

Multiplex Quantitative Analysis of Tumor-Infiltrating Lymphocytes and Immunotherapy Outcome in Metastatic Melanoma



Pok Fai Wong^{1,2}, Wei Wei³, James W. Smithy^{1,2}, Balazs Acs^{1,2}, Maria I. Toki^{1,2}, Kim R.M. Blenman⁴, Daniel Zelterman³, Harriet M. Kluger^{2,5}, and David L. Rimm^{1,2,5}

Abstract

Purpose: Because durable response to programmed cell death 1 (PD-1) inhibition is limited to a subset of melanoma patients, new predictive biomarkers could have clinical utility. We hypothesize that pretreatment tumor-infiltrating lymphocyte (TIL) profiles could be associated with response.

Experimental Design: Pretreatment whole tissue sections from 94 melanoma patients treated with anti-PD-1 therapy were profiled by multiplex immunofluorescence to perform TIL quantification (CD4, CD8, CD20) and assess TIL activation (CD3, GZMB, Ki67). Two independent image analysis technologies were used: inForm (PerkinElmer) to determine cell counts, and AQUA to measure protein by quantitative immunofluorescence (QIF). TIL parameters by both methodologies were correlated with objective response or disease control rate (ORR/DCR) by RECIST 1.1 and survival outcome.

Results: Pretreatment lymphocytic infiltration, by cell counts or QIF, was significantly higher in complete or partial response than in stable or progressive disease, particularly for CD8 ($P < 0.0001$). Neither TIL activation nor dormancy was associated with outcome. CD8 associations with progression-free survival (HR > 3) were independently significant in multivariable analyses and accounted for similar CD3 associations in anti-PD-1-treated patients. CD8 was not associated with melanoma prognosis in the absence of immunotherapy. Predictive performance of CD8 cell count (and QIF) had an area under the ROC curve above 0.75 (ORR/DCR), which reached 0.83 for ipilimumab plus nivolumab.

Conclusions: Pretreatment lymphocytic infiltration is associated with anti-PD-1 response in metastatic melanoma. Quantitative TIL analysis has potential for application in digital precision immuno-oncology as an "indicative" companion diagnostic.

Introduction

Although metastatic melanoma is the leading cause of skin cancer mortality globally, a new immunotherapy paradigm has been established by immune checkpoint blockade, with the median overall survival increasing from ~9 months before 2011 to now greater than 3 years (1–4). Immune checkpoints are expressed on the tumor-infiltrating lymphocyte (TIL) population and include programmed cell death 1 (PD-1), which is the target of nivolumab and pembrolizumab; and cytotoxic T-lymphocyte associated protein 4 (CTLA-4), which is targeted by ipilimumab (5). Although responses to these drugs are impressive, benefit is restricted to approximately 40% of metastatic melanoma patients treated with anti-PD-1 therapy (6). Despite

the push toward precision medicine, there are no approved predictive strategies for immunotherapy in melanoma (1, 6). With the recent approval of anti-PD-1 therapy for the management of melanoma, lung cancer, and other malignancies, there is an urgent need for robust predictive biomarkers to inform clinical decision-making (7–9).

Here, we investigate functional states of the tumor immune microenvironment for the prediction of anti-PD-1 response in metastatic melanoma. Because TILs are the major cellular target of anti-PD-1 therapy, we hypothesized that pretreatment TIL profiles would be associated with immunotherapy outcome. A randomized-controlled trial is necessary to statistically test for interaction and prove a predictive biomarker. Furthermore, it is no longer ethically possible to have a placebo arm. Instead, we tested a matched, historic cohort predating immunotherapy to determine absence of prognostic value in the presence of an association between the biomarker and treatment outcome.

We propose introduction of the term "indicative biomarker" in this setting to define a new category for a biomarker that is associated with treatment outcome, but is independent of disease prognosis in a control cohort of historic patients that predated the approval of the given therapy. Indicative value is demonstrated when (1) the HR is statistically significant in the treatment cohort and is not significant in the control cohort; or (2) the HR is statistically significant in both the treatment and control cohorts, but the respective 95% confidence intervals (CI) do not significantly overlap. The former characteristic is purely indicative, and the latter is both prognostic and indicative. The nomenclature

¹Department of Pathology, Yale School of Medicine, New Haven, Connecticut.

²Yale Cancer Center, Yale School of Medicine, New Haven, Connecticut.

³Department of Biostatistics, Yale School of Public Health, New Haven, Connecticut.

⁴Department of Dermatology, Yale School of Medicine, New Haven, Connecticut.

⁵Section of Medical Oncology, Department of Internal Medicine, Yale School of Medicine, New Haven, Connecticut.

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

Corresponding Author: David L. Rimm, Yale School of Medicine, 310 Cedar St, BML 116, PO Box 208023, New Haven, CT 06520. Phone: 203-737-4204; Fax: 203-737-5089; E-mail: david.rimm@yale.edu

doi: 10.1158/1078-0432.CCR-18-2652

©2019 American Association for Cancer Research.

Translational Relevance

Despite an increasing need for predictive biomarkers for immune checkpoint therapy, they can only be proven by a statistical test for interaction in randomized placebo-controlled trials, which are ethically prohibited in melanoma after approval of this class of drugs. Nonetheless, we determined the clinical significance of tumor-infiltrating lymphocyte quantification (CD4, CD8, CD20) and activation (CD3, GZMB, Ki67) for prediction of melanoma immunotherapy outcome by two independent methodologies. CD8, by both techniques, was significantly associated with response, independent of baseline variables, in a retrospective anti-PD-1-treated melanoma cohort, but not associated with melanoma prognosis in a historic cohort predating immunotherapy. We introduce the term "indicative biomarker" because CD8 is indicative of specific immunotherapy outcome and not associated with prognosis. This new category allows distinction from truly predictive biomarkers, and importantly, facilitates development of companion diagnostics for established therapies where new predictive biomarkers by definition are statistically impossible.

reflects such a biomarker is indicative of specific treatment outcome that is separate from disease prognosis in a context where a statistical interaction test cannot be performed. This new category allows distinction from truly predictive biomarkers, and importantly, facilitates development of companion diagnostics for established therapies where new predictive biomarkers by definition are statistically impossible.

Our candidate biomarkers were tested using multiplex immunofluorescence panels to (1) perform TIL quantification of helper T cells by CD4, cytotoxic T cells by CD8, and B cells by CD20; and (2) assess TIL activation by identifying T cells by CD3, cytolytic activity by granzyme B (GZMB), and proliferation by Ki67. These assays were performed on an anti-PD-1-treated cohort and a historic cohort of patients seen before the advent of immunotherapy to exclude prognostic value. Because predictive value is statistically impossible to prove in this setting, we propose the term "indicative biomarker."

Materials and Methods

Patient cohort

The study cohort consists of a retrospective collection of 94 melanoma patients treated with anti-PD-1 therapy at Yale Cancer Center between 2011 and 2017. Patients with uveal melanoma were excluded (10). Pretreatment formalin-fixed, paraffin-embedded (FFPE) specimens from Yale Pathology archives were reviewed by a board-certified pathologist. Clinicopathologic data were collected from clinical records and pathology reports; the data cutoff date was September 1, 2017. RECIST 1.1 were used to classify best overall response as complete response (CR), partial response (PR), stable disease (SD), or progressive disease (PD), and to determine objective response rate (ORR; CR/PR), disease control rate (DCR; CR/PR/SD), and progression-free survival (PFS; ref. 11). A historic cohort of 100 untreated melanoma patients was used as the control group. Cohort characteristics are described in Table 1. All patients provided written-informed

Table 1. Clinicopathologic characteristics of the melanoma cohort treated with anti-PD-1 therapy and the untreated melanoma cohort

Characteristic	Anti-PD-1 patients, N (%)	ORR (CR/PR), N (%)	DCR (CR/PR/SD), N (%)	Untreated patients, N (%)
Overall	94 (100)	42 (45)	65 (69)	100 (100)
Age (y)				
<65	54 (57)	24 (44)	40 (74)	44 (44)
≥65	40 (43)	18 (45)	25 (63)	56 (56)
Sex				
Male	57 (61)	27 (47)	38 (67)	60 (60)
Female	37 (39)	15 (41)	27 (73)	40 (40)
Treatment				
Pembrolizumab	35 (37)	16 (46)	25 (71)	0
Nivolumab	14 (15)	5 (36)	7 (50)	0
Ipilimumab plus nivolumab	45 (48)	21 (47)	33 (73)	0
Prior immune checkpoint blockade				
Yes	29 (31)	11 (38)	19 (66)	0
No	65 (69)	31 (48)	46 (71)	100 (100)
Mutation status				
BRAF	27 (29)	9 (33)	17 (63)	NA
NRAS	15 (16)	7 (47)	10 (67)	NA
KIT	2 (2)	1 (50)	2 (100)	NA
None detected	50 (53)	25 (50)	36 (72)	NA
Stage at diagnosis				
I	21 (22)	12 (57)	17 (81)	42 (42)
II	19 (20)	9 (47)	12 (63)	54 (54)
III	29 (31)	11 (38)	18 (62)	1 (1)
IV	15 (16)	5 (33)	11 (73)	1 (1)
Not available	10 (11)	5 (50)	7 (70)	2 (2)

Abbreviation: NA, not available.

consent or waiver of consent. The study was approved by the Yale Human Investigation Committee protocol #9505008219 and conducted in accordance with the Declaration of Helsinki.

Multiplex immunofluorescence TIL quantification and activation panels

FFPE whole tissue sections were processed for 5-plex immunofluorescence with simultaneous detection of markers by isotype-specific antibodies as previously described (12). The protocol is detailed in the Supplementary Material.

Image analysis by two independent methods: cell counts versus quantitative immunofluorescence

Cell counts were determined by the pattern recognition software, inForm Tissue Finder (PerkinElmer), on multispectral images acquired using a Vectra 3 system (PerkinElmer) as previously described (13). Multispectral images were decomposed into their various components by spectral unmixing using a digital spectral library consisting of spectral profiles of each of the fluorophores. Automated tissue segmentation identified tumor and stroma regions. Cell segmentation within these regions identified individual cells and respective nuclei, cytoplasm, and membrane components using signal in the nucleus and membrane as internal and external cell borders, then cells were phenotyped for marker expression. Cell counts for each melanoma case were calculated in terms of the number of cells positive for the marker of interest as a percentage of the cell population in which it was measured. Protein expression of the various markers was determined by the AQUA method of quantitative immunofluorescence (QIF) on fluorescence images acquired using a PM-2000 system (Navigate BioPharma) as previously described (14).

A total compartment, consisting of all cells, or a CD3 compartment was generated by automated processing and thresholding of the DAPI signal or CD3 signal, respectively. QIF scores were calculated by dividing the summated pixel intensities for the marker of interest by the area of the compartment in which it was measured (14). Overall QIF scores were derived for each melanoma case by averaging scores from each field of view.

Statistical analysis

Statistical comparisons for cell counts and QIF scores were made by unpaired *t* test or ANOVA followed by Tukey test for multiple comparisons as appropriate. Joinpoint regression (NCL, Bethesda, MD) determines statistically significant thresholds based on the data distribution without any input from outcome or other variables, and was used to objectively define low and high status for the measured TIL parameters (15). Kaplan–Meier estimates of survival functions were computed, and comparisons were made by the log-rank test. Multivariable Cox proportional hazards models included age, sex, mutation status, stage, treatment, and prior immune checkpoint blockade as covariates (16–19). ROC curves were constructed from logistic regression models for the prediction of ORR or DCR. All statistical tests were two-sided, and statistical significance was defined as $P < 0.05$. Statistical analyses were performed using GraphPad Prism 7 (GraphPad Software) and JMP Pro 13 (SAS Institute). The sample size of 94 patients had at least 80% power at $P = 0.05$ to detect a difference in means of 0.59 standard deviations in each TIL parameter for responders (CR/PR) versus nonresponders (SD/PD).

Results

Correlation between cell counts and quantitative immunofluorescence

In situ quantification of tissue biomarkers can be performed by counting cells with expression of a biomarker over a predefined threshold or by quantitative protein expression levels per unit area. These are different parameters and may differ in clinical significance. The relationship between these two types of parameters was assessed for the six markers (Supplementary Fig. S1A). Cell counts and QIF exhibited a positive correlation that was best for the most abundant cell types or markers, CD8 ($R^2 = 0.78$), CD3 ($R^2 = 0.62$), GZMB ($R^2 = 0.70$), and Ki67 ($R^2 = 0.75$), which had broad distributions of values (Supplementary Fig. S1A). This direct proportionality deteriorated for CD4 ($R^2 = 0.28$) where cell counts clustered under 15% with 0.9 relative frequency, and collapsed for CD20 ($R^2 = 0.018$) where cell counts clustered under 5% with 0.9 relative frequency. This is expected because quantitative per unit area measurement methods become less accurate as events decrease per unit area. Finally, there was no correlation between different markers, which confirmed their independence (Supplementary Fig. S1B and S1C).

Best overall response by RECIST and TIL parameters

Pretreatment whole tissue sections from 94 melanoma patients treated with anti-PD-1 therapy were profiled with two multiplex immunofluorescence panels to perform TIL quantification (CD4, CD8, CD20; Fig. 1A) and assess TIL activation (CD3, GZMB, Ki67; Fig. 1B). For TIL activation, we used criteria previously described for lung cancer, classifying three major states of the tumor immune microenvironment: immune desert (CD3 low),

TIL dormancy (CD3 high, Ki67 and GZMB low), and TIL activation (CD3 high, Ki67 and/or GZMB high; Fig. 1B; ref. 20). The TIL quantification and activation parameters were analyzed in relation to specimen-specific variables and best overall response defined by RECIST 1.1 (11). There were no significant differences associated with sex nor mutation status of melanoma patients in any TIL parameters by cell counts or QIF (Supplementary Fig. S2). From the TIL quantification panel, CD4 cell counts in CR/PR were higher than in PD ($P = 0.024$), whereas CD20 cell counts were not correlated with response (Fig. 1C). Notably, CD8 cell counts in CR/PR were 2-fold higher than in SD, and 4-fold higher than in PD ($P < 0.0001$; Fig. 1C). From the TIL activation panel, CD3 cell counts in CR/PR versus SD versus PD were in the same ratio of 4:2:1 ($P < 0.0001$) as observed for CD8, which is consistent with the fact that CD8 cells are a subset of CD3 cells (Fig. 1D). Neither cytolytic nor proliferative CD3 cell counts were associated with response (Fig. 1D). These findings were corroborated by the QIF data (Supplementary Fig. S3) and further analyses on ORR and DCR (Supplementary Fig. S4), which revealed similar trends.

Survival outcome and TIL parameters

For survival analysis, the continuous TIL parameters were dichotomized into low and high statuses objectively defined by Joinpoint regression (15), which determines statistically significant thresholds based on the data distribution without any input from outcome or other variables (Supplementary Fig. S5). Therefore, this approach represents the standardized derivation of a threshold that is a fundamental characteristic of the population data. From the TIL activation panel, high CD3 cell count was associated with prolonged survival ($P = 0.0002$; Fig. 2A). Neither cytolytic nor proliferative CD3 cell counts alone were associated with survival (Fig. 2A). The combined status of CD3, GZMB, and Ki67 was then used in a survival analysis comparing the three tumor immune microenvironment states described above. When the immune infiltration (CD3 high) category was stratified according to GZMB alone, Ki67 alone, or both GZMB and Ki67, no significant survival differences between TIL dormancy and TIL activation were found (all $P > 0.05$; Fig. 2B). Indeed, significant survival advantage was attributed to immune infiltration independent of the absence or presence of TIL activation (all $P < 0.0015$; Fig. 2B). From the TIL quantification panel, survival was not associated with CD4 cell count, and a marginal difference was observed for CD20 cell count (Fig. 3A).

High CD8 cell count was associated with prolonged survival of anti-PD-1-treated melanoma patients ($P < 0.0001$; Fig. 3A). To distinguish this from prognostic value, we assessed outcome in a historic cohort of melanoma patients with known disease-specific survival in place of a placebo arm. In contrast to the treated patients, survival of untreated melanoma patients was not associated with CD8 cell count (Fig. 3B). Indeed, Cox regressions confirmed that CD8 was significant in anti-PD-1 patients with an HR of 3.35 (95% CI, 1.89–6.24) and not in untreated patients (HR = 1.22; 95% CI, 0.48–3.48; Table 2). Multivariable analyses also revealed significant CD8 survival associations (HR > 3; $P < 0.0025$) independent of age, sex, mutation, stage, treatment, and prior immune checkpoint blockade, which accounted for similar CD3 survival associations as expected (Table 2). Again, these results were corroborated by the QIF data which showed similar profiles in relation to survival (Supplementary Figs. S6 and

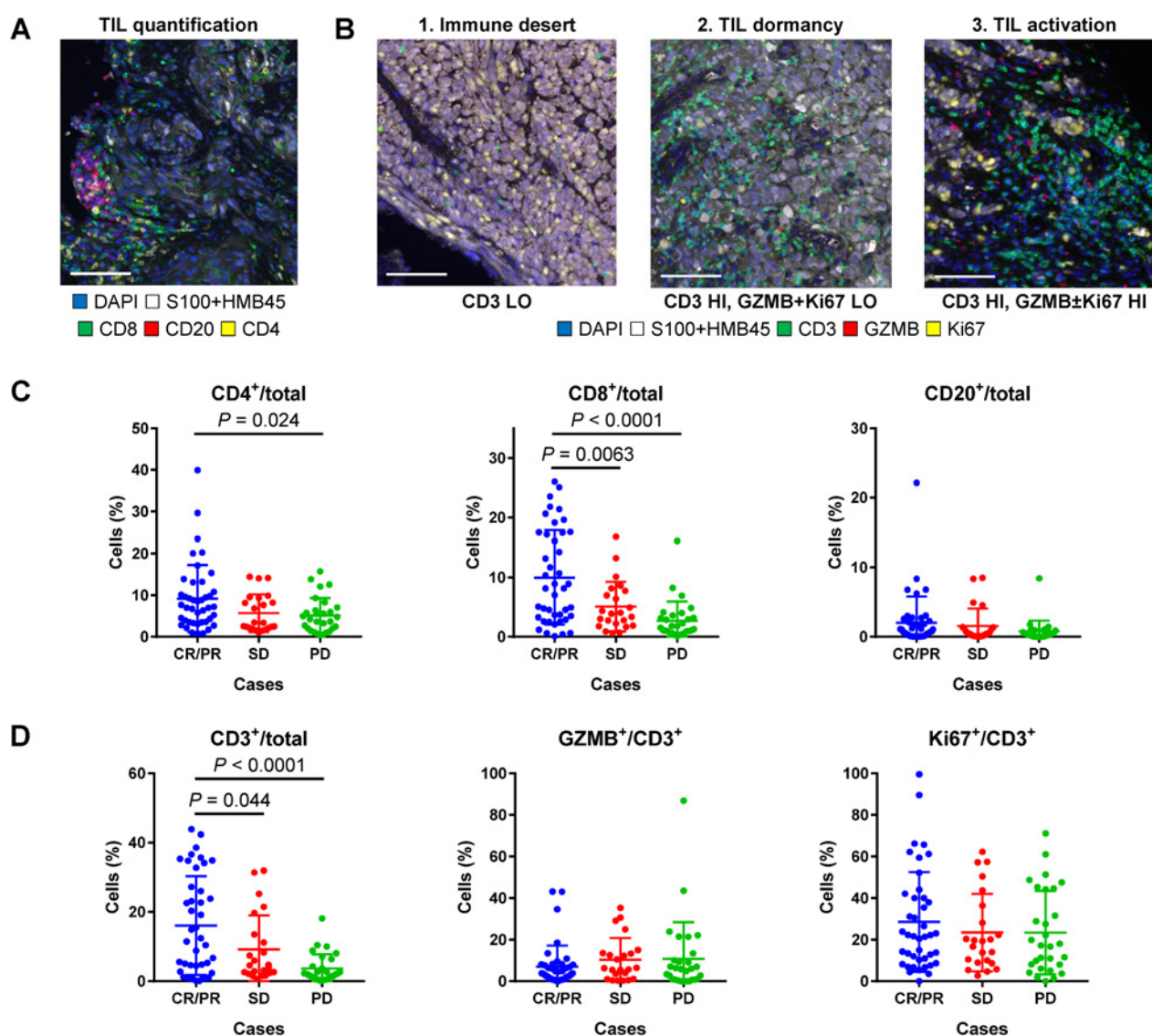


Figure 1.

RECIST categories of melanoma patients treated with anti-PD-1 therapy and TIL parameters by cell counts. Representative multispectral immunofluorescence images of TIL quantification (CD4, CD8, CD20; **A**) and the three major states of the tumor immune microenvironment (**B**) in melanoma: immune desert (CD3 low), TIL dormancy (CD3 high, Ki67 and GZMB low), and TIL activation (CD3 high, Ki67 and/or GZMB high; magnification, $\times 200$; scale bar, 100 μm). TIL quantification (**C**) and TIL activation (**D**) parameters by cell counts per RECIST categories of best overall response. Data are presented as mean with standard deviation (error bars). Abbreviations: HI, high; LO, low.

S7; Supplementary Table S1). To prove that this biomarker can be applied in the Clinical Laboratory Improvement Act-certified laboratory setting, CD8 cell count was performed by conventional chromogenic IHC in a tissue microarray format (Supplementary Fig. S8), which reproduced similar results (HR = 2.60; 95% CI, 1.11–7.08) to that obtained by both fluorescent methods. Furthermore, survival analysis by treatment group revealed similar trends (Supplementary Tables S2 and S3).

Predictive performance of CD8

Although this study is not a randomized clinical trial, ROC curves can still be constructed from logistic regression models for the prediction of anti-PD-1 response in terms of ORR and DCR.

Therefore, the predictive performances of CD8 cell count and QIF were assessed by the AUC. CD8 cell count achieved a favorable predictive performance where the AUC for ORR was 0.75 (95% CI, 0.65–0.85; $P < 0.0001$) and for DCR was 0.78 (95% CI, 0.68–0.87; $P < 0.0001$; Fig. 4). For dual therapy, CD8 cell count reached AUCs of 0.83 (95% CI, 0.70–0.95; $P = 0.0002$) and 0.83 (95% CI, 0.68–0.99; $P = 0.0007$) for ORR and DCR, respectively (Fig. 4). Similarly, CD8 QIF exhibited a favorable predictive performance with AUCs of 0.72 (95% CI, 0.61–0.83; $P = 0.0003$) for ORR and 0.77 (95% CI, 0.67–0.87; $P < 0.0001$) for DCR, which increased for dual therapy to 0.77 (95% CI, 0.63–0.91; $P = 0.0020$) and 0.81 (95% CI, 0.66–0.97; $P = 0.0015$), respectively (Supplementary Fig. S9).

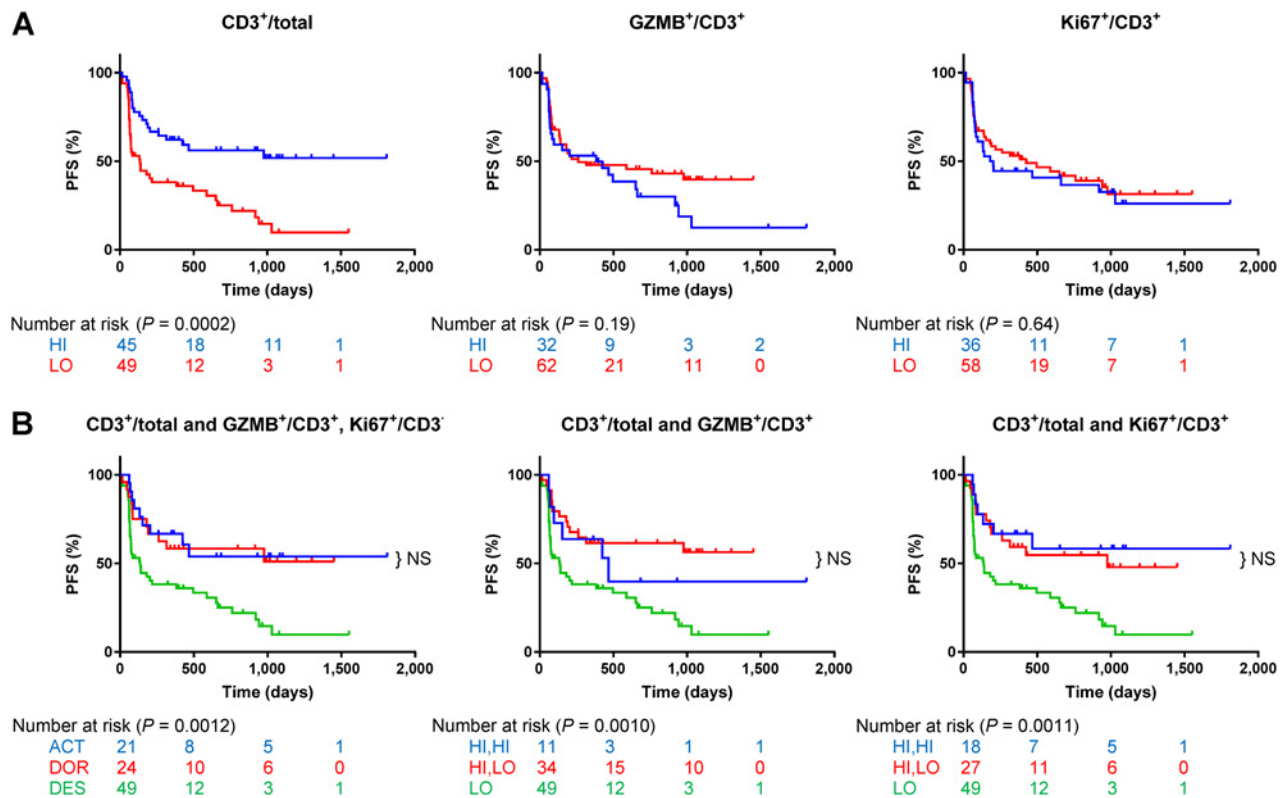


Figure 2. TIL activation parameters by cell counts and survival of melanoma patients treated with anti-PD-1 therapy. Kaplan-Meier analysis of PFS according to TIL activation (CD3, GZMB, Ki67; **A**) parameters by cell counts, and the three states of the tumor immune microenvironment (**B**): immune desert (CD3 low), TIL dormancy (CD3 high, Ki67 and GZMB low), and TIL activation (CD3 high, Ki67 and/or GZMB high). The immune infiltration (CD3 high) category was stratified according to GZMB alone, Ki67 alone, or both GZMB and Ki67. Low and high statuses were objectively defined using thresholds determined by Joinpoint regression (see Materials and Methods). Abbreviations: ACT, TIL activation; DES, immune desert; DOR, TIL dormancy; HI, high; LO, low; NS, not significant.

Discussion

Here, we determine the clinical significance of pretreatment TIL activation status (CD3, GZMB, Ki67) and three additional lymphocytic subpopulations (CD4, CD8, CD20) according to both *in situ* cell counts and protein expression by QIF in relation to immunotherapy outcome in metastatic melanoma. Although the two quantitative methods are different, cell counts correlated with QIF and revealed concordant associations with response and survival. Pretreatment lymphocytic infiltration was significantly higher in CR/PR than in SD/PD. Significant CD8 associations with survival were independent of age, sex, mutation, stage, treatment, and prior immune checkpoint blockade, and also accounted for similar CD3 survival associations (16–19).

Although this study attempts to rigorously investigate multiplex TIL profiling and melanoma immunotherapy outcome, there are a number of limitations. The most significant limitation is the fact that this or any biomarker cannot be proven to be predictive if it is not included in the original successful trial of a given therapy. That is, biomarker data of placebo patients are required in a statistical test for interaction to prove a predictive biomarker; however, a placebo arm is unethical after a successful initial trial. Theoretically, tissue could be obtained retrospectively from prior trials, but that is challenging and dependent on availability. Therefore, this and all posttrial predictive biomarker studies are

limited by the same statistical requirement. To circumvent this problem, and not dilute the term "predictive," we propose a new category called an "indicative" biomarker. To prove indicative value, we analyzed a similar historic melanoma cohort that predates the approval of anti-PD-1 therapy. Although this is an imperfect solution given that we assess disease-specific survival rather than PFS, it allows the evaluation of prognostic value in a similar setting. Indicative value is inferred if the biomarker is associated with outcome in the treated arm (or in this case, cohort), but the biomarker-positive and -negative groups within the placebo arm (or cohort) are comparable in outcome. This is demonstrated in Fig. 3B, where CD8 does not appear to have prognostic value as assessed by disease-specific survival. This is in contrast to TIL studies and CD8 mRNA expression studies that have shown prognostic value (21, 22). Those studies do not specifically address the prognostic value of the protein, in contrast to our quantitative efforts which fail to find prognostic value. We claim here that CD8 has indicative value and may be useful as a clinical assay to determine likelihood of response to anti-PD-1 therapy in melanoma. Further investigation of this biomarker is planned, especially in the adjuvant setting where only 1 in 5 treated melanoma patients benefit from anti-PD-1 therapy (9). Additional biomarkers in combination with CD8 may be required for successful clinical implementation because a number of responders had low CD8.

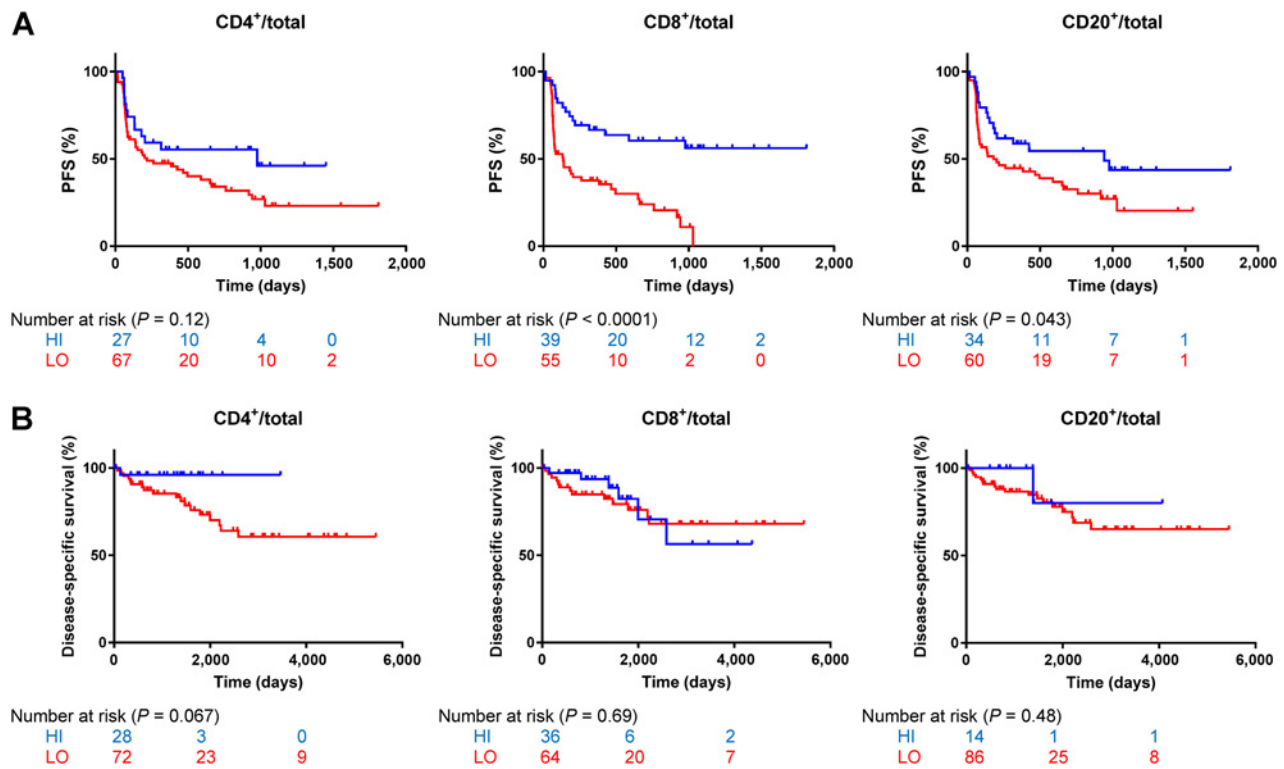


Figure 3. TIL quantification parameters by cell counts and survival of melanoma patients treated with anti-PD-1 therapy and untreated melanoma patients. Kaplan-Meier analysis of PFS of anti-PD-1-treated melanoma patients (A) and disease-specific survival of untreated melanoma patients (B) according to TIL quantification (CD4, CD8, CD20) parameters by cell counts. Low and high statuses were objectively defined using thresholds determined by Joinpoint regression (see Materials and Methods). Abbreviations: HI, high; LO, low.

This is a single-institutional retrospective study with a modest sample size, even though all available cases at Yale were collected at the time of the study. Although the hypothesis of this study is well-substantiated in the literature, an independent validation cohort would be ideal. We look forward to prospective application of these assays or similar in future clinical trials. Our TIL profiling methodologies used quantitative fluorescence imaging systems for increased accuracy and the ability to compare measurement methods. The finding that CD8 has indicative value can be easily translated to current conventional tests. Although this observation is in apparent contradiction to previous studies that have reported prognostic value of TIL

grade in melanoma, those studies did not subclassify the T cells nor use molecular markers (23, 24). Our data suggest that the TIL prognostic value may be driven by CD4 cells (Fig. 3B). However, CD4 is expressed by helper T cells as well as regulatory T cells, which require additional markers for their identification. Work using artificial intelligence technologies for the assessment of TILs and immunotherapy outcome is also underway.

The finding that pretreatment lymphocytic infiltration (CD3 and CD8) is the primary determinant of immunotherapy outcome is consistent with previous literature (25, 26). Whereas previous studies involved anti-PD-1 monotherapy, our work

Table 2. Univariable and multivariable Cox regression analyses for survival of melanoma patients and TIL parameters by cell counts

Variable (LO/HI)	Untreated patients		Anti-PD-1 patients							
	Univariable analysis		Univariable analysis		Trivariable analysis for TIL activation		Multivariable ^a analysis per variable		Multivariable ^a analysis for CD3 and CD8	
	HR (95% CI)	P value	HR (95% CI)	P value	HR (95% CI)	P value	HR (95% CI)	P value	HR (95% CI)	P value
CD4 ⁺ /total	5.39 (1.10-97.3)	0.036	1.62 (0.90-3.13)	0.11			1.44 (0.78-2.85)	0.26		
CD8 ⁺ /total	1.22 (0.48-3.48)	0.69	3.35 (1.89-6.24)	<0.0001			3.74 (2.01-7.33)	<0.0001	3.37 (1.53-7.81)	0.0022
CD20 ⁺ /total	2.06 (0.42-37.3)	0.44	1.78 (1.03-3.22)	0.039			1.65 (0.90-3.16)	0.11		
CD3 ⁺ /total			2.66 (1.56-4.66)	0.0003	2.65 (1.52-4.74)	0.0005	2.42 (1.36-4.43)	0.0024	1.17 (0.55-2.52)	0.69
GZMB ⁺ /CD3 ⁺			0.71 (0.42-1.20)	0.20	0.95 (0.55-1.66)	0.84	0.76 (0.43-1.38)	0.39		
Ki67 ⁺ /CD3 ⁺			0.88 (0.53-1.50)	0.64	0.84 (0.49-1.45)	0.52	1.03 (0.57-1.95)	0.92		

NOTE: Bold font denotes statistical significance where $P < 0.05$.

Abbreviations: HI, high; LO, low.

^aCox proportional hazards model included age, sex, mutation status, stage, treatment, and prior immune checkpoint blockade as covariates.

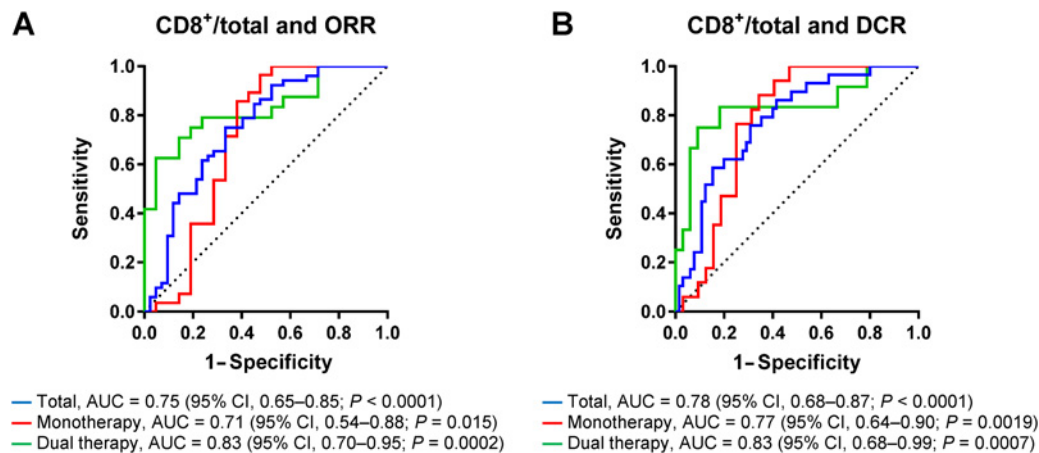


Figure 4.

ROC curve analysis of CD8 by cell counts for the prediction of anti-PD-1 ORR or DCR in melanoma. ROC curves constructed from logistic regression models for the prediction of anti-PD-1 response in terms of ORR (**A**) and DCR (**B**) for the total cohort, monotherapy (pembrolizumab or nivolumab), or dual therapy (ipilimumab plus nivolumab). AUC of 0.50 represents performance of random chance (line of identity, dotted); 1.00 represents perfect predictive performance. P values indicate probability that the AUC is significantly different from 0.50.

extends this knowledge to dual therapy and addresses both measures of clinical response, ORR and DCR. Furthermore, mutation status was not associated with any TIL parameter (27, 28). Multivariable analyses provided unique insights including the redundancy of CD3 in the presence of CD8, which is statistical evidence for CD8 cells being the functional subset of T cells in the observed outcome associations.

The use of two distinct and independent image analysis technologies permitted a technical evaluation of different systems of measurement, cell counts and QIF, and their concordance increases the confidence in the result. The AQUA method of QIF calculates the cumulative signal intensity per unit compartment area as an effective measure of protein expression, which is fundamentally different from counts of digitally phenotyped cells (29). Overall, the two methodologies produced similar findings, suggesting shared biological relevance. However, the correlation between cell counts and QIF declined for the rarer markers, consistent with the fact that the accuracy of QIF by AQUA depends on confluent tissue compartments, which does not apply to low frequency objects. Furthermore, cell counts are expressed in intuitive absolute units and achieved a higher predictive performance than QIF; therefore, it may have a greater potential for clinical utility. Notably, our work shows that the indicative value of CD8 cell count is also easily achieved using conventional chromogenic IHC (Supplementary Fig. S8).

In summary, this study shows the clinical significance of independent TIL subpopulations in relation to immunotherapy outcome in metastatic melanoma. Pretreatment lymphocytic infiltration, by *in situ* cell counts or QIF of protein expression, is significantly associated with melanoma anti-PD-1 response. Similar multiplex analysis of the tumor immune microenvironment has the potential for application as a companion diagnostic in next-generation precision immunotherapy. Furthermore, conversion of this assay to routinely used CD8 IHC tests may offer clinicians valuable information when choosing among therapeutic options in the absence of other defined methods for patient stratification.

Disclosure of Potential Conflicts of Interest

H.M. Kluger reports receiving commercial research grants from Merck, Bristol-Myers Squibb, and Apexigen, and is a consultant/advisory board member for Alexion, Corvus, Nektar, Biodesix, Genentech, Merck, Celldex, Pfizer, Iovance, and Immunocore. D.L. Rimm is an employee of Bristol-Myers Squibb, NanoString, Ultivue, and Biocept; reports receiving other commercial research support from AstraZeneca, Cepheid, Navigate Biopharma, NextCure, Lilly, Ultivue, and Perkin Elmer; and is a consultant/advisory board member for AstraZeneca, Amgen, Cell Signaling Technology, Cepheid, Daiichi Sankyo, Merck, and In Vitro/Konica Minolta.

No potential conflicts of interest were disclosed by the other authors.

Disclaimer

The funding sources had no role in study design; collection, analysis, and interpretation of data; preparation of the article; or the decision to submit for publication.

Authors' Contributions

Conception and design: P.F. Wong, H.M. Kluger, D.L. Rimm

Development of methodology: P.F. Wong, K.R.M. Blenman, D. Zelterman, D.L. Rimm

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): P.F. Wong, J.W. Smithy, B. Acs, H.M. Kluger

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): P.F. Wong, W. Wei, B. Acs, K.R.M. Blenman, D. Zelterman, H.M. Kluger, D.L. Rimm

Writing, review, and/or revision of the manuscript: P.F. Wong, W. Wei, J.W. Smithy, B. Acs, M.I. Toki, K.R.M. Blenman, H.M. Kluger, D.L. Rimm

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): P.F. Wong, M.I. Toki, K.R.M. Blenman, D.L. Rimm

Study supervision: P.F. Wong, D.L. Rimm

Acknowledgments

The authors thank Lori A. Charette and the staff of Yale Pathology Tissue Services for expert histology services. B. Acs was supported by the Fulbright Program and the Rosztoczy Foundation Scholarship Program. This work is based on the PhD dissertation research of Dr. Pok Fai Wong as a Gruber Science Fellow at Yale University.

This work was supported by funds from Navigate BioPharma (Novartis subsidiary), Yale SPORE in Lung Cancer and Yale Cancer Center to D.L. Rimm; R01 CA227473, K24CA172123, and P50 CA121974 to H.M. Kluger; the Melanoma Research Alliance Young Investigator Award Program to

K.R.M. Blenman; and the Gruber Science Fellowship to P.F. Wong from the Gruber Foundation.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked

advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received August 13, 2018; revised October 19, 2018; accepted January 3, 2019; published first January 7, 2019.

References

- Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med* 2012;366:2443–54.
- Weber JS, D'Angelo SP, Minor D, Hodi FS, Gutzmer R, Neyns B, et al. Nivolumab versus chemotherapy in patients with advanced melanoma who progressed after anti-CTLA-4 treatment (CheckMate 037): a randomised, controlled, open-label, phase 3 trial. *Lancet Oncol* 2015;16:375–84.
- Larkin J, Chiarion-Sileni V, Gonzalez R, Grob JJ, Cowey CL, Lao CD, et al. Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. *N Engl J Med* 2015;373:23–34.
- Wolchok JD, Chiarion-Sileni V, Gonzalez R, Rutkowski P, Grob JJ, Cowey CL, et al. Overall survival with combined nivolumab and ipilimumab in advanced melanoma. *N Engl J Med* 2017;377:1345–56.
- Robert C, Schachter J, Long GV, Arance A, Grob JJ, Mortier L, et al. Pembrolizumab versus ipilimumab in advanced melanoma. *N Engl J Med* 2015;372:2521–32.
- Ribas A, Hamid O, Daud A, Hodi FS, Wolchok JD, Kefford R, et al. Association of pembrolizumab with tumor response and survival among patients with advanced melanoma. *JAMA* 2016;315:1600–9.
- Weber J, Mandala M, Del Vecchio M, Gogas HJ, Arance AM, Cowey CL, et al. Adjuvant nivolumab versus ipilimumab in resected stage III or IV melanoma. *N Engl J Med* 2017;377:1824–35.
- Abdel-Wahab N, Shah M, Suarez-Almazor ME. Adverse events associated with immune checkpoint blockade in patients with cancer: a systematic review of case reports. *PLoS One* 2016;11:e0160221.
- Eggermont AMM, Blank CU, Mandala M, Long GV, Atkinson V, Dalle S, et al. Adjuvant pembrolizumab versus placebo in resected stage III melanoma. *N Engl J Med* 2018;378:1789–801.
- Qin Y, Petaccia de Macedo M, Reuben A, Forget MA, Haymaker C, Bernatchez C, et al. Parallel profiling of immune infiltrate subsets in uveal melanoma versus cutaneous melanoma unveils similarities and differences: a pilot study. *Oncoimmunology* 2017;6:e1321187.
- Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;45:228–47.
- Brown JR, Wimberly H, Lannin DR, Nixon C, Rimm DL, Bossuyt V. Multiplexed quantitative analysis of CD3, CD8, and CD20 predicts response to neoadjuvant chemotherapy in breast cancer. *Clin Cancer Res* 2014;20:5995–6005.
- Huang W, Hennrick K, Drew S. A colorful future of quantitative pathology: validation of Vectra technology using chromogenic multiplexed immunohistochemistry and prostate tissue microarrays. *Hum Pathol* 2013;44:29–38.
- Camp RL, Chung GG, Rimm DL. Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nat Med* 2002;8:1323–7.
- Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. *Stat Med* 2000;19:335–51.
- Eton O, Legha SS, Moon TE, Buzaid AC, Papadopoulos NE, Plager C, et al. Prognostic factors for survival of patients treated systemically for disseminated melanoma. *J Clin Oncol* 1998;16:1103–11.
- Manola J, Atkins M, Ibrahim J, Kirkwood J. Prognostic factors in metastatic melanoma: a pooled analysis of Eastern Cooperative Oncology Group trials. *J Clin Oncol* 2000;18:3782–93.
- Neuman HB, Patel A, Ishill N, Hanlon C, Brady MS, Halpern AC, et al. A single-institution validation of the AJCC staging system for stage IV melanoma. *Ann Surg Oncol* 2008;15:2034–41.
- Joose A, Collette S, Suci S, Nijsten T, Patel PM, Keilholz U, et al. Sex is an independent prognostic indicator for survival and relapse/progression-free survival in metastasized stage III to IV melanoma: a pooled analysis of five European Organisation for Research and Treatment of Cancer randomized controlled trials. *J Clin Oncol* 2013;31:2337–46.
- Gettinger SN, Choi J, Mani N, Sanmamed MF, Datar I, Sowell R, et al. A dormant TIL phenotype defines non-small cell lung carcinomas sensitive to immune checkpoint blockers. *Nat Commun* 2018;9:3196.
- Piras F, Colombari R, Minerba L, Murtas D, Floris C, Maxia C, et al. The predictive value of CD8, CD4, CD68, and human leukocyte antigen-D-related cells in the prognosis of cutaneous malignant melanoma with vertical growth phase. *Cancer* 2005;104:1246–54.
- Danilova L, Wang H, Sunshine J, Kaunitz GJ, Cottrell TR, Xu H, et al. Association of PD-1/PD-L axis expression with cytolytic activity, mutational load, and prognosis in melanoma and other solid tumors. *Proc Natl Acad Sci U S A* 2016;113:E7769–77.
- Azimi F, Scolyer RA, Rumcheva P, Moncrieff M, Murali R, McCarthy SW, et al. Tumor-infiltrating lymphocyte grade is an independent predictor of sentinel lymph node status and survival in patients with cutaneous melanoma. *J Clin Oncol* 2012;30:2678–83.
- Thomas NE, Busam KJ, From L, Kricke A, Armstrong BK, Anton-Culver H, et al. Tumor-infiltrating lymphocyte grade in primary melanomas is independently associated with melanoma-specific survival in the population-based genes, environment and melanoma study. *J Clin Oncol* 2013;31:4252–9.
- Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJ, Robert L, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 2014;515:568–71.
- Chen PL, Roh W, Reuben A, Cooper ZA, Spencer CN, Prieto PA, et al. Analysis of immune signatures in longitudinal tumor samples yields insight into biomarkers of response and mechanisms of resistance to immune checkpoint blockade. *Cancer Discov* 2016;6:827–37.
- Cancer Genome Atlas Network. Genomic classification of cutaneous melanoma. *Cell* 2015;161:1681–96.
- Krauthammer M, Kong Y, Bacchicocchi A, Evans P, Pomputtpong N, Wu C, et al. Exome sequencing identifies recurrent mutations in NF1 and RASopathy genes in sun-exposed melanomas. *Nat Genet* 2015;47:996–1002.
- Carvajal-Hausdorf DE, Schalper KA, Neumeister VM, Rimm DL. Quantitative measurement of cancer tissue biomarkers in the lab and in the clinic. *Lab Invest* 2015;95:385–96.