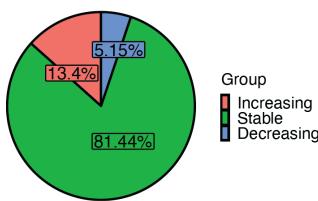


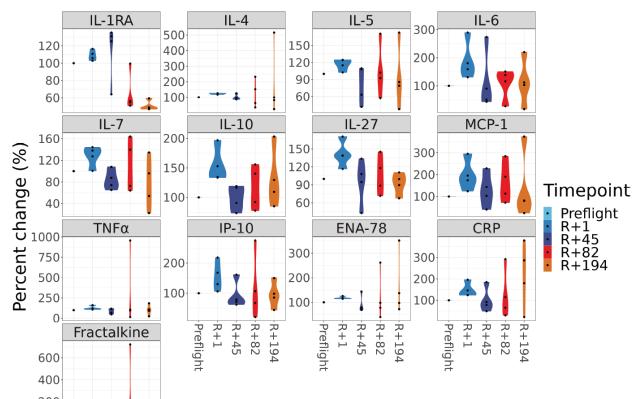
1 **Supplementary Figures**

2 **Supplementary Figure 1**

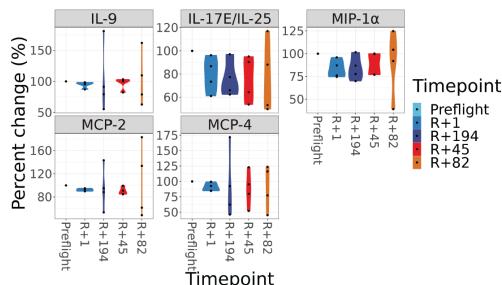
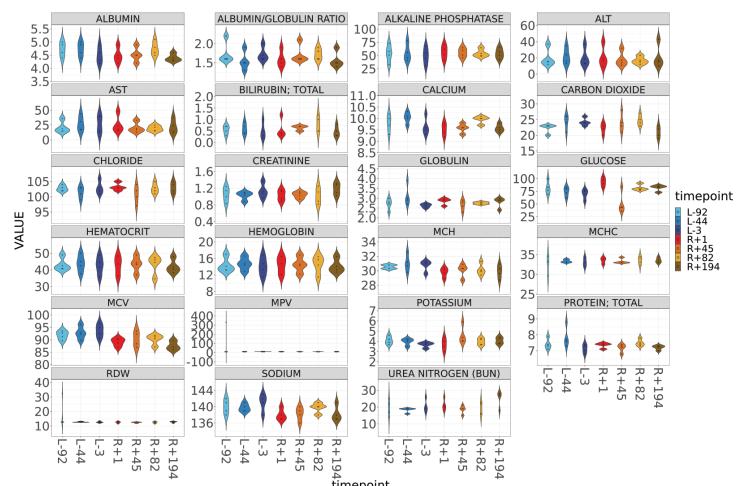
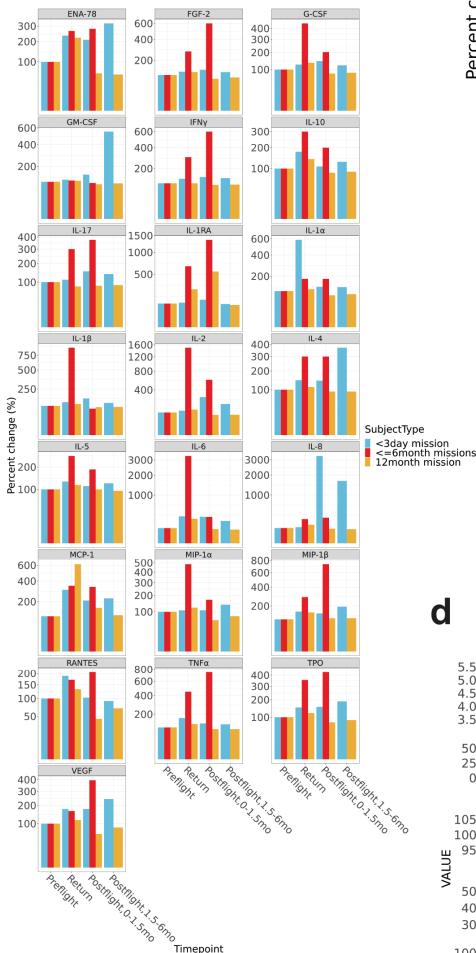
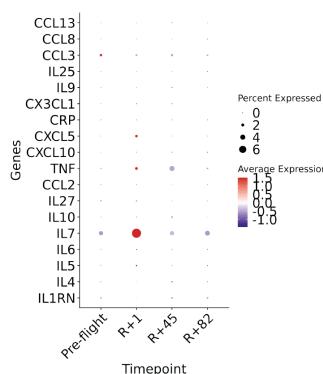
3 a, Pie chart of significantly increased, significantly decreased, and stable BCPs. b, The
4 percentage change of the significantly increased BCPs (Top) and decreased BCPs (Bottom) over
5 time. Each dot represents the value of each crew (Wilcoxon-rank sum test, p-value < 0.05, two-
6 sided). c, Percent change of the selected biochemical profiles in the <3 day (i4 mission), < 6
7 months (28 astronauts), and 1 year (NASA Twins study). d, Expression profiles of the
8 comprehensive metabolic blood panel (CMP) over time. Each dot represents the value of each
9 crew. e, Dot plot of 18 significantly changed BCPs (from Supplementary Fig. 1b) in pseudo-bulk
10 PBMC. Source data are provided as a Source Data file.

a**b**

Significantly up

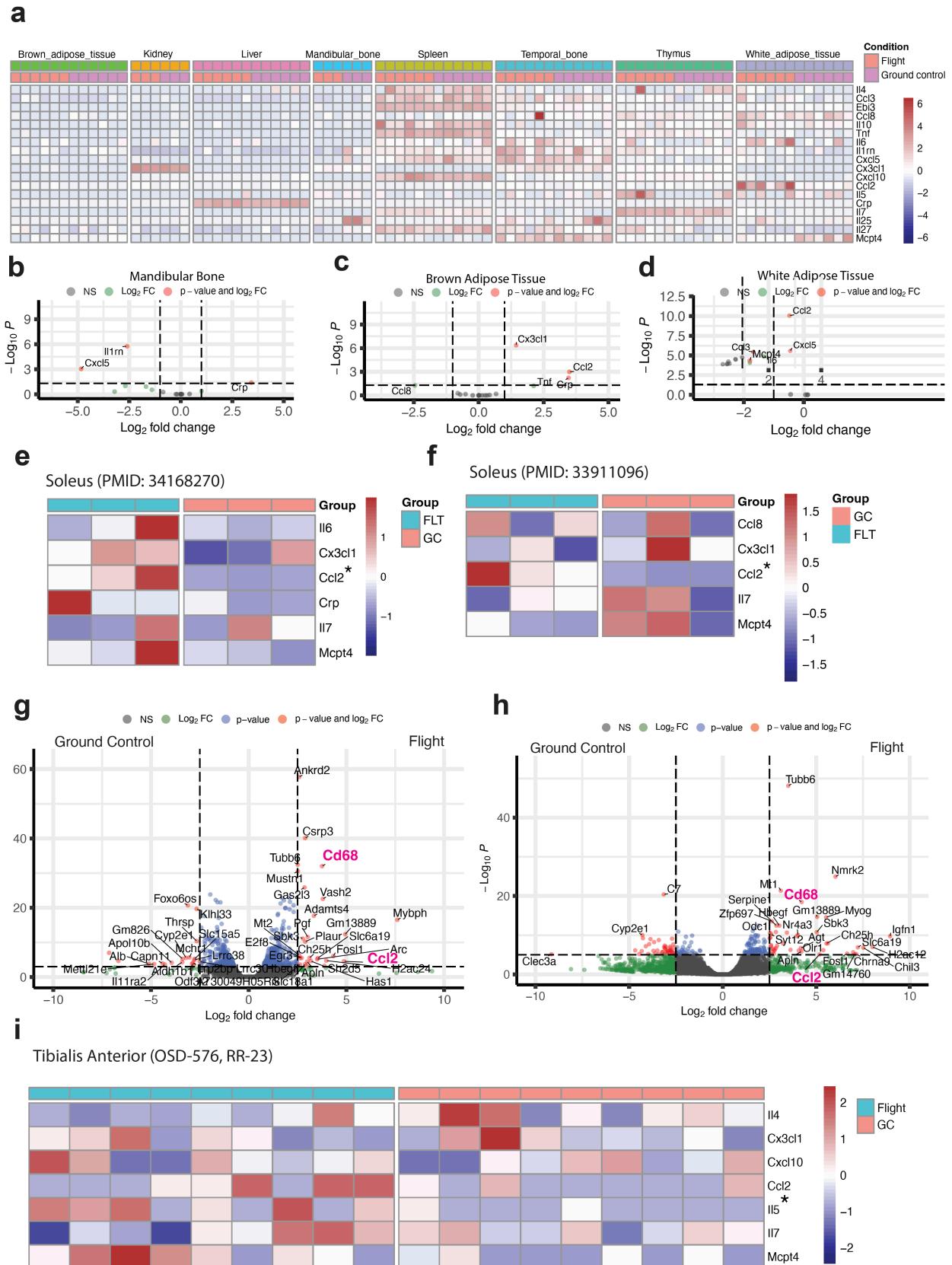


Significantly down

**d****c****e**

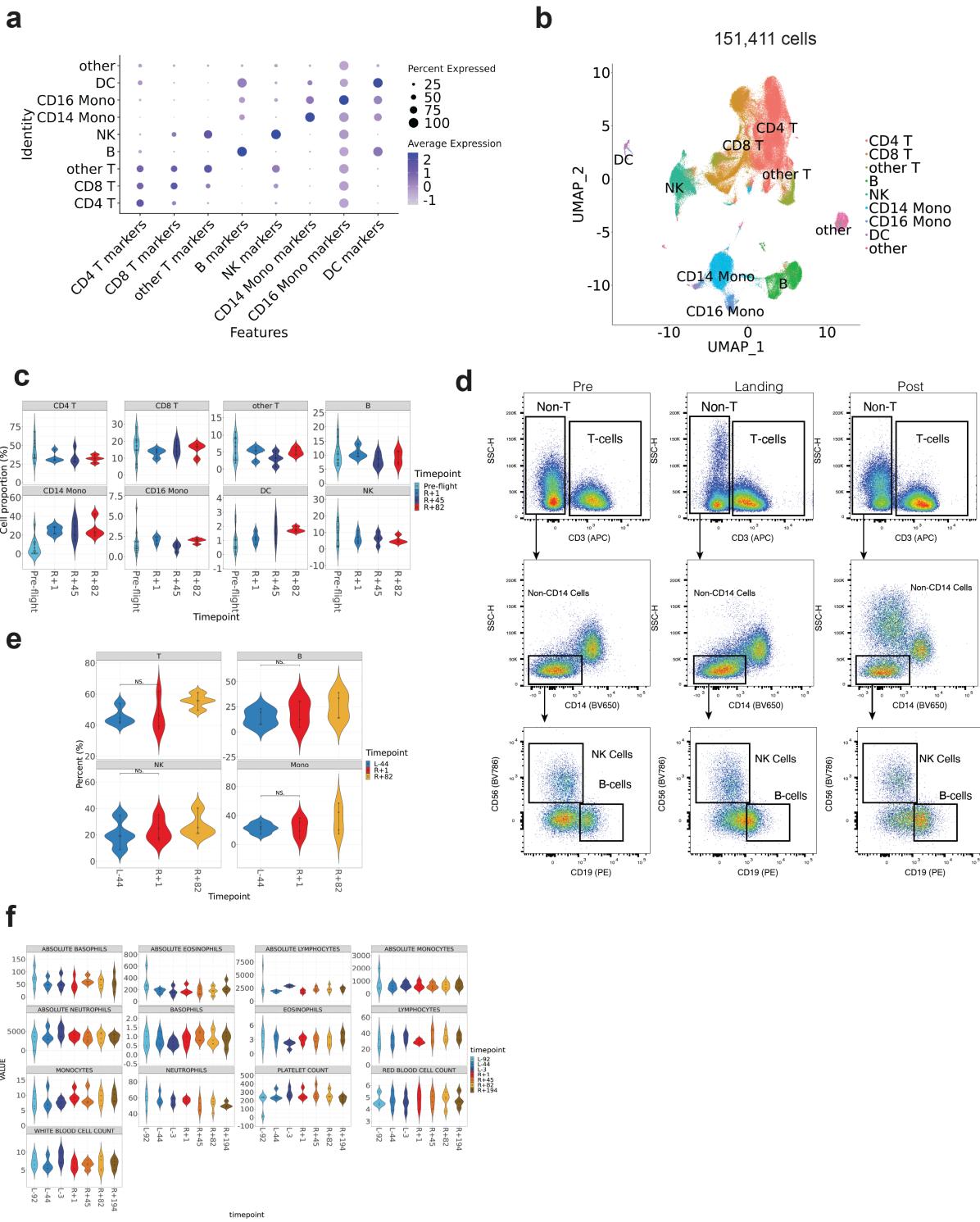
12 **Supplementary Figure 2**

13 a, Heatmap of the transcriptomic expression in mouse tissues of the cytokines found significantly
14 changed in the serum of the i4 crew. b-d Volcano plots of the mouse tissues in which the
15 cytokines shown in panel (a) were found differentially expressed after landing. e-f, Heatmaps
16 with the expression of the cytokines in panel (a) in the soleus muscle comparing flight (FLT) to
17 ground control (GC) in the datasets PMID 34168270 (e) and PMID:33911096 (f). g-h, Volcano
18 plots comparing gene expression between flight and ground control in the soleus from PMID
19 34168270 (g) and PMID:33911096 (h). i. Heatmap of the expression in the tibialis anterior
20 muscle (OSD-576) of the cytokines found significantly changed in the serum of the i4 crew.
21 Wald test (two-sided) was used to identify differentially expressed genes, with raw p-values, (*)
22 means p-value <0.05 in panels e,f, and i. Source data are provided as a Source Data file.



24 **Supplementary Figure 3**

25 a, Activity score of PBMC subpopulation markers in the annotated i4 PBMC. b, UMAP of
26 151,411 cells derived from single-nuclei multi-ome (GEX+ATAC) with subpopulation
27 annotation. c, Cellular profiles of i4 PBMC over time calculated from single-nuclei multi-ome
28 data. Each dot represents each crew. d, FACS gating of T cell, B cell, NK cell, and monocytes
29 from the i4 PBMC. e, Cellular profiles of i4 PBMC over time calculated from FACS. Paired
30 Student's t-test. f, CBC profiles of i4 crews over time. Each dot represents each crew. Source
31 data are provided as a Source Data file.



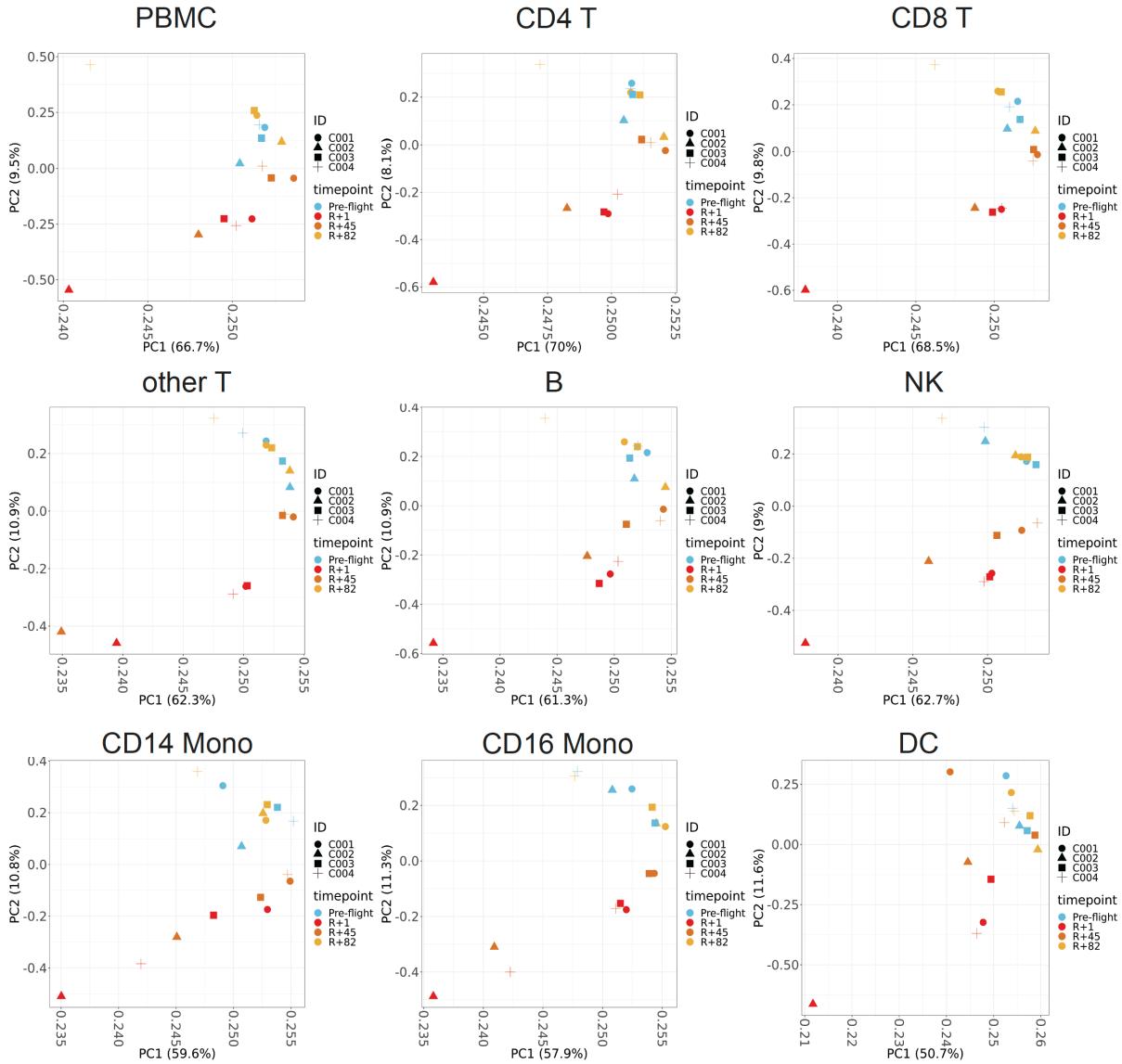
32

33

34

35 **Supplementary Figure 4**

36 Principal component analysis of single-nuclei GEX of the i4 immune cells. Shape represents ID.
 37 Color represents timepoints. Source data are provided as a Source Data file.



38

39

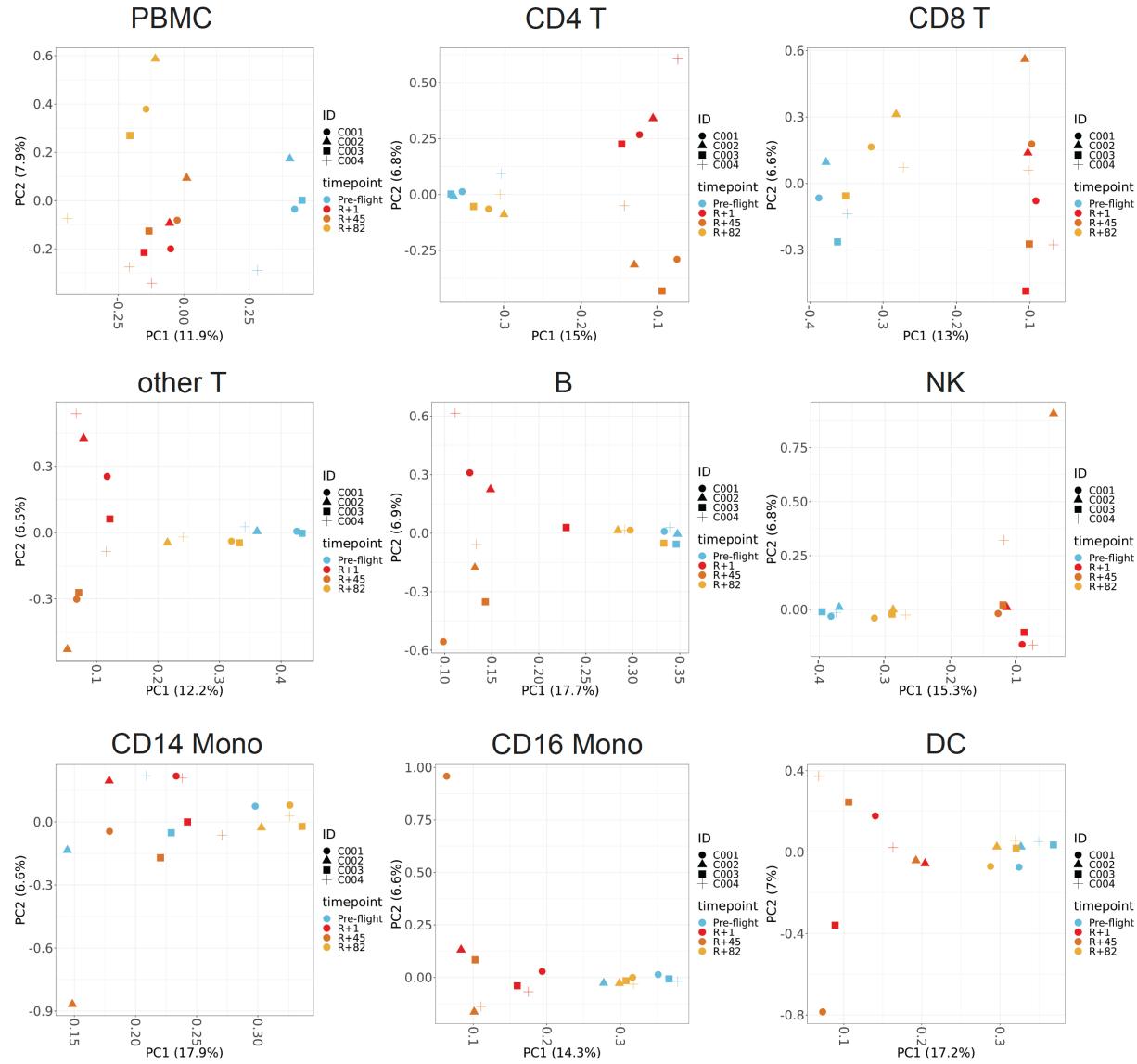
40

41

42

43 **Supplementary Figure 5**

44 Principal component analysis of single-nuclei ATAC of the i4 immune cells. Shape represents
 45 ID. Color represents timepoints. Source data are provided as a Source Data file.



46

47

48

49

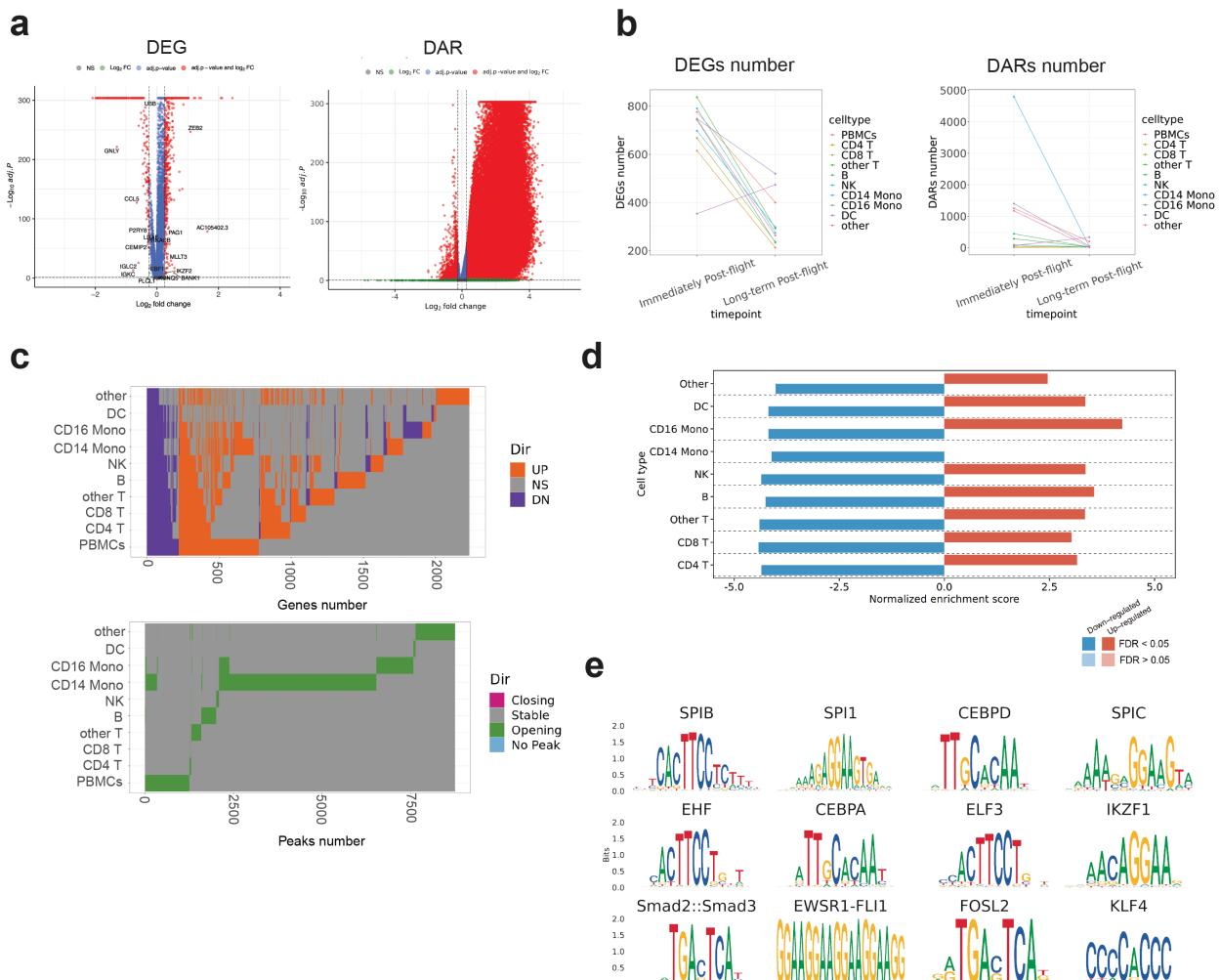
50

51

52 **Supplementary Figure 6**

53 a, Volcano plot of the pseudo-bulk PBMCs DEGs (Left) and DARs (Right). A Wilcoxon rank-
 54 sum test to identify differentially expressed genes between clusters, with raw p-values adjusted
 55 for multiple testing using the Bonferroni correction to control the family-wise error rate (FWER).
 56 b, The number of DEGs and DARs identified post-flight (R+1) and the long-term post-flights
 57 (R+45 and R+82) from PBMC and subpopulations. c, Overlap of up- and down-regulated DEGs
 58 and DARs among PBMC and subpopulations. d, GSEA analysis the i4 immune cells DEGs with
 59 the ‘spaceflight signatures in the i4 astronauts’ DEGs. A one-sided permutation-based test to
 60 determine the significance of gene set enrichment, with raw p-values adjusted for multiple
 61 testing using the Benjamini-Hochberg procedure to control the false discovery rate (FDR). e,
 62 Top 12 enriched DNA motifs of PBMC that are over-represented in a set of peaks that are
 63 differentially accessible between R+1 and pre-flight. Source data are provided as a Source Data
 64 file.

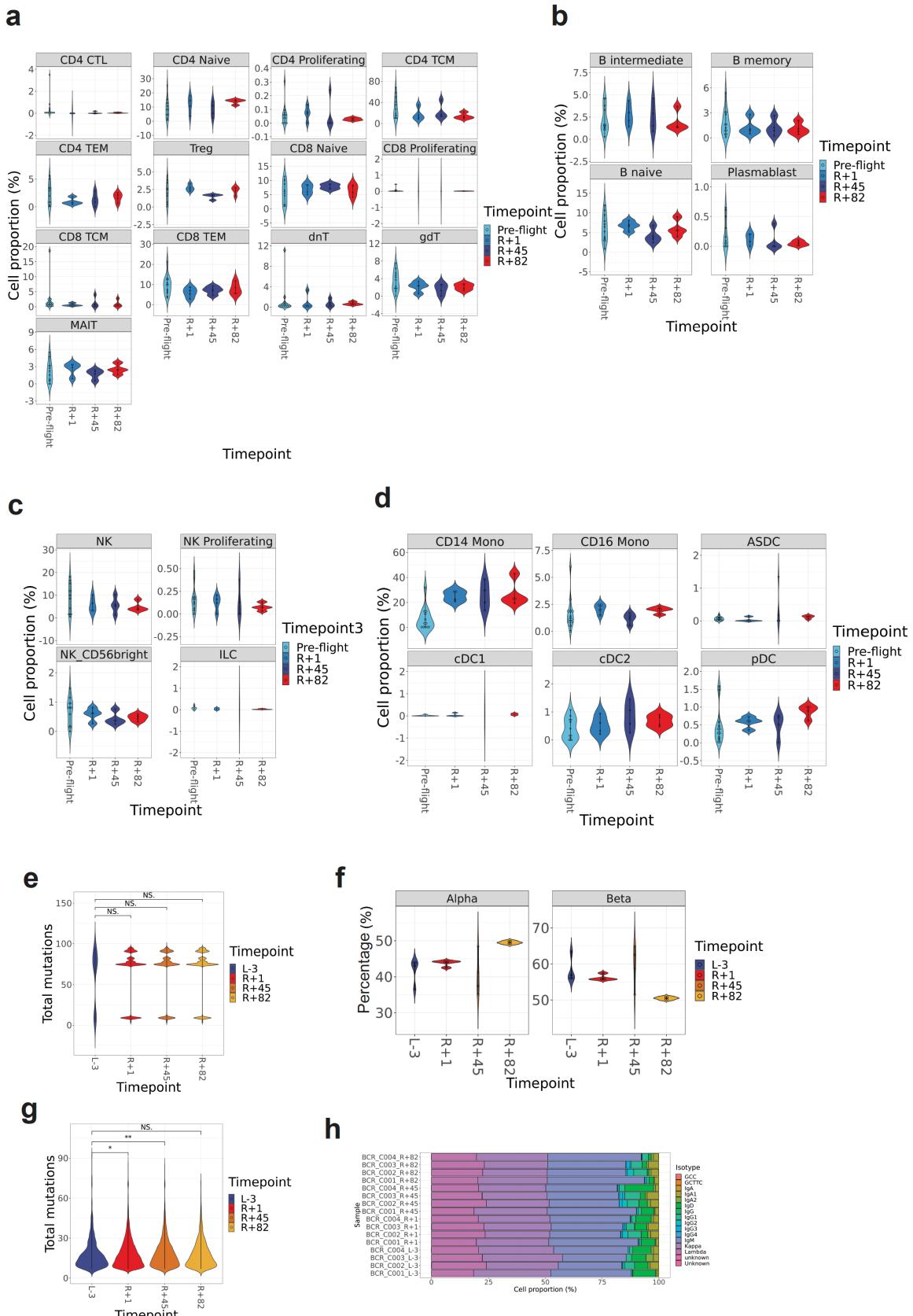
65



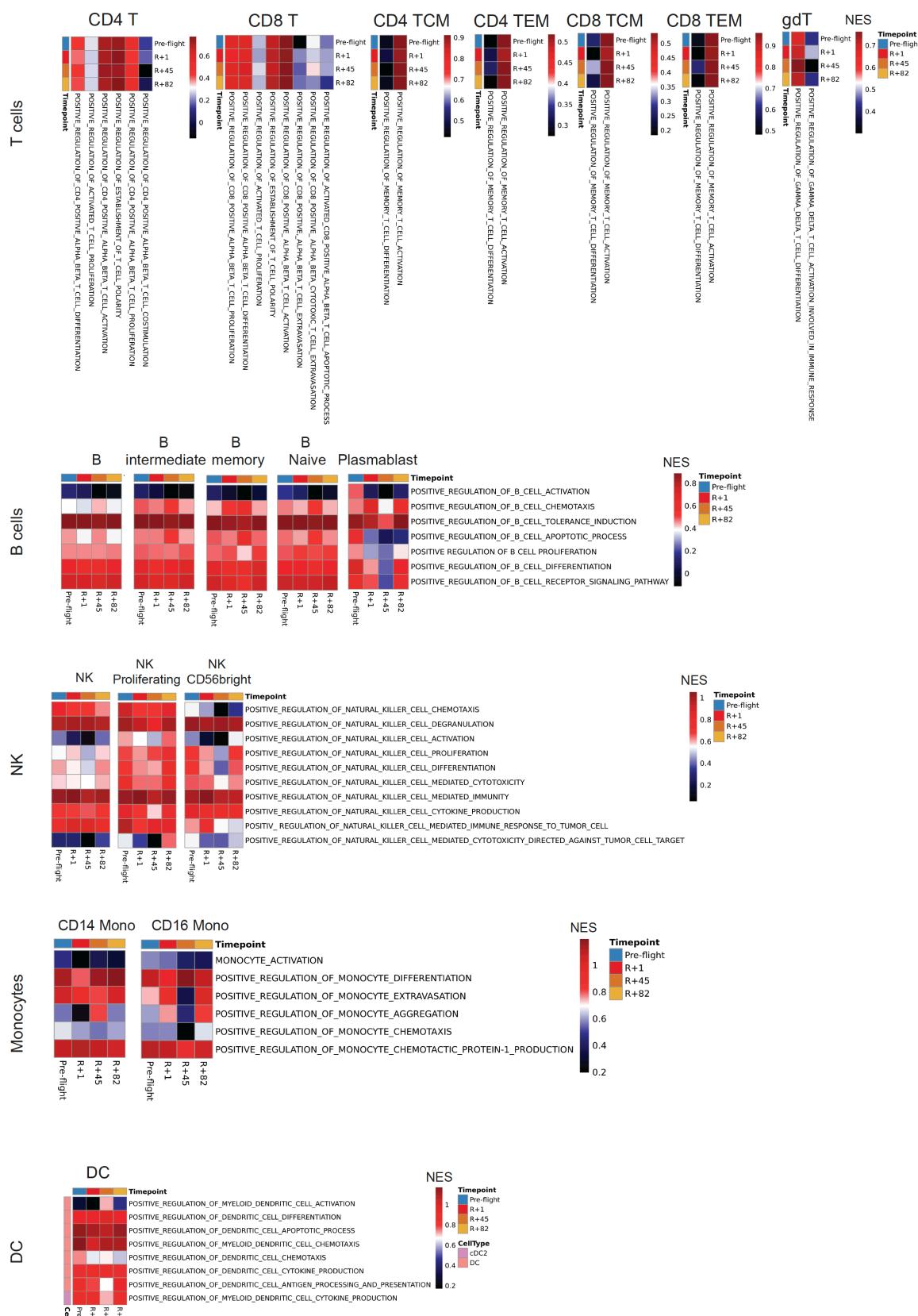
66

67

68 **Supplementary Figure 7**
69 a, Cellular profiles of T cells from single-nuclei multi-ome data over time. Each dot represents
70 each crew. b, Cellular profiles of B cells from single-nuclei multi-ome data over time. Each dot
71 represents each crew. c, Cellular profiles of NK cells from single-nuclei multi-ome data over
72 time. Each dot represents each crew. d, Cellular profiles of i4 monocytes and DCs from single-
73 nuclei multi-ome data over time. Each dot represents each crew. e, The number of total
74 mutations in TCR from L-3 to R+82 (Wilcoxon-rank sum test, p-value < 0.05). f, The
75 percentage of gene types of TCR from L-3 to R+82. g, The number of total mutations in BCR
76 from L-3 to R+82 (Wilcoxon-rank sum test, *p-value < 0.05, ** p-value < 0.01). h, The
77 percentage of antibody isotypes of BCR from L-3 to R+82. Source data are provided as a Source
78 Data file.

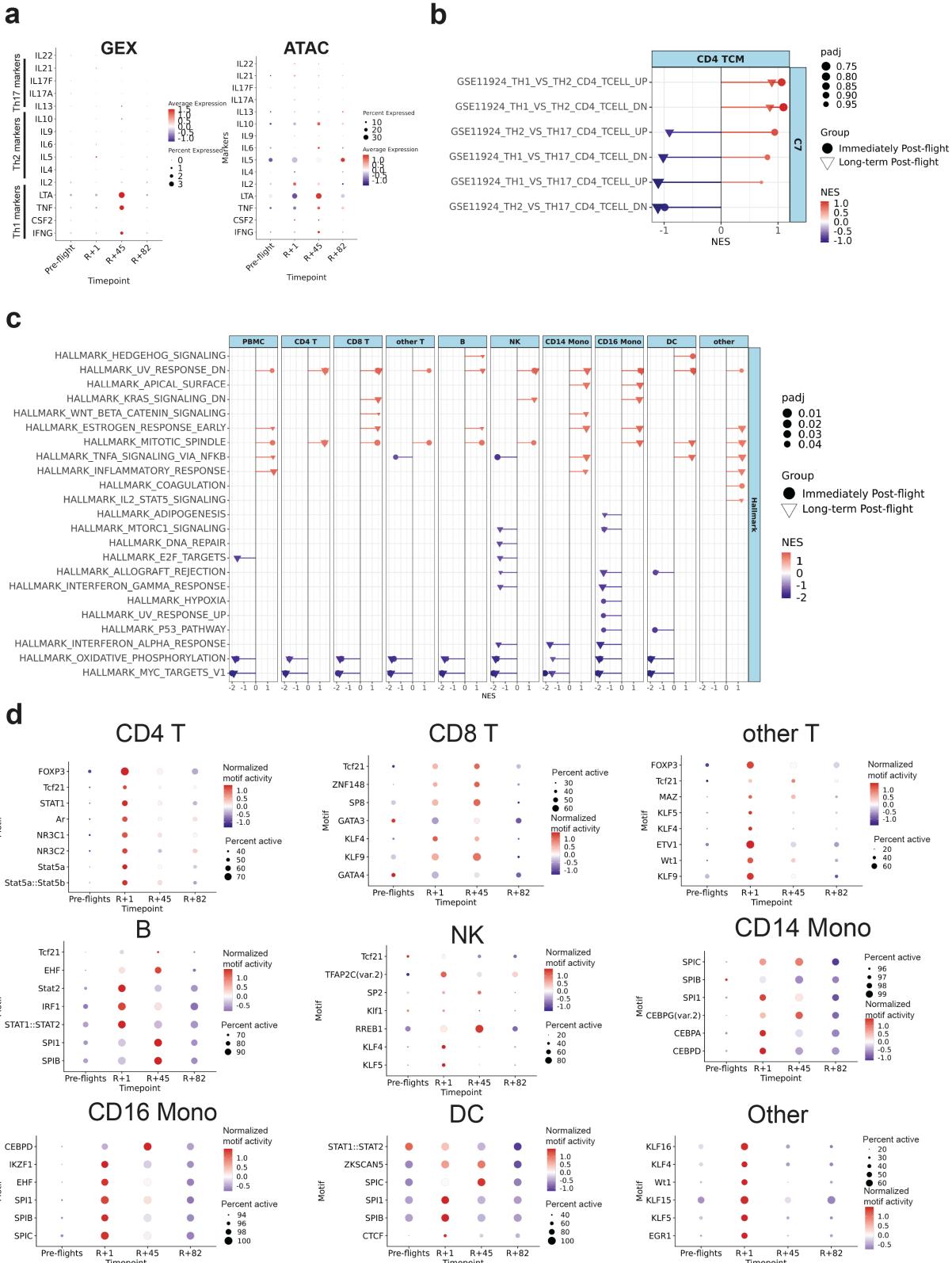


80 **Supplementary Figure 8**
81 Enrichment score of the relevant immune function pathways of T cell (top row), B cell (2nd
82 row), NK cell (3rd row), monocytes (4th row), and dendritic cell (bottom row) subpopulations
83 before (Pre-flight: L-92, L-44, L-3) and after spaceflight (Post-flight: R+1, R+45, R+82). Source
84 data are provided as a Source Data file.



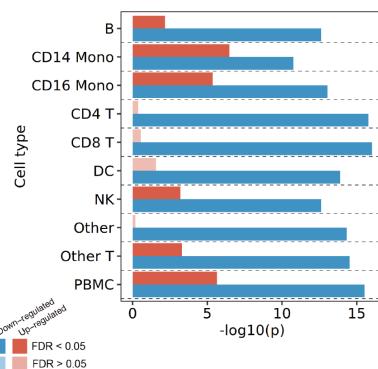
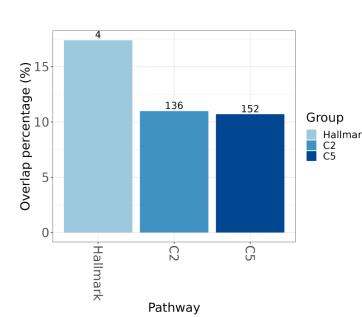
86 **Supplementary Figure 9**

87 a, Dot plots of Th1-, Th2-, and Th17-secreted cytokines in CD4 TCM cells over time (Left: gene
88 expression, Right: ATAC derived gene expression). b, GSEA of CD4 TCM cells with the
89 MSigDB C7 GSE1924 for Th1, Th2, Th17 CD4 T cell gene expression signatures. A one-sided
90 permutation-based test to determine the significance of gene set enrichment, with raw p-values
91 adjusted for multiple testing using the Benjamini-Hochberg procedure to control the false
92 discovery rate (FDR). c, GSEA of PBMC and subpopulations with MSigDB Hallmark pathway
93 (padj < 0.05). A one-sided permutation-based test to determine the significance of gene set
94 enrichment, with raw p-values adjusted for multiple testing using the Benjamini-Hochberg
95 procedure to control the false discovery rate (FDR). d, Activity scores of top enriched motifs
96 from PBMC subpopulations. Source data are provided as a Source Data file.



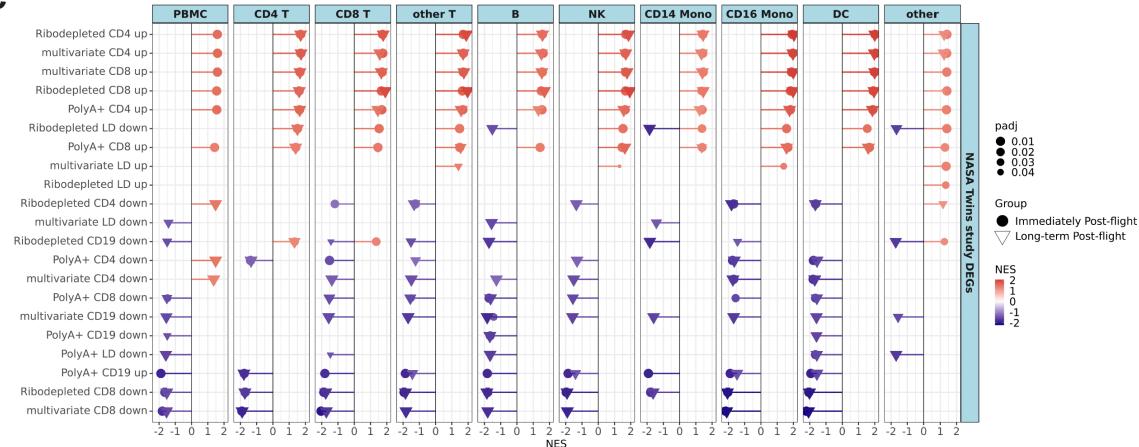
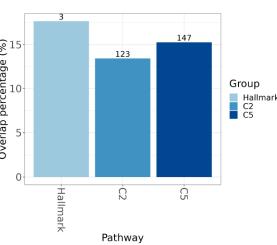
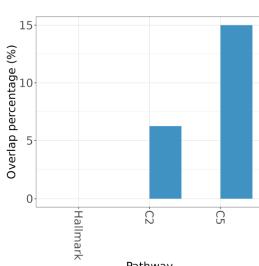
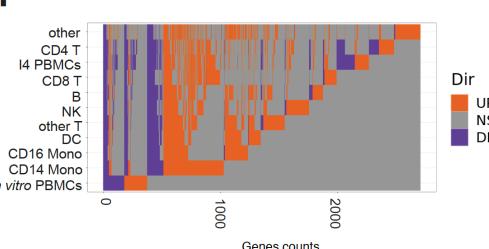
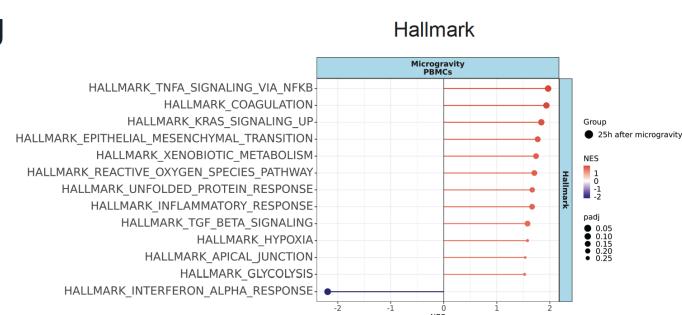
98 **Supplementary Figure 10**

99 a, Fisher's exact test of 'spaceflight signature in mice' with 'spaceflight signature in
100 Inspiration4 astronauts'. P-values for the overlaps were adjusted by multiple tests using
101 Benjamini-Hochberg method. b, Overlap percentage of GSEA of the 'spaceflight signatures in
102 mice' with i4. The number on the bar represents the number of overlapping pathways (Fisher's
103 exact test. padj; Hallmark: 3.844985e-02, C2: 3.317221e-47, C5: 2.341219e-83). P-values for the
104 overlaps were adjusted by multiple tests using Benjamini-Hochberg method. c, GSEA of i4
105 PBMCs and subpopulations at the immediately post-flight (R+1) and long-term post-flights
106 (R+45 and R+82) with up-regulated and down-regulated DEGs of NASA Twin study (padj <
107 0.05). LD: lymphocyte depleted. A one-sided permutation-based test to determine the
108 significance of gene set enrichment, with raw p-values adjusted for multiple testing using the
109 Benjamini-Hochberg procedure to control the false discovery rate (FDR). d, Overlap percentage
110 of GSEA of the i4 spatial transcriptomics with the i4 immune cells. The number on the bar
111 represents the number of overlapping pathways. Fisher's exact test. P-values for the overlaps
112 were adjusted by multiple tests using Benjamini-Hochberg method. e, Overlap percentage of
113 GSEA of the i4 EVP and plasma proteomics with the i4 immune cells. The number on the bar
114 represents the number of overlapping pathways. Fisher's exact test. P-values for the overlaps
115 were adjusted by multiple tests using Benjamini-Hochberg method. f, Altuna plot represents the
116 overlap of up-regulated DEGs (Orange) and down-regulated DEGs (Purple) from i4 PBMCs and
117 subpopulations and the core 375 DEGs of *in vitro* microgravity simulated PBMCs. g, Gene set
118 enrichment analysis of 375 core DEGs of *in vitro* microgravity simulated DEGs with the
119 MSigDB hallmark (Top, padj < 0.3). A one-sided permutation-based test to determine the
120 significance of gene set enrichment, with raw p-values adjusted for multiple testing using the
121 Benjamini-Hochberg procedure to control the false discovery rate (FDR). Source data are
122 provided as a Source Data file.

a**b**

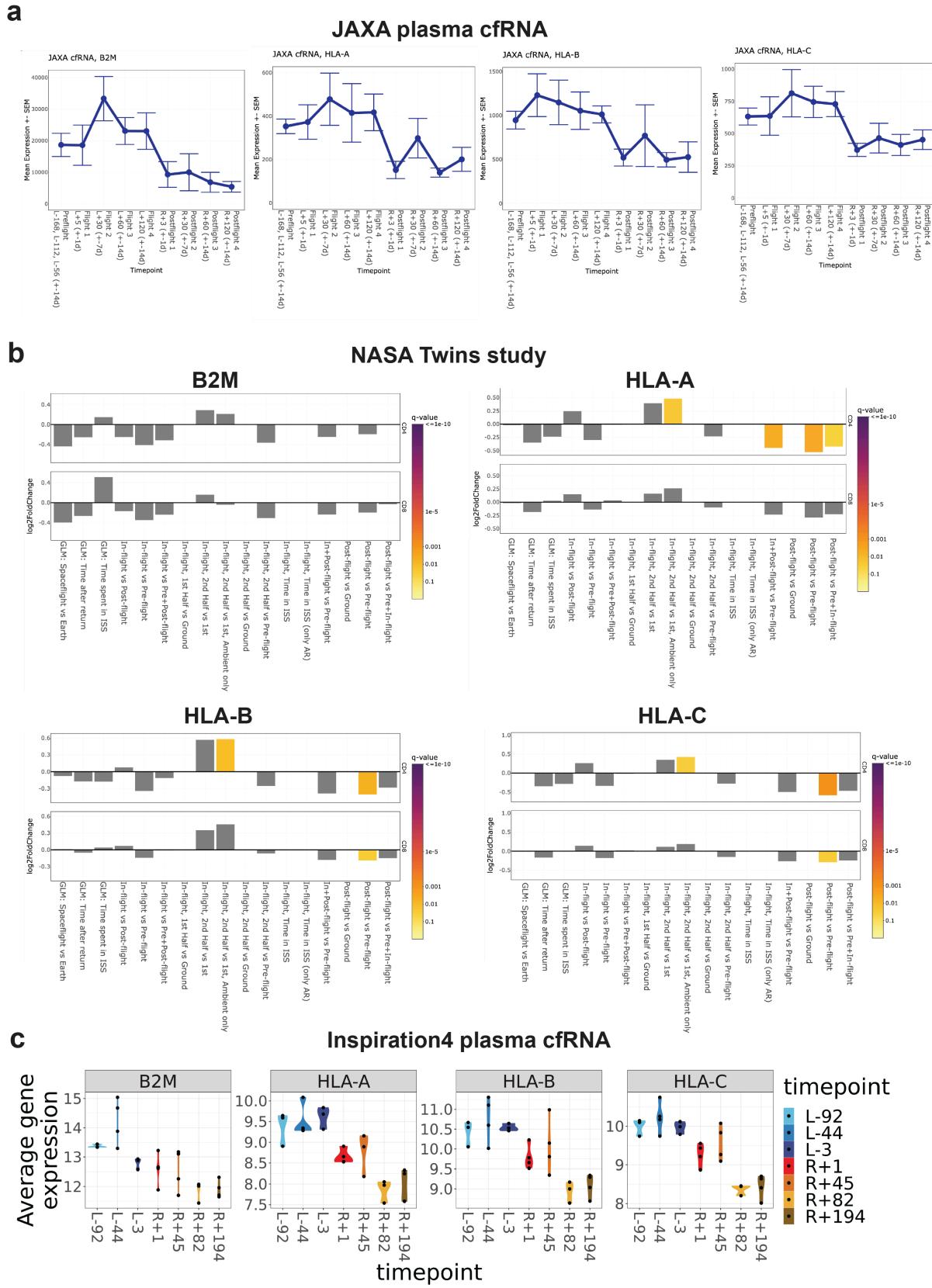
Representative overlapping pathways

OXIDATIVE_PHOSPHORYLATION
MYC_TARGETS_V1
UV_RESPONSE_DN
MITOTIC_SPINDLE
TCF21_TARGETS_2_DN
UV_RESPONSE_VIA_ERCC3_COMMON_DN
INFECTIOUS_DISEASE
VIRAL_GENE_EXPRESSION
RESPIRATORY_ELECTRON_TRANSPORT
INFLUENZA_INFECTION
RIBOSOME
TRANSLATION

c**d****e****f****g**

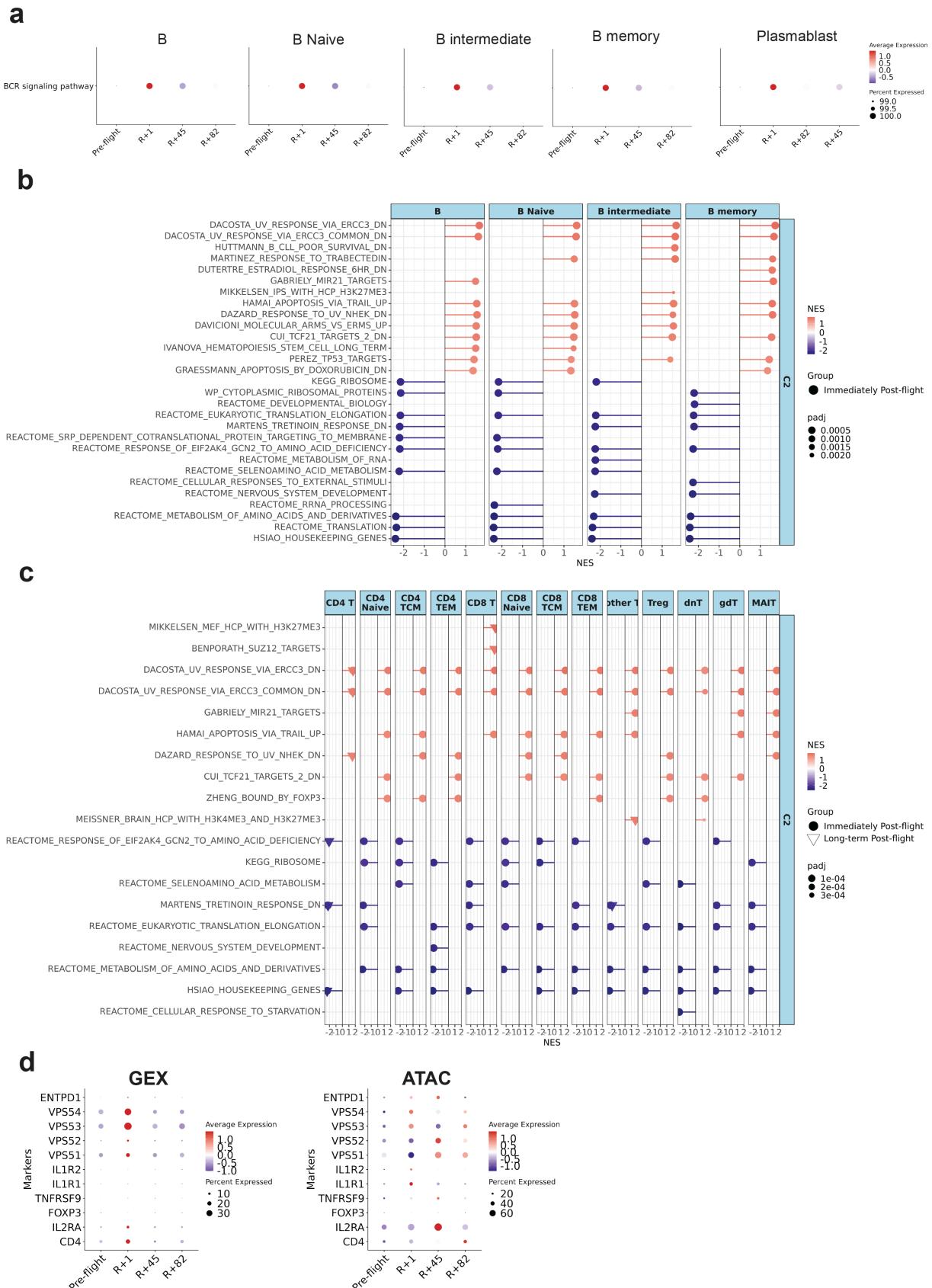
124 **Supplementary Figure 11**

125 a, MHC class I gene expression in JAXA CFE plasma cfRNA. Error bar represents the SEM. b,
126 log2FoldChange and q-value of MHC class I genes in CD4 T and CD8 T cells of NASA Twins
127 study. Negative binomial test with adjustment only within each comparison separately. c, MHC
128 class I gene expression in Inspiration4 plasma cfRNA. Source data are provided as a Source Data
129 file.



131 **Supplementary Figure 12**

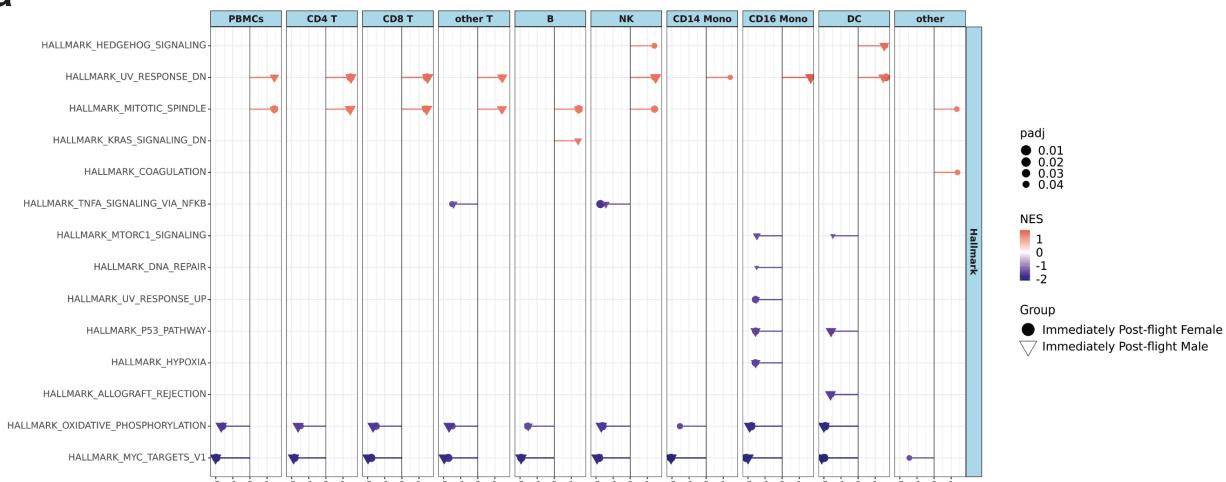
132 a, Dot plots of BCR signaling pathway scores in B cells over time. b, GSEA of B cells with the
133 MSigDB hallmark and C2 pathway (C2: top7 of NES > 0 and NES < 0, padj < 0.05). A one-sided
134 permutation-based test to determine the significance of gene set enrichment, with raw p-values
135 adjusted for multiple testing using the Benjamini-Hochberg procedure to control the false
136 discovery rate (FDR). c, GSEA of T cells with the MSigDB hallmark and C2 pathway (C2: top7
137 of NES > 0 and NES < 0, padj < 0.05). A one-sided permutation-based test to determine the
138 significance of gene set enrichment, with raw p-values adjusted for multiple testing using the
139 Benjamini-Hochberg procedure to control the false discovery rate (FDR). d, Dot plot of Treg
140 markers and Treg activation markers in T cells (Left: gene expression. Right: ATAC derived
141 gene expression). Source data are provided as a Source Data file.



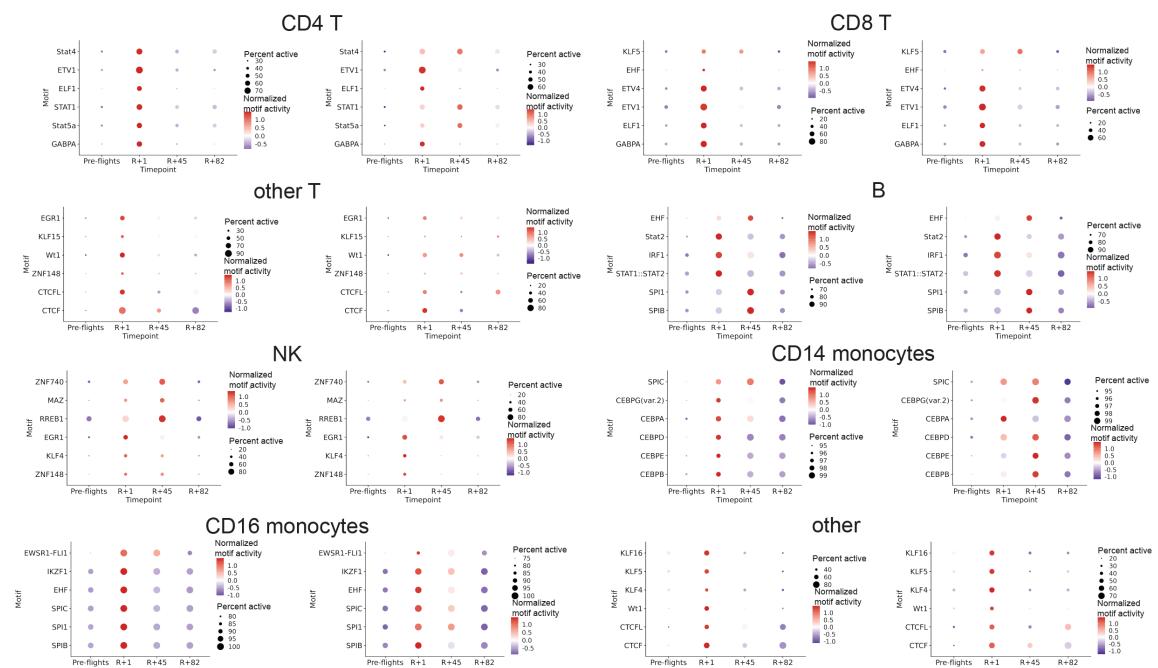
143 **Supplementary Figure 13**

144 a, GSEA of PBMC and subpopulations from females and males at R+1 with the MSigDB
 145 hallmark ($\text{padj} < 0.2$). A one-sided permutation-based test to determine the significance of gene
 146 set enrichment, with raw p-values adjusted for multiple testing using the Benjamini-Hochberg
 147 procedure to control the false discovery rate (FDR). b, Activity scores of top enriched motifs
 148 from PBMC subpopulations separated by sex. Source data are provided as a Source Data file.

a



b



149

Female

Male

Female

Male

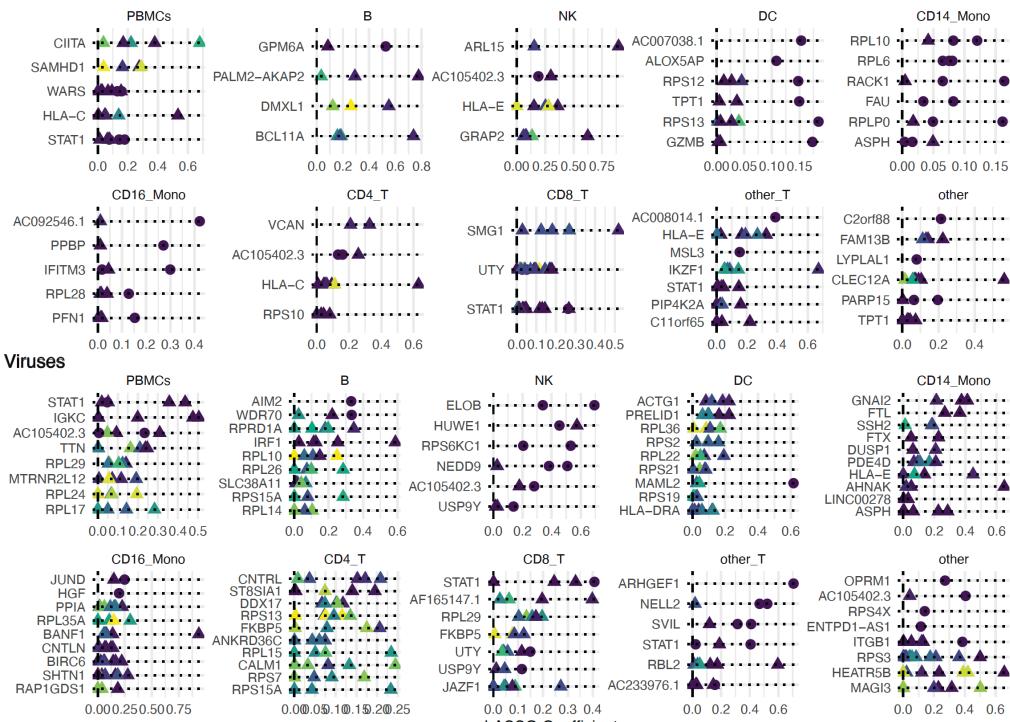
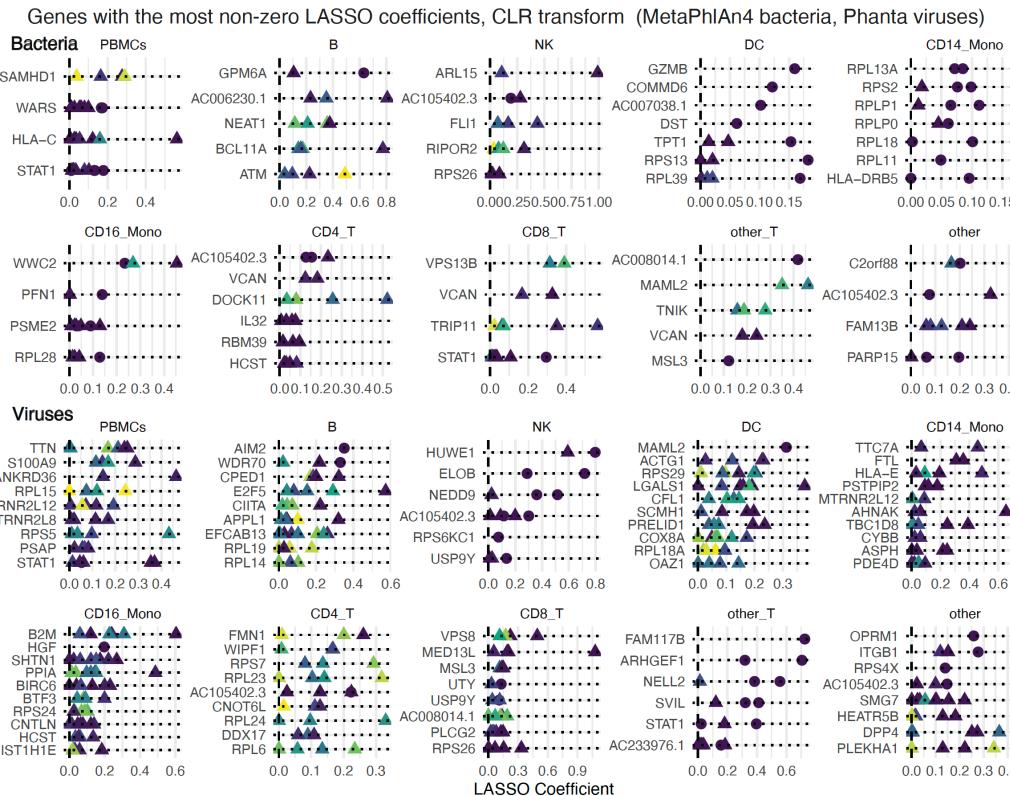
150

151

152 **Supplementary Figure 14.** Comparing association stability across taxonomic classifiers. a, Log-
153 transformed associations on MetaPhlAn4-classified bacterial species and Phanta-classified
154 viruses. Each point in the plot bodies represent a different bacterial species (top) or viral genus
155 (bottom). The y-axes describe different human genes. For each cell type, we ranked genes with
156 non-zero LASSO coefficients first by the number of Bonferroni < 0.2 findings, then by the total
157 number of nominally associated ($p\text{-value} < 0.05$) microbial features (bacteria or viruses). We
158 report up to ten human genes per sub-panel. b, This panel contains the same ranking and plotting
159 strategy as panel a, except the associations were computed on CLR-transformed data. Lasso
160 regression and the mixed effect linear regression approach were used for $p\text{-value}$ estimation
161 (two-sided). Bonferroni correction was used to adjust for multiple hypothesis testing. Source data
162 are provided through the github link.

a Genes with the most non-zero LASSO coefficients, log transform (MetaPhlAn4 bacteria, Phanta viruses)

Bacteria

**b**

165 **Supplementary Figure 15.** Alternative methods for exploring microbiome-immune
166 associations. a, We computed LASSO and mixed modeling associations using Center-Log-Ratio
167 instead of log-transformed data. This was done to compare that our results were stable under
168 multiple compositional data analysis methods. The three bars in each sub-panel correspond to the
169 number of associations in the “real” (log-transformed) data versus CLR data and the overlap
170 therein at different stringency levels in controlling for false positives. b, The human genes, per
171 cell type, with the greatest number of microbial associations that themselves had low or
172 Bonferroni-significant p-values. Each point in the plot bodies represents a different bacterial
173 species (top) or viral genus (bottom). For each cell type, we ranked genes with non-zero LASSO
174 coefficients first by the number of Bonferroni < 0.2 findings, then by the total number of
175 nominally associated ($p\text{-value} < 0.05$) microbial features (bacteria or viruses). We report up to
176 ten human genes per sub-panel. Lasso regression and the mixed effect linear regression approach
177 were used for p-value estimation (two-sided). Bonferroni correction was used to adjust for
178 multiple hypothesis testing. Source data are provided through the github link.

