Introduction

Tumors use multiple mechanisms to suppress adaptive immune responses directed at antigens expressed in the tumor microenvironment. In this context, chronic and persistent antigen stimulation results in increased expression of programmed cell death 1 (PD-1) on CD8+ T cells infiltrating neoplastic tissue (1). The net result of signaling through this receptor is an attenuation of the cytotoxic and cytokine-producing capacity of these cells, leading to ineffective antitumor immune responses (1). Targeted inhibition of the PD-1 pathway has shown excellent efficacy in several human tumors (2–4); however, many patients do not respond, and the cellular and molecular mechanisms underlying this clinical heterogeneity are only beginning to be elucidated. A biomarker that accurately predicts clinical response to anti-PD-1 therapy is critical in order to appropriately select patients for this therapy in the face of multiple emerging treatment options for metastatic cancer.

It is becoming increasingly clear that the immune composition in tumors is markedly different from that observed in peripheral blood (5, 6). Thus, robust biomarkers that predict response to immunotherapy will most likely be derived from tumor tissue. Quantification of programmed cell death ligand 1 (PD-L1) and PD-1 expression in tumors by routine IHC has been used in an attempt to predict response to anti–PD-1 therapy (7–11). These studies have yielded provocative results; however, constraints in the number of markers able to be simultaneously assessed and the inherent difficulties in quantifying staining intensity have limited the potential of this approach. In the current study, we used multiparameter flow cytometry to comprehensively analyze the tumor immune microenvironment prior to anti–PD-1 therapy. Using this approach, we quantified the accumulation of a unique immune cell population that robustly predicts response to this treatment. In functional experiments, we show that this cell subset represents a partially exhausted tumor-infiltrating T cell and that treatment with anti–PD-1 effectively activates these cells in tumors.
Results

Because anti–PD-1 therapy is thought to directly target PD-1–expressing T cells, we hypothesized that the relative abundance of these cells within tumors would predict response to treatment. To test this hypothesis, we performed multiparameter flow cytometry on freshly isolated metastatic melanoma tumor samples prior to anti–PD-1 therapy. We assessed CD45, CD3, CD4, CD8, and FOXP3 expression to quantify the relative percentages of CD4+ effector T cells (Teff) (CD45+CD3+CD4+FOXP3−), CD4+ Tregs (CD45+CD3+CD4+FOXP3+), and CD8+ cytotoxic T lymphocytes (CTLs) (CD45+CD3−CD4−CD8+) infiltrating tumors (Supplemental Figure 1; supplemental material available online with this article; doi:10.1172/JCI87324DS1). In addition, we quantified PD-1, PD-L1, cytotoxic T lymphocyte–associated protein 4 (CTLA-4), and MHC class II (HLA-antigen DR; HLA-DR) expression on each of these subsets. Patients were then treated with anti–PD-1 monotherapy, and clinical outcome data were collected. The therapeutic response was evaluated using response evaluation criteria in solid tumor malignancies (Response Evaluation Criteria In Solid Tumors [RECIST], version 1.1). Patients with a complete response (CR) or a partial response (PR) were considered responders, and those with stable or progressive disease were considered nonresponders. Multivariate analysis revealed that CTL expression of CTLA-4 was the single parameter that showed a statistically significant association with a clinical response (Supplemental Figure 2). Further analysis of CTLA-4–expressing CTLs revealed that this subset of CD8+ T cells expressed the highest levels of PD-1 (Supplemental Figure 2). Representative flow cytometric plots and clinical images of a nonresponder and a responder to anti–PD-1 therapy are shown in Figure 1A. The immunophenotype of a nonresponder clearly showed the presence of tumor-infiltrating CD8+ CTLs; however, these cells expressed low levels of CTLA-4 and PD-1. In contrast, patients who responded to this treatment had relatively increased percentages of CTLA-4hiPD-1hi cells within the CTL gate.

In a discovery cohort of 20 patients, RR and PFS were significantly correlated with the relative abundance of CTLA-4hiPD-1hi CD8+ tumor-infiltrating CTLs (Figure 1B). For patients with more than 20% CTLA-4hiPD-1hi CTLs, the PFS was 31.6 months versus 9.6 months for tumors that contained 20% or fewer CTLA-4hiPD-1hi cells within the CTL gate (*P = 0.017, log-rank test). The objective response rate (CR + PR) was 0% (0 of 6) in the low CTLA-4hiPD-1hi CTL group versus 85.7% (12 of 14) in the greater than 20% CTLA-4hiPD-1hi CTL group. Having identified a potential biomarker for response to anti–PD-1 therapy, we set out to prospectively validate our results in an independent cohort of patients (n = 20). In this validation cohort, patients with 20% or fewer tumor-infiltrating CTLA-4hiPD-1hi CTLs had a PFS of
apy for human melanoma (11). This assay used an immunohistochemical approach to both quantify lymphocytes and discern the invasive tumor margin. We have recently performed this assay on 21 patients with metastatic melanoma (Supplemental Figure 5). Using 1,500 CD8\(^+\) cells/mm\(^2\) as a cutoff, in our hands, this assay had a positive predictive value (PPV) of 57% and a negative predictive value (NPV) of 100%. This is comparable to our flow cytometry–based assay, which had a PPV of 82% and an NPV of 100%. However, using the immunohistochemical approach, 8 of the 21 samples (38%) were not able to be evaluated, because the invasive tumor margin could not be reliably discerned. The majority of these samples were smaller biopsy specimens (core biopsies and skin punch biopsies). Because small tumor specimens are the only samples that can be obtained from many patients with skin, liver, lung, or bone metastasis, the inability to evaluate these samples is a major limitation of the immunohistochemical assay.

Tumor-infiltrating CTLA-4\(^{hi}\)PD-1\(^{hi}\) CTLs have been shown to contain the majority of tumor-antigen–specific T cells (5, 6). These cells are only detectable in tumors (i.e., not peripheral blood) and have been suggested to have an “exhausted” phenotype (5). In order to determine the functional capacity of tumor-infiltrating CTLA-4\(^{hi}\)PD-1\(^{hi}\) CTLs in our patient population, we prepared single-cell suspensions from metastatic melanoma tumors and stimulated these cells with PMA and ionomycin. Intracellular cytokine production was then assessed by flow cytometry. When compared with

9.9 months versus 15.9 months for the high CTLA-4\(^{hi}\)PD-1\(^{hi}\) CTL group (Figure 1C). The RR was 78.6% for the high CTLA-4\(^{hi}\)PD-1\(^{hi}\) CTL group versus 0% (0 of 6) for the low CTLA-4\(^{hi}\)PD-1\(^{hi}\) group. No objective clinical responses were observed in patients with infiltration of 20% or fewer CTLA-4\(^{hi}\)PD-1\(^{hi}\) CTLs into their tumors (Figure 2). The percentages of CTLA-4\(^{hi}\)PD-1\(^{hi}\) CTLs were independent of biopsy site or prior therapy (Figure 2). In addition, response rates were independent of other tumor-infiltrating T cell populations, including Tregs (Figure 2). The increased proportions of CTLA-4\(^{hi}\)PD-1\(^{hi}\) CTLs also correlated with overall survival (OS); however, this did not achieve statistical significance during the relatively short study period (Supplemental Figure 3). Clinical response strongly correlated with the abundance of CTLA-4\(^{hi}\)PD-1\(^{hi}\) double-positive cells within the CD8\(^{+}\) gate (with a concomitant reduction in CTLA-4\(^{lo}\)PD-1\(^{lo}\) double-negative cells) and only minimally with CTLA-4\(^{lo}\)PD-1\(^{hi}\) single-positive cells (Supplemental Figure 4). Taken together, our data suggest that the relative abundance of CTLA-4\(^{hi}\)PD-1\(^{hi}\) CTLs infiltrating metastatic melanoma strongly correlates with clinical response to anti-PD-1 therapy. It is important to acknowledge, however, that given the small sample size of our study, the CI (especially of the low CTLA-4\(^{hi}\)PD-1\(^{hi}\) group) is wide and may be similar to the response rate reported for PDL-1–negative tumors (8–10).

In a previous report, the abundance of CD8\(^{+}\) T cells located at the invasive tumor margin predicted response to anti-PD-1 therapy for human melanoma (11). This assay used an immunohistochemical approach to both quantify lymphocytes and discern the invasive tumor margin. We have recently performed this assay on 21 patients with metastatic melanoma (Supplemental Figure 5). Using 1,500 CD8\(^{+}\) cells/mm\(^2\) as a cutoff, in our hands, this assay had a positive predictive value (PPV) of 57% and a negative predictive value (NPV) of 100%. This is comparable to our flow cytometry–based assay, which had a PPV of 82% and an NPV of 100%. However, using the immunohistochemical approach, 8 of the 21 samples (38%) were not able to be evaluated, because the invasive tumor margin could not be reliably discerned. The majority of these samples were smaller biopsy specimens (core biopsies and skin punch biopsies). Because small tumor specimens are the only samples that can be obtained from many patients with skin, liver, lung, or bone metastasis, the inability to evaluate these samples is a major limitation of the immunohistochemical assay.

Tumor-infiltrating CTLA-4\(^{hi}\)PD-1\(^{hi}\) CTLs have been shown to contain the majority of tumor-antigen–specific T cells (5, 6). These cells are only detectable in tumors (i.e., not peripheral blood) and have been suggested to have an “exhausted” phenotype (5). In order to determine the functional capacity of tumor-infiltrating CTLA-4\(^{hi}\)PD-1\(^{hi}\) CTLs in our patient population, we prepared single-cell suspensions from metastatic melanoma tumors and stimulated these cells with PMA and ionomycin. Intracellular cytokine production was then assessed by flow cytometry. When compared with

Figure 2. Clinical characteristics and CTL profiles of patients with metastatic melanoma who responded or did not respond to anti–PD-1 therapy.

Histogram showing flow cytometric quantification of the percentages of CTLA4\(^{hi}\)PD-1\(^{hi}\) tumor-infiltrating T cells in the CD8\(^{+}\) CTL gate versus characteristics of individual patients. Responders included patients with tumor target lesions that met RECIST 1.1 criteria for a CR (>99% reduction in the target lesions) or a PR (≥30% reduction in target lesions). Nonresponders included patients with tumor target lesions that met RECIST 1.1 criteria for progressive (≥20% increase in the target lesions) or stable disease (<30% reduction or <20% increase in tumor target lesions). n = 40 patients. Tregs were defined as FOXP3\(^{hi}\)CTLA-4\(^{hi}\) cells within the live CD45\(^{+}\)CD3\(^{+}\)CD4\(^{+}\) gate. V, validation; D, discovery; I, ipilimumab; T, targeted therapy with BRAF/MEK inhibitors; Liv, liver; LN, lymph node; Mets, metastasis; S, skin; Ova, ovary.
CTL-4negPD-1lo cells, CTLA-4 hiPD-1hi CTLs contained equivalent percentages in patients who responded to anti–PD-1 therapy. CTLs isolated during therapy had increased percentages of HLA-DR is a robust activation marker for human CTLs (14–16). Significant reduction in the CD4/CD8 ratio (Figure 4). Expression of HLA-DR is a robust activation marker for human CTLs (14–16). In this model, specific functional properties such as IL-2 and TNF-α production are lost first, whereas IFN-γ production is more resistant to attenuation. Thus, partially exhausted CD8+ T cells have been characterized as cells that are capable of producing IFN-γ but lack the ability to produce TNF-α and IL-2, a phenotype consistent with tumor-infiltrating CTLA-4hiPD-1hi CTLs found in high percentages in patients who responded to anti-PD-1 therapy.

Given that the abundance of CTLA-4hiPD-1hi CTLs in metastatic tumors predicts the response to anti-PD-1 therapy and that the functional properties of this cell population is consistent with a partially exhausted T cell phenotype, we set out to determine whether treatment with anti-PD-1 reverses CD8+ T cell exhaustion in tumors. To test this, we immunophenotyped metastatic lesions during anti-PD-1 therapy and compared these with pretreatment samples. Treatment with anti-PD-1 resulted in a marked increase in the percentage of tumor-infiltrating CD8+ T cells, with a concomitant decline in the percentage of CD4+ T cells, resulting in a significant reduction in the CD4/CD8 ratio (Figure 4). Expression of HLA-DR is a robust activation marker for human CTLs (14–16). Thus, we assessed HLA-DR expression on tumor-infiltrating T cells before beginning therapy and after the initiation of therapy. When compared with pretreatment samples, tumor-infiltrating CTLs isolated during therapy had increased percentages of HLA-DR38 cells and increased HLA-DR mean fluorescence intensity (Figure 4). Interestingly, there was no difference in Treg percentages or activation (as indicated by either CTLA-4 or HLA-DR expression) between pretreatment and treatment samples (Figure 4). These results suggest that treatment with anti-PD-1 preferentially activates CD8+ CTLs within the tumor microenvironment.

**Discussion**

Taken together, our data support the notion that the relative abundance of partially exhausted tumor-infiltrating CTLs correlates with response to anti–PD-1 therapy. In addition, treatment with anti-PD-1 results in increased activation of this cell population. Thus, the presence of the “target” cell population within tumors is highly predictive of the response to therapy. In addition to measuring the abundance of CD8+ T cells at the invasive tumor margin (as described above), previous studies have attempted to predict clinical responses by assaying for PD-L1 expression on tumor cells using traditional immunohistochemical approaches (7–10, 17). These attempts have yielded variable results. PD-L1 is expressed on many hematopoietic and nonhematopoietic cells within the tumor microenvironment, and in many cases, there is a wide range of staining intensity. Both of these variables make distinctions between relative degrees of tumor PD-L1 expression difficult to discern. In addition, traditional staining is performed on relatively thin sections of tumor and therefore assesses only a small portion of the specimen. In contrast, flow cytometry allows for evaluation of the entire tumor, enabling a more comprehensive assessment of lesions. Flow cytometry is readily performed on small tumor specimens (core and skin punch biopsies), a setting in which the reliable detection of tumor margins can be challenging. However, it should be noted that flow cytometry of nonhematopoietic tissue is difficult to standardize for clinical use, given the inherent variability in enzymatic tissue digestion and the requirement for fresh (nonpreserved) samples. Future studies focused on developing strategies for the robust quantification of partially exhausted CTLs at either the RNA or DNA (i.e., epigenetic) levels using preserved tumor tissue may help to circumvent these issues.

Interestingly, a recent study involved the use of flow cytometry to show that T cell numbers increase in tumors following PD-1 blockade (18). These results are consistent with our findings; however, this study was not focused on defining biomarkers that predict response to therapy and thus did not identify the partially exhausted CTL population elucidated in our work. High expression of both CTLA-4 and PD-1 on tumor-infiltrating CTLs was instrumental in identifying our biomarker subset (Supplemental Figure 2 and Supplemental Figure 4), providing mechanistic insight into how anti-PD-1 therapy results in tumor regression in select groups of patients. Our results are supported by those of a previous study that found high tumor CTLA-4 expression to be predictive of response to PD-L1 blockade (19). In our assay, patients who had more than 20% of total tumor-infiltrating CD8+ T cells that coexisted high levels of PD-1 and CTLA-4 had an increased likelihood of responding to anti-PD-1 therapy. However, response rates were more variable the closer the percentages were to 20% (Figure 2). This suggests that there is
a subset of patients who have partially exhausted CTLs that hover around the threshold required to achieve a response. It is interesting to speculate that clinical strategies aimed to induce or augment these cells prior to therapy may result in better response rates in these patients. Furthermore, because patients with low percentages of CTLA-4hiPD-1hi CTLs are less likely to respond to anti–PD-1 monotherapy, the risk-benefit profile in this patient subset may favor combination therapy with multiple checkpoint inhibitors. Further clinical investigation is required to discern whether the relative abundance of tumor-infiltrating partially exhausted CTLs can be used to make meaningful clinical therapeutic decisions. In addition, the utility of this biomarker in cancers other than melanoma remains to be determined.

Methods

Study design and cohorts. Between January 2014 and January 2016, 120 patients with advanced melanoma were treated with the anti–PD-1 Ab pembrolizumab or nivolumab. Of these patients, 42 had accessible tumors and were amenable to having a research biopsy performed for flow cytometric analysis. Of these, 2 patients were excluded because of inadequate numbers of cells for analysis. An inadequate number of cells prior to therapy may result in better response rates in these patients. Furthermore, because patients with low percentages of CTLA-4hiPD-1hi CTLs are less likely to respond to anti–PD-1 monotherapy, the risk-benefit profile in this patient subset may favor combination therapy with multiple checkpoint inhibitors. Further clinical investigation is required to discern whether the relative abundance of tumor-infiltrating partially exhausted CTLs can be used to make meaningful clinical therapeutic decisions. In addition, the utility of this biomarker in cancers other than melanoma remains to be determined.

Treatment outcome groups and efficacy analysis. Two treatment outcome groups, responders and nonresponders, were defined using radiologic imaging following anti–PD-1 treatment. Responders included patients with tumor target lesions that met RECIST 1.1 criteria for a CR (≥99% reduction in the target lesions) or a PR (≥30% reduction in target lesions). Nonresponders included patients with tumor target lesions that met RECIST 1.1 criteria for progressive (≥20% increase in the target lesions) or stable disease (<30% reduction or <20% increase in tumor target lesions). Available efficacy and immunological data as of January 2016 were included in all the analyses. The efficacy analysis included 2 endpoints: (1) best overall response was defined as the best tumor response from the start of treatment to the time of disease progression or death; and (2) PFS was defined as the interval between the date of enrollment and the date of progression or death (or the last clinic visit date for which the patient was known not to have had radiological or clinical progression). OS was calculated as the time from the date of enrollment to the time of death or to the last known date that the patient was known to be alive. The best overall response was determined from investigator-reported data according to RECIST 1.1 criteria. The Rosenblum laboratory team was blinded to all clinical data (including treatment response) throughout all studies.

Tumor sample procurement. Patients’ biopsies of metastatic lesions were obtained within 30 days of starting the treatment. Tumor biopsies were performed with either a 16- or 18-gauge needle or a 4-mm punch tool.

Flow cytometric analyses. Multiparameter flow cytometry was performed on pretreatment and treatment samples obtained from metastatic tumors as previously described (20). Freshly isolated samples were minced and digested overnight with buffer consisting of collagen-type IV (4188; Worthington Biochemical Corp.), DNAse (SDN25-1G; Sigma-Aldrich), 10% FBS, 1% HEPES, and 1% penicillin-streptomycin in RPMI medium. Single-cell suspensions were double filtered, centrifuged, and counted. For intracellular cytokine analysis, digested tumor cell suspensions were stimulated with PMA and ionomycin for 4 hours as previously described (20). Approximately 2 × 10⁶ cells were stained with multiple fluorochrome-conjugated mAbs. The following Abs were used (all from eBioscience, unless otherwise stated): anti-human CD3 (anti-hCD3) (UCHT1); anti-hCD8 (RPA-T8); anti-hCD45 (HI30); anti-CD4 (SK3); anti–FOXP3 (PCH101); anti-CTLA-4 (14D3); anti–PD-1 (EH12.2H7; BioLegend); anti–HLA-DR (LN3); anti–PD-L1 (MIH1); and LIVE/DEAD Fixable Aqua Dead Cell Stain (Life Technologies, Thermo Fisher Scientific). Data were acquired by an LSRFortessa (BD Biosciences) and analyzed using FlowJo software (Tree Star Inc.).

Flow cytometric standardization and gating strategy. All the samples were fresh and acquired by the LSRFortessa at different time points. To standardize voltages over time, Sphero Ultra Rainbow
Beads (Spherotech) were used to calibrate and normalize to baseline intensity. Gates were determined using both isotype control Ab staining and an internal negative control cell population (i.e., PD-1 and CTLA-4 expression on CD3+ cells).

Statistics. PFS and OS curves were constructed with the Kaplan-Meier method and compared with the Maentel-Henzel log-rank test using SPSS, version 23 (IBM Corp.). Progression was recorded on the date of scans showing progression or on the date clinical progression or death was noted. All tests were 2-sided, with P values of less than 0.05 considered statistically significant. Patients with tumors that had fewer than 20% CTLA-4 hiPD-1 hi cells were noted to have infrequent responses in initial testing, and this threshold was used in the discovery and validation cohorts. For flow cytometric data, significance was assessed using an unpaired, 2-tailed Student’s t test. In all figures showing quantification of flow cytometric data, the mean value is visually depicted. Statistical analysis was performed using GraphPad Prism, version 6 (GraphPad Software).

Study approval. Biopsies of melanoma tumors were performed after patients provided informed consent under the UCSF Committee on Human Research Protocol 138510. Informed consent was obtained from all patients for biopsies of metastatic lesions within 30 days of starting the treatment.

Author contributions AID and MDR designed the research studies. AID, KL, IKG, MLP, RSR, PMS, and K Taravati conducted experiments. AID, KL, MLP, RSR, and K Tsai acquired data. AID, MDR, KL, MLP, JH, RSR, K Taravani, K Tsai, AN, LN, MDA, APA, MHP, RHP, IVL, and MFK analyzed data. AID, MDR, KL, and MLP wrote the manuscript.

Acknowledgments

This work was supported by grants awarded to AID from the Helen Diller Comprehensive Cancer Center and the Amoroso and Juels-gaard/Cook Funds.

Address correspondence to: Michael D. Rosenblum, 513 Parras Avenue, San Francisco, California 94143, USA. Phone: 415.476.1685; E-mail: Michael.Rosenblum@ucsf.edu.