

Accounting for B-cell Behavior and Sampling Bias Predicts Anti-PD-L1 Response in Bladder Cancer

Ilya A. Dyugay^{1,2}, Daniil K. Lukyanov^{1,2}, Maria A. Turchaninova^{2,3}, Ekaterina O. Serebrovskaya^{2,3}, Ekaterina A. Bryushkova^{2,3,4}, Andrew R. Zaretsky^{2,3}, Oybek Khalmurzaev⁵, Vsevolod B. Matveev⁵, Mikhail Shugay^{2,3}, Pavel V. Shelyakin^{2,3}, and Dmitriy M. Chudakov^{1,2,3}



ABSTRACT

Cancer immunotherapy is predominantly based on T cell-centric approaches. At the same time, the adaptive immune response in the tumor environment also includes clonally produced immunoglobulins and clonal effector/memory B cells that participate in antigen-specific decisions through their interactions with T cells. Here, we investigated the role of infiltrating B cells in bladder cancer via patient dataset analysis of intratumoral immunoglobulin repertoires. We showed that the IgG1/IgA ratio is a prognostic indicator for several subtypes of bladder cancer and for the whole IMVigor210 anti-PD-L1 immunotherapy study cohort. A high IgG1/IgA ratio associated with the prominence of a cytotoxic gene signature, T-cell receptor signaling, and IL21-mediated signaling. Immunoglobulin repertoire analysis indicated that effector B-cell function, rather than clonally produced antibodies, was involved in antitumor responses. From the T-

cell side, we normalized a cytotoxic signature against the extent of immune cell infiltration to neutralize the artificial sampling-based variability in immune gene expression. Resulting metrics reflected proportion of cytotoxic cells among tumor-infiltrating immune cells and improved prediction of anti-PD-L1 responses. At the same time, the IgG1/IgA ratio remained an independent prognostic factor. Integration of the B-cell, natural killer cell, and T-cell signatures allowed for the most accurate prediction of anti-PD-L1 therapy responses. On the basis of these findings, we developed a predictor called PRedictive MolecUlar Signature (PRIMUS), which outperformed PD-L1 expression scores and known gene signatures. Overall, PRIMUS allows for reliable identification of responders among patients with muscle-invasive urothelial carcinoma, including the subcohort with the low-infiltrated “desert” tumor phenotype.

Introduction

Tumor-infiltrating B cells and intratumorally produced immunoglobulins (Ig) play important roles in the tumor microenvironment (TME) and response to immunotherapy (1–5). IgG antibodies produced by intratumoral B cells may drive antibody-dependent cellular cytotoxicity (ADCC) and enhance antigen presentation by dendritic cells (6–8). B cells are efficient antigen-specific antigen presenters that modulate the behavior of Th cells (9–11). Data on the role of B cells in the bladder cancer TME remain somewhat contradictory and indicate that B-cell infiltration may be associated with not only increased tumor invasiveness (12), but also with

increased T-cell infiltration, colocalization of CD4⁺ T cells and B cells, and antigen presentation (13). The complexity of interpreting these findings arises from the fact that the roles of different clonal B-cell populations may differ in terms of their antigen specificity, propensity to produce clonal antibodies or present antigens, and interactions with T cells, which may have immunosuppressive or cytotoxicity-inducing features (5).

The functionality of antibodies, determined by their isotype, can also affect antitumor responses, including potential immunosuppressive effects (14) and inflammatory processes promoted by immune complexes (5, 15). For instance, high intratumoral IgA in bladder cancer associates with shorter patient survival (16). The relative prevalence of antibodies with a certain isotype may also reflect the cytokine milieu in the ongoing antitumor response, as well as the general polarization of B cells (3). In melanoma, a high IgG1/IgA ratio (IgG1/IgA1+A2 gene expression ratio) and large IgG1 clonal expansions [which mainly reflect the presence of clonal IgG1-producing plasma cells in RNA sequencing (RNA-seq) data] associate with a favorable prognosis (17). This suggests that cytotoxic, tumor-specific antibody production is one of the major mechanisms of melanoma immune surveillance via ADCC and antibody-dependent cellular phagocytosis (7, 18). In contrast, for KRAS-mutated lung adenocarcinoma, where a high IgG1/IgA ratio also associates with better prognosis, high clonality of the IgG1 repertoire does not associate with longer survival (19). These observations suggest the existence of alternative explanations for the association of a better prognosis with a high IgG1/IgA ratio, such as B cell-mediated antigen presentation (9–11) or direct B-cell cytotoxicity (20). Alternatively, because a lower IgG1/IgA ratio also implies relatively higher abundance of IgA antibodies, the negative immunosuppressive (14) or pro-inflammatory (15) roles of the latter immunoglobulin isotype could contribute to a better prognosis for patients with a high intratumoral IgG1/IgA ratio.

¹Center of Life Sciences, Skolkovo Institute of Science and Technology, Moscow, Russia. ²Genomics of Adaptive Immunity Department, Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia. ³Institute of Translational Medicine, Pirogov Russian National Research Medical University, Moscow, Russia. ⁴Molecular Biology Department, Lomonosov Moscow State University, Moscow, Russia. ⁵Department of Urology, Federal State Budgetary Institution “N.N. Blokhin National Medical Research Center of Oncology” of the Ministry of Health of the Russian Federation, Moscow, Russia.

Note: Supplementary data for this article are available at Cancer Immunology Research Online (<http://cancerimmunolres.aacrjournals.org/>).

Corresponding Author: Dmitriy M. Chudakov, Genomics of Adaptive Immunity, IBCH RAS, Miklukho-Maklaya, 16/10, Moscow 117997, Russia. Phone: 7 (495) 335-01-00; E-mail: chudakovdm@gmail.com

Cancer Immunol Res 2022;10:343–53

doi: 10.1158/2326-6066.CIR-21-0489

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 International (CC BY-NC-ND).

©2022 The Authors; Published by the American Association for Cancer Research

Here, we investigated the architecture of the intratumoral immunoglobulin repertoire in distinct subcohorts of patients with bladder cancer. We showed that a high IgG1/IgA ratio associated with longer patient survival, response to anti-PD-L1 therapy, higher prevalence of cytotoxic processes among infiltrating immune cells, increased T-cell receptor (TCR) signaling, and a more prominent IL21-mediated signaling signature. The IgG1/IgA ratio was an independent prognostic factor, suggesting an active role for tumor-infiltrating B cells. Because IgG1/IgA is a ratio of immune gene expression, this metric does not depend on tissue sampling nor the extent of tumor infiltration and reflects the proportion of IgG1- versus IgA-switched B cells and plasma cells in the sample. We further normalized a CD8⁺ effector T-cell gene signature against CD45-correlated pan-leukocyte genes. The resulting metric reflected the relative activity of cytotoxic processes among tumor-infiltrating immune cells, rather than the extent of tumor infiltration by immune cells. Similar to the IgG1/IgA ratio, this metric does not depend on tissue sampling. The normalization increased the prognostic and predictive power of CD8⁺ effector T-cell gene signature. We next integrated the B-cell, natural killer (NK) cell, and T-cell signatures and developed a tumor RNA-seq-based predictor of anti-PD-L1 therapy responses in muscle-invasive urothelial carcinoma. The resulting predictor achieved superior sensitivity compared with PD-L1 expression scores or existing gene signatures, allowing for reliable identification of responders even within the “desert” patient subcohort, which we analyzed as a holdout dataset. Feature interaction analysis further supported an important role for interaction between T and B cells in response to anti-PD-L1 immunotherapy. General relevance of the model was validated in an independent, non-immunotherapy treated The Cancer Genome Atlas (TCGA) bladder cancer cohort (BLCA). Altogether, our results revealed an unrealized potential for the rational prediction of response to cancer immunotherapy.

Materials and Methods

Dataset analysis

Patient data from TCGA BLCA included 412 cases, and RNA-seq data were available for 408 cases (21). Cases contained data on both tumor and healthy tissues. For 6 patients, multiple replicates of tumor samples were present. There were 142 patients with basal squamous, 142 patients with luminal papillary, 78 with luminal infiltrated, 26 with luminal, and 20 with neuronal molecular subtypes. We downloaded FPKM-UQ (fragments per kilobase of transcript per million mapped reads upper quartile) files from the GDC data portal (21) for 433 tumor samples, including replicates, and we used only one randomly selected replicate for each patient. The data were then transformed to transcripts per million (TPM) according to the formula (22):

$$TPM_i = \frac{FPKM_i}{\sum_j FPKM_j} \cdot 10^6$$

Patient data from the IMVigor210 clinical trial (NCT02108652) were downloaded from European Genome-Phenome Archive (accession number EGAS00001002556). This clinical trial consisted of participants with locally advanced or metastatic urothelial bladder cancer who were treatment-naïve and ineligible for cisplatin-containing chemotherapy ($n = 119$ patients) or participants who had progressed during or following a prior platinum-based chemotherapy regimen ($n = 310$ patients) and included RNA-seq tissue transcriptomic data for 345 patients. Patients were classified into three immune phenotypes: desert ($n = 95$), excluded ($n = 164$), and inflamed ($n = 88$). All patients in IMVigor210 cohort had mea-

surable disease at baseline. The SP142 Ventana IHC assay was used in this trial. The scoring system (IC_A) indicates the percentage of PD-L1⁺ immune cells in a given tumor area (23, 24): IC0, <1%; IC1, 1%–5%; IC2/3, >5%. RECIST v1.1 was used to assess response to therapy. Abundances of transcripts from bulk RNA-seq data were quantified using Kallisto (25).

Normalization on pan-leukocyte genes

Gene expression in both TCGA and IMVigor210 datasets was normalized similarly to Teltsh and colleagues (26), with modifications. We first selected an immune-normalized gene set (INGS): a group of genes with a Spearman correlation coefficient with *PTPRC* (CD45) >0.9. Next, the sample-specific normalization factor (f_{INGS}) was calculated for each sample as the averaged expression of genes from INGS, and then the first normalization was performed. The normalization coefficient for genes included in INGS avoided self-normalization and was calculated as the averaged expression of the remaining genes. We selected genes from INGS for which the ratio between the coefficient of variation before and after the first normalization was <0.8 and used those genes as the final INGS. The second normalization was performed using the final INGS. Final genes for INGS were selected independently for TCGA and IMVigor210 datasets. TCGA final INGS included genes: *MPEG1*, *EVI2B*, *IL10RA*, *GPR65*, *WDFY4*, *CD53*, *ARHGAP9*, *CD48*, *CD84*, *CYTIP*, *RHOH*, *SAMSN1*, *CD3E*, *SLAMF6*, *DOCK2*, *SLA*, *ITGB2*, *SNX20*, *MNDA*, *CYBB*, *CXorf21*, *ITGAL*, *BTK*, *P2RY10*, *IL21R*, *PTPN22*, *TRAC*, *SLAMF1*, *ITK*, *LCP2*, *SPN*, *SASH3*, *CD2*, *PTPRC*, *NCKAP1L*, *PTPN7*, *SH2D1A*, and *PLEK*. IMVigor210 final INGS included genes: *DOCK2*, *IRF8*, *NCKAP1L*, *ARHGAP15*, *CD48*, *ITGAL*, *SAMSN1*, *ZC3H12D*, *CD226*, *P2RY10*, *CD53*, *WDFY4*, *IL10RA*, *PYHIN1*, *ICOS*, *ITGA4*, *AOAH*, *PTPN22*, *TRAC*, *CYTIP*, *CD2*, *INSY2B*, *ITK*, *SPN*, *SLAMF1*, *STAT4*, *PTPRC*, *IKZF1*, *SLFN12L*, *SLAMF6*, *CD3E*, *GPR65*.

Differential expression analysis

State-of-the-art methods for differential expression analysis are not applicable for normalized TPM values used in our work. To perform differential expression analysis, we were restricted to use a nonparametric approach. Differential expression analysis was performed using the Mann-Whitney *U* test to find differentially expressed genes in two categories of TCGA-BLCA patients: those with high and low IgG1/IgA ratio tertiles. IgA expression measurements were calculated as a sum of *IGHA1* and *IGHA2* genes. The values for low and high IgG1/IgA tertiles for the basal squamous subcohort ($n = 47$) were 0.52 and 1.35, respectively, and for the whole TCGA-BLCA cohort ($n = 136$), the corresponding values were 0.3 and 0.96. Obtained *P* values were adjusted with the Benjamini-Hochberg method (27). Fold change was calculated for each gene as the ratio of the median expression in the two samples. Pathway enrichment analysis was performed using slimGO tool with annotation dataset of biological processes (28).

Clonality analysis

We obtained immunoglobulin heavy chain (IGH) complementary-determining region 3 (CDR) repertoires for TCGA-BLCA and IMVigor210 patients from raw RNA-Seq data using MiXCR 3.0 (29) in RNA-seq mode. Data were prefiltered on the basis of 15-mer nucleotide matches to V/J segments of immune receptors. V and J sequences of *IGH*, *IGK*, *IGL*, *TRA*, *TRB*, *TRG*, and *TRD* receptors were taken from IMGT database, and all possible 15-mers were extracted. Samples with less than 300 IGH CDR3-related reads were omitted from the analysis. Included in the analysis were 217 patients from the whole TCGA cohort, including 89 patients from the basal squamous subcohort, and

228 patients in the IMVigor210 dataset. Immunoglobulin CDR3 repertoires were downsampled to 300 randomly chosen reads for normalization purposes. Immunoglobulin clonality was calculated as $(1 - \text{normalized Shannon-Wiener index; ref. 30})$ using VDJtools (31) software.

Gene signatures

Gene signatures were calculated for TCGA-BLCA and IMVigor210 patients as the first principal component of principal component analysis (PCA) performed with z-score-transformed expression of input genes. Calculations were performed with PCA from the Python scikit-learn library. Genes included in the CD8 signature (32): *CD8A*, *CXCL10*, *CXCL9*, *GZMA*, *GZMB*, *IFNG*, *PRF1*, and *TBX21*. The refined CD8 signature included: *CD8A*, *CXCL9*, *CXCL10*, and *GZMB*. The NK signature included: *KLRC2*, *KLRC3*, and *KLRC4*.

Random forest model

Random forest modeling, one of the universal machine-learning algorithms that can model response prediction by fitting training data based on different input features, was performed with the RandomForestClassifier from the Python scikit-learn library. During training, 15% of the IMVigor210 data (45 patients: 10 responders, 35 non-responders) were selected as a holdout set, and the remaining 85% (253 patients: 58 responders, 195 nonresponders) were selected as training data, preserving the proportions of response or non-response to immunotherapy. Hyperparameters of the model (i.e., maximal amount of samples in the leaf and tree depth) were optimized with RandomizedSearchCV and GridSearchCV from the Python scikit-learn library using 5-fold cross-validation. The number of estimators in the model was 50. The model was trained using the F1 score as a measure of quality:

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

The model performance was evaluated with five-inner/five-outer nested cross-validations. This approach, unlike regular cross-validation, assumes to fit the model using two nested loops, where the inner loop is used for hyperparameter optimization and model selection, as with regular cross-validation, whereas the outer loop is used to split the data into training and test datasets to estimate the performance. We decided to use nested cross-validation because of the small size of the available data points, unbiased generalization performance estimation, and prevention of selection bias.

Validation of the PRIMUS model

The performance of our Predictive Molecular Signature (PRIMUS) model was compared with the support vector machine (SVM)-based model with a linear kernel trained on the same training data as PRIMUS. We used SVM realization from the Python scikit-learn library. PRIMUS is a random forest-based model which is slightly prone to overfit because the high depth of decision trees in the ensemble can result in an overcomplicated model and incur unnecessary variance (33, 34). On the other hand, SVM with linear kernel is a simple model that is well designed for discriminating linearly separable data and is unlikely to overfit complex data due to high variance of the model. The variance of the PRIMUS model can be interpreted as the difference between training and test set quality metrics (35). We compared the difference between training and test set quality metrics for the PRIMUS and SVM model to detect overfitting of the PRIMUS model.

We also applied PRIMUS to the IMVigor210 data (all patients) to explore input feature importance and interactions using SHAP (36), a game theory approach to explain machine-learning model and understand the decision-making process by quantifying the contribution that each feature brings to the prediction made by the model.

Statistical analysis

Survival analysis was performed with the lifelines (37) Python library. Survival plots were created using the Kaplan–Meier estimator. Significance of the difference between two survival curves was estimated with a log-rank test. For comparing multiple survival curves, multivariate log-rank test was used. Cox proportional hazards models were fitted on either features or features with an interaction value, which occurred when the effect of an independent variable on a dependent variable changed, depending on the value of other independent variables. The relative reliability of models was estimated by the Akaike information criterion (38) and concordance index (39). The Cox area under the receiver operating curve (AUROC) was calculated with the Python scikit-learn library. For multiple comparisons, correction was performed using the Benjamini–Hochberg procedure (27). Group comparison in boxplots was performed with the Kruskal–Wallis test. All statistical calculations were performed using Python. $P < 0.05$ was considered statistically significant.

Results

Immunoglobulin isotype composition and clonality

We investigated BLCA patient cohorts from TCGA with distinct mRNA expression-based molecular subtypes to identify patients more likely to have a favorable prognosis while exhibiting a high intratumoral IgG1/IgA ratio (Fig. 1A; Supplementary Data S1). The basal squamous and luminal-infiltrated molecular subtypes showed the most significant correlation with patient survival when stratified on the basis of IgG1/IgA ratio, with a high IgG1/IgA ratio associated with better survival versus a low IgG1/IgA ratio (Fig. 1B; Supplementary Fig. S1). Basal squamous tumors are generally characterized by high lymphocytic infiltration, including CD8⁺ T cells (40) and Th1 cells (41). Thus, one of the possible explanations for this result could be the association of IgG1 isotype switching with a type 1 immune response (3, 42, 43).

Next, we studied the clonality of IgG repertoires extracted from BLCA RNA-seq data using MiXCR (29). We observed that high clonality (i.e., the presence of large clonal immunoglobulin expansions) was not significantly associated with prognosis for all patients, nor for patients with basal squamous tumors (Supplementary Fig. S2A and S2B). However, the combination of IgG1/IgA ratio and immunoglobulin clonality showed high prognostic value, with the best prognosis associated with high IgG1/IgA and low immunoglobulin clonality (Fig. 1C). These results suggest that high clonal antibody production does not efficiently contribute to immune surveillance of bladder cancer, which contrasts with melanoma but is similar to observations in KRAS-mutated lung adenocarcinoma. This observation, thus, refutes the role of antigen-specific, IgG1-mediated ADCC as a basis for association of a high IgG1/IgA ratio with longer survival in basal squamous bladder cancer and indicates that this association is likely attributable to other B-cell functions.

A high IgG1/IgA expression ratio associates with a cytotoxic immune signature

Next, we aimed to identify the immune processes associated with a high IgG1/IgA expression ratio. To this end, we divided patients with

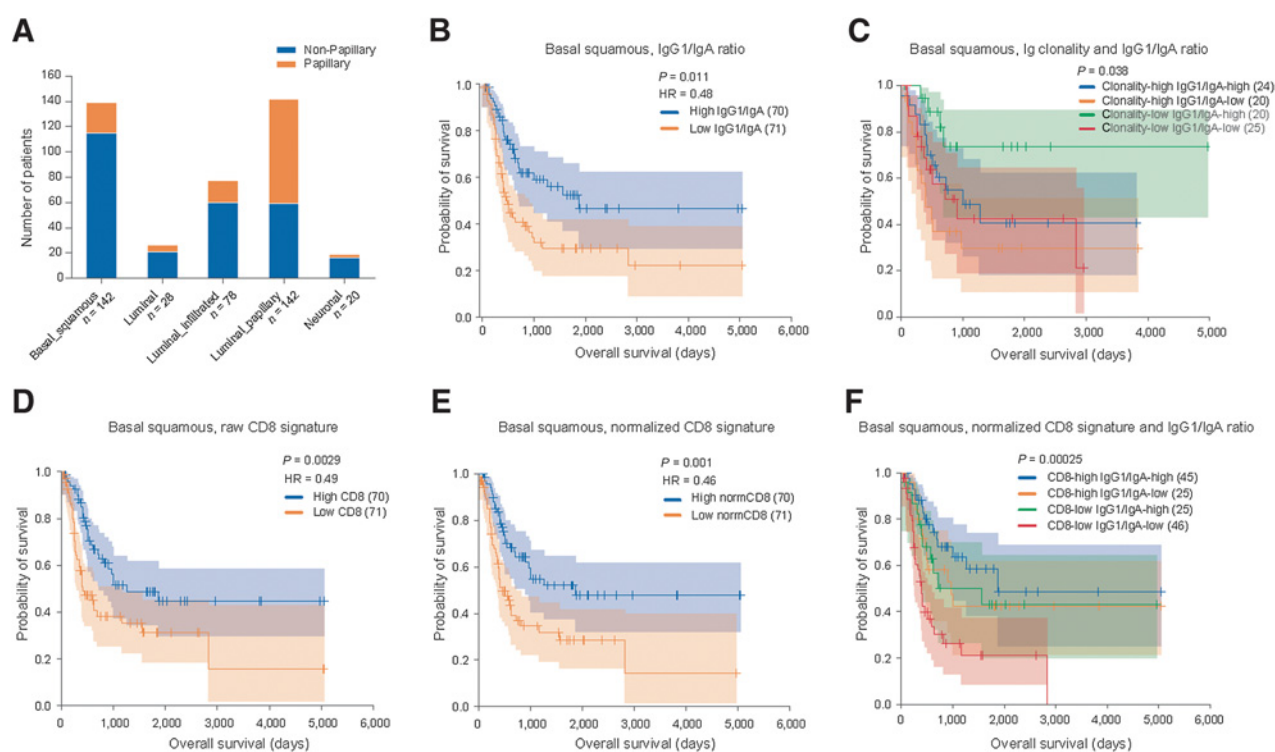


Figure 1.

Immunoglobulin repertoire features and the normalized CD8⁺ T-cell signature predict survival in basal squamous BLCA-TCGA patients. **A**, Relative overlap of the histologic (papillary, $n = 132$ patients); (non-papillary, $n = 271$ patients) and mRNA expression-based molecular subtypes of bladder cancer from TCGA. **B–F**, Kaplan-Meier curves showing overall survival for TCGA patients with basal squamous bladder cancer with high and low IgG1/IgA expression ratio (**B**); a combination of high or low IgG1/IgA expression ratio and high or low Ig clonality (high clonality-high ratio vs. high clonality-low ratio: $P = 0.16$; high clonality-high ratio vs. low clonality-high ratio: $P = 0.07$; high clonality-high ratio vs. low clonality-low ratio: $P = 0.63$; high clonality-low ratio vs. low clonality-high ratio: $P = 0.004$; high clonality-low ratio vs. low clonality-low ratio: $P = 0.35$; low clonality-high ratio vs. low clonality-low ratio: $P = 0.04$, log-rank test; **C**); high or low CD8⁺ T-cell signature (raw; **D**); high or low normalized CD8⁺ T-cell signature (normalization against pan-leukocyte genes; **E**); and a combination of high and low IgG1/IgA expression ratios and normalized CD8⁺ T-cell signature (high signature-high ratio vs. high signature-low ratio: $P = 0.26$; high signature-high ratio vs. low signature-high ratio: $P = 0.27$; high signature-high ratio vs. low signature-low ratio: $P = 5 \times 10^{-5}$; high signature-low ratio vs. low signature-high ratio: $P = 1.0$; high signature-low ratio vs. low signature-low ratio: $P = 0.02$; low signature-high ratio vs. low signature-low ratio: $P = 0.03$, log-rank test; **F**). Patient cohorts are divided on the basis of median, with total number of patients shown in parentheses.

basal squamous bladder cancer into tertiles based on IgG1/IgA ratio and compared the differential gene expression in patients from high versus low IgG1/IgA tertiles. Pathway analysis of the normalized genes overexpressed in the high IgG1/IgA subcohort showed enrichment of TCR signaling, CD8⁺ T-cell activation, NK cell-mediated cytotoxicity (*CXCL9*, *CXCL10*, *CD8A*, *CD8B*, *GZMA*, *GZMB*, *PRF1*, *TBX21*, *IFNG*, *KLRC2*, *GNLY*), IL21-mediated signaling, immune checkpoints (*CTLA4*, *LAG3*, *PDCD1*), B-cell receptor signaling, phagocytosis, and apoptosis (Supplementary Fig. S3; Supplementary Table S1 for the full list of positively differentially expressed genes). The association between a high IgG1/IgA ratio and increased expression of cytotoxic genes suggests a possible correlation between the type 1 T-cell responses and IgG1 isotype switching. IL21 is known to exert an antitumor effect by enhancing and supporting CD8⁺ T-cell responses (44), and IL21 produced by follicular Th cells promotes B-cell proliferation and maturation (45). Similar analysis of the full TCGA BLCA patient cohort showed more cytotoxic genes positively associated with a high IgG1/IgA ratio, along with more prominent expression of costimulatory CD80, increased IL21-mediated signaling, checkpoint regulation, Fcγ receptor signaling, and receptor-mediated phagocytosis and endocytosis (Supplementary Fig. S3).

Prognostic value of a normalized CD8⁺ effector T-cell signature and IgG1/IgA ratio

A study has proposed a CD8⁺ effector T cell-associated gene signature for tumors with an inflamed histologic phenotype that associates with response to anti-PD-L1 immunotherapy. The genes for this signature include *CD8A*, *CXCL9*, *CXCL10*, *GZMA*, *GZMB*, *IFNG*, *PRF1*, and *TBX21* (32). When we applied this raw CD8⁺ T-cell signature to the basal squamous TCGA BLCA subcohort, high expression of the signature associated with better patient survival (**Fig. 1D**). We hypothesized that the relative activity of distinct processes among tumor-infiltrating immune cells may prevail over the infiltration level *per se* in terms of prognostic and predictive value. We also reasoned that the estimation of relative activity of immune processes would neutralize the artificial variability in immune gene expression resulting from random tissue sampling. To estimate the relative activity of distinct processes in tumor-infiltrating immune cells independently of both the extent of tumor infiltration and of the sampling randomness, we normalized gene expression against *CD45*-correlated pan-leukocyte genes, similar to the previously reported immFocus approach (26). Using this normalized CD8⁺ T-cell signature resulted in a more accurate association with prognosis (**Fig. 1E**). The combination of the IgG1/IgA ratio and normalized CD8⁺ T-cell signature

had greater prognostic value compared with the CD8⁺ T-cell signature alone, with the best prognosis associated with a high IgG1/IgA ratio and high normalized CD8⁺ signature (Fig. 1F). This observation suggests that the IgG1/IgA isotype ratio has independent prognostic value and is not merely a passive biomarker of type 1 T-cell responses. Multivariate Cox proportional hazards regression analysis confirmed the independent contribution of the IgG1/IgA ratio to predicting

overall survival of patients with basal squamous bladder cancer (Supplementary Table S2).

Predicting survival and response to immunotherapy

We next evaluated the applicability of the normalized CD8⁺ T-cell signature and IgG1/IgA ratio to predict the individual responses to anti-PD-L1 immunotherapy using data from the IMVigor210 phase II

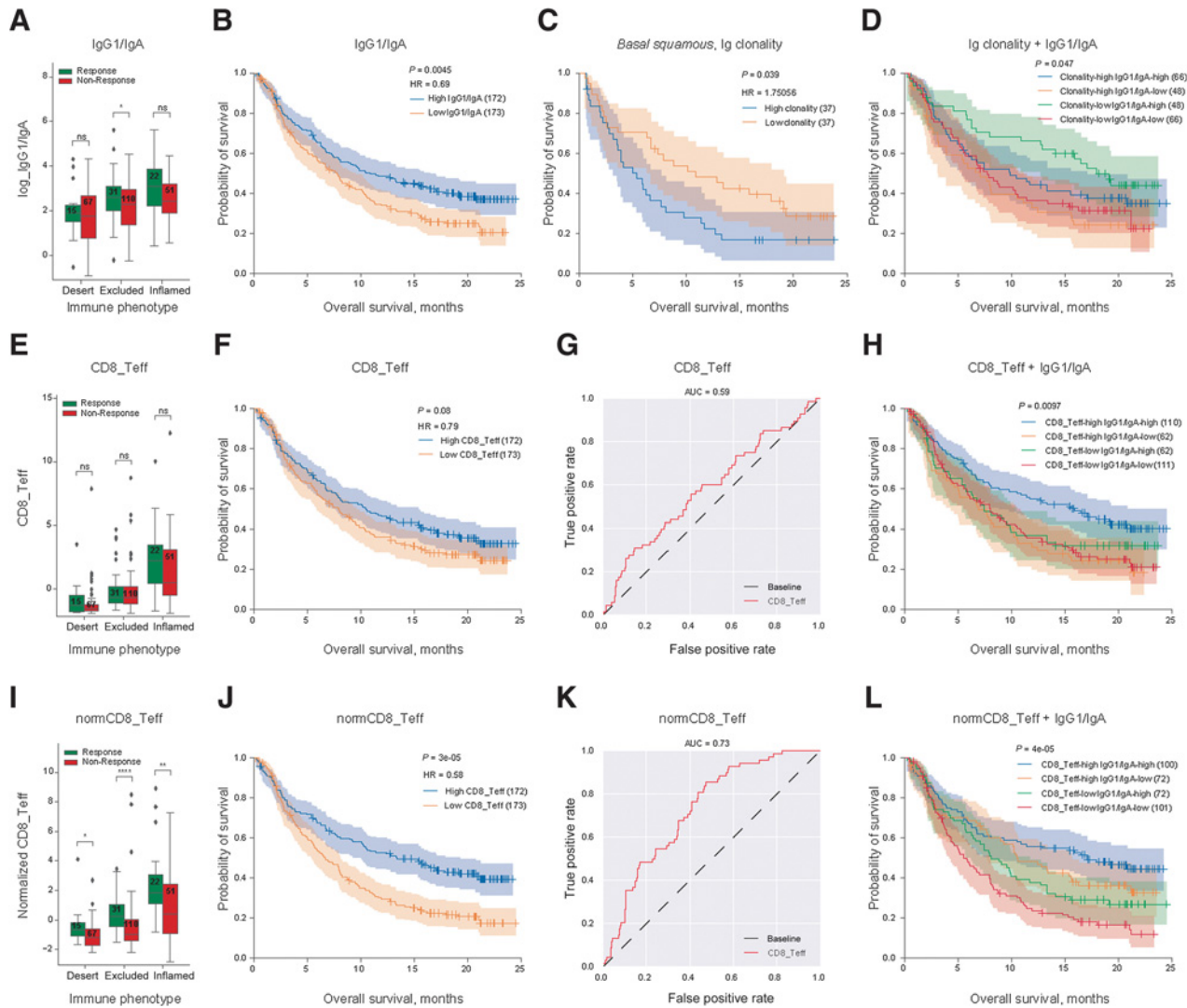


Figure 2.

Predictive and prognostic role of different immune features in anti-PD-L1 immunotherapy of bladder cancer for the IMVigor210 cohort. **A-D**, Predictive and prognostic role of the IgG1/IgA ratio and immunoglobulin clonality. **D**, High clonality-high ratio versus high clonality-low ratio: $P = 0.16$; high clonality-high ratio versus low clonality-high ratio: $P = 0.17$; high clonality-high ratio versus low clonality-low ratio: $P = 0.4$; high clonality-low ratio versus low clonality-high ratio: $P = 0.01$; high clonality-low ratio versus low clonality-low ratio: $P = 0.53$; low clonality-high ratio versus low clonality-low ratio: $P = 0.02$, log-rank test. **E-H**, Predictive and prognostic role of the raw CD8⁺ T-cell signature. **H**, High signature-high ratio versus high signature-low ratio: $P = 0.004$; high signature-high ratio versus low signature-high ratio: $P = 0.11$; high signature-high ratio versus low signature-low ratio: $P = 0.005$; high signature-low ratio versus low signature-high ratio: $P = 0.32$; high signature-low ratio versus low signature-low ratio: $P = 0.56$; low signature-high ratio versus low signature-low ratio: $P = 0.55$, log-rank test. **I-L**, Predictive and prognostic role of the normalized CD8⁺ T-cell signature (normalization against pan-leukocyte genes). **L**, high signature-high ratio versus high signature-low ratio: $P = 0.28$; high signature-high ratio versus low signature-high ratio: $P = 0.02$; high signature-high ratio versus low signature-low ratio: $P < 1 \times 10^{-5}$; high signature-low ratio versus low signature-high ratio: $P = 0.21$; high signature-low ratio versus low signature-low ratio: $P = 0.0006$; low signature-high ratio versus low signature-low ratio: $P = 0.03$, log-rank test. Boxplots in **A**, **E**, and **I** show associations with response to anti-PD-L1 immunotherapy for different tumor immune phenotypes. Median, bottom quartile, top quartile, and interquartile range are shown. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$. Kaplan-Meier plots show overall survival of patients with different immune features either alone (**B**, **C**, **F**, and **J**) or in combination (**D**, **H**, and **L**). **G** and **K**, AUROC showing discriminative ability of the raw and normalized CD8⁺ T-cell signature to diagnose patients who would benefit from atezolizumab immunotherapy. Total number of patients shown in parentheses. ns, non-significant. Except for **C**, data are shown for the whole IMVigor210 cohort. AUC, area under the curve.

clinical trial of atezolizumab (anti-PD-L1) in muscle-invasive urothelial carcinoma (32). This cohort includes patients with distinct immune phenotypes that can be distinguished on the basis of biopsy histology. The “inflamed” phenotype is characterized by the abundant infiltration of CD4⁺ and CD8⁺ T cells into the tumor parenchyma, often accompanied by other immune cells, including immunosuppressive subtypes. The “excluded” phenotype is characterized by the localization of multitudes of immune cells in the tumor stroma surrounding nests of tumor cells. Finally, the “desert” phenotype is characterized by a non-inflamed TME, with few or no T cells in either the parenchyma or stroma regions of the tumor (46, 47). The IgG1/IgA ratio alone had significant predictive value for response to atezolizumab only for the excluded immune phenotype (Fig. 2A), and significant prognostic value for overall survival of treated patients for the whole IMVigor210 cohort (Fig. 2B). A significant association of high immunoglobulin clonality with negative prognosis was seen for the basal squamous subcohort of treated patients (Fig. 2C; ref. 48). Low immunoglobulin clonality combined with a high IgG1/IgA ratio associated with the best prognosis for the whole IMVigor210 cohort (Fig. 2D), as was seen for the basal squamous subset of TCGA patients (Fig. 1C).

The raw CD8⁺ T-cell signature was poorly predictive of response (Fig. 2E–H), but normalization against pan-leukocyte genes improved predictive power (Fig. 2I–K). The combination of the normalized CD8⁺ T-cell signature with IgG1/IgA ratio yielded the best prognostic value (Fig. 2L). Multivariate Cox proportional hazards regression analysis again showed a prominent and independent contribution of the IgG1/IgA ratio and normalized CD8 signature to the prognosis for ImVigor210 patients (Supplementary Table S2).

Integrative predictive modeling of response to anti-PD-L1 immunotherapy

IHC measurement of PD-L1 expression in tumor samples is currently used to identify patients who may have a higher chance of responding to immunotherapy. The IMVigor210 trial measured PD-L1 using antibodies to the PD-L1 C-terminus, and the scoring system (IC_A) was calculated as the percentage of PD-L1⁺ immune cells in a given tumor area (23, 24). The cutoff for first-line therapy is

5% (IC2/3), but the predictive value of the PD-L1 scoring was relatively low (Fig. 3; refs. 49–51).

To develop an improved gene expression-based predictor for rational patient stratification, we used a random forest model, with the performance evaluated via a nested cross-validation approach (Fig. 4; ref. 52). This approach allowed us to overcome the overly optimistic evaluation of the model performance introduced by hyperparameter optimization during the model selection procedure. First, to set a baseline for our model’s performance to compare with subsequent iterations during the refinement process, the model was trained using the raw CD8⁺ T-cell signature. The resulting AUROC, representing the accuracy of predictive modeling of response, was similar to the one built on the signature values (compare Fig. 2G with Fig. 4A), which did not differ from random patient selection. The mean nested cross-validation F1 score (reflects the balance between precision and recall, whereby values closer to 1 are better), was 0.344 ± 0.06 . Next, we trained the model using the normalized CD8⁺ T-cell signature. The AUROC was greater than the AUROC for the raw signature, and similar to the AUROC of the normalized signature used without the model (compare Fig. 2K with Fig. 4D). The mean nested cross-validation F1 score was 0.465 ± 0.05 .

Before selecting the final set of input features, we further refined the normalized CD8⁺ T-cell signature by fitting the model for each signature gene separately rather than in combination. On the basis of feature importance for the resulting model, four genes were selected: *CXCL9*, *CXCL10*, *CD8A*, *GZMB*. To appropriately estimate feature importance, we also analyzed them for the presence of multicollinearity. For each variable, we calculated a variance inflation factor (estimates how much each predictor’s variance is inflated under a multicollinearity condition) of <5 , which indicates that we had no multicollinear input parameters. We also excluded several parameters that demonstrated no significant predictive value, including immunoglobulin clonality, *IL21R*, and *CD80*.

Finally, we integrated the refined normalized CD8⁺ T-cell signature (*CXCL9*, *CXCL10*, *CD8A*, *GZMB*), the IgG1/IgA isotype ratio, and features associated with a high IgG1/IgA ratio for TCGA-BLCA patients that included the KLRC (killer cell lectin-like receptor)-NK signature (*KLRC2*, *KLRC3*, *KLRC4*), *IL21* and *GPLY* expression (all

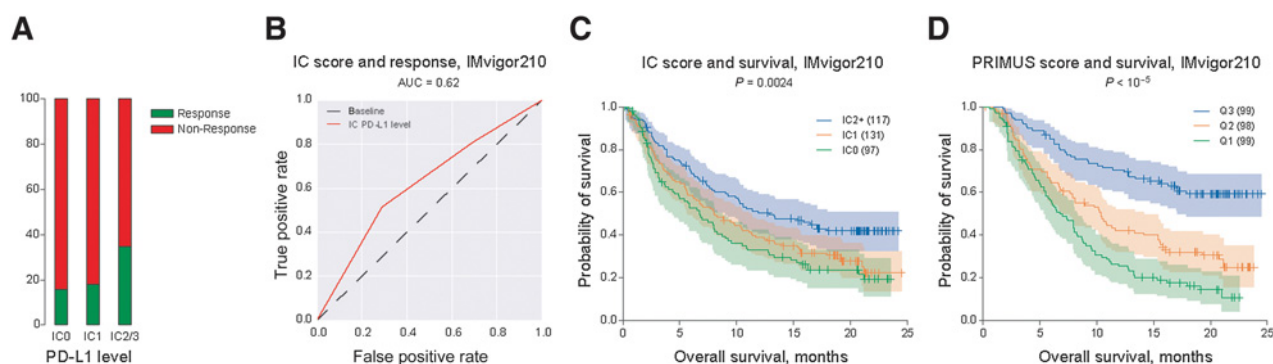


Figure 3.

Prognostic value of IHC PD-L1 measurements among tumor-infiltrating immune cells in the IMVigor210 cohort. **A**, Percentage of responders and nonresponders among IMVigor210 patients, divided on the basis of the abundance of PD-L1⁺ immune cells. Responders/nonresponders ratio: IC0 13/70, IC1 20/92, IC2+ 35/66. **B**, ROC representing the discriminative ability of PD-L1⁺ immune cell counts in biopsies to identify patients who would benefit from atezolizumab immunotherapy. AUC, area under the curve. **C**, Kaplan-Meier survival plots for patients with different counts of PD-L1⁺ immune cells. IC2+ versus IC1: $P = 0.02$; IC2+ versus IC0: $P = 0.0007$; IC1 versus IC0: $P = 0.27$, log-rank test. **D**, Kaplan-Meier survival plots for patients divided according to predicted probability of response by PRIMUS. Patients were divided into tertiles (quantile1/2/3). Q3 versus Q2: $P = 6 \times 10^{-5}$; Q3 versus Q1: $P < 1 \times 10^{-5}$; Q2 versus Q1: $P = 0.008$, log-rank test. This panel formally includes the data used for training the model. **C** and **D**, Total number of patients shown in parentheses.

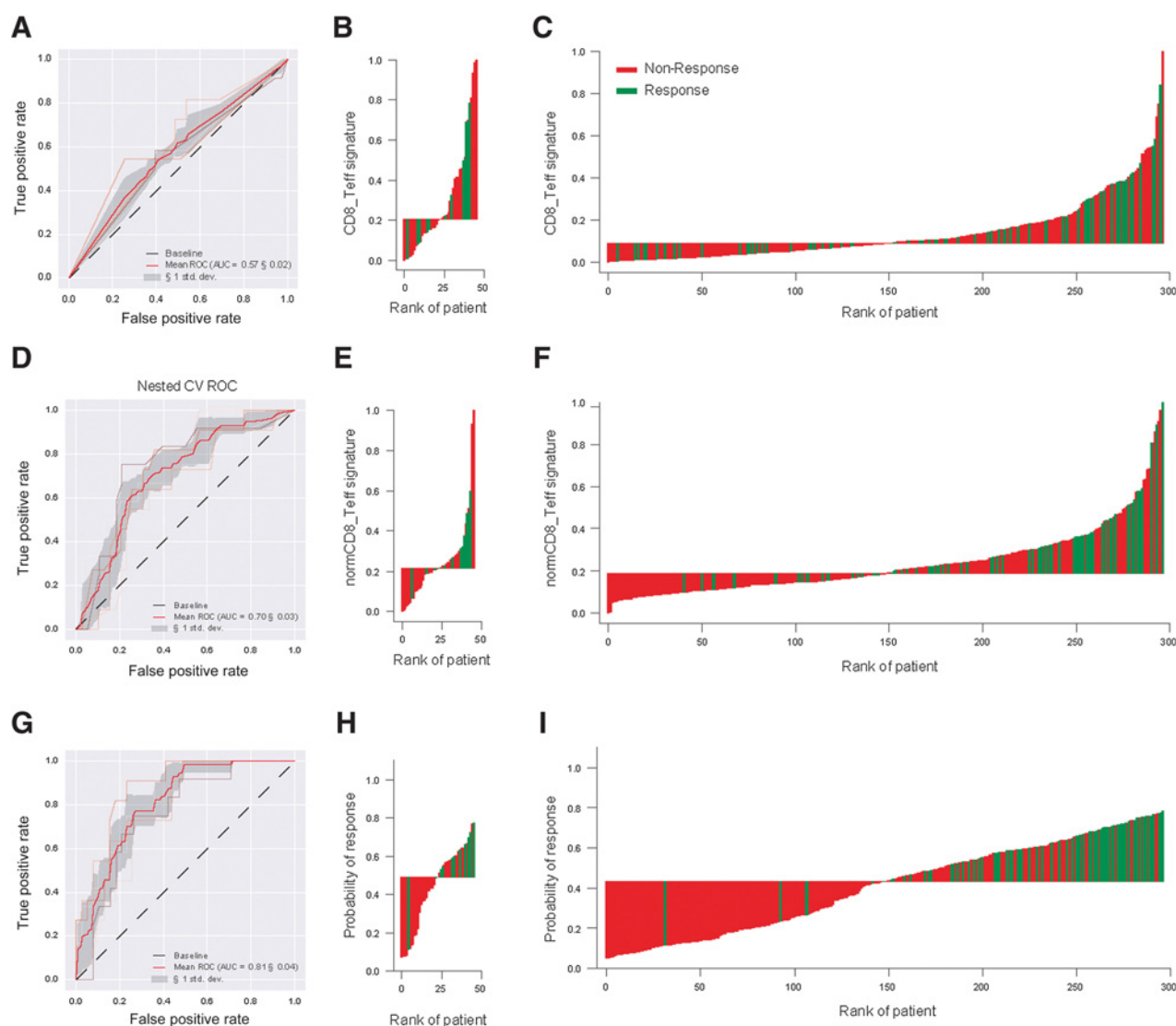


Figure 4. Predictive value of the raw and normalized CD8⁺ T-cell signatures and the integrative random forest model on the IMVigor210 cohort. **A-C**, Raw CD8⁺ T-cell signature based on expression of *CD8A*, *CXCL10*, *CXCL9*, *GZMA*, *GZMB*, *IFNG*, *PRFI*, *TBX21*. **D-F**, Normalized CD8⁺ T-cell signature. **G-I**, Model based on a refined normalized CD8⁺ T-cell signature combined with the normalized KLRC-NK signature, normalized *IL21*, and *GNLY* expression. AUROC (**A**, **D**, and **G**) shows the discriminative ability of the signature to identify patients who would benefit from atezolizumab immunotherapy. Waterfall charts show distributions of responders and nonresponders according to corresponding ranked feature values for holdout sets (**B**, **E**, and **H**) and the whole patient cohort (**C**, **F**, and **I**). **I** formally includes data used for training the model.

normalized against pan-leukocyte genes), as well as non-normalized *TGFB1* expression (45). The performance of the resulting PRIMUS model (see Materials and Methods), which had a mean nested cross-validation F1 score 0.495 ± 0.05 , was superior to that of the model trained on the normalized CD8⁺ T-cell signature (**Fig. 4D-I**). PRIMUS allowed for superior prediction of response and yielded greater prognostic value compared with PD-L1 IHC (**Fig. 3B-D**; **Fig. 4G**). PRIMUS also outperformed prognoses based on the normalized CD8⁺ T-cell signature or the combination of high/low IgG1/IgA isotype ratio and normalized signature (compare **Fig. 2I-L** with **Fig. 3D**). To confirm the validity of our model, we compared the performance of PRIMUS with another machine-learning algorithm, the SVM-based model with a linear kernel. The results of the SMV model were

comparable with those obtained with PRIMUS, with a mean nested cross-validation F1 score of 0.45 ± 0.05 . The difference between training and test set quality metrics for PRIMUS and SVM-based models were: AUROC: 0.034 and 0.057 and F1: 0.071 and 0.01, respectively. This indicated that the PRIMUS model was not overfitted.

Feature importance and interaction analysis

We next applied PRIMUS to the IMVigor210 data to explore the importance of, and interactions between, input features using SHAP (36). First, we compared the contribution of our input variables to randomly generated numbers. We expected features selected for the final iteration of PRIMUS to have higher feature importance for response compared with randomly generated numbers. Indeed, each

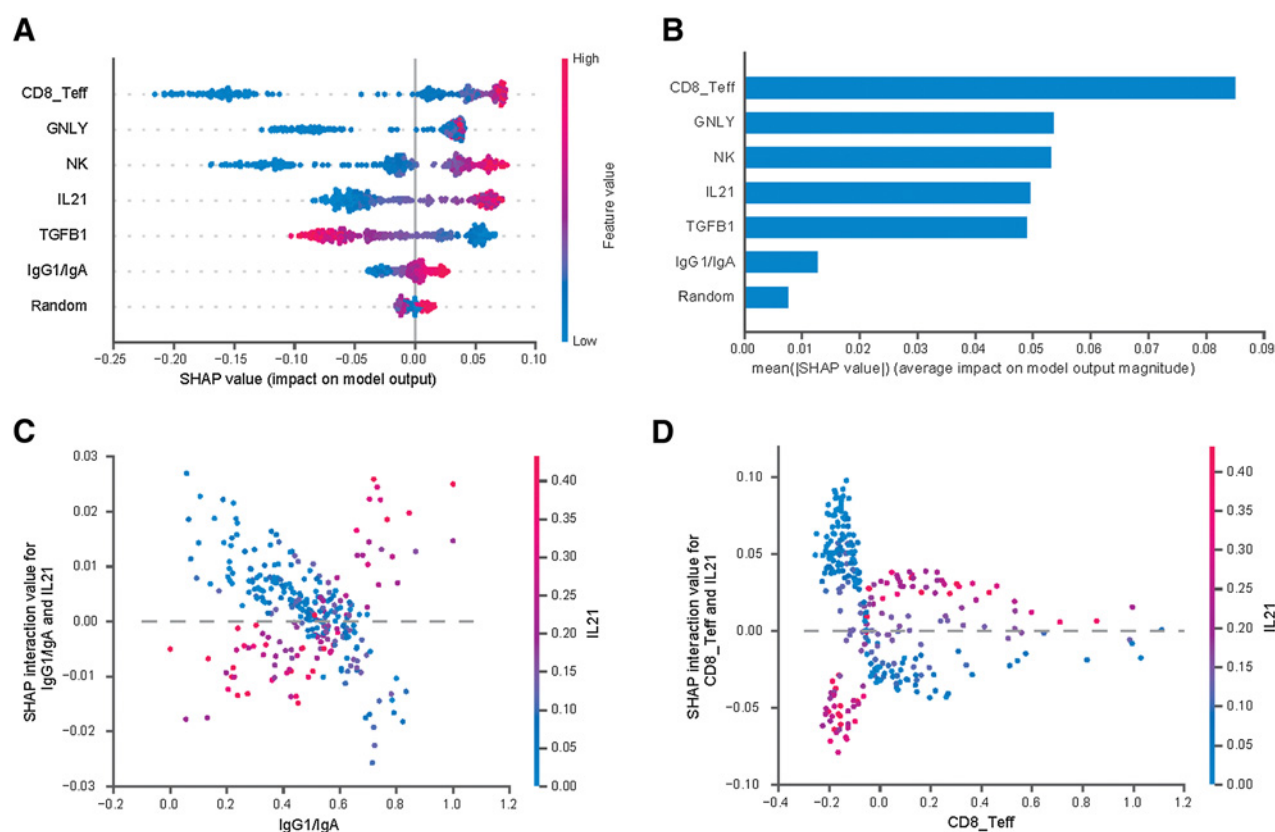


Figure 5. PRIMUS feature importance and interactions. **A** and **B**, PRIMUS feature importance compared with randomly generated numbers estimated with SHAP. **C**, Impact of the interaction between IgG1/IgA ratio and *IL21* expression estimated with SHAP. **D**, Impact of the interaction between refined normalized CD8⁺ T-cell signature and normalized *IL21* expression estimated with SHAP.

PRIMUS variable showed higher feature importance versus randomly generated numbers, with the normalized CD8⁺ T-cell signature having the most significant association with response (Fig. 5A and B). We also identified interactions between the variables. Specifically, a high IgG1/IgA ratio was more predictive of response combined to high *IL21* expression (Fig. 5C), supporting the importance of B-cell differentiation and isotype switching to IgG1 for patient responses. An interaction between high expression of the CD8⁺ T-cell signature and high *IL21* expression was observed (Fig. 5D). We also observed that IgG1/IgA ratio was efficient for response prediction only at the highest values, in contrast to the CD8⁺ T-cell signatures, which showed significant discriminative power across all values (Fig. 5A; Supplementary Fig. S4).

An integrative predictor demonstrates high efficiency for the “desert” phenotype

Among the three tumor immune phenotypes (inflamed, excluded, and desert), treatment response is especially difficult to predict in the “desert” subgroup (53). A CD8⁺ T-cell signature, TGFβ response signature, and tumor mutational burden all fail to predict response in patients with this immune phenotype in the IMVigor210 study (32). In contrast, PRIMUS predicted responses in immune “desert” patients from the IMVigor210 trial (Fig. 6A). However, we recognize a caveat with these data—although our model was generally protected against overfitting, this analysis formally included data used for model

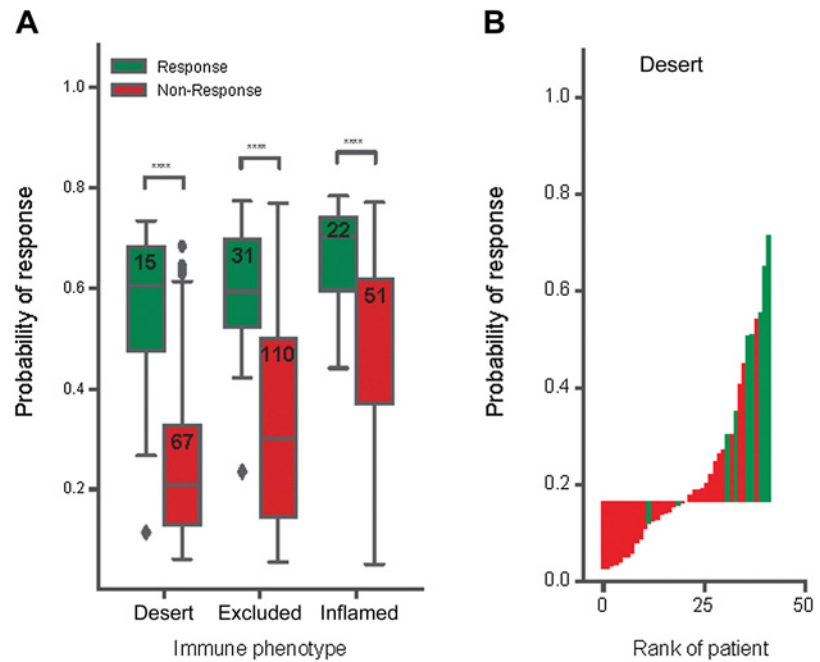
training. To further verify that the PRIMUS model could efficiently predict response among “desert” tumor phenotypes, we trained PRIMUS using the IMVigor210 cohort, with 41 patients with a “desert” phenotype designated as a holdout set. PRIMUS successfully predicted response for these patients (Fig. 6B), thus, demonstrating the utility of PRIMUS for predicting the response to anti-PD-L1/PD1 immunotherapy.

Model validation in an independent dataset

Independent, publicly available datasets that provide transcriptional profiles of patients treated with atezolizumab or other anti-PD-L1 drugs along with relevant clinical data are difficult to acquire. It can be assumed that factors influencing the survival of patients after immunotherapy partially overlap with immune-related parameters that determine the ability of the immune system to control tumor development (54). From this perspective, the assessment of PRIMUS’s ability to predict overall survival in cohorts of patients with bladder cancer who did not receive immunotherapy could alternatively be used to indirectly confirm the model’s performance. On the basis of this logic, we verified the model on the BLCA-TCGA dataset. We performed quantile normalization of the combined dataset, retrained PRIMUS on the normalized data from the IMVigor210 cohort, and then applied the resulting model to the normalized TCGA data. A significant association between the output score of the algorithm and patient

Figure 6.

Performance of PRIMUS for patients with an immune “desert” phenotype, IMVigor210 cohort. **A**, Boxplots show probability of response to anti-PD-L1 immunotherapy for different tumor histologic immune phenotypes, whole IMVigor210 cohort. Median, bottom quartile, top quartile, and interquartile range are shown. ****, $P < 0.0001$. Plotted data include those used for training the model. **B**, Predicted probability of response in a desert holdout set of 41 patients divided by median.



survival was seen for basal squamous subcohort and also for the full TCGA cohort (Supplementary Fig. S5), which confirmed the general relevance of the model for bladder cancer.

Discussion

Several efforts are underway to develop rational prognostic and predictive signatures for patients with BLCA based on the analysis of differentially expressed genes (55–58) or IHC-based classification systems (59). However, most do not take into account the stochastic aspects of tumor sampling. Any tumor sample from a given patient will most likely not fully represent the entirety of the heterogeneous tumor tissue (60, 61). From this point of view, the accuracy of an immune gene signature calculation is intrinsically limited by the stochastic nature of tissue sampling, where the observed variability in immune gene expression results from the abundance of immune cells that infiltrate the particular tissue fragment being sampled. There are few studies of specific biological processes that might be involved in survival and immunotherapy response for patients with bladder cancer that could help in developing rational prediction algorithms (62).

There is supported rationale for using CD8⁺ T cell-specific gene expression (63) and IHC (64) features to predict treatment response. However, an effective predictor of response should take into account diverse components of the immune system, and that such predictors may differ for different types of cancer (65). Multiple studies have demonstrated substantial involvement of CD4⁺ T cells (66, 67), NK cells (68), and B cells (1–5) in cancer immunosurveillance and immunotherapy responses (1, 2). In our approach, we combined T-cell, NK-cell, and B-cell parameters of the TME and accounted for the relative representation of various immune processes. Normalization of the CD8⁺ T-cell signature against pan-leukocyte genes led to improvement in predicting treatment responses and survival after immunotherapy. In parallel, the IgG1/IgA ratio, a parameter that also does not depend on the extent of tumor infiltration and reflects B-cell behavior, independently and prominently improved the prognostic utility of our approach.

By combining rationally preselected parameters, including the normalized CD8⁺ T-cell signature, IgG1/IgA ratio, and a limited number of genes involved in NK-cell responses and T-cell/B-cell interactions, we were able to develop PRIMUS, a model that efficiently predicted response to anti-PD-L1 immunotherapy in muscle-invasive urothelial carcinoma, including tumors with the immune “desert” phenotype. Pending further validation, we hope to pursue clinical implementation of our approach and its derivatives in the near future.

Our results demonstrate the potential for predicting responses to immunotherapy using transcriptomic data. By building on a deeper understanding of the immune processes underlying an effective antitumor response and using relevant statistical approaches, we can make further progress in developing predictors of response to a given therapy or combinations thereof.

Authors’ Disclosures

This work is supported by a grant from the Ministry of Science and Higher Education of the Russian Federation no. 075-15-2020-807.

Authors’ Contributions

I.A. Dyugay: Data curation, investigation, methodology, writing—original draft, writing—review and editing. **D.K. Lukyanov:** Investigation, methodology. **M.A. Turchaninova:** Investigation. **E.O. Serebrovskaya:** Investigation, writing—review and editing. **E.A. Bryushkova:** Investigation. **A.R. Zaretsky:** Conceptualization, investigation, writing—review and editing. **O. Khalmurzaev:** Conceptualization, investigation. **V.B. Matveev:** Conceptualization, investigation, writing—review and editing. **M. Shugay:** Data curation, supervision, methodology. **P.V. Shelyakin:** Data curation, supervision, validation, investigation, methodology. **D.M. Chudakov:** Conceptualization, supervision, funding acquisition, investigation, writing—original draft, project administration, writing—review and editing.

Acknowledgments

We are grateful to Michael Eisenstein for his valuable help in editing the article.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received June 21, 2021; revised August 18, 2021; accepted January 6, 2022; published first January 10, 2022.

References

1. Helmink BA, Reddy SM, Gao J, Zhang S, Basar R, Thakur R, et al. B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* 2020; 577:549–55.
2. Cabrita R, Lauss M, Sanna A, Donia M, Skaarup Larsen M, Mitra S, et al. Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature* 2020;577:561–5.
3. Sautès-Fridman C, Petitprez F, Calderaro J, Fridman WH. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat Rev Cancer* 2019; 19:307–25.
4. Petitprez F, de Reyniès A, Keung EZ, Chen TW-W, Sun C-M, Calderaro J, et al. B cells are associated with survival and immunotherapy response in sarcoma. *Nature* 2020;577:556–60.
5. Sharonov GV, Serebrovskaya EO, Yuzhakova DV, Britanova OV, Chudakov DM. B cells, plasma cells and antibody repertoires in the tumour microenvironment. *Nat Rev Immunol* 2020;20:294–307.
6. Scott AM, Wolchok JD, Old LJ. Antibody therapy of cancer. *Nat Rev Cancer* 2012;12:278–87.
7. Ochoa MC, Minute L, Rodriguez I, Garasa S, Perez-Ruiz E, Inogés S, et al. Antibody-dependent cell cytotoxicity: immunotherapy strategies enhancing effector NK cells. *Immunol Cell Biol* 2017;95:347–55.
8. Carmi Y, Spitzer MH, Linde IL, Burt BM, Prestwood TR, Perlman N, et al. Allogeneic IgG combined with dendritic cell stimuli induce antitumour T-cell immunity. *Nature* 2015;521:99–104.
9. Rossetti RAM, Lorenzi NPC, Yokochi K, Rosa MBS de F, Benevides L, Margarido PFR, et al. B lymphocytes can be activated to act as antigen presenting cells to promote anti-tumor responses. *PLoS One* 2018;13:e0199034.
10. Bruno TC, Ebner PJ, Moore BL, Squalls OG, Waugh KA, Eruslanov EB, et al. Antigen-presenting intratumoral B cells affect CD4+ TIL phenotypes in non-small cell lung cancer patients. *Cancer Immunol Res* 2017; 5:898–907.
11. Rivera A, Chen C-C, Ron N, Dougherty JP, Ron Y. Role of B cells as antigen-presenting cells *in vivo* revisited: antigen-specific B cells are essential for T cell expansion in lymph nodes and for systemic T cell responses to low antigen concentrations. *Int Immunol* 2001;13:1583–93.
12. Ou Z, Wang Y, Liu L, Li L, Yeh S, Qi L, et al. Tumor microenvironment B cells increase bladder cancer metastasis via modulation of the IL-8/androgen receptor (AR)/MMPs signals. *Oncotarget* 2015;6:26065–78.
13. Jiang Q, Fu Q, Chang Y, Liu Z, Zhang J, Xu L, et al. CD19+ tumor-infiltrating B-cells prime CD4+ T-cell immunity and predict platinum-based chemotherapy efficacy in muscle-invasive bladder cancer. *Cancer Immunol Immunother* 2019; 68:45–56.
14. Monteiro RC. The role of IgA and IgA Fc receptors as anti-inflammatory agents. *J Clin Immunol* 2010;30:S61–4.
15. Hansen IS, Baeten DLP, den Dunnen J. The inflammatory function of human IgA. *Cell Mol Life Sci* 2019;76:1041–55.
16. Welinder C, Jirstrom K, Lehn S, Nodin B, Marko-Varga G, Blixt O, et al. Intratumour IgA1 is common in cancer and is correlated with poor prognosis in bladder cancer. *Heliyon* 2016;2:e00143.
17. Bolotin DA, Poslavsky S, Davydov AN, Frenkel FE, Fanchi L, Zolotareva OI, et al. Antigen receptor repertoire profiling from RNA-seq data. *Nat Biotechnol* 2017; 35:908–11.
18. Gilbert AE, Karagiannis P, Dodev T, Koers A, Lacy K, Josephs DH, et al. Monitoring the systemic human memory B cell compartment of melanoma patients for anti-tumor IgG antibodies. *PLoS One* 2011;6:e19330.
19. Isaeva OI, Sharonov GV, Serebrovskaya EO, Turchaninova MA, Zaretsky AR, Shugay M, et al. Intratumoral immunoglobulin isotypes predict survival in lung adenocarcinoma subtypes. *J Immunother Cancer* 2019;7:279.
20. Shi JY, Gao Q, Wang ZC, Zhou J, Wang XY, Min ZH, et al. Margin-infiltrating CD20+ B cells display an atypical memory phenotype and correlate with favorable prognosis in hepatocellular carcinoma. *Clin Cancer Res* 2013;19:5994–6005.
21. TCGA BLCA. Available from: <https://portal.gdc.cancer.gov/projects/TCGA-BLCA>.
22. Zhao S, Ye Z, Stanton R. Misuse of RPKM or TPM normalization when comparing across samples and sequencing protocols. *RNA* 2020;26:903–9.
23. Eckstein M, Cimadamore A, Hartmann A, Lopez-Beltran A, Cheng L, Scarpelli M, et al. PD-L1 assessment in urothelial carcinoma: a practical approach. *Ann Transl Med* 2019;7:690.
24. FDA. FDA limits the use of Tecentriq and Keytruda for some urothelial cancer patients; 2018. Available from: <https://www.fda.gov/drugs/resources-information-approved-drugs/fda-limits-use-tecentriq-and-keytruda-some-urothelial-cancer-patients>.
25. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 2016;34:525–7.
26. Telsh O, Porgador A, Rubin E. Extracting tumor tissue immune status from expression profiles: correlating renal cancer prognosis with tumor-associated immunome. *Oncotarget* 2015;6:33191–205.
27. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* 2001;125:279–84.
28. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res* 2019;47:D419–26.
29. Bolotin DA, Poslavsky S, Mitrophanov I, Shugay M, Mamedov IZ, Putintseva EV, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods* 2015;12:380–1.
30. Tumei PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJM, Robert L, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 2014;515:568–71.
31. Shugay M, Bagaev DV, Turchaninova MA, Bolotin DA, Britanova OV, Putintseva EV, et al. VDJtools: unifying post-analysis of T cell receptor repertoires. *PLoS Comput Biol* 2015;11:e1004503.
32. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* 2018;554:544–8.
33. Segal MR. Machine learning benchmarks and random forest regression; 2004. Available from: <https://escholarship.org/uc/item/35x3v9t4>.
34. Hastie T, Tibshirani R, Friedman J. Random forests. In: Hastie T, Tibshirani R, Friedman J, editors. *The elements of statistical learning: data mining, inference, and prediction*. New York: Springer; 2009. p. 587–604.
35. Hastie T, Tibshirani R, Friedman J. Overview of supervised learning. In: Hastie T, Tibshirani R, Friedman J, editors. *The elements of statistical learning: data mining, inference, and prediction*. New York: Springer; 2009. p. 9–41.
36. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell* 2020;2:56–67.
37. lifelines — lifelines 0.25.6 documentation. Available from: <https://lifelines.readthedocs.io/en/latest/>.
38. Akaike H. A new look at the statistical model identification. *IEEE Trans Autom Control* 1974;19:716–23.
39. Lin LIK. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 1989;45:255–68.
40. Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, et al. Comprehensive molecular characterization of muscle-invasive bladder cancer. *Cell* 2017;171:540–56.
41. Vidotto T, Nersesian S, Graham C, Siemens DR, Koti M. DNA damage repair gene mutations and their association with tumor immune regulatory gene expression in muscle invasive bladder cancer subtypes. *J Immunother Cancer* 2019;7:148.
42. Zhu W, Germain C, Liu Z, Sebastian Y, Devi P, Knockaert S, et al. A high density of tertiary lymphoid structure B cells in lung tumors is associated with increased CD4+ T cell receptor repertoire clonality. *Oncoimmunology* 2015;4:e1051922.
43. Pitzalis C, Jones GW, Bombardieri M, Jones SA. Ectopic lymphoid-like structures in infection, cancer and autoimmunity. *Nat Rev Immunol* 2014;14:447–62.
44. Moroz A, Eppolito C, Li Q, Tao J, Clegg CH, Shrikant PA. IL-21 enhances and sustains CD8+ T cell responses to achieve durable tumor immunity: comparative evaluation of IL-2, IL-15, and IL-21. *J Immunol* 2004;173:900–9.
45. Spolski R, Leonard WJ. IL-21 and T follicular helper cells. *Int Immunol* 2010;22: 7–12.
46. Chen DS, Mellman I. Elements of cancer immunity and the cancer-immune set point. *Nature* 2017;541:321–30.
47. Lane RS, Lund AW. Non-hematopoietic control of peripheral tissue T cell responses: implications for solid tumors. *Front Immunol* 2018;9:2662.
48. Kim J, Kwiatkowski D, McConkey DJ, Meeks JJ, Freeman SS, Bellmunt J, et al. The cancer genome atlas expression subtypes stratify response to checkpoint inhibition in advanced urothelial cancer and identify a subset of patients with high survival probability. *Eur Urol* 2019;75:961–4.

49. Fusi A, Festino L, Botti G, Masucci G, Melero I, Lorigan P, et al. PD-L1 expression as a potential predictive biomarker. *Lancet Oncol* 2015;16:1285–7.
50. Patel SP, Kurzrock R. PD-L1 expression as a predictive biomarker in cancer immunotherapy. *Mol Cancer Ther* 2015;14:847–56.
51. Festino L, Botti G, Lorigan P, Masucci GV, Hipp JD, Horak CE, et al. Cancer treatment with anti-PD-1/PD-L1 agents: is PD-L1 expression a biomarker for patient selection? *Drugs* 2016;76:925–45.
52. Cawley GC, Talbot NLC. On over-fitting in model selection and subsequent selection bias in performance evaluation; 29.
53. Herbst RS, Soria J-C, Kowanetz M, Fine GD, Hamid O, Gordon MS, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* 2014;515:563–7.
54. Masucci GV, Cesano A, Hawtin R, Janetzki S, Zhang J, Kirsch I, et al. Validation of biomarkers to predict response to immunotherapy in cancer: volume I — pre-analytical and analytical validation. *J Immunother Cancer* 2016;4:76.
55. Qiu H, Hu X, He C, Yu B, Li Y, Li J. Identification and validation of an individualized prognostic signature of bladder cancer based on seven immune related genes. *Front Genet* 2020;11:12.
56. Cheng S, Jiang Z, Xiao J, Guo H, Wang Z, Wang Y. The prognostic value of six survival-related genes in bladder cancer. *Cell Death Discov* 2020;6:58.
57. Quan J, Zhang W, Yu C, Bai Y, Cui J, Lv J, et al. Bioinformatic identification of prognostic indicators in bladder cancer. *Biomark Med* 2020;14:1243–54.
58. Chen X, Jin Y, Gong L, He D, Cheng Y, Xiao M, et al. Bioinformatics analysis finds immune gene markers related to the prognosis of bladder cancer. *Front Genet* 2020;11:607.
59. Yuk HD, Jeong CW, Kwak C, Kim HH, Moon KC, Ku JH. Clinical outcomes of muscle invasive bladder cancer according to the BASQ classification. *BMC Cancer* 2019;19:897.
60. Cyll K, Ersvaer E, Vlatkovic L, Pradhan M, Kildal W, Kjær MA, et al. Tumour heterogeneity poses a significant challenge to cancer biomarker research. *Br J Cancer* 2017;117:367–75.
61. Munari E, Zamboni G, Marconi M, Sommaggio M, Brunelli M, Martignoni G, et al. PD-L1 expression heterogeneity in non-small cell lung cancer: evaluation of small biopsies reliability. *Oncotarget* 2017;8:90123–31.
62. Cao R, Yuan L, Ma B, Wang G, Qiu W, Tian Y. An EMT-related gene signature for the prognosis of human bladder cancer. *J Cell Mol Med* 2020;24:605–17.
63. Fridman WH, Zitvogel L, Sautès-Fridman C, Kroemer G. The immune contexture in cancer prognosis and treatment. *Nat Rev Clin Oncol* 2017;14:717–34.
64. Ilie M, Hofman V, Dietel M, Soria J-C, Hofman P. Assessment of the PD-L1 status by immunohistochemistry: challenges and perspectives for therapeutic strategies in lung cancer patients. *Virchows Arch* 2016;468:511–25.
65. Bruni D, Angell HK, Galon J. The immune contexture and Immunoscore in cancer prognosis and therapeutic efficacy. *Nat Rev Cancer* 2020;20:662–80.
66. Kreiter S, Vormehr M, van de Roemer N, Diken M, Löwer M, Diekmann J, et al. Mutant MHC class II epitopes drive therapeutic immune responses to cancer. *Nature* 2015;520:692–6.
67. Borst J, Ahrends T, Båbala N, Melief CJM, Kastenmüller W. CD4 + T cell help in cancer immunology and immunotherapy. *Nat Rev Immunol* 2018;18:635–47.
68. André P, Denis C, Soulas C, Bourbon-Caillet C, Lopez J, Arnoux T, et al. Anti-NKG2A mAb is a checkpoint inhibitor that promotes anti-tumor immunity by unleashing both T and NK cells. *Cell* 2018;175:1731–43.