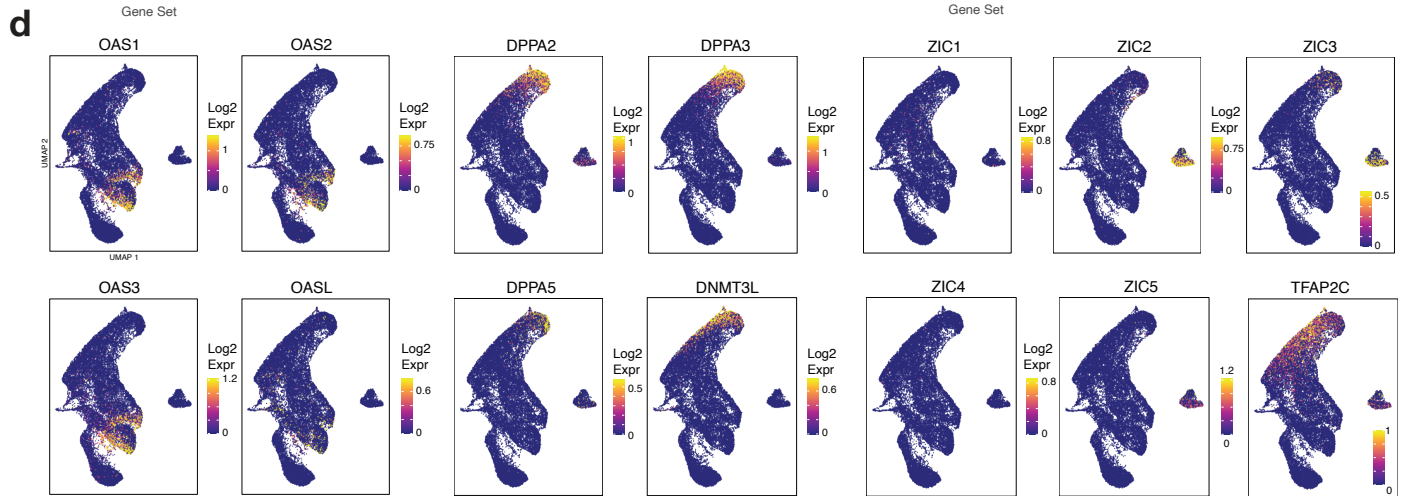
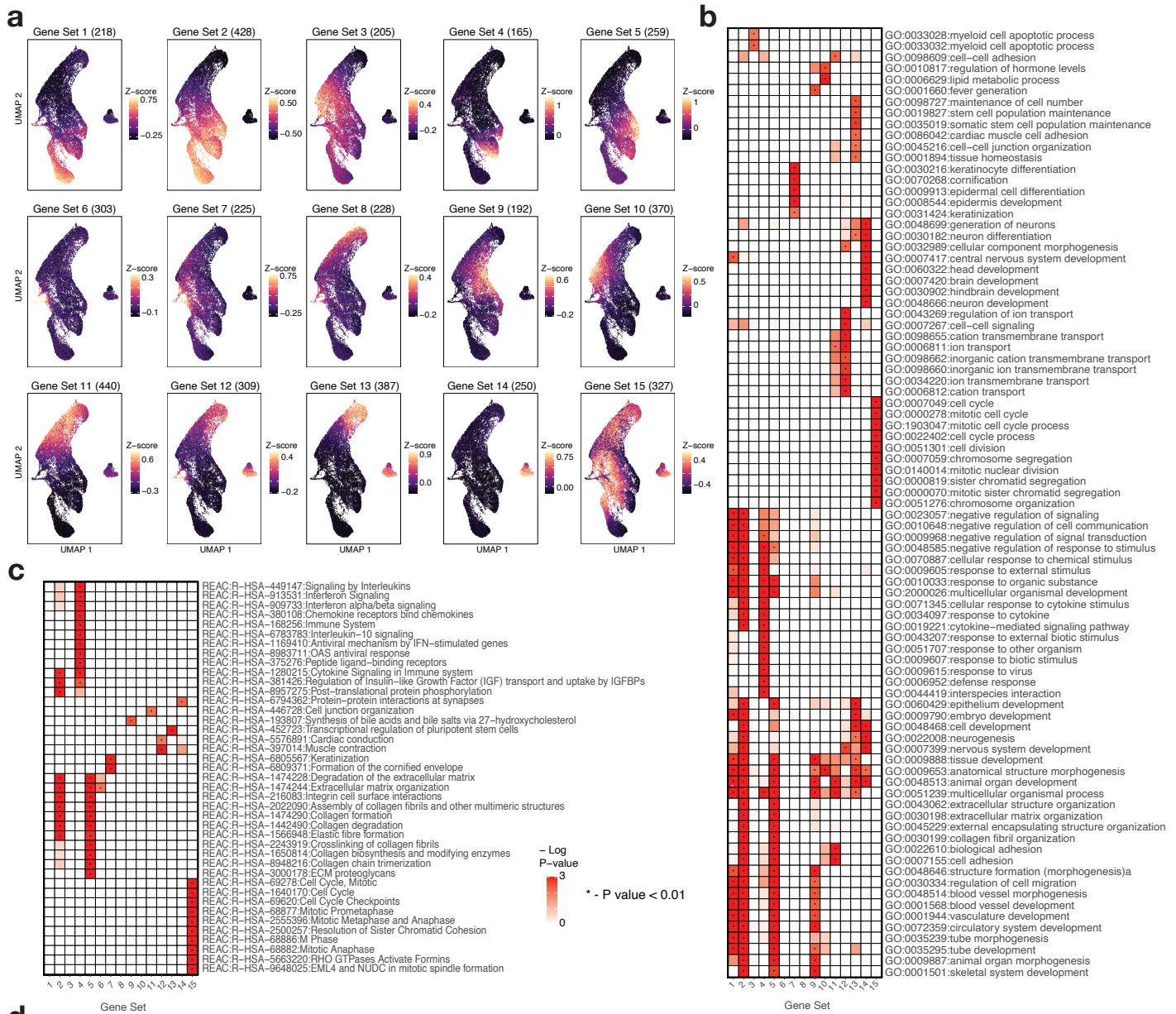


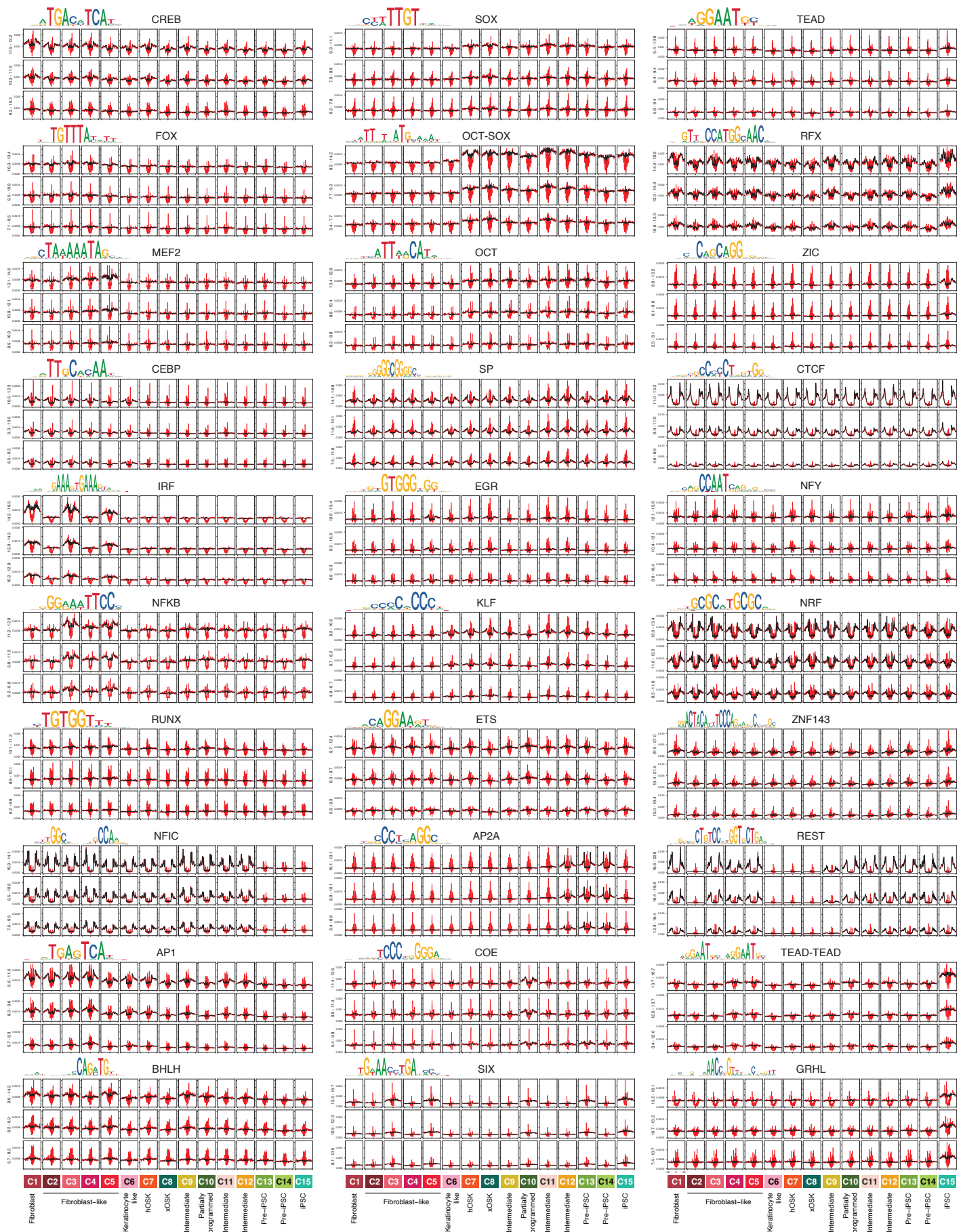
Extended Fig 1:

a) Quality control for scATAC-seq samples b) Quality control for scRNA-seq samples



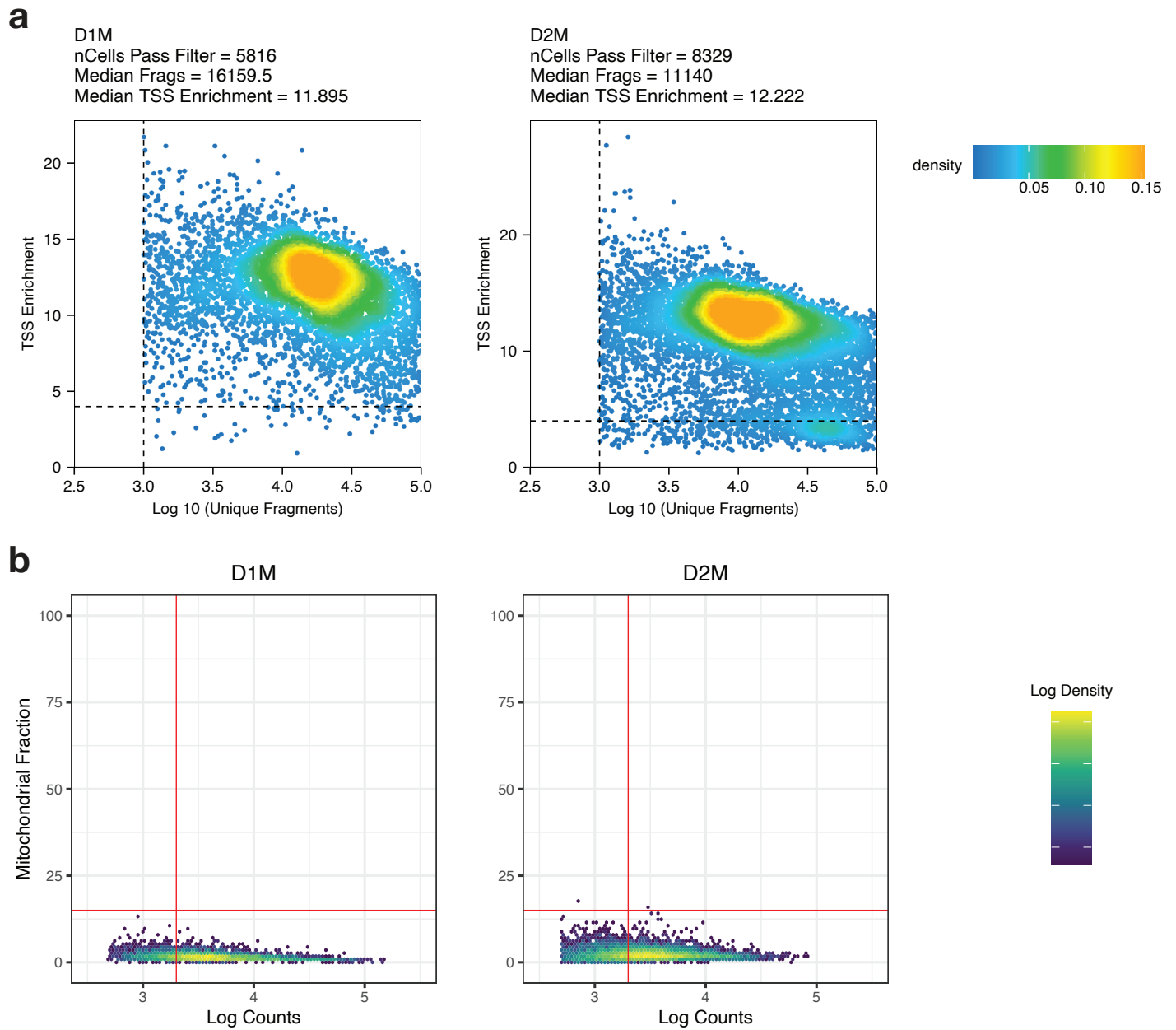
Extended Fig 2:

a) Gene expression Z-scores for 15 gene sets based on clustering of scRNA-seq data. Number of genes in each set indicated in parenthesis. These gene sets are distinct from the gene sets obtained by finding genes linked to peaks within the 20 peak sets **b)** Gene Ontology Biological Process terms enrichment matrix across the 15 gene sets **c)** Reactome (REAC) terms enrichment matrix across the 15 gene sets **d)** scRNA-seq gene expression plots for genes from the OAS, DPPA, and ZIC families and TFAP2C.



Extended Fig 3:

Virtual footprinting for all 30 motifs obtained by inserting their instances stratified by log-odds scores into random background sequences and averaging predicted profile probability distributions with (red) and without (black) bias for each cell state.



Extended Fig 4:

a) Quality control for ATAC-seq component of single nucleus multiome samples **b)** Quality control for RNA-seq component of single nucleus multiome samples