Received: 24 January 2023 | Accepted: 26 January 2023

DOI: 10.1096/fba.2023-00009

RESEARCH ARTICLE

FASEB BioAdvances WILEY

# B cell-dependent subtypes and treatment-based immune correlates to survival in stage 3 and 4 lung adenocarcinomas

Susan Raju Paul[1,2,3] | Ivan Valiev[4] | Skylar E. Korek[1,2] | Vladimir Zyrin[4] | Diana Shamsutdinova[4] | Olga Gancharova[4] | Alexander Zaitsev[4] | Ekaterina Nuzhdina[4] | Diane L. Davies[5] | Ibiayi Dagogo-Jack[2,6] | Felix Frenkel[4] | Jessica H. Brown[4] | Joshua M. Hess[1] | Sarah Viet[7] | Jason L. Petersen[7] | Cameron D. Wright[5] | Harald C. Ott[5] | Hugh G. Auchincloss[5] | Ashok Muniappan[5] | Toshihiro Shioda[3,6] | Michael Lanuti[5] | Christel M. Davis[7] | Erik A. Ehli[7] | Yin P. Hung[3,8] | Mari Mino-Kenudson[3,6,8] | Maria Tsiper[4] | Ann E. Sluder[1,2] | Patrick M. Reeves[1,2,3] | Nikita Kotlov[4] | Alexander Bagaev[4] | Ravshan Ataullakhanov[4] | Mark C. Poznansky[1,2,3]

[1]Vaccine and Immunotherapy Center, Massachusetts General Hospital, Charlestown, Massachusetts, USA

[2]Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA

[3]Harvard Medical School, Boston, Massachusetts, USA

[4]BostonGene, Waltham, Massachusetts, USA

[5]Department of Thoracic Surgery, Massachusetts General Hospital, Boston, Massachusetts, USA

[6]Cancer Center, Massachusetts General Hospital, Boston, Massachusetts, USA

[7]Avera Institute of Human Genetics, Sioux Falls, South Dakota, USA

[8]Department of Pathology, Massachusetts General Hospital, Boston, Massachusetts, USA

## Abstract

Lung cancer is the leading cause of cancer-related deaths worldwide. Surgery and chemoradiation are the standard of care in early stages of non-small cell lung cancer (NSCLC), while immunotherapy is the standard of care in late-stage NSCLC. The immune composition of the tumor microenvironment (TME) is recognized as an indicator for responsiveness to immunotherapy, although much remains unknown about its role in responsiveness to surgery or chemoradiation. In this pilot study, we characterized the NSCLC TME using mass cytometry (CyTOF) and bulk RNA sequencing (RNA-Seq) with deconvolution of RNA-Seq being performed by Kassandra, a recently published deconvolution tool. Stratification of patients based on the intratumoral abundance of B cells identified that the B-cell rich patient group had increased expression of CXCL13 and greater abundance of $PD1^+$ CD8 T cells. The presence of B cells and $PD1^+$ CD8 T cells correlated positively with the presence of intratumoral tertiary lymphoid structures (TLS). We then assessed the predictive and prognostic utility of these cell types and TLS

**Abbreviations:** CS, cell suspension; CyTOF, cytometry by time of flight; DC, dendritic cell; FFPE, formalin-fixed paraffin-embedded; FFT, fresh frozen tissue; GEO, Gene Expression Omnibus; H&E, hematoxylin and eosin; ICI, immune checkpoint inhibitors; IgH, immunoglobulin heavy chain; IHC, immunohistochemistry; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; MDSC, myeloid-derived suppressor cell; MGH, Massachusetts General Hospital; NSCLC, non-small cell lung cancer; OS, overall survival; PBMC, peripheral blood mononuclear cells; PD1, programmed cell death protein-1; PD-L1, programmed cell death ligand-1; PFS, progression free survival; RNA-Seq, bulk RNA sequencing; TCGA, the cancer genome atlas; TIL, tumor infiltrating leukocytes; TLS, tertiary lymphoid structures; TME, tumor microenvironment; TPM, transcripts per million; TRA, T-cell receptor alpha; t-SNE, t-distributed stochastic neighbor embedding; VIC, samples from MGH.

Susan Raju Paul and Ivan Valiev contributed equally to this manuscript.

Mark C. Poznansky and Ravshan Ataullakhanov contributed equally to this manuscript.

**Correspondence**
Mark C. Poznansky, Vaccine and Immunotherapy Center, Massachusetts General Hospital, Bldg 149, 13th Street, Room 5.245, Charlestown, MA 02129, USA.
Email: mpoznansky@mgh.harvard.edu

Ravshan Ataullakhanov, BostonGene, Waltham, MA, USA.
Email: ravshan.ataullakhanov@bostongene.com

within publicly available stage 3 and 4 lung adenocarcinoma (LUAD) RNA-Seq datasets. As previously described by others, pre-treatment expression of intratumoral 12-chemokine TLS gene signature is associated with progression free survival (PFS) in patients who receive treatment with immune checkpoint inhibitors (ICI). Notably and unexpectedly pre-treatment percentages of intratumoral B cells are associated with PFS in patients who receive surgery, chemotherapy, or radiation. Further studies to confirm these findings would allow for more effective patient selection for both ICI and non-ICI treatments.

**KEYWORDS**
B cells, deconvolution, immunology, immunotherapy, NSCLC, TLS

## 1 | INTRODUCTION

Lung cancer remains the major cause of cancer-related deaths both globally and within the USA, accounting for 12% of all new cancers and 22% of cancer-related deaths.[1] Non-small cell lung cancer (NSCLC), the most common histopathological variant of lung cancer, accounts for 85% of the newly diagnosed cases each year and includes two predominant subtypes, adenocarcinoma and squamous cell carcinoma.[1] Lung adenocarcinoma (LUAD) accounts for about 40% of all lung cancer.[1] A combination of surgical resection, radiation, and chemotherapy is the standard of care in early stages of the disease, and immunotherapy improves patient outcomes in advanced stages of the disease.[2] The 5-year survival rate when combining all stages of NSCLC is among the lowest in cancers, even with treatment.[1] Hence it is crucial to identify better pre-treatment predictive and prognostic indicators that would allow for better selection of treatment plans for each patient based on their unique tumor microenvironment (TME).

Tumors expressing higher levels of programmed cell death ligand-1 (PD-L1) are known to be more responsive to treatment with immune checkpoint inhibitors (ICI); however, the predictive power of the commonly used FDA-approved biomarker for ICI response in NSCLC, PD-L1 expression identified by immunohistochemistry (IHC), is limited.[3,4] Multiple immune cell populations have been hypothesized as more robust predictors of response to ICI. Tumor infiltrating leukocytes (TIL), particularly CD8 T cells, are suggested in this context.[5–8] However, their location in the TME poses a challenge in using CD8 T cells as a reliable biomarker in routine clinical practice. A B-cell-derived gene expression signature has also been shown to be correlated with response to treatment with ICI in a study reported by Varn et al.[9] While there is literature on identifying responders to ICI, there is limited research on cellular and cytokine elements of the pre-treatment TME and whether they could be prognostic indicators of outcome in surgery, chemoradiation or immunotherapy. Understanding and developing pre-treatment immune markers and correlates will allow for better selection of patients for each treatment (surgery, chemoradiation, and/or immunotherapy) and could help improve the 5-year survival rate in lung cancer.

We proposed that an in-depth analysis of the TME would allow for a better understanding of immune subtypes of NSCLC and also identify pre-treatment prognostic indicators. While single-cell proteomic analysis of tumors provides this information, tissue dissociation alters the composition and functional state of the TME.[10–12] Alternatively, bulk RNA Sequencing (RNA-Seq) can be performed on fresh frozen and formalin-fixed paraffin-embedded (FFPE) tissues. Cell composition of the TME is then reconstructed using one of many cell deconvolution algorithms, which are often developed for specific tumor tissues and have limited capacity to identify cell subpopulations.[13,14]

In this pilot study, we evaluate the cell composition of the TME in a small cohort of NSCLC, using mass cytometry (CyTOF)[15–17] and by deconvolution of RNA-Seq using Kassandra, a recently published deconvolution tool,[18–20] and assess the predictive and prognostic utility of select cell types within publicly available datasets of LUAD. As previously described by others, the pre-treatment expression of intratumoral 12-chemokine tertiary lymphoid structure (TLS) gene signature is associated with progression free survival (PFS) in patients who receive treatment with ICI. Interestingly, in our study we found that pre-treatment percentages of intratumoral B cells in the TME are associated with PFS in patients who receive surgery, chemotherapy, or radiation. Validating these findings on a larger dataset would allow us to identify appropriate treatment options based on pre-treatment expression of the immune markers. Further studies would also be of benefit to look at prognostic indicators in surgery and chemoradiation, independently.

## 2 | METHODS

### 2.1 | Subjects and samples

The study was approved by the Institutional Review Board at Massachusetts General Hospital (MGH). Tumor tissue was obtained from patients diagnosed with NSCLC who underwent surgical resection at MGH between 2017 and 2019. To avoid metal contamination that would affect CyTOF analysis, active smokers and patients who underwent chemotherapy with cisplatin or had a recent MRI with contrast agent were excluded. Of the 56 samples collected, 11 had sufficient tissue for both CyTOF and RNA-Seq. CyTOF and RNA-Seq were performed on the same cell suspension (CS) for all 11 patients, and RNA-Seq alone was performed on fresh frozen tissue (FFT) for 9 of the 11 patients (Figure 1). For eight of the eleven samples we were able to procure hematoxylin and eosin (H&E) and IHC images of a portion of the tumor. Seven out of eleven samples utilized in this study were previously described by Zaitsev et al.[20] in validating the deconvolution tool, Kassandra.

### 2.2 | Tumor tissue processing

Resected tissue from each patient was immediately collected into a conical tube containing Leibowitz-15 media (Life Technologies Corporation; Grand Island, New York, USA). Dark black patches that appeared necrotic were removed, and the remaining tissue was weighed. The tissue was dissected into small pieces of approximately 30 mg weight using microdissection scissors. Random pieces of tissue weighing a minimum of 30 mg and a maximum of 90 mg were collected per patient and stored in aliquots of 30 mg/500 µL of RNA*later* (Invitrogen, Thermo Fisher Scientific; Waltham, Massachusetts, USA). These were used as FFT for RNA-seq. The remaining tissue was treated using the protocol referred to in Quatromoni et al.[21] Using microdissection scissors, the tissue was dissected into 1 $mm^3$ pieces and digested at 37°C for 1.5 h in L-15 media containing collagenase I, collagenase II, collagenase IV, elastase, and RNAse-free DNAse (Worthington Biochemical Corporation; Lakewood, New Jersey, USA; Roche Diagnostics; Germany) at 25 mL of digestion media per 0.5 mg of tissue. Digested samples were then passed through a 70 µm filter and homogenized using a plunger. Erythrocytes were lysed using multi-species red blood cell lysis buffer (eBioscience, Thermo Fisher Scientific; Waltham, Massachusetts, USA). Post-lysis, the cell suspension (CS) was passed through a 40 µm filter and counted using trypan blue staining. Cells (2.5 million) were separated for CyTOF labelling. Another 2.5 million cells were resuspended in RNA*later* at 500,000 cells/350 µL of RNA*later*. CS and FFT samples in RNA*later* were stored at 4°C overnight and frozen to −80°C the following day.

### 2.3 | Mass cytometry

#### 2.3.1 | Sample preparation

Pre-conjugated antibodies were obtained from Fluidigm Inc. (South San Francisco, CA) and the Lederer Lab at Brigham

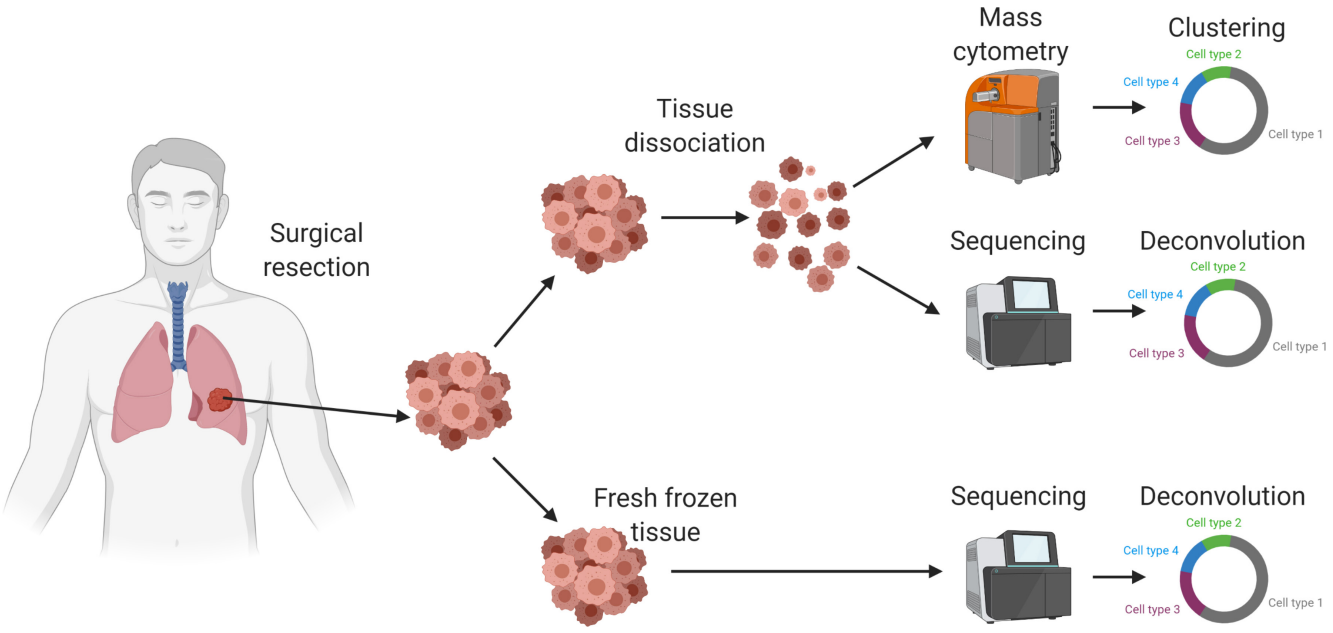


FIGURE 1 Overall schematic for sample collection and analysis.

and Women's Hospital (Boston, MA) (Table S1). A reference sample was created for the study to assess the day-to-day variation in staining using peripheral blood mononuclear cells (PBMC) isolated from a leukopak. The PBMCs were frozen in multiple aliquots and stored in liquid nitrogen in media containing 90% fetal bovine serum (FBS) and 10% dimethyl sulfoxide (DMSO). CS samples were labeled for CyTOF on the same day as surgical resection of the tumor. The reference sample was labeled with CD45 on 165Ho and was added into each tumor sample at a 1:20 ratio of reference-to-tumor cells. Then, cells were labeled for CyTOF with cisplatin (viability reagent) and surface antibodies, fixed with paraformaldehyde, permeabilized, and labeled with intracellular antibodies. The following day, the sample was intercalated with an iridium-based DNA intercalator as a cell identifier and data was acquired at a rate of 200–400 events per second on the Helios mass cytometer (Fluidigm Inc.; South San Francisco, CA). An unlabeled sample was also acquired to check for any metal contamination in the sample.

### 2.3.2 | Data analysis

The overall pipeline for the identification of major cell populations and subpopulations is shown in Figure S1. Detailed data analysis is included in Supplementary Methods. In brief, all 11 CyTOF samples were gated using FlowJo 10.0.7 (FlowJo LLC; Ashland, Oregon, USA) to remove the reference PBMCs and thereafter to identify viable cells from the patient (Figure S2A). Ten major cell populations were identified in each sample by FlowSOM[22] (version 1.20.0) clustering. All unlabeled cells and cells with non-specific annotation were grouped under "Other." CD90 was not labeled in VIC-017 and VIC-019; hence, fibroblasts and endothelial cells were not identified in these patients.

Subpopulations were defined for four major cell populations. Prior to this, batch effect correction was performed by quantile-based normalization within the individual cell populations. Post-normalization, each cell population was independently clustered using FlowSOM to identify 14, 12, 8, and 11 subpopulations for CD8 T cells, CD4 T cells, B cells, and macrophages/monocytes/dendritic cells (DC), respectively, each with a unique phenotypic profile.

## 2.4 | RNA sequencing

### 2.4.1 | RNA extraction, library preparation, and sequencing

Tumor CS samples were prepared using Rneasy Micro and/or Rneasy Plus Micro kits (Qiagen; Hilden, Germany). FFT samples (30 mg each) were prepared using the AllPrep DNA/RNA Mini kit (Qiagen; Hilden, Germany). Total RNA amounts and degradation values for CS and FFT were obtained using RNA 6000 Nano Chip run on a 2100 BioAnalyzer or RNA screening tape run on a 2200 TapeStation (Agilent; Santa Clara, California, USA). RNA-Seq libraries were prepared using the TruSeq RNA Library Prep Kit v2 (Illumina, Inc; San Diego, California, USA). Library products were purified by an AMPure XP (Beckman Coulter Genomics, Indianapolis, Indiana, USA) bead purification and enriched with standard polymerase chain reaction to create a final cDNA library. Final library quality control was carried out by evaluating the fragment size on a DNA1000 chip run on a 2100 BioAnalyzer or high sensitivity D1000 ScreenTape run on a 2200 TapeStation (Agilent; Santa Clara, California, USA). The concentration of each library was determined by quantitative PCR by the KAPA Library Quantification Kit for Next Generation Sequencing (KAPA Biosystems; Woburn, Massachusetts, USA) prior to sequencing.

Libraries were normalized to 2 nmol/L in 10 mM Tris-Cl, pH 8.5 with 0.1% Tween 20, then pooled evenly to yield approximately 100 million paired-end reads for each sample. The pooled libraries were denatured with 0.05 N NaOH and diluted to 20 pmol/L. Cluster generation of the denatured libraries was performed according to the manufacturer's specifications (Illumina, Inc; San Diego, CA) utilizing the HiSeq PE Cluster Kit v2 and/or v4 chemistry and flow cells. Libraries were clustered appropriately with a 1% PhiX spike-in. Sequencing-by-synthesis (SBS) was performed on a HiSeq2500 utilizing appropriate chemistry with paired-end 101 bp reads. Sequence-read data were processed and converted to FASTQ format for downstream analysis by Illumina BaseSpace analysis software, FASTQ Generation v1.0.0.

### 2.4.2 | Data analysis

Next-generation sequencing (NGS) quality control analysis was performed using FastQC (v0.11.5 http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), FastQ Screen[23] (v0.11.1), RSeQC[24] (v3.0.0) and MultiQC[25] (v1.6). Reads were aligned using Kallisto[26] (v0.46.0) to Gencode (v23). Noncoding, coding histones, and mitochondrial transcripts were removed, resulting in 20,062 transcripts for downstream analysis. Gene expressions were quantified as transcripts per million (TPM) with log2 transformation. Deconvolution of RNA-Seq was performed using Kassandra.[18–20]

MiXCR[27] (v2.1.7) was applied to quantify BCR/TCR repertoires from RNA-Seq. Post-processing of data was done on Python.

## 2.5 | TCGA and GEO datasets

Data from The Cancer Genome Atlas (TCGA) LUAD,[28] GSE135222 (Jung et al.[29]) and GSE126044 (Cho et al.[30]) were used in this manuscript. For the TCGA data, clinical data, including survival data, was downloaded from the GDC TCGA data portal[31] (MC3 dataset). Transcriptomic data were downloaded from the USCS XENA portal (https://xena.ucsc.edu/) as TPM units. Sample IDs were unified to patient IDs (first 12 characters). Patients with more than 1 tumor RNA-Seq sample or missing clinical annotation were removed. GSE135222 (Jung et al.[29]) and GSE126044 (Cho et al.[30]) were downloaded from Gene Expression Omnibus (GEO) as SRA archives. Annotation was downloaded for GEO, which included immunotherapy response data for GSE126044 (Cho et al.[30]). For GSE135222 (Jung et al.[29]), additional supplementary information that included survival and immunotherapy response data was extracted from the corresponding publication.

## 2.6 | Histopathology: H&E and IHC

For H&E staining, 5-micron-thick FFPE slides 5-micron were washed twice for 5 min each in xylene, 100% ethanol, and 95% ethanol, followed by a brief rinse in water and a 15-min incubation in hematoxylin. After a brief water rinse post-incubation, slides were dipped in 0.25% hydrochloric acid-ethanol, followed by brief dip in 1% lithium carbonate, a water rinse, and a 5-min incubation in eosin Y. Slides were then briefly rinsed twice each with 95% ethanol, 100% ethanol, and xylene and were cover-slipped.

IHC for CD3 and CD20 was performed on 5-micron-thick formalin-fixed paraffin-embedded whole tissue sections using the following antibodies (clone, catalog number, dilution, vendor): CD3 (polyclonal, A0452, 1:400, Dako, Agilent Technologies; Santa Clara, California, USA) and CD20 (monoclonal/L26, M0755, 1:140, Dako, Agilent Technologies; Santa Clara, California, USA), each using DIVA retrieval in the decloaker (Biocare Medical; Pacheco, California, USA) and a polymer detection system (Dako, Agilent Technologies; Santa Clara, California, USA), with counterstains using dab and hematoxylin.

### 2.6.1 | Data analysis

A pathologist conducted a blinded examination of FFPE H&E slides using Aperio ImageScope software (Aperio, Leica; Vista, CA, USA). H&E and IHC slides were manually aligned by the pathologist. Semi-quantitative lymphocytic and fibrotic scores were calculated using a modification of a previously described scoring system.[32–34] The 5-grade system comprises scores: 0 (characteristic is missing), 1 (minimal presence of a characteristic), 2 (characteristic is moderately pronounced), 3 (characteristic is pronounced), and 4 (characteristic is extremely pronounced). In IHC, the percentage of $CD20^+$ and $CD3^+$ cells were calculated in the tumor in VIC samples.

The presence of intratumoral TLS was assessed on H&E-stained slides and by IHC in full-face tumor sections using a previously published grading system.[35–38] On H&E, total amount of TLS included TLS in all 3 stages of development, namely, aggregates (clusters of lymphocytes), primary follicles (circles of lymphocytes with reticular stroma and without germinal center formation), or secondary follicles (primary follicles with germinal centers). On IHC, TLS were identified as dense round-shaped clusters of $CD20^+$ cells surrounded by $CD3^+$ cells.[37]

## 2.7 | Statistical analysis and data visualization

### 2.7.1 | Analysis

For statistical analyses and plotting, python (version 3.7.5) and R (version 4.0.2) were used. Data analysis in R was done using the following packages: dplyr, stringi, emmeans, arm, pwr, PRROC, car, and annotables. Median comparison was done using the Wilcoxon signed-rank test for paired samples and the Mann–Whitney test for unpaired samples. For survival analyses, Kaplan–Meier curves, log-rank tests, and Cox regression were performed using package Lifelines[39] (version 0.25.9). Spearman's test was used for correlation analysis, except in CyTOF data normalization when Pearson's correlation was used. For t-SNE plots, Barnes–Hut implementation in python package Multicore tSNE was used.[40] Generalized linear models were constructed by the binomial family with logit link function.

### 2.7.2 | Plotting

Figure 1 was created in BioRender (https://biorender.com/). Survival curves were drawn on Python using the package Lifelines (version 0.25.9). Figure S1 was drawn in Inkscape version 1.0. All other plots were created using python packages, matplotlib version 3.3.0, and seaborn version 0.9.0. Formatting was done in Inkscape for Linux version 1.0.

# 3 | RESULTS

## 3.1 | Subjects

To analyze the TME in NSCLC, we collected tissue samples of surgically resected tumors from 11 NSCLC patients (labeled as VIC). Six out of eleven patients were males, and the mean age of the VIC cohort was 70.8 years with a standard deviation of 9.6 years. Histopathological analysis revealed that nine out of eleven cases were lung adenocarcinoma while two out of eleven were lung squamous cell carcinomas (LUSC) (Table S2).

## 3.2 | Immune and non-immune components of the TME in NSCLC as identified by mass cytometry

CS samples were first analyzed by CyTOF. Viable single cells were identified through manual gating (Figure S2A) and clustered using FlowSOM to identify 11 major cell populations: $CD20^+$ B cells, $CD20^-$ B cells, CD4 T cells, CD8 T cells, NK cells, neutrophils, macrophages/monocytes/DC, myeloid-derived suppressor cells (MDSC), fibroblasts, endothelial cells, tumor, and others (expressing non-canonical markers). These major cell populations are represented in a t-distributed stochastic neighbor embedding (t-SNE) plot in Figure 2A. The expression of canonical markers within the t-SNE coordinates is shown in Figure 2B. The abundance of the major cell populations in the TME in each sample is shown in Figure 2C. B cells, CD4 T cells, CD8 T cells, and macrophages were independently clustered to identify subpopulations (Figure S2B–D).

No differences were observed in the abundance of the immune populations between LUAD (9/11) and LUSC (2/11) samples. In general, samples with greater abundance of B cells ($CD20^+$ and $CD20^-$) had lower abundance of innate cell population namely macrophage, monocyte, and dendritic cells (Figures 2C and S2D).

## 3.3 | Deconvolution of RNA-Seq data by Kassandra correlated with mass cytometry

In most studies involving human samples, fresh frozen or formalin fixed tumor samples are analyzed routinely by bulk RNA-Seq. To identify immune cells from this data, RNA-Seq data needs to be deconvolved. In our study, we used the recently published deconvolution tool, Kassandra.[18–20] In the context of overall correlation of all cell populations within the tumor, in addition to the cell populations assessed by Zaitsev et al.,[20] we also looked at B-cell subpopulations and separated monocytes from macrophages and dendritic cells. Samples with <75% viable cells in CyTOF (Figure S3A) were excluded in this comparative analysis, as dead cells are known to have increased non-specific binding of antibodies. The Kassandra-predicted cell percentage correlated to the abundance identified by CyTOF (Spearman correlation coefficient (rho) = 0.862, $p < 0.001$, Figures 2D,E and S3B). Additionally, in CS samples the Kassandra-predicted abundance of T cells and $CD20^+$ B cells correlated respectively with the expression levels of T-cell receptor alpha (TRA) and immunoglobulin heavy chain (IgH) determined by RNA-Seq (Spearman rho = 0.96 and 1 respectively, $p < 0.001$, Figure 2F).

## 3.4 | B-cell abundance in the TME stratified NSCLC into two subtypes

New evidence continues to emerge that B cells in the TME play a critical role in NSCLC.[9,41] Therefore, we assessed the B-cell composition of the VIC cohort. The samples in the cohort distinctly separated into two groups (B-cell rich and B-cell poor) based on the median abundance of B cells in CS (Figure 3A, right). We observed the same trend for these samples when their FFT RNA-Seq data was deconvolved (Figure 3A, left). B-cell rich samples in FFT had elevated levels of the transcripts for CXCL13 and of $PD1^+$ CD8 T cells ($p < 0.07$ and $p < 0.2$ respectively, Figure 3B, upper panel). CXCL13 is a chemokine that is produced predominately by $PD1^+$ CD8 T cells[42] (Figure S4).

Since the VIC dataset was small ($n = 11$), we decided to see if these subtypes were also seen in larger publicly available datasets of NSCLC. LUAD samples from the TCGA dataset ($n = 479$) showed a similar trend. Similar to VIC samples, samples with higher abundance of B cells from the TCGA dataset had elevated levels of transcripts for CXCL13 and $PD1^+$ CD8 T cells ($p < 0.001$, Figure 3B, lower panel). The percentage of $PD1^+$ CD8 T cells predicted by Kassandra correlated well with the expression of CXCL13 in both VIC and TCGA datasets (Spearman's rho = 0.709 to 0.999, $p < 0.001$, Figure 3B).

## 3.5 | Predictive and prognostic correlates in LUAD

IHC staining of VIC samples with anti-CD20 showed clusters of $CD20^+$ cells (Figure 4A). Based on this observation and the differential abundance of B cell, $PD1^+$ CD8 T cell and transcripts for CXCL13 in the VIC cohort and the publicly available datasets, we decided to evaluate for tertiary lymphoid structures (TLS) which are abundant in B cells and T cells.[43,44] TLS were identified in FFT

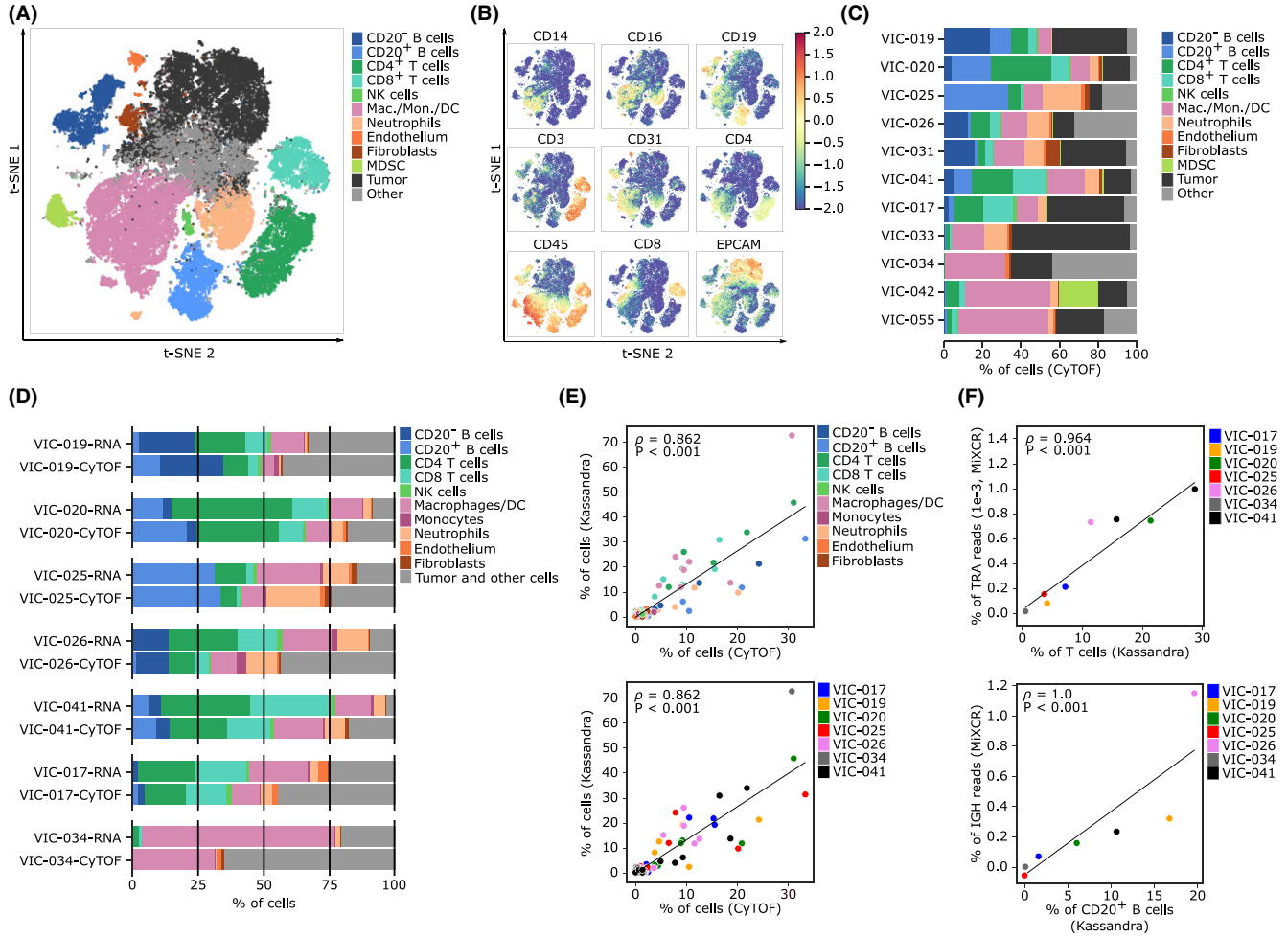


**FIGURE 2** (A–C) CyTOF-based immune profiles of VIC samples ($n = 11$). (A) t-SNE plot depicting cell populations identified by FlowSOM. (B) t-SNE map with signal intensity of canonical markers. (C) Stacked bar graph of major cell population percentages in 11 tumor samples. (D–F) Validation of Kassandra deconvolution on RNA-Seq using CyTOF and histopathology of VIC samples ($n = 7$). (D) Comparative stacked bar graph representation of the cellular composition of 7 CS CyTOF and CS RNA-Seq samples. CyTOF data is represented as a percent of viable cells and deconvolved CS RNA-Seq data is represented as a percent of all cells. (E) Scatterplots depicting the percentages of cell populations derived from CyTOF and CS RNA-Seq. The upper plot is color coded for cell types, while the lower plot is color coded for patients. (F) Scatterplots depicting the expression of TRA reads against Kassandra deconvolved percentage of T cells (upper plot), and the percentage of IgH reads against Kassandra deconvolved percentage of $CD20^+$ B cells (lower plot) from CS RNA-Seq.

samples through H&E and IHC, and in RNA-Seq data using a previously described 12-chemokine TLS gene expression signature.[45,46] The TLS identified through H&E and IHC moderately correlated with the 12-chemokine TLS gene signature (rho = 0.6, $p > 0.05$, Figure 4B). Interestingly, there was a high correlation between the 12-chemokine TLS gene signature and the percentage of B cells (rho = 0.81, $p = 0.015$, Figure 4C), while there was a moderate correlation between $PD1^+$ CD8 T cells and the 12-chemokine TLS gene signature (rho = 0.595, $p > 0.05$, Figure 4C).

Recent literature points to an increasing role of B cells[9] and TLS[45,47] as robust predictors of response to ICI. Due to the small number of samples in our cohort, we could not assess this. Hence the 12-chemokine TLS gene signature, CD8 T cells, and B cells were assessed as a predictor of response to ICI in two GEO datasets (GSE135222 and GSE126044). When compared to known predictors of response to ICI[48–55] (Supplementary Table. 3), we found that the abundance of intratumoral CD8 T cells and 12-chemokine TLS gene signature, individually, performed just as well as a predictive marker of response to ICI (AUC = 0.78 and 0.79) in stage 3 and 4 LUAD, while the abundance of B cells in the TME was a poor predictor of response to ICI (Figures 5A and S5).

In addition to looking at the predictive power of these cell populations, we also performed a survival analysis based on the individual abundance of B cells, CD8 T cells, and 12-chemokine TLS gene signature within stage 3 and stage 4 LUAD. Survival analysis conducted on GSE135222

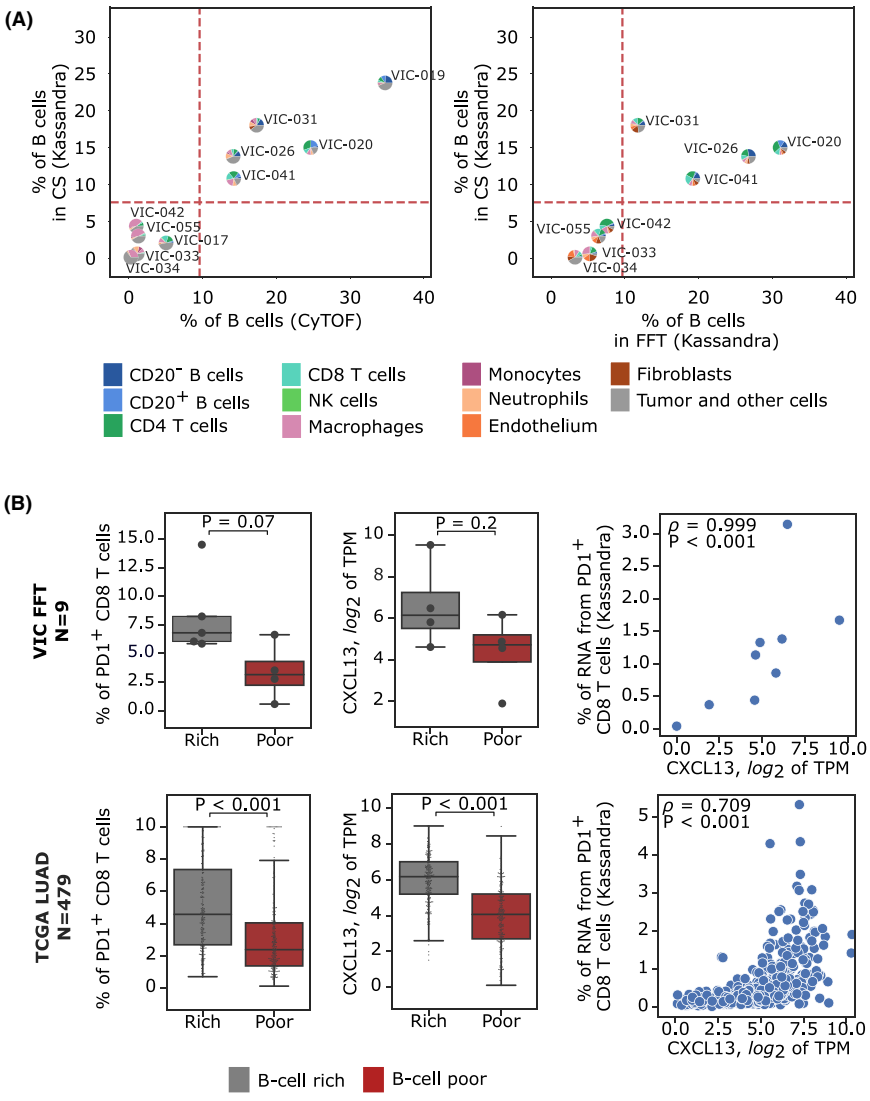


**FIGURE 3** Immune subtypes of NSCLC. (A) Scatterplots depicting the percentage of B cells in CyTOF (left) and the percentage of B cells in FFT RNA-Seq (right) against the percentage of B cells in Kassandra deconvolution of CS in VIC samples; patients stratified into B-cell rich and B-cell poor groups based on median abundance (red dotted line). Pie plots depict percentages of all cell populations as identified by CyTOF (left) and by Kassandra deconvolution of FFT RNA-Seq (right). (B) Percentage of $PD1^+$ CD8 T cells identified by Kassandra (left) and expression of CXCL13 identified by RNA-Seq (middle) in B-cell rich and B-cell poor groups in VIC FFT (first row) and TCGA LUAD (second row) datasets. Scatterplots show Spearman's correlation between the percentage of $PD1^+$ CD8 T cells and CXCL13 expression for each of the corresponding datasets (right).

indicated that the pre-treatment expression of intratumoral 12-chemokine TLS gene signature is associated with progression free survival (PFS) in patients who receive treatment with ICI for Stage 3 and 4 LUAD (Figures 5B and S6). Survival analysis conducted on the TCGA LUAD dataset indicated that the pre-treatment percentage of intratumoral B cells is associated with PFS in patients who receive surgery, chemotherapy or radiation, for Stage 3 and 4 LUAD, regardless of pre-treatment levels of immune checkpoint markers in tumor (Figures 5C and S7). When we looked at all stages of LUAD, we found a similar trend that the pre-treatment percentage of intratumoral B cells is associated with overall survival (OS) in patients who receive surgery, chemotherapy, or radiation (Figure S8).

## 4 | DISCUSSION

In this study, we describe the composition of the TME in a small cohort of patients with NSCLC using CyTOF and RNA-Seq. Intratumoral immune profiles described by CyTOF were then used to confirm predicted immune profiles generated through deconvolution of RNA-Seq using Kassandra, a recently described deconvolution algorithm with enhanced capabilities in identifying cell populations across all cancers. Subsequent analysis segregated the patients into two groups based on the abundance of intratumoral B cells, confirmed that TLS are comparable to known predictors of ICI response in NSCLC and identified that pre-treatment levels of B cells in patients treated with surgery and/or chemoradiation have associations to survival in LUAD.

While bulk RNA-Seq, which is routinely used in clinical research, provides an unbiased profile of the whole transcriptome, it does not provide data for single cells nor does it evaluate protein expression. Deconvolution algorithms enable us to deconvolve cell populations from RNA-Seq data; however, most of these algorithms have technical constraints in how the noise in the data is addressed and most are unable to deconvolve innate

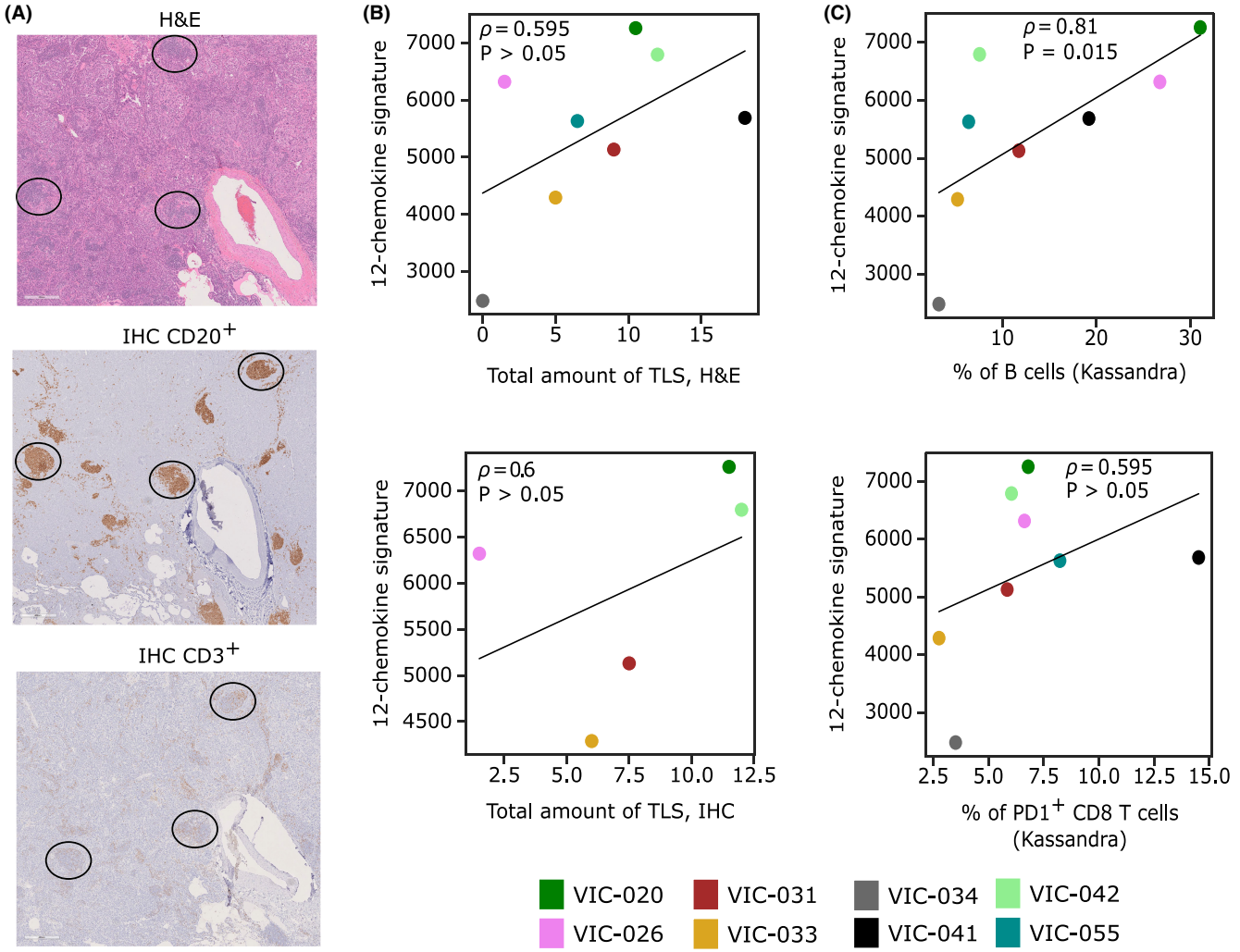


**FIGURE 4** Assessment of TLS as a predictive and prognostic marker. (A) Manually aligned regions of interest on H&E (top panel), $CD20^+$ IHC (middle) and $CD3^+$ IHC (bottom) for patient VIC-019, with TLS circled in each. (B) Scatterplots depicting correlation between 12-chemokine signature and TLS on H&E (top) and IHC (bottom) slides in VIC FFT samples. (C) Scatterplots depicting correlation between 12-chemokine signature and Kassandra deconvolved percentage of B cells (top) and $PD1^+$ CD8 T cells (bottom) in VIC FFT samples.

immune cell types.[13,14] Recently, the RNA-Seq deconvolution tool Kassandra was developed to deconvolve RNA-Seq data from most tumors and identify subpopulations of T cells, B cells, and monocytes.[18–20] The comparison of cell populations identified by CyTOF and Kassandra was carried out by creating a single cell suspension of the tumor which had to be processed fresh on the day of surgical resection. Consequently, CyTOF data required batch correction. Quantile-based normalization used routinely in next-generation sequencing data analysis was repurposed to remove batch variations in CyTOF. Furthermore, while performing this assessment we had to address inherent differences between the two techniques. For example, historically in cytometry data analysis, dead cells are eliminated from downstream analysis, as these cells often bind antibodies non-specifically. However, RNA from all cells, including dead cells, is assessed in RNA-Seq. Hence, comparison of results between both techniques required exclusion of samples with a high proportion of dead cells and debris. The study is limited by the low sample size and its heterogeneity. The number of samples included was constrained by the large quantity of tissue required and other inclusion criteria, which also extended the time to two years to procure the 11 samples used in the VIC cohort. In addition to the canonical cell populations assessed in the correlation of CyTOF to RNA-Seq by Zaitsev et al.,[20] we confirmed a good correlation between Kassandra and CyTOF in the VIC cohort when assessing B cell subpopulations namely $CD20^+$ and $CD20^-$ B cells, and when assessing monocytes separate from macrophages and dendritic cells (Figure 2D,E). For further validation of the deconvolved T cells and B cells, we also assessed the correlation of these populations to TCR and IgH which should only be associated with T cells and B cells, respectively. There was a high correlation between the variables (Figure 2F). Therefore,

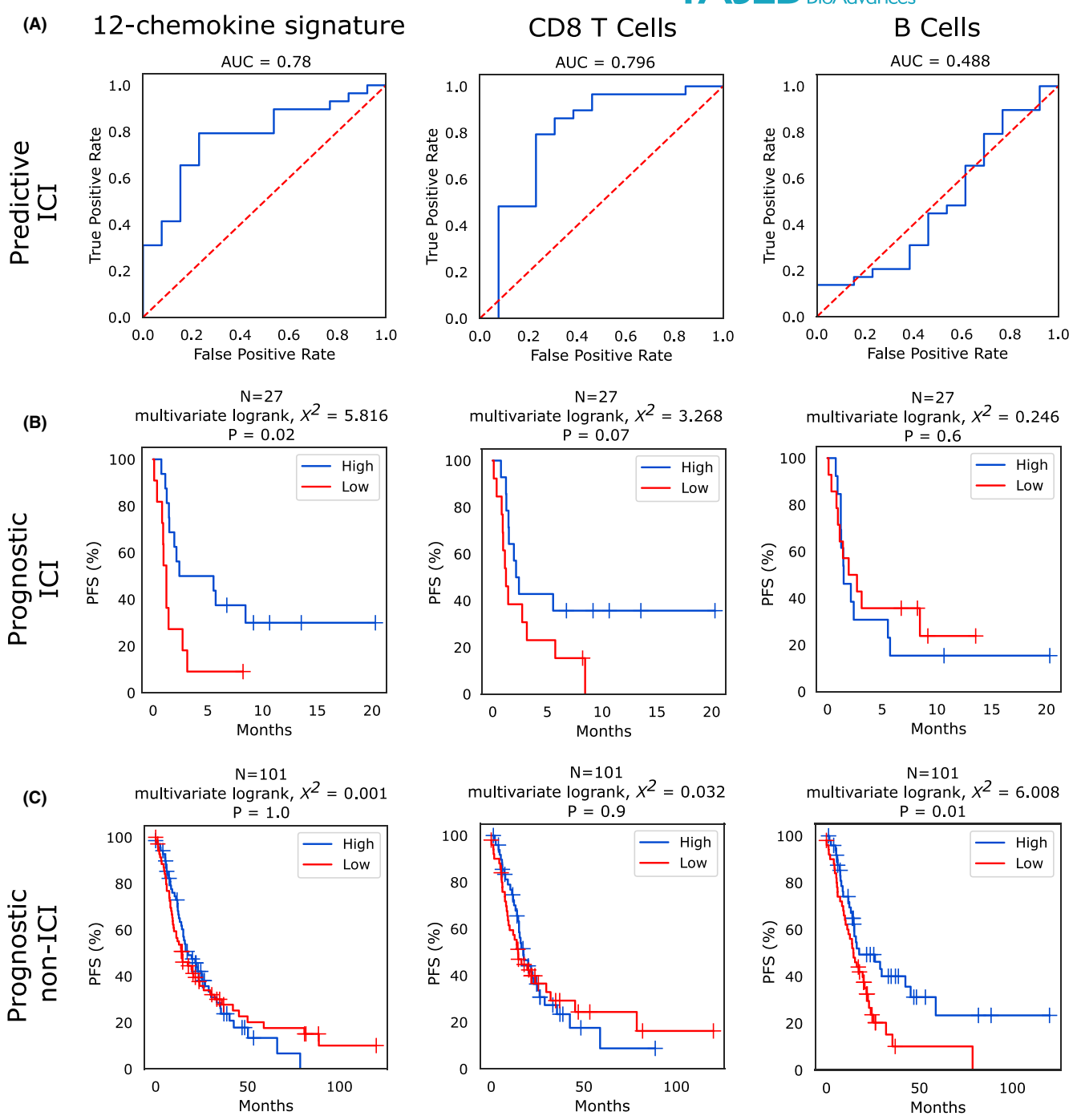


**FIGURE 5** (A) Receiver operating characteristics curves for 12-chemokine TLS gene signature, CD8 T cells, and B cells in predicting response to ICI, in stage 3 and 4 NSCLC, in GSE135222 and GSE126044 datasets ($n = 40$). (B) Kaplan–Meier survival curves (up to 20 months post diagnosis) when patients were divided into two groups based on the median of the 12-chemokine signature and the median percentage of B cells and CD8 T cells in the GSE135222 dataset. (C) Kaplan–Meier survival curves (up to 125 months post diagnosis) when patients were divided into two groups based on the median of the 12-chemokine signature and the median percentage of B cells and CD8 T cells in the TCGA LUAD dataset (stage 3 and stage 4 LUAD only).

we decided to continue to use Kassandra in deconvolving cells populations from the external datasets.

Informed by recent advances in the understanding of B cells in cancer,[9,56–59] we interrogated the VIC and external datasets for patterns in intratumoral B-cell distribution and identified a B-cell rich and B-cell poor immune subtype in NSCLC. Various studies have linked the increased abundance of B cells in the TME to improved OS in multiple cancers, including NSCLC.[56,57,60–62] Transcripts for CXCL13, a chemokine which is known to be a B-cell chemoattractant, was also higher in the B-cell rich subtype of NSCLC. Thommen et al.[42] reported

that intratumoral CD8 T cells expressing high levels of PD1 in tumors have a unique transcriptional phenotype and are one of the cell types that produce the chemokine, CXCL13. In our study we confirmed that there was a statistically significant correlation between the abundance of $PD1^+$ CD8 T cells and the expression of CXCL13, and that $PD1^+$ CD8 T cells and CXCL13 were higher in the B-cell rich subtype. CXCL13 and $PD1^+$ CD8 T cells also play a key role in the formation of TLS,[63] which among other cells also contain B cells. TLS, which are identified within several cancers, are typically associated with favorable outcomes.[45,64] In early-stage NSCLC, TLS are associated with a low rate of recurrence.[65] Additionally, in some cancers the presence of TLS correlates with a better response to neoadjuvant or immunotherapy.[66–72] In TLS, which are found primarily near tumor margins, antigen presentation results in B-cell development within germinal centers.[45,56,66,67,73] TLS consist of T cells, mature dendritic cells, and B cells, surrounded by high endothelial venules and TLS architecture is regulated by B cells.[43,44] There are a number of markers used in the identification of TLS, including a 12-chemokine TLS gene signature composed of CCL2, CCL3, CCL4, CCL5, CCL8, CCL18, CCL19, CCL21, CXCL9, CXCL10, CXCL11, and CXCL13.[45,46] In our study, we observed a statistically significant positive correlation between the 12-chemokine TLS gene signature and B cells, which is not surprising as CXCL13 is a gene included in the 12-chemokine TLS gene signature. Although there was a positive correlation between the 12-chemokine TLS gene signature and $PD1^+$ CD8 T cells, within our dataset this was not statistically significant. One possible explanation is that the 12-chemokine TLS gene signature takes into account all stages of TLS development while $PD1^+$ CD8 T cells secrete CXCL13, which functions primarily during the formation of TLS.

Through our assessment of publicly available datasets, we confirmed that the 12-chemokine TLS gene signature and CD8 T cells, regardless of the expression of PD1, are comparable to other known predictors of ICI response.[45,74,75] B cells, CD8 T cells, and TLS are increasingly being studied as predictors of ICI response.[45–47] TLS are associated with good clinical outcomes in most cancers, including NSCLC.[45,74,75] Stromal CD8 T cells are also suggested as strong predictors of response to ICI.[5–8,76] Our findings were similar in this regard. When we divided the patients in the ICI dataset based on their 12-chemokine TLS gene signature, B cells, or CD8 T cells, those with a higher expression of the 12-chemokine TLS gene signature had greater PFS. Similarly, association with improved PFS was also seen in patients with elevated CD8 T cells in their TME regardless of PD1 expression on cytotoxic T cells. However, using our analytical approach B cells were not identified as predictive or prognostic indicators for ICI.

Surgery, radiation, and chemotherapy are commonly used in treating NSCLC, especially in the early stages of therapy. The role of the TME or peripheral immune system in outcomes with these non-ICI treatments has not been widely studied. In gastric cancer, preoperative systemic inflammatory index developed from circulating counts of neutrophils, lymphocytes, and platelets indicated postoperative outcomes.[77] The preoperative neutrophil to lymphocyte ratio, and platelet to lymphocyte ratio, also correlates with outcome in gastric cancer.[78,79] Hence it is of interest to know how and if pre-treatment immune infiltration affects outcome when providing non-ICI treatment, namely surgery and/or chemoradiation. Ojlert et al.[80] has reported that a rich adaptive immune TME in LUAD indicated better prognosis after surgery. Helmink et al.[67] found that a higher abundance of B cells correlated with OS in Stage 3 melanoma. To our knowledge, our assessment is the first of its kind for non-ICI treatments, in LUAD. The TCGA LUAD dataset that we analyzed includes patients that received non-ICI treatment for LUAD. Using the RNA sequencing data from this cohort we identified that pre-treatment B cells in the TME correlate with improved PFS in all stages of LUAD.

In the VIC cohort, we found that the 12-chemokine TLS signature had a positive correlation to intratumoral B cells (Figure 4C). However, when we assessed the predictive and prognostic role of the 12-chemokine TLS signature and B cells in publicly available datasets, the 12-chemokine TLS signature and B cells did not track similarly (Figure 5). One possible reason for this is the difference in the abundance of $PD1^+$ CD8 T cells in the B-cell rich subtypes between the VIC cohort and the TCGA LUAD cohort (Figure 3B). $PD1^+$ CD8 T cells secrete the B-cell chemoattractant, CXCL13, which recruits B cells into the TME and plays a role in the formation of TLS. While all samples in the VIC cohort that belong to the B-cell rich subtype express higher levels of $PD1^+$ CD8 T cells, there are some samples in the TCGA LUAD B-cell rich subtype that express lower levels of $PD1^+$ CD8 T cells when compared to the B-cell poor subtype. This difference in the abundance of $PD1^+$ CD8 T cells could indicate that some of the TCGA samples have B cells in their TME that are not only recruited via TLS but are rather directly recruited into the TME. This could be a reason why the 12-chemokine TLS gene signature and percentage of B cells do not track similarly when we assessed them as predictive and prognostic indicators in our LUAD datasets.

In conclusion, we confirmed the presence of two immune subtypes of NSCLC based on the intratumoral abundance of B cells. TLS, which contain T cells and B cells, were then described based on a previously identified 12-chemokine TLS signature. As observed by others, CD8 T cells and TLS were comparable predictors of response

to ICI in publicly available datasets. In particular, our retrospective study of the TCGA dataset identified that pretreatment levels of B cells in LUAD could possibly serve as a new prognostic indicator of PFS for LUAD patients receiving surgery or chemoradiation regardless of the stage of the disease. This supports the established view that immune subpopulations could be used as prognostic indicators in effective patient selection for appropriate treatment (ICI, surgery, or chemoradiation) and that this is worthy of further exploration in larger-scale prospective studies.

## AUTHOR CONTRIBUTIONS

SRP, MT, AES, PMR, AB, RA, and MCP contributed to conception of the study and study design. SRP, SK, DD, SV, JLP, TS, CD, CW, HO, HA, AM, ML, YH, MM, and PMR contributed to data collection. SRP, IV, VZ, DS, OG, AZ, EN, FF, PMR, NK, AB, RA, and MCP contributed to data analysis and interpretation. SRP, IV, ID, JHB, JH, EAE, TS, YH, AES, PMR, NK, AB, RA, and MCP contributed to writing and editing of the manuscript. SRP, IV, JHB, AES, and MCP contributed to critical review of manuscript.

## ACKNOWLEDGMENTS

The authors would like to thank Nicole Brousaides and Heather Cahill for technical assistance in staining histopathology slides and in acquiring CyTOF data, respectively. The results shown in this manuscript are in part based upon data generated by the TCGA Research Network (https://www.cancer.gov/tcga).

## DISCLOSURES

Ivan Valiev, Vladimir Zyrin, Diana Shamsutdinova, Olga Gancharova, Alexander Zaitsev, Ekaterina Nuzhdina, Nikita Kotlov, Alexander Bagaev, and Ravshan Ataullakhanov are employees of BostonGene and inventors on patents related to this work. Jessica H. Brown, and Maria Tsiper are employees of BostonGene. Felix Frenkel was an employee of BostonGene at the time this study was conducted. Ibiayi Dagogo-Jack has received honoraria from American Lung Association, Aptitude Health, ASCO Post, Onc Live, DAVA Oncology, Creative Education Concepts, Total Health Conferencing and Medscape; is a consultant/advisor for AstraZeneca, Bayer, Bristol Myers Squibb, Boerhinger-Ingelheim, BostonGene, Genentech, Janssen, Novocure, Pfizer, Sanofi/Genzyme, Syros, Xcovery; and has received research funding from Array, Genentech, Pfizer, Novartis. Cameron D. Wright is a consultant for Bayer. Michael Lanuti is a consultant for AstraZeneca and Lovance Biotherapeutics. Yin P. Hung has received royalties from Elsevier. Mari Mino-Kenudson is a consultant for AstraZeneca, SANOFI, BMS and Janssen Oncology; and has received loyalty from Elsevier. Patrick M. Reeves holds patents through Massachusetts General Hospital, is a member of the ImmunoScape scientific advisory board, has received speaking honoraria from Fluidigm, and has previously consulted for Merck. Mark C. Poznansky holds patents through Massachusetts General Hospital, is a consultant to BostonGene and a scientific co-founder of Voltron Therapeutics Inc. All other authors declare no conflict of interest relating to this work.

## DATA AVAILABILITY STATEMENT

The data that support the findings will be freely available in the BostonGene repository at https://github.com/BostonGene following the date of publication.

## REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020;70(1):7-30. doi:10.3322/caac.21590
2. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature*. 2018;553(7689):446-454. doi:10.1038/nature25183
3. Brueckl WM, Ficker JH, Zeitler G. Clinically relevant prognostic and predictive markers for immune-checkpoint-inhibitor (ICI) therapy in non-small cell lung cancer (NSCLC). *BMC Cancer*. 2020;20(1):1-16. doi:10.1186/s12885-020-07690-8
4. Schoenfeld AJ, Rizvi H, Bandlamudi C, et al. Clinical and molecular correlates of PD-L1 expression in patients with lung adenocarcinomas. *Ann Oncol*. 2020;31(5):599-608. doi:10.1016/j.annonc.2020.01.065
5. Geng Y, Shao Y, He W, et al. Prognostic role of tumor-infiltrating lymphocytes in lung cancer: a meta-analysis. *Cell Physiol Biochem*. 2015;37(4):1560-1571. doi:10.1159/000438523
6. Wang L, Yue HU, Wang S, Shen J, Wang X. Biomarkers of immunotherapy in non-small cell lung cancer (review). *Oncol Lett*. 2020;20(5):1-14. doi:10.3892/OL.2020.11999
7. Durgeau A, Virk Y, Corgnac S, Mami-Chouaib F. Recent advances in targeting CD8 T-cell immunity for more effective cancer immunotherapy. *Front Immunol*. 2018;9:14. doi:10.3389/fimmu.2018.00014
8. Tumeh PC, Harview CL, Yearley JH, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature*. 2014;515(7528):568-571. doi:10.1038/nature13954
9. Varn FS, Wang Y, Cheng C. A B cell-derived gene expression signature associates with an immunologically active tumor microenvironment and response to immune checkpoint blockade therapy. *Onco Targets Ther*. 2019;8(1):1-10. doi:10.1080/2162402X.2018.1513440
10. Denisenko E, Guo BB, Jones M, et al. Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. *Genome Biol*. 2020;21(1):1-25. doi:10.1186/s13059-020-02048-6
11. van den Brink SC, Sage F, Vértesy Á, et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat Methods*. 2017;14(10):935-936. doi:10.1038/nmeth.4437
12. Adam M, Potter AS, Potter SS. Psychrophilic proteases dramatically reduce single-cell RNA-seq artifacts: a molecular atlas of kidney development. *Development (Cambridge)*. 2017;144(19):3625-3632. doi:10.1242/dev.151142

13. Becht E, Giraldo NA, Lacroix L, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol*. 2016;17(1):1-20. doi:10.1186/s13059-016-1070-5
14. Şenbabaoğlu Y, Gejman RS, Winer AG, et al. Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol*. 2016;17(1):1-25. doi:10.1186/s13059-016-1092-z
15. Bandura DR, Baranov VI, Ornatsky OI, et al. Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Anal Chem*. 2009;81(16):6813-6822. doi:10.1021/ac901049w
16. Leelatian N, Doxie DB, Greenplate AR, et al. Single cell analysis of human tissues and solid tumors with mass cytometry. *Cytometry B Clin Cytom*. 2017;92(1):68-78. doi:10.1002/cyto.b.21481
17. Reeves PM, Sluder AE, Paul SR, Scholzen A, Kashiwagi S, Poznansky MC. Application and utility of mass cytometry in vaccine development. *FASEB J*. 2018;32(1):5-15. doi:10.1096/fj.201700325R
18. Kotlov N, Bagaev A, Revuelta M, et al. Clinical and biological subtypes of B-cell lymphoma revealed by microenvironmental signatures. *Cancer Discov*. 2021;11(6):1468-1489. doi:10.1158/2159-8290.CD-20-0839
19. Luginbuhl AJ, Johnson JM, Harshyne LA, et al. Tadalafil enhances immune signatures in response to Neoadjuvant Nivolumab in Resectable head and neck squamous cell carcinoma. *Clin Cancer Res*. 2022;28(5):915-927. doi:10.1158/1078-0432.CCR-21-1816
20. Zaitsev A, Chelushkin M, Dyikanov D, et al. Precise reconstruction of the TME using bulk RNA-seq and a machine learning algorithm trained on artificial transcriptomes. *Cancer Cell*. 2022;40(8):879-894.e16. doi:10.1016/j.ccell.2022.07.006
21. Quatromoni JG, Singhal S, Bhojnagarwala P, Hancock WW, Albelda SM, Eruslanov E. An optimized disaggregation method for human lung tumors that preserves the phenotype and function of the immune cells. *J Leukoc Biol*. 2015;97(1):201-209. doi:10.1189/jlb.5ta0814-373
22. van Gassen S, Callebaut B, van Helden MJ, et al. FlowSOM: using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A*. 2015;87(7):636-645. doi:10.1002/cyto.a.22625
23. Wingett SW, Andrews S. Fastq screen: a tool for multi-genome mapping and quality control [version 1; referees: 3 approved, 1 approved with reservations]. *F1000Res*. 2018;7(May):1-13. doi:10.12688/f1000research.15931.1
24. Wang L, Wang S, Li W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics*. 2012;28(16):2184-2185. doi:10.1093/bioinformatics/bts356
25. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. 2016;32(19):3047-3048. doi:10.1093/bioinformatics/btw354
26. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 2016;34(5):525-527. doi:10.1038/nbt.3519
27. Bolotin DA, Poslavsky S, Mitrophanov I, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods*. 2015;12(5):380-381. doi:10.1038/nmeth.3364
28. Collisson EA, Campbell JD, Brooks AN, et al. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014;511(7511):543-550. doi:10.1038/nature13385
29. Jung H, Kim HS, Kim JY, et al. DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. *Nat Commun*. 2019;10(1):1-12. doi:10.1038/s41467-019-12159-9
30. Cho JW, Hong MH, Ha SJ, et al. Genome-wide identification of differentially methylated promoters and enhancers associated with response to anti-PD-1 therapy in non-small cell lung cancer. *Exp Mol Med*. 2020;52(9):1550-1563. doi:10.1038/s12276-020-00493-8
31. Ellrott K, Bailey MH, Saksena G, et al. Scalable Open Science approach for mutation calling of tumor exomes using multiple genomic pipelines. *Cell Syst*. 2018;6(3):271-281.e7. doi:10.1016/j.cels.2018.03.002
32. Lo RC, Kim H. Histopathological evaluation of liver fibrosis and cirrhosis regression. *Clin Mol Hepatol*. 2017;23(4):302-307. doi:10.3350/cmh.2017.0078
33. Fuchs TL, Sioson L, Sheen A, et al. Assessment of tumor-infiltrating lymphocytes using international TILs working group (ITWG) system is a strong predictor of overall survival in colorectal carcinoma: a study of 1034 patients. *Am J Surg Pathol*. 2020;44(4):536-544. doi:10.1097/PAS.0000000000001409
34. Li Y, Wei Y, Tang W, et al. Association between the degree of fibrosis in fibrotic focus and the unfavorable clinicopathological prognostic features of breast cancer. *PeerJ*. 2019;2019(11):1-18. doi:10.7717/peerj.8067
35. Murakami J, Shimizu Y, Kashi Y, et al. Functional B-cell response in intrahepatic lymphoid follicles in chronic hepatitis C. *Hepatology*. 1999;30(1):143-150. doi:10.1002/hep.510300107
36. Finkin S, Yuan D, Stein I, et al. Ectopic lymphoid structures function as microniches for tumor progenitor cells in hepatocellular carcinoma. *Nat Immunol*. 2015;16(12):1235-1244. doi:10.1038/ni.3290
37. Buisseret L, Desmedt C, Garaud S, et al. Reliability of tumor-infiltrating lymphocyte and tertiary lymphoid structure assessment in human breast cancer. *Mod Pathol*. 2017;30(9):1204-1212. doi:10.1038/modpathol.2017.43
38. Figenschau SL, Fismen S, Fenton KA, Fenton C, Mortensen ES. Tertiary lymphoid structures are associated with higher tumor grade in primary operable breast cancer patients. *BMC Cancer*. 2015;15(1):1-11. doi:10.1186/s12885-015-1116-1
39. Davidson-Pilon C. Lifelines: survival analysis in python. *J Open Source Softw*. 2019;4(40):1317. doi:10.21105/joss.01317
40. van der Maaten L. Accelerating t-SNE using tree-based algorithms. *J Mach Learn Res*. 2015;15:3221-3245.
41. Leong TL, Bryant VL. B cells in lung cancer-not just a bystander cell: a literature review. *Transl Lung Cancer Res*. 2021;10(6):2830-2841. doi:10.21037/tlcr-20-788
42. Thommen DS, Koelzer VH, Herzig P, et al. A transcriptionally and functionally distinct pd-1 + cd8 + t cell pool with predictive potential in non-small-cell lung cancer treated with pd-1 blockade. *Nat Med*. 2018;24(7):994-1004. doi:10.1038/s41591-018-0057-z

43. Ansel KM, Ngo VN, Hyman PL, et al. A chemokine-driven positive feedback loop organizes lymphoid follicles. *Nature*. 2000;406(6793):309-314. doi:10.1038/35018581
44. Germain C, Gnjatic S, Dieu-Nosjean MC. Tertiary lymphoid structure-associated B cells are key players in anti-tumor immunity. *Front Immunol*. 2015;6(FEB):1-14. doi:10.3389/fimmu.2015.00067
45. Sautès-Fridman C, Petitprez F, Calderaro J, Fridman WH. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat Rev Cancer*. 2019;19(6):307-325. doi:10.1038/s41568-019-0144-6
46. Coppola D, Nebozhyn M, Khalil F, et al. Unique ectopic lymph node-like structures present in human primary colorectal carcinoma are identified by immune gene array profiling. *Am J Pathol*. 2011;179(1):37-45. doi:10.1016/j.ajpath.2011.03.007
47. Vanhersecke L, Brunet M, Guégan JP, et al. Mature tertiary lymphoid structures predict immune checkpoint inhibitor efficacy in solid tumors independently of PD-L1 expression. *Nat Cancer*. 2021;2(8):794-802. doi:10.1038/s43018-021-00232-6
48. Huang AC, Orlowski RJ, Xu X, et al. A single dose of neoadjuvant PD-1 blockade predicts clinical outcomes in resectable melanoma. *Nat Med*. 2019;25(3):454-461. doi:10.1038/s41591-019-0357-y
49. Ayers M, Lunceford J, Nebozhyn M, et al. IFN-γ–related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Investig*. 2017;127(8):2930-2940. doi:10.1172/JCI91190
50. Fehrenbacher L, Spira A, Ballinger M, et al. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. *Lancet*. 2016;387(10030):1837-1846. doi:10.1016/S0140-6736(16)00587-0
51. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*. 2015;160(1–2):48-61. doi:10.1016/j.cell.2014.12.033
52. Davoli T, Uno H, Wooten EC, Elledge SJ. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science (1979)*. 2017;355(6322):eaaf8399. doi:10.1126/science.aaf8399
53. Shukla SA, Bachireddy P, Schilling B, et al. Cancer-germline antigen expression discriminates clinical outcome to CTLA-4 blockade. *Cell*. 2018;173(3):624-633.e8. doi:10.1016/j.cell.2018.03.026
54. Hennequin A, Derangère V, Boidot R, et al. Tumor infiltration by Tbet+ effector T cells and CD20+ B cells is associated with survival in gastric cancer patients. *Onco Targets Ther*. 2016;5(2):e1054598. doi:10.1080/2162402X.2015.1054598
55. Gu-Trantien C, Loi S, Garaud S, et al. CD4+ follicular helper T cell infiltration predicts breast cancer survival. *J Clin Investig*. 2013;123(7):2873-2892. doi:10.1172/JCI67428
56. Patel AJ, Richter A, Drayson MT, Middleton GW. The role of B lymphocytes in the immuno-biology of non-small-cell lung cancer. *Cancer Immunol Immunother*. 2020;69(3):325-342. doi:10.1007/s00262-019-02461-2
57. Al-Shibli KI, Donnem T, Al-Saad S, Persson M, Bremnes RM, Busund LT. Prognostic effect of epithelial and stromal lymphocyte infiltration in non-small cell lung cancer. *Clin Cancer Res*. 2008;14(16):5220-5227. doi:10.1158/1078-0432.CCR-08-0133
58. Kinoshita T, Muramatsu R, Fujita T, et al. Prognostic value of tumor-infiltrating lymphocytes differs depending on histological type and smoking habit in completely resected non-small-cell lung cancer. *Ann Oncol*. 2016;27(11):2117-2123. doi:10.1093/annonc/mdw319
59. Pelletier MP, Edwardes MDDB, Michel RP, Halwani F, Morin JE. Prognostic markers in resectable non-small cell lung cancer: a multivariate analysis. *Can J Surg*. 2001;44(3):180-188.
60. Stankovic B, Bjørhovde HAK, Skarshaug R, et al. Immune cell composition in human non-small cell lung cancer. *Front Immunol*. 2018;9(February):3101. doi:10.3389/fimmu.2018.03101
61. Hernández-Prieto S, Romera A, Ferrer M, et al. A 50-gene signature is a novel scoring system for tumor-infiltrating immune cells with strong correlation with clinical outcome of stage I/II non-small cell lung cancer. *Clin Transl Oncol*. 2015;17(4):330-338. doi:10.1007/s12094-014-1235-1
62. Wouters MCA, Nelson BH. Prognostic significance of tumor-infiltrating B cells and plasma cells in human cancer. *Clin Cancer Res*. 2018;24(24):6125-6135. doi:10.1158/1078-0432.CCR-18-1481
63. Workel HH, Lubbers JM, Arnold R, et al. A transcriptionally distinct CXCL13þCD103þCD8þ T-cell population is associated with B-cell recruitment and Neoantigen load in human cancer. *Cancer Immunol Res*. 2019;7(5):784-796. doi:10.1158/2326-6066.CIR-18-0517
64. Salmon H, Remark R, Gnjatic S, Merad M. Host tissue determinants of tumour immunity. *Nat Rev Cancer*. 2019;19(4):215-227. doi:10.1038/s41568-019-0125-9
65. Gottlin EB, Bentley RC, Campa MJ, Pisetsky DS, Herndon JE, Patz EF. The association of intratumoral germinal centers with early-stage non-small cell lung cancer. *J Thorac Oncol*. 2011;6(10):1687-1690. doi:10.1097/JTO.0b013e3182217bec
66. Cabrita R, Lauss M, Sanna A, et al. Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature*. 2020;577(7791):561-565. doi:10.1038/s41586-019-1914-8
67. Helmink BA, Reddy SM, Gao J, et al. B cells and tertiary lymphoid structures promote immunotherapy response. *Nature*. 2020;577(7791):549-555. doi:10.1038/s41586-019-1922-8
68. Denkert C, von Minckwitz G, Darb-Esfahani S, et al. Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet Oncol*. 2018;19(1):40-50. doi:10.1016/S1470-2045(17)30904-X
69. Song IH, Heo SH, Bang WS, et al. Predictive value of tertiary lymphoid structures assessed by high endothelial Venule counts in the Neoadjuvant setting of triple-negative breast cancer. *Cancer Res Treat*. 2017;49(2):399-407. doi:10.4143/crt.2016.215
70. Remark R, Lupo A, Alifano M, et al. Immune contexture and histological response after neoadjuvant chemotherapy predict clinical outcome of lung cancer patients. *Oncoimmunology*. 2016;5(12):e1255394. doi:10.1080/2162402X.2016.1255394
71. Lee HJ, Kim JY, Park IA, et al. Prognostic significance of tumor-infiltrating lymphocytes and the tertiary lymphoid structures in HER2-positive breast cancer treated with adjuvant trastuzumab. *Am J Clin Pathol*. 2015;144(2):278-288. doi:10.1309/AJCPIXUYDVZ0RZ3G
72. Martinet L, le Guellec S, Filleron T, et al. High endothelial venules (HEVs) in human melanoma lesions: major gateways for tumor-infiltrating lymphocytes. *Onco Targets Ther*. 2012;1(6):829-839. doi:10.4161/onci.20492

73. Petitprez F, de Reyniès A, Keung EZ, et al. B cells are associated with survival and immunotherapy response in sarcoma. *Nature*. 2020;577(7791):556-560. doi:10.1038/s41586-019-1906-8
74. Tokunaga R, Nakagawa S, Sakamoto Y, et al. 12-chemokine signature, a predictor of tumor recurrence in colorectal cancer. *Int J Cancer*. 2020;147(2):532-541. doi:10.1002/ijc.32982
75. Munoz-Erazo L, Rhodes JL, Marion VC, Kemp RA. Tertiary lymphoid structures in cancer – considerations for patient prognosis. *Cell Mol Immunol*. 2020;17(6):570-575. doi:10.1038/s41423-020-0457-0
76. Li F, Li C, Cai X, et al. The association between CD8+ tumor-infiltrating lymphocytes and the clinical outcome of cancer immunotherapy: a systematic review and meta-analysis. *EClinicalMedicine*. 2021;41:101134. doi:10.1016/j.eclinm.2021.101134
77. Inoue H, Kosuga T, Kubota T, et al. Significance of a preoperative systemic immune-inflammation index as a predictor of postoperative survival outcomes in gastric cancer. *World J Surg Oncol*. 2021;19(1):173. doi:10.1186/s12957-021-02286-3
78. Zhang X, Zhao W, Yu Y, et al. Clinicopathological and prognostic significance of platelet-lymphocyte ratio (PLR) in gastric cancer: an updated meta-analysis. *World J Surg Oncol*. 2020;18(1):191. doi:10.1186/s12957-020-01952-2
79. Kim EY, Song KY. The preoperative and the postoperative neutrophil-to-lymphocyte ratios both predict prognosis in gastric cancer patients. *World J Surg Oncol*. 2020;18(1):293. doi:10.1186/s12957-020-02059-4
80. Öjlert ÅK, Halvorsen AR, Nebdal D, et al. The immune microenvironment in non-small cell lung cancer is predictive of prognosis after surgery. *Mol Oncol*. 2019;13(5):1166-1179. doi:10.1002/1878-0261.12475


## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Raju Paul S, Valiev I, Korek SE, et al. B cell-dependent subtypes and treatment-based immune correlates to survival in stage 3 and 4 lung adenocarcinomas. *FASEB BioAdvances*. 2023;5:156-170. doi:10.1096/fba.2023-00009