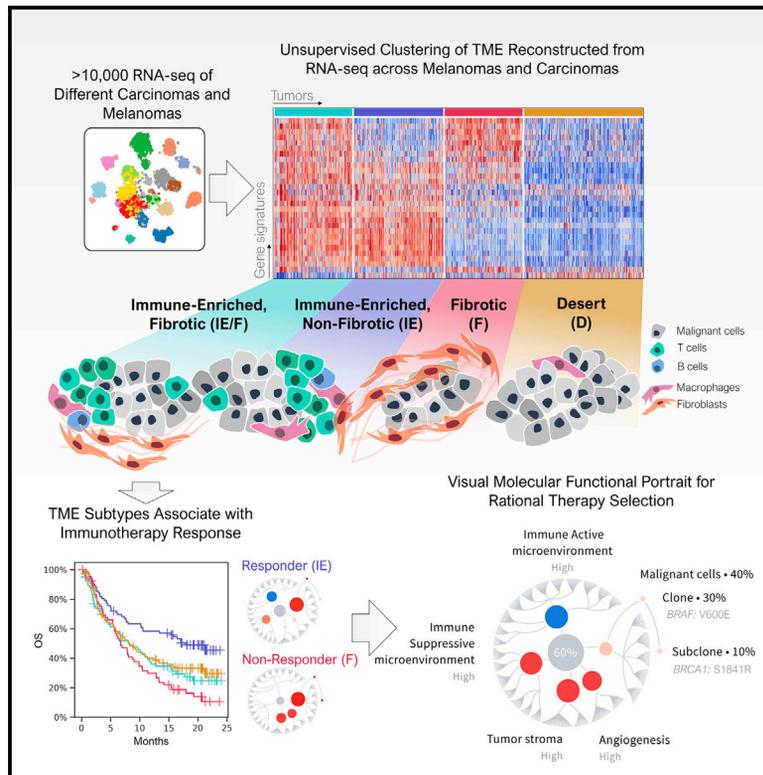


Conserved pan-cancer microenvironment subtypes predict response to immunotherapy

Graphical abstract



Authors

Alexander Bagaev, Nikita Kotlov, Krystle Nomie, ..., Maria Tsiper, Ravshan Ataullakhanov, Nathan Fowler

Correspondence

ravshan.ataullakhanov@bostongene.com (R.A.), nfowler@mdanderson.org (N.F.)

In brief

Bagaev et al. identify four tumor microenvironment (TME) subtypes that are conserved across diverse cancers and correlate with immunotherapy response in melanoma, bladder, and gastric cancers. A visual tool revealing the TME subtypes integrated with targetable genomic alterations provides a planetary view of each tumor that can aid in oncology clinical decision making.

Highlights

- Development of a holistic transcriptomic-based TME classification platform
- Detection of four immune/fibrotic TME subtypes conserved in a broad array of cancers
- The four TME subtypes are predictive of response to immunotherapy in multiple cancers
- Integration of genomics and transcriptomics into a visual tool with a planetary view



Article

Conserved pan-cancer microenvironment subtypes predict response to immunotherapy

Alexander Bagaev,^{1,3} Nikita Kotlov,^{1,3} Krystle Nomie,¹ Viktor Svekolkina,¹ Azamat Gafurov,¹ Olga Isaeva,¹ Nikita Osokin,¹ Ivan Kozlov,¹ Felix Frenkel,¹ Olga Gancharova,¹ Nava Almog,¹ Maria Tsiper,¹ Ravshan Ataulakhov,^{1,*} and Nathan Fowler^{1,2,4,*}

¹BostonGene, Waltham, MA 02453, USA

²Department of Lymphoma and Myeloma, Unit 0429, the University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA

³These authors contributed equally

⁴Lead contact

*Correspondence: ravshan.ataullakhanov@bostongene.com (R.A.), nfowler@mdanderson.org (N.F.)

<https://doi.org/10.1016/j.ccell.2021.04.014>

SUMMARY

The clinical use of molecular targeted therapy is rapidly evolving but has primarily focused on genomic alterations. Transcriptomic analysis offers an opportunity to dissect the complexity of tumors, including the tumor microenvironment (TME), a crucial mediator of cancer progression and therapeutic outcome. TME classification by transcriptomic analysis of >10,000 cancer patients identifies four distinct TME subtypes conserved across 20 different cancers. The TME subtypes correlate with patient response to immunotherapy in multiple cancers, with patients possessing immune-favorable TME subtypes benefiting the most from immunotherapy. Thus, the TME subtypes act as a generalized immunotherapy biomarker across many cancer types due to the inclusion of malignant and microenvironment components. A visual tool integrating transcriptomic and genomic data provides a global tumor portrait, describing the tumor framework, mutational load, immune composition, anti-tumor immunity, and immunosuppressive escape mechanisms. Integrative analyses plus visualization may aid in biomarker discovery and the personalization of therapeutic regimens.

INTRODUCTION

Precision medicine has the potential to revolutionize cancer care as accumulating evidence suggests that patients who receive personalized therapy have better clinical outcomes (Schwaeberle et al., 2016). The genomic characterization of tumors has become more common in both clinical trials and in standard care. Despite the growing acceptance of genomic analysis as a part of clinical decision making, genomic characterization often largely entails the utilization of targeted panels comprising only a limited number of genes that capture a fraction of oncogenic alterations (Suwinski et al., 2019; Castellanos et al., 2017; McCabe et al., 2019). Transcriptomic analysis offers an additional opportunity to dissect the complexity and heterogeneity of tumors and to discover new biomarkers that can be used to develop novel therapeutic strategies (Cieřlik and Chinnaiyan, 2018). Whole-exome sequencing (WES) and RNA sequencing (RNA-seq) accompanied by conventional pathological, immunohistochemical, and clinical tests offer a multifaceted view of tumor characteristics and can potentially lead to the further identification and optimization of individual cancer patient therapy (El-Deiry et al., 2019; Koepfel et al., 2018; Malone et al., 2020). Nevertheless, large-scale exome and transcriptome sequencing provide thousands of parameters, which is often overwhelming and unus-

tainable for routine treatment decision support (Koepfel et al., 2018; Schwarze et al., 2020; Singer et al., 2019; Wadapurkar and Vyas, 2018).

The tumor microenvironment (TME) plays a significant role in clinical outcomes and response to therapy. By exerting pro- and anti-tumorigenic actions, tumor-infiltrating immune cells can profoundly influence tumor progression and the success of anti-cancer therapies (Fridman et al., 2012; Chen and Mellman, 2017). Cancer-associated fibroblasts (CAFs) as well as angiogenic signals from stromal cells have been shown to affect outcomes (Liu et al., 2019b; Pommier and Fearon, 2016; Schaaf et al., 2018; Tao et al., 2017). Deciphering the tumor-immune microenvironment profile of a cancer can improve the tailoring of targeted and immunotherapeutic strategies. However, to date, comprehensive and integrated genomic and transcriptomic analyses that assess the tumor and the TME as a whole remain rare, are often unstructured, and lack efficient and useful models. Here, we developed an accessible transcriptomic analysis platform for TME classification to identify functional gene sets defined by cellular populations, signaling signatures, TME processes, and cancer cell properties that ensure comprehensive characterization of both the tumor and its TME. Finally, transcriptomic and genomic data integration were visualized, showing a planetary view of the tumor through a comprehensive tumor portrait.



RESULTS

Establishment of TME gene expression signatures

To classify TMEs using a transcriptomic-based analytical platform, knowledge-based functional gene expression signatures (Fges) representing the major functional components and immune, stromal, and other cellular populations of the tumor were selected using published literature (STAR Methods, Figure 1A, Table S1), ultimately creating a holistic approach thoroughly describing the TME within one single model. T cell activity and antigen-presentation Fges were chosen according to the principles of the cancer immunogram (Blank et al., 2016) and cancer immunity cycle (Chen and Mellman, 2013). Fges that characterize the stromal compartments (e.g., angiogenesis, CAFs), immune suppression via macrophages and myeloid-derived suppressor cells, and Fges related to cancer cell properties (Sanchez-Vega et al., 2018) were included. Ultimately, a manually curated list of 29 Fges was created covering known cellular and functional TME properties (STAR Methods, Table S1; Figures 1A and 1B). Each Fges was refined to include only genes associated with a specific cell type or biological process (Table S2, Figures 1C, S1A, and S1B).

The unique expression patterns of these genes were validated by cross correlation of gene expression within each signature using RNA-seq analysis of tissue samples from The Cancer Genome Atlas (TCGA), International Cancer Genome Consortium (ICGC), or Genotype-Tissue Expression (GTEx) databases (Figure 1A). Specifically, the relative expression level of each Fges was calculated by single-sample gene set enrichment analysis (ssGSEA), which was applied across data types to score samples based on gene set expression (Aran et al., 2017; Charoentong et al., 2017). To confirm the cell type-specific expression patterns of each of the selected genes, additional gene selection was performed based on 4,212 RNA-seq profiles of sorted immune and stromal cell subpopulations across multiple public datasets collected from FANTOM5 (FANTOM Consortium and the RIKEN PMI and CLST (DGT) et al., 2014), ENCODE (Raney et al., 2011; Consortium and The ENCODE Project Consortium, 2004), BLUEPRINT, IRIS, and the Gene Expression Omnibus (GEO) (Edgar et al., 2002) (Figure 1A). As expected, the cytotoxic cell signature was predominantly expressed in sorted CD8⁺ T cells or natural killer (NK) cells, and Fges related to extracellular matrix formation were primarily enriched in fibroblasts. The final Fges were highly cell type specific, with high expression scores for cell types associated with each signature, and low scores for non-target cell populations (Mann-Whitney p value $< 10^{-100}$) (Figure S1B), showing effective segregation of the 4,212 cell profiles (Figure 1D). The expression of the tumor proliferation rate signature, which consisted of cell cycle- and tumor progression-associated genes, showed strong correlation and a progressive increase in malignant melanomas compared with normal tissue and nevi (Figure 1E). The final Fges showed a high level of correlation with The Molecular Signatures Database (MsigDB) (Liberzon et al., 2015; Subramanian et al., 2005) collection of hallmarks and canonical pathways and other previously published signatures (false discovery rate [FDR] < 0.05) (Table S3) (Charoentong et al., 2017; Cancer Genome Atlas Network, 2015; Şenbabaoglu et al., 2016; Bindea et al., 2013).

The expression patterns of the 29 Fges were further characterized using a large TCGA melanoma dataset (Figure 1F). Pearson correlation analysis of the 29 Fges revealed two major groups bound by positive correlations that were co-activated or co-repressed across the dataset. One group comprised major T cell Fges such as those responsible for trafficking, T-helper 1 (Th1), and regulatory T cell (Treg) phenotypes and effector cell function, as well as NK cell-, major histocompatibility complex (MHC) class II expression-, checkpoint inhibition-, and macrophage-associated Fges. Collectively, these Fges encompassed the immune compartment of the TME, including both anti-tumor-immune and tumor-promoting processes (Figure 1F). The second group consisted of Fges defining stromal components such as CAFs, matrix, and matrix remodeling, as well as endothelium, angiogenesis, and pro-tumor cytokine expression. Remarkably, these two groups (immune and stromal compartments) showed no correlation with cancer cell Fges such as the tumor proliferation rate and EMT Fges or the oncogenic activity of cell specific signaling pathways such as PI3K, VEGF, EGFR, p53, and MAPK using the PROGENY algorithm (Pathway RespOnsive GENes) (Schubert et al., 2018) (Figure S1C). As expected, NF- κ B, JAK-STAT, and TRAIL signaling activity correlated with the presence of active Fges of cytotoxic T cells, while transforming growth factor beta (TGF- β) signaling activity correlated with the abundance of CAFs, major producers of TGF- β (Erdogan and Webb, 2017). The survival probabilities were compared between patient groups divided by the median of the most representative Fges from each group, and the group consisting of immune-associated Fges strongly correlated with increased survival (Figures S1D and S1E). Lastly, the 29 Fges were correlated with other tumor and microenvironment classification methods such as CIBERSORT (Newman et al., 2015), the microenvironment cell populations-counter (MCP-counter) deconvolution algorithm (Becht et al., 2016), Şenbabaoglu et al. signatures (Şenbabaoglu et al., 2016), and the six TCGA immune subtypes (Thorsson et al., 2018) (Figure S1F), with positive correlation of the stromal and lymphocytic Fges with the MCP-counter algorithm and Şenbabaoglu et al. signatures, showing the connections among the different approaches.

Four melanoma microenvironment subtypes were revealed by unsupervised analysis of the TME Fges

Significant progress has been recently achieved using immunotherapy and therapeutic vaccines (Ott et al., 2017) for cutaneous melanoma (Weiss et al., 2019; Cuevas and Daud, 2018), emphasizing the importance of the immune microenvironment in this cancer. To classify the melanoma TME using this curated list of 29 Fges, their expression patterns were assessed across approximately 470 melanoma tumors (TCGA-SKCM [Skin Cutaneous Melanoma]; Cancer Genome Atlas Network, 2015) using unsupervised dense Louvain clustering based on their ssGSEA scores (Blondel et al., 2008). This analysis revealed that the molecular profiles of melanomas can be clustered into four distinct microenvironments termed (1) immune-enriched, fibrotic (IE/F); (2) immune-enriched, non-fibrotic (IE); (3) fibrotic (F); and (4) immune-depleted (D) (Figures 2A, 2B, S2A, and S2B, Table S4). Classification using the 29 Fges resulted in four distinct TME subtypes characterized by the presence of an immune-active

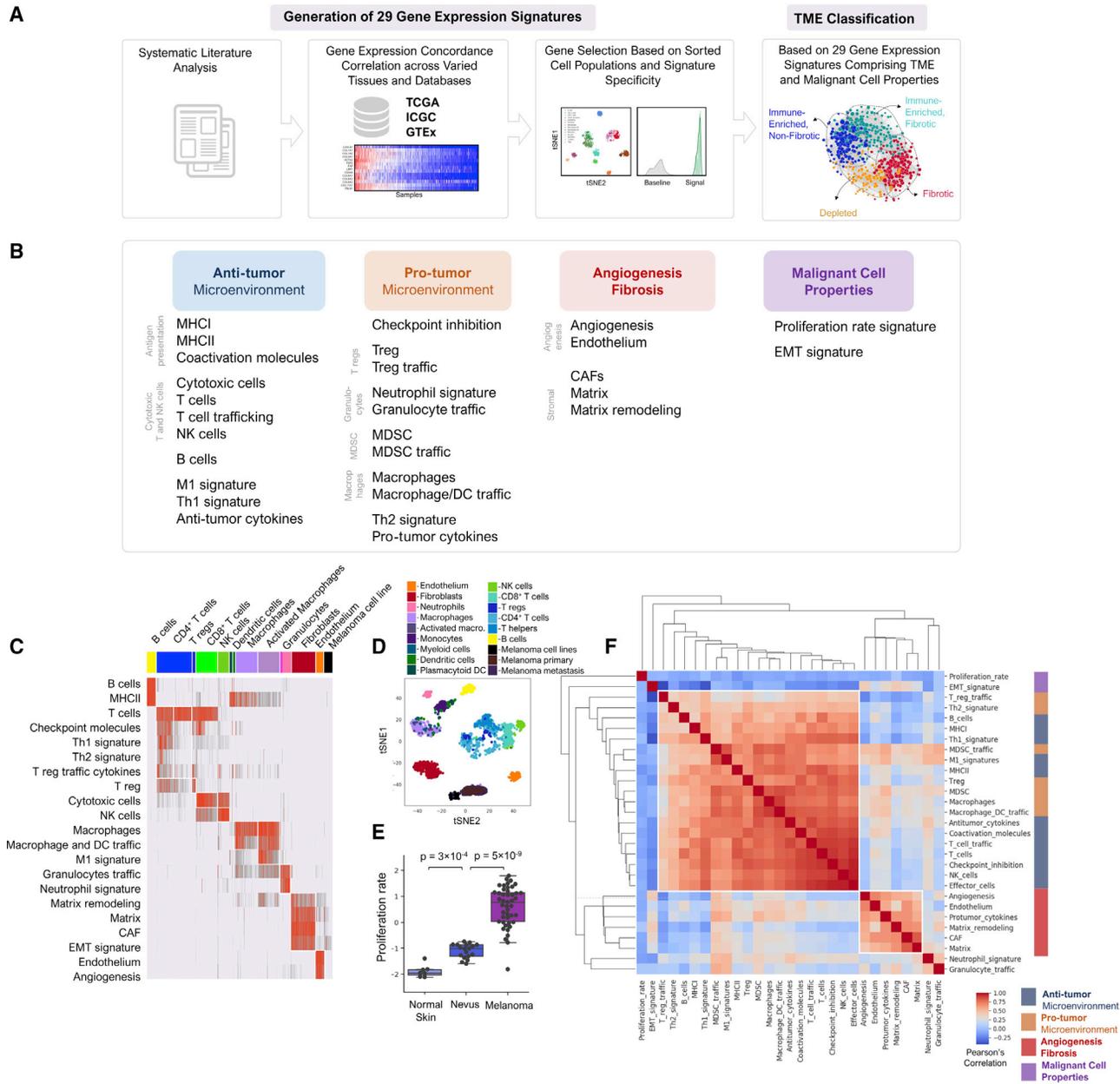


Figure 1. Generation of the 29 Fges utilized for transcriptomic-based TME classification

(A) Schematic overview of the workflow employed to generate the TME classification platform.

(B) The 29 Fges included in each functional group.

(C) Heatmap of the scores of the manually curated Fges (y axis) related to the cellular composition and functional properties measured in 3,462 samples of purified cell populations (x axis).

(D) t-Distributed stochastic neighbor embedding (tSNE) projection of the purified cell samples in the space of the signature scores.

(E) Box plots showing the proliferation rate signature expression scores in normal (n = 7), nevus (n = 18), and malignant tissue (n = 45). In the box plots, the upper whisker indicates the maximum value or 75th percentile +1.5 interquartile range (IQR); the lower whisker indicates the minimum value or 25th percentile -1.5 IQR.

(F) Pearson correlation between gene signature scores of 470 TCGA cutaneous melanoma (TCGA-SKCM) tumor samples.

See also Figure S1, Tables S1, S2, and S3.

or immunosuppressive microenvironment and tumor stroma, including CAFs and vascularization, compared with conventional classification methods that often result in fewer subtypes (Şenbabaoğlu et al., 2016). These TME subtypes varied significantly based on the expression of the 29 Fges, with each

subtype differing by the activity of at least six Fges ($p < 10^{-7}$; $FDR < 10^{-5}$) (Figures 2A and S3B). Differences between the TME subtypes distinguished by our approach were also observed when analyzing previously reported signatures (Chen and Mellman, 2017; Şenbabaoğlu et al., 2016), by cell type using

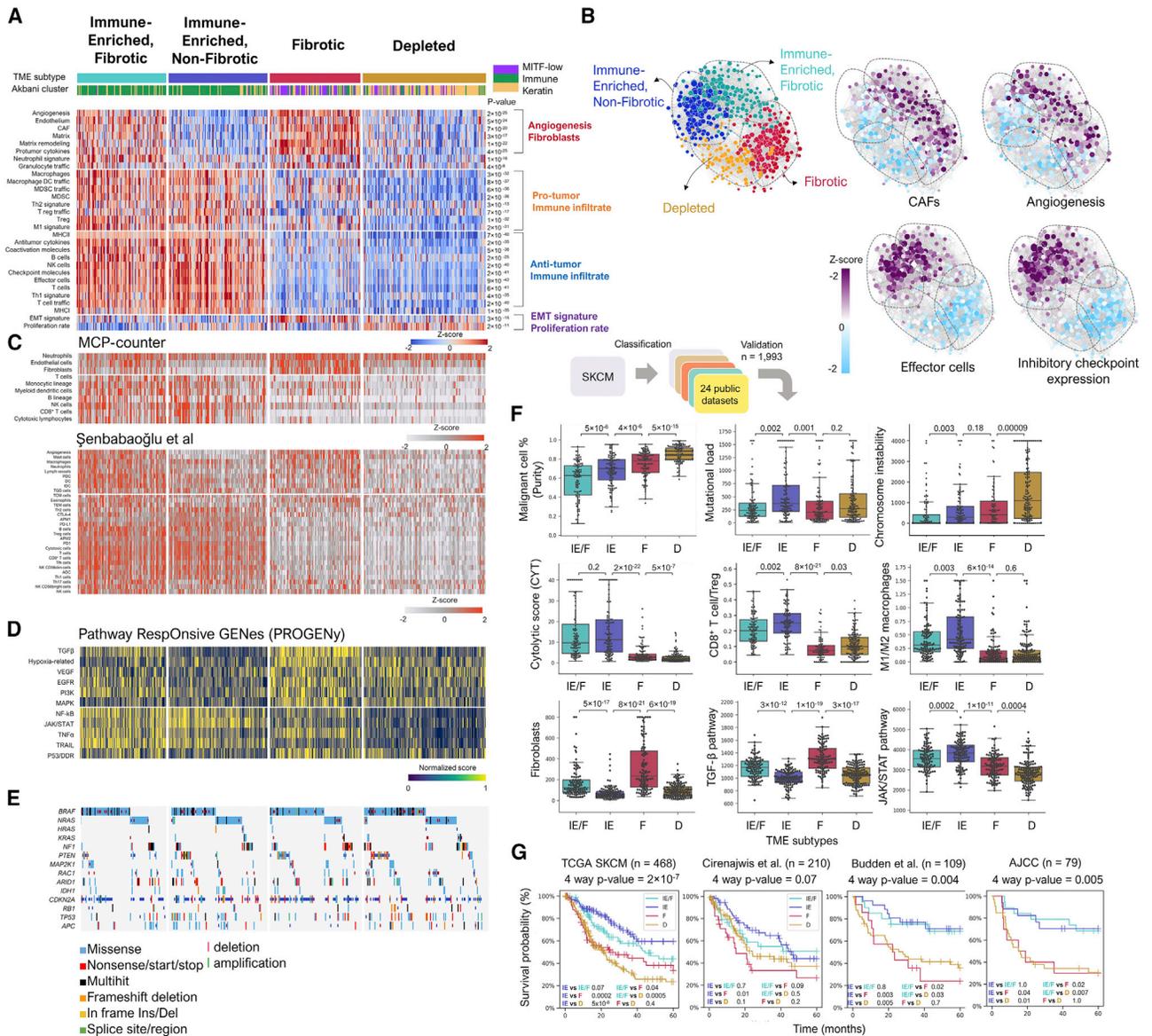


Figure 2. The four distinct TME subtypes identified in melanoma

(A) Heatmap of 468 TCGA cutaneous melanomas (TCGA-SKCM) classified into four distinct TME subtypes based on unsupervised dense clustering of the 29 Figs. p values were calculated with the chi-square test. Additional annotation includes melanoma expression subtypes "MITF-low," "immune," and "keratin" taken from TCGA.

(B) Correlation-based graph network of tumor samples representing observed dense clusters revealed by unsupervised dense Louvain clustering (left).

(C) Heatmap of gene signature scores derived from the MCP-counter cell deconvolution algorithm or [Şenbabaoğlu et al. \(2016\)](#) for the TCGA-SKCM dataset (n = 468).

(D) Relative signaling pathway activity scores in tumor cells measured from RNA-seq by PROGENy.

(E) Oncoplot of genomic alterations found in frequently mutated genes in melanoma.

(F) Box plots showing differences in tumor cellularity, total mutational burden, CIN score, cytolytic score (CYT), CD8⁺ T cells/Treg and M1/M2 macrophage ratios, fibroblast signature, and TGF-β and JAK/STAT pathway signaling activity calculated via PROGENy across the four TME subtypes for the TCGA-SKCM dataset.

(G) OS of melanoma patients stratified by TME subtype classification. In the box plots, the upper whisker indicates the maximum value or 75th percentile +1.5 IQR; the lower whisker indicates the minimum value or 25th percentile -1.5 IQR.

See also [Figures S2](#) and [S3](#), [Table S4](#).

the MCP-counter cell deconvolution method ([Figure 2C](#)) ([Becht et al., 2016](#)), and by malignant signaling pathway activities calculated with PROGENy ([Figure 2D](#)) ([Schubert et al.,](#)

[2018](#)). Previous signatures that did not consider stromal components did not allow for the independent identification of all four TME subtypes ([Figures S2C–S2F](#)) and instead only found three

TME subtypes primarily differing by inflammation levels. Genomic alterations frequently found in melanoma were also not consistently associated with the specific TME subtypes (Figure 2E).

The immune-inflamed IE/F melanomas were characterized by the elevated expression of Fges associated with angiogenesis and with CAF activation. The IE melanomas were distinguished by high levels of immune infiltrate and significantly increased cytolytic score (Rooney et al., 2015) and demonstrated a more immune-active microenvironment compared with the IE/F subtype. TME subtype IE melanomas also had the highest mutational load ($p < 0.001$) and highest ratios of CD8⁺ T cells/Tregs ($p < 0.001$) and M1/M2 macrophages ($p < 0.01$) in comparison with the other subtypes (Figure 2F). Analysis of activated signaling pathways revealed increased JAK/STAT pathway activation in IE melanomas, which may be associated with increased T cell activity (Figure 2F). The two additional TME subtypes F and D possessed minimal or completely lacked leukocyte/lymphocyte infiltration, with subtype D containing the highest malignant cell percentage (Figure 2F), correlating with previously described immune-desert phenotypes (Chen and Mellman, 2017; Galon and Bruni, 2019). Chromosomal instability (CIN) measured as the frequency of genome duplications or deletions (Ock et al., 2017) was increased in subtype D melanoma, suggesting a possible association of increased copy-number alterations (CNAs) with the TME subtype ($p < 0.001$) (Figure 2F). In contrast, melanomas classified as subtype F showed elevated expression of Fges related to angiogenesis (Figure 2A) and increased CAFs (Figure 2F). Fibroblasts are strong immune suppressors and TME remodelers via secretion of TGF- β (Liu et al., 2019b; Chakravarthy et al., 2018; Zhuang et al., 2015). Indeed, the TGF- β signaling pathway (Figure 2F) as well as the EMT transition and the cancer cell metastasis Fges were significantly upregulated in the melanoma tumors classified as subtype F (Figure 2A) and correlated with shortened patient survival to standard-of-care therapies (Figure 2G). The patients with subtype IE melanomas had significantly longer overall survival (OS) and progression-free survival (PFS) compared with patients with subtypes F and D (Figure 2G), with subtype F demonstrating the worst OS.

The same tumor classification system was then validated using 23 independent melanoma datasets, the largest collected melanoma transcriptomic dataset ($n = 1,993$ including TCGA-SKCM), and the same four TME subtypes were revealed (Figures S3A and S3B), with similar prognostic significance noted in the selected melanoma cohorts with available clinical annotation summarized using the entire dataset of 915 melanomas combined and individually (Figures 2G, S3C, and S3D). Again, melanomas classified as subtypes F and D were associated with decreased survival to various chemotherapy/standard-of-care regimens, including late-stage disease (Figures S3C and S3D). The observed prognostic significance was validated utilizing independent cohorts of patients with clinical annotation (Figure 2G).

The histology phenotypes related to the gene expression patterns in each TME subtype were investigated through analysis of representative melanoma tumor specimens from TCGA-SKCM ($n = 62$) to validate the TME subtypes histologically. Indeed, an increased abundance of lymphocytes in melanoma subtype IE and the high presence of fibroblasts in melanoma subtype F

were found (Figure 3A; Table S5), which was further confirmed using an independent melanoma dataset (Figure 3B, GEO: GSE8401) (Xu et al., 2008). Malignant cell content was the highest in subtype D (Figures 3A and 3B). The gene signature expression pattern among the four different TME subtypes histologically correlated with the tumor abundance and spatial organization of lymphocytes and stroma (Figure 3C). For example, lymphocytes were highly abundant in the TME subtypes IE/F and IE melanomas. Subtype IE/F melanomas had higher stromal content than subtype IE melanomas, whereas subtype F was highly fibrotic with dense collagen formation (Figure 3C). The majority of solid tumors display three major immunological phenotypes, termed immune inflamed, immune excluded, and immune desert (Hegde et al., 2016; Herbst et al., 2014); therefore, our TME subtypes were categorized into these three classifications. Quantitative histopathological review of TCGA melanomas revealed that subtype IE but not subtype IE/F melanomas were enriched with an immune-inflamed histological phenotype characterized by abundant lymphocyte infiltration (Figure 3D), which was also observed with a large, independent bladder cancer cohort ($n = 372$) (Mariathasan et al., 2018) (Figure 3D). The majority of subtype IE/F melanomas were characterized by an excluded phenotype (Figure 3D), and the subtype D melanomas were primarily characterized as desert (Figure 3D). Subtype IE, but not the subtype IE/F melanomas, had increased PD-L1 expression, both on lymphocytes ($p < 1 \times 10^{-10}$) and on tumor cells ($p = 5 \times 10^{-5}$) as measured by immunohistochemistry (Figure 3E), suggesting an environment whose anti-tumor immunity could be activated with anti-PD1/PD-L1 immune checkpoint blockade therapies. Lastly, the analysis of mapped TCGA digitized hematoxylin-eosin (H&E)-stained images (Saltz et al., 2018) showed the percentage of tumor-infiltrating lymphocytes to be significantly higher in the IE/F and IE TME subtypes of melanoma and bladder cancer (Figure 3F), providing additional large-scale histological validation of the TME subtypes.

The four TME subtypes are conserved across a broad array of cancers

Transcriptomic data from additional cancer types were analyzed to determine whether the identified four TME subtypes are also present in other cancers. The same gene expression-based classification system was applied to a cohort of 8,024 TCGA tumors. Due to significant intrinsic differences in cellular composition, expression analysis yielded distinct nosology-specific subsets (Figure 4A). To remove tissue-type-specific effects, we normalized gene signature values by median-score transformation within each cancer type (Figures S4A–S4C). Following normalization, tumors of different tissue origin formed a uniform single set in UMAP (Uniform Manifold Approximation and Projection) analysis (Figures 4A, 4B, and S4A). Normalization retained the relationship and correlations between different TME modules, resulting in the identification of the same four major TME subtypes found in melanoma within 24 various cancers ($n = 8,024$ tumors; Figures 4C and 4D). Independent validation of the classification system was performed on another pan-cancer cohort of 1,800 ExpO project adenocarcinomas where expression was measured by microarrays (GEO: GSE2109). Pearson correlation analysis of the 29 Fges across the 8,024 TCGA tumors revealed two major groups bound by positive

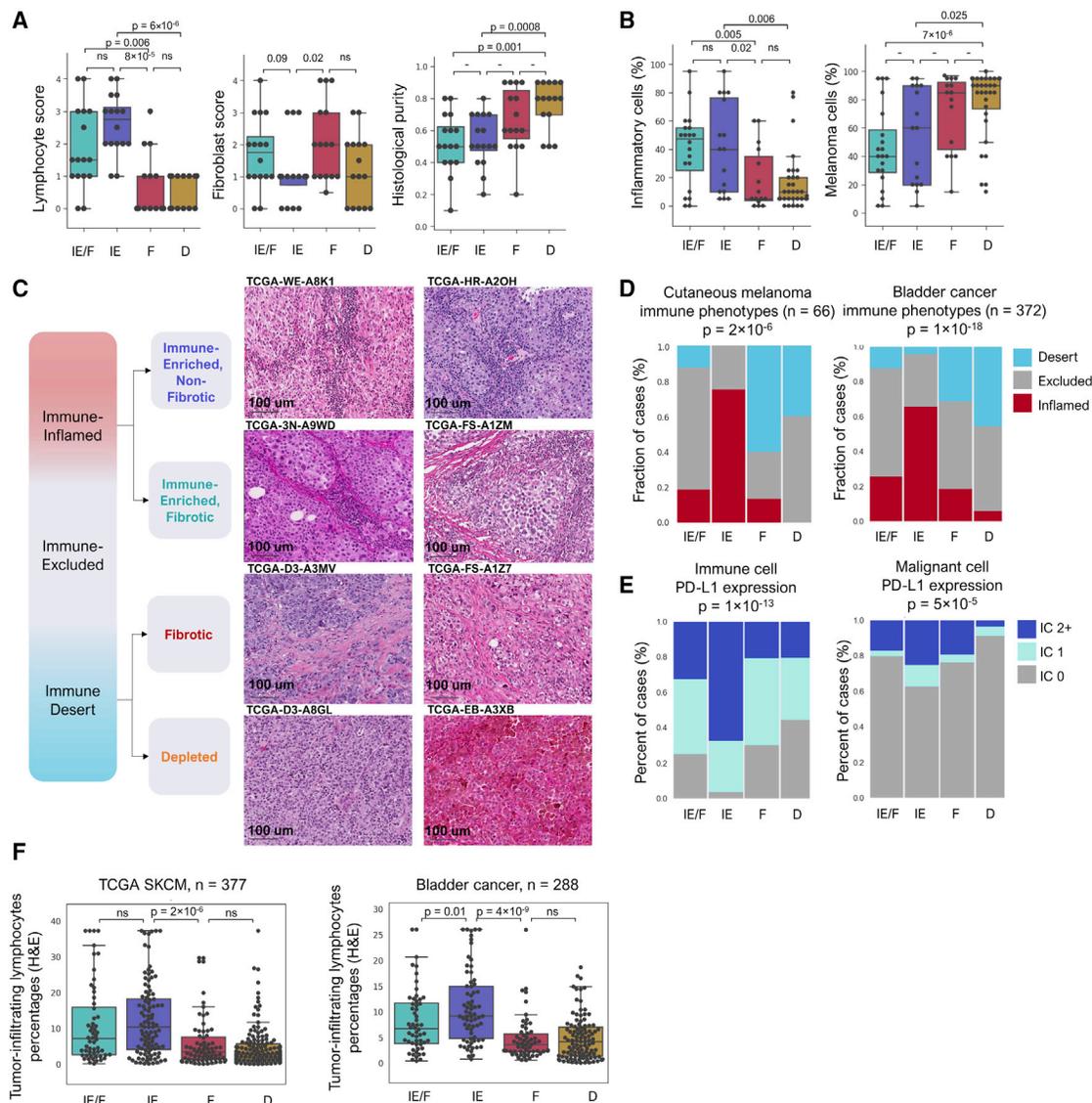


Figure 3. The four TME subtypes correlate with tumor histology in melanoma and bladder cancer

(A) Box plots depicting relative lymphocyte, fibroblast, and malignant cell percentage scores by TME subtype of randomly selected TCGA-SKCM samples (n = 10–16/TME subtype). Lymphocyte and fibroblast abundance was scored from 0 to 4 by manual examination of H&E by a pathologist.

(B) Box plots showing percentages of inflammatory and malignant cells per TME subtype of metastatic melanoma (GEO: GSE8401) as determined by histopathological review.

(C) Representative TCGA H&E histological images of the TME subtypes and their association with three major histologically defined TME phenotypes: immune inflamed, immune excluded, and immune desert. Scale bar denotes 100 μ m.

(D) Percentage of histologically defined TCGA-SKCM melanomas (n = 62) and bladder cancers per TME subtype.

(E) Percentage of bladder cancers with PD-L1 expression measured by relative IC score (PD-L1 positivity cell percentage, IC0 < 1%, IC1 = 1–5%, IC2+ > 5%) on immune cells (left) and malignant cells (right) per TME subtype.

(F) Box plots showing percentages of tumor-infiltrating lymphocytes as determined by a pathologist in the listed cancer types. In all box plots, the upper whisker indicates the maximum value or 75th percentile +1.5 IQR; the lower whisker indicates the minimum value or 25th percentile –1.5 IQR.

See also Table S5.

correlations, showing that the 29 Fges initially investigated in melanoma can be employed for TME classification across different cancer types (Figure S4D). Tumors segregated in an unsupervised manner into the same four TME subtypes as in the 8,024 TCGA tumors and the 1,995 melanomas (Figure S4E, Table S6), indicating that this expression signature-based TME

classification system can be broadly applied at the pan-cancer level.

The association between the TME subtypes and the intrinsic biology of different cancer types and viral-driven cancers was also examined as viruses induce immune escape. Subtype IE was enriched in EBV-positive gastro-esophageal cancer

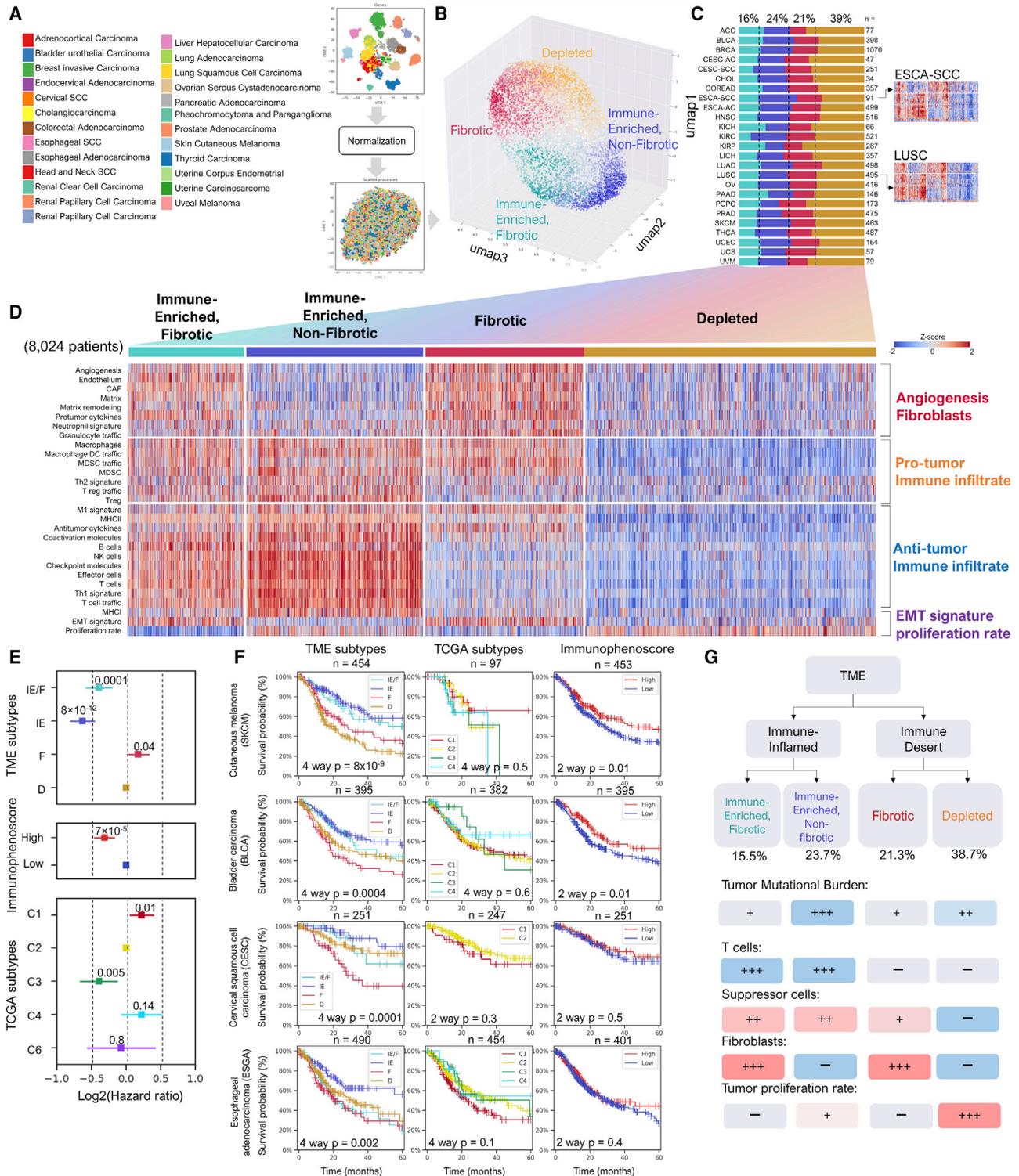


Figure 4. The four TME subtypes are conserved across a broad array of cancers

(A) tSNE analysis of the 29 Fgcs scores across TCGA tumors. The dots represent individual tumor samples, and the colors represent the datasets (cancer types) from TCGA.

(B) 3D UMAP projection of cancer patients per TME subtype (subnetworks) based on unsupervised dense clustering.

(C) Bar graphs depicting segregation of the carcinomas into the four TME subtypes with representative heatmaps of unsupervised dense clustering of the 29 Fgcs shown for esophageal cancer squamous cell carcinoma (ESCA-SCC) and lung squamous cell carcinoma (LUSC).

(D) Heatmap of 8,024 TCGA carcinomas segregated into the four TME subtypes by unsupervised dense clustering based on the intensity of the 29 Fgcs.

(legend continued on next page)

(chi-square test, $p = 7 \times 10^{-6}$) and in HPV-positive head and neck carcinomas (chi-square test, $p = 2 \times 10^{-8}$) (Figure S5A), reflecting the expected association of viral-driven cancers with an inflamed TME. Next, we analyzed the T cell repertoire (TCR) and B cell repertoire (BCR) recovered from the RNA-seq of multiple cancer types. TCR and BCR richness was significantly higher in the IE/F and IE subtypes, and T cell and B cell diversity (Shannon entropy index) was greatest in subtype IE (Figure S5B). Colorectal and gastric cancers with high microsatellite instability (MSI) have superior survival compared with tumors with low MSI due to increased immunogenicity by infiltrating lymphocytes and macrophages (Giannakis et al., 2016). Indeed, the majority of hypermutated, including MSI and/or *polymerase E* (*POLE*)-mutated, colorectal and gastric cancers, were classified as subtype IE but not IE/F, differentiating between two types of inflammation (Figure S5C). The conserved TME subtypes were assessed across the different TCGA molecular subclasses of specific adenocarcinomas, and the TCGA molecular subclasses displayed marked differences in distribution across the four TME subtypes (Figures S5D–S5L), showing the link between molecular and TME classification systems and similarities among different classification systems.

Although the four TME subtypes were identified across multiple cancers, the significance of this TME subtyping needed to be further evaluated; therefore, the prognostic value of the TME subtypes was assessed at the pan-cancer level and within individual cancer types (Figures 4E and 4F). The multi-variate Cox proportional hazard model controlling for cancer type and sex showed a strong positive correlation with survival for subtypes IE/F and IE and a negative correlation with subtype F ($p = 0.04$). The majority of cancers with subtype IE TMEs correlated with superior prognosis ($p = 8 \times 10^{-12}$, Figure 4E). The stromal subtype F was associated with inferior prognosis and shortened patient survival across various carcinomas ($p < 0.04$). The immune-depleted subtype D was also linked to inferior prognosis; however, this phenomenon appeared to be cancer type specific. For example, in cervical cancer (cervical squamous cell carcinoma [CESC]), no survival difference was observed between the subtype D tumors and immune-enriched tumors (Figures 4F and S6). Subtype IE, but not IE/F, was widely associated with better survival in comparison with the other TME subtypes analyzed within a cancer type, with indolent cancer types (prostate adenocarcinoma, pheochromocytoma, and thyroid carcinoma) excluded from the survival analysis due to $<5\%$ events at 5 years (Figures 4F, S6, and S7A). The inclusion of in-depth covariate analysis of multiple clinical parameters known to be important to the prognosis of each cancer type analyzed found that TME subtyping based on the Fges strongly correlated with survival in the majority of cancer types independent of the other cancer-specific clinical factors (Figures S7B and S7C). For example, TME subtype IE significantly associated with survival ($p = 0.001$) independently of age and sex in cutaneous melanoma (Figure S7B). In bladder cancer, TME subtypes IE, IE/F, and D

significantly correlated with survival independently of the other covariates (Figure S7B). Lastly, TME subtype IE strongly correlated with survival independently of MSI status, Epstein-Barr virus (EBV) infection, and stage ($p = 0.0016$) in esophagogastric cancer.

Ultimately, this combined analysis of the Fges patterns across various cancer types created our classification system, which divided the TME into the four subtypes (Figure 4G, Tables S4 and S6) that significantly correlated with OS and PFS in comparison with other pan-cancer classification approaches such as the immunophenoscore (Charoentong et al., 2017) and the six TCGA immune TME subtypes (Thorsson et al., 2018) (Figure 4E). For example, the immune-enriched cluster C3 identified in the TCGA classification model, without normalization, showed a generalized positive correlation with survival (Figure 4E); however, when analyzed within each cancer type separately, our TME classification model correlated more significantly with OS in melanoma, bladder cancer, cervical carcinoma, gastroesophageal adenocarcinoma (Figures 4F and S7B), and others (Figures S6, S7C–S7E).

The four conserved TME subtypes correlate with immunotherapy efficacy

Response to immunotherapy, and specifically to immune checkpoint blockade, has been previously linked with T cell infiltration (Engelhard et al., 2018; Tang et al., 2016), high tumor mutation burden (TMB) (Hellmann et al., 2019; Rizvi et al., 2015), neoantigen burden (Łuksza et al., 2017), and tumor antigenicity (Chowell et al., 2018; Zaretsky et al., 2016). To determine whether this TME classification system can be utilized as a tool to predict response to immune checkpoint blockade, the responses to immune checkpoint inhibitors were investigated across the four different TME subtypes. In two independent skin cutaneous melanoma cohorts ($n = 58$) (Nathanson et al., 2017; Snyder et al., 2014; Van Allen et al., 2015) treated with anti-CTLA-4 therapy, patients were classified into the four TME subtypes (Figure 5A), which significantly correlated with response to ipilimumab. The percentage of responders (R + long-term survivors [LS]) to anti-CTLA-4 therapy in the immune, non-fibrotic TME subtype IE was 82% in contrast to only 10% of subtype F (Figure 5B). In both cohorts, OS following anti-CTLA-4 was the longest in TME subtype IE (Figures 5C and S8A). Similar findings were observed with three independent anti-CTLA-4 naive cohorts of melanoma patients ($n = 114$) treated with anti-PD1 therapy (Figure 5D) (Gide et al., 2019; Hugo et al., 2016; Liu et al., 2019a). Overall response was significantly higher in patients with TME subtype IE (75%) compared with subtype F (10%) (Figure 5E), with prolonged PFS and OS also noted in TME subtype IE in combined and individual cohort analysis (Figures 5F, S8B, and S8C). The association of the TME subtypes with immunotherapy response was further analyzed using independent cohorts of anti-PD-L1 or -PD1-treated patients with bladder ($n = 346$ patients)

(E) Log-hazard ratios with deviation for different tumor classification systems using Cox's proportional regression hazard model for 8,042 patients with carcinomas controlled for cancer type.

(F) OS of TCGA patients with the denoted cancers according to TME subtype, TCGA subtype, and immunophenoscore.

(G) Schematic description of the features associated with the four pan-cancer TME subtypes.

See also Figures S4–S7, Table S6.

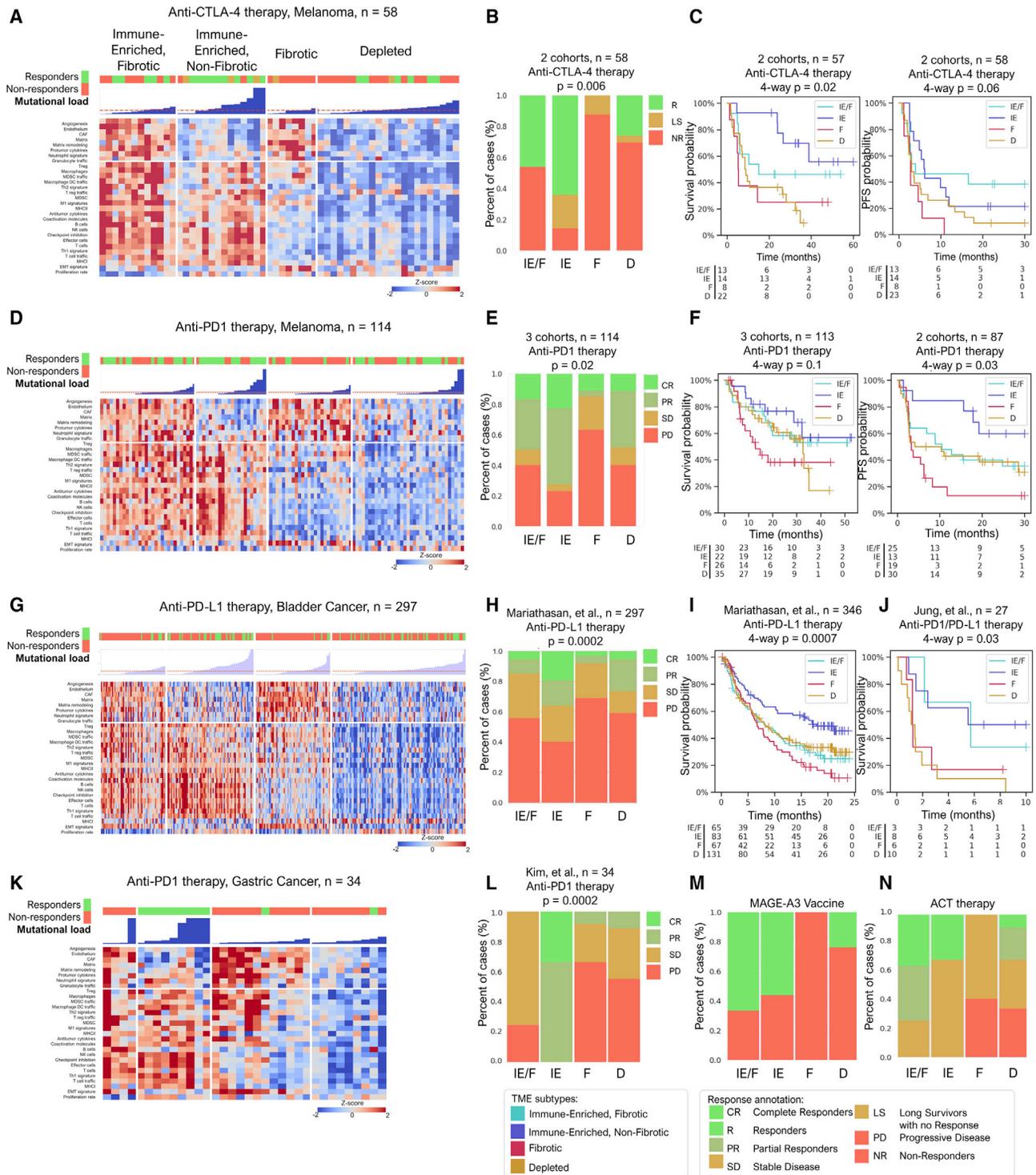


Figure 5. The four TME subtypes correlate with response to immunotherapy

(A) Heatmap of the 29 Fg genes showing TME subtype classification of melanoma biopsies collected prior to anti-CTLA-4 therapy (Van Allen/dbGaP: phs000452 and Nathanson/SRA: SRP067586, n = 58).

(B) Percentages of responders (R) and LS who did not respond to anti-CTLA-4 therapy, and non-responders (NRs) treated with anti-CTLA-4 therapy across the four TME subtypes.

(C) OS and PFS of patients with different TME subtypes treated with anti-CTLA-4 therapy.

(D) Heatmap of the 29 Fg genes for TME subtype classification across skin, non-acral melanoma biopsies collected prior to anti-PD1 therapy from three independent datasets (Liu/phs000452, Hugo/GEO: GSE78220/GEO: GSE96619, Gide/ENA: ERP105482, total n = 114).

(legend continued on next page)

(Mariathasan et al., 2018), lung cancer (n = 27) (Jung et al., 2019), and gastric cancer (n = 34) (Figures 5G–5L and S8D). Over 38% of bladder cancer patients with TME subtype IE were responders (partial responder [PR], complete responder [CR], and stable disease [SD] with PFS >6 months) (Figure 5H). In contrast, the immunosuppressive subtype F bladder cancer patients had significantly lower response rates to immunotherapy (<10% ORR; PD and SD with PFS <6 months). In bladder cancer, immunotherapy-treated patients in subtype F also displayed an inferior prognosis compared with the other TME subtypes (Figures 5I and 5J). Bladder and lung cancer patients with TME subtype IE demonstrated the longest OS (Figures 5I and 5J). Notably, in gastric cancer, TME subtype IE was only composed of responders (CR and PR) to anti-PD1 therapy (Figure 5L). Next, the ability of the TME subtypes to delineate responses to additional immune-based therapies such as therapeutic vaccination against melanoma-associated antigen A3 (anti-MAGE-A3) (Ulloa-Montoya et al., 2013) and adoptive cell therapy (Lauss et al., 2017) was tested, with responses strongly correlating with the immune-enriched TME subtypes IE/F and IE (Figures 5M, 5N, and S8E), suggesting that the TME classification system can be applied to diverse immune-based therapies as a potential biomarker of response.

High TMB has been shown to correlate with immune checkpoint blockade response (Hellmann et al., 2019; Rizvi et al., 2015); therefore, the predictive potential of the TMB in concert with the TME classification system was assessed. The effect of the TMB in predicting immunotherapy response was determined by dividing the patients belonging to anti-CTLA-4 or -PD1 studies into groups with low, medium, and high TMB (Figure S8F). In the examined melanoma cohorts, no significant association was observed between TMB and therapeutic response (Figures S8G–S8J), which contrasted with the significant correlations between TMB and response and survival observed in bladder cancer (Figures S8K and S8L). Strikingly, survival rates of IE bladder cancer patients did not differ between TMB-high and -low patients (Figure S8M). Specifically, Cox proportional hazards modeling showed a TMB-independent predictive association of TME subtype IE with survival (Figure S8K). TME subtype D bladder cancer tumors that were enriched with TMB-high tumors had a relatively high proportion of responders and increased OS (Figure S8M). These results suggest that immune-low tumors with high TMB can still benefit from immunotherapy in the absence of an immunosuppressive CAF (e.g., IE/F or F TME subtypes). The TME classification system may explain the effectiveness of immunotherapy in patients with low TMB, and these two classifica-

tion systems, TME subtyping and TMB analysis, can complement each other as immunotherapy biomarkers.

Both predictive and prognostic biomarkers and patient outcome stratification models, including expression signatures, the expression of 1–2 genes, scoring systems, as well as microenvironment and molecular classification systems, are being developed for immunotherapy across many cancers, including melanoma and bladder cancer. To determine the universal predictive and prognostic potential of TME subtyping in comparison with other methods, these potential stratification methods were analyzed in the anti-CTLA-4- and anti-PD1-treated melanoma cohorts as well as the anti-PD-L1-treated bladder cancer cohort. To conduct these analyses, previously published scoring systems or signatures were extrapolated, and an example of this extrapolation is shown in Figure S9A for the TCGA TME types (Thorsson et al., 2018) with further application to the TCGA-SKCM, which produced an area under the curve (AUC) of greater than 0.99, showing that we could accurately reproduce the TCGA classification method in our datasets. With this ability to classify datasets with these different stratification methods, these methods were applied across the immunotherapy datasets Figure S9B. Next, survival was assessed in the anti-CTLA-4- and anti-PD1-treated melanoma cohorts as well as the anti-PD-L1-treated bladder cancer cohort using TCGA TME typing (Thorsson et al., 2018), immunophenoscore (Charoentong et al., 2017), TMB, PD-L1 (Topalian et al., 2012) and interferon (IFN) γ expression (Ayers et al., 2017), and TCGA molecular subtypes (Cancer Genome Atlas Network, 2015). Significant stratification in survival in all cohorts was not observed across all three cohorts (Figure S9C), in contrast with TME subtyping, which meaningfully and significantly stratified the survival of melanoma, bladder, lung, and gastric cancer patients treated with immunotherapy (Figure 5). Indeed, the TME subtypes had the ability to more clearly delineate survival to immune checkpoint blockade compared with TMB and immunophenoscore (Charoentong et al., 2017), the six TCGA immune TME subtypes (Thorsson et al., 2018), PD-L1 (Topalian et al., 2012) and IFN γ expression (Ayers et al., 2017), and TCGA molecular subtypes for melanoma (Cancer Genome Atlas Network, 2015) and bladder cancer (Jemal et al., 2014) (Figures S9A–S9C). Specifically, TCGA subtypes did not correlate with survival in melanoma patients treated with anti-CTLA-4 ($p = 0.7$) and anti-PD1 therapy ($p = 0.2$) and in bladder cancer patients treated with anti-PD-L1 ($p = 0.5$). Immunophenoscore had a weak correlation ($p = 0.1$) with survival in anti-CTLA-4-treated melanoma patients

(E) Percentages of CR, PR, SD, and progressed disease (PD) patients treated with anti-PD1 therapy from three melanoma datasets segregated by TME subtype (total n = 114).

(F) OS and PFS of melanoma patients with different TME subtypes treated with anti-PD1 therapy.

(G) Heatmap of the 29 Fges across bladder cancer biopsies collected before anti-PD-L1 therapy (n = 297).

(H) Percentages of CR, PR, SD and PD among bladder cancer patients treated with anti-PD-L1 therapy across the four TME subtypes.

(I) OS of the bladder cancer patients treated with anti-PD-L1 therapy across the four different TME subtypes.

(J) OS of lung cancer patients treated with anti-PD-L1 therapy and segregated into the four different TME subtypes.

(K) Heatmap of the 29 Fges across gastric cancer biopsies collected before anti-PD1 therapy (n = 34).

(L) Percentages of CR, PR, SD, and PD among gastric cancer patients treated with anti-PD1 therapy across the four TME subtypes.

(M) Percentages of responders (green) and NRs (red) to MAGE-A3 vaccine across the defined TME subtypes.

(N) Percentages of CR, PR, SD, and PD melanoma patients treated with ACT across the four TME subtypes.

See also Figures S8–S10.

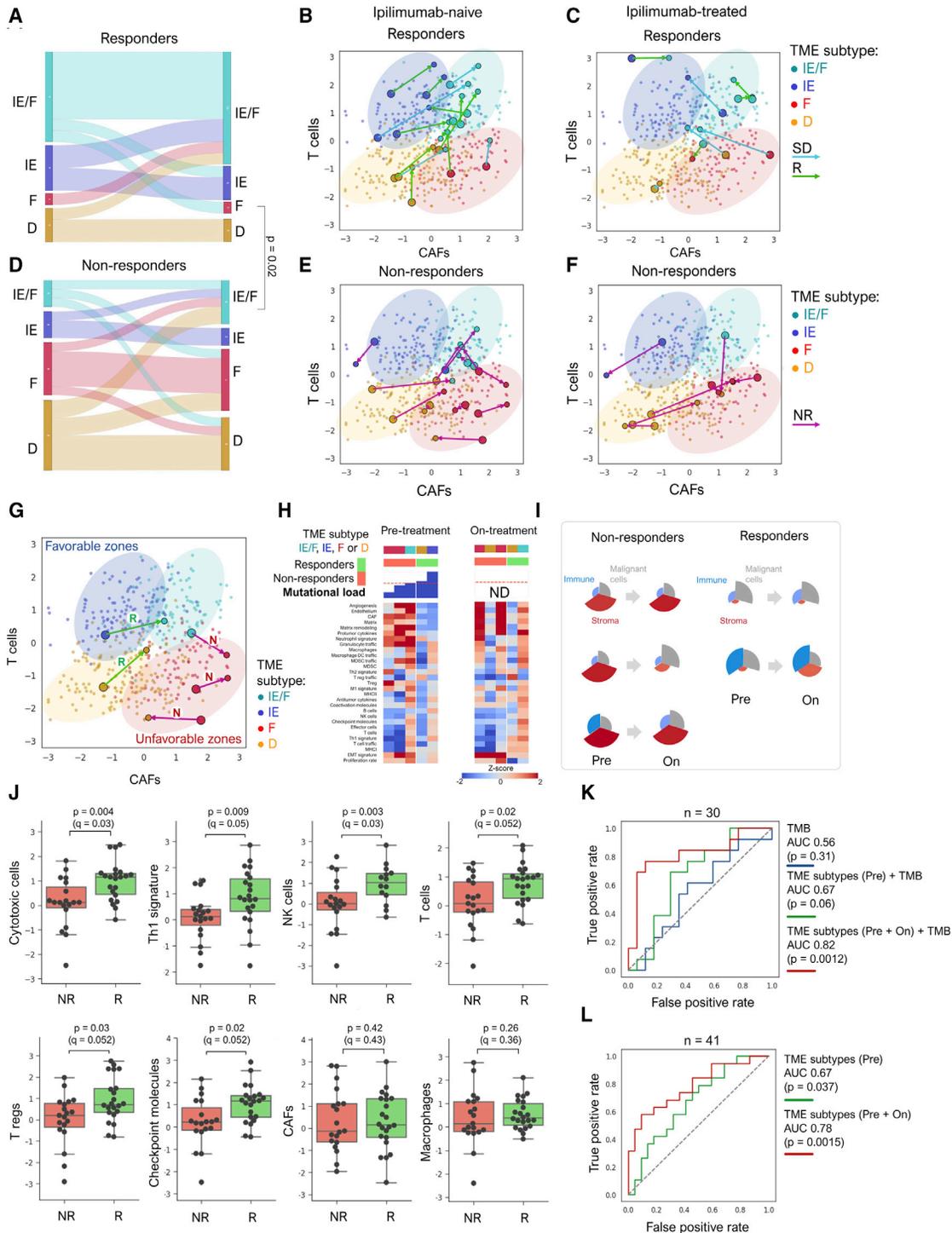


Figure 6. Dynamic changes within the tumor following treatment improves response prediction

(A) Sankey diagram showing the flow/changes of the TME subtypes of the responders (n = 16) to anti-PD1 therapy pre-treatment and on treatment. (B and C) TME changes of ipilimumab-naïve and -treated responders to anti-PD1 therapy. Patient tumor samples classified into TME subtypes before and after treatment are depicted based on the number of T cells (y axis) and CAFs (x axis). (D) Sankey diagram showing the flow/changes of the TME subtypes of the NRs (n = 20) to anti-PD1 therapy pre-treatment and on treatment. (E and F) TME changes of NRs to anti-PD1 therapy. Patient tumor samples classified into TME subtypes before and after treatment are depicted based on the number of T cells (y axis) and CAFs (x axis). (G) Dynamic changes of the tumor microenvironment of five melanoma patients treated with anti-PD1 therapy. Patient tumor samples classified into TME subtypes before and after treatment are depicted based on the number of T cells (y axis) and CAFs (x axis).

(legend continued on next page)

(Figure S9C). IFN γ expression and the TCGA molecular subtypes significantly correlated with survival in melanoma ($p = 0.04$ and 0.06 and 0.004 and 0.09 , respectively) but not bladder cancer ($p = 0.3$ and 0.5 , respectively). TMB did not correlate with survival to anti-CTLA-4 and anti-PD1 in melanoma ($p = 0.5$ and 0.4 , respectively) but TMB did correlate with survival to anti-PD-L1 therapy in bladder cancer ($p = 8 \times 10^{-5}$, Figure S9C), leading to the Cox regression analysis depicting that immune-enriched and high TMB burden (Figure S8K) are independent significant markers for response in bladder cancer. Additionally, high TMB and low TMB within the immune-enriched subtype respond equally to immunotherapy (Figure S8M). Not only were these classification methods compared, inflamed, immune-excluded, and immune-desert types (Chen and Mellman, 2017; Cristescu et al., 2018) were assessed in the anti-PD-L1-treated bladder cancer cohort, and these types did not significantly correlate with survival or response to therapy (Figure S9D). Lastly, patient stratification based on MHC1 and MHCII expression (Rodig et al., 2018; Liu et al., 2019a), CXCL9 expression (Qu et al., 2020), and the immuno-predictive score (IMPRES) (Auslander et al., 2018) was evaluated in the three cohorts (Figure S9E). Again, the expression of these genes and signatures did not significantly correlate with survival in all of the cohorts. For example, MHC1 and II expression only correlated with survival in the anti-CTLA-4 melanoma cohort ($p = 0.01$ and 0.001) but not the anti-PD1-treated melanoma cohort ($p = 0.1$ for both) or the anti-PD-L1-treated bladder cancer cohort ($p = 0.2$ and 0.3). CXCL9 expression did not significantly correlate with survival in the anti-PD1 melanoma cohort ($p = 0.08$). The IMPRES score did not delineate survival in the anti-CTLA-4 melanoma and anti-PD-L1 bladder cancer cohorts ($p = 0.1$ and 1.0 , respectively). MHC1 and MHCII expression and CXCL9 expression are markers of active inflammation and correlate with the TME subtypes (Figure S9F). Overall, several of these biomarkers correlated with response in individual cohorts, similarly to the TME subtype; however, none of these biomarkers correlated with response across all the cohorts, highlighting the importance of the development of multi-parametric biomarkers such as TME subtyping (Figures S10A and S10B). Additionally, the TME subtypes correlated with PD-L1 and CXCL9 expression and the IFN γ signature and expression; however, the TME subtypes acted as biomarkers that can be generalized across many cancer types by taking into account numerous microenvironmental factors. Overall, these comparative analyses definitively demonstrate the ability of this TME classification platform to correlate with survival and reveal its

potential as a universal biomarker across multiple cancers and immunotherapies.

Dynamic evolution of the TME revealed in response to immunotherapy

The TME directly influences the efficacy of immune checkpoint blockade; however, immunotherapies can also affect and alter the TME (Gide et al., 2019; Riaz et al., 2017). Therefore, tracking TME evolution can provide important insights into immune checkpoint blockade response. In melanoma cohorts where tumor evolution could be evaluated, we analyzed the bulk RNA-seq of pre-treatment and on-treatment biopsies (1–6 weeks post treatment initiation) collected from a cohort of melanoma patients treated with anti-PD1 immunotherapies (Figures 6A–6G). Specifically, melanoma patients were segregated into anti-CTLA-4-naive and anti-CTLA-4-progressed melanoma patients due to reported marked differences in their baseline immunosuppression profiles (Riaz et al., 2017). Linear regression analysis of the TME dynamics of anti-CTLA-4-naive melanoma patients treated with anti-PD1 showed an overall pre-response (high effect size) trend toward the IE/F or IE subtypes, and an overall anti-response trend toward the F or D subtypes. Remaining as the TME subtype F ($p = 0.0525$) or IE/F ($p = 0.053$) before and during treatment was slightly significant (Figure S11). TME evolution dynamics were specifically visualized using the T cell and CAF Fges—two major determinants of the four TME subtypes—highlighting immunotherapy-favorable and -unfavorable zones (Figures 6B–6G). Of these two groups, those who then responded to anti-PD1 therapy possessed primarily subtypes IE/F and IE (13 out of 16 responders) that remained unchanged on treatment or evolved toward the immune-enriched environments (Figures 6A–6C). This evolution was visualized by the dynamic tracks of responders on the T cells versus CAFs map arrowed toward the immune-favorable zones of IE and IE/F types based on increasing T cell signature activity. In contrast, the TMEs of the majority of non-responders to anti-PD1 therapy appeared to maintain or move toward the immune-unfavorable TME ($n = 14$ out of 20 non-responders; Figure 6D, $p = 0.02$, non-responders versus responders) and became increasingly immune depleted and fibrotic (Figures 6D–6F). These evolutionary patterns were supported by additional TME tracking of one of the melanoma patient cohorts treated with anti-PD1 immunotherapy ($n = 5$, Figures 6G–6I) (Garcia-Diaz et al., 2017). Two melanoma patients with pre-treatment TME subtype F tumors did not respond to immune checkpoint blockade and remained in the subtype F after treatment (Figure 6G). In general, tumors from melanoma patients who responded to immunotherapy evolved to the immune-enriched TME subtypes IE and IE/F, whereas the

(H) Heatmap showing the expression patterns of the 29 Fges before and 6 weeks after anti-PD1 treatment for the same patients presented in (A) across TME subtypes and with TMB for the pre-treatment biopsies.

(I) Pie charts representing the changes to the microenvironments of these five patients post-anti-PD1 therapy.

(J) Box plots showing fold changes in the T cells, CAF, immune checkpoint, and proliferation rate Fges after the treatment for ipilimumab-naive and -treated patients followed by anti-PD1 treatment. Unassociated p values and FDR-adjusted p values (q values) are denoted. In the box plots, the upper whisker indicates the maximum value or 75th percentile +1.5 IQR; the lower whisker indicates the minimum value or 25th percentile –1.5 IQR.

(K) Receiver operating characteristics of the anti-PD1 treatment prediction model based on the TMB or the addition of the change in TME subtype before and after treatment of 30 melanoma patients.

(L) Receiver operating characteristics of the anti-PD1 treatment prediction model based on the change in TME subtype before and after treatment of 30 melanoma patients. p values are assigned to the AUC calculations.

See also Figure S11.

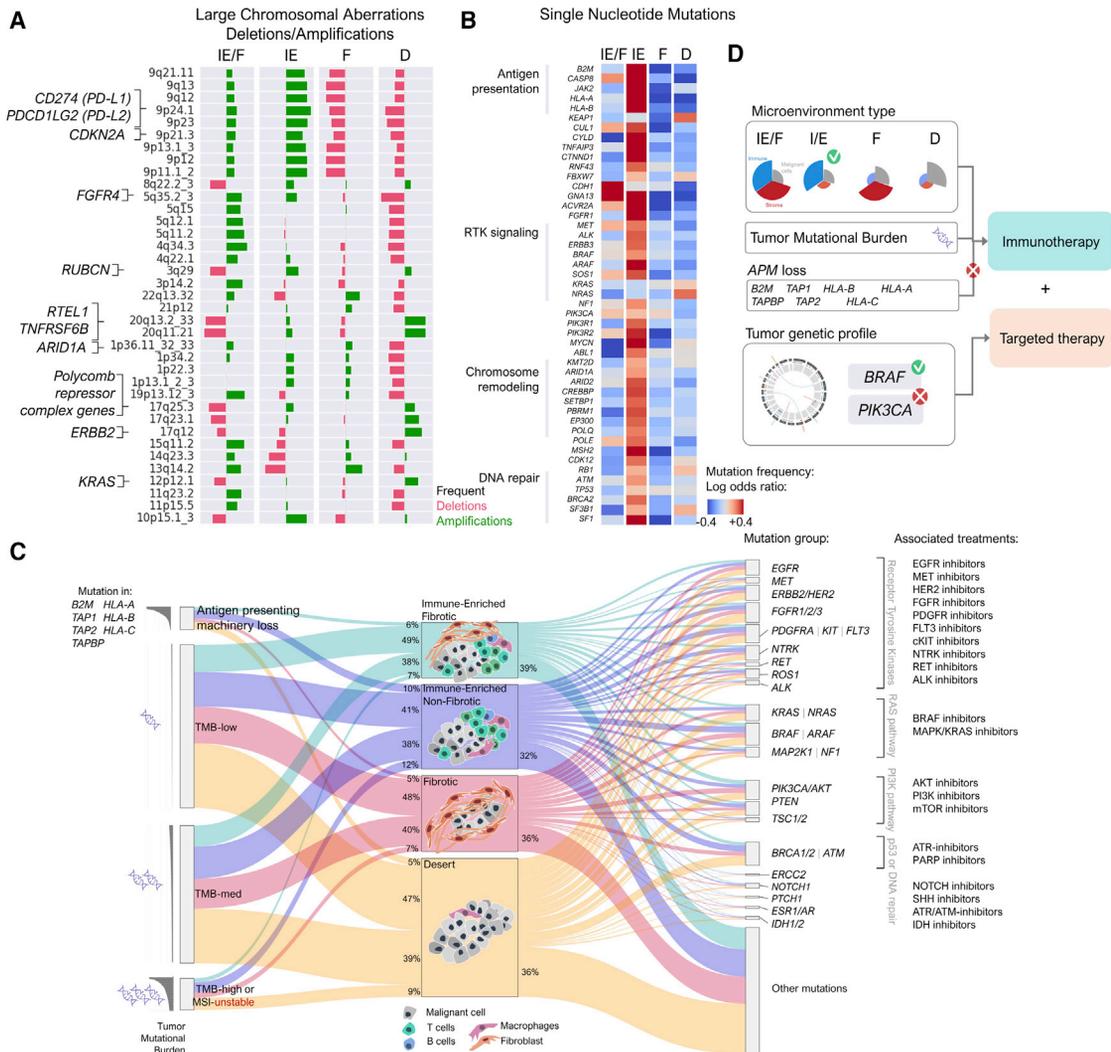


Figure 7. Genomic alterations that correlate with TME subtype for rational therapeutic design

(A) Statistically significant differences in the enrichment of copy-number variants (deletions or amplifications) presented as fold odds ratio of copy-number event in cytoband, controlled for tumor histology across the four TME subtypes (FDR < 0.01).
 (B) Heatmap showing statistically significant enrichment of mutations in tumor types displayed as fold enrichment log10 odds ratio. Significance was evaluated by the Cochran-Mantel-Haenszel chi-square test to account for tumor histology.
 (C) Sankey plot showing antigenicity and TMB (left) per TME subtype linked to mutation group (right) across TCGA patients at the pan-cancer level.
 (D) Schema for the rational selection of combinatorial anti-cancer treatment based on (1) the analysis of the TME subtype and tumor antigenicity to select appropriate immunotherapy options; and (2) TME-independent analysis of tumor genomic alterations to select appropriate targeted therapy.
 See also [Figure S12](#).

non-responders retained the immune unfavorable subtype F (Figure 6I).

The expression of T cell- and NK cell-associated Fges, Tregs, and immune checkpoint expression significantly increased in the responding tumors over time (Figure 6J, FDR-adjusted p values [q values] ranging from 0.03 to 0.052), suggesting activation of the TME and increased immune cell infiltration in response to immune checkpoint blockade. We determined the predictive value added by assessing pre- and on-treatment biopsies in addition to TMB, the common biomarker of immune checkpoint blockade response. Notably, using a logistic regression model, TMB alone was not able to predict the response to anti-PD1 immunotherapy compared with the addition of TME subtype changes before and

during treatment (AUC, 0.56; $p = 0.31$ versus 0.82; $p = 0.0012$; Figure 6K). TMB could not discern a potential responder versus non-responder (AUC = 0.56), with essentially only a 50% chance of correctly predicting the response, whereas incorporating the TME classification both before and on treatment improved the probability of correctly separating responders from non-responders with more than 80% accuracy. In an enlarged cohort (Riaz et al., 2017; Garcia-Diaz et al., 2017; Rizvi et al., 2015) with no TMB data available, TME classification utilizing pre- and on-treatment biopsies still had improved predictive potential compared with TME classification using pre-treatment biopsies alone (AUC, 0.78; $p = 0.0015$ versus 0.67; $p = 0.037$; Figure 6L), supporting the utilization of these two classification methods,

TME subtyping and TMB status, in collaboration and the addition of the information gained by on-treatment biopsies to fully develop computational systems to accurately predict immunotherapy response.

TME subtypes associate with cancer genomic alterations

Genomic alterations and oncogenic signaling within the tumor have been shown to affect anti-tumor immunity and TME activity (Rooney et al., 2015; Spranger et al., 2015); therefore, links between tumor mutations and TME classification were investigated. The patterns of multiple specific CNAs in chromosome regions encoding immunoregulatory and oncogenic genes were associated with each TME subtype (Figure 7A). For example, IE carcinomas were enriched with *CD274* (PD-L1) and *PDCD1LG2* (PD-L2) gene amplifications (9p24), suggesting that tumor-reactive lymphocytes may increase the expression of these immune checkpoints, dampening anti-tumor immunity. The anti-apoptotic gene *BCL2L1* (BCL-xL; 20q11.21), was amplified in TME subtype IE carcinomas, in contrast to its common deletion in subtype F carcinomas (Figure 7A). TME subtype IE carcinomas had a relatively high frequency of alterations in genes regulating DNA mismatch repair such as *MSH2* and histone modifications such as *CREBBP*, *KMT2A*, and *PBRM1* (Miao et al., 2018; Pan et al., 2018). Additional immune escape mechanisms observed in TME subtype IE carcinomas may be linked to mutations in genes associated with antigen-presentation machinery or the interferon-inducible pathway (*CASP8*) (Figure 7B). TME subtype D carcinoma was enriched in *CDK6* (8p11) and *BCL2L1* gene amplifications, associated with increased cellular proliferation (Figure 7A) as well as *RTEL1* (20q.13), a regulator of telomere elongation. Polycomb Repressive Complex-associated genes (PRC.17q25) (Pirrotta, 1998; Grossniklaus and Paro, 2014) were frequently lost in TME subtype IE/F but amplified in TME subtype D, suggesting different modes of transcriptional regulation of this TME subtype via chromatin remodeling. TME subtype D carcinomas were also enriched with *KEAP1* gene mutations, suggesting altered regulation of oxidative stress, chemoresistance, and enhanced tumor cell growth, potentially leading to the decreased presence of immune cells by the NRF2-KEAP1 pathway (Wu et al., 2019).

The genomic alterations associated with each TME subtype were further dissected across the 8,024 TCGA melanomas and adenocarcinomas, primarily focusing on different levels of antigenicity measured by TMB (left, Figure 7C) as well as major established clinically actionable and targetable mutations (Thorsson et al., 2018; Sanchez-Vega et al., 2018) (right, Figure 7C). Moreover, patients possessing alterations in genes resulting in loss of MHC class I machinery were assessed as a distinct group because these patients would not respond to the immunotherapy due to the lack of antigen presentation (Montesio et al., 2021) (Figure 7C). Approximately 4%–10% of patients per subtype had these mutations, with the highest percentage in the IE subtype and the lowest in the F subtype (Figures 7B and 7C), which correlates with previously observed immunoreactive and suppressive properties of those subtypes. Ultimately, this analysis shows that transcriptomic-based TME classification coupled with genomics analysis can be exploited for combinatorial therapeutic design (Figure 7D).

At the pan-cancer level, no clear enrichment of alterations was observed in any of the TME subtypes, potentially due to imbalances in the TCGA cohorts and different frequencies of the alterations per cancer type (Figures 7C and S12A). Interestingly, actionable alterations were grouped into different genomic classes by supervised clustering of the tumor mutational profile (Figures S12A and 7C). The majority of carcinomas belonged to only one genomic group, forming a diagonal of enriched potentially actionable alterations (Figure S12A), indicating which targeted pathway inhibitors would be potentially active against the tumors regardless of the TME cluster. Genomic alterations and copy-number variations in genes associated with oncogenic signaling, DNA repair, and cell cycle regulation (Thorsson et al., 2018) per cancer type were also examined (Figure S12B), revealing potential associations of mutational status of clinically actionable genes with TME. Interestingly, *EGFR* alterations in head and neck ($p = 2 \times 10^{-6}$) and lung cancers ($p = 0.006$) (Figures S12B and S12C) and *KRAS/NRAS* in colorectal ($p = 2 \times 10^{-6}$) and renal adenocarcinomas ($p = 0.02$) were associated with the F and D TME subtypes, whereas *PIK3CA* alterations in lung squamous cell carcinoma ($p = 0.03$) and stomach adenocarcinoma ($p = 0.008$) were associated with the inflamed IE subtype. Breast cancer tumors with *PTEN* loss also showed a more inflamed IE subtype ($p = 0.004$). *BRCA1/2* loss in endometrial carcinoma was associated with TME subtype IE ($p = 0.01$). This analysis provides an overview of clinically actionable mutational associations and TME subtypes, complementing each other in potential therapeutic decision making for cancer types.

A visual tool based on integrated transcriptomics and genomics for personalized therapy selection

To visualize the integrated analysis of genomic alterations with gene expression patterns in both malignant cells and the TME for each tumor sample, we generated a planetary schema termed Molecular Functional (MF) Portrait (Figure 8A). The MF Portrait was separately generated for each tumor sample based on its unique characteristics, ultimately representing a personalized tumor map. The entire portrait includes qualitative and quantitative descriptions of modules built based on our Fges (Figure 1B). The TME gene expression patterns were generated as modules associated with endothelial, stromal, or immune cells, which were further categorized as either pro- or anti-tumor. Each MF Portrait module represents the scores of the 29 Fges, with the size of each module corresponding to the intensity of the normalized ssGSEA score, and the colors denoting pro- (red) or anti-cancer activity (blue) (Figure 8A). The relative percentage of the malignant cell and TME compartments was calculated based on the purity of each tumor sample using WES (Figure 8A). In the MF Portrait, potentially targetable genomic alterations in the cancer cells related to different processes such as cell proliferation, oncogenic signaling, EMT, angiogenesis, and anti-tumor immunity are listed and prioritized. For example, in the MF Portrait of a melanoma tumor, clinically actionable mutations in only *BRAF* and *BRCA2* were found out of the approximately 50 prioritized genes known to be involved in melanoma tumorigenicity and progression (Figure 8A, red letters). Also, the TME subtypes were integrated into the MF Portrait visual tool to aid in the rational design of therapeutic strategies relevant to each of the four TME subtypes (Figure 8A). For

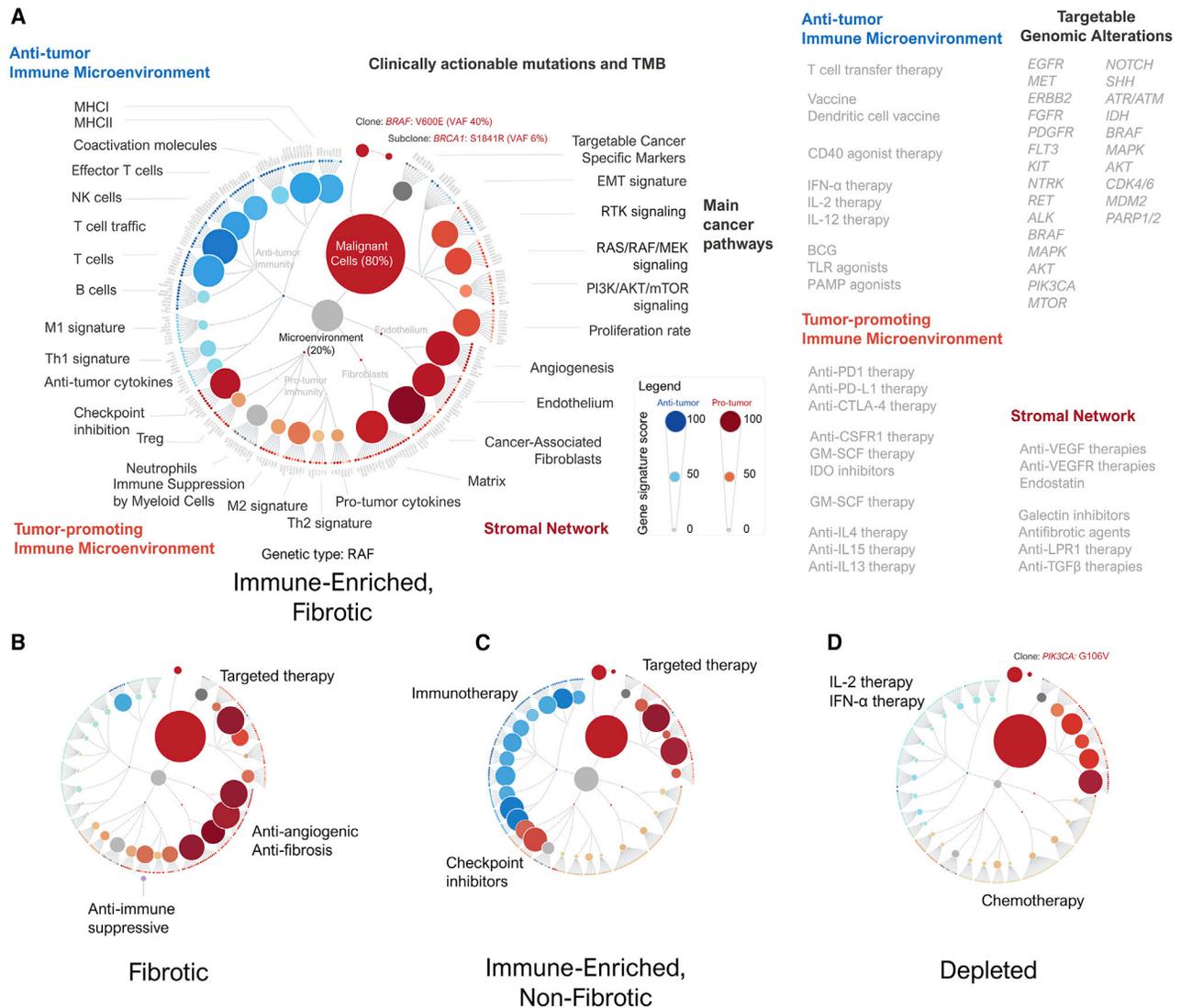


Figure 8. A visual tool for rational therapeutic design in oncology

(A) MF Portrait of a melanoma tumor with TME subtype IE/F. Potential targetable genes (around the circle), signaling pathways, and cellular processes related to each of 29 Fges. Two potentially actionable mutations were identified as listed in red: *BRAF* V600E and *BRCA1* S1841R. On the right is a list of potential therapies for each tumor subcategory.

(B) MF Portrait of a melanoma tumor with TME subtype F.

(C) Melanoma MF Portrait of TME subtype IE.

(D) MF Portrait of a melanoma tumor with TME subtype D.

example, the analyzed melanoma tumor shown in Figure 8A was classified as the IE/F TME subtype, suggesting that immune checkpoint blockade may promote anti-tumor immunity in this immune-rich tumor in combination with inhibition of immunosuppressive stromal signaling. To support rational design of combinatorial therapeutic strategies and patient selection to clinical trials, each analyzed tumor is assigned a TME subtype using transcriptomics and then integrated with the genomics analyses as shown in Figures 8B–8D, providing a tool (<http://science.bostongene.com/tumor-portrait/>) with potential clinical utility for clinicians making treatment decisions for a broad array of cancer patients.

DISCUSSION

Major advances and breakthroughs in the field of precision medicine have relied solely on genomic analyses. Nevertheless, the role and clinical impact of transcriptomic analysis have recently emerged (Cieslik and Chinnaiyan, 2018; Rodon et al., 2019; Sailer et al., 2019; Borad et al., 2016; Vaske et al., 2019) as comprehensive analysis of RNA transcripts can also significantly contribute to tissue deconvolution, detailing BCR/TCR rearrangements, and neoantigen identification. Additionally, integrated genomic and transcriptomic analysis outperforms single-omics analysis (Rodon et al., 2019; Robinson et al., 2017; Beaubier et al., 2019).

Although the importance of the TME in various cancers is well established, comprehensive analyses based on combined genomic and transcriptomic data of the tumor and its TME remain lacking. Clearly understanding the cancer type- and treatment response-specific variations in the TME may elucidate the mechanisms underlying therapeutic resistance to both targeted therapies and immunotherapies as well as identify novel therapeutic strategies. Hence, we developed a multi-omics and robust analytical platform to classify, reconstruct, and visualize the entire tumor composition.

We defined four distinct TME subtypes utilizing unsupervised analysis of approximately 2,000 melanoma patients. Similar to prior reports, we found our subtypes separated into clusters defined by unique immune cell compositions (Petitprez et al., 2020). However, based on our curated list of Fges, we identified TME subtypes defined by unique stromal, vascular, and cytokine expression patterns. These four TME subtypes were conserved across at least 20 additional cancers in over 8,500 tumor samples, revealing the intrinsic connection of the biological processes described by the Fges of each cancer tissue. Interestingly, the conservation of these TME subtypes in myriad cancers shows that the biological processes, stroma, and immunological activity of the TME in different cancers are markedly similar. Although cancer cells are unique, immune relationships appeared to be similar across all the patients, even those with different cancer types, enabling TME classification into four simple TME subtypes. Importantly, the TME subtypes significantly correlated with patient OS or PFS and had higher correlation rates compared with other pan-cancer classification approaches such as immunophenoscore (Charoentong et al., 2017), the six TCGA immune TME subtypes (Thorsson et al., 2018), PD-L1 (Topalian et al., 2012) and IFN γ expression (Ayers et al., 2017), TCGA molecular subtypes for melanoma (Cancer Genome Atlas Network, 2015) and bladder cancer (Jemal et al., 2014), CXCL9 expression (Qu et al., 2020), MHCII and II expression (Rodig et al., 2018; Liu et al., 2019a), and the IMPRES score (Auslander et al., 2018). Notably, our TME classification was the only platform that consistently and significantly correlated with survival after immune checkpoint blockade in both melanoma and bladder cancer. The immune subtypes described by the TCGA consortium were obtained without tissue normalization (Thorsson et al., 2018), revealing differences between the biology of cancer types rather than conserved TME subtypes, leading to the classification of greater than 90% of renal and thyroid carcinomas as the inflammatory phenotype and 80% of colorectal cancers as the wound-healing subtype. Normalization of the RNA-seq expression values enabled pan-cancer TME classification that can easily be applied to all cancer patients.

The observed TME subtypes share multiple similarities with the clusters identified in other studies, reflecting or expanding those patterns. The stromal component refines the previously classified immune or T cell-enriched clusters in melanoma or renal cancer (Cancer Genome Atlas Network, 2015; Şenbaobaoğlu et al., 2016) into the IE/F and IE subtypes, which are strikingly different by immunosuppressive profile, mutational burden, and tumor genetics. Our conserved TME subtypes were remarkably similar to the recent clusters identified in sarcoma (Petitprez et al., 2020), where type C showed similar patterns to our fibrotic

subtype and type E displayed similarities to our immune-enriched, non-fibrotic subtype. This tissue-independent TME clustering transforms the ideas of the cancer immunogram (Blank et al., 2016) or cancer-immune set points (Chen and Mellman, 2017), which suggest numerous possible options to a limited number of well-conserved TME subtypes.

Although immune checkpoint inhibition has revolutionized cancer care, durable responses are still observed only in a minority of patients, sometimes at the cost of severe toxicities (Wolchok et al., 2017). Therefore, the *a priori* identification of responders would improve clinical outcome and is critically needed. Here, we show that TME classification can further stratify patient responses to immunotherapy, with melanoma patients possessing F and D TMEs experiencing the worst outcomes to immune checkpoint blockade. High TMB and immune-favorable TME type IE led to the prediction of patients who would benefit the most from immunotherapy. TME classification before and on treatment led to significantly better prediction of response to immunotherapy compared with TMB alone. Interestingly, we clearly separated inflammation into good and bad inflammation during TME classification. Specifically, although IE/F was enriched with active immune cells, no clear benefits were observed in response to immunotherapy, and subtype F, enriched with fibroblasts, was the most suppressive TME. Taken together, these findings suggest that the basis of immune suppression lies not with the immune cell activity but instead with the presence and activity of stromal cells, indicating that simultaneous stromal signaling suppression (e.g., anti-TGF- β antibody) with immune checkpoint blockade may be a beneficial therapeutic strategy for cancer patients with fibrotic TMEs.

Although we identified genetic aberrations associated with the different TME clusters, key tumor driver mutations observed in melanoma and in several other cancers were not consistently associated with the specific TME subtypes at the pan-cancer level, although clinically actionable mutations were associated with the different TME subtypes in specific cancers such as *EGFR* mutations in head and neck and lung cancers. Nevertheless, this overall limited association between clinically actionable genomic alterations and the TME subtypes at the pan-cancer level supports future strategies integrating multi-platform analysis of the TME and tumor genomics.

Despite genomics and transcriptomics becoming more accessible and widespread over the last decade, real-time and comprehensive analyses that can be applied in a clinical setting are still lacking, primarily due to the absence of an analytical infrastructure and easy-to-use tools for clinical decision support (Freimuth et al., 2017). To address this need, we developed an intuitive visual tool that integrates multi-omics datasets and generates a comprehensive portrait of each tumor, based on its unique composition of genomic alterations, gene signature expression patterns, and cellular content, providing a global summary of all the potentially targetable alterations and mechanisms that characterize each tumor, leading to the rational selection of anti-cancer treatments. Continued development of such multi-compartment analytic strategies, combined with intuitive visualization tools and the addition of clinical trials designed to demonstrate the benefits of targeting cancer transcriptomics, will further improve the implementation and use of advanced personalized therapeutic strategies across the oncology field.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Tissue samples
- METHOD DETAILS
 - Tumor datasets
 - TCGA molecular and clinical data
 - Bioinformatics
 - Development of the gene expression signatures (Fges)
 - Statistical analysis of alterations and CNV
 - Data visualization
- QUANTIFICATION AND STATISTICAL ANALYSIS
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.ccell.2021.04.014>.

ACKNOWLEDGMENTS

We thank Yuriy Frolov for his contribution to the visual tool. We thank Linghua Wang (University of Texas MD Anderson Cancer Center) and Leandro Cerchietti (Weill Cornell Medicine) for their insightful comments on this manuscript.

AUTHOR CONTRIBUTIONS

N.F. and R.A. conceived and jointly supervised the study. N.F., R.A., A.B., and N.K. conceived of the experiments. A.B., N.K., and O.I. created and validated the gene expression signatures. A.B., N.K., V.S., A.G., and F.F. conducted the data analyses and raw sequencing data processing. O.G. performed histological assessment of the tissue slides. N.O., and I.K. created the Web-based MF Portrait visual tool. A.B., N.K., and K.N. contributed to figure and table generation for the manuscript. A.B., N.K., K.N., N.A., M.T., R.A., and N.F. wrote and revised the manuscript and prepared the figures.

DECLARATION OF INTERESTS

N.F. is the Chief Medical Officer of BostonGene, Corp., and a professor at the University of Texas MD Anderson Cancer Center. This research was funded by BostonGene, Corp., and all authors were employees thereof at the time the study was performed. The authors declare no other competing financial interests. A.B., F.F., N.K., and R.A. are inventors on patents related to the TME classification platform and the visual tool.

Received: May 20, 2020

Revised: September 14, 2020

Accepted: April 23, 2021

Published: May 20, 2021

REFERENCES

Van Allen, E.M., Miao, D., Schilling, B., Shukla, S.A., Blank, C., Zimmer, L., Sucker, A., Hillen, U., Foppen, M.H.G., Goldinger, S.M., et al. (2015). Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* 350, 207–211.

Aran, D., Sirota, M., and Butte, A.J. (2015). Systematic pan-cancer analysis of tumour purity. *Nat. Commun.* 6, 8971.

Aran, D., Hu, Z., and Butte, A.J. (2017). xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 18, 220.

Augustine, C.K., Jung, S.-H., Sohn, I., Yoo, J.S., Yoshimoto, Y., Olson, J.A., Jr., Friedman, H.S., Ali-Osman, F., and Tyler, D.S. (2010). Gene expression signatures as a guide to treatment strategies for in-transit metastatic melanoma. *Mol. Cancer Ther.* 9, 779–790.

Auslander, N., Zhang, G., Lee, J.S., Frederick, D.T., Miao, B., Moll, T., Tian, T., Wei, Z., Madan, S., Sullivan, R.J., et al. (2018). Robust prediction of response to immune checkpoint blockade therapy in metastatic melanoma. *Nat. Med.* 24, 1545–1549.

Ayers, M., Lunceford, J., Nebozhyn, M., Murphy, E., Loboda, A., Kaufman, D.R., Albright, A., Cheng, J.D., Peter Kang, S., Shankaran, V., et al. (2017). IFN- γ -related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Invest.* 127, 2930–2940.

Badal, B., Solovyov, A., Di Cecilia, S., Chan, J.M., Chang, L.-W., Iqbal, R., Aydin, I.T., Rajan, G.S., Chen, C., Abbate, F., et al. (2017). Transcriptional dissection of melanoma identifies a high-risk subtype underlying TP53 family genes and epigenome deregulation. *JCI Insight* 2, e92102.

Bailey, M.H., Tokheim, C., Porta-Pardo, E., Sengupta, S., Bertrand, D., Weerasinghe, A., Colaprico, A., Wendl, M.C., Kim, J., Reardon, B., et al. (2018). Comprehensive characterization of cancer driver genes and mutations. *Cell* 174, 1034–1035.

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., et al. (2013). NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* 41, D991–D995.

Beaubier, N., Bontrager, M., Huether, R., Igartua, C., Lau, D., Tell, R., Bobe, A.M., Bush, S., Chang, A.L., Hoskinson, D.C., et al. (2019). Integrated genomic profiling expands clinical options for patients with cancer. *Nat. Biotechnol.* 37, 1351–1360.

Becht, E., Giraldo, N.A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., Selves, J., Laurent-Puig, P., Sautès-Fridman, C., Fridman, W.H., et al. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* 17, 218.

Bergmann, E.A., Chen, B.-J., Arora, K., Vacic, V., and Zody, M.C. (2016). Conpair: concordance and contamination estimator for matched tumour-normal pairs. *Bioinformatics* 32, 3196–3198.

Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenaus, A.C., Angell, H., Fredriksen, T., Lafontaine, L., Berger, A., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 39, 782–795.

Blank, C.U., Haanen, J.B., Ribas, A., and Schumacher, T.N. (2016). CANCER IMMUNOLOGY. The “cancer immunogram. *Science* 352, 658–660.

Blondel, V.D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech. Theor. Exp.* 2008, P10008.

Bogunovic, D., O’Neill, D.W., Belitskaya-Levy, I., Vacic, V., Yu, Y.-L., Adams, S., Darvishian, F., Berman, R., Shapiro, R., Pavlick, A.C., et al. (2009). Immune profile and mitotic index of metastatic melanoma lesions enhance clinical staging in predicting patient survival. *Proc. Natl. Acad. Sci. U S A* 106, 20429–20434.

Bonneville, R., Krook, M.A., Kautto, E.A., Miya, J., Wing, M.R., Chen, H.-Z., Reeser, J.W., Yu, L., and Roychowdhury, S. (2017). Landscape of microsatellite instability across 39 cancer types. *JCO Precis. Oncol.* 2017, PO.17.00073.

Borad, M.J., Egan, J.B., Condjella, R.M., Liang, W.S., Fonseca, R., Ritacca, N.R., McCullough, A.E., Barrett, M.T., Hunt, K.S., Champion, M.D., et al. (2016). Clinical implementation of integrated genomic profiling in patients with advanced cancers. *Sci. Rep.* 6, 25.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 34, 525–527.

Budden, T., Davey, R.J., Vilain, R.E., Ashton, K.A., Braye, S.G., Beveridge, N.J., and Bowden, N.A. (2016). Repair of UVB-induced DNA damage is

- reduced in melanoma due to low XPC and global genome repair. *Oncotarget* 7, 60940–60953.
- Cancer Genome Atlas Network (2012). Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487, 330–337.
- Cancer Genome Atlas Network (2015). Genomic classification of cutaneous melanoma. *Cell* 161, 1681–1696.
- Cancer Genome Atlas Research Network (2014). Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* 513, 202–209.
- Cancer Genome Atlas Research Network, Analysis Working Group, Asan University, BC Cancer Agency, Brigham and Women's Hospital, Broad Institute, Brown University, Case Western Reserve University, Dana-Farber Cancer Institute, Duke University, Greater Poland Cancer Centre, et al. (2017). Integrated genomic characterization of oesophageal carcinoma. *Nature* 541, 169–175.
- Castellanos, E., Gel, B., Rosas, I., Tornero, E., Santin, S., Pluvinet, R., Velasco, J., Sumoy, L., Del Valle, J., Perucho, M., et al. (2017). A comprehensive custom panel design for routine hereditary cancer testing: preserving control, improving diagnostics and revealing a complex variation landscape. *Sci. Rep.* 7, 39348.
- Chakravarthy, A., Khan, L., Bensler, N.P., Bose, P., and De Carvalho, D.D. (2018). TGF- β -associated extracellular matrix genes link cancer-associated fibroblasts to immune evasion and immunotherapy failure. *Nat. Commun.* 9, 4692.
- Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., Hackl, H., and Trajanoski, Z. (2017). Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep* 18, 248–262.
- Chen, D.S., and Mellman, I. (2013). Oncology meets immunology: the cancer-immunity cycle. *Immunity* 39, 1–10.
- Chen, D.S., and Mellman, I. (2017). Elements of cancer immunity and the cancer-immune set point. *Nature* 541, 321–330.
- Chowell, D., Morris, L.G.T., Grigg, C.M., Weber, J.K., Samstein, R.M., Makarov, V., Kuo, F., Kendall, S.M., Requena, D., Riaz, N., et al. (2018). Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science* 359, 582–587.
- Cieslik, M., and Chinnaiyan, A.M. (2018). Cancer transcriptome profiling at the juncture of clinical translation. *Nat. Rev. Genet.* 19, 93–109.
- Cirenajwis, H., Ekedahl, H., Lauss, M., Harbst, K., Carneiro, A., Enoqsson, J., Rosengren, F., Werner-Hartman, L., Törngren, T., Kvist, A., et al. (2015). Molecular stratification of metastatic melanoma using gene expression profiling: prediction of survival outcome and benefit from molecular targeted therapy. *Oncotarget* 6, 12297–12309.
- Consortium, T.E.P.; The ENCODE Project Consortium (2004). The ENCODE (ENCyclopedia of DNA elements) project. *Science* 306, 636–640.
- Cristescu, R., Mogg, R., Ayers, M., Albright, A., Murphy, E., Yearley, J., Sher, X., Liu, X.Q., Lu, H., Nebozhyn, M., et al. (2018). Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science* 362, eaar3593.
- Cuevas, L.M., and Daud, A.I. (2018). Immunotherapy for melanoma. *Semin. Cutan. Med. Surg.* 37, 127–131.
- Davidson-Pilon, C., Kalderstam, J., Kuhn, B., Zivich, P., Fiore-Gartland, A., Moneda, L., Parij, A., Stark, K., Anton, S., Besson, L., et al. (2018). CamDavidsonPilon/lifelines, p. v0.14.6.
- Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30, 207–210.
- El-Deiry, W.S., Goldberg, R.M., Lenz, H.-J., Shields, A.F., Gibney, G.T., Tan, A.R., Brown, J., Eisenberg, B., Heath, E.I., Phuphanich, S., et al. (2019). The current state of molecular testing in the treatment of patients with solid tumors, 2019. *CA Cancer J. Clin.* 69, 305–343.
- Elliott, K., Bailey, M.H., Saksena, G., Covington, K.R., Kandoth, C., Stewart, C., Hess, J., Ma, S., Chiotti, K.E., McLellan, M., et al. (2018). Scalable open science approach for mutation calling of tumor exomes using multiple genomic pipelines. *Cell Syst.* 6, 271–281.e7.
- Engelhard, V.H., Rodriguez, A.B., Mauldin, I.S., Woods, A.N., Peske, J.D., and Slingluff, C.L., Jr. (2018). Immune cell infiltration and tertiary lymphoid structures as determinants of antitumor immunity. *J. Immunol.* 200, 432–442.
- Erdogan, B., and Webb, D.J. (2017). Cancer-associated fibroblasts modulate growth factor signaling and extracellular matrix remodeling to regulate tumor metastasis. *Biochem. Soc. Trans.* 45, 229–236.
- Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048.
- FANTOM Consortium and the RIKEN PMI and CLST (DGT), Forrest, A.R.R., Kawaji, H., Rehli, M., Baillie, J.K., de Hoon, M.J.L., Haberle, V., Lassmann, T., Kulakovskiy, I.V., Lizio, M., et al. (2014). A promoter-level mammalian expression atlas. *Nature* 507, 462–470.
- Favero, F., Joshi, T., Marquard, A.M., Birkbak, N.J., Krzystanek, M., Li, Q., Szallasi, Z., and Eklund, A.C. (2015). Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann Oncol* 26, 64–70.
- Frankish, A., Diekhans, M., Ferreira, A.M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* 47, D766–D773.
- Freimuth, R.R., Formea, C.M., Hoffman, J.M., Matey, E., Peterson, J.F., and Boyce, R.D. (2017). Implementing genomic clinical decision support for drug-based precision medicine. *CPT Pharmacometrics Syst. Pharmacol.* 6, 153–155.
- Fridman, W.H., Pagès, F., Sautès-Fridman, C., and Galon, J. (2012). The immune contexture in human tumours: impact on clinical outcome. *Nat. Rev. Cancer* 12, 298–306.
- Galon, J., and Bruni, D. (2019). Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat. Rev. Drug Discov.* 18, 197–218.
- Garcia-Diaz, A., Shin, D.S., Moreno, B.H., Saco, J., Escuin-Ordinas, H., Rodriguez, G.A., Zaretsky, J.M., Sun, L., Hugo, W., Wang, X., et al. (2017). Interferon receptor signaling pathways regulating PD-L1 and PD-L2 expression. *Cell Rep.* 19, 1189–1201.
- Gautier, L., Cope, L., Bolstad, B.M., and Irizarry, R.A. (2004). affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 20, 307–315.
- Giannakis, M., Mu, X.J., Shukla, S.A., Qian, Z.R., Cohen, O., Nishihara, R., Bahl, S., Cao, Y., Amin-Mansour, A., Yamauchi, M., et al. (2016). Genomic correlates of immune-cell infiltrates in colorectal carcinoma. *Cell Rep.* 17, 1206.
- Gide, T.N., Quek, C., Menzies, A.M., Tasker, A.T., Shang, P., Holst, J., Madore, J., Lim, S.Y., Velickovic, R., Wongchenko, M., et al. (2019). Distinct immune cell populations define response to anti-PD-1 monotherapy and anti-PD-1/anti-CTLA-4 combined therapy. *Cancer Cell* 35, 238–255.e6.
- Goldman, M.J., Craft, B., Hastie, M., Repečka, K., McDade, F., Kamath, A., Banerjee, A., Luo, Y., Rogers, D., Brooks, A.N., et al. (2020). Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* 38, 675–678.
- Grossniklaus, U., and Paro, R. (2014). Transcriptional silencing by polycomb-group proteins. *Cold Spring Harb. Perspect. Biol.* 6, a019331.
- Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 14, 7.
- Hao, H., Xiao, D., Pan, J., Qu, J., Egger, M., Waigel, S., Sanders, M.A.G., Zacharias, W., Rai, S.N., and McMasters, K.M. (2017). Sentinel lymph node genes to predict prognosis in node-positive melanoma patients. *Ann. Surg. Oncol.* 24, 108–116.
- Hayward, N.K., Wilmott, J.S., Waddell, N., Johansson, P.A., Field, M.A., Nones, K., Patch, A.-M., Kakavand, H., Alexandrov, L.B., Burke, H., et al. (2017). Whole-genome landscapes of major melanoma subtypes. *Nature* 545, 175–180.
- Hegde, P.S., Karanikas, V., and Evers, S. (2016). The where, the when, and the how of immune monitoring for cancer immunotherapies in the era of checkpoint inhibition. *Clin. Cancer Res.* 22, 1865–1874.
- Hellmann, M.D., Callahan, M.K., Awad, M.M., Calvo, E., Ascierto, P.A., Atmaca, A., Rizvi, N.A., Hirsch, F.R., Selvaggi, G., Szustakowski, J.D., et al.

- (2019). Tumor mutational burden and efficacy of nivolumab monotherapy and in combination with ipilimumab in small-cell lung cancer. *Cancer Cell* 35, 329.
- Herbst, R.S., Soria, J.-C., Kowanetz, M., Fine, G.D., Hamid, O., Gordon, M.S., Sosman, J.A., McDermott, D.F., Powderly, J.D., Gettinger, S.N., et al. (2014). Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* 515, 563–567.
- Hugo, W., Zaretsky, J.M., Sun, L., Song, C., Moreno, B.H., Hu-Lieskovan, S., Berent-Maoz, B., Pang, J., Chmielowski, B., Cherry, G., et al. (2016). Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* 165, 35–44.
- Jayawardana, K., Schramm, S.-J., Haydu, L., Thompson, J.F., Scolyer, R.A., Mann, G.J., Müller, S., and Yang, J.Y.H. (2015). Determination of prognosis in metastatic melanoma through integration of clinico-pathologic, mutation, mRNA, microRNA, and protein information. *Int. J. Cancer* 136, 863–874.
- Jemal, A., Forbes, S.A., Goebell, P.J., M.A.K., Gui, Y., Hurst, C.D., Platt, F.M., Taylor, C.F., Knowles, M.A., Lindgren, D., et al. (2014). Comprehensive molecular characterization of urothelial bladder carcinoma. *Nature* 507, 315–322.
- Jönsson, G., Busch, C., Knappskog, S., Geisler, J., Miletic, H., Ringné, M., Lillehaug, J.R., Borg, A., and Lønning, P.E. (2010). Gene expression profiling-based identification of molecular subtypes in stage IV melanomas with different clinical outcome. *Clin. Cancer Res.* 16, 3356–3367.
- Jung, H., Kim, H.S., Kim, J.Y., Sun, J.-M., Ahn, J.S., Ahn, M.-J., Park, K., Esteller, M., Lee, S.-H., and Choi, J.K. (2019). DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. *Nat. Commun.* 10, 4278.
- Kim, S.T., Cristescu, R., Bass, A.J., Kim, K.-M., Odegaard, J.I., Kim, K., Liu, X.Q., Sher, X., Jung, H., Lee, M., et al. (2018). Comprehensive molecular characterization of clinical responses to PD-1 inhibition in metastatic gastric cancer. *Nat. Med.* 24, 1449–1458.
- Koeppl, F., Bobard, A., Lefebvre, C., Pedrero, M., Deloger, M., Boursin, Y., Richon, C., Chen-Min-Tao, R., Robert, G., Meurice, G., et al. (2018). Added value of whole-exome and transcriptome sequencing for clinical molecular screenings of advanced cancer patients with solid tumors. *Cancer J.* 24, 153–162.
- Kunz, M., Löffler-Wirth, H., Dannemann, M., Willscher, E., Doose, G., Kelso, J., Kottek, T., Nickel, B., Hopp, L., Landsberg, J., et al. (2018). RNA-seq analysis identifies different transcriptomic types and developmental trajectories of primary melanomas. *Oncogene* 37, 6136–6151.
- Lauss, M., Donia, M., Harbst, K., Andersen, R., Mitra, S., Rosengren, F., Salim, M., Vallon-Christersson, J., Törngren, T., Kvist, A., et al. (2017). Mutational and putative neoantigen load predict clinical benefit of adoptive T cell therapy in melanoma. *Nat. Commun.* 8, 1738.
- Leinonen, R., Sugawara, H., and Shumway, M.; International Nucleotide Sequence Database Collaboration (2011). The sequence read archive. *Nucleic Acids Res.* 39, D19–D21.
- Liang, W.S., Hendricks, W., Kiefer, J., Schmidt, J., Sekar, S., Carpten, J., Craig, D.W., Adkins, J., Cuyugan, L., Manojlovic, Z., et al. (2017). Integrated genomic analyses reveal frequent TERT aberrations in acral melanoma. *Genome Res.* 27, 524–532.
- Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 1, 417–425.
- Liu, D., Schilling, B., Liu, D., Sucker, A., Livingstone, E., Jerby-Amon, L., Zimmer, L., Gutzmer, R., Satzger, I., Loquai, C., et al. (2019a). Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nat. Med.* 25, 1916–1927.
- Liu, T., Han, C., Wang, S., Fang, P., Ma, Z., Xu, L., and Yin, R. (2019b). Cancer-associated fibroblasts: an emerging target of anti-cancer immunotherapy. *J. Hematol. Oncol.* 12, 86.
- Łuksza, M., Riaz, N., Makarov, V., Balachandran, V.P., Hellmann, M.D., Solovoyov, A., Rizvi, N.A., Merghoub, T., Levine, A.J., Chan, T.A., et al. (2017). A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. *Nature* 551, 517–520.
- Malone, E.R., Oliva, M., Sabatini, P.J.B., Stockley, T.L., and Siu, L.L. (2020). Molecular profiling for precision cancer therapies. *Genome Med.* 12, 8.
- Mantel, N., and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 22, 719–748.
- Mariathasan, S., Turley, S.J., Nickles, D., Castiglioni, A., Yuen, K., Wang, Y., Kadel, E.E., III, Koeppen, H., Astarita, J.L., Cubas, R., et al. (2018). TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* 554, 544–548.
- Matplotlib (n.d.) A 2D Graphics Environment - IEEE Journals & Magazine.
- McCabe, M.J., Gauthier, M.-E.A., Chan, C.-L., Thompson, T.J., De Sousa, S.M.C., Puttick, C., Grady, J.P., Gayevskiy, V., Tao, J., Ying, K., et al. (2019). Development and validation of a targeted gene sequencing panel for application to disparate cancers. *Sci. Rep.* 9, 17052.
- Mermel, C.H., Schumacher, S.E., Hill, B., Meyerson, M.L., Beroukhi, R., and Getz, G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* 12, R41.
- McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R., Thormann, A., Flicek, P., and Cunningham, F. (2016). The ensembl variant effect predictor. *Genome Biol* 17, 122.
- Miao, D., Margolis, C.A., Gao, W., Voss, M.H., Li, W., Martini, D.J., Norton, C., Bossé, D., Wankowicz, S.M., Cullen, D., et al. (2018). Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. *Science* 359, 801–806.
- Montesio, M., Murugesan, K., Dexter, J.X., Sharaf, R., Sanchez, N., Guria, A., Minker, M., Li, G., Fisher, V., Sokol, E.S., et al. (2021). Somatic HLA class I loss is a widespread mechanism of immune evasion which refines the use of tumor mutational burden as a biomarker of checkpoint inhibitor response. *Cancer Discov.* 11, 282–292.
- Nathanson, T., Ahuja, A., Rubinsteyn, A., Aksoy, B.A., Hellmann, M.D., Miao, D., Van Allen, E., Merghoub, T., Wolchok, J.D., Snyder, A., et al. (2017). Somatic mutations and neoepitope homology in melanomas treated with CTLA-4 blockade. *Cancer Immunol. Res.* 5, 84–91.
- Newman, A.M., Liu, C.L., Green, M.R., Gentles, A.J., Feng, W., Xu, Y., Hoang, C.D., Diehn, M., and Alizadeh, A.A. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12, 453–457.
- Ock, C.-Y., Hwang, J.-E., Keam, B., Kim, S.-B., Shim, J.-J., Jang, H.-J., Park, S., Sohn, B.H., Cha, M., Ajani, J.A., et al. (2017). Genomic landscape associated with potential response to anti-CTLA-4 treatment in cancers. *Nat. Commun.* 8, 1050.
- Ott, P.A., Hu, Z., Keskin, D.B., Shukla, S.A., Sun, J., Bozym, D.J., Zhang, W., Luoma, A., Giobbie-Hurder, A., Peter, L., et al. (2017). An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* 547, 217–221.
- Pan, D., Kobayashi, A., Jiang, P., Ferrari de Andrade, L., Tay, R.E., Luoma, A.M., Tsoucas, D., Qiu, X., Lim, K., Rao, P., et al. (2018). A major chromatin regulator determines resistance of tumor cells to T cell-mediated killing. *Science* 359, 770–775.
- Petitprez, F., de Reyniès, A., Keung, E.Z., Chen, T.W.-W., Sun, C.-M., Calderaro, J., Jeng, Y.-M., Hsiao, L.-P., Lacroix, L., Bougouin, A., et al. (2020). B cells are associated with survival and immunotherapy response in sarcoma. *Nature* 577, 556–560.
- Pirrotta, V. (1998). Polycombing the genome: PcG, trxG, and chromatin silencing. *Cell* 93, 333–336.
- Pomeranz Krummel, D., Nasti, T., Izar, B., Xu, M., Lowder, L., Press, R., Kaluzova, M., Kallay, L., Rupji, M., Burnham, A., et al. (2019). EXTH-12. Radiation enhances melanoma response to immunotherapy and synergizes with benzodiazepines to promote anti-tumor activity. *Neuro. Oncol.* 21, vi84.
- Pommier, A., and Fearon, D.T. (2016). Disruption of anti-tumor T cell responses by cancer-associated fibroblasts. *Resist. Targeted Anti Cancer Ther.* 77–98.
- Qu, Y., Wen, J., Thomas, G., Yang, W., Prior, W., He, W., Sundar, P., Wang, X., Potluri, S., and Salek-Ardakani, S. (2020). Baseline frequency of inflammatory Cxcl9-expressing tumor-associated macrophages predicts response to avelumab treatment. *Cell Rep.* 32, 108115.

- Raney, B.J., Cline, M.S., Rosenbloom, K.R., Dreszer, T.R., Learned, K., Barber, G.P., Meyer, L.R., Sloan, C.A., Malladi, V.S., Roskin, K.M., et al. (2011). ENCODE whole-genome data in the UCSC genome browser (2011 update). *Nucleic Acids Res.* **39**, D871–D875.
- Raskin, L., Fullen, D.R., Giordano, T.J., Thomas, D.G., Frohm, M.L., Cha, K.B., Ahn, J., Mukherjee, B., Johnson, T.M., and Gruber, S.B. (2013). Transcriptome profiling identifies HMG2 as a biomarker of melanoma progression and prognosis. *J Invest Dermatol* **133**, 2585–2592.
- Riaz, N., Havel, J.J., Makarov, V., Desrichard, A., Urba, W.J., Sims, J.S., Hodi, F.S., Martin-Algarra, S., Mandal, R., Sharfman, W.H., et al. (2017). Tumor and microenvironment evolution during immunotherapy with nivolumab. *Cell* **171**, 934–949.e16.
- Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47.
- Rizvi, N.A., Hellmann, M.D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J.J., Lee, W., Yuan, J., Wong, P., Ho, T.S., et al. (2015). Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124–128.
- Robinson, D.R., Wu, Y.-M., Lonigro, R.J., Vats, P., Cobain, E., Everett, J., Cao, X., Rabbani, E., Kumar-Sinha, C., Raymond, V., et al. (2017). Integrative clinical genomics of metastatic cancer. *Nature* **548**, 297–303.
- Rodig, S.J., Gusenleitner, D., Jackson, D.G., Gjini, E., Giobbie-Hurder, A., Jin, C., Chang, H., Lovitch, S.B., Horak, C., Weber, J.S., et al. (2018). MHC proteins confer differential sensitivity to CTLA-4 and PD-1 blockade in untreated metastatic melanoma. *Sci. Transl. Med.* **10**, eaar3342.
- Rodon, J., Soria, J.-C., Berger, R., Miller, W.H., Rubin, E., Kugel, A., Tsimberidou, A., Saintigny, P., Ackersstein, A., Braña, I., et al. (2019). Genomic and transcriptomic profiling expands precision cancer medicine: the WINTHER trial. *Nat. Med.* **25**, 751–758.
- Rooney, M.S., Shukla, S.A., Wu, C.J., Getz, G., and Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48–61.
- Sailer, V., Eng, K.W., Zhang, T., Bareja, R., Pisapia, D.J., Sigaras, A., Bhinder, B., Romanel, A., Wilkes, D., Sticca, E., et al. (2019). Integrative molecular analysis of patients with advanced and metastatic cancer. *JCO Precis. Oncol.* **3**, PO.19.00047.
- Saltz, J., Gupta, R., Hou, L., Kurc, T., Singh, P., Nguyen, V., Samarasinghe, D., Shroyer, K.R., Zhao, T., Batiste, R., et al. (2018). Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Rep.* **23**, 181–193.e7.
- Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W.K., Luna, A., La, K.C., Dimitriadou, S., Liu, D.L., Kantheti, H.S., Saghatian, S., et al. (2018). Oncogenic signaling pathways in the cancer genome atlas. *Cell* **173**, 321–337.e10.
- Saunders, C.T., Wong, W.S., Swamy, S., Becq, J., Murray, L.J., and Cheetham, R.K. (2012). Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* **28**, 1811–1817.
- Schaaf, M.B., Garg, A.D., and Agostinis, P. (2018). Defining the role of the tumor vasculature in antitumor immunity and immunotherapy. *Cell Death Dis.* **9**, 115.
- Schubert, M., Klinger, B., Klünemann, M., Sieber, A., Uhlitz, F., Sauer, S., Garnett, M.J., Blüthgen, N., and Saez-Rodriguez, J. (2018). Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat. Commun.* **9**, 20.
- Schwaederle, M., Zhao, M., Lee, J.J., Lazar, V., Leyland-Jones, B., Schilsky, R.L., Mendelsohn, J., and Kurzrock, R. (2016). Association of biomarker-based treatment strategies with response rates and progression-free survival in refractory malignant neoplasms: a meta-analysis. *JAMA Oncol.* **2**, 1452–1459.
- Schwarze, K., Buchanan, J., Fermont, J.M., Dreau, H., Tilley, M.W., Taylor, J.M., Antoniou, P., Knight, S.J.L., Camps, C., Pentony, M.M., et al. (2020). The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom. *Genet. Med.* **22**, 85–94.
- Şenbabaoğlu, Y., Gejman, R.S., Winer, A.G., Liu, M., Van Allen, E.M., de Velasco, G., Miao, D., Ostrovskaya, I., Drill, E., Luna, A., et al. (2016). Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol.* **17**, 231.
- Singer, J., Irmisch, A., Ruscheweyh, H.-J., Singer, F., Toussaint, N.C., Levesque, M.P., Stekhoven, D.J., and Beerenwinkel, N. (2019). Bioinformatics for precision oncology. *Brief. Bioinform.* **20**, 778–788.
- Sjödahl, G., Lauss, M., Lövgren, K., Chebil, G., Gudjonsson, S., Veerla, S., Patschan, O., Aine, M., Fernö, M., Ringnér, M., et al. (2012). *Clin. Cancer Res.* **18**, 3377–3386.
- Snyder, A., Makarov, V., Merghoub, T., Yuan, J., Zaretsky, J.M., Desrichard, A., Walsh, L.A., Postow, M.A., Wong, P., Ho, T.S., et al. (2014). Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* **371**, 2189–2199.
- Spranger, S., Bao, R., and Gajewski, T.F. (2015). Melanoma-intrinsic β -catenin signalling prevents anti-tumour immunity. *Nature* **523**, 231–235.
- Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U S A* **102**, 15545–15550.
- Suwinski, P., Ong, C., Ling, M.H.T., Poh, Y.M., Khan, A.M., and Ong, H.S. (2019). Advancing personalized medicine through the application of whole exome sequencing and big data analytics. *Front. Genet.* **10**, 49.
- Szolek, A., Schubert, B., Mohr, C., Sturm, M., Feldhahn, M., and Kohlbacher, O. (2014). OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics* **30**, 3310–3316.
- Tang, H., Wang, Y., Chlewicki, L.K., Zhang, Y., Guo, J., Liang, W., Wang, J., Wang, X., and Fu, Y.-X. (2016). Facilitating T cell infiltration in tumor microenvironment overcomes resistance to PD-L1 blockade. *Cancer Cell* **30**, 500.
- Tao, L., Huang, G., Song, H., Chen, Y., and Chen, L. (2017). Cancer associated fibroblasts: an essential role in the tumor microenvironment. *Oncol. Lett.* **14**, 2611–2620.
- Thorsson, V., Gibbs, D.L., Brown, S.D., Wolf, D., Bortone, D.S., Ou Yang, T.-H., Porta-Pardo, E., Gao, G.F., Plaisier, C.L., Eddy, J.A., et al. (2018). The immune landscape of cancer. *Immunity* **48**, 812–830.e14.
- Topalian, S.L., Hodi, F.S., Brahmer, J.R., Gettinger, S.N., Smith, D.C., McDermott, D.F., Powderly, J.D., Carvajal, R.D., Sosman, J.A., Atkins, M.B., et al. (2012). Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N. Engl. J. Med.* **366**, 2443–2454.
- Uhlen, M., Karlsson, M.J., Zhong, W., Tebani, A., Pou, C., Mikes, J., Lakshmikanth, T., Forsström, B., Edfors, F., Odeberg, J., et al. (2019). A genome-wide transcriptomic analysis of protein-coding genes in human blood cells. *Science* **366**, eaax9198.
- Ulloa-Montoya, F., Louahed, J., Dizier, B., Gruselle, O., Spiessens, B., Lehmann, F.F., Suci, S., Kruit, W.H.J., Eggemont, A.M.M., Vansteenkiste, J., et al. (2013). Predictive gene signature in MAGE-A3 antigen-specific cancer immunotherapy. *J. Clin. Oncol.* **31**, 2388–2395.
- Vaske, O.M., Bjork, I., Salama, S.R., Beale, H., Tayi Shah, A., Sanders, L., Pfeil, J., Lam, D.L., Learned, K., Durbin, A., et al. (2019). Comparative tumor RNA sequencing analysis for difficult-to-treat pediatric and young adult patients with cancer. *JAMA Netw. Open* **2**, e1913968.
- Vivian, J., Rao, A.A., Nothaft, F.A., Ketchum, C., Armstrong, J., Novak, A., Pfeil, J., Narkizian, J., Deran, A.D., Musselman-Brown, A., et al. (2017). Toil enables reproducible, open source, big biomedical data analyses. *Nat. Biotechnol.* **35**, 314–316.
- Wadapurkar, R.M., and Vyas, R. (2018). Computational analysis of next generation sequencing data and its applications in clinical oncology. *Inform. Med. Unlocked* **11**, 75–82.
- Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184–2185.

Waskom, M., Botvinnik, O., O'Kane, D., Hobson, P., Lukauskas, S., Gemperline, D.C., Augspurger, T., Halchenko, Y., Cole, J.B., Warmenhoven, J., et al. (2017). Mwaskom/Seaborn, p. v0.8.1.

Weiss, S.A., Wolchok, J.D., and Sznol, M. (2019). Immunotherapy of melanoma: facts and hopes. *Clin. Cancer Res.* 25, 5191–5201.

Wolchok, J.D., Chiarion-Sileni, V., Gonzalez, R., Rutkowski, P., Grob, J.-J., Cowey, C.L., Lao, C.D., Wagstaff, J., Schadendorf, D., Ferrucci, P.F., et al. (2017). Overall survival with combined nivolumab and ipilimumab in advanced melanoma. *N. Engl. J. Med.* 377, 1345–1356.

Wu, C., MacLeod, I., and Su, A. (2012). BioGPS and MyGene.info: organizing online, gene-centric information. *Nucleic Acids Res.* 41, D561–D565.

Wu, S., Lu, H., and Bai, Y. (2019). Nrf2 in cancers: a double-edged sword. *Cancer Med.* 8, 2252–2267.

Xu, L., Shen, S.S., Hoshida, Y., Subramanian, A., Ross, K., Brunet, J.-P., Wagner, S.N., Ramaswamy, S., Mesirov, J.P., and Hynes, R.O. (2008). Gene expression changes in an animal melanoma model correlate with aggressiveness of human melanoma metastases. *Mol. Cancer Res.* 6, 760–769.

Zaretsky, J.M., Garcia-Diaz, A., Shin, D.S., Escuin-Ordinas, H., Hugo, W., Hu-Lieskovan, S., Torrejon, D.Y., Abril-Rodriguez, G., Sandoval, S., Barthly, L., et al. (2016). Mutations associated with acquired resistance to PD-1 blockade in melanoma. *N. Engl. J. Med.* 375, 819–829.

Zhuang, J., Lu, Q., Shen, B., Huang, X., Shen, L., Zheng, X., Huang, R., Yan, J., and Guo, H. (2015). TGF β 1 secreted by cancer-associated fibroblasts induces epithelial-mesenchymal transition of bladder cancer cells through lncRNA-ZEB2NAT. *Sci. Rep.* 5, 11924.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
Melanoma tissue samples	TCGA (https://portal.gdc.cancer.gov/)	See Table S5
Deposited Data		
Adenocarcinomas	expO Project	GEO: GSE2109
Metastatic melanoma-Cirenajwis	Cirenajwis et al. (2015)	GEO: GSE65904
Melanoma-Budden	Budden et al. (2016)	GEO: GSE59455
Anti-PD1-treated metastatic melanoma-Liu	Liu et al., 2019a	dbGaP: phs000452
Anti-PD1-treated advanced melanoma-Riaz	Riaz et al. (2017)	GEO: GSE91061
Melanoma-Hao	Hao et al. (2017)	GEO: GSE43081
Anti-PD1+ anti-CTLA-4 or anti-PD1-treated melanoma-Gide	Gide et al. (2019)	ENA: ERP105482
Melanoma-AJCC_1 and _2	Hayward et al. (2017) ; Jayawardana et al. (2015)	GEO: GSE54467, GEO: GSE80435
Melanoma-Xu	Xu et al. (2008)	GEO: GSE8401
Melanoma and nevus-Kunz	Kunz et al. (2018)	GEO: GSE112509
Nevus and melanoma-Badal	Badal et al. (2017)	GEO: GSE98394
MAGE-3-treated melanoma-Ulloa-Montoya	Ulloa-Montoya et al. (2013)	GEO: GSE35640
Melanoma-Raskin	Raskin et al., 2013	GEO: GSE15605
Melanoma-Jonsson	Jönsson et al. (2010)	GEO: GSE22153
Metastatic melanoma-Augustine	Augustine et al. (2010)	GEO: GSE19293
Metastatic melanoma-Bogunovic	Bogunovic et al. (2009)	GEO: GSE19234
Anti-CTLA-4-treated melanoma-Van Allen	Van Allen et al. (2015)	dbGaP: phs000452
Anti-CTLA-4, anti-CTLA-4+ anti-PD1, anti-PD1-treated melanoma-Auslander	Auslander et al. (2018)	GEO: GSE115821
Anti-CTLA-4, anti-PD1-treated melanoma-Liang	Liang et al. (2017)	dbGaP: phs001036
Anti-PD1-treated melanoma-Hugo	Hugo et al. (2016) ; Garcia-Diaz et al. (2017)	GEO: GSE78220, GEO: GSE96619
ACT-treated melanoma-Lauss	Lauss et al. (2017)	GEO: GSE100797
Anti-CTLA-4-treated melanoma-Nathanson	Nathanson et al. (2017)	SRA: SRP067586
Anti-PD-1-treated melanoma-Khan	Pomeranz Krummel et al. (2019)	GEO: GSE131521, SRP198996
Anti-PDL-1-treated bladder cancer-Mariathasan	Mariathasan et al. (2018)	EGA: EGAS00001002556
Anti-PD-L1 or anti-PD1-treated lung cancer-Jung	Jung et al. (2019)	GEO: GSE135222
Anti-PD-L1-treated gastric cancer-Kim	Kim et al. (2018)	ENA: ERP107734, SRA: PRJEB25780
human reference genome GRCh38	Genome Reference Consortium	http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/
Software and Algorithms		
Codes used for ssGSEA scoring affy 1.52.0	This paper Gautier et al., 2004	https://github.com/BostonGene RRID: SCR_012835; http://www.bioconductor.org/packages/release/bioc/html/affy.html

(Continued on next page)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
limma R	Ritchie et al., 2015	RRID: SCR_010943; http://bioconductor.org/packages/release/bioc/html/limma.html
MANTIS	Bonneville et al. (2017)	https://github.com/OSU-SRLab/MANTIS
Immunophenoscore	Charoentong et al. (2017)	https://tcia.at/tools/toolsMain
FastQC v0.11.1 and v0.11.5	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
RSeQC v3.0.0	Wang et al., 2012	http://rseqc.sourceforge.net/
MultiQC v1.6	Ewels et al., 2016	https://github.com/ewels/MultiQC
Conpair algorithm	Bergmann et al. (2016)	https://github.com/nygenome/conpair
FilterByTile/BBMap v37.90	NA	https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/bbmap-guide/
BWA v0.7.17	NA	https://github.com/lh3/bwa/releases/tag/v0.7.17
Picard's v2.6.0	http://broadinstitute.github.io/picard/	http://broadinstitute.github.io/picard/
GATK v3.8.1	NA	https://gatk.broadinstitute.org/hc/en-us
Strelka v2.9	Saunders et al., 2012	https://github.com/Illumina/strelka
Variant Effect Predictor v92.1	McLaren et al., 2016	https://uswest.ensembl.org/info/docs/tools/vep/index.html
Sequenza v2.1.2	Favero et al., 2015	https://github.com/cran/sequenza
Kallisto v0.42.4	Bray et al., 2016	https://pachterlab.github.io/kallisto/
GENCODE v23 transcripts 69	Frankish et al., 2019	https://www.gencodegenes.org/
MIXCR version 2.1.7	https://mixcr.readthedocs.io/en/master/	https://mixcr.readthedocs.io/en/master/
CamDavidsonPilon/lifelines: v0.14.6 (Version v0.14.6)	doi: 10.5281/zenodo.4002777	https://github.com/CamDavidsonPilon/lifelines
CIBERSORT	Newman et al. (2015)	https://cibersort.stanford.edu/
MCP-counter R package	Becht et al. (2016)	https://github.com/ebecht/MCPcounter
ssGSEA	Subramanian et al. (2005)	http://software.broadinstitute.org/gsea
PROGENY	Schubert et al. (2018)	https://saezlab.github.io/progeny/
Louvain community detection algorithm	Blondel et al. (2008)	https://github.com/taynaud/python-louvain
Cochran-Mantel-Haenszel test (R stats package v3.4.4)	Mantel and Haenszel, 1959	NA
GISTIC 2.0.23 analysis	Mermel et al., 2011	https://software.broadinstitute.org/cancer/cga/gistic
Ordinal regression model (R clm from package ordinal, v2018.4.19)	NA	https://github.com/runehaubo/ordinal
lsmeans R package (v2.27.62)	NA	https://github.com/rvlenth/lsmeans
spectral co-clustering algorithms (Spectral Co-clustering model from scikit-learn 0.22.2)	NA	https://scikit-learn.org/stable/
Python-matplotlib (3.0.3)	NA	https://github.com/matplotlib/matplotlib
python-seaborn (v0.7.1) mwaskom/seaborn: v0.8.1 Zenodo	Waskom et al. (2017)	https://doi.org/10.5281/zenodo.883859
UMAP package	NA	https://github.com/lmcinnes/umap
Other		
Molecular Functional Portrait Visual Tool	This paper	https://science.bostongene.com/tumor-portrait/

RESOURCE AVAILABILITY

Lead contact

Further information and requests should be directed to and will be fulfilled by the Lead Contact, Nathan Fowler (nfowler@mdanderson.org).

Materials availability

This study did not generate new unique reagents.

Data and code availability

This study did not produce any raw data. Processed data, including reanalysis of published datasets showing the Fges scores are available in the web-based tool at <http://science.bostongene.com/tumor-portrait/>. Scripts used to generate results are available at <https://github.com/BostonGene>. The MF Portrait visualization tool can be queried and visualized at <http://science.bostongene.com/tumor-portrait/>.

Accessions for the datasets used in this study include the following: phs000178 (TCGA), GSE2109 (adenocarcinomas), GSE65904 (Cirenajwis et al., 2015), GSE59455 (Budden et al., 2016), GSE98394 (Badal et al., 2017), phs001036 (Liang et al., 2017), GSE91061 (Riaz et al., 2017), GSE43081 (Hao et al., 2017), ERP105482 (Gide et al., 2019), GSE54467 and GSE80435 (AJCC_1 and AJCC_2 (Hayward et al., 2017; Jayawardana et al., 2015), GSE8401 (Xu et al., 2008), GSE112509 (Kunz et al., 2018), GSE35640 (Ulloa-Montoya et al., 2013), GSE15605 (Raskin et al., 2013), GSE22153 (Jönsson et al., 2010), GSE19293 (Augustine et al., 2010), GSE19234 (Bogunovic et al., 2009), phs000452 (Liu et al., 2019a; Van Allen et al., 2015), GSE115821 (Auslander et al., 2018), GSE78220 and GSE96619 (Hugo et al., 2016; Garcia-Diaz et al., 2017), GSE100797 (Lauss et al., 2017), SRP067586 (Nathanson et al., 2017), GSE131521/SRP198996 (Pomeranz Krummel et al., 2019), EGAS00001002556 (Mariathasan et al., 2018), GSE135222 (Jung et al., 2019), ENA: ERP107734/SRA: PRJEB25780 (Kim et al., 2018).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Tissue samples

Randomly selected melanoma H&E tissue images per TME subtype (total n = 69) were uploaded from TCGA (<https://portal.gdc.cancer.gov/>), and after QC, n = 66 were analyzed by a histopathologist. The sample ID numbers are listed in Table S5.

METHOD DETAILS

Tumor datasets

Melanoma dataset collection

We collected 24 public melanoma datasets from the GEO and SRA databases, including the TCGA skin cutaneous melanoma project (TCGA-SKCM) (total n samples = 2,016) (Cancer Genome Atlas Network, 2015). Out of 2,016 samples, 23 samples did not pass quality control due to one of the following reasons: PCA outlier, low correlation with others within the cohort (<0.8 for Affymetrix platforms, < 0.65 for Illumina platforms), low coverage and low phred scores for the RNA-seq, high non-human tissue contamination (>3%), or high percentage of duplicates (>80%), totaling 1,993 samples. Nevus samples (n = 50) (Badal et al., 2017) (Kunz et al., 2018) were included for comparison, yielding the final cohort of 2,043 samples. These datasets were processed and then harmonized as shown in Figures S13A–S13F. Overall, the melanoma cohorts underwent rigorous annotation, including information regarding the original tissue storage method (fresh-frozen versus FFPE), laboratory source, RNA extraction (total RNA and poly(A) and RNA-seq protocols provided in Table S5). We found that almost all the RNA-seq datasets subjected to the same RNA extraction and library preparation methods could be analyzed as combined for further analysis (i.e., without batch effects). RNA-seq datasets (n = 12) were combined into 5 RNA-seq datasets controlling for batch effects using PCA projections and gene expression correlation analysis. Next, Fges scores using ssGSEA were calculated for all 5 RNA-seq datasets as well as the microarray datasets and then median transformed (Figures S13C and S13D). Finally, the scores from all datasets were combined and analyzed together. Representative normalization steps for 3 Fges are presented in Figure S13D, with normalization occurring for all the Fges. The process intensity had relative scores that allowed for the analysis across different datasets. This procedure was followed for all datasets of various cancer types and is further detailed below.

Patient clinical information (e.g., therapy, OS, PFS, DFS, DMFS, response, recurrence, prior therapies, primary/metastatic site, LN/cutaneous/acral/mucosal/other location, primary type (cutaneous/acral/mucosal/other), histologic subtype, Breslow depth, ulceration) and tumor sample molecular data (e.g., mutation/neoantigen load, specific mutation status, and histological data) from all the cohorts were manually curated and harmonized.

For each RNA-seq cohort, RNA enrichment type (total or poly(A)), sample source (FF/Fresh or FFPE) and library prep protocol were curated. The Liang et al. cohort (Liang et al., 2017) (phs001036) was split into two cohorts based on RNA enrichment type (total vs poly(A)). HLA haplotype analysis, SNP concordance and correlation of expression showed that 5 patient samples from GSE96619 (Garcia-Diaz et al., 2017) were the same patients from the Hugo cohort (Hugo et al., 2016) (Pt 5 = SRR5343917 + SRR5343918 post treatment; Pt 10 = SRR5343921 + SRR5343922 post treatment; Pt 12 + SRR5343924 + SRR5343923 pre-treatment - potential timing mix up; Pt 15 = SRR5343919 + SRR5343920; Pt 23 = SRR5343925 + SRR5343926 post treatment). We extended the Hugo et al. cohort (Hugo et al., 2016) (GSE78220, GSE96619) with the 5 non-duplicate samples. GSE54467 and GSE80435 samples were split by platforms resulting in two cohorts: AJCC_1 (GPL6884) and AJCC_2 (GPL10558) (Jayawardana et al., 2015; Hayward et al., 2017). Clinical annotations for AJCC cohorts were matched with the MELA-AU ICGC project annotations by MELA id.

Raw and processed microarray data were downloaded from GEO. Expression was re-processed from raw files, if possible, using affygcRMA and oligo R packages. All affymetrix datasets with available CEL files were re-normalized using the gcRMA package with

default parameters. Illumina array data were downloaded from GEO as is. Next, probes were converted into genes using 1 probe with the highest mean values in the cohort per gene.

RNA-seq expression for cohorts where raw fastq files were available: Auslander (GSE115821) (Auslander et al., 2018); Badal (GSE98394) (Badal et al., 2017); Gide (ERP105482) (Gide et al., 2019); Hugo (GSE78220, GSE96619) (Hugo et al., 2016; Garcia-Diaz et al., 2017); Khan (GSE131521) (Pomeranz Krummel et al., 2019), Kunz (GSE112509) (Kunz et al., 2018), Liang (phs001036) (Liang et al., 2017), Nathanson (SRP067586) (Nathanson et al., 2017), Riaz (GSE91061) (Riaz et al., 2017), TCGA-SKCM (Cancer Genome Atlas Network, 2015), Van Allen (phs000452) (Van Allen et al., 2015; Liu et al., 2019a) was re-calculated using a unified RNA-seq pipeline (Vivian et al., 2017; <https://xenabrowser.net/>). Cohorts with the same RNA extraction type, sample source and similar sequencing protocols were combined into large cohort groups and processed together controlling for batch effects in the space of the Fges (Figure S13). RNA-seq cohorts with no available raw data files at the time of analysis were downloaded from the corresponding supplemental files. The RNA-seq data belonging to the Liu and Lauss cohorts were downloaded from their supplementary files. Transcript values were summarized and converted to HUGO Symbols, and FPKM values were converted to TPM-like values by normalizing to 1M transcripts. Each transcriptomic microarray or RNA-seq dataset with no raw data available was assigned to a unique cohort group (Table S5).

Immunotherapy-associated dataset collection

We collected datasets of patients treated with immunotherapies with available transcriptomics data: 10 melanoma, one lung cancer (Jung, n = 27) (Jung et al., 2019) one bladder cancer (Mariathasan, n = 346) (Mariathasan et al., 2018), and one gastric cancer (Kim, n = 34) (Kim et al., 2018) datasets (Table S5). For the Kim et al. cohort, several samples were excluded due to sample mismatches with WES (and probably annotation): PB-16-047, PB-16-048, PB-16-049, PB-16-051, PB-16-052, PB-16-055, PB-16-056, PB-16-057. Samples PB-16-043, PB-16-066 excluded due to low protein coding coverage (<10M reads). Sample PB-16-054 was excluded because it was outlier with PCA. In total, 11 samples were excluded from the Kim et al. cohort. Additionally, we included transcriptomic cohorts of patients treated with MAGE-A3 vaccine (Ulloa-Montoya et al., 2013) and adoptive cell therapy (Lauss et al., 2017) (Table S5).

Melanoma immunotherapy datasets

We curated samples annotated with pre/on/post treatment, response (bCT-based or progression-free survival based), Recist, OS, PFS, time to therapy start, and previous treatment with immune checkpoint therapies as shown in Figure S13. We selected only pre-treatment samples collected less than 200 days before the start of therapy. For anti-CTLA-4-treated patients from Van Allen et al. (n = 40) (Van Allen et al., 2015) and Nathanson et al. (n = 20) (Nathanson et al., 2017) cohorts, only non-acral or mucosal samples were analyzed (Table S5). For cohorts of patients treated with anti-PD1 therapy, we selected pre-treatment samples of patients with cutaneous melanomas which received anti-CTLA4 or anti-PD1 therapy within less than 200 days after biopsy. The Riaz et al. cohort (Riaz et al., 2017) was excluded because only 15 samples remained after filtering non-cutaneous and anti-CTLA-4 progressed samples. The Auslander et al. cohort (Auslander et al., 2018) was excluded due to low patient number and the high prevalence of non-responders. The GSE131521 (Pomeranz Krummel et al., 2019) cohort was excluded because it primarily consisted of brain metastases and almost completely was classified as TME subtype D. The Liang et al. cohort (Liang et al., 2017) was excluded due to all samples being acral melanomas. Finally, the Gide, Liu, Hugo cohort samples were combined and analyzed as shown in Figure S13 (Table S5).

TCGA molecular and clinical data

TCGA analysis

Overall, 25 TCGA solid tumors (carcinomas) ACC, BLCA, BRCA, CESC, CHOL, COAD, READ, ESCA, STAD, HNSC, KICH, KIRC, KIRP, LIHC, LUAD, LUSC, OV, PAAD, PCPG, PRAD, SKCM, THCA, UCEC, UCS, UVM were used in the study. Clinical and mutation data were downloaded from the GDC TCGA data portal (MC3 dataset) (Ellrott et al., 2018). TCGA immune subtypes were retrieved from the immune landscape publication. Transcriptomic data were downloaded from the USCS XENA portal <https://xena.ucsc.edu/> as TPM units. Sample IDs were unified to patient IDs (first 12 characters). Patients with more than 1 tumor RNA-Seq sample or missing clinical annotation were removed. In total, 7,979 samples were analyzed. The samples from the 25 TCGA projects were re-organized into 25 histological subtypes based on ICD10 histological code and divergence on PCA and UMAP projections. COAD and READ were combined into colorectal adenocarcinomas (COREAD) (Cancer Genome Atlas Network, 2012). The CESC project was divided into cervical squamous cell carcinomas (CESC_SCC) and adenocarcinomas (CESC_AC). The ESCA project was divided into squamous cell carcinomas (ESCA_SCC) and adenocarcinomas, which were combined with STAD (gastric cancer samples) into esophagogastric adenocarcinomas (ESGA_AC) (Cancer Genome Atlas Research Network et al., 2017). Tumor cellularity (purity) estimation was determined as purity estimations (CPE) as previously described (Aran et al., 2015). Tumor-infiltrating nonmalignant cell number was calculated as 1 - tumor cellularity. Viral abundance was obtained from a previous publication (Rooney et al., 2015). Samples with more than 1 viral RPM were considered positive. MSI status was calculated using MANTIS (Bonnevillie et al., 2017). Samples with MANTIS Score >0.4 were considered as MSI-positive. MSI-positive and *POLE* mutants (Cancer Genome Atlas Research Network, 2014; Cancer Genome Atlas Network, 2012) were combined to compare “hypermutated” sample percentages across the clusters. The Immunophenoscore (IPS) (anti-PD1) for the TCGA melanoma cohorts was obtained from <https://tcia.at/> (Charoentong et al., 2017). TCGA immune clusters C1-C6, signature values and TCGA individual tumor subtypes were curated from (Thorsson et al., 2018).

Melanoma TCGA dataset (SKCM)

In the original TCGA SKCM publication (Cancer Genome Atlas Network, 2015), only 329 samples were classified into transcriptomic clusters. We utilized KNN with $K = 25$ neighbors in the space of top 500 STD genes to classify all other samples ($n = 141$). Top frequently altered genes in melanoma were selected for comparison between clusters: [BRAF, NRAS, HRAS, NF1, TP53, APC, CDKN2A] (Cancer Genome Atlas Network, 2015).

Histological examination of the SKCM TCGA samples

We randomly selected approximately 17 samples of each TME subtype ($n = 69$). A pathologist performed blind re-examination of diagnostic formalin-fixed paraffin-embedded slides of the tissue provided by the TCGA data portal (<https://portal.gdc.cancer.gov/>). Poor quality histology slides were excluded from analysis ($n = 3$ slides), totaling 66 analyzed slides (Table S5). Semi-quantitative pathological assessment was performed on the tissue slides. Pathological analysis focused on the description of tumor immune inflammation calculating lymphocytic score and stromal organization of tissue by measuring relative fibroblast score. The semi-quantitative scores were calculated using a 5-grade system (0–4). Tumor purity was also estimated histopathologically as the visually detected percent of tumor cells to all cells in the sample slide. Representative images were taken from slides of appropriate quality for each TME subtype using pathology sideviewing software Aperio ImageScope on magnification 200x.

TME cluster comparison with publicly available scores and publicly available clusters within the TCGA dataset

PRAD, PCPG, THCA projects were excluded from survival comparison in Figure 4E due to low percentage (<5%) of 5-year events (deaths, Figure S7A). Immunophenoscore (IPS) was split by median (Charoentong et al., 2017). TCGA pan-cancer clusters (Thorsson et al., 2018) were distributed unevenly in each project. Low populated clusters in each cancer type (less than 10 samples) were excluded from all survival analyses (Figures 4F and S6) for better visualization. Also, because of the disproportion of IS clusters within TCGA projects, in HR comparison in Figure 4E, we excluded TCGA projects (THCA and KIRC) where the majority of the samples (>85%) were classified as a single cluster (Figure S7D).

Bioinformatics

NGS data quality control

Quality control of all NGS samples was performed using FastQC v0.11.5, FastQ Screen v0.11.1, RSeQC v3.0.0, MultiQC v1.6. HLAs were genotypes from RNA-seq or WES using OptiType (Szelek et al., 2014). Sample correspondence was checked using HLA comparison and the conpair algorithm (Bergmann et al., 2016).

WES processing

Alignment: low quality reads were filtered using FilterByTile/BBMap v37.90 and aligned to human reference genome GRCh38 (GRCh38.d1.vd1 assembly) using BWA v0.7.17. Duplicate reads were removed using Picard's v2.6.0 MarkDuplicates, indels were realigned by IndelRealigner and recalibrated by BaseRecalibrator and ApplyBQSR (last three tools from GATK v3.8.1). Variant calling: Both germline and somatic single nucleotide variations (sSNVs), small insertions and deletions were all detected using Strelka v2.9. All variants, insertions and deletions were annotated using Variant Effect Predictor v92.1. Copy number alterations were evaluated with a customized version of Sequenza v2.1.2.

Tumor mutation load (TMB) calculations

Non-synonymous coding mutations were used to calculate mutation load. WES mutation burden was calculated as (Mutation load/50). Melanoma samples with <100 mutations were considered low-mutated, samples with >500 mutations were considered highly mutated and the rest were medium-mutated. In the Mariathasan et al. cohort (Mariathasan et al., 2018), we split mutations by top quartile of mutation burden or by median. For pan-cancer analysis within TCGA, we performed median transformation within cohort groups.

RNA-seq processing

RNA-seq reads were aligned using Kallisto v0.42.4 to GENCODE v23 transcripts 69 with default parameters. The protein-coding transcripts, IGH/K/L- and TCR-related transcripts were retained, and the noncoding RNA, histone- and mitochondria-related transcripts were removed, resulting in 20,062 protein coding genes. Gene expression was quantified as transcripts per million (TPM) and log₂-transformed (Goldman et al., 2020).

T cell receptor/B cell receptor repertoire profiling

MIXCR version 2.1.7 was used to analyze the RNA-seq samples. Single clonotypes were grouped into clones with unique VDJ combination and identical CDR3 nucleotide sequences. For B-cells, the clones were further aggregated into clone groups if the VDJ combination was the same and if the CDR3 nucleotide sequences differed no more than 1 nt.

Survival analysis

Survival differences were assessed using log rank test CamDavidsonPilon/lifelines: v0.14.6 (Version v0.14.6) (Davidson-Pilon et al., 2018). OS and PFS for the SKCM TCGA project (Cancer Genome Atlas Network, 2015) were corrected to sample collection time (e.g., submitted tumor diagnosis). Samples with PFS time < collection time were not used. Survival analysis for melanoma cohorts without systematic treatment or on standard of care therapies was performed on cohorts with typical survival curves: TCGA-SKCM (Cancer Genome Atlas Network, 2015), Cirenajwis (Cirenajwis et al., 2015), AJCC_1 (Hayward et al., 2017), Xu (Xu et al., 2008), and Bogunovic and Jönsson (Bogunovic et al., 2009; Jönsson et al., 2010). Immunotherapy-treated cohorts were excluded from the pan-melanoma survival analysis.

For survival analysis amongst biomarkers, single-variate Cox regression modeling was conducted controlling for cohort variability for the melanoma datasets. Only our TME subtypes, melanoma (Cancer Genome Atlas Network, 2015) and bladder cancer Lund

clusters (Sjödahl et al., 2012) were analyzed as categorical variables. All numeric variables were median-transformed unless otherwise specified. Kaplan-Meier comparative survival analyses were also conducted. To extract the TCGA types for survival comparisons across the melanoma and bladder cancer cohorts, the TCGA gene signatures were curated from Thorsson et al. (2018), and the gene names were updated using mygene v.3.0.0 (Wu et al., 2012). Signature values were calculated using ssGSEA (Figure S9A). Next, a random forest-based model was trained and validated on the TCGA SKCM dataset and applied to the other immunotherapy-treated melanoma cohorts (Gide et al., 2019; Hugo et al., 2016; Liu et al., 2019a; Nathanson et al., 2017; Van Allen et al., 2015) with melanoma-scaled signature values. The same approach was applied to the TCGA bladder cancer cohort and applied to the immunotherapy-treated bladder cancer cohorts (Mariathasan et al., 2018). The Immunophenoscore (Charoentong et al., 2017) values were curated from the TCI portal (<https://tcia.at/tools/toolsMain>). TMB was calculated as described elsewhere in the manuscript and log₁₀ transformed. The interferon γ signature was obtained from Thorsson et al. (2018) and calculated using ssGSEA. We reproduced the IMPRES score as previously described (Auslander et al., 2018), which was applied to all the analyzed cohorts. For MHC1 and MHCII, the corresponding Fges were applied. CXCL9 and PD-L1 expression was analyzed as single gene expression in log₂ scale. The F subtype was used as an intercept for each cancer type (Figure S10A). Other continuous variables were normalized (median-scaled), clipped to (-2; 2) for visualization and projected into [0–1] for better hazard ratio comparability within each other and with categorical variables. Estimated marginal means analysis was applied to extract intercept-independent information for each variable (R package emmeans; Figure S11).

Transcriptomic signatures

Cell abundance was measured using CIBERSORT with the LM22 matrix (Newman et al., 2015) and the MCP-counter R package (Becht et al., 2016). Expression signatures from (Şenbabaoğlu et al., 2016) were calculated using ssGSEA. Pathways activity scores (N = 11) were calculated using PROGENy (Schubert et al., 2018). CYT score was calculated as previously described (Rooney et al., 2015). Chromosomal instability score (CIN) was calculated as previously described (Ock et al., 2017). The TCGA melanoma types were curated from (Cancer Genome Atlas Network, 2015). Samples with unknown class were typed using the k-nearest neighbors algorithm with n-nearest neighbors = 25 in the space of the 500 most variable genes in TCGA-SKCM cohort.

Development of the gene expression signatures (Fges)

Purified cell type compendium

We collected 4,212 RNA-seq gene expression datasets from purified cell populations, including normal and melanoma cells from public data sources (GEO (Barrett et al., 2013), SRA (Leinonen et al., 2011), ENA, Array Express, Protein Atlas (Uhlen et al., 2019), Blueprint, ImmPort (import.org), to create a cell compendium (Table S2). We included datasets using the following criteria: isolated from human tissue, poly(A) or total RNA-seq performed with read length higher than 31 bp, having at least 4M of coding read counts, passed quality control by FASTQC and no contamination detected (<2%). The compendium included sorted T cells, CD4⁺ T cells, T-regs, T helper cells, CD8⁺ T cells, NK cells, benign B cells, granulocytes of different types, neutrophils, macrophages, monocytes, myeloid cells, dendritic cells, plasmacytoid dendritic cells, vascular endothelial cells, and fibroblasts. “Activated macrophages” cells were selected from samples of macrophages activated with bacteria or LPS according to annotation. Sorted melanoma tumor cells and cell lines were also included for comparison.

Fges describing the TME and tumor properties

We developed Fges of TME cellular phenotypes, cellular states, physiological and pathological processes and signaling pathways using a combination of Fges and literature curation. By integrating knowledge from multiple publications, we manually curated each signature to include only those genes that are exclusively expressed in the defined cell type or specifically associated with a particular biological process as described in the text. The Fges defining unique cell types were validated on the RNA-seq compendium of purified immune and tumor cell populations n = 7,011 using tSNE projections and Mann-Whitney tests. Signature scores were calculated by in-house python implementation of the ssGSEA (Hänzelmann et al., 2013). Then, the intensities were median-scaled (median-centered and MAD-scaled) for all the samples within the cohort groups. Signature correlation analysis was performed using Pearson’s correlation. TCGA SKCM samples were used to estimate correlating signatures.

Louvain clustering

We calculated the Pearson correlation [-1, 1] between all TCGA SKCM samples in the space of the 26 process intensities (normalized ssGSEA enrichment scores). Next, the distance matrix was converted into a graph where each sample formed a node and two nodes formed an edge with the weight equal to their Pearson correlation. All edges with weight <0.45 were removed. The Louvain community detection algorithm (Blondel et al., 2008) was applied to calculate graph partitioning into clusters with default parameters. To mathematically determine the optimal weight threshold for observed clusters, minimum David Bolduin, maximum Calinski Harabasz and Silhouette scores were employed (Figure S2). Separations with low populated clusters (<5% of samples) were not considered, and suboptimal separations were observed with more than 4 clusters/subtypes. The best four-cluster separation was selected, and the clusters were labeled as TME Subtypes IE/F, IE, F, and D. Ultimately, the same procedure was applied to the pan-cancer cohort. The melanoma clusters were visualized in cytoscape (v3.4.0) (10.1101/gr.1239303.metabolite). Nodes were organized using “Perforce force directed layout” (default spring coefficient = 1×10^{-5} , number of iterations = 100). Node size represents the number of its neighbors (adjacent edges). Node color corresponds to tumor subtype unless otherwise mentioned. Other melanoma cohorts were classified using KNN with K = 35. Bladder cancer (Mariathasan et al., 2018) was classified using a KNN model trained on TCGA BLCA samples; lung cancer (Jung et al., 2019) was classified using a KNN model trained on TCGA LUAD samples; the EXPO dataset GSE2108 was classified using a KNN model trained on pan-cancer TCGA cohort samples.

Statistical analysis of alterations and CNV

Mutation associations

To investigate the association of mutations in driver genes with clusters, we used the Cochran-Mantel-Haenszel test (R stats package v3.4.4), controlling for tumor histological subtypes, to test the independence of the immune clusters and gene mutation counts. A total of 509 genes were selected by combining TCGA PanCancer Atlas Driver Mutation Working Group CGAT list (Bailey et al., 2018) and genes from the immune processes gene sets. Resulting p values were FDR corrected using Benjamini-Hochberg procedure, and genes odds ratios are reported.

CNA association

To limit the number of segments with CNAs for testing, we performed GISTIC v23 analysis on the pan-cancer TCGA cohort. To estimate prevalence of copy number variations in clusters, ordinal regression was used. Copy number events in cytobands per sample were averaged in genes that passed 0.95 GISTIC significance threshold. Then, each level of averaged event was converted into ordinal scale, preserving increasing order of alterations from deletion to multiplication. Ordinal regression model (R *clm* from package *ordinal*, v2018.4.19) with proportional odds link function was fitted using these data for each cytoband, with cluster as a categorical variable and controlling for histological type. Marginal effects of each cluster were calculated using contrasts implemented in *lsmeans* R package (v2.27.62), with FDR Benjamini-Hochberg correction of effects p values. Cluster effects on CNA in proportional odds scale was plotted as horizontal barplots.

Alteration clusters

A list of actionable genes was obtained from (Sanchez-Vega et al., 2018), resulting in the total of 5,116 patients from TCGA having at least one mutation in a selected gene. SNVs and CNAs from selected genes were clustered using spectral co-clustering algorithms (Spectral Co-clustering model from *scikit-learn* 0.22.2) to obtain diagonal matrices of actionable clusters. Then, actionable clusters were manually curated to refine groups of genes belonging to pathways.

Data visualization

Images and graphics

Python-matplotlib (v1.5.1) (Matplotlib, n.d.) and python-seaborn (v0.7.1) mwwaskom/seaborn: v0.8.1 Zenodo (Waskom et al., 2017) were employed. The Pearson correlation was used as the default similarity metric (unless otherwise mentioned) for correlation matrices. Hierarchical clustering was performed using complete linkage and Euclidean distance for correlation matrices clustering. UMAP projection was performed using the UMAP package <https://github.com/lmcinnes/umap> and visualized with matplotlib.

Visualization of the Molecular Functional portrait

The portraits were visualized as a graph based structure using Mathematica 11 standard packages (Wolfram Research, Champaign, IL, USA). A node size that describes an intensity of a process in a particular patient was based on a normalized score calculated for process intensity. A distribution of ssGSEA scores for each process was mapped to the range of (0,1) by a cumulative distribution function (CDF) within the corresponding TCGA cohort. Driver mutations influencing therapeutic and prognostic outcomes were depicted in the tumor properties group as the “mutation status” node, representing a total number of nonsynonymous mutations found in the patient tumor. The upper genes arising from this node show only the most important recurrent mutations. The “mutation status” node size was also transformed to the range of (0,1) by CDF from the corresponding cohort distribution.

All the processes were labeled as either anti-tumor or pro-tumor. Anti-tumor processes were colored in a blue gradient, pro-tumor processes were colored in a burgundy gradient. The color shades represent process intensity. Gene nodes have a fixed size and are colored using the same method as the processes. The size of the “malignant cell” node as well as the “non-malignant microenvironment” nodes were visualized based on the tumor purity.

QUANTIFICATION AND STATISTICAL ANALYSIS

The statistical information for the experiments are detailed in the text, figure legends, and figures.

Chi-square test was used to check associations between 2 categorical variables with more than 2 categories. The Pearson's correlation was used as a Fges similarity measure. The Mann-Whitney U test was used for statistical analysis comparing non-categorical values between groups. For survival analysis, statistical analysis between Kaplan-Meier curves was performed using the log rank test (pairwise and/or multivariate). Linear regression analysis was applied to assess TME evolution and dynamics. Log rank score test was used for Cox proportional hazards models. LN-transformed Hazard ratios were shown unless specified. If other statistical tests were used, it is reported in the figure legends. Significance values correspond to p values, q values or FDR as follows: ns \geq 0.05, * <0.05, ** <0.01, *** <0.001, **** <0.0001. In the box plots, the upper whisker indicates the maximum value or 75th percentile +1.5 IQR; the lower whisker indicates the minimum value or 25th percentile -1.5 IQR.

ADDITIONAL RESOURCES

Data visualization tool generated <http://science.bostongene.com/tumor-portrait/>.