Fig. S1

A Gene count

B UMI count

C (Doublet / Singlet)

D N6 N9 N13 N14 N20 N31 S1 S2 S3 S4 S5 WM HD

E (HD / WM) B cell, Plasma cell, T cell, NK cell, HSPC, Myeloid cell, Erythroid cell

F 1 2 3 4 5 6 7 — Erythroid, CMP, HSPC

G CD34 SPINK2 CD52 FLT3 ACTG1 CD74 ZFAS1 ENO1 AIF1 VIM TMSB4X FCER1A FCGR2A GATA2 SOX4 CD38 ITGA2B CSF2RB PDLIM1 TFRC GYPA GYPB GATA1 ALAS2 ENG HBB HBD HBA1 HBA2 AHSP — Erythroid, CMP, HSPC

H (Erythroid / CMP / HSPC) HD WM

I Pseudotime, HD/WM, Pseudosome

J RNA splicing; ribonucleoprotein complex biogenesis; RNA catabolic process; translational initiation; mRNA catabolic process; mRNA splicing, via spliceosome; regulation of chromosome organisation — -log10(P-value)

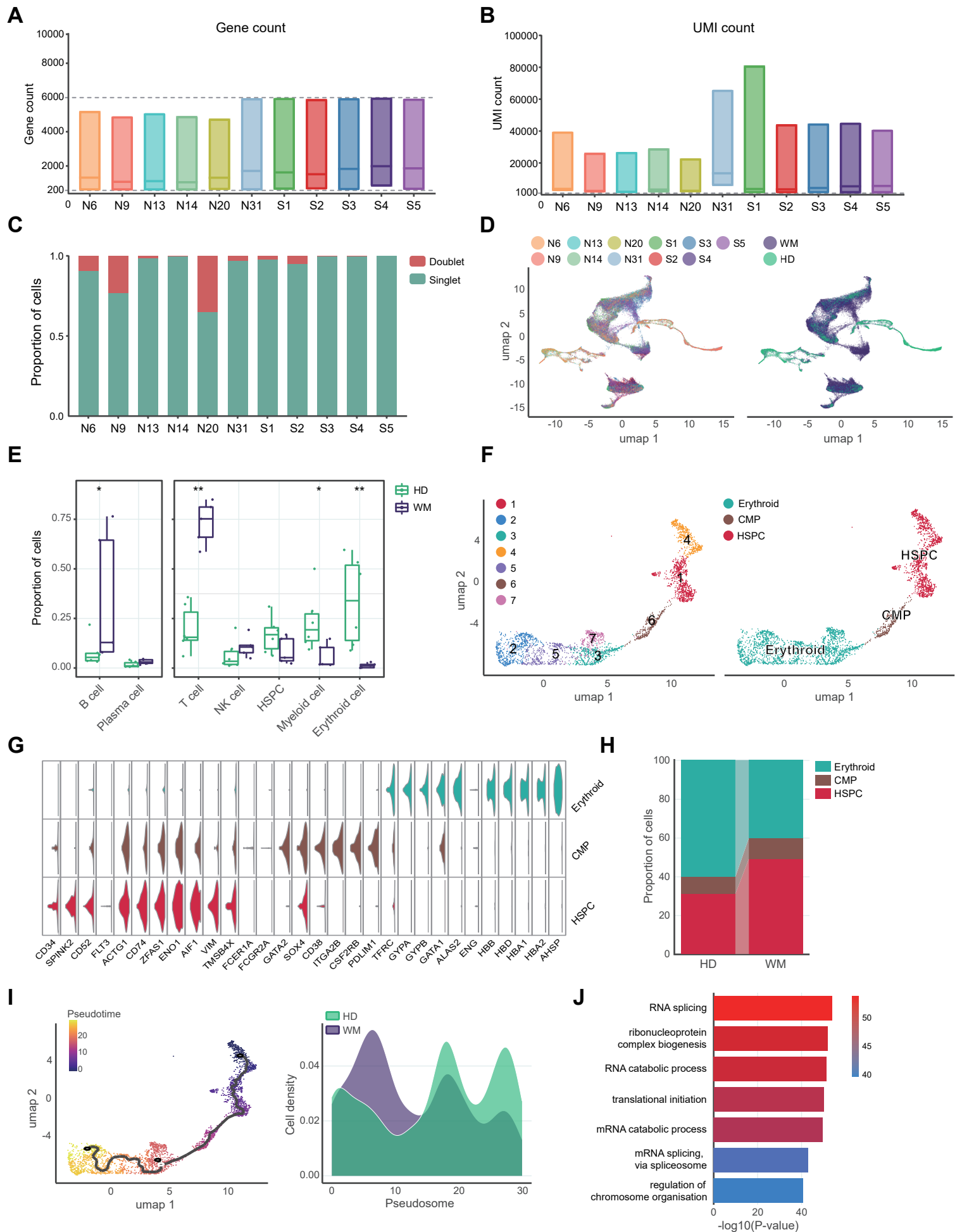**Fig. S1 | Quality control and visualization of scRNA-seq data**

(A-B) Box plots showing filtering criteria for sequenced cells, including gene count and unique molecular identifier (UMI) count. The average values, maximum and minimum values of data after filtration are shown through the box plots.

(C) Bar charts showing the fraction of doublet cells filtered by DoubletDecon.

(D) UMAP plots showing cell origins by color, individual origin (left panel), and WM or HD origin (right panel).

(E) Boxplots illustrating the proportion of cells identified in BMNCs of HD and WM. The proportion of immune cells was calculated after normalizing the differential proportion of WM cells. *$P < 0.05$, **$P < 0.01$, unpaired Wilcoxon test between HDs and WM patients.

(F) UMAP projections of subclustered erythroid cells and HSPCs (left), and of cell type annotations (right), labeled in different colors.

(G) Violin plots showing the expression of marker genes in three cell subtypes.

(H) Stacked bar chart showing the proportions of each cell subtype in HD and WM, unpaired Wilcoxon test for erythroid cells in HDs and WM patients, $P<0.05$.

(I) UMAP plot of the pseudotime-ordered erythroid cells and HSPCs, from HD and WM (left). The cell density distribution, by cell origin, is shown at the right of the figure.

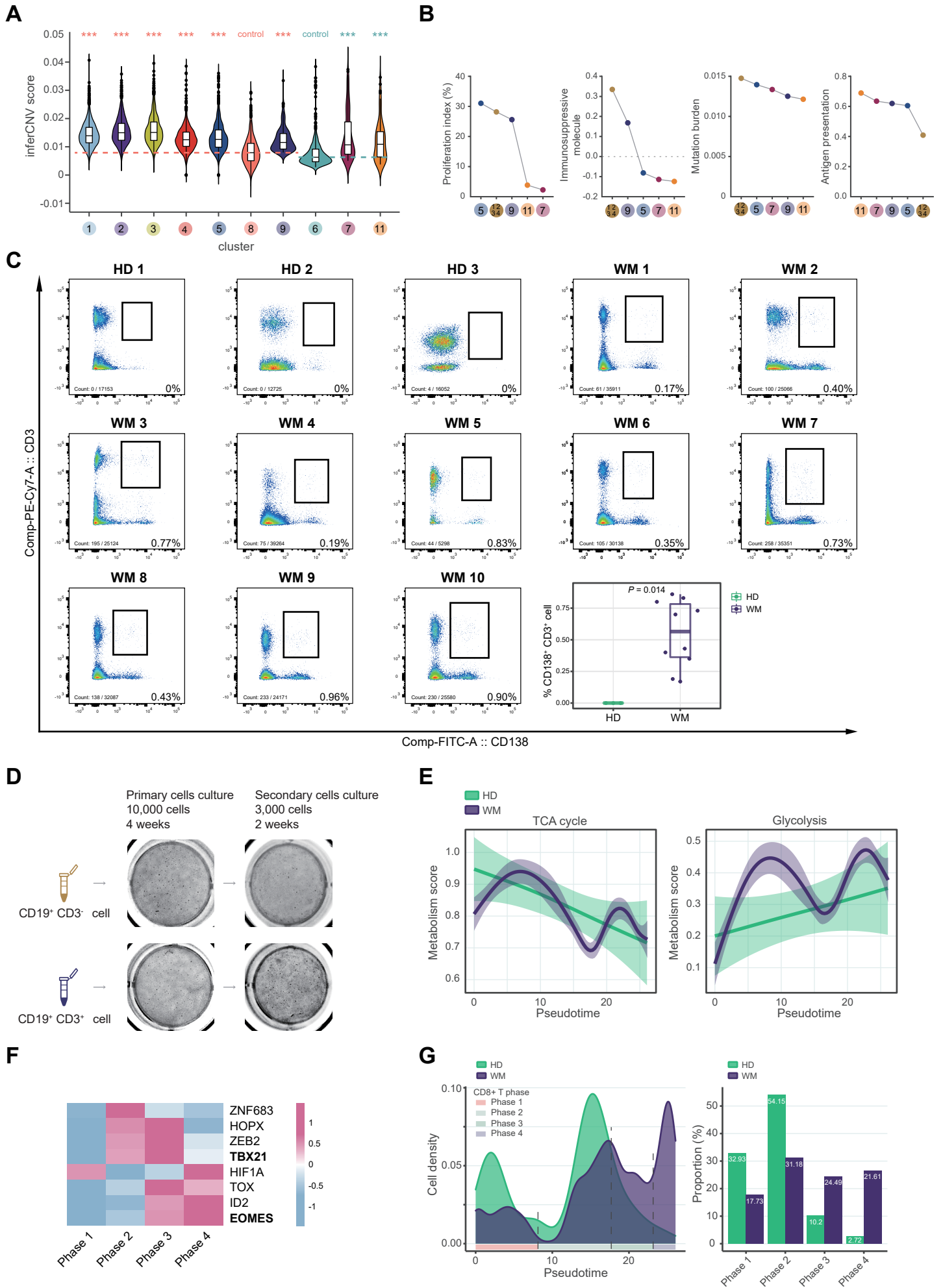(J) Bar chart showing the enrichment of GO pathways, based on the DEGs of HSPCs between WM and HD.

Fig. S2

Fig. S2 | Tumor heterogeneity and T cells state in WM

(A) Violin plots showing the inferred copy number variation (CNV) scores among 10 subclusters of CD79$^+$ cells. CD19$^+$ cells and CD138$^+$ cells were compared separately with the control group by unpaired Wilcoxon test. ***P < 0.001.

(B) Line charts representing the median of proliferation, immunosuppressive molecule, mutation burden, and antigen presentation scores among 5 malignant WM subpopulations (sub-C1 to sub-C4, sub-C5, sub-C7, sub-C9, and sub-C11).

(C) Density dot plots of flow cytometry analysis displaying the population of CD138$^+$CD3$^+$ cells derived from bone marrow aspirates of 3 healthy donors (HDs) and 10 WM patients. Boxplot showing the statistical result with unpaired Wilcoxon test between HDs and WM patients (bottom right).

(D) Colony formation assay of CD19$^+$CD3$^+$ cells and CD19$^+$CD3$^-$ cells in WM.

(E) Two-dimensional plots showing the expression scores for genes related to the TCA cycle and glycolysis, in HD (green) and WM (purple) samples, along with the pseudotime.

(F) Heatmap indicating the expression of transcriptional factors (TF) in 4-phase CD8$^+$ T cells.

(G) Density line plot showing the cell distribution of HD and WM along with the pseudotime, by state (left). Histogram displaying the cell proportion of 4-phase CD8$^+$ T cells in HD and WM samples (right, Chi-square test between HDs and WM patients, P < 0.001).