nature medicine

High-dimensional single-cell analysis predicts response to anti-PD-1 immunotherapy

Carsten Krieg[1,6], Malgorzata Nowicka[2,3], Silvia Guglietta[4], Sabrina Schindler[5], Felix J Hartmann[1], Lukas M Weber[2,3], Reinhard Dummer[5], Mark D Robinson[2,3], Mitchell P Levesque[5,7] & Burkhard Becher[1,7]

Immune-checkpoint blockade has revolutionized cancer therapy. In particular, inhibition of programmed cell death protein 1 (PD-1) has been found to be effective for the treatment of metastatic melanoma and other cancers. Despite a dramatic increase in progression-free survival, a large proportion of patients do not show durable responses. Therefore, predictive biomarkers of a clinical response are urgently needed. Here we used high-dimensional single-cell mass cytometry and a bioinformatics pipeline for the in-depth characterization of the immune cell subsets in the peripheral blood of patients with stage IV melanoma before and after 12 weeks of anti-PD-1 immunotherapy. During therapy, we observed a clear response to immunotherapy in the T cell compartment. However, before commencing therapy, a strong predictor of progression-free and overall survival in response to anti-PD-1 immunotherapy was the frequency of CD14$^+$CD16$^-$HLA-DR^{hi} monocytes. We confirmed this by conventional flow cytometry in an independent, blinded validation cohort, and we propose that the frequency of monocytes in PBMCs may serve in clinical decision support.

Immunotherapy with anti-PD-1 aims to block the interaction of tumor-reactive T cells with PD-1 ligands (PD-L1 and PD-L2), which are expressed on various cell types, including leukocytes and the tumor cells themselves[1]. Clinical trials for PD-1 and PD-L1 blockade for patients with advanced melanoma have demonstrated consistent therapeutic responses, thus prompting their application to several other cancers[2–8].

Despite these encouraging results, clinical outcomes remain highly variable; only a fraction of patients show durable responses, some with early progression and others with a late response, whereas the majority of treated patients show no beneficial clinical response[2,9]. Reliable criteria to discriminate responders from nonresponders before the initiation of treatment are urgently needed. Predictive biomarkers would allow for the selection of patients who are more likely to respond and to provide potential nonresponders with alternative, perhaps more efficacious, therapeutic options. Some recent reports have shown the use of single-cell analysis to evaluate the expression of PD-1 and of downstream signaling molecules on tumor-infiltrating and circulating CD8$^+$ T cells with the aim of identifying such predictive biomarkers[10]. However, these approaches are hampered by the limited accessibility of patient material, the small number of parameters, overfitting due to the absence of independent cohorts for validation, and lack of systematic, unbiased bioinformatics pipelines, which has resulted in a paucity of predictive biomarkers to date[11].

Here we used peripheral blood mononuclear cells (PBMCs) from patients with metastatic melanoma before and during therapy as a readily accessible and minimally invasive biopsy to probe immune signatures associated with responsiveness to anti-PD-1 immunotherapy[12]. High-dimensional, single-cell mass cytometry was used along with optimized immune marker panels and a customized, interactive bioinformatics pipeline to generate a thorough analysis of the peripheral blood immune cells in an effort to identify a responsiveness-associated predictive signature.

RESULTS

Stratification of therapy response using mass cytometry

We performed the initial analysis with 40 cryopreserved PBMC samples isolated from the blood of a cohort of 20 patients with melanoma before and after initiation of anti-PD-1 immunotherapy, as well as 20 samples from ten age- and sex-matched healthy donors (total $n = 60$). Baseline samples and samples obtained after 12 weeks of anti-PD1 therapy originated from the same patients (**Supplementary Fig. 1** and **Supplementary Table 1**).

For the 'cytometry by time of flight' (CyTOF) analysis, frozen PBMCs were thawed and stained (**Supplementary Fig. 1** and **Supplementary Table 2**) using three separate and partially overlapping mass cytometry panels—one for the phenotypic characterization of lymphocytes, one for T cell function and one specifically for the in-depth characterization of myeloid cells. The first staining panel contained 30 leukocyte markers to identify all of the major immune cell populations and cover all stages of T cell differentiation and activation (**Supplementary Table 2**). After acquisition of the samples, each sample was de-barcoded using

[1]Institute of Experimental Immunology, University of Zurich, Zurich, Switzerland. [2]Institute of Molecular Life Sciences, University of Zurich, Zurich, Switzerland. [3]Swiss Institute of Bioinformatics (SIB), University of Zurich, Zurich, Switzerland. [4]Department of Experimental Oncology, European Institute of Oncology, Milan, Italy. [5]Department of Dermatology, University Hospital Zurich, Zurich, Switzerland. [6]Present address: Department of Microbiology and Immunology and Department of Dermatology, Hollings Cancer Center, Medical University of South Carolina, Charleston, South Carolina, USA. [7]These authors jointly directed this work. Correspondence should be addressed to C.K. (kriegc@musc.edu), M.P.L. (mitchell.levesque@usz.ch) or B.B. (becher@immunology.uzh.ch).

Received 1 May 2017; accepted 20 November 2017; published online 8 January 2018; doi:10.1038/nm.4466

Boolean gating. Staining quality was evaluated by defining a biological positive and negative control (**Supplementary Fig. 2**).

After data pre-processing (Methods), we performed hierarchical clustering on normalized (per batch) median marker expression on live cells that expressed the leukocyte common antigen (CD45) for every patient before and after therapy initiation (**Fig. 1a**). Our analysis showed two major clades of samples—one (left branch of dendrogram) that contained 15 samples, of which all were responders, and another (right branch) that consisted of 25 samples, 18 of which (72%) were nonresponders (**Fig. 1a**). Thus, normalized median marker expression was sufficient to separate most of the responders from the nonresponders. This unbiased clustering approach stratified the patients into responders and nonresponders before therapy, which encouraged us to perform a more in-depth analysis of the data set.

Altered T cell memory before therapy in responders

We began our comparative analysis between responders and nonresponders, before and after therapy initiation, by investigating the

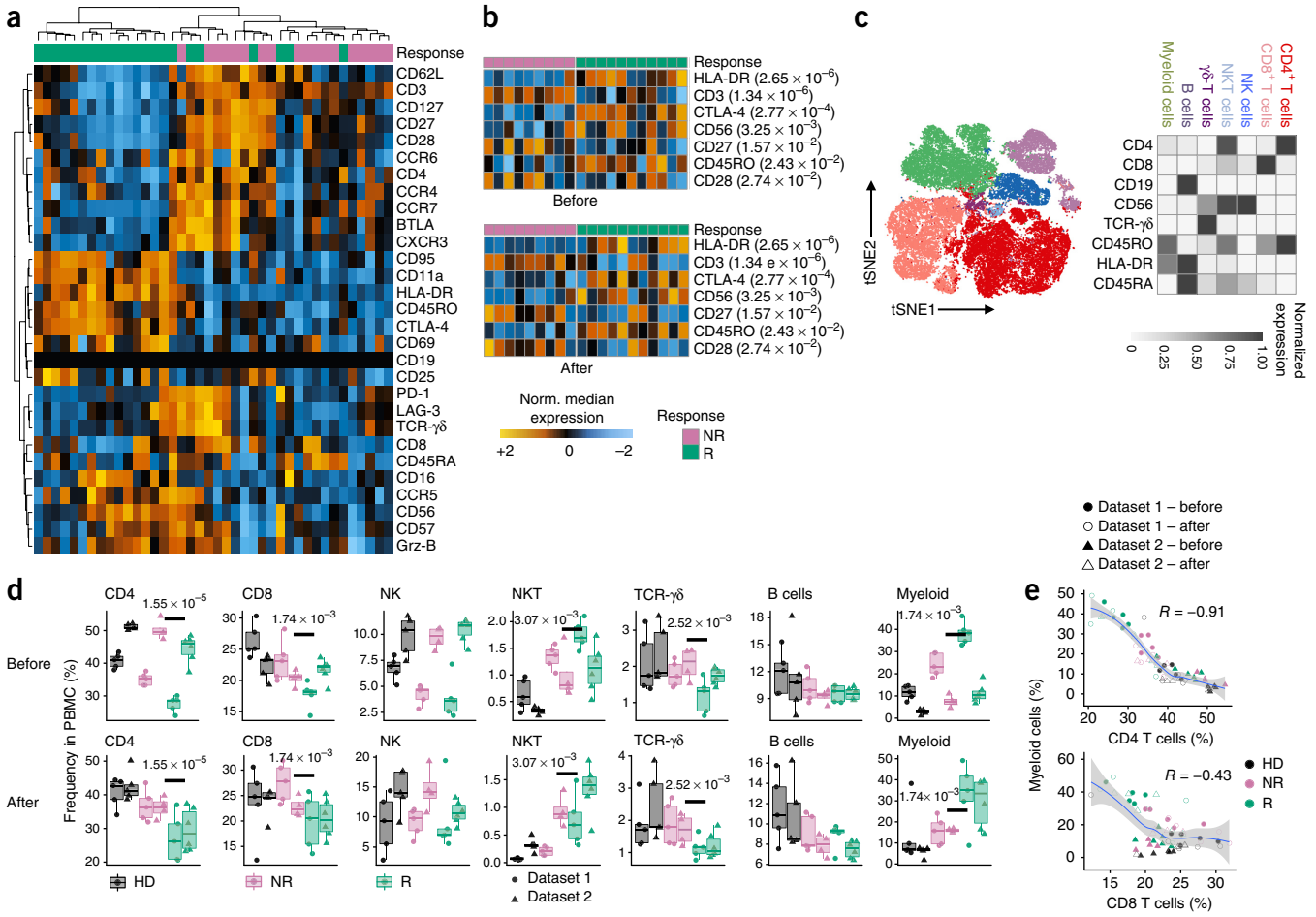

Figure 1 Stratification of responders and nonresponders, and identification of differences in immune cell populations using mass cytometry. (**a**) Heat map of median arcsinh-transformed marker expression normalized per batch to a mean of 0 and a s.d. of 1. Orange indicates relative marker overexpression, blue indicates relative underexpression. The median expression was calculated on single, live CD45$^+$ cells from thawed PBMC samples from patients. The dendrograms for markers (rows) and samples (columns) were constructed with hierarchical clustering (Euclidean distance, Ward linkage). Bars at the top of the heat map represent individual samples from responders (green) and nonresponders (pink). Each column represents one patient sample from one time point (HD, $n = 20$ (not shown); patients before treatment, $n = 20$; patients after treatment, $n = 20$; total samples, $n = 60$). (**b**) Heat map of median arcsinh-transformed marker expression normalized per batch to a mean of 0 and a s.d. of 1 for markers that were significantly (adjusted p-value < 0.05) differentially expressed between responders and nonresponders before and 12 weeks after therapy initiation, in pre-processed live single cells as in **a**. Colored bars at the top of the heat map represent individual samples from responders (green) and nonresponders (pink). Numbers in brackets indicate adjusted p-values. R, responders; NR, nonresponders. (**c**) Exemplified tSNE visualization of 45,000 merged events from nonresponders (NR, $n = 10$) or responders (R, $n = 10$), before or after therapy and from healthy donors (HD, $n = 10$) from data set 1 (discovery cohort;total $n = 30$; for discovery cohort design see **Supplementary Fig. 1**). Cells are colored according to the cluster they were assigned to using the FlowSOM algorithm and manual annotation. The heat map represents the median arcsinh-transformed marker expression normalized to a 0 to 1 range of respective markers within the seven cellular clusters from data set 1 and was used to annotate clusters. (**d**) Direct comparison of cluster frequencies in healthy donors (HD, black), nonresponders (NR, pink) and responders (R, green) in combined data sets (dataset 1 and dataset 2, $n = 30$ each, total $n = 60$; **Supplementary Fig. 1**). Numbers indicate adjusted p-values. (**e**) Relationship between CD4 or CD8 T cell frequency and myeloid cell frequency. R indicates the Spearman correlation between these quantities ($n = 60$). Box plots represent the interquartile range (IQR), with the horizontal line indicating the median. Whiskers extend to the farthest data point within a maximum of 1.5× IQR. All p-values were calculated using two-sided t-tests and were corrected for the multiple comparison using the Benjamini–Hochberg adjustment.

global (measured based on all of the cells) median expression of the 29 leukocyte markers (**Fig. 1b**). Higher expression of HLA-DR, CTLA-4, CD56 and CD45RO, and lower amounts of CD3, CD27 and CD28, were observed in responders than in nonresponders before therapy.

Next we sought to identify which cell populations best described the differences between responders and nonresponders in terms of cell frequency. Markers used for subsequent cell clustering were selected using the principal component analysis (PCA) informativeness score[13], and cells were clustered using the FlowSOM algorithm[14,15] with consensus clustering, separately for each of the measurements rounds. A two-dimensional *t*-stochastic neighbor embedding (tSNE) projection[16] was used for visualization, along with heat maps of normalized (0–1) median marker expression in each cluster (**Fig. 1c**). On the basis of the median marker intensities observed in the clusters, we manually annotated the seven major cell populations (CD4 T cells, CD8 T cells, NK cells, NKT cells, B cells, γδ T cells and myeloid cells) in both measurement rounds. We subsequently examined differences in frequencies of the identified clusters between the three groups (healthy donors (HD), nonresponders (NR) and responders (R)) (**Fig. 1d**) using generalized mixed models (see workflow[17] and **Supplementary Methods**). Of note, the barcoding allowed the model to track the patients and match baseline and on-treatment samples for the analysis. In responders, the frequency of $CD4^+$ T cells and $CD8^+$ T cells was lower, whereas the frequency of $CD19^-HLA\text{-}DR^+$ myeloid cells was significantly higher (adjusted *p*-values = 1.55×10^{-5}, 1.74×10^{-3}, and 1.74×10^{-3}, respectively) than that in the nonresponders (before and after the start of treatment; **Fig. 1d**,**e**). We also observed a higher frequency of NKT cells and a lower frequency of γδ T cells (adjusted *p*-values = 3.07×10^{-3} and 2.52×10^{-3}, respectively) in responders than in nonresponders at both time points. To evaluate whether the immune status in the blood of responders might be reflected in the tumor environment, we analyzed 23 matched tumor biopsies within our cohort. Specifically, we assessed the infiltration by T cells (CD3, CD8 and CD4), the number of phagocytes (CD68 or CD163), as well as the expression of PD-L1 (**Supplementary Fig. 3a**,**b**). As expected from their response to checkpoint blockade, we found that patients who were classified as responders had higher numbers of infiltrating $CD4^+$ T cells and $CD8^+$ T cells (**Supplementary Fig. 3a**).

Because T cells are described to be the major target of anti-PD-1 immunotherapy, and given the altered T cell composition in responders before immunotherapy, we next compared the median marker expression on $CD4^+$ and $CD8^+$ T cells (isolated *in silico* from the sets displayed in **Fig. 1c**,**d**) between nonresponders and responders before and after the initiation of therapy. Global differential marker expression on merged samples in $CD4^+$ T cells in responders showed an upregulation of the activation markers CTLA-4, HLA-DR, CD69 and BTLA before and after therapy (**Fig. 2a**). Similarly, $CD8^+$ T cells in responders showed a higher expression of CD45RO, CTLA-4, CD62L, CD69, CD11a and CCR4 (**Fig. 2b**). This finding, together with the enrichment of $CD8^+$ T cells in the tumor biopsies, suggests a higher migratory capacity of $CD8^+$ T cells in patients classified as responders.

To determine whether there were differences in T cell subpopulations, we re-ran the FlowSOM analysis on the extracted $CD4^+$ T cells and $CD8^+$ T cells and subdivided these populations into $CD45RO^-CD62L^+$ naive cells, $CD45RO^-CD62L^-$ effector (TE) cells, $CD45RO^+CD62L^-$ effector memory (EM) cells, $CD45RO^+CD62L^+$ central memory (CM) cells or $CD127^-CD25^+$ regulatory T (T_{reg}) cells by manual annotation[17]. We then compared the frequencies of resultant T cell subclusters between responders and nonresponders before and 12 weeks after the start of therapy (**Fig. 2c**,**d**). The patients who eventually responded to therapy showed a significantly lower frequency of circulating $CD4^+$ EM T cells, as well as a lower frequency of $CD8^+$ naive T cells at baseline and after the start of treatment (adjusted *p*-values = 8.21×10^{-3}, 6.95×10^{-3}, respectively). Additionally, the $CD8^+$ T cell subpopulation of responders had a higher frequency of CM T cells before and after treatment initiation than the $CD8^+$ T cell subpopulation of nonresponders. To refine our analysis on $CD8^+$ T cells, which have previously been associated with a response to anti-PD-1 immunotherapy[18,19], we analyzed the phenotype of $CD8^+$ T cell clusters at higher resolution using 100 FlowSOM clusters before and after therapy initiation (**Supplementary Fig. 4**). Only two clusters were found to be differentially abundant before therapy, whereas ten clusters were expanded in responders after the start of therapy, which clearly shows a therapy-induced proliferative burst. Cluster 16 ($CD45RO^+CD27^+HLA\text{-}DR^+$) closely resembles the $CD8^+$ CM T cell subpopulation, which previously has been described to expand during anti-PD1 immunotherapy[18]. In agreement with previous research[18], we also found an increase in T_{reg} cells in patients with cancer as compared to that in healthy controls (**Fig. 2c**).

Anti-PD-1 treatment alters T cell properties

To compare the functional properties of T cells between nonresponders and responders, we designed a second mass cytometry panel to investigate cytokine production (**Supplementary Fig. 5**) in polyclonally activated cells. PBMCs were processed as described above. To analyze T cell function and T cell phenotype independently, we first used cell surface marker expression (CD45RO, CD45RA, CCR7, CD28, CD127, CD69 and CD25) to define T cell subpopulations (as in **Fig. 2c**,**d**). Independently of the surface markers, we extracted the frequencies of T cells that expressed cytokines (IL-2, IL-4, IL-10, IL-13, IL-17A, GM-CSF, TNF-α, IFN-γ and granzyme (Grz)-B), PD-1 or CTLA-4 (**Figs. 3** and **4**). Of note, we found a modest increase in the frequencies of $CD4^+$ T cells that expressed IL-4, Grz-B, IFN-γ or GM-CSF and of $CD8^+$ T cells that expressed CTLA-4, Grz-B or IL-13 in responders versus nonresponders before therapy (data not shown). However, after the start of therapy, there were higher numbers of $CD4^+$ T cells expressing PD-1, IL-4, IFN-γ, IL-10, IL-17A and Grz-B in responders than in nonresponders (**Fig. 3a–d**). For $CD8^+$ T cells, an upregulation of CTLA-4 and granzyme B was detected in responders as compared to that in nonresponders (**Fig. 4a–d**). We next created a binary matrix containing all possible cytokine and PD-1 and CTLA-4 combinations (based on marker-specific cutoffs; see **Supplementary Methods** for further details) in $CD4^+$ or $CD8^+$ T cells. Cytokine combination groups (CCGs) that were found to be differentially abundant between responders and nonresponders are displayed for $CD4^+$ T cells or $CD8^+$ T cells in **Figures 3c** and **4c**, respectively. Using this approach, we found three CCGs in the $CD4^+$ T cell subsets that were observed at a higher frequency in responders than in nonresponders (**Fig. 3c**). For $CD8^+$ T cells, we found 13 CCGs that were expanded, and one CCG that was reduced, in responders when compared to nonresponders (**Fig. 4c**). These CCGs were then linked back and displayed in the context of the initial T cell subpopulation. Among the enlarged CCGs in the $CD4^+$ T cells, the most common signature was $CTLA\text{-}4^+$, $granzyme\text{-}B^+$, $TNF\text{-}\alpha^+$ and $PD\text{-}1^+$; thus, the phenotype of these CCGs resembled that of CM T cells (**Fig. 3d**). As for the $CD8^+$ T cells, CCG 1, which was the only one downregulated in frequency when comparing responders to nonresponders, expressed no cytokines, and its cell surface phenotype resembled that of TE cells (**Fig. 4d**). The only other cell population resembling TE cells

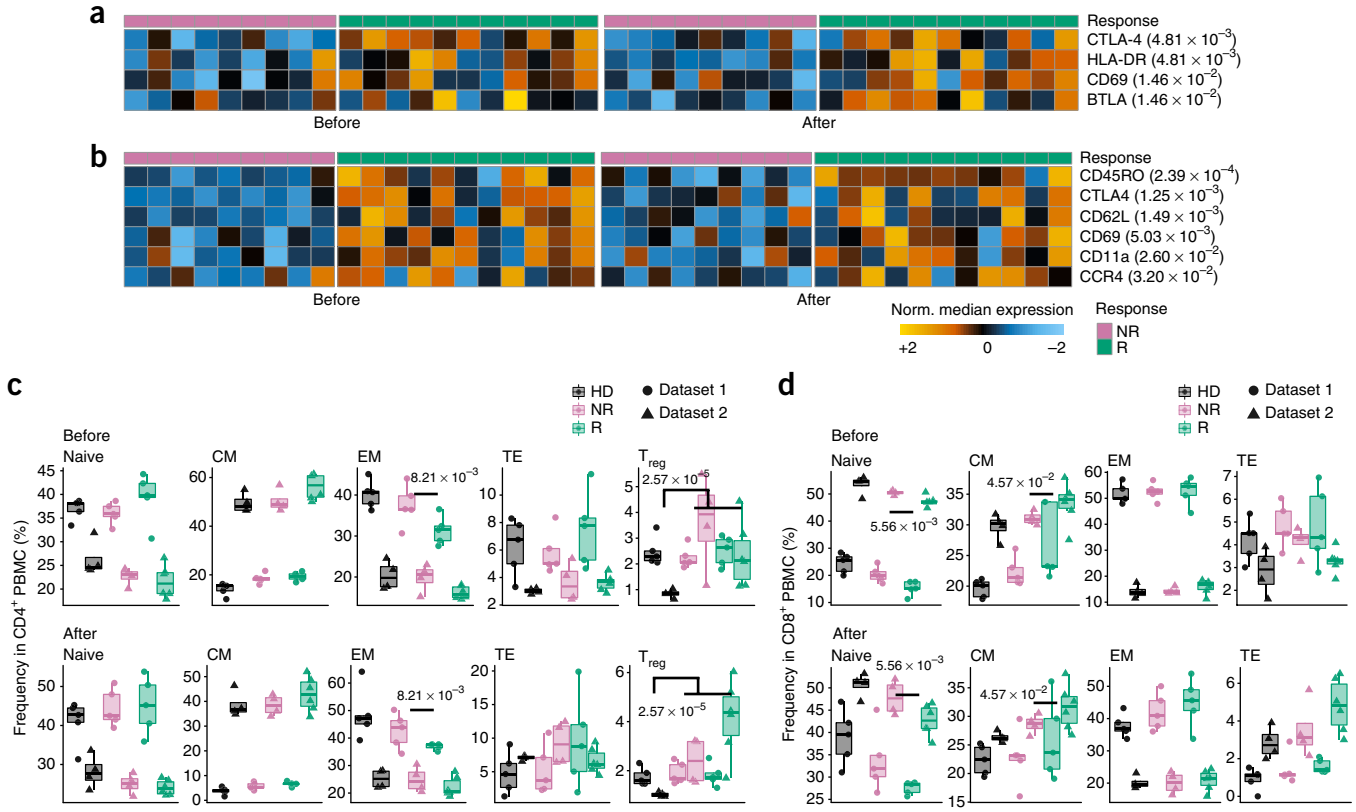

Figure 2 Differences in T cell activation status and in the frequency of the T cell subpopulations before and after 12 weeks of anti-PD-1 therapy in responders and nonresponders. (**a**,**b**) Heat map of per-batch-normalized (a mean of 0 and a s.d. of 1) median expression for markers that were significantly (adjusted p-value < 0.05) differentially expressed in CD4$^+$ T cells (**a**) and CD8$^+$ T cells (**b**) between responders (green bar on top; $n = 11$) and nonresponders (pink bar on top; $n = 9$) before and after the initiation of therapy. Numbers in parentheses show adjusted p-values. (**c**,**d**) FlowSOM was used to generate the indicated CD4$^+$ (**c**) and CD8$^+$ (**d**) T cell subpopulations, and the resulting cluster frequencies are shown. Numbers indicate the significance-adjusted level of differences in cell frequencies between nonresponders ($n = 18$) and responders ($n = 22$) before and after treatment, except for T_{reg} cells, for which the difference between healthy donors (HD; $n = 18$, one sample excluded) and patients with cancer (R and NR) was tested (dataset 1 and dataset 2, $n = 30$ each). Box plots represent the IQR, with the horizontal line indicating the median. Whiskers extend to the farthest data point within a maximum of 1.5× IQR. All p-values were calculated using two-sided t-tests and were corrected for the multiple comparison using the Benjamini–Hochberg adjustment.

was CCG 14, which was granzyme-B$^+$, TNF-α$^+$, IFN-γ$^+$ and IL-2$^+$. The remaining 12 CCGs were expanded and were positive for at least two marker and cytokine combinations; all 12 were CTLA4$^+$ and displayed a CM and TM or EM surface phenotype (**Fig. 4d**). CTLA4$^+$ CCGs could be separated into four groups that were distinguishable by their cytokine and PD-1 expression profiles: CCGs 4, 5, 12, 13, 21 and 28 were granzyme-B$^+$, CCGs 44 and 45 were TNF-α$^+$ or IL-2$^+$, CCGs 22, 30 and 36 produced IL-2 or IL-17, and CCG 43 neither produced cytokines nor expressed PD-1. Surface phenotype and cytokine profiles of the previously identified Grz-B$^+$ CCGs overlapped with previously published CD45RA$^-$Grz-B$^+$CD8$^+$ T cell populations that were expanded during anti-PD-1 immunotherapy[18,19].

Myeloid cell frequencies predict anti-PD-1 responsiveness

Because we found higher frequencies of myeloid cells in anti-PD-1 therapy responders before therapy (**Fig. 1d**), we interrogated a third myeloid-centric panel (**Supplementary Fig. 6**). FlowSOM analysis was used to separate seven subpopulations, which were annotated as T cells, B cells, NK cells, CD14$^+$ (CD11b$^+$HLA-DR^{hi}) myeloid cells, CD14$^-$ (CD11b$^+$HLA-DR^{lo}) myeloid cells, classical CD1c$^+$CD11c$^+$HLA-DR$^+$ dendritic cells (cDCs) and plasmacytoid dendritic cells (CD123$^+$CD303$^+$HLA-DR$^+$CD11c$^-$; pDCs). For the annotated clusters, cell frequencies were calculated in each sample, and the composition of the individual samples was plotted (**Fig. 5a** and **Supplementary Fig. 7**). As already shown in **Figure 1d**, a significantly lower frequency of T cells (p-value = 1.59×10^{-3}) and a higher frequency of CD14$^+$ myeloid cells was observed in responders than in nonresponders (p-value = 5.82×10^{-3}) (**Supplementary Fig. 7**). Although we found that the number of CD14$^-$ myeloid cells was indeed higher in patients with cancer than in healthy donors, we did not observe differences in this cell population between responders and nonresponders. Next, to better characterize the myeloid cells, we extracted live myeloid cells by manually gating out T cells (CD3$^+$) and B cells (CD19$^+$) cells and excluding NK cells (CD7$^+$ and CD56$^+$) from further analysis. Unsupervised clustering of normalized median marker expression in myeloid cells again separated patients into two distinct clusters, with one clade being mostly composed of nonresponders (12/14; 86%) and the other clade being mostly composed of responders (19/25; 76%) (**Fig. 5b**).

We next searched for changes in global median marker expression between nonresponders and responders before and 12 weeks after the start of therapy, and we found that 16 markers were significantly higher in the myeloid compartment of responders as compared to that in nonresponders (**Fig. 5c**). Of note, the enhanced frequency of IFN-γ-producing T cells correlated with the expansion of the

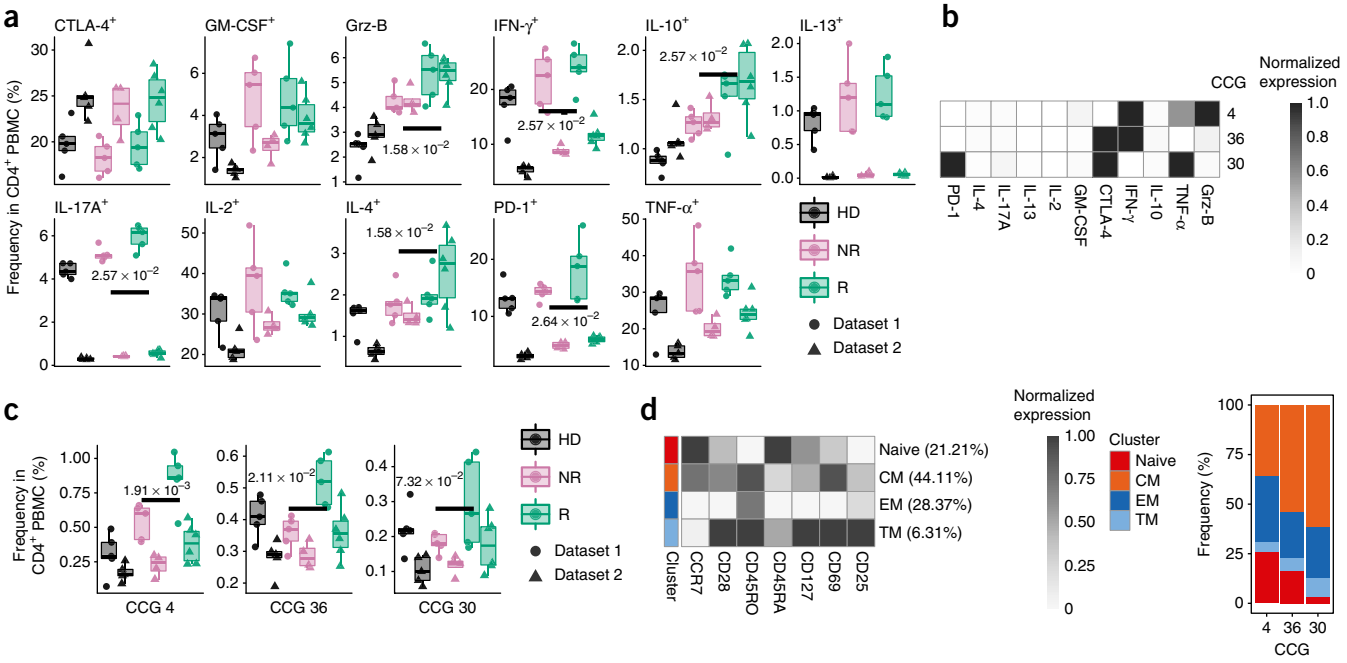

Figure 3 Increased activation in CD4+ T cells after the initiation of immunotherapy in responders. PBMC collected from patients after the initiation of therapy were cultured for 4 h with the stimulants PMA and ionomycin. (**a**) Frequencies of CD4+ T cells expressing PD-1, CTLA-4 or the indicated cytokines in the PBMC of responders (green, $n = 11$) and nonresponders (pink, $n = 9$). Healthy subjects (HD; $n = 10$) served as controls. Numbers indicate the significance level of differences in cell frequencies between the NR and R groups after the start of treatment (adjusted p-values). (**b**) A matrix combining all cytokine-positive cells into CCGs for CD4+ T cells was created, and cells were sorted into this matrix using FlowSOM. Shown are the significantly different CCGs after comparing responders to nonresponders. (**c**) Frequencies of significantly different CCGs between responders and nonresponders from the matrix in **b** (numbers indicate adjusted p-values). (**d**) Heat map (left) and phenotypic composition (right) of the significantly different CCGs for CD4+ T cells. Box plots represent the IQR, with the horizontal line indicating the median. Whiskers extend to the farthest data point within a maximum of 1.5× IQR. All p-values were calculated using two-sided t-tests and were corrected for the multiple comparison using the Benjamini–Hochberg adjustment.

myeloid compartment and PD-L1 expression after therapy initiation (**Supplementary Fig. 8**). Next, FlowSOM analysis was used to subdivide the myeloid compartment into four major clusters, which were annotated as classical monocytes (CD14+CD16−HLA-DR^{hi}), CD14− myeloid cells, pDCs and cDCs (**Fig. 6a**). For higher resolution and to identify a core myeloid signature, 100 FlowSOM clusters were run on all cells, and CD14+ myeloid cells were extracted. Clusters 8, 9 and 18 differed in the expression of some markers, such as CD14, CD11b and PD-L1, but were still elevated in responders versus nonresponders, showing that they were a relatively homogenous cell population (**Supplementary Fig. 9**).

Identification of a monocyte signature using CellCnn

FlowSOM analysis allowed us to identify and characterize CD14+CD16b−HLA-DR^{hi} monocytes as being elevated in responders as compared to that in nonresponders before anti-PD-1 immunotherapy. To identify a core myeloid signature within CD14+CD16b−HLA-DR^{hi} cells that would allow us to predict responsiveness to anti-PD-1 immunotherapy without prior assumptions, we used the machine-learning algorithm CellCnn, which is based on a representation learning approach using convolutional neural networks and is designed to detect rare cell populations associated with disease status[20]. In a data-driven way, CellCnn automatically 'learns' combinations of markers ('filters', which do not need to correspond to known populations), whose presence or frequency discriminates between two groups. We ran CellCnn on all cells of the baseline samples with staining panel 3 and identified a signature population with a relative abundance of 4.8% ± 2.0% (mean ± s.d.) in responders as compared to 2.4% ± 1.5% in nonresponders (**Supplementary Fig. 10**). Although the variability within each group was relatively large, we found the difference in abundance to be statistically significant ($P < 0.01$). In terms of marker expression, we found that this automatically detected population contained a core signature of CD14+CD33+HLA-DR^{hi}ICAM-1+CD64+CD141+ CD86+CD11c+CD38+PD-L1+CD11b+ monocytes (**Supplementary Fig. 10**). Back-projection of this cell population to the tSNE map, independently generated from all cells in panel 3, showed a notable overlap with the CD14+CD16−HLA-DR^{hi} cluster (**Supplementary Fig. 11**).

Characterization of the myeloid compartment by RNA sequencing

To determine whether there were cell-intrinsic changes in the monocyte signature, we performed RNA sequencing (RNA-seq) analysis on sorted CD14+CD16−HLA-DR^{hi} cells from healthy donors, nonresponders and responders at baseline. The most notable differences in gene expression were observed between patients with cancer and healthy donors. In particular, we found an enrichment in genes associated with metabolism, migration and inflammation in patients with cancer (**Fig. 6b**). We found no significant difference in gene expression between CD14+CD16−HLA-DR^{hi} cells of nonresponders and responders, with the exception of the chemokine-encoding gene *CXCL2*, whose expression was increased in responders

(**Supplementary Fig. 12**). These data suggest that the frequency of circulating myeloid cells before therapy rather than gene expression patterns or monocyte polarization correlates with responsiveness to anti-PD1 immunotherapy.

Signature validation by Citrus and flow cytometry

To independently validate the computational results, we used Citrus—a clustering-based supervised algorithm that identifies stratifying signatures—to compare the identified cell types and marker expression

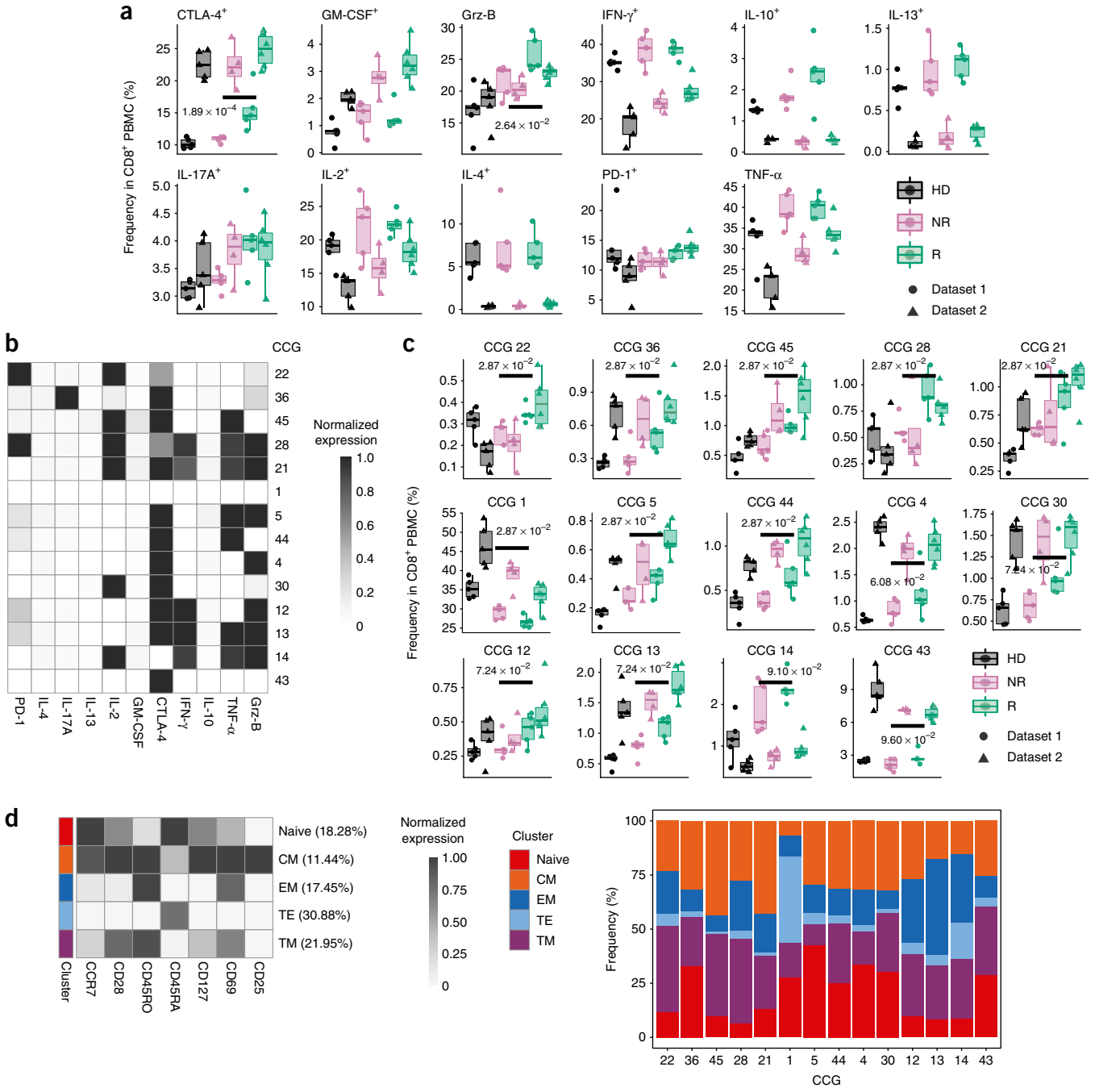

Figure 4 Increased activation in CD8+ T cells after the initiation of immunotherapy in responders. PBMCs collected from patients after the initiation of therapy were cultured for 4 h with PMA and ionomycin. (**a**) Frequencies of CD8+ T cells expressing PD-1, CTLA-4 or the indicated cytokines in the PBMC of responders (green, $n = 11$) and nonresponders (pink, $n = 9$). Healthy subjects (HD, $n = 10$) served as controls. Numbers indicate the adjusted significance level of differences in cell frequencies between the NR and R groups after treatment initiation. (**b**) A matrix combining all cytokine-positive cells into CCGs for CD8+ T cells was created, and cells were sorted into this matrix using FlowSOM. Significantly different CCGs after comparing responders to nonresponders are shown. (**c**) Frequencies of significantly different CCGs between responders and nonresponders from the matrix in **b** (numbers show respective adjusted *p*-values). (**d**) Heat map (left) and phenotypic composition (right) of the significantly different CCGs for CD8+ T cells. Box plots represent the IQR, with the horizontal line indicating the median. Whiskers extend to the farthest data point within a maximum of 1.5× IQR. All *p*-values were calculated using two-sided *t*-tests and were corrected for the multiple comparison using the Benjamini–Hochberg adjustment.

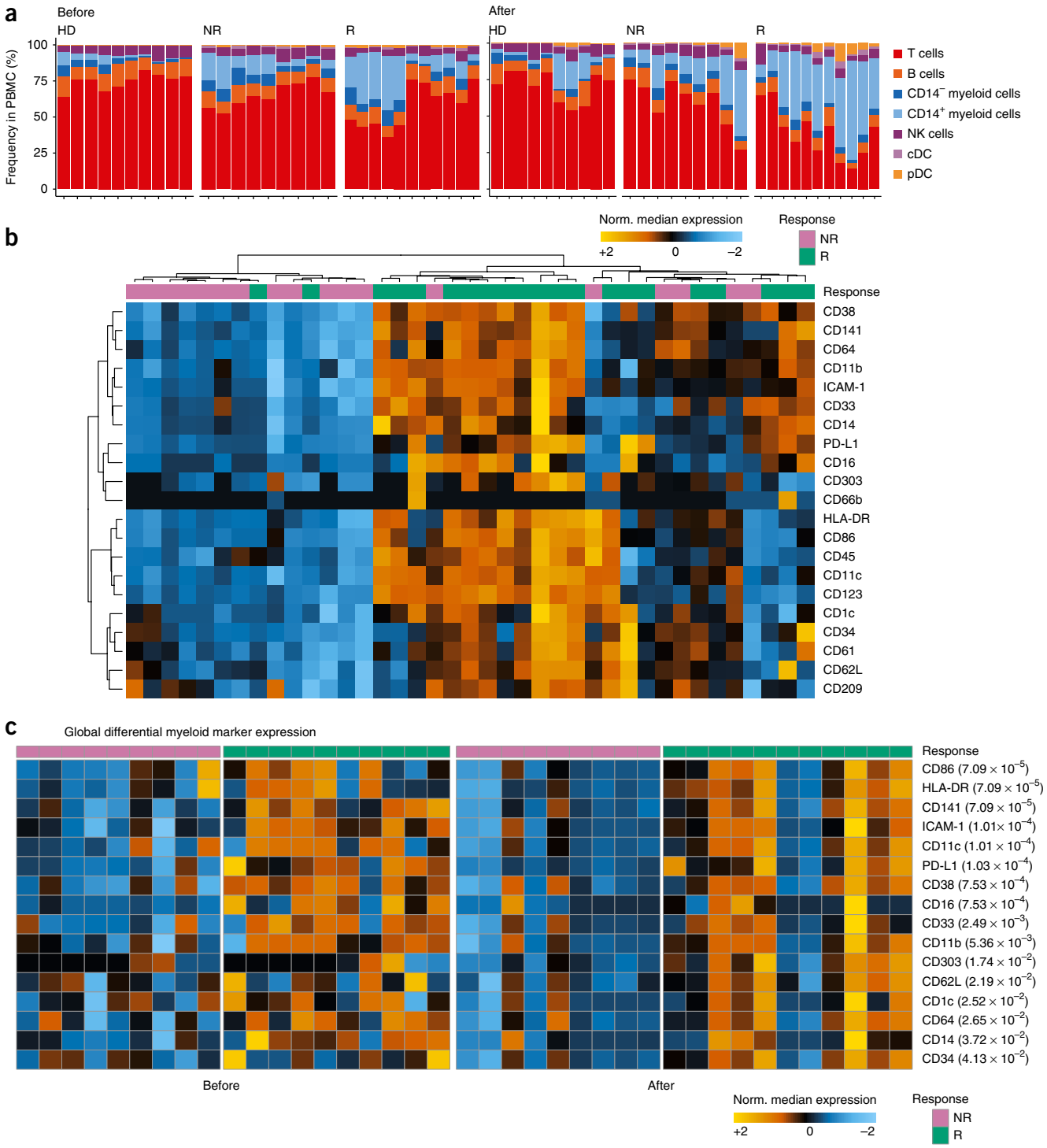

Figure 5 Patient stratification based on myeloid cell markers and expansion of classical monocytes in responders. (**a**) Cell population composition in the PBMCs of healthy donors (HD, $n = 20$), nonresponders (NR, $n = 18$) and responders (R, $n = 22$) before and after the initiation of anti-PD-1 therapy. (**b**) Heat map for patient stratification on the basis of normalized median marker expression, with dendrograms for expression markers (rows) and samples (columns). Bars at the top of the heat map represent individual samples from responders (green) and nonresponders (pink) from the two time points i.e., before or after therapy ($n = 39$; one baseline sample with a cell count <50 was excluded from further differential expression analysis). (**c**) Heat map of normalized median expression of markers that were significantly (adjusted p-value < 0.05) differentially expressed between responders and nonresponders before and 12 weeks after therapy initiation, as in **Figure 1c**, in the myeloid compartment ($CD3^{-}CD19^{-}$) ($n = 39$). Bars at the top of the heat maps represent individual samples from responders (green) and nonresponders (pink). Numbers in brackets show adjusted p-values.

differences that could distinguish nonresponders from responders before therapy (**Supplementary Figs. 13** and **14**). Citrus independently confirmed the lower frequency observed in the T cell compartment and the increase in cells of the myeloid compartment before therapy in the T cell and myeloid panel, respectively.

To facilitate the translation of our observations into clinical practice, we designed a flow-cytometry-based validation panel using a reduced number of markers. We selected a combination of markers that were significantly differentially expressed in **Figures 1b** and **6c** and markers that defined the cellular composition in the blood (**Supplementary Fig. 15**). A blinded validation was performed on PBMCs from a second independent cohort of 31 patients with melanoma, which contained 15 responders and 16 nonresponders before anti-PD-1 therapy (**Supplementary Table 3**).

The data confirmed the lower frequency of T cells ($CD3^+CD56^-$; $P = 9.52 \times 10^{-3}$) and the higher frequency of $CD14^+$ monocytes

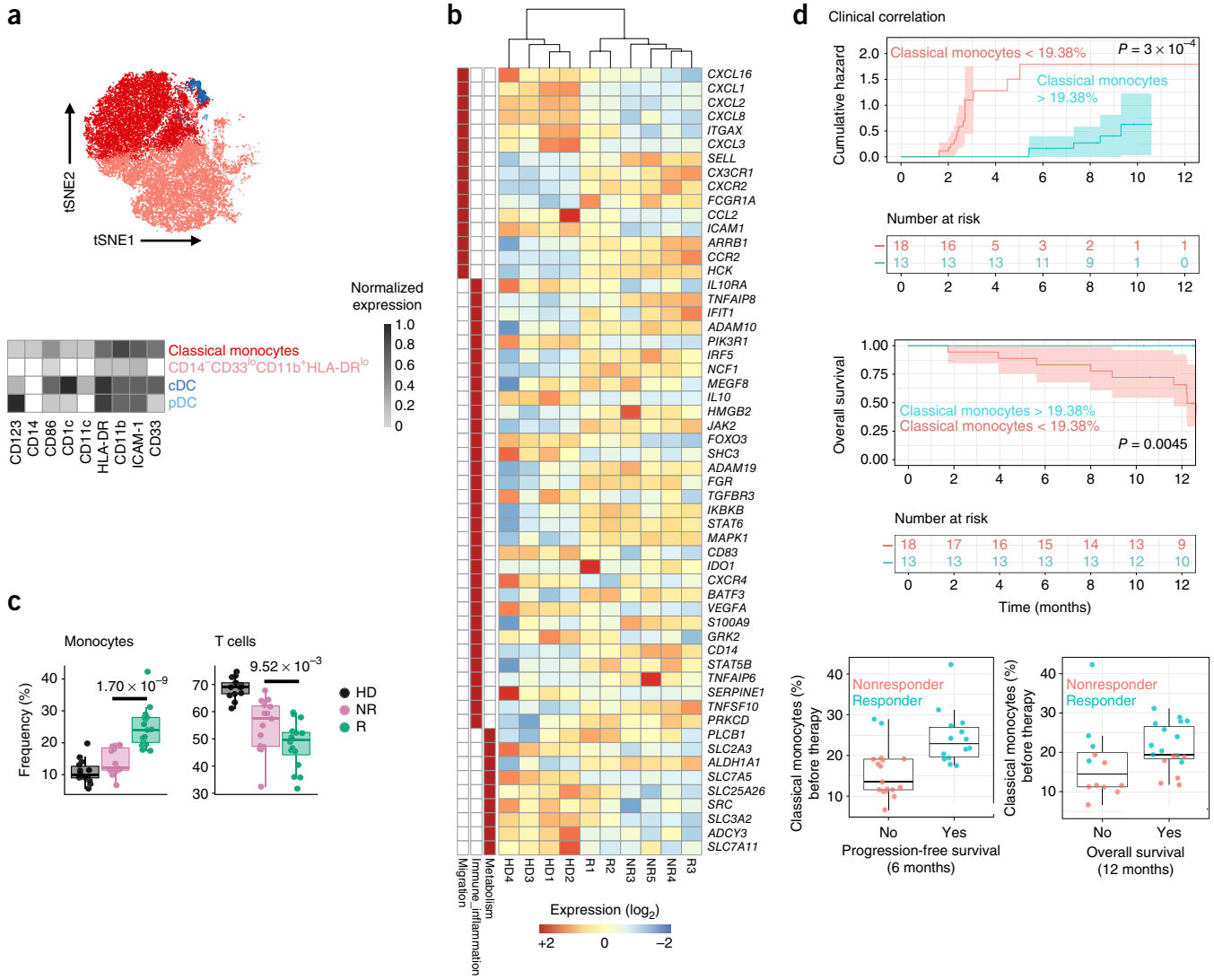

Figure 6 Enhanced activation of classical monocytes in responders and validation by conventional flow cytometry. (**a**) Left, tSNE visualization of FlowSOM-generated myeloid clusters ($CD3^-CD19^-$) in merged data (for CyTOF data analysis see **Supplementary Methods**) from nonresponders (NR, $n = 10$) and responders (R, $n = 10$) before and after therapy and from healthy donors (HD, $n = 30$) from data set 1. $CD7^+$ and $CD56^+$ cells were excluded from analysis. Right, heat map representing the expression of the indicated markers within the cellular clusters. (**b**) Heat map of 56 differentially expressed genes, after RNA-seq analysis of $CD14^+CD16^-HLA\text{-}DR^{hi}$ monocytes that were sorted form healthy donors, responders and nonresponders, comparing HD to patients with cancer (R and NR, $n = 4$ for each group). Expression was quantified with Salmon and are represented as a fold change relative to the median value of gene expression. Differential expression was determined using edgeR. The three columns at the bottom indicate sorting of the genes according to their function (migration, inflammation and metabolism, source: MSigDB). (**c**) Validation of results on a second independent cohort of 31 patients using flow cytometry and the CD3, CD4, CD11b, CD14, CD19, CD16, CD33, CD56 and HLA-DR markers (HD black, NR red, R green). Numbers show adjusted *p*-values, and all *p*-values were calculated using a two-sided *t*-test. Controlling for multiple comparison was done using the Benjamin–Hochberg approach. (**d**) Cumulative hazard and overall survival in patients with monocyte frequencies greater than (blue) or less than (red) 19.38% over time (months). All patients in this study were analyzed ($n = 51$), and *p*-values were determined by the log-rank test. The box plots represent monocyte frequency at baseline against progression-free survival (PFS) at 6 months and overall survival (OS, $P = 0.00092$ by Wilcoxon test) at 12 months in nonresponders (red) and responders (blue, $P = 0.027$ by Wilcoxon test). Box plots represent the IQR with the horizontal line indicating the median. Whiskers extend to the farthest data point within a maximum of 1.5× IQR. Unless otherwise indicated, all *p*-values were calculated using two-sided *t*-tests and were corrected for the multiple comparison using the Benjamini–Hochberg adjustment.

($CD3^-CD19^-CD14^+CD16^-HLA\text{-}DR^+$; $P = 1.70 \times 10^{-9}$) before therapy in responders (**Fig. 6c**). To visualize and assess the survival benefit conferred by a higher frequency of classical monocytes before treatment, we calculated the optimal cutoff point in monocyte frequency, which best stratified responders and nonresponders. The calculated cutoff of 19.38% was then used to compute a cumulative hazard function for the groups of monocytes that were present at a high or low frequency (**Fig. 6d**). The resulting plot shows a clear difference in hazard ratios between patients who had a high frequency or a low frequency of classical monocytes at baseline. Our model thus indicates that a classical monocyte frequency >19.38%, before anti-PD1 therapy initiation, is predictive of a better treatment response and patient survival (**Fig. 6d**).

Monocytes associate with progression-free survival

Finally, using a Cox proportional-hazards model, we assessed the prognostic value of 53 standard clinical parameters (**Supplementary Fig. 16**) plus the frequency of classical monocytes (all measured at the baseline) with progression-free survival (PFS) in all of the 51 patients evaluated in this study. Multiple variables were significant (p-value $< 3.36 \times 10^{-2}$) in the univariate analysis (**Supplementary Fig. 17**). There were no additional variables that were robustly associated with PFS in the multivariate analysis (including all factors with p-values < 0.05 from the univariate analysis), other than immature granulocytes and classical monocytes (**Supplementary Fig. 16b**). The same 53 parameters were tested for association with response (nonresponders versus responders) using linear models (LMs) for continuous parameters and generalized linear models (GLMs) for the binary parameters, and these analyses confirmed classical monocytes to be different (data not shown).

DISCUSSION

The successful therapeutic responses in patients with advanced melanoma have encouraged the application of anti-PD-1 immunotherapy to several other cancers[2,3,5–7,9,21–23]. Despite increasing overall survival in 33–40% of patients with melanoma, anti-PD1 treatment has not been found to be effective in the majority of treated patients and has resulted in disease progression at a median follow-up of 21 months in only 25% of patients[24,25]. Moreover, given the broadening of its application, we can anticipate that the rate of patients who are nonresponsive to, or relapsed after, anti-PD-1 therapy will further increase. In this context, the identification of biomarkers allowing for the discrimination of responders and nonresponders before therapy initiation may tailor the application of this treatment to only those patients that are likely to benefit from it, while providing alternative treatments to the patients that are unlikely to show a response. Using single-cell mass cytometry combined with clustering and regression analyses, we searched for differential immune signatures in responders versus nonresponders before therapy. Besides a modest alteration of the lymphocyte compartment before therapy, i.e., lymphopenia of $CD4^+$ and $CD8^+$ T cells, γδ T cells and a slight elevation of NKT cells, we found the frequency of classical $CD14^+CD16^-CD33^+HLA\text{-}DR^{hi}$ monocytes to predict responsiveness to anti-PD-1 immunotherapy.

In recent years, the role of myeloid cells in cancer has been extensively debated, and numerous studies have addressed the role of the so-called myeloid-derived suppressor cells, which have been shown, mainly in preclinical models, to arise during chronic inflammation and cancer[26]. However, the phenotypic, morphological and functional heterogeneity of these cells generates confusion when investigating their roles in anti-cancer immune responses in the clinic. It has been proposed that high frequencies of myeloid cells with immunosuppressive features, defined as $CD33^+CD11b^+HLA\text{-}DR^{lo/-}$ cells, may lead to T cell dysfunction and failure to respond to immunotherapy. Accordingly, a reduction of suppressive myeloid cells correlates with an increase in the objective clinical responses and long-term survival[27–31]. Here we used clustering to define cell populations and found that the frequency of a cell population resembling $CD33^{low}CD11b^+HLA\text{-}DR^{lo}$ myeloid cells showed no differences between responders and nonresponders and did not change during therapy. However, the responding patients' classical monocytes ($CD14^+CD16^-$) had higher amounts of migration and activation markers, such as ICAM-1 and HLA-DR, suggesting that monocytes may sustain the development of an effective anti-tumor immune response during anti-PD-1 immunotherapy, similarly to what has been described for $CD14^-CD16^+$ monocytes during anti-CTLA-4 treatment[32]. Further support for a role of monocytes in anti-tumor immune responses comes from a study in which untreated patients with melanoma with the highest tumor burden harbored dysregulated intermediate ($CD14^+CD16^+$) and nonclassical monocytes ($CD14^-CD16^+$) that were characterized by decreased expression of HLA-DR and inflammatory markers[33]. Moreover, the upregulation of PD-L1 on the monocytes from responders before therapy is likely the result of an ongoing immune response against the tumors and elevated levels of IFN-γ, which directly increases PD-L1 expression[34,35].

Consistent with the notion that the presence of activated classical monocytes may be a prerequisite for a successful response during anti-PD-1 immunotherapy, we reported higher frequencies of central memory T cells and NKT cells in circulation and a more activated ($CTLA\text{-}4^+$, $TNF\text{-}\alpha^+$, $PD\text{-}1^+$, $Grz\text{-}B^+$ and $IL\text{-}2^+$) T cell compartment after therapy initiation in responding patients. This is consistent with the observation that after 3 weeks of anti-PD-1 therapy, T cell proliferation peaks but can still be observed after 12 weeks[18].

Given the shift of frequency from naive to central memory T cells in responders before therapy, and the increase in CTLA-4, IFN-γ, IL-17A, Grz-B and PD-1 after therapy start, our findings confirm that anti-PD-1 immunotherapy supports functionally activated T cells. This is consistent with recent research showing that higher levels of CTLA-4 on intra-tumoral T cells correlates with a better response to anti-PD-1 treatment and that resistance to anti-PD-1 immunotherapy is associated with defects in the pathways of antigen presentation and interferon receptor signaling[34,36]. Indeed, besides being a regulator during T cell expansion, CTLA-4 is also a marker of activated T cells[37]. We found a consistent reduction of T cells in the peripheral blood of responders as compared to that in the blood of nonresponders (**Fig. 2c**,**d**). This phenomenon may be due to their enhanced ability to migrate to the tumor site[38]. Indeed, we also found an upregulation of CD11a in the $CD8^+$ T cell compartment of responder patients, which has been shown to be essential for migration to lymph nodes and distal sites[37,39].

The goal of this study was to interrogate the PBMC compartment for signatures that were predictive of an anti-PD-1 response. We are aware of the variability between the two measurements in the discovery cohort, but by applying our custom bioinformatic workflow, we observed a significant difference in classical monocyte frequencies. Moreover, by using traditional fluorescence flow cytometry analysis, we could confirm this difference in monocyte frequency in responding versus nonresponding patients in a second, blinded, independent cohort of patients undergoing immunotherapy.

In addition to the increase in myeloid cells before therapy initiation, we observed a reduction in T cell frequency to correspond with treatment response and overall survival. The tendency for an increased number of tumor-infiltrating lymphocytes in responders supports the notion that myeloid cell frequency increases when T cells move

from the blood into the tumor, which is consistent with previous studies[36,38]. This phenomenon is—in all likelihood—the result of the increased (although insufficient) antitumor response in patients at baseline, who are more likely to become responders. Indeed, IFN-γ produced by tumor-specific NK cells and T cells can trigger myeloid cell egress from the bone marrow[40,41]. In addition, IFN-γ has been shown to enhance myeloid cell expansion[42]. Altogether, we have provided evidence for a response-associated immune signature in patients with metastatic melanoma undergoing anti-PD-1 immunotherapy.

Future studies entailing a systematic, prospective collection of paired blood and tumor samples are needed to confirm these signatures in larger, multicenter cohorts of patients with melanoma, as well as in patients with other cancer types for which anti-PD-1 treatment has been approved. A prediction signature might then be directly used in clinical practice to stratify patients before initiating immunotherapy.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank V. Tosevski and T.M. Brodie (mass cytometry core facility, University of Zurich), A. Langer (Department of Dermatology, University of Zurich), and C. Beisel and K. Eschbach (Genomics Facility, ETH Basel) for excellent technical assistance and N. Nunes, B. Chatterjee, E. Terskikh, and C. Gujer (all from the Institute of Experimental Immunology, University Zurich), A. Zollinger (Swiss Institute of Bioinformatics, Lausanne), all members of the COST Action BM1404 Mye-EUNITER (http://www.mye-euniter.eu/), and P. Cheng (University of Zurich) for discussions. We also thank C. Guglietta for graphical design and layout. This work received funding from the University Research Priority Program (URPP) in Translational Cancer Research (C.K.), the Swiss National Science Foundation (grants 310030_146130 and 316030_150768; B.B.), the European Union FP7 project ATECT (B.B.), and the European Training Network MELGEN (M.P.L.).

AUTHOR CONTRIBUTIONS

C.K., M.P.L., and B.B. conceived the study and analyzed data; C.K., S.G., and B.B. designed and performed the experiments; F.J.H. and S.G. assisted with the experiments; S.S., R.D., and M.P.L. provided clinical samples and performed statistical analyses of clinical parameters; R.D. and M.P.L. analyzed histology; M.N., L.M.W., and M.D.R. provided analysis algorithms and analyzed data; C.K. and S.G. wrote the manuscript; M.P.L., M.D.R., and B.B. edited the manuscript; and all authors read and gave final approval to submit the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Topalian, S.L., Drake, C.G. & Pardoll, D.M. Targeting the PD-1/B7-H1(PD-L1) pathway to activate antitumor immunity. *Curr. Opin. Immunol.* **24**, 207–212 (2012).
2. Topalian, S.L. *et al.* Safety, activity and immune correlates of anti-PD-1 antibody in cancer. *N. Engl. J. Med.* **366**, 2443–2454 (2012).
3. Powles, T. *et al.* MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. *Nature* **515**, 558–562 (2014).
4. Brahmer, J. *et al.* Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. *N. Engl. J. Med.* **373**, 123–135 (2015).
5. Motzer, R.J. *et al.* Nivolumab versus everolimus in advanced renal cell carcinoma. *N. Engl. J. Med.* **373**, 1803–1813 (2015).
6. Rizvi, N.A. *et al.* Activity and safety of nivolumab, an anti-PD-1 immune checkpoint inhibitor, for patients with advanced, refractory squamous non-small-cell lung cancer (CheckMate 063): a phase 2, single-arm trial. *Lancet Oncol.* **16**, 257–265 (2015).
7. Ansell, S.M. *et al.* PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. *N. Engl. J. Med.* **372**, 311–319 (2015).
8. Center for Drug Evaluation Research. *Approved Drugs—Hematology/Oncology (Cancer) Approvals & Safety Notifications* (US Food and Drug Administration, 2016).
9. Hamid, O. *et al.* Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. *N. Engl. J. Med.* **369**, 134–144 (2013).
10. Nishino, M., Ramaiya, N.H., Hatabu, H. & Hodi, F.S. Monitoring immune-checkpoint blockade: response evaluation and biomarker development. *Nat. Rev. Clin. Oncol.* **14**, 655–668 (2017).
11. Wistuba-Hamprecht, K. *et al.* Establishing high-dimensional immune signatures from peripheral blood via mass cytometry in a discovery cohort of stage IV melanoma patients. *J. Immunol.* **198**, 927–936 (2017).
12. Pérez-Callejo, D., Romero, A., Provencio, M. & Torrente, M. Liquid-biopsy-based biomarkers in non-small-cell lung cancer for diagnosis and treatment monitoring. *Transl. Lung Cancer Res.* **5**, 455–465 (2016).
13. Levine, J.H. *et al.* Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell* **162**, 184–197 (2015).
14. Van Gassen, S. *et al.* FlowSOM: using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A* **87**, 636–645 (2015).
15. Weber, L.M. & Robinson, M. D. Comparison of clustering methods for high-dimensional single-cell flow and mass cytometry data. *Cytometry A* **89**, 1084–1096 (2016).
16. Maaten, L.V.D. & Hinton, G. Visualizing data using tSNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
17. Nowicka, M. *et al.* CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets. *F1000Res.* **6**, 748 (2017).
18. Huang, A.C. *et al.* T cell invigoration to tumor burden ratio associated with anti-PD-1 response. *Nature* **545**, 60–65 (2017).
19. Kamphorst, A.O. *et al.* Proliferation of PD-1$^+$ CD8 T cells in peripheral blood after PD-1-targeted therapy in lung cancer patients. *Proc. Natl. Acad. Sci. USA* **114**, 4993–4998 (2017).
20. Arvaniti, E. & Claassen, M. Sensitive detection of rare disease-associated cell subsets via representation learning. *Nat. Commun.* **8**, 14825 (2017).
21. Brahmer, J.R. *et al.* Phase 1 study of single-agent anti-programmed-death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics and immunologic correlates. *J. Clin. Oncol.* **28**, 3167–3175 (2010).
22. Brahmer, J.R. *et al.* Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N. Engl. J. Med.* **366**, 2455–2465 (2012).
23. Weber, J.S. *et al.* Nivolumab versus chemotherapy in patients with advanced melanoma who progressed after anti-CTLA-4 treatment (CheckMate 037): a randomized, controlled, open-label, phase 3 trial. *Lancet Oncol.* **16**, 375–384 (2015).
24. Robert, C. *et al.* Pembrolizumab versus ipilimumab in advanced melanoma. *N. Engl. J. Med.* **372**, 2521–2532 (2015).
25. Ribas, A. *et al.* Association of pembrolizumab with tumor response and survival among patients with advanced melanoma. *J. Am. Med. Assoc.* **315**, 1600–1609 (2016).
26. Ostrand-Rosenberg, S. & Sinha, P. Myeloid-derived suppressor cells: linking inflammation and cancer. *J. Immunol.* **182**, 4499–4506 (2009).
27. Gebhardt, C. *et al.* Myeloid cells and related chronic inflammatory factors as novel predictive markers in melanoma treatment with ipilimumab. *Clin. Cancer Res.* **21**, 5453–5459 (2015).
28. Meyer, C. *et al.* Frequencies of circulating MDSC correlate with clinical outcome of melanoma patients treated with ipilimumab. *Cancer Immunol. Immunother.* **63**, 247–257 (2014).
29. Sade-Feldman, M. *et al.* Clinical significance of circulating CD33$^+$CD11b$^+$HLA-DR$^-$ myeloid cells in patients with stage IV melanoma treated with ipilimumab. *Clin. Cancer Res.* **22**, 5661–5672 (2016).
30. Komohara, Y., Jinushi, M. & Takeya, M. Clinical significance of macrophage heterogeneity in human malignant tumors. *Cancer Sci.* **105**, 1–8 (2014).
31. Zhang, Q.-W. *et al.* Prognostic significance of tumor-associated macrophages in solid tumor: a meta-analysis of the literature. *PLoS One* **7**, e50946 (2012).
32. Romano, E. *et al.* Ipilimumab-dependent cell-mediated cytotoxicity of regulatory T cells *ex vivo* by nonclassical monocytes in melanoma patients. *Proc. Natl. Acad. Sci. USA* **112**, 6140–6145 (2015).
33. Chavan, R. *et al.* Untreated stage IV melanoma patients exhibit abnormal monocyte phenotypes and decreased functional capacity. *Cancer Immunol. Res.* **2**, 241–248 (2014).
34. Zaretsky, J.M. *et al.* Mutations associated with acquired resistance to PD-1 blockade in melanoma. *N. Engl. J. Med.* **375**, 819–829 (2016).
35. 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* **4**, e1008824 (2015).
36. Daud, A.I. *et al.* Tumor immune profiling predicts response to anti-PD-1 therapy in human melanoma. *J. Clin. Invest.* **126**, 3447–3452 (2016).
37. Herbst, R.S. *et al.* Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* **515**, 563–567 (2014).
38. Kluger, H.M. *et al.* Characterization of PD-L1 expression and associated T cell infiltrates in metastatic melanoma samples from variable anatomic sites. *Clin. Cancer Res.* **21**, 3052–3060 (2015).
39. von Andrian, U.H. & Mempel, T.R. Homing and cellular traffic in lymph nodes. *Nat. Rev. Immunol.* **3**, 867–878 (2003).
40. Takizawa, H., Regoes, R.R., Boddupalli, C.S., Bonhoeffer, S. & Manz, M.G. Dynamic variation in cycling of hematopoietic stem cells in steady state and inflammation. *J. Exp. Med.* **208**, 273–284 (2011).
41. Nagai, Y. *et al.* Toll-like receptors on hematopoietic progenitor cells stimulate innate immune system replenishment. *Immunity* **24**, 801–812 (2006).
42. Caux, C., Moreau, I., Saeland, S. & Banchereau, J. Interferon-γ enhances factor-dependent myeloid proliferation of human CD34$^+$ hematopoietic progenitor cells. *Blood* **79**, 2628–2635 (1992).

ONLINE METHODS

Patient samples. Fifty-one cryopreserved peripheral blood mononuclear cells (PBMC) samples from patients with melanoma (time since prior ipilimumab treatment: median, 84 d; range, 23–162 d; average, 87.3 d) were provided by the Department of Dermatology, University Hospital Zurich, Switzerland (**Supplementary Fig. 16** and **Supplementary Tables 1** and **3**). From 20 of these patients, called the discovery cohort, samples before and about 12 weeks after anti-PD-1 immunotherapy initiation were used ($n = 40$). Patients were treated with 3 mg per kg body weight (mg/kg) nivolumab every 2 weeks or 2 mg/kg pembrolizumab every 3 weeks for 12 weeks, at which time their clinical status was assessed again. Response was defined as the patient's disease control rate (DCR) in the course of treatment. That is, the responder group comprised every patient who showed signs of clinical benefit within the first 15 weeks of treatment, which included a partial response (PR), a complete response (CR) and stable disease (SD), thus better capturing 'real-world patients'. The nonresponder group included every patient who discontinued treatment due to disease progression or who showed signs of progression within the first 15 weeks of treatment. Progression was defined as a measurable increase in tumor size, the presence of new metastatic sites or the need to treat the patient with a secondary treatment, such as radiotherapy. The validation cohort was composed of 31 patients with melanoma. Samples from the validation cohort were collected before the initiation of anti-PD-1 therapy. Age- and sex-matched PBMC from healthy donors were isolated from the blood provided by the Red Cross Blood Bank, Zurich, Switzerland. All human biological samples were collected after written informed consent of the patients was obtained and with approval of the local ethics committee (Kantonale Ethikkommission Zürich, KEK-ZH authorization Nr. 2014-0425) in accordance to 'good clinical practice' (GCP) guidelines and the Declaration of Helsinki.

Stimulations, stainings and mass cytometry data acquisition. PBMC stimulations, staining and data acquisition by mass cytometry were performed as described previously[43]. Frozen PBMCs were used in this retrospective study to balance cohorts in terms of response and to reduce batch effects through a unique barcoding strategy. Data were stored using the Flow Repository[44], which can be accessed under at https://flowrepository.org/experiments/1124. The complete methods are included in the **Supplementary Methods**.

Statistical analysis. Data acquired by mass cytometry were normalized using the standalone MATLAB normalizer (version 2013b)[45], marker expression was controlled in FlowJo (version10.1r5), and patient samples were debarcoded using Boolean gating. For further analysis we developed a customized R workflow to discover different biomarkers when comparing marker expression between responders and nonresponders. The workflow is described in the **Supplementary Methods** (CyTOF data analysis), and the R code can be accessed at https://github.com/gosianow/carsten_cytof_code. Additional analyses to identify distinguishing cell subsets were done using CellCnn (**Supplementary Methods** (CellCnn analysis); code available at https://github.com/lmweber/PD1_analysis_CellCnn).

Transcriptomics analysis. Representative samples ($n = 4$ each) of responders, nonresponders and healthy donors were selected from archival samples stored in the dermatology biobank according to the same clinical criteria used in the discovery and validation cohorts for CyTOF and FACS analysis. $CD14^+CD16^-HLA\text{-}DR^{hi}Lin^-$ ($CD3^-CD4^-CD19^-CD45RO^-$) monocytes were sorted from frozen PBMC obtained from blood samples of healthy donors, responders and nonresponders at baseline. RNA was isolated using the RNA Plus Micro kit (Qiagen). We used the SMARTer Stranded Total RNA-seq Pico kit (Clontech) for RNA preparations. Samples were sequenced on an Illumina MySeq instrument. Reads in FASTQ format from the 12 samples were quantified at the transcript level using Salmon[46] against an Ensembl catalog, aggregated to the gene level using tximport[47] and delivered to edgeR[48] for the differential expression analysis using the GLM functionality[49]. To determine differential expression, a false discovery rate (FDR) cutoff of 5% and minimum fold change of 1.5 was used. Gene groups were selected from MSigDB (http://software.broadinstitute.org/gsea/msigdb). Due to the gender bias in responders versus nonresponders, 14 genes from the Y chromosome were removed before differential expression analysis. The RNA-seq data can be found under ArrayExpress accession E-MTAB-6214.

Validation by flow cytometry. Validation of the CyTOF data was done by using a combination of markers with significantly different expression from the initial discovery mass cytometry approach and by markers that defined the cellular composition in blood using flow cytometry (**Supplementary Fig. 9**). A set of PBMCs was analyzed in a blinded fashion from a second, independent cohort of 31 patients with melanoma, which consisted of 15 responders and 16 nonresponders, before anti-PD-1 therapy. The panel is described in the **Supplementary Methods** (validation by flow cytometry). At least 100,000 live cells were acquired using Diva software on a Fortessa flow cytometer (BD) and analyzed using FlowJo software (TriStar). The data are available at https://flowrepository.org/experiments/1124. From the FlowJo data, the frequencies of $CD3^+$ T cells and $CD14^+CD16^-HLA\text{-}DR^{hi}$ monocytes were extracted from the three groups (responders, nonresponders and healthy donors). For statistical testing, we applied a GLM and cut-point calculations as described in the **Supplementary Methods** (validation by flow cytometry).

Life Sciences Reporting Summary. Further information on experimental design is available in the **Life Sciences Reporting Summary**.

Data availability statement. The mass cytometry and flow cytometry data are available at https://flowrepository.org/experiments/1124. The workflow R code can be accessed at https://github.com/gosianow/carsten_cytof_code. Additional analyses code to identify distinguishing cell subsets using CellCnn can be found at https://github.com/lmweber/PD1_analysis_CellCnn. RNA-seq data can be found at ArrayExpress under accession number E-MTAB-6214.

43. Hartmann, F.J. *et al.* High-dimensional single-cell analysis reveals the immune signature of narcolepsy. *J. Exp. Med.* **213**, 2621–2633 (2016).
44. Spidlen, J. & Brinkman, R.R. Use FlowRepository to share your clinical data upon study publication. *Cytometry B Clin. Cytom.* (2016).
45. Finck, R. *et al.* Normalization of mass cytometry data with bead standards. *Cytometry A* **83**, 483–494 (2013).
46. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
47. Soneson, C., Love, M.I. & Robinson, M.D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res.* **4**, 1521 (2015).
48. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
49. McCarthy, D.J., Chen, Y. & Smyth, G.K. Differential expression analysis of multifactor RNA-seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).

doi:10.1038/nm.4466

nature research

Corresponding author(s): Carsten Krieg, Mitchell P. Levesque, Burkhard Becher

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

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

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

Experimental design

1. Sample size

Describe how sample size was determined.

Manuscript: pages 3 (Supplementary Table 1) and 9 (Supplementary Table 3 and Figures 15+16) and page 21 (Online Methods).
Supplement: page 1 (patient samples), Figure 15 and Cox model (Figure 16).
There was no analysis to predetermine sample size. Samples were obtained as a retrospective, exploratory study based archival samples collected through the routine biobanking program in the Department of Dermatology. While the sample size is small, we validated the results from the discovery cohort (number of analyzed biopsies=60) in a second, independent, blinded cohort (n=31).

2. Data exclusions

Describe any data exclusions.

Supplement: page 2: samples with less than 50 cells were excluded....

3. Replication

Describe whether the experimental findings were reliably reproduced.

For the discovery cohort 2 independent batches were prepared and run. Significant differences were confirmed by using an independent validation cohort and conventional flow cytometry.

4. Randomization

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

Manuscript: page 3: barcoding to reduce sample variation; page 9 third paragraph (validation of cellular immune signature by Citrus and conventional flow cytometry). Supplement: page 2 (CyTOF data analysis).
Samples were randomly allocated to batches prior to measurement.

5. Blinding

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

Samples used in the flow cytometry validation cohort were blinded.

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

Nature Medicine: doi:10.1038/nm.4466

6. Statistical parameters

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

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

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

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	FlowJo v10.1, customized R workflow, available under https://github.com/gosianow/carsten_cytof_code with the following setup: R version 3.3.0 with the following Bioconductor packages from version 3.3: Rtsne_0.13, pheatmap_1.0.8, multcomp_1.4-7, lme4_1.1-13, FlowSOM_1.4.0, flowCore_1.38.2, ComplexHeatmap_1.10.2, ConsensusClusterPlus_1.36.0, citrus_0.08, and ggplot2_2.2.1. For the RNA-seq analyses, we used R 3.4.0, edgeR 3.18.1 and tximport 1.4.0

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

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	None

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	A total of >80 commercially available (Fluidigm) or custom labeled (BD, Biolegend, RnD) antibodies for mass cytometry are listed in supplementary table 2 (page 9 of Supplement). Antibodies were tested and test results are displayed in supplementary figures 2, 5, and 6. Further antibodies for conventional flow cytometry and sorting are listed on page 6 of the supplement.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	No eukaryotic cell lines were used
b. Describe the method of cell line authentication used.	No eukaryotic cell lines were used
c. Report whether the cell lines were tested for mycoplasma contamination.	No eukaryotic cell lines were used
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	*Provide a rationale for the use of commonly misidentified cell lines OR state that no commonly misidentified cell lines were used.*

Nature Medicine: doi:10.1038/nm.4466

Animals and human research participants

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

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	No animals were used

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	All of the clinical variables (i.e., covariates within both cohorts) are described in supplementary figure 16+17, which includes a univariate analysis of each feature's association with PFS at the primary clinical endpoint, as well as a multivariate analysis. A summary of all clinical patient and healthy donor data can be found in Supplementary tables 1 and 3 and Supplementary Figure 15.

Nature Medicine: doi:10.1038/nm.4466