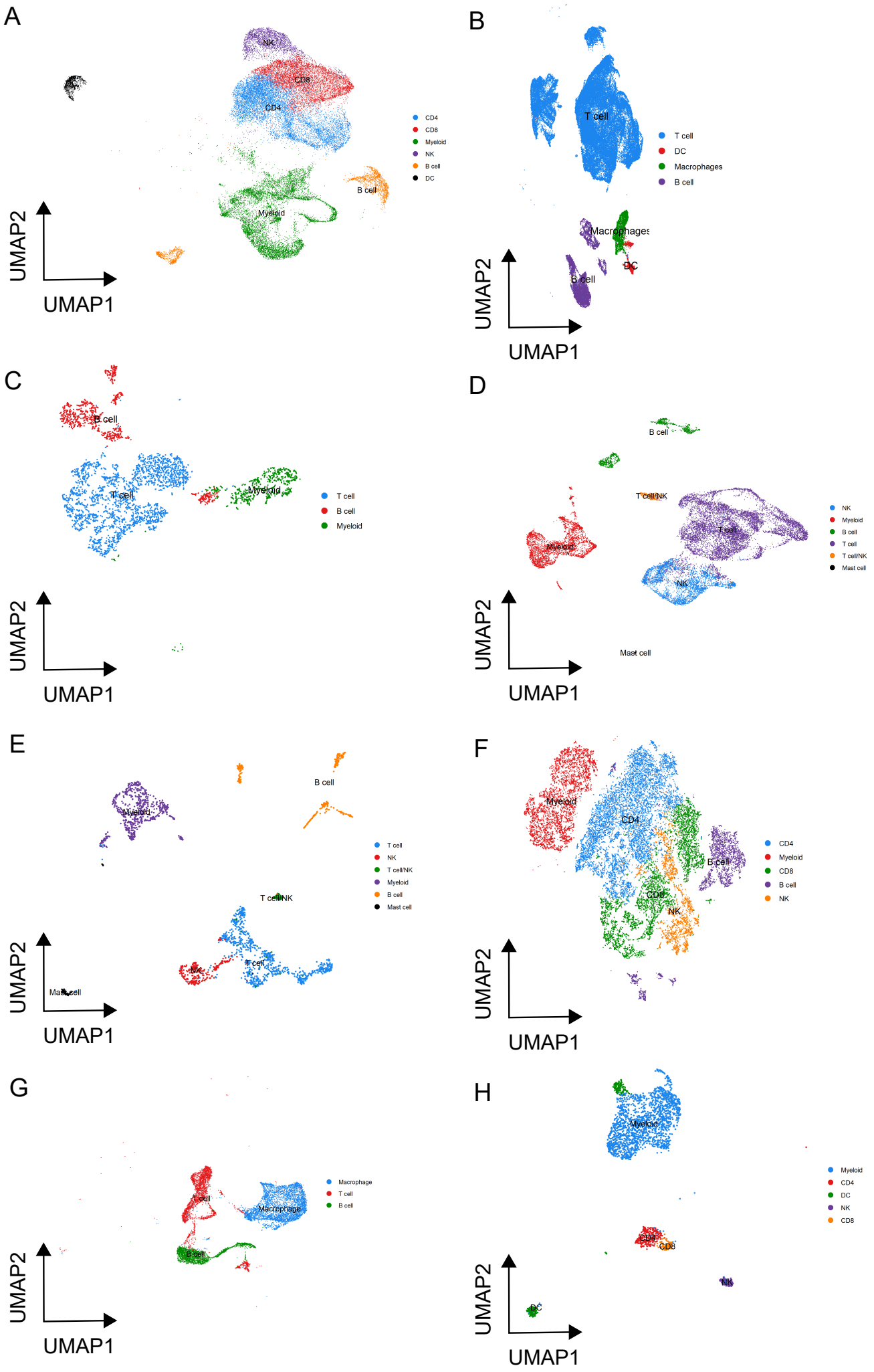
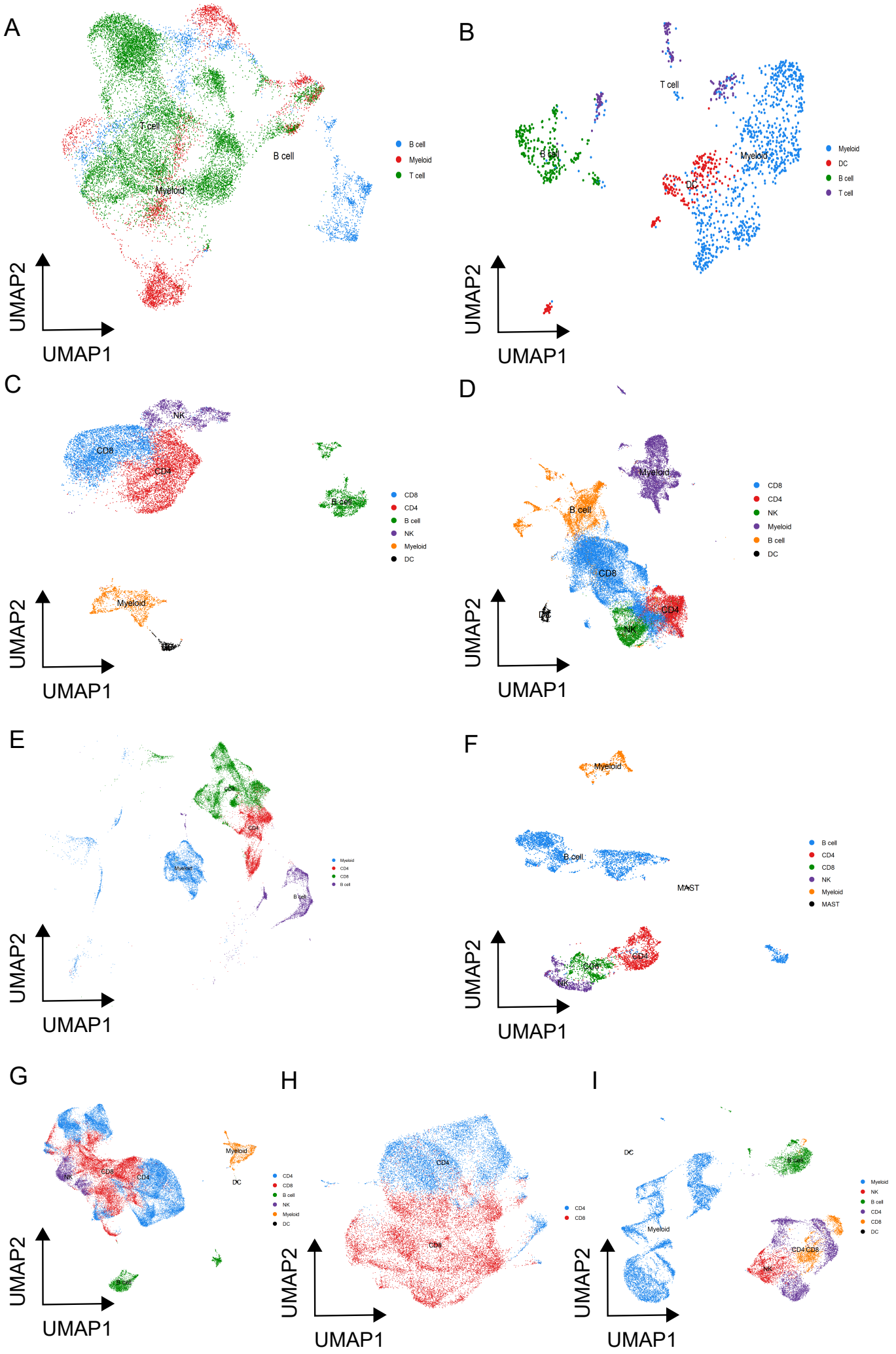
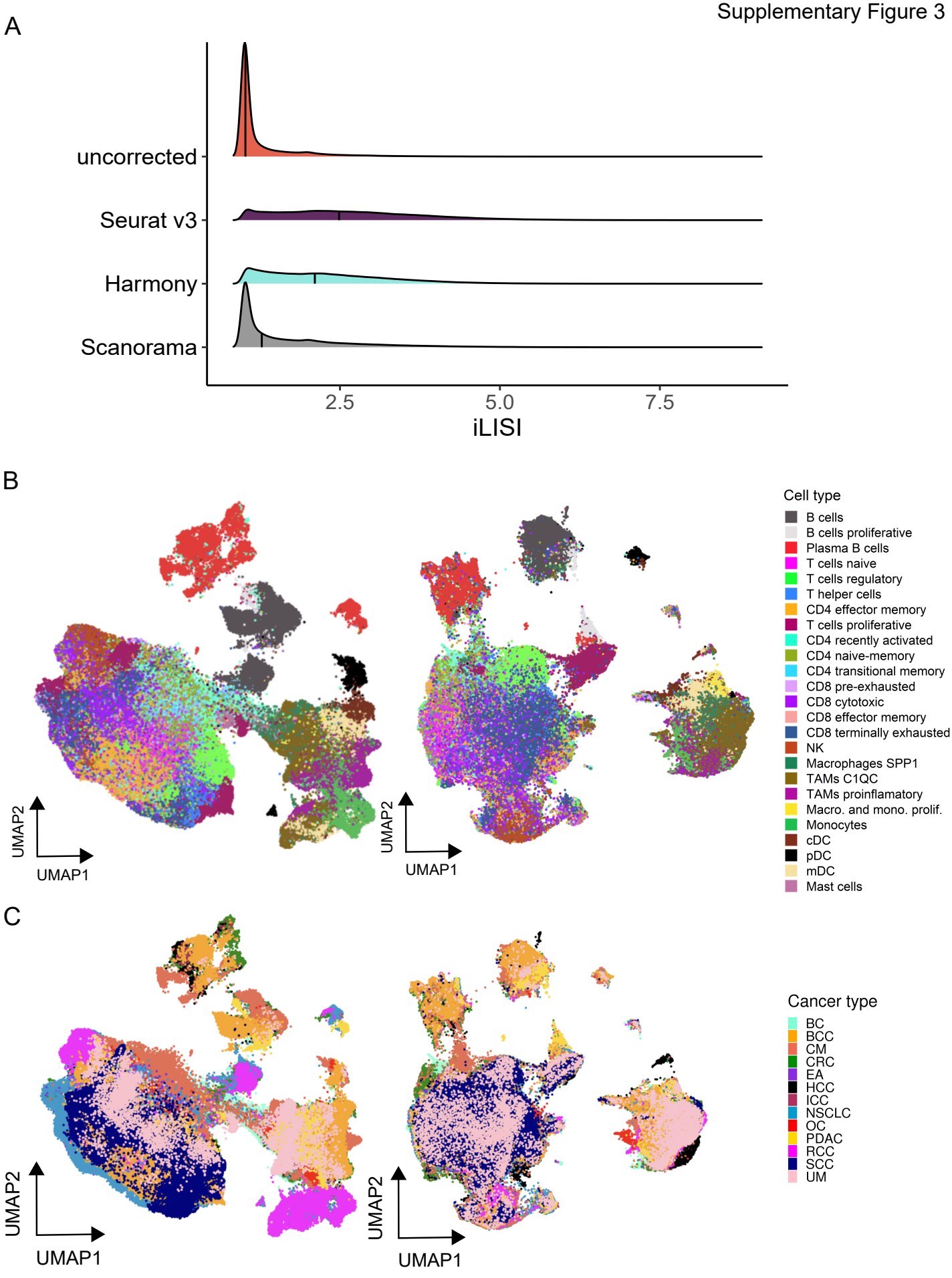

A
NK
CD8
CD4
B cell
Myeloid
CD4
CD8
Myeloid
NK
B cell
DC
UMAP2
UMAP1
B
T cell
Macrophages
DC
B cell
T cell
DC
Macrophages
B cell
UMAP2
UMAP1
C
B cell
T cell
Myeloid
T cell
B cell
Myeloid
UMAP2
UMAP1
D
B cell
T cell/NK
Myeloid
T cell
NK
Mast cell
NK
Myeloid
B cell
T cell
T cell/NK
Mast cell
UMAP2
UMAP1
E
Myeloid
B cell
T cell/NK
NK
T cell
Mast cell
T cell
NK
T cell/NK
Myeloid
B cell
Mast cell
UMAP2
UMAP1
F
Myeloid
CD4
B cell
CD8
NK
CD4
Myeloid
CD8
B cell
NK
UMAP2
UMAP1
G
T cell
Macrophage
B cell
Macrophage
T cell
B cell
UMAP2
UMAP1
H
Myeloid
CD4
CD8
NK
DC
Myeloid
CD4
DC
NK
CD8
UMAP2
UMAP1

**Supplementary Figure 1 | Original annotation of the datasets prior to integration.** Azizi et al. 2018 (**A**), Yost et al. 2019 (**B**), Lichun et al. 2019 (**C**), Zhang et al. 2019 (**D, E**), Lee et al. 2020 (**F**), Peng et al. 2019 (**G**), Schelker et al. 2017 (**H**).


A
T cell
B cell
Myeloid
B cell
Myeloid
T cell
UMAP2
UMAP1
B
T cell
B cell
Myeloid
DC
Myeloid
DC
B cell
T cell
UMAP2
UMAP1
C
NK
CD8
CD4
B cell
CD8
CD4
B cell
NK
Myeloid
DC
Myeloid
UMAP2
UMAP1
D
Myeloid
B cell
CD8
DC
CD4
NK
CD8
CD4
NK
Myeloid
B cell
DC
UMAP2
UMAP1
E
CD8
CD4
Myeloid
B cell
Myeloid
CD4
CD8
B cell
UMAP2
UMAP1
F
Myeloid
B cell
MAST
CD4
CD8
NK
B cell
CD4
CD8
NK
Myeloid
MAST
UMAP2
UMAP1
G
NK
CD8
CD4
Myeloid
DC
B cell
CD4
CD8
B cell
NK
Myeloid
DC
UMAP2
UMAP1
H
CD4
CD8
CD4
CD8
UMAP2
UMAP1
I
DC
Myeloid
CD4 CD8
NK
Myeloid
NK
B cell
CD4
CD8
DC
UMAP2
UMAP1

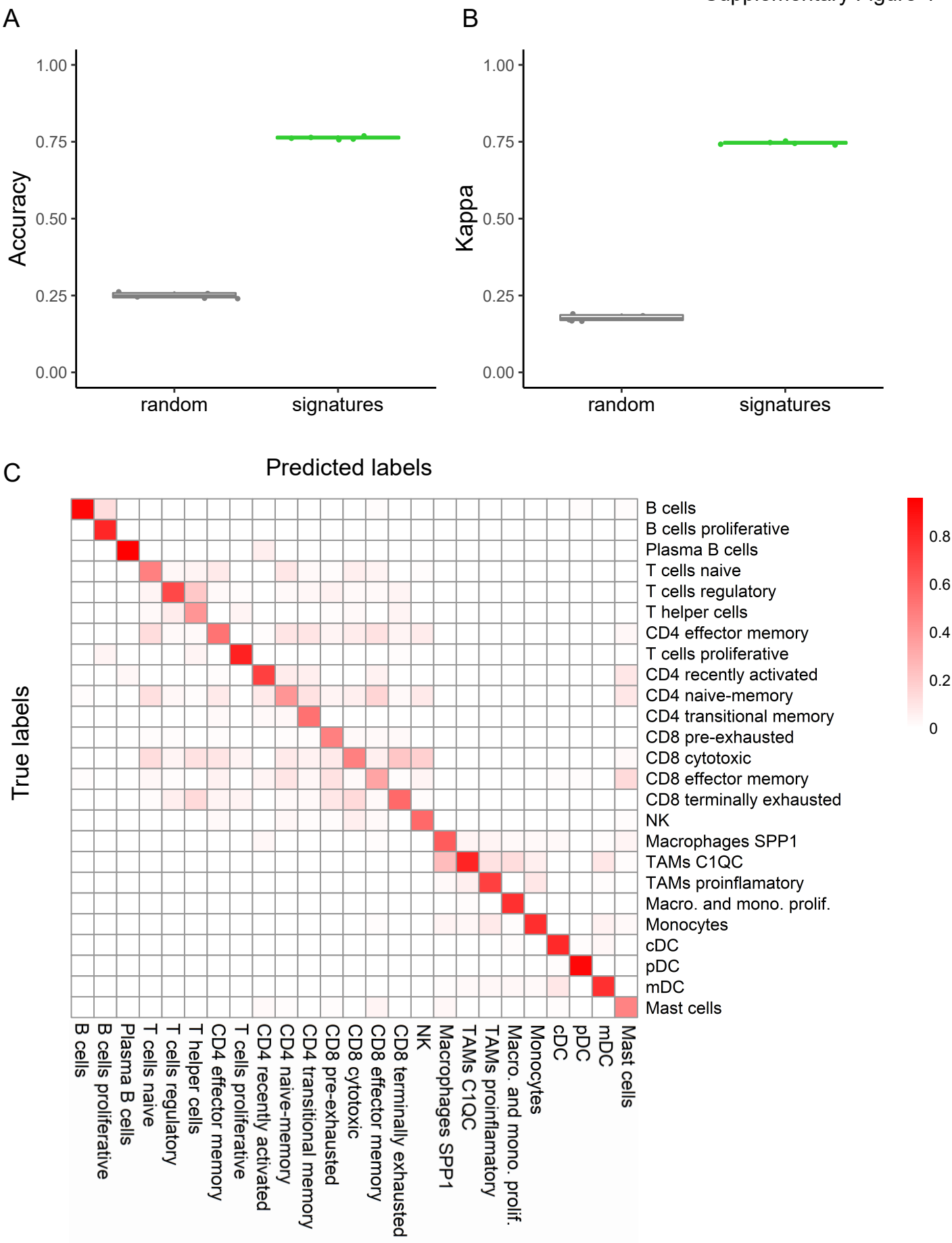
**Supplementary Figure 2 | Original annotation of the datasets prior to integration.** Lambrechts et al. 2018 (**A**), Lavin et al. 2017 (**B**), Sade-Feldman et al. 2018 (**C**), Li et al. 2019 (**D**), Durante et al. 2020 (**E**), Wu et al. 2020 CRC (**F**), NSCLC (**G**), EA (**H**), RCC (**I**).

Supplementary Figure 3



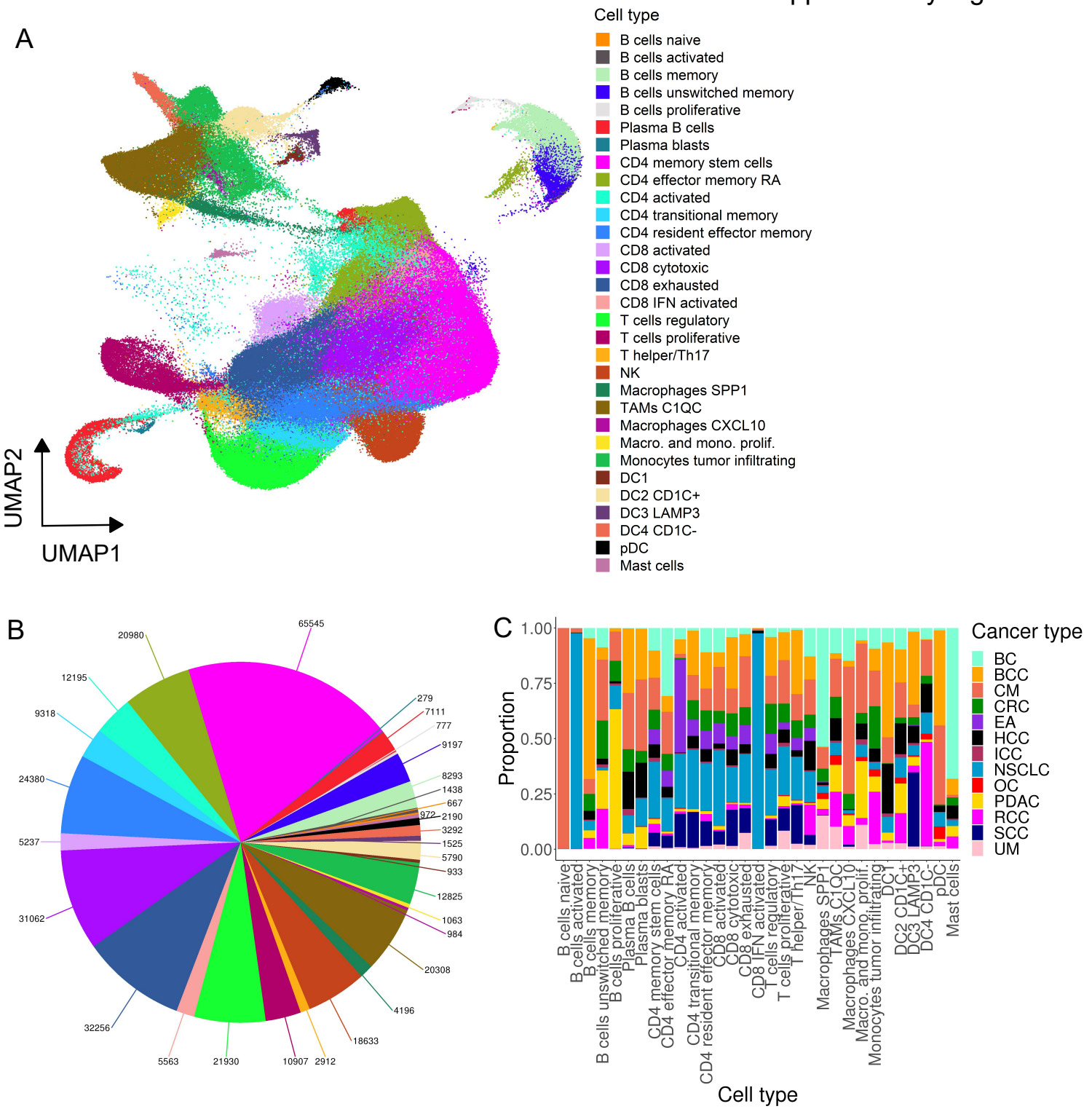
**Supplementary Figure 3 | Comparing data integration tools.** (**A**) iLISI score to quantify the integration success of the three methods, Seurat v3, Harmony and Scanorama, against the uncorrected dataset. (**B,C**) UMAP representation of the datasets integrated with Scanorama (left) or Harmony (right) and colored by celltype/state as identified in the Seurat integration (**B**) or by cancer type (**C**).

Supplementary Figure 4



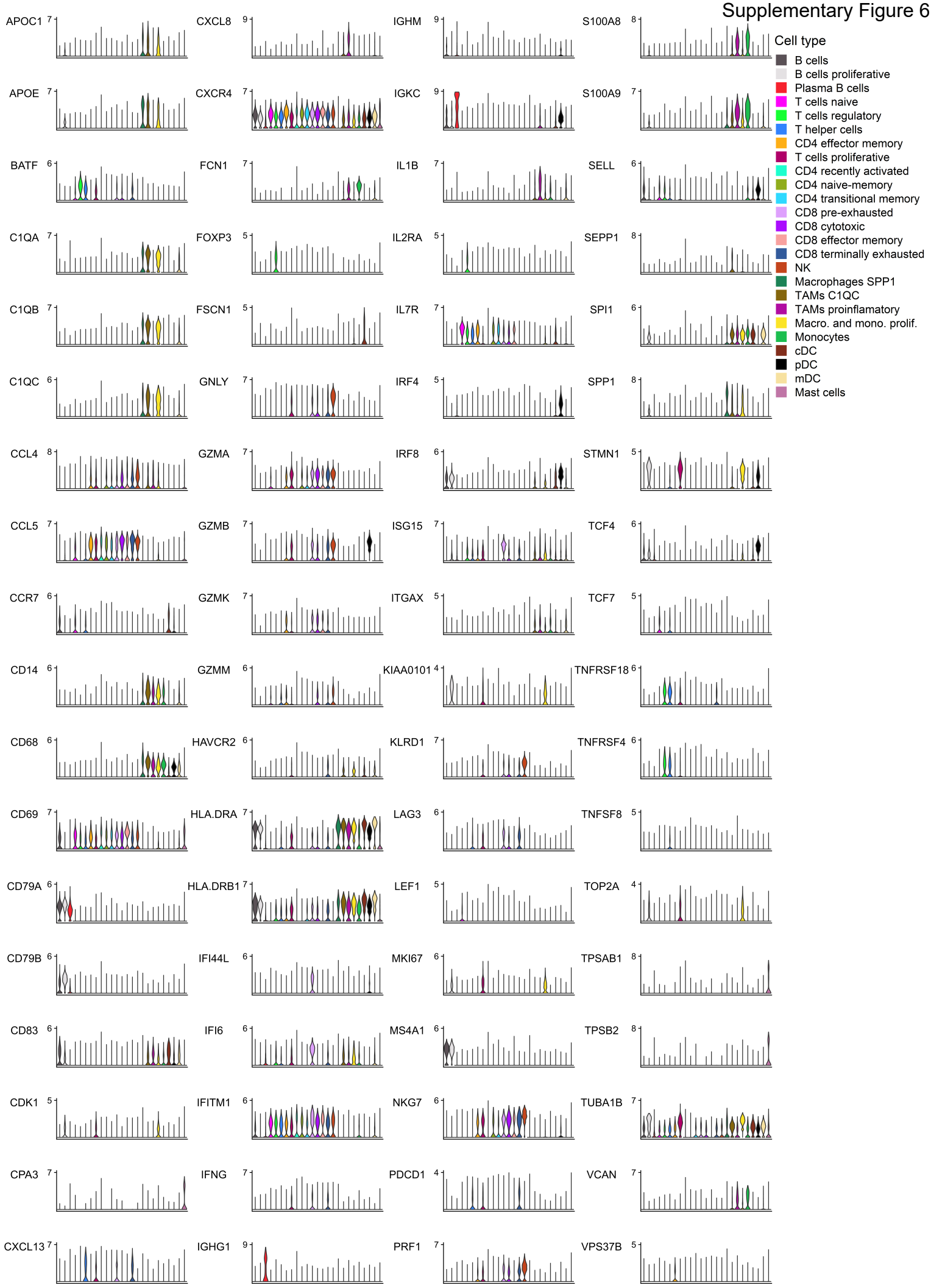
**Supplementary Figure 4 | Cell type prediction using a random forest classifier.** (**A**) Accuracy and (**B**) kappa statistic using cell type-specific or random signatures as features. (**C**) Confusion matrix displaying the average probability that a cell of a cell type X (rows) is classified as a cell type Y (columns).

# Supplementary Figure 5



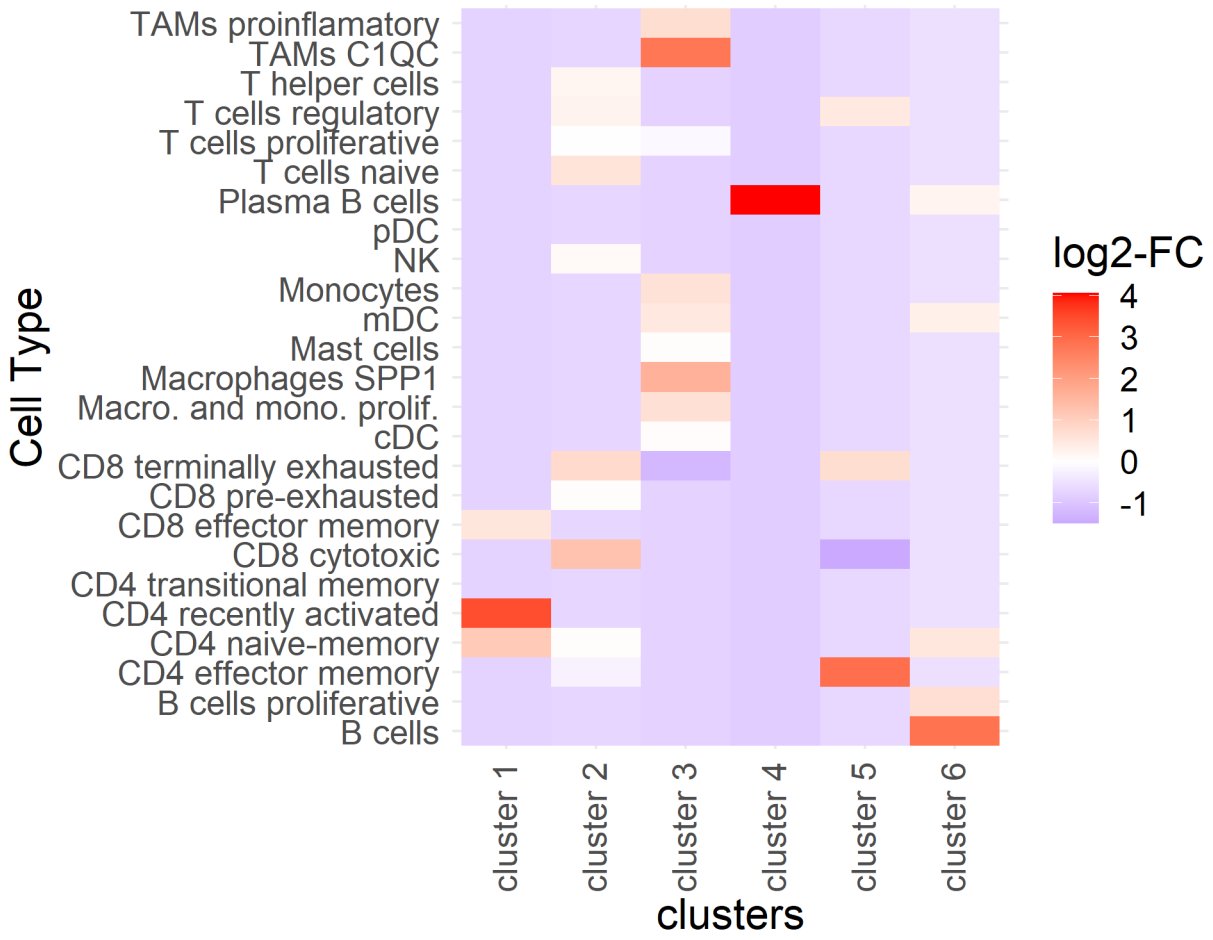
**Supplementary Figure 5 | Characterization of higher resolution clusters**. (**A**) UMAP of 317,111 immune cells from 13 cancer types colored by their 31 annotated cell types and states. (**B**) Total number of cells of each immune cell type/state; color code as in (A). (**C**) Proportions of cancer type for each cell type/state. (**D**) Number of unique patients representing each cell type/state in the tumor immune cell atlas.

Supplementary Figure 6



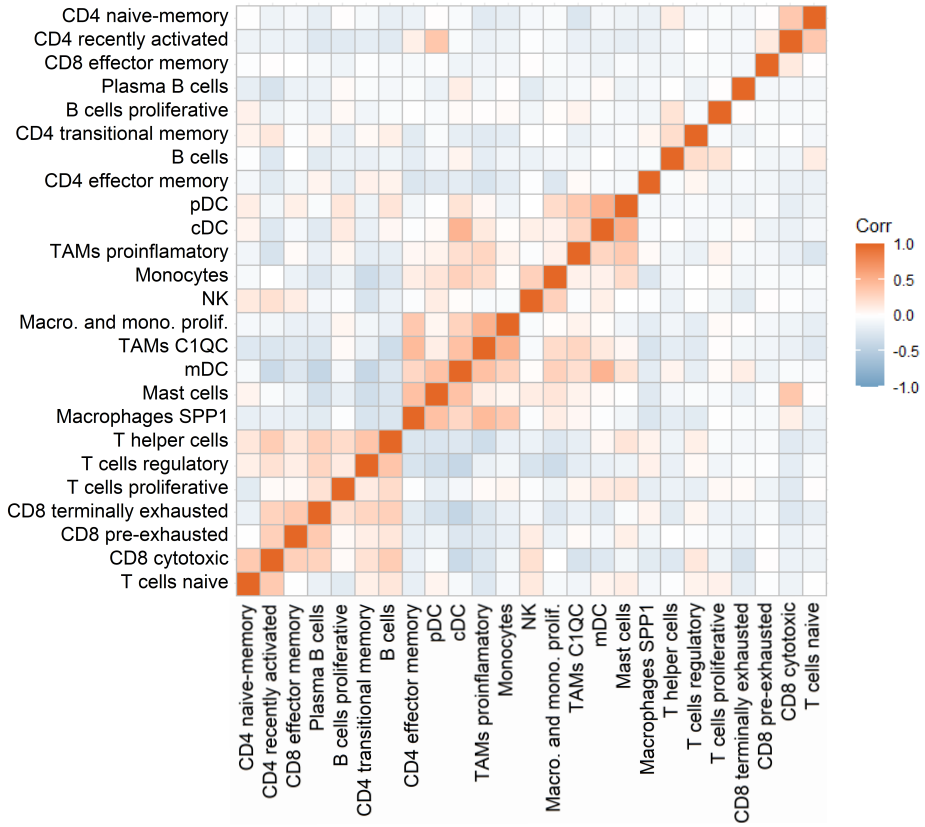
**Supplementary Figure 6 | Expression of selected markers across all cell types.**

Supplementary Figure 7

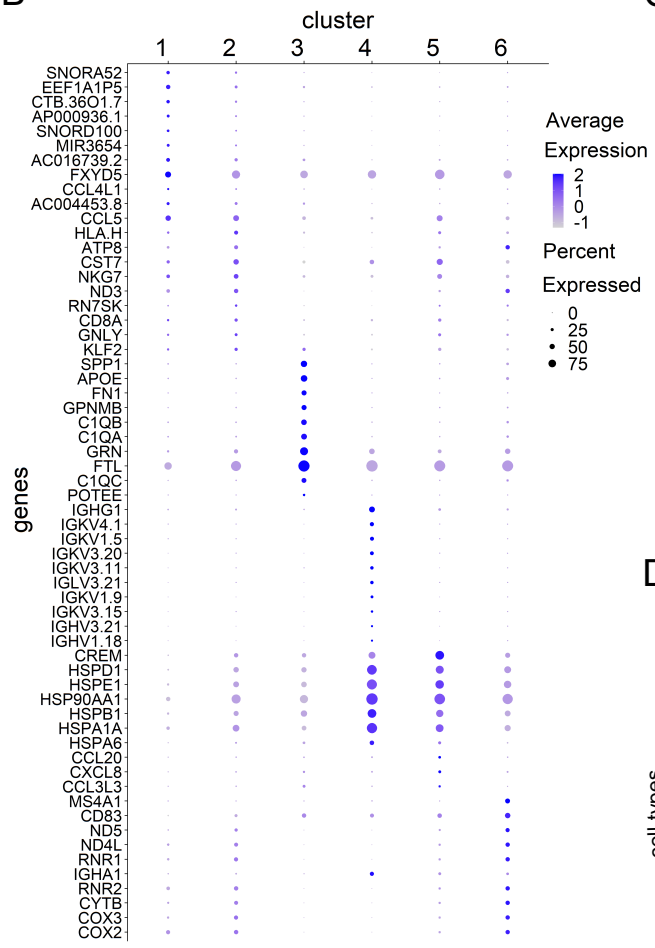


**Supplementary Figure 7 | scCODA results.** Heatmap comparing the log2 fold-change calculated with scCODA for all clusters and cell types.

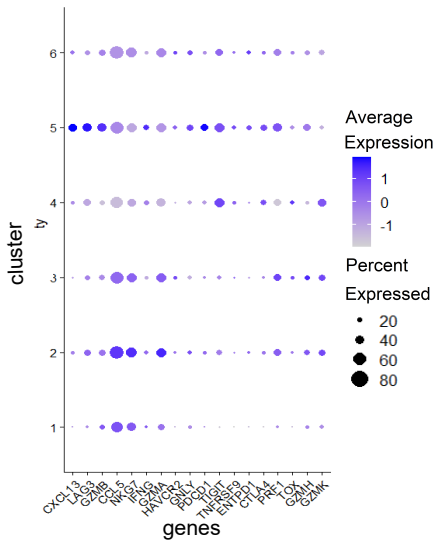
A


CD4 naive-memory
CD4 recently activated
CD8 effector memory
Plasma B cells
B cells proliferative
CD4 transitional memory
B cells
CD4 effector memory
pDC
cDC
TAMs proinflamatory
Monocytes
NK
Macro. and mono. prolif.
TAMs C1QC
mDC
Mast cells
Macrophages SPP1
T helper cells
T cells regulatory
T cells proliferative
CD8 terminally exhausted
CD8 pre-exhausted
CD8 cytotoxic
T cells naive
Corr
1.0
0.5
0.0
-0.5
-1.0

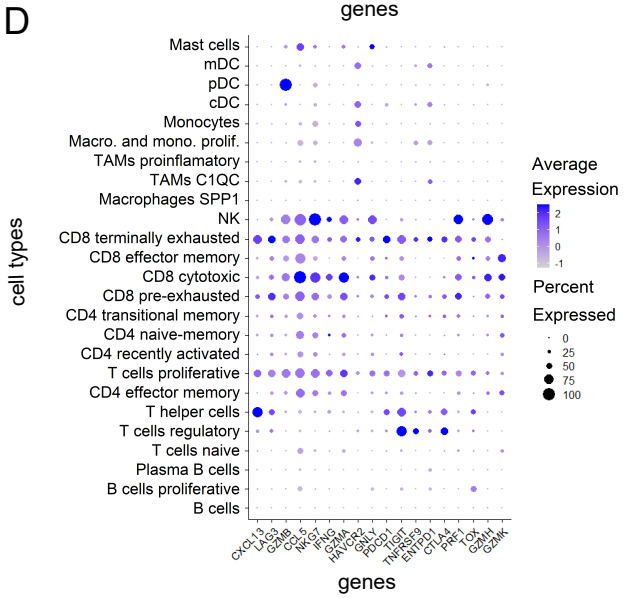
B


cluster
1
2
3
4
5
6
genes
Average Expression
Percent Expressed

C

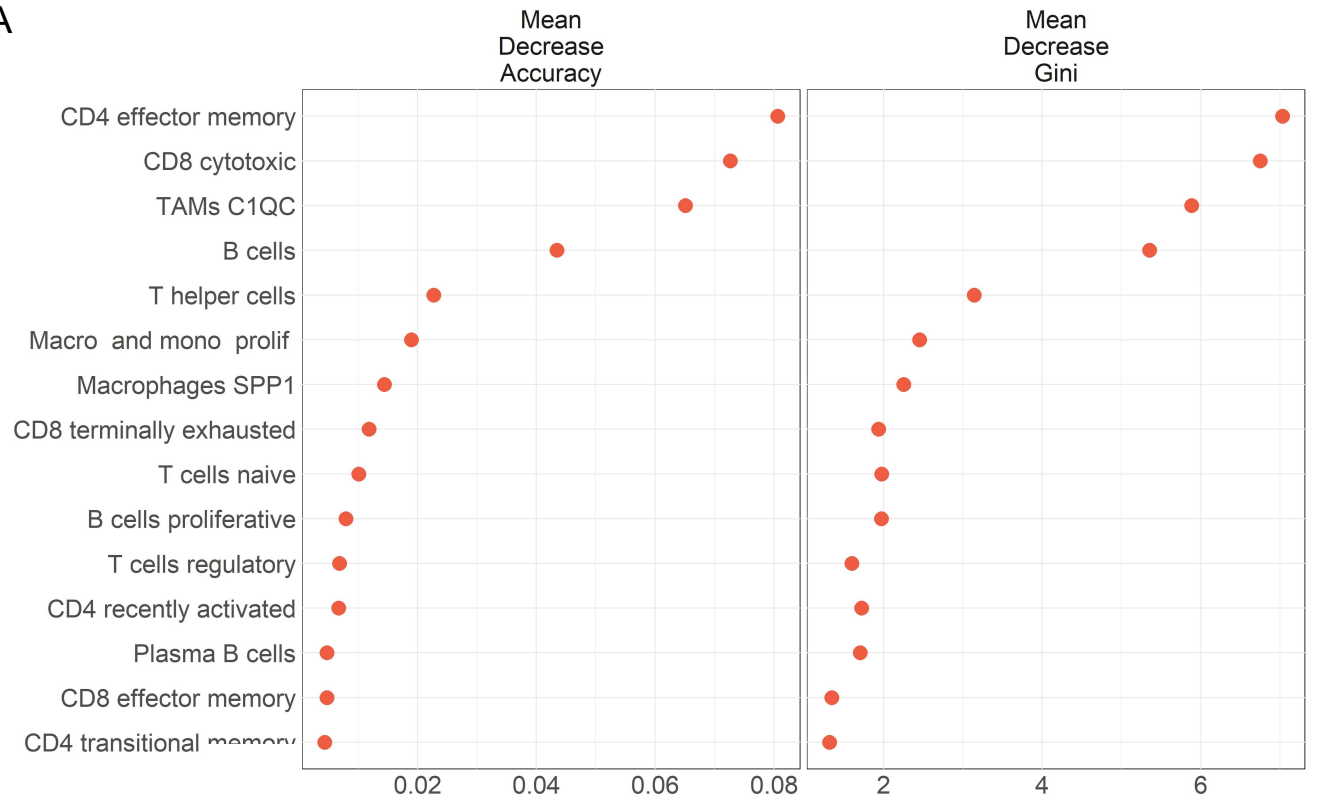

cluster
genes
Average Expression
Percent Expressed

D

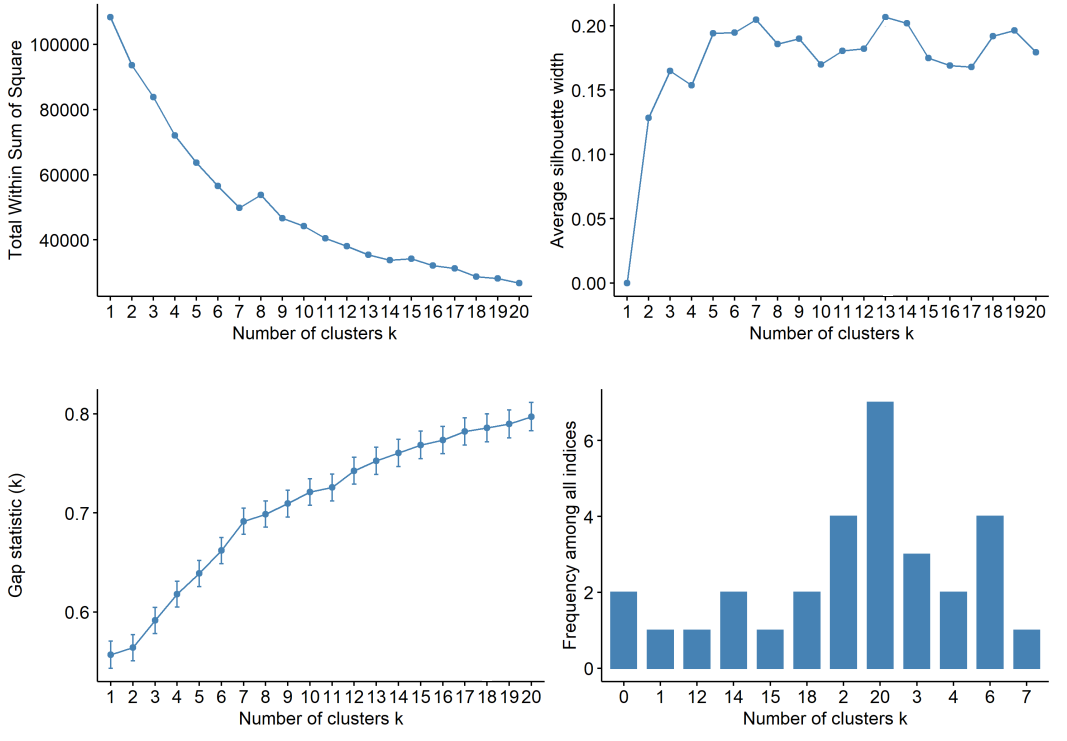

cell types
genes
Average Expression
Percent Expressed

**Supplementary Figure 8 | Immune cell states correlate in composition frequencies and enable to cluster patients into immune subtypes.** (**A**) Correlation matrix of the 25 immune cell states across patients. Co-occurring and mutually exclusive relationships are quantified by Pearson's correlation (positive: red; negative: blue). (**B-D**) Dotplots depicting the percentage of cells that express and the expression level of cluster marker/exhaustion genes. Plots display the top 10 marker genes for each immune cluster (C1-6, B), the expression of exhaustion markers in CD8 T-cells split by immune cluster (C1-6, C) and the expression of exhaustion marker in all cell states in cluster 5 (C5, D).

A

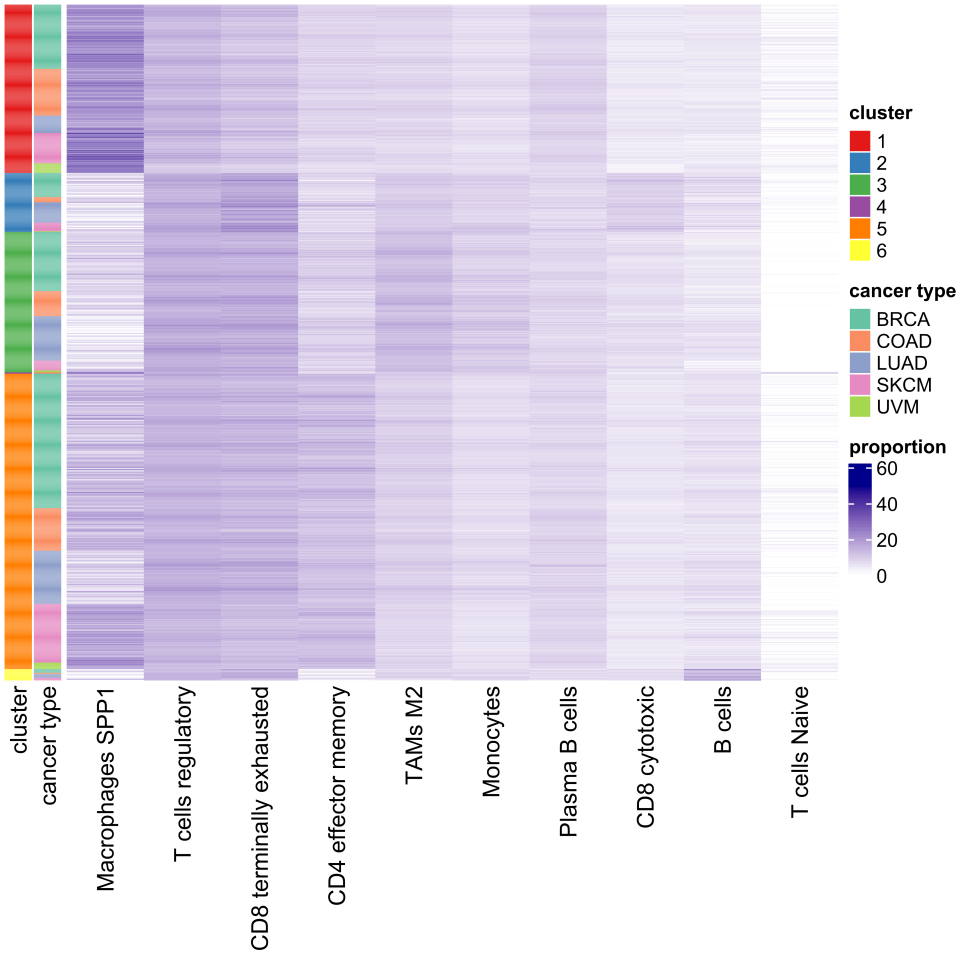

Mean Decrease Accuracy
Mean Decrease Gini
CD4 effector memory
CD8 cytotoxic
TAMs C1QC
B cells
T helper cells
Macro and mono prolif
Macrophages SPP1
CD8 terminally exhausted
T cells naive
B cells proliferative
T cells regulatory
CD4 recently activated
Plasma B cells
CD8 effector memory
CD4 transitional memory
0.02
0.04
0.06
0.08
2
4
6

B

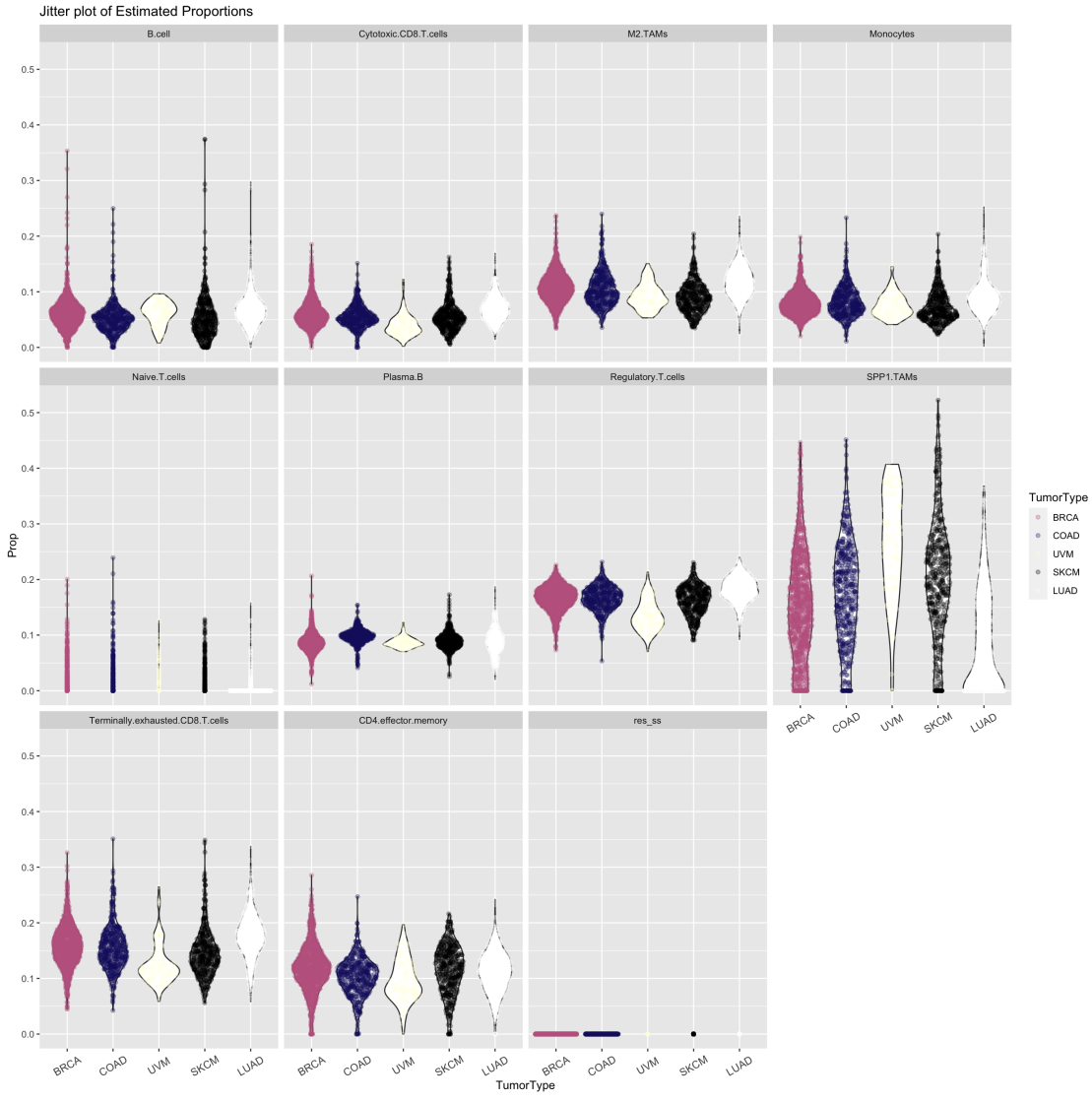

Total Within Sum of Square
Number of clusters k
Average silhouette width
Number of clusters k
Gap statistic (k)
Number of clusters k
Frequency among all indices
Number of clusters k

**Supplementary Figure 9 | Patient stratification.** (A) Random Forest top 15 most significant variables**.** (B) Different parameters to determine the best number of clusters, k.

A

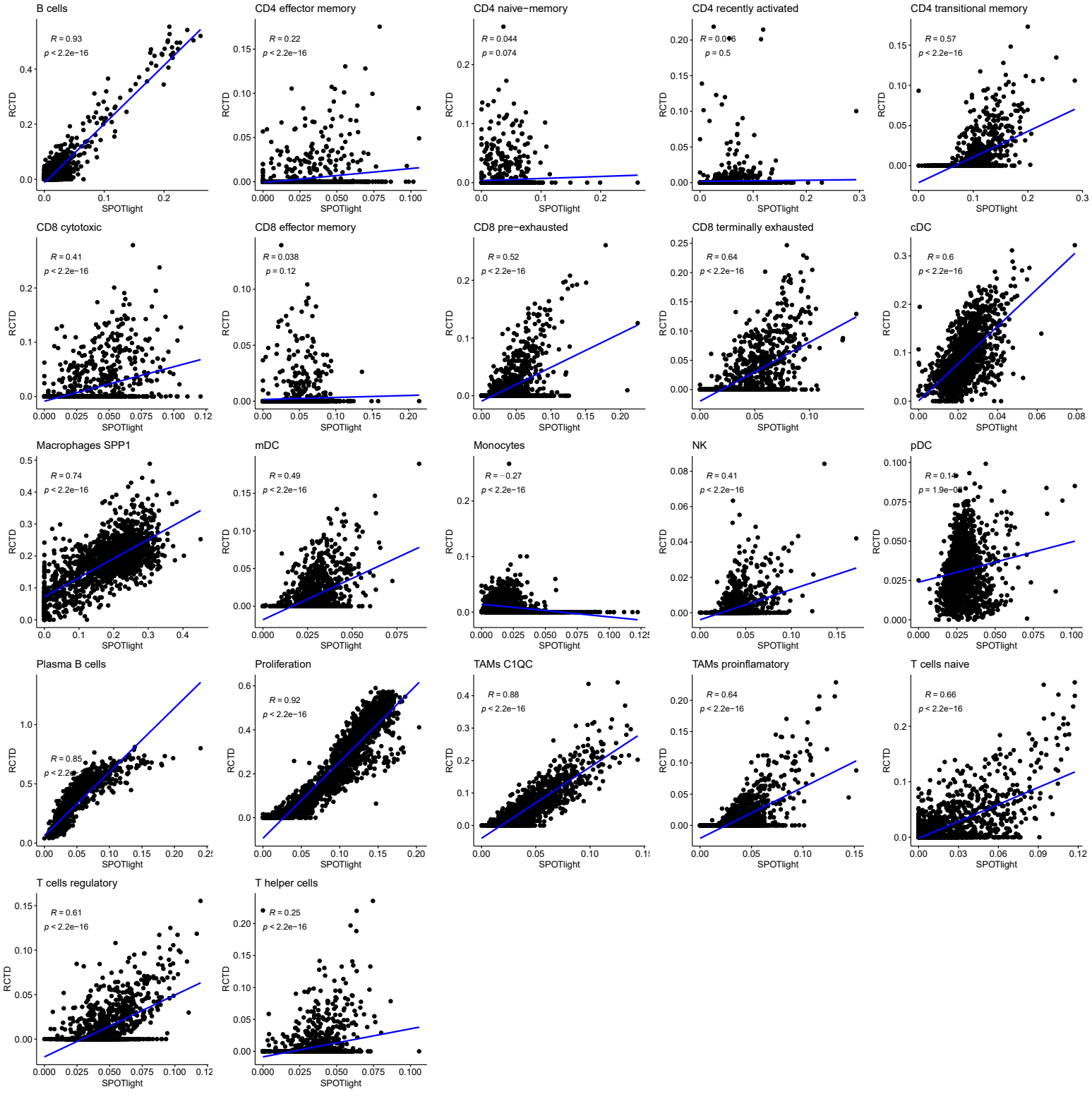

cluster
1
2
3
4
5
6
cancer type
BRCA
COAD
LUAD
SKCM
UVM
proportion
60
40
20
0
cluster
cancer type
Macrophages SPP1
T cells regulatory
CD8 terminally exhausted
CD4 effector memory
TAMs M2
Monocytes
Plasma B cells
CD8 cytotoxic
B cells
T cells Naive

B


Jitter plot of Estimated Proportions
B.cell
Cytotoxic.CD8.T.cells
M2.TAMs
Monocytes
Naive.T.cells
Plasma.B
Regulatory.T.cells
SPP1.TAMs
Terminally.exhausted.CD8.T.cells
CD4.effector.memory
res_ss
Prop
TumorType
BRCA
COAD
UVM
SKCM
LUAD

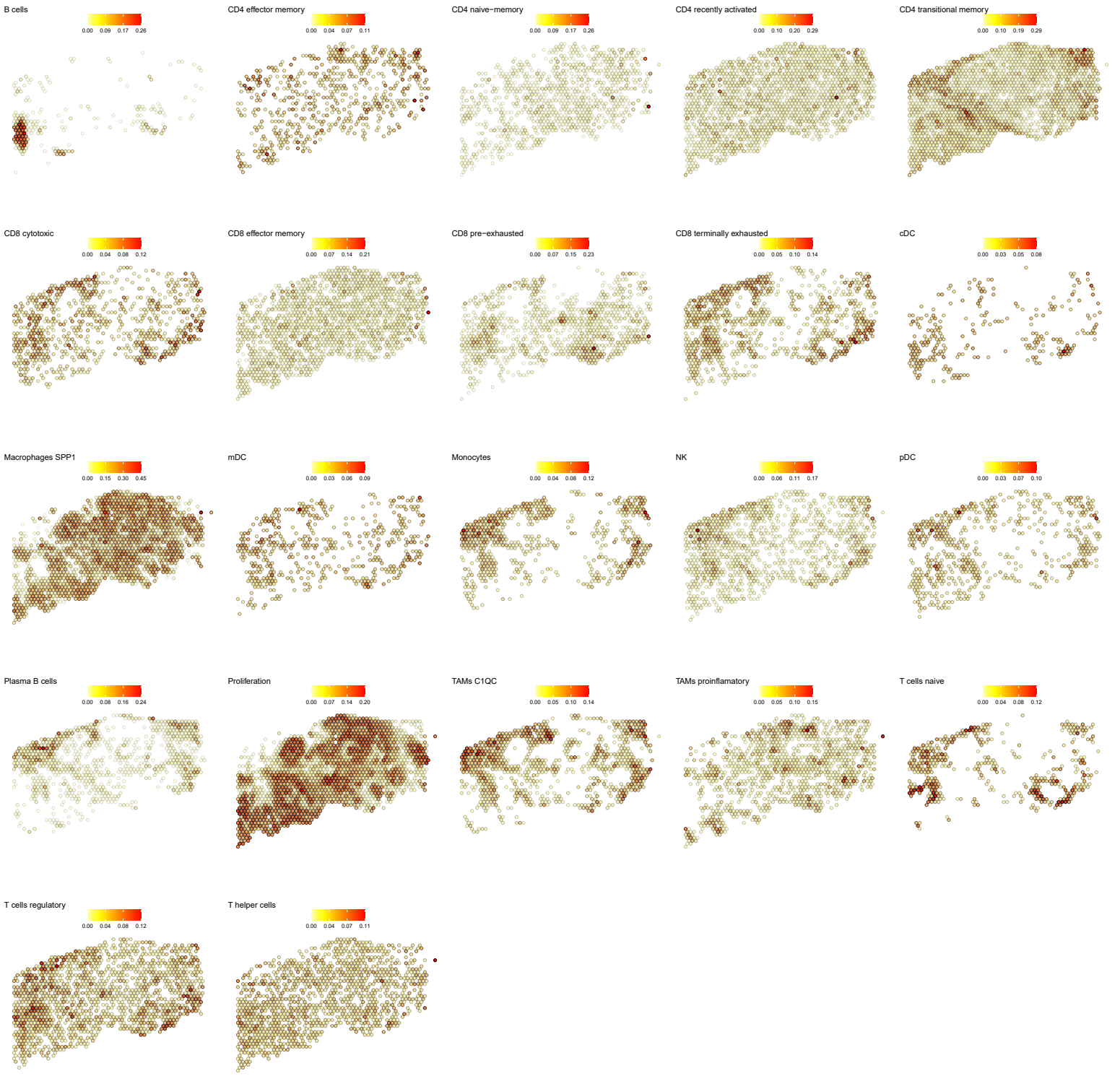
**Supplementary Figure 10 | Deconvolution of bulk RNA-seq TCGA data using the single-cell reference.** (**A**) Heatmap of the predicted cell type proportions for 1222 breast invasive carcinoma (BRCA), 594 lung adenocarcinoma (LUAD), 472 skin cutaneous melanoma (SKCM), and 521 colon adenocarcinoma (COAD) TCGA patient samples colored by cancer type and their immune cluster assigned through the Random Forest classifier. (**B**) Variability in the deconvoluted cell type proportions per cancer type.

Supplementary Figure 11



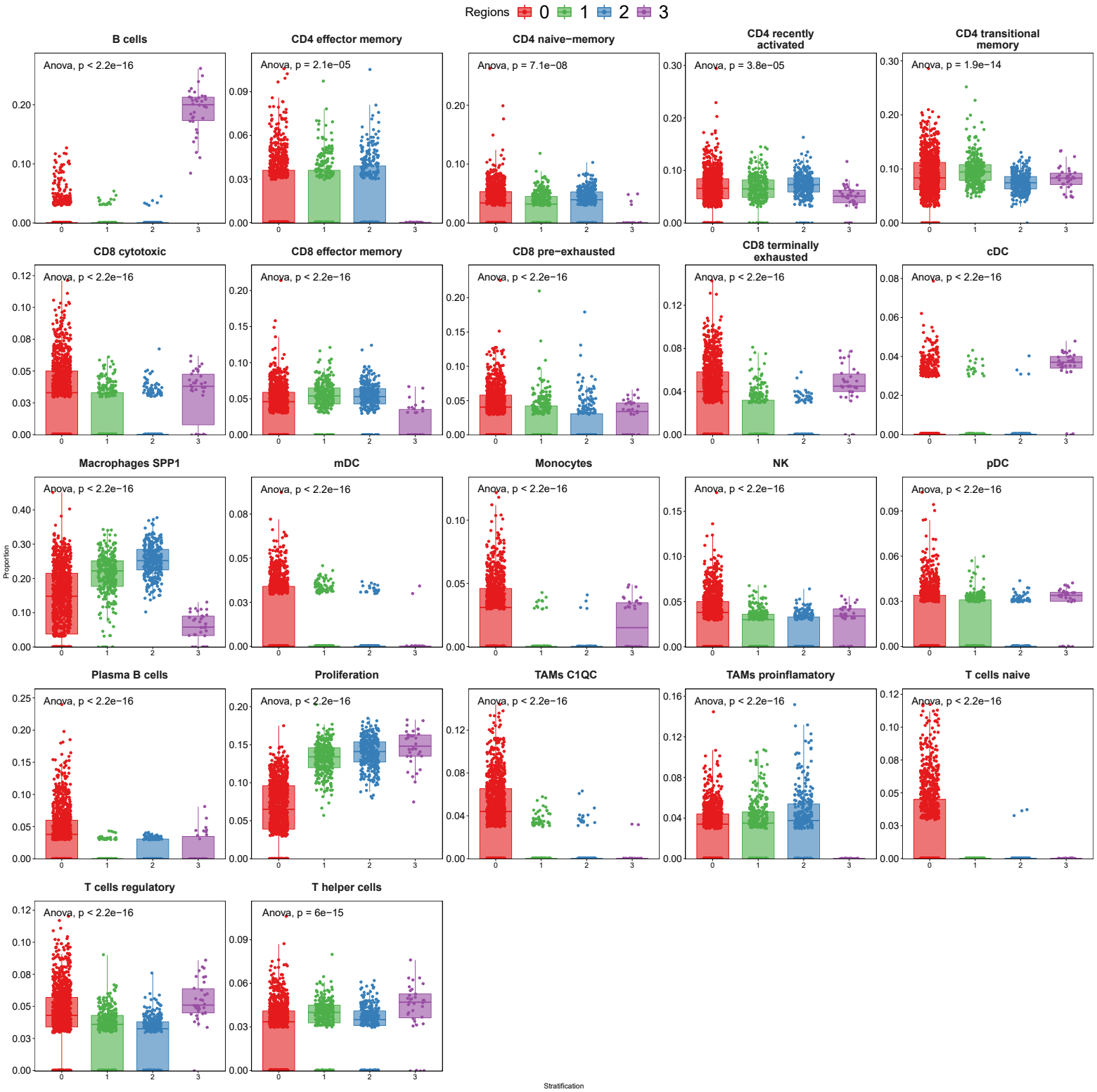
**Supplementary Figure 11 | SPOTlight and RCTD ST deconvolution.** Correlation analysis of predicted cell type/state level within ST spot using RCTD or SPOTlight based deconvolution (Pearson's correlation; sample 16142, SCC).

Supplementary Figure 12



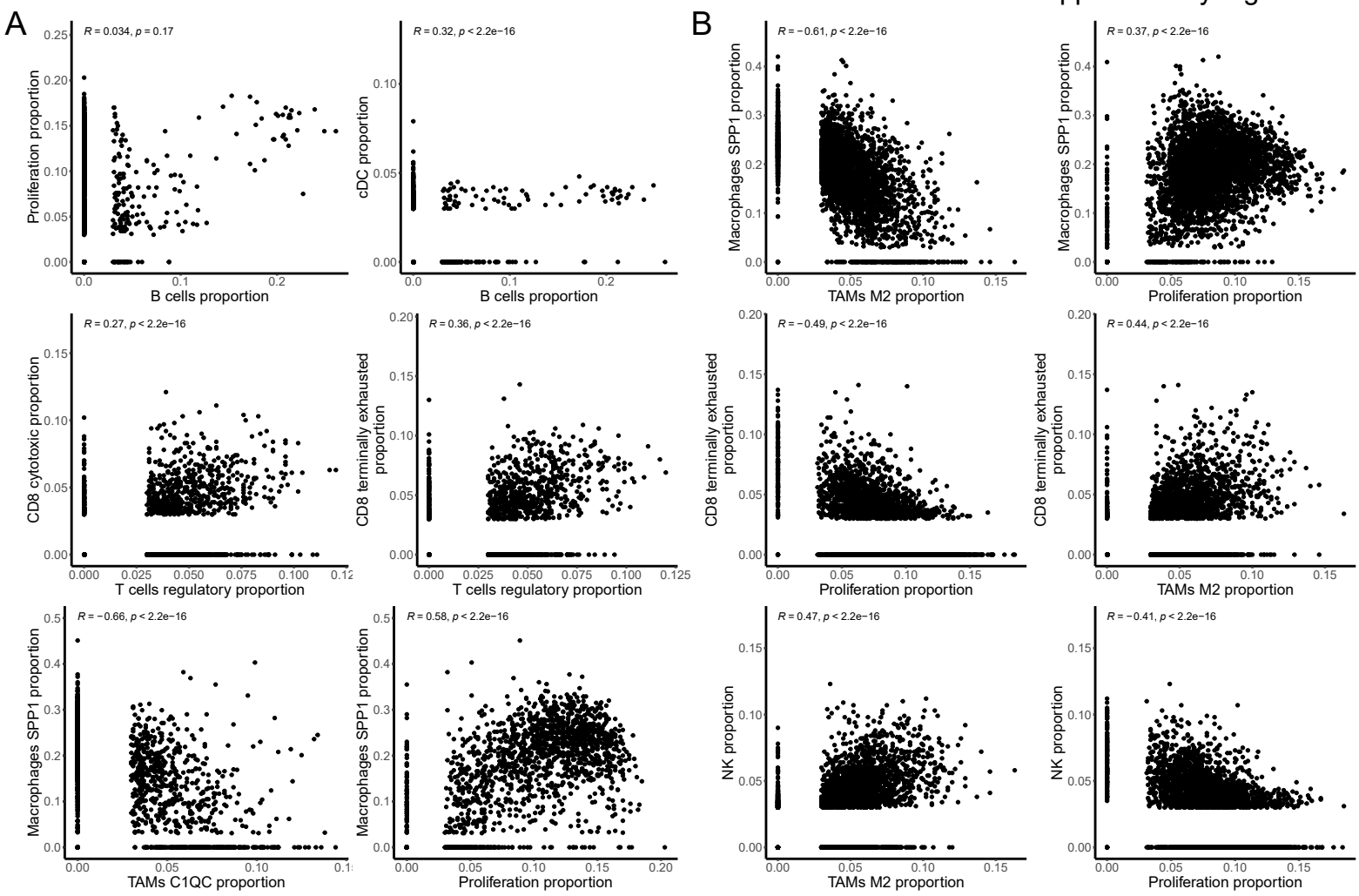
**Supplementary Figure 12 |** Sample 161429 (SCC) predicted cell type/state proportion within each spot showing spatially differential immune patterns.

Supplementary Figure 13

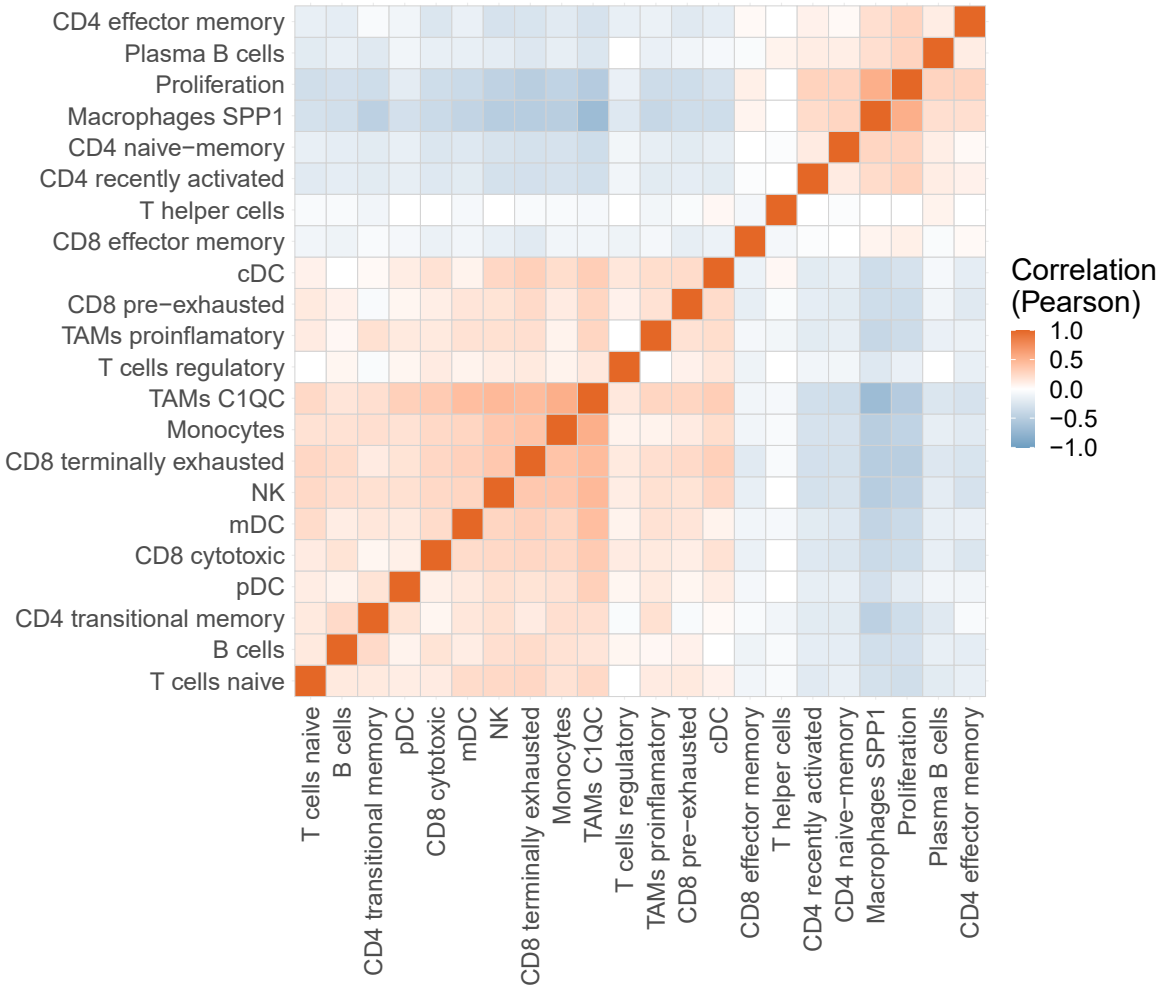


**Supplementary Figure 13 |** Sample 161429 (SCC) box plots of the cell types/states of interest presenting differential proportions of cell types found between regions. Differences between tumor sub-regions suggest a differential tumor immune microenvironment between heterogeneous tumor regions.

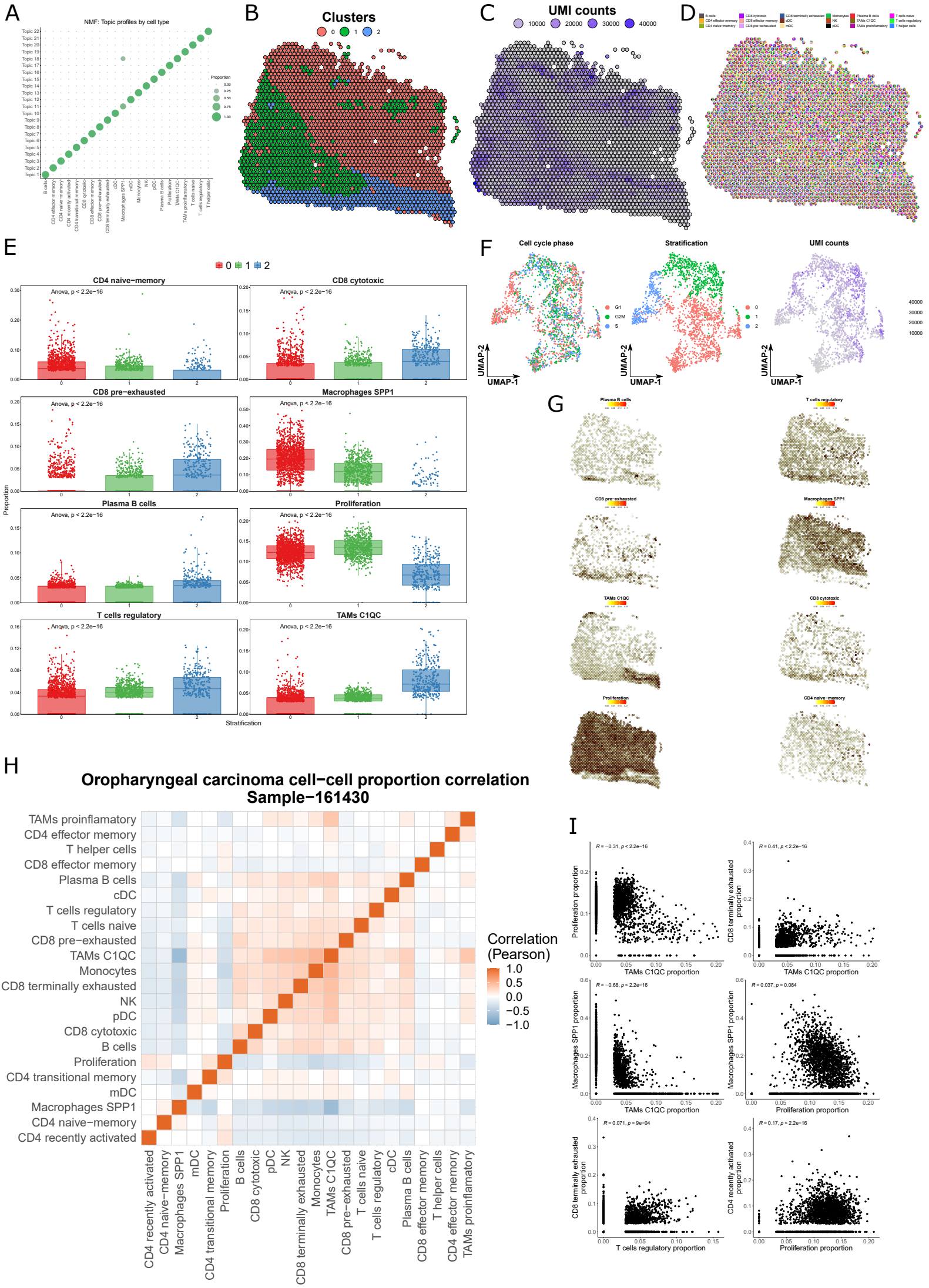
Supplementary Figure 14




C

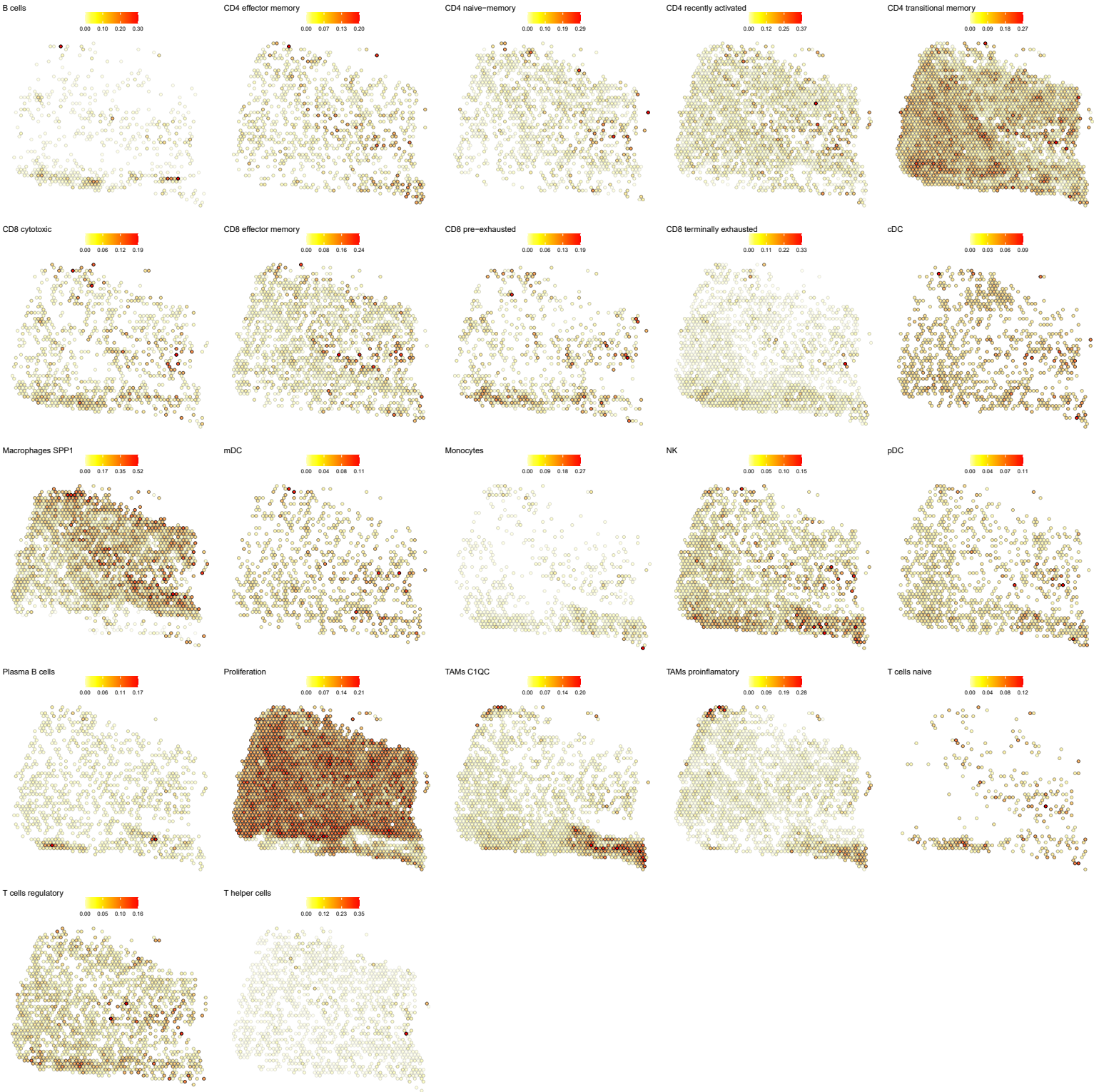


**Supplementary Figure 14 | (A)** Scatter plots of proportions of (anti-) correlating cell type pairs identified in Fig. 5H (oropharyngeal carcinoma). **(B)** Scatter plots of proportions of (anti-) correlating cell type pairs identified in C (breast carcinoma). **(C)** Clustered correlation matrix between the predicted cell-type proportions providing evidence for immune cell type co-localization in tumors.


A
NMF: Topic profiles by cell type
B
Clusters
C
UMI counts
D
E
CD4 naive−memory
CD8 cytotoxic
CD8 pre−exhausted
Macrophages SPP1
Plasma B cells
Proliferation
T cells regulatory
TAMs C1QC
Stratification
F
Cell cycle phase
Stratification
UMI counts
G
H
Oropharyngeal carcinoma cell−cell proportion correlation
Sample−161430
Correlation (Pearson)
I

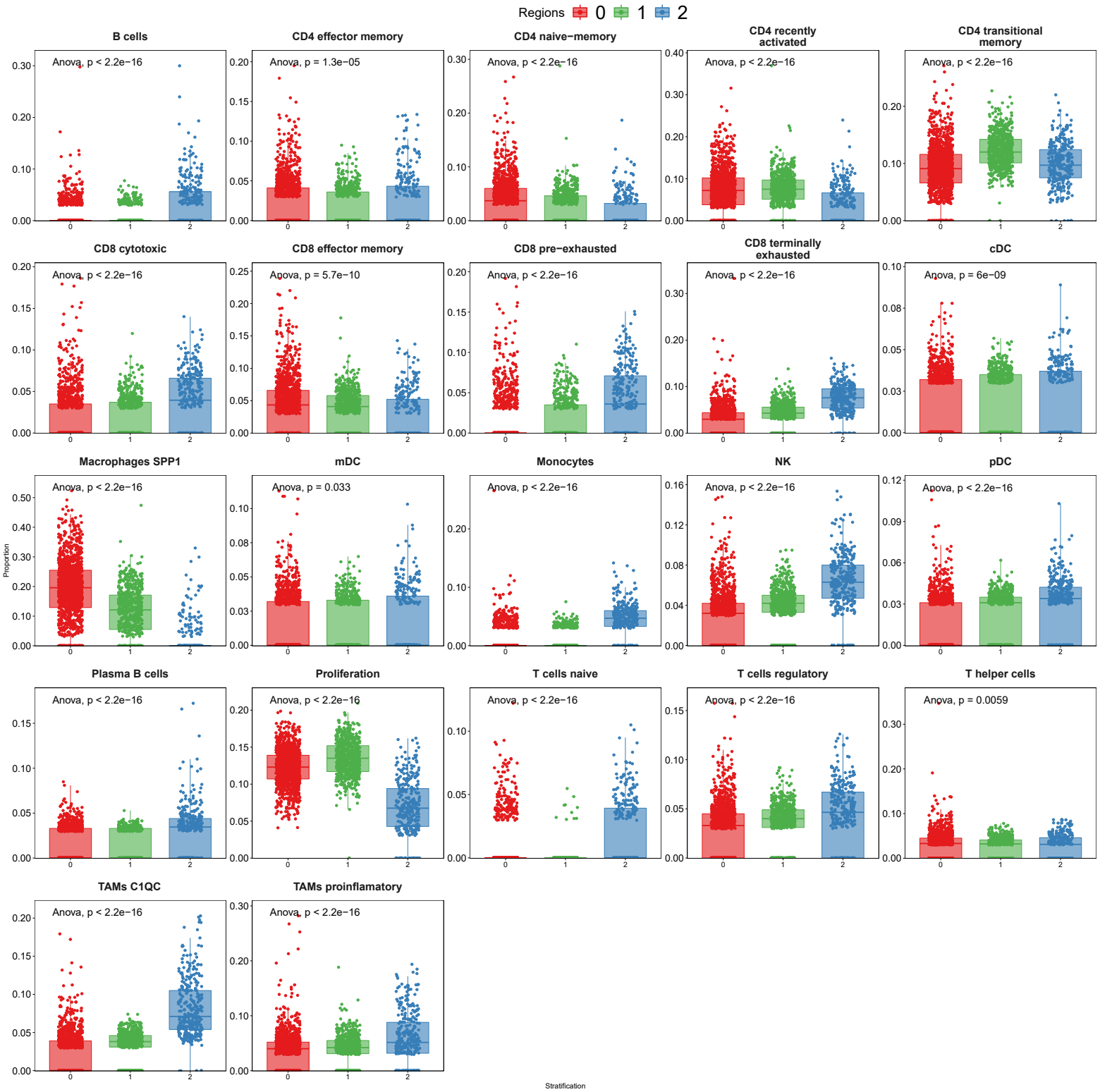
**Supplementary Figure 15 | Spatial mapping of the reference immune cell types using ST oropharyngeal squamous cell carcinoma (SCC) section ID 161430. (A)** Cell type specific topic profiles presenting a high topic/cell type specificity. **(B)** ST profiled section of a SCC primary tumor. Tissue stratification according to unsupervised clustering. **(C)** The number of unique molecular identifiers (UMI) recovered from each spot indicate the areas transcriptionally most active. **(D)** Pie chart representation showing proportions (per ST spot) of *SPOTlight* predicted immune cells based on the single-cell immune reference atlas. To visualize spatially variable cell types, only immune cell types present in <75% of the spots are displayed. **(E)** UMAP embedding of ST spots presenting the cell cycle phase (left), cluster identity (middle), and UMI counts (right) for each spot. **(F)** Box plots displaying significant differences in cell type proportion of clusters (ANOVA test). **(G)** Location and proportion of significantly differentially located cell types in the SCC section. **(H)** Clustered correlation matrix between the predicted cell type proportions identifying co-localization (red) and exclusive (blue) immune distribution patterns. **(I)** Scatter plots of (anti-) correlating cell type pairs identified in (H).

# Supplementary Figure 16

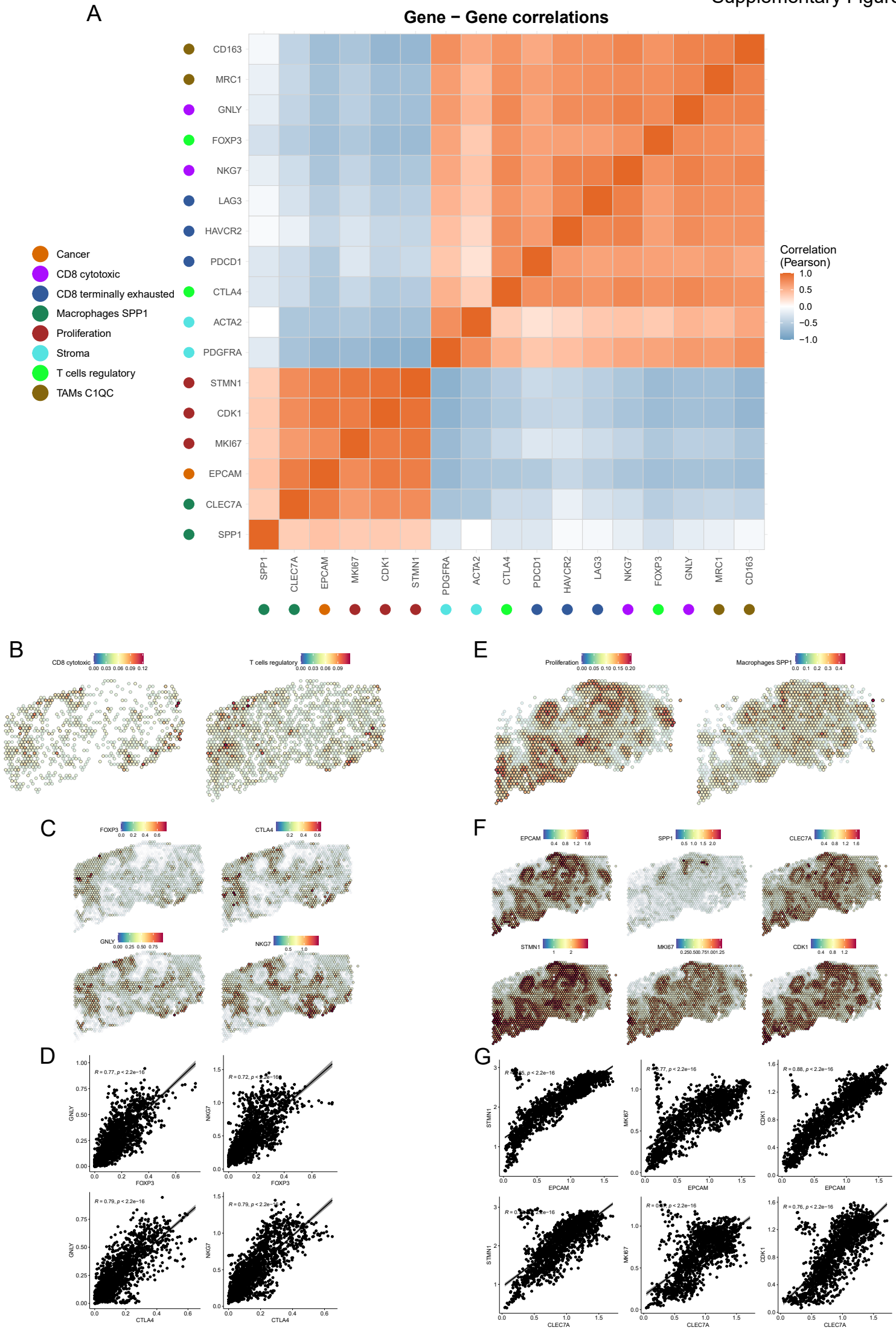


**Supplementary Figure 16 |** Sample 161430 predicted cell type/state proportion within each spot showing spatially differential immune patterns.

Supplementary Figure 17

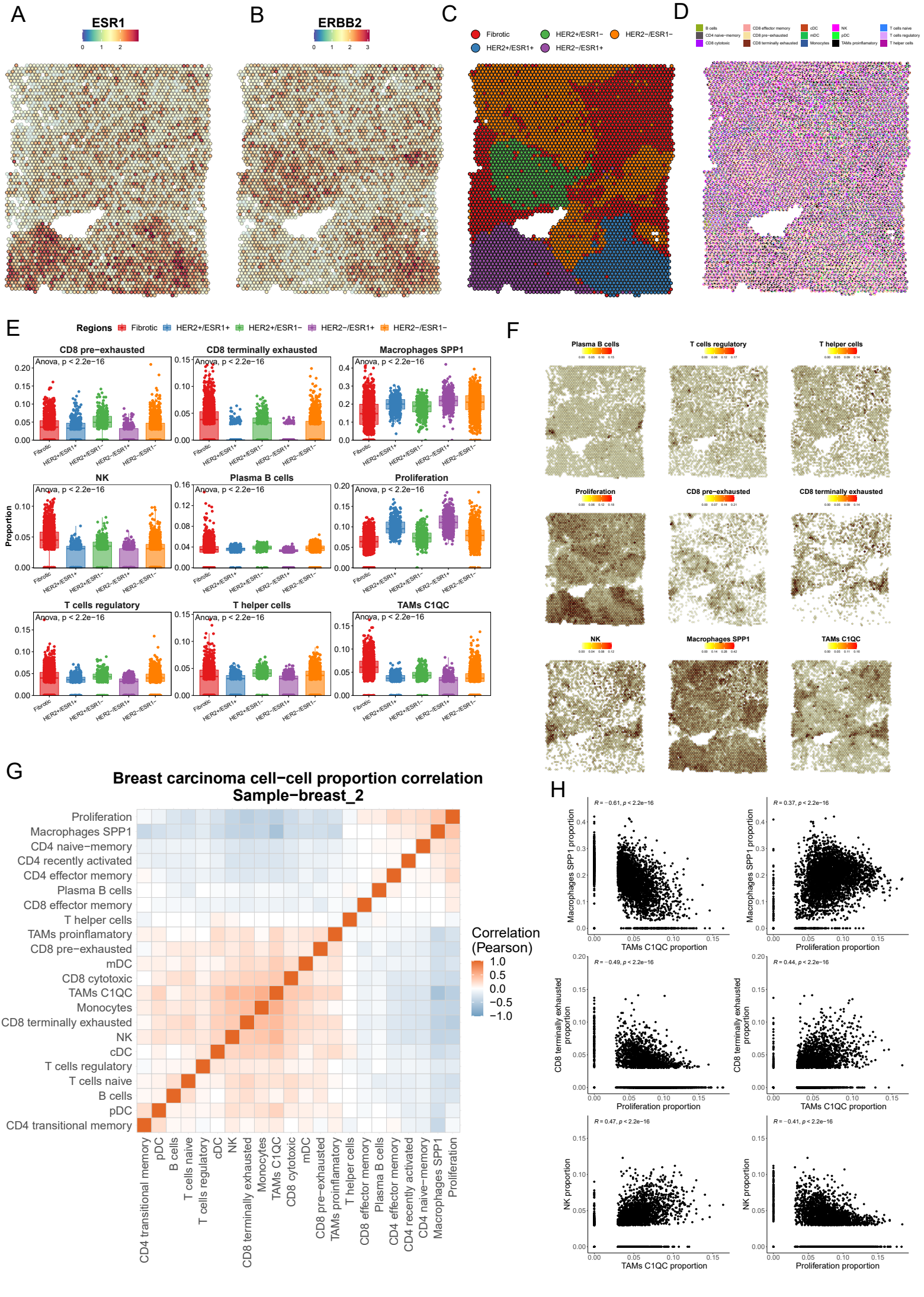


**Supplementary Figure 17 |** Sample 161430 box plots of the cell types/states of interest presenting differential proportions of cell types found between regions. Differences between tumor sub-regions suggest a differential tumor immune microenvironment between heterogeneous tumor regions.


A
Gene − Gene correlations
CD163
MRC1
GNLY
FOXP3
NKG7
LAG3
HAVCR2
PDCD1
CTLA4
ACTA2
PDGFRA
STMN1
CDK1
MKI67
EPCAM
CLEC7A
SPP1
Cancer
CD8 cytotoxic
CD8 terminally exhausted
Macrophages SPP1
Proliferation
Stroma
T cells regulatory
TAMs C1QC
Correlation (Pearson)
1.0
0.5
0.0
−0.5
−1.0
B
CD8 cytotoxic
T cells regulatory
C
FOXP3
CTLA4
GNLY
NKG7
D
R = 0.77, p < 2.2e−16
R = 0.72, p < 2.2e−16
R = 0.79, p < 2.2e−16
R = 0.79, p < 2.2e−16
E
Proliferation
Macrophages SPP1
F
EPCAM
SPP1
CLEC7A
STMN1
MKI67
CDK1
G
R = 0.85, p < 2.2e−16
R = 0.77, p < 2.2e−16
R = 0.88, p < 2.2e−16
R = 0.76, p < 2.2e−16

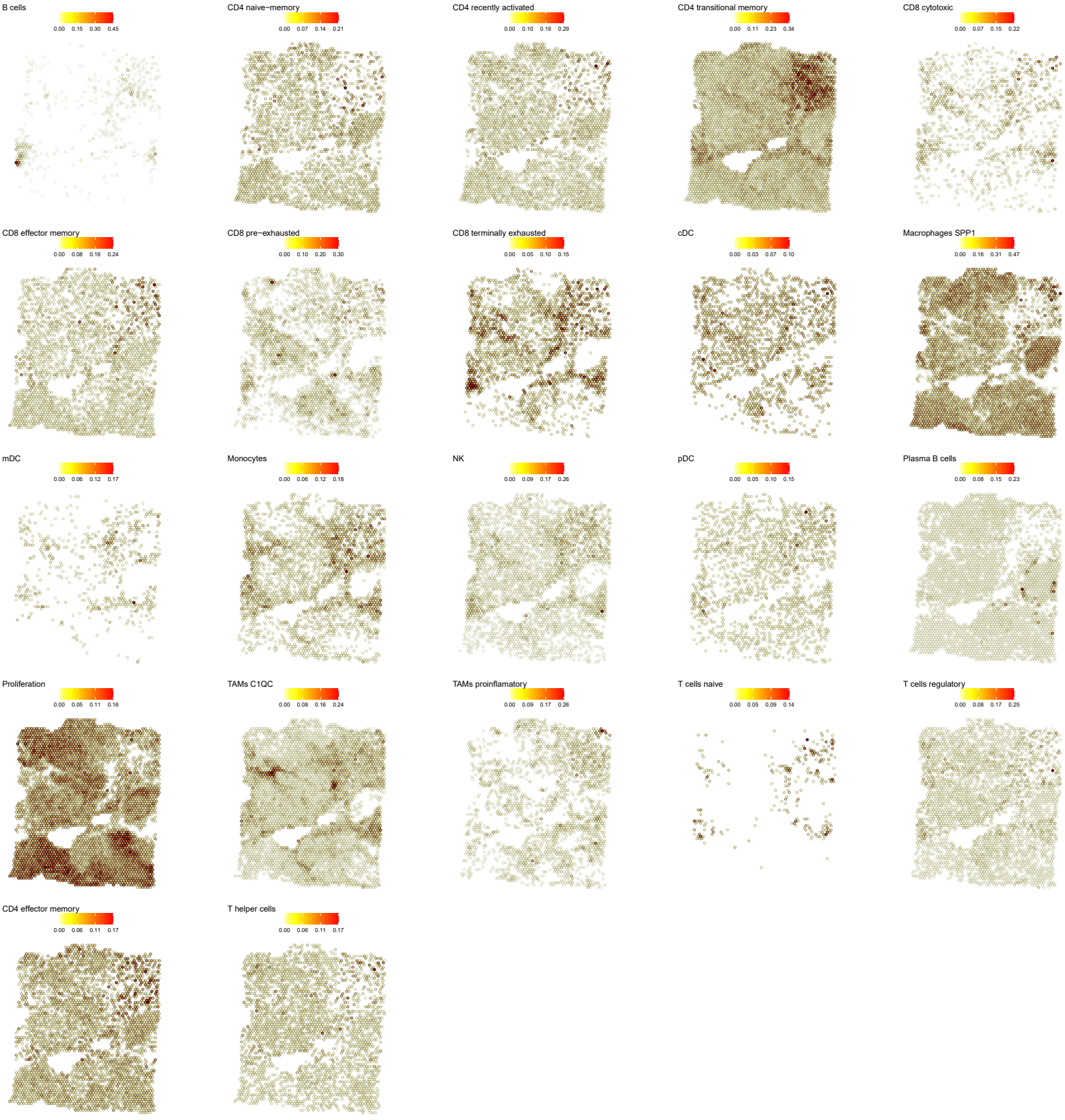
**Supplementary Figure 18 | Digital ISH SCC (161429) following denoising with MAGIC.** (**A**) Gene - Gene Pearson's correlation matrix between canonical marker genes of epithelial cancer, stromal fibroblasts, proliferation and immune cells of interest - CD8 cytotoxic, CD8 terminally exhausted, T cells regulatory, SPP1 and C1QC TAMs. (**B**) Cell type predicted proportion using SPOTlight. (**C**) Canonical marker genes for CD8 cytotoxic (*GNLY, NKG7*) and regulatory T-cells (*FOCP3, HAVCR2*). (**D**) Scatter plot and correlation between canonical markers for CD8 cytotoxic and regulatory T-cells. (**E**) Cell type predicted proportion using SPOTlight. (**F**) Canonical marker genes for cancer (*EPCAM*), proliferation (*STMN1*, *MKI67, CDK1*) and SPP1 macrophages (*SPP1, CLEC7A*). (**G**) Scatter plot and correlation between proliferation markers and the canonical markers for cancer and SPP1 macrophages.

Supplementary Figure 19



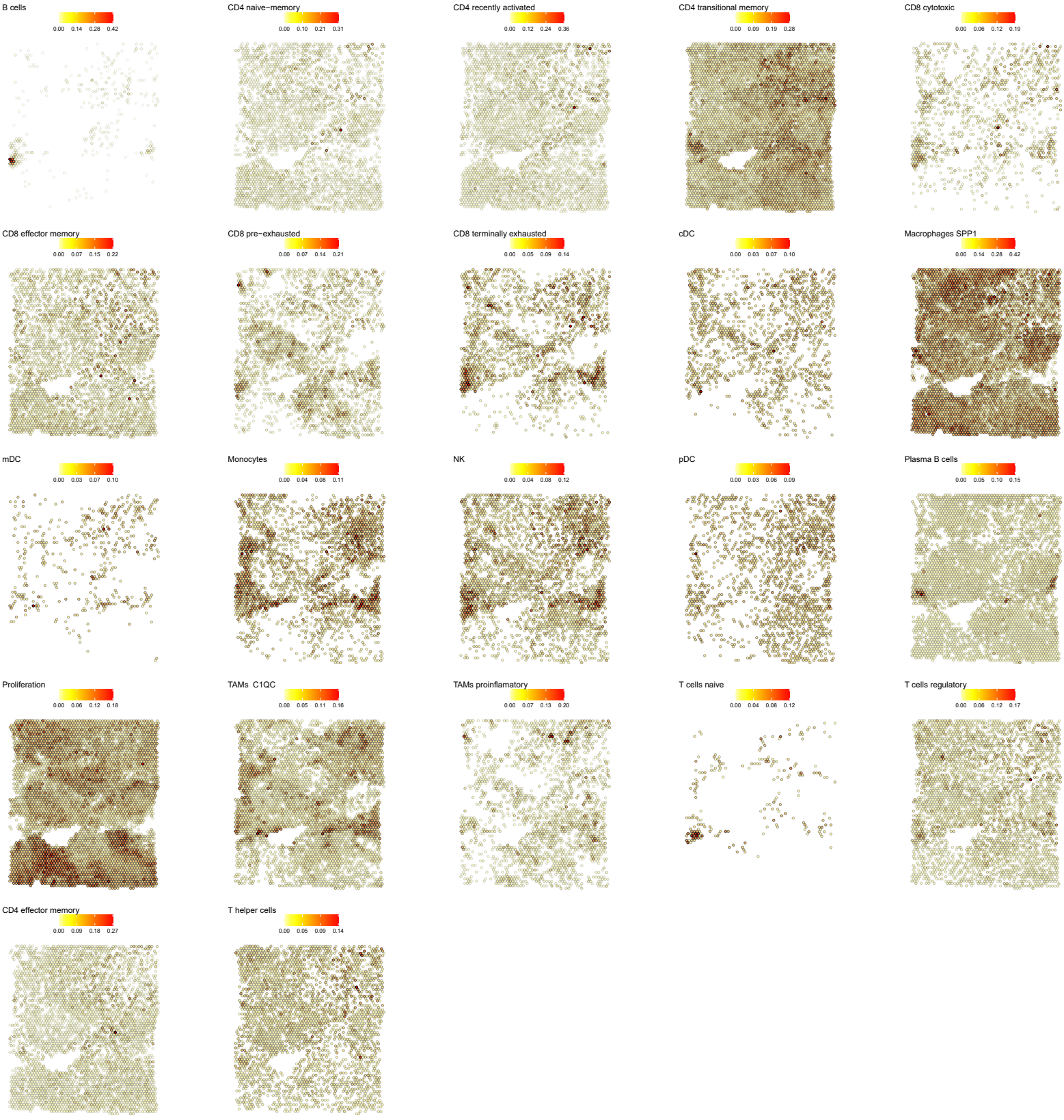
**Supplementary Figure 19 | Tumor immune reference mapping ST section from a ductal breast carcinoma (replicate slice 2). (A,B)** Estrogen receptor (ESR1, **A**) and HER2 receptor (ERBB2, **B**) gene expression levels on the ST section, indicating profound regionality of the expression. **(C)** Tissue stratification and labeling according to unsupervised clustering. **(D)** Pie chart representation showing proportions (per ST spot) of *SPOTlight* predicted immune cells based on the single-cell immune reference atlas. To visualize spatially variable cell types, only immune cell types present in <75% of the spots are displayed. **(E)** Box plots of significantly differentially localized cell type proportions between the clusters (ANOVA test). Differences between tumor areas (i.e. HER+ and ER+) are observed, suggesting differential tumor microenvironments of the tumor subclones. **(F)** Location and proportion of immune cell types with local enrichment **(G)** Clustered correlation matrix between the predicted cell-type proportions providing evidence for immune cell type co-localization in tumors. **(H)** Scatter plots of proportions of (anti-) correlating cell type pairs identified in (G).

# Supplementary Figure 20



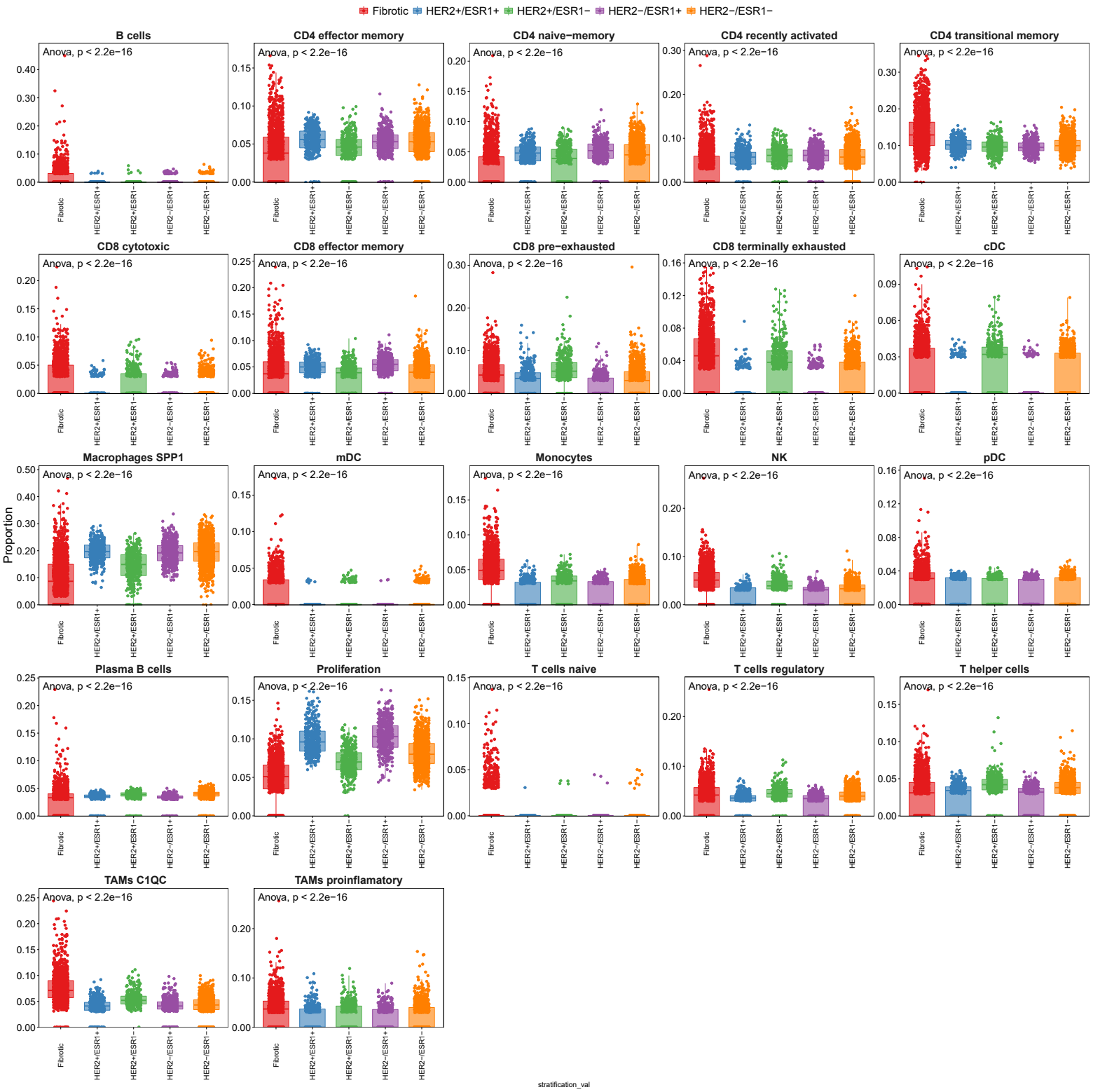
**Supplementary Figure 20 |** Predicted cell type/state proportion within each spot showing spatially differential immune patterns for the breast ductal carcinoma slice 1.

# Supplementary Figure 21



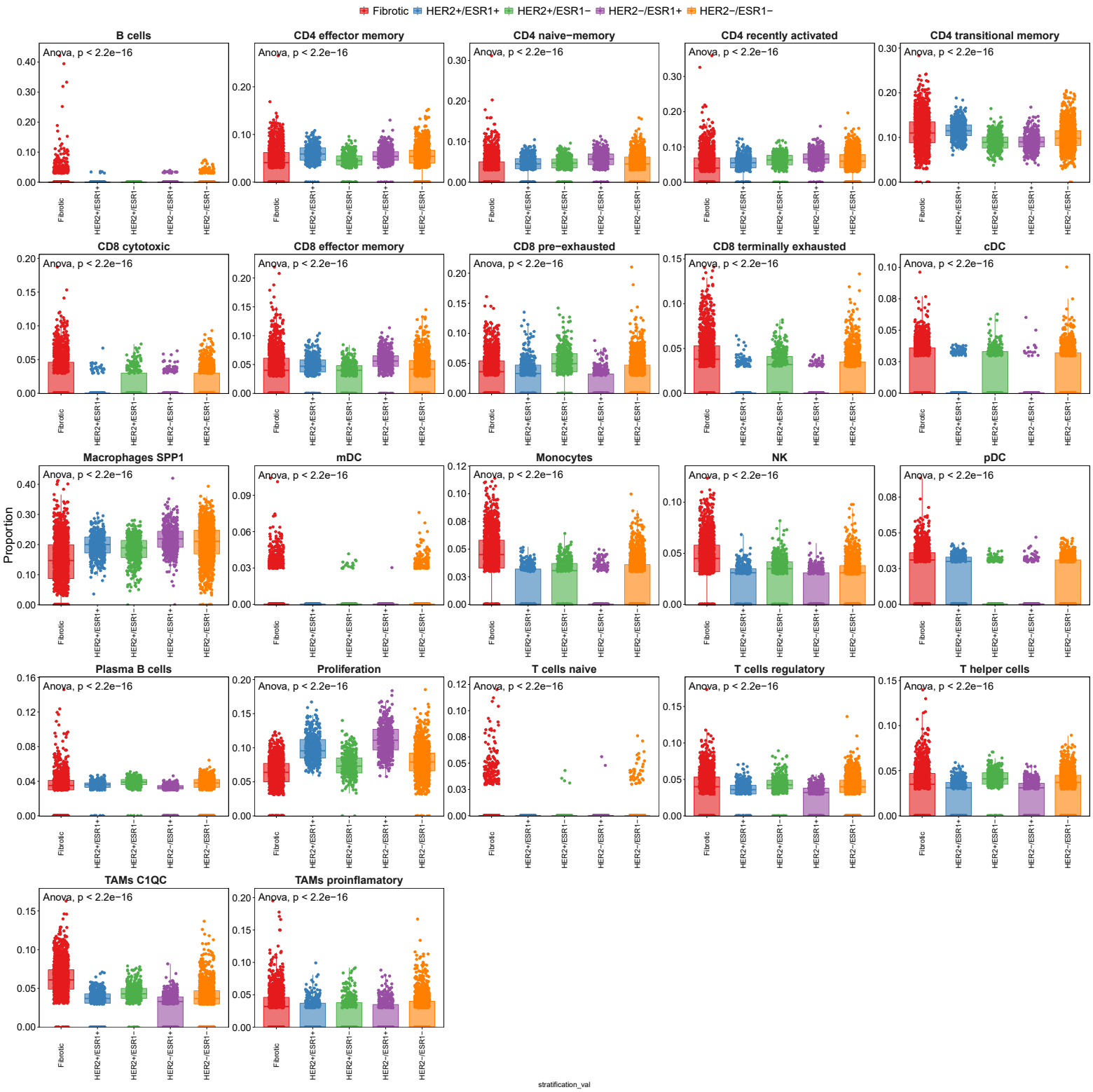
**Supplementary Figure 21 |** Predicted cell type/state proportion within each spot showing spatially differential immune patterns for the breast ductal carcinoma sequential slice 2 replicate.

Supplementary Figure 22



**Supplementary Figure 22 |** Box plots of the cell types/states of interest presenting differential proportions of cell types found between regions for the breast ductal carcinoma slice 1. Differences between tumor sub-regions suggest a differential tumor immune microenvironment between heterogeneous tumor regions.

Supplementary Figure 23



**Supplementary Figure 23 |** Box plots of the cell types/states of interest presenting differential proportions of cell types found between regions for the breast ductal carcinoma sequential slice 2 replicate. Differences between tumor sub-regions suggest a differential tumor immune microenvironment between heterogeneous tumor regions.