The Effects of Huntington's Disease on Learning

Shota Weaver, Winchester High School, Winchester, MA, 01890 & BioScience Project, Wakefield, MA, 01880

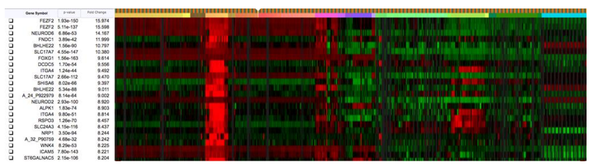
Introduction:

- Huntington's Disease is a disorder which causes the death of nerve cells in the brain. This disease is inherited and is known to cause a range of symptoms including behavioral and movement disorders. Ultimately as the disease progresses, it will cause a patient to lose their ability to move and/or speak (http://hdsa.org/what-is-hd/)
- Because this project is on the effects of the Huntington's Disease on learning, the hippocampus is the region of the brain investigated.

Method:

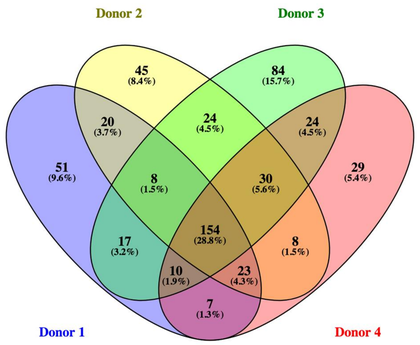
- The source for the gene expression data is The Brain Atlas (http://www.brain-map.org/). Using the differential search in the database allowed for a direct comparison of gene expression between the target structure (in this case the hippocampus) and the contrast structure (gray matter). The data used for this project was collected from four of the six donor brains:H0351.2001, H0351.1009, H0351.1015, H0351.1012. Individual donors will be referred to as Donors 1-4 respectively. The heat map, under Gene Expression Profiling, was obtained from six donors provided at the database.
- Venny 2.1 (http://bioinfogp.cnb.csic.es/tools/venny/) was used to find the common genes within the four donors and these results were later used for cluster analysis.
- PhythonAnywhere (https://www.pythonanywhere.com) is a cloud-based IDE. This was used to run a code in Python to calculate various statistics of the fold change values.
- Cluster analysis was done using DAVID (https://david.ncifcrf.gov/).
- The String Database (https://string-db.org/) was used to determine any genes related to the candidate genes obtained from DAVID.
- Gene Weaver (www.geneweaver.org) was used to determine if any of the genes from the STRING datase has any association with learning.

Part 1: Gene Expression Profiling



The heat map above shows the gene expression patterns for six donor brains. This is useful because it gives an idea of the areas of the hippocampus that are the same across six donor brains and the areas which differ. The areas in black indicate that compared to the control (the gray matter), the certain gene (shown in the y-axis) has an average gene expression in that particular part of the hippocampus (the x-axis). Red indicates that the gene is over-expressed and Green indicates that the gene is under-expressed in that particular part of the hippocampus.

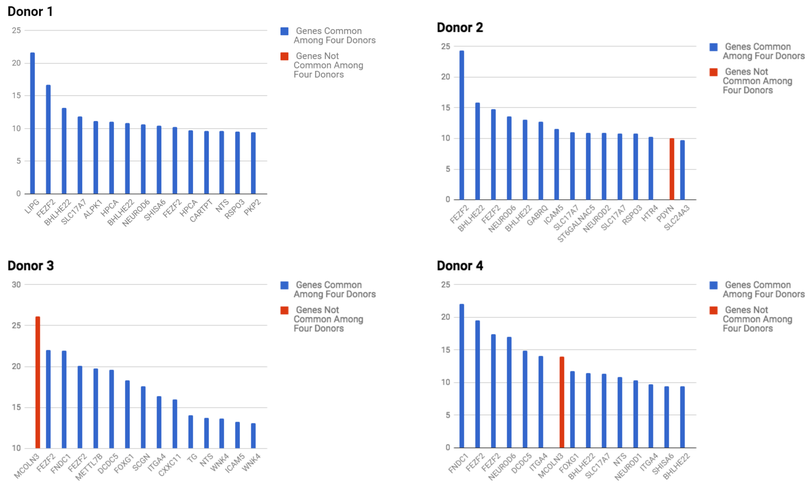
Part 2: Common Genes



The resulting Venn Diagram after compiling the gene names indicate that 28.8% of the genes were common among the four donors. This represents154 common genes which ill be later used for cluster analysis.

Part 3: Top 15 Genes with the Highest Expression

The charts below show the top 15 genes and their fold-change value for the four donor brains. Fold Change in this case refers to the difference in gene expression in the hippocampus relative to the gray matter. Let's take Donor 1 as an example: we see that the gene expression for LIPG is approximately 22 times greater in the hippocampus than in the gray matter. The blue bars in the charts are the genes which are common in all four of the donors and the bars in red are the genes unique to the individual donor brain.

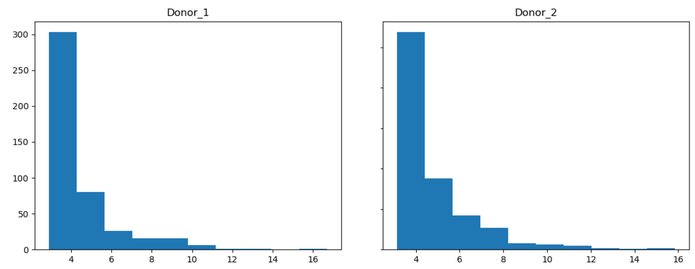


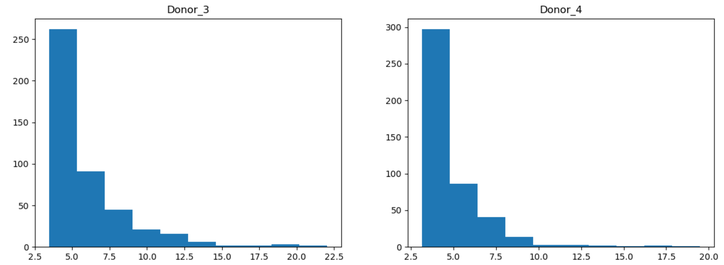
The common genes which were consistently in the top 15 for at least three of the donors were: FEZF2: transcription regulation, BHLHE22:brain development, SLC17A7:neurotransmitter transport/development, NEUROD6:trancription neuronal differentiation/development, and NTS:neuronal signaling. Of these, only FEZF2 was in the top 15 genes with the highest fold change values for all four donor brains.

Part 4: Statistical Analysis:

```
stats
          Donor_1     Donor_2     Donor_3     Donor_4
count  450.000000  450.000000  450.000000  450.000000
mean     4.350840    4.744116    5.893109    4.797896
std      1.842343    1.914717    2.964043    2.088734
min      2.881000    3.124000    3.441000    3.129000
25%      3.178000    3.454000    3.954250    3.472000
50%      3.742500    4.025500    4.751500    4.069500
75%      4.684250    5.348250    6.698000    5.348750
max     16.717000   15.858000   22.051000   19.497000
variance
Donor_1    3.394227
Donor_2    3.666143
Donor_3    8.785553
Donor_4    4.362812
dtype: float64
>>>
```

The means of the fold change values for the genes expressed were relatively constant among all four donor brains (averaging to about 5). The maximum fold change values were somewhere around 18 and the minimum fold change values were around 5 for each of the donors. This approximates a range of about 13 for each donor. Considering that the standard deviation (labeled as std above) is around 2, it is apparent that there is not very much dispersion among the fold change values. This can lead to the conclusion that the majority of the genes expressed should have a fold change value close to the mean which is 5. This can be further illustrated in the histograms shown below (histograms of Donor 3 and 4 are shown on top right) comparing the fold change values on the x-axis with the frequency of that values in the y-axis for all four donors. Note that the histograms are NOT depicting the genes that are most abundant, but the frequency of the fold change value themselves.



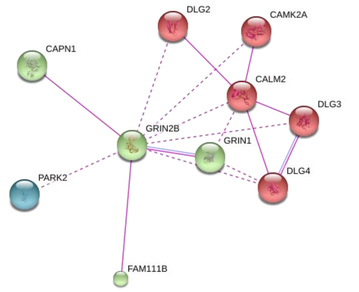


Notice how all of the histograms are right skewed and there are fewer genes with high fold change values. The majority of the genes have a much lower fold change value.

Part 5: Gene Classification

GRIN2B	glutamate ionotropic receptor NMDA type subunit 2B(GRIN2B)	Related Genes	Homo sapiens
BIOCARTA	Nitric Oxide Signaling Pathway,		
GOTERM_BP_DIRECT	MAPK cascade, transport, glutamate receptor signaling pathway, chemical synaptic transmission, learning or memory, ion transmembrane transport, ionotropic glutamate receptor signaling pathway, positive regulation of GTPase activity, response to ethanol, ephrin receptor signaling pathway,		
GOTERM_CC_DIRECT	intracellular, plasma membrane, integral component of plasma membrane, cell surface, NMDA selective glutamate receptor complex, cell junction, neuron projection, postsynaptic membrane,		
GOTERM_MF_DIRECT	NMDA glutamate receptor activity, Ras guanyl-nucleotide exchange factor activity, extracellular-glutamate-gated ion channel activity, protein binding, zinc ion binding, glycine binding,		
INTERPRO	Ionotropic glutamate receptor, NMDA receptor, Extracellular ligand-binding receptor, Glutamate [NMDA] receptor, epsilon subunit, C-terminal, Glutamate receptor, L-glutamate/glycine-binding,		
KEGG_PATHWAY	Ras signaling pathway, Rap1 signaling pathway, cAMP signaling pathway, Neuroactive ligand-receptor interaction, Circadian entrainment, Long-term potentiation, Glutamatergic synapse, Dopaminergic synapse, Alzheimer's disease, Amyotrophic lateral sclerosis (ALS), Huntington's disease, Cocaine addiction, Amphetamine addiction, Nicotine addiction, Alcoholism, Systemic lupus erythematosus,		
OMIM_DISEASE	Mental retardation, autosomal dominant 6, Epileptic encephalopathy, early infantile, 27,		
SMART	PBPe, SM00918,		
UP_KEYWORDS	3D-structure, Calcium, Cell junction, Cell membrane, Chromosomal rearrangement, Complete proteome, Disease mutation, Disulfide bond, Epilepsy, Glycoprotein, Ion channel, Ion transport, Ligand-gated ion channel, Magnesium, Membrane, Mental retardation, Metal-binding, Phosphoprotein, Polymorphism, Postsynaptic cell membrane, Receptor, Reference proteome, Signal, Synapse, Transmembrane, Transmembrane helix, Transport, Zinc,		
UP_SEQ_FEATURE	chain:Glutamate [NMDA] receptor subunit epsilon-2, compositionally biased region:Poly-His, glycosylation site:N-linked (GlcNAc...), helix, modified residue, sequence conflict, sequence variant, short sequence motif:PDZ-binding, signal peptide, site:Functional determinant of NMDA receptors, strand, topological domain:Cytoplasmic, topological domain:Extracellular, transmembrane region, turn,		

After filtering through any genes of interest in the list of genes common across the four donor brains, GRIN2B was singled out as the one gene associated with Huntington's Disease (shown above).



A network analysis of protein interactions based on experimental evidence using GRIN2B as the seed node revealed multiple other genes/proteins linked to GRIN2B (shown above). Among these, two were also notably associated with Huntington's Disease: DLG4: signaling at neuoronal synapses and GRIN1.:Glutamate receptor

DLG4	discs large MAGUK scaffold protein 4(DLG4)	Related Genes	Nomascus leucogenys
GOTERM_BP_DIRECT	protein complex assembly, protein localization to synapse, establishment of protein localization,		
GOTERM_CC_DIRECT	cortical cytoskeleton, postsynaptic membrane,		
INTERPRO	Src homology-3 domain, PDZ domain, Guanylate kinase, Guanylate kinase/L-type calcium channel, Membrane-associated guanylate kinase (MAGUK) scaffold protein, PDZ-associated domain of NMDA receptors, Membrane-associated guanylate kinase (MAGUK), PEST domain, N-terminal, Guanylate kinase, conserved site, P-loop containing nucleoside triphosphate hydrolase,		
KEGG_PATHWAY	Hippo signaling pathway, Glutamatergic synapse, Huntington's disease, Cocaine addiction,		
SMART	GuKc, PDZ, SH3, SM01272,		
UP_KEYWORDS	Complete proteome, Reference proteome, SH3 domain, Signal,		

GRIN1	glutamate ionotropic receptor NMDA type subunit 1(GRIN1)	Related Genes	Ovis aries
GOTERM_CC_DIRECT	integral component of membrane,		
INTERPRO	Extracellular ligand-binding receptor,		
KEGG_PATHWAY	Ras signaling pathway, Rap1 signaling pathway, Calcium signaling pathway, cAMP signaling pathway, Neuroactive ligand-receptor interaction, Circadian entrainment, Long-term potentiation, Glutamatergic synapse, Alzheimer's disease, Amyotrophic lateral sclerosis (ALS), Huntington's disease, Cocaine addiction, Amphetamine addiction, Nicotine addiction, Alcoholism,		
UP_KEYWORDS	Complete proteome, Membrane, Reference proteome, Transmembrane, Transmembrane helix,		

Final Results and Conclusion:

A final step was used to determine if any of the three genes associated with Huntington's Disease,: GRIN2B, DLG4, and GRIN1, have been previously associated with learning. Results from the GeneWeaver database supports that all three genes have been linked to maze learning and associative learning (below).

GRIN2B: Maze Learning, Discrimination Learning, Associative Learning, Memory
DLG4: Maze Learning, Associative Learning, Thinking
GRIN1: Discrimination Learning, Maze Learning, Associative Learning, Memory

Given that three genes in the interaction network have bee associated with learning and linked to Huntington's Disease, the other genes identified in the interaction network may be good targets for follow up study.