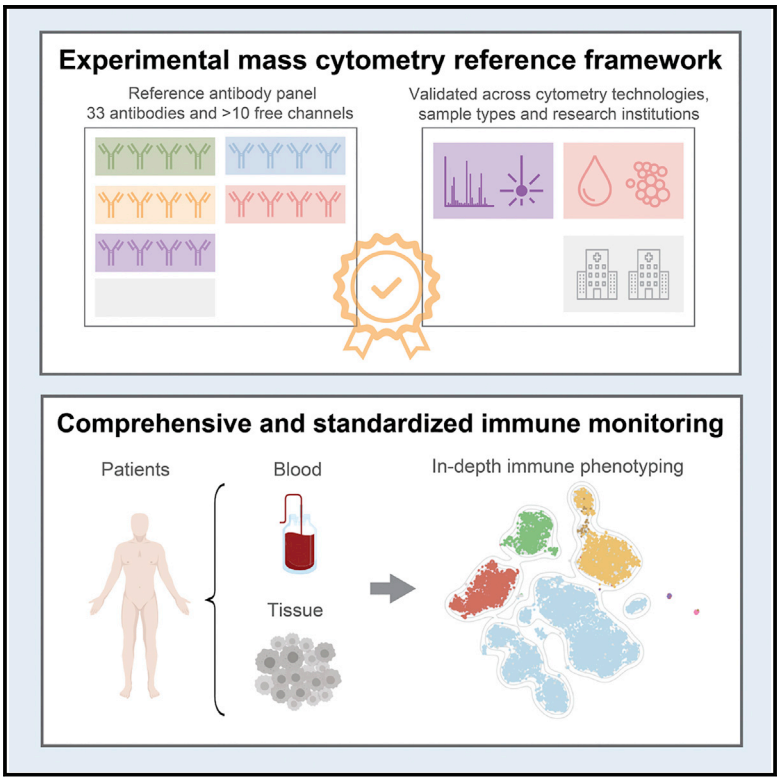
Cell Reports

Resource

Comprehensive Immune Monitoring of Clinical Trials to Advance Human Immunotherapy

Graphical Abstract



Authors

Felix J. Hartmann, Joel Babdor, Pier Federico Gherardini, ..., Everett Meyer, Matthew H. Spitzer, Sean C. Bendall

Correspondence

matthew.spitzer@ucsf.edu (M.H.S.), bendall@stanford.edu (S.C.B.)

In Brief

Hartmann et al. provide an experimental framework to identify and characterize all major human immune cell lineages in a single assay using mass cytometry (CyTOF). This validated and readily available workflow ensures comprehensive immunophenotypic analysis, improves data comparability, and allows identification of disease-associated immune signatures and biomarkers for human immunotherapy.

Highlights

- Single assay to identify and characterize all major human immune cell lineages
- Readily available and extensively validated antibody panel
- Additional (>10) targets can be added to meet specific hypotheses
- Allows identification of disease-associated immune signatures and biomarkers

Hartmann et al., 2019, Cell Reports *28*, 819–831
July 16, 2019
https://doi.org/10.1016/j.celrep.2019.06.049

CellPress

Cell Reports
Resource

Comprehensive Immune Monitoring of Clinical Trials to Advance Human Immunotherapy

Felix J. Hartmann,[1] Joel Babdor,[2] Pier Federico Gherardini,[3] El-Ad D. Amir,[4] Kyle Jones,[5] Bita Sahaf,[6] Diana M. Marquez,[2] Peter Krutzik,[7] Erika O'Donnell,[7] Natalia Sigal,[8] Holden T. Maecker,[8] Everett Meyer,[9] Matthew H. Spitzer,[2,3,10,*] and Sean C. Bendall[1,3,11,*]

[1]Department of Pathology, School of Medicine, Stanford University, Palo Alto, CA 94305, USA
[2]Departments of Otolaryngology-Head and Neck Surgery and Microbiology and Immunology, University of California, San Francisco, San Francisco, CA 94143, USA
[3]Parker Institute for Cancer Immunotherapy, San Francisco, CA 94125, USA
[4]Astrolabe Diagnostics, Inc., Fort Lee, NJ 07024, USA
[5]Department of Orofacial Sciences, University of California, San Francisco, San Francisco, CA 94143, USA
[6]Cancer Correlative Science Unit, Cancer Institute, School of Medicine, Stanford University, Palo Alto, CA 94305, USA
[7]Primity Bio, Fremont, CA 94538, USA
[8]Department of Microbiology and Immunology, Stanford University, Palo Alto, CA 94305, USA
[9]Cellular Therapy Facility, Blood and Marrow Transplantation, School of Medicine, Stanford University, Palo Alto, CA 94305, USA
[10]Chan Zuckerberg Biohub, San Francisco, CA 94158, USA
[11]Lead Contact
*Correspondence: matthew.spitzer@ucsf.edu (M.H.S.), bendall@stanford.edu (S.C.B.)
https://doi.org/10.1016/j.celrep.2019.06.049

SUMMARY

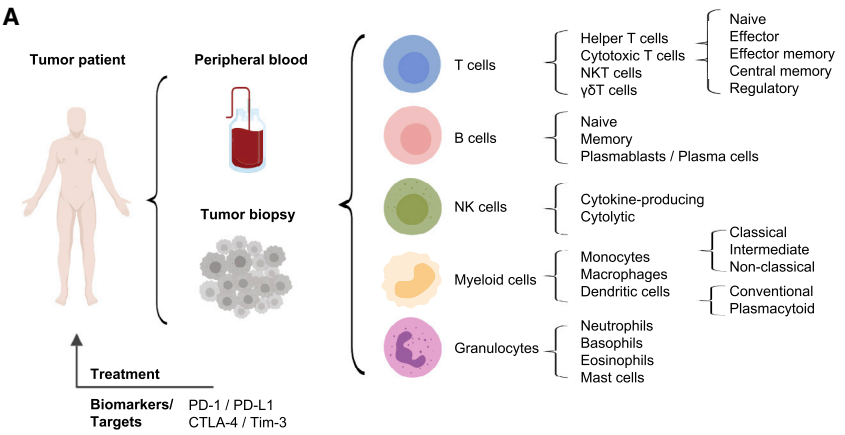
The success of immunotherapy has led to a myriad of clinical trials accompanied by efforts to gain mechanistic insight and identify predictive signatures for personalization. However, many immune monitoring technologies face investigator bias, missing unanticipated cellular responses in limited clinical material. We present here a mass cytometry (CyTOF) workflow for standardized, systems-level biomarker discovery in immunotherapy trials. To broadly enumerate immune cell identity and activity, we established and extensively assessed a reference panel of 33 antibodies to cover major cell subsets, simultaneously quantifying activation and immune checkpoint molecules in a single assay. This assay enumerates $\geq$98% of peripheral immune cells with $\geq$4 positively identifying antigens. Robustness and reproducibility are demonstrated on multiple samples types, across two research centers and by orthogonal measurements. Using automated analysis, we identify stratifying immune signatures in bone marrow transplantation-associated graft-versus-host disease. Together, this validated workflow ensures comprehensive immunophenotypic analysis and data comparability and will accelerate biomarker discovery.

INTRODUCTION

Treating cancer via modulation of the immune system has recently shown curative clinical benefit in multiple types of cancer for which conventional chemotherapy has not worked. Three of the most widely employed strategies are hematopoietic stem cell transplantation, immune checkpoint blockade (Ribas and Wolchok, 2018), and adoptive transfer of chimeric antigen receptor (CAR) T cells (June et al., 2018), although many other approaches are being developed. To further investigate the immunotherapeutic potential of all approaches and combinations thereof, thousands of clinical trials are currently being planned and conducted (Farkona et al., 2016).

Many immunotherapy trials are accompanied by immune monitoring, which can provide crucial insights into immune cell behavior at both population and single-cell levels. Comprehensive phenotyping of immune populations aids in the elucidation of the cellular mechanisms underlying newly developed therapeutic approaches. It can also identify the presence of cellular and molecular signatures that stratify patients into distinct risk groups and/or help to predict clinical responses to therapy. The tremendous complexity and heterogeneity of the human immune system necessitates the use of single-cell technologies for its analysis. Although flow cytometry has traditionally been the mainstay for such immune-monitoring applications, the advent of mass cytometry (i.e., cytometry by time-of-flight [CyTOF]) now provides an opportunity to simultaneously quantify more molecular features while reducing signal overlap and background noise (Bandura et al., 2009; Bendall et al., 2011). The high-dimensional capabilities of mass cytometry enable the identification of a wide array of immune populations and cellular states in a single assay, which not only allows comprehensive immune monitoring of small sample quantities but also across millions of cells from large groups of patients (Spitzer and Nolan, 2016).

Before using mass cytometry for immune cell phenotyping in clinical trials, rigorous validation studies must be performed to establish proper experimental and analytical workflows. Although such specialized workflows have been developed independently at some research institutions, published studies using these methods are typically not comparable because



B

T cells

CD3	Total T cells
CD4	CD4+ T cells
CD8a	CD8+ T cells
CD25	Treg, activation
CD27	Maturation
CD38	Maturation
CD45RA	Naive/memory cells
CD45RO	Naive/memory cells
CD56	NKT cells, T cell subsets
CD127	T cell subsets
CD152 (CTLA4)	Checkpoint
CD197 (CCR7)	Naive/memory cells
CD279 (PD-1)	Checkpoint
FoxP3	Treg cells
HLA-DR	Activation
T-bet	T cell subsets
gdTCR	gd T cells
TCRVa24-Ja18	iNKT cells
Tim-3	Checkpoint

B cells

CD19	Total B cells
CD25	Activation
CD27	Naive/memory
CD38	Naive/memory
HLA-DR	Antigen presentation

General

CD45, CD235ab, CD61, DNA, live/dead

NK cells

CD8a	NK cell subsets
CD16	NK cell subsets
CD38	Activation
CD56	NK cells subsets
T-bet	Maturation

Myeloid cells

CD11b	Monocytes, macrophages
CD11c	Monocytes, macrophages, DCs
CD14	Monocyte subsets, macrophages
CD16	Monocyte subsets
CD33	Total myeloid
CD123	pDCs
CD274 (PD-L1)	T cell checkpoint ligand
FceR1 a	pDCs, DCs
HLA-DR	Antigen presentation

Granulocytes

CD11b	Neutrophils
CD16	Neutrophils
CD117	Mast cells
CD123	Basophils
FceR1a	Basophils

Free channels

≥10 channels	Open for customization

Figure 1. Comprehensive Assessment of Immune Composition for Clinical Research in Cancer Immunotherapy

(A) Common sample types anticipated from tumor patients include peripheral blood samples and tumor biopsies. Within these samples, immune cell lineages and respective subpopulations are indicated. Although more subsets can be delineated, these populations were chosen as a reference set of interest for comprehensive immunophenotyping. In addition to population identification, important clinical targets and currently available biomarkers are of high interest.

(B) Antigens were selected based on their relevance for population and subpopulation identification or for defining important activation and/or maturation stages.

For additional information on clones, dilutions, and metal-assignments, see Table S1 and the Key Resources Table.

each workflow uses distinct antibody panels to identify different target immune cell populations. Furthermore, the scope of a given study is often limited to specific target populations hypothesized to be of importance instead of broadly surveying all immune cell subsets in a given sample. This approach likely biases the analysis and overlooks unanticipated, potentially novel effects on other immune cell populations. Further, the unbiased analysis of such studies requires researchers to establish dedicated analytical frameworks to more effectively mine the high-dimensional datasets generated by mass cytometry (Arvaniti and Claassen, 2017; Bruggner et al., 2014; Nowicka et al., 2017).

To address these issues, we here present a mass cytometry-based experimental workflow for comprehensive immune monitoring of cancer immunotherapy clinical trials. The proposed reference antibody panel used in this workflow is comprised of a readily available, established, and validated set of 33 surface and intracellular antibodies, enabling the robust identification of key immune cell populations and cell states in a single assay. We achieved assignment of 98% of peripheral immune cells by positivity of four or more antigens. Importantly, the design facilitates the space for additional (≥10) targets without disruption of the core reference panel to address experiment-specific hypotheses, providing an unprecedented level of flexibility and customization compared with other workflows. Exemplifying this ability, we identify additional B-cell maturation states and characterize myeloid cell heterogeneity across matched primary tumors and lymph node metastases, suggesting tissue-dependent expression of co-stimulatory molecules (CD86). Finally, we demonstrate the utility of this framework by monitoring immune cell reconstitution and identifying disease-associated immune signatures using an automated pipeline after bone marrow transplantation (BMT) in leukemia patients (n = 15). Together, this workflow provides a standardized immune monitoring approach that can greatly improve understanding of key molecular and cellular factors that can influence and predict therapeutic success and failure, providing biomarkers to improve the application of next-generation treatments.

RESULTS

Comprehensive Phenotyping for Human Immunotherapy Trials

To build a comprehensive human immunophenotyping panel for a single-pass analysis, we took a cell-lineage–agnostic approach to maximize coverage of all immune populations expected in biological specimens (i.e., peripheral blood and tissue) from immunotherapy trials. As such, we first selected the major immune cell lineages and their subsets that would be ideal to detect in human cancer samples. This list comprises T cells, B cells, natural killer (NK) cells, and various myeloid and

granulocyte populations, thus covering all major immune cell lineages typically found (Figure 1A).

The importance of T cells in cancer has been well-established and is illustrated by the clinical success of CAR–T-cell therapies (June et al., 2018) and checkpoint-blockade approaches (Ribas and Wolchok, 2018). Therefore, in addition to identifying T cells and their functionally diverse subsets, determination of the expression levels of checkpoint-related molecules, such as PD-1, CTLA-4, and TIM-3, as well as receptors, such as PD-L1, is critical. One specific T-cell subset of high interest included in the panel is regulatory T (Treg) cells. Treg cells are able to suppress T-cell responses against self-antigens as well as anti-tumor T-cell responses and are often associated with poor prognosis (Tanaka and Sakaguchi, 2017).

Besides T cells, functional heterogeneity also exists within other compartments, including NK cells. Traditionally, $CD56^{high}CD16^{-}$ are thought to be the main producers of an array of cytokines, whereas $CD56^{low}CD16^{+}$ NK cells exhibit increased cytolytic activity (Björklund et al., 2016; Cooper et al., 2001). Likewise, multiple functionally diverse myeloid populations have been identified (Villani et al., 2017; Wong et al., 2012), some of which have been correlated with therapeutic success in immunotherapy (Krieg et al., 2018).

To detect and analyze the immune cell populations listed in Figure 1A, we identified a combination of surface and intracellular proteins that characterize these immune cell lineages and their functional states (Figure 1B) and selected a panel of 33 anti-human heavy-metal–conjugated monoclonal antibodies targeting these epitopes (Table S1; Key Resources Table). Allocation of antibodies to specific heavy-metal isotopes followed mass cytometry specific panel design considerations, including allocation of low-abundance targets to higher-sensitivity channels and minimization of potential spectral overlap (Takahashi et al., 2017). Further, all antibodies were titrated to maximize separation of positive and negative populations and to minimize sources of non-specific signal in adjacent channels (Table S1; Key Resources Table).

Importantly, given the high-dimensional capabilities of mass cytometry, the proposed panel does not exhaust the full range of metal isotopes commonly used in mass cytometry experiments, which allows for the inclusion of 10 or more additional antibodies to further customize the panel toward more-specific hypotheses. This antibody panel, therefore, provides the backbone needed to comprehensively and robustly identify all major immune cell populations in patient samples from immunotherapy clinical trials while allowing further customization.

Analysis of Immune Composition and Activation State

Having defined the range of immune cell populations and proteins to be analyzed, we used this panel to stain cryopreserved peripheral blood mononuclear cells (PBMCs) from healthy donors. Stained samples were acquired on a CyTOF mass cytometer, and data were normalized using bead standards (see Method Details). Samples were pre-gated on single, DNA^{+}, live, $CD45^{+}$ non-platelet, and non-erythrocyte cells (Figure S1A). Next, we used a sequential gating approach for initial data exploration and to identify the major immune populations within these samples (Figure 2A). All major immune cell lineages could be readily identified using a series of lineage defining surface proteins and calculated frequencies were found to be within known ranges (Brodin and Davis, 2017) (Figure 2B). Importantly, using the proposed gating strategy, we were able to assign 98.4% ± 0.3% (median ± SEM) of pre-gated cells to a specific immune lineage. Remaining cells are likely unassigned because of the strict cutoffs inherent to biaxial gating and could be identified using high-dimensional approaches as shown below.

Our panel enabled the identification of multiple immune cell subpopulations. For example, T cells could be further subdivided into $CD4^{+}$ T helper (Th) cells, $CD8^{+}$ T cells, natural killer T (NKT) cells, and $\gamma\delta$ T cells (Figures 2A and 2C). Additionally, using the differential expression patterns of CD27, CD45RA, CD45RO, and CCR7, several maturation and antigen-experience states of T cells, such as naive, effector, effector memory, and central memory, could be discriminated (Sallusto et al., 2004) (Figure 2C). Treg cells were identified through high expression of the interleukin-2 receptor alpha chain (CD25), by low to negative levels of the IL-7 receptor CD127, and via expression of the lineage-defining transcription factor FoxP3. We tested multiple staining conditions to obtain optimal intracellular staining quality for FoxP3, given its importance for Treg cell identification (Figures S1B and S1C).

Aside from T cells, other immune cell lineages could be subdivided into various functional subsets (Figure 2C). Specifically, we were able to discriminate between various stages of B-cell maturation via CD27 and CD38 expression, multiple functionally distinct monocyte subsets based on their expression of CD14 and CD16, and NK cell subsets based on their combinatorial expression of CD16 and CD56.

In addition to immune cell composition, several other cellular features could be evaluated using this antibody panel (Figure 2D). CD25, HLA-DR, and CD38 allowed determination of the activation state of T cells, whereas Ki-67 expression identified actively proliferating cells across multiple cell types. Importantly, expression levels of the immune checkpoint-related molecules PD-1, PD-L1, CTLA-4, and TIM-3 could be assessed on all cells. Taken together, the highly optimized approach proposed here for immune monitoring allowed us to comprehensively assess both immune composition and cell activation states, simultaneously.

Reliability and Robustness across Different Analysis Conditions

To assess the reliability and robustness of this immunophenotypic antibody panel in obtaining comprehensive population enumeration, we calculated the number of detected antigens on each individual cell. We found that 99.8% ± 0.1% (median ± SEM) of live cells were positive for at least four or more antigens in our panel (Figures 3A, S2A, and S2B). The same was true for virtually all individual immune cell lineages, demonstrating the antibody panel's ability to further subdivide these populations (Figure 2B). Although certain antigens might be downregulated in specific diseases, in contexts with substantial cell activation, this number will likely increase as additional proteins become expressed. Importantly, expression of a board range of proteins ensures that all major immune lineages differ from each other by expression of multiple proteins (Figure S2C), indicating that cell identification does not depend on a single antigen given an appropriate gating strategy or by using clustering approaches, as discussed below.

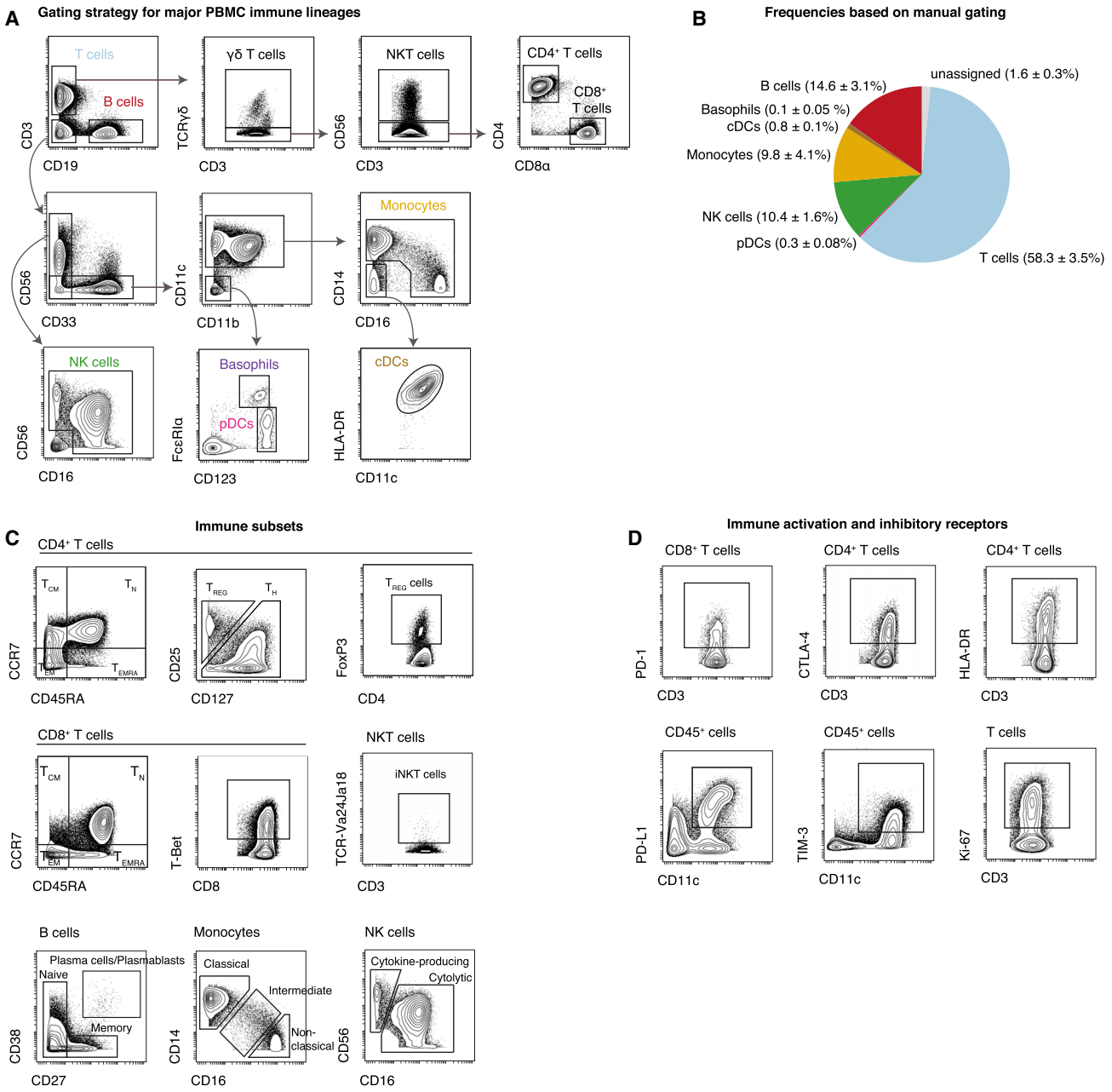


Figure 2. Data Exploration and Identification of Immune Cell Subsets in Peripheral Blood

PBMCs were stained with the indicated set of antibodies (see Table S1) and analyzed by mass cytometry.

(A) Cells were pre-gated as non-beads, DNA$^+$, single, live, CD45$^+$, CD235ab/CD61$^-$, non-neutrophils (see Figure S1). The major immune lineages and certain subsets are identified through the indicated series of gating steps.

(B) Median frequencies ± SEM in PBMCs from healthy donors (n = 5).

(C) Exemplary identification of immune cell subsets, pre-gated on the indicated populations. Treg cells can be identified as CD25^{high} CD127^{low}, FoxP3^{pos}, or a combination thereof.

(D) Assessment of expression levels of important checkpoint and activation molecules on various immune cell populations. Expression was induced by stimulating cells with anti-CD3, anti-CD28-coated beads for 2 days.

To assess the robustness of the selected panel across different research institutions, aliquots of PBMC samples obtained from the same blood draw of five healthy donors were distributed to two research centers, where sample staining was performed by the respective researchers, using separate reagents. Stained samples were then acquired on the two respective mass cytometers present in those laboratories. Immune cell frequencies were centrally determined through manual gating and compared

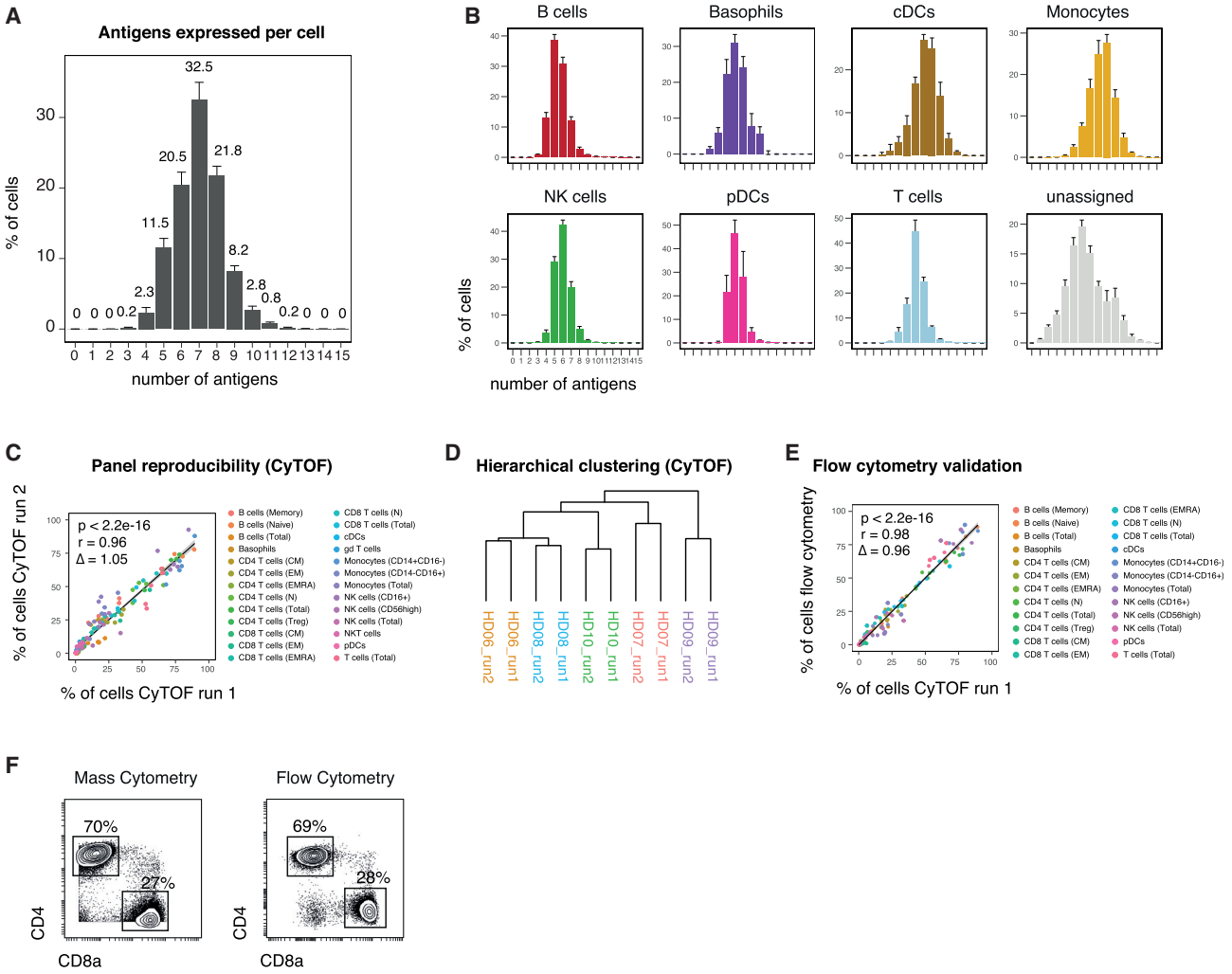


Figure 3. Reproducible Assessments of Immune Composition across Independent Analyses

PBMCs from healthy donors (n = 5) were analyzed in two research centers. Immune cell populations were identified through serial gating as before (see Figure 2).

(A) Median number of positive antigens per cell, based on manually determined cutoffs (see Figure S2A). Numbers indicate median frequency of total pre-gated cells. Error bars represent SEM.

(B) Median number of positive antigens per cell as in (A), stratified by immune cell lineage.

(C) Different PBMC aliquots of the same donors (n = 5) were stained and acquired by mass cytometry in two different research institutes. Frequencies of immune lineages were determined through serial gating. Linear regression line is shown in black with the 95% confidence intervals (CIs, shaded). Coefficients, p values, and slope Δ were calculated based on data from all donors.

(D) Hierarchical clustering of samples from two independent mass cytometry runs based on frequencies as in (C).

(E) PBMCs aliquots of the same donors as in (C) were stained and acquired by flow cytometry, employing four separate staining reactions. Frequencies of immune lineages were determined through serial gating and plotted against the frequencies determined from mass cytometry as in (C). Linear regression line is shown in black with the 95% CIs (shaded). Coefficients, p values, and slope Δ were calculated based on data from all five donors.

(F) Exemplary biaxial plots and frequencies of $CD4^+$ and $CD8^+$ T cell subsets within one donor (HD08), as determined by mass cytometry (left) and flow cytometry (right).

between the individual runs (Figure 3C). We found strong agreement (r = 0.96) between the manually gated immune cell populations from the two independent runs. This correlation was found over a broad range of frequencies and was not dependent on highly abundant populations (Figure S2D). Further, frequency-based hierarchical clustering grouped aliquots from the same donor run on different CyTOF analyzers together, thus confirming the data reproducibility between different the study centers (Figure 3D). Additional aliquots of the same PBMCs were run by flow cytometry, employing four independent antibody panels focusing on separate immune cell populations. Importantly, we obtained strong agreement (r = 0.98) between immune cell populations over a broad range of frequencies analyzed with either flow cytometry or CyTOF (Figures 3E, 3F, and S2D).

Lastly, to accommodate for a wide variety of immune sample collection techniques, we assessed the effect of sample fixation

before surface staining and analysis with mass cytometry (Figures S2E–S2G). We calculated fold changes of the 95th percentile between unfixed and paraformaldehyde (PFA)-fixed cells for each antigen. Although the majority of the antigens were not overtly altered in their dynamic ranges, a subset of antigens (including CCR7 and CD11b) showed decreased staining on previously fixed cells. However, manually gated immune cell frequencies from live-stained cells versus cells fixed with PFA before surface staining were nevertheless highly correlated ($r = 0.94$). As before, hierarchical clustering confirmed an overall highly similar immune profile between fixed and unfixed samples. Together, these data demonstrate the robustness and reproducibility of this mass cytometry-based analysis across multiple study centers and staining conditions as well as strong correlation with the historical gold standard, fluorescence-based flow cytometry.

Data Visualization and Population Identification Using Automated Approaches

Thus far, we used a defined sequential gating strategy to identify major immune cell populations, a method that is widely used by researchers and founded in empirical biological knowledge. However, with the increase in simultaneously acquired parameters, it is progressively infeasible to manually identify populations in highly multiplexed datasets, making computational approaches, such as clustering, extremely advantageous (Chester and Maecker, 2015; Mair et al., 2016; Saeys et al., 2016; Spitzer and Nolan, 2016).

To enable initial exploration, high-dimensional data are often projected into a lower-dimensional space interpretable by humans using dimensionality-reduction algorithms. These lower-dimensional maps give an immediate overview of data structure and the presence of various populations. One method that has become increasingly popular is t-Distributed Stochastic Neighbor Embedding (tSNE) (Amir et al., 2013; van der Maaten and Hinton, 2008). We visualized PBMC data from five healthy donors using tSNE and assigned cells to unique colors by overlaying the results of our manual gating (Figures 4A and S3A). Manual gating and separation by tSNE appeared in high concordance, demonstrating consistent results with one another. Another hybrid approach that allows the visualization and comparison of multidimensional datasets are scaffold maps (Spitzer et al., 2015). Scaffold maps groups similar cells into clusters, which are then visualized based on their similarity with manually determined (e.g., gated) landmark nodes. We built a reference scaffold map using healthy donor PBMCs and used manual gating to define landmark nodes, which represented all major immune cell populations identified in our mass cytometry data (Figure 4B). This method allows for comparison with other samples, such as tumor-infiltrating leukocyte populations from tissue biopsies of cancer patients, which can then be mapped onto this reference map and compared through visual inspection or statistical methods (Spitzer et al., 2017) (Figure 4C).

Besides scaffold, a multitude of other high-dimensional clustering algorithms have been reported (Weber and Robinson, 2016). Although many of these algorithms now have graphical user interfaces (e.g., cytofkit [Chen et al., 2016] and Cytosplore [van Unen et al., 2017]), comprehensive and reproducible analysis methods for large groups of samples remains a challenge that often requires basic familiarity with programming languages.

Recent automated commercial solutions have been developed to analyze large multidimensional datasets. We here employed the Astrolabe platform (Astrolabe Diagnostics, Inc.), which uses the flow self-organization map (FlowSOM) algorithm (Van Gassen et al., 2015), followed by a labeling step that automatically assigns cells to pre-selected and biologically known immune cell lineages (Figures 4D and 4E). Depending on the required resolution, these populations can be further subdivided, again using unsupervised FlowSOM-based clustering (Figure S3B). Using the Matthews correlation coefficient (MCC; see Method Details) to compare lineage assignments between manual gating and clustering, we found good correlations for all major leukocyte populations (Figures 4F, S3C, and S3D), with minor disagreements for basophils (present here at extremely low levels) and NK cell subsets (Figure S3E).

In summary, a variety of automated methodologies can be applied to the high-dimensional datasets generated using our proposed antibody panel, thus allowing the exploration, visualization, and comparison of single samples or sample groups to ultimately gain novel biological insights in a hypothesis-free and comprehensive approach.

Identifying Disease-Associated Immune Signatures Following BMT

One scenario in which comprehensive immunophenotyping, without prior knowledge of the system composition, is crucial is hematopoietic reconstitution in leukemia patients following BMT. We collected PBMC samples from 15 individuals, sampled at multiple time points after BMT, for a total of 28 samples (see Table S2). Of these patients, a small subset suffered from graft versus host disease (GvHD, $n = 3$), whereas most other patients did not experience such complications (Figure 5A). To monitor immune reconstitution and to identify potential GvHD-associated immune signatures, we applied the above outlined mass cytometry–based experimental and analytic workflow.

After staining and acquisition, we used the Astrolabe platform to identify the major immune populations and their subsets. Annotated clustering identified 30 immune cell subsets spanning the major immune lineages (Figure 5B). tSNE dimensionality reduction was used to give an immediate overview of various reconstituted populations (Figure 5C). Across all samples, immune reconstitution was dominated by T cells (29.0% ± 7.5%) and monocytes (27.0% ± 5.4%), followed by NK cells (7.1% ± 0.9%) and human leukocyte antigen–DR isotype $(HLA\text{-}DR)^{-}CD16^{+}$ cells (Figure 5E). B cells (dominated by $CD27^{-}$ B cells) were present at lower frequency (2.5% ± 1.3%).

Exploring the biological significance of patient-to-patient variation in their immune composition, we investigated whether immune cell proportions stratify among patients with different clinical outcomes, e.g., the occurrence of GvHD. To compare between patients with or without GvHD, we calculated the fold changes (FCs), p values, and false discovery rates (FDRs), correcting for multiple-hypothesis testing (see Method Details). This approach identified a reduction in two immune cell populations as a potential immune-signature of failed engraftment/occurrence of GvHD in this cohort (Figure 5F). First, $CD27^{-}$

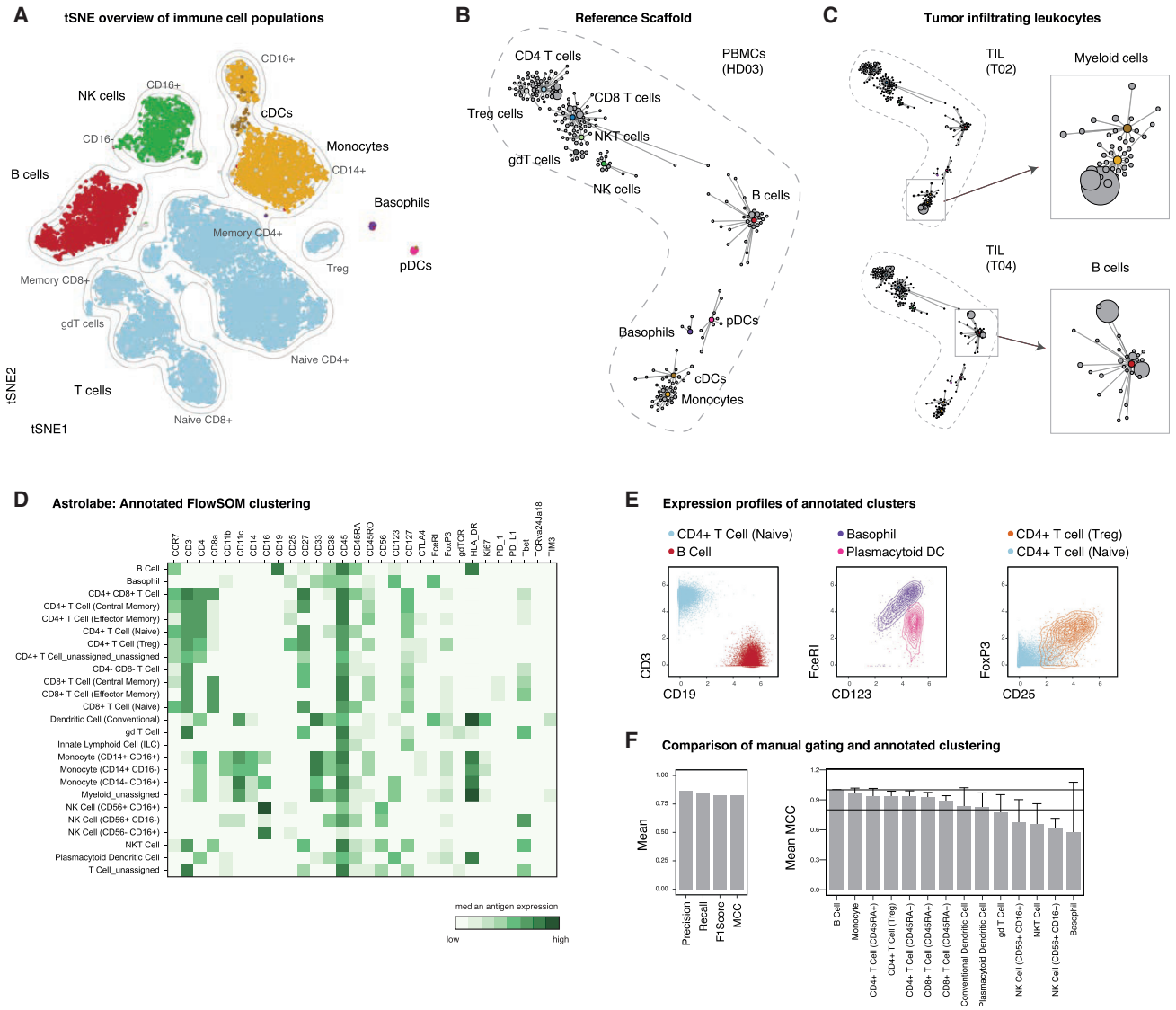


Figure 4. Automated Data Visualization and Population Identification

PBMCs from healthy subjects (n = 5) and tumor biopsies from cancer patients (n = 5) were analyzed by mass cytometry using the reference panel (see Table S1).

(A) Data from all healthy donors was randomly subsampled to 20,000 cells and subjected to tSNE dimensionality reduction. Cells are colored by their immune cell lineage assignment from manual gating. Grey indicates cells unassigned by manual gating.

(B) A reference scaffold map of PBMC data was created using manually gated landmarks (colored) and all antigens for the clustering analysis. Inter-cluster connections were used to create the graph but are not depicted here. Shown is one representative sample (HD03).

(C) Pre-gated, $CD45^+$ cells from tumor samples were mapped onto the reference scaffold. Maps from two patients are shown (left). Enlarged examples of modulated immune cell populations are pointed out (right).

(D) PBMC data as above were clustered and automatically annotated using the Astrolabe platform. Shown are median expression levels of all antigens for all clusters.

(E) Exemplary expression profiles of immune cell populations as determined by Astrolabe (HD06).

(F) Mean precision, recall, F1score, and Matthews correlation coefficient (MCC; see Method Details) between manual lineage assignments and FlowSOM-based clustering for all donors and populations (left). Mean MCC for all donors stratified by population (right). Two horizontal lines indicate MCC = 1 (maximum agreement) and MCC = 0.8, respectively. Error bars represent SEM.

B cells were reduced in patients with GvHD (0.44% ± 0.21% versus 3.33% ± 2.2%, p = 0.0046, FDR = 0.069; Figure 5G). In addition, patients with GvHD displayed lower frequencies of naive $CD4^+$ T cells (0.09% ± 0.03% versus 0.3% ± 0.3%, p = 0.0036, FDR = 0.069). Lastly, although the comprehensive assessment of a broad range of immune cell populations was necessary to identify these stratifying populations, once their identity is known, manual gating can again be used to confirm their reduction in patients with GvHD (Figure 5H). In summary, this demonstrates the utility of the outlined framework to perform

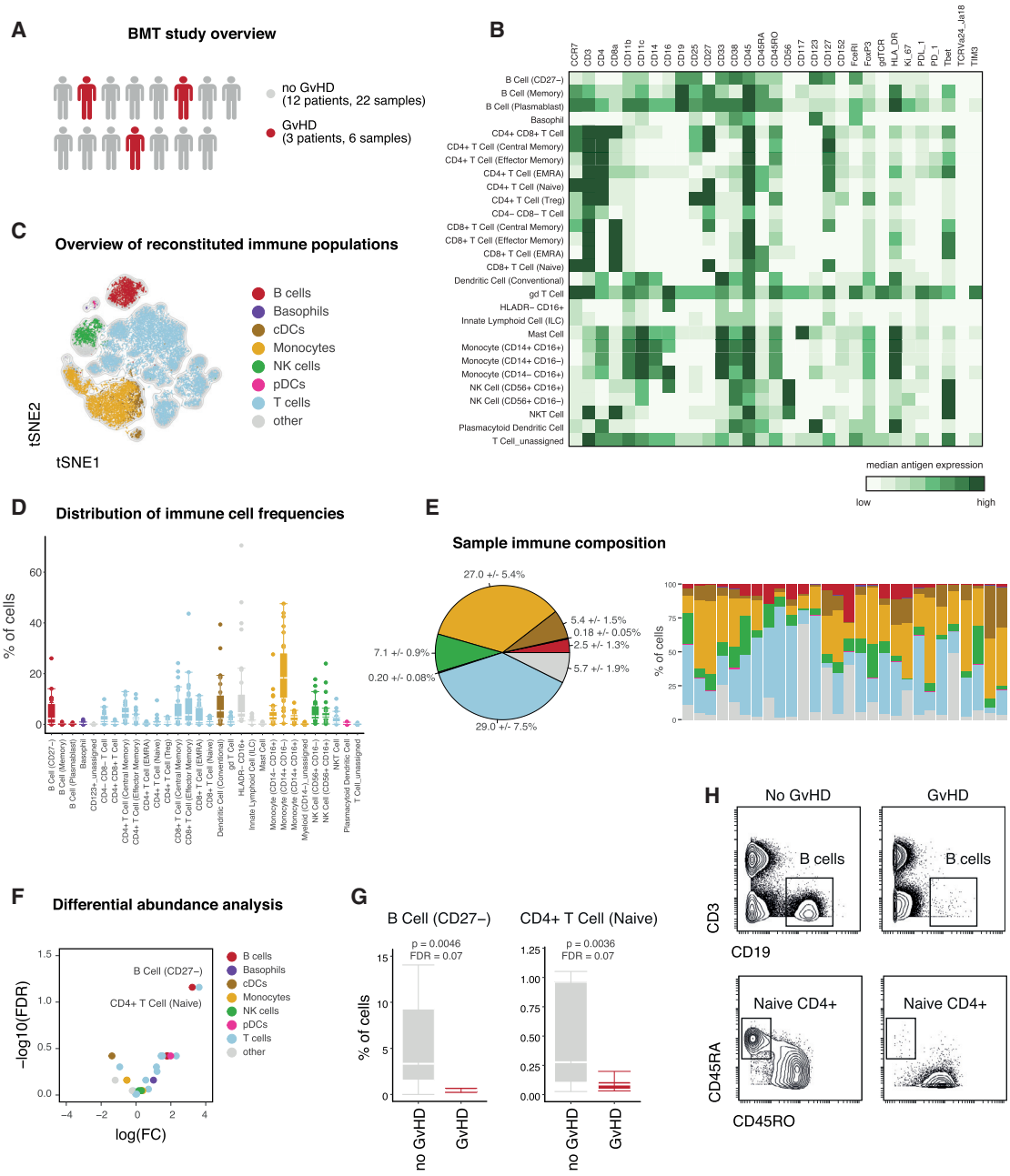


Figure 5. Identification of Disease-Associated Immune Signatures Following Bone Marrow Transplantation

(A) Following tumor therapy, patients (n = 15; Table S2) underwent bone marrow transplantation. Peripheral blood samples were collected and subsequently stained with the described reference panel and analyzed by mass cytometry.

(B) Data were uploaded to the Astrolabe platform, clustered, and automatically annotated. Exemplary heatmap of one patient depicting the median protein expression levels across all populations identified through clustering.

(C) The 20,000 randomly subsampled cells of one patient were subjected to tSNE dimensionality reduction. Color-assignments represent different immune lineages as identified through annotated clustering.

(D) Clustering-derived frequencies of immune populations for all samples in this study (n = 28). Boxplots depict the interquartile range (IQR) with a horizontal line representing the median. Whiskers extend to the farthest data point within a maximum of 1.5× IQR. Points represent individual samples.

(E) Frequencies of immune cell subpopulations were combined into frequencies for major immune cell lineages and color-coded as in (C). Pie chart depicting the median frequencies ± SEM of all major immune lineages across all samples (left). Immune composition for all analyzed samples (n = 28; right). (F) FDR and fold change (FC) of immune cell frequencies in patients with or without GvHD.

(legend continued on next page)

clinically relevant monitoring of immune perturbations in a medical setting. Employing this approach, treatment-, disease-, or time-dependent, immunological responses can be assessed in a straightforward and comprehensive manner to discover novel biomarkers and immune signatures.

Extendibility and Flexibility of the Reference Assay Framework

Given the proposed application of this workflow to a diverse array of studies, an important feature of the immunophenotypic antibody panel is that it does not exhaust the full range of available lanthanide isotopes available for use by mass cytometry. Up to 10 antibodies or more, depending on the availability of newly developed reagents, can be added to the described reference panel without modification. We illustrated this ability to customize the panel in two separate scenarios, focusing on different leukocyte populations (see Table S3). We targeted up to 10 additional antigens with mass-tagged antibodies, stained, and acquired samples from multiple donors with these antibodies in addition to the reference panel.

First, we included an additional 10 antibodies to further distinguish B-cell maturation as well as co-stimulatory molecule and isotype expression (Kaminski et al., 2012) (Figures S4A–S4C). Together with the immunophenotypic reference panel, these additional antigens enabled the identification of multiple additional B-cell subpopulations, including plasma cells and several stages of isotype-switched naive and memory B cells.

Further, we set focus on tissue-resident myeloid cell subpopulations by including antibodies against molecules associated with dendritic cells (DCs), neutrophils, monocytes, and macrophages and their activation or co-stimulatory states (Figure 6A). For that analysis, we included tumor biopsies (n = 4) and matched, metastatic lymph node samples (n = 2) from patients with squamous cell carcinomas (see Table S2).

Using a combination of tSNE visualization and FlowSOM-clustering, multiple myeloid subpopulations could be distinguished (Figures 6B and 6C). This included previously unresolved subsets of cDCs ($CD141^+$ cDC1 and $CD1c^+$ cDC2) as well as different populations of monocyte/macrophage cells, hereafter referred to as monocyte-derived cells (MDCs). In addition to their identification, these subsets could also be analyzed for differential expression of many subset associated proteins (CCR2, CD244, CD172ab, and CD206) and costimulatory molecules (CD40 and CD86), which have been shown reflect the activation state and propensity to provide co-stimulation to T cells (Figure 6D).

Although preliminary, because of the limited number of samples, this panel extension enabled us to compare between cells isolated from primary tumors and lymph node metastases. Frequencies of defined subpopulations were comparable between tumors and lymph nodes (Figures 6E, S4D, and S4E). We further compared the expression of costimulatory molecules on myeloid populations from lymph nodes and tumors and observed a trend toward increased CD86 expression on all three MDC subsets isolated from lymph node metastases compared with the respective primary tumors (Figures 6F and 6G). In summary, these results demonstrate that, while retaining the ability to cover all immune populations across a variety of tissues and collection conditions (Figures 1, 2, 3, 4, and 5), the proposed immune reference workflow provides the flexibility to further increase the resolution of the analysis toward a specific immune population or scientific hypothesis.

DISCUSSION

In this study, we established a reference panel of 33 anti-human antibodies for mass cytometry that can easily be incorporated into routine immunophenotyping studies in the context of cancer immunotherapy. The selected target antigens are distributed broadly across immune cell types and thus ensure that all major immune cell lineages and various functional subsets can be identified robustly and unambiguously. Apart from proteins essential for the identification of immune cell populations, we also included antibodies against targets that can be used to assess functional states, e.g., proliferative activity or expression levels of immune checkpoint-related molecules, such as CTLA-4, Tim-3, PD-1, and PD-L1, some of which have already been proposed as candidate biomarkers in cancer immunotherapy (Patel and Kurzrock, 2015).

We validated the panel using various sample types from healthy donors or cancer patients, including PMBCs and biopsies of tumor tissue or lymph nodes. In all cases, we were able to identify the major immune cell lineages as well as their functionally diverse subsets and cell states. Additionally, these samples were collected and analyzed by different researchers across various research institutions, underwent different preprocessing protocols, and were stained and acquired at multiple locations (Leipold et al., 2018). Notwithstanding, we obtained highly correlated results from each of the patient samples analyzed, regardless of pre-staining processing or location where the samples were analyzed. Additionally, immune-cell frequencies derived from flow cytometry methods strongly correlated with our mass cytometry results (Bendall et al., 2011), further validating the proposed workflow.

It should be noted that the proposed reference panel focuses on major immune cell populations and well-established subpopulations. However, to date, there is no comprehensive consensus regarding cell-type definitions and annotations, and many immune cell populations can be further subdivided, depending on the use of additional antigens. To account for this, and to allow customization of the antibody panel for specific research needs, the proposed immunophenotypic reference antibody panel does not exhaust the full range of available analysis channels, and additional antibodies can easily be added. Importantly, the absence of spectral overlap between different analysis channels in mass cytometry allows the straightforward

(G) Comparison of differentially abundant immune cell frequencies in patients with or without GvHD. Boxplots depict the IQR with a horizontal line representing the median.

(H) Confirmation of reduced abundance of B cells (top) and naive $CD4^+$ T cells (bottom) in an exemplary patient with (right) and without (left) GvHD. Examples of B cells were pre-gated on single, live, $CD45^+$ cells. Examples of naive $CD4^+$ T cells were pre-gated on single, live, $CD4^+$ T cells.

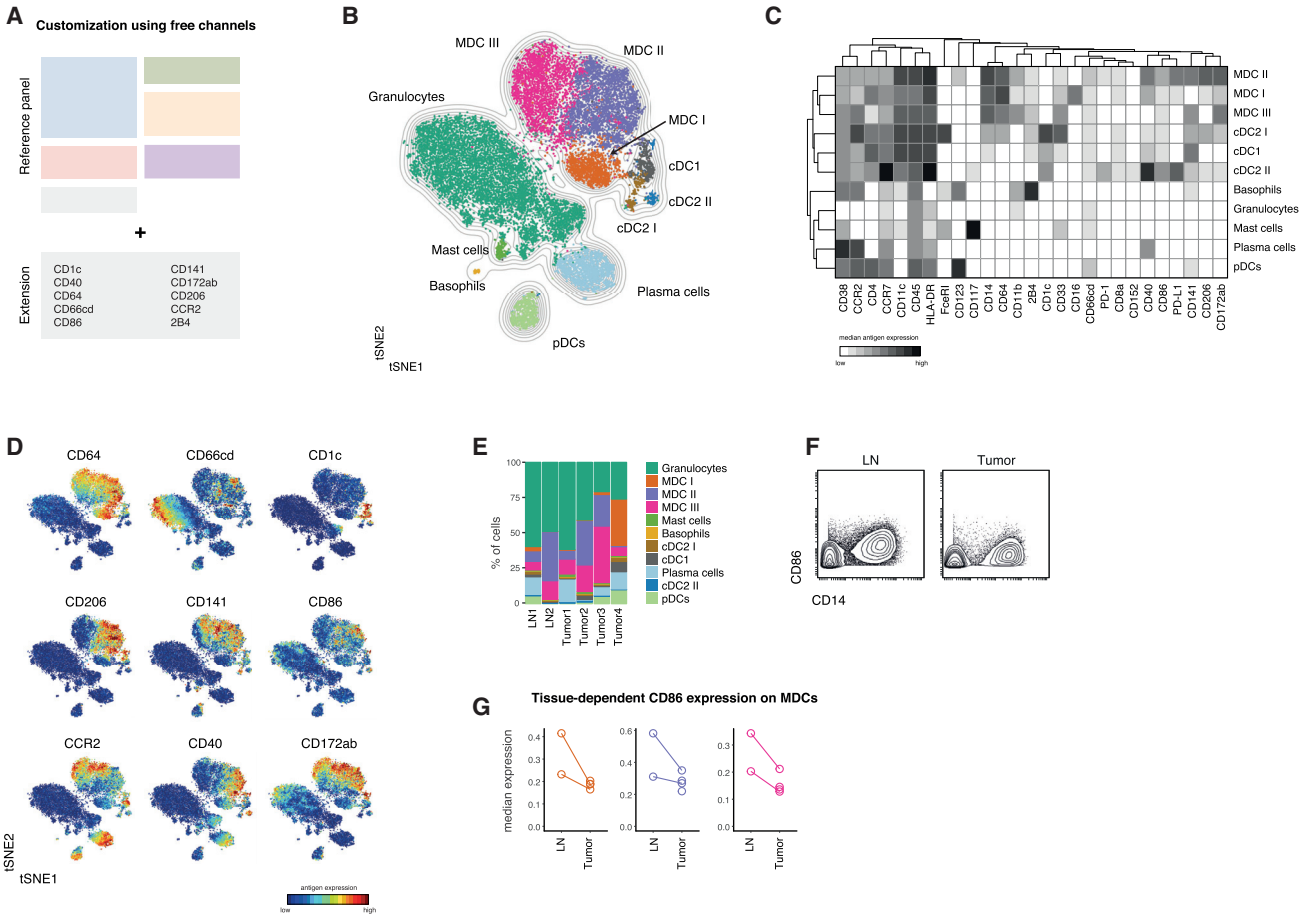


Figure 6. Flexibility of the Proposed Framework Enables Augmented Exploration of Heterogeneous Populations

(A) Antibodies targeting additional antigens of interest were conjugated to non-occupied heavy metal isotopes (see Table S3). Cells from lymph node biopsies (n = 2) and tumor biopsies (n = 4) of patients with head and neck carcinoma (see Table S2) were stained with these antibodies in combination with the reference set.

(B) Data were pre-gated on single, live, $CD45^{+}CD3^{-}CD19^{-}CD7^{-}CD56^{-}$ to exclude T cells, most B cells, and NK cells. To create a tSNE overview, data from all samples were randomly subsampled to 20,000 cells with equal contribution from all samples. Cells are colored by their FlowSOM-based cluster-assignment. Grey lines indicate the density distribution of the tSNE map.

(C) Cluster-based median expression levels for all population-relevant antigens used in the tSNE and FlowSOM analysis.

(D) Protein expression levels of all additional antigens are overlaid as a color-dimension onto the tSNE map.

(E) Frequencies of FlowSOM-based clusters as in (B) and (C) in all samples.

(F) Exemplary CD86 expression levels on total MDCs ($CD14^{+}$ cells) in cells derived from a lymph node metastasis (left) and primary tumor (right) of the same patient.

(G) Median CD86 expression levels (arcsinh-transformed and percentile normalized) on MDC subsets from lymph nodes and tumors. Lines connect different tissues of the same patients.

addition of further antibodies. We illustrated this flexibility by including additional antibodies in the panel that were specific for B cells or myeloid cell subsets and activation states; however, other cell populations or combinations thereof could be targeted. Furthermore, fixation and permeabilization procedures are already integrated in this framework, thus allowing the incorporation of additional intracellular antibodies without having to modify the employed staining protocol. Currently, up to 10 channels can be customized, and making use of alternate antibody-heavy metal-conjugation protocols, such as direct binding of cisplatin to partially reduced antibodies (Mei et al., 2016), this number can be further increased. Open channels also ensure compatibility of the panel with fixed or live-cell barcoding approaches (Hartmann et al., 2018; Mei et al., 2015; Zunder et al., 2015). These approaches help to eliminate technical variability and increase sample comparability, which is especially valuable when using clinical samples obtained from different research studies. Our framework, therefore, further contributes to the standardization and quality control of mass cytometry experimentation, which builds upon the already published reports of using bead-based normalization protocols (Finck et al., 2013) and the addition of reference cells to increase comparability between different experiments (Kleinsteuber et al., 2016).

With minor exceptions, this panel is exclusively comprised of commercially available, off-the-shelf reagents, minimizing conjugation batch differences across time and different study sites. Implementation of such standardized experimental workflows

and antibody panels has been proposed for flow cytometry (Finak et al., 2016; Maecker et al., 2012) but has not yet been implemented for analogous studies using mass cytometry. In addition to assessing a broad and defined set of immune cell populations, our proposed workflow will allow valuable cross-trial comparisons and would simplify and enhance meta-analyses as recently proposed (Hu et al., 2018).

Further, we have demonstrated that results generated using this workflow are amenable to a variety of data analysis approaches. Major immune cell subsets and established subpopulations can be identified using a series of two-dimensional gates using the proposed manual-gating scheme (Figure 2). However, many alternative approaches exist and, especially for the comprehensive exploration of high-dimensional datasets, automated data analysis methods are advantageous (Newell and Cheng, 2016; Saeys et al., 2016). We used multiple, semi-automated algorithmic analyses approaches, including tSNE dimensionality reduction (Amir et al., 2013; van der Maaten and Hinton, 2008) as well as clustering and visualization through Scaffold maps (Spitzer et al., 2015). Alternatively, data could be visualized using force-directed layouts (Samusik et al., 2016) or uniform manifold approximation and projection (UMAP) dimension reduction (Becht et al., 2018; McInnes and Healy, 2018). Other approaches dedicated to identifying differential immune cell frequencies in groups of samples can be applied, including approaches relying on statistical comparisons of cluster frequencies (Bruggner et al., 2014; Spitzer et al., 2017), convolutional neutral networks (Arvaniti and Claassen, 2017), empirical Bayes moderated tests (Weber et al., 2019), or hyperspheres (Lun et al., 2017). Because these approaches typically require advanced computational skills, we additionally demonstrate compatibility of this experimental workflow with a fully-automated, commercial analysis platform (http://astrolabediagnostics.com) to perform a systems-level analysis of immune cell reconstitution after BMT (Lakshmikanth et al., 2017; Stern et al., 2018) and to identify factors associated with the development of acute GvHD (Stikvoort et al., 2017). Albeit preliminary, because of the limited sample number and potential confounding factors, we demonstrated the utility of our framework to identify such disease-associated cellular immune signatures in a clinical cohort. Importantly, we are currently employing the described methodology to investigate the longitudinal influence of modified grafts in this scenario. In addition, this framework is already being applied to multiple studies in the field of immunotherapy research, including the study of DC vaccination approaches in combination with checkpoint inhibition (Nowicki et al., 2018).

In summary, we have established and extensively validated an experimental framework for comprehensive immunophenotyping. Although the initial scope of this panel was its application to clinical research in the field of cancer immunotherapy, its broad assessment of immune cell states and populations would be a valuable approach for research in other fields, such as infectious disease (Bengsch et al., 2018; Newell et al., 2012), vaccine development (Pejoski et al., 2016), and assessment of autoimmunity (Hartmann et al., 2016; Rao et al., 2017). This study demonstrates this platform's broad applicability and provides examples of how it will accelerate and improve immune monitoring of patients enrolled in clinical trials. Altogether, by taking our cell-agnostic approach to immune monitoring, laying out a unified protocol and panel for comprehensive analysis, this study democratizes the elucidation of therapeutic mechanisms and discovery of immune cell signatures and biomarkers.

STAR★METHODS

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

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human subjects
- METHOD DETAILS
  - Panel design and heavy-metal conjugation of antibodies
  - Mass cytometry workflow
  - Flow cytometry
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Data normalization and gating
  - Data visualization and analysis
  - Statistical analysis
  - Visualization
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.06.049.

ACKNOWLEDGMENTS

We thank D.R. Glass, D. Mrdjen, and J.P. Oliviera for discussions and comments. F.J.H. was supported by the EMBO Long-Term Fellowship 1141-2017, the Novartis Foundation for Medical-Biological Research, and the Swiss National Science Foundation (SNF Early Postdoc Mobility P2ZHP3_171741). K.B.J. is supported by KL2-TR001870. M.H.S. was supported by the Chan Zuckerberg Biohub and the NIH (1S10OD018040-01 and DP5OD023056). S.C.B. was supported by the Damon Runyon Cancer Research Foundation DRG-2017-09; the NIH 1DP2OD022550-01, 1R01AG056287-01, 1R01AG057915-01, 1-R00-GM104148-01, 1U24CA224309-01, 5U19AI116484-02, and U19 AI104209; The Bill and Melinda Gates Foundation OPP1113682; and a Translational Research Award from the Stanford Cancer Institute. The Parker Institute for Cancer Immunotherapy provided core funding for this study to M.H.S. and S.C.B.

AUTHOR CONTRIBUTIONS

Conceptualization, F.J.H., M.H.S., and S.C.B.; Methodology, F.J.H. and B.S.; Validation, F.J.H., J.B., N.S., P.F.G., B.S., P.K., and E.O.; Formal Analysis, F.J.H. and E.-A.D.A.; Investigation, F.J.H. and J.B.; Resources, D.M., K.B.J., H.T.M., M.H.S., and S.C.B.; Writing – Original Draft, F.J.H.; Writing – Review & Editing, F.J.H., J.B., P.F.G., K.B.J., M.H.S., and S.C.B.; Visualization, F.J.H. and E.-A.D.A.; Supervision, M.H.S. and S.C.B.; Project Administration, M.H.S. and S.C.B.; Funding Acquisition, F.J.H., M.H.S., and S.C.B.

DECLARATION OF INTERESTS

E.-A.D.A. is a co-founder of Astrolabe Diagnostics, Inc. The other authors declare no competing interests.

Received: December 18, 2018
Revised: May 6, 2019
Accepted: June 12, 2019
Published: July 16, 2019

REFERENCES

Amir, A.D., Davis, K.L., Tadmor, M.D., Simonds, E.F., Levine, J.H., Bendall, S.C., Shenfeld, D.K., Krishnaswamy, S., Nolan, G.P., and Pe'er, D. (2013). viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. Nat. Biotechnol. *31*, 545–552.

Arvaniti, E., and Claassen, M. (2017). Sensitive detection of rare disease-associated cell subsets via representation learning. Nat. Commun. *8*, 14825.

Bandura, D.R., Baranov, V.I., Ornatsky, O.I., Antonov, A., Kinach, R., Lou, X., Pavlov, S., Vorobiev, S., Dick, J.E., and Tanner, S.D. (2009). Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. Anal. Chem. *81*, 6813–6822.

Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol. *37*, 38–44.

Bendall, S.C., Simonds, E.F., Qiu, P., Amir, A.D., Krutzik, P.O., Finck, R., Bruggner, R.V., Melamed, R., Trejo, A., Ornatsky, O.I., et al. (2011). Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science *332*, 687–696.

Bengsch, B., Ohtani, T., Khan, O., Setty, M., Manne, S., O'Brien, S., Gherardini, P.F., Herati, R.S., Huang, A.C., Chang, K.-M., et al. (2018). Epigenomic-Guided Mass Cytometry Profiling Reveals Disease-Specific Features of Exhausted CD8 T Cells. Immunity *48*, 1029–1045.e5.

Björklund, Å.K., Forkel, M., Picelli, S., Konya, V., Theorell, J., Friberg, D., Sandberg, R., and Mjösberg, J. (2016). The heterogeneity of human CD127(+) innate lymphoid cells revealed by single-cell RNA sequencing. Nat. Immunol. *17*, 451–460.

Boughorbel, S., Jarray, F., and El-Anbari, M. (2017). Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. PLoS ONE *12*, e0177678.

Brodin, P., and Davis, M.M. (2017). Human immune system variation. Nat. Rev. Immunol. *17*, 21–29.

Bruggner, R.V., Bodenmiller, B., Dill, D.L., Tibshirani, R.J., and Nolan, G.P. (2014). Automated identification of stratifying signatures in cellular subpopulations. Proc. Natl. Acad. Sci. USA *111*, E2770–E2777.

Chen, H., Lau, M.C., Wong, M.T., Newell, E.W., Poidinger, M., and Chen, J. (2016). Cytofkit: A Bioconductor Package for an Integrated Mass Cytometry Data Analysis Pipeline. PLoS Comput. Biol. *12*, e1005112.

Chester, C., and Maecker, H.T. (2015). Algorithmic tools for mining high-dimensional cytometry data. J. Immunol. *195*, 773–779.

Cooper, M.A., Fehniger, T.A., and Caligiuri, M.A. (2001). The biology of human natural killer-cell subsets. Trends Immunol. *22*, 633–640.

Ellis, B., Haaland, P., Hahne, F., Meur, N.Le, and Gopalakrishnan, N. (2009). FlowCore: basic structures for flow cytometry data (BioConductor). https://bioconductor.org/packages/release/bioc/vignettes/flowCore/inst/doc/HowTo-flowCore.pdf.

Farkona, S., Diamandis, E.P., and Blasutig, I.M. (2016). Cancer immunotherapy: the beginning of the end of cancer? BMC Med. *14*, 73.

Finak, G., Langweiler, M., Jaimes, M., Malek, M., Taghiyar, J., Korin, Y., Raddassi, K., Devine, L., Obermoser, G., Pekalski, M.L., et al. (2016). Standardizing flow cytometry immunophenotyping analysis from the Human ImmunoPhenotyping Consortium. Sci. Rep. *6*, 20686.

Finck, R., Simonds, E.F., Jager, A., Krishnaswamy, S., Sachs, K., Fantl, W., Pe'er, D., Nolan, G.P., and Bendall, S.C. (2013). Normalization of mass cytometry data with bead standards. Cytometry A *83*, 483–494.

Hartmann, F.J., Bernard-Valnet, R., Quériault, C., Mrdjen, D., Weber, L.M., Galli, E., Krieg, C., Robinson, M.D., Nguyen, X.-H., Dauvilliers, Y., et al. (2016). High-dimensional single-cell analysis reveals the immune signature of narcolepsy. J. Exp. Med. *213*, 2621–2633.

Hartmann, F.J., Simonds, E.F., and Bendall, S.C. (2018). A Universal Live Cell Barcoding-Platform for Multiplexed Human Single Cell Analysis. Sci. Rep. *8*, 10770.

Hartmann, F.J., Simonds, E.F., Vivanco, N., Bruce, T., Borges, L., Nolan, G.P., Spitzer, M.H., and Bendall, S.C. (2019). Scalable conjugation and characterization of immunoglobulins with stable mass isotope reporters for single-cell mass cytometry analysis. Methods Mol. Biol. *1989*, 55–81.

Hu, Z., Jujjavarapu, C., Hughey, J.J., Andorf, S., Lee, H.-C., Gherardini, P.F., Spitzer, M.H., Thomas, C.G., Campbell, J., Dunn, P., et al. (2018). MetaCyto: a tool for automated meta-analysis of mass and flow cytometry Data. Cell Rep. *24*, 1377–1388.

June, C.H., O'Connor, R.S., Kawalekar, O.U., Ghassemi, S., and Milone, M.C. (2018). CAR T cell immunotherapy for human cancer. Science *359*, 1361–1365.

Kaminski, D.A., Wei, C., Qian, Y., Rosenberg, A.F., and Sanz, I. (2012). Advances in human B cell phenotypic profiling. Front. Immunol. *3*, 302.

Kleinsteuber, K., Corleis, B., Rashidi, N., Nchinda, N., Lisanti, A., Cho, J.L., Medoff, B.D., Kwon, D., and Walker, B.D. (2016). Standardization and quality control for high-dimensional mass cytometry studies of human samples. Cytometry A *89*, 903–913.

Kotecha, N., Krutzik, P.O., and Irish, J.M. (2010). Web-based analysis and publication of flow cytometry experiments. Curr. Protoc. Cytom. *Chapter 10*, Unit10.17.

Krieg, C., Nowicka, M., Guglietta, S., Schindler, S., Hartmann, F.J., Weber, L.M., Dummer, R., Robinson, M.D., Levesque, M.P., and Becher, B. (2018). High-dimensional single-cell analysis predicts response to anti-PD-1 immunotherapy. Nat. Med. *24*, 144–153.

Krijthe, J.H. (2015). RTSNE: T-distributed stochastic neighbor embedding using Barnes-Hut implementation (R Foundation for Statistical Computing). https://github.com/jkrijthe/Rtsne.

Lakshmikanth, T., Olin, A., Chen, Y., Mikes, J., Fredlund, E., Remberger, M., Omazic, B., and Brodin, P. (2017). Mass Cytometry and topological data analysis reveal immune parameters associated with complications after allogeneic stem cell transplantation. Cell Rep. *20*, 2238–2250.

Leipold, M.D., Obermoser, G., Fenwick, C., Kleinstuber, K., Rashidi, N., McNevin, J.P., Nau, A.N., Wagar, L.E., Rozot, V., Davis, M.M., et al. (2018). Comparison of CyTOF assays across sites: results of a six-center pilot study. J. Immunol. Methods *453*, 37–43.

Lun, A.T.L., Richard, A.C., and Marioni, J.C. (2017). Testing for differential abundance in mass cytometry data. Nat. Methods *14*, 707–709.

Maecker, H.T., McCoy, J.P., and Nussenblatt, R. (2012). Standardizing immunophenotyping for the Human Immunology Project. Nat. Rev. Immunol. *12*, 191–200.

Mair, F., Hartmann, F.J., Mrdjen, D., Tosevski, V., Krieg, C., and Becher, B. (2016). The end of gating? An introduction to automated analysis of high dimensional cytometry data. Eur. J. Immunol. *46*, 34–43.

Matthews, B.W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochim. Biophys. Acta *405*, 442–451.

McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. *40*, 4288–4297.

McInnes, L., and Healy, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction (ArXiv). https://arxiv.org/abs/1802.03426.

Mei, H.E., Leipold, M.D., Schulz, A.R., Chester, C., and Maecker, H.T. (2015). Barcoding of live human peripheral blood mononuclear cells for multiplexed mass cytometry. J. Immunol. *194*, 2022–2031.

Mei, H.E., Leipold, M.D., and Maecker, H.T. (2016). Platinum-conjugated antibodies for application in mass cytometry. Cytometry A *89*, 292–300.

Newell, E.W., and Cheng, Y. (2016). Mass cytometry: blessed with the curse of dimensionality. Nat. Immunol. *17*, 890–895.

Newell, E.W., Sigal, N., Bendall, S.C., Nolan, G.P., and Davis, M.M. (2012). Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8+ T cell phenotypes. Immunity *36*, 142–152.

Nowicka, M., Krieg, C., Weber, L.M., Hartmann, F.J., Guglietta, S., Becher, B., Levesque, M.P., and Robinson, M.D. (2017). CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets. F1000Res. *6*, 748.

Nowicki, T.S., Berent-Maoz, B., Cheung-Lau, G., Huang, R.R., Wang, X., Tsoi, J., Kaplan-Lefko, P., Cabrera, P., Tran, J., Pang, J., et al. (2018). A pilot trial of the combination of transgenic NY-ESO-1-reactive adoptive cellular therapy with dendritic cell vaccination with or without ipilimumab. Clin. Cancer Res. *25*, 2096–2108.

Patel, S.P., and Kurzrock, R. (2015). PD-L1 expression as a predictive biomarker in cancer immunotherapy. Mol. Cancer Ther. *14*, 847–856.

Pejoski, D., Tchitchek, N., Rodriguez Pozo, A., Elhmouzi-Younes, J., Yousfi-Bogniaho, R., Rogez-Kreuz, C., Clayette, P., Dereuddre-Bosquet, N., Lévy, Y., Cosma, A., et al. (2016). Identification of vaccine-altered circulating B cell phenotypes using mass cytometry and a two-step clustering analysis. J. Immunol. *196*, 4814–4831.

R Development Core Team (2008). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing). https://www.gbif.org/en/tool/81287/r-a-language-and-environment-for-statistical-computing.

Rao, D.A., Gurish, M.F., Marshall, J.L., Slowikowski, K., Fonseka, C.Y., Liu, Y., Donlin, L.T., Henderson, L.A., Wei, K., Mizoguchi, F., et al. (2017). Pathologically expanded peripheral T helper cell subset drives B cells in rheumatoid arthritis. Nature *542*, 110–114.

Ribas, A., and Wolchok, J.D. (2018). Cancer immunotherapy using checkpoint blockade. Science *359*, 1350–1355.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Saeys, Y., Van Gassen, S., and Lambrecht, B.N. (2016). Computational flow cytometry: helping to make sense of high-dimensional immunology data. Nat. Rev. Immunol. *16*, 449–462.

Sallusto, F., Geginat, J., and Lanzavecchia, A. (2004). Central memory and effector memory T cell subsets: function, generation, and maintenance. Annu. Rev. Immunol. *22*, 745–763.

Samusik, N., Good, Z., Spitzer, M.H., Davis, K.L., and Nolan, G.P. (2016). Automated mapping of phenotype space with single-cell data. Nat. Methods *13*, 493–496.

Spitzer, M.H., and Nolan, G.P. (2016). Mass cytometry: single cells, many features. Cell *165*, 780–791.

Spitzer, M.H., Gherardini, P.F., Fragiadakis, G.K., Bhattacharya, N., Yuan, R.T., Hotson, A.N., Finck, R., Carmi, Y., Zunder, E.R., Fantl, W.J., et al. (2015). Immunology. an interactive reference framework for modeling a dynamic immune system. Science *349*, 1259425.

Spitzer, M.H., Carmi, Y., Reticker-Flynn, N.E., Kwek, S.S., Madhireddy, D., Martins, M.M., Gherardini, P.F., Prestwood, T.R., Chabon, J., Bendall, S.C., et al. (2017). Systemic immunity is required for effective cancer immunotherapy. Cell *168*, 487–502.e15.

Stern, L., McGuire, H., Avdic, S., Rizzetto, S., Fazekas de St Groth, B., Luciani, F., Slobedman, B., and Blyth, E. (2018). Mass cytometry for the assessment of immune reconstitution after hematopoietic stem cell transplantation. Front. Immunol. *9*, 1672.

Stikvoort, A., Chen, Y., Rådestad, E., Törlén, J., Lakshmikanth, T., Björklund, A., Mikes, J., Achour, A., Gertow, J., Sundberg, B., et al. (2017). Combining flow and mass cytometry in the search for biomarkers in chronic graft-versus-host disease. Front. Immunol. *8*, 717.

Takahashi, C., Au-Yeung, A., Fuh, F., Ramirez-Montagut, T., Bolen, C., Mathews, W., and O'Gorman, W.E. (2017). Mass cytometry panel optimization through the designed distribution of signal interference. Cytometry A *91*, 39–47.

Tanaka, A., and Sakaguchi, S. (2017). Regulatory T cells in cancer immunotherapy. Cell Res. *27*, 109–118.

van der Maaten, L., and Hinton, G. (2008). Visualizing data using t-SNE. J. Mach. Learn. Res. *9*, 2579–2605.

Van Gassen, S., Callebaut, B., Van Helden, M.J., Lambrecht, B.N., Demeester, P., Dhaene, T., and Saeys, Y. (2015). FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. Cytometry A *87*, 636–645.

van Unen, V., Höllt, T., Pezzotti, N., Li, N., Reinders, M.J.T., Eisemann, E., Koning, F., Vilanova, A., and Lelieveldt, B.P.F. (2017). Visual analysis of mass cytometry data by hierarchical stochastic neighbour embedding reveals rare cell types. Nat. Commun. *8*, 1740.

Villani, A.-C., Satija, R., Reynolds, G., Sarkizova, S., Shekhar, K., Fletcher, J., Griesbeck, M., Butler, A., Zheng, S., Lazo, S., et al. (2017). Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science *356*, eaah4573.

Weber, L.M., and Robinson, M.D. (2016). Comparison of clustering methods for high-dimensional single-cell flow and mass cytometry data. Cytometry A *89*, 1084–1096.

Weber, L.M., Nowicka, M., Soneson, C., and Robinson, M.D. (2019). diffcyt: Differential discovery in high-dimensional cytometry via high-resolution clustering. Commun Biol *2*, 183.

Wickham, H. (2016). ggplot2: Elegent Graphics for Data Analysis (Springer-Verlag).

Wong, K.L., Yeap, W.H., Tai, J.J.Y., Ong, S.M., Dang, T.M., and Wong, S.C. (2012). The three human monocyte subsets: implications for health and disease. Immunol. Res. *53*, 41–57.

Zunder, E.R., Finck, R., Behbehani, G.K., Amir, A.D., Krishnaswamy, S., Gonzalez, V.D., Lorang, C.G., Bjornson, Z., Spitzer, M.H., Bodenmiller, B., et al. (2015). Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm. Nat. Protoc. *10*, 316–333.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies for reference panel		
Anti-human CD45 89Y (clone H130)	Fluidigm	Cat# 3089003B, RRID:AB_2661851
Anti-human CD235ab (clone HIR2)	BioLegend	Cat# 306602, RRID:AB_314620
Anti-human CD61 (clone VI-PL2)	BioLegend	Cat# 336402, RRID:AB_1227584
Anti-human CD3 (clone UCHT1)	BioLegend	Cat# 300402, RRID:AB_314055
Anti-human CD19 142Nd (clone HIB19)	Fluidigm	Cat# 3142001B, RRID:AB_2651155
Anti-human CD117 143Nd (clone 104D2)	Fluidigm	Cat# 3143001B, RRID: N/A
Anti-human CD11b 144Nd (clone IRCF44)	Fluidigm	Cat# 3144001B, RRID:AB_2714152
Anti-human CD4 145Nd (clone RPA-T4)	Fluidigm	Cat# 3145001B, RRID:AB_2661789
Anti-human CD8a 14Nd (clone RPA-T8)	Fluidigm	Cat# 3146001B, RRID:AB_2687641
Anti-human CD11c 147Sm (clone BU15)	Fluidigm	Cat# 3147008B, RRID:AB_2687850
Anti-human CD14 148Nd (clone RMO52)	Fluidigm	Cat# 3148010B, RRID: N/A
Anti-human FceRI 150Nd (clone AER-37/CRA-1)	Fluidigm	Cat# 3150027B, RRID: N/A
Anti-human CD123 151Eu (clone 6H6)	Fluidigm	Cat# 3151001, RRID:AB_2661794
Anti-human TCRgd 152Sm (clone 11F2)	Fluidigm	Cat# 3152008B, RRID: N/A
Anti-human CD45RA 153Eu (clone HI100)	Fluidigm	Cat# 3153001B, RRID: N/A
Anti-human Tim-3 154Sm (clone F38-2E2)	Fluidigm	Cat# 3153008B, RRID:AB_2687644
Anti-human PD-L1 156Gd (clone 29E.2A3)	Fluidigm	Cat# 3156026B, RRID: N/A
Anti-human CD27 158Gd (clone L128)	Fluidigm	Cat# 3155001B, RRID:AB_2687645
Anti-human Tbet 160Gd (clone 4B10)	Fluidigm	Cat# 3160010B, RRID: N/A
Anti-human CD152 161Dy (clone 14D3)	Fluidigm	Cat# 3161004B, RRID: N/A
Anti-human FoxP3 162Dy (clone PCH101)	Fluidigm	Cat# 3162011A, RRID:AB_2687650
Anti-human CD33 163Dy (clone WM53)	Fluidigm	Cat# 3163023, RRID:AB_2687857
Anti-human CD45RO 164Dy (clone UCHL1)	Fluidigm	Cat# 3164007B, RRID: N/A
Anti-human CD127 165Ho (clone A019D5)	Fluidigm	Cat# 3165008B, RRID: N/A
Anti-human CCR7 167Er (clone G043H7)	Fluidigm	Cat# 3167009A, RRID: N/A
Anti-human Ki-67 168Er (clone B56)	Fluidigm	Cat# 3168007B, RRID:AB_2800467
Anti-human CD25 169Tm (clone 2A3)	Fluidigm	Cat# 3169003B, RRID:AB_2661806
Anti-human TCRVa24-Ja18 170Er (clone 6B11)	Fluidigm	Cat# 3170015B, RRID: N/A
Anti-human CD38 172Yb (clone HIT2)	Fluidigm	Cat# 3144014B, RRID:AB_2687640
Anti-human HLA-DR 174Yb (clone L243)	Fluidigm	Cat# 3174001B, RRID:AB_266539
Anti-human PD-1 175Lu (clone EH12.2H7)	Fluidigm	Cat# 3175008B, RRID: N/A
Anti-human CD56 176Yb (clone NCAM16.2)	Fluidigm	Cat# 3176008B, RRID:AB_2661813
Anti-human CD16 209Bi (clone 3G8)	Fluidigm	Cat# 3209002B, RRID:AB_2756431
Biological Samples		
PBMCs from healthy subjects	Stanford blood center	https://stanfordbloodcenter.org/
PBMCs from healthy subjects	Parker Institute for Cancer Immunotherapy	https://www.parkerici.org/
Tumor biopsies from cancer patients	UCSF	https://www.ucsf.edu/
PBMCs from bone marrow transplant patients	Stanford	https://www.stanford.edu/
Chemicals, Peptides, and Recombinant Proteins		
Sodium heparin	Sigma-Aldrich	Cat# H4784
Benzonase	Sigma-Aldrich	Cat# E1014
Cisplatin	Fluidigm	Cat# 201064

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
0.1 uM centrifugal filter	Millipore	Cat# UFC30VV00
Intercalator-Ir	Fluidigm	Cat# 201192B
Calibration Beads, 151/153Eu	Fluidigm	Cat# 201073
Calibration Beads, EQ™ Four Element	Fluidigm	Cat# 201078
Antibody Stabilizer	Candor Bioscience	Cat# 131 050
Critical Commercial Assays		
eBioscience Foxp3 / Transcription Factor Staining Buffer Set	Thermo Fisher Scientific	Cat# 00-5523-00
MaxPar conjugation set	Fluidigm	Cat# N/A
Deposited Data		
Dataset accession numbers FR-FCM-Z249 and FR-FCM-Z244	Flowrepository	https://flowrepository.org
Software and Algorithms		
Cytobank analysis software	Kotecha et al., 2010	https://www.cytobank.org, RRID:SCR_014043
R environment	R Development Core Team, 2008	https://www.r-project.org/, RRID:SCR_001905
Rtsne	van der Maaten and Hinton, 2008	https://github.com/jkrijthe/Rtsne, RRID:SCR_016900
statisticalScaffold R package	Spitzer et al., 2017	https://github.com/SpitzerLab/statisticalScaffold, RRID: N/A
Vortex	Samusik et al., 2016	https://github.com/nolanlab/vortex, RRID:SCR_017047
Astrolabe	N/A	https://www.astrolabediagnostics.com/, RRID: N/A
Normalizer	Finck et al., 2013	https://github.com/nolanlab/bead-normalization, RRID: N/A
Other		
CyTOF2 mass cytometer	Fluidigm	Cat# N/A

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Sean C. Bendall (bendall@stanford.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human subjects

PBMC samples

All samples from human subjects (see Table S2) were obtained and experimental procedures were carried out in accordance with the guidelines of the Stanford Institutional Review Board (IRB). Written informed consent was obtained from all subjects. For healthy donors, fresh whole human blood in heparin collection tubes or leukoreduction system (LSR) chamber contents (Terumo BCT) were obtained via the Stanford Blood Center. Samples from BMT patients were drawn on 30 and 90 post BMT. PBMCs were isolated via Ficoll (GE Healthcare) density gradient centrifugation, resuspended in fetal bovine serum (FBS, Omega Scientific) supplemented with 10% DMSO (Sigma) and stored in liquid nitrogen.

Tissue samples

Tissue samples (see Table S2) were collected fresh shortly after surgery and transported for processing on ice in transport medium (Leibovitz's L-15 medium supplemented with 6 g/L glucose and 15 mM HEPES buffer). Tumor samples were then finely minced and placed into tumor dissociation buffer (transport medium, 2% fetal bovine serum (FBS), 5mg/ml collagenase IV, 0.1 mg/ml DNase I) for 45 min at 37°C with gentle rotation. Following dissociation, cells were filtered through a 70 μm filter, centrifuged at 500 g for 5 min at 4°C, and resuspended in PBS with 5 mM EDTA. Cells were then mixed with viability buffer (PBS, 5 mM EDTA, 50 μM cisplatin) for 60 s at room temperature, quenched with wash buffer (PBS, 5 mM EDTA, 0.5% bovine serum albumin), centrifuged at 500 g for 5 minutes at 4°C, resuspended again in wash buffer, and fixed in 1.6% PFA for 10 min at RT. After fixation, cells were centrifuged at 600 g for 5 min at 4°C, rinsed with wash buffer, centrifuged again at 600 g for 5 min at 4°C, resuspended in freezing medium (PBS, 10% DMSO, 0.5% bovine serum albumin), and frozen at −80°C until staining.

METHOD DETAILS

Panel design and heavy-metal conjugation of antibodies

Most antibodies were obtained pre-conjugated to heavy-metal isotopes from Fluidigm. When possible, targets were allocated to specific heavy-metal isotopes following the sensitivity of the mass cytometers (e.g., placing lower abundance targets on higher sensitivity channels) and to avoid problems with potential spectral overlap as outlined previously (Takahashi et al., 2017). Where needed, in-house conjugations were performed using the MaxPar X8 antibody-labeling kit (Fluidigm) following an optimized and updated protocol (Hartmann et al., 2019). In short, antibody buffer exchange was performed by washing 100 μg of antibody with R buffer (Fluidigm) using a 50 kDa MWCO microfilter (Millipore) and centrifuging for 10 min, 12′000 g at RT. Antibodies were then reduced with 100 μL of 4 mM TCEP (Thermo Fisher) for 30 min at 37°C and washed two times with C buffer (Fluidigm). Metal chelation was performed by adding lanthanide metal solutions (final 0.05 M) to MaxPar chelating polymers in L-buffer (both Fluidigm) and incubating for 40 min at RT. Metal-loaded polymers were washed twice with L-buffer using a 3 kDa MWCO microfilter (Millipore) by centrifuging for 30 min, 12′000 g at RT. Partially reduced antibodies and metal-loaded polymers were incubated together for 60-120 min at 37°C. Conjugated antibodies were washed four times with 400 μL W buffer (Fluidigm) and collected by two centrifugations (2 min, 1'000 g, RT) with 50 μL of W buffer into an inverted column in a fresh 1.6 mL collection tube. Protein content was assessed by NanoDrop (Thermo Fisher) measurement, antibody stabilization buffer (Candor Bioscience) was added to a final volume of at least 50 v/v % and antibodies were stored at 4°C.

Mass cytometry workflow

Sample preparation

Cryopreserved PBMC and tumor biopsy samples where thawed into 10 mL of cold cell culture medium (RPMI-1640 (life technologies), 10% FBS, 1x L-glutamine, 1x penicillin/streptomycin (Thermo Fisher)) supplemented with 20 U/ml sodium heparin and 0.025 U/ml benzonase (Sigma) and washed once (250 g, 4°C).

Cellular barcoding

Where indicated, samples where barcoded and combined into a composite sample before surface staining. Barcoding was performed employing either a palladium-based barcoding approach applicable to fixed cells (Zunder et al., 2015) or a live cell barcoding methodology involving antibodies against the surface molecules beta-2-microglobulin and a sodium-potassium pump (CD298) as described (Hartmann et al., 2018).

Viability staining

Cisplatin (Sigma) was resuspended to 100mM in DMSO, pre-conditioned for 48 h at 37°C and stored at −20°C. Viability staining was performed by resuspending the sample in 1 mL of PBS and adding cisplatin to a final concentration of 500 nM, followed by incubation for 5 min at RT and washing with CSM. Where indicated, cells were fixed with 1.6% PFA in PBS for 10 min at RT and washed twice with cell staining medium (CSM: PBS with 0.5% BSA and 0.02% sodium azide (all Sigma)) before staining. In case live cell barcoding was employed, viability assessment was performed by substituting cisplatin with DCED-palladium (Sigma) and following the protocol as described here.

Antibody staining

Cell-surface antibody master-mix (2x) was prepared by adding appropriate dilutions of all cell-surface antibodies (Table S1; Key Resources Table) into 50 μL CSM per sample. If samples contained more than 3×10^6 cells, antibody volume (but not total CSM volume) was increased accordingly (e.g., 2-fold for up to 6×10^6 cells). The antibody master-mix was then filtered through a pre-wetted 0.1 μm spin-column (Millipore) to remove antibody aggregates and 50 μL were added to the sample resuspended in 50 μL of CSM. After incubation for 30 min at RT, cells were washed once with CSM. For intracellular staining, cells were fixed using the FoxP3 / transcription factor staining buffer set (Thermo Fisher Scientific) to fix for 1 h at RT. After fixation, samples were washed once with CSM and once with 1x permeabilization buffer (Thermo Fisher Scientific) by centrifugation for 5 min, 600 g at 4°C. Intracellular antibody master-mix (2x) was prepared analogously to the surface antibody mix by adding appropriate dilutions of all intracellular antibodies (see Table S1 and Key Resources Table) into 50 ul permeabilization buffer per sample. 50 μL of 2x antibody master mix was added to the samples in 50 μL permeabilization buffer and incubated for 1 h at RT. Cells were washed once with permeabilization buffer and once with CSM. Finally, samples were resuspended in intercalation solution (1.6% PFA in PBS and 0.5 μM iridium-intercalator (Fluidigm)) for 20 min at RT or overnight at 4°C.

Data acquisition

Before acquisition, samples were washed once in CSM and twice in ddH_2O and filtered through a cell strainer (Falcon). Cells were then resuspended at 1×10^6 cells/mL in ddH_2O supplemented with 1x EQ four element calibration beads (Fluidigm) and acquired on a CyTOF2 mass cytometer (Fluidigm).

Flow cytometry

PBMC samples were thawed as described above and subsequently treated with Fc blocking reagent (BioLegend) for 10 min at 4°C. Antibody cocktails were then added for 30 min and incubated at 4°C. All samples were washed with PBS containing BSA (0.5%), then fixed with 1.6% PFA for 10 min at RT. Finally, the samples were washed and analyzed on an LSRII flow cytometer (BD Biosciences) equipped with 405, 488, 561, and 640nm lasers.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data normalization and gating

After acquisition, data from acquired samples was bead-normalized using MATLAB-based software (Finck et al., 2013). Barcoded cells were assigned back to their initial samples using MATLAB-based debarcoding software (Zunder et al., 2015). Normalized data was then uploaded onto the Cytobank analysis platform (Kotecha et al., 2010) to perform initial gating and population identification using the indicated gating schemes (Figures 2 and S1).

Data visualization and analysis

For further downstream analysis, pre-gated data was imported into the R environment (R Development Core Team, 2008) using the flowCore package (Ellis et al., 2009). Data was transformed with an inverse hyperbolic sine (arcsinh) transformation using a cofactor of 5 and normalized to the 99.5th percentile of each respective channel before downstream tSNE and Scaffold analysis. Visualization of samples by tSNE dimensionality reduction was calculated using the Rtnse package (Krijthe, 2015) with default parameters: perplexity = 30, theta = 0.5, max_iter = 1000 using the indicated channels.

To build a reference scaffold, bead and percentile-normalized data from live, CD45^{+}, single, non-neutrophil cells was imported into the statisticalScaffold package (Spitzer et al., 2017). All available channels were used to build the reference maps. All population-relevant antigens were included in the clustering analysis. Astrolabe analysis was carried out by uploading bead-normalized data. Single-cell data was clustered using the FlowSOM R package (Van Gassen et al., 2015). Cell subset definitions follow (Finak et al., 2016; Maecker et al., 2012). Cluster labeling, method implementation, and visualization were done through the Astrolabe Cytometry Platform (Astrolabe Diagnostics, Inc.).

Statistical analysis

Cell frequencies are reported as medians unless stated otherwise. Standard error of median was calculated in R using bootstrapping with 1000-fold resampling. For frequency correlations between different centers and technologies, manually gated frequencies of cell populations were compared by linear regression using the lm() function. Hierarchical clustering using the R function hclust() was performed using the same frequency matrix.

To compare manual gating with automated clustering we employed the Matthews correlation coefficient (MCC) (Boughorbel et al., 2017; Matthews, 1975) which takes into account true and false positives as well as negatives and expresses these results in a single coefficient. A coefficient of +1 represents perfect agreement.

Differential abundance analysis for identifying GvHD-associated immune signatures was done through the Astrolabe platform using the edgeR R package (McCarthy et al., 2012; Robinson et al., 2010) following the method outlined in Lun et al. (2017). Samples from both time points were pooled for this analysis.

Visualization

Plots were created using the ggplot2 R package (Wickham, 2016). Schematic representations were created with biorender (https://biorender.com/). Figures were prepared in Illustrator (Adobe).

DATA AND CODE AVAILABILITY

Single-cell mass cytometry datasets for validating the reference panel across two research centers (FR-FCM-Z249) and for identifying stratifying populations in bone-marrow transplanted patients (FR-FCM-Z244) are available at flowrepository.org.

Cell Reports, Volume *28*

Supplemental Information

Comprehensive Immune Monitoring of Clinical Trials to Advance Human Immunotherapy

Felix J. Hartmann, Joel Babdor, Pier Federico Gherardini, El-Ad D. Amir, Kyle Jones, Bita Sahaf, Diana M. Marquez, Peter Krutzik, Erika O'Donnell, Natalia Sigal, Holden T. Maecker, Everett Meyer, Matthew H. Spitzer, and Sean C. Bendall

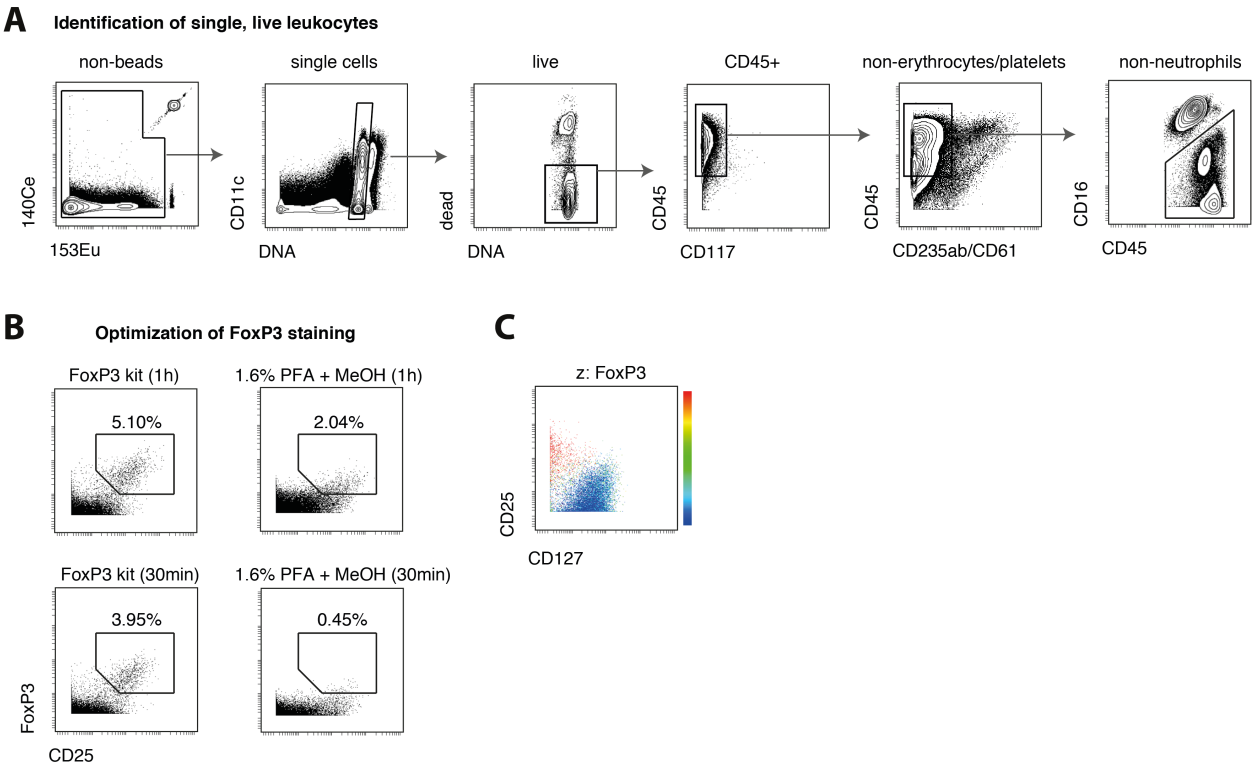


Figure S1. Identification of single, live leukocytes and optimization of FoxP3 staining. **Related to Figure 2.**

PBMCs or whole blood of healthy donors was stained with the reference panel as outlined in Table S1. (A) Serial gates used to identify single, live cells. Firstly, remaining beads are excluded and next, DNA positive events are identified. CD11c is used to discriminate duplets from monocyte populations which show higher DNA staining. Next, cisplatin-positive (dead) cells are excluded and CD45+ (non-mast cells) are selected. Potentially remaining erythrocytes, platelets and neutrophils are excluded. Representative examples shown here may be from different samples to demonstrate the utility of each gating step. (B) PBMCs from healthy subjects were first stained for surface antigen and then fixed using different fixation schemes. For the FoxP3 fixation permeabilization kit (left), cells were fixed with 1x fixation buffer for 1 h at RT. Intracellular staining was performed for 1 h (top) or 30 min (bottom) in 1x permeabilization buffer. Alternatively, cells were fixed with 1.6% PFA (right) for 10 min at RT and permeabilized with 100% cold MeOH for 10 min on ice. Cells were then stained intracellularly for 1 h (top) or 30 min (bottom) in CSM. Shown are $CD4^+$ T cells. (C) Cells are colored by FoxP3 intensity as well as CD127 and CD25 which can be used to identify Treg cells.

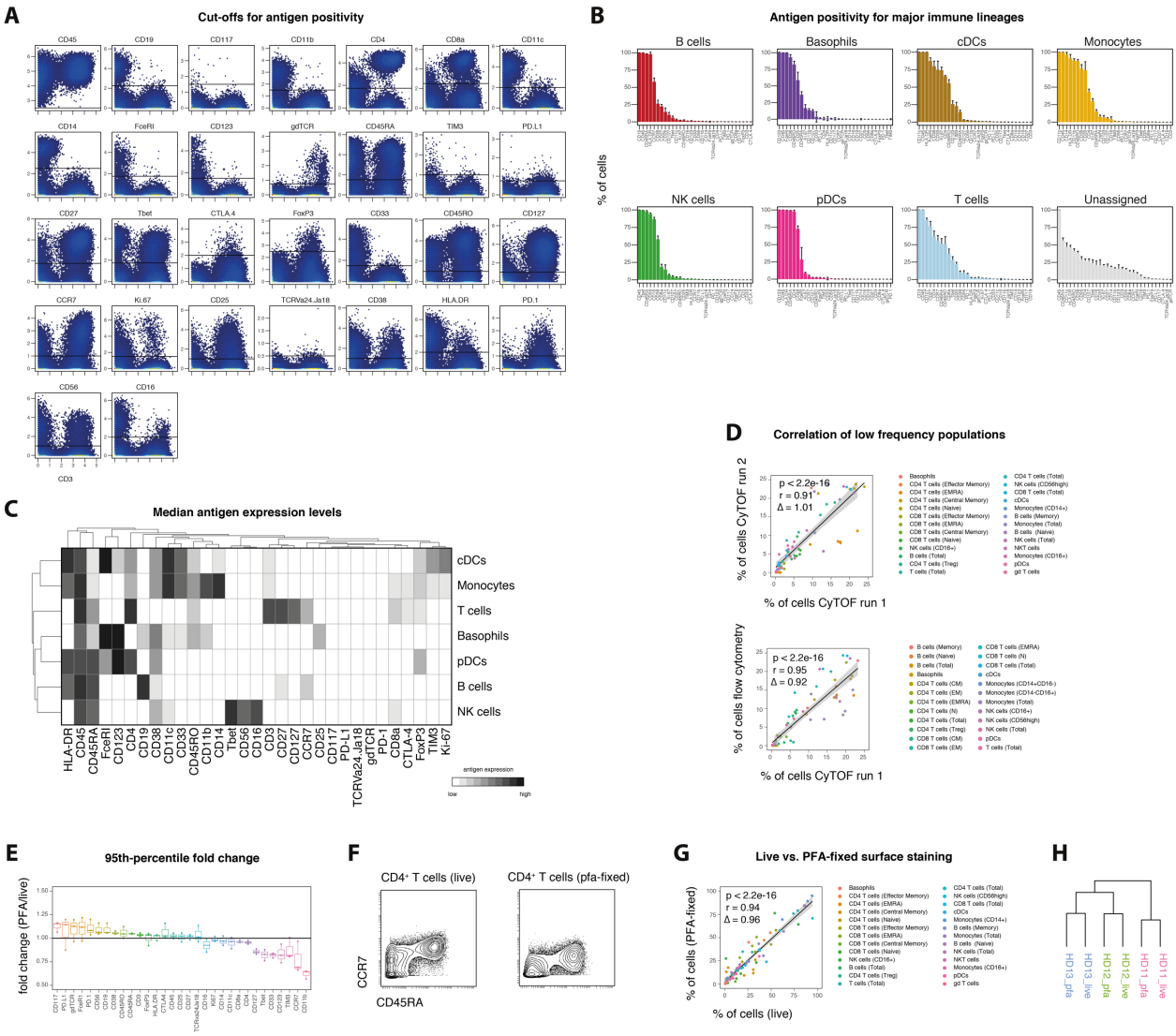


Figure S2. Robustness of immune cell lineage identification. Related to Figure 3. PBMCs from healthy donors (n = 5) were stained with the reference panel and analyzed by mass cytometry. Cells were pre-gated as single, live, CD45^{+}, non-neutrophils. (A) Cutoffs (black horizontal lines) used to determine positivity of antigen expression. (B) Median percentage of cells in each immune cell lineage which is positive for the indicated antigen based on the cutoffs as in A. Error bars represent s.e.m. (C) Median antigen expression levels on manually gated immune cell lineages. (D) Immune cell frequencies (only populations below 25% of total cells are shown) determined by two different CyTOF analyses (top) or by CyTOF and flow cytometry (bottom). Frequencies of immune lineages were determined through serial gating. Linear regression line is shown in black with the 95% confidence intervals (CI, shaded). Coefficients, p-values and slope Δ were calculated based on data from all donors. (E-H) PBMCs from healthy donors (n = 3) were stained directly with surface antibodies or fixed with 1.6% PFA for 10 min at RT before

surface staining. (E) To assess the influence of PFA-fixation on the maximum dynamic range, we first calculated the 95th percentile of each marker and subsequently calculated the respective fold change (FC) for all markers. The black horizontal line represents FC = 1. Boxplots depict the interquartile range (IQR) with a horizontal line representing the median. Whiskers extend to the farthest data point within a maximum of 1.5x IQR. Points represent individual samples. (F) Exemplary biaxial plots showing reduced CCR7 and comparable CD45RA staining intensities on fixed T cell populations. (G) Frequencies of immune cell populations were determined through serial gating and plotted against each other. Linear regression line is shown in black with the 95% confidence intervals (CI) shaded. Coefficients, p-values and slope Δ were calculated based on data from all donors. (H) Hierarchical clustering of unfixed or fixed samples based on frequencies as in G.

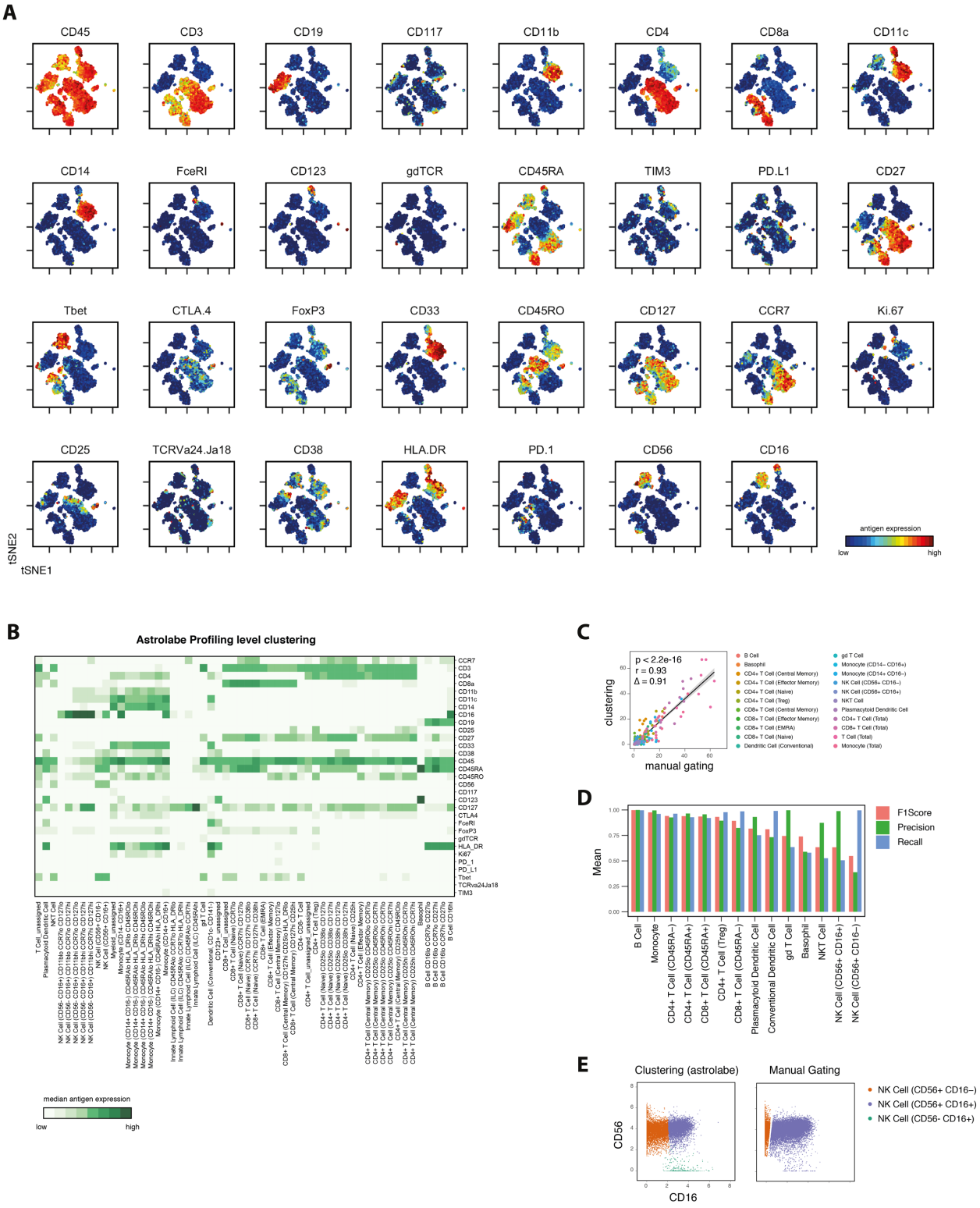


Figure S3. Automated profiling of immune cell populations. Related to Figure 4. (A) tSNE visualizations were produced from randomly subsampled data (20'000 cells) of five healthy donors. Indicated antigen expression levels are overlaid as a color dimension. (B) PBMC data from healthy donors was clustered and automatically annotated

using the Astrolabe platform. Shown are median expression levels of all antigens across all clusters on the Profiling level. (C) Frequencies of immune cell populations determined by manual gating or through Astrolabe clustering. Linear regression line is shown in black with the 95% confidence intervals (CI) shaded. Coefficients, p-values and slope Δ were calculated based on data from all donors. (D) Mean precision, recall and F1score between manual lineage assignments and Astrolabe-based clustering across different populations. (E) Example of differential thresholds employed in clustering (left) and manual gating (right).

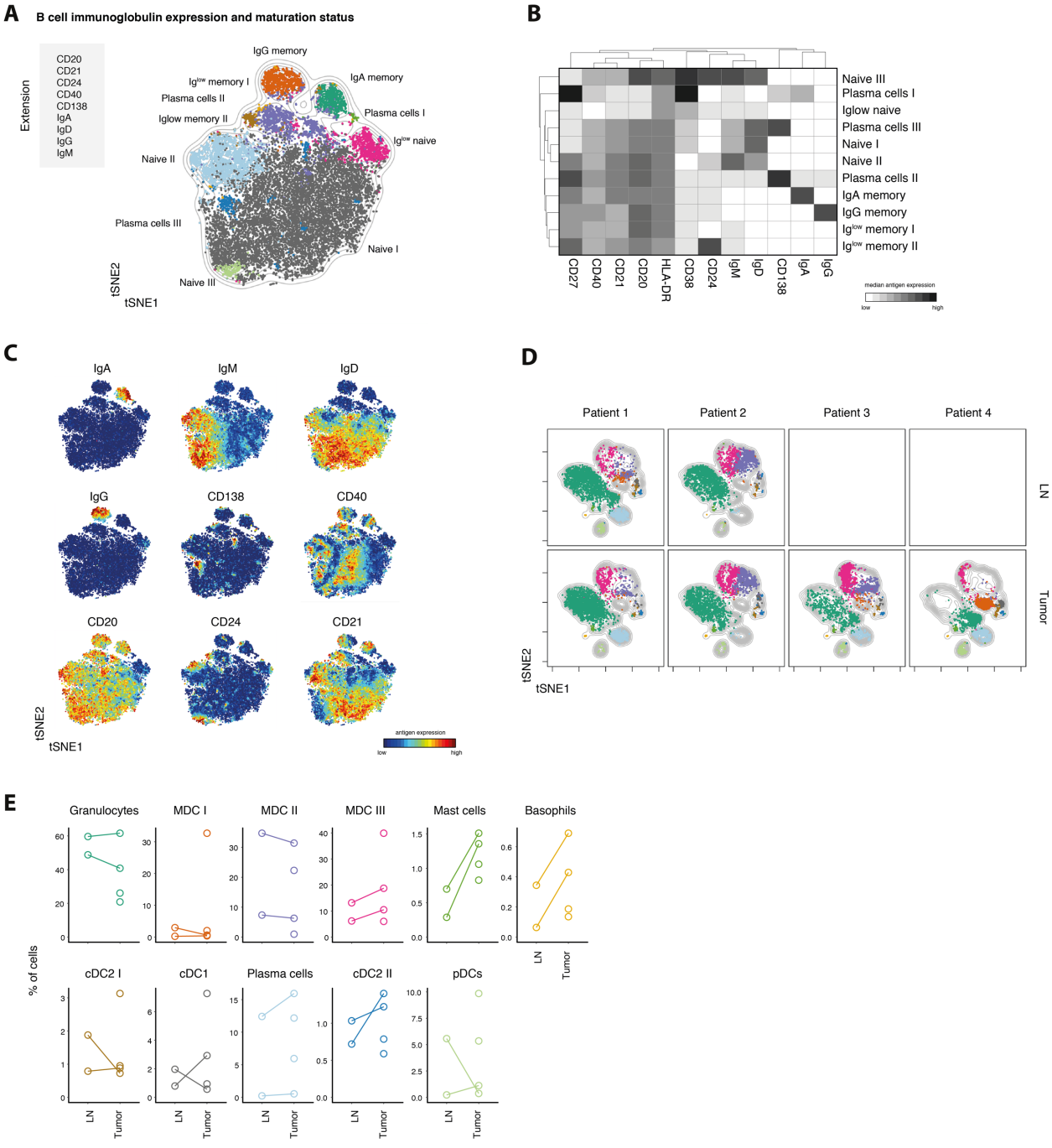


Figure S4. Exploration of heterogeneous populations through panel extensions. Related to Figure 6. (A) Antibodies targeting additional antigens of interest (left) were conjugated to non-occupied heavy metal isotopes. PBMCs from healthy donors (see Table S2) were stained with these antibodies in combination with the reference set. Data was pre-gated to exclude T cells and NK cells. To create a tSNE overview, data from all samples was randomly subsampled to 21'000 cells with equal contribution from all samples. Cells are colored by their FlowSOM-based cluster-assignment. Grey lines indicate the density distribution of the tSNE map. (B) Cluster-based

median expression levels for all population relevant antigens used in the tSNE and FlowSOM analysis. (C) Protein expression levels of all additional antigens are overlaid as a color-dimension onto the tSNE map. (D-E) Antibodies targeting additional myeloid antigens were conjugated to non-occupied heavy metal isotopes. Cells from lymph node biopsies (n = 2) and tumor biopsies (n = 4) of patients with head and neck carcinoma (see Table S2) were stained with these antibodies in combination with the reference set. (D) tSNE overview with colors denoting FlowSOM-cluster assignments stratified by sample. (E) Frequencies of FlowSOM-based clusters as in D in all samples. Lines connect different tissues of the same patients.

Table S1. Reference panel of anti-human antibodies for mass cytometry. Related to Figure1.

Isotope	Element	Marker	Clone	Staining step	Dilution factor	Catalogue number
89	Y	CD45	H130	surface	100	3089003B
102	Pd					
104	Pd					
105	Pd					
106	Pd					
108	Pd					
110	Pd					
113	In					
115	In					
139	La	CD235ab/CD61	HIR2/VI-PL2	surface		custom
140	Ce					
141	Pr	CD3	UCHT1	surface		custom
142	Nd	CD19	HIB19	surface	100	3142001B
143	Nd	CD117	104D2	surface	100	3143001B
144	Nd	CD11b	IRCF44	surface	400	3144001B
145	Nd	CD4	RPA-T4	surface	200	3145001B
146	Nd	CD8a	RPA-T8	surface	1600	3146001B
147	Sm	CD11c	BU15	surface	200	3147008B
148	Nd	CD14	RMO52	surface	1600	3148010B
149	Sm					
150	Nd	FceRI	AER-37 (CRA-1)	surface	100	3150027B
151	Eu	CD123	6H6	surface	100	3151001B
152	Sm	gdTCR	11F2	surface	100	3152008B
153	Eu	CD45RA	HI100	surface	100	3153001B
154	Sm	TIM3	F38-2E2	surface	100	3154010B
155	Gd					
156	Gd	PD-L1 (CD274)	29E.2A3	surface	100	3156026B
157	Gd					
158	Gd	CD27	L128	surface	400	3158010B
159	Tb					
160	Gd	Tbet	4B10	intracellular	100	3160010B
161	Dy	CD152 (CTLA-4)	14D3	intracellular	800	3161004B
162	Dy	FoxP3	PCH101	intracellular	400	3162011A
163	Dy	CD33	WM53	surface	100	3163023B
164	Dy	CD45RO	UCHL1	surface	200	3164007B
165	Ho	CD127	A019D5	surface	100	3165008B
166	Er					
167	Er	CCR7 (CD197)	G043H7	surface	100	3167009A
168	Er	Ki-67	B56	intracellular	400	3168007B
169	Tm	CD25	2A3	surface	100	3169003B
170	Er	TCR Va24-Ja18	6B11	intracellular	100	3170015B
171	Yb					
172	Yb	CD38	HIT2	surface	400	3172007B
173	Yb					
174	Yb	HLA-DR	L243	surface	100	3174001B
175	Lu	PD-1	EH12.2H7	surface	100	3175008B
176	Yb	CD56	NCAM16.2	surface	100	3176008B
191	Ir	DNA				
191	Ir	DNA				
194	Pt					
195	Pt	Live/dead				
196	Pt					
198	Pt					
209	Bi	CD16	3G8	surface	100	3209002B

Table S2. Donor characteristics. Related to Figure 3, Figure 4, Figure 5 and Figure 6.

Healthy donors

Donor_ID	Tissues_analyzed	Diagnosis	Age	Sex
HD01	PBMC	Healthy Donor	NA	NA
HD02	PBMC	Healthy Donor	NA	NA
HD03	PBMC	Healthy Donor	NA	NA
HD04	PBMC	Healthy Donor	NA	NA
HD05	PBMC	Healthy Donor	NA	NA
HD06	PBMC	Healthy Donor	39	M
HD07	PBMC	Healthy Donor	35	M
HD08	PBMC	Healthy Donor	46	F
HD09	PBMC	Healthy Donor	31	F
HD10	PBMC	Healthy Donor	54	F
HD11	PBMC	Healthy Donor	NA	NA
HD12	PBMC	Healthy Donor	NA	NA
HD13	PBMC	Healthy Donor	NA	NA

Tumor

Donor_ID	Tissues_analyzed	Diagnosis	Age	Sex	Tumor_site	Tumor	p16_Istus	Locoregional LN Met	AJCC Stage
T01	Tumor	SCC	57	M	Tonsil	Primary	Positive	Yes	pT2N2b
T02	Tumor, Met	SCC (poor)	78	F	Tongue	Recurrent	NA	Yes	pT3N2c
T03	Tumor, Met	SCC (moderate)	83	F	Gingiva	Primary	NA	Yes	pT4aN2b
T04	Tumor	SCC (moderate)	69	F	Floor of mouth	Primary	NA	Yes	pT4aN2c
T05	Tumor	SCC (basaloid)	55	M	Supraglottis	Recurrent	NA; prior was positive	NA	pT4aNx
T06	Tumor	SCC (moderate to poor)	70	M	Ventral tongue / Floor of	Primary	NA	Yes	pT2N1
T07	Tumor	Warthin Tumor	76	M	Parotid Gland	Primary	NA	No	NA

BMT patients

Donor_ID	Tissues_analyzed	Diagnosis	Age	Sex	aGVHD	GVHD_onset	cGVHD	SAE	CMV (D/R)	Conditioning	GVHD PPx	Graft_type	Sampling
BMT01	PBMC	AML	58	F	None	NA	None	E.coli, bacteremia	+/+	Bu/Cy	Tacrolimus, MTX	PBMSC	day_30, day_90
BMT02	PBMC	AML	57	M	None	NA	None	Coccoides infection, CMV	+/+	Bu/Cy	Tacrolimus, MTX	PBMSC	day_30, day_90
BMT03	PBMC	B cell ALL	31	F	None	NA	None	Tacrolimus associated	-/+	TBI/VP-16	Tacrolimus, MTX	PBMSC	day_30, day_90
BMT04	PBMC	AML	39	F	None	NA	None	Clostridium bacteremia,	+/+	TBI/VP-16	Tacrolimus, MTX	PBMSC	day_30, day_90
BMT05	PBMC	AML	58	F	None	NA	None	None	+/+	Bu/Cy	Tacrolimus, MTX	PBMSC	day_30, day_90
BMT06	PBMC	FLT3+ AML	42	M	None	NA	None	Mild VOD	+/-	Bu/Cy	Sirolimus	CD34/Treg/Tcon	day_30, day_90
BMT07	PBMC	ETP-ALL	34	M	None	NA	None	Herpes labialis	+/+	TBI/Cy/VP-16	Sirolimus	CD34/Treg/Tcon	day_30, day_90
BMT08	PBMC	ALL	20	M	None	NA	None	Acute cholecystitis	+/+	Bu/Cy	Tacrolimus	CD34/Treg/Tcon	day_30, day_90
BMT09	PBMC	CML blast crisis	54	M	None	NA	None	Coronavirus URI	+/-	Bu/Cy	Tacrolimus	CD34/Treg/Tcon	day_30, day_90
BMT10	PBMC	AML	53	M	None	NA	None	Unspecified transaminitis	-/-	Bu/Cy	Tacrolimus	CD34/Treg/Tcon	day_30, day_90
BMT11	PBMC	MF	56	M	None	NA	None	Streptritis bacteremia	+/-	Bu/Cy	Tacrolimus	CD34/Treg/Tcon	day_30, day_90
BMT12	PBMC	Refractory NHL	49	M	None	NA	None	CoNS bacteremia	-/-	BCNU/VP-16/Cy	Sirolimus	CD34/Treg/Tcon	day_30, day_90
BMT13	PBMC	AML	61	F	Yes, grade 3 skin and liver	day_22	Mild day_474	CMV reactivation	+/-	TBI/Cy/VP-16	None	CD34/Treg(frozen)/T	day_30, day_90
BMT14	PBMC	Gamma delta NHL	45	F	Yes, grade 1 skin	day_30	Mild day_145	Gram negative bacteremia	+/+	BCNU/VP-16/Cy	None	CD34/Treg(frozen)/T	day_30, day_90
BMT15	PBMC	AML	51	M	Yes, grade 1 skin	day_24	None	Acute diverticulitis, HHV6	+/+	TBI/Cy/VP-16	None	CD34/Treg(frozen)/T	day_30, day_90

Table S3. Extension panels. Related to Figure 6.

B cell immunoglobulin expression and maturation status

Isotope	Element	Marker	Clone	Staining step
113	In	CD40	5C3	surface
115	In	CD20	2H7	surface
140	Ce	IgM	G20-127	surface
149	Sm	IgA	IS11-8E10	surface
155	Gd	IgD	IA6-2	surface
157	Gd	CD138	DL-101	surface
159	Tb	IgG	M1310G05	surface
166	Er	CD24	ML5	surface
171	Yb	CD21	LT21	surface
173	Yb	CD40L	24-31	surface

Myeloid cell diversity across multiple tissues

Isotope	Element	Marker	Clone	Staining step
113	In	CD66cd	YTH71.3	surface
115	In	CD7	M-T701	surface
140	Ce	CD86	IT2.2	surface
149	Sm	CD1c	L161	surface
155	Gd	CD64	10.1	surface
157	Gd	CD206	15-2	surface
159	Tb	2B4	M2B4	surface
164	Dy	CD172ab	SE5A5	surface
166	Er	CD141	1A4	surface
171	Yb	CD40	5C3	surface
173	Yb	CCR2	K036C2	surface