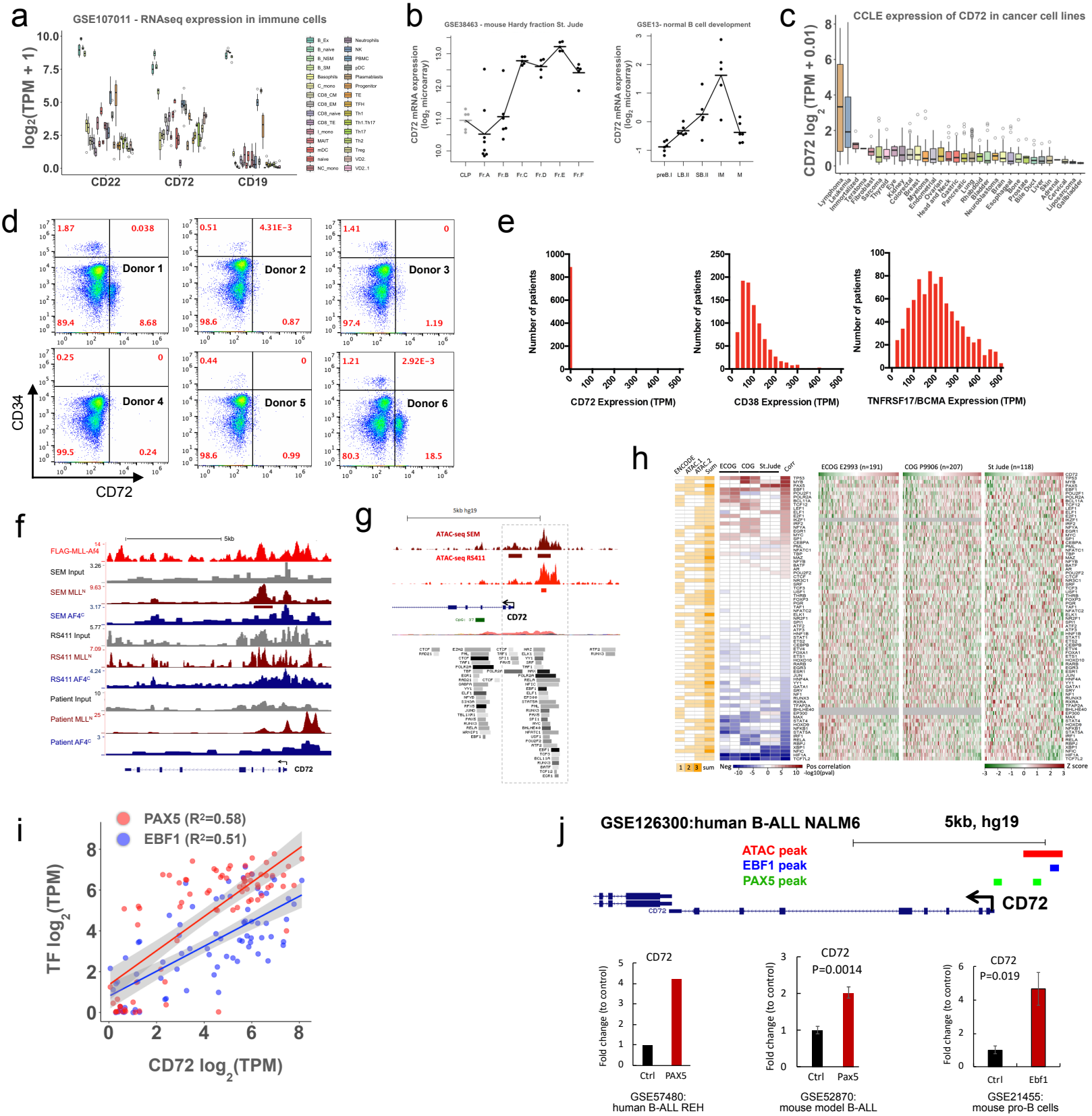
Supplementary Figure S4



Supplementary Figure S4: Transcriptome analysis of CD72 in normal and malignant cells
(a) Transcript abundance of CD22 and CD19, and CD72, in 29 different immune cell types measured by RNAseq (Human Protein Atlas Database, GSE107011, https://www.proteinatlas.org). **(b)** Transcript abundance of CD72 in B cell precursor populations from common lymphoid progenitor (CLP) through Hardy stage F (mouse Hardy Fraction, St. Jude, GSE38463), and normal B-cell development (GSE13). **(c)** Transcript abundance of CD72 in malignant cell lines (n=1461; CCLE, accessed October 14th 2019). TPM, transcript per million mapped reads. **(d)** Cell surface abundance of CD72 and CD34 on primary patient apheresis samples. **(e)** Transcript abundance of CD72, CD38, and BCMA cell surface receptors in malignant plasma cells (research.themmrf.org). **(f)** ChIP-seq tracks focused on the CD72 genomic locus, showing differential binding peaks for the N-terminus of the MLL protein, or the C-terminus of the AF4 protein, in MLL-AF4 patient samples and cell lines (GSE83671; GSE38403; GSE84116). (g) ATAC-seq signature of CD72 genomic locus in the SEM and RS411 cell lines (MLLr B-ALL) with candidate transcription factors and their binding sites (GSE117865).**(h)** Heat maps displaying Pearson correlations for 75 candidate transcription factors and their positive and negative correlation with CD72 transcript abundance in three B-ALL patient cohorts (ECOG E2993, n=191; COG P9906, n=207; St. Jude, n=118). Transcription factors ranked high to low -log10(p-value), with most likely CD72 gene locus interacting transcription factors at the top. Predicted transcription factor binding at the CD72 locus is shown in the left most heat map according the ENCODE ChIP-seq database as well as ATAC-seq in the SEM and RS411 cell lines (GSE117865), along with summation of all binding evidence (Sum column). **(i)** Correlation plot comparing CD72 gene expression to PAX5 and EBF1 gene expression in malignant cell lines of B-cell origin (n=77; CCLE, access October 14th 2019). **(j)** Combined ATAC-seq and ChIP-seq tracks in the NALM6 B-ALL cell line focused on the CD72 genomic locus. EBF1 and PAX5 binding peaks are highlighted. Fold expression change of CD72 in different cell lines engineered to over-express PAX5 or EBF1 transcription factors.