Community Detection & String DB neighbor HPO association.

This is based off analysis done in:
check-for-community-hpos-in-neighbor-protein-hpos.py
Results are found in:
CommunityDetection/community-hpos-in-neighbor-protein-hpos.csv

The purpose of this analysis is to check if HPO terms found with community detection on the MyGene2 dataset are already known to be associated with the genes found in their respective communities or known to be associated with genes/proteins that interact with the gene in their community.
Gene/protein interaction are defined based on String DB interaction network.
I am using version 11 of String DB.

Community 6
4/4 MyGene2 HPO terms are known to be associated with gene/proteins interacting with genes in the community.

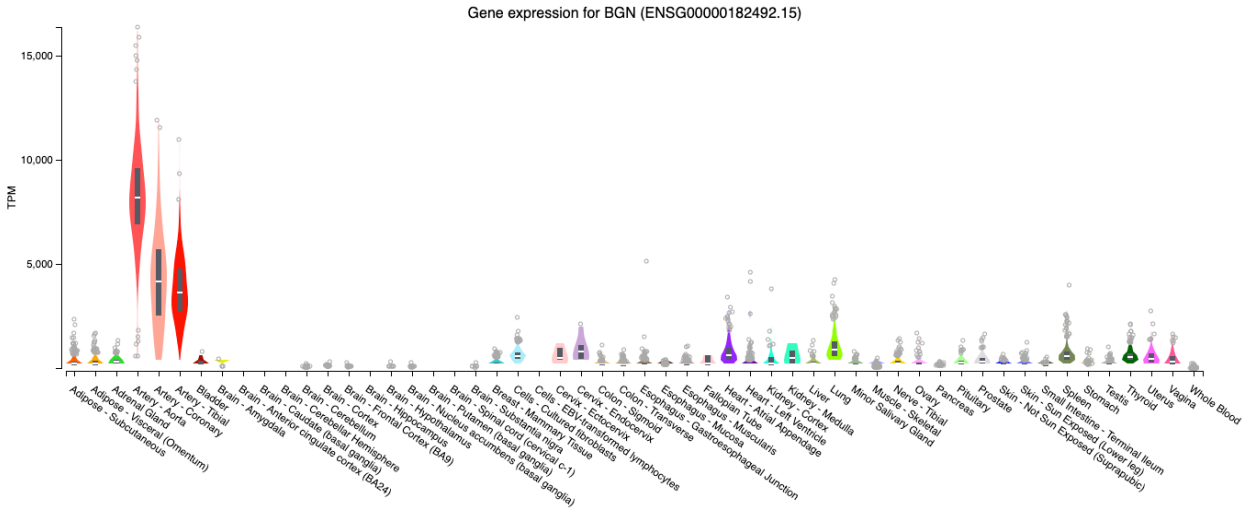
Community 8
4/5 HPO term have known associations with related genes/proteins. The exception is HP:0000456 (Bifid nasal tip). But, there is only one case one case of Bifid nasal tip in MyGene2.

Diagnosis
New phenotype

BGN
Gene Card
Small leucine-rich proteoglycan (SLRP) family of proteins. The encoded preproprotein is proteolytically processed to generate the mature protein, which plays a role in bone growth, muscle development and regeneration, and collagen fibril assembly in multiple tissues.

Gnomad
o/e = 0.2 (0.08 - 0.64)

GTEx



Community 9

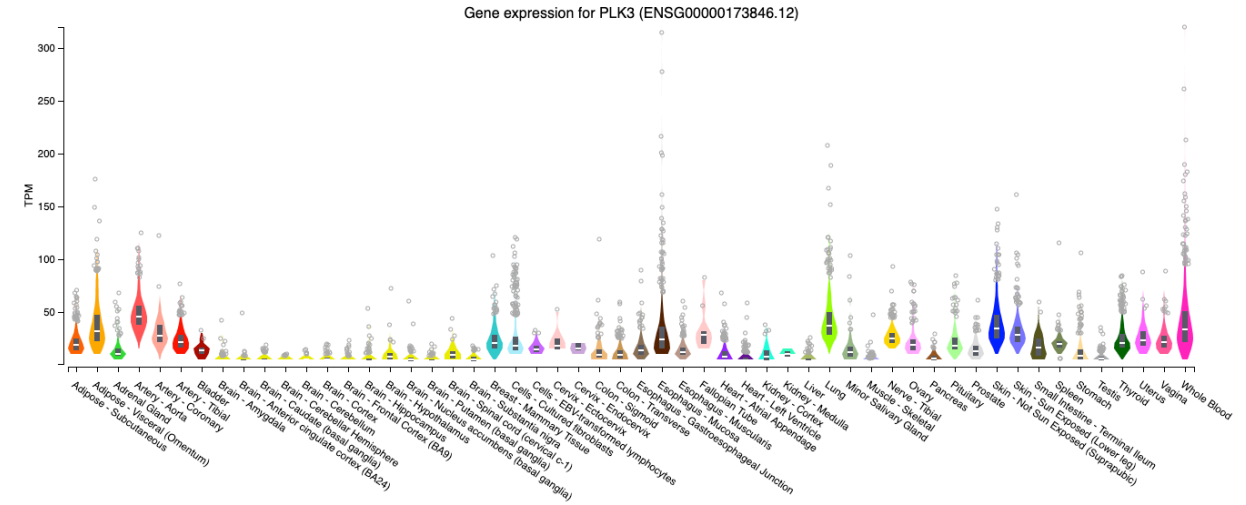
2/3 HPO term have known associations with related genes/proteins. The exception is HP:0030430 (Neuroma). There are 2 cases of Neuroma in MyGene2, with no over lapping genes.

Diagnosis
Neurofibromatosis 1-Like Syndrome
Parkinson's Disease: unexplained

PLK3

The protein encoded by this gene is a member of the highly conserved polo-like kinase family of serine/threonine kinases. Members of this family are characterized by an amino-terminal kinase domain and a carboxy-terminal bipartite polo box domain that functions as a substrate-binding motif and a cellular localization signal.

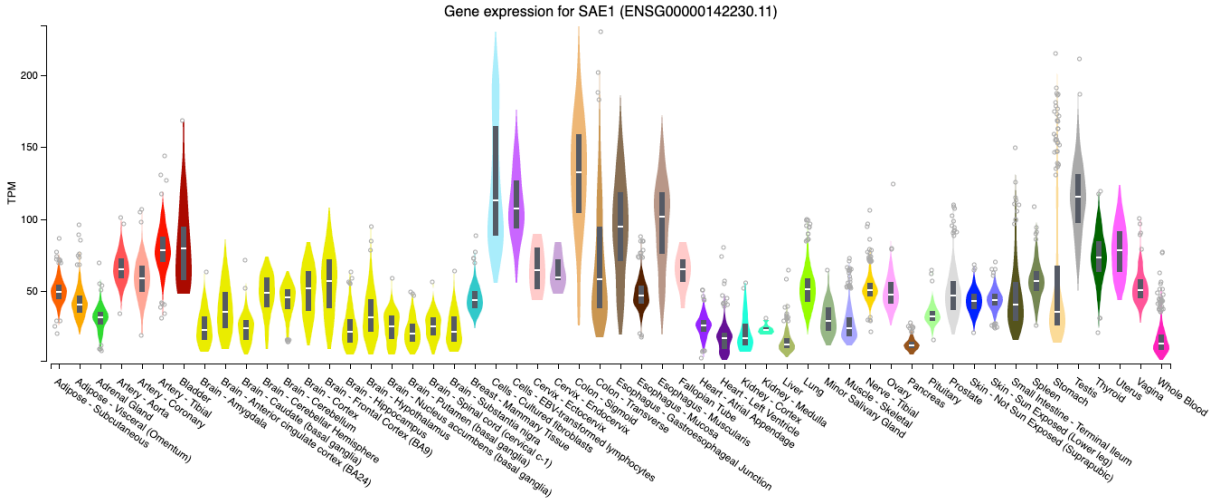
o/e = 0.46 (0.31 - 0.71)



SAE1

Posttranslational modification of proteins by the addition of the small protein SUMO (see SUMO1; MIM 601912), or sumoylation, regulates protein structure and intracellular localization.

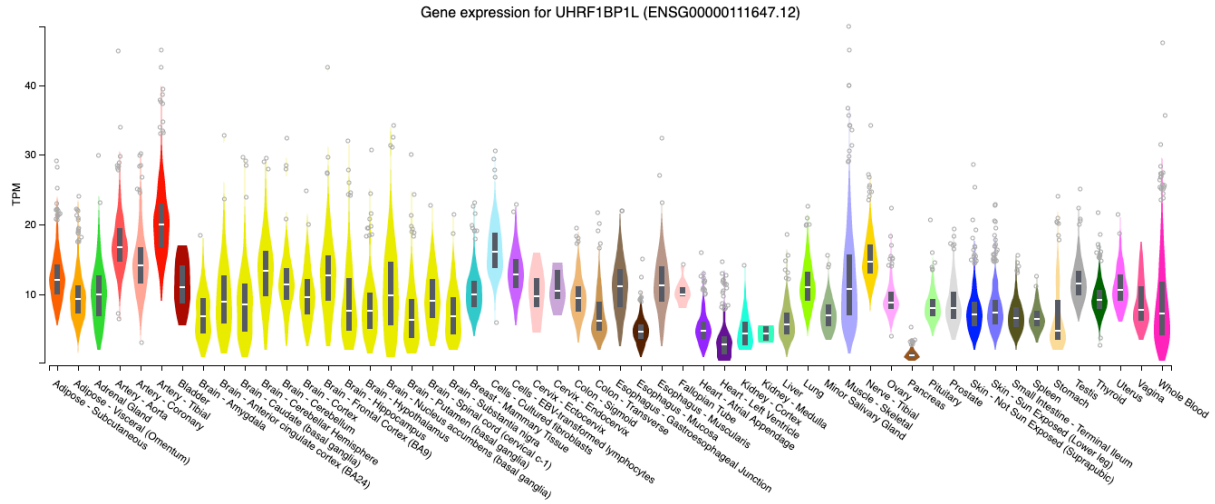
o/e = 0.06 (0.02 - 0.27)



UHRF1BP1L

UHRF1BP1L (UHRF1 Binding Protein 1 Like) is a Protein Coding gene. Diseases associated with UHRF1BP1L include Leukodystrophy, Hypomyelinating, 6. An important paralog of this gene is UHRF1BP1.

o/e = 0.27 (0.19 - 0.41)



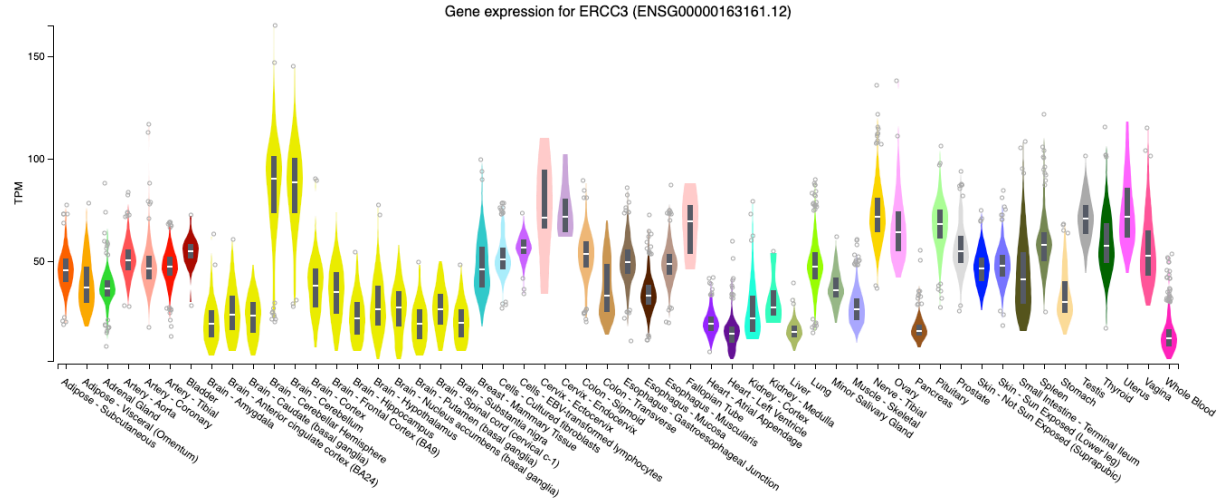
Community 10

3/4 HPO terms have known associations with related genes/proteins. The exception is HP:0030127 (Endometriosis, *growth of endometrial tissue outside the uterus*). There is only one case of Endometriosis in MyGene2

Diagnosis:
unknown, unexplained

ERCC3
This gene encodes an ATP-dependent DNA helicase that functions in nucleotide excision repair

o/e = 0.56 (0.4 - 0.79)



UVRAG
This gene complements the ultraviolet sensitivity of xeroderma pigmentosum group C cells and encodes a protein with a C2 domain. The protein activates the Beclin1-PI(3)KC3 complex, promoting autophagy and suppressing the proliferation and tumorigenicity of human colon cancer cells.

pLI = 0.99
o/e = 0.14 (0.07 - 0.29)