



- (51) **International Patent Classification:**
G01N 33/50 (2006.01) G01N 33/574 (2006.01)
C12Q 1/68 (2018.01)
- (21) **International Application Number:**
PCT/US2023/060573
- (22) **International Filing Date:**
12 January 2023 (12.01.2023)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
63/299,377 13 January 2022 (13.01.2022) US
- (71) **Applicants:** WASHINGTON UNIVERSITY [US/US];
One Brookings Drive, St. Louis, Missouri 63130 (US).
THE BOARD OF TRUSTEES OF THE LELAND
STANFORD JUNIOR UNIVERSITY [US/US]; Building
170, Third Floor, Main Quad, P.O. Box 20386, Stanford,
California 94305 (US).
- (72) **Inventors:** CHAUDHURI, Aadel; One Brookings Drive,
St. Louis, Missouri 63130 (US). NEWMAN, Aaron; 265
Campus Drive, Palo Alto, California 94305 (US).

(74) **Agent:** MCCAY, Michael; Campus Box 8013-0537-01,
660 South Euclid, St. Louis, Missouri 63110 (US).

(81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) **Title:** LIQUID BIOPSY ANALYSIS OF CELLULAR STATES TO PREDICT IMMUNOTHERAPY TOXICITY

(57) **Abstract:** Methods are disclosed for predicting a likelihood of developing a severe immune-related adverse event (irAE) associated with the administration of an immunotherapy in a melanoma patient based on abundances of activated CD4 memory T cells and/or diversities of T cell receptors (TCR) within a peripheral blood sample obtained from the patient.

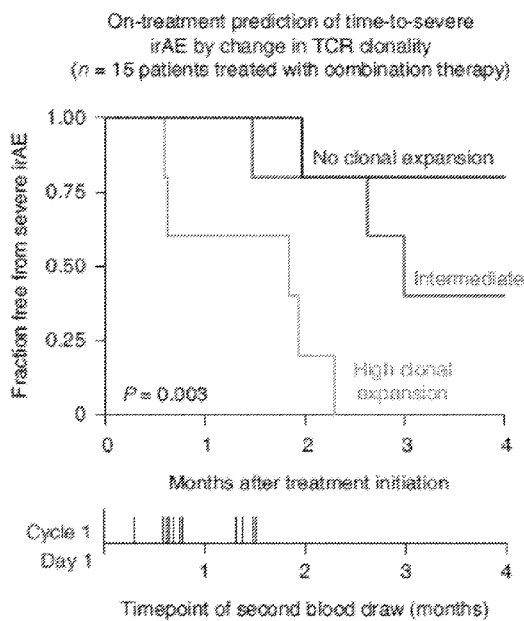


FIG. 5D



Published:

— *with international search report (Art. 21(3))*

TITLE OF THE INVENTION**LIQUID BIOPSY ANALYSIS OF CELLULAR STATES TO PREDICT
IMMUNOTHERAPY TOXICITY****CROSS-REFERENCE TO RELATED APPLICATIONS**

5 This application claims priority from U.S. Provisional Application Serial No. 63/299,377 filed on January 13, 2022, which is incorporated herein by reference in its entirety.

**STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR
DEVELOPMENT**

10 This invention was made with government support under CA187192 and CA238711 awarded by the National Institutes of Health. The government has certain rights in the invention.

MATERIAL INCORPORATED-BY-REFERENCE

Not applicable.

15 FIELD OF THE INVENTION

The present disclosure generally relates to methods for predicting immunotherapy toxicity in patients.

BACKGROUND OF THE INVENTION

20 Although ICIs have revolutionized cancer treatment, approximately 10–60% of ICI-treated patients with melanoma currently develop severe immune-related toxicities, with the rate of toxicity closely linked to the specific therapy administered. Also known as irAEs, ICI-induced toxicities impact a range of organ systems, including the lungs, liver, heart, skin, pituitary gland, and gastrointestinal tract, and can be associated with substantial morbidity requiring
25 urgent medical intervention. Such morbidities can lead to the suspension of anticancer treatment, and in the most severe cases, death. The biological drivers of irAEs are poorly characterized and there is no method in standard clinical practice to identify which patients are at the highest risk for developing them.

Accordingly, several groups have investigated potential biomarkers of ICI-

induced toxicity based on blood or tumor analysis. However, these studies have generally been focused on early on-treatment prediction or single organ systems, with only modest performance for predicting irAEs in the pretreatment setting independent of the affected organ system. Recently, a candidate
5 pneumonitis-only irAE biomarker using tumor immunohistochemistry was reported; however, this biomarker was indirectly identified from The Cancer Genome Atlas, which lacks toxicity annotations, and was evaluated in a case-control setting without the inclusion of low-grade irAEs. Another group identified a single-nucleotide polymorphism within the gene encoding microRNA-146a that
10 was associated with severe irAE development. Still, other groups have identified ICI response biomarkers without examining irAEs.

Given the considerable heterogeneity of ICI-induced irAEs, including variation in their timing, severity, and location, determining the factors that cause them has remained challenging. Pre-existing autoantibodies, autoreactive tissue-
15 resident T cells, and T cells with specificity for viral antigens stemming from chronic viral infection have all been implicated in irAEs. Changes in the gut microbiome leading to increased colonic interleukin-1 β expression were also recently reported in ICI-induced colitis. Given these observations, several groups have investigated parallels between irAEs and autoimmune disease. Indeed,
20 case reports have shown that ICIs can cause frank autoimmunity, suggesting that irAEs could represent subclinical autoimmunity in a subset of patients. However, whether a common immunological state precedes distinct manifestations of ICI-induced toxicity is unknown.

SUMMARY OF THE INVENTION

25 Among the various aspects of the present disclosure is the provision of methods and compositions for the prediction of the likelihood of developing a severe immune-related adverse event (irAE) in a patient receiving immunotherapy based on a biomarker derived from a peripheral blood sample obtained from the patient prior to receiving the immunotherapy. In one aspect,
30 disclosed methods include obtaining a peripheral blood sample from a subject prior to receiving an immunotherapy treatment and quantifying an abundance of activated CD4 memory T cells and a diversity of T cell receptors (TCR) in the peripheral blood sample. Preferred methods additionally include classifying the

patient as likely to develop a severe irAR if the abundance of activated CD4 memory T cells in combination with the diversity of T cell receptors (TCR) exceeds a threshold (sometimes referred to herein as a model index). The threshold can be determined using a model index that identifies levels of

5 activated CD4 memory T cells and TCR diversity and provides a range of values that represent a ceiling beyond which the patient is susceptible to irAR. In one aspect, a value of the model index (the combination of CD4 memory T cells and TCR diversity values) that exceeds a predetermined threshold is predictive of a more severe irAR. In an additional aspect, the method further includes

10 determining the threshold value by reference to known clinical standards. In another aspect, the disclosed methods include determining the abundance of activated CD4 memory T cells and the diversity of T cell receptors (TCR) using at least one of: bulk RNA-sequencing (CIBERSORTx and MiXCR), mass cytometry by time of flight (CyTOF), immunoSEQ® TCR-β profiling, droplet-

15 based scRNA-sequencing and scTCR-sequencing, and targeted RNA-sequencing using an RNA panel targeted to activated CD4 memory T cells.

In other aspects of the present disclosure, methods for predicting a likelihood of developing a severe immune-related adverse event (irAE) in a patient receiving an immunotherapy are disclosed. In one aspect, the methods

20 include obtaining a first peripheral blood sample from a subject prior to receiving an immunotherapy treatment and a second peripheral blood sample subsequent to the administration of the immunotherapy. The disclosed methods in these other aspects include quantifying a first TCR diversity level from the first peripheral blood sample and a second TCR diversity level from the second

25 peripheral blood sample. The disclosed methods further methods include obtaining a degree of TCR expansion by subtracting the first TCR diversity level from the second TCR diversity level. The disclosed methods further include classifying the patient as likely to develop severe irAR if the degree of TCR expansion exceeds a threshold value. In one aspect, the methods include

30 predicting a time of onset of the severe irAR based on the degree of TCR expansion, wherein a higher degree of TCR expansion is predictive of an earlier onset of severe irAR. In one aspect, the methods include determining the diversity levels of T cell receptors (TCR) using at least one of: bulk RNA-sequencing (CIBERSORTx and MiXCR), mass cytometry by time of flight

(CyTOF), immunoSEQ® TCR-β profiling, droplet-based scRNA-sequencing and scTCR-sequencing, and targeted RNA-sequencing using an RNA panel targeted to activated CD4 memory T cells.

Other objects and features will be in part apparent and in part pointed out
5 hereinafter.

DESCRIPTION OF THE DRAWINGS

Those of skill in the art will understand that the drawings, described below, are for illustrative purposes only. The drawings are not intended to limit the scope of the present teachings in any way.

10 FIG. 1 is a schematic of a study schema described in the present disclosure, including an overview of patients included in this study, a summary of their irAE status, exclusion criteria, and downstream analyses that were performed. Among 78 total eligible patients, 71 were evaluable for irAE analysis after exclusion criteria were applied.

15 FIG. 2A is a set of color-coded charts representing the characteristics of the single-cell discovery cohort from FIG. 1, including the highest irAE grade experienced and durable clinical response status after the start of immunotherapy.

FIG. 2B is a UMAP chart of a viSNE projection of peripheral blood cells
20 analyzed by CyTOF. t-SNE, t-distributed stochastic neighbor embedding.

FIG. 2C is a (Left) heatmap showing the relative abundance of 20 cell states identified by CyTOF in 18 patients, grouped by future irAE status, as well as (Right) a graph showing the association of cell state abundance with severe irAE development. Statistical significance was determined by a two-sided,
25 unpaired Wilcoxon rank-sum test and expressed as directional $-\log_{10}$ P values. For associations with no severe irAE, $-\log_{10}$ P values were multiplied by -1 . Q values were determined by the Benjamini–Hochberg method.

FIG. 2D is a graph of the frequencies of CD4 TEM cells (CyTOF) in the pretreatment peripheral blood of patients stratified by future irAE status (no
30 severe irAE, n= 10 patients; severe irAE, n= 8 patients). The box center lines, box bounds, and whiskers denote the medians, first and third quartiles, and minimum and maximum values, respectively. Statistical significance was

determined by a two-sided, unpaired Wilcoxon rank-sum test.

FIG. 3A is a UMAP of peripheral blood cells profiled by scRNA-seq from 13 patients coanalyzed by CyTOF (FIG. 2A), colored by cell type, patient, and state (n= 32). T/NKT, NK-like T cells.

5 FIG. 3B is a UMAP of cell state abundances (scRNA-seq) versus future irAE status and CD4 TEM cell frequencies (CyTOF). The former was quantified by a two-sided, unpaired Wilcoxon rank-sum test and expressed as $-\log_{10} P$ values. For associations with no severe irAE, $-\log_{10} P$ values were multiplied by -1 . CD4 T cell states 5 and 3 are indicated together as CD4 T 5 + 3.

10 FIG. 3C is a heatmap of DEGs ($P_{\text{adj}} < 0.05$) between CD4 T cell states 5 and 3 and other CD4 T cell states. Within each state, the columns represent the mean expression from individual patients converted to z-scores.

FIG. 3D is a set of 2 graphs of (Left) frequencies of candidate activated and resting subsets of CD4 T 5 + 3 cell states in 13 patients stratified by no
15 severe (n= 7) and severe (n= 6) irAE status. Activation markers with counts per million (CPM) > 0 were considered expressed. Significance was determined by a two-sided, unpaired Wilcoxon rank-sum test, as well as (Right) a receiver operating characteristic curve plot showing the performance of the CD4 T 5 + 3 subsets (from the left panel) for predicting severe irAE development. NS, not
20 significant.

FIG. 3E is a graph showing pretreatment TCR clonotype diversity within each T cell state, total T cells, CD8 T cells, CD4 T cells, and activated versus
resting CD4 T 5 + 3 cells (defined as in FIG. 3D), grouped by future irAE status. TCR diversity was calculated for all patients with at least 100 TCR clones (n= 9).
25 States are ordered by the AUC between TCR diversity and severe irAE status.

FIG. 3F is a color-coded chart of the mean expression of key lineage and activation genes in CD4 T cell states. States within the box are consistent with
TEM and TEM-like phenotypes. The box center lines, box bounds, and whiskers indicate the medians, first and third quartiles, and minimum and maximum
30 values, respectively.

FIG. 4A is a graph showing the association between pretreatment peripheral blood leukocyte composition (CIBERSORTx) and severe irAE

development in bulk cohort 1 (n= 26 patients) and bulk cohort 2 (n= 27 patients) (FIG. 1). Significance was determined by two-sided, unpaired Wilcoxon rank-sum test and expressed as $-\log_{10}$ P values. For associations with no severe irAE, $-\log_{10}$ P values were multiplied by -1 .

5 FIG. 4B is a graph showing TCR clonotype diversity (Shannon entropy) in both bulk cohorts (n= 53 patients), stratified by future irAE status (no severe irAE, n= 36; severe irAE, n= 17). The box center lines, box bounds, and whiskers denote the medians, first and third quartiles, and minimum and maximum values, respectively. Significance was determined by a two-sided, unpaired Wilcoxon
10 rank-sum test.

 FIG. 4C is a graph showing the development of a composite model for the prediction of severe irAEs, integrating activated CD4 TM cell abundance and TCR clonotype diversity from pretreatment peripheral blood transcriptomes, with model scores trained on bulk cohort 1 and shown across both cohorts. The cut-
15 point for high/ low scores was optimized using Youden's J statistic on bulk cohort 1.

 FIG. 4D is a set of 2 graphs of (Left) a ROC plot showing composite model performance in bulk cohort 2 (held-out validation), whether applied to all patients (both therapies, n= 27), combination therapy patients (n= 11) or PD-1
20 monotherapy patients (n= 16), as well as (Right) a ROC plot showing composite model performance in bulk cohorts 1 and 2, whether trained on PD-1 patients (n= 29) and tested on combination therapy patients (n= 24) or vice versa. The AUC is shown for each ROC curve.

 FIG. 4E is a graph of composite model scores for all bulk cohort patients
25 (n= 53) after model training for severe irAE development with LOOCV (FIG. 13), grouped by the highest irAE grade per patient. The box center lines, box bounds, and whiskers indicate the medians, first and third quartiles, and minimum and maximum values within $1.5\times$ the interquartile range of the box limits, respectively. Statistical significance was determined by a Kruskal-Wallis test.

30 FIG. 5A is a graph showing pretreatment prediction of time-to-severe irAE onset in patients treated with combination therapy. The cut-point was optimized using composite model scores trained with LOOCV. Only patients from bulk cohorts 1 and 2 who did not experience early progression were analyzed (n= 23).

Statistical significance was assessed by a two-sided log-rank test.

FIG. 5B is a set of two graphs showing TCR clonal dynamics in relation to severe irAE development in patients treated with combination therapy. Left: Change in TCR clonality from baseline after initiation of combination therapy as measured by $1 - \text{Pielou's evenness}$, with future irAE status indicated by color. Right: Same as the left but showing change in clonality according to future irAE status. Significance was determined by a two-sided, unpaired Wilcoxon rank-sum test.

FIG. 5C is a graph showing enrichment of a CD4 T 5 + 3 gene signature in CD4 T cells from pretreatment PBMC samples obtained from 3 patients analyzed in FIG. 5B, all of whom developed severe irAEs and showed TCR clonal expansion after ICI initiation (FIG. 15D). The box center lines, box bounds, and whiskers indicate the medians, first and third quartiles and minimum and maximum values, respectively. The points denote cells profiled by scRNA-seq and annotated either by Azimuth (CD4 naive, $n = 245$ cells; CD4 TCM, $n = 320$ cells) or by their clonal persistence from baseline to early on-treatment time points (persistent CD4, $n = 190$ cells). The most persistent CD4 clonotypes in this analysis showed evidence of clonal expansion (FIG. 15F and G). Significance was determined relative to persistent cells by a two-sided, unpaired Wilcoxon rank-sum test. ssGSEA, single-sample GSEA.

FIG. 5D is a graph of the differences in freedom from severe irAE stratified by the degree of TCR clonal expansion after initiating combination therapy, as measured by the change in $1 - \text{Pielou's evenness}$. Patients were grouped into the following tertiles: no clonal expansion ($n = 5$), intermediate ($n = 5$), and high clonal expansion ($n = 5$). Statistical significance was assessed by a two-sided log-rank test.

FIG. 6 is a color-coded chart showing the large-scale assessment of circulating leukocytes in autoimmune diseases. Enrichment of circulating leukocyte levels in two autoimmune disorders relative to healthy controls. Leukocyte composition was determined by CIBERSORTx. Significance was determined by a two-sided, unpaired Wilcoxon rank-sum test and integrative meta z-score. Details of the analytical workflow and underlying datasets are provided in FIG. 16.

FIG. 7A is a UMAP representation of pretreatment peripheral blood leukocytes profiled by droplet-based scRNA-seq (10x Genomics) from 13 patients with metastatic melanoma, colored by major cell lineages, severe irAE status, TCR expression by scV(D)J-seq, and BCR expression by scV(D)J-seq
5 (related to FIG. 3A).

FIG. 7B is a schematic of the unsupervised hierarchical clustering (average linkage) of the mean log2 transcriptome per CD4 T cell cluster identified from scRNA-seq data.

FIG. 7C is a dot plot showing the average expression of key activation
10 (HLA-DX, MKI67) and lineage markers (SELL, CCR7) in CD4 T cell clusters.

FIG. 7D is a graph of the unsupervised hierarchical clustering (average linkage) of the mean log2 transcriptome per CD4 T cell cluster identified from scRNA-seq data showing all pairwise combinations of scRNA-seq clusters within each of the major cell types analyzed (B cells, CD4 T cells, CD8 T cells, NK
15 cells, monocytes). Across 82 possible pairwise combinations, CD4 T 5 + 3 achieved the highest Spearman correlation against CD4 TEM levels enumerated by CyTOF and the strongest association with severe irAE development. Cells annotated as 'T/NKT' were collapsed into CD8 T cells.

FIG. 7E is a graph of the unsupervised hierarchical clustering (average
20 linkage) of the mean log2 transcriptome per CD4 T cell cluster identified from scRNA-seq data showing all pairwise combinations ranked by the mean of each feature following unit variance normalization (mean of 0 and standard deviation of 1). In this analysis, the $-\log_{10}$ P-value for the association with severe irAE (two-sided, unpaired Wilcoxon rank sum test) was normalized to unit variance
25 without considering the direction of the association.

FIG. 8A shows UMAP projections of scRNA-seq data generated in this work, embedded and labeled by Azimuth using a reference PBMC atlas of 162k cells profiled by scRNA-seq and 228 antibodies.

FIG. 8B is a confusion matrix showing the agreement between phenotypic
30 labels determined by marker genes and unsupervised clustering (rows; related to FIG. 3A and FIG. 7A) versus reference-guided annotation with Azimuth (columns). In total, 85% of single cells assigned to a major lineage group by Azimuth (B cells, CD4 T, CD8 T, NK cells, monocytes) were assigned to the

same identity by canonical marker gene assessment. Given the absence of NKT cells in the reference atlas used for Azimuth, the T/NKT cluster defined by unsupervised analysis was relabeled as CD8 T cells.

FIG. 8C is a graph of the same analysis as in FIG. 3B but shown for all 27 phenotypic states identified by Azimuth. Among these states, CD4 TEM was most associated with severe irAE and CyTOF-enumerated CD4 TEM. A population combining CD4 TEM and CD4 Proliferating states was also strongly associated with severe irAE. The latter showed the highest expression of HLA-DX and the lowest expression of SELL (panel d), consistent with an activated CD4 TEM phenotype.

FIG. 8D is a dot plot depicting key activation and lineage markers among CD4 T cell states annotated by Azimuth.

FIG. 8E is a set of violin plots showing protein expression levels imputed by Azimuth using antibody-derived tag (ADT) data, supporting the combination of CD4 TEM and CD4 Proliferating states shown in FIG. 8C and F.

FIG. 8F is a grid showing the performance of top-ranking cell subsets identified by Azimuth and unsupervised clustering for the prediction of severe irAEs. The combined CD4 T 5 + 3 clusters (FIG. 3B) were more associated with severe irAE and CyTOF than the top-ranking reference-guided population (FIG. 3C). Statistical significance was calculated using a two-sided, unpaired Wilcoxon rank sum test. Data in all panels shown are from the 13 samples profiled by scRNA-seq in FIG. 3.

FIG. 9A is a graph showing the association between severe irAE development and pretreatment levels of T cell states identified by unsupervised clustering (left) and memory-like T cell states identified by Azimuth (right) in 13 PBMC samples profiled by scRNA-seq (FIG. 1 and 3A). Activated cells were defined as those expressing HLA-DX or MKI67 (CPM > 0); resting cells were defined by the absence of HLA-DX and MKI67 expression (CPM = 0).

FIG. 9B is a set of graphs showing an analysis of activated, resting, and parental T cell subsets in relation to severe irAE development. Left: Association between severe irAE development and pretreatment levels of memory T cell subsets, total CD4 and CD8 T cells, and total T cells quantified by CyTOF, for all 18 patients analyzed in the single-cell discovery cohort (FIG. 1 and 2A).

Activated phenotypes were defined as CD38+ or HLA-DR+ or Ki67+. Resting phenotypes were defined as CD38- HLA-DR- Ki67-. Right: ROC plot showing the performance of activated and resting CD4 TEM subsets (left panel) for predicting severe irAE development. Cell fractions were assessed relative to total PBMC content. Statistical significance in a, b was determined by a two-sided, unpaired Wilcoxon rank sum test and nominal $-\log_{10}$ P-values are displayed. $-\log_{10}$ P-values were further multiplied by -1 for associations with no severe irAE.

FIG. 10A is a schematic showing the key TCR diversity measures and the impact of cell abundance, TCR richness, and distinct clonal repertoires on such measures. Hypothetical CD4 naïve and TEM cell subsets are shown as examples. Triangles depicting differences in magnitude are not drawn to scale.

FIG. 10B is a graph of mean Shannon entropy versus mean clonality ($1 - \text{Pielou's evenness}$) for each CD4 T cell state identified by unsupervised clustering of scRNA-seq data. CD4 T 5 + 3 (FIG. 3B and C), a TEM state enriched for activated cells, shows elevated clonality relative to other CD4 states, as expected for this phenotype, while also showing higher diversity (Shannon entropy), indicating elevated richness.

FIG. 10C is a schematic showing the distribution of EM-like CD4 T cell states (from FIG. 3F) with available scTCR clonotype data.

FIG. 10D is a set of graphs showing the association between severe irAE development and TCR diversity (Shannon entropy) in pseudo-bulk T cells from pretreatment blood, shown for all T cell states identified by scRNA-seq (left) and after the removal of the EM-like states indicated in FIG. 10C (no severe irAE, $n=5$ patients; severe irAE, $n=4$ patients). Bounds of the box and whiskers indicate medians, 1st and 3rd quartiles, and minimum and maximum values, respectively.

FIG. 10E is a graph showing the same association as in FIG. 10D but shown for EM-like states alone. Bounds of the box and whiskers indicate medians, 1st and 3rd quartiles, and minimum and maximum values, respectively.

FIG. 10F is a graph showing the area under the curve (AUC) for the association between pretreatment peripheral TCR diversity (Shannon entropy) and severe irAE development, shown for all combinations of the constituent cell

states in e, including the combined CD4 T 5 + 3 cluster after restricting to activated cells (CPM > 0 for HLA-DX or MKI67). Of note, no other combination of activated EM-like states achieved an AUC > 0.85 in this analysis.

5 FIG. 10G is a graph showing BCR clonotype diversity (Shannon entropy), shown for each B cell state identified by unsupervised clustering (FIG. 3A). In FIG. 10B and D-F, only patients with at least 100 TCR clones were analyzed (n= 9). The same patients were analyzed in FIG. 10G for consistency. Bounds of the box and whiskers indicate medians, 1st and 3rd quartiles, and minimum and maximum values, respectively.

10 FIG. 11A is a graph showing the expression of developmentally-regulated marker genes in major CD4 T cell subsets from the LM22 signature matrix (MAS5 normalized), showing that the LM22 reference signature for activated CD4 memory T cells has a TEM profile.

15 FIG. 11B is a graph showing CIBERSORTx versus mass cytometry for the enumeration of activated CD4 memory T cells in the pretreatment peripheral blood of 17 metastatic melanoma patients. A linear regression line with 95% confidence band is shown. Concordance and significance were determined by Pearson r and a two-sided t-test, respectively. While activated CD4 memory T cells quantitated by CyTOF were defined by CD38 expression in this plot, other
20 activated CD4 TEM subsets were also significantly correlated with CIBERSORTx (FIG. 11C).

FIG. 11C is a cross-correlation plot of lymphocyte subset frequencies determined by CyTOF and CIBERSORTx. Act., Activated.

25 FIG. 11D is a cross-correlation plot showing the correlation between activated CD4 memory T cell levels inferred by CIBERSORTx and 14 memory T cell states profiled by CyTOF, including CD38+ activated subsets manually gated within each population, in PBMCs from 17 metastatic melanoma patients.

30 FIG. 11E is a scatter plot depicting the global correlation of lymphocyte subsets enumerated by CIBERSORTx and flow cytometry in peripheral blood samples from five healthy subjects profiled by bulk RNA-seq. A linear regression line with 95% confidence band is shown. Concordance and significance were determined by Pearson r and a two-sided t-test, respectively. As monocytes were variably underestimated by cytometry compared to complete blood counts,

all results in b–e are expressed as a function of total lymphocytes.

FIG. 11F is a graph showing the distribution of activated CD4 memory T cell levels quantitated by CyTOF (CD38+, HLA-DR+ or Ki67+ CD4 TEM cells, n= 28 patients), scRNA-seq (HLA-DX+ or MKI67+ cells within CD4 T clusters 5 and 3, n= 13 patients), and CIBERSORTx (n= 60 patients) across all irAE-evaluable samples profiled by each modality in this work. Box center lines, bounds of the box, and whiskers indicate medians, 1st and 3rd quartiles, and minimum and maximum values, respectively. Statistical significance was determined by a Kruskal-Wallis test. n.s., not significant ($P > 0.05$).

FIG. 12A is a graph showing an association between baseline bulk TCR diversity and the highest irAE grade observed for each patient in bulk cohorts 1 and 2, shown for Shannon entropy and stratified by therapy type. Patients treated with combination therapy are stratified by future irAE status: no severe irAE (n= 10) versus severe irAE (n= 14 patients) (left) and irAE grade (right): 0/1 (n= 3), 2 (n= 7), 3 (n= 12), and 4 (n= 2). Two-group comparisons were assessed by a two-sided, unpaired Wilcoxon rank sum test. n.s., not significant ($P > 0.05$). Linear regression was applied to evaluate the median value of each measure grouped by irAE grade (insets). The significance of linear concordance was determined by a two-sided t-test. Grades 0 and 1 reflect no toxicity and asymptomatic toxicity, respectively, and were combined. The box center lines, bounds of the box, and whiskers denote medians, 1st and 3rd quartiles, and minimum and maximum values within $1.5 \times$ IQR (interquartile range) of the box limits, respectively.

FIG. 12B is a graph showing the association between baseline bulk TCR diversity and the highest irAE grade observed for each patient in bulk cohorts 1 and 2, shown for the Gini-Simpson index and stratified by therapy type. Patients treated with combination therapy are stratified by future irAE status: no severe irAE (n= 10) versus severe irAE (n= 14 patients) (left) and irAE grade (right): 0/1 (n= 3), 2 (n= 7), 3 (n= 12), and 4 (n= 2). Two-group comparisons were assessed by a two-sided, unpaired Wilcoxon rank sum test. n.s., not significant ($P > 0.05$). Linear regression was applied to evaluate the median value of each measure grouped by irAE grade (insets). The significance of linear concordance was determined by a two-sided t-test. Grades 0 and 1 reflect no toxicity and

FIG. 13C is a plot showing the performance of the composite model versus other candidate pretreatment factors for predicting severe irAE development. The composite model was trained in bulk cohort 1 (BC1) and validated in bulk cohort 2 (BC2) or vice versa, as indicated.

5 FIG. 13D is a graph showing the performance of the composite model trained on bulk cohort 1 for predicting severe irAEs in different patient subgroups from bulk cohort 2. DCB, durable clinical benefit; NDB, no durable clinical benefit; GI, gastrointestinal.

10 FIG. 13E is a graph showing composite model scores determined by LOOCV for all bulk cohort patients treated with combination therapy (n= 24), stratified by future irAE grade: 0/1 (n= 3), 2 (n= 7), 3 (n= 12), and 4 (n= 2). Center lines, bounds of the box, and whiskers indicate medians, 1st and 3rd quartiles, and minimum and maximum values within $1.5 \times$ IQR (interquartile range) of the box limits, respectively. Statistical significance was determined by a Kruskal-
15 Wallis test.

FIG. 13F is a graph showing model performance for predicting grade 2 +, 3 +, or 4 irAE development in combination therapy patients using the scores in FIG. 13E.

20 FIG. 13G is a graph showing composite model scores determined by LOOCV in both bulk cohorts (n= 53 patients) versus the number of symptomatic irAEs (grade 2 +) per patient. Center lines, bounds of the box, and whiskers indicate medians, 1st and 3rd quartiles, and minimum and maximum values within $1.5 \times$ IQR (interquartile range) of the box limits, respectively. Statistical significance was determined by a Kruskal-Wallis test.

25 FIG. 13H is a graph showing composite model scores determined by LOOCV in both bulk cohorts (n= 53 patients) versus the number of organ system toxicities per patient. Center lines, bounds of the box, and whiskers indicate medians, 1st and 3rd quartiles, and minimum and maximum values within $1.5 \times$ IQR (interquartile range) of the box limits, respectively. Statistical significance
30 was determined by a Kruskal-Wallis test.

FIG. 13I is a plot showing the distribution of irAEs across patients and organ systems. Patients from bulk cohorts 1 and 2 are organized by decreasing composite model scores determined via LOOCV. The line distinguishing high/low

scores was optimized using LOOCV.

FIG. 13J is a graph showing the fraction of patients in both bulk cohorts that developed irAEs in at least 2 organ systems versus those that did not, stratified by the threshold in FIG. 13I. Significance was determined by a two-sided Fisher's exact test.

FIG. 14 is a set of graphs showing composite model performance for predicting time to severe irAE in validation bulk cohort 2.

FIG. 14A is a graph showing a–c, Kaplan-Meier analysis for freedom from severe irAE in bulk cohort 2 for patients treated with combination or PD1 immune checkpoint blockade (a), combination therapy (b), or PD1 monotherapy (c), stratified by the composite model score. Statistical significance was calculated by a two-sided log-rank test. In all panels, training was performed in bulk cohort 1, and the cut-point predicting severe irAE was optimized for bulk cohort 1 using Youden's J statistic. Notably, the analyses in a–c were landmarked between treatment initiation and three months following treatment initiation, with all severe irAEs occurring within this period. The Kaplan-Meier plots are shown out to four months given the extended follow-up of patients that did not develop any severe irAE.

FIG. 14A is a graph showing Kaplan-Meier analysis for freedom from severe irAE in bulk cohort 2 for patients treated with combination or PD1 immune checkpoint blockade, stratified by the composite model score. Statistical significance was calculated by a two-sided log-rank test. In all panels, training was performed in bulk cohort 1, and the cut-point predicting severe irAE was optimized for bulk cohort 1 using Youden's J statistic. Notably, the analyses were landmarked between treatment initiation and three months following treatment initiation, with all severe irAEs occurring within this period. The Kaplan-Meier plots are shown out to four months given the extended follow-up of patients that did not develop any severe irAE.

FIG. 14B is a graph showing Kaplan-Meier analysis for freedom from severe irAE in bulk cohort 2 for patients treated with combination therapy, stratified by the composite model score. Statistical significance was calculated by a two-sided log-rank test. In all panels, training was performed in bulk cohort 1, and the cut-point predicting severe irAE was optimized for bulk cohort 1 using

Youden's J statistic. Notably, the analyses in FIG. 14A-C were landmarked between treatment initiation and three months following treatment initiation, with all severe irAEs occurring within this period. The Kaplan-Meier plots are shown out to four months given the extended follow-up of patients that did not develop any severe irAE.

FIG. 14C is a graph showing Kaplan-Meier analysis for freedom from severe irAE in bulk cohort 2 for patients treated with PD1 monotherapy, stratified by the composite model score. Statistical significance was calculated by a two-sided log-rank test. In all panels, training was performed in bulk cohort 1, and the cut-point predicting severe irAE was optimized for bulk cohort 1 using Youden's J statistic. Notably, the analyses in a-c were landmarked between treatment initiation and three months following treatment initiation, with all severe irAEs occurring within this period. The Kaplan-Meier plots are shown out to four months given the extended follow-up of patients that did not develop any severe irAE.

FIG. 15A is a graph showing evenness (Pielou's index) of TCR repertoires assembled by MiXCR (bulk RNA-seq) and immunoSEQ® (genomic DNA) from paired pretreatment PBMC samples (n= 15 combination therapy patients). Concordance and significance were determined by Spearman ρ and a two-sided t-test, respectively.

FIG. 15B is a graph similar to that in FIG. 5B but showing clonality for each pre- and on-treatment PBMC sample. Statistical significance was determined by a two-sided, paired Wilcoxon rank sum test. ns, not significant ($P > 0.05$).

FIG. 15C is a graph showing the fraction of pretreatment peripheral blood TCR clonotypes detected on-treatment in 15 combination therapy patients, stratified by no severe (n= 6) and severe (n= 9) irAE status. Clonotypes with matching productive CDR3 β -chain nucleotide sequences were considered identical. Center lines, bounds of the box, and whiskers indicate medians, 1st and 3rd quartiles, and minimum and maximum values, respectively. Significance was determined by a two-sided, unpaired Wilcoxon rank sum test.

FIG. 15D is a schematic showing persistent T cell clones identified by immunoSEQ® were cross-referenced with scTCR-seq and scRNA-seq data of

pretreatment PBMCs from the same three patients (YUALOE, YUNANCY, YUHONEY), all of whom received combination therapy and developed severe ICI-induced toxicity.

FIG. 15E is a dot plot showing log₂ expression of key lineage and activation markers across major T cell states annotated by Azimuth along with persistent clones classified into CD4 and CD8 T cells.

FIG. 15F is a graph showing an aggregate change from baseline in the productive frequencies of persistent clonotypes, stratified by lineage (n= 2 cell types) and patient (n= 3). The sum of the difference in productive frequencies (on-treatment % – pretreatment %) was calculated from immunoSEQ® data. Bars denote mean +/- SD.

FIG. 15G is a set of graphs showing peripheral blood TCR-β profiling with immunoSEQ®. Top: Change in bulk TCR clonality from baseline (Fig. 5b). Bottom: Same as FIG. 15F but showing the underlying clonotypes, where circle size is proportional to pretreatment clone frequency (immunoSEQ®).

FIG. 15H is a graph similar to that seen in FIG. 5D but restricted to blood draws taken cycle 1 day 1 of combination therapy and < 1 month later (n=7 patients).

FIG. 16 is a schema of a large-scale assessment of peripheral blood leukocytes in autoimmune disorders versus healthy controls. Schema describing the workflow and statistical meta-analysis for evaluating the enrichment of individual circulating leukocyte subsets in autoimmune disorders relative to healthy controls (FIG. 6). In brief, CIBERSORTx was applied to enumerate 15 leukocyte subsets in bulk RNA-seq or microarray profiles of peripheral blood samples from patients with either systemic lupus erythematosus^{57–59} (SLE; n= 239) or inflammatory bowel disease (IBD; n= 348) compared to healthy controls. For each dataset and cell subset, a two-sided, unpaired Wilcoxon rank sum test was applied to assess the difference in relative abundance between healthy and disease phenotypes. Results were subsequently combined across studies by meta-z statistics (Meth

FIG. 17A is a schematic showing a–e, Gating hierarchies and staining results for CD4 T cell subsets and NKT cells profiled by CyTOF from pretreatment PBMCs. All CD4 T cell subsets except T regulatory cells (Tregs)

were gated analogously for CD8 T cells. TCM, central memory T cell; TEM, effector memory T cell; EMRA, CD45RA+ terminally differentiated effector memory T cell.

FIG. 17B is a schematic showing gating hierarchies and staining results for activated vs. resting CD4 TEM cells profiled by CyTOF from pretreatment PBMCs.

FIG. 17C is a schematic showing gating hierarchies and staining results for monocyte subsets profiled by CyTOF from pretreatment PBMCs.

FIG. 17D is a schematic showing gating hierarchies and staining results for B cell subsets profiled by CyTOF from pretreatment PBMCs.

FIG. 17E is a schematic showing gating hierarchies and staining results for NK cell subsets profiled by CyTOF from pretreatment PBMCs.

FIG. 18 is a set of graphs showing a comparison of automated and manual cell state quantitation from CyTOF data.

FIG. 18A is a scatter plot showing the concordance in frequencies between automated gating (Astrolabe) and manual gating for the indicated peripheral blood cell types. Concordance was assessed by Pearson correlation and linear regression (95% confidence band is shown). A two-sided t-test was used to assess statistical significance. Data are from patients analyzed by CyTOF in FIG. 1 (n = 18).

FIG. 18B is a scatter plot similar to that seen in FIG. 18A but for CD4 TEM cells. A representative gating scheme for CD4 TEM is provided in FIG. 7A. Concordance was assessed by Pearson correlation and linear regression (95% confidence band is shown). A two-sided t-test was used to assess statistical significance. Data in are from patients analyzed by CyTOF in FIG. 1 (n = 18).

FIG. 18C is a graph showing the association of pretreatment CD4 TEM abundance with severe irAE development when expressed as a fraction of total PBMCs, total T cells, or CD4 T cells. Box center lines, bounds of the box, and whiskers denote medians, 1st and 3rd quartiles, and minimum and maximum values, respectively. Data in are from patients analyzed by CyTOF in FIG. 1 (n = 18).

FIG. 19 is a schematic representing the methods described in the current

disclosure.

DETAILED DESCRIPTION OF THE INVENTION

Severe immune-related adverse events (irAEs) occur in ~60% of melanoma patients treated with combination immune checkpoint inhibitors (ICIs) and cause treatment-related morbidity and mortality. However, there is no reliable way to predict the development or timing of severe irAEs.

Pre-treatment and on-treatment analysis of cellular states and T cell receptors predict immunotherapy toxicity onset and timing. We specifically combine the abundance of activated CD4 T effector memory cells and the diversity of the T cell receptor repertoire in peripheral blood to yield a composite biomarker predictive of immunotherapy toxicity. Clonal expansion from pre- to on-treatment predicts the timing of severe toxicity. A targeted RNA sequencing panel enables this analysis in a practical and cost-effective manner.

Immunotherapy toxicities (immune-related adverse events) can be severe, dangerous, life-threatening, and deadly. We have no way in practice to predict them reliably, early or pre-treatment. Doing so would facilitate toxicity anticipation, earlier intervention, and more personalized and precise administration of immunotherapy.

In various aspects, methods of predicting immunotherapy toxicity in patients are disclosed. The disclosed methods are based on the discovery that two factors derived from the analysis of peripheral blood samples comprising activated CD4 memory T cell abundance levels and bulk TCR diversity strongly correlate with severe immunity-related adverse event (irAE) development. In various aspects, a liquid biopsy method of predicting immunotherapy toxicity in patients is disclosed that includes obtaining a peripheral blood sample from a subject prior to receiving an immunotherapy treatment. In various aspects, the method further includes quantifying an abundance of activated CD4 memory T cells and a diversity of T cell receptors (TCR) within the peripheral blood sample. In some aspects, the method further includes determining a model index predictive of the likelihood of the patient developing severe irAR, in which the model index comprises a combination of the abundance of activated CD4 memory T cells and a diversity of T cell receptors (TCR). The method further includes classifying the patient as likely to develop a severe irAR if the value of

the model index exceeds a threshold value. In some aspects, the method further comprises predicting the severity of the irAR based on the value of the model index, wherein a higher value of the model index is predictive of a more severe irAR. The threshold for a higher value of the model index can be determined
5 empirically or by reference to known clinical standards.

In various other aspects, methods of predicting immunotherapy toxicity in patients are disclosed that are based on the degree of TCR expansion, defined herein as the increase in the diversity of TCRs over an early period of immunotherapy relative to the pre-treatment diversity of TCRs. The methods
10 include obtaining a first peripheral blood sample from the patient prior to initiation of an immunotherapy and obtaining a second peripheral blood sample early in the administration of an immunotherapy to the patient. The methods further include obtaining a first TCR diversity from the first peripheral blood sample and a second TCR diversity from the second peripheral blood sample. The methods
15 further include subtracting the first TCR diversity from the second TCR diversity to obtain a degree of TCR expansion. In some aspects, the methods further include classifying the patient as likely to develop severe irAR if the degree of TCR expansion exceeds a threshold value. In some aspects, the methods further include predicting the time of onset of the severe irAR based on the
20 degree of TCR expansion.

MOLECULAR ENGINEERING

The following definitions and methods are provided to better define the present invention and to guide those of ordinary skill in the art in the practice of the present invention. Unless otherwise noted, terms are to be understood
25 according to conventional usage by those of ordinary skill in the relevant art.

The terms "heterologous DNA sequence", "exogenous DNA segment" or "heterologous nucleic acid," as used herein, each refers to a sequence that originates from a source foreign to the particular host cell or, if from the same source, is modified from its original form. Thus, a heterologous gene in a host
30 cell includes a gene that is endogenous to the particular host cell but has been modified through, for example, the use of DNA shuffling. The terms also include non-naturally occurring multiple copies of a naturally occurring DNA sequence. Thus, the terms refer to a DNA segment that is foreign or heterologous to the

cell, or homologous to the cell but in a position within the host cell nucleic acid in which the element is not ordinarily found. Exogenous DNA segments are expressed to yield exogenous polypeptides. A "homologous" DNA sequence is a DNA sequence that is naturally associated with a host cell into which it is
5 introduced.

Expression vector, expression construct, plasmid, or recombinant DNA construct is generally understood to refer to a nucleic acid that has been generated via human intervention, including by recombinant means or direct chemical synthesis, with a series of specified nucleic acid elements that permit
10 transcription or translation of a particular nucleic acid in, for example, a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector can include a nucleic acid to be transcribed operably linked to a promoter.

A "promoter" is generally understood as a nucleic acid control sequence
15 that directs the transcription of a nucleic acid. An inducible promoter is generally understood as a promoter that mediates the transcription of an operably linked gene in response to a particular stimulus. A promoter can include necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA element. A promoter can optionally
20 include distal enhancer or repressor elements, which can be located as many as several thousand base pairs from the start site of transcription.

A "transcribable nucleic acid molecule" as used herein refers to any nucleic acid molecule capable of being transcribed into an RNA molecule. Methods are known for introducing constructs into a cell in such a manner that
25 the transcribable nucleic acid molecule is transcribed into a functional mRNA molecule that is translated and therefore expressed as a protein product. Constructs may also be constructed to be capable of expressing antisense RNA molecules, in order to inhibit the translation of a specific RNA molecule of interest. For the practice of the present disclosure, conventional compositions
30 and methods for preparing and using constructs and host cells are well known to one skilled in the art (see e.g., Sambrook and Russel (2006) Condensed Protocols from Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, ISBN-10: 0879697717; Ausubel et al. (2002) Short Protocols

in Molecular Biology, 5th ed., Current Protocols, ISBN-10: 0471250929; Sambrook and Russel (2001) Molecular Cloning: A Laboratory Manual, 3d ed., Cold Spring Harbor Laboratory Press, ISBN-10: 0879695773; Elhai, J. and Wolk, C. P. 1988. Methods in Enzymology 167, 747-754).

5 The "transcription start site" or "initiation site" is the position surrounding the first nucleotide that is part of the transcribed sequence, which is also defined as position +1. All other sequences of the gene and its controlling regions may be numbered relative to this initiation site. Downstream sequences (i.e., further protein encoding sequences in the 3' direction) can be denominated positive,
10 while upstream sequences (mostly of the controlling regions in the 5' direction) are denominated negative.

"Operably-linked" or "functionally linked" refers preferably to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a regulatory DNA
15 sequence is said to be "operably linked to" or "associated with" a DNA sequence that codes for an RNA or a polypeptide if the two sequences are situated such that the regulatory DNA sequence affects expression of the coding DNA sequence (i.e., that the coding sequence or functional RNA is under the transcriptional control of the promoter). Coding sequences can be operably-
20 linked to regulatory sequences in sense or antisense orientation. The two nucleic acid molecules may be part of a single contiguous nucleic acid molecule and may be adjacent. For example, a promoter is operably linked to a gene of interest if the promoter regulates or mediates transcription of the gene of interest in a cell.

25 A "construct" is generally understood as any recombinant nucleic acid molecule such as a plasmid, cosmid, virus, autonomously replicating nucleic acid molecule, phage, or linear or circular single-stranded or double-stranded DNA or RNA nucleic acid molecule, derived from any source, capable of genomic integration or autonomous replication, comprising a nucleic acid molecule where
30 one or more nucleic acid molecule has been operably linked.

A construct of the present disclosure can contain a promoter operably linked to a transcribable nucleic acid molecule operably linked to a 3' transcription termination nucleic acid molecule. Constructs can also include, but

are not limited to, additional regulatory nucleic acid molecules from, e.g., the 3'-untranslated region (3' UTR). Constructs can include but are not limited to the 5' untranslated regions (5' UTR) of an mRNA nucleic acid molecule which can play an important role in translation initiation and can also be a genetic component in an expression construct. These additional upstream and downstream regulatory nucleic acid molecules may be derived from a source that is native or heterologous with respect to the other elements present on the promoter construct.

The term "transformation" refers to the transfer of a nucleic acid fragment into the genome of a host cell, resulting in genetically stable inheritance. Host cells containing the transformed nucleic acid fragments are referred to as "transgenic" cells, and organisms comprising transgenic cells are referred to as "transgenic organisms".

"Transformed," "transgenic," and "recombinant" refer to a host cell or organism such as a bacterium, cyanobacterium, animal, or plant into which a heterologous nucleic acid molecule has been introduced. The nucleic acid molecule can be stably integrated into the genome as generally known in the art and disclosed (Sambrook 1989; Innis 1995; Gelfand 1995; Innis & Gelfand 1999). Known methods of PCR include, but are not limited to, methods using paired primers, nested primers, single specific primers, degenerate primers, gene-specific primers, vector-specific primers, partially mismatched primers, and the like. The term "untransformed" refers to normal cells that have not been through the transformation process.

"Wild-type" refers to a virus or organism found in nature without any known mutation.

Design, generation, and testing of the variant nucleotides, and their encoded polypeptides, having the above required percent identities, and retaining a required activity of the expressed protein is within the skill of the art. For example, directed evolution and rapid isolation of mutants can be according to methods described in references including, but not limited to, Link et al. (2007) *Nature Reviews* 5(9), 680-688; Sanger et al. (1991) *Gene* 97(1), 119-123; Ghadessy et al. (2001) *Proc Natl Acad Sci USA* 98(8) 4552-4557. Thus, one skilled in the art could generate a large number of nucleotide and/or polypeptide

variants having, for example, at least 95-99% identity to the reference sequence described herein and screen such for desired phenotypes according to methods routine in the art.

Nucleotide and/or amino acid sequence identity percent (%) is understood
5 as the percentage of nucleotide or amino acid residues that are identical with
nucleotide or amino acid residues in a candidate sequence in comparison to a
reference sequence when the two sequences are aligned. To determine percent
identity, sequences are aligned and if necessary, gaps are introduced to achieve
the maximum percent sequence identity. Sequence alignment procedures to
10 determine percent identity are well known to those of skill in the art. Often
publicly available computer software such as BLAST, BLAST2, ALIGN2, or
Megalign (DNASTAR) software is used to align sequences. Those skilled in the
art can determine appropriate parameters for measuring alignment, including
any algorithms needed to achieve maximal alignment over the full length of the
15 sequences being compared. When sequences are aligned, the percent
sequence identity of a given sequence A to, with, or against a given sequence B
(which can alternatively be phrased as a given sequence A that has or
comprises a certain percent sequence identity to, with, or against a given
sequence B) can be calculated as: percent sequence identity = $X/Y100$, where X
20 is the number of residues scored as identical matches by the sequence
alignment program's or algorithm's alignment of A and B and Y is the total
number of residues in B. If the length of sequence A is not equal to the length of
sequence B, the percent sequence identity of A to B will not equal the percent
sequence identity of B to A.

25 Generally, conservative substitutions can be made at any position so long
as the required activity is retained. So-called conservative exchanges can be
carried out in which the amino acid which is replaced has a similar property as
the original amino acid, for example, the exchange of Glu by Asp, Gln by Asn,
Val by Ile, Leu by Ile, and Ser by Thr. For example, amino acids with similar
30 properties can be Aliphatic amino acids (e.g., Glycine, Alanine, Valine, Leucine,
Isoleucine); Hydroxyl or sulfur/selenium-containing amino acids (e.g., Serine,
Cysteine, Selenocysteine, Threonine, Methionine); Cyclic amino acids (e.g.,
Proline); Aromatic amino acids (e.g., Phenylalanine, Tyrosine, Tryptophan);
Basic amino acids (e.g., Histidine, Lysine, Arginine); or Acidic and their Amide

(e.g., Aspartate, Glutamate, Asparagine, Glutamine). Deletion is the replacement of an amino acid by a direct bond. Positions for deletions include the termini of a polypeptide and linkages between individual protein domains. Insertions are introductions of amino acids into the polypeptide chain, a direct bond formally
5 being replaced by one or more amino acids. Amino acid sequences can be modulated with the help of art-known computer simulation programs that can produce a polypeptide with, for example, improved activity or altered regulation. Based on these artificially generated polypeptide sequences, a corresponding nucleic acid molecule coding for such a modulated polypeptide can be
10 synthesized in-vitro using the specific codon-usage of the desired host cell.

“Highly stringent hybridization conditions” are defined as hybridization at 65 °C in a 6 X SSC buffer (*i.e.*, 0.9 M sodium chloride and 0.09 M sodium citrate). Given these conditions, a determination can be made as to whether a given set of sequences will hybridize by calculating the melting temperature (T_m)
15 of a DNA duplex between the two sequences. If a particular duplex has a melting temperature lower than 65°C in the salt conditions of a 6 X SSC, then the two sequences will not hybridize. On the other hand, if the melting temperature is above 65 °C in the same salt conditions, then the sequences will hybridize. In general, the melting temperature for any hybridized DNA:DNA
20 sequence can be determined using the following formula: $T_m = 81.5 \text{ °C} + 16.6(\log_{10}[\text{Na}^+]) + 0.41(\text{fraction G/C content}) - 0.63(\% \text{ formamide}) - (600/l)$. Furthermore, the T_m of a DNA:DNA hybrid is decreased by 1-1.5°C for every 1% decrease in nucleotide identity (see e.g., Sambrook and Russel, 2006).

Host cells can be transformed using a variety of standard techniques
25 known to the art (see, e.g., Sambrook and Russel (2006) Condensed Protocols from Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, ISBN-10: 0879697717; Ausubel et al. (2002) Short Protocols in Molecular Biology, 5th ed., Current Protocols, ISBN-10: 0471250929; Sambrook and Russel (2001) Molecular Cloning: A Laboratory Manual, 3d ed., Cold Spring
30 Harbor Laboratory Press, ISBN-10: 0879695773; Elhai, J. and Wolk, C. P. 1988. Methods in Enzymology 167, 747-754). Such techniques include, but are not limited to, viral infection, calcium phosphate transfection, liposome-mediated transfection, microprojectile-mediated delivery, receptor-mediated uptake, cell

fusion, electroporation, and the like. The transfected cells can be selected and propagated to provide recombinant host cells that comprise the expression vector stably integrated in the host cell genome.

Conservative Substitutions I	
<u>Side Chain Characteristic</u>	<u>Amino Acid</u>
Aliphatic Non-polar	G A P I L V
Polar-uncharged	C S T M N Q
Polar-charged	D E K R
Aromatic	H F W Y
Other	N Q D E

Conservative Substitutions II	
<u>Side Chain Characteristic</u>	<u>Amino Acid</u>
<u>Non-polar (hydrophobic)</u>	
A. Aliphatic:	A L I V P
B. Aromatic:	F W
C. Sulfur-containing:	M
D. Borderline:	G
<u>Uncharged-polar</u>	
A. Hydroxyl:	S T Y
B. Amides:	N Q
C. Sulfhydryl:	C
D. Borderline:	G
Positively Charged (Basic):	K R H
Negatively Charged (Acidic):	D E

Conservative Substitutions III	
<u>Original Residue</u>	<u>Exemplary Substitution</u>
Ala (A)	Val, Leu, Ile
Arg (R)	Lys, Gln, Asn
Asn (N)	Gln, His, Lys, Arg
Asp (D)	Glu
Cys (C)	Ser
Gln (Q)	Asn
Glu (E)	Asp
His (H)	Asn, Gln, Lys, Arg Leu, Val, Met, Ala,
Ile (I)	Phe,
Leu (L)	Ile, Val, Met, Ala, Phe
Lys (K)	Arg, Gln, Asn
Met(M)	Leu, Phe, Ile
Phe (F)	Leu, Val, Ile, Ala
Pro (P)	Gly

Ser (S)	Thr
Thr (T)	Ser
Trp(W)	Tyr, Phe
Tyr (Y)	Trp, Phe, Tur, Ser
Val (V)	Ile, Leu, Met, Phe, Ala

Exemplary nucleic acids which may be introduced to a host cell include, for example, DNA sequences or genes from another species, or even genes or sequences which originate with or are present in the same species but are

5 incorporated into recipient cells by genetic engineering methods. The term “exogenous” is also intended to refer to genes that are not normally present in the cell being transformed, or perhaps simply not present in the form, structure, *etc.*, as found in the transforming DNA segment or gene, or genes which are normally present and that one desires to express in a manner that differs from

10 the natural expression pattern, *e.g.*, to over-express. Thus, the term “exogenous” gene or DNA is intended to refer to any gene or DNA segment that is introduced into a recipient cell, regardless of whether a similar gene may already be present in such a cell. The type of DNA included in the exogenous

15 DNA can include DNA that is already present in the cell, DNA from another individual of the same type of organism, DNA from a different organism, or a DNA generated externally, such as a DNA sequence containing an antisense message of a gene, or a DNA sequence encoding a synthetic or modified version of a gene.

Host strains developed according to the approaches described herein can

20 be evaluated by a number of means known in the art (*see e.g.*, Studier (2005) *Protein Expr Purif.* 41(1), 207–234; Gellissen, ed. (2005) *Production of Recombinant Proteins: Novel Microbial and Eukaryotic Expression Systems*, Wiley-VCH, ISBN-10: 3527310363; Baneyx (2004) *Protein Expression Technologies*, Taylor & Francis, ISBN-10: 0954523253).

25 Methods of down-regulation or silencing genes are known in the art. For example, expressed protein activity can be down-regulated or eliminated using antisense oligonucleotides, protein aptamers, nucleotide aptamers, and RNA interference (RNAi) (*e.g.*, small interfering RNAs (siRNA), short hairpin RNA (shRNA), and micro RNAs (miRNA) (*see e.g.*, Fanning and Symonds (2006)

30 *Handb Exp Pharmacol.* 173, 289-303G, describing hammerhead ribozymes and

small hairpin RNA; Helene, C., et al. (1992) *Ann. N.Y. Acad. Sci.* 660, 27-36; Maher (1992) *Bioassays* 14(12): 807-15, describing targeting deoxyribonucleotide sequences; Lee et al. (2006) *Curr Opin Chem Biol.* 10, 1-8, describing aptamers; Reynolds et al. (2004) *Nature Biotechnology* 22(3), 326 – 5 330, describing RNAi; Pushparaj and Melendez (2006) *Clinical and Experimental Pharmacology and Physiology* 33(5-6), 504-510, describing RNAi; Dillon et al. (2005) *Annual Review of Physiology* 67, 147-173, describing RNAi; Dykxhoorn and Lieberman (2005) *Annual Review of Medicine* 56, 401-423, describing RNAi). RNAi molecules are commercially available from a variety of sources 10 (e.g., Ambion, TX; Sigma Aldrich, MO; Invitrogen). Several siRNA molecule design programs using a variety of algorithms are known to the art (see e.g., Cenix algorithm, Ambion; BLOCK-iT™ RNAi Designer, Invitrogen; siRNA Whitehead Institute Design Tools, Bioinformatics & Research Computing). Traits influential in defining optimal siRNA sequences include G/C content at the 15 termini of the siRNAs, T_m of specific internal domains of the siRNA, siRNA length, position of the target sequence within the CDS (coding region), and nucleotide content of the 3' overhangs.

Definitions and methods described herein are provided to better define the present disclosure and to guide those of ordinary skill in the art in the 20 practice of the present disclosure. Unless otherwise noted, terms are to be understood according to conventional usage by those of ordinary skill in the relevant art.

In some embodiments, numbers expressing quantities of ingredients, properties such as molecular weight, reaction conditions, and so forth, used to 25 describe and claim certain embodiments of the present disclosure are to be understood as being modified in some instances by the term “about.” In some embodiments, the term “about” is used to indicate that a value includes the standard deviation of the mean for the device or method being employed to determine the value. In some embodiments, the numerical parameters set forth 30 in the written description and attached claims are approximations that can vary depending upon the desired properties sought to be obtained by a particular embodiment. In some embodiments, the numerical parameters should be construed in light of the number of reported significant digits and by applying ordinary rounding techniques. Notwithstanding that the numerical ranges and

parameters setting forth the broad scope of some embodiments of the present disclosure are approximations, the numerical values set forth in the specific examples are reported as precisely as practicable. The numerical values presented in some embodiments of the present disclosure may contain certain
5 errors necessarily resulting from the standard deviation found in their respective testing measurements. The recitation of ranges of values herein is merely intended to serve as a shorthand method of referring individually to each separate value falling within the range. Unless otherwise indicated herein, each individual value is incorporated into the specification as if it were individually
10 recited herein.

In some embodiments, the terms “a” and “an” and “the” and similar references used in the context of describing a particular embodiment (especially in the context of certain of the following claims) can be construed to cover both the singular and the plural, unless specifically noted otherwise. In some
15 embodiments, the term “or” as used herein, including the claims, is used to mean “and/or” unless explicitly indicated to refer to alternatives only or the alternatives are mutually exclusive.

The terms “comprise,” “have” and “include” are open-ended linking verbs. Any forms or tenses of one or more of these verbs, such as “comprises,”
20 “comprising,” “has,” “having,” “includes” and “including,” are also open-ended. For example, any method that “comprises,” “has” or “includes” one or more steps is not limited to possessing only those one or more steps and can also cover other unlisted steps. Similarly, any composition or device that “comprises,” “has” or “includes” one or more features is not limited to possessing only those one or
25 more features and can cover other unlisted features.

All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g. “such as”) provided with respect to certain embodiments herein is intended merely to better
30 illuminate the present disclosure and does not pose a limitation on the scope of the present disclosure otherwise claimed. No language in the specification should be construed as indicating any non-claimed element essential to the practice of the present disclosure.

Groupings of alternative elements or embodiments of the present disclosure disclosed herein are not to be construed as limitations. Each group member can be referred to and claimed individually or in any combination with other members of the group or other elements found herein. One or more members of a group can be included in, or deleted from, a group for reasons of convenience or patentability. When any such inclusion or deletion occurs, the specification is herein deemed to contain the group as modified thus fulfilling the written description of all Markush groups used in the appended claims.

As will be appreciated based upon the foregoing specification, the above-described aspects of the disclosure may be implemented using computer programming or engineering techniques including computer software, firmware, hardware, or any combination or subset thereof. Any such resulting program, having computer-readable code means, may be embodied or provided within one or more computer-readable media, thereby making a computer program product, i.e., an article of manufacture, according to the discussed aspects of the disclosure. The computer-readable media may be, for example, but is not limited to, a fixed (hard) drive, diskette, optical disk, magnetic tape, semiconductor memory such as read-only memory (ROM), and/or any transmitting/receiving media, such as the Internet or other communication network or link. The article of manufacture containing the computer code may be made and/or used by executing the code directly from one medium, by copying the code from one medium to another medium, or by transmitting the code over a network.

These computer programs (also known as programs, software, software applications, "apps", or code) include machine instructions for a programmable processor, and can be implemented in a high-level procedural and/or object-oriented programming language, and/or in assembly/machine language. As used herein, the terms "machine-readable medium" and/or "computer-readable medium" refers to any computer program product, apparatus and/or device (e.g., magnetic discs, optical disks, memory, Programmable Logic Devices (PLDs)) used to provide machine instructions and/or data to a programmable processor, including a machine-readable medium that receives machine instructions as a machine-readable signal. The "machine-readable medium" and "computer-readable medium," however, do not include transitory signals. The term "machine-readable signal" refers to any signal used to provide machine

instructions and/or data to a programmable processor.

As used herein, a processor may include any programmable system including systems using micro-controllers, reduced instruction set circuits (RISC), application specific integrated circuits (ASICs), logic circuits, and any
5 other circuit or processor capable of executing the functions described herein. The above examples are examples only, and are thus not intended to limit in any way the definition and/or meaning of the term "processor."

As used herein, the terms "software" and "firmware" are interchangeable and include any computer program stored in memory for execution by a
10 processor, including RAM memory, ROM memory, EPROM memory, EEPROM memory, and non-volatile RAM (NVRAM) memory. The above memory types are examples only and are thus not limiting as to the types of memory usable for the storage of a computer program.

In one aspect, a computer program is provided, and the program is
15 embodied on a computer-readable medium. In one aspect, the system is executed on a single computer system, without requiring a connection to a server computer. In a further aspect, the system is being run in a Windows® environment (Windows is a registered trademark of Microsoft Corporation, Redmond, Washington). In yet another aspect, the system is run on a mainframe
20 environment and a UNIX® server environment (UNIX is a registered trademark of X/Open Company Limited located in Reading, Berkshire, United Kingdom). The application is flexible and designed to run in various different environments without compromising any major functionality.

In some aspects, the system includes multiple components distributed
25 among a plurality of computing devices. One or more components may be in the form of computer-executable instructions embodied in a computer-readable medium. The systems and processes are not limited to the specific aspects described herein. In addition, components of each system and each process can be practiced independent and separate from other components and processes
30 described herein. Each component and process can also be used in combination with other assembly packages and processes. The present aspects may enhance the functionality and functioning of computers and/or computer systems.

All publications, patents, patent applications, and other references cited in this application are incorporated herein by reference in their entirety for all purposes to the same extent as if each individual publication, patent, patent application, or other reference was specifically and individually indicated to be
5 incorporated by reference in its entirety for all purposes. Citation of a reference herein shall not be construed as an admission that such is prior art to the present disclosure.

Having described the present disclosure in detail, it will be apparent that modifications, variations, and equivalent embodiments are possible without
10 departing the scope of the present disclosure defined in the appended claims. Furthermore, it should be appreciated that all examples in the present disclosure are provided as non-limiting examples.

EXAMPLES

The following non-limiting examples are provided to further illustrate the
15 present disclosure. It should be appreciated by those of skill in the art that the techniques disclosed in the examples that follow represent approaches the inventors have found function well in the practice of the present disclosure, and thus can be considered to constitute examples of modes for its practice. However, those of skill in the art should, in light of the present disclosure,
20 appreciate that many changes can be made in the specific embodiments that are disclosed and still obtain a like or similar result without departing from the spirit and scope of the present disclosure.

EXAMPLE 1: T CELL CHARACTERISTICS ASSOCIATED WITH TOXICITY TO IMMUNE CHECKPOINT BLOCKADE IN PATIENTS WITH MELANOMA

25 The following example describes methods for predicting the likelihood of developing a severe immune-related adverse event (irAE) associated with the administration of an immunotherapy in a melanoma patient.

Abstract:

30 Severe immune-related adverse events (irAEs) occur in up to 60% of patients with melanoma treated with immune checkpoint inhibitors (ICIs). However, it is unknown whether a common baseline immunological state precedes irAE development. Here mass cytometry by time of flight, single-cell RNA sequencing, single-cell V(D)J sequencing, bulk RNA sequencing, and bulk T cell receptor

(TCR) sequencing were applied to study peripheral blood samples from patients with melanoma treated with anti-PD-1 monotherapy or anti-PD-1 and anti-CTLA-4 combination ICIs. By analyzing 93 pre- and early on-ICI blood samples and 3 patient cohorts (n= 27, 26, and 18), it was found that 2 pretreatment factors in circulation—activated CD4 memory T cell abundance and TCR diversity—are associated with severe irAE development regardless of organ system involvement. We also explored on-treatment changes in TCR clonality among patients receiving combination therapy were also explored and the findings were linked to the severity and timing of irAE onset. These results demonstrate circulating T cell characteristics associated with ICI-induced toxicity, with implications for improved diagnostics and clinical management.

Results:

In this study, immunological features in the peripheral blood associated with ICI-induced toxicity in patients with metastatic melanoma were systematically evaluated. Across distinct single-cell and bulk profiling modalities, common T cell features linked to the development of severe irAEs within three months of treatment initiation were identified. These features were independent of key clinical variables, including durable clinical response and treatment with anti-PD-1 monotherapy or anti-PD-1 and anti-CTLA-4 combination therapy. Leveraging these findings, predictive models of irAE development were developed, and explored their utility for pretreatment and early on-treatment identification of ICI-induced toxicity was explored.

Clinical cohort characteristics.

To study candidate risk factors associated with severe (grade 3+) irAE development, 78 patients with metastatic melanoma were identified, 71 of whom were evaluable after exclusion criteria were applied (Fig. 1). Among these patients, 33 were treated with anti-PD-1 monotherapy, 38 were treated with anti-PD-1 plus anti-CTLA-4 combination therapy and 90% had no previous ICI history. All patients were monitored closely during and after ICI treatment for irAE development (median follow-up time of 14.9 months; median time to grade 3+ irAE of 1.5 months). Most patients experienced one or more irAEs, ranging from mild (grade 1) to life-threatening (grade 4) and affecting diverse organ systems, which were classified by board-certified clinicians according to

standardized criteria (CTCAE v.5.0). 71 patients were stratified into 3 nonoverlapping cohorts, a single-cell discovery cohort, and a larger bulk cohort divided into training and validation sets (FIG. 1).

Determinants of severe irAEs from pretreatment blood.

5 Pretreatment peripheral blood was analyzed to identify cellular determinants of severe irAEs using mass cytometry.

10 First, high-dimensional single-cell profiling of pretreatment peripheral blood samples from 18 patients (single-cell discovery cohort, FIGS. 1 and 2A) was performed, of which 8 patients experienced severe irAEs after treatment initiation. By applying mass cytometry by time of flight (CyTOF) to profile 35 leukocyte markers in each sample, 20 distinct subpopulations from nearly 800,000 evaluable cells were analyzed, encompassing 7 major mononuclear lineages (B cells, plasmablasts, CD4 and CD8 T cells, natural killer (NK) cells, natural killer T (NKT) cells, monocytes) (FIGS. 2B, 2C, 7, and 8A). Next, each 15 subpopulation was interrogated with respect to severe irAE outcomes (Fig. 2c). Of all subpopulations, only CD4 effector memory T (TEM) cells were significant after multiple hypothesis corrections, with higher levels in pretreatment blood associated with severe irAE development ($P=0.0002$; $Q=0.004$; FIG. 2C, 2D, 7B, 8B, and 8C). To corroborate this finding, we examined the same peripheral blood 20 samples from 13 patients using 5' droplet-based 10x Chromium single-cell RNA sequencing (scRNA-seq) paired with single-cell V(D)J sequencing (scV(D)J-seq) of TCR and B cell receptor (BCR) clonotypes. After quality control (FIG. 7A), the 5' assay yielded 24,807 cells and 7 major lineages classified on the basis of canonical marker gene expression (FIG. 3A). Employing unsupervised 25 clustering, 32 distinct transcriptional states across the 7 cell types were identified (FIG. 3A). Then the association between cell state abundance and the development of severe irAEs was calculated. Remarkably, across these 32 cell states, it was found that CD4 T cell state 5, which lacks expression of CCR7 and SELL (CD62L) and is consistent with CD4 TEM cells, was most strongly 30 associated with severe irAE development (nominal $P=0.05$, two-sided, unpaired Wilcoxon rank-sum test; FIG. 3B). This state was also most correlated with CD4 TEM levels measured by CyTOF (FIG. 3B). When considering the joint probability of this result via permutation testing, an empirical P value of 0.003

was calculated. Further analysis revealed that CD4 T cell state 3, which is closely related to state 5 by unsupervised hierarchical clustering (FIG. 7B), also showed an expression profile consistent with CD4 TEM (FIG. 3C and 7C). When combined with state 5, the resulting cluster (CD4 T 5+3) was more significantly associated with severe irAE development and CD4 TEM levels enumerated by CyTOF (FIG. 3B). In fact, across all 82 possible pairwise combinations of cell states within each major cell type, CD4 T 5+3 achieved both the highest Spearman correlation against CD4 TEM levels enumerated by CyTOF and the strongest association with severe irAE development (FIG. 7D and 7E).

Differential gene expression analysis against other CD4 T cell states revealed that CD4 T 5 and 3 are enriched for markers of activated effector cells including HLA-DRA, MKI67, TNFRSF4 (OX40), CCL5, and IL32, and depleted in markers of TCM cells (SELL/CD62L) and naive T cells (CCR7, TCF7) (FIG. 3C and 7C). Using Seurat Azimuth for reference-guided cell labeling, it was confirmed that CD4 TEM cells are most associated with severe irAE and most similar to the CD4 T 5+3 population identified by de novo analysis (FIG. 8). Moreover, when the CD4 T 5+3 population was subdivided into activated and resting subsets based on the expression of canonical activation markers (HLA-DX, MKI67), the activated subset showed the strongest association with severe irAE development ($P=0.002$, two-sided, unpaired Wilcoxon rank-sum test; FIG. 3D and 9A). This finding was verified using reference-guided annotation with Azimuth and with CyTOF (FIG. 9A and 9B), suggesting that activated CD4 TEM cells preferentially underlie severe ICI toxicity.

Given this observation, it was wondered whether pretreatment TCR diversity in activated CD4 TEM cells might also correlate with severe ICI toxicity. Indeed, single-cell TCR clonotype diversity (Shannon entropy) of activated CD4 T 5+3 cells was elevated in patients who experienced severe irAEs (area under the receiver operating characteristic curve (AUC)=0.90, $P=0.05$; FIG. 3E). This suggests that TCR richness, defined as the number of unique clonotypes within a sample and a key component of diversity metrics including Shannon entropy, eclipses the loss of diversity resulting from clonal expansion when activated CD4 TEM cells are quantified relative to total peripheral blood mononuclear cells (PBMCs) (FIG. 10A and B). In other words, among total PBMCs, the TCR richness of activated CD4 TEM cells underlies an overall increase in

pretreatment TCR diversity in patients destined to develop severe irAE. Notably, definitions of clonotype diversity that incorporate richness have substantial precedent in previous literature, including studies of circulating and tumor-infiltrating T cells, providing a strong foundation for their application in this work.

5 While this association between TCR diversity and severe irAE development was diminished or absent in other T cell subpopulations when combining all evaluable T cells, a striking trend was observed between bulk TCR diversity in pretreatment samples and severe irAE development (AUC=0.80; FIG. 3E). Moreover, this association was primarily attributable to
10 CD4 T cells with an effector memory profile (low CCR7 and SELL) (FIG. 3F and 10C-F). In contrast, differences in peripheral blood BCR diversity linked to severe irAE development were less pronounced (FIG. 10G). Collectively, these findings suggest that a more diverse TCR repertoire at baseline in CD4 TEM cells, broadly reflected in bulk peripheral blood, is associated with the
15 development of severe ICI toxicity.

Extended analysis of T cell features associated with irAEs.

 Having identified candidate pretreatment determinants of severe irAE development, the findings were verified in a larger independent group of patients. Based on sample size estimates, bulk RNA sequencing (bulk RNA-seq)
20 was applied to pretreatment peripheral blood samples from 53 additional patients with metastatic melanoma spanning two cohorts (n=26 and 27) treated with single-agent (anti-PD-1, n=29) or combination-agent (anti-PD-1 and anti-CTLA-4, n=24) checkpoint blockade (FIG. 1). To assess circulating immunological features in bulk transcriptomic profiles, CIBERSORTx, a machine learning
25 approach for the enumeration of cell subsets from bulk tissue expression profiles, and MiXCR, a computational approach for V(D)J clonotype assembly and quantitation from bulk RNA-seq data, were applied. By direct comparison to cytometry assays, the accuracy of CIBERSORTx for deconvolution of major blood lineages was confirmed, including the specificity of an activated CD4
30 memory T cell (TM) signature for activated CD4 TEM cells using peripheral blood from 17 patients with melanoma (CyTOF) (FIG. 11). Remarkably, of 13 PBMC subsets evaluable by CIBERSORTx, only activated CD4 TM cell levels were associated with severe irAE development (FIG. 4A; P<0.025, HR=8.3 and

14.8 for combination and PD-1, respectively; FIG. 14B and C) or across cohorts by LOOCV (P=0.0028 and HR=12.2 for combination therapy, FIG. 5A; P=0.03 and HR=9.0 for PD-1 therapy). The model also predicted time-to-severe irAE in multivariable models independently of therapy type, age, sex, and other key parameters.

Peripheral TCR clonal expansion linked to severe irAEs.

Previous case reports of patients with melanoma experiencing deadly ICI-mediated toxicity have shown evidence of clonally expanded self- or virus-reactive T cells in the affected tissue, linking self- and pathogen-recognizing T cell clones to lethal toxicity. Accordingly, it was hypothesized that pretreatment TCR clonotypes in peripheral blood might show a greater propensity to expand in patients destined to develop severe irAE after ICI treatment initiation. To examine this, immunoSEQ was applied to profile bulk TCR- β repertoires in paired pretreatment and early on-treatment PBMC samples collected from 15 patients with metastatic melanoma treated with combination therapy. Using a TCR clonality index that is robust to variation in the number of clones captured (Pielou's evenness), significant concordance between MiXCR (bulk RNA-seq) and immunoSEQ (DNA) was confirmed in pretreatment samples from these 15 patients, underscoring the integrity of the composite model in bulk cohorts 1 and 2 (FIG. 15A). TCR clonal expansion (that is, clonal dominance) was then assessed after treatment initiation, as measured by an increase in $1 - \text{Pielou's evenness}$. In support of the hypothesis, both significantly increased TCR clonal expansion and persistence of baseline clones were observed in patients who developed severe irAE compared to those who did not (FIG. 5B, 15B, 15C). In severe irAE patients for whom scRNA-seq and single-cell TCR sequencing (scTCR-seq) (n=3) was performed, preferential expansion of the activated CD4 TEM compartment was observed among clones detected in both blood draws (FIG. 15D-G). Moreover, persistent CD4 T cell clones were highly enriched for the CD4 T 5+3 population identified by scRNA-seq analysis (FIG. 5C and 15E).

Whether the degree of TCR clonal expansion early on-treatment correlated with the timing of severe irAE development was additionally explored. Indeed, whether assessed in tertiles by log-rank test or by rank via Cox proportional hazards regression, patients with a greater magnitude of TCR clonal

expansion developed severe irAE sooner ($P=0.003$, log-rank test; FIG. 5D).

These results were significant independently of the time between blood draws and when restricting the analysis to on-treatment blood draws obtained within one month of cycle 1 ICI (FIG. 15H).

5 *Circulating leukocytes in autoimmune disease.*

Lastly, whether the baseline peripheral blood profile of patients at risk for severe irAE development parallels clinical autoimmunity was asked. To this end, CIBERSORTx was applied to examine 15 leukocyte subsets in bulk peripheral blood transcriptomes spanning 6 studies and 587 patients with either systemic lupus erythematosus (SLE) or inflammatory bowel disease (IBD) relative to 191 healthy controls. Using a meta-analytical framework to integrate P values across studies and pathologies (FIG. 16), it was found that circulating activated CD4 TM cells were most significantly associated with autoimmune disorders relative to healthy individuals (FIG. 6). These data suggest that severe irAEs might represent a subclinical or latent autoimmune state that is clinically unmasked on ICI administration, in line with recent case reports and multi-institutional data showing that patients with autoimmunity treated with immune checkpoint blockade have the propensity to experience flares in their autoimmune symptoms.

20 Discussion

In this study, two baseline features—activated CD4 TM cell abundance and a more clonally diverse TCR repertoire in the peripheral blood— were identified as promising determinants of ICI-induced irAEs in patients with metastatic melanoma. Although previous studies have linked (1) activated T cells and clonally expanded TCRs in postmortem tissue to fatal irAEs (myocarditis, encephalitis) and (2) effector CD4 T cells to organ-specific irAEs (destructive thyroiditis, hepatitis), this work extends the scope of these findings to pretreatment T cell characteristics of irAE development in diverse organ systems. Integration of these features into a composite model predicted greater risk for severe irAEs and demonstrated sufficient granularity to distinguish different irAE grades and burdens.

A striking correlation was also identified between early T cell clonal expansion and the timing of severe irAE onset in patients treated with combination therapy. Future studies are needed to further characterize this

finding and elucidate the relative contributions of CD4 and CD8 T cells to irAE-associated clonal dynamics.

Consistent with the possibility of a common immunological mechanism underlying both irAE development and autoimmunity, elevated levels of activated CD4 TM cells in patients with SLE or IBD were additionally observed. While it is reasonable to predict that patients with previous autoimmunity would be enriched for higher activated CD4 T cell levels and higher rates of severe irAE from ICI, none of the patients in the cohort had documented pre-existing autoimmunity. Moreover, such patients may develop compensatory immune regulatory mechanisms before starting ICI that change their baseline irAE risk. Nevertheless, it is important to study this connection in greater detail in future studies and determine whether circulating activated CD4 TM cells exhibit an increased propensity for recognizing self-antigens in patients at risk for severe ICI toxicity. Indeed, the risk of flare is greater in patients with autoimmune disease treated with combination immunotherapy, particularly those with gastrointestinal or rheumatological conditions. More reliably identifying these at-risk patients during ICI decision-making could improve their outcomes.

This study has several limitations. First, it employed a retrospective design using banked clinical samples. Second, patients received either anti-PD-1 monotherapy or anti-PD-1 plus anti-CTLA-4 combination therapy, which are associated with different risk profiles for severe irAE development. Third, while most irAEs occur within the first three months of ICI treatment initiation, a subset can occur later. Whether the findings generalize to late-onset irAEs will need to be investigated since the median time-to-severe irAE development in our cohorts was 6.4 weeks (consistent with clinical trial data), with no irAEs occurring beyond 3 months. Fourth, the timing of on-treatment peripheral blood collection during immunotherapy with respect to treatment initiation was not homogeneous. Finally, it is yet unclear whether the findings will generalize to ICI-related irAE risk in other cancer types.

Future studies should address these limitations, along with a greater application of single-cell profiling both before and early during immunotherapy. In addition, it will be important to confirm our findings in larger multi-institutional cohorts and assess whether the circulating immunological determinants of ICI-induced toxicity vary based on the organs most likely to be involved. If

prospectively validated, these findings could facilitate treatment adaptation to improve the risk profile of immune checkpoint blockade, with implications for the prediction and potential prevention of ICI-mediated toxicities.

Methods

5 *Study design and participants.*

The samples analyzed in this study were collected with informed consent for research use and were approved by the Yale University School of Medicine and Washington University School of Medicine institutional review boards, in accordance with the Declaration of Helsinki (2013) as part of the observational
10 registry studies focusing on melanoma. Eligible patients aged >18 years with metastatic melanoma were treated with ICI treatment consisting of either anti-PD-1 blockade (nivolumab or pembrolizumab) or combination immune checkpoint blockade (anti-PD-1 (nivolumab) and anti-CTLA-4 (ipilimumab); FIG. 1). Ninety percent of patients were naive to any previous immune checkpoint
15 blockade at the time of pretreatment blood collection. All patients underwent routine clinical assessments for irAEs and responses by board-certified medical oncologists. Surveillance occurred before each cycle of ICI treatment (approximately every 3 weeks), and in several cases, more frequently (for example, by inpatient medical staff in patients admitted to the hospital for severe
20 irAEs). It also continued, when applicable, after completion of the treatment course. All irAEs were classified according to the United States Health and Human Services Common Terminology Criteria for Adverse Events (CTCAE) v.5.0, with grades ≥ 2 and ≥ 3 considered symptomatic and severe, respectively. Within and across patient cohorts, irAEs spanned diverse organ systems
25 including the gastrointestinal tract, skin, liver, pituitary, thyroid, adrenal, musculoskeletal, ocular, pancreatic, and cardiac systems (FIG. 13I). Three patients experienced a systemic inflammatory syndrome related to ICI administration (YUGIM, YUHERN, and YUTORY). All severe irAEs occurred within three months of ICI initiation, a landmark period during which no patients
30 in this cohort died. The response was scored as durable clinical benefit, no durable benefit, or not evaluable as defined previously. Three cohorts of patients were identified who met the aforementioned eligibility criteria and had pretreatment PBMC samples collected just before the first cycle of anti-PD-1 or combination ICI administration (median 0d; range 0–2 months). PBMCs from

each cohort (pretreatment for all patients and pre/ on-treatment pairs for 15 patients) were analyzed as depicted in FIG. 1.

Blood collection and processing.

Peripheral blood specimens were collected in K2EDTA Vacutainer tubes
5 (Becton Dickinson) and processed within 1h of phlebotomy. PBMC extraction was by either an ammonium chloride or Lymphoprep (STEMCELL Technologies) protocol. The Lymphoprep protocol was applied according to the manufacturer's instructions. With the ammonium chloride protocol, 4–8ml of blood was mixed
10 with 20ml of cold ammonium chloride lysing buffer (0.1M of ammonium chloride, 0.01M of Tris-HCl) and incubated for 5min at room temperature. Cells were then centrifuged at 300g for 5min and washed with 5 ml of cold PBS. PBMC samples were cryopreserved in 10% dimethyl sulfoxide/90% FBS. Cryovials were placed in Nalgene Mr. Frosty containers (Thermo Fisher Scientific) for 24h, then stored in liquid nitrogen until cellular and RNA processing for expression analysis.

15 *Mass cytometry.*

Metal-conjugated antibodies were either purchased pre-conjugated from Fluidigm or purchased purified from BioLegend, Thermo Fisher Scientific, or Cell Signaling Technology and subsequently conjugated to metals using Maxpar Antibody Labeling Kits (Fluidigm) according to the manufacturer's instructions..
20 PBMCs from each of the 28 patients were prepared for CyTOF. Cryopreserved cell suspensions were first thawed by holding cryovials in a 37 °C water bath for 1–2min without submerging the cap. Subsequently, 1–3×10⁶ PBMCs in single-cell suspension were incubated with Human TruStain FcX (BioLegend) at room temperature for 10min to block nonspecific antibody binding, followed by
25 incubation with metal-conjugated antibodies against cell surface molecules for 20min on ice. Cells were also incubated with Cell-ID Cisplatin (Fluidigm) according to the manufacturer's instructions to identify viable cells. After treatment with intracellular fixation and permeabilization buffers (Thermo Fisher Scientific), cells were incubated with metal-conjugated antibodies against
30 intracellular proteins. Cells were then washed and stained with Cell-ID Intercalator-Ir (Fluidigm) diluted in PBS containing 1.6% paraformaldehyde (Electron Microscopy Sciences) and stored at 4 °C until acquisition. After a wash step, sample acquisition was then performed using the Helios System (Fluidigm) at an event rate of <400s⁻¹.

To reduce technical variation between samples, Ce beads were used in each sample and the files were normalized together using Bead Normalizer v0.3 (<https://github.com/nolanlab/bead-normalization/wiki/Installing-the-Normalizer>). To further minimize technical variability, the sample processing and acquisition
5 batches were limited to four, the same reagent lots were used across all samples, and no major adjustments were made to Helios calibration. It was also noted that Astrolabe does not compare numerical intensities between samples; rather it analyzes each sample separately, with the assumption that a given subset is the same whether the underlying marker intensities are shifted or not.
10 Thus, the platform has been reported to be resistant to batch effects.

Mass cytometry data analysis.

CyTOF data were initially analyzed with Cytobank v8.0 and v8.1 (Beckman Coulter) using the FlowSOM algorithm for hierarchical cluster optimization and the viSNE algorithm (5,000 iterations, perplexity=100) for
15 visualization of high-dimensional data. Subsequent cell subpopulation identification and data visualization were performed using the Astrolabe Cytometry Platform v3.6 and v4.0 (Astrolabe), which leverages the Ek'Balam algorithm⁷¹, a knowledge-based hierarchical annotation strategy coupled with unsupervised clustering, for automated labeling of cell subpopulations. In total,
20 20 cell subpopulations spanning major mononuclear lineages in peripheral blood were identified and quantified. For each patient sample, cell subpopulation levels were normalized to sum to 1, with unclassifiable cells based on protein marker expression excluded from the analysis. To corroborate Astrolabe, Cytobank was used to perform blinded manual gating of major cell populations including CD4
25 TEM cells (FIG. 17, 18A, and 18B). The total abundance of CD4 TEM cells, whether calculated as a fraction of total PBMCs or circulating T cells, but not as a fraction of CD4 T cells, was significantly associated with severe irAE development (FIG. 18C).

Flow cytometry.

30 PBMCs collected from five healthy donors were analyzed by flow cytometry (FIG. 11E). Briefly, 2–5 million PBMC cells were treated with TruStain FcX Fc Receptor Blocking Solution (BioLegend) for 10min at room temperature to block Fc receptors and then stained with fluorophore-tagged surface antibodies for 30min at room temperature. The following antibodies were used to

stain the cells: FITC-conjugated anti-human CD45 (clone 2D1; BioLegend); AF700-conjugated anti-human CD3 (clone OKT3; BioLegend); APC-conjugated anti-human CD4 (clone OKT4; BioLegend); PE/Cy7-conjugated anti-human CD8 (clone SK1; BioLegend); APC-Cy7-conjugated anti-human CD19 (clone HIB19; BioLegend); PerCp/Cy5.5-conjugated anti-human CD14 (clone HCD14; BioLegend); and BV605-conjugated anti-human CD56 (clone 5.1H11; BioLegend). Cells were then washed twice with ice-cold PBS-based buffer (1 × PBS, 2% FBS, 1mM of EDTA) and stained with 4',6-diamidino-2-phenylindole (DAPI) (BioLegend) to evaluate cell viability. Antibody capture beads (BD Biosciences) were used to compensate each fluorophore in the experiment. Stained cells were analyzed by flow cytometry with operator assistance using a MoFlo Legacy instrument (Beckman Coulter) at the Siteman Flow Cytometry Core at the Washington University School of Medicine. After exclusion of DAPI-positive cells and putative doublets based on forward and side scatter analysis, major lymphocyte populations including B cells, CD4 T cells, CD8 T cells, and NK cells were enumerated as a percentage of total lymphocytes using FlowJo v.10 (FlowJo LLC).

scRNA-seq and scV(D)J-seq library preparation and sequencing.

Single-cell suspensions from PBMC samples were obtained as described above and prepared to a concentration of 700–1,200 viable cells μl^{-1} using a hemocytometer (Thermo Fisher Scientific) or Coulter Counter (Beckman Coulter Life Sciences) for cell counting, according to the manufacturer's instructions. Single-cell suspensions subsequently underwent library preparation for scRNA-seq with paired scV(D) J-seq for TCR and BCR clonotypes using the 5' transcriptome kit (10x Genomics) according to the manufacturer's instructions. Complementary DNA libraries were sequenced on a NovaSeq instrument (Illumina) with 2×92 base pair (bp) paired-end reads targeting a mean of 20,000 reads per cell.

scRNA-seq analysis (discovery cohort).

Raw scRNA-seq reads were barcode-deduplicated and aligned to the hg38 reference genome using Cell Ranger v.3.1.0, yielding sparse digital count matrices, which were analyzed to identify cell types and cellular states using Seurat v.3.1.5 or v.3.2.1 (ref. 72). Outlier cells were identified and removed based on the following criteria: (1) >25% mitochondrial content or (2) cells with

less than 100 or greater than 1,500–3,000 expressed genes, depending on sample-level distributions. After normalization (NormalizeData) and variable feature identification (FindVariableFeatures with $n=2,000$ features), FindIntegrationAnchors (dims=1:30) were applied to identify anchors and IntegrateData (with default parameters) to perform batch correction. Once integrated, principal component analysis (PCA) and uniform manifold approximation and projection (UMAP) were applied using the 2,000 most variable genes and the top 30 principal components. FindClusters was applied to identify cell types and cellular states with a resolution parameter set to 3, yielding 37 clusters.

All identified clusters were assigned to major cell lineages based on the expression of canonical marker genes: CD3D/CD3E^{hi}=T cells; CD8A/CD8B^{hi} and NKG7/GNLY^{lo}=CD8 T cells; non-CD8 T cells with high IL7R expression and low NKG7/GNLY=CD4 T cells; NKG7/GNLY^{hi} and CD3D/CD3E^{lo}=NK cells; CD14 or FCGR3A^{hi}=monocytes; FCER1A^{hi}=dendritic cells (DCs); MS4A1^{hi}=B cells; HBB^{hi}=red blood cells; PPB^{hi}=platelets. Cells with high expression of CD3D/E and GNLY/NKG7 that were not annotated as CD8/CD4 T cells were included in a T or NKT cell group, denoted T/NKT. Clusters annotated as red blood cells or platelets were omitted from further analysis. To assess the effective doublet rate, cellular barcodes were cross-referenced with single-cell BCR (scBCR) and TCR (scTCR) clonotypes. By determining (1) the percentage of non-T cells anomalously mapped to TCR clonotypes (denoted m) and (2) the frequency (that is, recovery rate) of annotated T cells with a matching scTCR clonotype (denoted f), we calculated an effective doublet rate (m/f) of 2.2%. The effective doublet rate calculated for scBCR clonotypes mapping to non-B cells was the same (also 2.2%). Since the effective doublet rate was reasonably low, all single cells with aberrant expression of TCR or BCR clonotypic sequences were eliminated. PCA, UMAP and FindClusters were then repeated as described above, yielding 32 clusters. Two red blood cell clusters, marked by very high HBB expression, remained and were removed from the analysis, followed by one final round of PCA, UMAP, and FindClusters, yielding a final set of 32 clusters (that is, states) and the low-dimensional embedding shown in FIG. 3A and 7A.

All 32 states were assessed for their association with severe irAE development (x-axis of FIG. 3B) and CD4 TEM abundance as measured by

CytoTOF (y-axis of FIG. 3B). Among them, CD4 T cluster 5 was most strongly correlated with both variables (FIG. 3B). To determine the statistical significance of this result, the joint probability of (1) being ranked first by each measure and (2) achieving a P value and Spearman correlation coefficient at least as strong as CD4 T cluster 5 was calculated. To calculate this probability empirically, a permutation scheme was implemented, where cell fractions associated with each scRNA-seq cluster were independently shuffled across all patient samples, then evaluated for (1) and (2) above. By repeating this process 10,000 times, an empirical P value of 0.003 was calculated for CD4 T cluster 5. A pairwise combinatorial analysis was also performed, restricting pairs of cell states to the same major cell type to maintain biological coherence (B cells, CD4 T cells, CD8 T cells, NK cells, monocytes) and compared each of 82 possible cell cluster combinations to CD4 TEM levels enumerated by CyTOF and severe irAE development (FIG. 7D and E). CD4 T cell clusters 5 and 3 emerged as the top-ranking pair. Using the abovementioned statistical approach, an empirical P value of 0.002 was calculated for this result. To identify the differentially expressed genes (DEGs) in FIG. 3C, Seurat FindMarkers were applied with default parameters to the CD4 T 5+3 population versus other CD4 T cell states.

To evaluate the relative utility of unsupervised clustering for delineating cellular determinants of irAE development, a reference-guided annotation framework within Seurat v.4.0.1 (Azimuth) was leveraged to project our scRNA-seq dataset onto a PBMC atlas of 161,764 cells spanning 6 major lineages and 27 finer-grained subsets defined with scRNA-seq and codetection of over 220 protein markers. First, the query dataset was preprocessed following the quality control steps described above, yielding 24,807 cells. The query dataset was then normalized by SCTransform, FindTransferAnchors was applied to the query and reference datasets using a precomputed supervised PCA transformation with 50 dimensions, and then MapQuery was applied to map the cell type labels and UMAP structure from the reference to the query dataset.

Among the 27 cell states identified by Azimuth (FIG. 8A), CD4 TEM was most strongly associated with severe irAE development and most correlated with CD4 TEM cells enumerated by CyTOF (FIG. 8C). Among two other CD4 TEM-like subsets identified by Azimuth (CD4 CTL, CD4 proliferating), CD4 proliferating showed the highest expression of HLA-DX and lowest expression of

SELL (FIG. 8D), which is consistent with an activated CD4 TEM phenotype. Additionally, when examining Azimuth-imputed protein expression from antibody-derived tag data, only CD4 TEM and CD4 proliferating states showed hallmarks of TEM cells (CD45RO^{hi}CD45RA^{lo}CD27^{lo}; FIG. 8E). Indeed, a population combining CD4 TEM and CD4 proliferating was most associated with severe irAE development (FIG. 8C). Hypergeometric testing was applied to assess overlap in cellular barcodes between the combined CD4 TEM +CD4 proliferating population (Azimuth) and states defined by de novo clustering. CD4 T 5+3 emerged as the top hit (Benjamini–Hochberg-adjusted $P=2.5 \times 10^{-7}$). Despite the strong overlap between unsupervised and supervised approaches, CD4 T 5+3 was more associated with severe irAE development and CyTOF than populations labeled by reference-guided annotation (FIG. 8F).

Bulk RNA-seq library preparation, sequencing, and quantification.

Cryopreserved cell suspensions were thawed as described above. RNA was subsequently extracted using the RNeasy PowerLyzer Tissue & Cells Kit (QIAGEN) and quality was assessed with a 2100 Bioanalyzer System (Agilent Technologies). All samples were sufficiently high quality for TruSeq RNA Exome analysis (DV200>30%) and were prepared using the TruSeq RNA Exome Kit (Illumina) according to the manufacturer's instructions. After hybrid capture, cDNA libraries were pooled and sequenced on a HiSeq 2500 instrument (Illumina) using 2×150bp paired-end reads with a target of 20–25 million reads per sample. Raw reads were quantified with Salmon v.0.12.0 using the GENCODE v.29 reference transcriptome; the following command line arguments were used with otherwise default parameters: --seqBias--gcBias--posBias--validateMappings--rangeFactorizationBins 4. Read counts were normalized to gene-level transcripts per million (TPM) using tximport v.1.10.1. Only samples with a mapping rate ≥60% and successful TCR assembly (see the V(D)J receptor profiling and clonotype analysis below) were included for further analysis, with the exception of 3 samples with mapping rates >40% (but <60%) and successful TCR assembly, which were included. In total, 53 sequenced samples (88%) in bulk cohorts 1 and 2 satisfied these criteria (FIG. 1).

Bulk RNA-seq deconvolution.

To determine leukocyte composition in bulk RNA-seq profiles of PBMCs, CIBERSORTx v.1.0.41 (<https://cibersortx.stanford.edu>) was applied with the

LM22 signature matrix to the TPM matrix of each cohort (FIG. 1). CIBERSORTx was separately applied with B-mode batch correction and no quantile normalization to each sequencing batch. LM22, which consists of highly optimized reference profiles for distinguishing 22 functionally defined human hematopoietic subsets, has been widely validated against flow cytometry for accurate enumeration of leukocyte subsets in whole blood and PBMCs, whether profiled by RNA-seq or microarray. CIBERSORTx and the performance of the LM22-activated CD4 TM cell profile were further corroborated in this work through gene expression analysis (CCR5, SELL, TCF7, and CD27; FIG. 11A) and comparison between CIBERSORTx, mass cytometry, flow cytometry and scRNA-seq using PBMC samples from patients with melanoma (FIG. 11B, 11C, 11D, 11E, and 11F). All LM22 subsets except the granulocyte and macrophage subsets were evaluated in this work (n=15; FIG. 4A), with their relative fractions renormalized to sum to 1 for each sample. While a total of 15 subsets were evaluated, 2 were sparsely detected by CIBERSORTx (regulatory T (Treg) cells, gamma delta T cells) and could not be assessed by the Wilcoxon rank-sum test in FIG. 4A.

V(D)J receptor profiling and clonotype analysis.

For the single-cell discovery cohort, raw scV(D)J-seq reads were mapped with Cell Ranger v.3.1.0 to reference refdata-cellranger-vdj-GRCh38-altsembl-4.0.0 and the resulting clonotype assemblies were downloaded from the Loupe V(D)J browser v.3.0.0 (10x Genomics). Given that activated TM cells arise from clonal expansion, the former is expected to have lower TCR diversity than their naive counterparts, provided that (1) cells from both populations are equally sampled (that is, their counts are equivalent) or (2) variation in total T cell counts is normalized out (FIG. 10A). However, by disregarding variation in total T cell frequency, such sampling ignores richness—the number of unique species (clonotypes) within a population and a key factor underlying immune repertoire diversity. As such, Shannon entropy was primarily used to characterize immune repertoire diversity in this work, an information theoretic metric that combines evenness and richness in a single measure (FIG. 10A). For each evaluable patient sample in the single-cell discovery cohort (FIG. 1 and 2A), the TCR clonotype repertoire was randomly sampled (without replacement) to equalize the number of evaluable PBMC cells across patients while addressing technical

variation in TCR recovery. To maximize the pool of TCR clones available for sampling, patients with <100 TCR clones were excluded (n=4; YUTAUR, YUTORY, YUHERN, and YUTHEA). We then calculated Shannon entropy (R package vegan v.2.5-6 (ref. 78)) relative to total PBMCs for each T cell subset and averaged the resulting values across 100 iterations of this procedure for the remaining 9 patients (FIG. 3E, 10B and 10D-F). Shannon entropy was analyzed as described above for scBCR clonotypes across IGK, IGL, and IGH chains in the same nine patients (FIG. 10G).

For bulk cohorts 1 and 2, after adapter sequence trimming using Skewer v.0.2.2, TCR clonotypes were assembled and quantitated with MiXCR v.3.0.125 using the following command: `mixcr align -p rna-seq -s hsa -O allowPartialAlignments=true data_R1.fastq.gz data_R2.fastq.gz alignments.vdjca`. For each patient sample, TCR clonotype diversity was measured in aggregate for TCR- α and TCR- β chains using Shannon entropy (R package vegan v.2.5-6) and compared between patients based on irAE severity (FIG. 4B, 4C, 12A, and 12C). The Gini–Simpson index was additionally applied, which was calculated using the R package immunarch v.0.6.5 (<https://doi.org/10.5281/zenodo.3367200>), to evaluate bulk TCR diversity according to irAE severity (FIG. 12B and 12D). Of note, TCR richness is a key component for calculating both Shannon entropy and the Gini–Simpson index.

Analysis of T cell clonal dynamics from bulk PBMCs.

Bulk TCR- β chain profiling was performed on paired pretreatment and early on-treatment PBMCs from 15 patients treated with combination ICIs. No patients had on-treatment peripheral blood collected after the onset of severe irAE. Genomic DNA was extracted using the DNeasy Blood & Tissue Kit (QIAGEN) and submitted for survey-resolution immunoSEQ (Adaptive Biotechnologies). Data from productive TCR- β chain rearrangements were exported using the immunoSEQ Analyzer online tool and evaluated for TCR- β repertoire richness and diversity using Pielou's evenness, with increased 1 – evenness associated with increased clonality. The Pielou's evenness results from immunoSEQ profiling were compared with bulk RNA-seq (MiXCR), which revealed concordance (FIG. 15A). We also verified that all pretreatment and on-treatment samples were properly paired by cross-comparison of TCR- β CDR3 sequences. Clonal expansion was inferred by analyzing the difference in

clonality, defined as $1 - \text{Pielou's evenness}$ in each sample, between paired on- and pretreatment time points (FIG. 5B and 15B). More specifically, to calculate the change in clonality from baseline, pretreatment clonality was subtracted from on-treatment clonality in a paired fashion, thereby normalizing all pretreatment samples to zero (FIG. 5B, left). The data from FIG. 5B were also analyzed without normalizing on-treatment samples to paired pretreatment samples in FIG. 15B.

To assess freedom from severe irAE, the degree of clonal expansion, denoted δ , was evenly divided into tertiles using the R package `dplyr` v.1.0.7 (FIG. 5D and 15H). This yielded the following groups: no clonal expansion, $\delta < 0$, $n=5$; intermediate, $0 < \delta < 0.009$, $n=5$; and high clonal expansion, $\delta > 0.009$, $n=5$. These thresholds were applied to the full immunoSEQ cohort ($n=15$; FIG. 5D) and to patients with blood samples obtained on ICI treatment day 1 and < 1 month later ($n=7$; FIG. 15H). Additionally, when represented in rank space, the degree of clonal expansion was significantly associated with time-to-severe irAE development in Cox regression models and was independent of the time between blood draws, the number of productive TCR clones detected, and the age and sex of each patient.

Analysis of persistent T cell clones.

Paired pretreatment peripheral blood scRNA-seq and scTCR-seq were performed for three patients (FIG. 5B) who experienced severe irAEs with variable levels of clonal expansion: YUALOE, YUNANCY, and YUHONEY (FIG. 5C and 15D, 15E, 15F, and 15G). Of note, samples from these three patients were not previously profiled by scRNA-seq or scV(D)J-seq in the single-cell discovery cohort. Sequencing libraries were generated and processed for quality control identically to those described in the single-cell discovery cohort. Mapping was performed with Cell Ranger v.5.0.1.

To analyze persistent clones—which were defined as productive TCR- β CDR3 nucleotide sequences shared between paired pretreatment and on-treatment blood samples—the immunoSEQ data was interrogated for shared clonotypes with at least 2 templates in 1 blood draw (pretreatment or on-treatment) and at least 1 template in the other blood draw (60% of all shared clones, on average). This allowed one to preferentially focus on persistent clones that either expanded or contracted. The resulting sequences were cross-

referenced with the TCR- β CDR3 nucleotide sequences from the pretreatment scTCR-seq libraries, which were further cross-referenced with scRNA-seq data and filtered for cells annotated as T cells by Azimuth (applied as described above) (FIG. 15D). In total, 1,504 single-cell transcriptomes with paired
5 immunoSEQ clonotype data were identified. Significant Spearman correlations between pretreatment single-cell and immunoSEQ TCR clonotype frequencies were observed for each patient ($p > 0.59$; P_0 and $CD8A/B = 0$ for CD4 T cells; $CD8A$ or $CD8B > 0$ and $CD4 = 0$ for CD8 T cells). In all, 69% of all cross-referenced clonotypes could be unambiguously labeled by this approach (FIG.
10 15E). For the plot shown in FIG. 5C, the mean \log_2 fold change between CD4 T 5 and 3 was calculated versus the remaining CD4 T cell clusters in the single-cell discovery cohort (FIG. 7B) and then the top 20 genes were selected for subsequent analysis. Enrichment of this gene set was determined using single-sample GSEA (R package *escape* v.1.0.1), which was applied to T cells labeled
15 by Azimuth or labeled as described above for persistent CD4/CD8 T cells. For the analysis shown in FIG. 15F and G, productive frequencies of persistent T cell clones measured by immunoSEQ were grouped into CD4 and CD8 T cells, with differences in productive frequencies displayed on a per-clonotype basis (FIG.
20 15G) or in aggregate (FIG. 15F) and compared to bulk clonal expansion from baseline (FIG. 5B).

Integrative models to predict irAE development.

Activated CD4 TM cell abundance and bulk TCR clonotype diversity were individually associated with severe irAE development (FIG. 4A and B). Accordingly, integrative modeling was explored as a means of improving
25 performance. While several techniques were assessed, including nonlinear modeling with random forests, logistic regression (*glm* in R) achieved comparable performance and was selected owing to the relative simplicity and robustness of a generalized linear model. Before training, each feature was tested in bulk cohorts 1 and 2 for outliers using the ROUT test with a false
30 discovery rate=10%. Of 88 data points (2 features \times 53 samples), 3 outliers were detected, all from activated CD4 TM cells in bulk cohort 1. Regardless of the training cohort, all detected outliers were invariably from among these three samples. Therefore, for each integrative model, the maximum fraction $\max F$ of activated CD4 TM cell levels was determined from among all non-outlier

samples in the training cohort. maxF was then used as a ceiling for all samples.

The composite model was trained to predict severe irAE (grade 3+) development in several ways. These include training on bulk cohort 1 and testing on held-out bulk cohort 2 (FIG. 4D, left); training on one therapy type and testing on another (FIG. 4D, right); and training across bulk cohorts using LOOCV. For all models assessed by LOOCV, the analysis was repeated n times, where n is the total number of patients. In each iteration, the model was trained on each patient except the i^{th} patient and evaluated on the held-out i^{th} patient. To mitigate overfitting when dividing patients into high and low groups by LOOCV, we applied Youden's J statistic was applied to determine the threshold that optimized sensitivity and specificity in each training cohort, then allocated the held-out i^{th} patient on the basis of this threshold.

Composite model scores were assessed by receiver operating characteristic (ROC) analysis. Models trained to discriminate severe from non-severe irAEs were used to predict the future development of severe irAE (FIG. 4D and 13A), irAE grade (FIG. 4C, 4E, 13B, and 13E), the number of irAE-impacted organ systems (FIG. 13H-J) and the time-to-severe irAE development (FIG. 5A and FIG. 14). They were also assessed in different patient subgroups (FIG. 4D and 13D) and compared to pathways and previously published biomarkers evaluated in bulk RNA-seq (FIG. 13C). Composite models were additionally validated at different irAE grade thresholds (FIG. 13F) and tested separately by therapy type to predict irAE development (FIG. 4D, 5A, 13A, 13D-F, 8B, 8C).

Assessment of circulating leukocyte composition in autoimmune disorders.

Peripheral blood gene expression datasets profiled by bulk RNA-seq or microarrays and spanning 239 patients with SLE, 348 patients with IBD, and 191 paired healthy controls, were downloaded from the Gene Expression Omnibus (GEO). RNA-seq data from Hung et al. were downloaded as a preprocessed expression matrix and TPM-normalized before analysis. Affymetrix microarray datasets ($n=5$) were downloaded as CEL files, MAS5-normalized (affy v.3.12 (ref. 82) in R), mapped to Entrez gene identifiers using a custom chip definition file specific to each platform (<http://brainarray.mbni.med.umich.edu/Brainarray/Database/CustomCDF/>) and converted to HUGO gene symbols. One dataset did not have raw CEL files available; instead, preprocessed expression

data were obtained from GEO. In cases for which multiple probe sets mapped to the same gene symbol, we selected the probe set with the highest mean log₂ expression across samples for further analysis. In the Palmer et al. dataset, some samples identified as controls were from subjects with Escherichia coli infection, celiac disease, or progression to Crohn's disease; these were excluded from the analysis. For replicate samples in the Carpintero et al. dataset, the most recent sample was selected. For the Peters et al. dataset, only pretreatment blood samples from patients with Crohn's disease (week 0) were further analyzed. CIBERSORTx49 was applied with LM22 to the Hung et al. bulk RNA-seq dataset as described above, while microarray datasets were either run with (1) quantile normalization and B-mode batch correction (non-HG-U133 platforms) or (2) quantile normalization and no batch correction (HG-U133 platforms). Leukocyte subsets were limited to mononuclear subsets found in peripheral blood (granulocytes and macrophages were omitted) and were renormalized to sum to one for each sample.

Within each dataset, a two-sided, unpaired Wilcoxon rank-sum test was applied to evaluate the levels of each leukocyte subset in peripheral blood between individuals with the disease and healthy controls from the same study (FIG. 16). The resulting P values were converted into two-sided z-scores while taking the directionality of the association into account. Within a given disease phenotype (SLE or IBD), z-scores were combined across datasets using Lipták's method weighted by sample size (FIG. 16). Lastly, SLE- and IBD-specific meta z-scores were combined via the Stouffer's method (FIG. 16), yielding a pan SLE/IBD meta z-score for each leukocyte subset (FIG. 6).

Candidate toxicity biomarkers from previous literature and pathway analysis.

The composite model was benchmarked against previously published irAE biomarkers and enriched pathways for severe irAE prediction (FIG. 13C). Each candidate biomarker was assessed separately in bulk cohorts 1 and 2 by determining the AUC by ROC analysis. The following pretreatment irAE biomarkers, which were measured by protein expression in previous literature, were assessed by RNA surrogates in the peripheral blood in this study: ADPGK and LCP1, which we evaluated individually and with bivariable linear regression; CD74 and GNAL15 expression; and the CYTOX score, which were evaluated as the geometric mean expression of genes encoding the same 11 cytokines

(CSF3, CSF2, CX3CL1, FGF2, IFNA2, IL12A, IL1A, IL1B, IL1RA, IL2, IL13). Separately, pre-ranked GSEA v.4.1.0 via GSEAPreranked v.7.1.0 was applied to identify the most irAE-enriched pathways in bulk cohorts 1 and 2 from the Molecular Signatures Database v7.4 hallmark pathway collection. As input, transcriptome-wide gene lists were defined for bulk cohorts 1 and 2 that were rank-ordered by log₂ fold change between patients who developed severe irAE and those who did not. Gene sets with $q < 0.25$ were considered statistically significant. The two most-enriched gene sets in patients with severe irAEs versus patients with no severe irAEs (MYC_TARGETS_V1; OXIDATIVE_PHOSPHORYLATION) were compared to the composite model in bulk cohorts 1 and 2 (FIG. 13C).

Statistics.

All statistical tests were two-sided unless stated otherwise. The Wilcoxon rank-sum test was used to assess statistical differences between the two groups. When assessing >2 groups simultaneously, the nonparametric Kruskal–Wallis test was used. The Benjamini–Hochberg method was applied for multiple hypothesis testing unless stated otherwise. A permutation scheme was implemented to assess scRNA-seq cluster correlation with severe irAE development and CyTOF CD4 TEM abundance as described above. A Fisher’s exact test was applied to assess statistical differences between two categorical variables. ROC analysis was performed to assess classification accuracy, which was quantified by AUC. The statistical significance of the AUC was determined by a two-sided z-test. Youden’s J statistic was used to identify the optimal cut-point after ROC analysis. Linear concordance was determined by Pearson (r) or Spearman (ρ) correlation and a two-sided t-test was used to assess whether the result was significantly nonzero. Kaplan–Meier and Cox regression analyses were used to assess covariates with respect to time-to-severe irAE. Significance levels and HRs for Kaplan–Meier analyses were determined using a two-sided log-rank test. The composite models and related analyses in FIG. 5A and 14 include patients from bulk cohorts 1 and 2 (FIG. 1) with the exception of two patients (YUDIME and YUMEDIC) who did not develop severe irAEs but experienced early disease progression leading to therapy switch before three months had elapsed. These two patients were included in other analyses since they each received 63d (2.1 months) of immune checkpoint blockade, a time

period within which 76% of all severe irAEs occurred in the patient population.

For Cox regressions, the results were analyzed based on the Wald statistic (z-score) and significance was assessed by the Wald test. The proportional hazards assumption was confirmed for each covariate included in a
5 Cox regression before analysis by evaluating the Schoenfeld residuals. Lipták's and Stouffer's methods were used for integrative statistical analyses, as appropriate. Sample size calculations for bulk cohorts 1 and 2 were performed using pwr v.1.3-0 in R86. In the single-cell discovery cohort, the association
10 between CD4 TEM cell abundance (CyTOF) and severe irAE development had an effect size of 1.99 (FIG. 2C and D). Bulk cohorts 1 and 2 were designed to satisfy this effect size requirement at $\alpha=0.05$ and $1-\beta=0.8$ while emphasizing specificity in bulk cohort 1 (number of patients without severe irAEs > number of patients with severe irAEs) and balance in bulk cohort 2 (number of patients without severe irAEs \approx number of patients with severe irAEs). All statistical
15 analyses were performed using R v.3.5.1+ or Prism 8+ (GraphPad Software).

CLAIMS

What is claimed is:

1. A method for predicting a likelihood of developing a severe immune-related adverse event (irAE) in a patient receiving an immunotherapy, the method comprising:
 - a. obtaining a peripheral blood sample from a subject prior to receiving an immunotherapy treatment;
 - b. quantifying an abundance of activated CD4 memory T cells and a diversity of T cell receptors (TCR) within the peripheral blood sample; and
 - c. classifying the patient as likely to develop a severe irAR if the abundance of activated CD4 memory T cells in combination with amounts of TCR exceeds a threshold value.
2. The method of claim 1, further comprising determining the threshold value by reference to known clinical standards.
3. The method of claim 1, wherein the abundance of activated CD4 memory T cells and the diversity of T cell receptors (TCR) are determined using at least one of: bulk RNA-sequencing (CIBERSORTx and MiXCR), mass cytometry by time of flight (CyTOF), immunoSEQ® TCR-β profiling, droplet-based scRNA-sequencing and scTCR-sequencing, and targeted RNA-sequencing using an RNA panel targeted to activated CD4 memory T cells.
4. A method for predicting a likelihood of developing a severe immune-related adverse event (irAE) in a patient receiving an immunotherapy, the method comprising:
 - a. obtaining a first peripheral blood sample from a subject prior to receiving an immunotherapy treatment and a second peripheral blood sample subsequent to the administration of the immunotherapy to the patient;

- b. quantifying a first TCR diversity level from the first peripheral blood sample and a second TCR diversity level from the second peripheral blood sample;
 - c. obtaining a degree of TCR expansion by subtracting the first TCR diversity level from the second TCR diversity level;
 - d. classifying the patient as likely to develop severe irAR if the degree of TCR expansion exceeds a threshold value.
5. The method of claim 4, further comprising predicting a time of onset of the severe irAR based on the degree of TCR expansion, wherein a higher degree of TCR expansion is predictive of an earlier onset of severe irAR.
6. The method of claim 4, wherein the first and second TCR diversities are determined using at least one of: bulk RNA-sequencing (CIBERSORTx and MiXCR), mass cytometry by time of flight (CyTOF), immunoSEQ® TCR- β profiling, droplet-based scRNA-sequencing and scTCR-sequencing, and targeted RNA-sequencing using an RNA panel targeted to activated CD4 memory T cells.

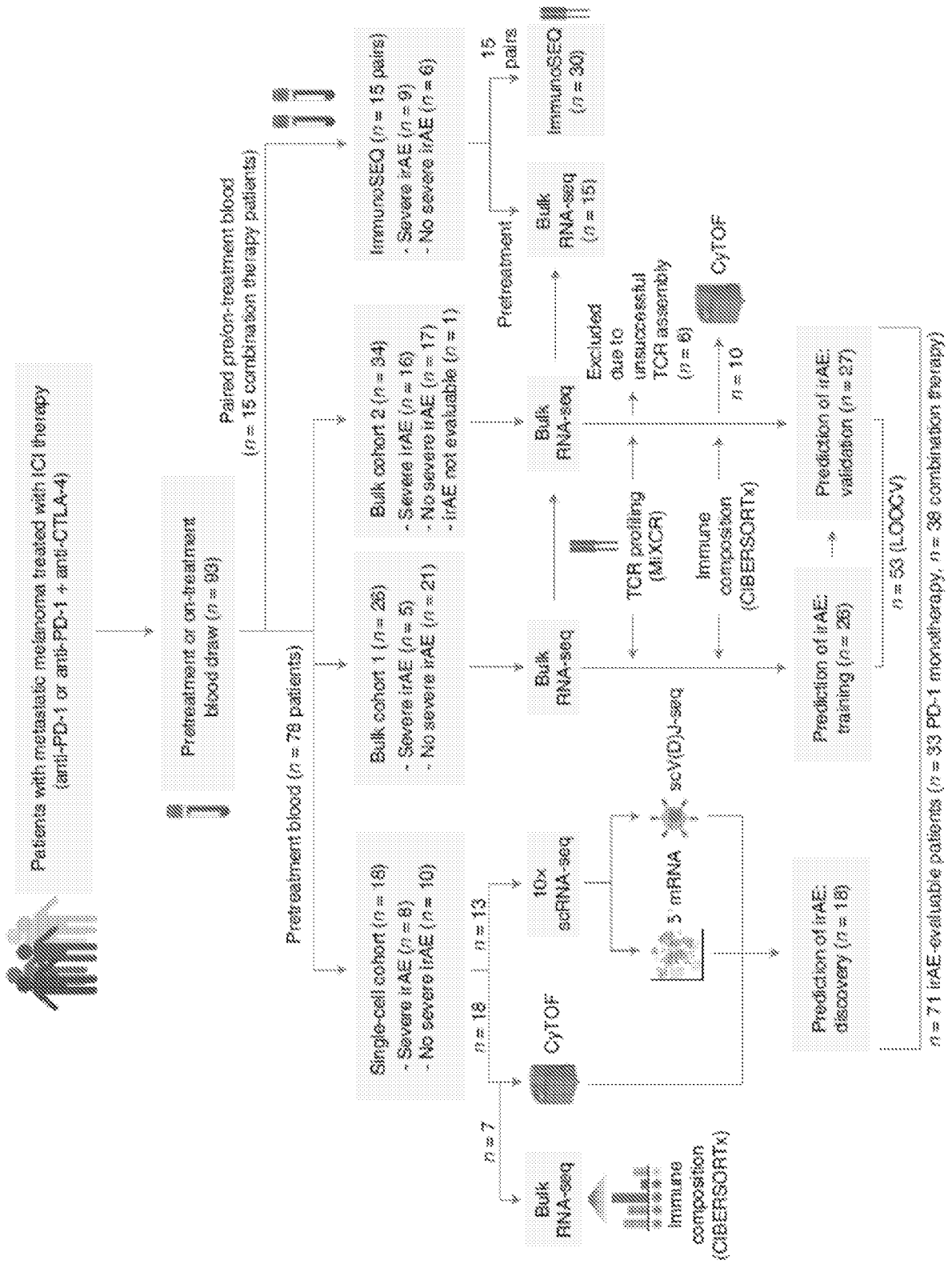


FIG. 1

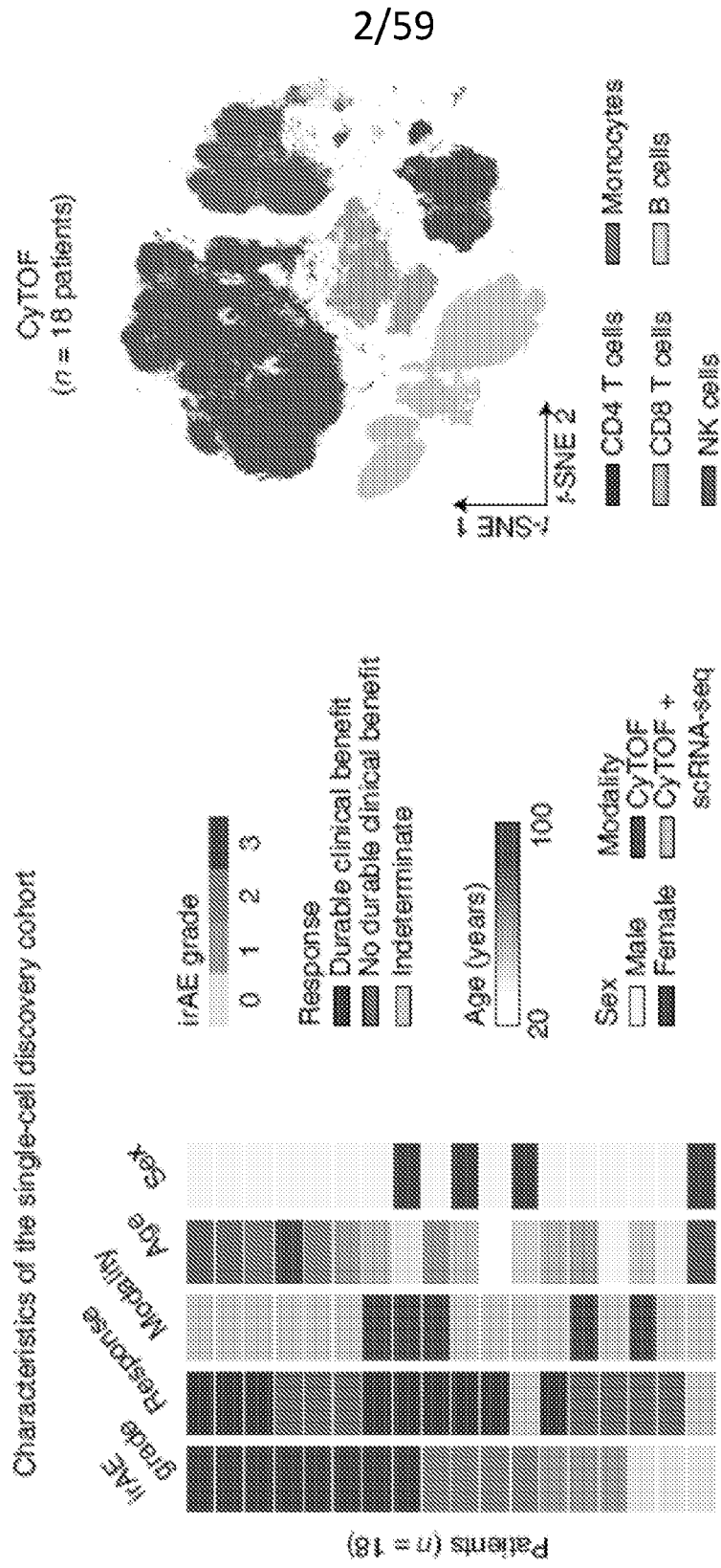


FIG. 2B

FIG. 2A

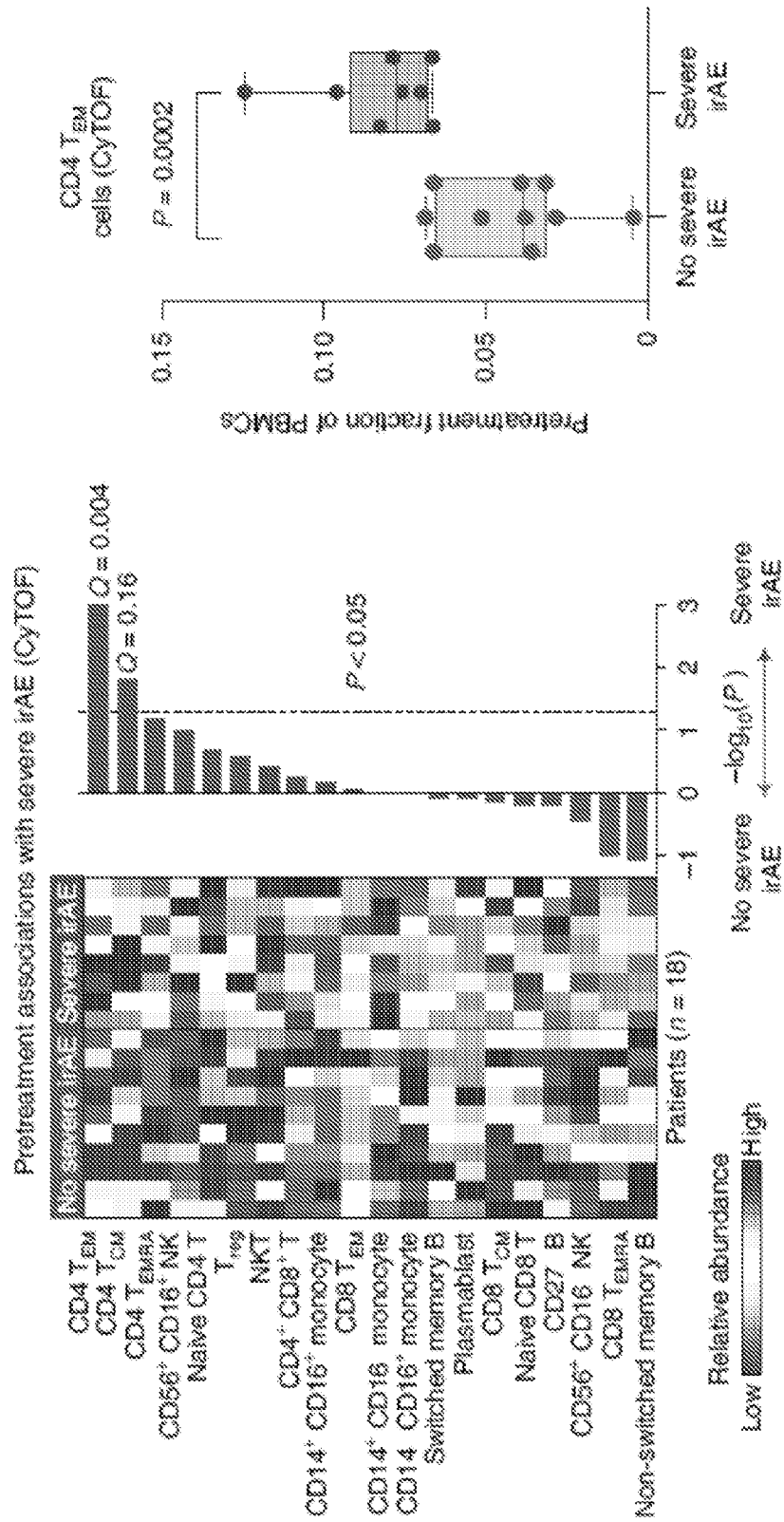


FIG. 2C

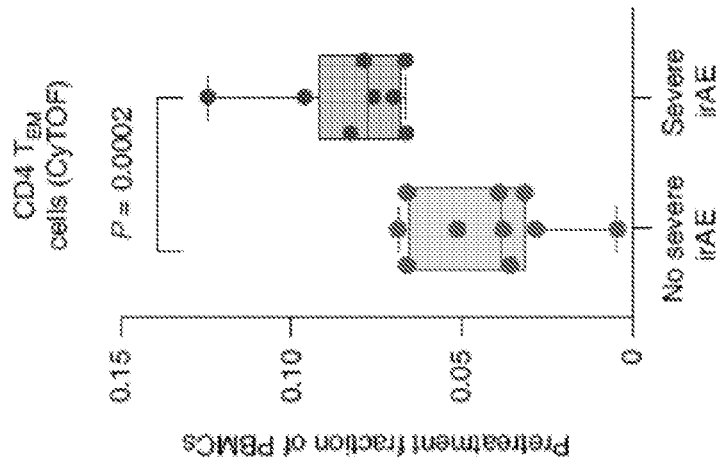


FIG. 2D

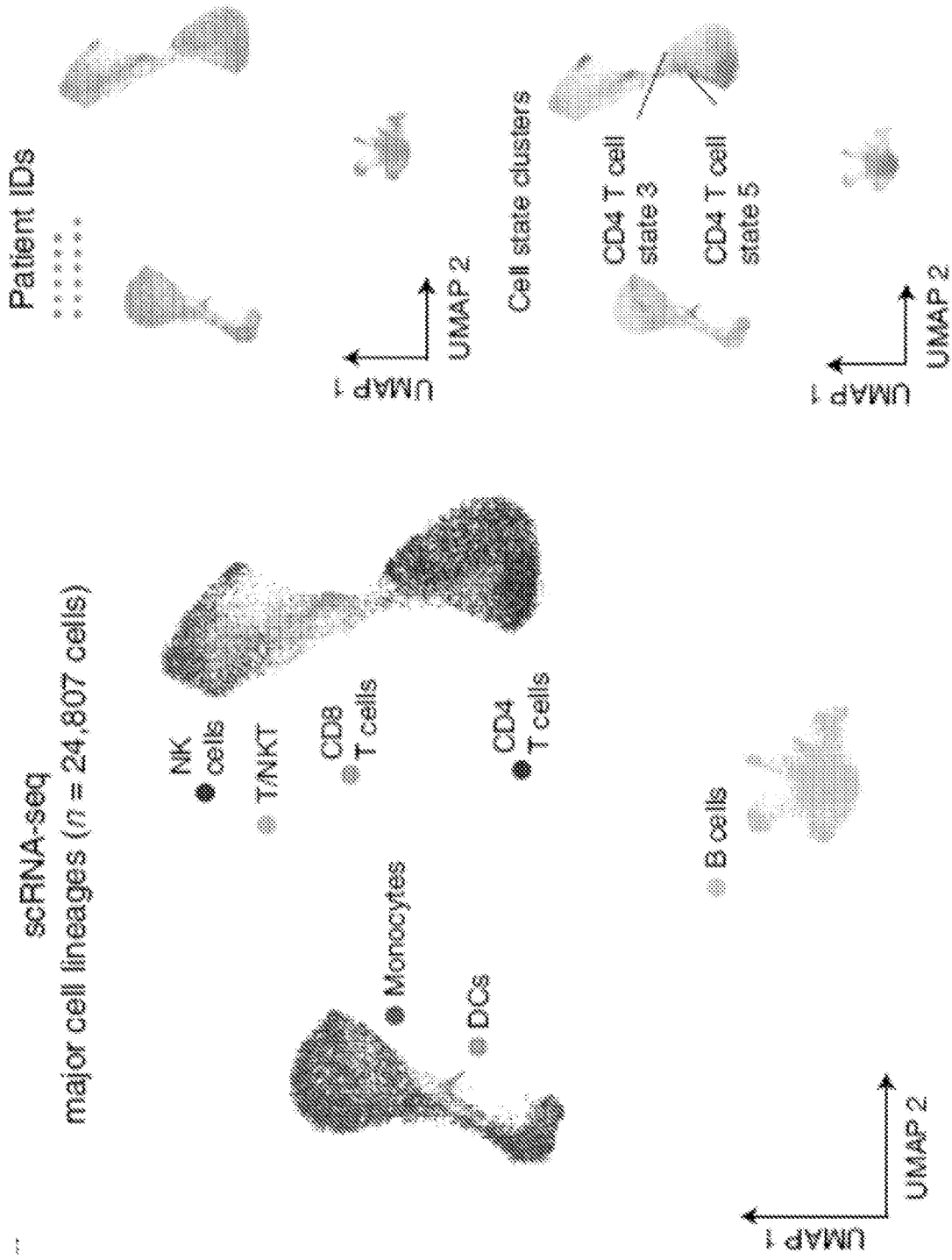


FIG. 3A

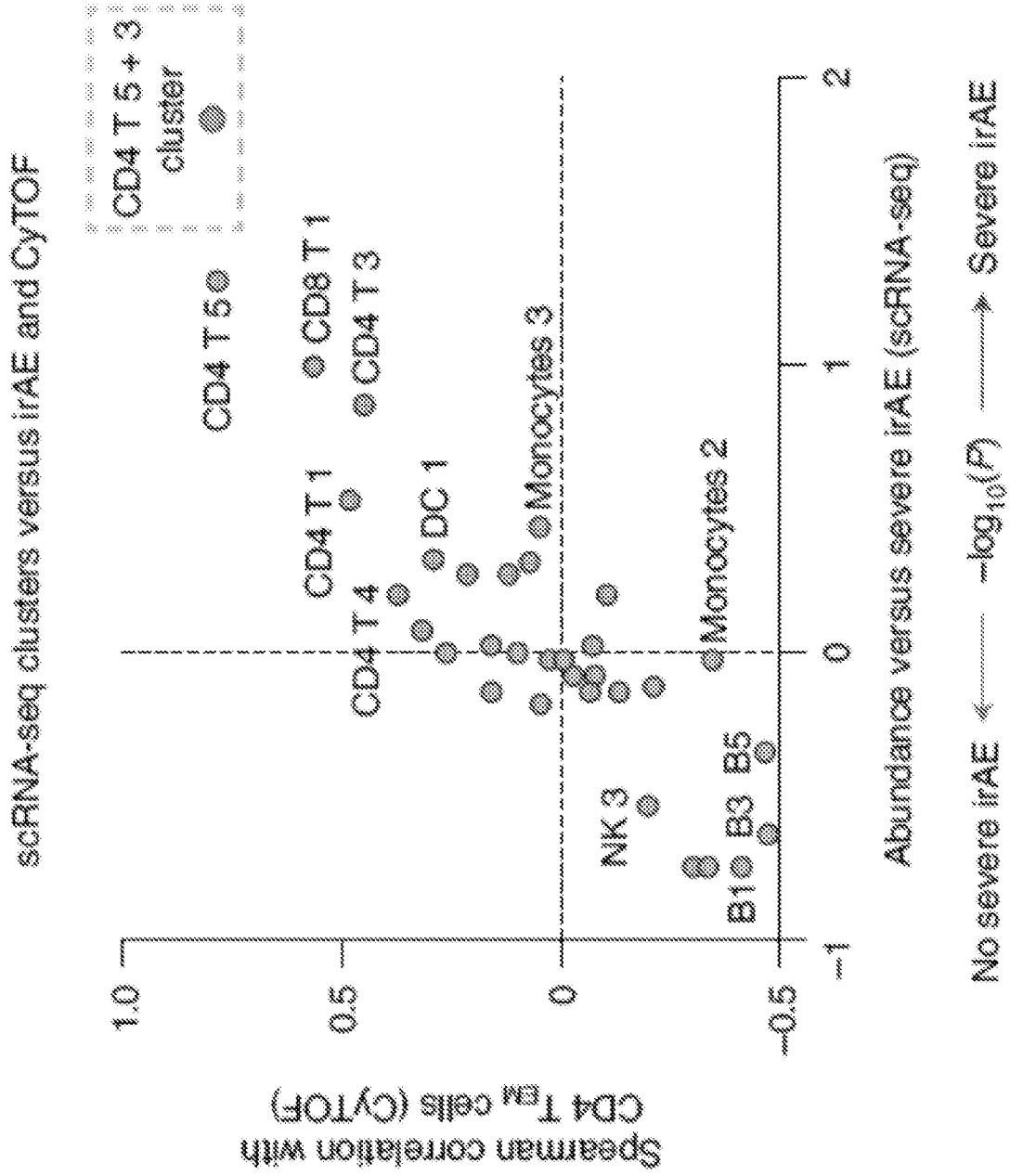


FIG. 3B

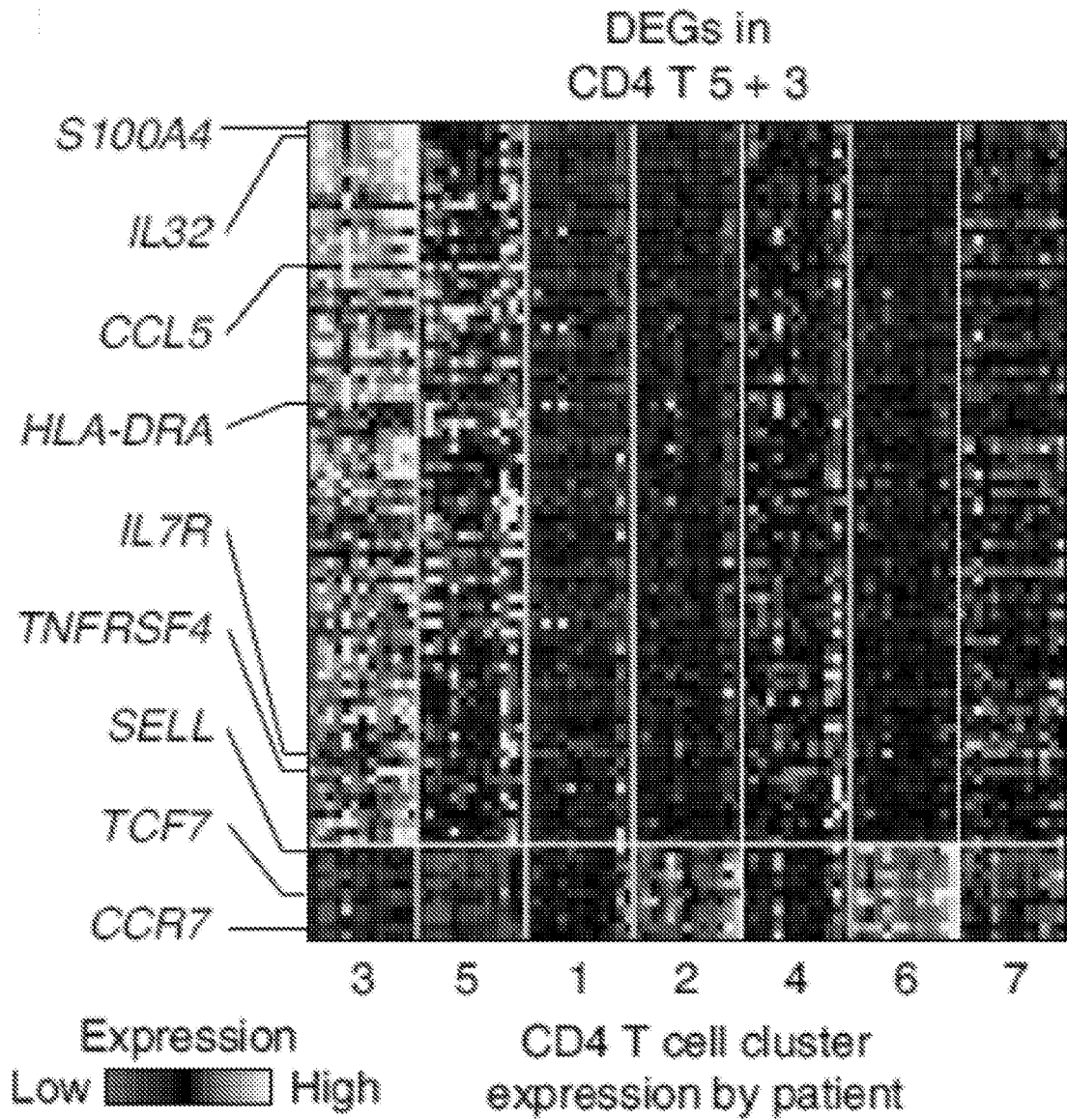


FIG. 3C

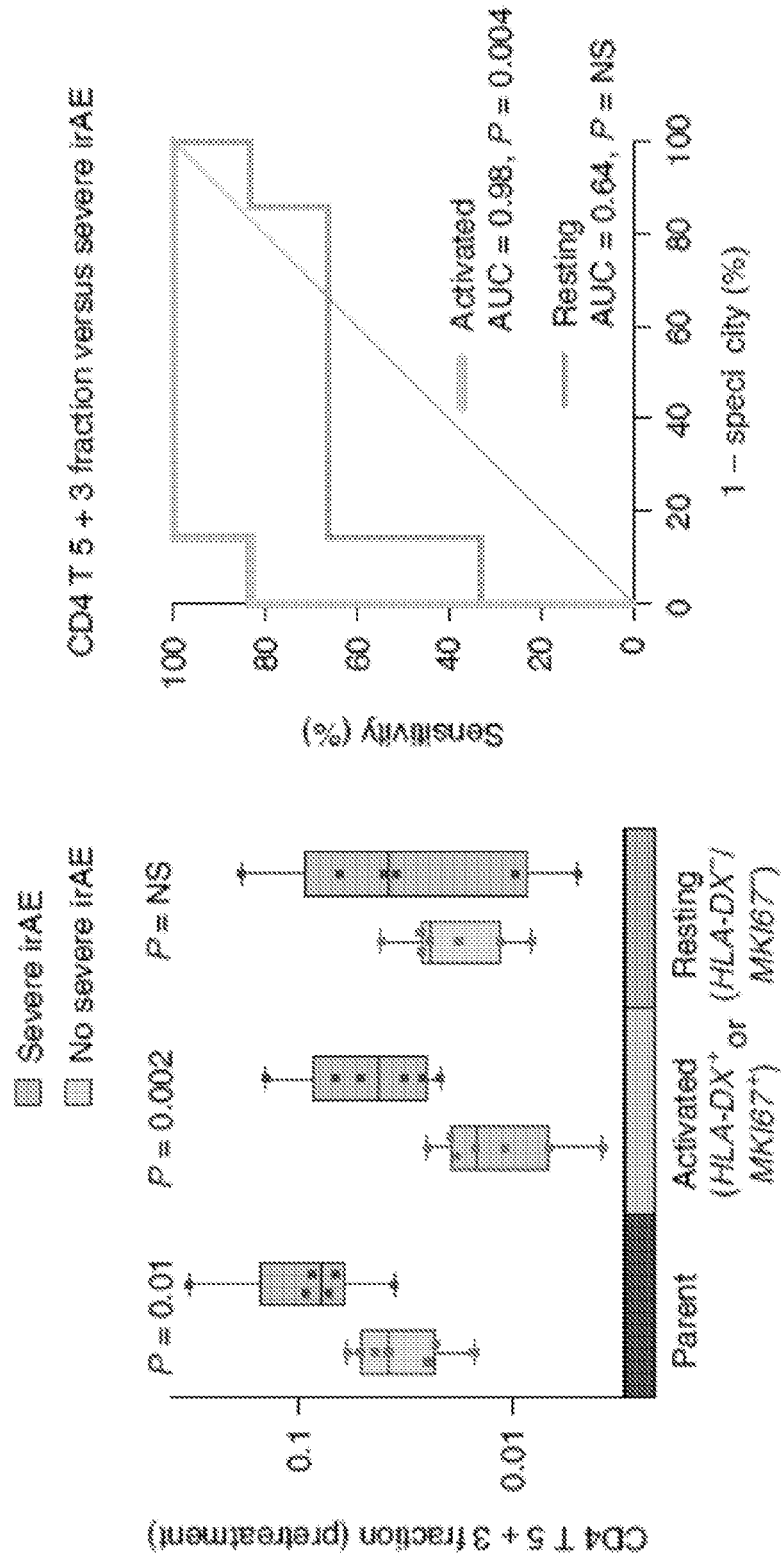


FIG. 3D

Expression of lineage and activation markers in each CD4 T cell cluster

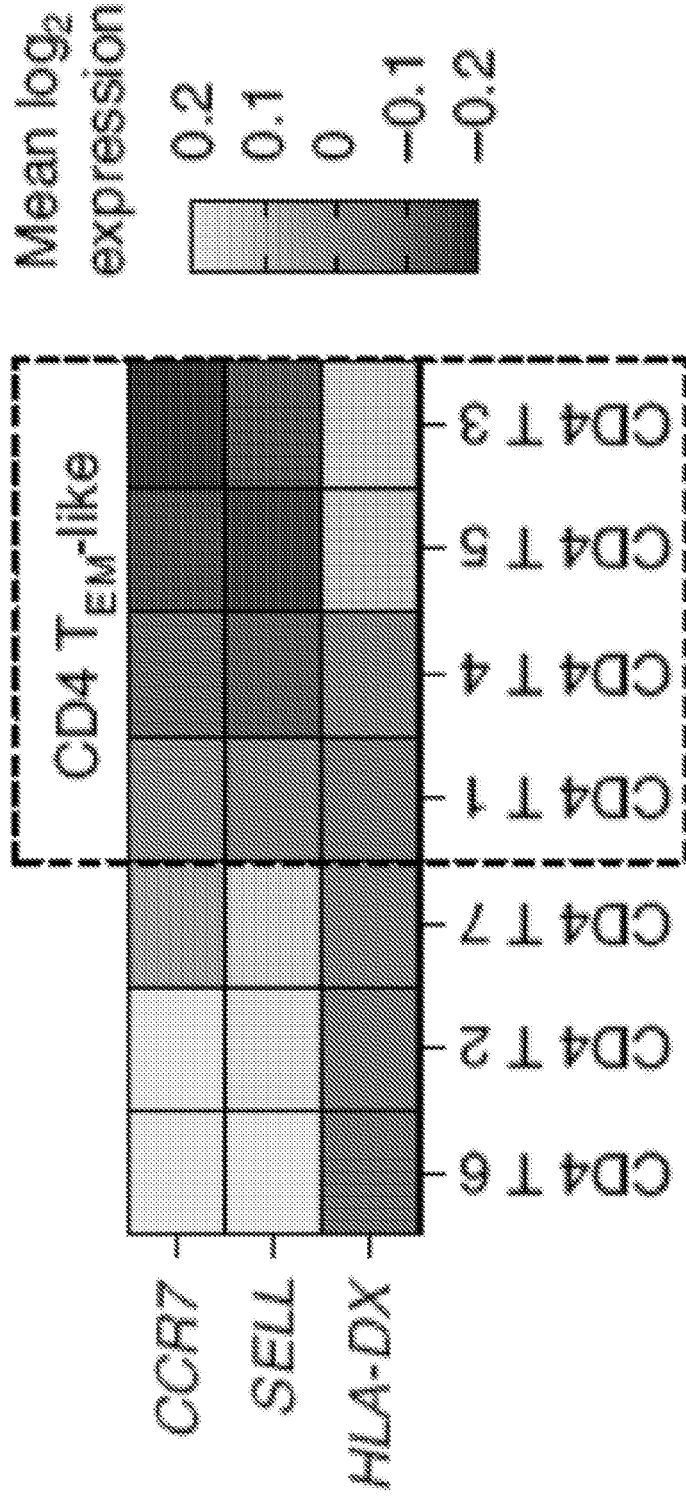


FIG. 3F

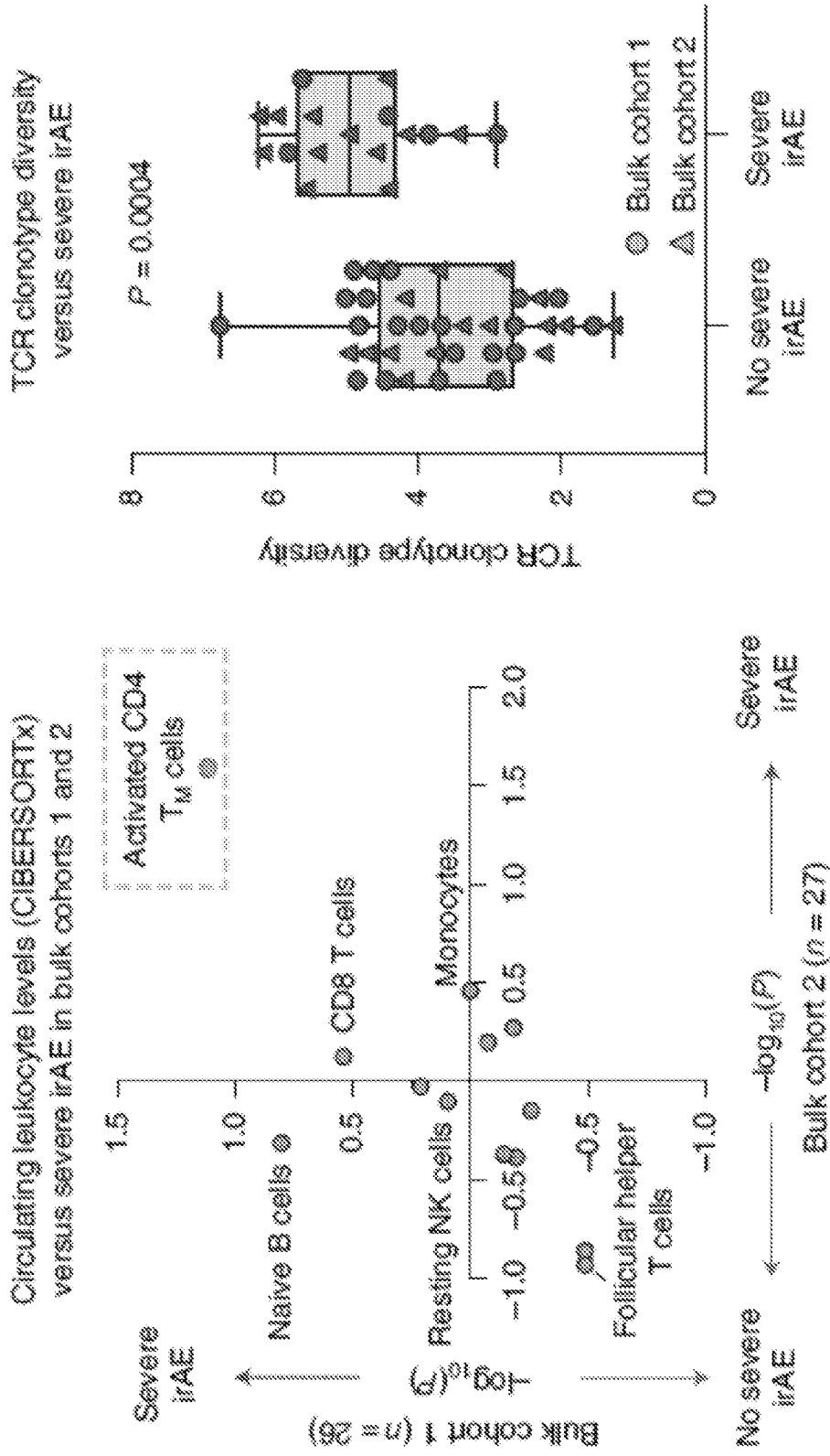


FIG. 4B

FIG. 4A

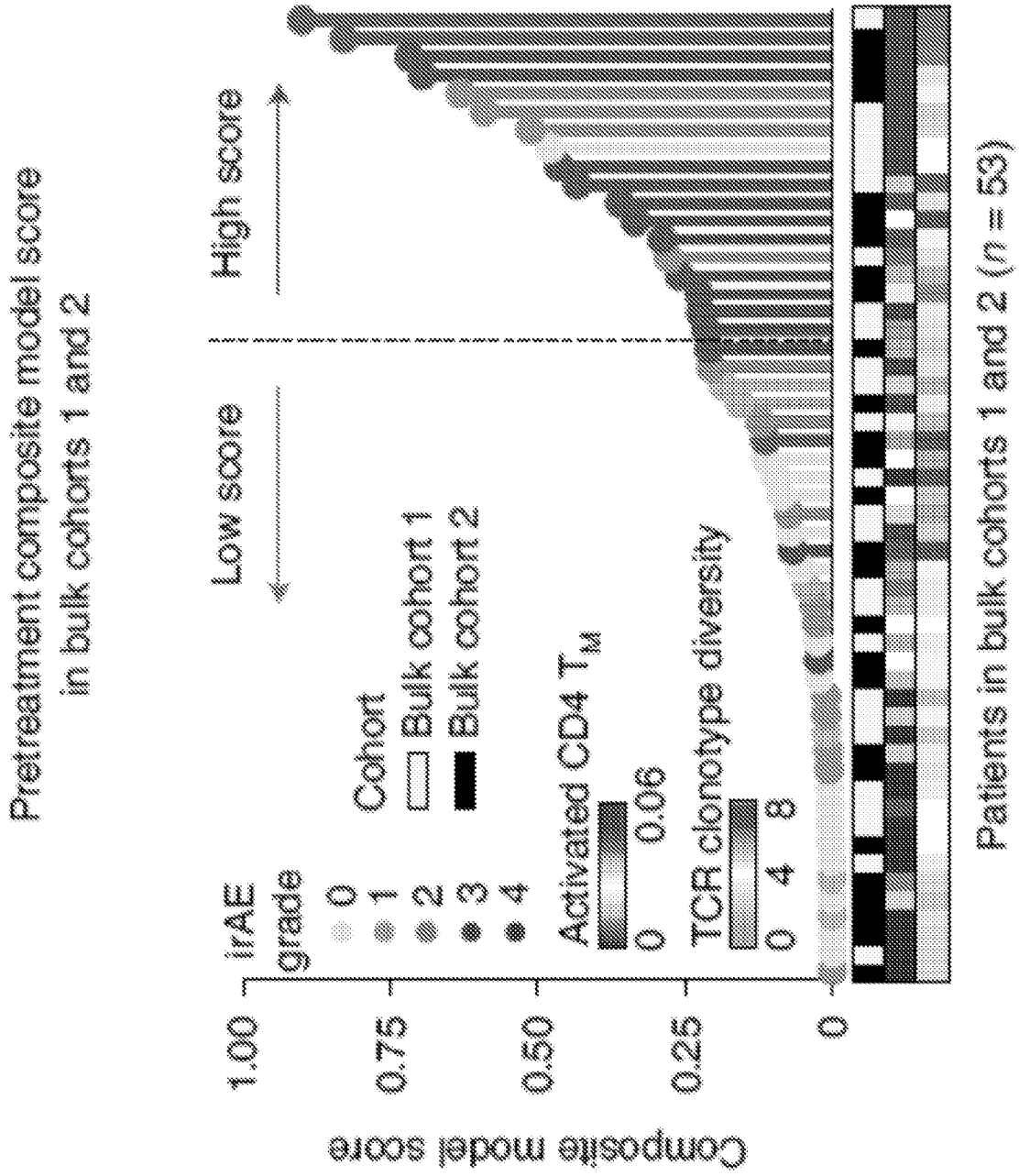


FIG. 4C

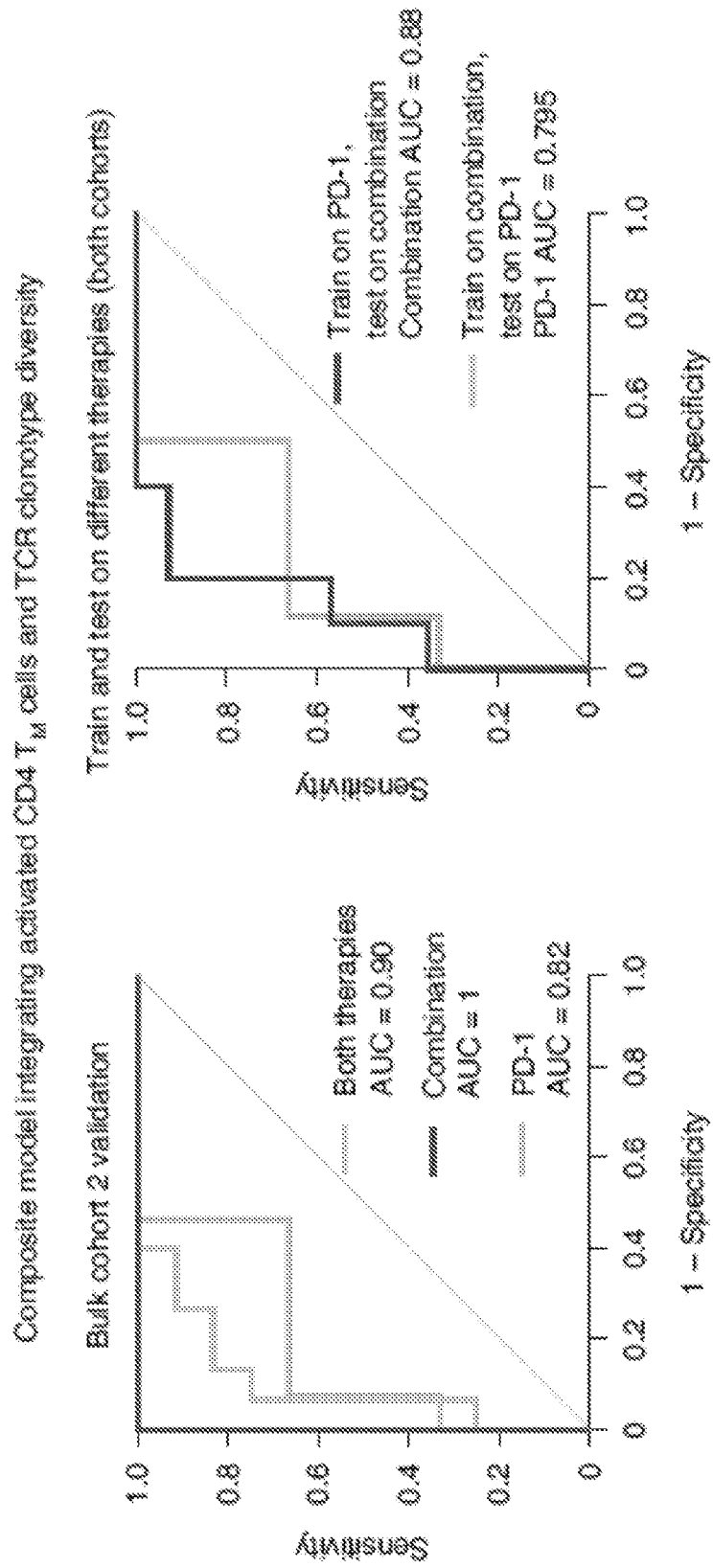


FIG. 4D

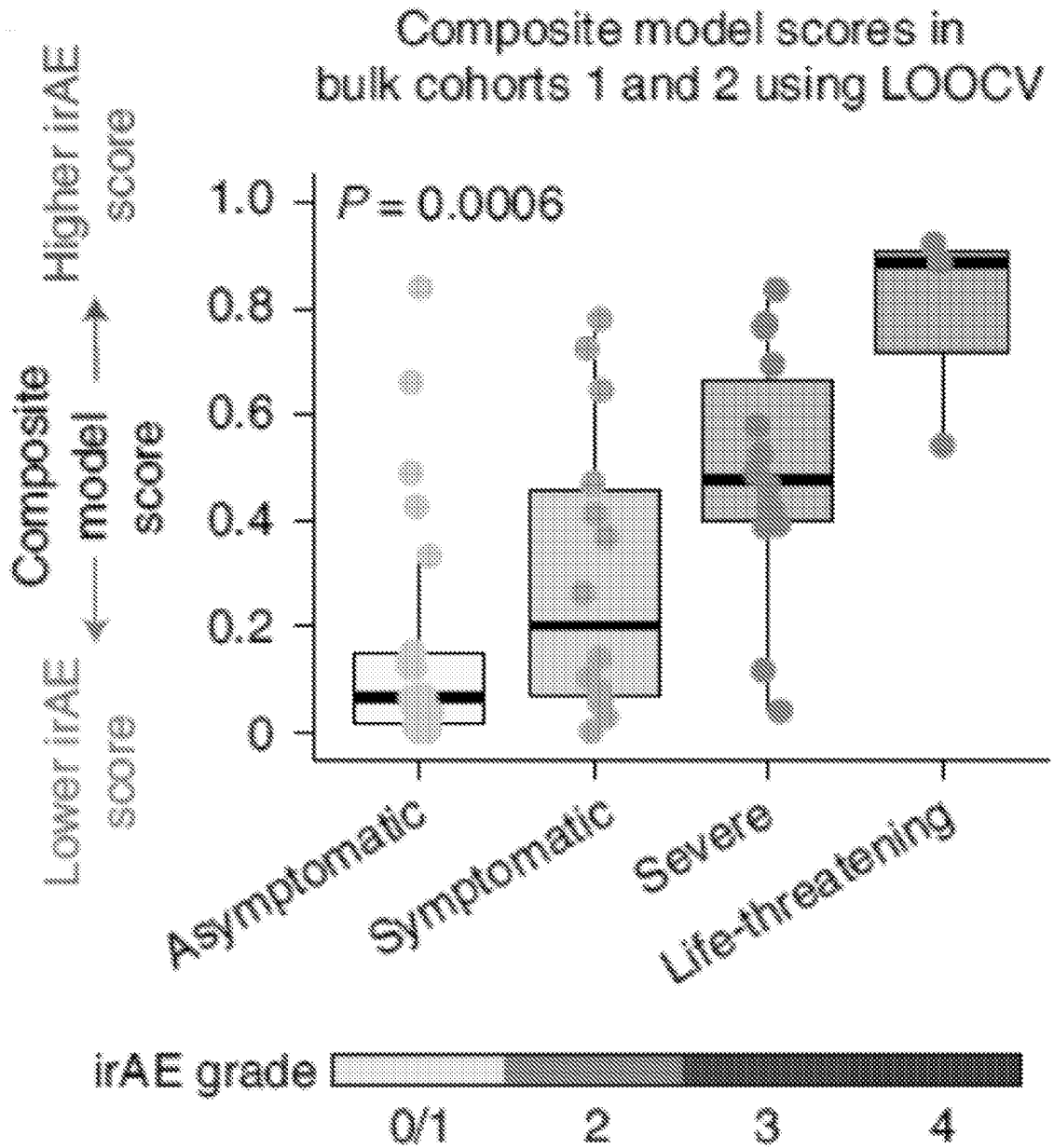
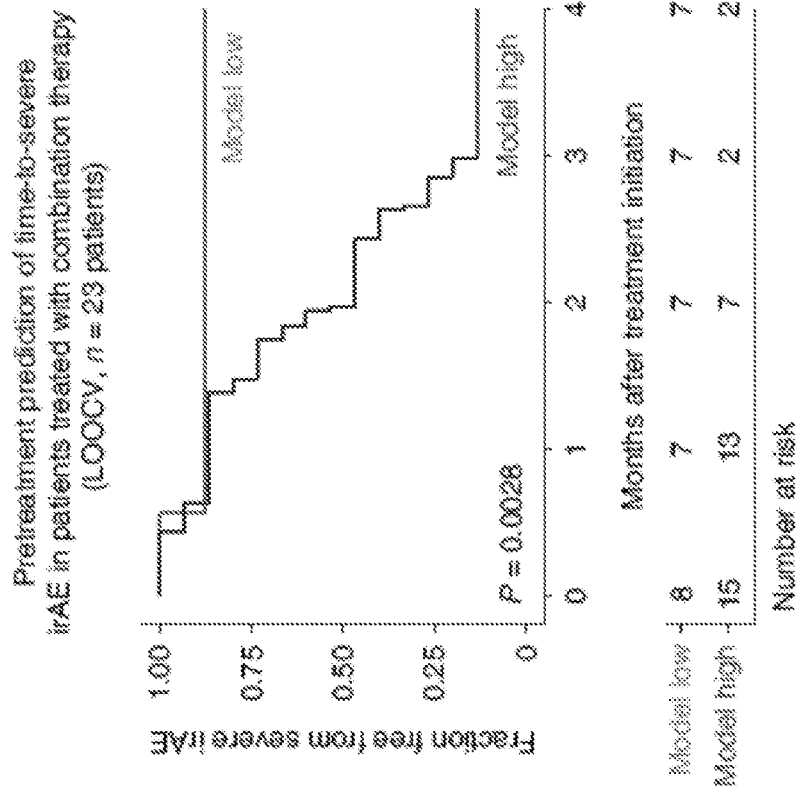
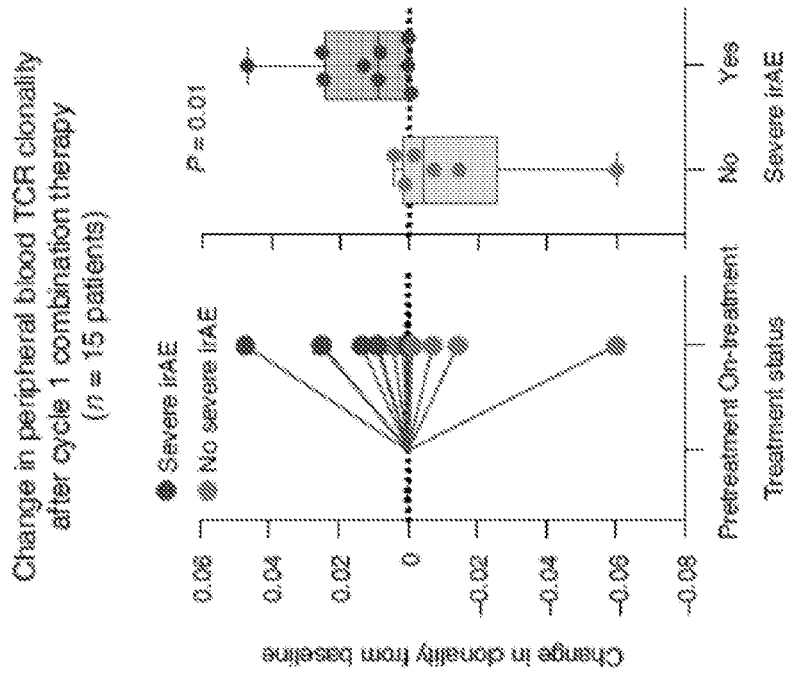


FIG. 4E



More clonal ← → More even

FIG. 5B

FIG. 5A

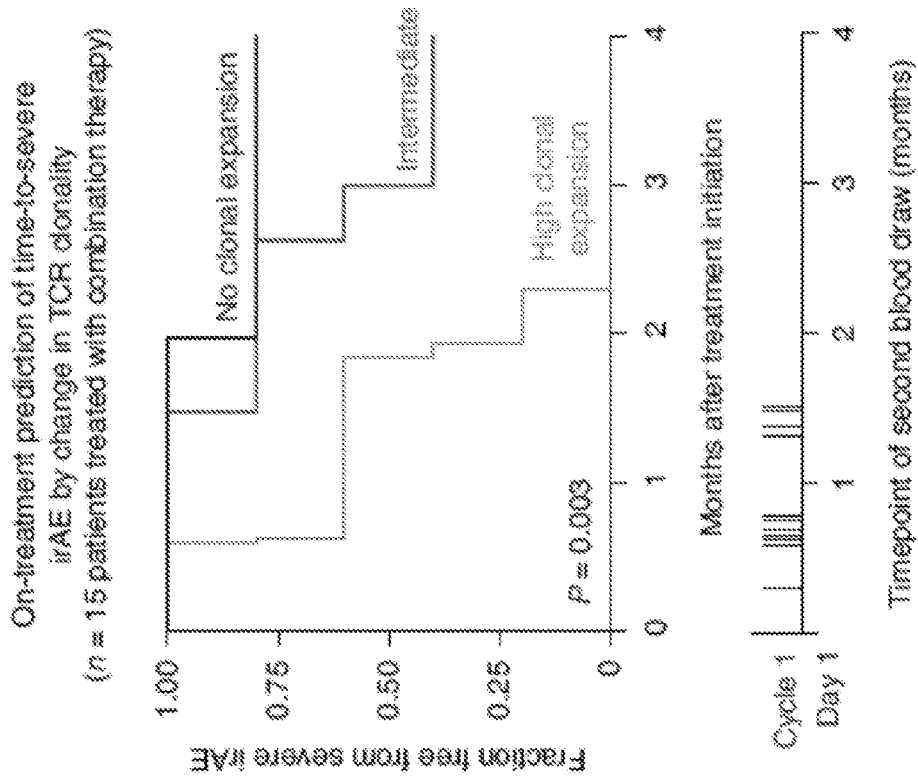


FIG. 5D

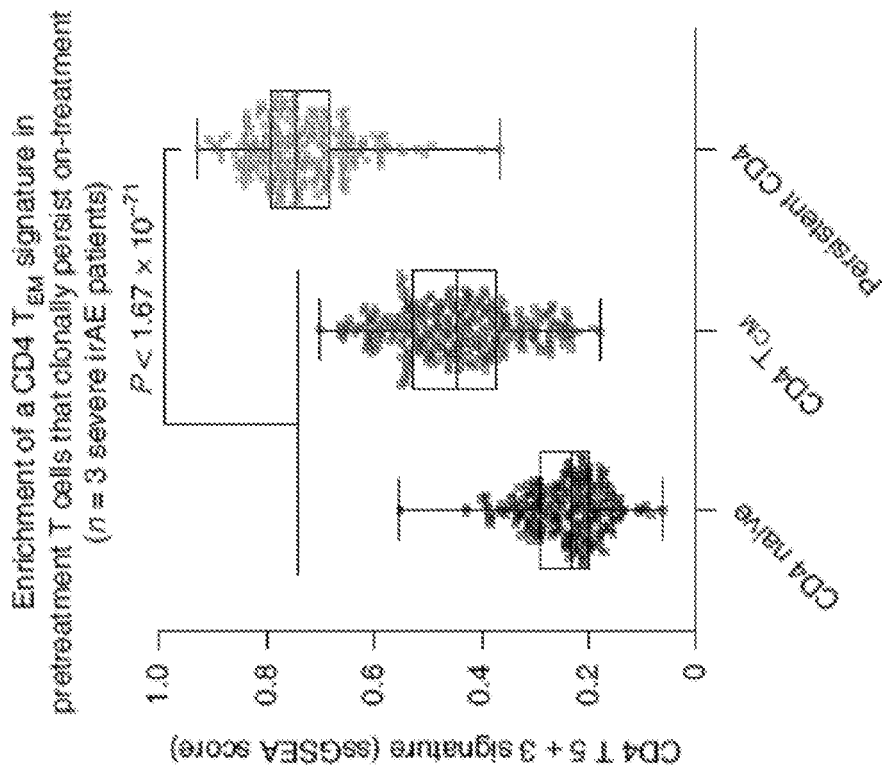


FIG. 5C

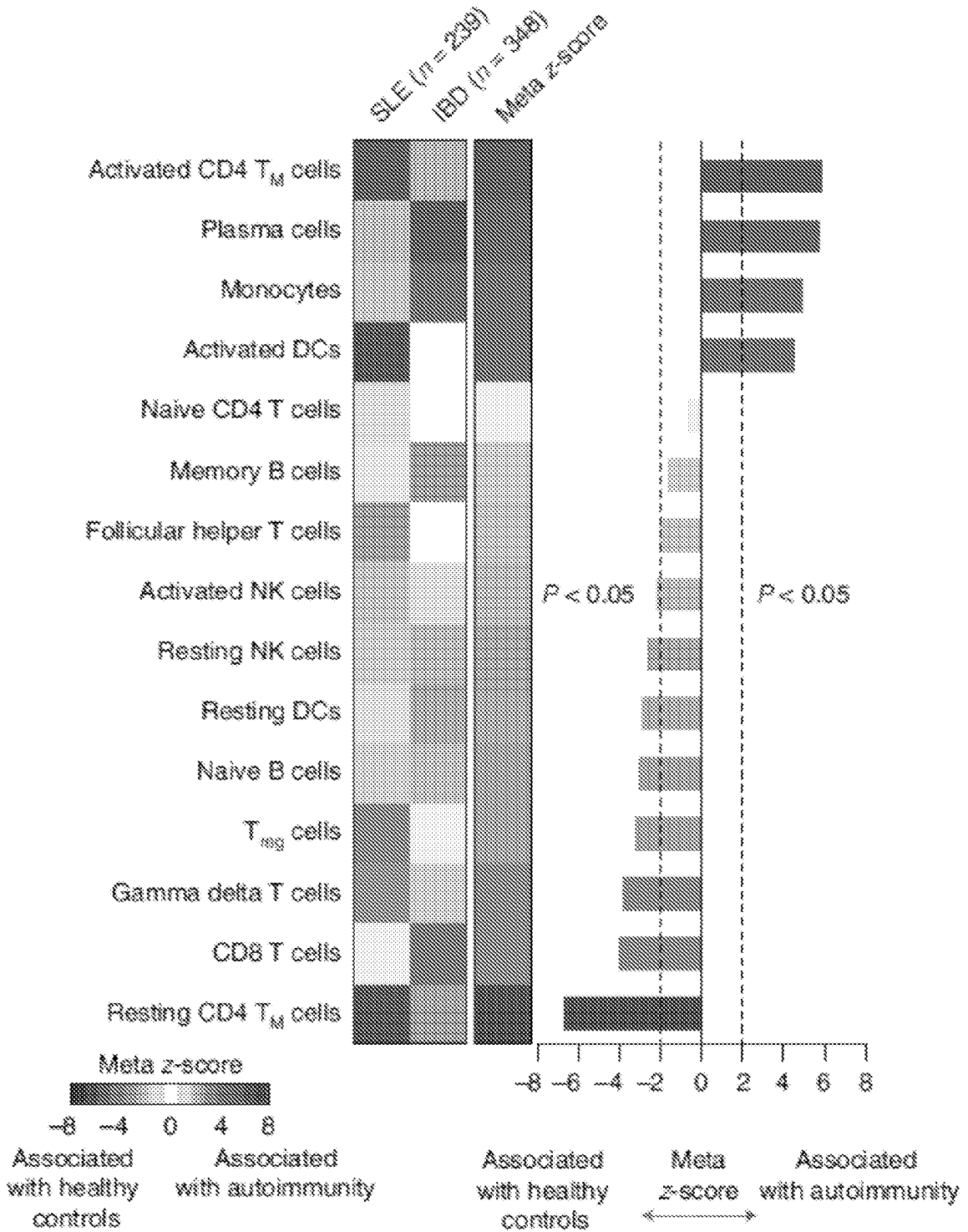


FIG. 6

FIG. 7A

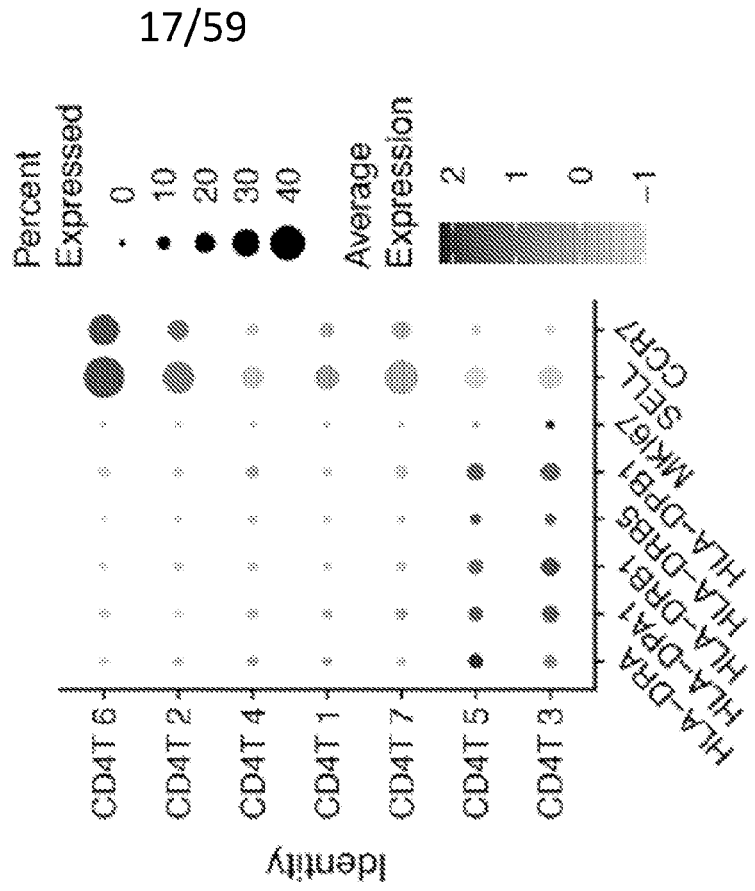


FIG. 7C

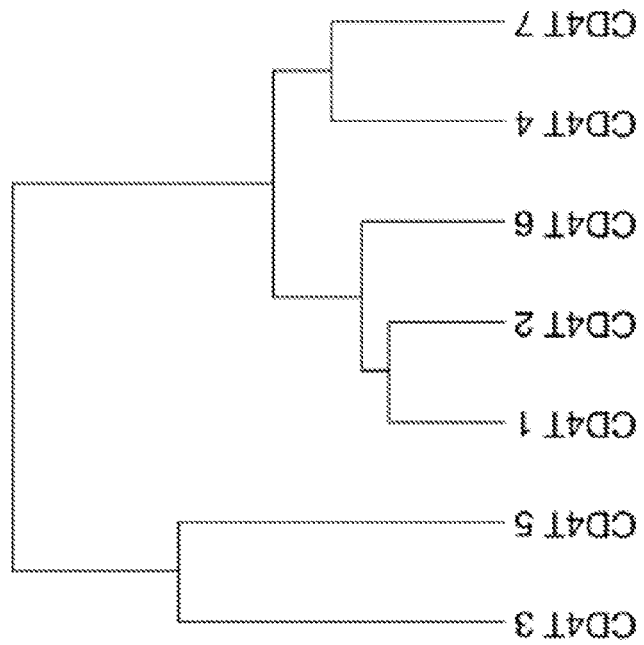


FIG. 7B

scRNA-seq clusters vs. irAE and CyTOF
(all pairwise comparisons within each cell type)

scRNA-seq clusters vs. irAE and CyTOF
(all pairwise comparisons within each cell type)

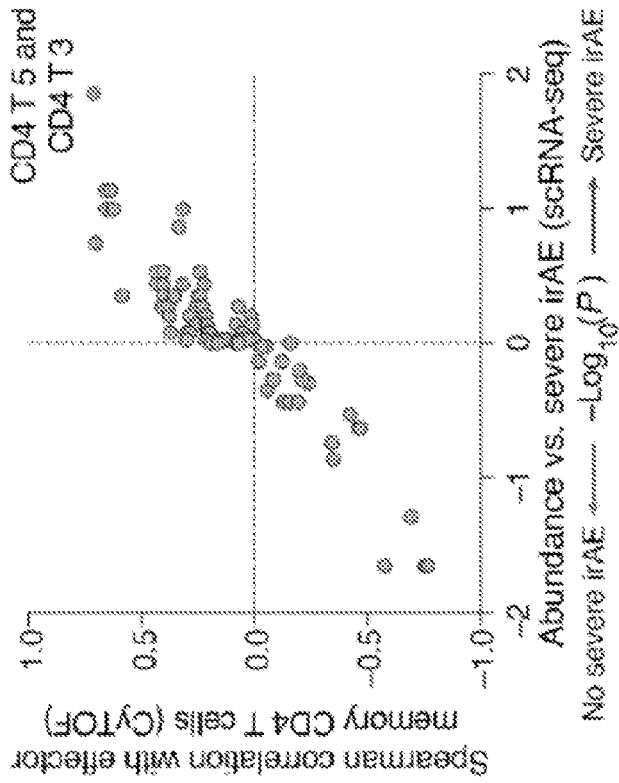
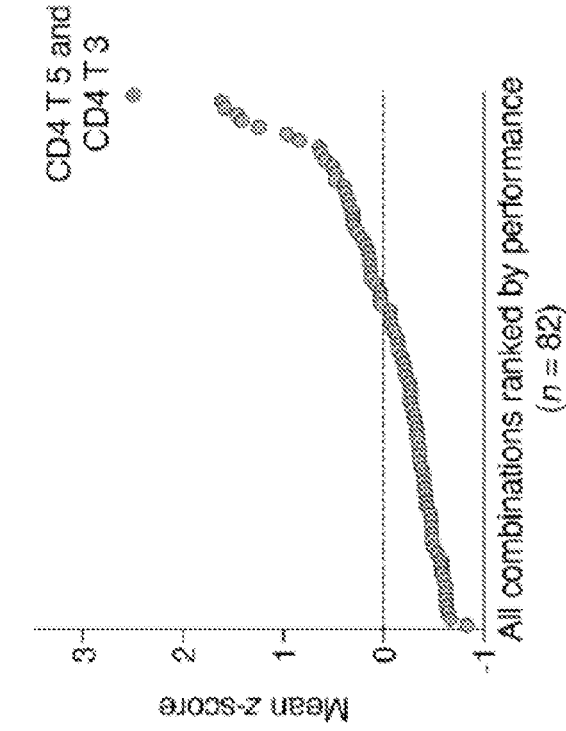


FIG. 7E

FIG. 7D

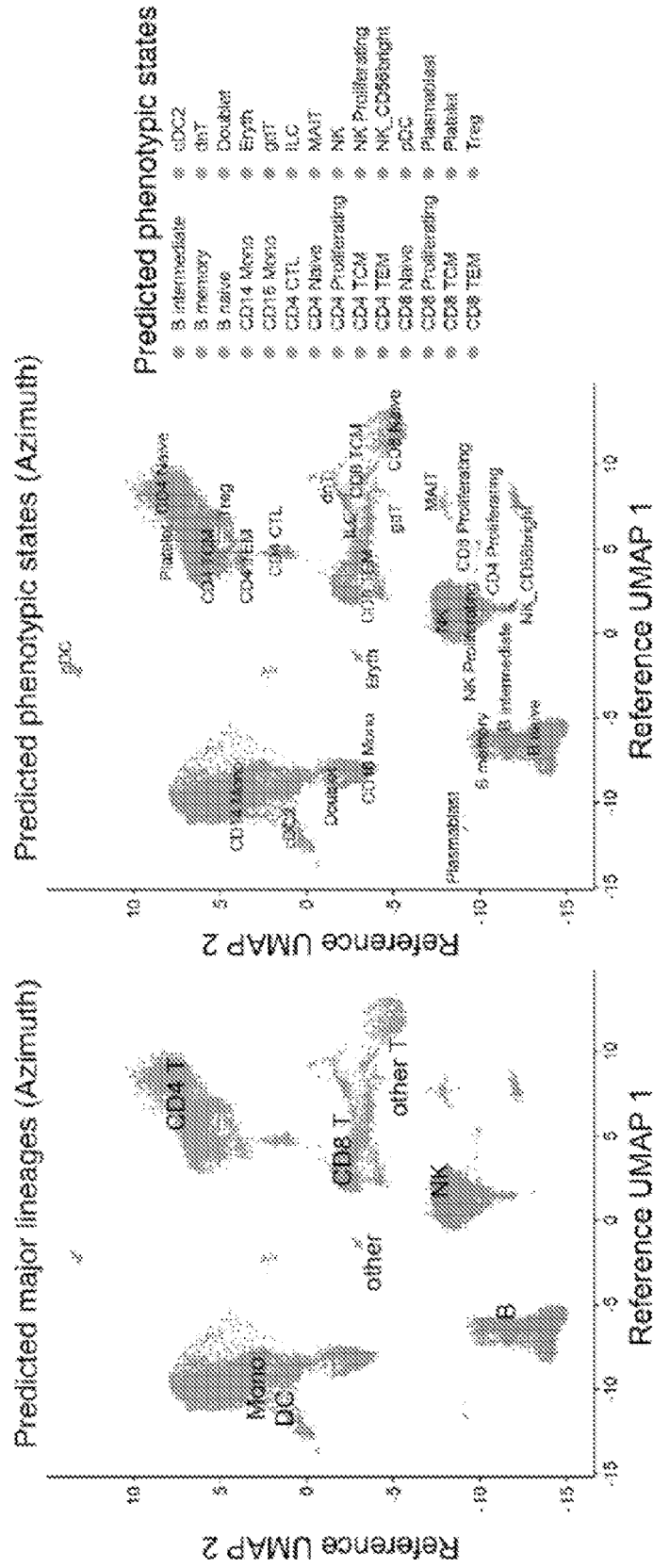


FIG. 8A

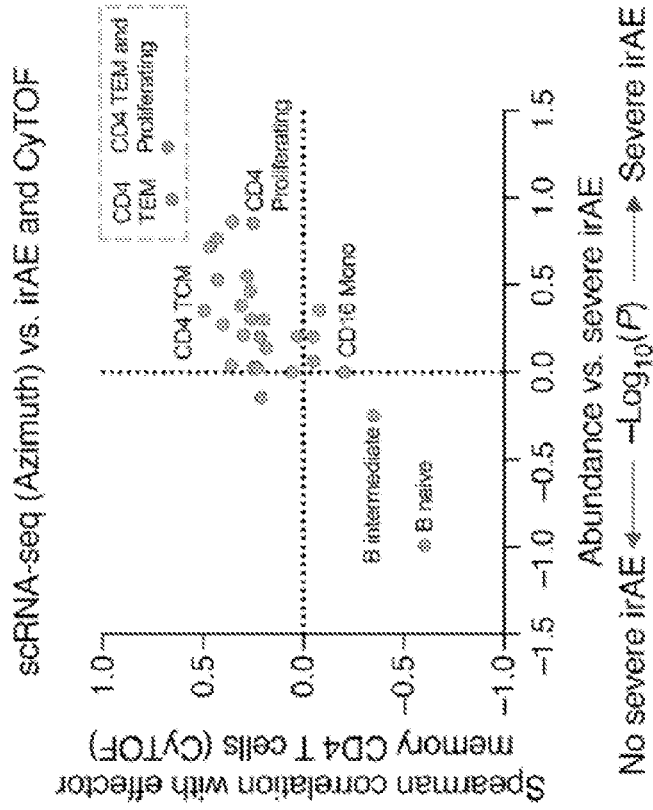


FIG. 8C

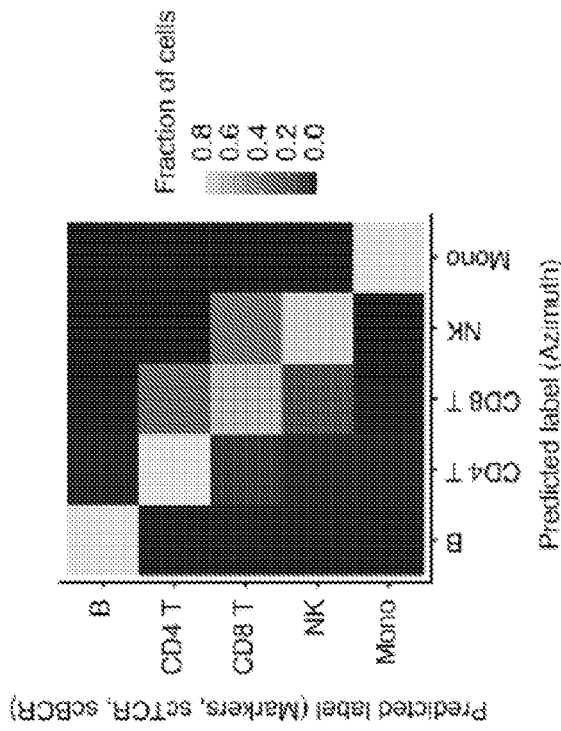


FIG. 8B

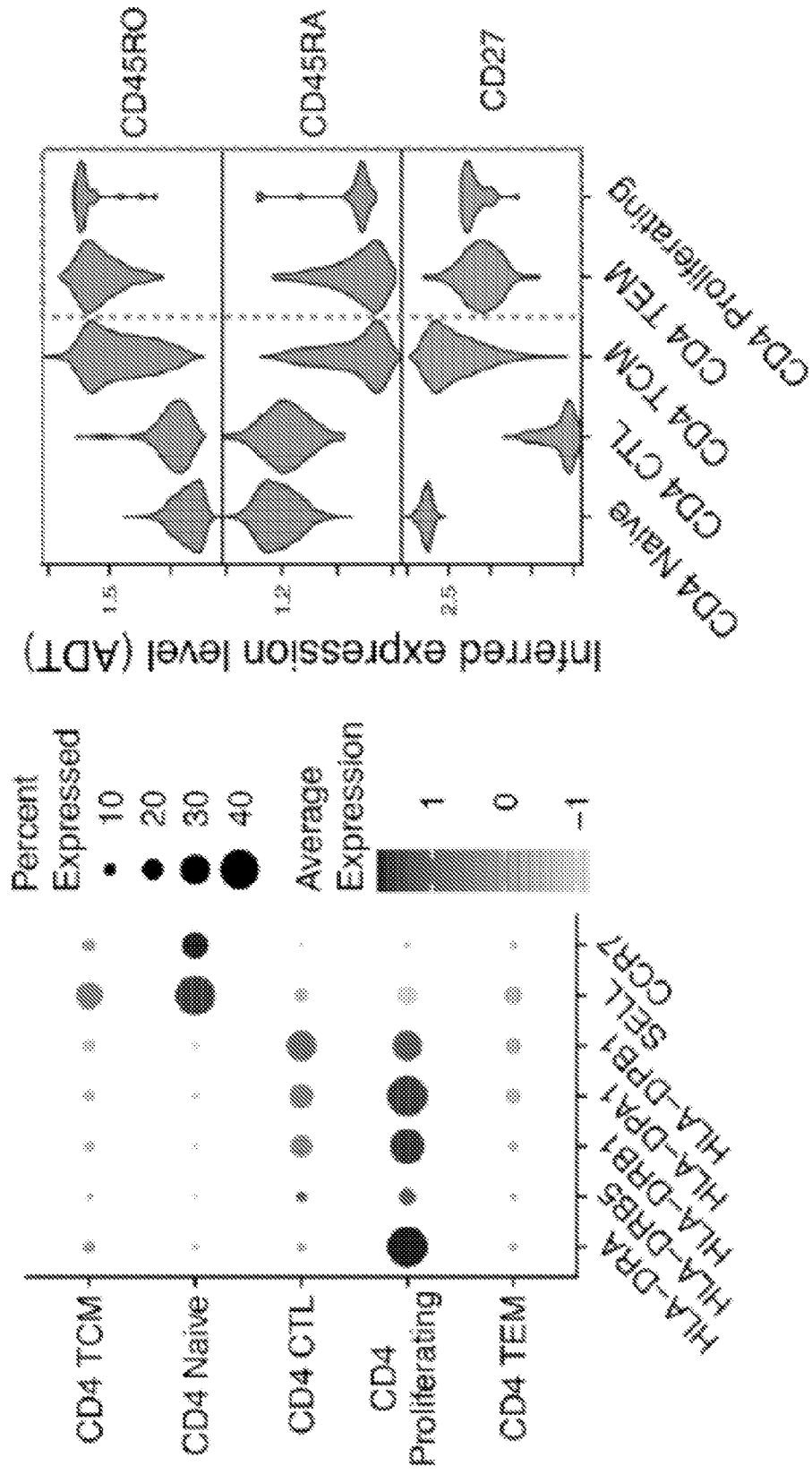


FIG. 8E

FIG. 8D

	CD4 TEM + Proliferating (Azimuth)	CD4 T 5 + 3 (Unsupervised)
Abundance vs. severe irAE (Nominal p-value)	0.05	0.01
Spearman correlation with CD4 TEM (CyTOF)	0.68	0.80

FIG. 8F

Association between pretreatment T cell states and severe irAE status
(scRNA-seq, n = 13)

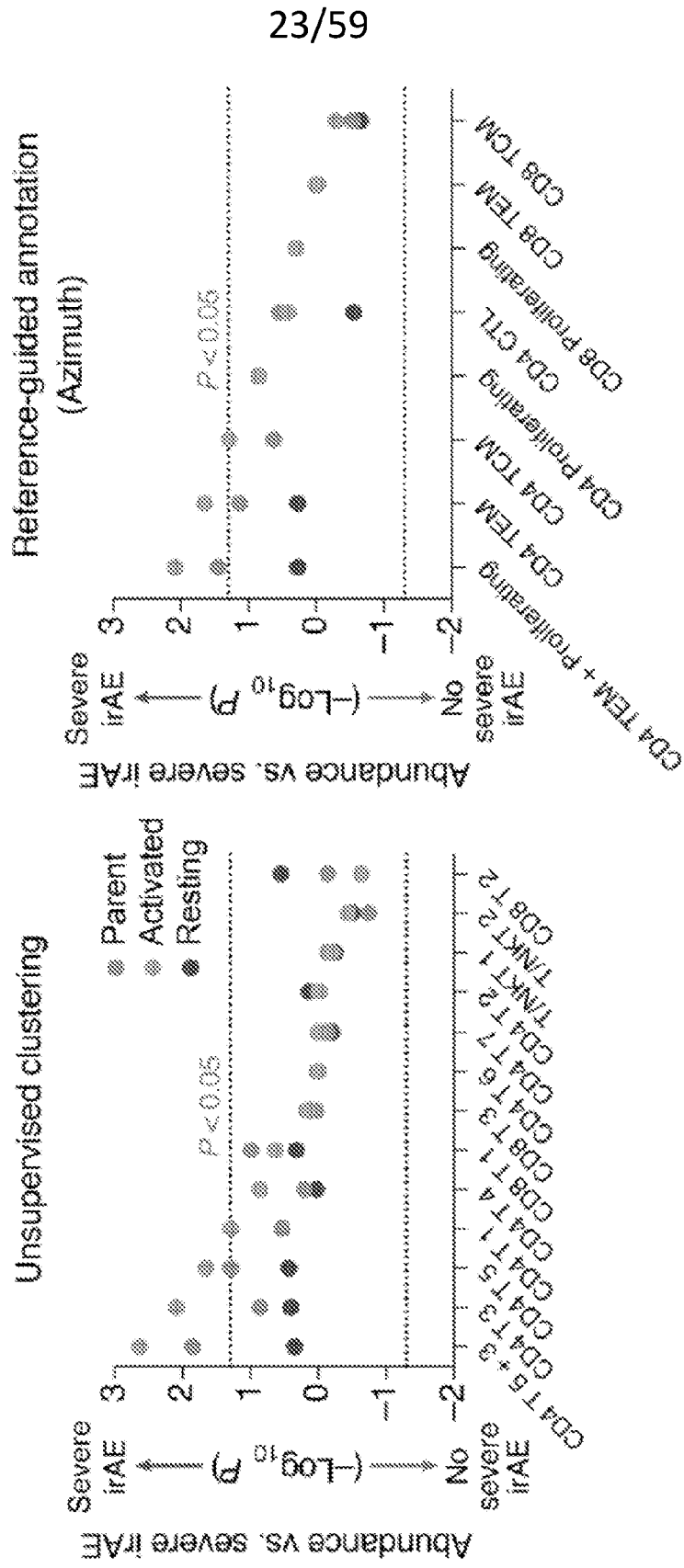


FIG. 9A

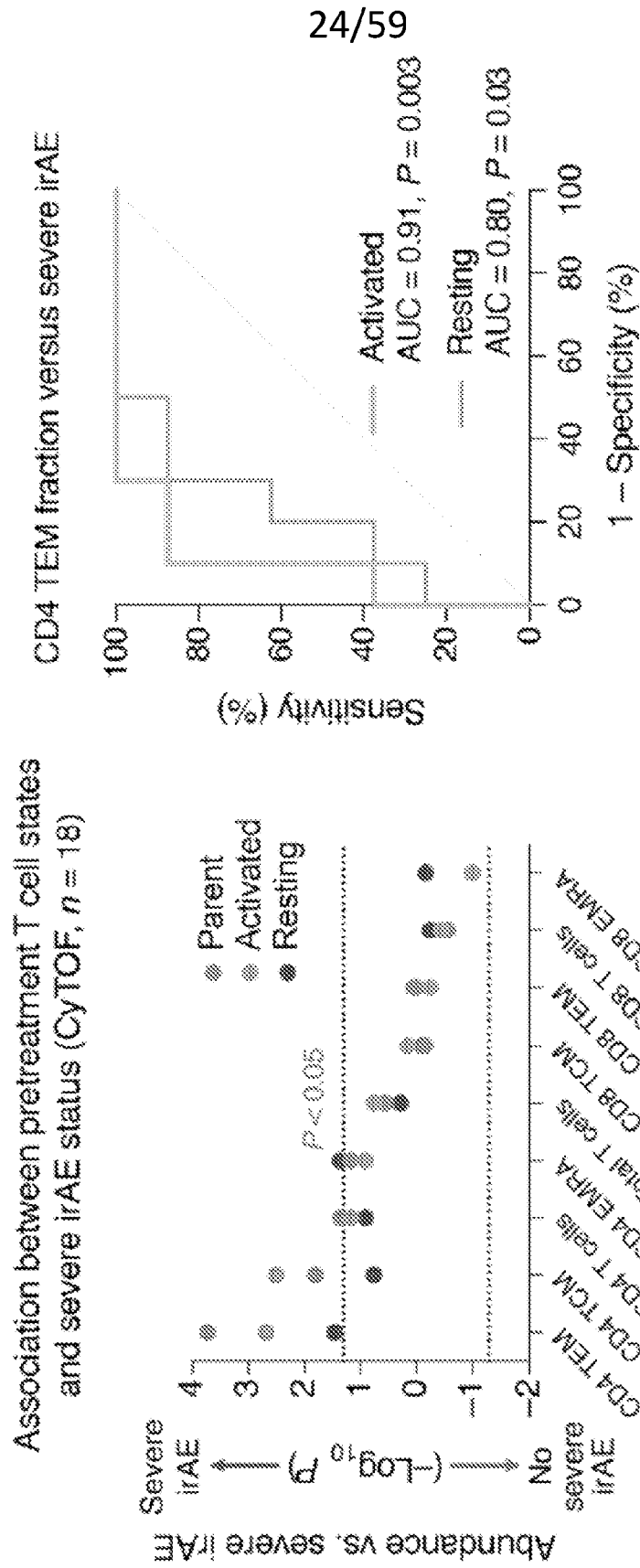


FIG. 9B

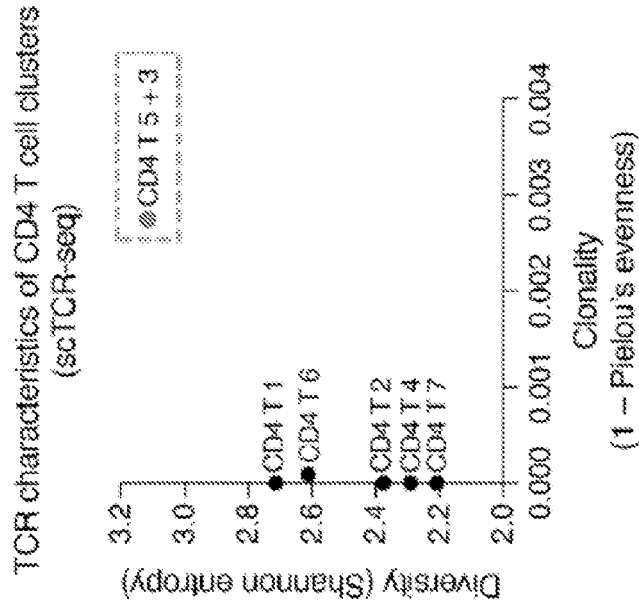


FIG. 10A

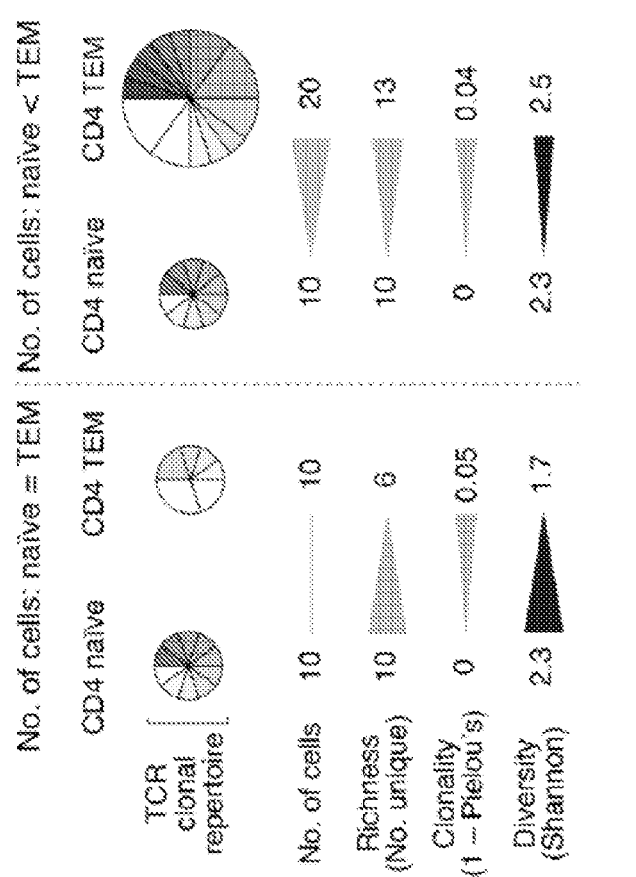


FIG. 10B

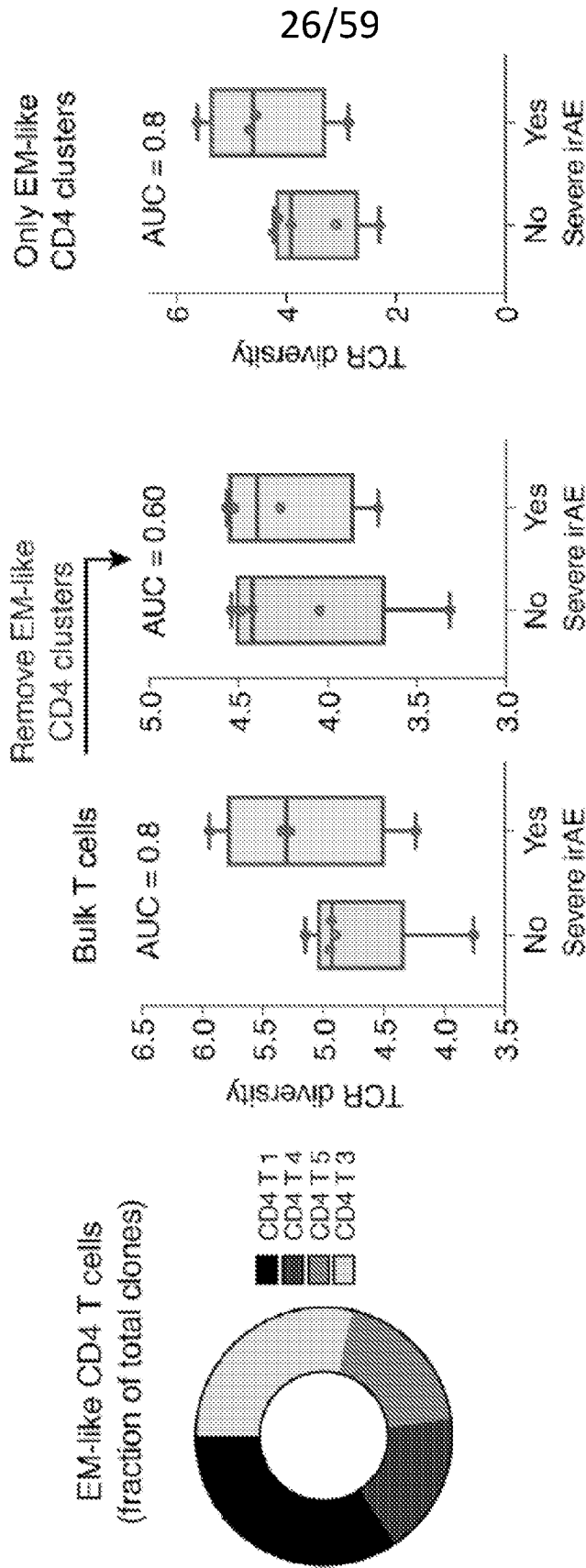


FIG. 10E

FIG. 10D

FIG. 10C

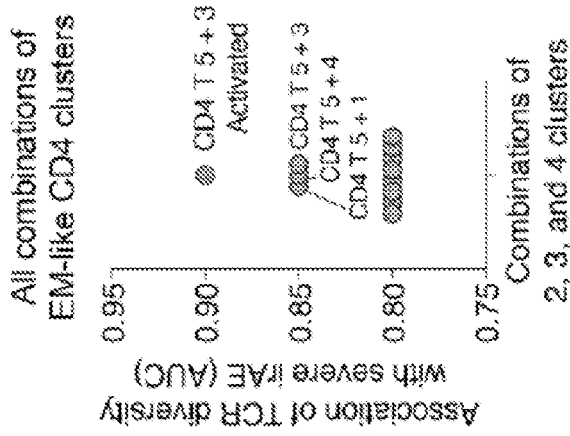
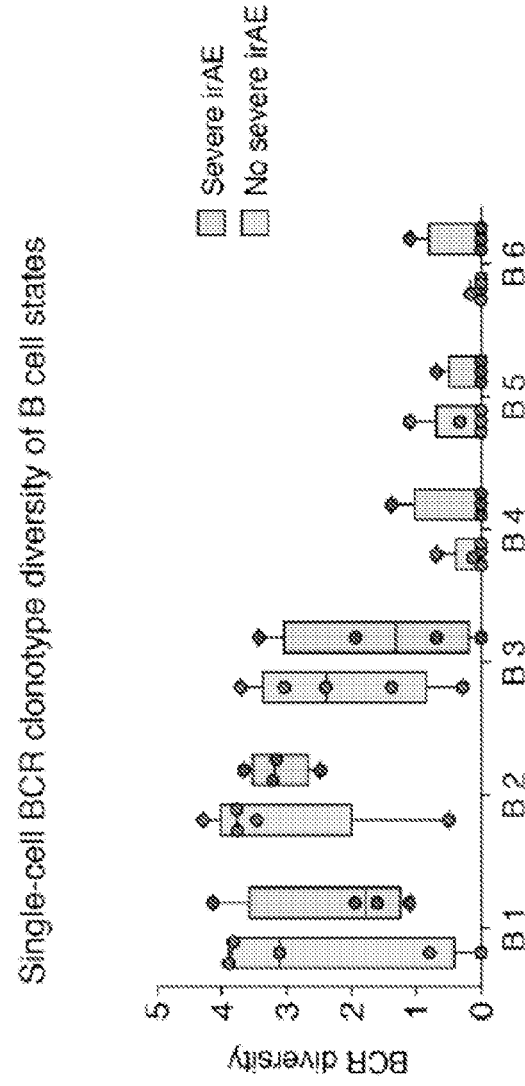


FIG. 10G

FIG. 10F

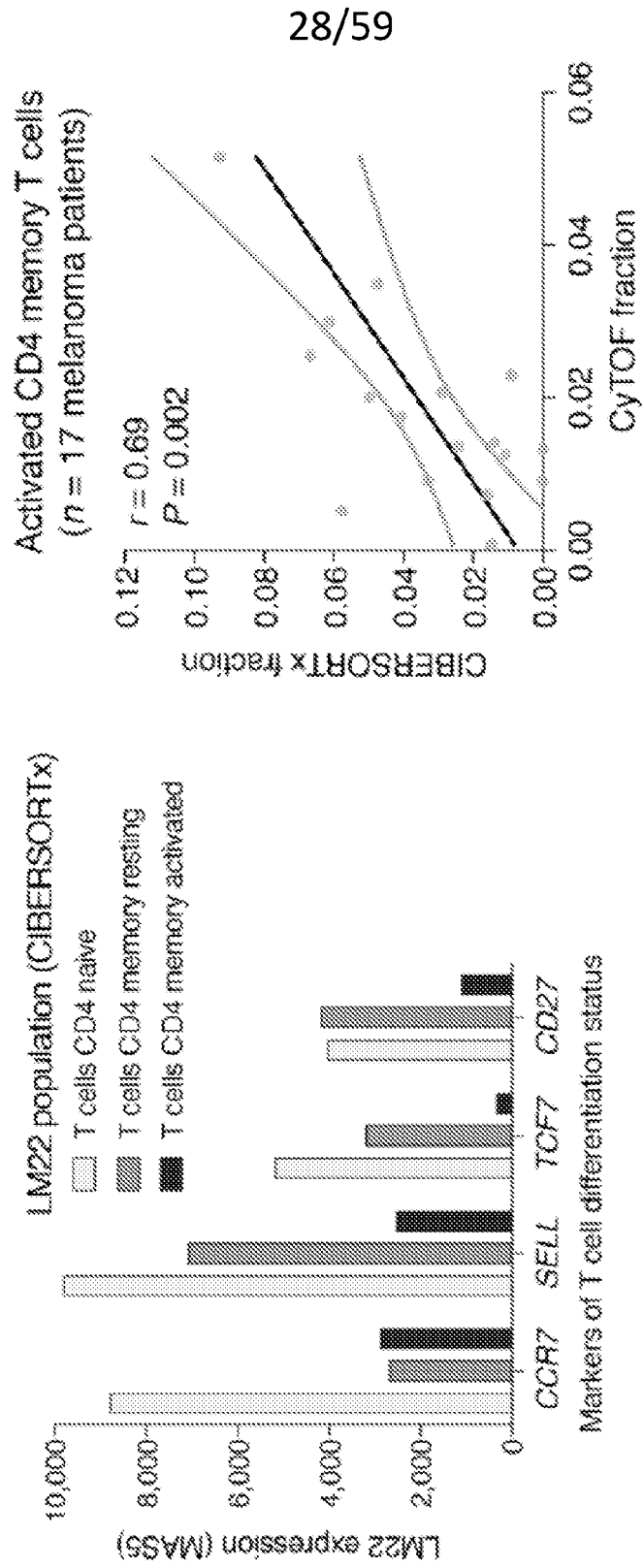


FIG. 11B

FIG. 11A

29/59

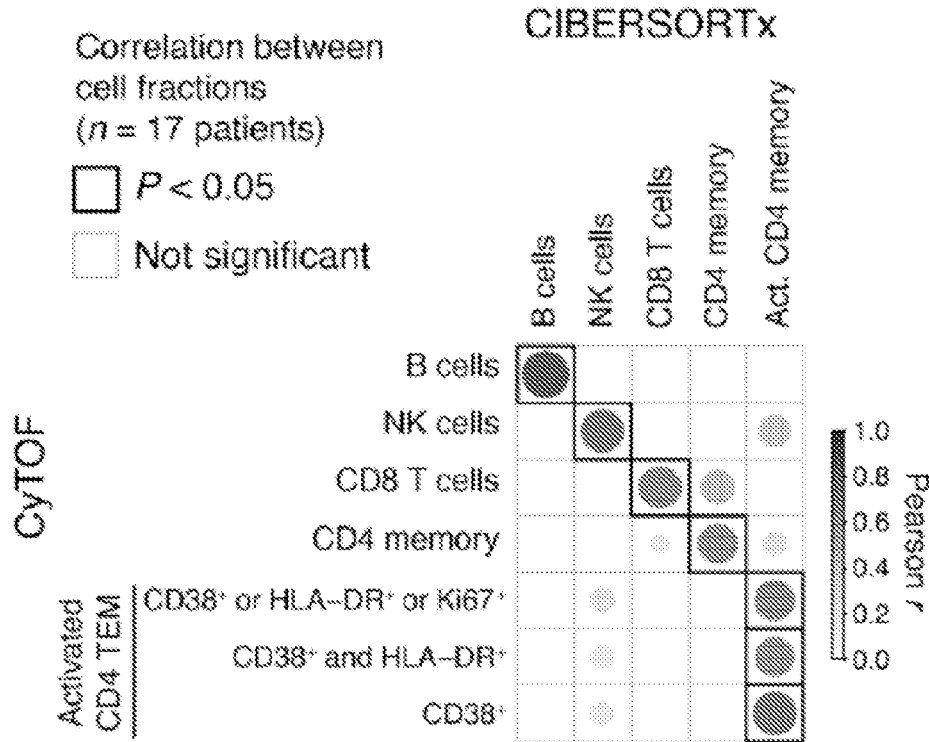


FIG. 11C

Correlation between activated CD4 memory T cell abundance (CIBERSORTx) and 14 memory T cell subsets (CyTOF) ($n = 17$ patients)

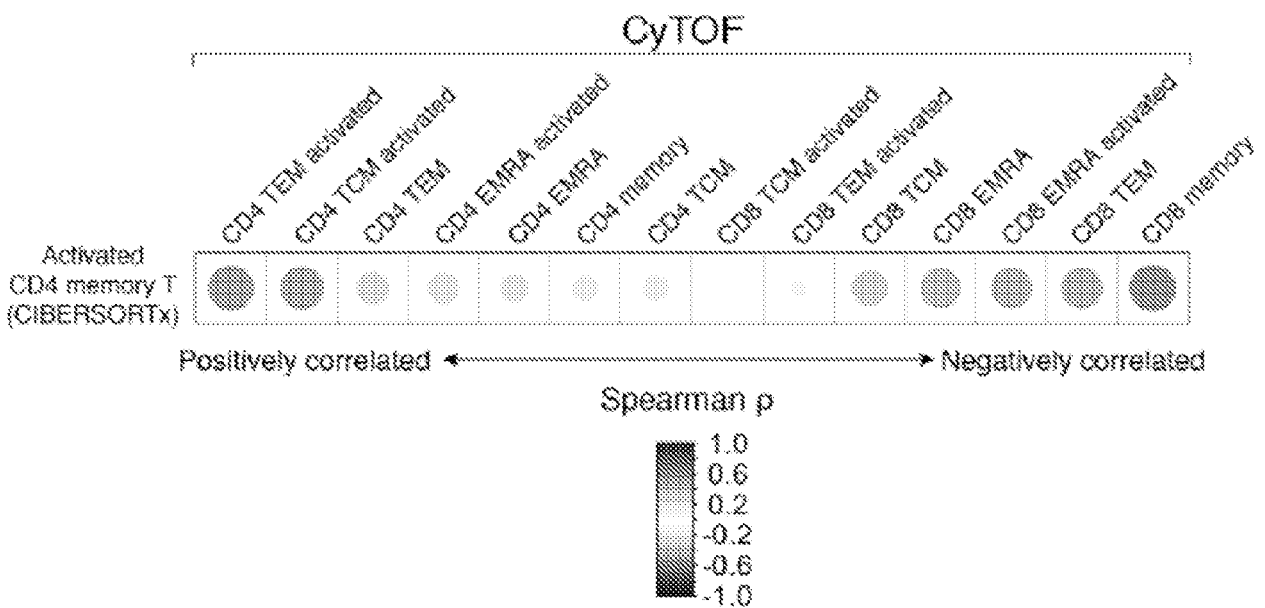


FIG. 11D

30/59

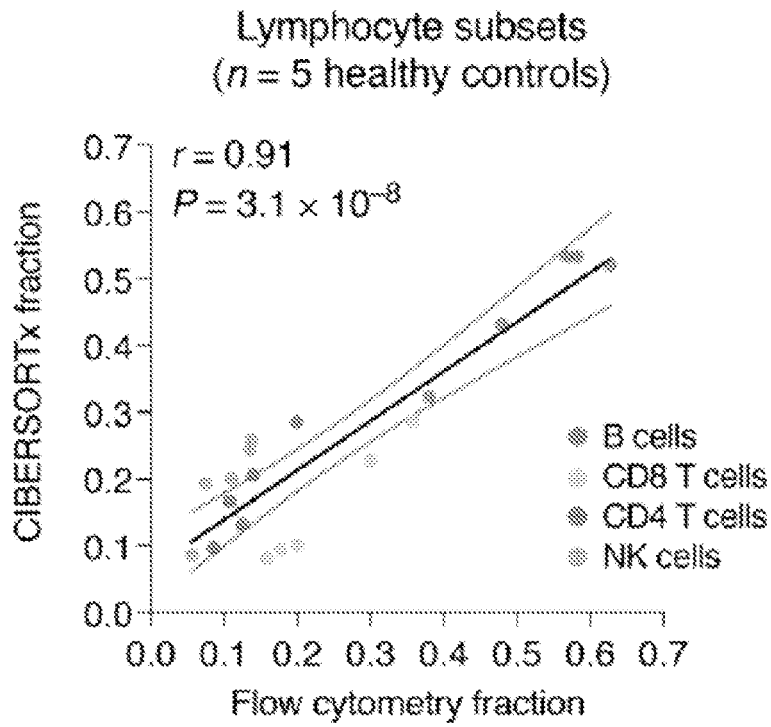


FIG. 11E

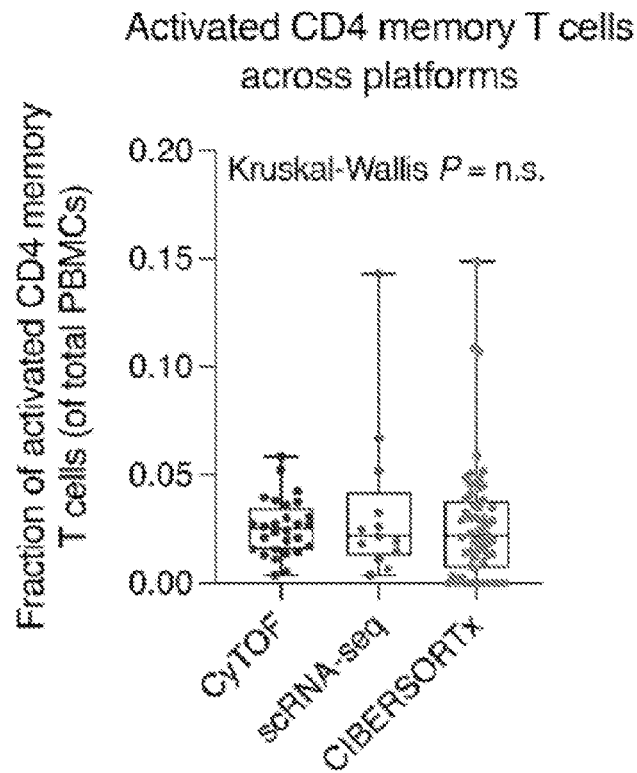


FIG. 11F

Combination therapy (Shannon entropy)

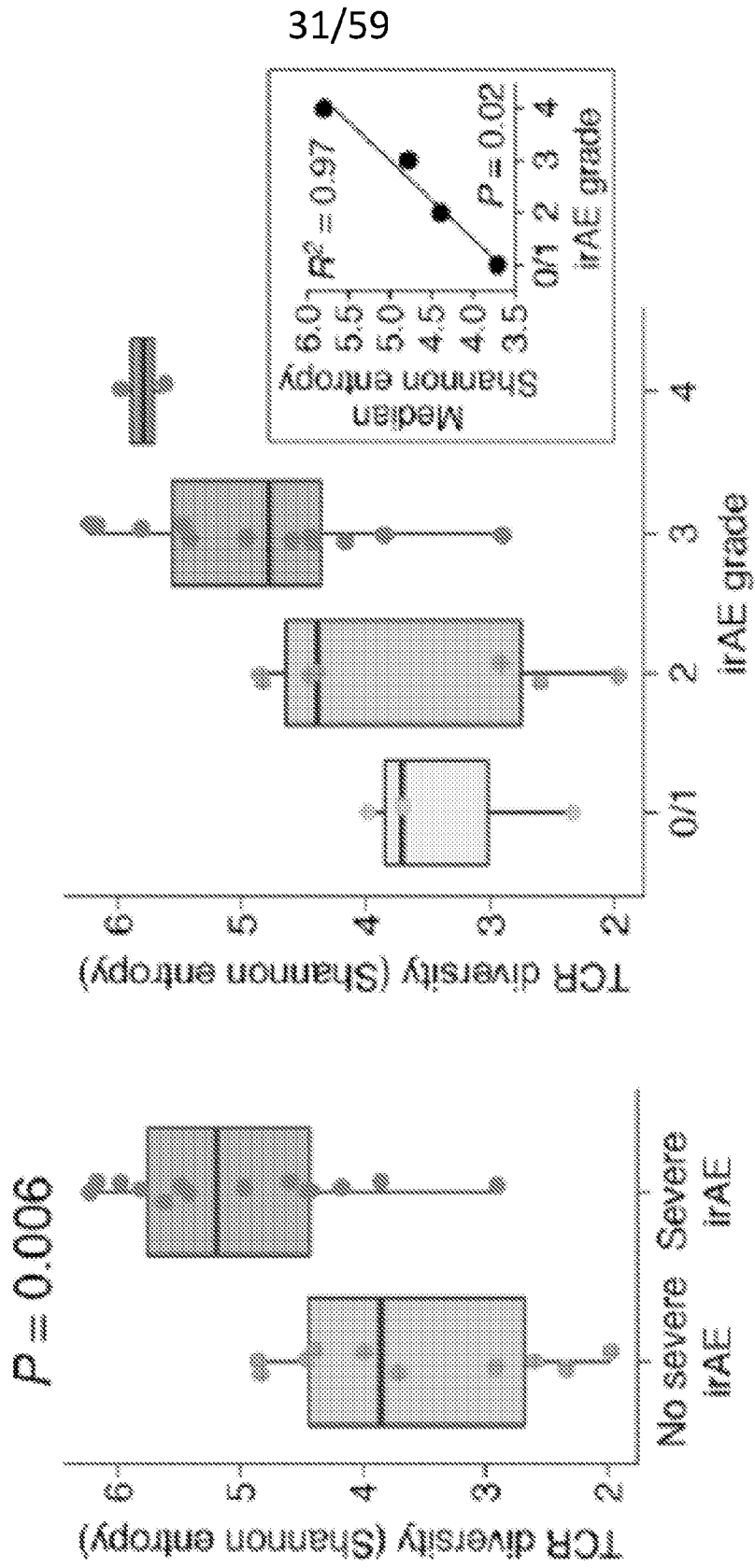


FIG. 12A

Combination therapy (Gini-Simpson index)

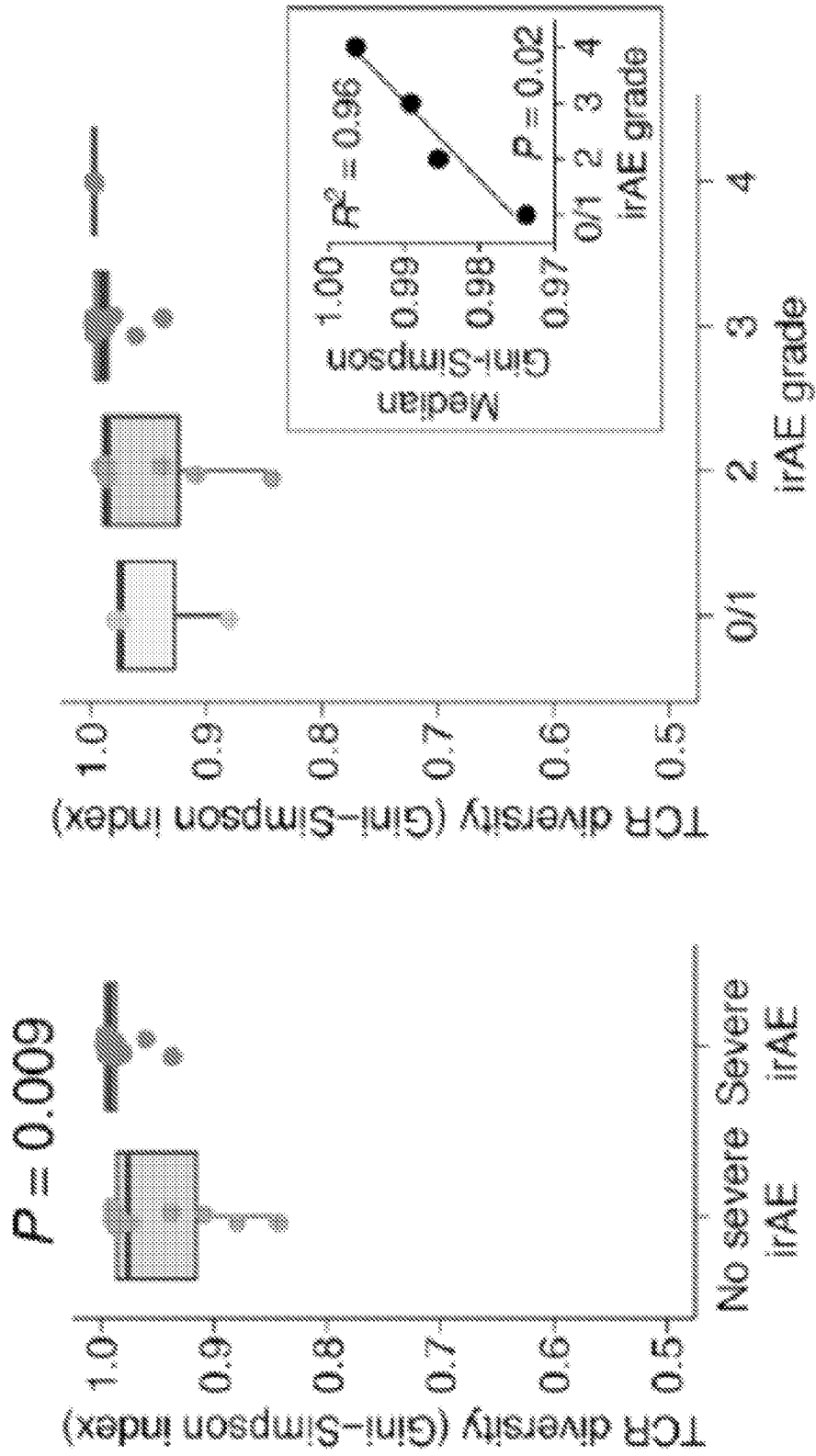


FIG. 12B

33/59

PD1 monotherapy (Shannon entropy)

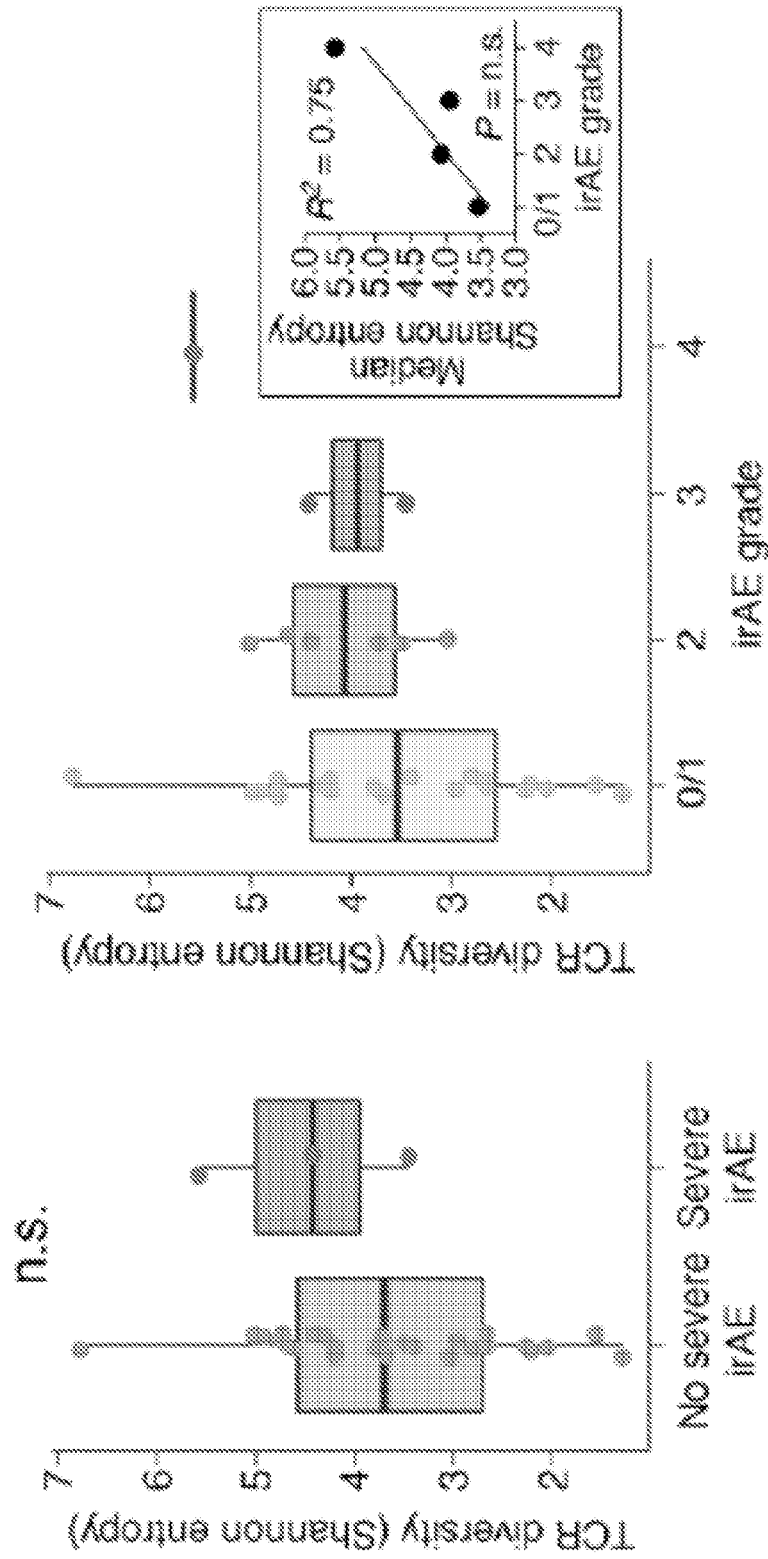


FIG. 12C

PD1 monotherapy (Gini-Simpson index)

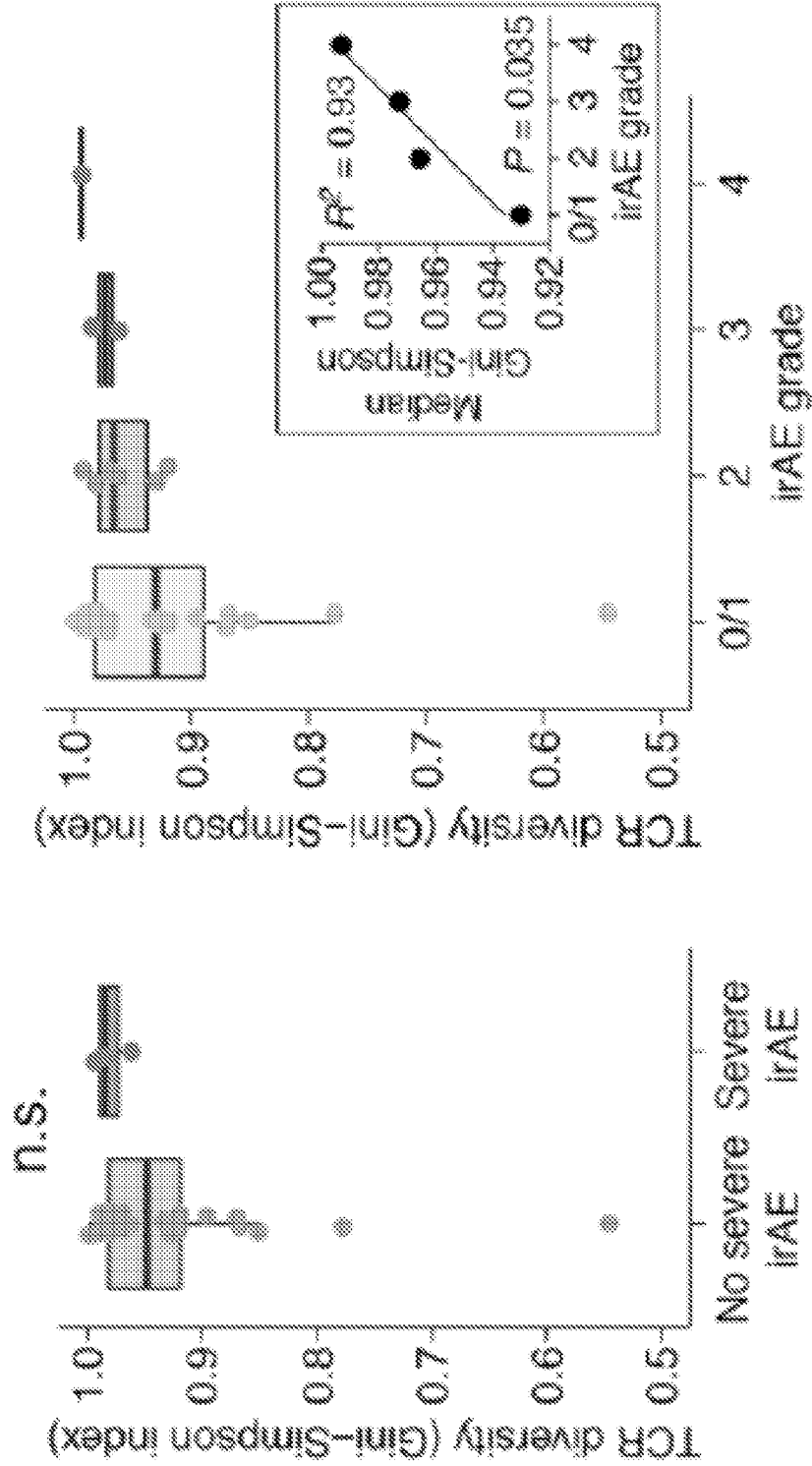


FIG. 12D

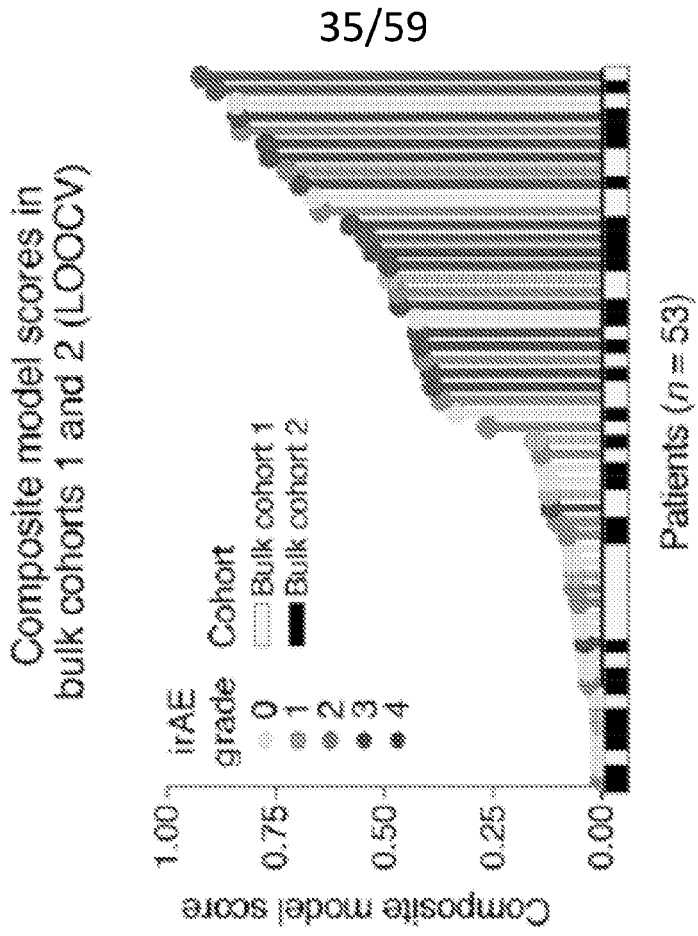


FIG. 13A

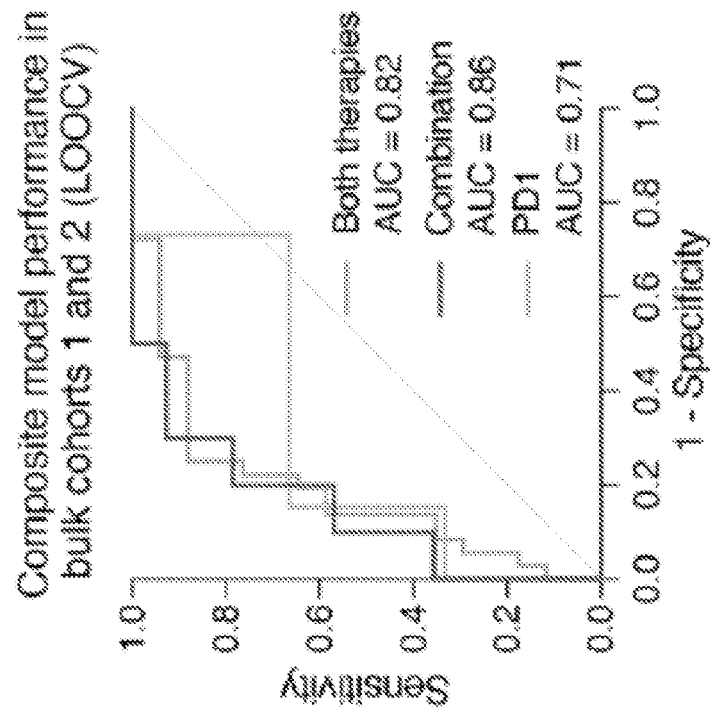


FIG. 13B

Candidate determinants of severe irAE in bulk cohorts 1 and 2

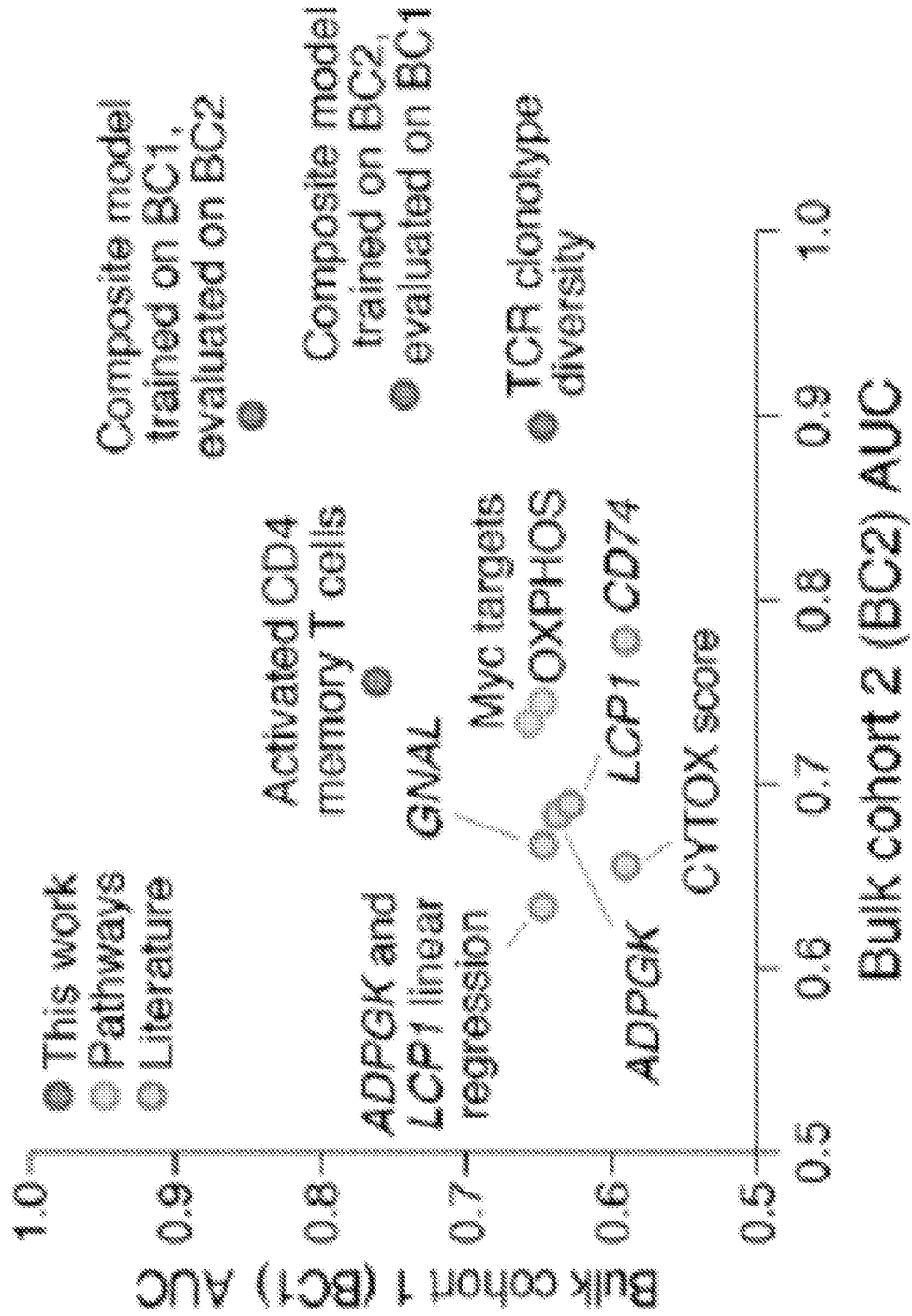


FIG. 13C

Performance of composite model for predicting severe irAEs in different patient subsets (bulk cohort 2)

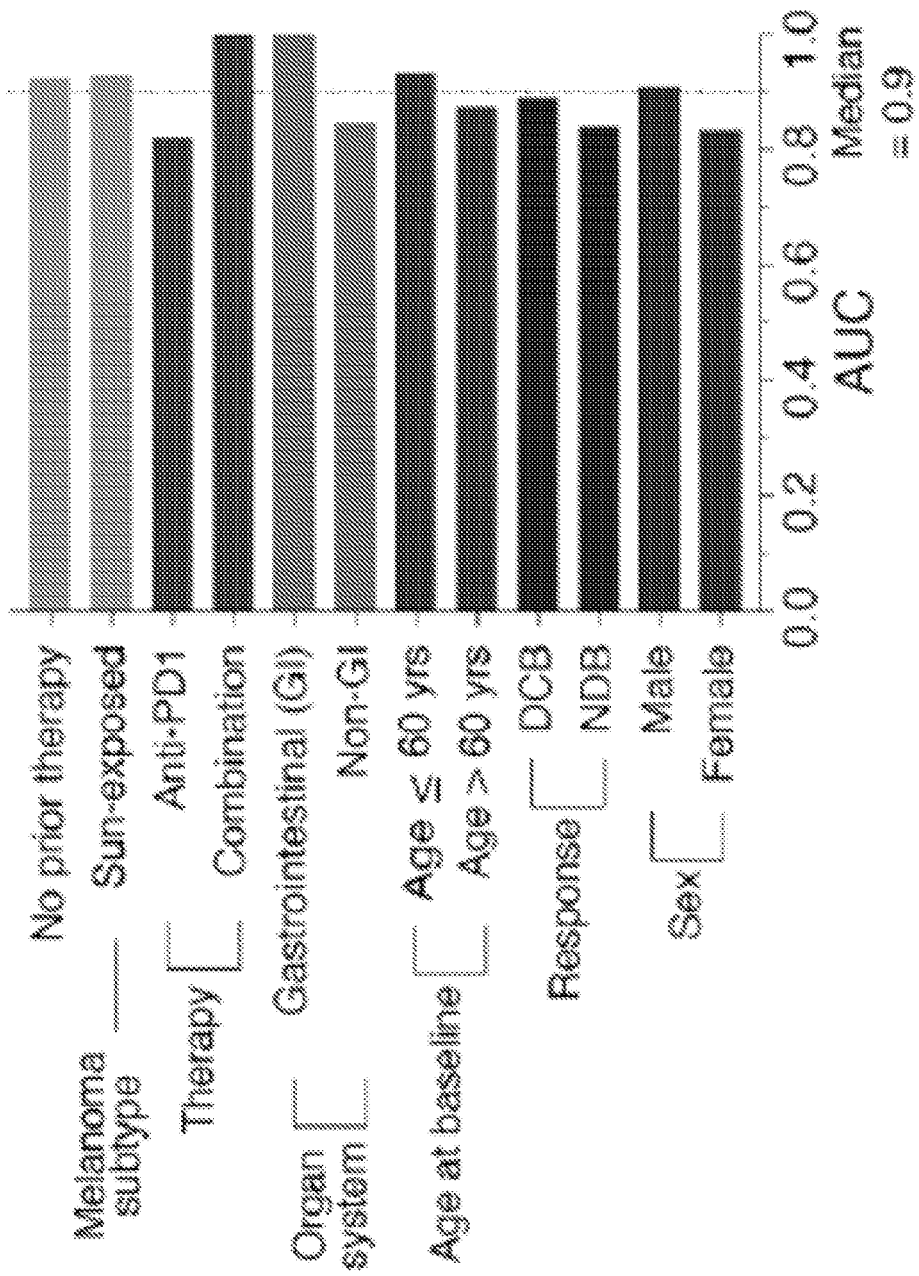


FIG. 13D

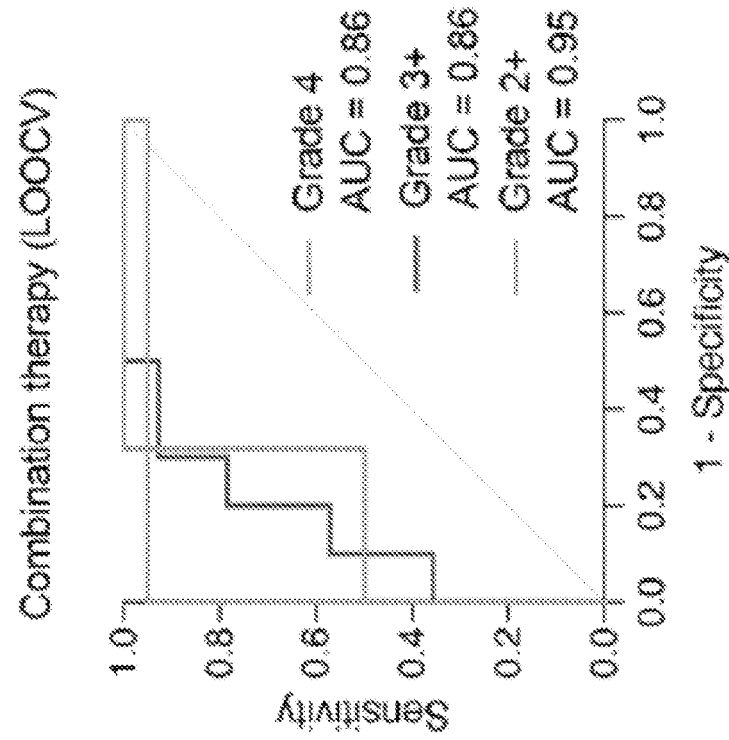


FIG. 13F

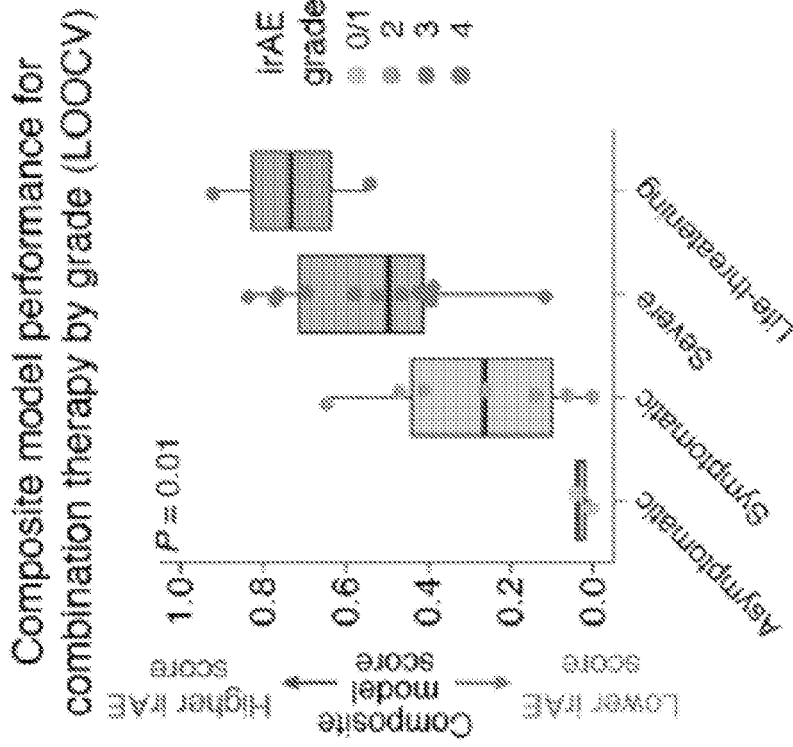


FIG. 13E

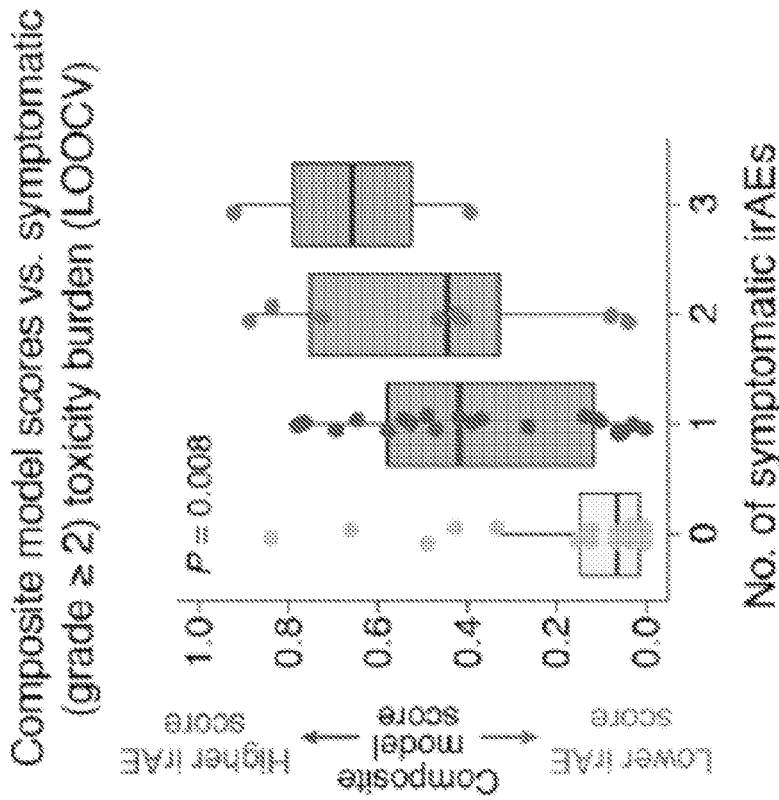
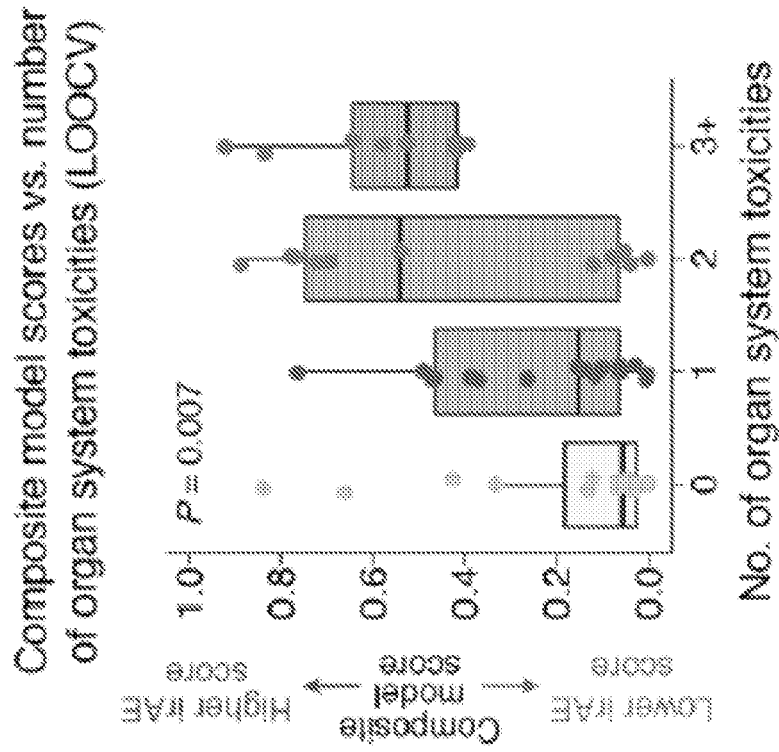


FIG. 13H

FIG. 13G

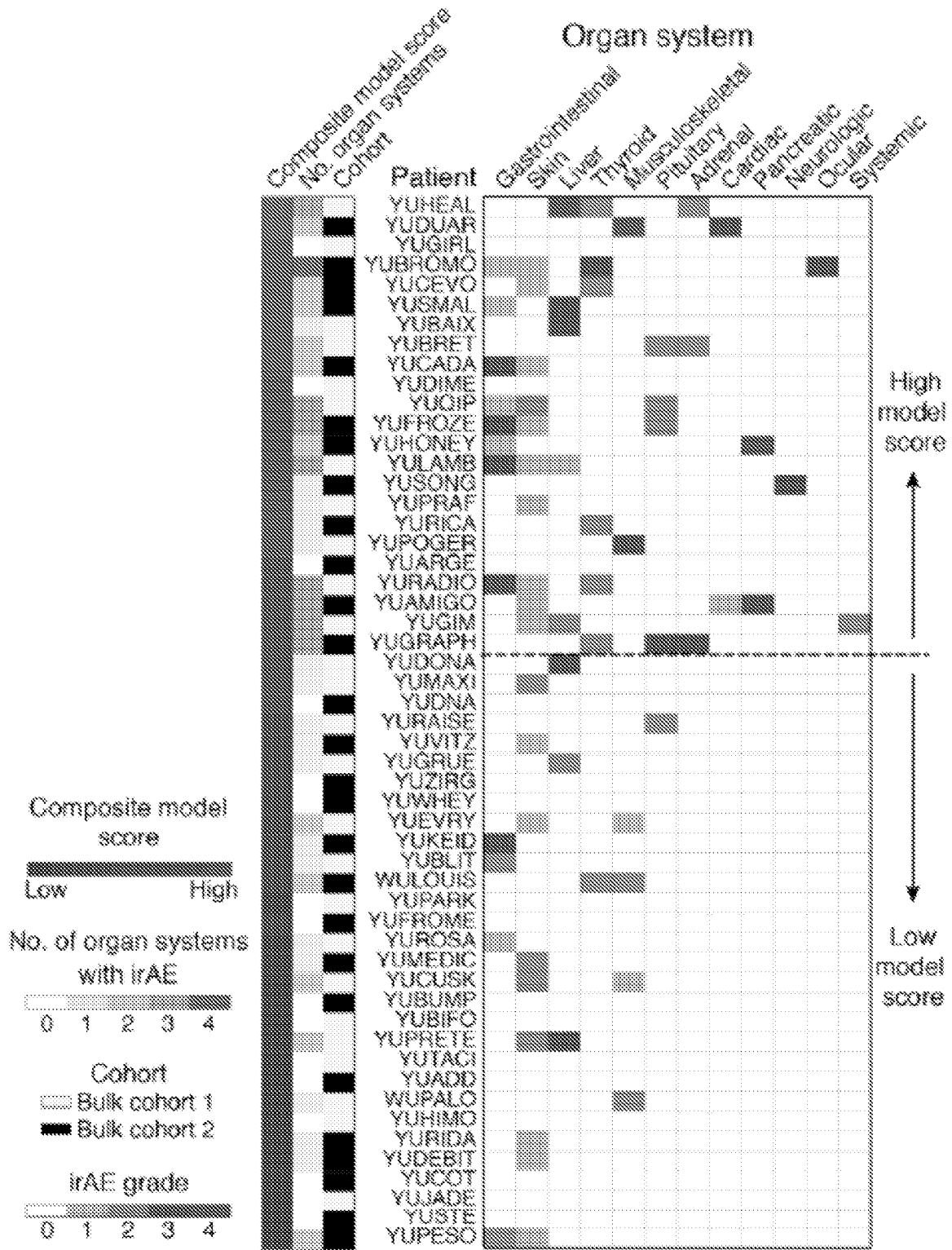


FIG. 131

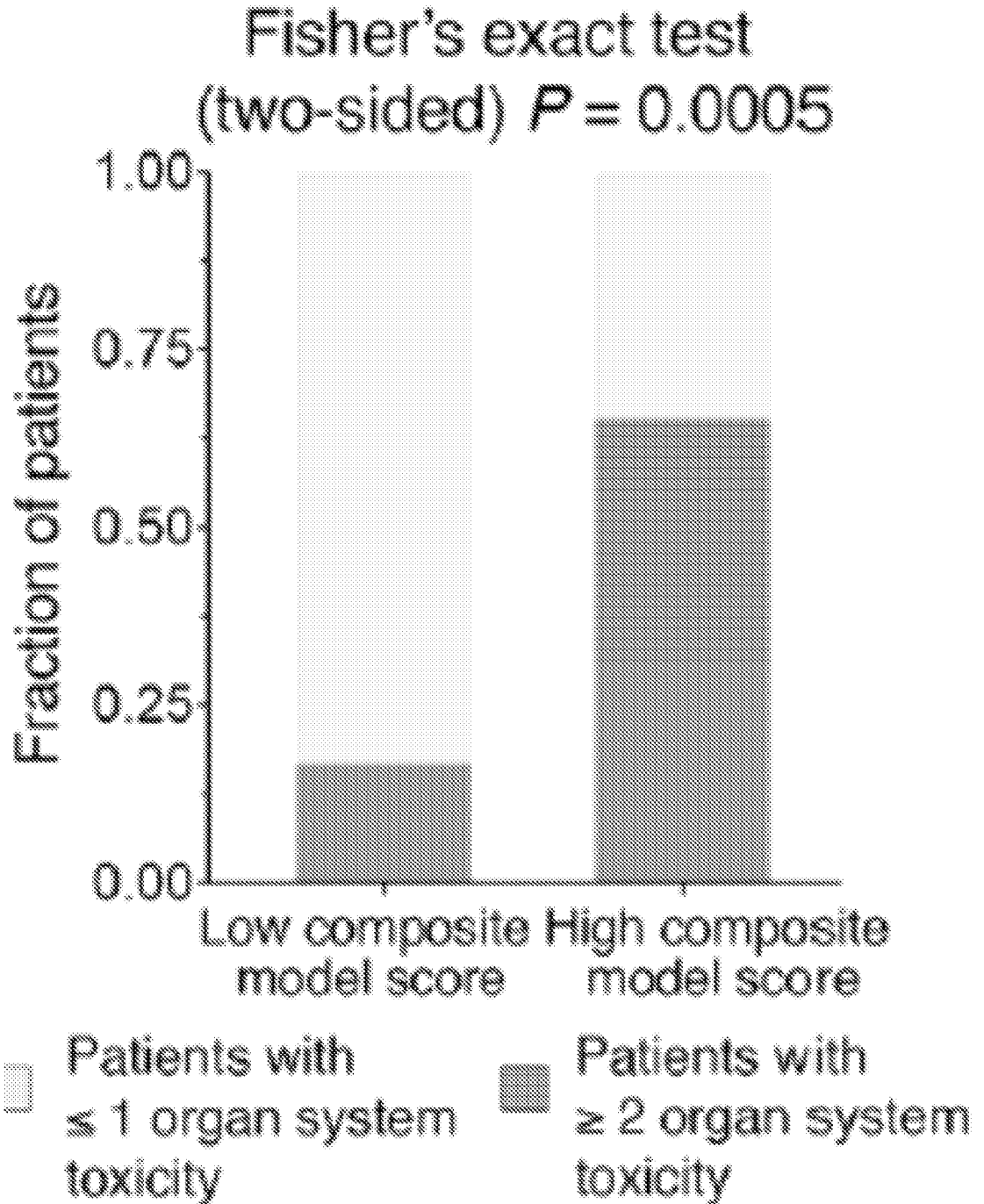


FIG. 13J

Pretreatment prediction of time to severe irAE in bulk cohort 2 (held-out validation, all patients)

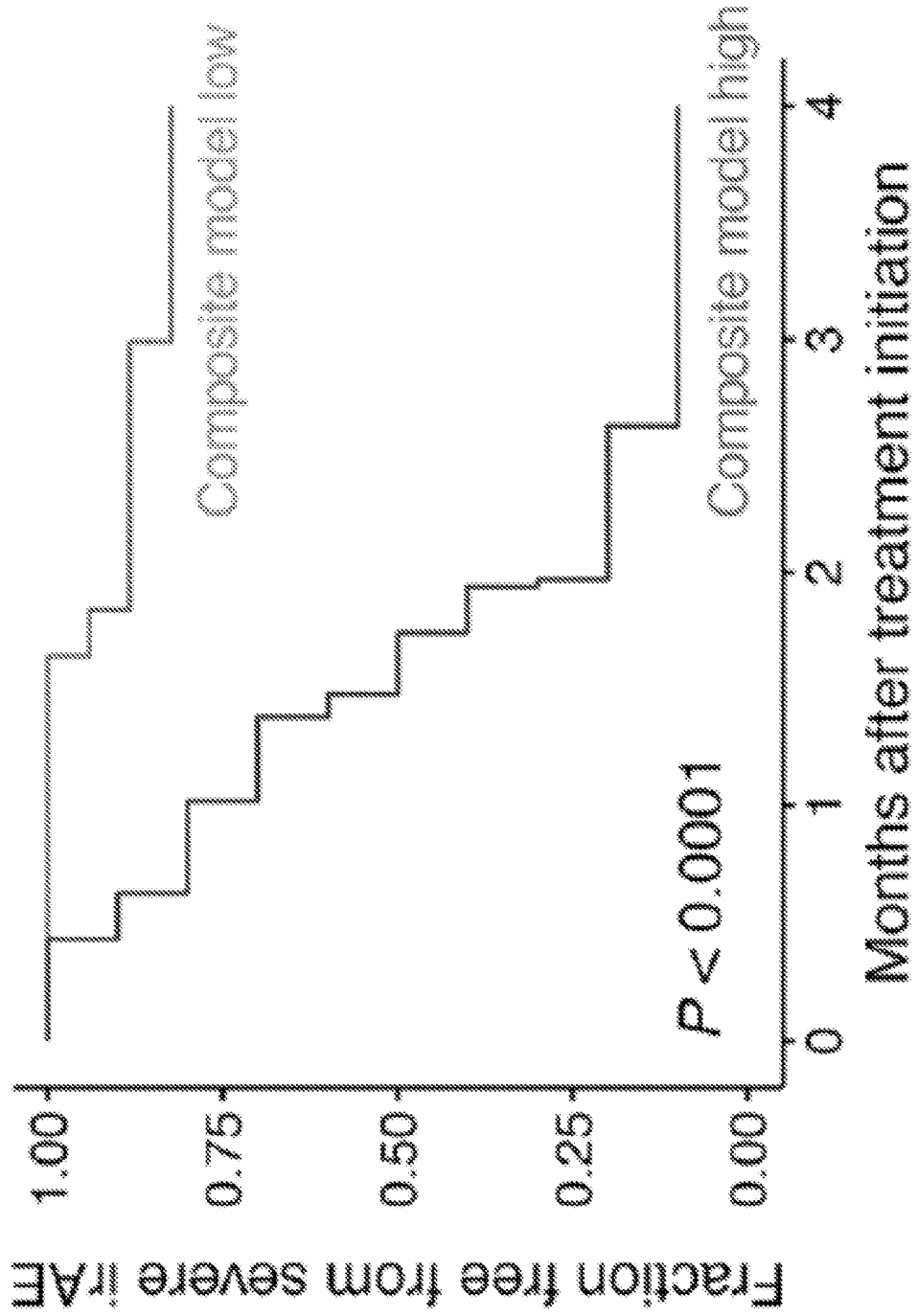


FIG. 14A

Pretreatment prediction of time to severe irAE in bulk cohort 2 (held-out validation, combination therapy)

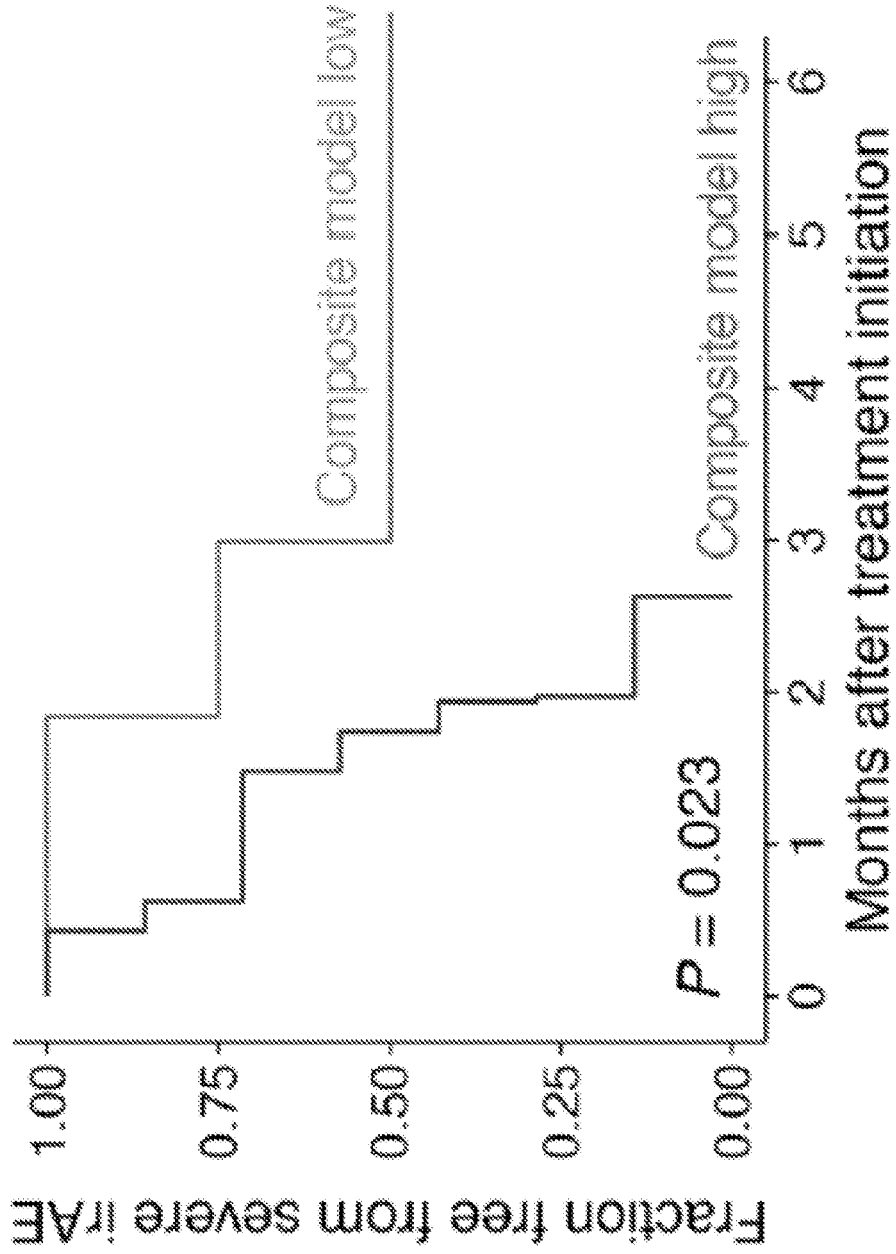


FIG. 14B

Pretreatment prediction of time to severe irAE in bulk cohort 2 (held-out validation, PD1 monotherapy)

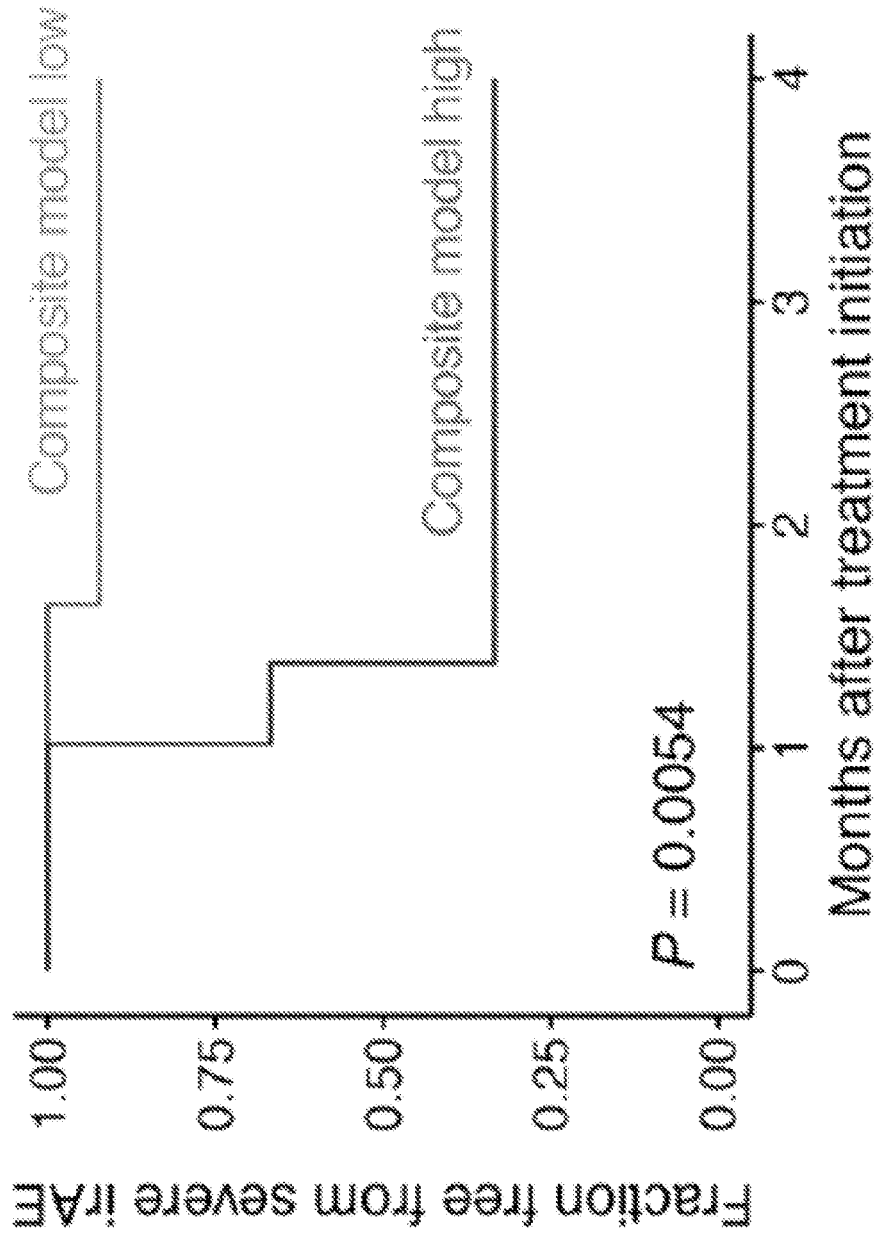


FIG. 14C

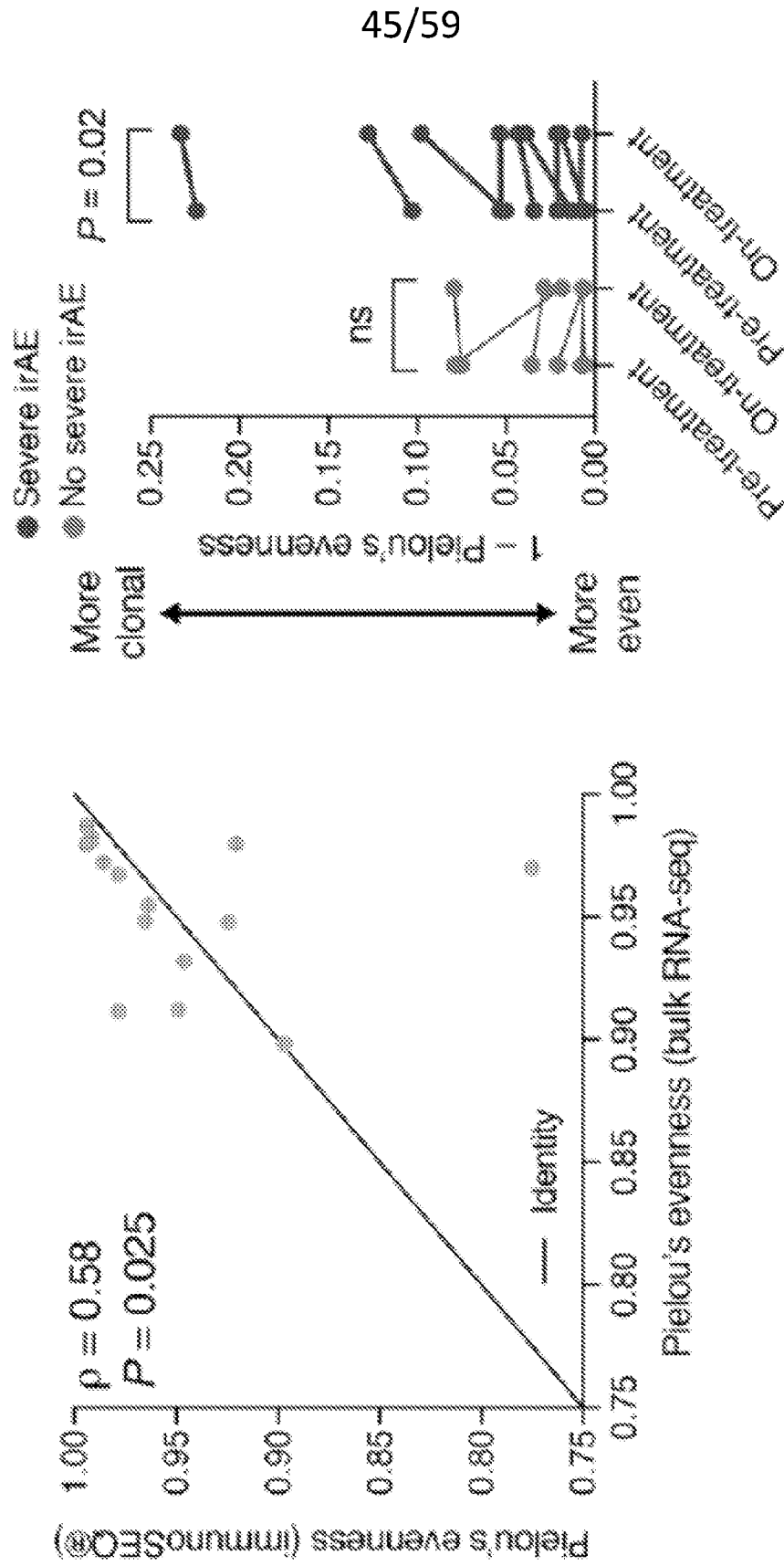


FIG. 15B

FIG. 15A

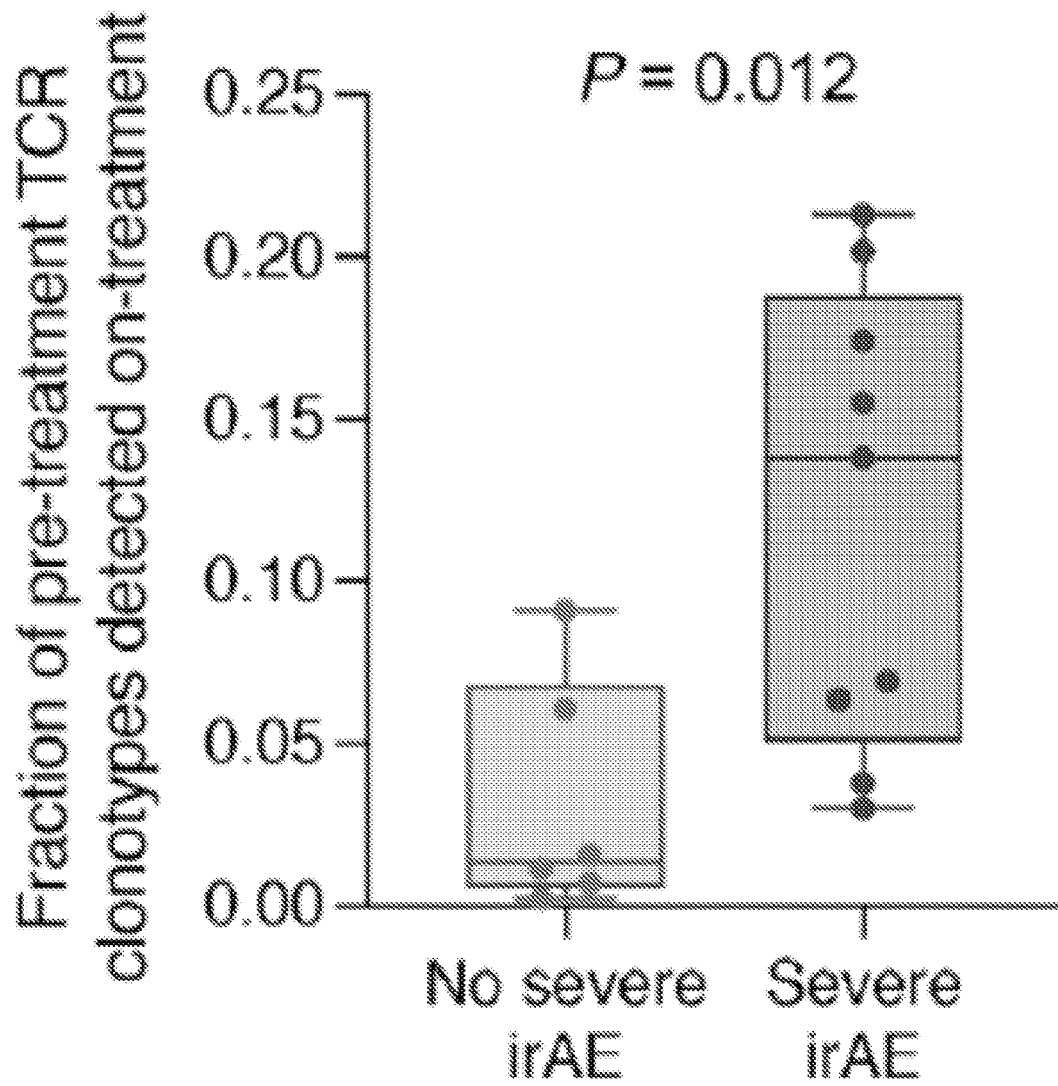


FIG. 15C

48/59

Preferential expansion in persistent CD4 T cell clonotypes

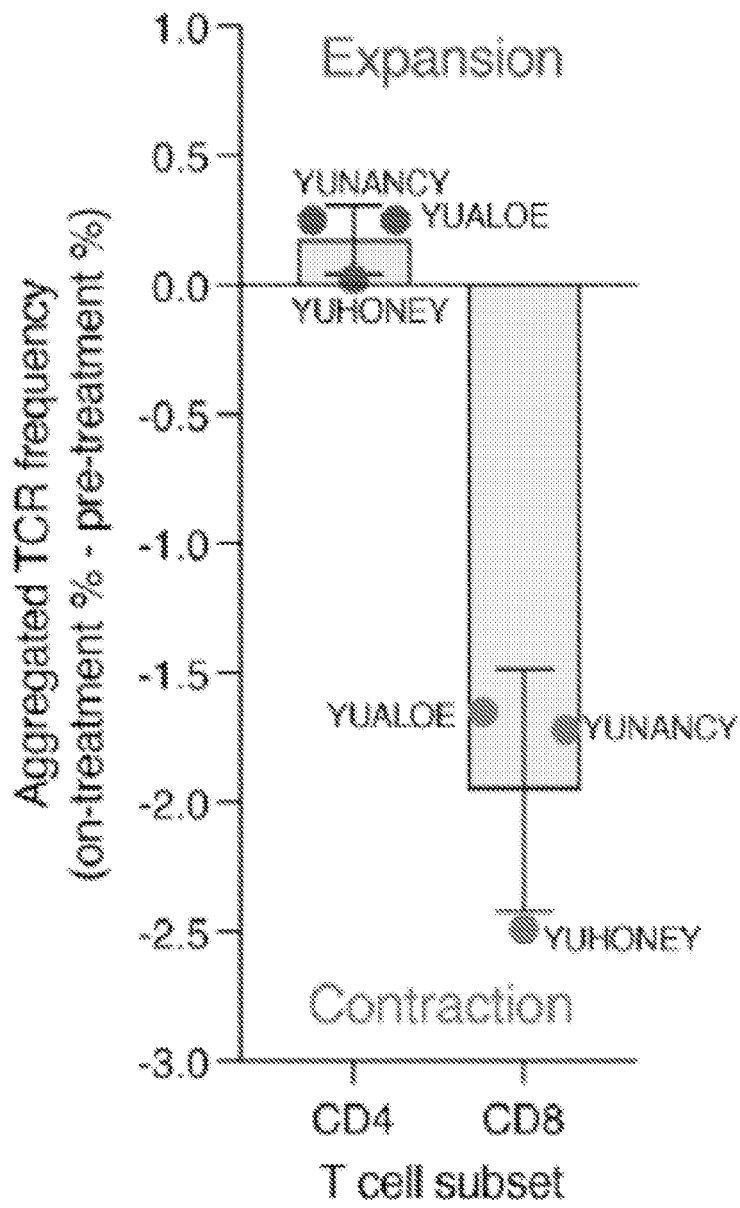


FIG. 15F

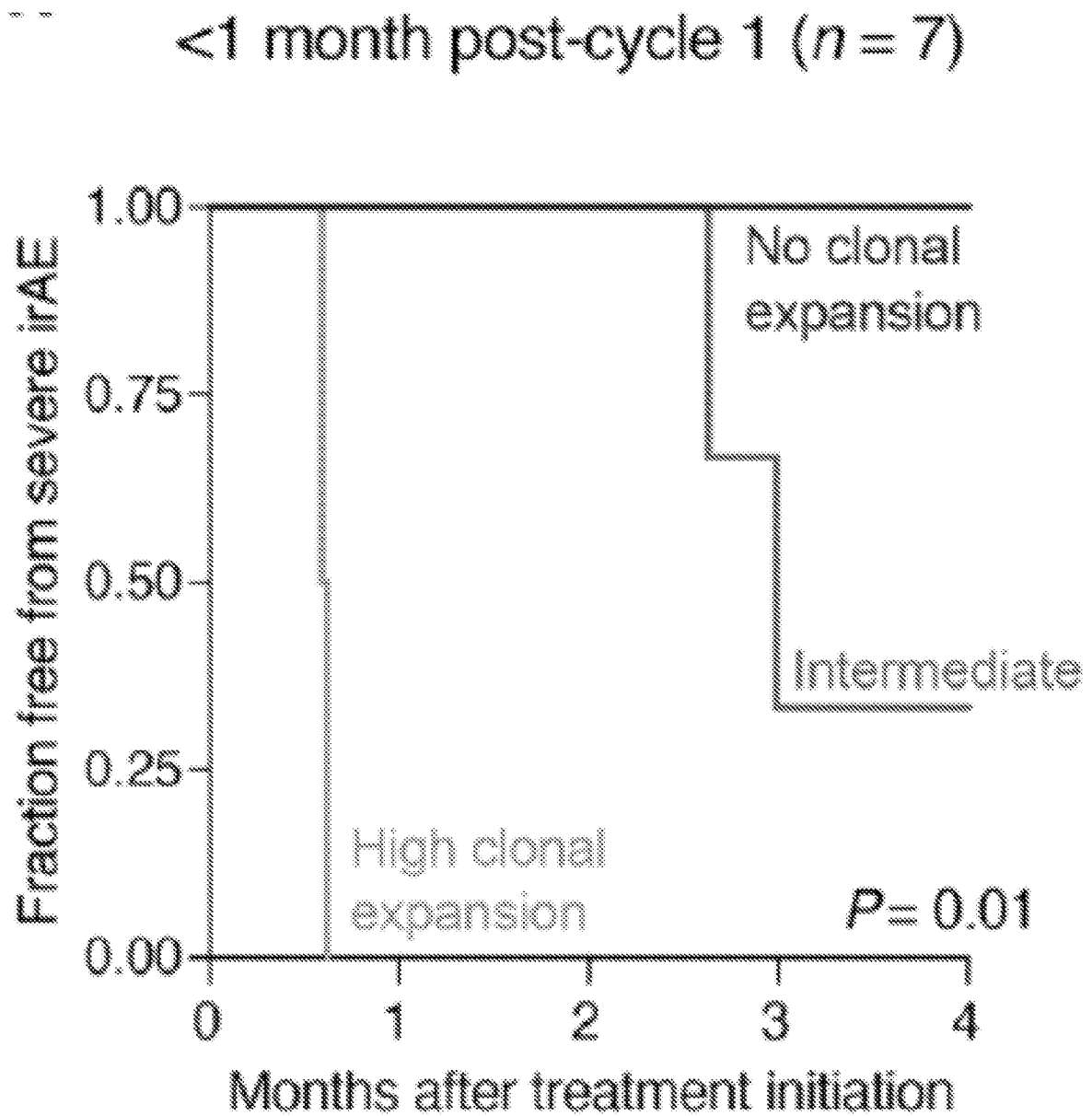


FIG. 15H

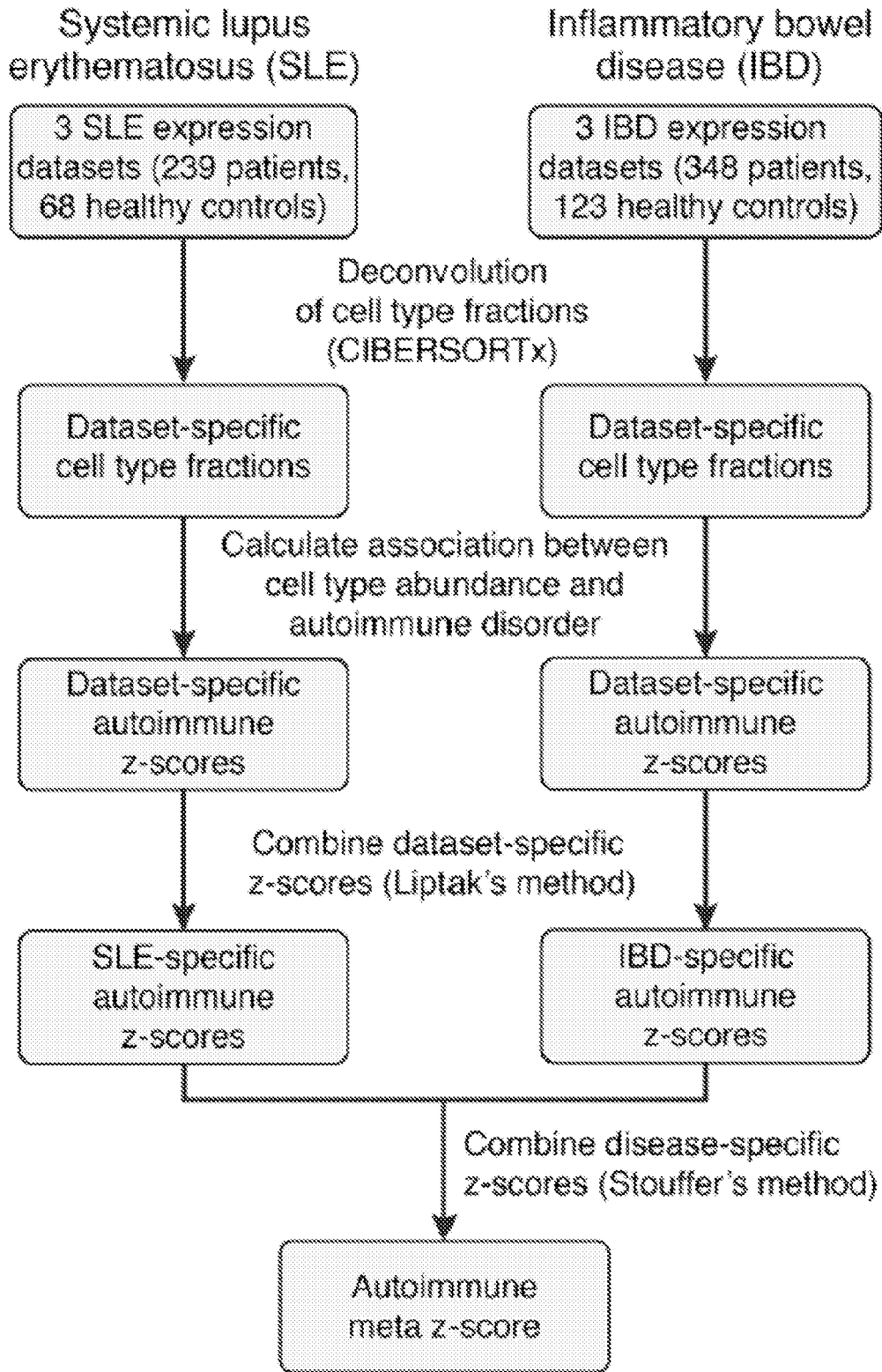


FIG. 16

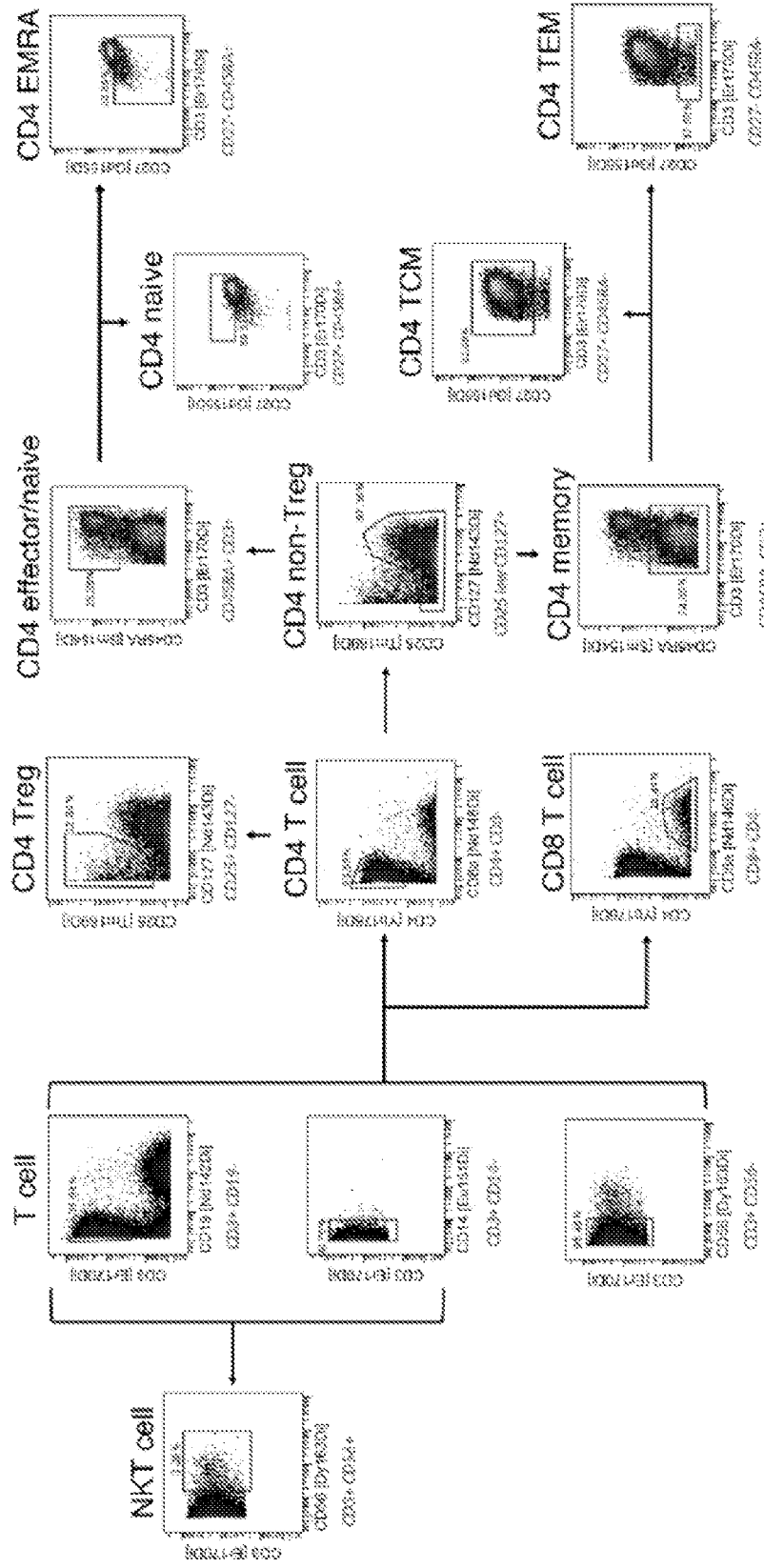


FIG. 17A

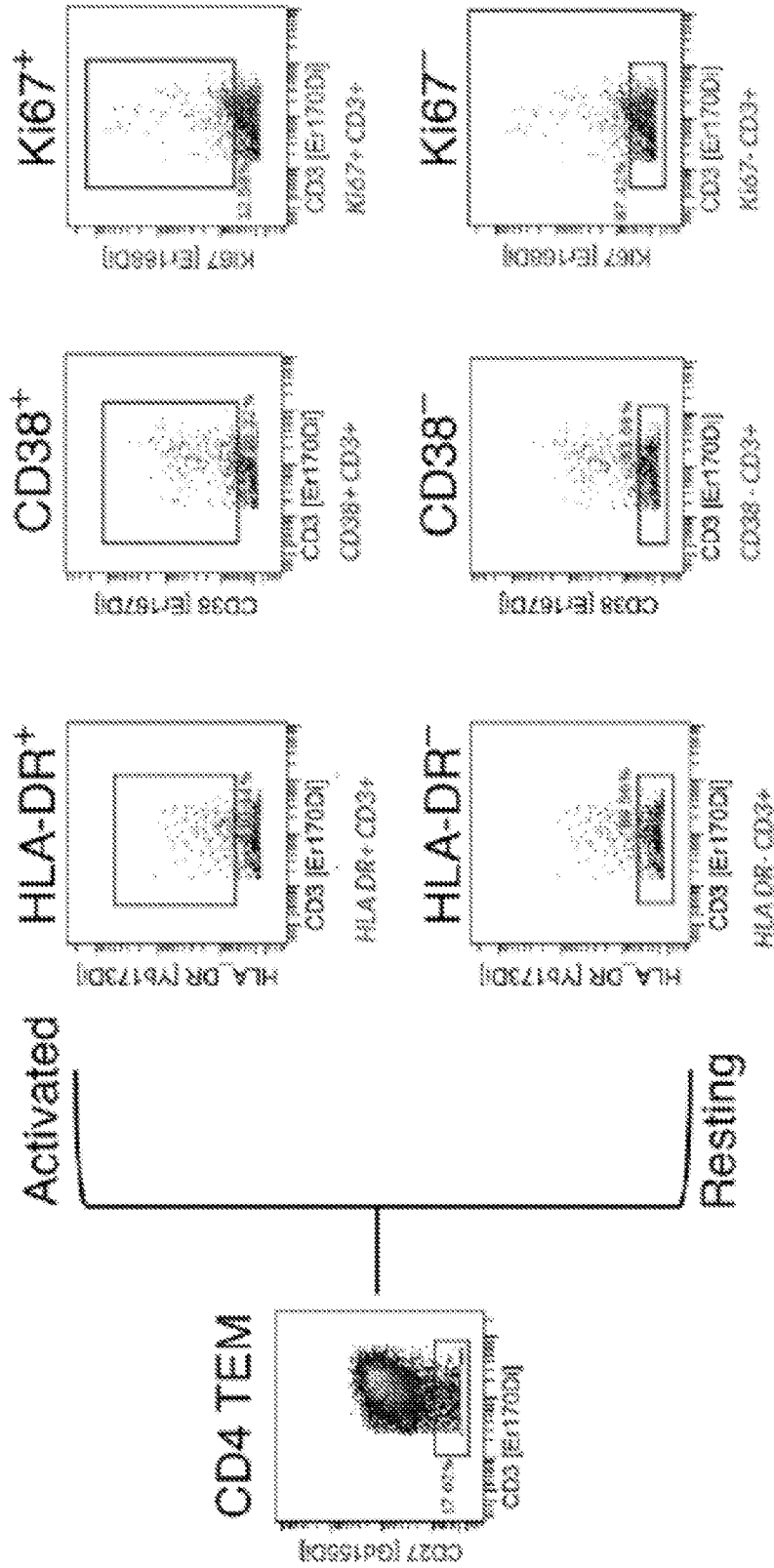


FIG. 17B

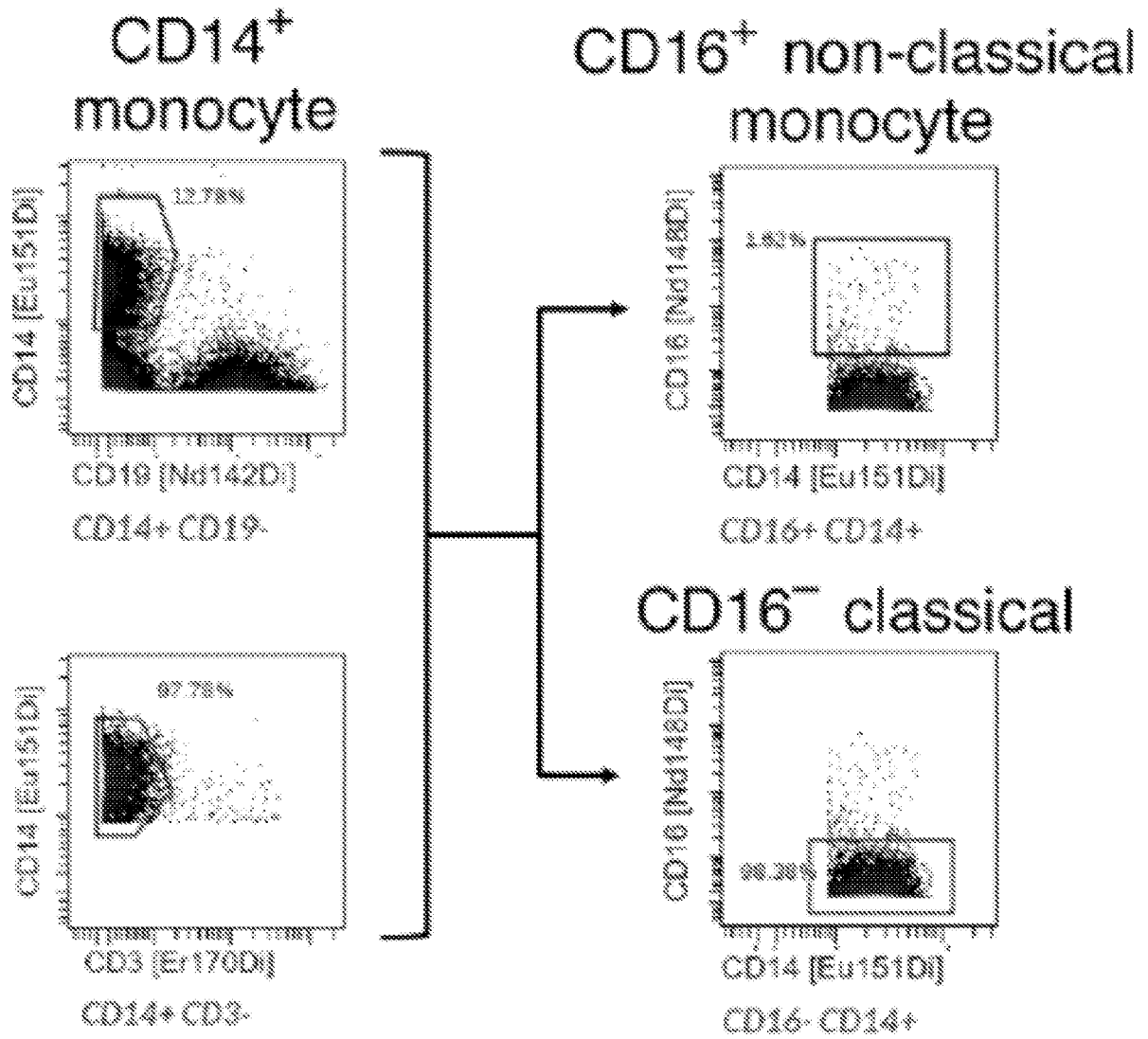


FIG. 17C

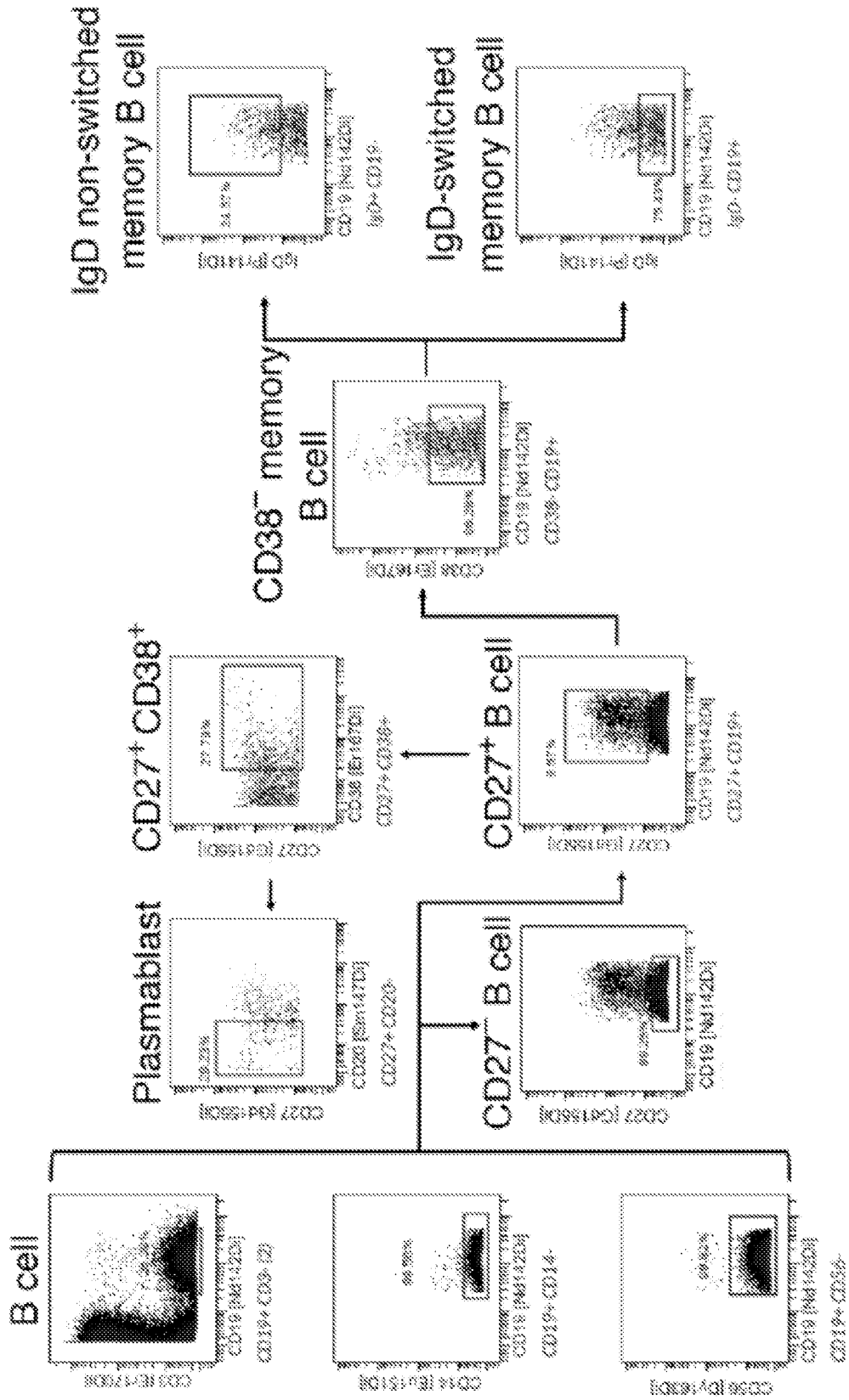


FIG. 17D

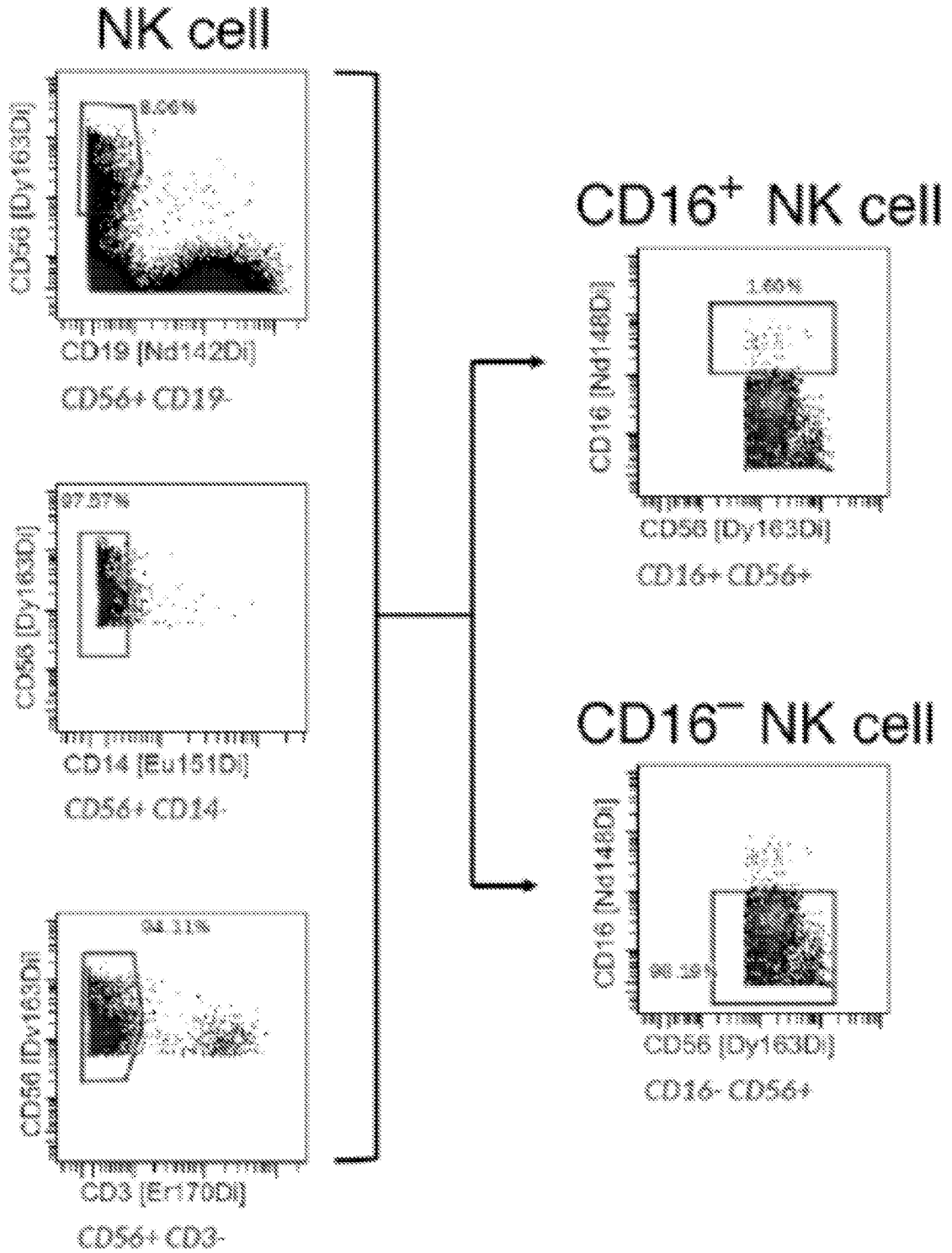


FIG. 17E

57/59

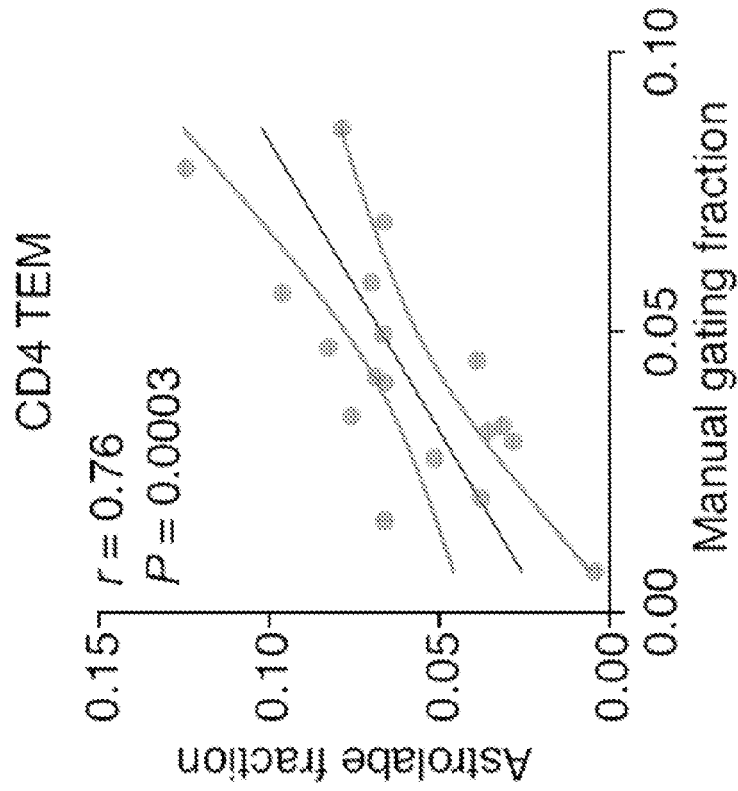


FIG. 18B

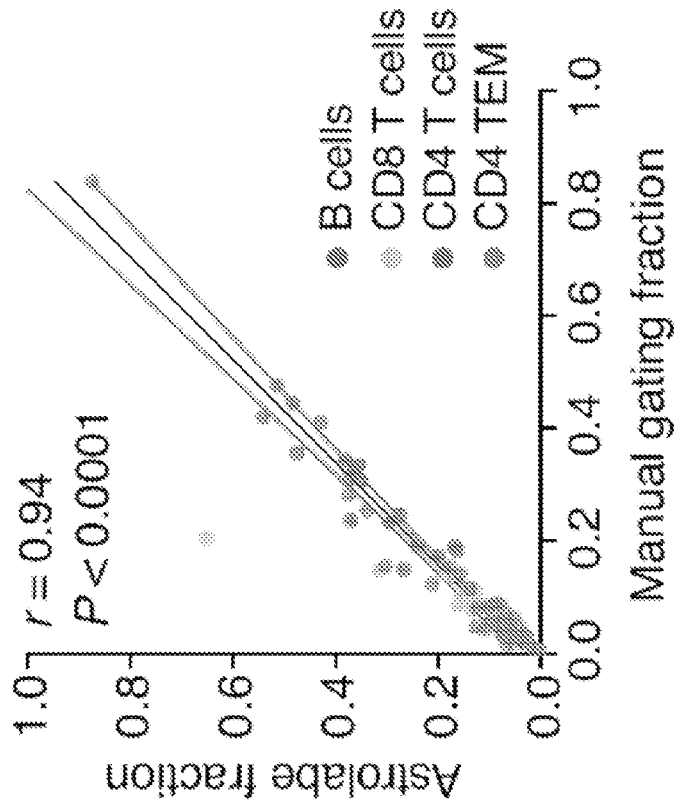


FIG. 18A

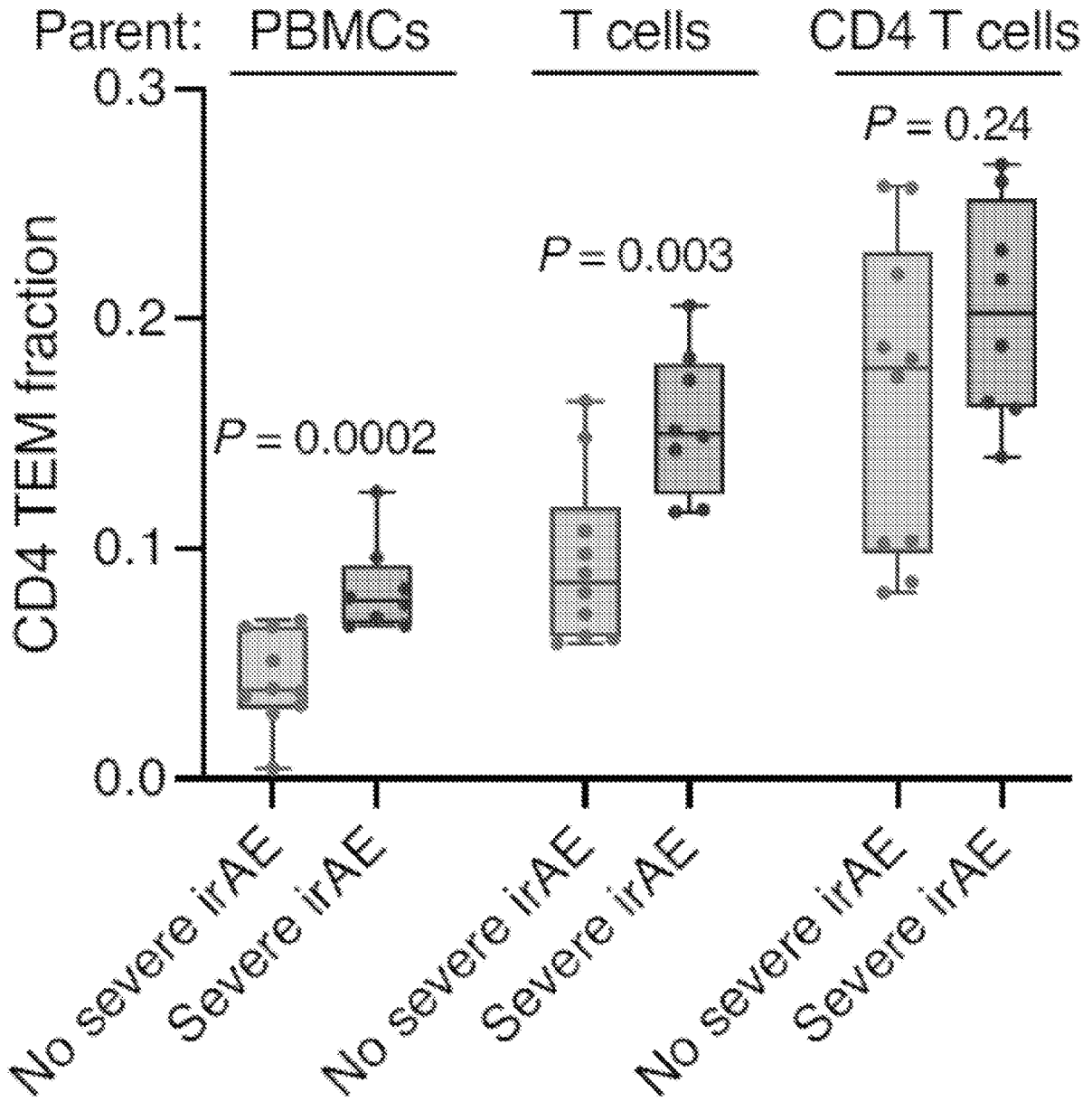


FIG. 18C

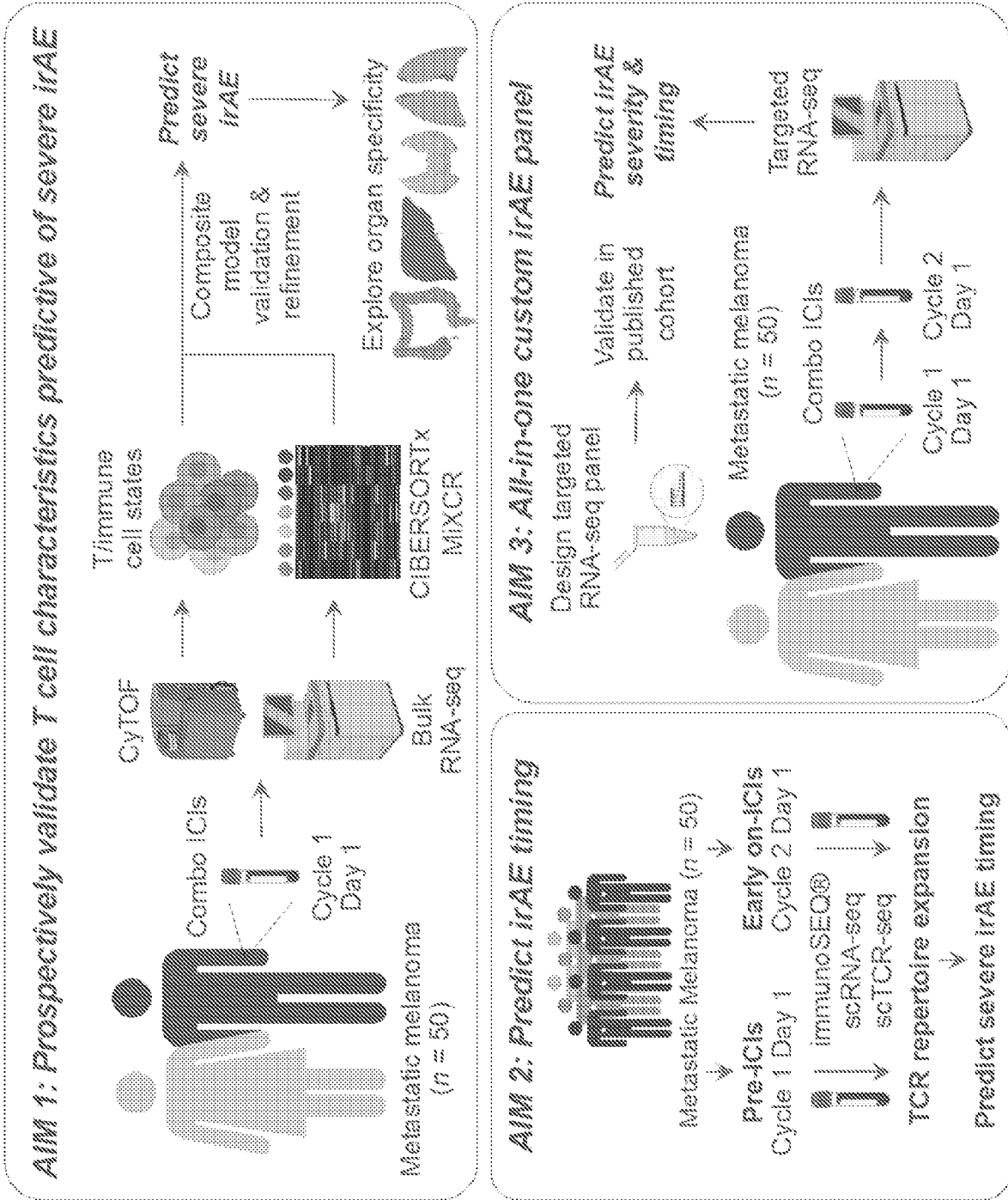


FIG. 19