nature medicine

ARTICLES

https://doi.org/10.1038/s41591-018-0136-1

Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response

Peng Jiang[1,2,10], Shengqing Gu[3,10], Deng Pan[4,5,10], Jingxin Fu[6], Avinash Sahu[1,2], Xihao Hu[1,2], Ziyi Li[6], Nicole Traugh[3], Xia Bu[3], Bo Li[1,2,9], Jun Liu[7], Gordon J. Freeman[3], Myles A. Brown[3,8], Kai W. Wucherpfennig[4,5,11]* and X. Shirley Liu[1,2,6,8,11]*

Cancer treatment by immune checkpoint blockade (ICB) can bring long-lasting clinical benefits, but only a fraction of patients respond to treatment. To predict ICB response, we developed TIDE, a computational method to model two primary mechanisms of tumor immune evasion: the induction of T cell dysfunction in tumors with high infiltration of cytotoxic T lymphocytes (CTL) and the prevention of T cell infiltration in tumors with low CTL level. We identified signatures of T cell dysfunction from large tumor cohorts by testing how the expression of each gene in tumors interacts with the CTL infiltration level to influence patient survival. We also modeled factors that exclude T cell infiltration into tumors using expression signatures from immunosuppressive cells. Using this framework and pre-treatment RNA-Seq or NanoString tumor expression profiles, TIDE predicted the outcome of melanoma patients treated with first-line anti-PD1 or anti-CTLA4 more accurately than other biomarkers such as *PD-L1* level and mutation load. TIDE also revealed new candidate ICB resistance regulators, such as *SERPINB9*, demonstrating utility for immunotherapy research.

Cancer immunotherapies by immune checkpoint blockade (ICB) aim to help the immune system recognize and attack cancer cells[1]. The primary targets of ICB treatment are programmed death-ligand 1 (PD-L1), programmed death 1 (PD1) and cytotoxic T-lymphocyte-associated protein 4 (CTLA4). Compared to conventional therapies, ICB can induce durable responses in patients with metastatic cancers. However, a significant limitation of ICB is that only one-third of patients respond to ICB in most cancer types tested[2]. Combination ICB therapies have shown improved outcomes but also result in more severe side effects than single-agent therapy[1]. Multiple factors can affect ICB effectiveness[2], including the degree of cytotoxic T cell infiltration[3], mutation or neo-antigen load[4], PD-L1 level[5], antigen presentation defects[6], interferon signaling[7], mismatch repair deficiency[8], tumor aneuploidy[9] and intestinal microbiota[10]. However, none of these factors is sufficient to achieve accurate outcome prediction[5], and identification of ICB response biomarkers and resistance regulators is a critical challenge in the field.

Gene expression biomarkers, such as Oncotype DX[11], MammaPrint[12] and Prosigna[13], have demonstrated clinical utility in predicting treatment benefits in breast cancer. We hypothesize that transcriptome signatures could also serve as reliable ICB biomarkers. Ideally, a large number of tumor molecular profiles together with the patient clinical outcome could be used to train a reliable multi-gene biomarker. However, current ICB clinical trials have gene expression profiles on only a small number of pre-treatment samples, which are insufficient to train robust prognostic biomarkers[3,14,15]. Alternatively, there are many public tumor profiling data sets from human and mouse models without immunotherapy, but which are informative regarding tumor immune escape. For example, recent analyses of TCGA and PRECOG data uncovered significant effects of tumor-infiltrating levels of different immune cell types on patient overall survival[16–18]. Predicting tumor response to ICB requires an understanding of how tumors escape the immune system. Therefore, the public tumor molecular profiles, even without ICB treatment, may still be valuable resources to model immune evasion and derive surrogate biomarkers of ICB response.

Recent work has revealed two distinct mechanisms of tumor immune evasion[19,20]. Some tumors have a high level of infiltration by cytotoxic T cells, but these T cells tend to be in a dysfunctional state. In other tumors, immunosuppressive factors may exclude T cells from infiltrating tumors[21]. Therefore, we developed a computational framework, Tumor Immune Dysfunction and Exclusion (TIDE), to identify factors that underlie these two mechanisms of tumor immune escape. TIDE integrated and modeled data from 189 human cancer studies, comprising a total of 33,197 samples. We hypothesized and validated that an accurate gene signature to model the tumor immune escape could serve as a reliable surrogate biomarker to predict ICB response. The web application, source code and analysis results of TIDE are available at http://tide.dfci.harvard.edu.

Results

A statistical interaction test identifies gene signatures of T cell dysfunction. Previous reports showed that cytotoxic T cells could infiltrate a subset of tumors, although they could still fail to control tumor growth if in a dysfunctional state[22]. We reasoned that by combining transcriptome profiles of treatment-naive tumors with

[1]Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA, USA. [2]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [3]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA. [4]Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston, MA, USA. [5]Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA, USA. [6]School of Life Science and Technology, Tongji University, Shanghai, China. [7]Department of Statistics, Harvard University, Cambridge, MA, USA. [8]Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA, USA. [9]Present address: Department of Bioinformatics, UT Southwestern, Dallas, TX, USA. [10]These authors contributed equally: Peng Jiang, Shengqing Gu, Deng Pan. [11]These authors jointly supervised: Kai W. Wucherpfennig, X. Shirley Liu. *e-mail: kai_wucherpfennig@dfci.harvard.edu; xsliu@jimmy.harvard.edu

patient survival outcome, we could identify known regulators of T cell dysfunction. For example, in the TCGA melanoma study, we used the average expression level of *CD8A*, *CD8B*, *GZMA*, *GZMB* and *PRF1* to estimate the cytotoxic T lymphocyte (CTL) level in a tumor[16]. Among metastatic melanoma tumors, a higher CTL level indicates a better patient survival, but only when *TGFB1* has a low expression level (Fig. 1a and Supplementary Fig. 1a). This observation corroborates the known role of the cytokine TGFβ (encoded by *TGFB1*) in promoting tumor immune escape and immunotherapy resistance[2,23,24].

In statistics, two variables interact if the effect of one variable depends on the other variable[25]. In the previous examples, the effect of CTL on survival outcome depends on the *TGFB1* level, which is a typical case of interaction between variables. The interaction of any two variables on survival outcome can be tested by a multiplication term in the Cox proportional hazard (Cox-PH) model[26] (Fig. 1b). The coefficient d of the multiplication term indicates the level of the interaction effect, and the Wald test can evaluate its statistical significance[26]. For example, the *TGFB1* expression level has a significantly antagonistic interaction with CTL level, indicating that a higher *TGFB1* level in tumors will decrease the beneficial association between CTL and overall survival (Supplementary Table 1a). In contrast to *TGFB1*, another gene *SOX10* expression level has a synergistic interaction with CTL level on patient survival outcome, indicating that a higher *SOX10* level in tumors will increase the beneficial association between CTL and survival (Supplementary Table 1b), which is consistent with the known function of *SOX10* to promote T cell-mediated tumor killing[27,28].

We aim to systematically identify genes such as *TGFB1* and *SOX10* that influence the function of cytotoxic T cells on patient survival outcome in cancer genomics data cohorts. Using the Cox-PH model, TIDE tests how the interaction between a candidate gene V and the CTL affects death hazard (estimated from survival) (Fig. 1b). The resulting T cell dysfunctional signature is a genome-wide vector, where the z score of each gene is the interaction coefficient d divided by its standard error (Supplementary Table 1). Genes with significant z scores are not restricted to genes expressed by T cells but could be expressed in cancer cells (for example, *SOX10*[27,28]) or different immune cells associated with T cell dysfunction. In the case of *TGFB1*, both cancer cells[29] and CD4$^+$ FOXP3$^+$ Treg cells[30] can express the cytokine TGFβ to inhibit T cell function.

To compute the T cell dysfunction scores in different cancer data sets, we collected hundreds of data sets from the TCGA[31], PRECOG[17] and METABRIC[32] databases, and focused on 73 that had a minimum of 50 samples with both tumor expression profiles on the genome scale and patient survival data (Supplementary Table 2a). Among the data sets, TIDE predicted different numbers of genes to interact with CTL with statistical significance. For example, the P-value distribution for genes in TCGA melanoma was skewed to the left, indicating many significant genes (Supplementary Fig. 1b). However, the peak of significant P values was absent in TCGA glioblastoma. This difference is likely due to differences in T cell infiltration, data quality or sample size. In five data sets, over 1% of genes have significant interaction with CTL to affect survival at a false discovery rate (FDR) cutoff of 0.1: melanoma, neuroblastoma, triple-negative breast cancer, endometrial cancer and acute myeloid leukemia (Supplementary Fig. 1b and Supplementary Table 2b). For visualization, genes with significant dysfunction scores (FDR < 0.1) in at least two cancer types are shown in Fig. 1c (Supplementary Table 3). Although some of the genes are known to regulate T cell-mediated tumor immunity, such as *PD-L1*, others are likely to be co-expressed with immune-suppressive genes.

The TIDE dysfunction scores are consistent with signatures of tumor immune evasion. We evaluated the quality of TIDE T cell dysfunction scores using published studies of tumor immune evasion in pre-clinical models (Supplementary Table 4). One shRNA screen identified positive or negative hit genes whose knockdown in T cells enhanced or decreased T cell accumulation in mouse tumors, respectively[33]. Gene expression profiles to study T cell dysfunction are also publicly available, including the transcriptome of exhausted CD8 T cells[34], activated regulatory T cells[35] and tumors with acquired ICB resistance[36]. The positive or negative hits are defined as genes upregulated or downregulated in the process of T cell dysfunction or ICB resistance, respectively (Supplementary Tables 4 and 5). We examined whether the TIDE T cell dysfunctional signatures give significantly different scores between positive and negative hit genes in these published studies. We found that TIDE dysfunction signatures averaged from the five clinical cohorts assign positive hits significantly higher dysfunction scores compared to the negative hits (Fig. 2a). Using receiver operating characteristic (ROC) curves, we found that averaging the TIDE dysfunction signatures from the five cohorts gave the best performance (Fig. 2b,c and Supplementary Fig. 2a), suggesting the average profile as a more robust dysfunctional signature.

Recent studies in mouse tumor models revealed two stages of T cell dysfunction[37,38]. While anti-PD1 treatment can revive the early-stage dysfunctional T cells, late-stage dysfunctional T cells are resistant to ICB reprogramming. The average profile of TIDE dysfunction signatures derived from the five cancer cohorts shows increasing correlation with the gene expression profiles of dysfunctional T cells in late stages[38] (Fig. 2d). This result suggests that the TIDE dysfunction signatures reflect the profiles at the late stage of T cell dysfunction. We also applied gene set enrichment analysis to analyze the functional enrichment of TIDE T cell dysfunction signatures. Immune pathways related to inflammatory and interferon response are highly enriched, while mTORC1 signaling[39], protein secretion[40] and glycolysis[41] that are known to promote CD8 T cell activation are consistently depleted (Supplementary Fig. 2b).

Immunosuppressive cell signatures predict immune escape by T cell exclusion. The previous section described T cell dysfunction signatures in tumors with high cytotoxic T cell infiltration. Next, we explored gene signatures of immune evasion through T cell exclusion in tumors with low T cell infiltration[19,20]. Several molecular mechanisms might explain the lack of T cell infiltration in the tumor, such as impaired priming of tumor-specific T cells or suppressive cells prohibiting T cell infiltration into the tumor[19,20]. To model the gene expression signature of T cell exclusion, we examined three cell types reported to restrict T cell infiltration in tumors, namely cancer-associated fibroblasts (CAFs), myeloid-derived suppressor cells (MDSCs) and the M2 subtype of tumor-associated macrophages (TAMs)[20]. We derived a genome-wide signature of T cell exclusion using expression profiles of these cell types from the Gene Expression Omnibus database[42] (Supplementary Table 4). In TCGA melanoma data, tumors whose expression profiles have a higher correlation with the MDSC, TAM or CAF signatures show a significantly lower level of CTLs (Fig. 3a). Moreover, using the average expression profile of MDSCs, TAMs and CAFs to model T cell exclusion, we observed a strong negative correlation between the T cell exclusion scores and the CTL levels across tumors (Fig. 3a). Moreover, the CTL levels and T cell exclusion scores were negatively correlated in all solid tumor data sets (Fig. 3b).

We further examined the associations between the gene signatures of T cell exclusion and T cell dysfunction. For each tumor, the enrichment of a signature is computed as the Pearson correlation between the tumor gene expression profile and the genome-wide scores of T cell exclusion or dysfunction signatures. In the five cancer types where we can identify significant T cell dysfunction scores, the level of T cell exclusion in a tumor inversely correlates with the level of T cell dysfunction (Fig. 3c and Supplementary Fig. 3a). We also calculated the signature enrichment based on

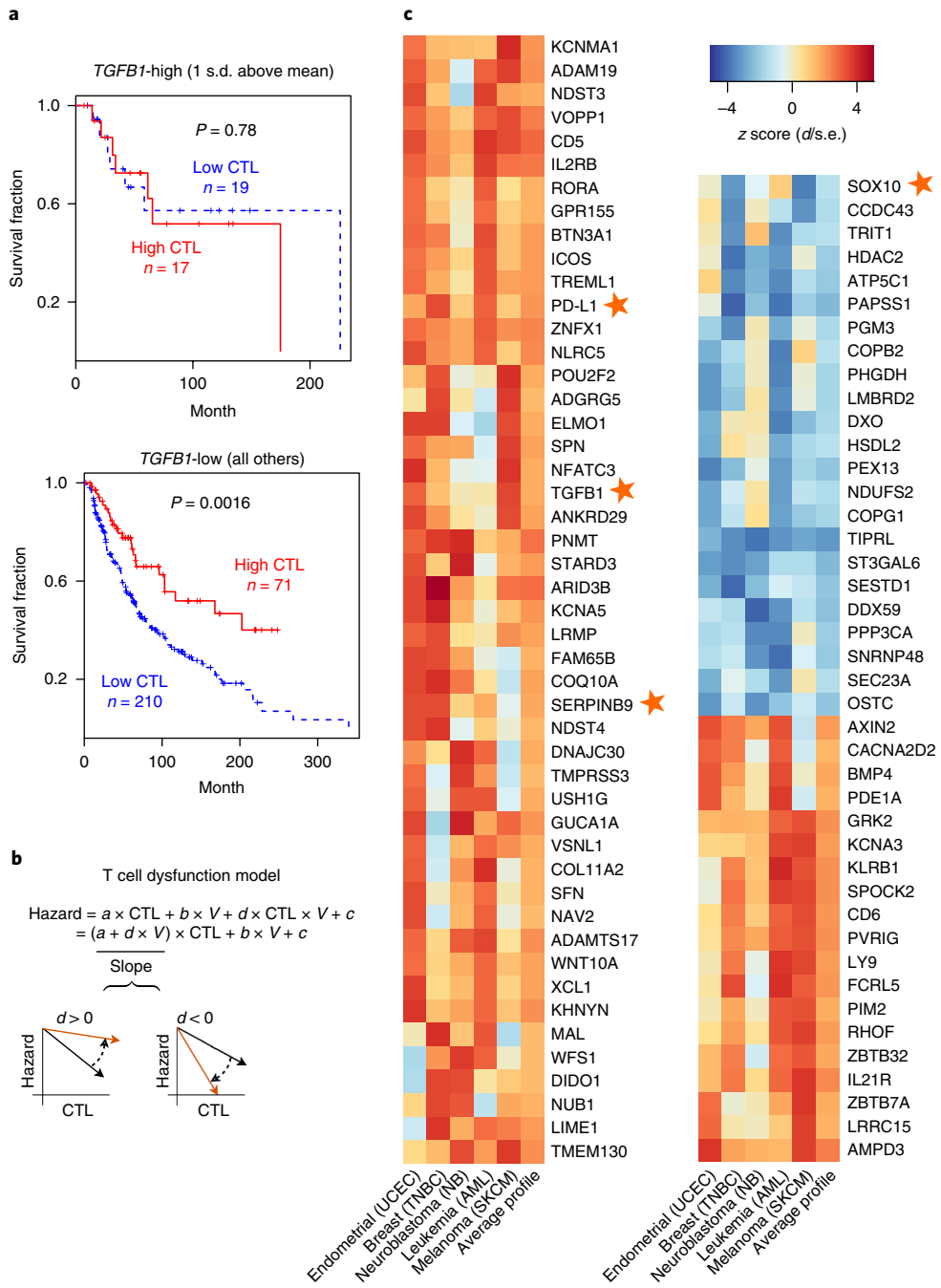


Fig. 1 | The interaction test identifies gene signatures of T cell dysfunction. a, The association between the CTL level and overall patient survival for melanoma tumors with different *TGFB1* levels. For each metastatic melanoma tumor in TCGA, the CTL infiltration level was estimated as the average expression level of *CD8A*, *CD8B*, *GZMA*, *GZMB* and *PRF1*. The association between the CTL level and overall survival was computed through the two-sided Wald test in the Cox-PH regression. Each Kaplan–Meier plot presents tumors in two groups: 'High CTL' (red) have above-average CTL values among all samples, while 'Low CTL' (blue) have values below average. Samples were split according to the *TGFB1* expression level to show the association between the CTL level and survival outcome. The top panel shows tumors with high *TGFB1* expression (one standard deviation above the average), while the bottom panels show the remaining samples. **b**, The interaction test in a Cox-PH regression to identify genes associated with the T cell dysfunction. The variable CTL represents the level of CTLs in each tumor. The variable *V* represents the status of a candidate gene. The coefficient *d* reflects the effect of interaction between the CTL and *V* on death hazard outcome estimated from the survival data. The graphs represent the association slopes between CTL and death hazard. The black and gold arrows represent the association slopes before and after increasing the level of *V*. **c**, Genes with significant T cell dysfunction scores in multiple cancer types. Five data sets, representing five cancer types, had more than 1% of genes passing the FDR threshold 0.1. We display the genes whose T cell dysfunction scores, defined as the *z* score of *d*/standard error (s.e.), are significantly positive or negative (two-sided Wald test *P* values corresponding to an FDR less than 0.1) in at least two cancer types. The orange stars indicate genes of special interest. The number of samples in each cohort is available in Supplementary Table 2b. UCEC, uterine corpus endometrial carcinoma, TNBC, triple-negative breast cancer; AML, acute myeloid leukemia; SKCM, skin cutaneous melanoma; NB, neuroblastoma.

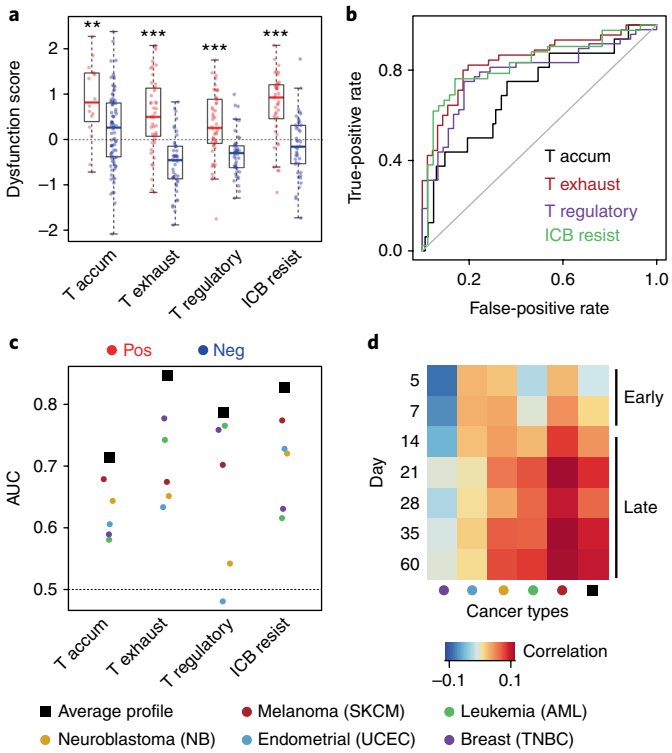


Fig. 2 | T cell dysfunction signatures are consistent with published signatures of tumor immune evasion. a, The consistency between T cell dysfunction signatures predicted by the interaction test and published gene signatures of tumor immune evasion. To evaluate the reliability of the T cell dysfunction gene scores, we collected four published gene signatures related to T cell dysfunction and immunotherapy resistance (Supplementary Table 4). We plotted the T cell dysfunction scores averaged across five cancer types (average profile in Fig. 1c) for the positive (red) and negative (blue) hits of each gene signature. The numbers of positive and negative hits for each signature are available in Supplementary Table 4. Within each group, the scattered dots represent all gene values, and the thick line represents the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). The whiskers encompass 1.5 times the interquartile range. The difference between positive and negative groups was compared through the two-sided Wilcoxon rank-sum test, and *P* values for signatures of 'T accum', 'T exhaust', 'T regulatory' and 'ICB resist' are 6.94×10^{-3}, 1.20×10^{-8}, 1.81×10^{-6} and 1.95×10^{-7}, respectively. The range of *P* values are labeled above each boxplot with asterisks (**$P<1\times10^{-2}$; ***$P<1\times10^{-3}$). T accum: shRNA screens for regulators of T cell accumulation in tumors; T exhaust: transcriptome of exhausted T cells; T regulatory: transcriptome of CD4 regulatory T cells. ICB resist: transcriptome of murine tumors that resist anti-CTLA4 checkpoint blockade. **b**, The ROC curves measuring the performance of the average T cell dysfunction scores (average profile in Fig. 1c) in predicting the positive and negative gene hits in each signature in **a**. **c**, The area under the ROC curve of the average profile of all five cancer types (black squares) and each of the individual cancer types SKCM, AML, NB, UCEC and TNBC with different dot colors. **d**, Pearson correlations between the T cell dysfunction scores and the expression profile of exhausted T cells. The correlations were computed across 12,498 genes shared between human and mouse signatures, for all pairwise combinations between five human cancer types and different time points in a mouse model of T cell exhaustion ('T exh fixed' in Supplementary Table 4).

the differential expression between tumor and normal samples across TCGA cancer types and observed similar negative correlations between T cell exclusion and T cell dysfunction (Fig. 3d and Supplementary Table 6). Kidney renal cell carcinoma (KIRC) has the highest CTL level and the highest enrichment of the T cell dysfunction signature (Fig. 3d and Supplementary Fig. 3b), while lung squamous carcinoma (LUSC) has the highest T cell exclusion signature (Fig. 3d and Supplementary Fig. 3c). Our results are consistent with previous reports of a high CTL level in KIRC and a low CTL level in LUSC[16]. These results suggest that the KIRC and LUSC tumors utilize distinct immune evasion strategies, with KIRC operating more through T cell dysfunction and LUSC through T cell exclusion. Previous studies reported paradoxical observations that in KIRC the degree of CD8 cytotoxic T cell infiltration is anti-correlated with survival benefits[43]. Our analysis revealed that KIRC tumors with higher CTL levels tend to have a stronger T cell dysfunction signature, which could impair the ability of cytotoxic T cells to kill cancer cells (Supplementary Fig. 3b).

TIDE signature predicts ICB response. In previous sections, we developed genome-wide expression signatures to measure the level of T cell dysfunction and T cell exclusion in tumors. We next examined whether integrating these two signatures could predict ICB clinical response. Among the five cancer types for which we could compute reliable TIDE signatures (Fig. 1c), only melanoma has publicly available data on tumor expression and clinical outcome of patients treated with anti-PD1[14] or anti-CTLA4[3], so it was the focus of our evaluation. We also evaluated TIDE on an anti-PD1 data set that profiled tumor expression profiles across four cancer types using the NanoString assay on a few hundred genes[15].

We classified the tumors as CTL-high if the expression levels of all CTL markers (*CD8A*, *CD8B*, *GZMA*, *GZMB* and *PRF1*) were higher than their average values in each data set and the remaining tumors as CTL low. In the CTL-high tumors, TIDE correlates the tumor expression data with the T cell dysfunction signature and predicts tumors with high correlation to T cell dysfunction as non-responders (Supplementary Fig. 4a). In CTL-low tumors, it has been reported that ICB can enhance the cytotoxic T cell infiltration[44,45], so patients with low tumor CTL might still derive clinical benefits from immunotherapies. Therefore, TIDE correlates the expression data for each tumor with the T cell exclusion signature in CTL-low tumors and predicts those with suppressive cells inhibiting T cell infiltration as non-responders (Supplementary Fig. 4a). Notably, the correlation between tumor expression profiles and TIDE signatures is a single value computed across all human genes (Supplementary Fig. 4b), and therefore not subject to multiple-hypotheses testing and less sensitive to the noise from individual expression or the TIDE signature value. For the pre-treatment transcriptome of each tumor, the Pearson correlation with either T cell dysfunction (in CTL-high tumors) or exclusion (in CTL-low tumors) signatures was defined as the TIDE prediction score (Fig. 4a–c). Correlations with T cell dysfunction or exclusion signatures may have different distributions (Supplementary Fig. 4c). Thus, when merging the prediction scores from two tumor CTL categories, we normalized the correlations by their standard deviations in the TCGA data. Finally, all tumors were ranked by their TIDE scores to predict their ICB response (Fig. 4a–c and Supplementary Fig. 4d).

To evaluate the prediction performance for ICB response, we used ROC to measure the true-positive rates against the false-positive rates at various thresholds of TIDE prediction scores (Fig. 4d–f). Compared to widely used ICB response biomarkers, tumor mutation load, *PD-L1* level and interferon gamma response[5,7], the TIDE signature achieved consistently better performance for both anti-PD1 and anti-CTLA4 therapies using both RNA-Seq and NanoString data (Fig. 4d–f and Supplementary Fig. 5a). We also compared TIDE with other ICB response signatures reported in the literature (Supplementary Table 7). Among all candidate biomarkers, we found the TIDE signature to be the best predictor for both anti-PD1 and anti-CTLA4 therapies (Fig. 4g–i and Supplementary Table 8a). The prediction performance of TIDE is also higher than

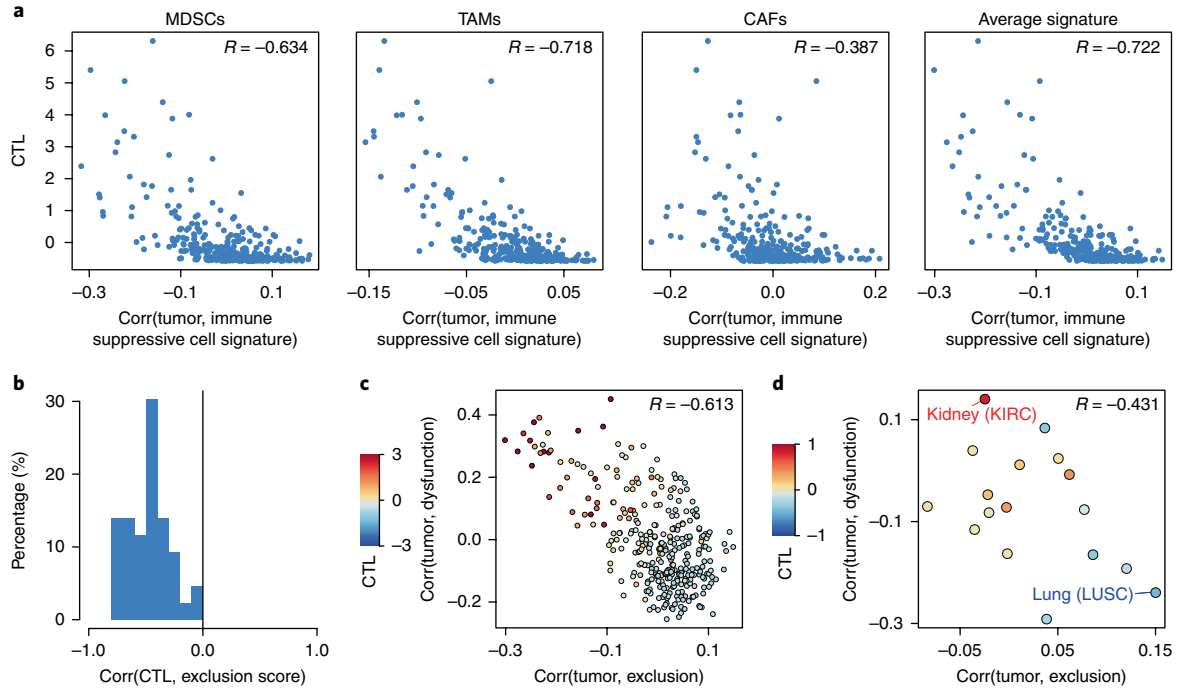


Fig. 3 | Immunosuppressive cell expression models gene signatures of T cell exclusion. a, Prediction of T cell exclusion scores for tumors using immunosuppressive cell signatures. For each metastatic tumor in the TCGA melanoma data set (blue dots, $n=317$), we computed the Pearson correlation between its expression profile and the expression signature of MDSCs, M2 TAMs or CAFs (Supplementary Table 4) or the average of the three expression signatures. In each graph, these values are plotted along the *x* axis. The *y* axis shows the CTL level for each tumor (average expression level of *CD8A*, *CD8B*, *GZMA*, *GZMB* and *PRF1*). The Pearson correlation (*R*) between the plotted values is shown in the upper right corner of each plot. The two-sided *t*-test *P* values for correlations in MDSCs, TAMs, CAFs and mean are 4.61×10^{-37}, 1.58×10^{-51}, 8.84×10^{-13} and 2.58×10^{-52}, respectively. **b**, A histogram of the correlations between the CTL levels and the T cell exclusion scores across tumors. The correlations analyzed in the histogram correspond to the *R* value in the top right corner of **a** (example of TCGA melanoma) across 43 solid tumor data sets. Gliomas are excluded because of low T cell infiltration levels in most gliomas[16]. **c**, Anti-correlation between T cell dysfunction scores and exclusion scores across TCGA melanoma tumors. For each metastatic melanoma tumor (colored dots, $n=317$), we computed the Pearson correlation between the sample's expression profile and the TIDE T cell dysfunction signature (*y* axis). The same computation was made between the tumor expression profile and the TIDE T cell exclusion signature (*x* axis). The Pearson correlation between the plotted values is shown in the upper right (two-sided *t*-test *P* value $=4.02\times10^{-34}$). The dot color indicates the CTL level in each tumor. **d**, Anti-correlation between T cell dysfunction scores and exclusion scores across TCGA cancer types. For each TCGA cancer type with normal control samples ($n=17$), we calculated the average expression difference between tumor versus normal samples. We then computed the Pearson correlation between that value and the TIDE T cell dysfunction signature (*y* axis). We also made the same calculation for the TIDE T cell exclusion signature (*x* axis). The Pearson correlation between the plotted values is shown in the upper right (one-sided *t*-test *P* value $=0.042$). The CTL level difference between tumor and normal samples is shown by the dot color.

the signatures of T cell dysfunction and ICB resistance discussed in Fig. 2a (Supplementary Fig. 5b). Meanwhile, the TIDE prediction performance is robust against modest variations of CTL cutoff in the definition of CTL-high or -low tumors (Supplementary Fig. 5c). Moreover, a higher tumor TIDE prediction score is associated not only with worse ICB response, but also with worse patient survival under anti-PD1 and anti-CTLA4 therapies (Fig. 4j–l). One explanation for the better performance of TIDE is that TIDE utilized both T cell dysfunction and exclusion signatures to model immune escape in tumors with different CTL levels, while other biomarkers consider only one aspect (Supplementary Fig. 5d–f and Supplementary Table 8b,c). Paradoxically, a previous computational method ImmunoPhenoScore claimed to have 100% accuracy in predicting ICB response in melanoma[46], but we observed considerably lower accuracy of ImmunoPhenoScore using the source codes provided by the authors (Fig. 4g–i).

Besides the anti-PD1 RNA-Seq cohort[14] analyzed in Fig. 4, a recent study generated RNA-Seq profiles on another melanoma cohort treated with anti-PD1[45]. We focused on 24 patients with genomics profiles (expression and mutation) of pre-treatment tumors and anti-PD1 as the first-line immunotherapy (without previous anti-CTLA4 therapy). While the TIDE prediction score has a similar accuracy to the mutation load (Supplementary Fig. 6a,b), it is significantly predictive of the patient overall survival ('Ipi naive' in Supplementary Fig. 6c), demonstrating its prognostic value. We noted that TIDE achieved a lower prediction performance in the Riaz study compared to its performance in the Hugo study (Fig. 4d versus 'Ipi naive' in Supplementary Fig. 6). A possible explanation is that the Riaz study[45] used the RECIST v1.1 criteria for disease progression, while the Hugo study[14] used the immune-related RECIST[47], which is more appropriate for predicting immunotherapy response. Further, TIDE is trained using data from ICB-naive tumors and thus not relevant in modeling the tumors that progressed after a first-line ICB[2] ('Ipi progressed' in Supplementary Fig. 6).

The TIDE dysfunction score predicts regulators of ICB resistance. We hypothesized that some genes with high scores in TIDE signatures might serve not only as biomarkers but also as ICB resistance regulators. We focused on the T cell dysfunction signature for genes regulating T cell dysfunction in tumors. As the T cell dysfunction scores were computed using the data from treatment-naive tumors, we utilized orthogonal data from a mouse

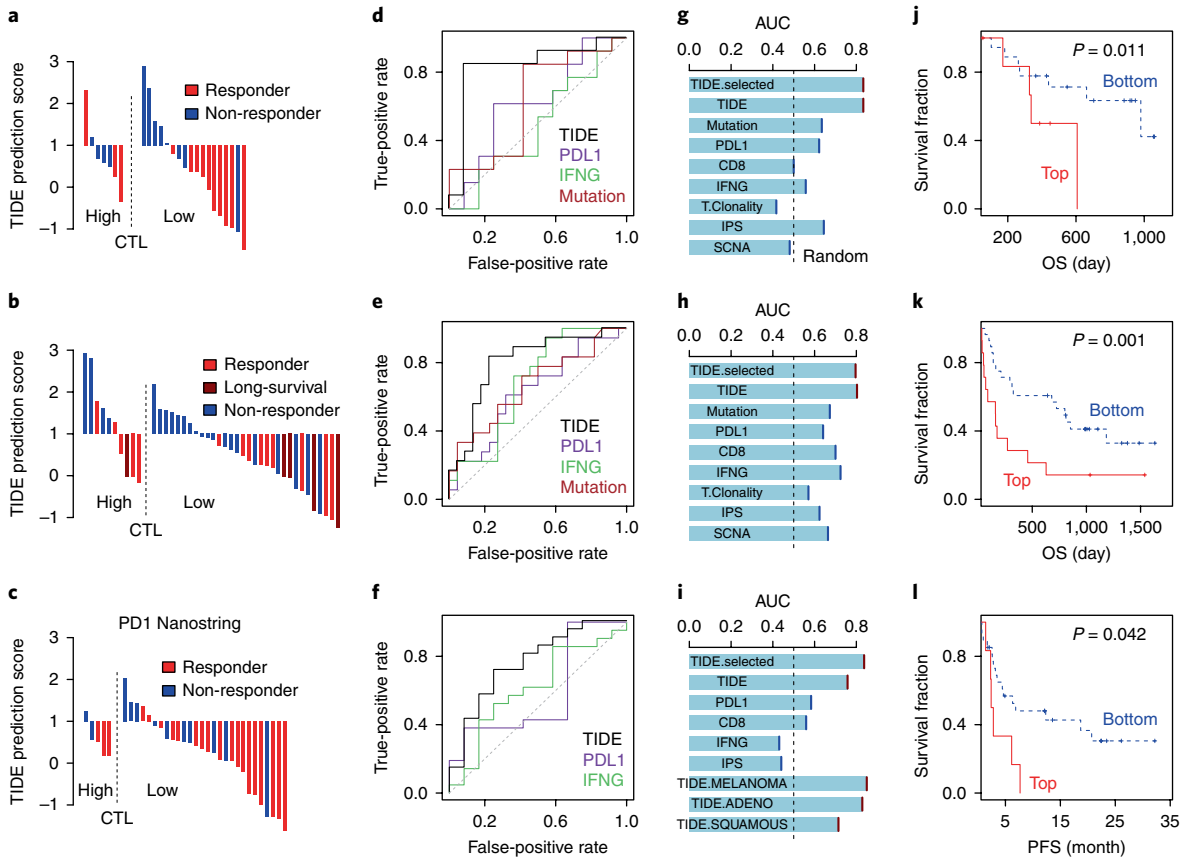


Fig. 4 | TIDE signatures predict ICB immunotherapy response. a, A waterfall plot of TIDE prediction scores across 25 melanoma tumors treated with anti-PD1[14]. The TIDE framework divided tumors into CTL-high or -low categories based on the expression level of CTL marker genes (Supplementary Fig. 4a). Red indicates a tumor that responded to therapy. Blue indicates a non-responder. In each category, we sorted tumors in descending order according to their TIDE prediction scores. **b**, A waterfall plot of TIDE prediction scores across 42 melanoma tumors treated with anti-CTLA4[3] in the same way as in **a**. Besides responders or non-responders, several patients are classified as long-survival in the original study due to the long overall survival time[3]. **c**, A waterfall plot of TIDE prediction scores across 33 tumors treated with anti-PD1[15] in the same way as in **a**. The gene expression profiles are measured by the NanoString platform. The 33 tumors comprise 9 melanoma, 12 lung adenocarcinoma, 9 lung squamous carcinoma and 3 head and neck tumors. **d**, ROC curves for the performance of the TIDE prediction score, *PD-L1* expression, interferon gamma (IFNG) response and total mutation load in predicting anti-PD1 response among 25 melanoma tumors in **a**. **e**, ROC curves for the performance of several signatures in predicting anti-CTLA4 response among 42 melanoma tumors in **b**. **f**, ROC curves for the performance of several signatures in predicting anti-PD1 response among 33 tumors in **c**. **g**, The area under the ROC curve (AUC) for several signatures in predicting anti-PD1 response among 25 melanoma tumors in **a**. The signatures are defined in Supplementary Table 7, with TIDE in dark red and other signatures in blue. Besides the genome-wide TIDE signature, 'TIDE.selected' is a variation focused on 770 genes with both high expression variation across tumors and significant values in the either T cell dysfunction or exclusion signatures. The performance of a random predictor (AUC = 0.5) is represented by the dashed line. **h**, AUC for signatures in predicting anti-CTLA4 response among 42 melanoma tumors in **b** in the same way as in **g**. **i**, AUC for signatures in predicting anti-PD1 response among 33 tumors in **c** in the same way as in **g**. TIDE AUC metrics are also shown separately for nine melanoma, twelve lung adenocarcinoma (Adeno) and nine lung squamous carcinoma (Squamous) tumors. **j**, Kaplan–Meier plots of overall survival (OS) for 25 melanoma patients (in **a**) treated with anti-PD1 with the top (>1) and bottom (<1) TIDE prediction scores. The *P* value was calculated by testing the association between TIDE prediction scores and overall survival with the two-sided Wald test in a Cox-PH regression. **k**, Kaplan–Meier plots of overall survival for 42 melanoma patients (in **b**) treated with anti-CTLA4 in the same way as in **j**. **l**, Kaplan–Meier plots of progression-free survival (PFS) for 33 patients (in **c**) treated with anti-PD1 in the same way as in **j**.

model of acquired anti-CTLA4 resistance to identify genes that are associated with ICB resistance[36]. We ranked genes with significant T cell dysfunction scores in Fig. 1c by the expression fold-change in the ICB-resistant tumors[36] and identified *Serpinb9* as the most upregulated gene (Fig. 5a,b). In ICB clinical cohorts, the *SERPINB9* expression level is consistently lower in responders than non-responders (Supplementary Fig. 7a). Moreover, *SERPINB9* expression alone is significantly associated with worse overall survival in two clinical studies of anti-CTLA4 therapy[3,48] (Fig. 5c, Supplementary Fig. 7b and Supplementary Table 9).

SERPINB9 is a member of the serine protease inhibitor (serpin) family. The encoded protein can inactivate granzyme B to protect lymphocytes (for example, T cells and natural killer cells) from granzyme that may leak from the granules[49]. It is normally expressed in cytotoxic lymphocytes, antigen-presenting cells and immune-privileged sites[50–52]. Meanwhile, a study using in vitro cell culture models reported that a high *SERPINB9* level in cancer cells resulted in resistance to T cell-mediated killing[53]. To infer which cell type in tumors is the potential source of a high *SERPINB9* level, we examined the Protein Atlas database of immunohistochemistry results for 15,000 genes in 20 cancer types[54]. The SERPINB9 protein is expressed at a higher level in cancer cells of melanoma and several other cancer types as compared to normal tissues (Supplementary Fig. 7c,d). Thus, *SERPINB9* may promote the resistance to T cell-mediated killing during ICB therapy through its high expression in cancer cells.

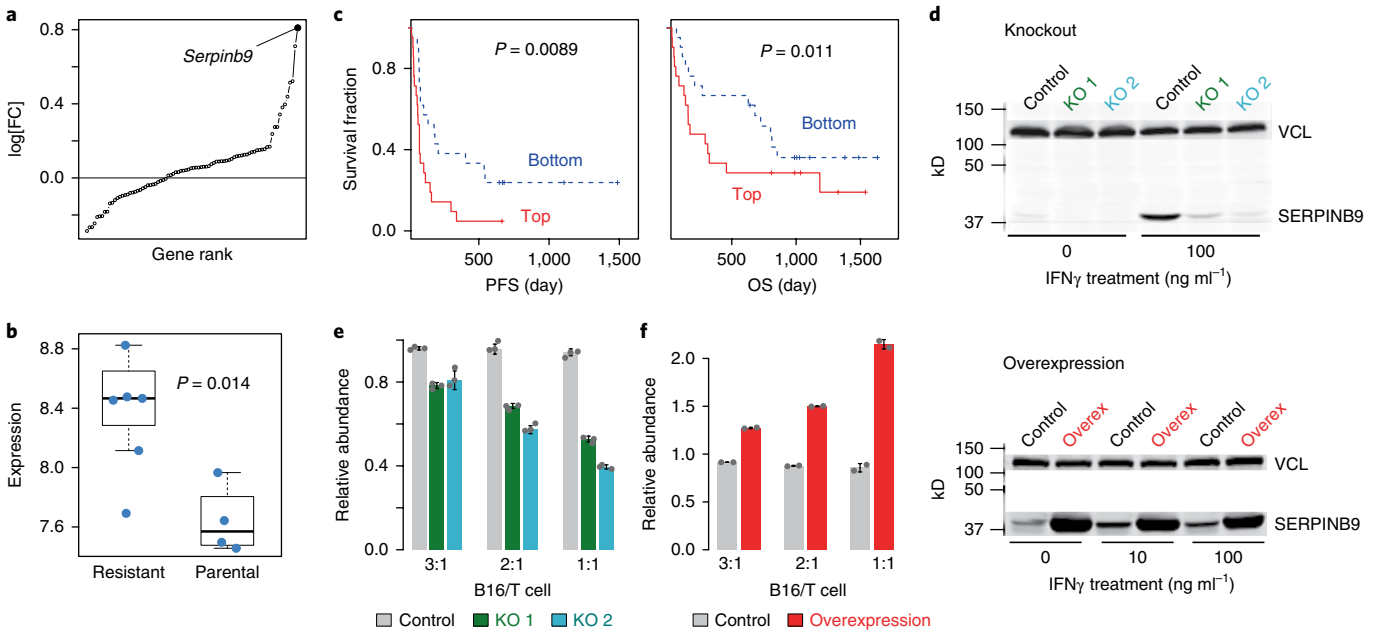


Fig. 5 | Validation of SERPINB9 as a regulator of tumor immune escape. a, The log-fold change (log[FC]) of expression between anti-CTLA4-resistant and parental B16 murine tumors for genes with significant T cell dysfunction scores in Fig. 1c. All genes are ranked increasingly with the top one labeled by name. **b**, The expression value of *Serpinb9* between anti-CTLA4-resistant and parental B16 tumors. Within each group, the scattered dots represent *Serpinb9* expression values ($n=6$ samples in the resistant group, $n=4$ samples in the parental group). The thick line represents the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). The whiskers encompass 1.5 times the interquartile range. The *P* value, testing the group difference, was calculated with the two-sided Wilcoxon rank-sum test. **c**, Kaplan–Meier plots of patients with top half and bottom half *SERPINB9* expression levels, using the data from an anti-CTLA4 study with 42 patients profiled[3]. Both progression-free survival and overall survival are shown. The association between *SERPINB9* expression and patient survival was tested by the two-sided Wald test in a Cox-PH regression (Supplementary Table 9). **d**, Western blot of SERPINB9 following genetic knockout and overexpression. For knockout (KO), there are two independent CRISPR guides targeting *Serpinb9* and a control non-targeting sequence. Cells were either untreated (left 3 lanes) or treated with 100 ng ml^{-1} IFNγ to induce *Serpinb9* expression (right 3 lanes). VCL is the loading control. For overexpression, the open reading frame of *Serpinb9* was cloned into the pEF1a plasmid and overexpressed in B16F10 cells. The protein level was compared between pEF1a backbone- and pEF1a-*Serpinb9*-transduced cells. The bands of related protein targets are cut and shown. All experiments have been repeated independently two times with similar results. **e**, The effect of *Serpinb9* knockout on T cell-mediated tumor killing. B16F10 cancer cells were co-cultured for three days with cytotoxic T cells at three B16F10 to T cell ratios (3:1, 2:1 or 1:1). Each CRISPR gRNA-transduced GFP positive cell line (Control, KO 1, KO 2) was mixed with the parental GFP-negative cell line at a 1:1 ratio. After co-culture, the ratio of edited GFP$^+$ cells to parental cells (GFP$^-$) was determined by flow cytometry. The bar plots present the median value among three cell-culture replicates with standard deviations as the error bars. The results of the two-sided Student *t*-test, comparing the difference between knockout and control conditions, are available in Supplementary Table 10. **f**, The effect of *Serpinb9* overexpression on T cell-mediated tumor killing. The effect of *Serpinb9* overexpression was examined. The bar plots present the median value among two cell-culture replicates with standard deviations as the error bars. The results of the two-sided Student *t*-test, comparing the difference between overexpression and control conditions, are available in Supplementary Table 10.

To validate *SERPINB9* function in cancer cells, we knocked out *Serpinb9* using CRISPR–Cas9 in the murine B16F10 melanoma cell line, which is the parental line of the anti-CTLA4-resistant tumor model previously discussed[36]. After knocking out *Serpinb9* using two different CRISPR guide RNAs (gRNAs), the protein level became undetectable (Fig. 5d and Supplementary Fig. 8). When co-cultured with Pmel-1 cytotoxic T cells, the *Serpinb9*-knockout B16F10 cells were more sensitive to T cell-mediated killing compared to control cells (Fig. 5e, Supplementary Fig. 9a and Supplementary Table 10). In contrast, B16F10 cells with *Serpinb9* overexpression were significantly more resistant to T cell-mediated killing compared to control cells (Fig. 5f, Supplementary Fig. 9b and Supplementary Table 10).

Notably, in B16F10 cells, the SERPINB9 protein level is significantly increased on treatment with IFNγ, a cytokine produced by cytotoxic T cells following antigen-specific activation[55] (Fig. 5d). This SERPINB9 induction following IFNγ treatment might be explained by the binding of IRF1, a transcription factor activated by IFNγ signaling[56], near the *Serpinb9* gene that is observed in public ChIP-Seq data sets (Supplementary Fig. 10a,b). In human melanoma tumors, the expression level of *SERPINB9* is highly correlated with *IRF1* on both bulk tumor and single-cell levels (Supplementary Fig. 10c,d). These results support that *SERPINB9* in cancer cells regulates resistance to T cell-mediated killing, which is essential for ICB response. Interestingly, the *SERPINB9* expression level is also consistently upregulated following pathogen infection in curated studies from the NCBI Gene Expression Omnibus[42] and Expression Atlas[57] databases (Supplementary Fig. 11 and Supplementary Table 11). This result indicates that *SERPINB9* is potentially a general regulator of immune evasion utilized by both tumors and pathogens.

Discussion

We developed a computational method called TIDE, which integrates the expression signatures of T cell dysfunction and T cell exclusion to model tumor immune evasion. The TIDE signatures, trained from treatment-naive tumor data, can predict ICB clinical response based on pre-treatment tumor profiles. Furthermore, TIDE predicted regulators of ICB resistance whose inhibition might improve ICB response. We experimentally validated the role of SERPINB9, an inhibitor of the cytotoxic lymphocyte protease GZMB, in tumor immune evasion, which is an essential process of ICB resistance. Although no small-molecule inhibitor of SERPINB9 is available, the Pfizer OASIS database indicates this protein as potentially druggable[58].

Of the 73 data sets analyzed in this study, only five gave statistically significant T cell dysfunction signatures from the interaction test (Supplementary Table 2). This result is partly because we considered only 48 out of the 73 data sets where a higher level of tumor-infiltrating cytotoxic T cells is correlated with better survival outcome. In some cancer types, such as renal cell carcinoma, which has a substantial level of CD8 T cell infiltration, higher CTL may not be associated with survival benefits[43]. Also, depending on the sample size or characteristics of specific data sets, there might not be any statistically significant genes interacting with CTL to influence survival. Since averaging signatures from the five data sets yielded a signature more robust than any individual signature, integrating additional cancer data sets in the future has the potential to further improve the robustness of the T cell dysfunction scores (Fig. 2c). With additional data, cancer-type-specific regulators may be identified on the basis of the biological variations of T cell dysfunction scores across different cancer types.

When using the TIDE model to predict ICB response, we determined a cutoff to classify tumors as CTL-high or CTL-low. We used the average expression of CTL markers (*CD8A*, *CD8B*, *GZMA*, *GZMB* and *PRF1*) across all tumors to determine the CTL threshold. However, if matched normal tissues are available, the CTL threshold could also be determined by comparing the CTL marker expression in tumors with marker expression in normal tissues. The TIDE signature consists of genome-wide scores of T cell dysfunction and exclusion. While a genome-wide transcriptome biomarker might be robust for ICB response prediction, RNA-Seq has not been widely adopted in the clinic. A smaller gene panel for qPCR or NanoString assays could be more clinically pragmatic. We demonstrated TIDE performance on an anti-PD1 response data set where baseline tumor expression was measured on the NanoString PanCancer panel (Fig. 4c).

One limitation of our study is that we focused primarily on gene expression biomarkers. However, other biomarker types can also predict T cell infiltration and ICB response. For example, beta-catenin protein level has a negative correlation with CTL in many cancer types (Supplementary Fig. 3d,e and Supplementary Table 12). Moreover, tumors initially responding to ICB may later acquire mutations, such as in *B2M*, *IFNGR1/2* and *JAK1/2* genes, to become ICB resistant[2]. However, our study focuses only on predicting intrinsic ICB resistance. Therefore, more data types and methods are necessary to model the immunotherapy efficacy comprehensively. It is possible that TIDE could be applied jointly with other types of biomarker to achieve a higher prediction performance.

To enable testing of TIDE by clinicians and the public, we created a web application for response prediction using transcriptome profiles at http://tide.dfci.harvard.edu. TIDE has the potential to help oncologists select patients who are more likely to benefit from ICB. It would be of significant interest to test the clinical utility of TIDE in ICB decision-making in a prospective clinical trial. New immune-oncology data are emerging at an increasingly rapid pace. We envision that computational modeling and data integration will be increasingly utilized to refine ICB response biomarkers and identify new immunotherapy targets.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41591-018-0136-1.

Received: 2 November 2017; Accepted: 27 June 2018;
Published online: 20 August 2018

References

1. Mahoney, K. M., Rennert, P. D. & Freeman, G. J. Combination cancer immunotherapy and new immunomodulatory targets. *Nat. Rev. Drug Discov.* **14**, 561–584 (2015).
2. Sharma, P., Hu-Lieskovan, S., Wargo, J. A. & Ribas, A. Primary, adaptive, and acquired resistance to cancer immunotherapy. *Cell* **168**, 707–723 (2017).
3. Van Allen, E. M. et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207–211 (2015).
4. Snyder, A. et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *New Engl. J. Med.* **371**, 2189–2199 (2014).
5. Nishino, M., Ramaiya, N. H., Hatabu, H. & Hodi, F. S. Monitoring immune-checkpoint blockade: response evaluation and biomarker development. *Nat. Rev. Clin. Oncol.* **14**, 655–668 (2017).
6. Zaretsky, J. M. et al. Mutations associated with acquired resistance to PD-1 blockade in melanoma. *N. Engl. J. Med.* **375**, 819–829 (2016).
7. Ayers, M. et al. IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Investig.* **127**, 2930–2940 (2017).
8. Le, D. T. et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N. Engl. J. Med.* **372**, 2509–2520 (2015).
9. Davoli, T., Uno, H., Wooten, E. C. & Elledge, S. J. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* **355**, eaaf8399 (2017).
10. Kroemer, G. & Zitvogel, L. Cancer immunotherapy in 2017: The breakthrough of the microbiota. *Nat. Rev. Immunol.* **18**, 87–88 (2018).
11. Paik, S. et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* **351**, 2817–2826 (2004).
12. van 't Veer, L. J. et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536 (2002).
13. Parker, J. S. et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009).
14. Hugo, W. et al. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* **165**, 35–44 (2016).
15. Prat, A. et al. Immune-related gene expression profiling after PD-1 blockade in non-small cell lung carcinoma, head and neck squamous cell carcinoma, and melanoma. *Cancer Res.* **77**, 3540–3550 (2017).
16. Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48–61 (2015).
17. Gentles, A. J. et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat. Med.* **21**, 938–945 (2015).
18. Li, B. et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biol.* **17**, 174 (2016).
19. Gajewski, T. F., Schreiber, H. & Fu, Y. X. Innate and adaptive immune cells in the tumor microenvironment. *Nat. Immunol.* **14**, 1014–1022 (2013).
20. Joyce, J. A. & Fearon, D. T. T cell exclusion, immune privilege, and the tumor microenvironment. *Science* **348**, 74–80 (2015).
21. Spranger, S. & Gajewski, T. F. Tumor-intrinsic oncogene pathways mediating immune avoidance. *Oncoimmunology* **5**, e1086862 (2016).
22. Wherry, E. J. & Kurachi, M. Molecular and cellular insights into T cell exhaustion. *Nat. Rev. Immunol.* **15**, 486–499 (2015).
23. Tauriello, D. V. F. et al. TGFβ drives immune evasion in genetically reconstituted colon cancer metastasis. *Nature* **554**, 538–543 (2018).
24. Mariathasan, S. et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* **554**, 544–548 (2018).
25. Freedman, D. *Statistical Models: Theory and Practice*. (Cambridge Univ. Press, Cambridge, 2009).
26. Kleinbaum, D. G. Survival analysis, a self-learning text. *Biom. J.* **40**, 107–108 (1998).
27. Patel, S. J. et al. Identification of essential genes for cancer immunotherapy. *Nature* **548**, 537–542 (2017).
28. Khong, H. T. & Rosenberg, S. A. The Waardenburg syndrome type 4 gene, *SOX10*, is a novel tumor-associated antigen identified in a patient with a dramatic response to immunotherapy. *Cancer Res.* **62**, 3020–3023 (2002).
29. Thomas, D. A. & Massague, J. TGF-β directly targets cytotoxic T cell functions during tumor evasion of immune surveillance. *Cancer Cell* **8**, 369–380 (2005).
30. Woo, E. Y. et al. Regulatory $CD4^+CD25^+$ T cells in tumors from patients with early-stage non-small cell lung cancer and late-stage ovarian cancer. *Cancer Res.* **61**, 4766–4772 (2001).
31. The Cancer Genome Atlas Research Network et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* **45**, 1113–1120 (2013).
32. Curtis, C. et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).
33. Zhou, P. et al. In vivo discovery of immunotherapy targets in the tumour microenvironment. *Nature* **506**, 52–57 (2014).
34. Giordano, M. et al. Molecular profiling of CD8 T cells in autochthonous melanoma identifies Maf as driver of exhaustion. *EMBO J.* **34**, 2042–2058 (2015).
35. Wakamatsu, E., Mathis, D. & Benoist, C. Convergent and divergent effects of costimulatory molecules in conventional and regulatory $CD4^+$ T cells. *Proc. Natl Acad. Sci. USA* **110**, 1023–1028 (2013).
36. Twyman-Saint Victor, C. et al. Radiation and dual checkpoint blockade activate non-redundant immune mechanisms in cancer. *Nature* **520**, 373–377 (2015).

37. Schietinger, A. et al. Tumor-specific T cell dysfunction is a dynamic antigen-driven differentiation program initiated early during tumorigenesis. *Immunity* **45**, 389–401 (2016).
38. Philip, M. et al. Chromatin states define tumour-specific T cell dysfunction and reprogramming. *Nature* **545**, 452–456 (2017).
39. Pollizzi, K. N. et al. mTORC1 and mTORC2 selectively regulate CD8$^+$ T cell differentiation. *J. Clin. Investig.* **125**, 2090–2108 (2015).
40. Huse, M., Lillemeier, B. F., Kuhns, M. S., Chen, D. S. & Davis, M. M. T cells use two directionally distinct pathways for cytokine secretion. *Nat. Immunol.* **7**, 247–255 (2006).
41. Frauwirth, K. A. et al. The CD28 signaling pathway regulates glucose metabolism. *Immunity* **16**, 769–777 (2002).
42. Barrett, T. et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* **41**, D991–D995 (2013).
43. Remark, R. et al. Characteristics and clinical impacts of the immune environments in colorectal and renal cell carcinoma lung metastases: influence of tumor origin. *Clin. Cancer Res.* **19**, 4079–4091 (2013).
44. Garcia-Diaz, A. et al. Interferon receptor signaling pathways regulating PD-L1 and PD-L2 expression. *Cell Rep.* **19**, 1189–1201 (2017).
45. Riaz, N. et al. Tumor and microenvironment evolution during immunotherapy with Nivolumab. *Cell* **171**, 934–949 e915 (2017).
46. Charoentong, P. et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint nlockade. *Cell Rep.* **18**, 248–262 (2017).
47. Wolchok, J. D. et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. *Clin. Cancer Res.* **15**, 7412–7420 (2009).
48. Nathanson, T. et al. Somatic mutations and neoepitope homology in melanomas treated with CTLA-4 blockade. *Cancer Immunol. Res.* **5**, 84–91 (2017).
49. Kaiserman, D. & Bird, P. I. Control of granzymes by serpins. *Cell Death Differ.* **17**, 586–595 (2010).
50. Hirst, C. E. et al. The intracellular granzyme B inhibitor, proteinase inhibitor 9, is up-regulated during accessory cell maturation and effector cell degranulation, and its overexpression enhances CTL potency. *J. Immunol.* **170**, 805–815 (2003).
51. Bladergroen, B. A. et al. The granzyme B inhibitor, protease inhibitor 9, is mainly expressed by dendritic cells and at immune-privileged sites. *J. Immunol.* **166**, 3218–3225 (2001).
52. Hirst, C. E. et al. Perforin-independent expression of granzyme B and proteinase inhibitor 9 in human testis and placenta suggests a role for granzyme B-mediated proteolysis in reproduction. *Mol. Hum. Reprod.* **7**, 1133–1142 (2001).
53. Medema, J. P. et al. Blockade of the granzyme B/perforin pathway through overexpression of the serine protease inhibitor PI-9/SPI-6 constitutes a mechanism for immune escape by tumors. *Proc. Natl Acad. Sci. USA* **98**, 11515–11520 (2001).
54. Uhlen, M. et al. A pathology atlas of the human cancer transcriptome. *Science* **357**, eaan2507 (2017).
55. Schoenborn, J. R. & Wilson, C. B. Regulation of interferon-gamma during innate and adaptive immune responses. *Adv. Immunol.* **96**, 41–101 (2007).
56. Taniguchi, T., Ogasawara, K., Takaoka, A. & Tanaka, N. IRF family of transcription factors as regulators of host defense. *Annu. Rev. Immunol.* **19**, 623–655 (2001).
57. Papatheodorou, I. et al. Expression Atlas: gene and protein expression across multiple studies and organisms. *Nucleic Acids Res.* **46**, D246–D251 (2018).
58. Fernandez-Banet, J. et al. OASIS: web-based platform for exploring cancer multi-omics data. *Nat Methods* **13**, 9–10 (2016).

Acknowledgements

The research was supported by the Cancer Immunologic Data Commons (1U24CA224316-01) grant of the National Cancer Institute (NCI), the Pathway to Independence Award (1K99CA218900-01) grant of NCI (to P.J.), the Specialized Center (1P50CA206963-01) grant of NCI (to G.J.F.), and the Breast Cancer Research Foundation (to X.S.L.). D.P. is a Cancer Research Institute/Robertson Foundation Fellow.

Author contributions

P.J., K.W.W. and X.S.L. designed the study and wrote the manuscript. P.J. carried out the computational works. S.G., D.P., Z.L. and N.T. carried out the experimental validation. P.J. and J.F. developed the website. A.S., X.H., X.B., B.L, J.L., G.J.F. and M.A.B. participated in discussions.

Competing interests

X.S.L. is a cofounder and board member of GV20 Oncotherapy, a scientific advisor of 3DMedCare and a paid consultant for Genentech. K.W.W. is a member of the scientific advisory board for TCR2 and Nextech; he serves as a consultant for Novartis. The laboratory of K.W.W. received sponsored research funding from Astellas Pharma Inc.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41591-018-0136-1.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to K.W.W. or X.S.L.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Data collection of clinical genomics studies. We collected cancer data sets with both patient survival durations and tumor gene expression profiles from the TCGA[31], PRECOG[17] and METABRIC[32] databases. If the clinical information is available, we separated the breast cancer data sets into the PAM50 (Prediction Analysis of Microarray 50, Prosigna) subtypes[13] of luminal A, luminal B, Her2 positive, basal and triple negative (a variation of basal subtype). This separation is because each PAM50 subtype has a distinct genomics profile[59] and degree of cytotoxic T cell infiltration[60]. Among all TCGA cancers, melanoma has two major tumor types annotated (that is, primary and metastatic); thus, we split melanoma profiles into primary and metastatic subtypes. The PRECOG database provided only survival duration information without other clinical factors; thus, we cannot perform subtype analysis. METABRIC is a breast cancer cohort, and we split all tumors according to the PAM50 subtypes (luminal A, luminal B, HER2, basal and triple negative).

To ensure the robustness of our analysis, we excluded the data sets from microarray platforms with fewer than 15,000 genes or without probes for cytotoxicity T cell markers (*CD8A*, *CD8B*, *GZMA*, *GZMB* and *PRF1*). Also, we included only data sets with more than 50 patients and 10% death rate because a low event number may undermine the reliability of Cox-PH survival regression[26]. Finally, 73 data sets from 3 databases passed our selection criteria (Supplementary Table 2a). The expression values of all genes are normalized by subtracting the mean values across all samples in a data set.

Statistical analysis. The interaction test in multivariate Cox-PH regression was applied to identify gene association with T cell dysfunction phenotype. In statistics, two variables interact if the effect of one variable depends on the status of the other, and a multiplication term in a multivariate linear model can test the interaction effect between two variables[25]. We applied the Cox-PH survival regression to test how the level of CTL interacts with other genes in the tumor to affect survival outcome. We solve a linear model Hazard $= a \times \mathrm{CTL} + b \times V + d \times \mathrm{CTL} \times V + c$ using the Cox-PH regression[26]. The CTL level is estimated through the bulk-tumor expression average of *CD8A*, *CD8B*, *GZMA*, *GZMB* and *PRF1*. In the Cox-PH model, the death hazard was estimated through the patient survival information. The variable V represents the expression level of a candidate gene in the test. Since we have selected data sets where CTL correlates with favorable survival outcome, the coefficient a is always negative. The association slope between CTL and Hazard is $a + d \times V$ (Fig. 1b). If the coefficient d is positive, a higher V level will flatten the slope between CTL and Hazard, indicating a reduced association between the cytotoxic T cell level and better survival outcome. If d is negative, a higher V level will sharpen the slope between CTL and Hazard, indicating an increased association between the cytotoxic T cell level and better survival outcome. The T cell dysfunction score for each gene is defined as the Wald test z score, which is the coefficient d divided by its standard error[26] (Fig. 1c and Supplementary Table 1). Of note, the thresholds shown in Fig. 1a and Supplementary Fig. 1a are used only to illustrate the principle of statistical interaction used by the model. When computing the T cell dysfunction scores through regression, we used the continuous variables without any thresholds. Also, we included clinical covariates, such as age, gender and stage (if available), in the regression to control for potential confounding factors.

To identify significant genes in the interaction test, we applied the Benjamini–Hochberg method to convert the two-sided Wald test P values to FDRs[61], and selected clinical data sets with more than 1% genes having an FDR smaller than 0.1. This procedure is equal to selecting data sets where the distribution of P values has a significant peak near zero[62]. For example, the P-value histogram computed using TCGA melanoma data has a spike near zero, indicating that a set of genes significantly interact with CTL to affect survival outcome (Supplementary Fig. 1b). In contrast, the result computed from glioblastoma data does not contain any genes with significant interactions (Supplementary Fig. 1b).

Performance comparison on predicting ICB response. We collected the RNA-Seq data in melanoma for anti-CTLA4[5] and anti-PD1[14] therapies with gene expression profiles for 25 and 42 pre-treatment tumors with complete clinical information, respectively. We collected the NanoString data for anti-PD1 therapies with gene expression profiles of 33 baseline tumors in four cancer types[15]. For each data set, we standardized the transcriptome data across patients by quantile-normalization, and further normalized the expression values of each gene by subtracting the average among all samples. Therefore, a zero value indicates the average expression.

To predict each tumor's potential to escape T cell-mediated killing, we first classified each tumor into CTL-high or CTL-low categories through the CTL marker expression levels (*CD8A*, *CD8B*, *GZMA*, *GZMB* and *PRF1*). Tumors with all positive values (higher than average) are classified as CTL-high, while the rest as CTL-low (Supplementary Fig. 4a). For the CTL-high tumors, we computed the Pearson correlation between tumor gene expression profiles and the T cell dysfunction signature (Supplementary Fig. 4b). For the CTL-low tumors, we computed the Pearson correlation between tumor gene expression profiles and the T cell exclusion signature (Supplementary Fig. 4b). The correlation with T cell dysfunction or exclusion signatures may have different distributions (Supplementary Fig. 4c). Therefore, to make the scale of Pearson correlations comparable between CTL-high and -low tumors, we normalized the correlation values within each sub-category through the standard deviation of correlations computed using the TCGA melanoma data. The scaled correlations were defined as TIDE prediction scores, representing the potential of tumor immune escape (Fig. 4a–c).

We also computed the response prediction from other biomarkers published in the literature. The predicted values of gene expression biomarkers (for example, *IFNG*, *CD8*, *PDL1* and *CRMA*) were the average values among all members defined by the original publications (Supplementary Table 7). The predicted values of ImmunoPhenoScore were computed using the source codes provided by the authors[46]. The predicted value of the tumor SCNA biomarker was downloaded from the original publication for the anti-CTLA4 data set[9] and provided by W. Hugo for the anti-PD1 data set[14].

The outcome predicted by all biomarkers is a range of values, instead of a binary outcome. For example, total mutation load, *CD8* expression level and TIDE prediction score all give one value for each patient tumor instead of a response classification. Therefore, we utilized the ROC curve, which plots the true-positive rates versus false-positive rates at various thresholds of biomarker values (Fig. 4d–f). The area under the ROC curve was used as the quality metric of prediction (Fig. 4g–i).

Gene selection for a focused TIDE signature. We select the most informative genes with both high variance across tumors and significant values in the TIDE signature. We selected 770 genes because that number is compatible with a NanoString platform that could be designed for a clinical assay. In the first step, we computed the standard deviation of expression values across samples for all genes in each TCGA cancer data set and selected 4,150 genes whose standard deviation is higher than the average of all genes in more 80% TCGA data sets. Next, we ranked the 4,150 genes by their maximum absolute values in the TIDE signatures of T cell dysfunction and exclusion. From this ranked list, we selected the top 770 genes, which is the maximum number that can fit on a NanoString assay. The column 'TIDE.selected' in Fig. 4g–i shows the TIDE performance on selected genes.

T cell killing assay based on co-culture between B16 and T cells. B16F10 cells were maintained in complete DMEM media (10% FBS and 50 U ml^{-1} of penicillin/streptomycin). B16F10-Cas9[63] cells were maintained in complete DMEM media with 2.5–5 μg ml^{-1} of blasticidin. CD8 T cells isolated from mice were cultured in complete RPMI 1640 media (10% FBS, 20 mM HEPES, 1 mM sodium pyruvate, 0.05 mM 2-mercaptoethanol, 2 mM L-glutamine and 50 U ml^{-1} streptomycin and penicillin). All cell lines are tested for mycoplasma contamination. Pmel-1 TCR transgenic mice were purchased from Jackson Laboratory (stock no. 005023).

CD8 T cells were isolated from spleen and lymph nodes from Pmel-1 TCR transgenic mice using the EasySep mouse CD8$^+$ T cell isolation kit (STEMCELL no. 19753) according to the manufacturer's protocol. Freshly isolated CD8 T cells were stimulated with anti-CD3/CD28 beads (ThermoFisher no. 11452D) at a bead to cell ratio of 1:2 to induce differentiation into an effector state. On day 3, recombinant mouse IL-2 (Biolegend, no. 575406) was added to the culture at 20 ng ml^{-1}. T cells were used for co-culture with B16F10 cells after at least six days of in vitro activation. Our animal experiments have complied with all relevant ethical regulations. The study protocol in this study was approved by the Institutional Care and Use Committee at Dana Farber Cancer Institute.

To knockout *Serpinb9*, CRISPR gRNA sequences targeting *Serpinb9* or non-targeting control were cloned into a PLKO3G-GFP vector and confirmed by sequencing. To overexpress *Serpinb9*, its cDNA was amplified, cloned into a pEF1a-puro vector and confirmed by sequencing. Knockout or overexpression constructs were co-transfected with pCMV-dR8.91 and pCMV-VSV-G (Addgene no. 8454) into HEK293T cells to generate lentivirus. Transfection was performed using TransIT-293 (Mirus, MIR2700) following the manufacturer's protocol. Lentivirus was collected 48 h later and stored at −80 °C. To generate *Serpinb9*-knockout cells, B16F10-Cas9 cells were infected with a lentivirus driving expression of a single gRNA overnight to inactivate *Serpinb9* genes individually. Infected cells were sorted on the basis of GFP expression by BD FACS Aria II. To generate *Serpinb9*-overexpressing cells, B16F10-Cas9 cells were infected with either pEF1a-puro backbone or pEF1a-puro-*Serpinb9*. Infected cells were selected by culturing with 2 μg ml^{-1} puromycin. Control (non-targeting gRNA or pEF1a-puro backbone), *Serpinb9*-deficient or *Serpinb9*-overexpressing B16F10-Cas9 cells were lysed and subjected to western blot analysis with the following antibodies: anti-SERPINB9 (Santa Cruz Biotechnology no. sc-57531) and anti-VCL (Sigma Aldrich no. V9264).

In a competition assay with *Serpinb9*-knockout cells, *Serpinb9*-deficient or non-targeting guide control B16F10-Cas9 cells (GFP positive) were mixed with control B16F10-Cas9 cells (GFP negative) at a 1:1 ratio. In a competition assay with *Serpinb9*-overexpressing cells, pEF1a control or pEF1a-*Serpinb9* B16F10-Cas9 cells (GFP negative) were mixed with control GFP-infected B16F10-Cas9 cells (GFP positive) at a 1:1 ratio. Mixed cells were stimulated with 10 ng ml^{-1} or 100 ng ml^{-1} of interferon gamma for 24 h to enhance MHC class I expression. These tumor cells were then co-cultured with in vitro-activated Pmel-1 T cells at different effector-to-target ratios in a 6-well plate. Tumor cells were plated at equal density in all wells, and T cells were added at 0, 1/3, 1/2 or 100% of the number of tumor cells. There are two or three cell-culture replicates for each condition. After a three-

day co-culture with T cells, the fold-change of *Serpinb9*-edited B16F10 cells was determined by FACS, comparing the percentage of *Serpinb9*-edited B16F10 cells to control B16F10 cells (Supplementary Fig. 9). T cells present in these cultures were gated out on the basis of antibodies specific for CD45 (APC–Cy7) (Biolegend, 103115).

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Software availability. The TIDE web application of response prediction is freely accessible with any modern web browser through http://tide.dfci.harvard.edu/. We will keep the website and tool operating and freely accessible for the foreseeable future. The source code for computing the T cell dysfunction score through the interaction test is available under GNU Public License v3 through GitHub at https://github.com/foreverdream2/dysfunction_interaction_test.

Data availability. Users can query our analysis results with gene names: http://tide.dfci.harvard.edu/query/. All of our processed input data, analysis output data and an example script to repeat our major results are available at http://tide.dfci.harvard.edu/download/.

References

59. Hoadley, K. A. et al. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell* **158**, 929–944 (2014).
60. Miyan, M., Schmidt-Mende, J., Kiessling, R., Poschke, I. & de Boniface, J. Differential tumor infiltration by T-cells characterizes intrinsic molecular subtypes in breast cancer. *J. Transl. Med.* **14**, 227 (2016).
61. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
62. Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl Acad. Sci. USA* **100**, 9440–9445 (2003).
63. Pan, D. et al. A major chromatin regulator determines resistance of tumor cells to T cell-mediated killing. *Science* **359**, 770–775 (2018).

nature research

Corresponding author(s): Xiaole Shirley Liu, Kai W. Wucherpfennig

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

   Describe how sample size was determined.

   The sample size of our computational analysis is determined by the public data we used. Thus, we have no choice and included the sample size in our figure legends.

   The sample size in the T-cell killing assay is determined through our previous experience: we are interested in at least 1.5 fold difference, according to our previous publication (Pan et al., 2018), two or three cell-culture replicates will be enough to detect this difference at >80% power, due to the low standard deviation of this assay.

2. Data exclusions

   Describe any data exclusions.

   No data were excluded in our analysis.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   All attempts at replications are successful for all experiments in our manuscript.

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups.

   For our computational work, randomization is not relevant since our analysis is based on public data cohorts.

   For the T-cell killing assay, randomization is also not relevant since all experiments are based on the same cell line (B16F10), with different genetic perturbation.

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   Blinding is not relevant with our computational analysis work.

   The T-cell killing assay was not blinded, but the robust phenotype of our result is based on objective measurements from the Flow Cytometry instead of any human estimation.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

Python 2.7.13 included in Anaconda 4.3.1
R version 3.4.1
FlowJo X 10.0

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

No unique materials were used.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

1. anti-SERPINB9: Santa Cruz Biotechnology #sc-57531, clone PI9-17, lot number: H0316, dilution 1:1000

SERPINB9 antibody (clone PI9-17) was validated in multiple cell lines according to manufacturer's documentation, and used in multiple other publications for detection of SERPINB9 protein (e.g. PMID 19956856). We also validated this antibody through comparing the protein levels between WT and Serpinb9-/- cells.

2. anti-VCL: Sigma Aldrich #V9264, clone hVIN-1, lot number: 124M4787V, dilution 1:1000

VCL antibody (clone hVIN-1) has been validated in manufacturer's documentation and multiple other publications (e.g. PMID 21159656).

3. CD45: APC-Cy7 Biolegend 103115, clone 30-F11, lot number: B185138, dilution 1:200

CD45 antibody (clone 30-F11) has been validated in manufacturer's documentation and multiple other publications (e.g. PMID: 29301958).

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

Both B16F10 and HEK293T cell lines were purchased from ATCC.

b. Describe the method of cell line authentication used.

Both B16F10 and HEK293T cell lines has been confirmed by short tandem repeat (STR) analysis.

c. Report whether the cell lines were tested for mycoplasma contamination.

Both B16F10 and HEK293T cell lines were tested negative for mycoplasma contamination.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

No commonly misidentified cell lines were used.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

We isolated T cells from Pmel1 TCR transgenic mice with C57BL6 background, 8-12 week old, male.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

This study does not involve human research participants.

nature research

Corresponding author(s): Xiaole Shirley Liu, Kai W. Wucherpfennig

☐ Initial submission ☐ Revised version ☒ Final submission

Flow Cytometry Reporting Summary

Form fields will expand as needed. Please do not leave fields blank.

Data presentation

For all flow cytometry data, confirm that:

☒ 1. The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ 2. The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ 3. All plots are contour plots with outliers or pseudocolor plots.

☒ 4. A numerical value for number of cells or percentage (with statistics) is provided.

Methodological details

5. Describe the sample preparation.	B16F10 cells (parental or with transgenic GFP expression) were co-cultured with CD8+ T cells derived from Pmel-1 TCR transgenic mice. After co-culture, cells were incubated with anti-CD45 and then analyzed by FACS.
6. Identify the instrument used for data collection.	Sony SP6800 Spectral Analyzer
7. Describe the software used to collect and analyze the flow cytometry data.	Data were collected using Sony SP6800 software and analyzed using Flowjo.
8. Describe the abundance of the relevant cell populations within post-sort fractions.	N/A
9. Describe the gating strategy used.	We first gated on DAPI- CD45- cells to exclude dead cells and hematopoietic cells. We then analyzed the percentage of GFP+ population in each sample, which is distinct from un-transduced parental control cells.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information. ☒

nature medicine

SUPPLEMENTARY INFORMATION

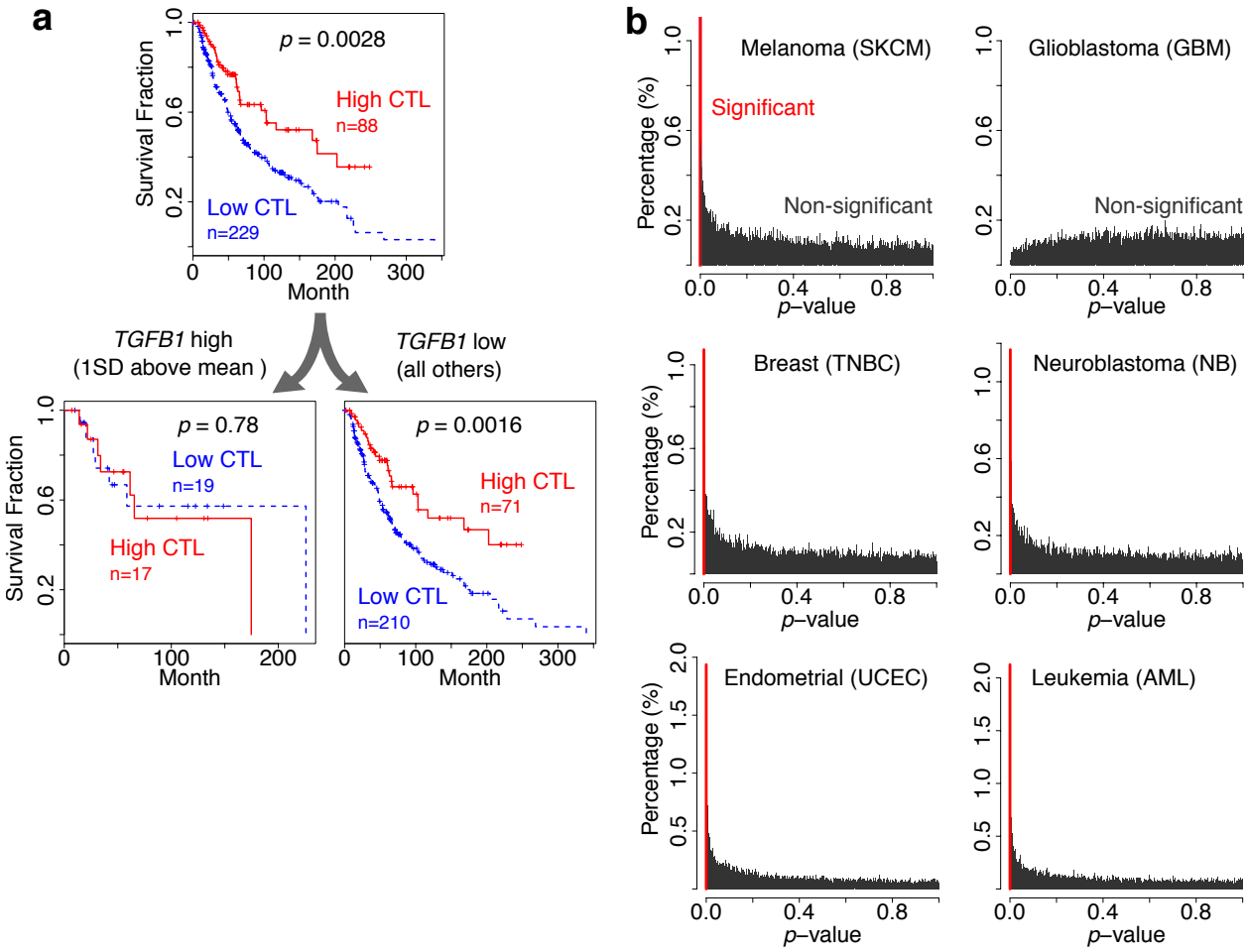
https://doi.org/10.1038/s41591-018-0136-1

In the format provided by the authors and unedited.

Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response

Peng Jiang[1,2,10], Shengqing Gu[3,10], Deng Pan[4,5,10], Jingxin Fu[6], Avinash Sahu[1,2], Xihao Hu[1,2], Ziyi Li[6], Nicole Traugh[3], Xia Bu[3], Bo Li[1,2,9], Jun Liu[7], Gordon J. Freeman[3], Myles A. Brown[3,8], Kai W. Wucherpfennig[4,5,11]* and X. Shirley Liu[1,2,6,8,11]*

[1]Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA, USA. [2]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [3]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA. [4]Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston, MA, USA. [5]Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA, USA. [6]School of Life Science and Technology, Tongji University, Shanghai, China. [7]Department of Statistics, Harvard University, Cambridge, MA, USA. [8]Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA, USA. [9]Present address: Department of Bioinformatics, UT Southwestern, Dallas, TX, USA. [10]These authors contributed equally: Peng Jiang, Shengqing Gu, Deng Pan.[11]These authors jointly supervised: Kai W. Wucherpfennig, X. Shirley Liu. *e-mail: kai_wucherpfennig@dfci.harvard.edu; xsliu@jimmy.harvard.edu

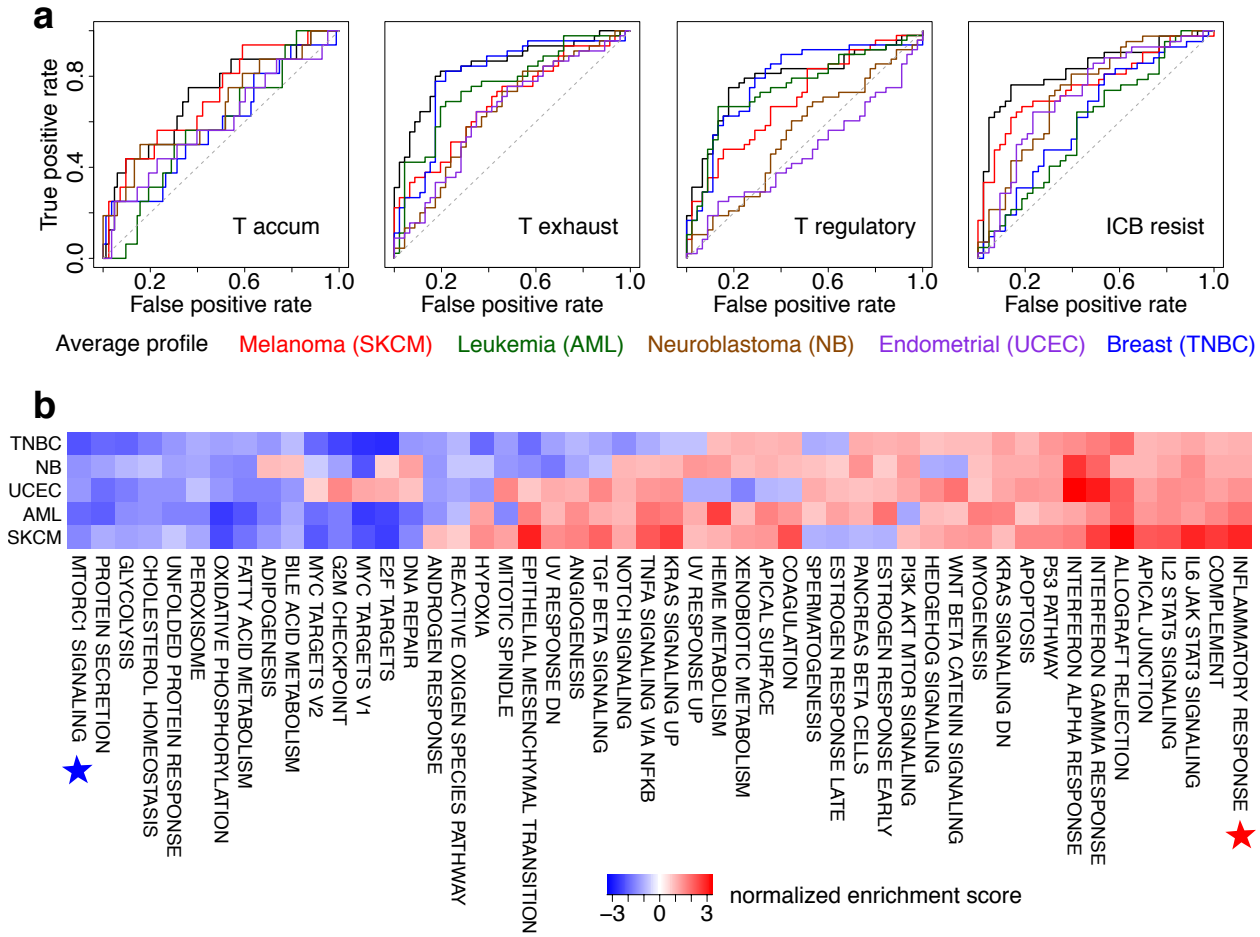


Supplementary Figure 1. The interaction test identifies gene signatures of T-cell dysfunction.

(a) The association between the cytotoxic T lymphocyte level (CTL) and overall survival for patients with metastatic melanoma tumors in TCGA. For each tumor, the CTL infiltration level was estimated as the average expression of *CD8A*, *CD8B*, *GZMA*, *GZMB*, and *PRF1*. The association between the CTL level and overall survival was computed through the two-sided Wald test in the Cox-PH regression. In the top part, samples were classified into two groups for each Kaplan Meier plot: "High CTL" (red) have above-average CTL values among all samples, while "Low CTL" samples (blue) are below average. In the lower part, samples were split according to the *TGFB1* expression to show the association between the CTL level and survival outcome. The left panel shows melanomas with High *TGFB1* expression (1 standard deviation above the average); Low *TGFB1* expression (the remaining samples) are plotted in the right panel.

(b) The distribution of test *p*-values in different cancer cohorts. The two-sided Wald test in Cox-PH regression was applied to compute the statistical significance of interactions between each gene and CTL level (Figure 1b and Supplementary Table 1). The Benjamini-Hochberg method was applied to convert the test *p*-values to false discovery rates (FDR)[61]. There are

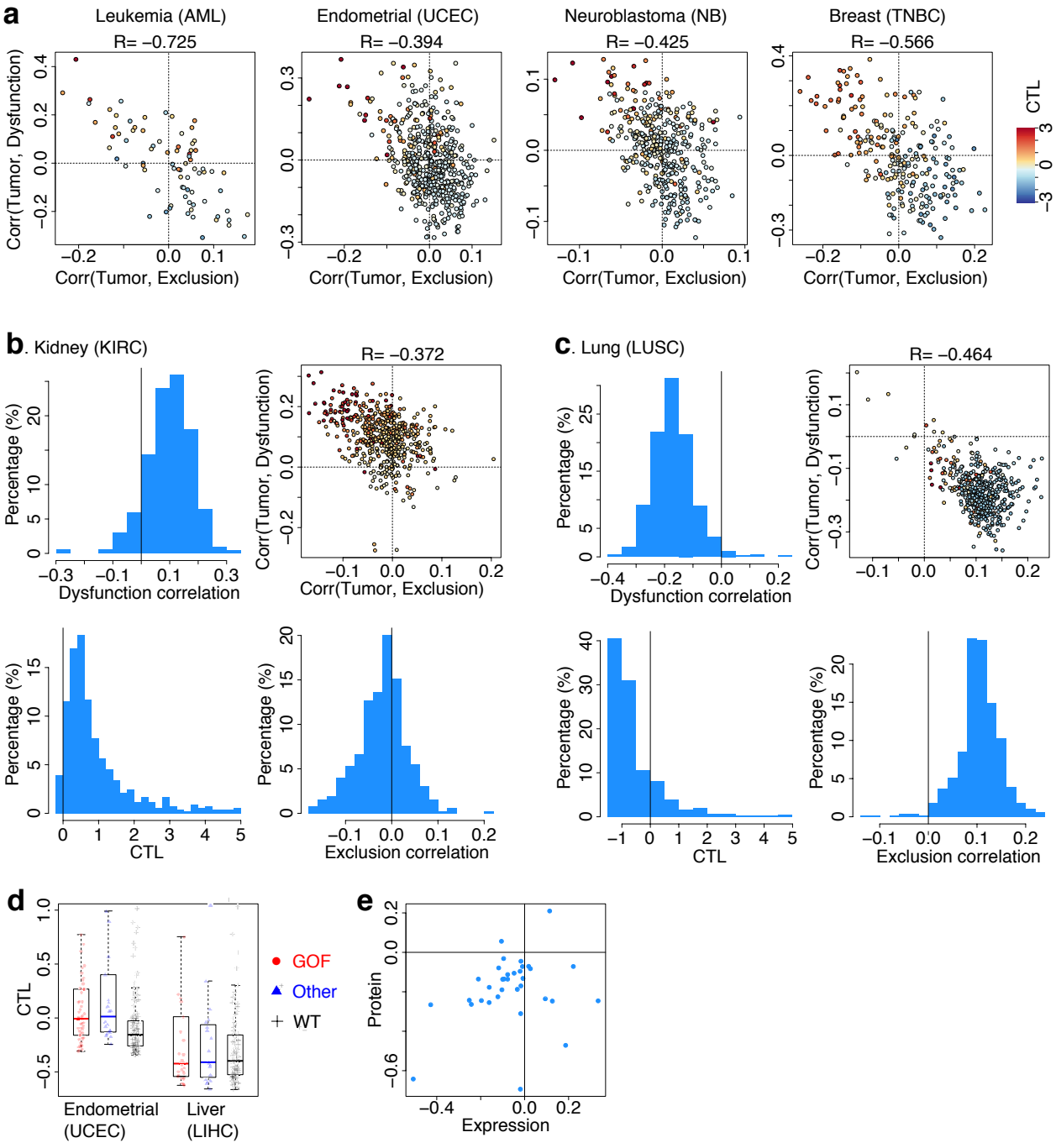
19511, 19241, 18091, 17649, 19716, and 20510 *p*-values for cancer types of SKCM, GBM, TNBC, NB, UCEC, and AML, respectively. The *p*-values deemed significant are shown in red (FDR < 0.1). The melanoma and other four cohorts met our criteria for more than 1% genes passing an FDR of 0.1. Glioblastoma did not meet the threshold.



Supplementary Figure 2. Analyses of T-cell dysfunction signatures.

(a) The receiver operating characteristic (ROC) curves for the performance of T-cell dysfunction scores in predicting top gene hits in known signatures of T-cell dysfunction and ICB resistance (Supplementary Table 4).

(b) Functional enrichment analysis for T-cell dysfunction signatures. We applied gene set enrichment analysis (GSEA)[64] using the hallmark gene category (a set of biological processes annotated by MSigDB)[65] to identify functions enriched within each T-cell dysfunction signature. The input to GSEA pre-ranked module is a ranked list of genes determined by T-cell dysfunction scores across all human genes. The GSEA analysis determines whether the hallmark gene sets annotated by MSigDB show statistically significant enrichment at either end of the ranking. The normalized enrichment scores are shown in heat maps. Stars highlighted the hallmark gene sets discussed in the manuscript.



Supplementary Figure 3. Analyses of T-cell exclusion signatures in different cancer types.

(a) Anti-correlation between T-cell dysfunction scores and exclusion scores across tumors in different cancer types. The figure is an extension of Figure 3c (the plot for melanoma). For each cancer type, the T-cell dysfunction and exclusion scores are computed as the Pearson correlations between the tumor expression profiles and TIDE signatures of T-cell dysfunction and exclusion, respectively. The association between dysfunction and exclusion scores are shown with 2D plot with CTL level of each tumor as the color of dots. The

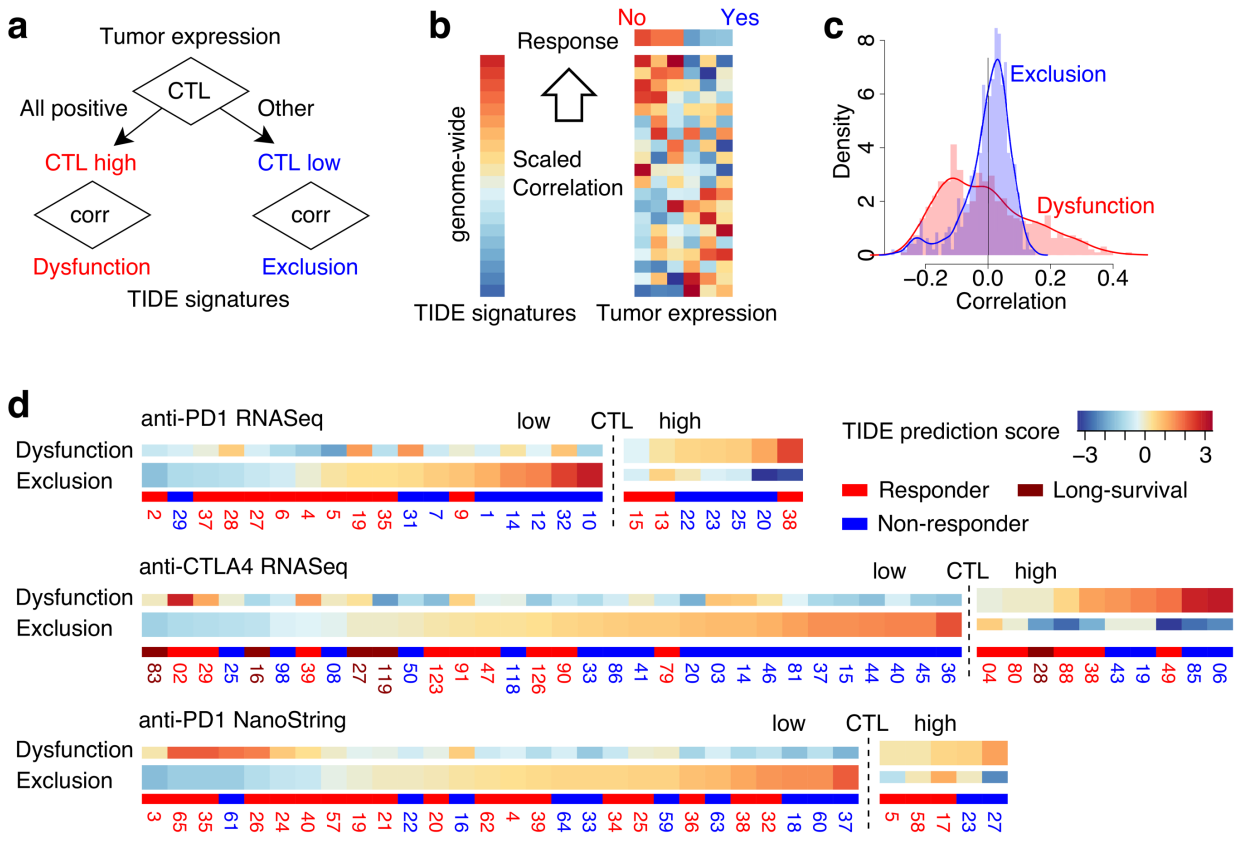
Pearson correlation between X and Y axis are shown after each cancer name. The number of samples in each cancer type is shown in Supplementary Table 2b. The two-sided t-test *p*-values of Pearson correlations for AML, UCEC, NB, and TNBC are 4.34e-14, 1.56e-21, 1.64e-18, and 3.98e-21, respectively. (TNBC: triple negative breast cancer, AML: acute myeloid leukemia, UCEC: uterine corpus endometrial carcinoma, NB: neuroblastoma).

(b) Signature analysis for kidney renal cell carcinoma (KIRC). Among all TCGA cancer types with normal samples profiled, KIRC has the highest enrichment of T-cell dysfunction signature (Figure 3d). The signature enrichment is shown as 2D plot on the level of individual tumors with the CTL level as the dot color across 531 tumors. The histograms of signature enrichment and CTL levels are also shown together. The two-sided t-test *p*-values of Pearson correlation is 7.96e-19.

(c) Signature analysis for lung squamous cell carcinoma (LUSC). Among all TCGA cancer types with normal samples profiled, LUSC has the highest enrichment of T-cell exclusion signature (Figure 3d). The signature enrichment is shown as panel b across 484 tumors. The two-sided t-test *p*-values of Pearson correlation is 3.47e-27.

(d) CTL level for tumors with different *CTNNB1* mutation status. One previous study reported beta-catenin signaling to prevent T-cell infiltration in melanoma tumors[66]. Therefore, we analyzed the association between the CTL level and *CTNNB1* (beta-catenin) status. Among TCGA cancer types, only endometrial and liver cancer have more than ten tumors harboring *CTNNB1* mutations. We classified these tumors according to the mutation status and showed the CTL levels by box-plots. For endometrial cancer, there are 52, 22, and 173 tumor samples in the GOF, Other, and WT groups, respectively. For liver cancer, there are 30, 19, and 146 tumor samples in the GOF, Other, and WT groups, respectively. Within each group, the scattered dots represent sample values, and the thick line represents the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). Whiskers encompass 1.5 times the inter-quartile range. GOF: gain of function mutation annotated by OncoKB[67], Other: other mutation types, WT: wild-type, UCEC: uterine corpus endometrial carcinoma, LIHC: liver hepatocellular carcinoma.

(e) 2D-scatter plot for the Pearson correlations between CTL level and *CTNNB1* molecular status in different cancer types. Each dot represents a TCGA cancer type. The X-axis represents the Pearson correlation between CTL and *CTNNB1* expression level across tumors in a cancer type. The Y-axis represents the Pearson correlation between CTL and CTNNB1 protein level across tumors in a cancer type. For each cancer cohort, the sample number in correlation computation is available in Supplementary Table 12. The correlation values for beta-catenin protein are negative in most cancer types.



Supplementary Figure 4. ICB response prediction in tumors with different CTL levels.

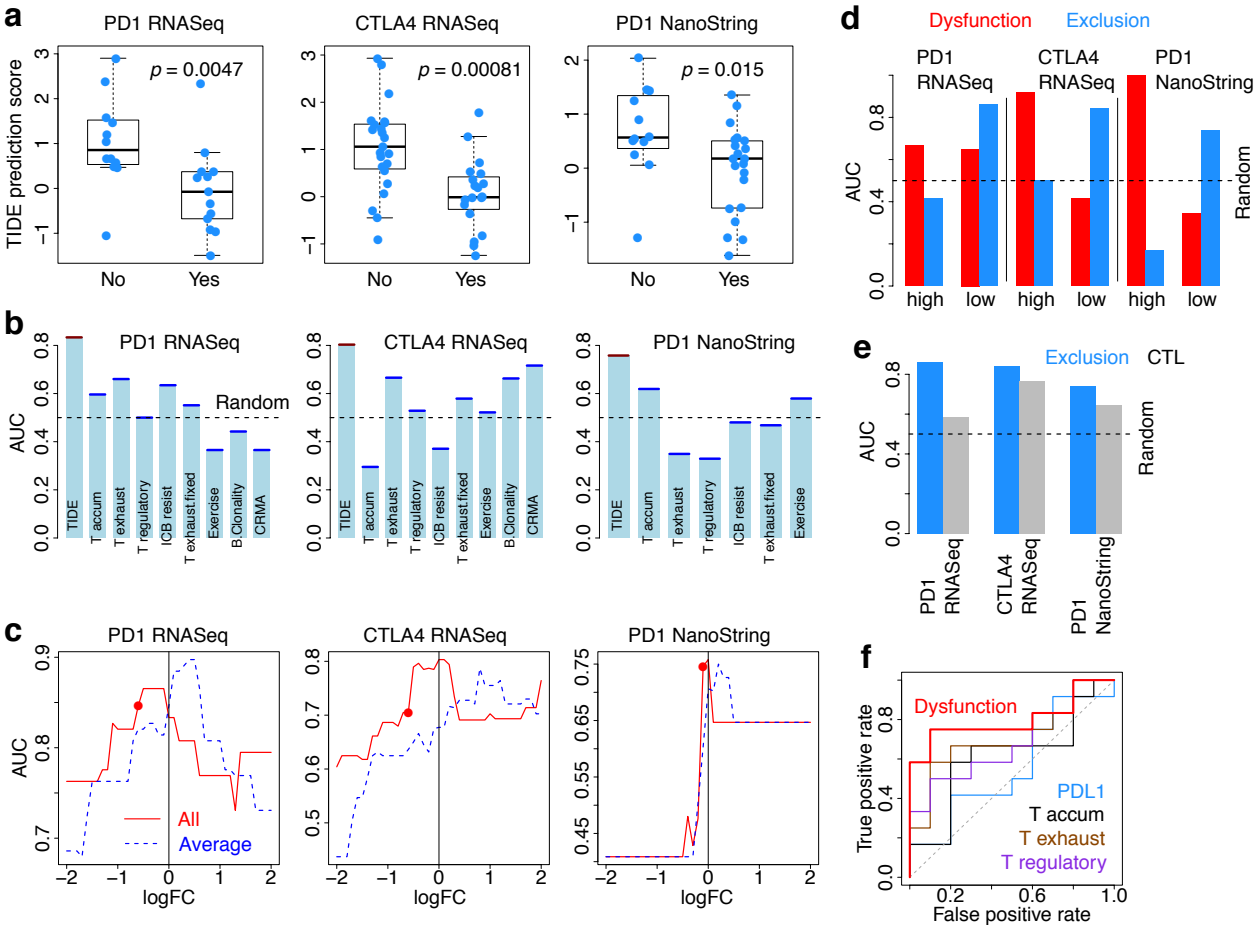
(a) Procedure to predict the immune evasion potential for tumors. All cancer samples are classified into cytotoxic T lymphocyte (CTL) high or low categories through the bulk tumor expression values of CTL markers (*CD8A*, *CD8B*, *GZMA*, *GZMB*, and *PRF1*). Tumors with all positive values (higher than average) are classified as CTL high, and other tumors are classified as CTL low. TIDE prediction scores of CTL-high tumors are computed as the Pearson correlation between tumor expression profiles and the T-cell dysfunction signature. TIDE prediction scores of CTL-low tumors are computed as the Pearson correlation between tumor expression profiles and the T-cell exclusion signature.

(b) Computation of genome-wide signature correlation. Each part of TIDE signature (e.g., dysfunction or exclusion) is a vector of scores across all human genes. The TIDE prediction score for each tumor is a genome-wide correlation between TIDE signatures and tumor gene expression profiles. The correlation was also scaled by the standard deviation of correlations computed using TCGA data of the same cancer type.

(c) Distribution of correlation between tumor expression profiles and TIDE signatures. For each TCGA metastatic melanoma tumor, we computed the Pearson correlation between tumor expression profile and T-cell dysfunction or exclusion signatures across 317 samples. The histogram and density estimation for both scores are plotted across tumors. Since the dysfunction and exclusion correlations have different distributions, we scaled these

correlations by their standard deviations in TCGA data to enable comparability when we merge the scores from CTL high and low tumors together.

(d) Heatmap of TIDE prediction scores for tumors in ICB cohorts. Tumors were divided into CTL high or low categories based on the expression level of CTL marker genes (panel a). Red indicates a tumor that responded to therapy. Blue indicates a non-responder. We sorted tumors in descending order in each category according to their TIDE prediction scores, which are scaled correlations between pre-treatment tumor expression profiles and TIDE signatures (panel b, c). The results are shown for the anti-PD1 cohort of 25 melanoma tumor profiles by RNA-Seq[14], anti-CTLA4 cohort of 42 melanoma tumor profiles by RNA-Seq[3], and anti-PD1 study of 33 tumor profiles (four cancer types) by NanoString platform[15].



Supplementary Figure 5. Performance of different signatures on predicting ICB response.

(a) TIDE prediction scores for responders (Yes) and non-responders (No) in ICB clinical cohorts[3,14,15]. The difference between two groups was tested by the two-sided Wilcoxon rank-sum test. For the CTLA4 RNASeq study[3], there are 19 and 23 patients in the "Yes" and "No" groups, respectively. For the PD1 RNASeq study[14], there are 13 and 12 patients in the "Yes" and "No" groups, respectively. For the PD1 NanoString study[15], there are 21 and 12 patients in the "Yes" and "No" groups, respectively. Within each group, the scattered dots represent sample values, and the thick line represents the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). Whiskers encompass 1.5 times the inter-quartile range.

(b) The area under ROC curve (AUC) for several gene signatures of T-cell dysfunction, immune response (Supplementary Table 4), and ICB response (Supplementary Table 7). The performance of random predictor (AUC = 0.5) was represented by the dotted line.

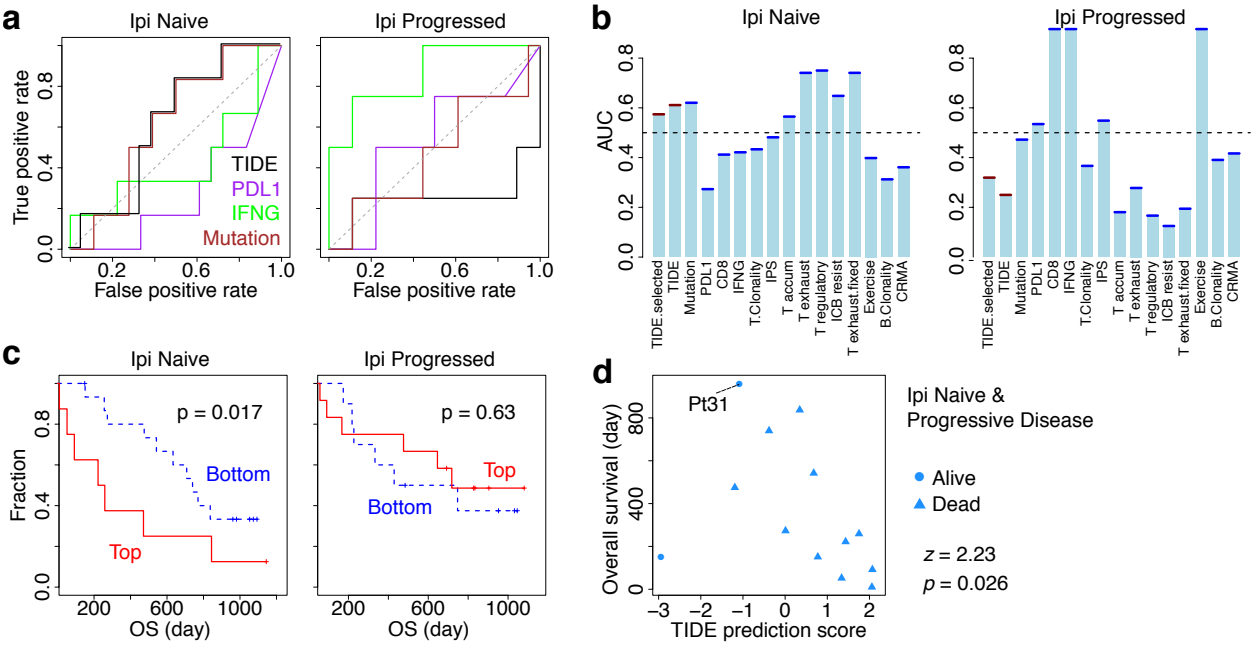
(c) The TIDE prediction performance with different CTL thresholds. We compared the AUC of predicting ICB response between two CTL cutoff types. In method one (All), we required all log-fold change (logFC) ratios between CTL genes and its average to be higher than a threshold value. In method two (Average), we required the average logFC of CTL genes to be higher than a threshold. Our current CTL cutoff corresponds to the method one (All) with

zero as the threshold. For method one, we also marked the AUC with median CTL value as the threshold.

(d) The performance of each TIDE signature (T-cell dysfunction or exclusion) in predicting ICB response of CTL-high and CTL-low tumors. For all datasets, the T-cell dysfunction signature always predicts the ICB outcome better than the T-cell exclusion signature among CTL-high tumors. In contrast, the T-cell exclusion signature always predicts the ICB outcome better than the T-cell dysfunction signature among CTL-low tumors.

(e) Performance comparison between the CTL level and T-cell exclusion signature within tumors with low CTL level. The performance of random predictor (AUC = 0.5) was represented by the dotted line.

(f) Performance comparison among signatures of T-cell dysfunction within tumors with high CTL level. All CTL-high tumors from three ICB cohorts are analyzed together due to the low sample size within each cohort (Figure 4a-c).



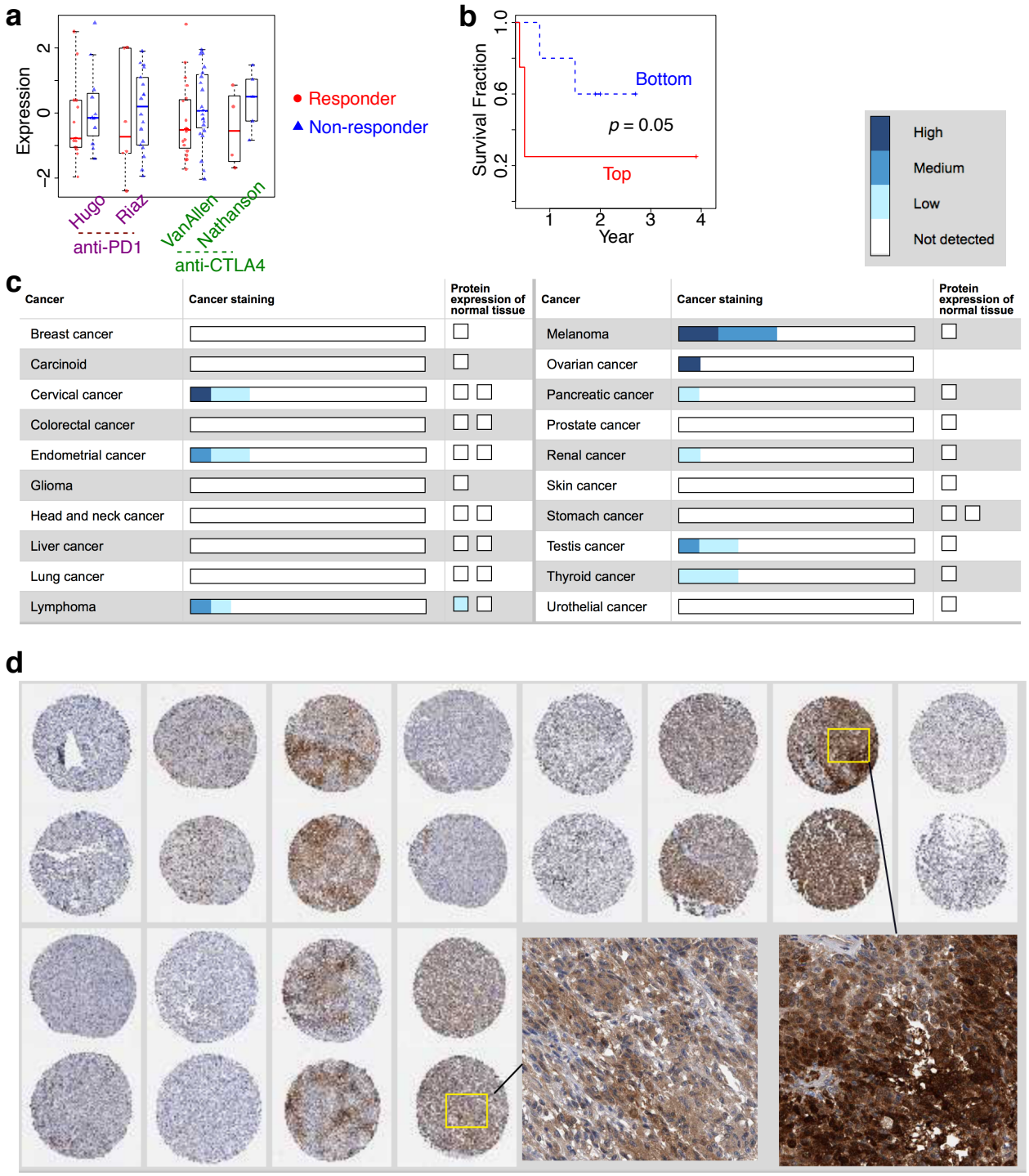
Supplementary Figure 6. Performance of different signatures on predicting anti-PD1 response in the study of Riaz and colleagues.

(a) ROC curves for the prediction performance of several biomarkers on an anti-PD1 RNA-Seq study[45]. There are 24 and 26 pre-treatment tumors with anti-PD1 as the first-line (Ipilimumab Naive) and second-line (Ipilimumab Progressed) immunotherapy, respectively.

(b) The area under ROC curve (AUC) for several gene signatures of T-cell dysfunction, immune response (Supplementary Table 4), and ICB response (Supplementary Table 7). The performance of random predictor (AUC = 0.5) was represented by the dotted line.

(c) The association between TIDE prediction scores and overall survival on the study in panel a. Patient tumors are classified into the top (>1) and bottom (<1) groups using the TIDE prediction scores. The association between TIDE prediction scores and overall survival was tested through the two-sided Wald test in a Cox-PH regression.

(d) Patients with progressive disease (PD) defined by RECIST v1.1 can have overall survival benefits from ICB. Among the 13 patients labeled as PD in the study of Riaz and colleagues, we plotted their overall survival durations on Y-axis and TIDE prediction scores on X-axis. The overall survival benefit of patients was significantly associated with TIDE scores (Cox-PH regression z-score = 2.23, two-sided Wald test *p*-value = 0.026). Therefore, RECIST v1.1 may underestimate the ICB clinical benefit. A potential explanation is that RECIST v1.1, developed for chemotherapeutic and targeted agents, measures the tumor shrinkage within a few weeks of initial therapy. Besides RECIST response patterns, ICB can lead to other patterns, such as an initial transient increase in tumor burden before a response due to anticancer immune response leading to tumor inflammation[47].



Supplementary Figure 7. Validation of *SERPINB9* as a regulator of tumor immune escape.

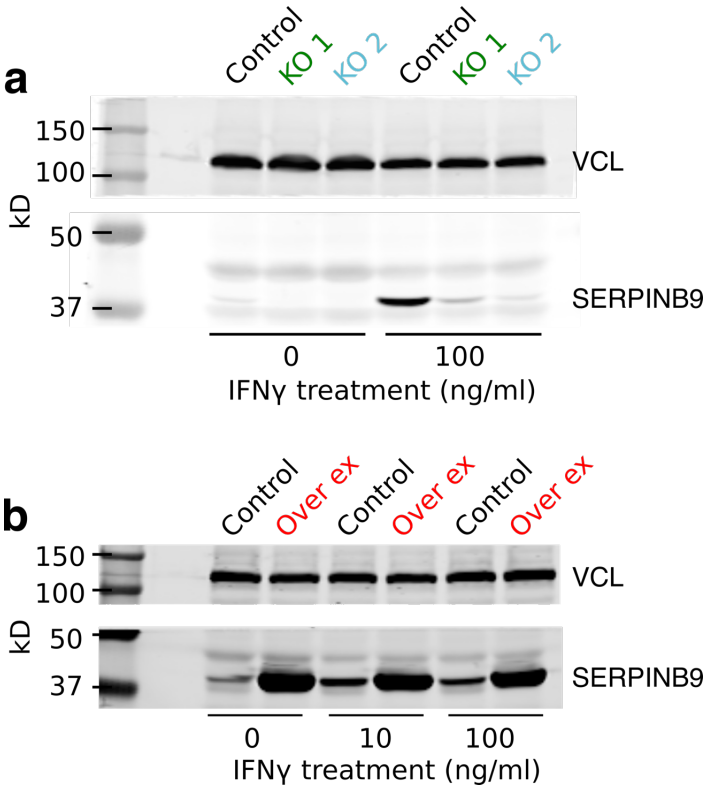
(a) The *SERPINB9* expression level in pre-treatment tumors of ICB responders and non-responders over four RNA-Seq datasets. There are 25, 24, 42, and 9 patients in the dataset of Hugo[14], Riaz[45], VanAllen[3], and Nathanson[48] cohorts, respectively. Within each group, the scattered dots represent sample values, and the thick line represents the median value. The

bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). Whiskers encompass 1.5 times the inter-quartile range.

(b) The overall survival for *SERPINB9* top and bottom half patients from an anti-CTLA4 study of 9 patients[48]. The association between the *SERPINB9* expression in pre-treatment tumors and patient overall survival under anti-CTLA4 treatment was evaluated through the two-sided Wald test in a Cox-PH regression.

(c) The SERPINB9 protein level in cancer. The Protein Atlas database provided the immunohistochemistry (IHC) protein staining for 15287 genes in 20 cancer types[54]. For each cancer type, the fraction of samples with protein expression levels high, medium, low, or not detected are provided by the blue color scales. The length of the bar represents the number of patient samples.

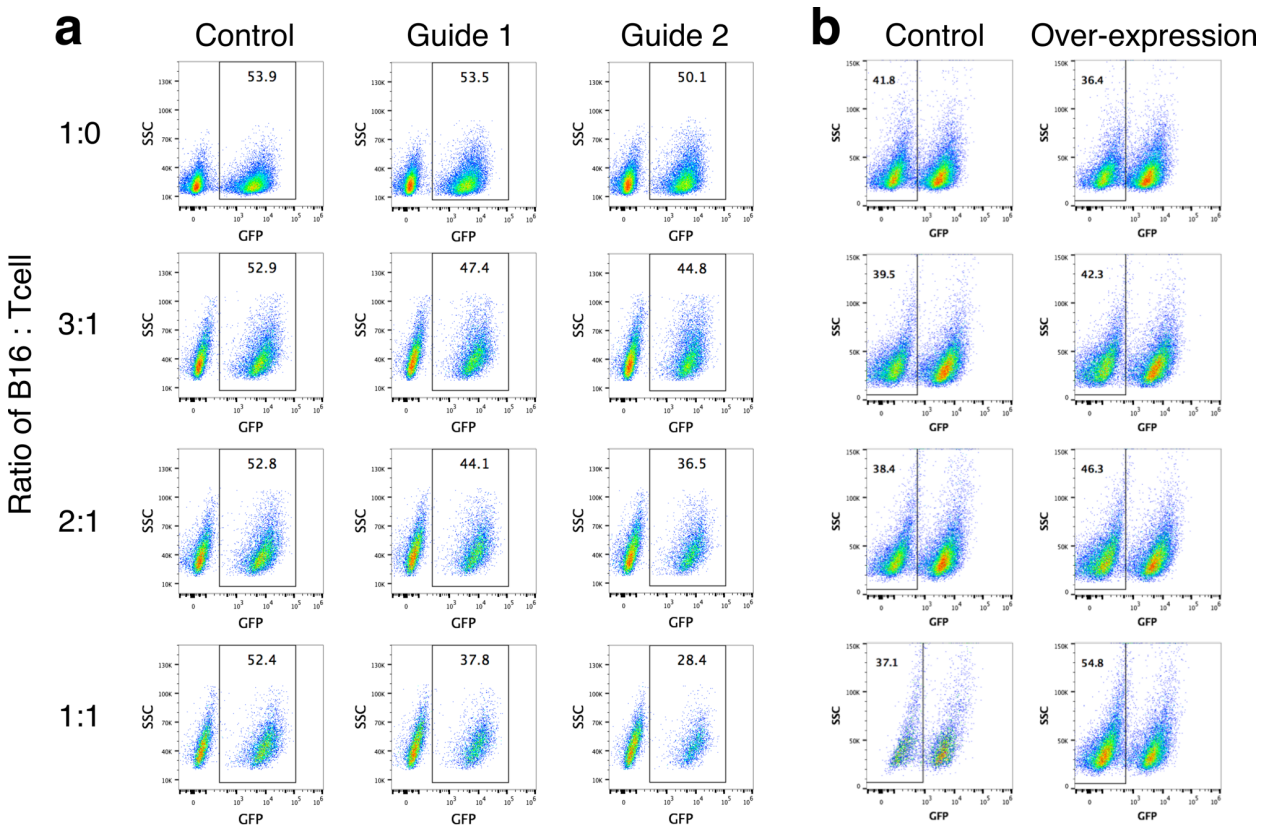
(d) The SERPINB9 protein staining data for melanoma tumors with two representative regions amplified for details. The cell nuclei are labeled by blue color, and the SERPINB9 protein is stained with brown color. According to the Protein Atlas annotation, the SERPINB9 protein is highly expressed at cancer cell cytoplasm, membrane, and nucleus.



Supplementary Figure 8. Original western-blot images for SERPINB9 protein levels. B16F10-Cas9 cells were lysed and subjected to Western blot analysis with the following antibodies: anti-SERPINB9 (Santa Cruz Biotechnology # sc-57531), and anti-VCL (Sigma Aldrich #V9264).

(a) SERPINB9 knockout (KO).

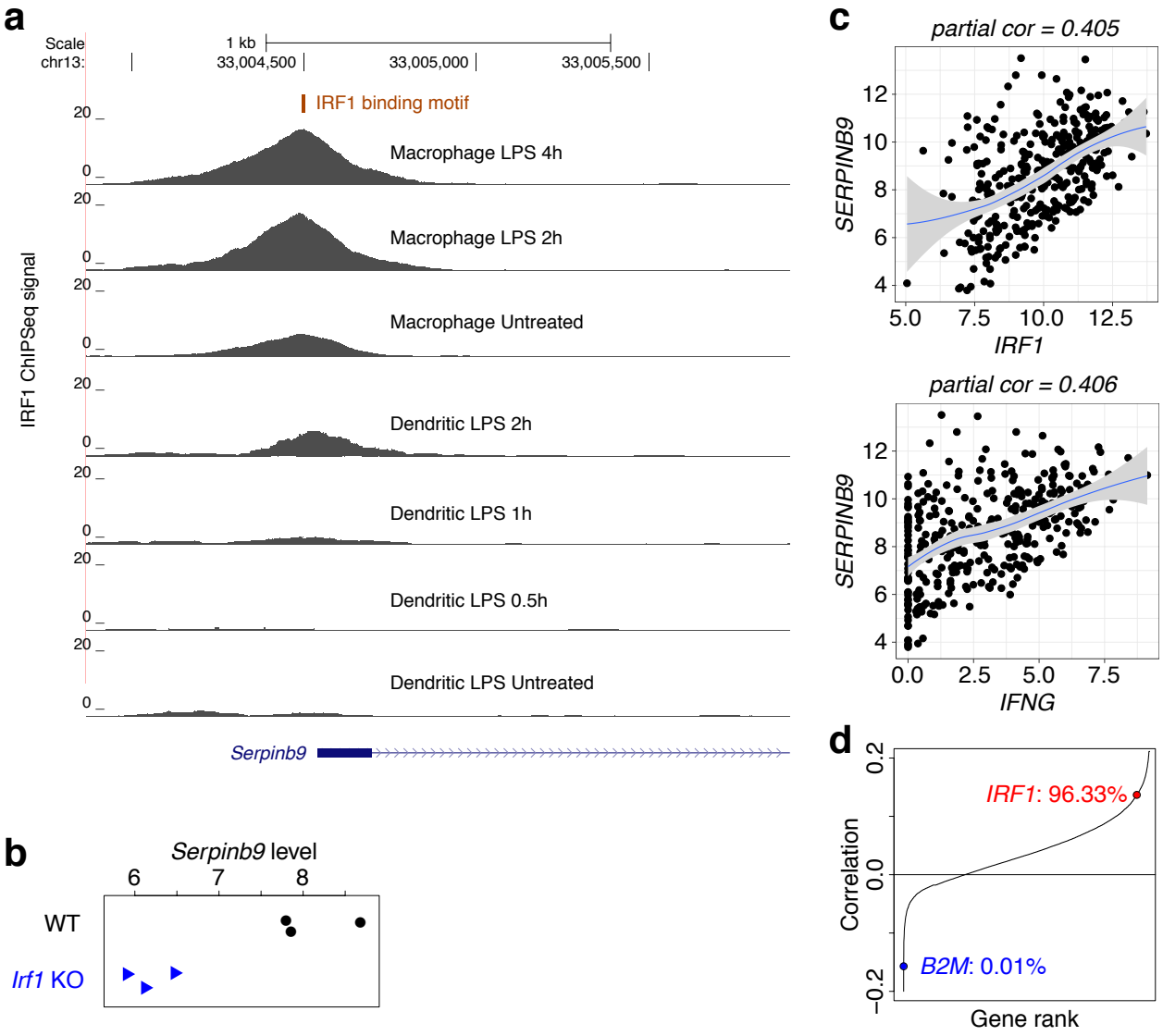
(b) SERPINB9 over-expression.



Supplementary Figure 9. SERPINB9 regulates the sensitivity of B16F10 cells to T-cell cytotoxicity. The density of B16F10 cells was the same for all conditions and T cells were added at varying numbers. Each group of mixed cells was cultured in the absence (B16: T cell = 1:0) or presence (B16: T cell = 3:1, 2:1, or 1:1) of Pmel-1 T cells targeting the gp100 antigen on B16 cells. After three days in culture, the percentage of GFP positive B16F10 cells was assessed by flow cytometry. T cells present in these cultures were gated out using CD45 antibodies.

(a) The relative fraction of *Serpinb9* knockout cells. B16F10-Cas9 cells were transduced with lentivirus co-expressing GFP and CRISPR guide RNAs (gRNA) targeting *Serpinb9* or non-targets. Each gRNA-transduced GFP positive cell line was mixed with the parental GFP negative cell line at a 1:1 ratio. For each group, the experiments were repeated in three independent cell cultures with similar results. One representative FACS plot with the median percentage value is shown.

(b) The relative fraction of *Serpinb9* over-expression cells. B16F10-Cas9 cells were transduced with lentivirus of pEF1a backbone or pEF1a-*Serpinb9* (GFP-negative, Puromycin is the selection marker) and mixed with GFP-positive cell line at a 1:1 ratio. For each group, the experiments were repeated in two independent cell cultures with similar results. One representative FACS plot is shown.



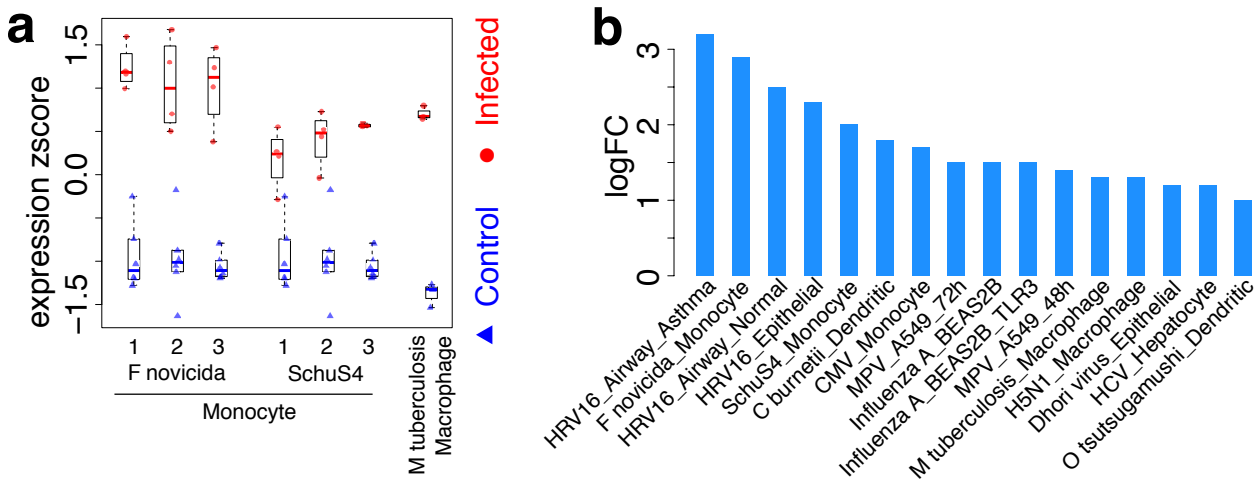
Supplementary Figure 10. Potential regulators of *Serpinb9* expression upon IFNγ treatment.

(a) IRF1 binding sites near *Serpinb9* promoter profiled by public studies in the GEO database. The Cistrome database[68] collected two IRF1 ChIP-Seq datasets that are profiled in mouse macrophage[69] and dendritic cell[70] types. The ChIP-Seq signal was shown around the transcription start site of *Serpinb9* together with the IRF1 binding motif site.

(b) The *Serpinb9* expression level after *Irf1* knockout in mice. A previous study generated RNA-Seq profiles in T regulatory cell for both wild-type (WT) and *Irf1* knockout (KO) cells[71]. The *Serpinb9* level was defined as the DESeq2 rlog value[72] computed from RNA-Seq data.

(c) Correlation between *SERPINB9* and *IRF1* expression levels across 317 metastatic melanoma tumors. This figure was generated by the TIMER server[18,73]. The log2(RSEM) values from TCGA RNA-Seq data are shown on X and Y axes for each gene with each dot as one tumor. The partial correlation between X and Y axes with tumor purity corrected is shown as the title.

(d) Pearson correlation of gene expression values between *SERPINB9* and other genes in melanoma across 1257 single-cell RNA-Seq profiles of cancer cells[74]. All genes are ordered by correlations with *SERPINB9* within the range of [-0.2, 0.2]. The correlation between *SERPINB9* and *IRF1* is one of the top positive correlations (ranked 96.33%). A negative control correlation is the one between *SERPINB9* and *B2M* (ranked 0.01%).



Supplementary Figure 11. *SERPINB9* expression change after pathogen infection. Both gene expression omnibus (GEO) database and Expression Atlas database provided gene expression profiles curated from public datasets. We searched these databases with keywords "*SERPINB9*" and "infection" and selected all human datasets that studied the gene expression change of host cells after pathogen infection.

(a) GEO results. The first study measured gene expression in human blood monocytes infected with Francisella tularensis and Schu S4 strain[75] using the HG-U133+2 microarray with three *SERPINB9* probes. The second study measured gene expression in THP-1 macrophage infected with Mycobacterium tuberculosis H37Rv strain[76]. The gene expression level of each probe was standardized to zero mean and one standard deviation (z-score). Within each group, the scattered dots represent sample values, and the thick line represents the median value. The bottom and top of the boxes are the 25th and 75th percentiles (interquartile range). Whiskers encompass 1.5 times the inter-quartile range.

(b) Expression-Atlas results. The source of each study is available in Supplementary Table 11. HRV: Human Rhino Virus; CMV: Cyto Megalo Virus; MPV: Meta Pneumo Virus; HCV: Hepatitis C Virus.

	Coef	Stderr	Z	Pr(>|z|)
Age	0.02	0.01	3.55	3.78E-04
Gender	0.02	0.17	0.12	9.07E-01
Stage	0.29	0.09	3.31	9.34E-04
CTL	-0.50	0.15	-3.32	9.05E-04
TGFB1	-0.10	0.10	-1.04	3.00E-01
CTL * *TGFB1*	0.11	0.03	3.47	5.18E-04

a. Antagonistic interaction

	Coef	Stderr	Z	Pr(>|z|)
Age	0.02	0.01	3.26	1.11E-03
Gender	0.03	0.17	0.15	8.80E-01
Stage	0.29	0.09	3.33	8.63E-04
CTL	-0.79	0.21	-3.79	1.51E-04
SOX10	-0.01	0.10	-0.11	9.10E-01
CTL * *SOX10*	-0.59	0.16	-3.69	2.23E-04

b. Synergistic interaction

Supplementary Table 1. The interactions between the cytotoxic T lymphocyte level and other candidate genes. We used the Cox-PH regression to test how does the expression level of a candidate gene interact with the cytotoxic T lymphocyte (CTL) level (average expression of *CD8A*, *CD8B*, *GZMA*, *GZMB*, and *PRF1*) to affect the patient overall survival using TCGA metastatic melanoma data. Clinical factors (e.g., age, gender and stage) are included as the background in regression. The statistical significance of coefficients was estimated by the two-sided Wald test in a Cox-PH regression using data from 317 samples.

(a) Antagonistic interaction between *TGFB1* and CTL. In an antagonistic interaction, a higher value of *TGFB1* in tumors will decrease the beneficial association between CTL and survival outcome.

(b) Synergistic interaction between *SOX10* and CTL. In a synergistic interaction, a higher value of *SOX10* in tumors will increase the beneficial association between CTL and survival outcome.

Database	Total	10% death	15K genes	50 patients	CTL profiled
TCGA	49	37	36	32	25
PRECOG	135	120	50	43	43
METABRIC	5	5	5	5	5

a. all datasets

Name	Database	Count	Description
SKCM	TCGA	317	Metastatic tumors of skin cutaneous melanoma[77].
UCEC		541	All tumors of uterine corpus endometrial carcinoma[78].
TNBC	METABRIC	233	Triple negative breast tumors[32].
AML	PRECOG	79	Acute myeloid leukemia profiled by U133+2.0 array[79].
NB		389	All neuroblastoma tumors[80].

b. significant datasets

Supplementary Table 2. Cancer gene expression datasets.

(a) Each column indicates the number of datasets that passed each filtering criterion. In TCGA, the same cancer type might be profiled by both RNA-Seq and microarray platforms, and we counted the datasets from different platforms separately. Breast tumors are analyzed separately in PAM50 subtypes, including luminal A, luminal B, Her2, basal, and triple negative (a variation of basal subtype). The TCGA melanoma cohort has two tumor categories including primary and metastatic subsets. **Total**: total number of datasets collected from the TCGA[31], PRECOG[17], and METABRIC[32] databases; **10% death**: the dataset should have more than 10% patient death for robust analysis in the Cox-PH regression; **15K genes**: The transcriptome profiling platform should include more than 15,000 genes; **50 patients**: the dataset should have more than 50 patients for robust analysis in the Cox-PH regression; **CTL profiled**: all cytotoxic T lymphocyte (CTL) markers, including *CD8A, CD8B, GZMA, GZMB,* and *PRF1*, should be included by the transcriptome profiling platform.

(b) The five datasets can predict a sufficient number of genes (>1% of all genes) with statistically significant *p*-values, false discovery rate (FDR) < 0.1, in the interaction test. All *p*-values are computed from the two-sided Wald test on the interaction covariates (Figure 1b and Supplementary Table 1). For each cohort, *p*-values are converted to FDR values through the Benjamini-Hochberg procedure[61]. (**Count**: number of patients profiled)

Supplementary Table 3. Significant T-cell dysfunction scores. All *p*-values are computed from the two-sided Wald test on the interaction covariates (Figure 1b and Supplementary Table 1). For each cohort, *p*-values are converted to FDR values through the Benjamini-Hochberg procedure[61]. For each of the five datasets in Figure 1c, we show all significant T-cell dysfunction scores (FDR < 0.1).

http://tide.dfci.harvard.edu/download/significant_dysfunction_scores.xlsx

Name	Description
T accum	In-vivo shRNA screen in mouse T cells to identify genes whose knockdown can increase the efficiency of T-cell accumulation in tumors[33]. The top hits are defined as genes with a median fold change greater than 2 in the primary screen and larger than one in the validation screen. The negative hits are defined as genes with median fold change smaller than one in the primary screen. In totally, there are 17 and 88 positive and negative hits, respectively.
T exhaust	Gene expression difference between exhausted CD8 T cells and activated CD8 T cells in a mouse model[34]. The positive and negative hits are defined as the top and bottom 50 genes ranked by the logFC of differential expression.
T regulatory	Gene expression change of CD4 regulatory T cells before and after activation[35]. The positive and negative hits are defined as the top and bottom 50 genes ranked by the logFC of differential gene expression.
ICB resist	Gene expression difference between anti-CTLA4 resistant mouse tumors and parental sensitive B16 tumors[36]. The positive and negative hits are defined as the top and bottom 50 genes ranked by the logFC of differential expression.
T exh Fixed	Gene expression features of exhausted CD8 T cells at different stages[38]. For each time point, we computed a logFC vector testing the gene expression difference between exhausted T cells and naïve T cells (analyzed in Figure 2d). We also computed an overall logFC vector testing the differential expression between late stage (after day 14) and early stage (day 5) (analyzed in Supplementary Figure 5b).
MDSC	Gene expression profiles of myeloid derived suppressor cells that can inhibit T-cell activation compared to monocytes sorted from peripheral blood mononuclear cells[81].
TAM M2/M1	Gene expression profile of M2 macrophages compared to M1 macrophages[82].
CAF	Gene expression profile of FAP+ cancer-associated fibroblasts compared to other cell types sorted from the same patients[83].

Supplementary Table 4. The definition of gene signatures related with tumor immune evasion.

Supplementary Table 5. The log-fold change of genes in each signature of tumor immune evasion. The last column "Hit" indicates the status of gene hits in our analysis. Pos: positive, Neg: negative.

URL: http://tide.dfci.harvard.edu/download/signatures.xlsx

	Name	CTL	Dysfunction	Exclusion	Count
KIRC	Kidney Renal Clear Cell Carcinoma	0.84	0.14	-0.02	19418
HNSC	Head And Neck Squamous Cell Carcinoma	0.41	-0.01	0.06	19426
ESCA	Esophageal Carcinoma	0.37	-0.07	0.00	19434
BRCA	Breast Invasive Carcinoma	0.23	-0.05	-0.02	19415
KIRP	Kidney Renal Papillary Cell Carcinoma	0.18	0.01	0.01	19383
CHOL	Cholangiocarcinoma	0.09	0.02	0.05	18895
UCEC	Uterine Corpus Endometrial Carcinoma	0.06	-0.07	-0.08	19446
STAD	Stomach Adenocarcinoma	0.05	-0.16	0.00	19442
KICH	Kidney Chromophobe	0.04	0.04	-0.04	19127
BLCA	Bladder Urothelial Carcinoma	0.03	-0.12	-0.04	19407
PRAD	Prostate Adenocarcinoma	0.01	-0.08	-0.02	19403
LIHC	Liver Hepatocellular Carcinoma	-0.08	-0.08	0.08	19340
LUAD	Lung Adenocarcinoma	-0.16	-0.19	0.12	19366
OV	Ovarian Serous Cystadenocarcinoma	-0.22	-0.29	0.04	17671
THCA	Thyroid Carcinoma	-0.34	0.08	0.04	19348
LUSC	Lung Squamous Cell Carcinoma	-0.43	-0.24	0.15	19411

Supplementary Table 6. Correlation with gene signatures of T-cell dysfunction and exclusion in tumors. For each TCGA cancer type with normal control samples, we calculated the average expression difference between tumor versus normal samples. We then computed the Pearson correlation between that value and the TIDE signatures of T-cell dysfunction and T-cell exclusion. The number of genes in the correlation computation is shown as the last column. The CTL level difference between tumor and normal samples is shown as the first column.

Signature	Description
Mutation	Total count of non-synonymous mutations in a tumor.
IFNG	Interferon gamma (IFNγ) response biomarkers of 6 genes including *IFNG, STAT1, IDO1, CXCL10, CXCL9*, and *HLA-DRA*[7].
CD8	Gene expression level of *CD8A + CD8B*.
IPS	Computational method Immunophenoscore to predict immunotherapy response from pre-treatment tumor expression profiles[46].
PDL1	An immunohistochemistry (IHC) biomarker approved by FDA[5]. In this study, we used the *PD-L1* gene expression as the IHC surrogate.
SCNA	Tumor somatic copy number alterations as biomarkers of immunotherapy response[9].
CRMA	Anti-CTLA4 resistance MAGE genes[84], including *MAGEA2, MAGEA2B, MAGEA3, MAGEA6*, and *MAGEA12*.
T clonality	The beta chain clonality of T cell receptor CDR3 sequence assembled from RNA-Seq reads[85]. Clonality = 1 - $\sum_{i=1}^{N} p_i \log \frac{1}{p_i} / \log N$ (p_i: the frequency of each receptor sequence).
B clonality	The immunoglobulin heavy chain clonality of B cell receptor sequence assembled from RNA-Seq reads[85].
Exercise	The tumor expression profiles of mice with regular running exercise compared with mice without exercise[86].

Supplementary Table 7. Biomarkers for response to immune checkpoint blockade.

Compare	AUC TIDE	Compare	AUC other	D-stat	p-value
PD1 RNA-Seq	0.83	PDL1	0.62	1.43	7.66E-02
		IFNG	0.51	2.06	1.95E-02
		Mutation	0.63	1.44	7.43E-02
		CD8	0.50	2.31	1.04E-02
		T. Clonality	0.58	1.52	6.37E-02
		IPS	0.64	1.33	9.25E-02
		SCNA	0.52	2.67	3.78E-03
CTLA4 RNA-Seq	0.80	PDL1	0.62	1.50	6.66E-02
		IFNG	0.67	1.08	1.41E-01
		Mutation	0.69	0.84	2.02E-01
		CD8	0.70	0.83	2.04E-01
		T. Clonality	0.43	3.34	4.17E-04
		IPS	0.61	1.80	3.59E-02
		SCNA	0.66	1.01	1.56E-01
PD1 NanoString	0.76	PDL1	0.42	2.33	9.96E-03
		IFNG	0.62	1.22	1.11E-01
		CD8	0.56	1.43	7.68E-02
		IPS	0.44	3.06	1.11E-03

a. Comparison among ICB biomarkers.

	AUC Exclusion	AUC CTL	D-stat	p-value
PD1 RNASeq	0.86	0.59	2.17	1.51E-02
CTLA4 RNASeq	0.84	0.77	0.91	1.81E-01
PD1 NanoString	0.74	0.64	0.88	1.89E-01

b. Comparison between T-cell exclusion signature and CTL.

Compare	AUC Dysfunction	Compare	AUC other	D-stat	p-value
T cell dysfunction	0.8	PDL1	0.44	2.95	1.58E-03
		T accum	0.62	1.50	6.67E-02
		T exhaust	0.71	0.66	2.53E-01
		T regulatory	0.68	0.93	1.75E-01

c. Comparison among T-cell dysfunction signatures.

Supplementary Table 8. Comparison of area under ROC curve (AUC) among biomarkers. In the bootstrap test, the D-stat is computed as (AUC1 – AUC2)/SD(AUC1 – AUC2), where the standard deviation (SD) of AUC difference was estimated through 2000 bootstrap replicates[87]. Then, the D-stat was compared to the normal distribution to compute the one-sided *p*-values.

(a) TIDE and other ICB biomarkers in Figure 4g-i.

(b) T-cell exclusion signatures and cytotoxic T lymphocyte level in Supplementary Figure 5e.

(c) T-cell dysfunction signatures in Supplementary Figure 5f.

	Coef	Stderr	Z	Pr(>|z|)
Age	-0.024	0.014	-1.701	8.902E-02
Gender	0.115	0.384	0.299	7.649E-01
Stage	0.251	0.249	1.006	3.145E-01
Pre-Therapy	-0.003	0.126	-0.023	9.816E-01
CTL	-0.590	0.191	-3.092	1.987E-03
SERPINB9	0.437	0.167	2.618	8.850E-03

a. Progression-free survival

	Coef	Stderr	Z	Pr(>|z|)
Age	0.001	0.017	0.087	9.309E-01
Gender	0.675	0.457	1.477	1.398E-01
Stage	0.277	0.276	1.003	3.157E-01
Pre-Therapy	0.149	0.131	1.138	2.551E-01
CTL	-0.682	0.201	-3.389	7.007E-04
SERPINB9	0.446	0.175	2.551	1.073E-02

b. Overall survival

Supplementary Table 9. High *SERPINB9* level correlates with short patient survival after ICB treatment. The association between the *SERPINB9* expression level (pre-treatment) and patient survival outcome after anti-CTLA4 treatment are tested by the Cox-PH regression using a public dataset with 42 patients[3]. Several clinical and tumor factors are included as background in the regression. **Pre-Therapy**: whether the patients were treated with other therapies before. **CTL**: the cytotoxic T lymphocyte level.

(a) Results with the progress free survival as end points.
(b) Results with the overall survival as end points.

B16 / T-cell	Condition	Median Perturb	Median Control	t-value	*p*-value
1:1	KO 1	0.528	0.942	-33.27	4.87E-06
	KO 2	0.395	0.942	-49.34	1.01E-06
2:1	KO 1	0.686	0.957	-18.72	4.79E-05
	KO 2	0.573	0.957	-22.02	2.52E-05
3:1	KO 1	0.783	0.961	-19.23	4.31E-05
	KO 2	0.808	0.961	-5.40	5.71E-03
1:1	Over expression	2.15	0.86	27.75	1.30E-03
2:1		1.50	0.88	130.44	5.88E-05
3:1		1.27	0.91	68.65	2.12E-04

Supplementary Table 10. The effect of *Serpinb9* perturbation on T-cell mediated tumor killing. The relative abundance values of tumor cells in *Serpinb9* perturbed and control conditions are compared through the two-sided student t-test. For the comparisons between knockout (KO) and control conditions, there are three cell-culture replicates for both conditions (t-test degree of freedom = 4). For the comparisons between over-expression and control conditions, there are two cell-culture replicates for both conditions (t-test degree of freedom = 2).

Accession	Pathogen	Cell	Condition	logFC	Adjusted *p*-value
GEOD-61141	HRV16	Airway	Asthma	3.2	1.53E-21
GEOD-12108	F novicida	Monocyte		2.9	6.63E-04
GEOD-61141	HRV16	Airway	Normal	2.5	6.00E-09
GEOD-11348	HRV16	Epithelial		2.3	1.23E-04
GEOD-12108	SchuS4	Monocyte		2	9.67E-04
GEOD-49016	C burnetii	Dendritic		1.8	2.70E-05
GEOD-11408	CMV	Monocyte		1.7	2.46E-05
GEOD-8961	MPV	A549	72h	1.5	6.50E-05
MEXP-1274	Influenza A	BEAS2B		1.5	6.40E-03
MEXP-1274	Influenza A	BEAS2B	TLR3	1.5	2.43E-02
GEOD-8961	MPV	A549	48h	1.4	3.08E-04
MEXP-3521	M tuberculosis	Macrophage		1.3	3.26E-02
GEOD-18816	H5N1	Macrophage		1.3	3.79E-02
GEOD-17400	Dhori virus	Epithelial		1.2	1.43E-06
GEOD-31455	HCV JFH1.2A	Hepatocyte		1.2	2.64E-03
GEOD-49016	O tsutsugamushi	Dendritic		1	9.68E-04

Supplementary Table 11. The *SERPINB9* expression change after pathogen infection from the Expression-Atlas database. The log2-fold change (logFC) and adjusted *p*-value of *SERPINB9* expression level between infected and uninfected conditions are shown. We also labeled the conditions for experiments with multiple measurements under different conditions. HRV: Human Rhino Virus; CMV: Cyto Megalo Virus; MPV: Meta Pneumo Virus; HCV: Hepatitis C Virus.

Cancer	Platform	Protein Corr	Expression Corr	Count
CHOL	RNASeq	-0.694	-0.019	30
THYM	RNASeq	-0.642	-0.507	87
GBM	RNASeq	-0.471	0.187	82
PRAD	RNASeq	-0.311	-0.018	351
TGCT	RNASeq	-0.265	-0.427	122
OV	Microarray	-0.264	-0.242	419
OV	RNASeq	-0.255	-0.161	241
PCPG	RNASeq	-0.247	0.127	82
GBMLGG	RNASeq	-0.246	0.335	517
ACC	RNASeq	-0.244	-0.196	46
KIRC	RNASeq	-0.243	-0.252	475
BLCA	RNASeq	-0.235	0.094	340
ESCA	RNASeq	-0.225	-0.122	125
UCS	RNASeq	-0.190	-0.103	48
CESC	RNASeq	-0.188	-0.033	171
SKCM	RNASeq	-0.177	-0.161	353
COAD	RNASeq	-0.170	-0.017	359
KIRP	RNASeq	-0.137	-0.100	215
LUAD	RNASeq	-0.136	-0.077	360
SARC	RNASeq	-0.136	-0.210	224
KICH	RNASeq	-0.135	-0.093	63
COADREAD	RNASeq	-0.132	-0.007	487
BRCA	RNASeq	-0.112	-0.077	887
UCEC	RNASeq	-0.106	-0.048	438
STES	RNASeq	-0.096	-0.020	461
READ	RNASeq	-0.083	0.027	128
THCA	RNASeq	-0.079	-0.118	222
STAD	RNASeq	-0.072	-0.010	336
LGG	RNASeq	-0.072	0.222	435
LUSC	RNASeq	-0.071	0.018	325
HNSC	RNASeq	-0.045	-0.016	212
PAAD	RNASeq	-0.031	-0.095	116
MESO	RNASeq	0.056	-0.105	63
DLBC	RNASeq	0.210	0.114	33

Supplementary Table 12. Pearson correlations between CTL level and *CTNNB1* molecular status. The Pearson correlations and number of samples are shown across TCGA cancer cohorts for both protein and expression levels.

Reference

64 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

65 Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell systems* **1**, 417-425, doi:10.1016/j.cels.2015.12.004 (2015).

66 Spranger, S., Bao, R. & Gajewski, T. F. Melanoma-intrinsic beta-catenin signalling prevents anti-tumour immunity. *Nature* **523**, 231-235, doi:10.1038/nature14404 (2015).

67 Chakravarty, D. *et al.* OncoKB: a precision oncology knowledge base. *JCO Precision Oncology* **1**, 1-16 (2017).

68 Mei, S. *et al.* Cistrome Data Browser: a data portal for ChIP-Seq and chromatin accessibility data in human and mouse. *Nucleic acids research* **45**, D658-D662, doi:10.1093/nar/gkw983 (2017).

69 Mancino, A. *et al.* A dual cis-regulatory code links IRF8 to constitutive and inducible gene expression in macrophages. *Genes & development* **29**, 394-408, doi:10.1101/gad.257592.114 (2015).

70 Garber, M. *et al.* A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Molecular cell* **47**, 810-822, doi:10.1016/j.molcel.2012.07.030 (2012).

71 Karwacz, K. *et al.* Critical role of IRF1 and BATF in forming chromatin landscape during type 1 regulatory cell differentiation. *Nat Immunol* **18**, 412-421, doi:10.1038/ni.3683 (2017).

72 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

73 Li, T. *et al.* TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer research* **77**, e108-e110, doi:10.1158/0008-5472.CAN-17-0307 (2017).

74 Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189-196, doi:10.1126/science.aad0501 (2016).

75 Butchar, J. P. *et al.* Microarray analysis of human monocytes infected with Francisella tularensis identifies new targets of host response subversion. *PloS one* **3**, e2924, doi:10.1371/journal.pone.0002924 (2008).

76 Verway, M. *et al.* Vitamin D induces interleukin-1beta expression: paracrine macrophage epithelial signaling controls M. tuberculosis infection. *PLoS Pathog* **9**, e1003407, doi:10.1371/journal.ppat.1003407 (2013).

77 Cancer Genome Atlas, N. Genomic Classification of Cutaneous Melanoma. *Cell* **161**, 1681-1696, doi:10.1016/j.cell.2015.05.044 (2015).

78 Cancer Genome Atlas Research, N. *et al.* Integrated genomic characterization of endometrial carcinoma. *Nature* **497**, 67-73, doi:10.1038/nature12113 (2013).

79 Metzeler, K. H. *et al.* An 86-probe-set gene-expression signature predicts survival in cytogenetically normal acute myeloid leukemia. *Blood* **112**, 4193-4201, doi:10.1182/blood-2008-02-134411 (2008).

80 Oberthuer, A. *et al.* Comparison of performance of one-color and two-color gene-expression analyses in predicting clinical endpoints of neuroblastoma patients. *Pharmacogenomics J* **10**, 258-266, doi:10.1038/tpj.2010.53 (2010).

81 Yaddanapudi, K. *et al.* MIF Is Necessary for Late-Stage Melanoma Patient MDSC Immune Suppression and Differentiation. *Cancer Immunol Res* **4**, 101-112, doi:10.1158/2326-6066.CIR-15-0070-T (2016).

82 Beyer, M. *et al.* High-resolution transcriptome of human macrophages. *PloS one* **7**, e45466, doi:10.1371/journal.pone.0045466 (2012).

83 Calon, A. *et al.* Dependency of colorectal cancer on a TGF-beta-driven program in stromal cells for metastasis initiation. *Cancer cell* **22**, 571-584, doi:10.1016/j.ccr.2012.08.013 (2012).

84 Shukla, S. A. *et al.* Cancer-Germline Antigen Expression Discriminates Clinical Outcome to CTLA-4 Blockade. *Cell* **173**, 624-633 e628, doi:10.1016/j.cell.2018.03.026 (2018).

85 Bolotin, D. A. *et al.* Antigen receptor repertoire profiling from RNA-seq data. *Nature biotechnology* **35**, 908-911, doi:10.1038/nbt.3979 (2017).

86 Pedersen, L. *et al.* Voluntary Running Suppresses Tumor Growth through Epinephrine- and IL-6-Dependent NK Cell Mobilization and Redistribution. *Cell metabolism* **23**, 554-562, doi:10.1016/j.cmet.2016.01.011 (2016).

87 Hanley, J. A. & McNeil, B. J. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* **148**, 839-843, doi:10.1148/radiology.148.3.6878708 (1983).