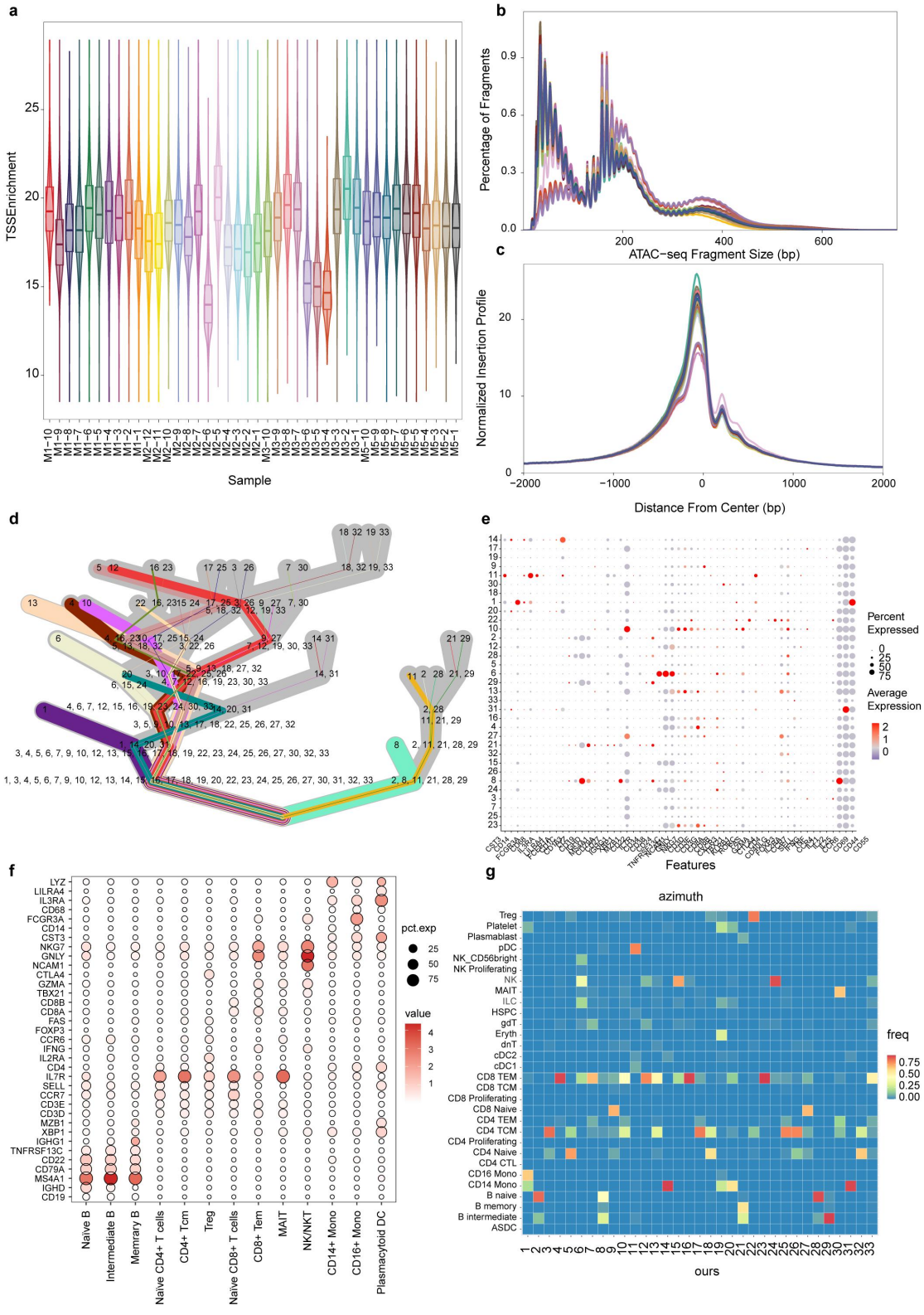
Table of Content



Extended Data Fig. 1 | Single-nucleus analysis quality control and cell type annotation for COVID-19 vaccine PBMC

a. Violin plots of transcription start site (TSS) enrichments for all snATAC cells from each sample.

b. Fragment lengths indicating nucleosomal periodicity in aggregated single-nucleus profiles from 10x snATAC-seq based samples.

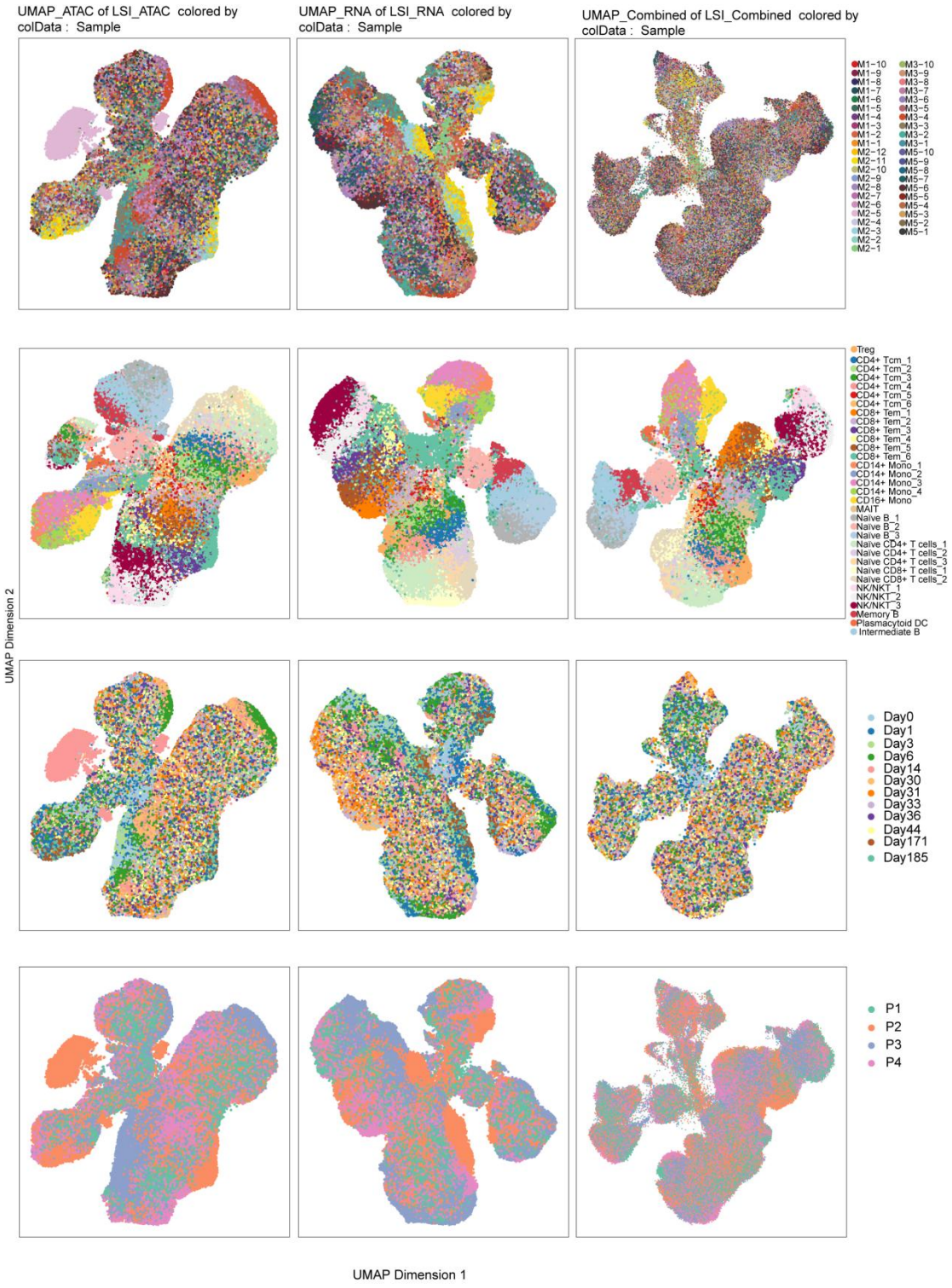
c. The enrichments of normalized Tn5 transposase insertions around the transcription start sites (TSSs) of 10x snATAC-seq based samples.

d. Hierarchical clustering visualization of the output from light HIPPO analysis.

e. Dot plots of gene expression for populations shown in Fig. 1a where the color is scaled by mean expression and the dot size is proportional to the percent of the population expressing the gene.

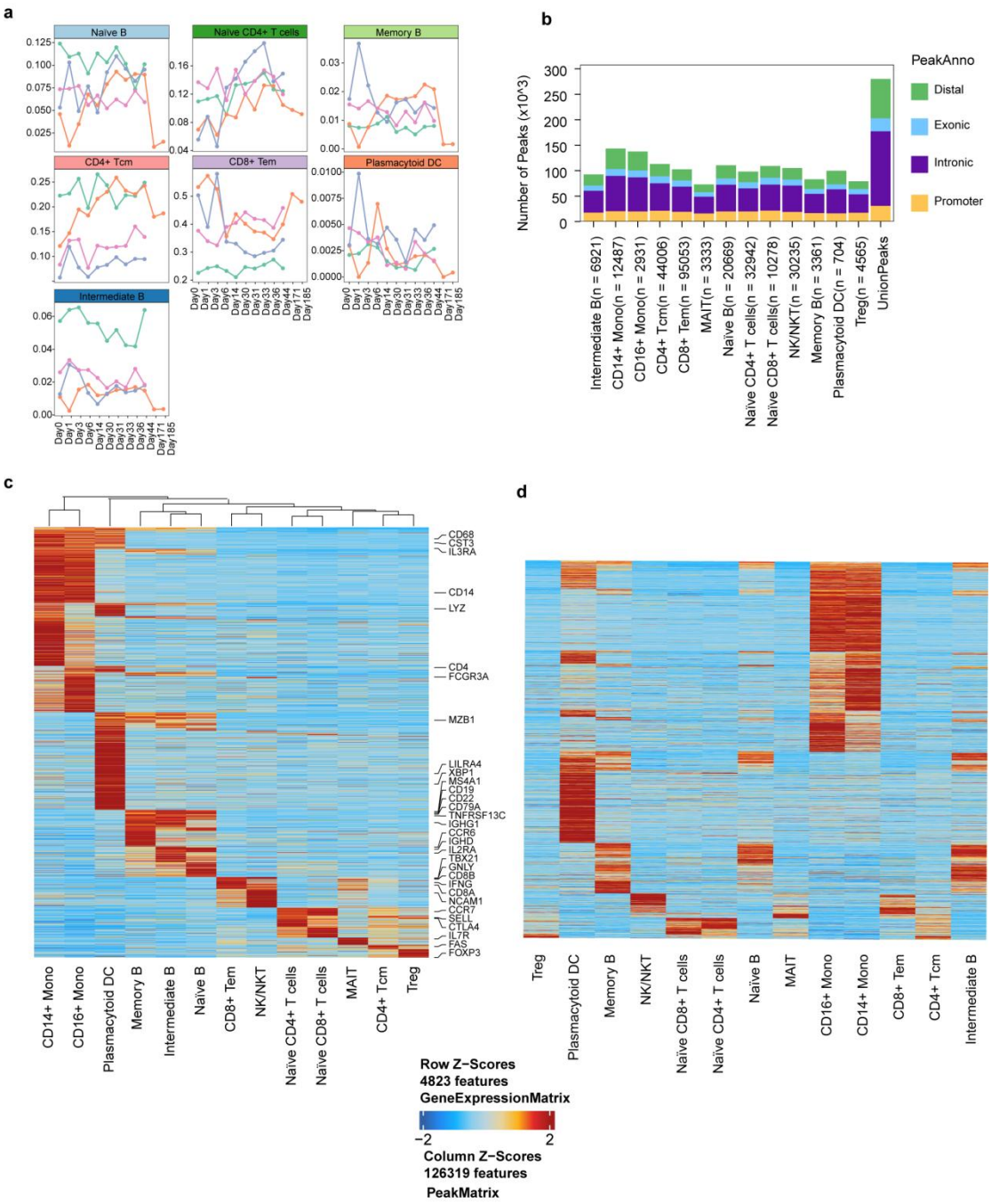
f. Dot plots of gene expression for populations shown in S1.d where the color is scaled by mean expression and the dot size is proportional to the percent of the population expressing the gene.

g. Tile plot showing the percentage concordance between COVID-19 PBMC annotation (x-axis) and Azimuth reference-based annotation tool (y-axis) (https://satijalab.org/azimuth/).



Extended Data Fig. 2 | Expanded view of matched snRNA-seq and snATAC-seq for 42 samples

UMAP visualizations where dots correspond to individual nuclei for nuclei profiled with snRNA-seq (left column), snATAC-seq (middle column), and integrated (right column), colored by sample, annotation, timepoints, and participants.


a
Naive B
Naive CD4+ T cells
Memory B
CD4+ Tcm
CD8+ Tem
Plasmacytoid DC
Intermediate B
b
Number of Peaks (x10^3)
PeakAnno
Distal
Exonic
Intronic
Promoter
Intermediate B(n = 6921)
CD14+ Mono(n = 12487)
CD16+ Mono(n = 2931)
CD4+ Tcm(n = 44006)
CD8+ Tem(n = 95053)
MAIT(n = 3333)
Naive B(n = 20669)
Naive CD4+ T cells(n = 32942)
Naive CD8+ T cells(n = 10278)
NK/NKT(n = 30235)
Memory B(n = 3361)
Plasmacytoid DC(n = 704)
Treg(n = 4565)
UnionPeaks
c
CD68
CST3
IL3RA
CD14
LYZ
CD4
FCGR3A
MZB1
LILRA4
XBP1
MS4A1
CD19
CD22
CD79A
TNFRSF13C
IGHG1
CCR6
IGHD
IL2RA
TBX21
GNLY
CD8B
IFNG
CD8A
NCAM1
CCR7
SELL
CTLA4
IL7R
FAS
FOXP3
CD14+ Mono
CD16+ Mono
Plasmacytoid DC
Memory B
Intermediate B
Naive B
CD8+ Tem
NK/NKT
Naive CD4+ T cells
Naive CD8+ T cells
MAIT
CD4+ Tcm
Treg
Row Z−Scores
4823 features
GeneExpressionMatrix
−2
2
Column Z−Scores
126319 features
PeakMatrix
d
Treg
Plasmacytoid DC
Memory B
NK/NKT
Naive CD8+ T cells
Naive CD4+ T cells
Naive B
MAIT
CD16+ Mono
CD14+ Mono
CD8+ Tem
CD4+ Tcm
Intermediate B

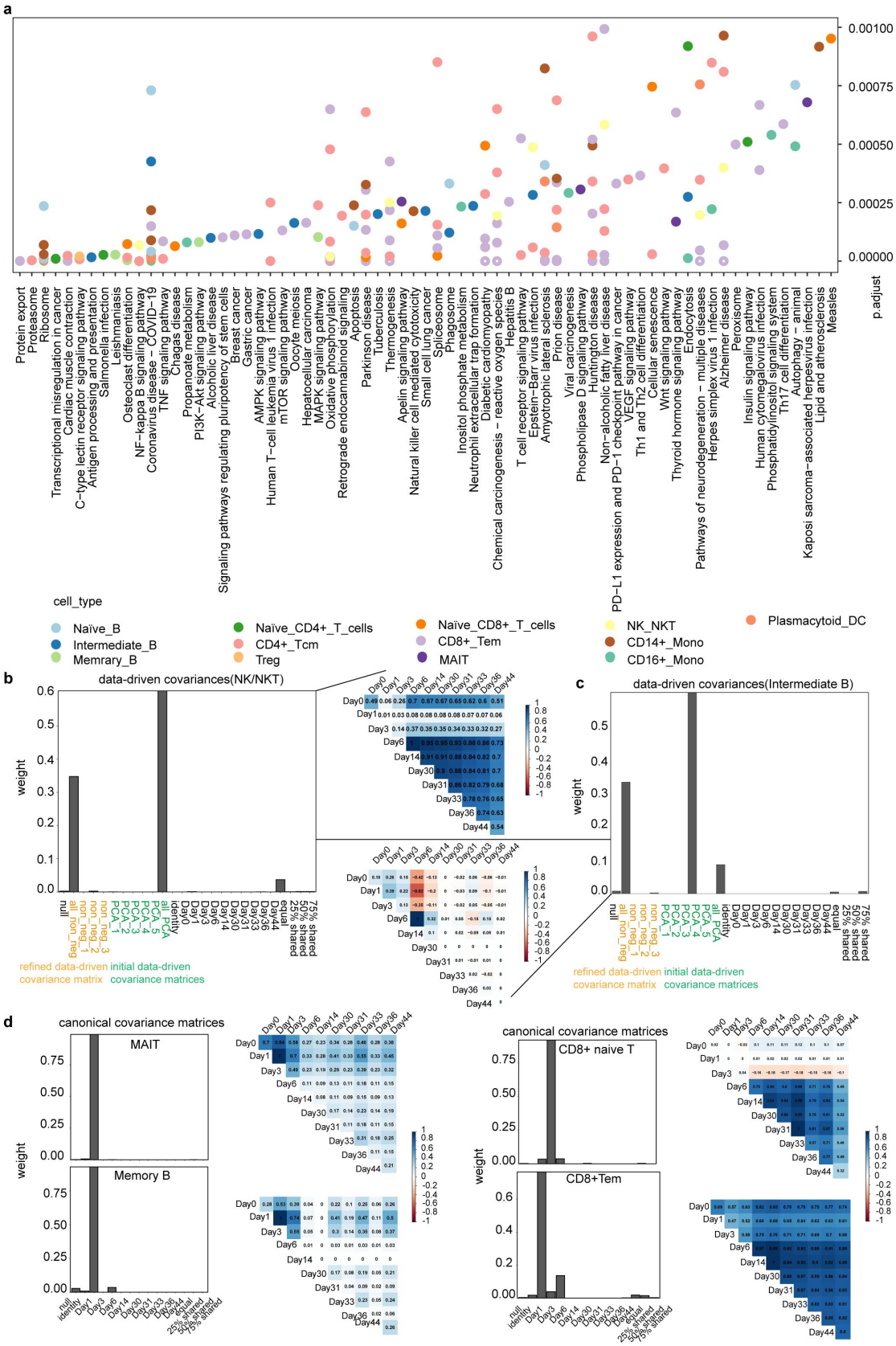
Extended Data Fig. 3 | Cell types display distinct gene regulatory processes

a. Line chart of the proportion of cell types, separated by time points and participants.

b. Number of peaks in each cell type. Peaks were annotated into distal, exonic, intronic and promoter.

c. Heatmap of PBMC cell-type marker genes (n = 5,121) across each cell type calculated from snRNA-seq gene expression. each row represents a unique marker gene. The color represents the normalized gene score of the marker genes in cell types.

d. Heatmap reflecting PBMC cell-type marker peaks that highlight CREs specific to only one or very limited cell types. each column represents an individual marker peak. The color represents the normalized marker peak accessibility in cell types.


a
p.adjust
0.00100
0.00075
0.00050
0.00025
0.00000
Protein export
Proteasome
Ribosome
Transcriptional misregulation in cancer
Cardiac muscle contraction
C−type lectin receptor signaling pathway
Antigen processing and presentation
Salmonella infection
Leishmaniasis
Osteoclast differentiation
NF−kappa B signaling pathway
Coronavirus disease − COVID−19
TNF signaling pathway
Chagas disease
Propanoate metabolism
PI3K−Akt signaling pathway
Alcoholic liver disease
Signaling pathways regulating pluripotency of stem cells
Breast cancer
Gastric cancer
AMPK signaling pathway
Human T−cell leukemia virus 1 infection
mTOR signaling pathway
Oocyte meiosis
Hepatocellular carcinoma
MAPK signaling pathway
Oxidative phosphorylation
Retrograde endocannabinoid signaling
Apoptosis
Parkinson disease
Tuberculosis
Thermogenesis
Apelin signaling pathway
Natural killer cell mediated cytotoxicity
Small cell lung cancer
Spliceosome
Phagosome
Inositol phosphate metabolism
Neutrophil extracellular trap formation
Diabetic cardiomyopathy
Chemical carcinogenesis − reactive oxygen species
Hepatitis B
T cell receptor signaling pathway
Epstein−Barr virus infection
Amyotrophic lateral sclerosis
Prion disease
Viral carcinogenesis
Phospholipase D signaling pathway
Huntington disease
Non−alcoholic fatty liver disease
PD−L1 expression and PD−1 checkpoint pathway in cancer
VEGF signaling pathway
Th1 and Th2 cell differentiation
Cellular senescence
Wnt signaling pathway
Thyroid hormone signaling pathway
Endocytosis
Pathways of neurodegeneration − multiple diseases
Herpes simplex virus 1 infection
Alzheimer disease
Peroxisome
Insulin signaling pathway
Human cytomegalovirus infection
Phosphatidylinositol signaling system
Th17 cell differentiation
Autophagy − animal
Kaposi sarcoma−associated herpesvirus infection
Lipid and atherosclerosis
Measles
cell_type
Naïve_B
Intermediate_B
Memrary_B
Naïve_CD4+_T_cells
CD4+_Tcm
Treg
Naïve_CD8+_T_cells
CD8+_Tem
MAIT
NK_NKT
CD14+_Mono
CD16+_Mono
Plasmacytoid_DC
b
data-driven covariances(NK/NKT)
weight
refined data-driven covariance matrix
initial data-driven covariance matrices
c
data-driven covariances(Intermediate B)
weight
refined data-driven covariance matrix
initial data-driven covariance matrices
d
canonical covariance matrices
MAIT
Memory B
weight
canonical covariance matrices
CD8+ naive T
CD8+Tem
weight

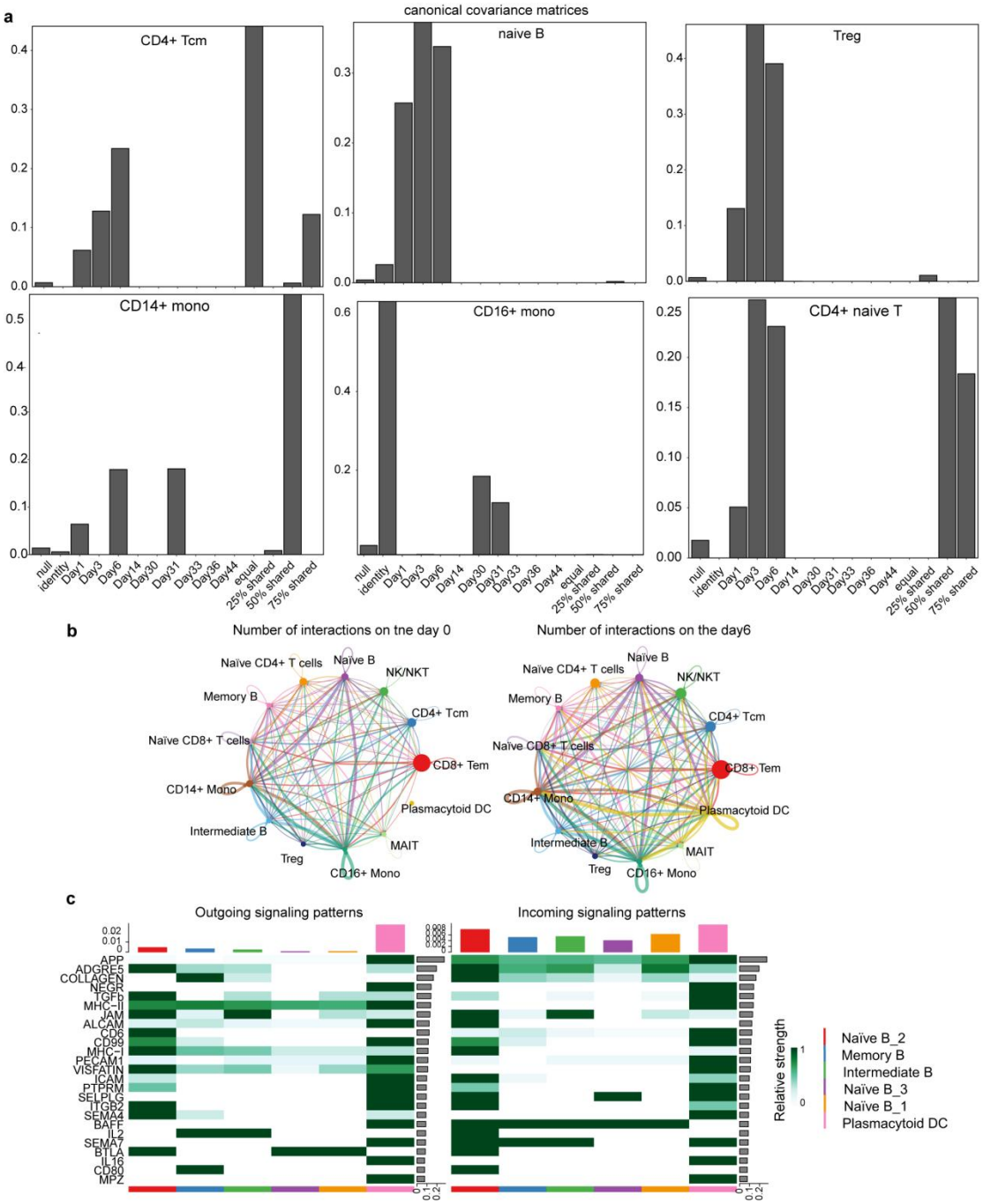
Extended Data Fig. 4 | Co-expression patterns across timepoints

a. Kyoto (KEGG) enrichment analysis for all gene sets derived from the co-expression analysis in each cell type. The values on the x-axis indicate the p adjust value (p.adjust-value of < = 0.001). Solid circles indicate that the fold enrichment ratio>=2/3.The numerator is the number of genes enriched on this GO entry,the denominator is the number of all entered genes for enrichment analysis.

b. Mashr estimates weight in NK cells for both the canonical matrix and the data-driven matrix, capturing the proportion of each effect pattern(left). Identity represents the independent effect across all conditions; Dayn indicates Dayn specific effects; equal_effects represent patterns that are equal in all conditions. Heatmap of the covariance matrix corresponding to the dominant mixture component identified by mash(right).

c. Mashr estimates weight in intermediate B cells for both the canonical matrix and the data-driven matrix(right). Heatmap of the covariance matrix corresponding to the dominant mixture component identified by mash(left).

d. Mashr estimates weight in MAIT, memory B, $CD8^+$ naive T and $CD8^+$ Tem.



Extended Data Fig. 5 | Shared and specific effects across timepoints

a. Mashr estimates wight in different cell types.

b. The number of ligand-receptor interactions going between the different cell types to, analyzed by CellChat.

c. Heatmap shows the relative strength of each signal pathway network for each cluster with both incoming and outgoing signaling patterns.

Extended Data Fig. 6 | Significant gene functional enrichment shared by the time points considered in all cell types

Heatmap showed differential pathways enriched in celltype according to GSVA score.

Extended Data Fig. 7 | Peak-to-gene links identified in the each participants

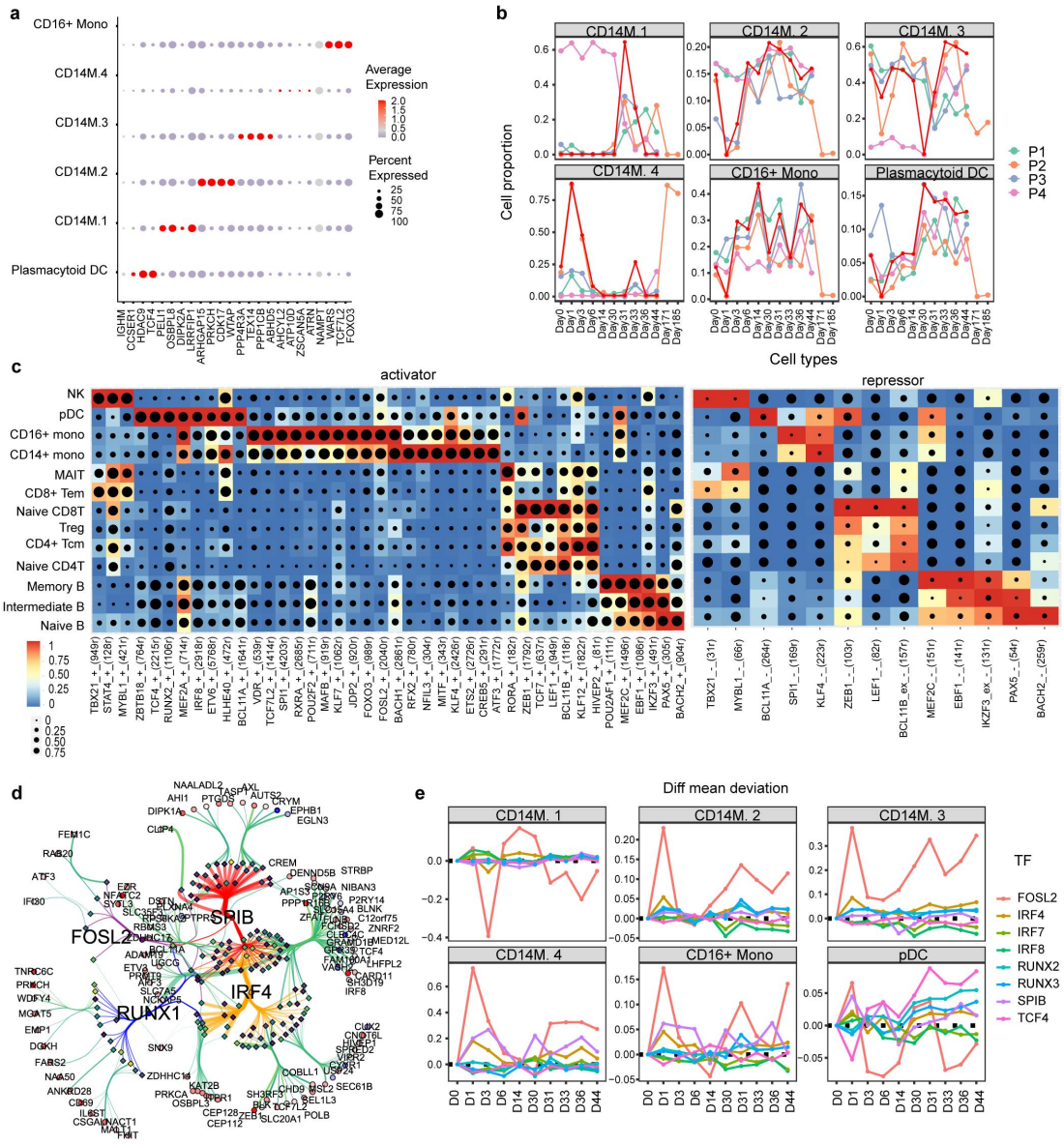
a. Histogram showing frequency of peak-to-gene correlations. We retained only peak-to-gene pairs with correlations greater than 0.45

b. Histogram showing the distribution of number of genes per distal peak.

c. Histogram showing the distribution of number of distal peaks per gene.

d. Row-scaled heatmaps of statistically significant distal peak-to-gene links in P1-P4, respectively. Each row represents the expression of a gene (right) correlated to the accessibility of a distal peak (left). The heatmaps represent peak-to-gene analyses using only cells from the indicated participant.

e. GO enrichment analysis of the 42 genes obtained in figure 3c.



Extended Data Fig. 8 | Regulational network for chromatin remodeling in myeloid sub-celltype

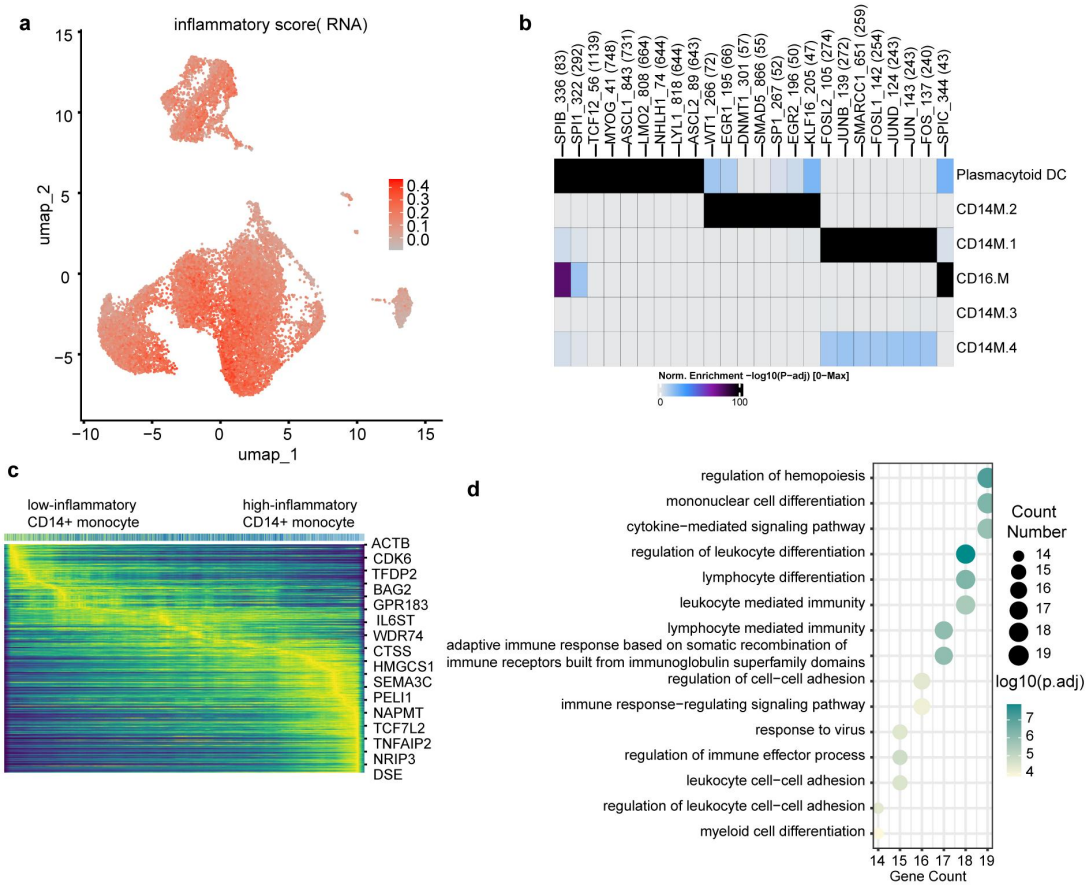
a. Dot plots of gene expression for myeloid sub-celltype where the color is scaled by mean expression and the dot size is proportional to the percent of the population expressing the gene.

b. Line chart of the cell proportion of six myeloid sub-celltype in four participants across 10 timepoints (P2 individual has 12 timepoints). Red dots indicate the geometric mean of the proportion of celltypes across the four participants at a time point.

c. Heatmap-dotplot showing TF expression of the eRegulon (TFs are linked to their target regions and these to their target genes) on a color scale and cell type specificity (RSS) of the eRegulon on a size scale.

d. Visualization of eGRN formed by SPIB, RUNX1, IRF4 and FOSL2 TF target nodes are restricted to highly variable genes and regions.

e. Line graph showing the difference in transcription factor (TF) accessibility during vaccination. Each line represents a TF.



Extended Data Fig. 9 | Gene transcription patterns of CD14⁺ monocytes

a. UMAP displaying inflammatory scores per cell.

b. Heatmap of transcription factor motifs enriched in marker peak of six myeloid sub-celltypes. The color represents the normalized motif enrichment score calculated in ArchR using HOMER with the hypergeometric test.

c. Heatmaps of gene expressions as a function of latent time. Color represents smoothed spliced counts.

d. GO enrichment results based on genes whose transcription mode was chromatin closing begins after transcriptional repression.