PNAS
www.pnas.org

Supplementary Information for

Single-cell analyses highlight the proinflammatory contribution of C1q-high monocytes to Behçet's disease

Wenjie Zheng[1†*], Xiaoman Wang[2†*], Jinjing Liu[1†], Xin Yu[1†], Lu Li[1,3], Heping Wang[2], Jijun Yu[4,5], Xiaoya Pei[2], Chaoran Li[1,6], Zhimian Wang[1], Menghao Zhang[1], Xiaofeng Zeng[1], Fengchun Zhang[1], Chenfei Wang[7],Hua Chen[1], Hou-Zao Chen[2*]

***Corresponding author**: Wenjie Zheng, Xiaoman Wang, Hou-Zao Chen
Email: zhengwj@pumch.cn, wangxm815@ibms.pumc.edu.cn, chenhouzao@ibms.cams.cn

This PDF file includes:

Supplementary Figures

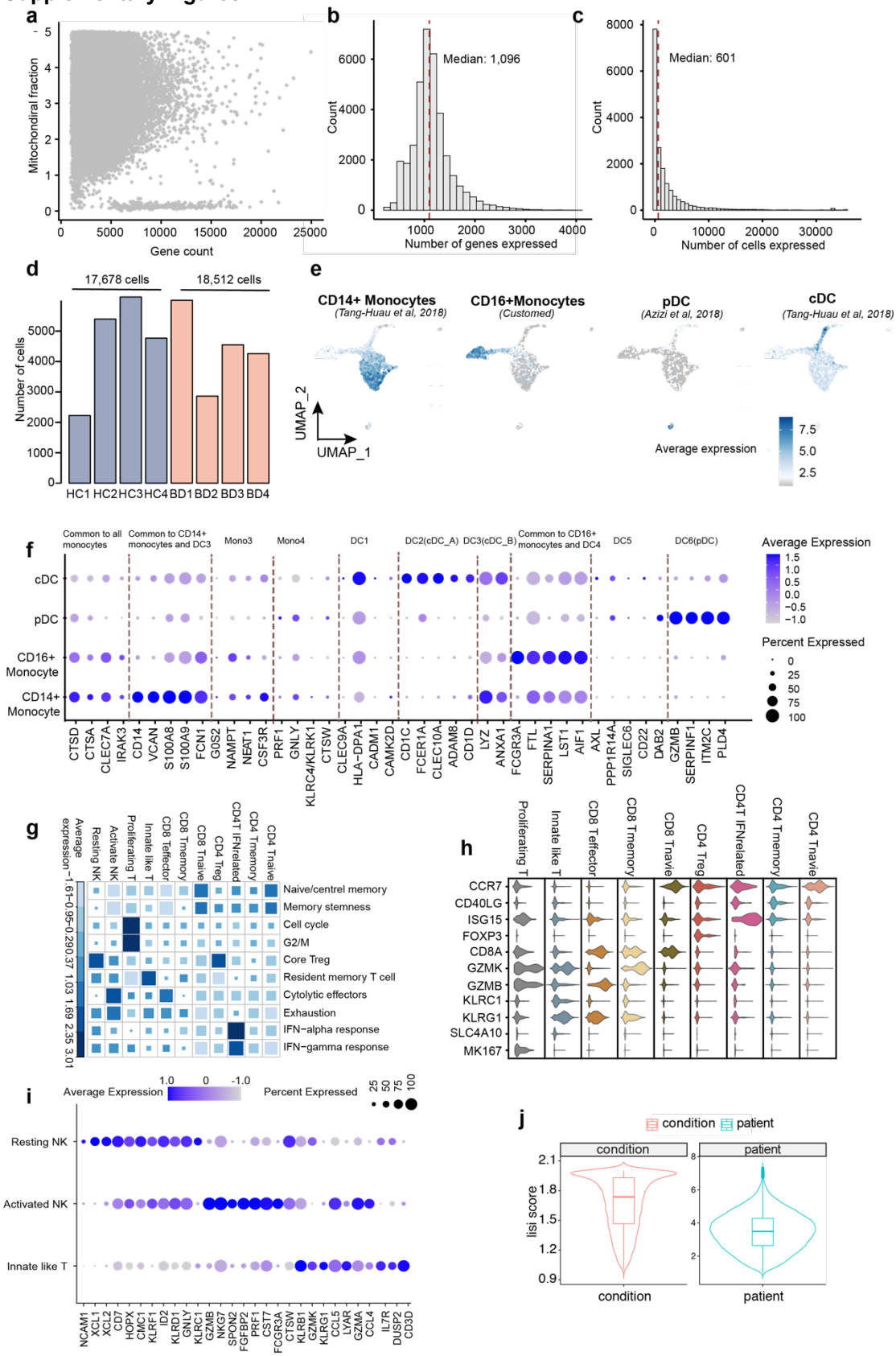

Figure S1 Quality control and cell type identification in scRNA-seq data of PBMCs. **(a)** Scatter plot showing the mitochondrial fractions in each cell (dot) after removing cells with a high mitochondrial fraction (cutoff, 5%). **(b)** Distribution of the number of genes detected in each cell in scRNA-seq data of PBMCs. **(c)** Distribution of the number of cells expressing the genes that were identified in the scRNA-seq data of PBMCs. **(d)** Measured cell number (Y-axis) in individual samples (X-axis). **(e)** UMAP plot showing myeloid populations colored by the average expression of four gene signatures (see *SI Materials and Methods*). **(f)** Dot plot showing the expression of known and canonical DC/monocyte subset markers (Villani, et al, Science, 2017) across myeloid cells. The size of each circle corresponds to the percentage of cells in the subtype expressing the gene, and the color represents the average expression. **(g)** Average expression (color and square size) of the ten published T cell-subtype gene signatures (see *SI Materials and Methods*).**(h)** Stacked violin plots showing the expression of canonical markers defining T cell subtypes. **(i)** Dot plot showing the expression of canonical markers among identified NK/innate-like T-cell subsets. The size of each circle corresponds to the percentage of cells in the subtype expressing the gene, and the color represents the average expression. (**j**) Violin plot showing the local inverse Simpson's index (LISI) scores across all cells in the PBMC scRNA-seq data for condition and patient batches, respectively.

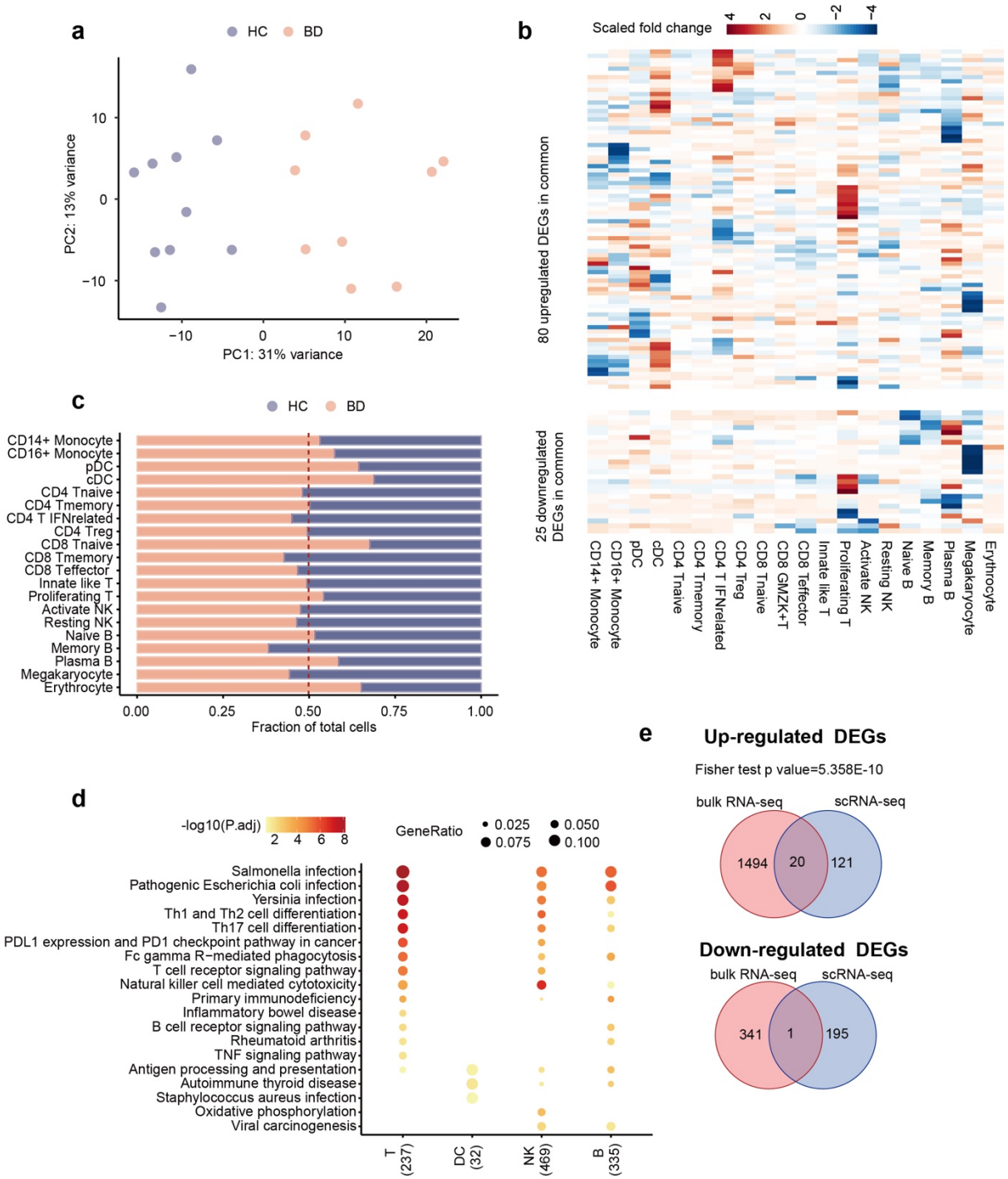

Figure S2 Transcriptional changes in immune cells in PBMCs from BD patients. **(a)** Principal component analysis (PCA) of 19 samples using all genes quantified by bulk RNA-seq. Patients (dots) are colored by condition (HC: blue; BD: red). **(b)** Heatmap visualizing the row-scaled fold change in common DEGs (same as the genes in Figure 2**b**-2**c**) in cell types identified in scRNA-seq data. The fold change is the ratio of the mean expression of the gene in BD versus HC. **(c)** The relative fraction of BD patients (red) and HCs (blue) for each cell subtype in scRNA-seq data. **(d)** KEGG analysis of DEGs (BD vs. HC) across the four main cell lineages. The top KEGG terms are colored by the Benjamini–Hochberg-corrected p values. The size represents the ratio of enriched genes in the pathway. **(e)** Venn diagrams of the overlap between DEGs identified by bulk RNA-seq (Figure 2**a**) and DEGs in monocyte scRNA-seq data (Figure

2**f**), with the upregulated genes in the upper panel and the downregulated genes in the lower panel (59,863 total genes detected in either dataset). The p values are from Fisher's test. PC: principal component.

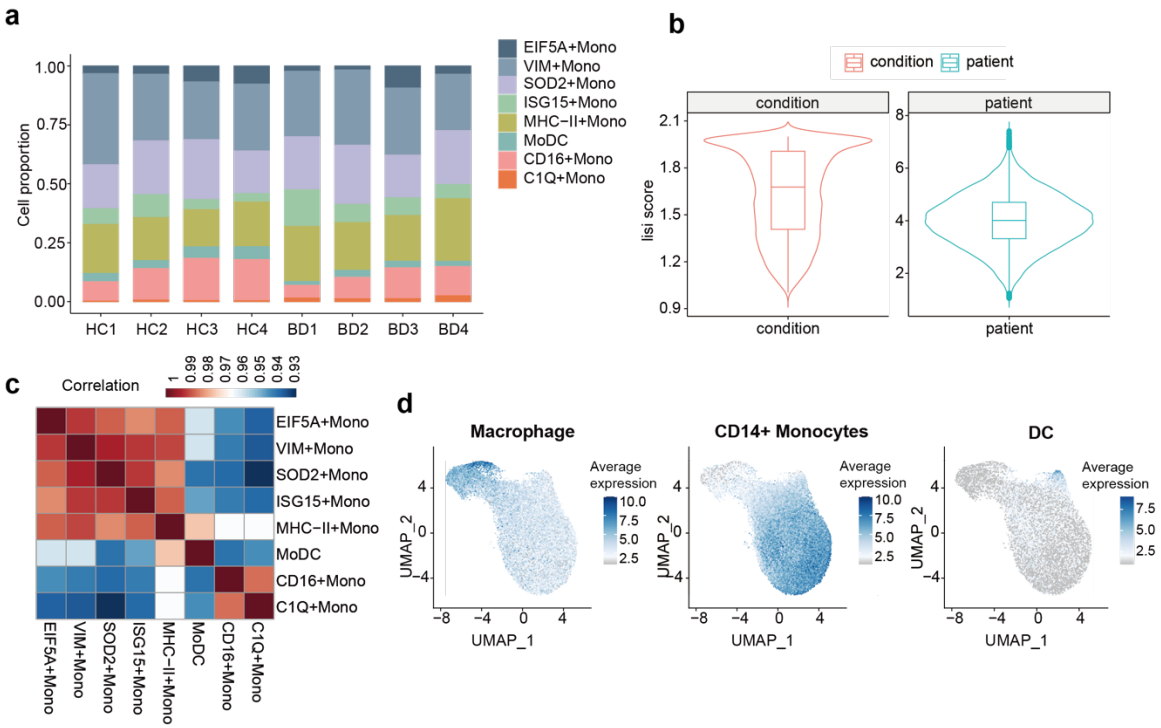

Figure S3 Identification of monocyte subtypes.
(a) The proportions of monocyte subtypes in individual patients. **(b)** Violin plot showing the local inverse Simpson's index (LISI) scores across all cells in the sorted monocyte scRNA-seq data for condition and patient batches, respectively. **(c)** Heatmap showing the Spearman correlations (colors) across all monocyte subtypes based on log-normalized average expression. **(d)** UMAP plot of all monocytes colored by the average expression of three published gene signatures (see *SI Materials and Methods*).

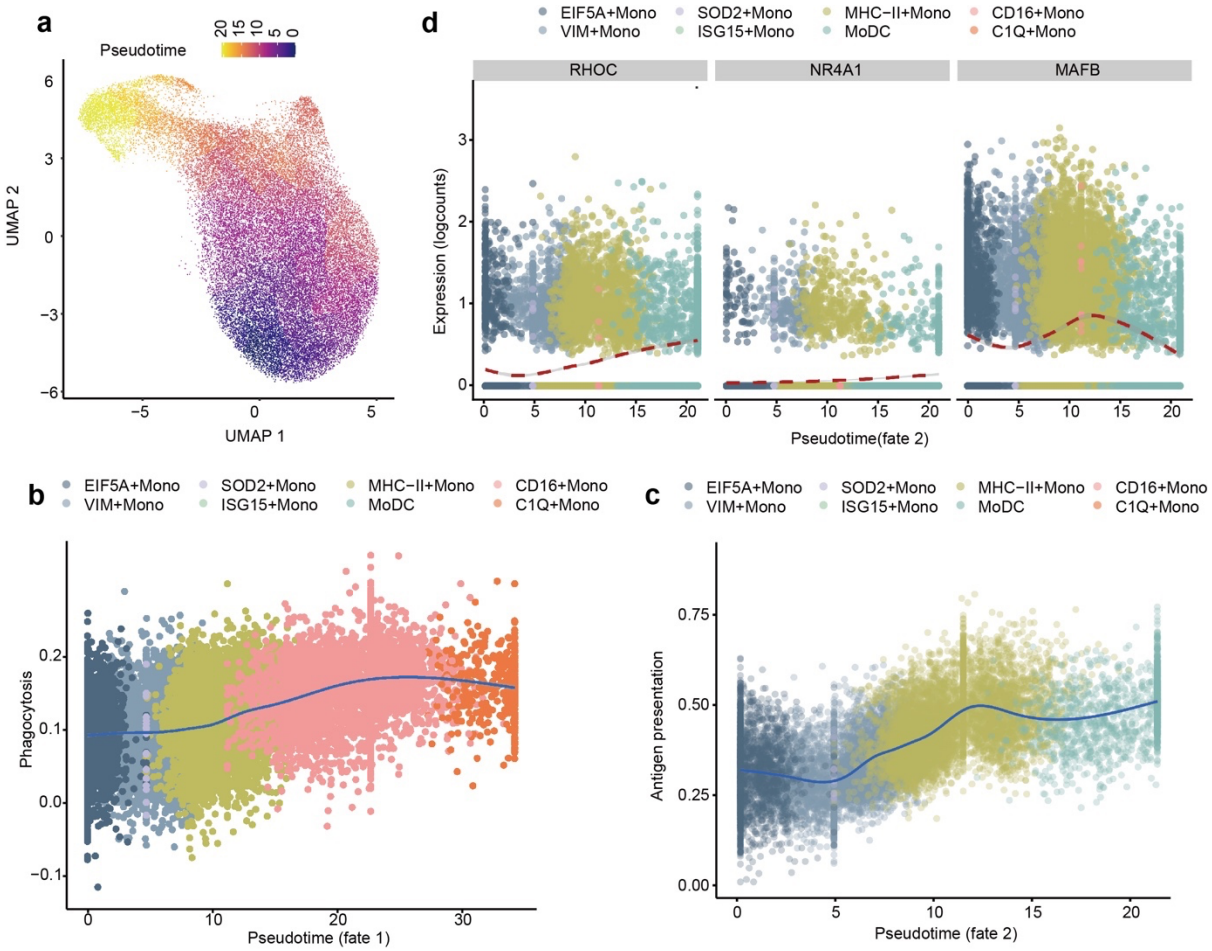

Figure S4 Pseudotime comparisons among all monocyte subtypes.
(a) UMAP plot showing the pseudotime (color) from the Monocle 3 algorithm. **(b)** The average expression of the phagocytosis pathway (see *SI Materials and Methods*) in each cell (dot) along the pseudotime of fate 1, colored by monocyte subtype. **(c)** The average expression of the antigen presentation pathway (see *SI Materials and Methods*) in each cell (dots) along the pseudotime of fate 2, colored by monocyte subtype. **(d)** The expression of known TFs that drive macrophage development along fate 2, colored by monocyte subtype.

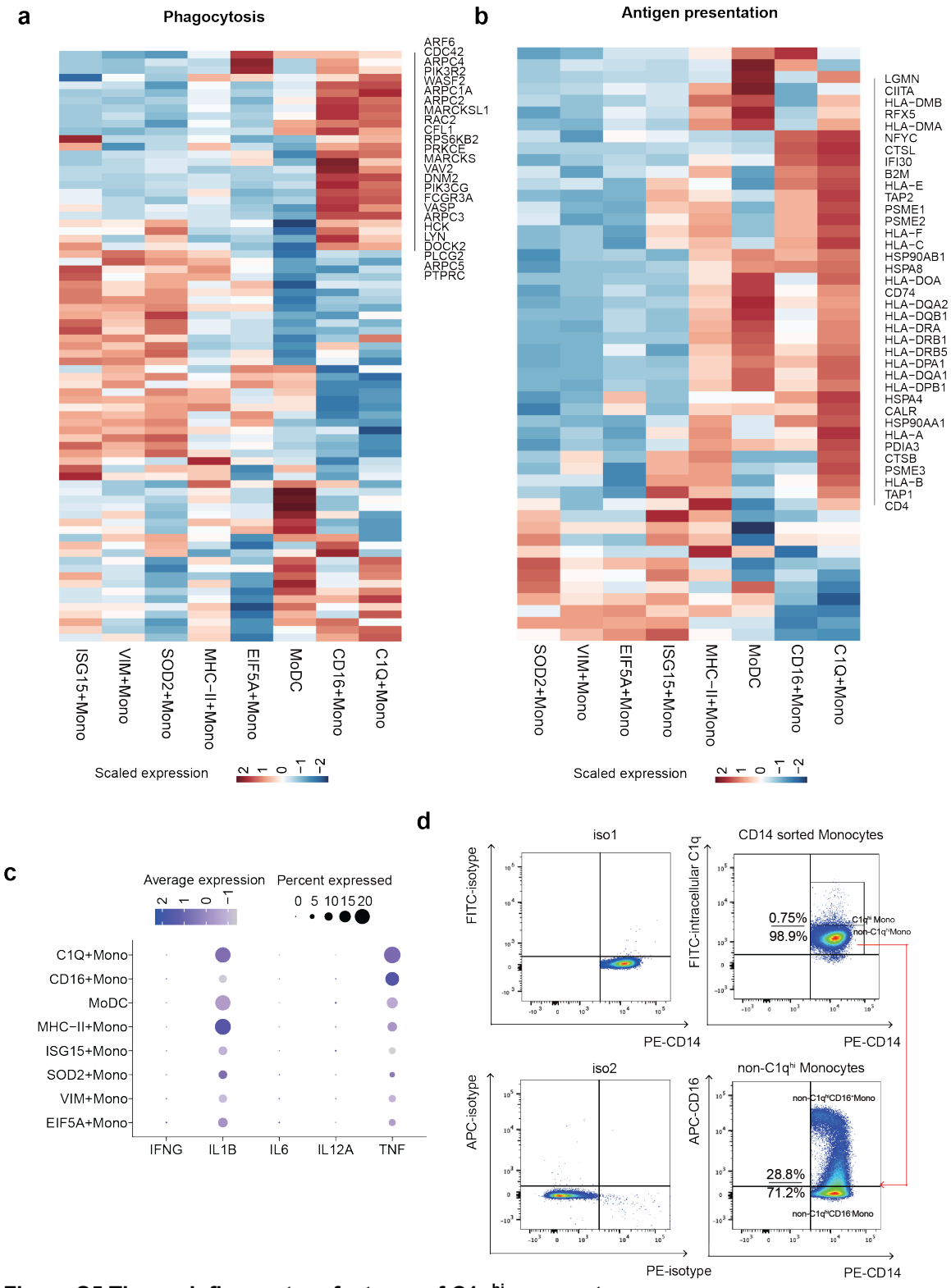

Figure S5 The proinflammatory features of C1q^{hi} monocytes.
(a-b) Heatmap showing the row-scaled average expression (color) of genes in the phagocytosis (**a**) and antigen presentation pathways (KEGG database) (**b**) across monocyte subtypes. Genes with an average expression of more than 0.1 across all detected cells are shown. **(c)** Dot plot showing the expression of BD-related cytokines across monocyte subtypes. The dot size

corresponds to the percentage of cells in the subtype expressing the gene, and the color represents the average expression level. **(d)** Representative flow cytometry plot of $C1q^{hi}$ monocytes, non-$C1q^{hi}CD16^{+}$ monocytes and non-$C1q^{hi}CD16^{-}$ monocytes.

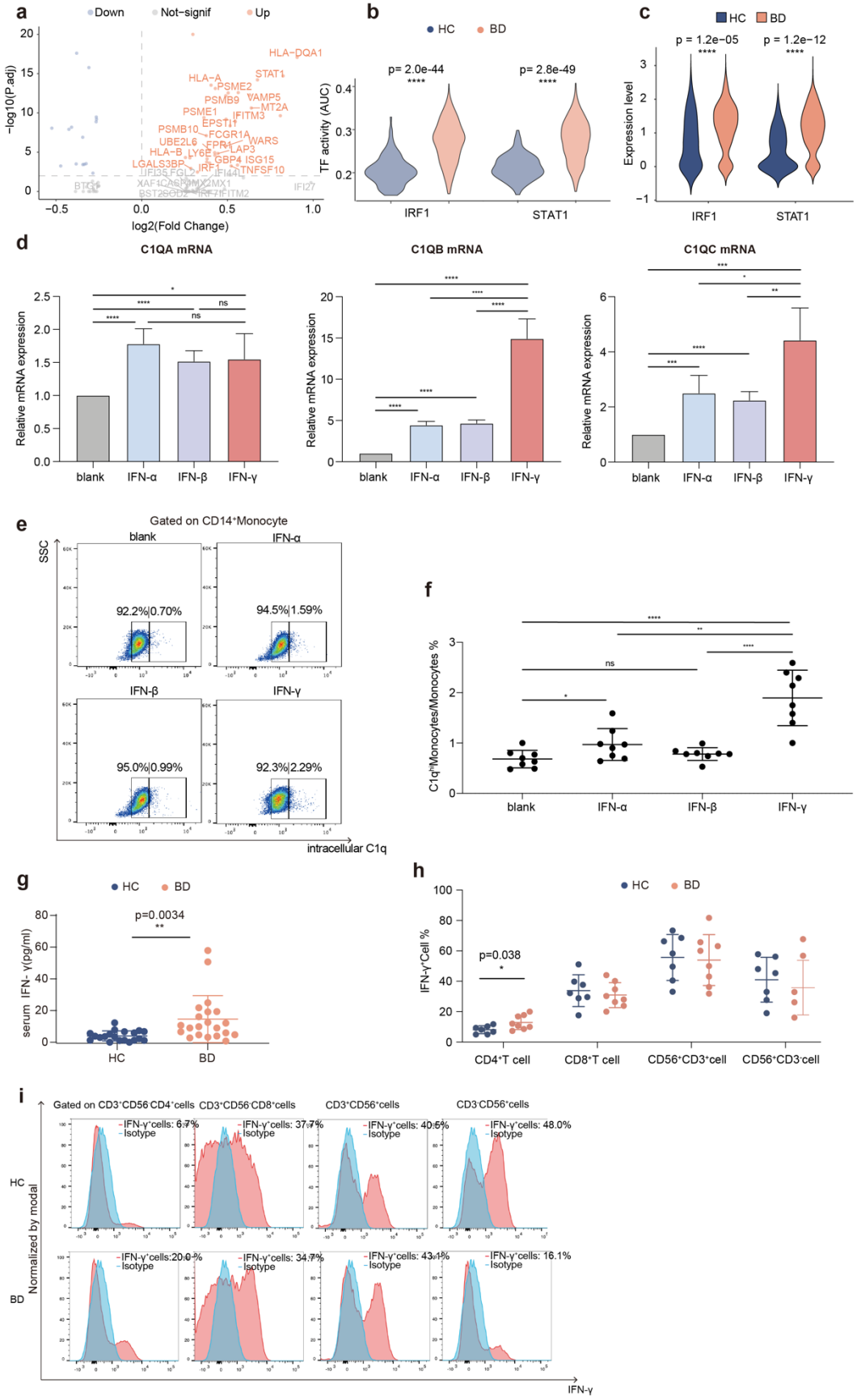
a
Down
Not-signif
Up
−log10(P.adj)
log2(Fold Change)
HLA-DQA1
STAT1
HLA-A
PSME2
PSMB9
VAMP5
MT2A
PSME1
IFITM3
EPSTI1
PSMB10
FCGR1A
UBE2L6
FPR1
LAP3
WARS
HLA-B
LY6E
LGALS3BP
GBP4
ISG15
IRF1
TNFSF10
IFI44L
IFI27
b
HC
BD
p= 2.0e-44
p= 2.8e-49
TF activity (AUC)
IRF1
STAT1
c
HC
BD
p = 1.2e−05
p = 1.2e−12
Expression level
IRF1
STAT1
d
C1QA mRNA
C1QB mRNA
C1QC mRNA
Relative mRNA expression
blank
IFN-α
IFN-β
IFN-γ
ns
e
Gated on CD14⁺Monocyte
blank
IFN-α
IFN-β
IFN-γ
SSC
92.2%|0.70%
94.5%|1.59%
95.0%|0.99%
92.3%|2.29%
intracellular C1q
f
C1q⁺Monocytes/Monocytes %
blank
IFN-α
IFN-β
IFN-γ
ns
g
HC
BD
p=0.0034
serum IFN- γ(pg/ml)
HC
BD
h
HC
BD
p=0.038
IFN-γ⁺Cell %
CD4⁺T cell
CD8⁺T cell
CD56⁺CD3⁺cell
CD56⁺CD3⁻cell
i
Gated on CD3⁺CD56⁻CD4⁺cells
CD3⁺CD56⁻CD8⁺cells
CD3⁺CD56⁺cells
CD3⁻CD56⁺cells
IFN-γ⁺cells: 6.7%
IFN-γ⁺cells: 37.7%
IFN-γ⁺cells: 40.5%
IFN-γ⁺cells: 48.0%
IFN-γ⁺cells:20.0%
IFN-γ⁺cells: 34.7%
IFN-γ⁺cells: 43.1%
IFN-γ⁺cells: 16.1%
Isotype
HC
BD
Normalized by modal
IFN-γ

Figure S6 Activated IFN-γ pathway in C1q^{hi} monocytes from BD patients.
(a) Volcano plot showing DEGs in C1q^{hi} monocytes between the BD and HC groups (adjusted p<0.05, fold change $\geqslant$ 0.25). Significant genes are colored red, and genes within IFN-γ pathways are labeled. **(b)** Violin plot comparing the difference in the area under the curve (AUC, from the SCENIC algorithm) for the top two TFs (Figure 6**b**) between BD patients (blue) and HCs (red). **(c)** Violin plot showing the expression of predicted TFs (IRF1 and STAT1) in C1q^{hi} monocytes of BD patients and HCs. **(d)** Monocytes were stimulated with IFN-α, IFN-β, IFN-γ or blank control in the concentration of 200 U/ml for 6 h, and the relative mRNA expression (left, C1QA; middle, C1QB; right, C1QC) to β-actin was measured using RT-qPCR. **(e, f)** Representative flow cytometry plot **(e)** and graph **(f)** displaying the proportions of C1q^{hi} monocytes after 12 h treatment with IFN-α, IFN-β or IFN-γ (200U/ml) (n=8). **(g)** Enzyme-linked immunosorbent assay (ELISA) showing IFN-γ concentration in serum from BD patients and HCs (n=21 in BD, n=20 in HC). **(h-i)** Representative histograms (**i**) and statisitical graph (**h**) of flow cytometry data showing the proportion of IFN-γ-positive cells in CD4$^+$ T cells, CD8$^+$ T cells, NK cells (CD3$^-$CD56$^+$), and innate-like T cells (CD3$^+$CD56$^+$) from BD patients (n=8) and HC (n=7) after 5 h of PMA (phorbol 12-myristate 13-acetate)/ionomycin stimulation. The Wilcoxon test (Figure S6**a-c**) and independent-samples t-test (Figure S6**d, f-h**) were applied. *, p<0.05; **, p<0.01; ***, p<0.001; ****, p<0.0001.

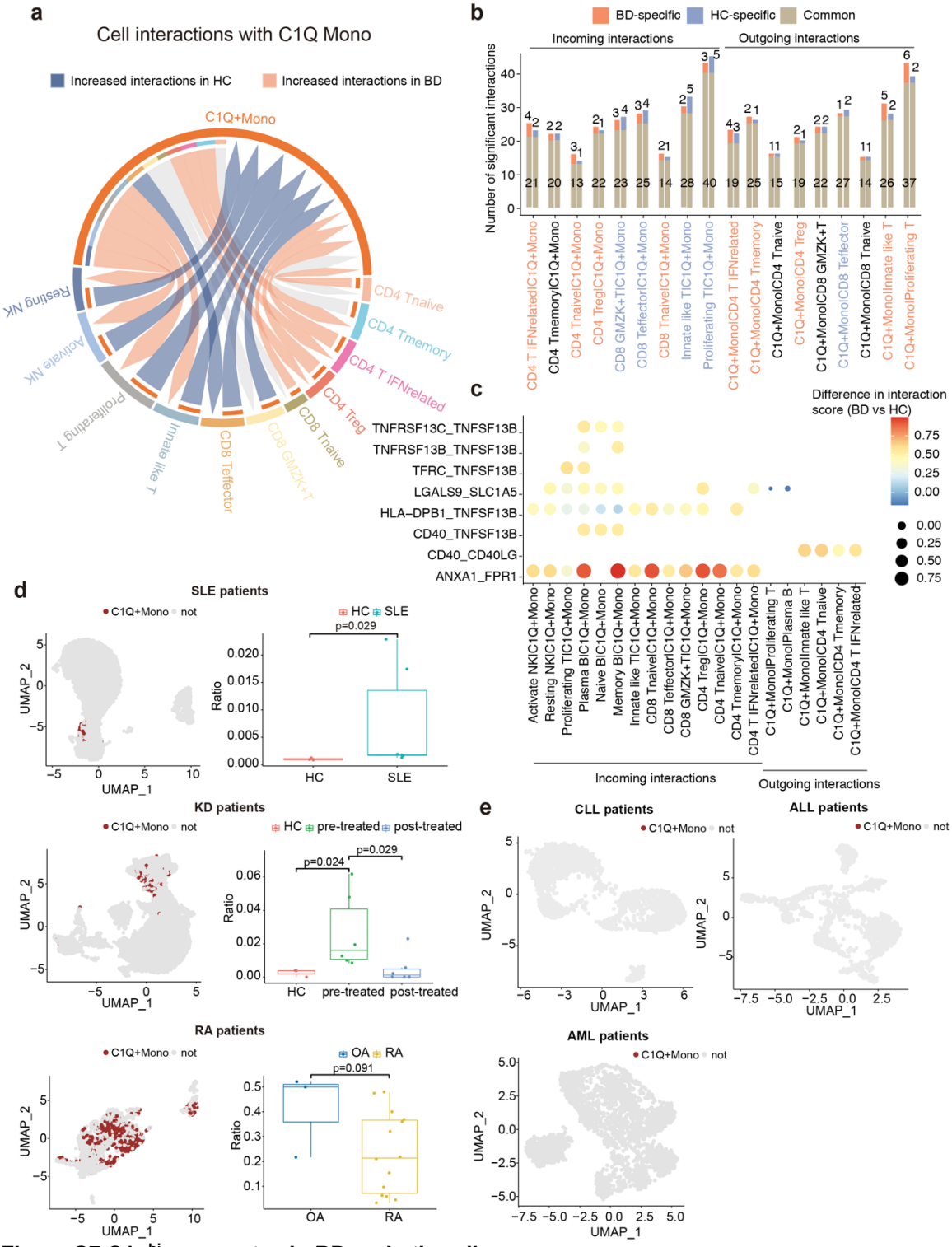

Figure S7 C1q^{hi} monocytes in BD and other diseases.

(a) Circos plot depicting changes in putative receptor-ligand interactions between C1q^{hi} monocytes and T/NK cells. The number of significant interactions was inferred by CellPhoneDB, and the color indicates an increase in the BD (red) or HC (blue) group. Arrows represent outgoing or incoming interactions (Outgoing interactions: the sum of ligands from C1q^{hi} monocytes that interact with receptors on certain cell types; incoming interactions are the opposite). **(b)** Numbers of significant interactions between C1q^{hi} monocytes and T cells (pink, only significant in BD;

purple, only significant in HC; grey, significant in both conditions). **(c)** Dot plot displaying significant ligand-receptor interactions (Y-axis) between C1q^{hi} monocytes and other immune cells. C1q^{hi} monocytes expressing receptors are defined as incoming interactions, and those expressing ligands are defined as outgoing interactions. Circle size and color denote the difference in interaction scores (inferred by CellPhoneDB) in interacting populations. **(d)** UMAP plots revealed C1q^{hi} monocytes in inflammatory diseases, including SLE, RA and KD with intravenous immunoglobulin therapy. Boxplots showed the ratios of C1q^{hi} monocytes to total monocytes for individual patients. Dots represent patients. The Wilcoxon test were applied. **(e)** UMAP plots revealed that C1q^{hi} monocytes were not identified in blood cancers, including chronic lymphocytic leukemia (CLL), acute lymphocytic leukemia (ALL), acute myeloid leukemia (AML). SLE, systemic lupus erythematosus; RA, rheumatoid arthritis; KD, Kawasaki disease; *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$.

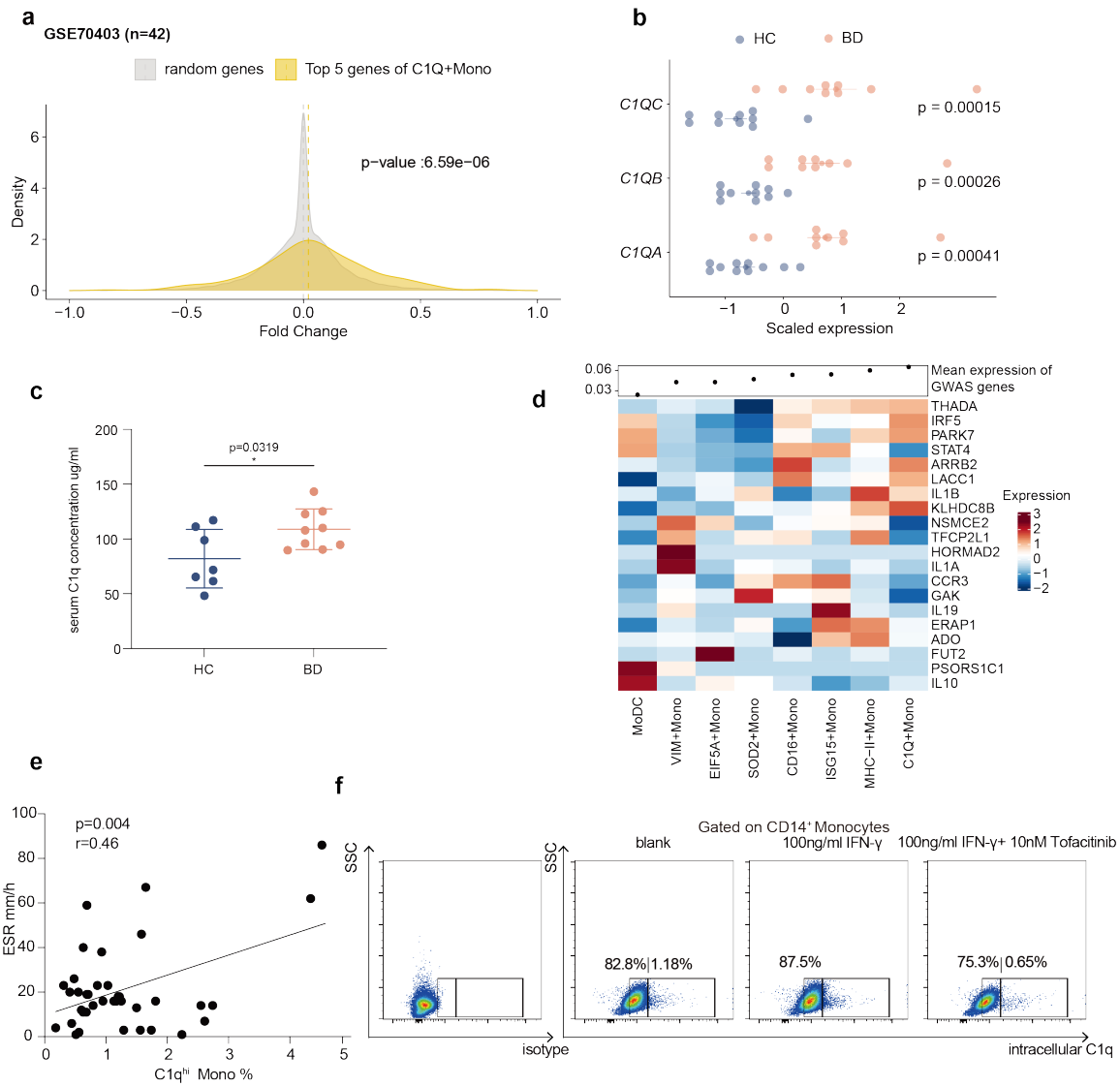

Figure S8 Clinical relevance of C1q^{hi} monocyte.

(a) Density plot showing the distribution of top five highly expressed genes in C1Q Monos and 100 randomly selected genes in the GSE70403 cohort (n=42). The p values were calculated by the Kolmogorov–Smirnov test. **(b)** Comparison of the z-score-scaled expression (X-axis) of C1q genes between BD and HC samples in bulk RNA-seq of our in-house cohort. The Wilcoxon test was applied. **(c)** The concentration of C1q in serum from the BD (n=9) and HC (n=7) groups, as measured by ELISA. **(d)** Heatmap showing the z-score-scaled expression of BD-associated risk genes in GWAS across monocyte subtypes (see *SI Materials and Methods*). The means of the displayed genes for individual subtypes are shown in dot plots. **(e)** Pearson correlation of C1q^{hi} monocyte proportion among all monocytes with erythrocyte sedimentation rate (ESR) in BD patients (n=38). **(f)** Representative graph of the flow cytometry results showing the proportion of C1q^{hi} monocytes among HC, IFN-γ-treated, IFN-γ- and tofacitinib-treated monocytes. The independent-sample t-test (Figure S7**c**) were applied to calculate the p value. *, p<0.05; **, p<0.01; ***, p<0.001.

SI Materials and Methods

Single-cell RNA library preparation and sequencing

PBMCs were isolated from blood by density gradient centrifugation and then resuspended in complete Dulbecco's Modified Eagle Medium (DMEM) (Sigma) or phosphate buffered saline (PBS). Monocytes were isolated from PBMCs with anti-CD14 microbeads (Miltenyi) according to the manufacturer's instructions. Sorted PBMCs or monocytes were washed and resuspended in PBS with 0.04% Bovine Serum Albumin (BSA), loaded into Chromium microfluidic chips with 3' chemistry, and barcoded with a 10× Chromium Controller (10X Genomics). RNA from the barcoded cells was subsequently reverse transcribed, and sequencing libraries were constructed with reagents from a Chromium Single Cell 3' v2 Reagent Kit (10X Genomics) according to the manufacturer's instructions. Sequencing was performed with an Illumina Novaseq 6000 according to the manufacturer's instructions (Illumina).

Bulk RNA library preparation and sequencing

Total RNA was extracted from PBMCs with TRIzol reagent. RNA purity was checked with a NanoPhotometer spectrophotometer (Implen), and RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies). A total of 1 μg RNA per sample was used as input for RNA sample preparation. Sequencing libraries were generated with a NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following the manufacturer's recommendations, and index codes were added to attribute sequences to each sample. PCR products were purified (AMPure XP system), and library quality was assessed on the Agilent Bioanalyzer 2100 system. Index-coded samples were clustered using TruSeq PE Cluster Kitv3-cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina HiSeq platform, and 125 bp paired-end reads were generated.

Preprocessing of scRNA-seq data

Raw data were processed to generate a count matrix using standard pipelines in the Cell Ranger Single-Cell Software Suite (v3.0.0). The reads were aligned against the GRCh38 human reference genome, filtered and subjected to unique molecular identifier (UMI) counting with the default parameters. The filtered feature matrixes were imported into Seurat (v3.1.5)(1) with the function *Read10*× to perform quality control and further exploration. For each sample, data normalization and variable feature identification were conducted separately with the default parameters. Cross-dataset anchors were then identified and used to correct for batch effects across samples. Mitochondrial percentages were calculated by the function *PercentageFeatureSet*. We retained cells with between 500 and 4,000 expressed genes and a mitochondrial percentages of less than 5%. After removing inferred doublets with DoubletFinder (v2.0) (2), a total of 36,190 qualified PBMCs and 39,385 monocytes remained for subsequent analyses. PCA was then performed based on 2,000 highly variable genes recognized by the function *FindVariableGenes*.

Cell clustering and annotation

Cell clustering was conducted based on the top 25 principal components using the graph-based clustering algorithm in the function *FindClusters* with a resolution of 0.4. UMAP was applied to visualize the identified clusters. We next used complementary approaches to annotate the cell clusters. In the first approach, highly differentially expressed genes in a certain cluster were identified by comparison with all other clusters using the Wilcoxon test in the function *FindAllMarkers* (min.pct=0.25, only.pos=T, and logfc.threshold=0.2). The cell clusters were assigned according to well-known cellular markers from the literature. In the second approach, we analyzed each cluster using the function *AddModuleScore* to estimate the average expression of published and well-established gene signatures, which were downloaded from the literature. In the third approach, we applied an unbiased cell type recognition method named SingleR (v10)(3), which leverages default reference transcriptomic datasets of known cell types for annotation, and assigned clusters based on the predicted cell type annotation. For PBMC scRNA-seq data, we utilized all the methods. For the sorted monocyte scRNA-seq data, we named cell clusters based

on the top DEGs from the first method only due to the limited signatures in the literature and databases in SingleR.
To test whether there were confounding factors in the scRNA-seq analyses, we calculated local inverse Simpson's index (LISI) scores across all cells for two potential batches (patients and HC/BD condition) according to previous studies (4, 5).

Bulk RNA-seq data analysis
Raw data in FASTQ format were first processed through in-house Perl scripts. In addition, the Q20, Q30, and GC contents of the clean data were calculated. All downstream analyses were based on clean, high-quality data. A reference genome index was built, and paired-end clean reads were aligned to the reference genome using HISAT2 (v2.0.5) (6). For quantification of gene expression levels, featureCounts (v1.5.0) (7) was applied to count the number of reads mapped to each gene.
The count matrix was input into DESeq2 (v1.30.0)(8) and fitted for a general linear model with a negative binomial distribution. Only genes with a detected count of more than 20 were retained. DEGs between the disease and control groups were identified by the functions *DESeq* and *lfcShrink* with the criteria log2 (fold change)>1 and an adjusted p <0.05 (Wald test and Bonferroni correction). PCA was performed for all qualified genes, and the results were visualized with the functions *vst* and *plotPCA*.
To identify the expression and fold change in these DEGs in the scRNA-seq data of PBMCs, we extracted the expression profile of each cluster from the integrated slot of a Seurat object by the function *AverageExpression*, and the fold change was defined as the ratio of the mean expression of the gene in BD versus HC. All genes identified in the scRNA-seq data overlapped with the DEGs in the bulk RNA-seq data analysis. Heatmaps were used to display the expression and fold change in these significantly upregulated and downregulated genes via pheatmap (v1.0.12).
To evaluate the immune cell proportions in each sample, we used CIBERSORT(9) to infer the relative abundance of 22 immune cells among PBMCs by the LM22 signature, with the TPM (transcript per million) gene matrix as input. LM22 signature matrix was the default reference in CIBERSORT, and was generated from human peripheral blood (9). Cell types with a total relative abundance of more than 0.01 across all samples were kept for downstream analyses. The Wilcoxon test was used to compare the differences between BD and HC samples as indicated in the figure legends.

Pathway enrichment analysis
The enriched pathways were assessed by hypergeometric testing in the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. The enrichments were performed by the function *compareCluster* in the clusterProfiler package (v3.0.4)(10). Significantly enriched pathways were determined with a cutoff of a Benjamini–Hochberg corrected p < 0.05.

Gene signature analysis
Gene signatures used in the monocyte scRNA-seq analysis were downloaded from the H (hallmark) gene sets and C2 (curated) gene sets of the MSigDB Collections(11), including the REACTOME_CELL_CYCLE, KEGG_FC_GAMMA_R_MEDIATED_PHAGOCYTOSIS, KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION, and HALLMARK_INTERFERON_GAMMA_RESPONSE signatures. Other signatures were collected from the previous studies(12). Monocytic cytokines were defined as follows: TNF, IL8, IL6, IL1B, IL1A, and IL12A. The function *AddModuleScore* in Seurat was employed to estimate the average expression of these established gene signatures.

Trajectory inference and identification of pseudotime-correlated genes
Three different algorithms, including diffusion map(13), TSCAN(14) and Slingshot(15), were used to infer the pseudotime ordering and trajectories of monocytes. The pseudotime was calculated by the function *quickPseudotime* with *use.dimred = "pca"* in the R package TSCAN (v1.28.0). The trajectories were built by the function *slingshot* in the R package Slingshot (v1.8.0), with the

setting *reducedDim = 'PCA'*. To visualize the results, we applied diffusion map to calculate diffusion components in the R package "destiny" v3.4.0(16) using the count matrix as input. Significantly correlated genes along fate 1 and fate 2 were recognized by the function *testPseudotime* with FDR<0.05 and logFC>0.1, and the log-normalized counts of the top 20 correlated genes were visualized with the function *plotHeatmap*. Nonoverlapping correlated genes along fate 1 or fate 2 were analyzed for the enrichment of KEGG pathways by the function *compareCluster* in clusterProfiler. Selected TFs were visualized along the two fates with the function *plotExpression* in scater (v1.18.6).

To verify the pseudotime ordering, we applied Monocle 3(17), which learned the cell trajectory by reverse graph embedding. The normalized expression matrix of highly variable genes identified by Seurat was used to create the Monocle object. We also loaded the results of the UMAP dimension reduction and cell clustering from Seurat into a new object. The trajectory was built by the functions learn_graph and order_cells with the default settings.

The starting point of the monocyte trajectory was set as SOD2 Monos, which were inferred as the starting cell type by Monocle3, CytoTRACE, and TSCAN algorithm. SOD2 Monos also exhibited lower expression of key genes (CSF1R, RHOC, and MAFB) linked to monocyte differentiation than other monocyte subtypes.

Transcription factor analysis

Differentially activated TFs in C1Q Monos between BD and HC were identified by SCENIC(18) tool. The raw count matrix of C1Q Monos was input and filtered for genes expressed in less than 1% of cells. We used GENIE3 algorithm to calculate the coexpression network, and the candidate TFs were identified by RcisTarget algorithm with the default parameters. The human v9 motif collection was used as the reference, and the "hg19-500 bp-upstream-7species.mc9nr" and "hg19-tss-centered-10 kb-7species.mc9nr" databases were downloaded from cisTarget. Then, the TF activity in each cell was scored by *AUCell*. C1Q Monos were split into 2 groups based on the disease state: BD or HC. The differentially activated TFs between the two groups were identified by the Wilcoxon test with the Benjamini–Hochberg correction.

Cell-cell interaction analysis

We used CellPhoneDB(19) to infer interactions between $C1q^{hi}$ monocytes and the main IFN-γ-producing cells (T/NK cells). The interaction strength between these two cell types was calculated based on the mean expression of ligand and receptor. The permutation test was used to determine the significance of the interaction pairs at $p<0.05$.

Public data collection and analysis

BD-associated GWAS risk loci were obtained from the GWAS catalog(20). The highly expressed genes (~200 genes) in C1Q Monos were calculated by *FindAllMarkers* as mentioned above. Public bulk RNA-seq data of blood samples in BD cohorts were downloaded from the GEO database (GSE17114, GSE165254, and GSE70403). Binomial regression models were used to build the links between genes within the top five highly expressed genes of C1Q Monos and disease status. The R package "pROC" was applied to generate receiver operating characteristic (ROC) curves and estimate the AUC. As the individual array of the GSE70403 cohort used HC samples as a control, the intensity value of each gene indicates the relative expression (or fold change) of BD samples compared to HC samples but not the absolute expression. Thus, we used the Kolmogorov–Smirnov test to determine the significance of differences in fold change among the top five genes and 500 randomly selected genes instead of the binomial regression models above.

The scRNA-seq datasets were downloaded from the GEO and ImmPort databases (KD, GSE168732; SLE, GSE135779; RA, SDY998; AML, GSE116256; CLL, GSE111014; and ALL, GSE132509). The top 20 highly expressed genes were defined as markers of C1Q Monos. We used SCENIC to identify the number of cells classified as C1Q Monos with the C1Q Mono markers as the input gene set. The threshold was automatically determined using *getThresholdSelected* with default parameters. Cells with scores higher than the selected thresholds were classified as C1Q Mono.

Flow cytometry and phospho-flow cytometry

PBMCs or monocytes were stained with fluorochrome-labeled antibodies for the following surface markers: CD14 (M5E2, BD Biosciences), CD16 (3G8, BD Biosciences), CD3 (HIT3a, BD Biosciences), CD56 (MEM-188, BioLegend), and CD4 (RPA-T4, BioLegend). For intracellular staining, cells were fixed and permeabilized according to the manufacturer's instructions (Cytofix/Cytoperm and Perm/Wash Buffer, BD Biosciences) and stained for 50 min on ice for C1q (polyclonal, Abcam), TNF-α (MAB11, BD Biosciences), IL-6 (MQ2-13A5, BD Biosciences) and IFN-γ (4S.B3, BD Biosciences). Brefeldin A was added before the intracellular staining of C1q. For intracellular phosphoprotein staining, cells were fixed and permeabilized according to the manufacturer's instructions (Cytofix Buffer and Phosflow Buffer I, BD Biosciences) and stained for 50 min on ice for phospho-STAT1 (polyclonal, Bioss Antibodies). The stained cells were immediately analyzed with a BD FACSAria II and FlowJo Software (Tree Star). Because the MFIs varied among different batches in phospho-flow cytometry, we normalized the MFI to an internal control each time according to a published protocol(21).

Cytokine secretion assay

Freshly isolated monocytes were suspended in DMEM and stimulated with LPS (20 ng/ml) for 4 h at 37°C and 5% CO_2. Then, surface and intracellular stainings were performed as mentioned above. The proportions of IL-6^+ and TNF-α^+ cells among monocytes were analyzed by flow cytometry.

Freshly isolated PBMCs were suspended in RPMI 1640 and stimulated with Cell Activation Cocktail with Brefeldin A (Biolegend) for 5 h at 37°C and 5% CO_2. Then, surface and intracellular stainings were performed as mentioned above. The proportions of IFN-γ^+ cells among $CD4^+$ T cells, $CD8^+$ T cells, $CD3^-CD56^+$ cells and $CD3^+CD56^+$ cells were analyzed by flow cytometry.

Phagocytosis test

Freshly isolated monocytes were suspended in 1 ml PBS at 5×10^5 cells/ml and incubated with 1 mg/ml TRITC-labeled 70 MW dextran at 4°C (negative control) or 37°C (blank control) for 60 min. Then, the cells were harvested and analyzed with a BD FACSAria II. The ΔMFI between the two conditions (37°C and 4°C) was employed to measure phagocytosis ability.

IFN-γ and tofacitinib stimulation test

Freshly isolated monocytes (5×10^5 cells/ml) were cultured in DMEM in a 24-well plate at 37°C and 5% CO_2, and IFN-α, IFN-β, IFN-γ (200 U/ml, PeproTech) was added for stimulation. In the drug treatment assay, IFN-γ (200U/ml, PeproTech) with or without tofacitinib (10nM, Selleck) was added for stimulation. Cells were harvested for PCR and flow cytometry analyses of C1q expression after 6 h and 12 h, respectively.

Enzyme-linked immunosorbent assay (ELISA)

Serum C1q (ab170246, Abcam) and IFN-γ (ab46025, Abcam) levels were determined in duplicate in 96-well half-area plates using a standard plate reader. The assays were performed according to the manufacturer's instructions (Abcam) with appropriate dilutions.

RNA extraction and reverse transcription-quantitative PCR (RT-qPCR)

Total RNA was extracted from monocytes stimulated with IFN-γ for 6 h using the TRIzol (Sigma) method, and 500 ng total RNA was reverse transcribed using a Fast All-in-One RT Kit (ES Science, China). Real-time PCR was performed in triplicate with SYBR Green Master Mix and a Roche LightCycler™ 480. The primer sequences were as follows: C1qA, forward: 5'-TCTGCACTGTACCCGGCTA-3' and reverse: 5'-CCCTGGTAAATGTGACCCTTTT-3'; C1qB, forward: 5'-ATGGGGCAGCATCCCAGTA-3' and reverse: 5'-CTCCCTTCTCTCCGAACTCAC-3'; C1qC, forward: 5'-CCAACCCGCAGGGAGATTATG-3' and reverse: 5'-CCGAGTTGACCTGATTGGTTTT-3'; and GAPDH, forward: 5'-GCGAGATCCCTCCAAAATCAA-3' and reverse: 5'-GTTCACACCCATGACGAACAT-3'. The results were normalized to GAPDH expression levels, and data were shown as the relative gene expression.

Statistical analysis

The data are summarized as the mean ± SD. To compare two groups, independent-sample t-tests or paired t-tests were used for normally distributed variables, and Wilcoxon rank-sum tests were used for non-normally distributed variables. To compare more than two groups, one-way analysis of variance (ANOVA) with the Tukey–Kramer post hoc test was used to compare data displaying a normal distribution and homogeneity of variance. A two-tailed $p<0.05$ was considered to indicate statistical significance; *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$. All statistical analyses were performed using SPSS v.17.0 and R v4.0.2.

SI References

1. T. Stuart *et al.*, Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e1821 (2019).
2. C. S. McGinnis, L. M. Murrow, Z. J. Gartner, DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* **8**, 329-337 e324 (2019).
3. D. Aran *et al.*, Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nature Immunology* **20**, 163-172 (2019).
4. H. T. N. Tran *et al.*, A benchmark of batch-effect correction methods for single-cell RNA sequencing data. *Genome Biol* **21**, 12 (2020).
5. I. Korsunsky *et al.*, Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* **16**, 1289-1296 (2019).
6. D. Kim, B. Langmead, S. L. Salzberg, HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**, 357-360 (2015).
7. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
8. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
9. A. M. Newman *et al.*, Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods* **12**, 453-457 (2015).
10. G. Yu, L.-G. Wang, Y. Han, Q.-Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287 (2012).
11. A. Subramanian *et al.*, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**, 15545-15550 (2005).
12. A. Leruste *et al.*, Clonally Expanded T Cells Reveal Immunogenicity of Rhabdoid Tumors. *Cancer Cell* **36**, 597-612.e598 (2019).
13. S. C. Bendall *et al.*, Single-Cell Mass Cytometry of Differential Immune and Drug Responses Across a Human Hematopoietic Continuum. *Science* **332**, 687-696 (2011).
14. Z. Ji, H. Ji, TSCAN: Pseudo-time reconstruction and evaluation in single-cell RNA-seq analysis. *Nucleic Acids Research* **44**, e117-e117 (2016).
15. K. Street *et al.*, Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).
16. P. Angerer *et al.*, destiny: diffusion maps for large-scale single-cell data in R. *Bioinformatics* **32**, 1241-1243 (2016).
17. C. Trapnell *et al.*, The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature Biotechnology* **32**, 381-386 (2014).
18. S. Aibar *et al.*, SCENIC: single-cell regulatory network inference and clustering. *Nature Methods* **14**, 1083-1086 (2017).
19. M. Efremova, M. Vento-Tormo, S. A. Teichmann, R. Vento-Tormo, CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* **15**, 1484-1506 (2020).
20. J. MacArthur *et al.*, The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* **45**, D896-D901 (2017).
21. D. Upreti, A. Pathak, S. K. P. Kung, Development of a standardized flow cytometric method to conduct longitudinal analyses of intracellular CD3ζ expression in patients with head and neck cancer. *Oncol Lett* **11**, 2199-2206 (2016).