Umair Khan

3/18/20

BME 230A

Benedict Paten, Josh Stuart

Relationship between Huntington's Disease Pathways and Diabetes and Alzheimer's

Abstract:

Studies have linked Huntington's disease with Diabetes and Alzheimer's based on biological constituents shared between the diseases. To further investigate the differential gene expression of Huntington's disease in relation to Diabetes and Alzheimer's, a gene set enrichment analysis will be done to observe enriched, or over-represented genes and how those genes relate together with metabolic pathways to influence mechanisms in Diabetes and Alzheimer's. This would suggest Huntington's disease pathway may have a biomarker for Diabetes and other neuropathological diseases. The results suggest certain cellular components and mechanisms involved in the HD pathway have a relationship in the onset of Diabetes and contribute to Alzheimer's.

Background:

Huntington's disease (HD) is a neurodegenerative disorder caused by a mutation consisting of CAG repeats in the Huntington protein. Multiple studies have linked HD to glucose metabolism alterations and diabetes along with other neurological syndromes specifically Alzheimer's primarily through observations of the mutant Huntington in vitro with insulin and differentially expressed proteins compared to controls in human brains[1]. Other research[2] on HD patients and normal controls suggests mutant Huntington affects glucose homeostasis determinants including insulin sensitivity and secretion as HD patients compared to the control group show a decrease in insulin sensitivity and an increase in insulin resistance implying that insulin secretion defect may lead to a failure to compensate insulin resistance in HD patients evaluated through homeostasis modeling. Furthermore, research from the Biomolecular Engineering Department at University of California, Santa Cruz investigated the Huntington protein's role in subcellular traffic as Huntington is a scaffold protein that connects different elements of the cytoskeleton including microtubules and intracellular vesicles derived from the Golgi Apparatus and Endoplasmic Reticulum[3]. In addition to this, human and cell models of the Huntington protein suggest that HD may alter processes dependent on the cytoskeleton with vesicular transport from the Golgi Apparatus and ER, impairing the subcellular traffic associated Alzheimer's and with the secretion of insulin by specifically affecting HLA/MHC transport (regulation of immune system) from the ER to the cell membrane[3]. Problems in processing antigens through intracellular components including GA and ER has been associated with Type 1 Diabetes[3]. This suggests HD may contribute to pathways related to the onset of Alzheimer's and Diabetes specifically those intracellular and subcellular traffic pathways pertaining towards endocytosis and exocytosis mechanisms.

Goal:

Observe shared genes to phenotypes and functions from HD pathway that may suggest alterations in glucose or other causes resulting in the onset of Diabetes.

Method:

To better understand the pathways affected by HD as well as other details regarding gene function and cellular components involved, a gene set enrichment analysis is performed using Python's GSEAPY wrapper for Enrichr. RNA-Seq profiles of human post-mortem prefrontal cortex in 20 HD patients and 49 normal individuals[4] were used and the top 5000 differentially expressed genes ranked by p-value (based on Fischer's exact test or hypergeometric tests) are used to create enriched sets of genes. Finding DE Genes was already performed using DESeq2 from source 4. Reference gene sets included the Gene Ontology for biological, molecular, and cellular components as well as WikiPathways, and KEGG libraries. The goal is to highlight key mechanisms, cellular components, and biological processes that are associated with cellular traffic and mechanisms related to Diabetes.

Results:

Below are the results from using Python's GSEAPY wrapper on Enrichr. The open source Enrichr tool was also used to verify the results and gain a more thorough analysis on all the given pathways and molecular functions. Simple GSEAPY implementation shown below.

```
import gseapy
import pandas as pd

DEGenes = pd.read_excel('S1 File.xlsx') # DE genes in HD

geneList = list(DEGenes['symbol'])
geneList = list(map(str,geneList))
geneList = [ x.strip() for x in geneList if x!= 'nan' and '2014' not in x ]

# Top 5000 genes by significance

geneList = geneList[:5000]


# Create file for testing with Enrichr

with open('enrichrInput.txt','w') as file:
    for i in geneList:
        file.write(i+'\n')


#Run Enrichr API using gseapy

enrich = gseapy.enrichr(gene_list = geneList,
                        gene_sets = ['KEGG_2019_Human','GO_Biological_Process_2018',
                                     'GO_Cellular_Component_2018','GO_Molecular_Function_2018',
                                    'WikiPathways_2019_Human'],
                        outdir = 'dat1',organism='human')

enrich.results
```

Figure 1. Jupyter Notebook Cell that runs the main GSEA using Enrichr API.

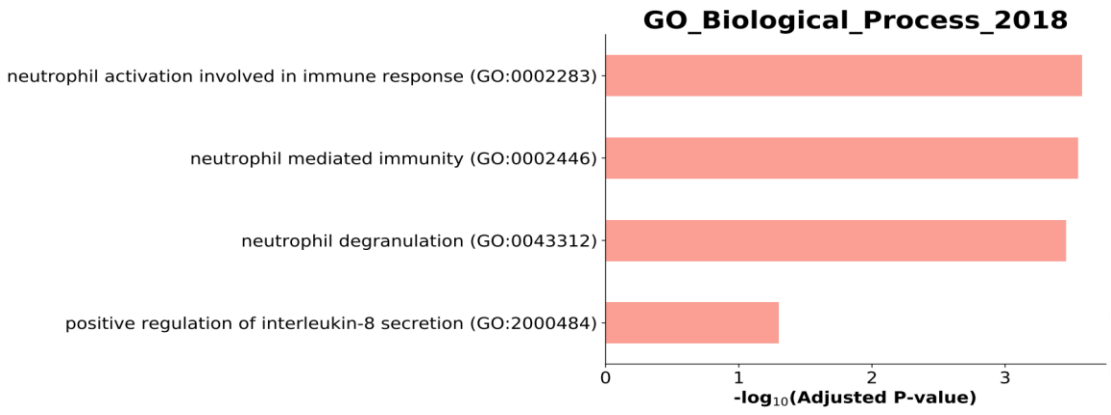

Figure 2. Gene Ontology for Biological Processes from my Python script. The top biological processes are related towards the immune system with key components consisting of neutrophil activation and positive regulation of interleukin-8 secretion. Neutrophils mediate secretion of proinflammatory substances through secreting granules by exocytosis on membrane bound secretory granules[5] and interleukin-8 secretion (signaling proteins secreted by cells) affects chemotaxis as well as being an important mediator for the innate immune system.

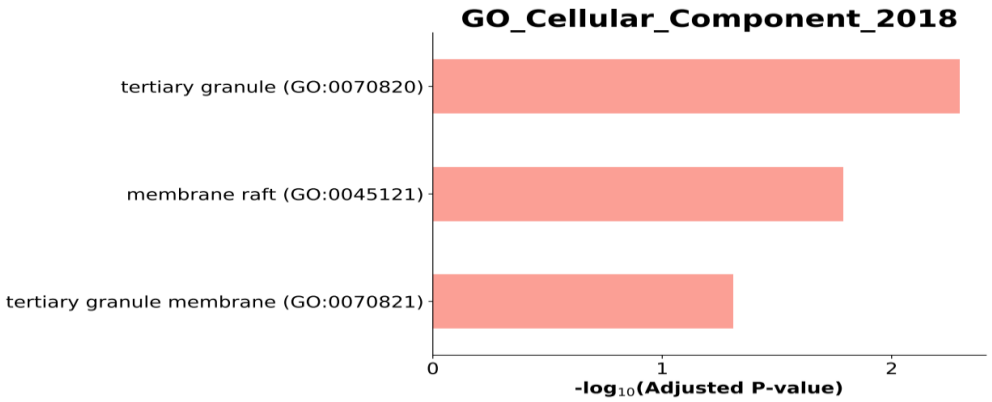

Figure 3. Gene Ontology for Cellular Components from my Python script. The top 3 cellular components involve the tertiary granule which is the extracellular matrix degrading enzymes and membrane receptors that are required for neutrophil extravasation (neutrophils responding during innate immune response) and the lipid raft (membrane raft), glycolipoprotein lipid microdomains contained in the plasma membrane that influences the assembly of signaling molecules and therefore membrane protein trafficking and thus regulating neurotransmission (Wikipedia).

GO_Molecular_Function_2018	cadherin binding involved in cell-cell adhesion (GO:0098641)	10		
GO_Molecular_Function_2018	core promoter proximal region DNA binding (GO:0001159)	11/22	0.	
GO_Molecular_Function_2018	RAGE receptor binding (GO:0050786)	6/9	0.0099167329402460	
GO_Molecular_Function_2018	oxidoreductase activity, acting on the CH-CH group of donors, NAD			
GO_Molecular_Function_2018	transcription regulatory region sequence-specific DNA binding (GO:			
GO_Molecular_Function_2018	BMP receptor binding (GO:0070700)	7/12	0.0141305283800600	
GO_Molecular_Function_2018	transcription factor activity, RNA polymerase II core promoter pro			
GO_Molecular_Function_2018	galactosyltransferase activity (GO:0008378)	6/10	0.01958430	
GO_Molecular_Function_2018	protein binding involved in cell-cell adhesion (GO:0098632)	10		
GO_Molecular_Function_2018	death receptor activity (GO:0005035)	11/24	0.0210862694567640	
GO_Molecular_Function_2018	tumor necrosis factor-activated receptor activity (GO:0005031)	11		
GO_Molecular_Function_2018	RNA polymerase II regulatory region sequence-specific DNA binding			
GO_Molecular_Function_2018	RNA polymerase II regulatory region DNA binding (GO:0001012)	63		
GO_Molecular_Function_2018	insulin-like growth factor I binding (GO:0031994)	7/13	0.	
GO_Molecular_Function_2018	phosphatidylcholine transporter activity (GO:0008525)	5/8	0.	
GO_Molecular_Function_2018	E-box binding (GO:0070888)	14/34	0.027723192729544183	1.
GO_Molecular_Function_2018	Wnt-activated receptor activity (GO:0042813)	10/22	0.02920198	
GO_Molecular_Function_2018	core promoter proximal region sequence-specific DNA binding (GO:00			
GO_Molecular_Function_2018	transforming growth factor beta-activated receptor activity (GO:00			
GO_Molecular_Function_2018	CD4 receptor binding (GO:0042609)	4/6	0.0374163947913132	
GO_Molecular_Function_2018	dopamine transmembrane transporter activity (GO:0005329)	4/		
GO_Molecular_Function_2018	chromo shadow domain binding (GO:0070087)	4/6	0.03741639	
GO_Molecular_Function_2018	insulin-like growth factor binding (GO:0005520)	7/14	0.03798242	
GO_Molecular_Function_2018	aldehyde dehydrogenase (NAD) activity (GO:0004029)	7/14	0.	
GO_Molecular_Function_2018	ATPase activity, coupled to transmembrane movement of substances (			
GO_Molecular_Function_2018	transcriptional activator activity, RNA polymerase II transcriptio			
GO_Molecular_Function_2018	amyloid-beta binding (GO:0001540)	18/49	0.0449362834558185	
GO_Molecular_Function_2018	peptidase inhibitor activity (GO:0030414)	14/36	0.04556790	

Figure 4. Gene Ontology Molecular Function from my Python script.

This contains a partial list molecular functions from the GSEA. Most of the molecular functions relate to transmembrane movement and cytokine signaling. The most significant molecular functions that affect Diabetes include insulin-like growth factor 1 binding, insulin-like growth factor binding, Wnt-activated receptor activity, transforming growth factor beta-activated receptor activity and T-cell signaling pathways all found in both the WikiPathways and GO libraries. Increased levels of Amyloid-beta binding and BMP contribute to the neurogenesis in Alzheimer's. Cadherin binding involved in cell-cell adhesion are also significant in cell plasticity, vascular changes, and neuroinflammation.

Analysis:

Gene Ontology's results on the biological process and cellular components suggest that HD may have a pathway associated with neutrophils and interluekins, where interluekin-8 serves as a chemoattractant for neutrophils and neutrophils regulate T cell function through dendritic cells. Given this along with the function of each cellular component that the Gene Ontology reported as well as the biological processes, HD may affect or perhaps even inihibit the exocytosis of tertiary granules on the plasma membrane level as well as the role of membrane rafts influencing membrane trafficking and thus nuerotransmission, supporting a potential relationship of HD to affect the intracellular immune response near the cell membrane affecting how T-cells are being activated nearby. The prevalence of Interluekin-8 may induce a worse inflammatory response as type 2 diabetes patients exhibit higher levels of Il-8 compared to normal patients and thus may be an essential marker in diabetes.

Furthermore, the most important molecular functions to consider that relate HD genes to Diabetes include Wnt-activated receptor activity. Wnt signaling pathways involve lipid metabolism, glucose homeostasis, insulin secretion and influencing T cell transcription factor. A mutation in LRP5 in this pathway is known to lead to Diabetes[6]. Thus, HD may have a role in this particular function, suggesting it may lead to defective Wnt signaling pathways that can lead to Diabetes. In addition, insulin-like growth factor binding along with the activity of coupled movement of substances across membranes (ATPs,cations,anions,Zinc,Ca) suggests HD has an influence on insulin along with transmembrane movement that may involve insulin. Insulin could also be stuck within the extracellular matrix of the tertiary granule preventing it from being secreted. The transforming growth factor (TGF-B1) cytokine and the terniary granule (extracellular matrix) from GO shows that HD involves a cross link with these two[7]. The corresponding results from GO along with additional research provide a strong relationship of a cross link (one polymer to another) of TGF-B1 and the terniary granule that appears significant in HD pathway as this cross link is associated with chronic renal failure which is increased in Diabetic patients[7] as well as the functional change in TGF-B1increasing diabetic nephropathy. Amyloid-beta binding is essential in forming extracellular plaques deposits of the B-amyloid peptide seen in Alzheimer's[8]. This contributes the neurotic tangles seen in Alzheimer's patient's brain. Cadherine cell-cell adhesion also contributes to processes in synaptic neurotransmission and synaptic vesicle recycling. Given the biological components highlight the extracellular matrix, these processes may be inhibited in Alzeihmer's thus contributing to a lack of synapses seen in the disease[9]. Another key process is BMP signaling which have been suggested to be regulators in neurogenesis. These growth factors contribute to neuronal cell development and research[10] shows the BMP6 is elevated in the hippocampus and accumulate around plaques. The biological components further support irregularities implicated in Alzeihmer's especially creating new synapses. This along with the three molecular functions analyzed suggest that HD pathway is enriched for pathways that contribute to the onset of Alzheihmer's.

Conclusion and Future Work:

Although the GO results shed light on key cellular components processes involved, more analysis of these described mechanisms should be considered, focusing on developing a network based GSEA. There are many tools that can do this online, and it may help understand the events taking place in biological processes to discover potential metabolic cascade pathways. Further grouping up biological processes and molecular functions by different gene set libraries may give a closer approximation of the most significant pathways. Linkage analysis would concretely justify if HD pathway is associated with Diabetes and Alzeihmer's as it would observe the inheritance of specific genes in respect to the chrosome location. By observing a genetic linkage mapping of differentially expressed genes of many patients and known biomarkers of each disease, we may be able to find relative frequencies that imply the distance of biomarkers. Thus, this would allow us to observe crossovers and and gene mutations that may contribute to pathways in one disease to the other. Although this analysis focuses on past research along with the gene set enrichment analysis, it still provides a valid reason to imply HD pathway contributing to other diseases as the common processes and molecular function are extremely overlapping in all three diseases being observed.

References

1. M. T. Montojo, M. Aganzo, and N. González, “Huntington's Disease and Diabetes: Chronological Sequence of its Association,” *Journal of Huntington's disease*, 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5676851/. [Accessed: 18-Mar-2020].

2. N. M. Lalić, “Glucose Homeostasis in Huntington Disease,” *Archives of Neurology*, 01-Apr-2008. [Online]. Available: https://jamanetwork.com/journals/jamaneurology/fullarticle/795466. [Accessed: 18-Mar-2020].

3. Chaves, G.; Stanley, J.; Pourmand, N. Mutant Huntingtin Affects Diabetes and Alzheimer’s Markers in Human and Cell Models of Huntington’s Disease. *Cells* 2019, *8*, 962.

4. A. Labadorf, A. G. Hoss, V. Lagomarsino, J. C. Latourelle, T. C. Hadzi, J. Bregu, M. E. MacDonald, J. F. Gusella, J.-F. Chen, S. Akbarian, Z. Weng, and R. H. Myers, “RNA Sequence Analysis of Human Huntington Disease Brain Reveals an Extensive Increase in Inflammatory and Developmental Gene Expression,” *PloS one*, 04-Dec-2015. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4670106/. [Accessed: 18-Mar-2020].

5. P. Lacy, “Mechanisms of degranulation in neutrophils,” *Allergy, asthma, and clinical immunology : official journal of the Canadian Society of Allergy and Clinical Immunology*, 15-Sep-2006. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2876182/. [Accessed: 18-Mar-2020].

6. T. Jin, “The WNT signalling pathway and diabetes mellitus,” *Diabetologia*, Oct-2008. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/18696049. [Accessed: 18-Mar-2020].

7. S. Chen and F. N. Ziyadeh, “Transforming Growth Factor-β Signal Transduction in the Pathogenesis of Diabetic Nephropathy,” *Contemporary Diabetes The Diabetic Kidney*, pp. 201–221.

8. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2813509/

9. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4868906/

10. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2978735/