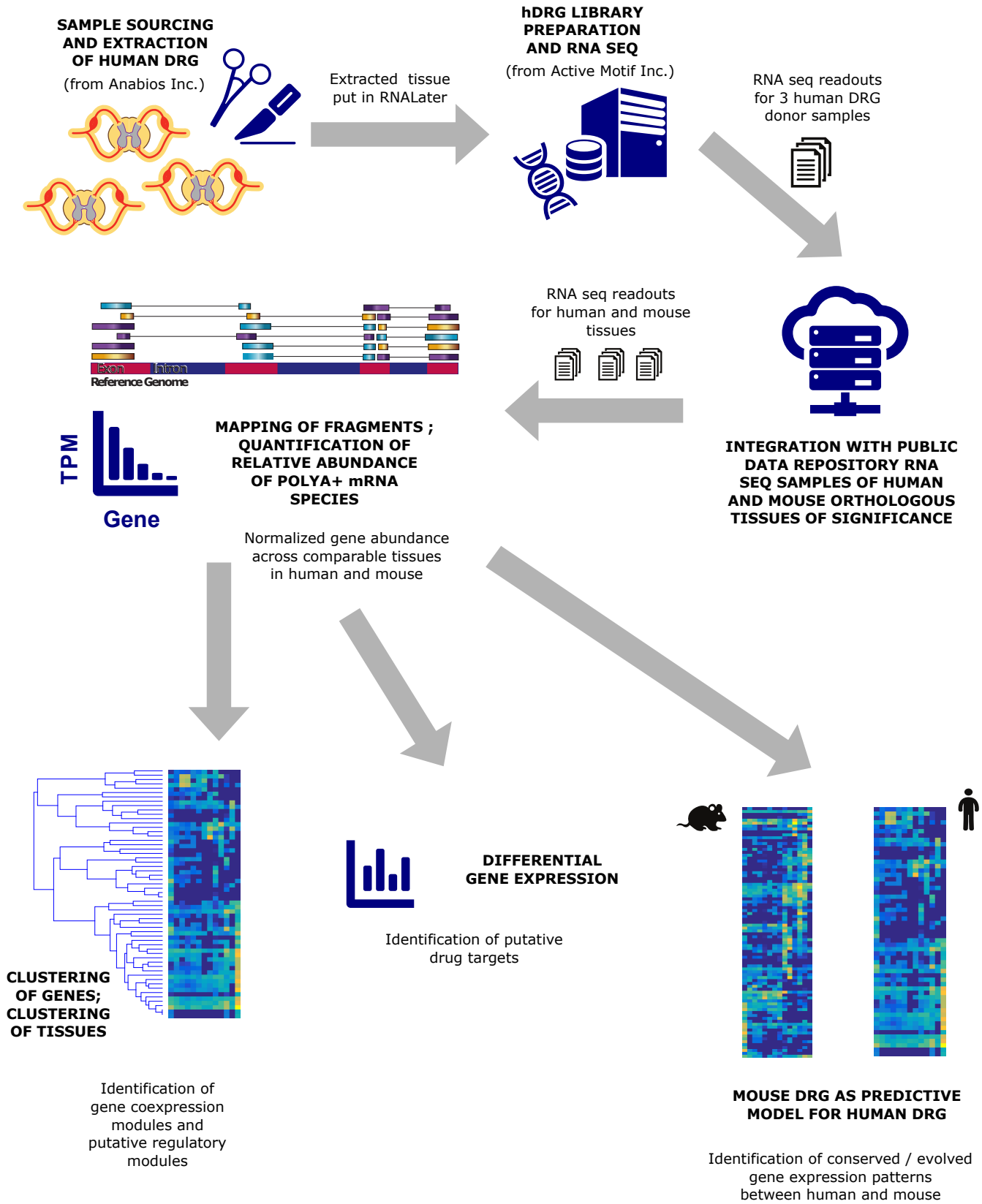
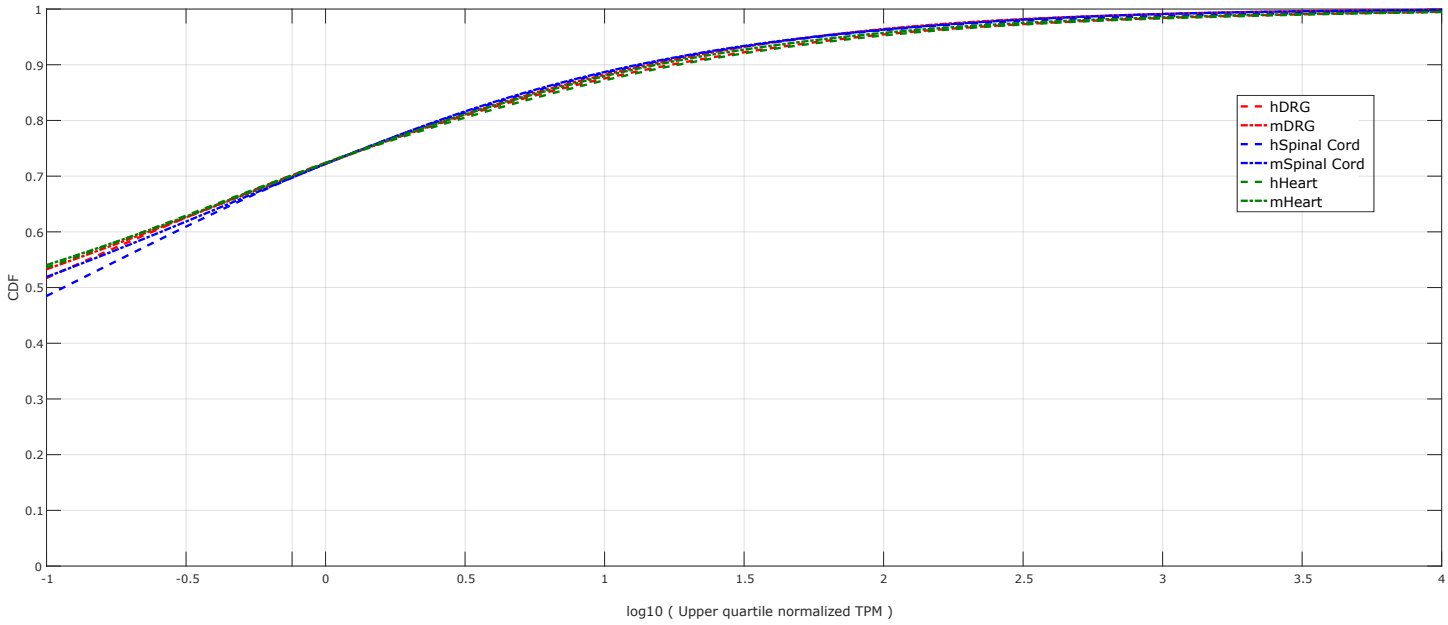
Supplementary figures

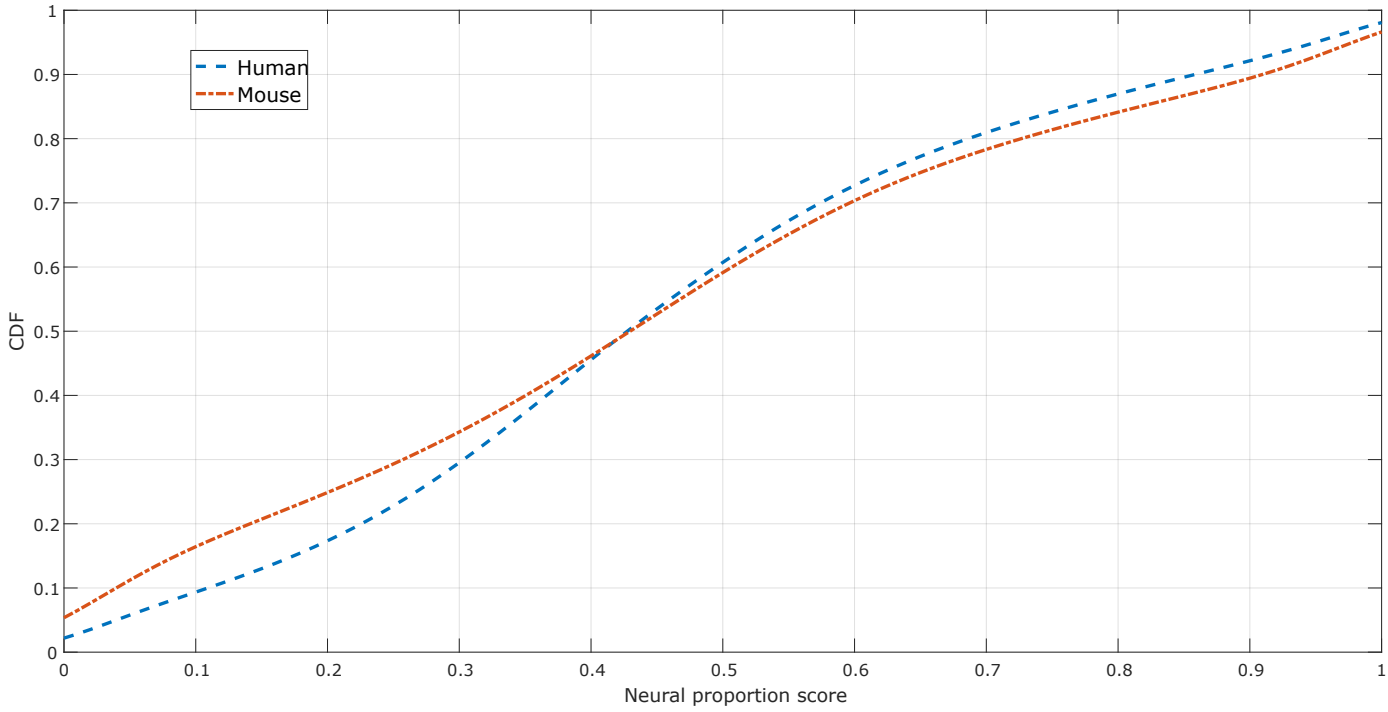


Supplementary Figure 1. Schema for generation and analysis of mRNA abundance profiles. The broad characterization of our work showing mRNA sequencing, RNA-seq mapping and quantification, gene expression normalization across samples and genes, and characterization of transcript profiles in terms of their tissue specificity, expression patterns across tissues, and conservation of expression pattern between human and preclinical mouse model is shown.

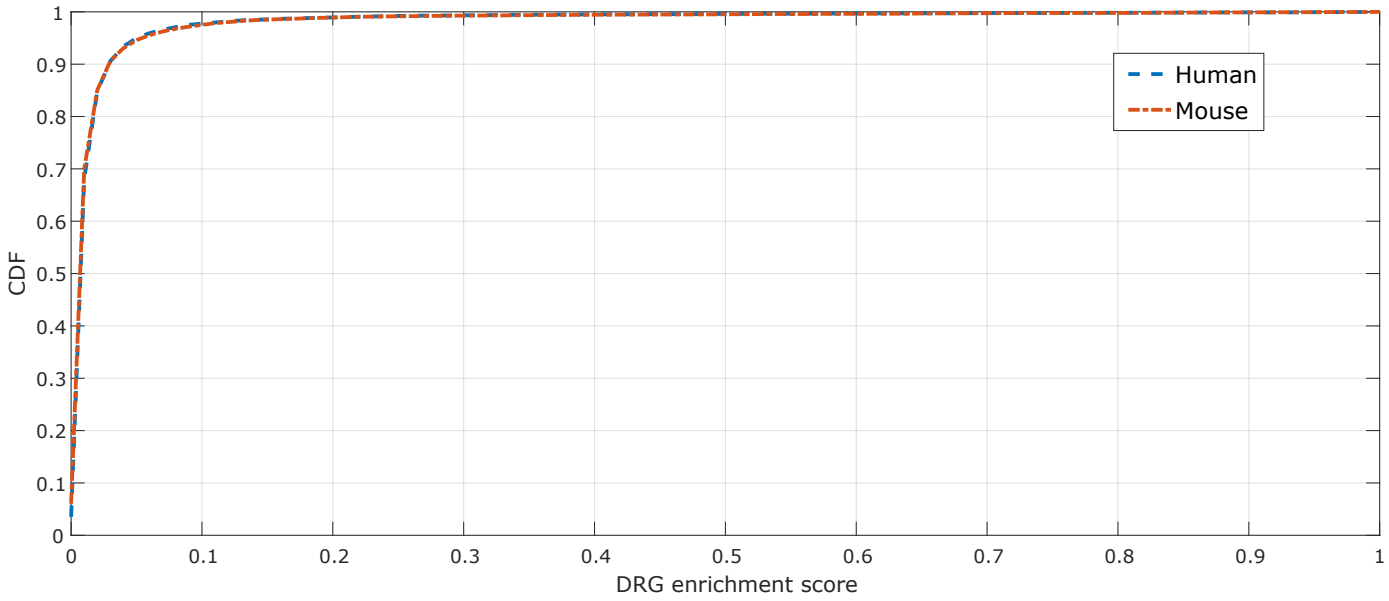


Supplementary Figure 2. Cumulative distribution function for gene uqTPMs of DRG and other analyzed tissues. CDFs of the relative abundance of coding genes (marked known in the Gencode annotation) across tissues (DRG, spinal cord, heart) and across species (human, mouse) track in similar ways above 0.75 uqTPM (-0.12 in logscale). Smoothing factor of 0.1 was added across all uqTPMs to avoid differences in the CDF at low abundance levels due to different sequencing depths or library preparation techniques.

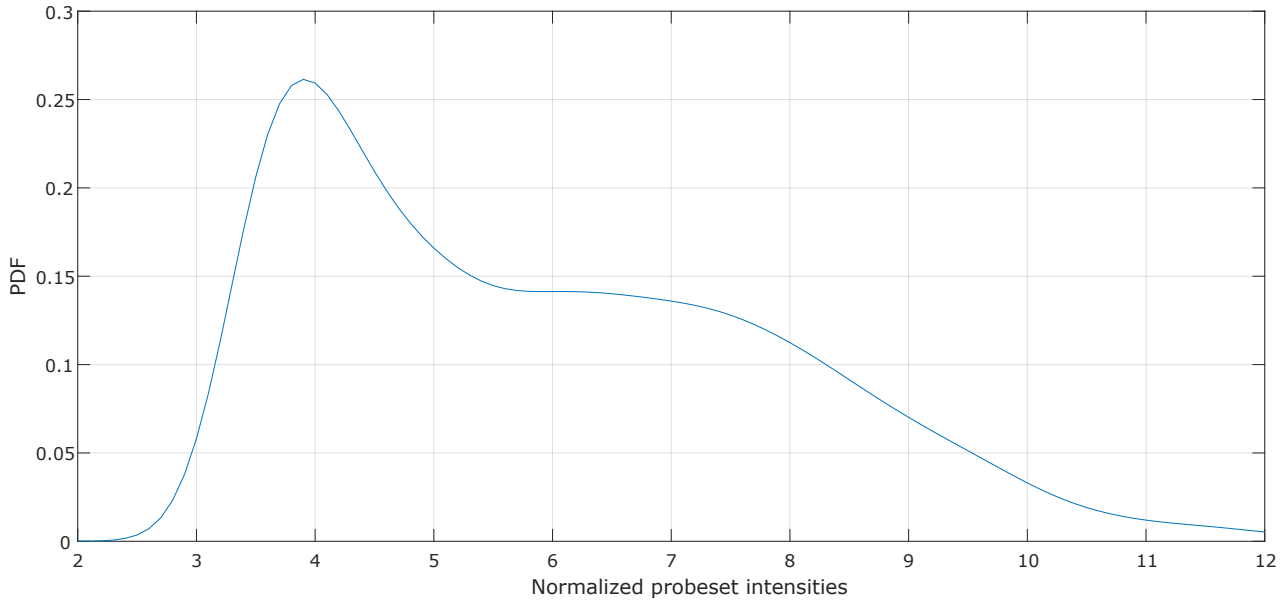
(A)



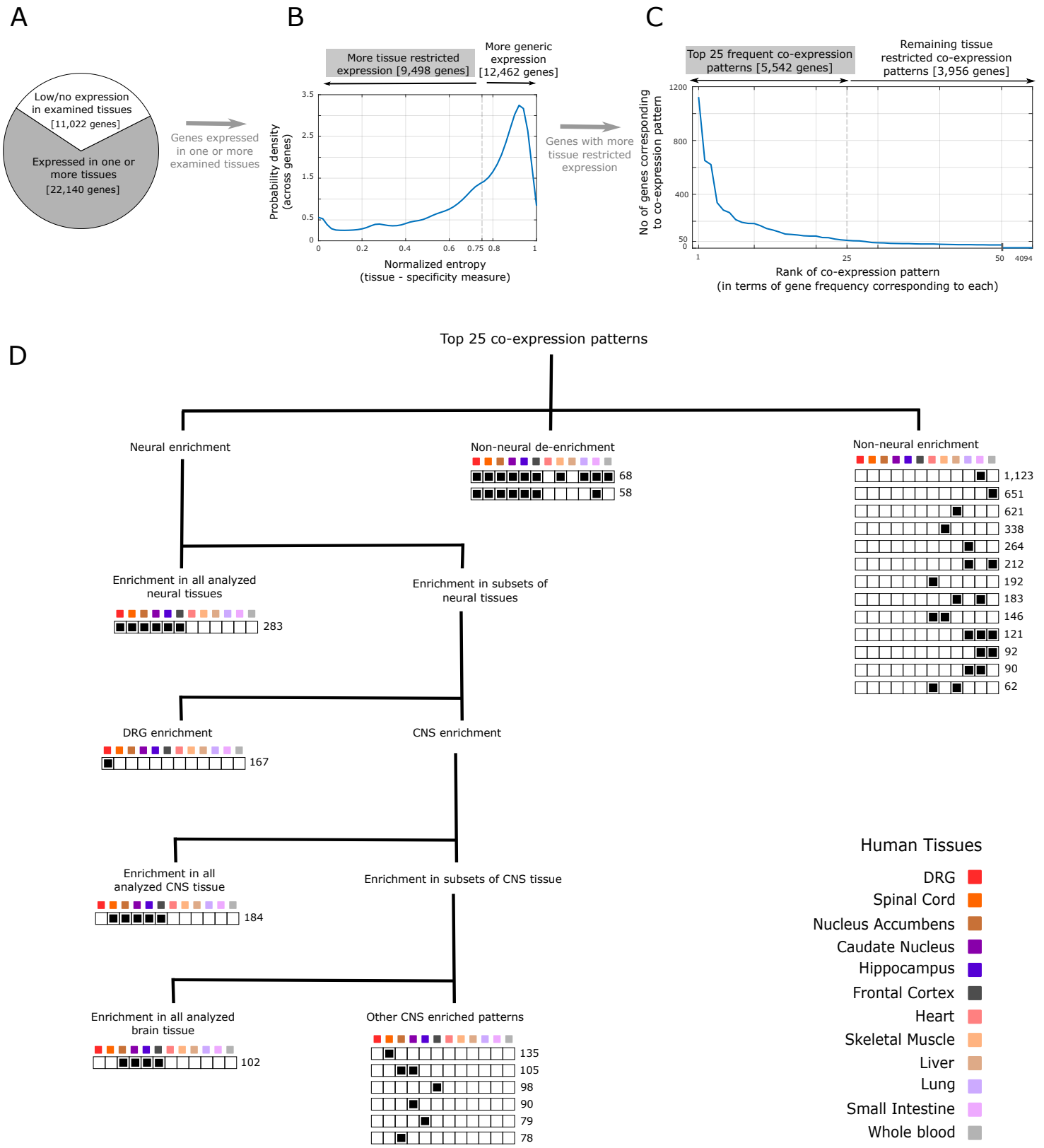
(B)



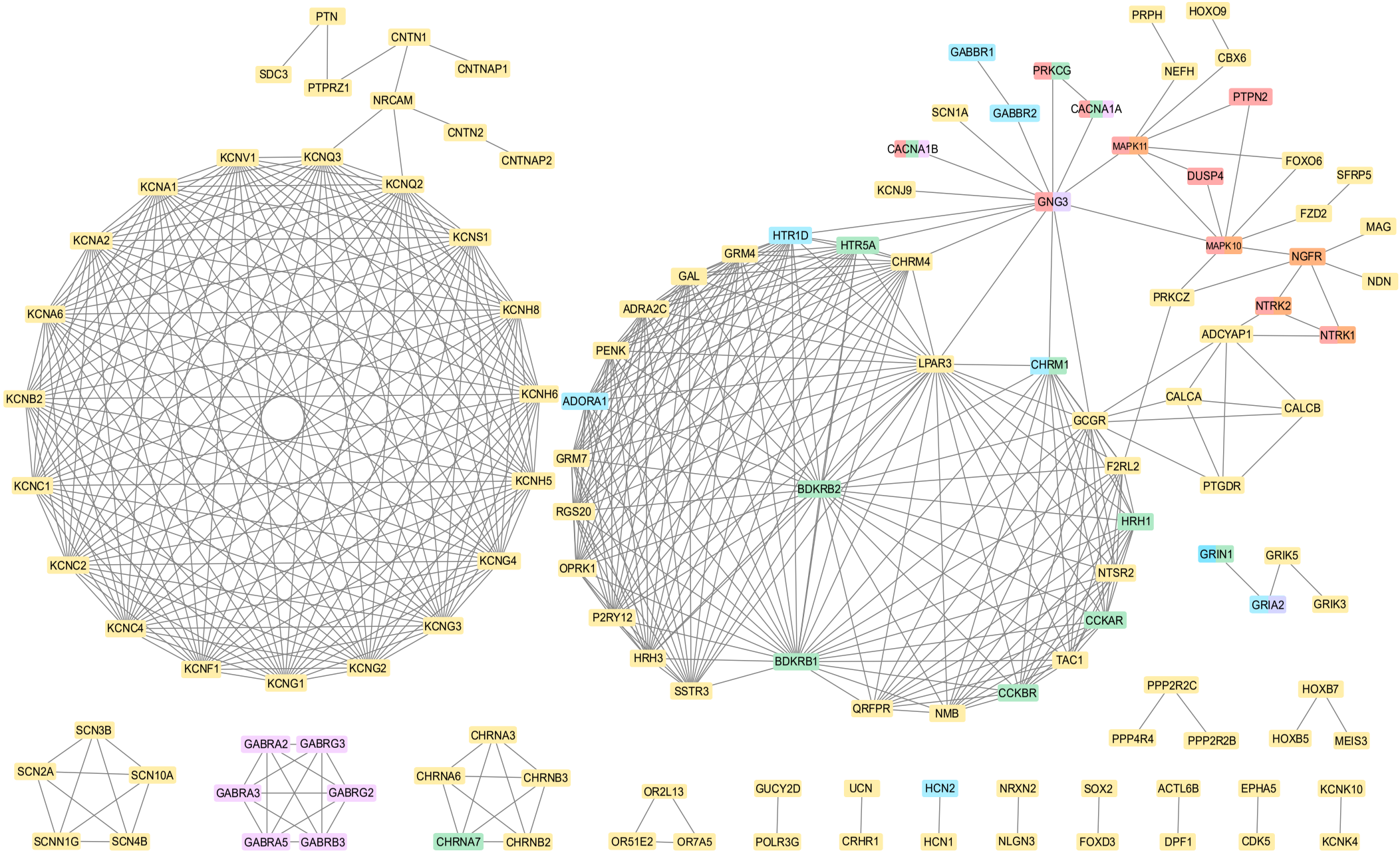
Supplementary Figure 3. CDFs for neural proportion and DRG enrichment. (A) CDFs for human and mouse neural proportion scores for coding genes differ between human and mouse, possibly due to the evolutionary divergence of the CNS transcriptome between human and mouse. **(B)** CDFs for human and mouse DRG enrichment scores for coding genes are very similar, suggesting a similar proportion of DRG-enriched genes in human and mouse.



Supplementary Figure 4. Estimated density function for normalized probeset values analyzed. Normalized probeset values for microarray datasets for human DRG, TG, cultured NHSCs, and cultured fibroblasts show a clear mode for low expressed / undetected probes near 4. A threshold between 5.0 and 6.0 allows for a classification of probesets based on the density function.



Supplementary Figure 5. Identification of common tissue-restricted human gene co-expression patterns in our analyzed tissue panel. (A) ~22,000 out of ~33,000 genes are expressed in at least one tissue based on our digital co-expression pattern **(B)** 9,498 genes showed tissue-restricted expression (based on a normalized entropy measure of tissue diversity) **(C)** 25 co-expression patterns are sufficient to account for 5,542 of these genes, in a power law–like distribution. **(D)** 25 most common co-expression patterns : grouped into 3 classes : enrichment patterns in neural tissues, enrichment patterns in non-neural tissues and de-enrichment patterns in non-neural tissues. The neural enrichment patternshave been further subdivided into groups.



Supplementary Figure 6. PPI subnetwork constructed from StringDB database Several PPI hubs and cliques were identified for products of genes that are strongly hDRG-enriched or weakly hDRG-enriched and a member of one of the profiled gene families we analyzed. We identified several pathways that overlap the gene products involved in our reconstructed PPI subnetwork, including the Neurotrophin Signalling Pathway (marked in Orange, detailed in Figure 7). Graph drawing was performed in Cytoscape.