin Oncol 2017, 35 15_suppl:11613
58. Chen H, Luthra R, Goswami RS, Singh RR, Roy-Chowdhuri S: Analysis of pre-analytic factors affecting the success of clinical next-generation sequencing of solid organ malignancies. Cancers 2015, 7:1699–1715
59. Akiyama Y, Kondou R, Iizuka A, Ohshima K, Urakami K, Nagashima T, Shimoda Y, Tanabe T, Ohnami S, Ohnami S, Kusuhara M, Mochizuki T, Yamaguchi K: Immune response-associated gene analysis of 1,000 cancer patients using whole-exome sequencing and gene expression profiling: Project HOPE. Biomed Res 2016, 37:233–242
60. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, Witteveen A, Glas A, Delahaye L, van der Velde T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH, Bernards R: A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002, 347:1999–2009
61. Mittempergher L, de Ronde JJ, Nieuwland M, Kerkhoven RM, Simon I, Rutgers EJ, Wessels LF, Van't Veer LJ: Gene expression profiles from formalin fixed paraffin embedded breast cancer tissue are largely comparable to fresh frozen matched tissue. PLoS One 2011, 6:e17163
62. Voduc D, Kenney C, Nielsen TO: Tissue microarrays in clinical oncology. Semin Radiat Oncol 2008, 18:89–97
63. Gyanchandani R, Lin Y, Lin HM, Cooper K, Normolle DP, Brufsky A, Fastuca M, Crosson W, Oesterreich S, Davidson NE, Bhargava R, Dabbs DJ, Lee AV: Intratumor heterogeneity affects gene expression profile test prognostic risk stratification in early breast cancer. Clin Cancer Res 2016, 22:5362–5369