

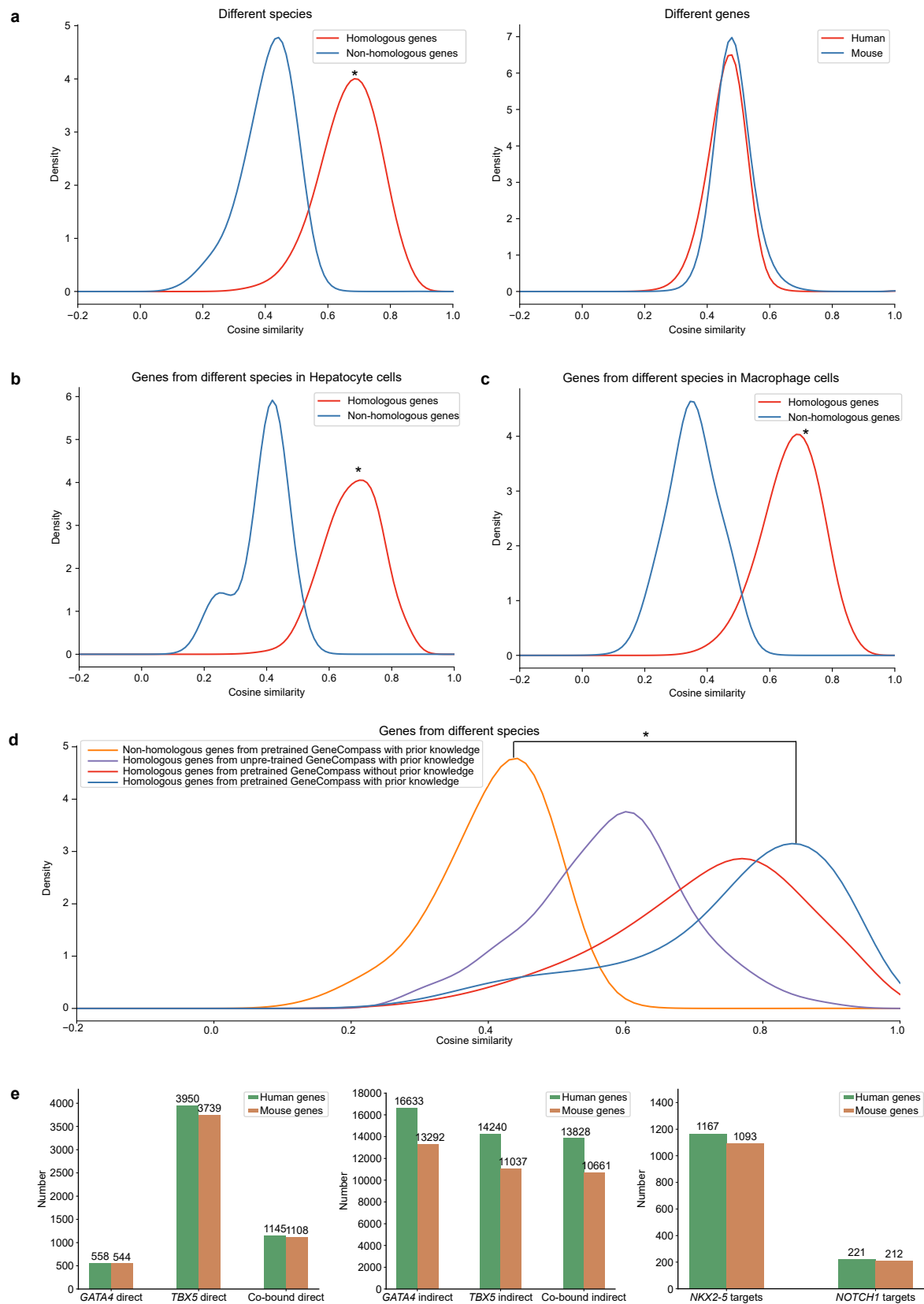


Fig. S2| Analysis of gene embedding and statistics of corresponding genes. **a**, Cosine similarity among different genes in the same species (left). Cosine similarity between homologous and non-homologous genes of different species (right). A more generalizable experiment were performed without restricting to specific cell types, including GABAergic neuron, Cardiomyocyte, Fibroblast, Erythroblast, and etc. **b**, Cosine similarity between homologous and non-homologous genes in Hepatocyte cells of different species. **c**, Cosine similarity between homologous and non-homologous genes in Macrophage cells of different species. **d**, Comparison among cosine similarity of non-homologous gene embeddings from the pretrained GeneCompass with prior knowledge and homologous gene embeddings which came from pretrained GeneCompass with prior knowledge, pretrained GeneCompass without prior knowledge, and un-pretrained GeneCompass with prior knowledge. **e**, Statistics of genes in the deletion of *GATA4* and *TBX5*. Mouse genes are obtained by homologous mapping. (* $P < 0.05$ wilcoxon-test, NS no significance)