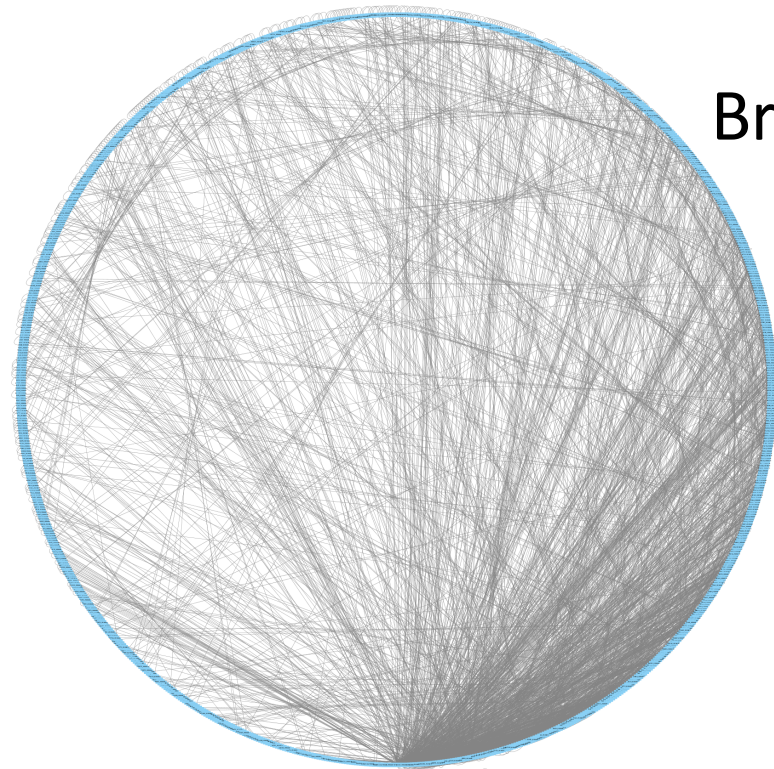


# Exploring in PPI Networks at the gene and transcript level in 16 tissues and cell types

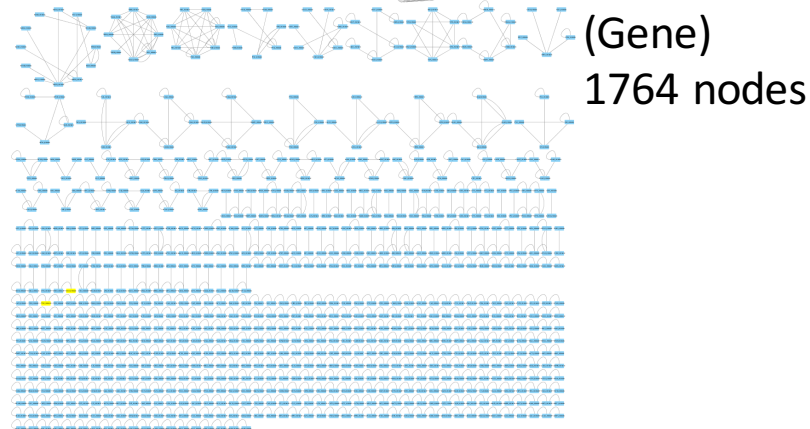
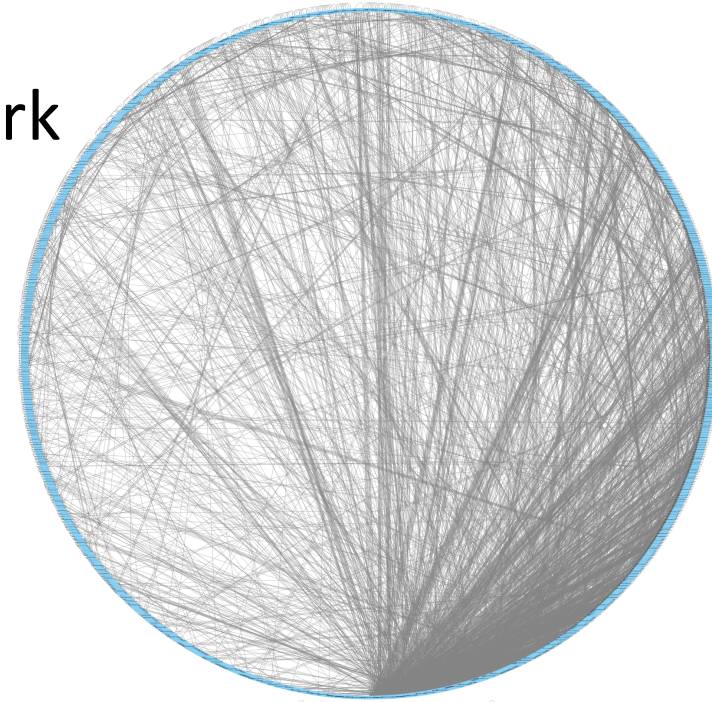
Daniel Burkhardt

with ANS + RK

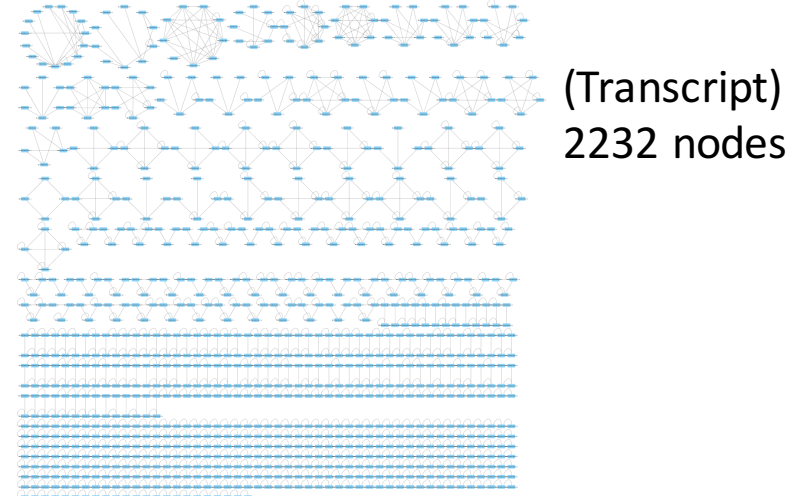
Question: Does using transcript level annotation of PPI networks give us more information about these networks?



Brain PPI Network



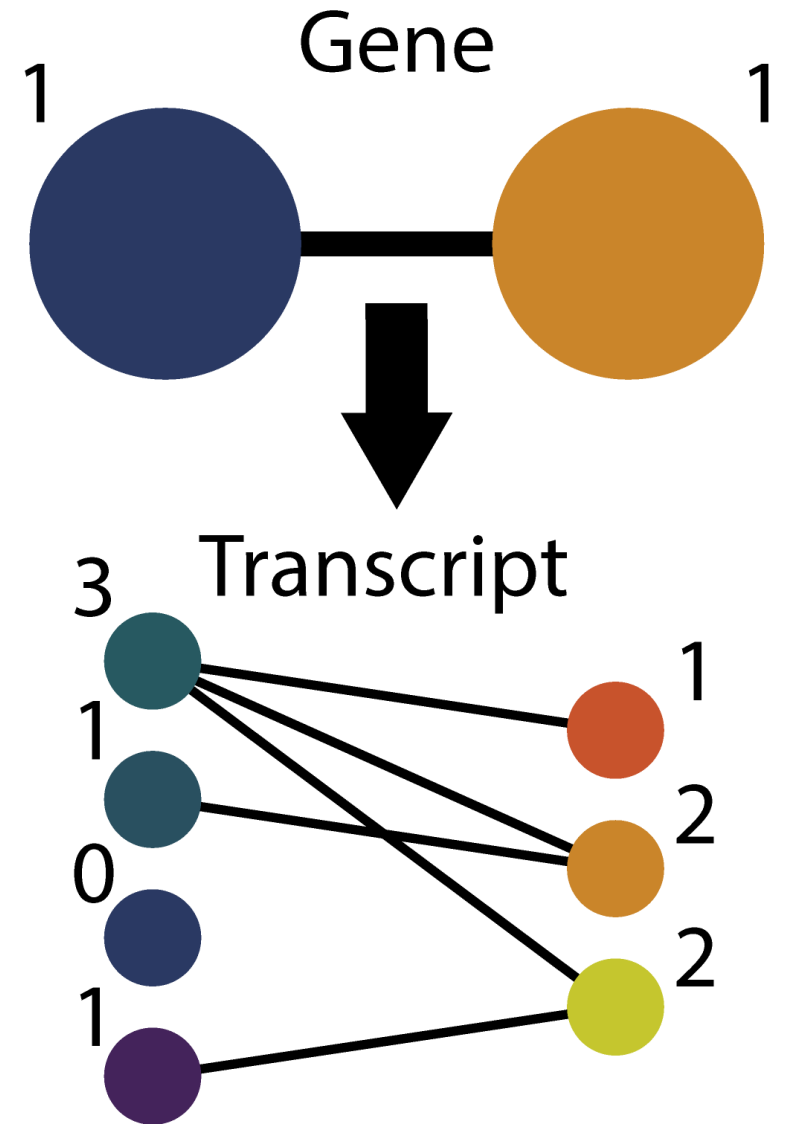
(Gene)  
1764 nodes



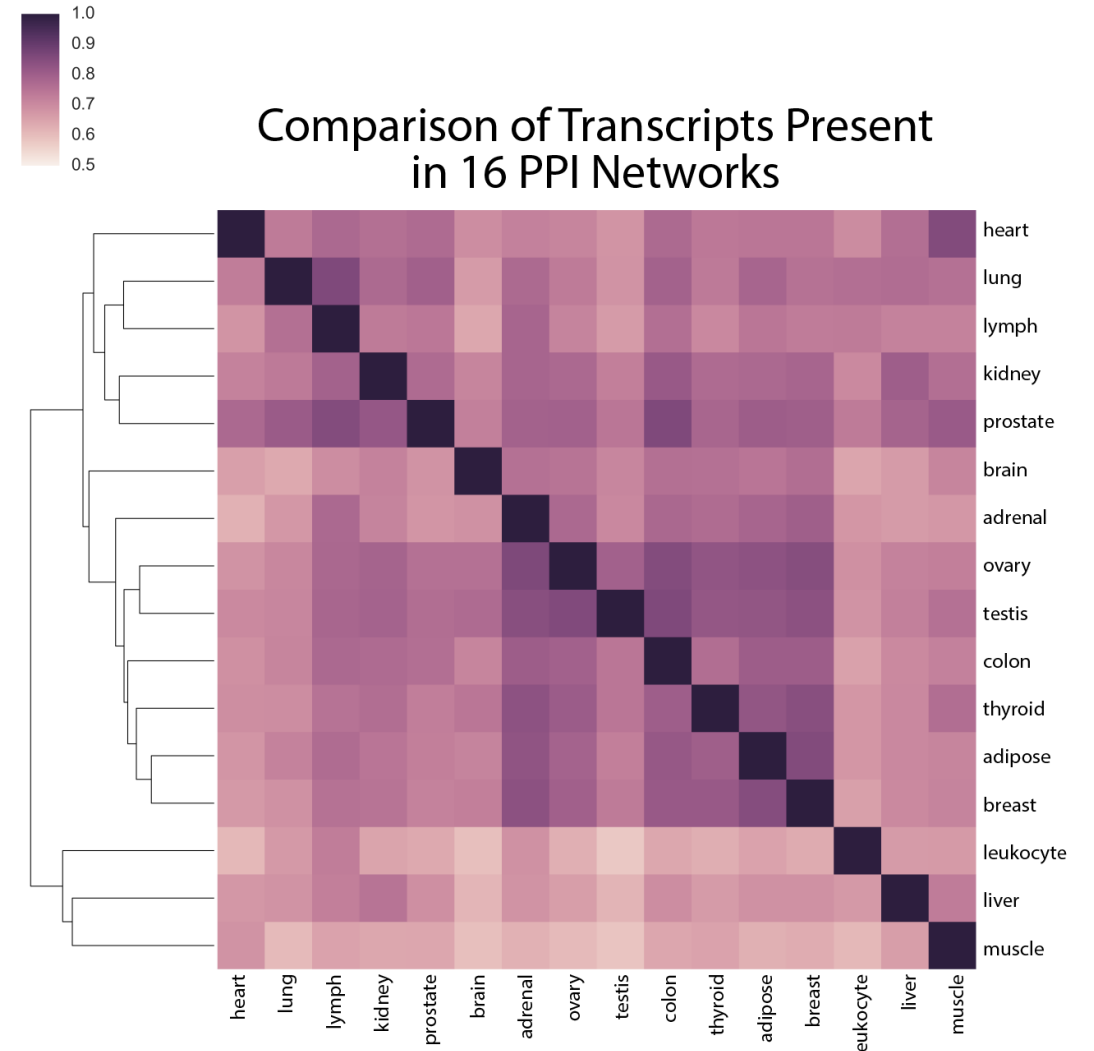
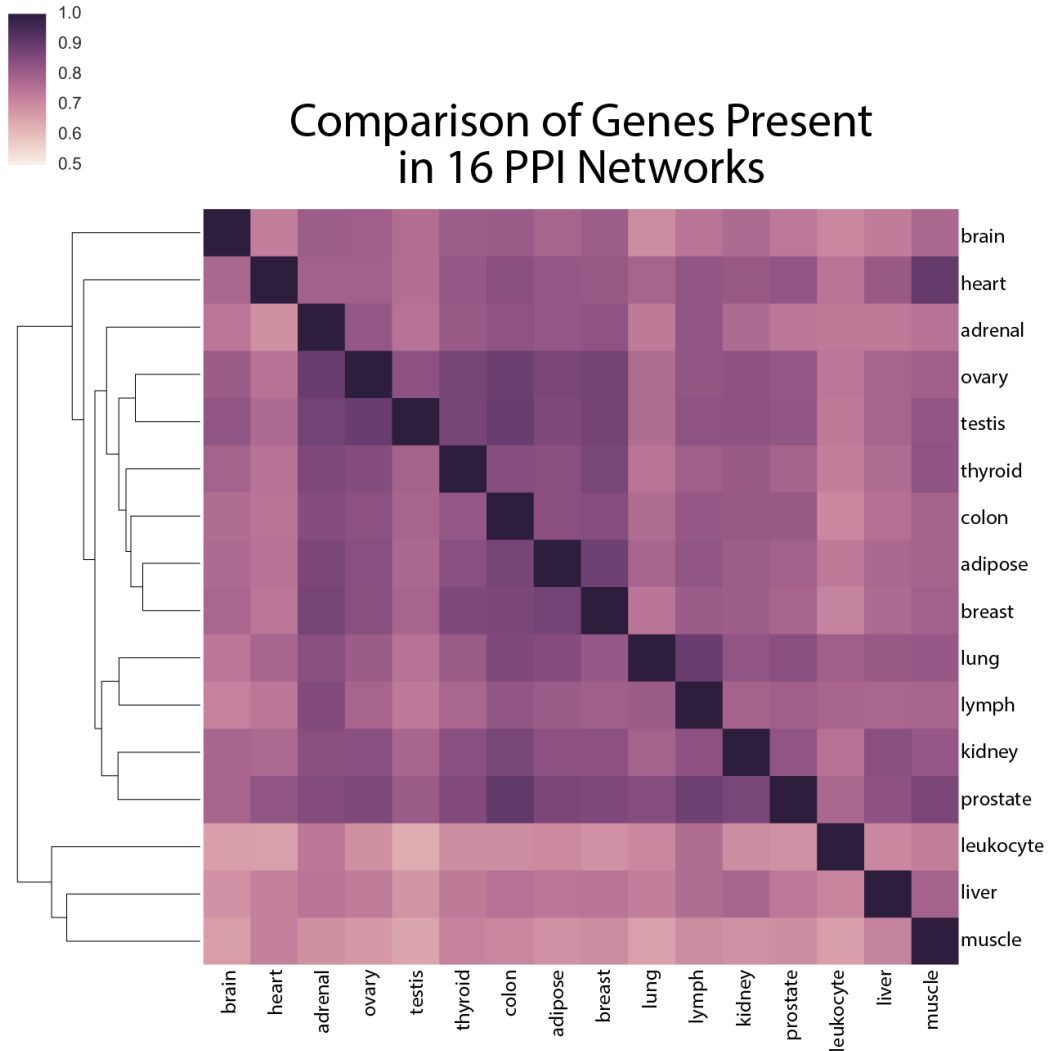
(Transcript)  
2232 nodes

# Workflow

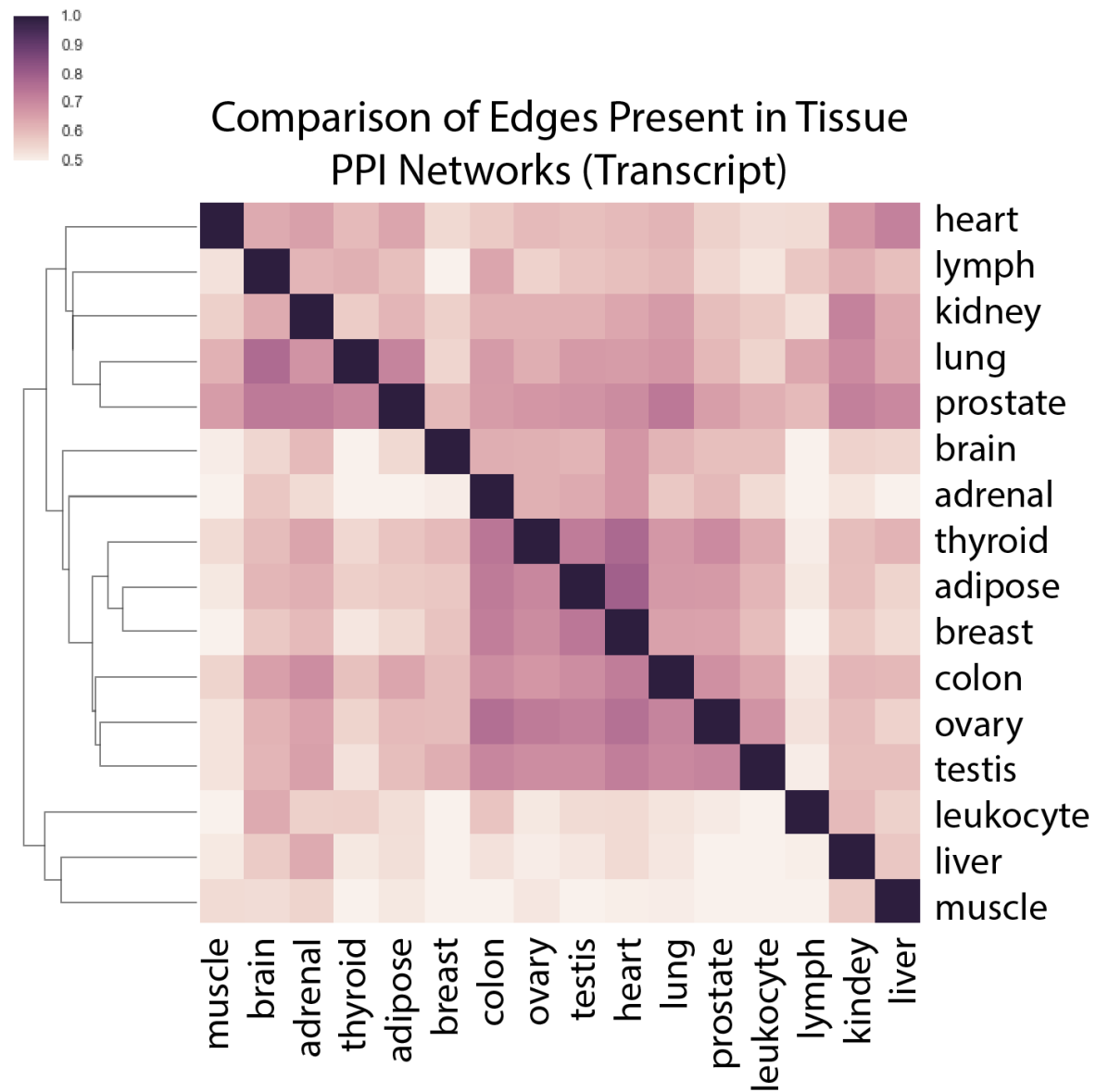
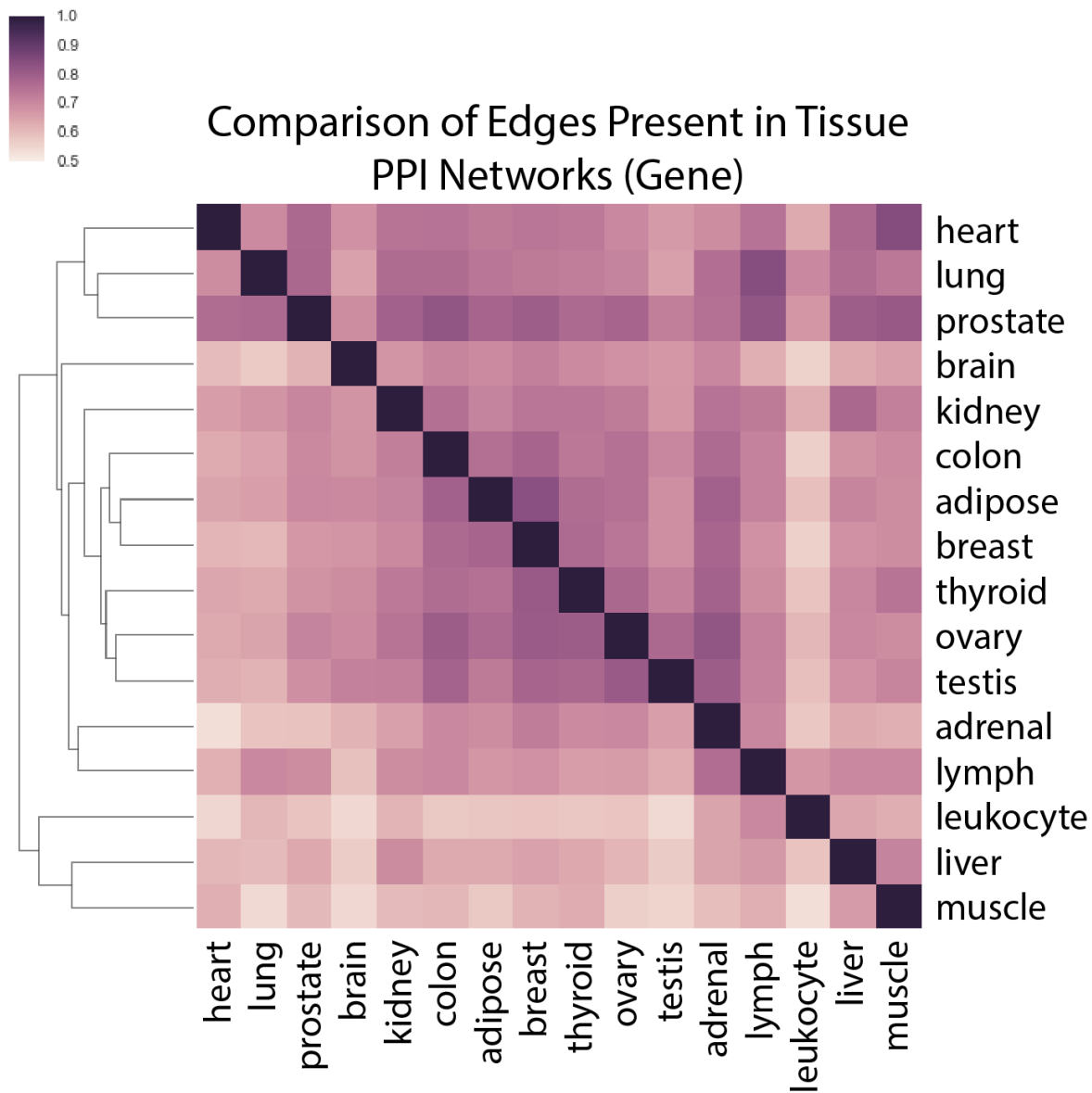
1. BodyMap 2.0 (RNA-seq – 16 tissues/cell types)
2. Gene-level and Transcript Level domain identification
3. Use HIPPIE to predict protein interactions
4. Build PPI Network



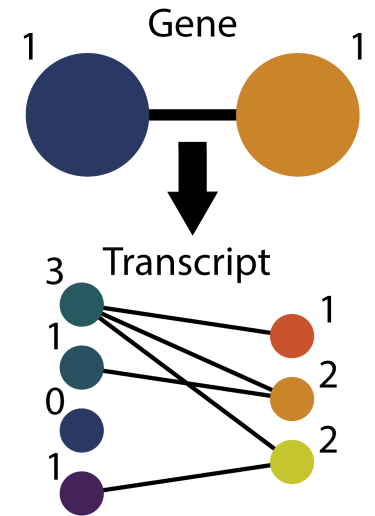
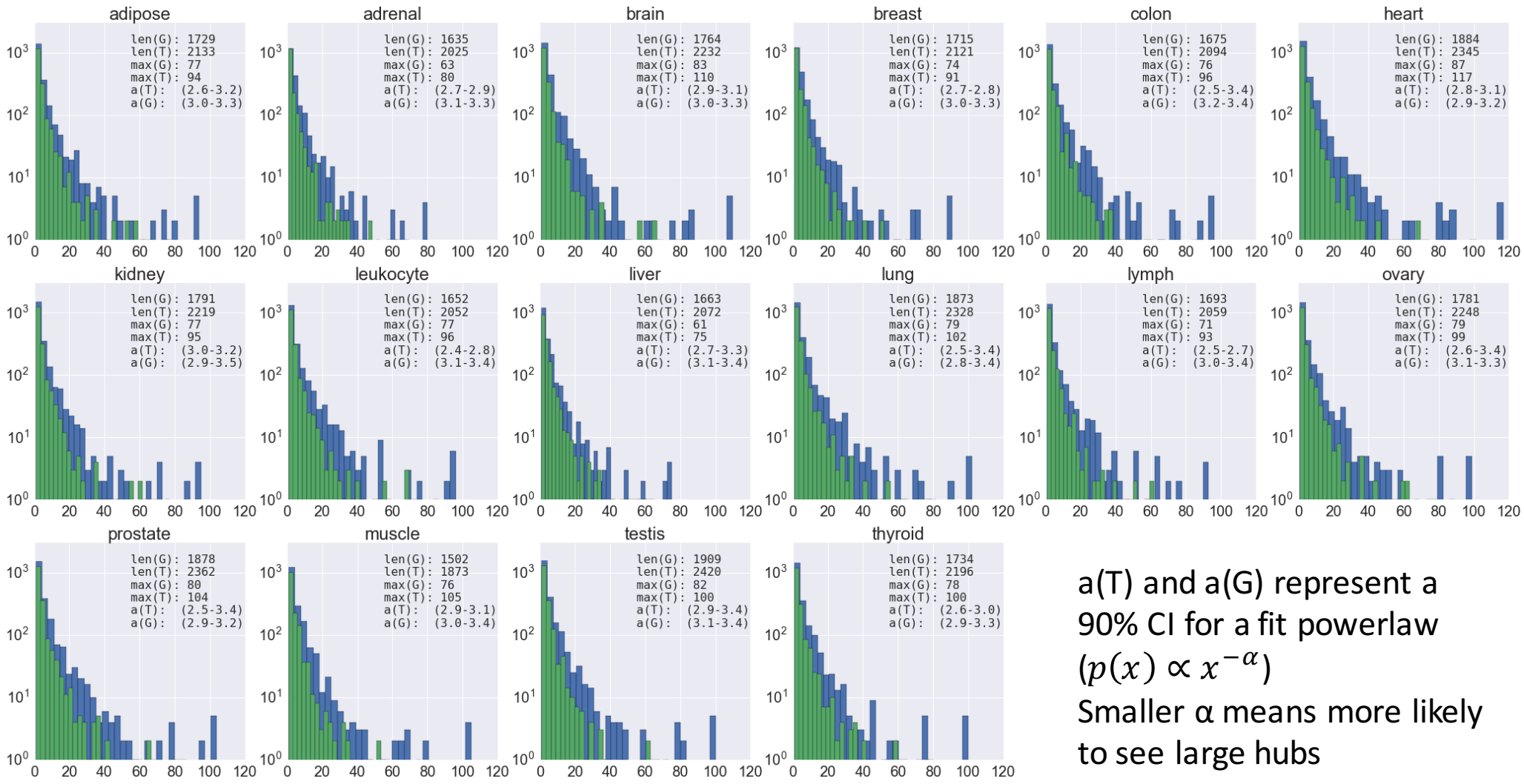
# Transcript networks are more diverse and leads to different clustering of tissues



# Edges are more diverse than nodes



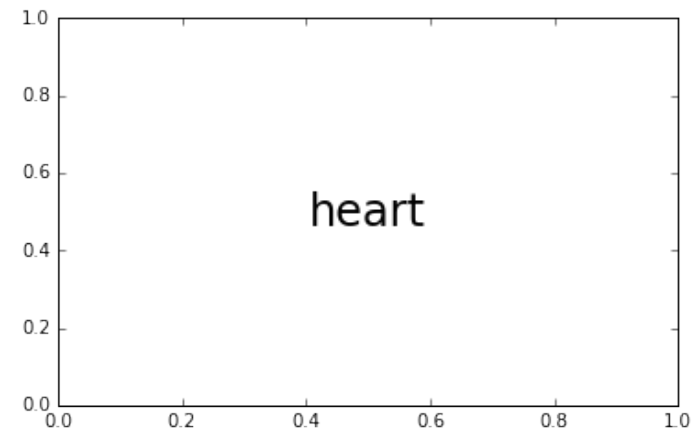
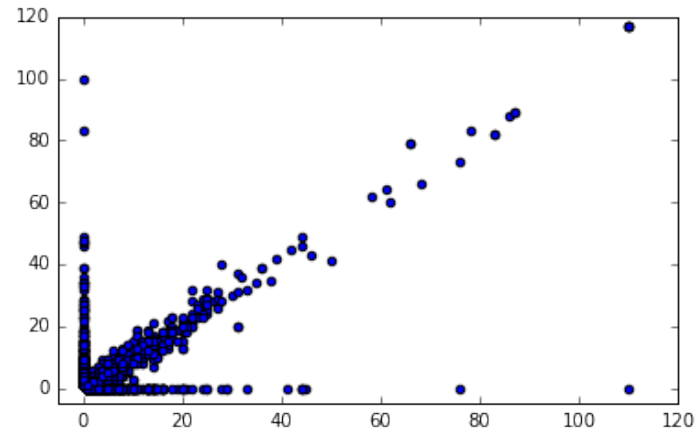
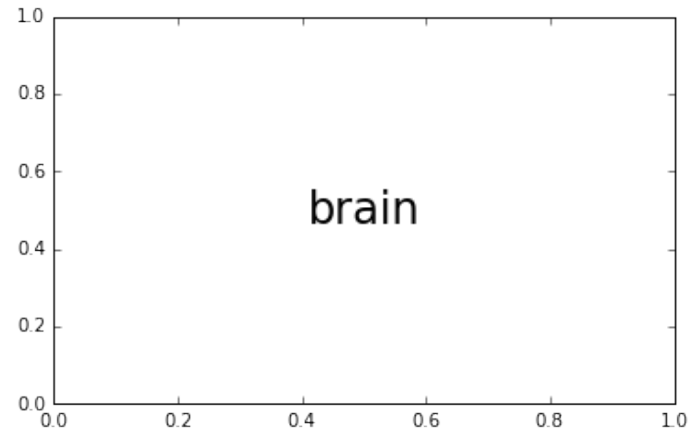
# More likely to see large hubs in transcript networks (blue) than gene (green)



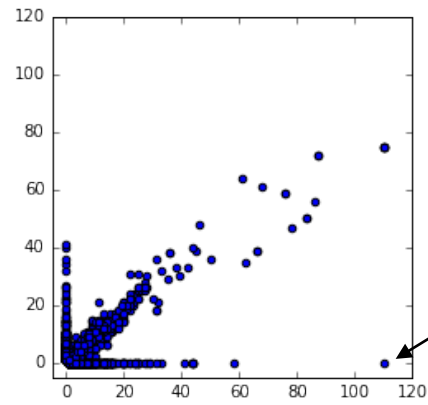
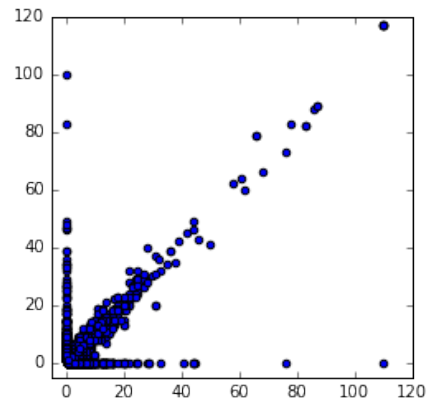
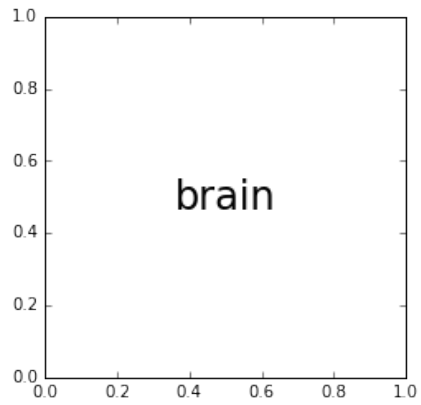
a(T) and a(G) represent a 90% CI for a fit powerlaw  $(p(x) \propto x^{-\alpha})$

Smaller  $\alpha$  means more likely to see large hubs

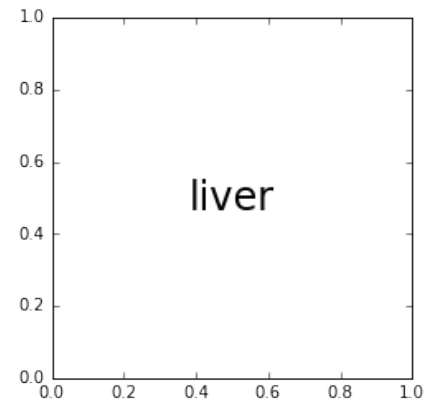
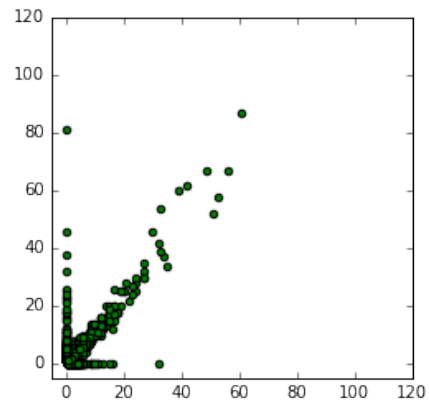
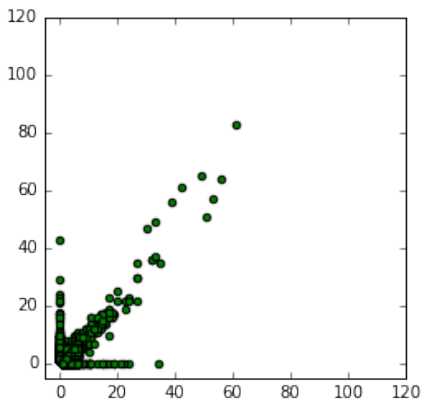
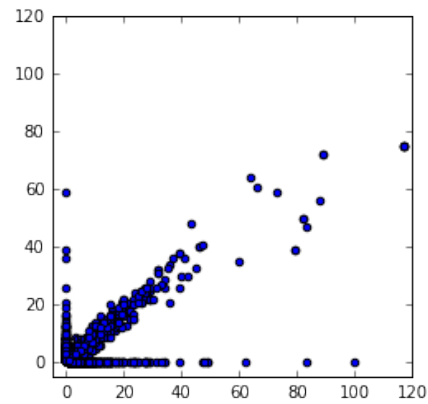
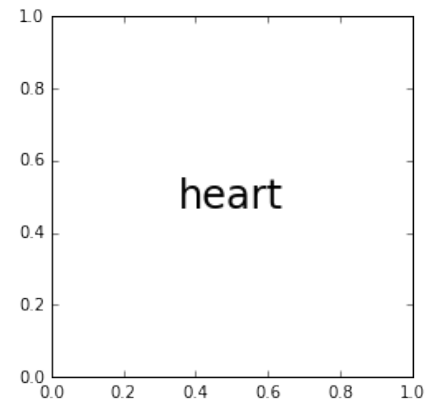
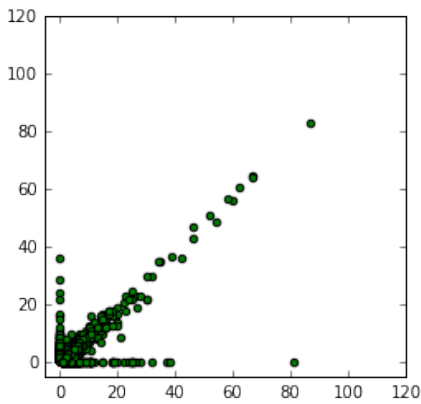
# Comparing degree of shared nodes



# Comparison of Brain Heart and Liver



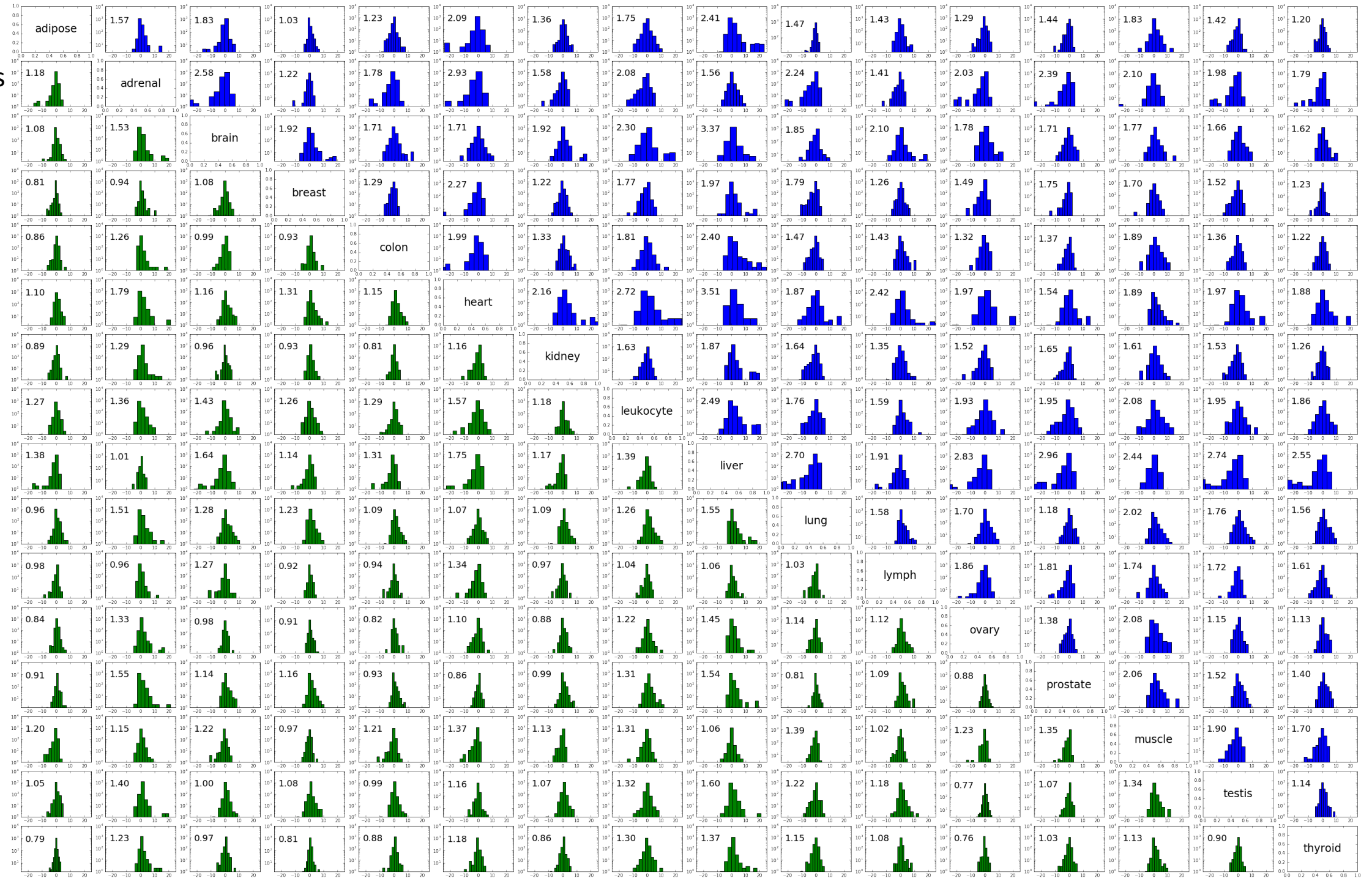
Brain-specific hub

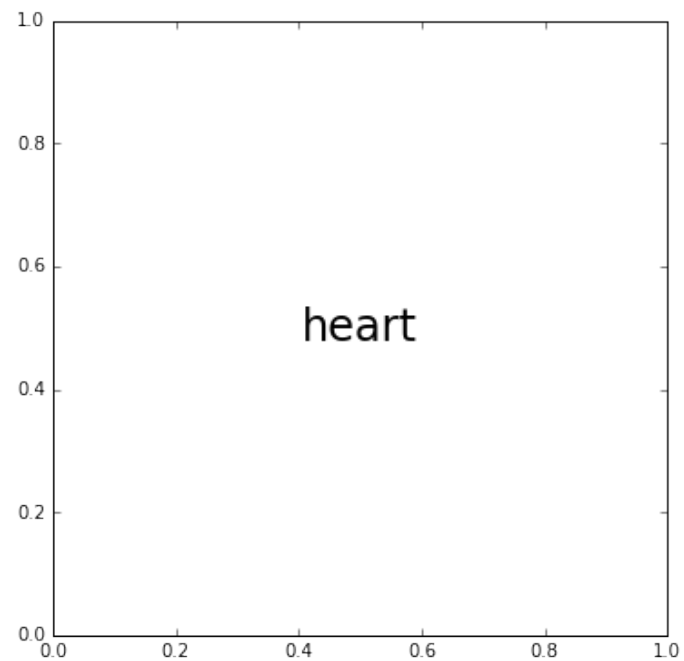
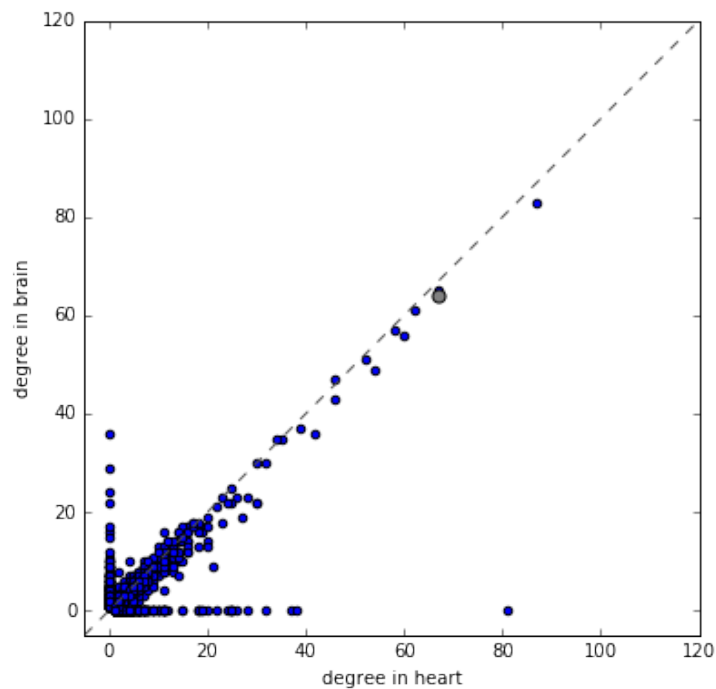
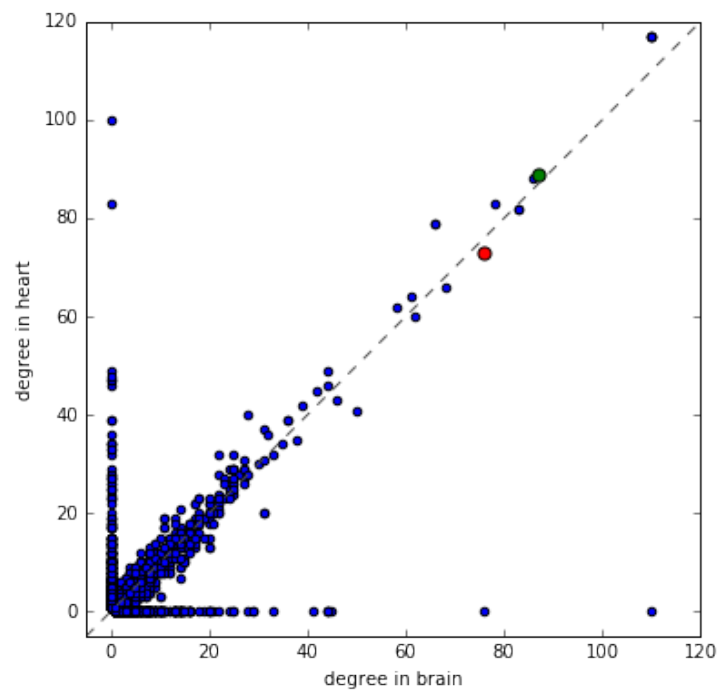
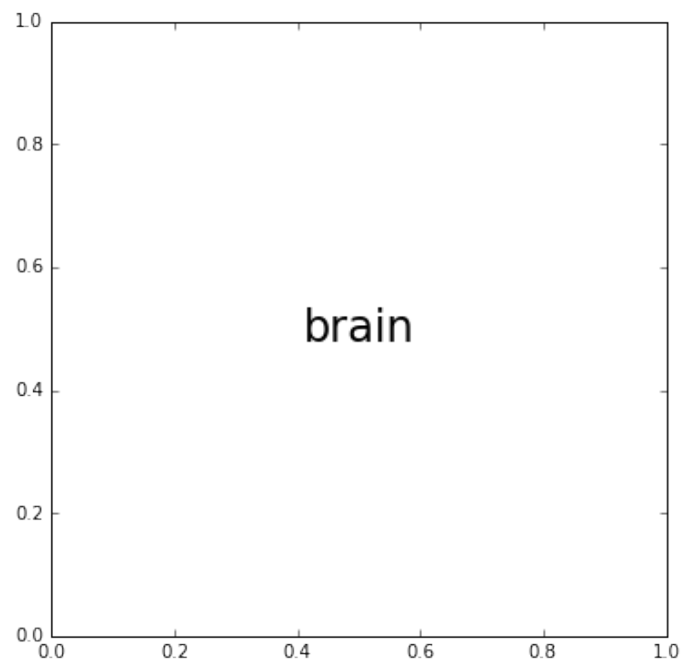






Inset  
numbers  
are  $s^2$

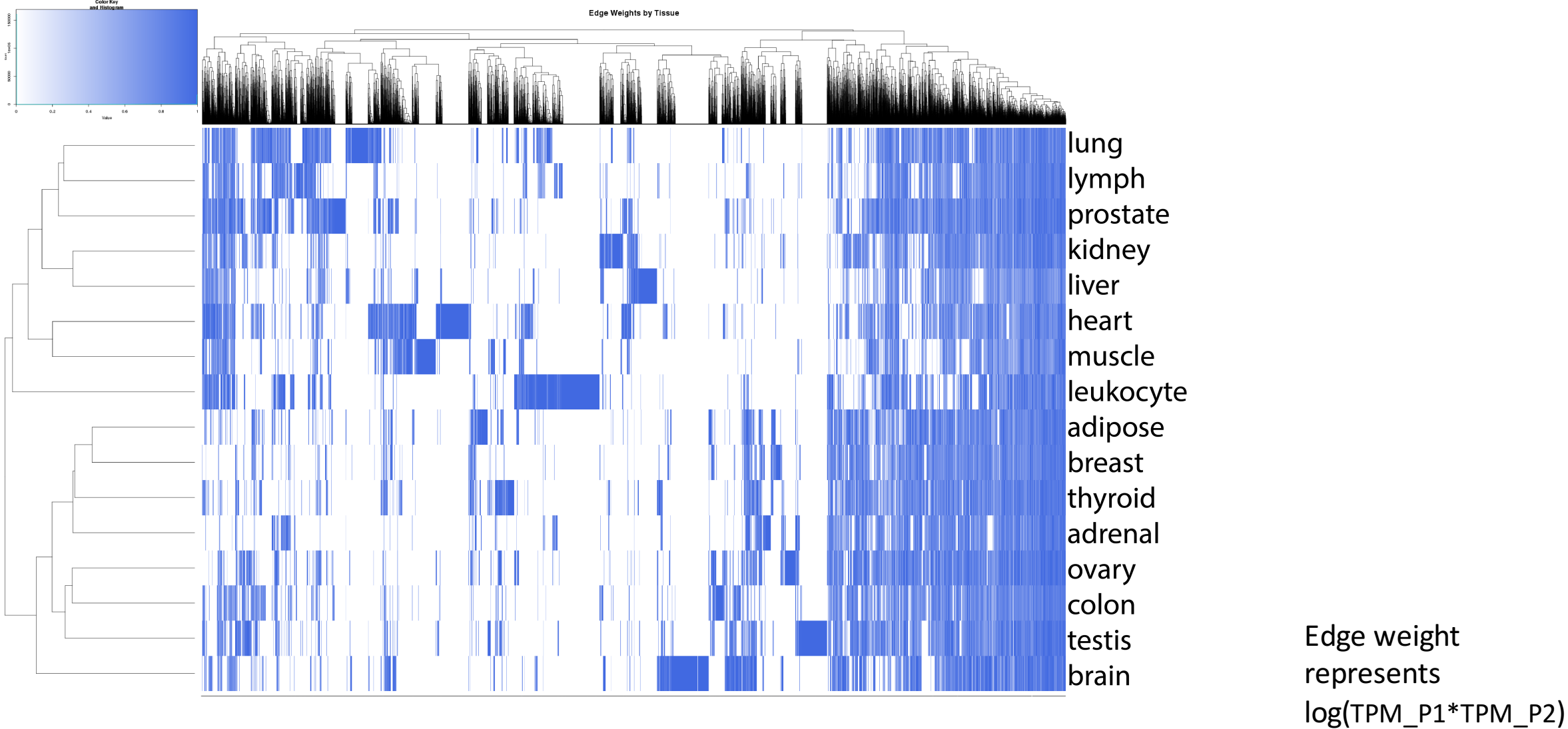




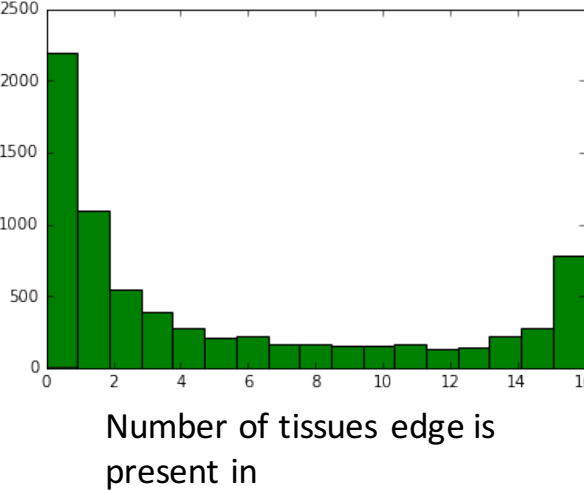
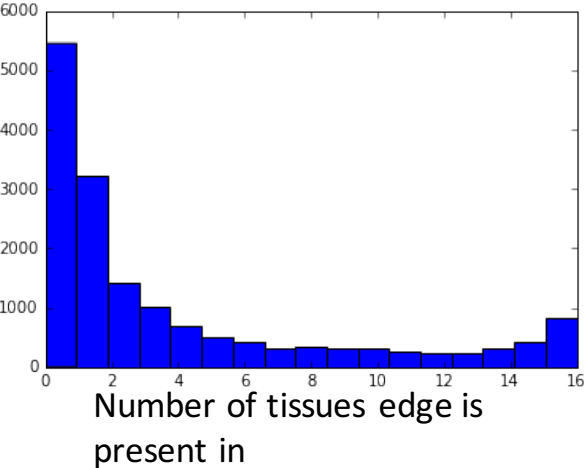
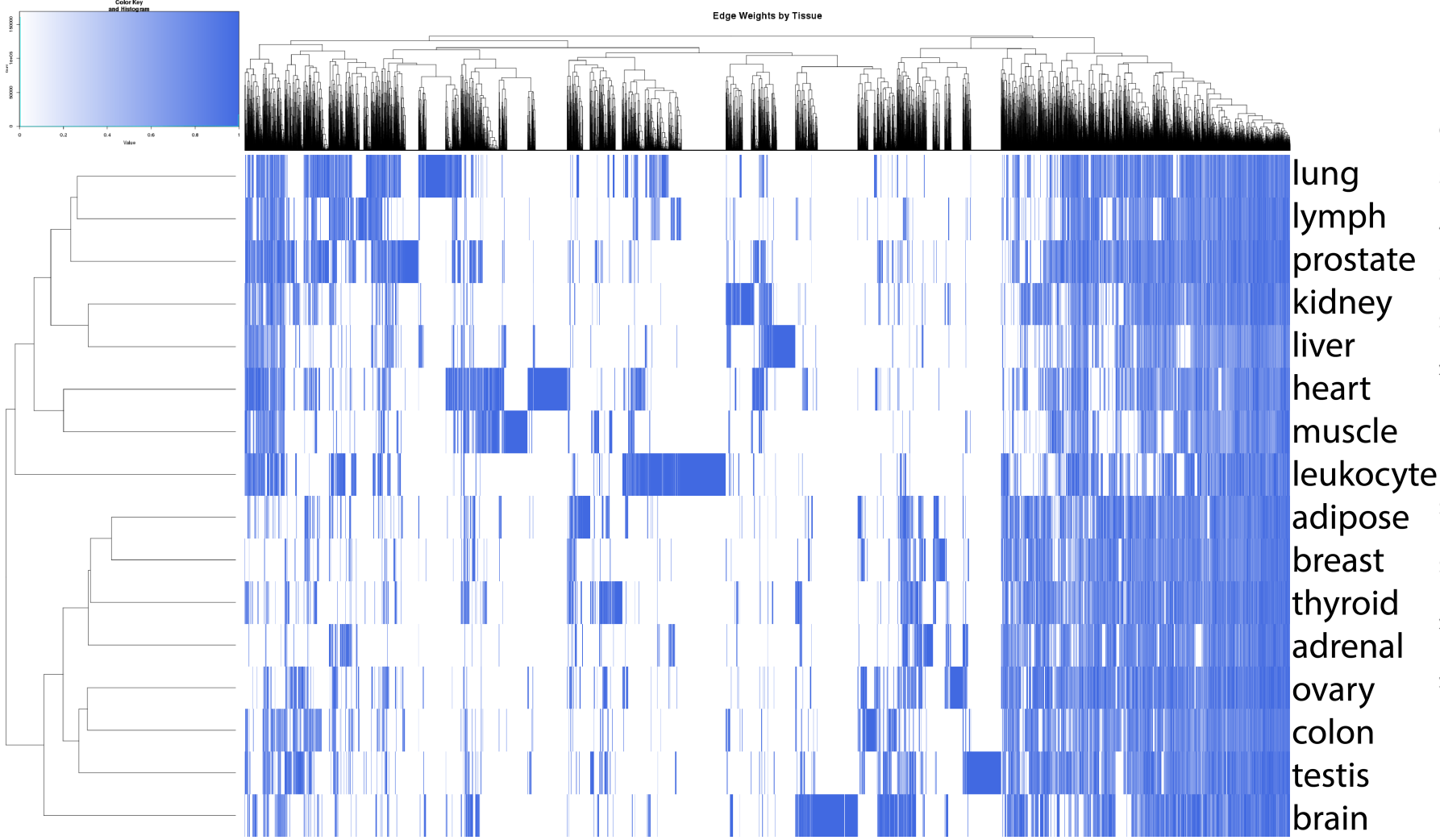
GRB2 (Growth factor receptor-bound protein 2)

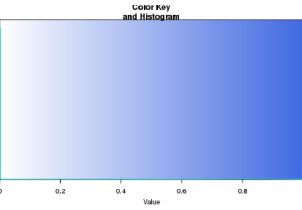
- link between cell surface growth factor receptors and the Ras signaling pathway
- Isoform 2 (green) has a 40AA deletion and prevents EGF activation of a RAS-responsive element
- Isoform 2 is also dominant negative over Isoform 1

# There exist populations of shared and tissue-specific edges

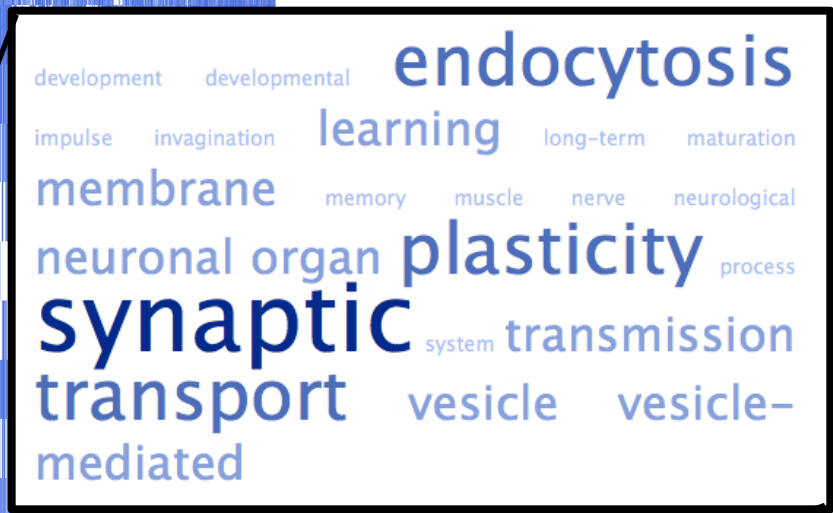
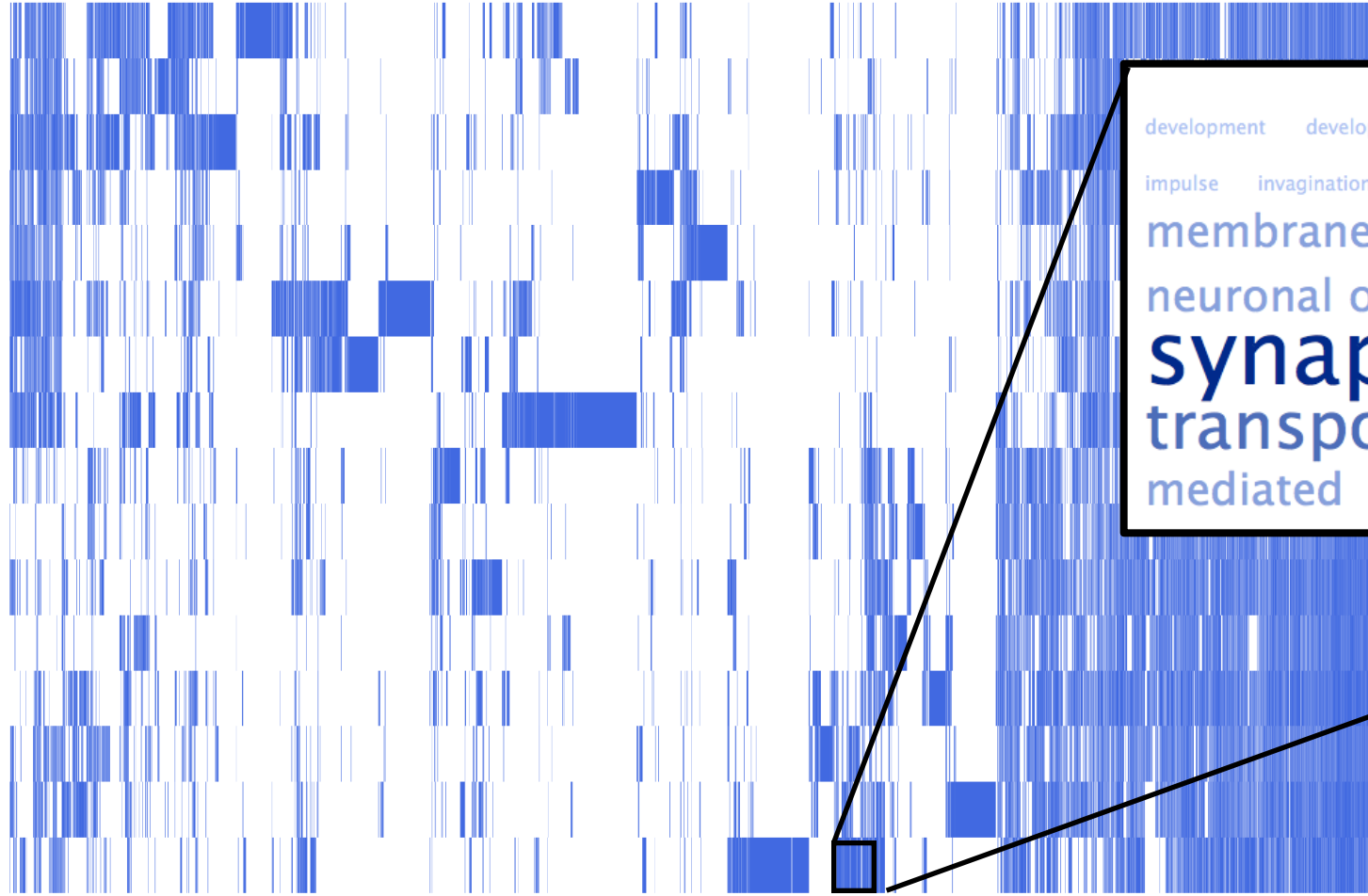
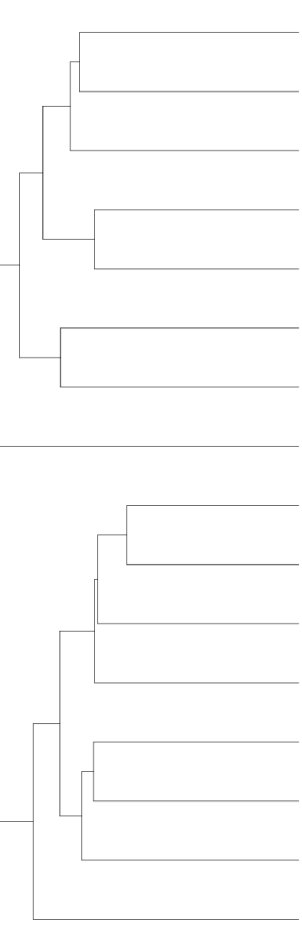
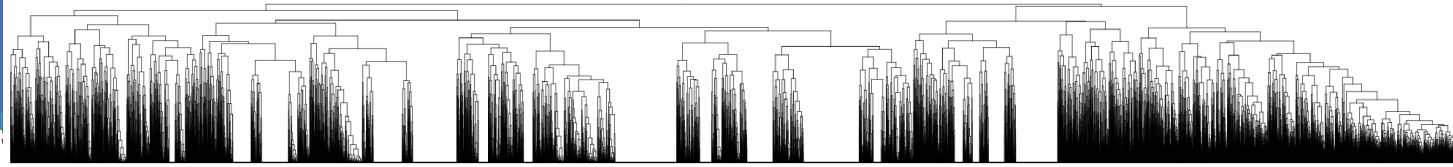


# There exist populations of shared and tissue-specific edges





Edge Weights by Tissue

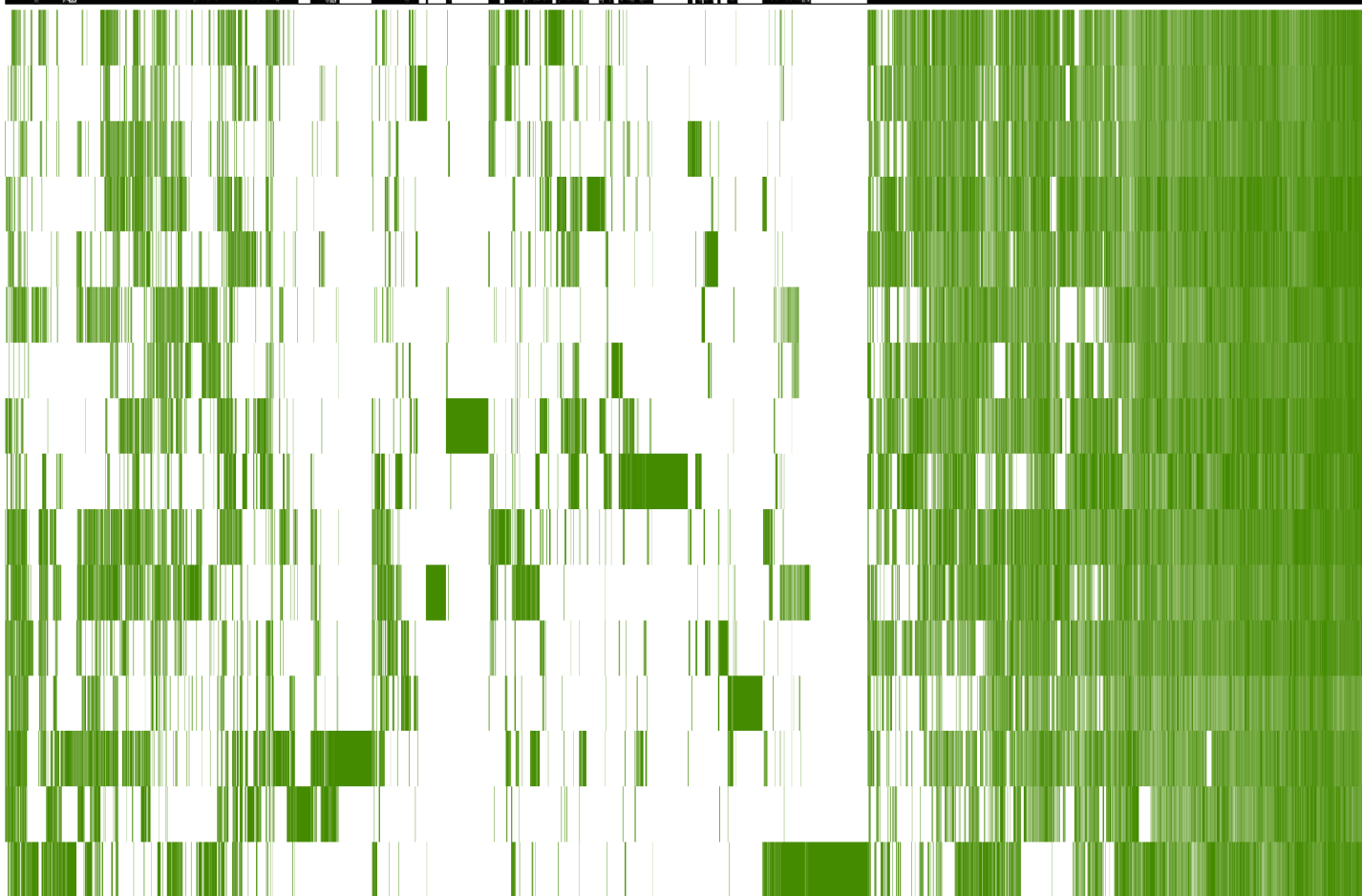
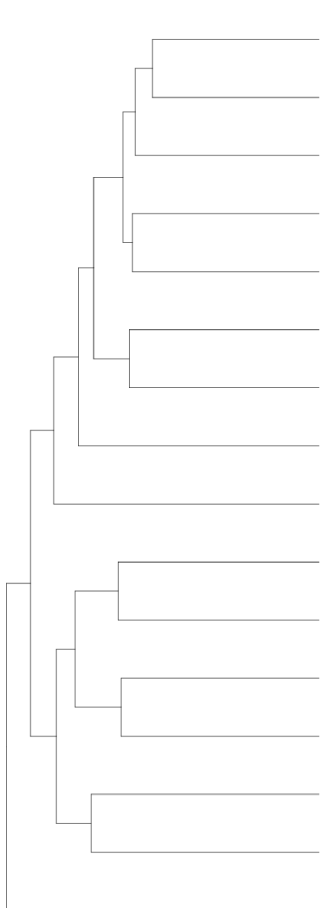
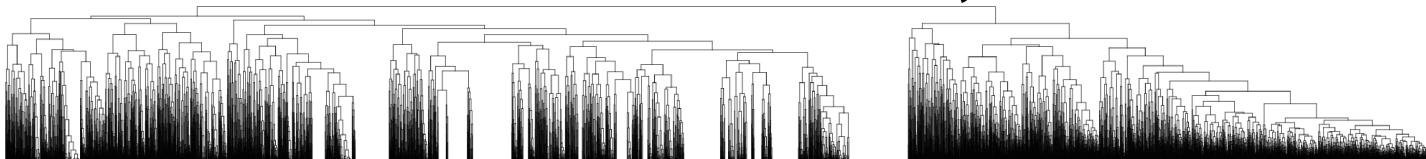
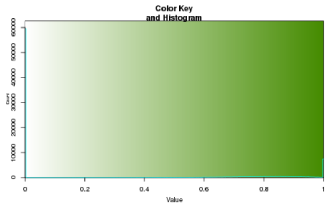


lung

breast  
 thyroid  
 adrenal  
 ovary  
 colon  
 testis  
 brain

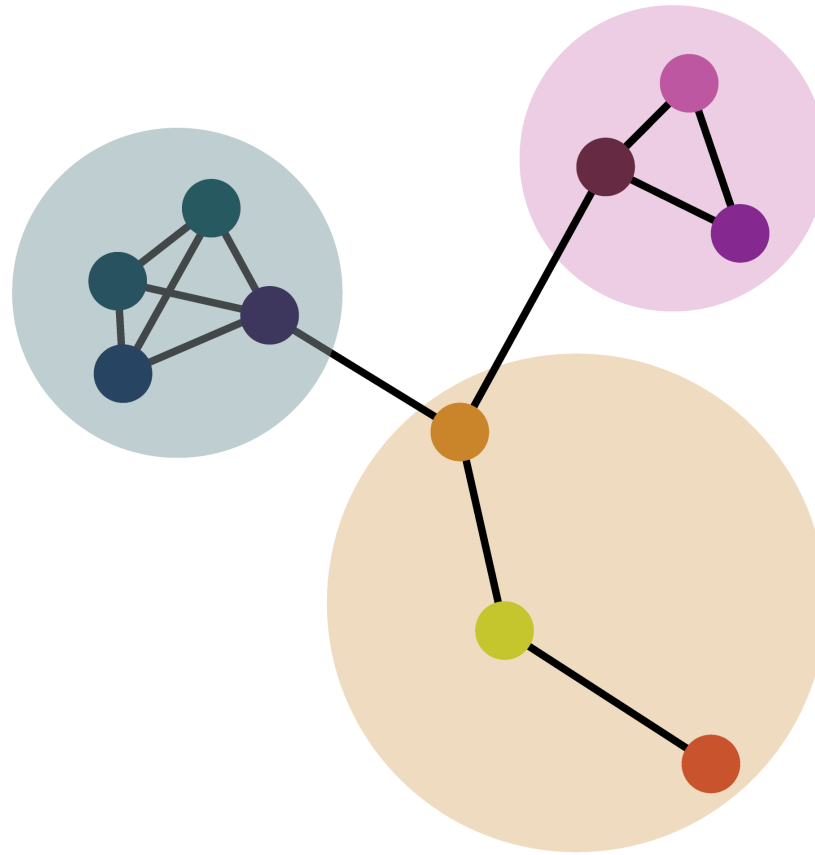
Edge weight represents  $\log(\text{TPM}_{P1} * \text{TPM}_{P2})$

# Presence of Genes by Tissue



adipose  
breast  
colon  
ovary  
thyroid  
lymph  
adrenal  
testis  
brain  
prostate  
lung  
kidney  
liver  
heart  
muscle  
leukocyte

What do we find when we look at gene / transcripts modules?





# Most shared modules

actin activation adhesion apoptosis  
assembly cascade **cell** cell-substrate cellular  
complex cytoskeleton **death** endocytosis  
epidermal extracellular factor filament-based formation **growth** gtpase  
junction kinase macromolecular macromolecule mechanical  
mediated **membrane** metabolic morphogenesis **negative**  
**organization** pathway **positive**  
postsynaptic potential process programmed proliferation  
**protein** ras **receptor**  
**regulation** signaling small  
synaptic transduction transition transport vesicle  
vesicle-mediated

acid actin activity binding biosynthetic catabolic  
cell cellular compound conjugation cytoskeleton death  
differentiation dna dna-dependent expression factor gene  
gene-specific ii localization macromolecule  
**metabolic** modification molecular mrna muscle  
**negative** nitrogen nucleic nucleobase nucleoside nucleotide  
organization polymerase **positive** process  
promoter **protein** reactions  
**regulation** response rna small  
splicing stimulus stress transcription  
transesterification via

acid activity adaptation amino apoptosis biosynthetic catabolic  
cell cellular chromatin deacetylation **death** dephosphorylation  
development differentiation dna-dependent gene gene-  
specific histone hypertrophy ii import intracellular kinase localization  
macromolecule **metabolic** microtubule-based  
modification muscle **negative** nucleus  
organization phosphate phosphorus phosphorylation polymerase  
**positive** process programmed promoter  
**protein** **regulation** response rna signal  
specific striated transcription transport

# Brain-specific modules

activity anaphase-promoting catabolic catalytic  
cell cellular complex-dependent cycle function involved  
ligase macromolecule metabolic  
mitotic modification modification-dependent  
molecular **negative** **positive**  
**process** proteasomal **protein**  
proteolysis **regulation** ubiquitin-  
dependent ubiquitin-protein  
ubiquitination

assembly biogenesis cell cellular  
**complex** cycle cytoskeleton differentiation  
division fission **macromolecular** microtubule  
microtubule-based mitosis mitotic movement neuron nuclear organelle  
**organization** phase polymerization process  
**protein** spindle subunit

# Heart-specific modules

activity binding cascade catabolic  
cellular chromatin conjugation covalent dna factor function  
histone i-kappab involved kinase ligase  
macromolecule metabolic **modification**  
modification-dependent molecular nf-kappab polyubiquitination  
**positive** postreplication **process** **protein**  
proteolysis **regulation** removal repair response  
small stress transcription ubiquitin-dependent ubiquitin-protein  
ubiquitination

adhesion blood cell-substrate collagen **development**  
ectoderm epidermis extracellular fibril matrix  
**morphogenesis** odontogenesis  
**organization** skin structure tissue vasculature vessel

# Future Directions

- Further investigate the roles of tissue-specific and degree-modulated hubs
- Examine functions of enriched tissue-specific edges in the transcript and gene networks
- Use Orthoclust to find perform PPI community analysis on tissues simultaneously