

IFN- γ -related mRNA profile predicts clinical response to PD-1 blockade

Mark Ayers,¹ Jared Lunceford,¹ Michael Nebozhyn,¹ Erin Murphy,¹ Andrey Loboda,¹ David R. Kaufman,¹ Andrew Albright,¹ Jonathan D. Cheng,¹ S. Peter Kang,¹ Veena Shankaran,² Sarina A. Piha-Paul,³ Jennifer Yearley,¹ Tanguy Y. Seiwert,⁴ Antoni Ribas,⁵ and Terrill K. McClanahan¹

¹Merck & Co. Inc., Kenilworth, New Jersey, USA. ²University of Washington, Seattle, Washington, USA. ³University of Texas MD Anderson Cancer Center, Houston, Texas, USA.

⁴University of Chicago, Chicago, Illinois, USA. ⁵UCLA, Los Angeles, California, USA.

Programmed death-1-directed (PD-1-directed) immune checkpoint blockade results in durable antitumor activity in many advanced malignancies. Recent studies suggest that IFN- γ is a critical driver of programmed death ligand-1 (PD-L1) expression in cancer and host cells, and baseline intratumoral T cell infiltration may improve response likelihood to anti-PD-1 therapies, including pembrolizumab. However, whether quantifying T cell-inflamed microenvironment is a useful pan-tumor determinant of PD-1-directed therapy response has not been rigorously evaluated. Here, we analyzed gene expression profiles (GEPs) using RNA from baseline tumor samples of pembrolizumab-treated patients. We identified immune-related signatures correlating with clinical benefit using a learn-and-confirm paradigm based on data from different clinical studies of pembrolizumab, starting with a small pilot of 19 melanoma patients and eventually defining a pan-tumor T cell-inflamed GEP in 220 patients with 9 cancers. Predictive value was independently confirmed and compared with that of PD-L1 immunohistochemistry in 96 patients with head and neck squamous cell carcinoma. The T cell-inflamed GEP contained IFN- γ -responsive genes related to antigen presentation, chemokine expression, cytotoxic activity, and adaptive immune resistance, and these features were necessary, but not always sufficient, for clinical benefit. The T cell-inflamed GEP has been developed into a clinical-grade assay that is currently being evaluated in ongoing pembrolizumab trials.

Introduction

Cancer cells have the ability to adapt to and avoid host immune checkpoints and surveillance as a means to promote their continued survival (1, 2). Although the exact molecular mechanisms of this adaptation are poorly understood, one of the key signaling pathways involves programmed death-1 (PD-1) and its ligands PD-L1 and PD-L2 (3–5). Results of recent clinical studies with anti-PD-1 monoclonal antibodies in patients with cancer have validated the concept that targeting the host immune checkpoint mechanism via PD-1 receptor blockade produces durable antitumor responses, resulting in improved progression-free survival (PFS) and overall survival (OS) (6). However, only a subset of patients benefit from PD-1 blockade. Elucidation of the underlying immunologic characteristics of the tumor

microenvironment associated with response and resistance will improve the identification of patients who will derive the most benefit from anti-PD-1 monotherapy and might reveal additional immunologic determinants that could be targeted in combination with PD-1 checkpoint blockade.

The mechanism by which the host immune system is capable of recognizing and eliminating dysplastic and neoplastic cells is the subject of ongoing investigation. Certain factors, such as exposure to ultraviolet radiation, smoking, and chronic viral infection, can damage critical genes involved in DNA replication and repair, resulting in abnormal cellular growth (7–9). In response, an orchestrated innate and adaptive antitumor immune response is initiated that may lead to the production of IFN- γ . IFN- γ is a key cytokine produced by activated T cells, as well as natural killer (NK) and NK T cells, in the tumor microenvironment, and it plays an important role in coordinating this process (10). However, the same IFN- γ signaling processes can ultimately induce feedback inhibition that compromises antitumor immunity (11). As part of this feedback loop, IFN- γ signaling enables the PD-1 signaling axis to become activated through direct upregulation of the ligands PD-L1 and PD-L2 in tumor, immune infiltrate, and stromal cells, which interact with PD-1 on tumor-infiltrating T cells to downregulate the cytotoxic response (12–14). In addition, IFN- γ can upregulate expression of other key immune suppressive molecules such as IDO1 within the tumor microenvironment (15). Tumor adaptation takes advantage of this delicate balance of positive and negative immune signaling factors, allowing the cancer to survive and progress.

Conflict of interest: M. Ayers, J. Lunceford, M. Nebozhyn, E. Murphy, A. Loboda, D.R. Kaufman, A. Albright, J.D. Cheng, S.P. Kang, J. Yearley, and T.K. McClanahan are employees of Merck & Co. Inc. S.P. Kang and T.K. McClanahan also disclose stock ownership in Merck & Co. Inc. M. Ayers, J. Lunceford, E. Murphy, A. Loboda, and T.K. McClanahan have a patent pending (system and methods for deriving gene signature biomarkers of response to PD-1 antagonists; PCT/US2015/064445), and S.P. Kang has a patent pending (MK-3475). V. Shankaran received a grant from Merck & Co. Inc. during the conduct of the study and has received grants from Amgen and Castle Biosciences. T.Y. Seiwert reports personal fees from Merck & Co. Inc. during the conduct of the study. A. Ribas has been a consultant for Merck & Co. Inc., with the honoraria paid to his institution.

Submitted: October 12, 2016; **Accepted:** May 11, 2017.

Reference information: *J Clin Invest.* 2017;127(8):2930–2940.

<https://doi.org/10.1172/JCI91190>.

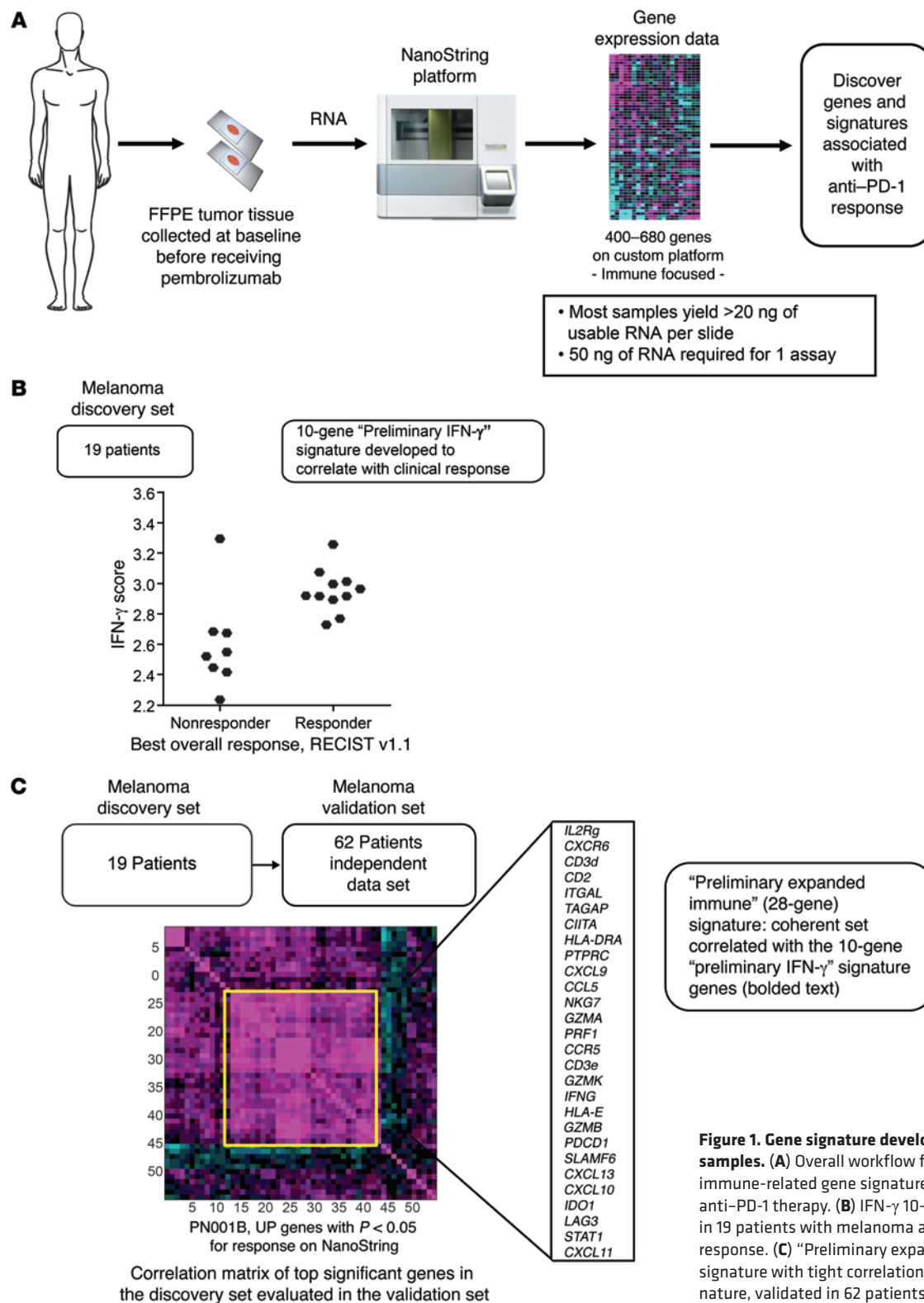


Figure 1. Gene signature development in melanoma samples. (A) Overall workflow for the development of immune-related gene signatures that predict response to anti-PD-1 therapy. (B) IFN- γ 10-gene signature evaluated in 19 patients with melanoma and association with response. (C) “Preliminary expanded immune” 28-gene signature with tight correlation to the IFN- γ 10-gene signature, validated in 62 patients with melanoma.

Recently, increased PD-L1 expression on the surface of tumor cells has been identified as an important determinant of response to immune therapies targeting the PD-1 axis (16). PD-L1 expression on immune cells in the tumor microenvironment has also been associated with treatment responses to such agents (17). Although the assessment of PD-L1 expression on tumor and immune cells

can be useful to predict clinical response to PD-1 checkpoint blockade, it offers only limited insight into the biology of the tumor-immune interface. In particular, PD-L1 expression might represent only a component of T cell-related biology that is relevant to a favorable tumor immune microenvironment. Newer genomic technologies can be used to evaluate complexities of tumor and

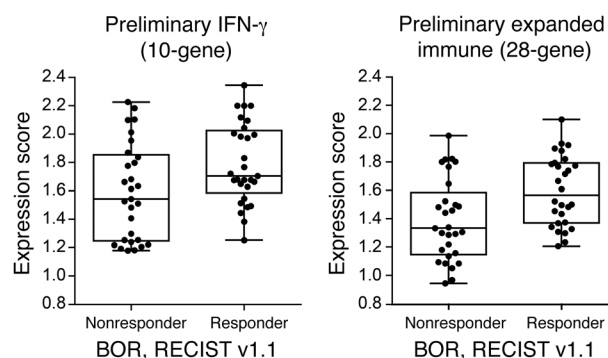


Figure 2. Box plots for the IFN- γ 10-gene and 28-gene expanded immune signatures and best overall response in 62 patients with melanoma with clinical outcomes under anti-PD-1 therapy.

host immune cell interactions within the tumor microenvironment, going beyond the measurement of single analytes such as PD-L1.

We specifically examined gene expression in the tumor microenvironment, using RNA isolated from formalin-fixed paraffin-embedded (FFPE) tumor tissue samples. Samples were obtained at baseline from patients undergoing treatment with pembrolizumab in clinical trials using multiple distinct tumor types. We report the rigorous, stepwise validation of the hypothesis that immune-related gene signatures can predict clinical response to PD-1 checkpoint blockade. Signatures related to IFN- γ signaling and activated T cell biology were initially delineated in a small pilot melanoma cohort, then confirmed and refined in a larger independent cohort of patients with melanoma. The cross-tumor predictive value of these signatures was demonstrated by testing in head and neck squamous cell carcinoma (HNSCC) and gastric cancer cohorts, followed by a modeling exercise to determine a final T cell-inflamed gene expression profile (GEP) and scoring algorithm that predicted response across 9 different cancer cohorts to arrive at a final signature, forming the basis of a clinical-grade assay for evaluation of clinical utility in select ongoing pembrolizumab clinical trials (18). The performance of this pan-tumor T cell-inflamed GEP compared favorably with that of immunohistochemistry (IHC) in an independent set of PD-L1-unselected patients with HNSCC and was shown to be robust despite any improvement using all genes on our profiling platform and all available data from the KEYNOTE-012 and KEYNOTE-028 clinical trials. Our data definitively confirm that a T cell-inflamed microenvironment, characterized by active IFN- γ signaling, cytotoxic effector molecules, antigen presentation, and T cell active cytokines, is a common feature of the biology of tumors that are responsive to PD-1 checkpoint blockade. Moreover, these data demonstrate that a focused set of genes can be used to identify this PD-1 checkpoint blockade-responsive biology and predict clinical response across a wide variety of tumor types.

Results

IFN- γ -related gene expression signatures predict response to PD-1 checkpoint blockade in melanoma. To define immune-related gene expression signatures associated with response to pembrolizumab, a set of baseline biopsies from 19 patients with metastatic melanoma enrolled in the KEYNOTE-001 study (NCT01295827, ClinicalTrials.gov) was used as a pilot data set. Pembrolizumab responders ($n = 8$) and nonresponders ($n = 11$) were identified using Response Evaluation Criteria In Solid Tumors, version 1.1 (RECIST v1.1) by central radiographic imaging vendor. A 1-sided t test was applied to rank top genes associated with responders and nonresponders, using a custom panel of 680 tumor- and immune-related genes on the NanoString nCounter platform (NanoString Technologies Inc.) (Figure 1A and data not shown). It was apparent that many of the top-ranked genes were directly linked to IFN- γ signaling and showed correlation with the expression of IFN- γ . A 10-gene “preliminary IFN- γ ” signature (*IFNG*, *STAT1*, *CCR5*, *CXCL9*, *CXCL10*, *CXCL11*, *IDO1*, *PRF1*, *GZMA*, and *MHCII HLA-DRA*) was constructed that was able to separate responders and nonresponders to anti-PD-1 therapy (pembrolizumab) among the 19 pilot data patients with melanoma (Figure 1B).

These findings were confirmed using expression data from an additional cohort of 62 patients with metastatic melanoma treated with pembrolizumab in KEYNOTE-001, of whom 51 were considered evaluable for objective response. Before unblinding of the clinical outcome information for these patients, genes from the preliminary IFN- γ signature were correlated with other genes on the custom 680-gene NanoString panel that had a nominal P value less than 0.05 for positive association with objective response and/or PFS on the initial set of 19 melanoma tumors. From these data, a larger expanded set of genes showing correlation with the preliminary IFN- γ signature was evident, and a 28-gene set, referred to as the “preliminary expanded immune” signature, was generated (Figure 1C). This 28-gene set encompassed genes related to cytolytic activity (e.g., granzyme A/B/K, *PRF1*), cytokines/chemokines for initiation of inflammation (*CXCR6*, *CXCL9*, *CCL5*, and *CCR5*), T cell markers (*CD3D*, *CD3E*, *CD2*, *IL2RG* [encoding IL-2R γ]), NK cell activity (*NKG7*, *HLA-E*), antigen presentation (*CIITA*, *HLA-DRA*), and additional immunomodulatory factors (*LAG3*, *IDO1*, *SLAMF6*) (Figure 2 and Table 1). These 2 preliminary signatures were then tested using prespecified hypotheses for their association with clinical outcome following treatment with pembrolizumab. Nominal 1-sided P values less than 0.05 were observed for both

Table 1. Independent results evaluating correlation between the gene signatures and response in the validation set in 62 patients with melanoma with clinical outcomes under anti-PD-1 therapy^A

Signature	BOR by RECIST ($n = 51$)	PFS by RECIST ($n = 62$)	OS ($n = 62$)
Preliminary IFN- γ signature	$P = 0.047$	$P = 0.016$	$P = 0.090$
Preliminary expanded immune signature	$P = 0.027$	$P = 0.015$	$P = 0.105$

Nominal 1-sided P values were derived from logistic regression (for BOR per RECIST v1.1) or Cox regression (for PFS and OS). ^ADevelopment of the expanded immune signature was performed in an unsupervised manner by individuals blinded to response data. BOR, best overall response; PFS, progression-free survival; OS, overall survival.

Table 2. IFN- γ and expanded immune gene signatures

IFN- γ	Expanded immune gene signature	
<i>IDO1</i>	<i>CD3D</i>	<i>IL2RG</i>
<i>CXCL10</i>	<i>IDO1</i>	<i>NKG7</i>
<i>CXCL9</i>	<i>CIITA</i>	<i>HLA-E</i>
<i>HLA-DRA</i>	<i>CD3E</i>	<i>CXCR6</i>
<i>STAT1</i>	<i>CCL5</i>	<i>LAG3</i>
<i>IFNG</i>	<i>GZMK</i>	<i>TAGAP</i>
	<i>CD2</i>	<i>CXCL10</i>
	<i>HLA-DRA</i>	<i>STAT1</i>
	<i>CXCL13</i>	<i>GZMB</i>

signatures for best overall response (BOR) and PFS (Table 1), confirming their predictive value in metastatic melanoma. These data indicate that IFN- γ signaling and the associated biology of T cell cytolytic activity, antigen presentation, and chemokine production are important components of a PD-1 checkpoint blockade-responsive immune microenvironment in melanoma.

Confirmatory testing and signature refinement across multiple cancer types demonstrates a common biology of responsiveness to PD-1 checkpoint blockade. Because activated Th1 and cytotoxic T cells are the likely direct effectors of checkpoint blockade-mediated tumor regression, we hypothesized that measures of this common biology would predict response to anti-PD-1 therapy regardless of tissue of origin. We therefore asked whether the biology of responsiveness to PD-1 checkpoint blockade that was observed in melanoma would extend to unrelated tumor types. After confirming that the 2 preliminary signatures in melanoma showed statistically significant associations with clinical outcome, both signatures were refined by removal of genes that did not reach a nominal 1-sided *P* value less than 0.05 for positive association with either objective response or PFS in the confirmatory set of 62 melanoma tumors, yielding refined “IFN- γ (6-gene)” and “expanded immune (18-gene)” signatures. Subsequently, as data from the KEYNOTE-012 (NCT01848834) study became available, these signatures were tested in biopsy specimens from pembrolizumab-treated patients with HNSCC or gastric cancer. RNA expression profiling data were available for tumors from 40 patients with HNSCC and were evaluable for objective response, with an objective response rate of 22.5%. There were 33 such patients with gastric cancer, with 22.2% responding. The refined signatures (Table 2) both showed statistically significant associations with improved clinical outcome at *P* < 0.05 for objective response and PFS for HNSCC (19) and at *P* < 0.10 and *P* < 0.05 for objective response and PFS, respectively, in gastric cancer (Figure 3, A and B, and Table 3). Receiver operating characteristics (ROCs) for response status over the range of the signature scores demonstrated the high discriminatory ability of the signatures (Figure 3, C and D). Areas under the ROC curves and their 95% CIs were 0.80 (0.61–0.95) for HNSCC (19) and 0.66 (0.47–0.83) for gastric cancer. To illustrate the potential clinical usefulness of these signatures to predict response to pembrolizumab, the Youden index-based cutoff values (20) for the IFN- γ 6-gene signature were determined and were similar for the 2 cancer types (1.882 HNSCC; 1.856 gastric). For

HNSCC, the associated positive predictive value (PPV, response rate above the cutoff) was 40.0% and the negative predictive value (NPV, nonresponse rate below the cutoff) was 95.0% (19). For gastric cancer, the PPV and NPV were 45% and 92%, respectively. The percentage of patients with tumors above these cutoffs was 50% for HNSCC and 61% for gastric cancer. These results validate the observations in the melanoma cohort and point to a common gene expression pattern in pretreatment tumor biopsy specimens that predicts post-treatment clinical outcomes to anti-PD-1 therapy across multiple tumor types. Moreover, these results suggest a potential for a high discriminatory value of IFN- γ -related gene signatures, enriching response rates to pembrolizumab (i.e., increasing PPV) while preserving a high NPV. However, we note that any formal evaluation of clinical utility depends on cutoffs selected for implementation of a clinical-grade diagnostic device in developmental trials and is dependent on the goals of those trials.

Having achieved proof-of-concept that common sets of IFN- γ and T cell-associated inflammatory genes can predict responsiveness to PD-1 blockade across different tumor types, we undertook a signature selection exercise with a larger and more diverse data set spanning 9 cancer types from pembrolizumab-treated patients in the KEYNOTE-012 (bladder, gastric, HNSCC, and triple-negative breast cancer) and KEYNOTE-028 (NCT02054806) (anal canal, biliary, colorectal, esophageal, and ovarian cancer) studies with a goal to identify a final pan-cancer predictive biomarker suitable for development as a diagnostic clinical test for use in pembrolizumab clinical trials. Penalized logistic regression was used to derive a final set of 18 genes, with weighting determined by the regression coefficients. Figure 4 shows a heatmap of unsupervised clustering of the data across 9 tumor types for these 18 genes, revealing the coordinated pattern of expression for many of the 18 genes in the T cell-inflamed GEP. All genes in the signature, except CD276, are positively associated with improved clinical outcome, as can be seen by the increased frequency of responders (denoted by “R” in Figure 4) in the region with generally larger expression levels of the genes. Across the set of patients in Figure 4, the Pearson correlation between the 18-gene score and the IFN- γ 6-gene signature score was 0.89. Within each cancer type,

Table 3. Correlation of signatures with BOR and PFS for HNSCC (43 total patients) and gastric cancer (33 patients) cohorts of KEYNOTE-012

Signature	Nominal 1-sided <i>P</i> value ^A	
	BOR ^B	PFS
Head and neck cohort	<i>n</i> = 40	<i>n</i> = 43
IFN- γ signature (6 genes)	0.005 ^C	<0.001 ^C
Expanded immune signature (18 genes)	0.015	<0.001
Gastric cohort	<i>n</i> = 33	<i>n</i> = 33
IFN- γ signature (6-gene)	0.077	0.032
Expanded immune signature (18-gene)	0.062	0.049

^AFrom logistic or Cox regression for overall response and PFS, respectively, with signature scores as a continuous variable. ^BBOR and PFS as assessed by investigator. ^CFrom Seiwert et al. (19). BOR, best overall response; PFS, progression-free survival.

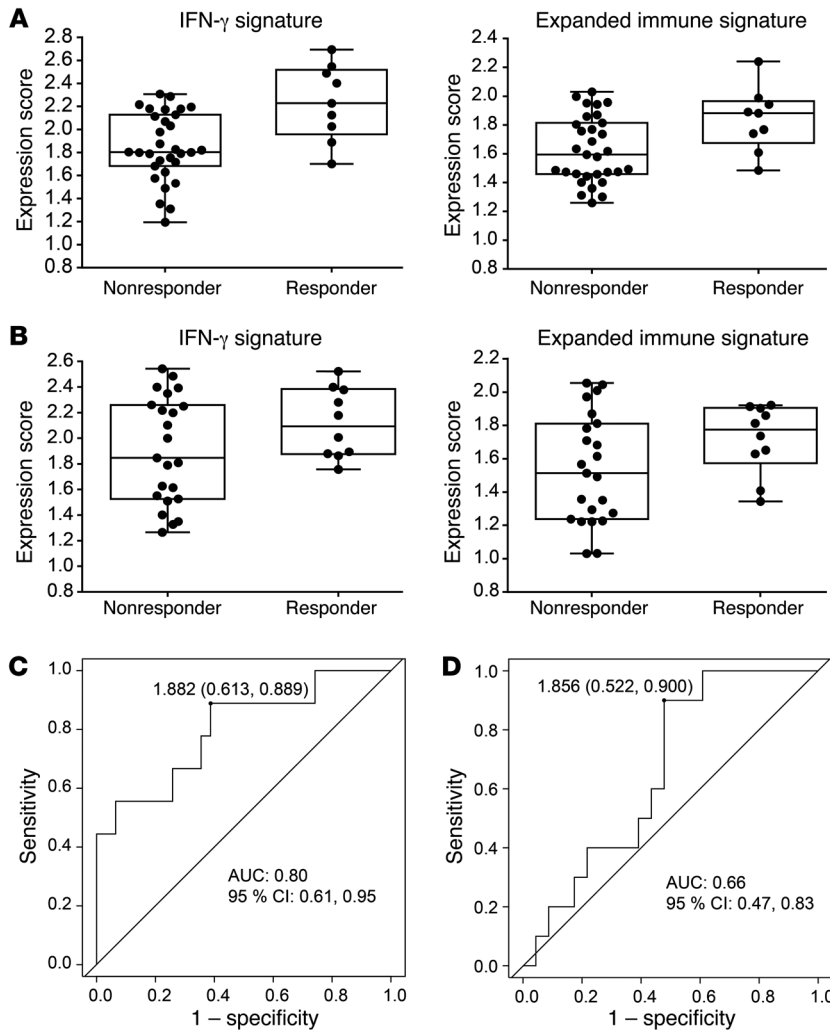


Figure 3. Confirmatory testing and signature refinement across multiple cancer types. (A and B) Confirmatory analyses of the IFN- γ and expanded immune signature scores for the HNSCC (43 total patients) (A) and gastric cancer (33 patients) (B) cohorts of KEYNOTE-012. (C and D) ROC curves of sensitivity and specificity for the HNSCC (C) and gastric cancer (D) cohorts of KEYNOTE-012.

that a multivariate approach using gene expression might offer benefits in sensitivity to detect responders to anti-PD-1 therapies.

Subsequent to the derivation of the final 18-gene signature, as a cross-check on the performance of the signature, the original KEYNOTE-012 and partial KEYNOTE-028 data used to train the signature were augmented with data of patients with additional cancers from KEYNOTE-028, roughly doubling the size of the original training data set, and an additional cross-validation exercise was conducted. These pooled data were randomly divided multiple times into two-thirds training and one-third testing, with elastic net predictors trained in the two-thirds portion of the data and the resultant area under the ROC curve of the new predictor compared with that of the 18-gene T cell-inflamed GEP in the holdout one-third of the data. No evidence of improvement was found (data not shown).

GEPs suggest discrete biologies of nonresponse to PD-1 blockade. Across multiple tumor types, we observed a general pattern of a lack of objective

response in patients whose tumors showed low expression levels across the genes, assumed to represent tumors without a T cell-inflamed phenotype. However, another category of nonresponders demonstrated clear evidence of IFN- γ and associated T cell inflammatory gene expression at baseline (Figure 6, A and B). Plotting the T cell-inflamed GEP compared with PFS yielded a similar observation (Figure 7). Patients whose tumors had scores less than -0.3 generally showed rapid disease progression, whereas a broad spectrum of progression times was observed for higher scores (Figure 7). These data suggest that the presence of a T cell-inflamed phenotype is necessary but not sufficient for clinical response to PD-1 checkpoint blockade. Moreover, these observations raise the possibility that distinct resistance mechanisms might be in play in tumors that lack T cell inflammation, compared with those that show evidence of an activated T cell infiltrate yet still show a lack of clinical response.

the area under the ROC curve was greater than 0.5, with an average area under the ROC curve of 0.75 across the 9 indications used to fit the model (data not shown). These data support the tumor type-independent applicability of a T cell-inflamed GEP that captures the biology of a T cell-inflamed microenvironment common to PD-1 checkpoint blockade responders.

To further evaluate the predictive usefulness of the 18-gene T cell-inflamed GEP, a PD-L1-unselected cohort of 96 patients with HNSCC from KEYNOTE-012 who were not involved in prior testing or the training of the signature was used to validate its performance in comparison with PD-L1 IHC. The 1-sided P value for association between the score and objective response was less than 0.001. Figure 5A shows a heatmap of the expression pattern for the 18 genes in these 96 patients with HNSCC, and Figure 5B shows the ROC curve for the T cell-inflamed GEP and PD-L1 expression by IHC, in which percentage staining using the number of PD-L1-positive cells (tumor cells, macrophages, lymphocytes) over total tumor cells was used as the scoring system. The area under the ROC curve was 0.75 for the T cell-inflamed GEP and 0.65 for PD-L1 by IHC. Although the all-comers HNSCC cohort was not sized/powerful for comparisons between areas under the ROC curves, a test comparing the areas under the ROC curves for equality had a 2-sided P value of 0.119. These results suggest

that a multivariate approach using gene expression might offer benefits in sensitivity to detect responders to anti-PD-1 therapies.

Discussion

Multigene immune signatures represent a robust means of capturing a complex, T cell-inflamed phenotype necessary for the clinical activity of PD-1-/PD-L1-directed monoclonal antibodies. In this proof-of-concept study, we assayed pretreatment tumor biopsy specimens, capturing broad measures of immune-related

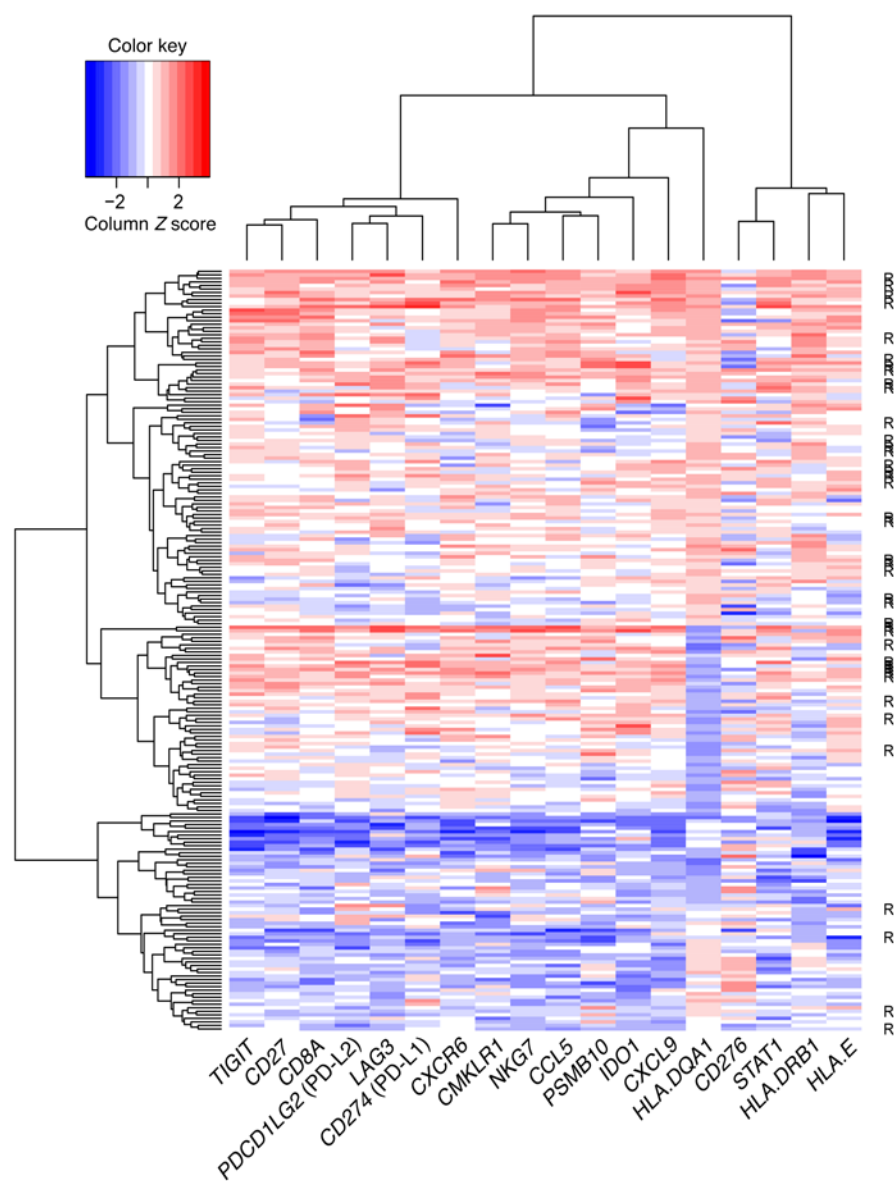


Figure 4. Heatmap for the final 18-gene T cell-inflamed GEP for 216 tumors from patients in KEYNOTE-012 and KEYNOTE-028 considered evaluable for objective response. Rows represent patients and columns genes. Expression levels have been standardized (centered and scaled) within columns for visualization. The “R” on the right side indicates whether the patient was a responder (by central imaging vendor in KEYNOTE-012 and by investigator assessment in KEYNOTE-028). The rows and columns have been grouped using unsupervised clustering.

gene expression using a straightforward, sensitive, and quantitative method compatible with tissue-limiting FFPE clinical samples. Through a rigorous, multistep validation process, we derived an 18-gene T cell-inflamed GEP that predicted response to pembrolizumab across multiple solid tumors. Our results suggest that immune GEPs such as the 18-gene profile, when assayed on a platform such as the NanoString nCounter platform, can define tumor type-independent dimensions of the tumor microenvironment relevant to predicting clinical outcome for agents targeting the PD-1/PD-L1 signaling pathway. Moreover, our data showed that such signatures might perform favorably compared with PD-L1 IHC in PD-L1-unselected populations. The 18-gene profile has undergone analytical validation on the NanoString platform and is under development as a clinical-grade diagnostic device, employed in a set of ongoing pembrolizumab trials (18).

Gene expression signatures that predicted clinical response to PD-1 checkpoint blockade were identified and refined in a stepwise manner through a number of independent tumor cohorts repre-

sented 10 different tumor types. A set of predictive gene signatures was initially defined and evaluated in baseline tumor biopsy specimens in patients with metastatic melanoma undergoing treatment with pembrolizumab and was subsequently shown, via rigorous independent statistical testing, to be associated with improved clinical outcome for patients with HNSCC or gastric cancer after pembrolizumab therapy. Objective response and PFS times were used as the focal clinical endpoints in this testing given that the source trials were all single-arm studies. The prognostic relationship between these signatures and survival under standard of care (chemotherapy) is the subject of ongoing investigation, and a full understanding of the improvement in overall survival (OS) relative to nonimmunotherapies will require randomized studies; however, the association between immune signatures and durable objective responses to anti-PD-1 treatments, which are responses uncharacteristic of standard of care, is expected to translate into clear increases in expected survival for anti-PD-1 therapy compared with standard of care in patients whose tumors exhibit the T cell-inflamed phenotype.

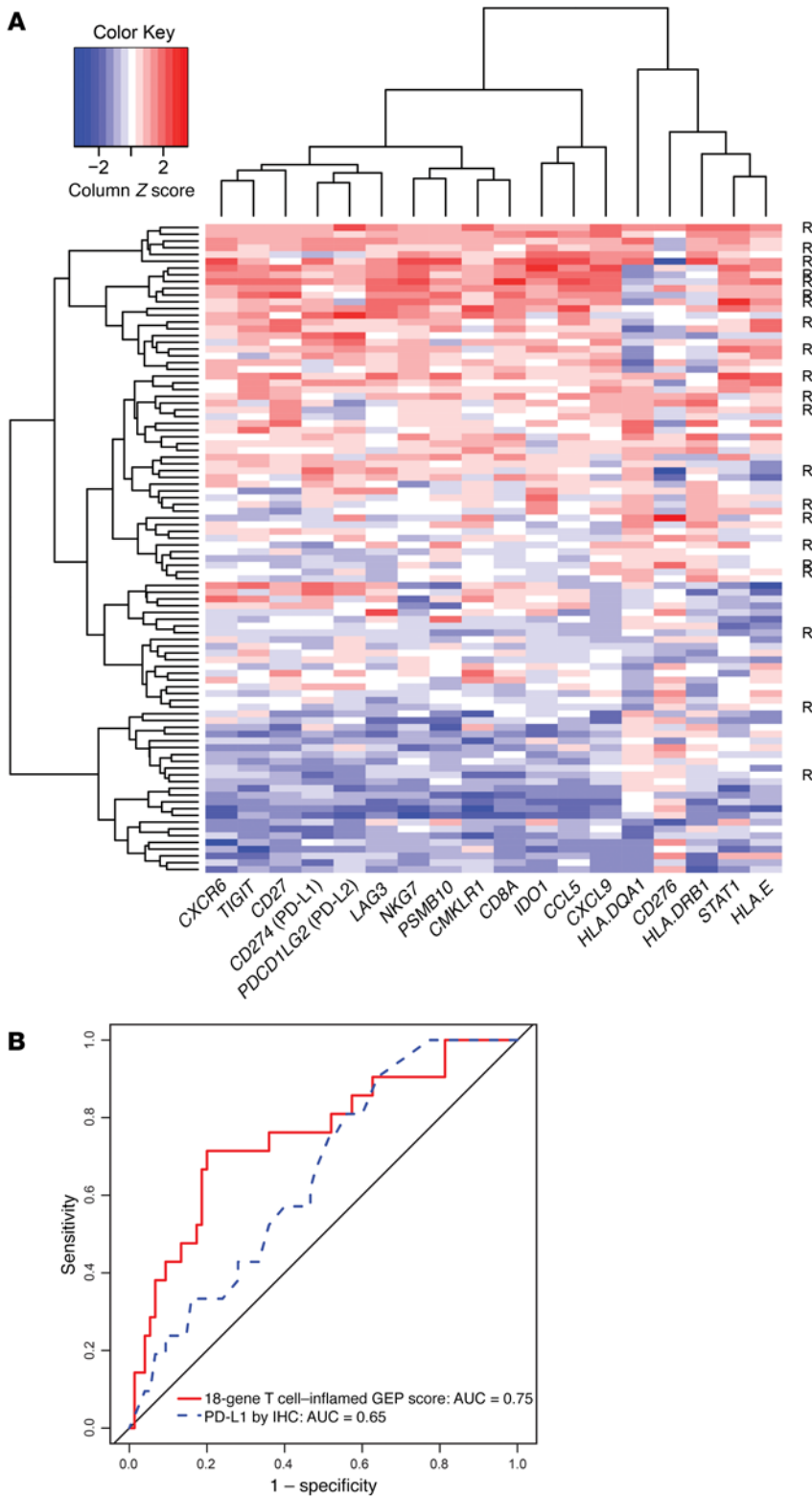


Figure 5. Validation of the final 18-gene T cell-inflamed GEP. (A) Heatmap of 18-gene T cell-inflamed GEP in 96 PD-L1-unselected patients with HNSCC from KEYNOTE-012. Expression levels have been standardized (centered and scaled) within columns for visualization. The “R” on the right side indicates whether the patient was a responder (by central imaging vendor). The rows and columns have been grouped using unsupervised clustering. (B) ROC curves comparing final 18-gene score with expression of PD-L1 as measured by IHC on tumor and inflammatory cells for a cohort of 96 PD-L1-unselected patients with HNSCC from KEYNOTE-012 considered evaluable for objective response by central imaging vendor.

large, independent cohort of PD-L1-unselected patients with HNSCC. In this cohort, the area under the ROC curve for the T cell-inflamed GEP compared favorably with PD-L1 IHC (0.75 vs. 0.65). Although both PD-L1 and the T cell-inflamed GEP reflect IFN- γ -inducible biology, these results signify the more robust nature of an assay that simultaneously measures multiple microenvironmental features, rather than a single analyte. Likewise, these data provide a much higher level of statistical validation across multiple tumor types than has previously been reported for single genes associated with the IFN- γ signaling pathway and response to PD-1 blockade (17), indicating that integrated multigene signatures might represent more robust predictive biomarkers than single genes.

For clinical implementation, a cutoff on the T cell-inflamed GEP must be selected to identify the subgroup in which meaningful improvement in response can be demonstrated relative to standard of care. The bar for enrichment may be higher in the first-line than in the second- and third-line settings. The latter may not require extremely high PPV to demonstrate convincing activity of the drug in patients with advanced metastatic disease who have limited or no treatment options.

Inspection of the T cell-inflamed GEP suggests that a subset of tumors from diverse tissues of origin exhibit at baseline an ongoing adaptive Th1 and cytotoxic CD8⁺ T cell response, including IFN- γ signaling, cytolytic activity, antigen presentation, and T cell trafficking, as well as inhibitory mechanisms that are evident in T cell homeostasis (Figure 8).

These data extend the observations of previous reports showing that CD8⁺ T lymphocytes and T cell active chemokines and T cell counterinhibitory molecules such as PD-1/PD-L1 and IDO1 are coregulated within tumors (4, 15, 21, 22) and confirm that response to anti-PD-1 blockade occurs primarily in patients with such a preexisting, intratumoral T cell adaptive immune response. Furthermore, these data indicate

The final 18-gene profile was derived through a cross-validated penalized regression modeling strategy in a large cohort of pembrolizumab-treated patients across 9 different tumor types. Although many of these patient cohorts had undergone enrichment for PD-L1 expression by pre-enrollment IHC screening, the predictive value of the signature was prospectively validated in a

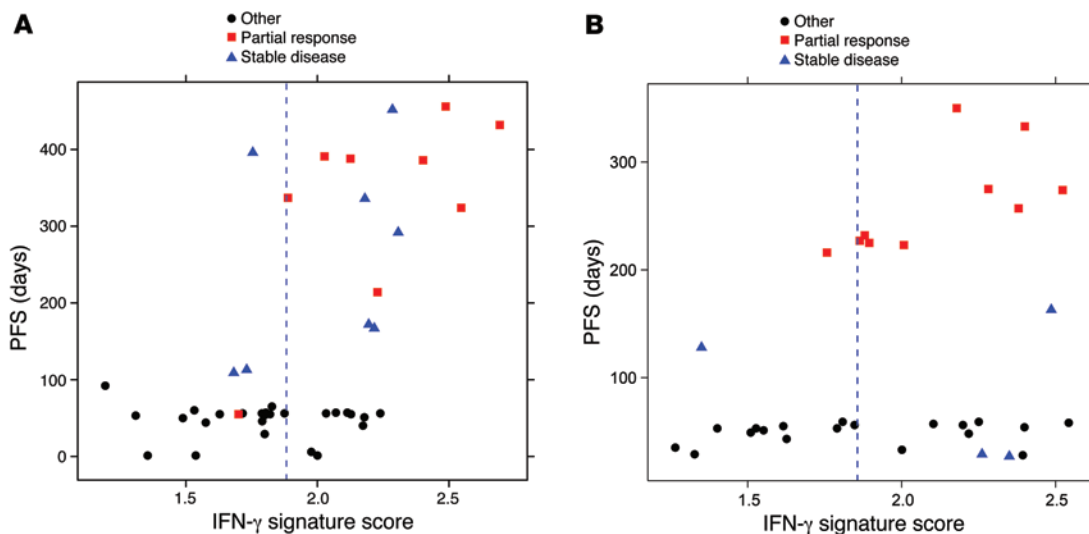


Figure 6. Relationship between increases in IFN- γ immune-related signature score and PFS in response to anti-PD-1 therapy for the HNSCC and gastric cancer cohorts of KEYNOTE-012. (A) Relationship in the HNSCC cohort (43 total patients). (B) Relationship in the gastric cancer cohort (33 patients). The cutoff associated with the Youden index is displayed in each figure.

that the PD-1 pathway likely represents a common mechanism for restraining such antitumor responses through the IFN- γ -mediated induction of PD-L1 and PD-L2 expression. These correlative gene signatures represent a novel method for capturing the complexity of the dynamic immune response to a tumor by distinguishing between tumors with preexisting inflammatory components and noninflamed tumors, a classification that is likely to be of high clinical relevance (23, 24).

The spectrum of genes associated with clinical response to PD-1 checkpoint blockade suggests a complex biology of immune cell homeostasis in anti-PD-1-responsive tumors. The production of IFN- γ by activated T cells, and perhaps other cells such as NK and NK T cells, activates a host of key downstream signaling molecules such as STAT1 and CMKLR1 in dendritic cells and macrophages. These cells in turn produce chemokines/chemoattractants, including CCL5 and CXCL9, which recruit additional CD8⁺ T cells. T cells upregulate activation-induced costimulatory receptors such as CD27 and produce an array of effector molecules, including IFN- γ , perforin, and granzymes. IFN- γ can induce upregulation of HLA molecules and other components of the immunoproteasome and antigen-presenting machinery such as PSMB10. However, IFN- γ also upregulates a host of checkpoint inhibitors, such as PD-L1 and PD-L2, on the surface of macrophages and dendritic and tumor cells. Other checkpoint molecules, such as IDO1, TIGIT, and LAG3, might also be homeostatically upregulated by T cell activation and IFN- γ signaling, and cooperate with PD-1 to restrain the antitumor immune response. The association of genes such as *HLA-E* and *NKG7* with clinical response suggests that additional aspects of NK cell biology or noncanonical T cell antigen recognition might be involved in the T cell-inflamed microenvironment. Overall, the genes identified suggest that multiple immune cell types contribute to a T cell-inflamed but immune checkpoint-restrained tumor microenvironment that can be successfully modulated by PD-1 checkpoint blockade.

Furthermore, our data suggest that nonresponders to anti-PD-1 therapy fall into 2 broad categories: those who lack a preexist-

ing antitumor immune response and those in whom an antitumor immune response has been generated, with associated infiltration of activated T cells into the tumor microenvironment. The reason for lack of clinical response in this subset of inflamed tumors is not clear, but it might be a result of effects of immunosuppressive pathways other than PD-1 within the tumor microenvironment, which could be T cell intrinsic or extrinsic. The observation that immunosuppressive molecules such as IDO1 and LAG3 are overexpressed in inflamed tumors suggests that increased anti-PD-1 efficacy in this class of tumors might be observed when combined with immune modulators targeting other suppressive pathways. For tumors lacking the inflamed phenotype, distinct strategies

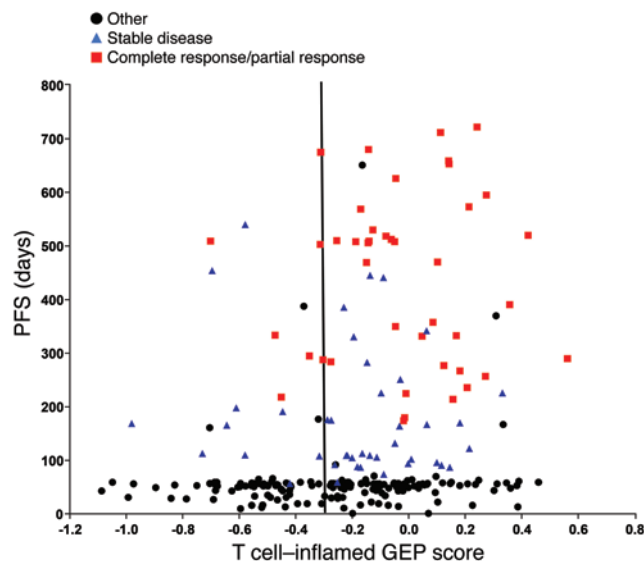
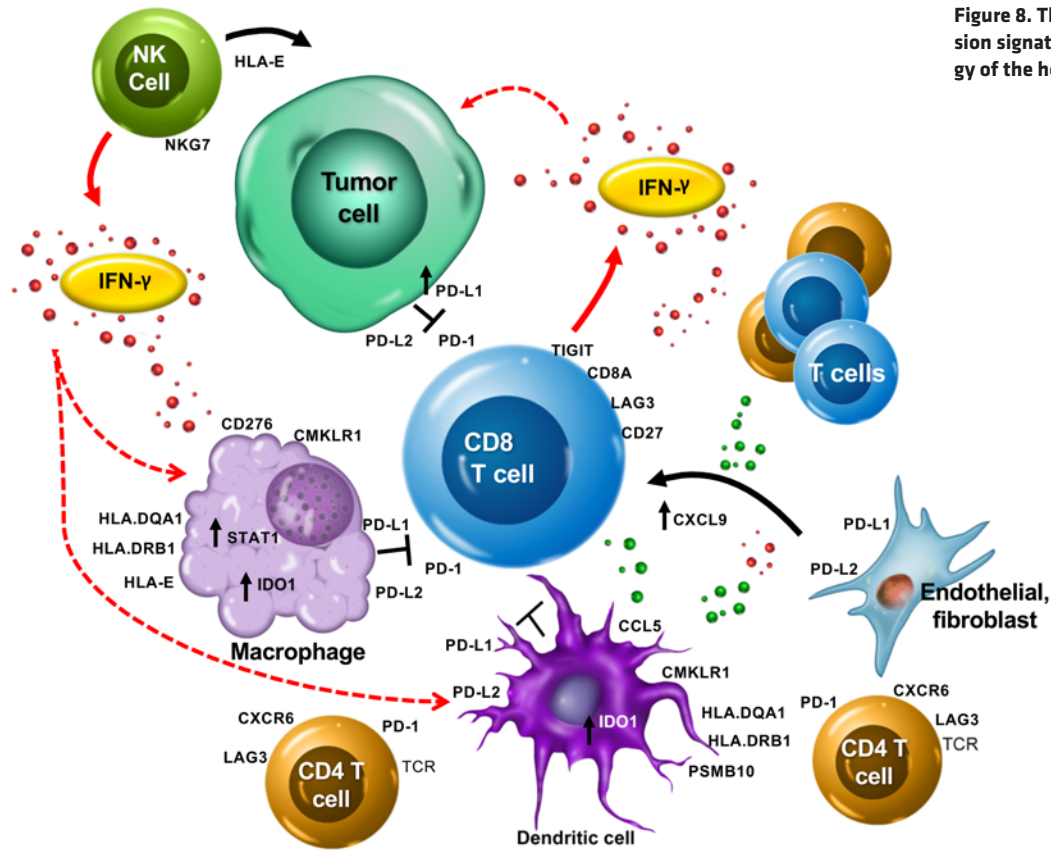


Figure 7. PFS time versus T cell-inflamed GEP score in 244 patients from KEYNOTE-012 and KEYNOTE-028 for the 9 cancer cohorts used to determine the T cell-inflamed GEP.

Figure 8. The T cell-inflamed gene expression signature highlights the complex biology of the host immune microenvironment.



to stimulate a host immune response against the tumor may be necessary. Further work is needed to elucidate the discrete mechanisms that inhibit T cell trafficking to T cell-poor tumors, which might include tumor cell-intrinsic mechanisms (e.g., WNT/ β -catenin pathway activation) (25), failure of T cell priming (amenable to reversal by tumor-specific vaccines, oncolytic viruses, or adoptive T cell therapy), or the activity of other cells in the tumor microenvironment, such as stromal cells (26) or vascular endothelial cells (27). Further studies using gene expression profiling might represent a powerful approach to dissecting the discrete biologies of resistance to PD-1 checkpoint blockade, potentially revealing new targets for converting anti-PD-1-resistant tumors to a state of responsiveness.

Methods

Quality assessment of clinical FFPE tissue specimens. The KEYNOTE-001, KEYNOTE-012, and KEYNOTE-028 trials were nonrandomized, open-label studies characterizing the safety, tolerability, and clinical activity of pembrolizumab in patients with locally advanced or metastatic cancers (KEYNOTE-001: melanoma, non-small cell lung cancer; KEYNOTE-012: HNSCC, gastric cancer, triple-negative breast cancer, bladder cancer; KEYNOTE-028: anal canal, biliary, colorectal, esophageal, and ovarian cancer). Pretreatment clinical specimens were collected as specified in the clinical protocols and consisted of FFPE blocks and/or cut tissue on slides. The melanoma cohort from KEYNOTE-001 and the HNSCC-expansion cohort of KEYNOTE-012 were all-comers cohorts (independent of PD-L1 expression), whereas the remaining cohorts from KEYNOTE-012 and all cohorts from KEYNOTE-028 were screened using an IHC assay for patients whose

tumors were positive ($\geq 1\%$ of cells) for PD-L1 expression either in tumor cells or in inflammatory cells of the surrounding stroma. When only an FFPE block was available, sections were cut at 5 μm onto positively charged slides. One cut section from each tissue sample was routinely stained with H&E and coverslipped for assessment of sample quality. Each H&E-stained slide was reviewed by a pathologist to evaluate the adequacy of tumor representation, the quality of tissue preservation, and whether significant artifacts relating to fixation, processing, or prefixation tissue handling were present. Specimens containing no tumor tissue or very minimal tumor tissue were excluded from analysis. The number of slides needed for RNA isolation from any given tissue sample was based on a combination of sample size on the slide and the proportion of the sample that consisted of tumor tissue; typically, RNA was isolated from 2–4 slides per case.

A PD-L1 clinical trial assay (PD-L1 IHC 22C3 pharmDx assay; Agilent Technologies) was subsequently developed and approved as a PD-L1 IHC companion diagnostic for use in non-small cell lung cancer in the United States (28, 29). The approved assay was used to stain the 96 patient HNSCC samples under the staining protocol described in the instructions of the commercial assay. Staining was scored using the number of PD-L1-positive cells (tumor cells, macrophages, lymphocytes) over total tumor cells, expressed as a percentage.

RNA isolation and gene expression analysis. Total RNA was isolated from 5- μm -thick FFPE sections of tumors fixed on positively charged slides using the Ambion RecoverAll kit for RNA isolation from FFPE tissue (Thermo Fisher Scientific) following the manufacturer's protocols. RNA was quantified using the NanoDrop ND1000 spectrophotometer (Thermo Fisher Scientific). Gene expression analysis was conducted on the NanoString nCounter gene expression

platform (NanoString Technologies). A custom code set consisting of a 680-gene panel related to T cell biology, immune regulation, and cellular markers of tumor-infiltrating lymphocytes and tumor-associated macrophages was used. Per sample, 50 ng of total RNA in a final volume of 5 μ l was mixed with a 3' biotinylated capture probe and a 5' reporter probe tagged with a fluorescent barcode from the custom gene expression code set. Probes and target transcripts were hybridized overnight at 65°C for 12–16 hours per the manufacturer's recommendations. Hybridized samples were run on the NanoString nCounter preparation station using the high-sensitivity protocol, in which excess capture and reporter probes were removed and transcript-specific ternary complexes were immobilized on a streptavidin-coated cartridge. The samples were scanned at maximum scan resolution on the nCounter Digital Analyzer.

Quantile normalization. Gene expression data for each individual sample were normalized by quantile normalization. Gene counts collected from the NanoString scanner were used as input variables with a reference distribution generated using a pool of counts from all samples and 680 genes (excluding data from positive and negative control probes). After performance of quantile normalization, a \log_{10} transformation was applied, and signature scores were calculated by averaging of the included genes for the IFN- γ (6-gene) and expanded immune (18-gene) signatures.

Housekeeping normalization. In anticipation of moving to a more limited set of genes that could serve as the basis of a focused clinical trial assay (a context not amenable to quantile normalization), a set of 11 genes showing low variance across a set of banked tumor samples from a variety of cancer types was identified and used to form a normalizing constant for each patient. The \log_{10} count of each gene on the platform was normalized by subtracting of the arithmetic mean of the \log_{10} counts of the housekeeping genes.

Statistics. Logistic regression modeling was used to conduct the hypothesis testing associated with best overall response (BOR), and a Cox model was used for testing of PFS and OS. PFS was defined as the time from the start of treatment to documented evidence of progressive disease or death. OS was defined as time from treatment initiation to death from any cause. For BOR, a responder was defined as a patient with "complete response" or "partial response" as determined by RECIST v1.1 criteria, assessed by central imaging vendor for melanoma and by investigator review for the HNSCC and gastric cohorts (centrally reviewed data were not as mature at the time of analysis). One-sided *P* values were calculated according to the hypothesized positive association between the immune signatures and improved clinical outcome under treatment. Calculation of the area under the ROC curve was used as a measure of discriminatory ability for the signature scores. The Youden index, a summary measure of the ROC curve (20), was used as an agnostic method for choosing an "optimal" cutoff on the signature scores to illustrate potential clinical usefulness. The NanoString platform was used as a discovery research assay rather than a clinical trial-grade assay.

Nominal *P* values were reported for signature testing. The gene signatures developed from the melanoma cohort were independently confirmed by prespecified hypothesis testing in 2 additional cancer types: HNSCC and gastric cancer. Afterward, a larger effort to define a focused gene signature and an optimal scoring for that signature was undertaken by combining of early versions of data from KEYNOTE-012 and KEYNOTE-028. A total of 220 patients with measurable disease

by investigator and RNA expression profiling from KEYNOTE-012 (25 bladder, 40 HNSCC, 30 triple-negative breast cancer, and 33 gastric) and KEYNOTE-028 (19 anal canal, 13 biliary tract, 18 colorectal, 18 esophageal, and 24 ovarian) were available at the time and were used in an elastic net penalized regression model (30) for objective response to select a final set of genes and their weighting. Fivefold cross-validation was used to select the penalty parameters, and 10 different random partitions of the data into 5-folds were used. The average of the estimated penalties from these 10 cross-validation runs was used as the final set of penalty parameters to determine the final gene set and weights. The penalized regression models adjusted for cancer type, performance status, and the interaction between cancer type and performance status. The final values of the regression coefficients on the genes that were not zeroed out by the penalty terms were used as the weights in the computation of the T cell-inflamed GEP score for the resulting 18 genes, not to be confused with the 18 genes of the expanded immune signature. The score was computed as the weighted sum of the housekeeping normalized values of the 18 genes.

Study approval. The original studies were conducted in accordance with the Declaration of Helsinki and the International Conference on Harmonization Good Clinical Practice guidelines and approved by relevant regulatory and independent ethics committees from each study's institution. All patients provided written informed consent before study entry.

Author contributions

MA participated in the concept, design, or planning of the study, acquired the data, and drafted the manuscript. JL analyzed the data, interpreted the results, and drafted the manuscript. MN analyzed the data and drafted the manuscript. EM acquired and analyzed the data. AL analyzed the data and interpreted the results. DRK analyzed the data, interpreted the results, and drafted the manuscript. AA acquired the data. JDC participated in the concept, design, or planning of the study, acquired the data, and interpreted the results. SPK acquired the data. VS interpreted the results. SAPP provided study materials and patients and interpreted the results. JY acquired the data. TYS provided study materials and patients, acquired the data, and interpreted the results. AR participated in the concept, design, or planning of the study, interpreted the results, and drafted the manuscript. TKM participated in the concept, design, or planning of the study, acquired and analyzed the data, interpreted the results, and drafted the manuscript. All authors critically reviewed iterations of the manuscript and approved the final draft for submission.

Acknowledgments

The authors thank Razvan Cristescu (Merck & Co. Inc.) for his contributions in the development of the manuscript, Robert Kastelein (Merck & Co. Inc.) for critical review of the manuscript, and Matthew Grzywacz and the ApotheCom oncology team (Yardley, Pennsylvania, USA) for assistance with manuscript editing (funded by Merck & Co. Inc.). Funding for this analysis was provided by Merck & Co. Inc.

Address correspondence to: Terrill K. McClanahan, Profiling & Expression, Biologics Discovery, Merck Research Laboratories, 901 California Avenue, Palo Alto, California 94304, USA. Phone: 650.496.1129; E-mail: terri.mcclanahan2@merck.com.

1. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer*. 2012;12(4):252–264.
2. Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. *Immunity*. 2013;39(1):1–10.
3. Keir ME, Butte MJ, Freeman GJ, Sharpe AH. PD-1 and its ligands in tolerance and immunity. *Annu Rev Immunol*. 2008;26:677–704.
4. Tumei PC, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature*. 2014;515(7528):568–571.
5. Sharpe AH, Wherry EJ, Ahmed R, Freeman GJ. The function of programmed cell death 1 and its ligands in regulating autoimmunity and infection. *Nat Immunol*. 2007;8(3):239–245.
6. Hamid O, et al. Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. *N Engl J Med*. 2013;369(2):134–144.
7. Zaidi MR, et al. Interferon- γ links ultraviolet radiation to melanomagenesis in mice. *Nature*. 2011;469(7331):548–553.
8. Rizvi NA, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*. 2015;348(6230):124–128.
9. Faridi R, Zahra A, Khan K, Idrees M. Oncogenic potential of Human Papillomavirus (HPV) and its relation with cervical cancer. *Virol J*. 2011;8:269.
10. Ikeda H, Old LJ, Schreiber RD. The roles of IFN gamma in protection against tumor development and cancer immunoeediting. *Cytokine Growth Factor Rev*. 2002;13(2):95–109.
11. Bald T, et al. Immune cell-poor melanomas benefit from PD-1 blockade after targeted type I IFN activation. *Cancer Discov*. 2014;4(6):674–687.
12. Liang SC, et al. Regulation of PD-1, PD-L1, and PD-L2 expression during normal and autoimmune responses. *Eur J Immunol*. 2003;33(10):2706–2716.
13. Abiko K, et al. IFN- γ from lymphocytes induces PD-L1 expression and promotes progression of ovarian cancer. *Br J Cancer*. 2015;112(9):1501–1509.
14. Bellucci R, et al. Interferon- γ -induced activation of JAK1 and JAK2 suppresses tumor cell susceptibility to NK cells through upregulation of PD-L1 expression. *Oncoimmunology*. 2015;4(6):e1008824.
15. Spranger S, et al. Up-regulation of PD-L1, IDO, and T(regs) in the melanoma tumor microenvironment is driven by CD8(+) T cells. *Sci Transl Med*. 2013;5(200):200ra116.
16. Garon EB, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med*. 2015;372(21):2018–2028.
17. Herbst RS, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature*. 2014;515(7528):563–567.
18. Wallden B, et al. Development and analytical performance of a molecular diagnostic for anti-PD1 response on the nCounter Dx Analysis System. *J Clin Oncol*. 2016;34(15 suppl):3034.
19. Seiwert TY, et al. Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): an open-label, multicentre, phase 1b trial. *Lancet Oncol*. 2016;17(7):956–965.
20. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3(1):32–35.
21. Harlin H, et al. Chemokine expression in melanoma metastases associated with CD8⁺ T-cell recruitment. *Cancer Res*. 2009;69(7):3077–3085.
22. Taube JM, et al. Colocalization of inflammatory response with B7-h1 expression in human melanocytic lesions supports an adaptive resistance mechanism of immune escape. *Sci Transl Med*. 2012;4(127):127ra37.
23. Sharma P, Allison JP. The future of immune checkpoint therapy. *Science*. 2015;348(6230):56–61.
24. Ribas A. Adaptive immune resistance: how cancer protects from immune attack. *Cancer Discov*. 2015;5(9):915–919.
25. Spranger S, Bao R, Gajewski TF. Melanoma-intrinsic β -catenin signalling prevents anti-tumour immunity. *Nature*. 2015;523(7559):231–235.
26. Feig C, et al. Targeting CXCL12 from FAP-expressing carcinoma-associated fibroblasts synergizes with anti-PD-L1 immunotherapy in pancreatic cancer. *Proc Natl Acad Sci U S A*. 2013;110(50):20212–20217.
27. Motz GT, et al. Tumor endothelium FasL establishes a selective immune barrier promoting tolerance in tumors. *Nat Med*. 2014;20(6):607–615.
28. Keytruda (pembrolizumab) for injection, for intravenous use [package insert]. Kenilworth, New Jersey, USA: Merck Sharp & Dohme Corp.; 2017.
29. Dako, an Agilent Technologies company, announces FDA approval of new companion diagnostic for lung cancer [press release]. Carpinteria, California, USA(5): Agilent Technologies; October 5, 2015.
30. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J Royal Statist Soc B*. 2005;67(part 2):301–320.