Universiteit Leiden
The Netherlands

Opleiding Informatica

Identifying Overlapping Processes of Alzheimer's and Huntington's Disease with a Protein-Protein Interaction Network Analysis

Aster Marthe de Boer

Supervisors:
Katy Wolstencroft & Lu Cao

BACHELOR THESIS

Leiden Institute of Advanced Computer Science (LIACS)
www.liacs.leidenuniv.nl

02/07/2023

Abstract

Alzheimer's disease (AD) and Huntington's disease (HD) are two neurodegenerative diseases, with similar characteristics such as aggregation of proteins and an inflammatory response in the brain. Protein-protein interaction networks (PPINs) can be used to find enriched biological processes within a group of interconnected proteins. This research looked at how two PPINs for AD and HD overlapped. The KEGG, Wikipathways and STRING database provided the data for these PPINs. The AD and HD PPINs were intersected and about two-thirds of the networks overlapped. Biological processes such as chemical synaptic transmission, autophagy, apoptosis and nervous system development were found to be enriched in the intersected network. In addition, a literature analysis revealed that these processes were also involved in AD and HD. Thus, this research provides insight into how PPIN analyses can contribute to finding overlapping processes of similar diseases. This information could also be valuable for studying shared drug targets.

Contents

1 Introduction

Alzheimer's disease (AD) and Huntington's disease (HD) are both neurodegenerative diseases. What makes these diseases similar is that misfolded protein aggregates are formed which cause a toxic reaction that leads to cell death of neurons. HD belongs to the category of Poly-Q diseases with the huntingtin protein as the main causal factor. AD belongs to a class of neurodegenerative diseases called tauopathies with the amyloid-β precursor protein (APP) as one of the main causal factors [TMP03]. Even though the genetic basis of AD and HD is different, they share molecular processes such as neuroinflammation, increased reactive oxygen species (ROS), and disrupted protein metabolisms [DKKK18]. It has been proposed that getting rid of the misfolded proteins by inducing autophagy in neurons might be a treatment option for both diseases [TMA+10].

1.1 Alzheimer's disease

Alzheimer's disease is a deadly neurodegenerative disease with dementia being the main clinical symptom. It is seen a lot in people who are older than 70 years, but there have also been genes identified that cause the disease at a younger age. The disease can nowadays be detected before any signs of memory loss start to show because three important biomarkers have been identified. These are amyloid-β ($A\beta$), phosphorylated tau, and neurodegeneration [SDSK+21].

Firstly, Aβ can form plaques in the brain when it is not broken down efficiently. These plaques can be found in AD patients. The protein it is derived from is called amyloid-β precursor protein (APP). This is a protein that is produced a lot by neurons, especially in the brain. When APP is cut at two different sites, you get three smaller proteins, one of which is Aβ [HHB+21]. Tau phosphorylation is probably also caused by Aβ plaques, but the pathway by which this happens is still unknown [SDSK+21].

In terms of genetics, it has been found that autosomal dominant variants of the APP gene, but also presenilin 1 (PSEN1) and presenilin 2 (PSEN2), can cause AD at a younger age [HHB+21]. One important gene that accounts for Alzheimer's in older people is the APOE e4 allele [SDSK+21]. Other risk genes have also been identified that become upregulated in the microglia cells as a response to the Aβ plaques [SLM+20]. The microglia, astroglia, and neurons are already affected before the disease is clinically visible [SDSK+21]. These cells, especially the microglia, also seem to be involved in a neuroinflammatory response [SLM+20].

1.2 Huntington's disease

Huntington's disease (HD) is characterized by brief involuntary movements, psychiatric conditions such as dementia, and usually also a family history of the disease. Patients commonly start having symptoms around 40 years old and live for another 15 to 20 years. The causal mutant gene for this autosomal dominant neurodegenerative disorder was discovered in 1993 and is called huntingtin (HTT). Because of this mutation, there is an additional tail of CAG repeats, which code for the protein glutamine, attached to the HTT protein. The length of this polyglutamine tail influences the severity of the disease. Sometimes children already display symptoms, when they have an exceptionally long sequence of glutamine repeats [VD98][Roo10][TGL19].

Both excessive movements and a decrease in movement are observed in patients with HD. Usually, people's cognitive abilities decrease over time, and patients display forms of memory loss. Especially memory of movements is affected [Roo10].

The exact function of the normal HTT protein is still unknown, but there is probably a relation between the protein's function and the disease symptoms [SH16][STW+98]. Current treatments for the disease focus on trying to lower the amount of HTT. This can be done in the DNA itself, or by targeting the transcription or translation products. For example, HTT mRNA can be targeted or the DNA sequence can be modified with CRISPR/Cas9. It has also been found that a lower amount of DNA repair genes such as MSH3 can decrease the rate of expansion of the CAG repeats in brain cells [TGL19].

1.3 Inflammation in neurodegenerative diseases

Similar processes seem to play a role in both AD and HD. For example, the HTT protein has been found to be overexpressed at aggregation sites in the brain of AD patients, although its specific role is not clear [STW+98].

Next to the formation of aggregates, there has been a lot of evidence that inflammation is what causes the eventual death of neurons in neurodegenerative diseases. These inflammation reactions start with the assembly of protein complexes called inflammasomes, which are formed as part of the innate immune response. They are assembled by pattern recognition receptors (PRRs), which can detect foreign proteins in the body. These inflammasomes can then activate, among others, the proteins caspase-1, IL-1beta, and IL-18 which promote inflammation in the Central Nervous System [VSLvL19]. When glutamate receptors are then activated, a cascade of reactions ensues which leads to excitotoxic cell death [Mat03].

Evidence for neuroinflammation was discovered for both diseases. Caspase-1, IL-1beta and IL-18 proteins have all been found to be overexpressed in AD patients, and the same goes for other proteins involved in the inflammation reaction caused by inflammasomes. Furthermore, the inflammasome NLRP3, which has been studied a lot, is activated by Aβ aggregates. In HD patients, the caspase-1 protein was found to be overexpressed, and inhibition of this protein delayed the disease progression [VSLvL19]. Moreover, excitotoxic cell death occurs in both AD and HD patients [Mat03].

1.4 Protein-protein interaction networks

Over the last few decades, it has become increasingly important to study not just individual proteins, but the whole system of protein-protein interactions within organisms [Ram10][BO04][MV07].

Protein-protein interaction networks (PPINs) are graph representations of the interactions between proteins in an organism. The nodes of the graph represent the proteins, and the edges represent the interaction types. These graphs can either be directed or undirected, depending on the type of interactions displayed. For example, when protein A activates protein B, there could be a directed edge from A to B. On the other hand, if A binds to B then B binds to A and there is an edge in two directions [BO04].

New interactions can be discovered in a number of ways. Firstly, with high-throughput experiments using for example micro-arrays or yeast two hybrid assays. Two disadvantages of this method are the amount of false positive results and difficulties in reproducing such experiments. Secondly, information from literature and databases that has accumulated over the years can be analyzed, with text mining methods for example. The last way of discovering new interactions is by computational methods, for example by using the structures of proteins to predict interactions [Ram10].

Several characteristics are used for analyzing PPINs such as node degree distribution, clustering coefficient, and characteristic path length [Ram10]. These will be described in more detail in the methodology section. Based on these characteristics, three different network topologies are in general distinguished: random networks, small-world networks, and scale-free networks [Ram10]. It has been discovered that most biological networks have a scale-free topology, which is characterized by hub nodes and clusters. Hub nodes have a lot more edges than the average amount per node and take a central position in the network [BO04].

Highly connected nodes are proteins that have a lot of interactions with each other [SSO+12]. When analyzing PPINs, the main goal is to find clusters of proteins that share characteristics. For example, a group of proteins can be expressed simultaneously, or at the same location in the cell, or they can have a similar function [KPA+12][SSO+12]. This means that they might be involved in the same biological process. Functional enrichment analysis of a cluster can be performed to find out which process this could be [SSO+12].

1.5 Research question

In previous research, a Huntington's disease PPIN consensus network was created using different databases, combining pathway and interaction information [Jia22][Hen23]. Using the workflow for creating this network as a template, a similar PPIN can be created for Alzheimer's disease. The two networks can be compared with each other and also merged, to look at overlapping proteins and processes. The research question of this thesis is therefore:

> **RQ**: How do the Alzheimer's disease consensus network and the Huntington's disease consensus network overlap?

1.6 Thesis overview

This section includes background information and the objective of this thesis. In Section 2, definitions can be found. Section 3 discusses related work about HD and AD PPINs. Section 4 describes the methodology of building and comparing the networks. Section 5 gives the results. The discussion can be found in Section 6 and the conclusion in Section 7. Figures are shown in Appendix A. Cytoscape files, code, collected data and additional figures are available at `https://git.liacs.nl/s2955199/alzheimerconsensus`. This thesis was written under the supervision of Katy Wolstencroft at the LIACS.

2 Definitions

The following abbreviations were used for the different networks and clusters.

- ADCN means Alzheimer's disease consensus network. It is a network consisting of protein interaction data from the KEGG, Wikipathways, and STRING databases. There are four versions, two where the STRING cutoff score was 0.4 and two where the STRING cutoff score was 0.7. Furthermore, two of them are expanded with 2000 proteins from the STRING database. The corresponding names are `ADCN-KWS-0.4`, `ADCN-KWS-0.7`, `ADCN-KWS-0.4-EXP`, `ADCN-KWS-0.7-EXP`.

- HDCN means Huntington's disease consensus network. `HDCN-NH` is the HDCN created by Nina Henninger [Hen23]. `HDCN-CJ` refers to the HDCN created by Chen Jiang [Jia22]. The STRING cutoff score for both of these was 0.4.

- The `ADCN-KWS-0.4` was intersected with the `HDCN-NH` to allow viewing of the overlap. This resulted in a merged network, called `ADCN∩HDCN`. These three networks were used for the final clustering and enrichment, and their clusters are called C1-A, C1-H and C1-AH. This means Cluster 1 from the `ADCN-KWS-0.4`, `HDCN-NH` and `ADCN∩HDCN`, respectively.

The databases and the software apps that were used are updated continuously, and the specific versions implemented for this research can be found in Table 1.

Table 1: Version information for data and software tools used for creating the ADCN and its analysis in Cytoscape.

Data/tools	**Version**
STRING database	11.5
KEGG hsa05010	8/11/2021
Wikipathways WP5124	29/5/2021
Cytoscape	3.9.1
Wikipathways (CysApp)	3.3.10
stringApp (CysApp)	2.0.1
NetworkAnalyzer (CysApp)	4.4.8
MCODE (CysApp)	2.0.2
BiNGO (CysApp)	3.0.5
Gene Ontology	1/4/2023

3 Related Work

3.1 Huntington's disease consensus networks

The Huntington's disease consensus network (HDCN) that will be used to compare to the Alzheimer's disease consensus network (ADCN) was created by Nina Henninger (N.H.) [Hen23]. This is a new version of the consensus network that was originally created by Chen Jiang (C.J.) [Jia22].

Firstly, C.J. created the workflow for combining the available knowledge of HD in a single PPIN. The KEGG, Reactome, Wikipathways, and STRING database were explored, but Reactome was omitted from the consensus network because no relevant information could be found. The Python package Bio.KEGG.REST was used to map KEGG identifiers to UniProt identifiers. Wikipathways identifiers were converted using the mapping tool on the UniProt website and the STRING proteins already contained a UniProt identifier. Next, with the Analyze function in Cytoscape, the network could be analyzed. After the network was divided into clusters, these clusters were analyzed with a functional enrichment analysis. From the results, it was concluded that regulation of cell communication, cellular response, oxidative phosphorylation, behavior, memory, and movement were all processes that were enriched in the network [Jia22].

Secondly, N.H. updated the network created by C.J. and compared this with another network created by Adam Labadorf et al. [LCM18]. For updating the C.J. network, only new STRING data had to be added since the other two databases had not been updated. The rest of the workflow stayed the same. It was discovered that the 5 proteins with the highest degree in the network were ACTB, GAPDH, AKT1, TP53 and MYC. Proteins with a high degree are called hub nodes. From these proteins, MYC was the only one that was also upregulated. Overexpression of this protein could suppress neurogenesis and incite neurodegeneration [Hen23].

3.2 Alzheimer's disease consensus networks

Three related research articles were found in which PPINs were created for AD. Firstly, Soler-López et al. [SLZL+11] created a network by using the OMIM database to identify 12 genes that were most likely causative for AD. Interaction networks have already provided the insight that genes that are involved in the same disease have significantly more interactions with each other than average. The researchers found that the shortest path in the interactome between these 12 genes was smaller than the average shortest path. With a yeast two-hybrid (Y2H) screening, which is a method for discovering protein-protein interactions, there were 66 new candidate proteins found that had a lot of interactions with the known Alzheimer's proteins. Combining this with information from the human interactome, a PPIN was created with 1704 nodes and 5881 edges.

Secondly, Karbalaei et al. [KART+18] found a relationship between AD and non-alcoholic fatty liver disease (NAFLD). They used the tool DisGeNET and found 189 overlapping genes that were expected to play a role in both diseases. With STRING, these genes were combined in a network and clusters were identified using Cytoscape. Several hub proteins, that had a lot of connections within the network, were identified. Using the OMIM database, the researchers concluded that these enriched proteins were related to diabetes and obesity. Thus, AD and NAFLD might be descendants of these diseases.

Thirdly, Calabrese et al. [CMM22] did a large comparison of the protein interaction networks of AD, HD, Parkinson's disease, and amyotrophic lateral sclerosis (ALS). The most well-known relation between these four neurodegenerative diseases is the formation of protein aggregates. To find out how the aggregation of proteins is related to these four diseases, the researchers created a PPIN, using data obtained from BioGRID, with the interactions between the 6 most relevant proteins. For AD, these were APP and Tau, and for HD this was HTT. Centrally located in this PPIN were the proteins Hsc70 and Hsp70 which also play a crucial role in preventing protein aggregation. Also, all 6 proteins that aggregate in the nervous tissue can cause subsequent cascades in the mitochondria that might lead to cell death. An interesting issue that the researchers raise is that interaction data between proteins is usually based on the interactions in healthy individuals, and to get a better idea of how proteins interact, a database with protein-protein interactions for specific diseases would be a solution.

4 Methods

This section includes the methods for building the ADCNs and merging the final ADCN with the `HDCN-NH`. An overview of the methodology can be found in Figure 1. The ADCNs were created in cooperation with Noria Yousufi (N.Y.). The basic framework used for the methodology was a review article by Koh et al. [KPA+12] on how to do a protein-protein interaction network analysis. The workflow for creating the ADCNs was inspired by the workflow for the `HDCN-CJ` [Jia22].

4.1 Creating the Alzheimer's consensus network

For creating the ADCN, the data from different databases first needed to be collected and preprocessed before it could be merged. In this section, these steps will be explained in detail.

4.1.1 Data gathering

The first step for creating a consensus network is gathering the data. Four different databases were explored, namely KEGG, STRING, Reactome, and Wikipathways. However, for Reactome, no relevant AD pathways were identified. The three other databases are described in the following paragraphs.

The Kyoto Encyclopedia of Genes and Genomes (KEGG) database was originally developed to map genes to biological pathways. Nowadays, it is not only used for genomics but also all other kinds of omics. The database consists entirely of KEGG objects with unique identifiers ranging from smaller biological molecules to disease pathways. The KEGG DISEASE database contains pathways of only human diseases, among which AD [KFT+17].

Wikipathways is a pathway database that contains transcriptomics, proteomics, and metabolomics information about various organisms. The pathways are updated through crowd-sourcing. So, many different contributors keep the Wikipathways database up to date. Its information-sharing values are based on the FAIR data principles, meaning the data is findable, accessible, interoperable, and reusable [SKH+18].

The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) mainly focuses on functional associations between proteins in all kinds of organisms. Proteins that play a part in the same biological process are said to be functionally associated. STRING uses several different methods to give each interaction an evidence score. Firstly, text-mining methods where PubMed abstracts are scanned to detect how often two proteins are mentioned in the same article. Secondly, interaction data is obtained from high-throughput experiments. Thirdly, computational predictions are used and lastly interaction evidence for multiple organisms when they share similar proteins [SGN+21].

In Cytoscape, stringApp is a useful plugin that can be used for all kinds of operations on STRING data. The information on diseases, tissues, and compartments is updated every week. With stringApp, networks can be expanded [DMGJ18].

The identifiers for the interaction data that were obtained from each database can be found in Table 2. The AD pathway obtained from KEGG was the only pathway found in the KEGG DISEASE database for Alzheimer's disease. There was also only one set of proteins for Alzheimer's disease in STRING. For Wikipathways, there were two pathways available, but one also contained miRNA information and this did not seem useful since the focus here lies on protein-protein interactions. For STRING, there was also just one identifier for AD, but here the number of proteins and interactions retrieved from this identifier could be customized.

Table 2: Identifiers of the data sources used for constructing the ADCNs and links to where they were retrieved.

Database	**Identifier for AD data**	**Link**
KEGG	hsa05010	`https://www.kegg.jp/pathway/hsa05010`
Wikipathways	WP5124	`https://www.wikipathways.org/pathways/WP5124.html`
STRING	DOID:10652	`https://string-db.org/`

4.1.2 Data preprocessing

The KEGG database has its own identifiers, so these first had to be converted to UniProt identifiers, which are the more conventional identifiers for proteins. Also, in the KEGG Alzheimer pathway data, some identifiers encoded multiple proteins. Therefore, a Python script was written called `kegg_data_converter.py` which splits these protein groups into single proteins and retrieves all protein identifiers. The final resulting network is saved as a `.gml` file, which can easily be imported into Cytoscape as a network. The script can be found at kegg_data_converter.py on git.liacs.nl.

The Wikipathways Alzheimer data also needed UniProt identifiers to be added. This was also done using a Python script written by N.Y. and can be found at Wikipathway_UniprotIDs_conversion. The identifiers in the resulting table were merged into the table of the Alzheimer network imported with the Wikipathways App in Cytoscape.

For STRING, the size of the Alzheimer's network could be customized, so the number of proteins chosen was set to the maximum in Cytoscape of 2000 proteins. Also, the confidence cutoff score could be varied. Lower confidence will give more information but also possibly more noise in the data. Higher confidence gives less false positive interactions but important information might be missing. So for comparison, two versions were made, one with a lower confidence cutoff of 0.4 and one with a higher confidence of 0.7, so that the differences could be compared.

As a next step, the KEGG network and Wikipathways network were converted to STRING networks by using a function from the Cytoscape stringApp [DMGJ18] called *Stringify*. The UniProt identifiers were used as keys for finding the corresponding protein in STRING. The exact settings for this can be found in Table 3.

The use of Stringify had a few reasons. Firstly, because this way, expression data of whether proteins were expressed in the nervous system became directly available which was useful for filtering later on. Secondly, not all UniProt identifiers in the STRING network were up to date with the most recent versions used for the KEGG and Wikipathways databases. Stringifying the KEGG and Wikipathways networks ensured the same identifiers for the same proteins. Lastly, the expansion of the network is done with the Cytoscape stringApp, which only expands proteins in the STRING database. So if the KEGG data and Wikipathways data were not converted to STRING networks, this data could also not be used for expansion.

Lastly, in each network, additional integer columns were added called `kegg_db`, `wiki_db`, and `string_db` with all values set to 1. This was done in order to be able to retrieve which columns came from which databases after the merge. The columns used for merging were the `stringdb::canonical name` columns, which are the UniProt identifier columns in a STRING network. To make sure there were no key columns that had a None value, the rows that did not have a `stringdb::canonical name` value were deleted. These were not proteins but compounds or processes present in the KEGG and Wikipathways networks.

Table 3: Parameter settings for stringifying KEGG and Wikipathways networks.

Parameter	**Setting**
Column for STRING query	uniprot
Include unmappable nodes	False
Map nodes to compounds	False
Species for the query	Homo Sapiens

4.1.3 Merging and expanding

The first merged network was created by merging the stringified KEGG, Wikipathways and STRING network all with a confidence score of 0.4. Secondly, the confidence of the stringified KEGG and Wikipathways networks was set to 0.7 with the StringApp *Change Confidence* tool. They were then merged with the STRING network with a confidence of 0.7. So, in total, two different merged networks were created, the `ADCN-KWS-0.4` and the `ADCN-KWS-0.7`. The settings for merging the networks can be found in Table 4.

Table 4: Parameter settings for merging the STRING, KEGG and Wikipathways networks.

Parameter	**Setting**
Operation	Union

Networks to merge (confidence 0.4/0.7)	STRING DOID:10652 (2000 proteins);
	Wikipathways WP5124 (stringified);
	KEGG hsa05010 (stringified)

Table 4 continued from previous page

Matching columns	stringdb::canonical name
How to merge columns	Default
Enable merging nodes/edges in the same network	True

Next, the `ADCN-KWS-0.4` and `ADCN-KWS-0.7`, were expanded with 2000 proteins each, to get even more information on the different key proteins involved in Alzheimer's disease. This is something that was not done for the two Huntington's disease networks, because they were built around only one protein, namely huntingtin. The settings for expansion can be found in Table 5. The resulting networks were the `ADCN-KWS-0.4-EXP` and the `ADCN-KWS-0.7-EXP`.

Table 5: Parameter settings for expanding the `ADCN-KWS-0.4` and `ADCN-KWS-0.7`.

Parameter	**Setting**
Number of interactors to expand network by	2000
Type of interactors to expand network by	Homo Sapiens
Selectivity of interactors	0.5

4.2 Automation of workflow

Cytoscape has a REST API called cyREST, via which workflows for creating networks can be automated [OMK+15]. This was done for the workflow for creating the ADCNs. Multiple libraries can connect to cyREST, and the one used here was py4cytoscape [OBND]. The automated workflow can be found at alzheimer_ppin_automated_workflow.ipynb, with an explanation of the different steps. The setup for this notebook was inspired by the ones found on the Cytoscape Automation GitHub page [Xin21].

The same steps are executed as described in the method above, and when it was tested the automated workflow gave the same networks as the ones created manually. However, STRING is updated a lot, so later runs will probably give slightly different results. Still, this kind of workflow is a practical tool and can also be adjusted to build other consensus networks. The workflow works only for a single STRING confidence score. So, in this case, it would have to be executed twice to get all four networks with confidence 0.4 and confidence 0.7.

4.3 Comparing and merging the two consensus networks

In total, four different Alzheimer's disease consensus networks were created; `ADCN-KWS-0.4`, `ADCN-KWS-0.7`, `ADCN-KWS-0.4-EXP`, and `ADCN-KWS-0.7-EXP`. In the second part of this research, these networks were analyzed and overlapped with the previously created `HDCN-NH`. Firstly, it had to be determined which networks would be used for the final clustering and enrichment analysis. Then, significantly enriched processes could be identified in the discovered clusters.

4.3.1 Network analysis

Networks can be analyzed with many parameters, which are described in Table 6. Retrieving the values of these parameters was done in Cytoscape by using the *Analyze Network* tool, with the setting ‘Analyze as Directed Graph?’ set to False.

Table 6: Parameters for analyzing PPINs. The definitions assume an undirected graph since the networks that are being analyzed are undirected. The definitions were adjusted from the review articles [Ram10][BO04][MV07][DH07].

Parameter	**Definition**
Node	A node is an element in the graph representing a single protein.
Edge	An edge is a connection between two nodes in the graph, representing a binding or association between two proteins.
Node degree	The node degree k is the number of edges that connect a node to other nodes.
Node degree distribution	The degree distribution P(k) is a probability distribution of the number of edges that a node has. It can help with distinguishing different network architectures.
Characteristic path length	The characteristic path length L is the average of the shortest paths between every pair of nodes in the network. The shortest path is the smallest amount of edges it takes to get from one node to another.
Average clustering coefficient	The clustering coefficient C of a node is a measure that indicates how well connected the neighbors of that node are to each other. So, a node that has highly connected neighbors will have a higher clustering coefficient. Taking the mean C for all nodes in the networks gives the average clustering coefficient.
Network diameter	The diameter of a network is the length of the longest shortest path between two nodes in the network.
Network density	The network density is a measure of how densely connected the network is.
Network heterogeneity	Network heterogeneity is a measure that indicates the differences in node degree. Large heterogeneity indicates the presence of hub nodes, which are nodes with a relatively large degree. Biological networks are usually heterogeneous.
Network centralization	The network centralization indicates the extent to which nodes differ in how central they lie in the network.
Betweenness	Betweenness indicates how central a node lies in the network. If many shortest paths pass through an edge, then its betweenness is relatively high.
Connected components	The number of connected components indicates how many isolated parts the network has. A graph in which all nodes are connected via edges has one single connected component. A graph with many isolated nodes has many connected components.

4.3.2 Clustering analysis

Different algorithms have been developed to find clusters within networks [KPA+12][SSO+12]. The most popular tool for clustering in Cytoscape is Molecular Complex Detection (MCODE) [BH03], which can detect clusters based on the topology of the network. It looks at how densely connected the neighbors of a node are and based on that identifies clusters within the network [KPA+12][SSO+12]. Some nodes however might not get allocated to any of the clusters. Also, something to be aware of is that MCODE can be sensitive to false positives in the protein-protein interactions [SSO+12]. Lastly, MCODE was specifically developed for finding molecular complexes, so there is a bias there in that these complexes will probably also share molecular functions [BH03].

Detecting similar clusters in the Alzheimer's and Huntington's disease networks might give insight into the processes that occur for both these neurodegenerative diseases. The parameter settings for the clustering in MCODE can be found in Table 7. Clustering was done for the `ADCN-KWS-0.4`, the `ADCN-KWS-0.4-EXP`, the `ADCN-KWS-0.7` and the `ADCN-KWS-0.7-EXP`.

Table 7: Parameter settings for clustering with MCODE.

Parameter	**Setting**
Find Clusters	In Whole Network
Include Loops	False
Degree Cutoff	2
Haircut	True
Fluff	False
Node Score Cutoff	0.2
K-Core	2
Max. Depth	100

4.3.3 Intersecting the ADCN and HDCN

After the network analysis and clustering analysis, only the `ADCN-KWS-0.4` was merged with the `HDCN-NH`. The reason for this was based on the results from the network and clustering analysis, which will be described in more detail in Section 5. Since the goal of this research was to observe the overlap between the networks, instead of the union, the intersection of the `ADCN-KWS-0.4` with the `HDCN-NH` was taken. Only first, the `HDCN-NH` was stringified, just like the KEGG and Wikipathways networks. This was done so that the proteins from KEGG and Wikipathways that were in the `HDCN-NH` also had a STRING identifier just like all the proteins in the two ADCNs. The settings for stringifying and merging can be found in Table 8 and Table 9, respectively.

The resulting network was the `ADCN∩HDCN`, which was also clustered with MCODE, using the same settings as in Table 7. Lastly, the `ADCN-KWS-0.4`, `HDCN-NH` and `ADCN∩HDCN` were used as networks for the final enrichment analysis.

Table 8: Parameter settings for stringifying the HDCN-NH. The name column was used for the query because it had no empty rows.

Parameter	**Setting**
Column for STRING query	name
Include unmappable nodes	True
Column for unmappable node labels	uniprot
Map nodes to compounds	False
Species for the query	Homo Sapiens

Table 9: Parameter settings for merging the ADCN-KWS-0.4 with the stringified HDCN-NH. This resulted in the ADCN∩HDCN

Parameter	**Setting**
Operation	Intersection
Networks to merge	ADCN-KWS-0.4; HDCN-NH
Matching columns	stringdb::canonical name
How to merge columns	Default
Enable merging nodes/edges in the same network	True

4.3.4 Enrichment analysis

The most common way to get enrichment information on a set of proteins is by using the Gene Ontology (GO) database [Con04]. This database is subdivided into three separate classes, the Cellular Component ontology, the Biological Process ontology, and the Molecular Function ontology. The Biological Process annotation of proteins can be used to find whether certain processes are enriched in a cluster of proteins [SSO+12][MHK05][DMGJ18].

For functional enrichment, the most used Cytoscape tool is BiNGO [MHK05]. It finds the GO terms that are most significantly enriched in a cluster and highlights them in a directed graph of the GO hierarchy. It looks at the chances of a certain term being found for a certain number of proteins in a set. In addition, statistical methods such as the Bonferroni correction or the Benjamini and Hochberg correction can be used to correct for false positives found in the enriched GO terms. Another tool that can be used is the stringApp feature for functional enrichment analysis of STRING data. Enrichment data of STRING proteins can be retrieved from the STRING database as a table in Cytoscape [DMGJ18].

The following enrichment was done for both the ADCN-KWS-0.4 and the HDCN-NH separately, and for the intersected ADCN∩HDCN. This way, the enriched processes in the separate networks could be compared to one another and the enriched processes in the ADCN∩HDCN. Both BiNGO and stringApp were used for the analysis.

The first step was to determine which clusters contained the proteins APP, APOE, MAPT, PSEN1, PSEN2 and HTT. These will be referred to as ‘disease proteins’. The reasons for choosing these proteins are as follows. APP and MAPT are the amyloid-β and tau proteins, from which the aggregates seem to play a key role in Alzheimer's disease progression. The APOE e4 allele is a causal factor for late-onset AD, while APP, PSEN1, and PSEN2 play a role in early-onset Alzheimer's disease [SDSK+21]. It was also interesting to see whether these proteins were present in the `HDCN-NH`. For HTT the question was whether it could be found in the `ADCN-KWS-0.4`.

As a second step, an enrichment analysis could be done for all clusters containing these disease proteins. Firstly, a STRING enrichment was retrieved for each cluster. The enrichment was filtered to display only biological processes and leave out redundant terms. Secondly, a BiNGO graph was created for each of the clusters. The parameter settings for this can be found in Table 10.

For the BiNGO analysis, the most recent version of the GO was downloaded, since the default in BiNGO is a much outdated version [KPA+12]. It can be downloaded at `http://geneontology.org/docs/download-ontology/#go_basic`. The version used for this research (data-version: releases/2023-04-01) can be found on git.liacs.nl at go-basic.obo. Also, the default STRING identifiers from the clusters could not be used in BiNGO, so the stringdb::canonical names of the cluster proteins had to be pasted manually in the parameter settings.

Table 10: Parameter settings for clustering with BiNGO.

Parameter	**Setting**
Retrieve data	Paste proteins from table (stringdb::canonical names)
Assess	Overrepresentation
Visualization	True
Statistical test	Hypergeometric test
Multiple testing correction	Benjamini & Hochberg False Discovery Rate (FDR) correction
Significance level	0.05
Categories to be visualized	Overrepresented categories after correction
Reference set	Use whole annotation as reference set
Ontology file	go-basic.obo
Namespace	biological_process
Organism/annotation	Homo sapiens
Discard evidence codes	-

5 Results

This section discusses the results of the network analysis, clustering, and enrichment of the networks. The Cytoscape session files with the full results can be found at cytoscape files. The images of the networks that were created can also be found in Appendix A.2 and at network images. The clusters of the `ADCN∩HDCN` can be found in Appendix A.3 and a full overview of the clusters is available at cluster images and cluster data.

5.1 Network analysis

The network images of the ADCNs can be found in Figures 2, 3, 4, 5. In these networks, the database distributions and the confidence that a protein is active in the nervous system are graphically visualized. Also, the six important disease proteins (APP, MAPT, APOE, PSEN1, PSEN2, HTT) that were the focus of the clustering analysis are enlarged. When looking at the four networks globally, they appear quite similar. An interesting result is that the huntingtin protein is also present in all these networks, but it was only obtained from the STRING database, and not the pathway databases. The other five proteins were, as expected, present in all three databases.

In the KEGG and Wikipathways networks, there was a smaller amount of nodes than in the STRING network, but more certainty that these proteins were involved in Alzheimer's disease pathways. Therefore, it was interesting to see how the information from the three databases overlapped in the ADCNs. The `ADCN-KWS-0.4` and `ADCN-KWS-0.7` both had the same amount of nodes and therefore the same database distribution, since the STRING confidence only changes the number of edges in a network. The database distribution of these two networks can be found in Table 11. Most nodes in the networks are from STRING, but there was a total of 136 proteins that overlapped between all three databases.

Table 11: Distribution of nodes over the KEGG, Wikipathways, and STRING database in the merged ADCNs. These amounts are the same for `ADCN-KWS-0.4` and `ADCN-KWS-0.7`.

Databases	**Node count**
KEGG	84
Wikipathways	1
STRING	1726
KEGG;Wikipathways	115
KEGG;STRING	41
Wikipathways;STRING	1
KEGG;Wikipathways;STRING	136

After the expansion of the two consensus networks, all four were analyzed using the *Network analysis* tool from stringApp. The analysis results can be found in Table 12. The analysis from the `HDCN-NH` was reused from [Hen23]. It can be seen that the number of nodes for the networks with confidence 0.4 and 0.7 is the same, but the number of edges in the `ADCN-KWS-0.4` is about three times as large as in the `ADCN-KWS-0.7`. Something that stands out is that the network parameter values for the `ADCN-KWS-0.4` and the `HDCN-NH` lie relatively close to each other compared to the values from the other networks. This is the first indication that the `ADCN-KWS-0.4` might be the best option for overlapping with the `HDCN-NH`.

Interestingly, the clustering coefficient of the `ADCN-KWS-0.4-EXP` is smaller than for the `ADCN-KWS-0.4`, while the clustering coefficient of the `ADCN-KWS-0.7-EXP` is larger than for the `ADCN-KWS-0.7`. Furthermore, the amount of connected components for the networks with a confidence of 0.7 is a lot higher than for the networks with a confidence of 0.4. This can also be seen in the network images because the `ADCN-KWS-0.7` and `ADCN-KWS-0.7-EXP` have a lot of isolated nodes that are not connected through any edges with the rest of the network. This indicates that these are all single components.

Table 12: Analysis results from the different ADCNs and the previously created HDCN-NH

Network	**ADCN-KWS-0.4**	**ADCN-KWS-0.4-EXP**	**ADCN-KWS-0.7**	**ADCN-KWS-0.7-EXP**	**HDCN-NH**
Node count	2104	4104	2104	4104	2023
Edge count	95608	256457	31810	96788	82753
Avg. neighbors	89.980	124.495	29.961	47.072	81.587
Diameter	6	5	8	8	5
Radius	3	3	4	4	3
Characteristic path length	2.282	2.286	3.014	2.903	2.334
Clustering coefficient	0.419	0.356	0.400	0.416	0.417
Network density	0.043	0.030	0.015	0.012	0.040
Network heterogeneity	0.975	0.885	1.059	0.962	0.958
Network centralization	0.371	0.313	0.119	0.112	0.336
Connected components	2	1	45	33	3

5.2 Clustering

Initially, five MCODE clustering analyses were performed for the five consensus networks. The number of clusters for each network was between about 30 and 60, which can be seen in Table 13. So, it was necessary to filter, and therefore only the clusters that contained either APP, MAPT, APOE, PSEN1, PSEN2, or HTT were looked at in more detail.

An overview of the clusters that contained the disease-related proteins, the number of nodes and edges, and the clustering score, can be seen in Table 13. The names of the clusters are ordered from the highest to lowest clustering score of the corresponding network. A higher clustering score means higher confidence that a group of proteins forms a complex.

A first important observation from these clustering results was that the expanded networks did not seem to give a lot of new information compared to the non-expanded networks. The goal of the expansion was to see if more relevant information could be gained, but from these results, it appears that this is not the case. Firstly, the ADCN networks with a confidence score of 0.7 both have as a first cluster one that contains MAPT, then one that contains APP and APOE, and one that contains PSEN1 and PSEN2. The only difference is that the clusters of the ADCN-KWS-0.7-EXP are larger, and none of them contained HTT. In the ADCN networks with confidence 0.4, the same kind of similarity can be seen between the expanded and non-expanded networks. Also, the ADCN-KWS-0.4-EXP had fewer clusters than the smaller ADCN-KWS-0.4. This corresponds to the lower clustering coefficient.

A further observation to make is that the disease proteins in the HDCN-NH were all spread over different clusters, while in the Alzheimer's disease networks, the clusters also contained combinations of these proteins. Also, no cluster was found that had a combination of huntingtin and one of the AD proteins.

Finally, because there were a lot of similarities between the clusters of the ADCNs, it was decided to only continue with the `ADCN-KWS-0.4` and overlap this with the `HDCN-NH`. This one had more clusters than the `ADCN-KWS-0.4-EXP`, but more edges than the two networks with confidence 0.7. Therefore, this choice was made, because it would be preferred to have a bit more noise in the data compared to potentially missing information.

Table 13: Cluster information on the clusters that contained the proteins APP, APOE, MAPT, PSEN1, PSEN2, or HTT.

Network	**Total #Clusters**	**Cluster**	**Disease proteins**	**Node count**	**Edge count**	**Clustering score**
ADCN-KWS-0.4	35	Cluster 1	APP, APOE, MAPT	209	8963	85.923
		Cluster 4	HTT	153	1149	15.039
		Cluster 8	PSEN1	108	488	9.121
		Cluster 19	PSEN2	15	27	3.857
ADCN-KWS-0.4-EXP	32	Cluster 2	MAPT, APOE	363	11538	63.182
		Cluster 3	APP	515	12766	49.14
		Cluster 5	PSEN1	449	3882	17.254
		Cluster 12	HTT	118	316	5.402
ADCN-KWS-0.7	42	Cluster 1	MAPT	64	2041	62.317
		Cluster 3	APP, APOE	71	599	17.086
		Cluster 9	PSEN1, PSEN2	201	890	8.65
		Cluster 27	HTT	22	33	3.143
ADCN-KWS-0.7-EXP	56	Cluster 1	MAPT	210	7565	69.005
		Cluster 4	APP, APOE	330	4274	25.915
		Cluster 14	PSEN1, PSEN2	214	632	5.897
HDCN-NH	30	Cluster 2	APOE	128	2482	38.882
		Cluster 3	APP	178	2309	25.605
		Cluster 5	PSEN1	169	1268	15.071
		Cluster 8	MAPT	83	324	7.902
		Cluster 9	HTT	138	561	7.723
		Cluster 19	PSEN2	60	124	4,203

5.3 ADCN-HDCN intersection

Before merging, the `HDCN-NH` was stringified, which can be seen in Figure 6. The original network can be found in [Hen23], but the only difference is that all nodes now have STRING properties. The number of nodes for the stringified network was still 2023 and the number of edges 82753. The STRING confidence score was 0.4. In the Figure, it can be seen that now the HTT protein was obtained from all three databases, while the AD proteins were only obtained from STRING.

After intersecting with the `HDCN-NH`, the `ADCN-KWS-0.4` was clustered with MCODE. The results can be seen in Table 14. Interestingly, the HTT and PSEN1 protein were found together in one cluster.

Table 14: Cluster information on the clusters in the ADCN∩HDCN that contained the proteins APP, APOE, MAPT, PSEN1, PSEN2 or HTT.

Network	Total #Clusters	Cluster	Disease proteins	Node count	Edge count	Clustering score
ADCN∩HDCN	32	Cluster 2	APOE	110	2372	43.541
		Cluster 4	APP	103	800	15.686
		Cluster 6	HTT, PSEN1	70	313	9,043
		Cluster 8	MAPT	47	140	6.087
		Cluster 29	PSEN2	13	17	2.833

The ADCN∩HDCN can be seen in Figure 7. In total, 1426 nodes and 58896 edges overlapped, so about two-thirds of the consensus networks. Included is part of a STRING enrichment analysis, displaying the eight processes with the lowest FDR. These were response to chemical, regulation of localization, regulation of cell death, nervous system development, positive regulation of signaling, localization, regulation of developmental process, and regulation of cellular component organization. In the overlapped network, it can be seen that the 6 disease proteins were all present in the intersect of the networks and partake in most of these processes. Only, the PSEN2 protein does not seem to play as central a role as the others.

5.4 Enrichment analysis

In total, 15 clusters were analyzed in more detail with a STRING enrichment and BiNGO enrichment analysis. These were the ones from the ADCN-KWS-0.4, HDCN-NH, and the ADCN∩HDCN that contained the proteins APP, MAPT, APOE, PSEN1, PSEN2, and HTT.

The STRING enrichment analysis ranks processes in order of their False Discovery Rate (FDR) value. A lower FDR means a lower chance that a protein classified as part of the biological process is not part of it. In the images of the clusters found in Appendix A.3 and at cluster images, the eight processes with the lowest FDR value are visualized in the cluster.

For the BiNGO analysis, the graphs that can be seen in the figures are displayed as hierarchical trees. They represent part of the Gene Ontology, with the nodes depicting in this case biological processes [MHK05]. The general processes are positioned at the top, and the more specific processes at the bottom. The white nodes are not enriched, but just connecting parts within the GO of Biological Processes. The node colors are ordered from light yellow to dark orange indicating a small to large significance that these nodes are enriched. The size of the nodes corresponds to how many proteins from the analyzed cluster have an annotation in the GO for that process. The graph is hierarchically ordered, so, if a node of certain color and size points to a node further down the hierarchy of the same color and size, then the node lowest in the hierarchy is the reason for that color and size. From these things, it can be interpreted that the large, dark orange nodes at the bottom of the BiNGO graphs are the most interesting to look at. This is what was done for all the clusters. Overall, the BiNGO graphs showed many of the same branches of the GO.

Cluster 2 ADCN∩HDCN (C2-AH)
The enrichment results of the C2-AH were very similar to Cluster 1 in the `ADCN-KWS-0.4` (C1-A) and Cluster 2 in the `HDCN-NH` (C2-H). They all contained the protein APOE but in the C1-A the APP and MAPT proteins were also present. The STRING enrichment of C2-AH can be found in Figure 8, and the other two at C1-A-STRING and C2-H-STRING. What is seen between all three in general is the involvement of the cluster proteins in *apoptosis*, *cell development*, *phosphorylation*, *cell signalling*, and *stress response*.

The BiNGO hierarchical network of the C2-AH can be seen in Figure 9. The ones for C1-A and C2-H can be found at C1-A-BiNGO and at C2-H-BiNGO. For all these clusters, the BiNGO graph was very large and widely spread out, meaning that a lot of branches from the GO were enriched. When inspecting the graphs more closely, one part at the bottom had some nodes involved in the *regulation of transcription of RNA polymerase II* that were a lot larger and darker colored than the other nodes at the same level of the GO hierarchy.

Cluster 4 ADCN∩HDCN (C4-AH)
The C4-AH contained the protein APP, just like Cluster 3 from the `HDCN-NH` (C3-H) and again the C1-A. For these clusters, the enriched processes from STRING appeared to be involved in *autophagy*, but also just like the C1-A in *response to stress*, *apoptosis*, and *phosphorylation*. The C4-AH can be seen in Figure 10 and the C3-H at C3-H-STRING.

The BiNGO graphs for the C3-H and C4-AH were both a bit less dense than the C1-A. This can also be seen when comparing Figure 11 and C3-H-BiNGO with C1-A-BiNGO. The C4-AH and C3-H both had a darker orange branch with nodes involved in *autophagy* and *protein deacetylation*. The largest difference was that the C3-H also had a very distinct branch involved in the *ubiquitin system*, which is another form of protein degradation next to autophagy.

Cluster 6 ADCN∩HDCN (C6-AH)
The C6-AH contained the proteins HTT and PSEN1. The STRING enrichment in Figure 12 showed that the proteins in this cluster were involved in trans-synaptic transmission, specifically *glutamatergic synaptic transmission*. The BiNGO graph in Figure 13 also highlights a branch of dark orange nodes involved in *chemical synaptic transmission*.

C4-A was the cluster from the `ADCN-KWS-0.4` that contained HTT and can be found at C4-A-STRING. For the STRING enrichment, three out of the four processes with the lowest FDR value are the same ones as those in Figure 12. These were *trans-synaptic signaling*, *regulation of biological quality*, and *behavior*. In the BiNGO graph of C4-A, which can be found at C4-A-BiNGO, the same noticeable branch of *chemical synaptic transmission* was seen as in Figure 13.

Cluster 8 of the `ADCN-KWS-0.4` (C8-A) contained PSEN1, and here, *regulation of cell death*, *localization* and *synaptic transmission* were observed among the STRING-enriched processes, see also C8-A-STRING. The BiNGO graph did not have one branch that stood out, but the *chemical synaptic transmission* branch and the one that ends at *regulation of transcription by RNA poly-*

merase II were again observed. The full graph can be observed at C8-A-BiNGO.

Cluster 5 from the HDCN-NH (C5-H) was the one that contained PSEN1. Here, the STRING enrichment showed again similar processes as before such as *response to stress* and *regulation of biological quality*, but also *DNA repair*, which was not seen in the other clusters. The BiNGO analysis confirmed this also with a DNA repair node at the bottom of the graph that was quite large and dark orange. The STRING and BiNGO analysis results can be found at C5-H-STRING and C5-H-BiNGO.

Cluster 9 from the HDCN-NH (C9-H) contained the HTT protein, and the STRING and BiNGO enrichment can be found at C9-H-STRING and C9-H-BiNGO. The cluster proteins were involved in *trans-synaptic signaling*, but also *microtubule-based movement* according to the STRING analysis. One larger branch in the BiNGO graph showed involvement in *nervous system development*, which was one of the STRING-enriched processes as well. Especially the microtubule-based movement was something that was not observed for the other clusters.

Cluster 8 ADCN∩HDCN (C8-AH)

The C8-AH was a smaller cluster that contained MAPT, of which the STRING analysis can be seen in Figure 14. The cluster proteins appear to be engaged in *chemical synaptic transmission*, *response to oxidative stress*, and *regulation of protein ubiquitination*, all processes that were seen for the other clusters as well.

The path that stood out the most in the BiNGO graph is seen in Figure 15. These are again the nodes involved in *chemical synaptic transmission*, confirming what is seen for the STRING analysis.

Cluster 8 from the HDCN-NH (C8-H) also included MAPT. The STRING enrichment and BiNGO analysis of this cluster, which can be viewed at C8-H-STRING and C8-H-BiNGO, also implied involvement of the cluster proteins in *chemical synaptic transmission* and *nervous system development.*

Cluster 29 ADCN∩HDCN (C29-AH)

The clusters containing PSEN2, which were C29-AH, C19-H and C19-A were very small, had a low clustering score and only resulted in enrichments of much lower significance than the other clusters. The results can all be found at cluster images. The only GO term that was enriched in the C29-AH was *intracellular signal transduction*, and this cluster gave a BiNGO graph of 0 nodes and edges, because no process was enriched above the threshold of 0.05.

6 Discussion

Four different versions of the ADCN were created but in the end, only the ADCN-KWS-0.4 was used for the overlap analysis. The reason was because the network analysis and clustering did not display any advantages of using a higher confidence of 0.7. Also, the expansions resulted in almost the same clusters as were found in the non-expanded networks.

In a review article by Ernhoefer et al. [EWH11], several shared pathways between Alzheimer's and Huntington's disease are listed and the implications for common drug targets are discussed. Interestingly, the five main overlapping processes between AD and HD were also found in the STRING and BiNGO enrichment results above as some of the most significantly enriched processes. The first main overlapping process was *neurotrophic factor-related abnormalities*, which means a dysfunction of proteins involved in neural development. Secondly, *post-translational modifications*, and phosphorylation is one of the most common post-translational modifications. The other three, *protein aggregation clearance mechanisms* such as autophagy and ubiquitin-dependent processes, *synaptic dysfunctioning*, and *apoptotic pathways* will also be discussed in more detail below.

Autophagy

Autophagy is a very important process for the breakdown of aggregated proteins. For both AD and HD, protein aggregation is a primary symptom. Research has shown that defects in the autophagic processes might play a causal role in multiple neurodegenerative diseases, including AD and HD. All forms of autophagy include the fusion of two vesicles, an autophagosome containing the aggregated or misfolded proteins, and a lysosome containing enzymes for the degradation of proteins [GLCL18].

In Alzheimer's disease, mutant PSEN1 is associated with impaired autophagy. PSEN1 aids in the merging of the autophagosome and lysosome. What appears to happen is that this process is disturbed, leading to an accumulation of autophagosomes. In Huntington's disease, aggregation of autophagosomes was also discovered, since HTT is involved in autophagosome formation [GLCL18].

Another protein linked to autophagy is beclin-1. It is very important for the initial formation of autophagosomes. The transcription of this protein was downregulated in AD patients. In mice, the overexpression of beclin-1 reduced the formation of Aβ aggregates [GLCL18][ABR+17]. The protein caspase-6 is important for cleaving both beclin-1 and HTT. In both AD and HD, this protein was upregulated. This might negatively influence the autophagic process that gets rid of protein aggregates in these neurodegenerative diseases [MLEH15].

Chemical synaptic transmission

Synapses are located at the ends of neurons and contain vesicles with neurotransmitters that can be released in the synaptic gap and transmit information from one neuron to another. Therefore, these synapses are crucial in communication between neurons [BK17].

Studies on both AD and HD showed a link between the release of the neurotransmitters dopamine and glutamate and disease progression. In HD, deficits in the release of these two neurotransmitters were found in transgenic mice [RDW+15]. In mouse models of AD, it was also discovered that there was a defect in the synaptic transmission of glutamatergic neurotransmitters [ZK23]. Another large review study found increased disease symptoms of AD because of a lower amount of dopaminergic neurotransmitters [PKW+19].

Under healthy conditions, the Aβ protein increases the chances at neurotransmitter release. In

AD patients, the probability of signal transduction from one neuron to another is decreased. Not only Aβ but also APP and presenilins are involved in synaptic transmission [BK17]. For the tau protein, it is less clear how it is involved in synaptic transmission, but one study by Moreno et al. [MCY+11] examined the effect of human tau on the squid giant synapse. A direct link was found between tau injection into the synapse and a decrease in the vesicle release of neurotransmitters.

Apoptosis and stress response
Excitotoxic cell death is observed for both AD and HD, as was also discussed in the introduction. The same goes for increased reactive oxygen species, which is part of the stress response in the body. So, finding these processes among the enriched ones in the clusters was consistent with the existing literature.

Further discussion
One last process that was not necessarily mentioned in the article by Ernhoefer et al., but which stood out especially in the BiNGO analysis results, was regulation of transcription by RNA polymerase II. It has been found that proteins involved in polyglutamine diseases, the category under which HD falls, are often associated with transcriptional processes [RO06].

The huntingtin protein might be a transcription factor itself. For one, many polyglutamine proteins are transcription factors. Also, changes in the mRNA were observed in people and mice with mutant HTT. In mouse models, there was an increased amount of RNA polymerase II subunit (RPB1) and it was found in aggregates of HTT [LCC03].

In Alzheimer's mouse models, RPB1 was in some cases mislocalized from the cell nucleus to the cytoplasm. Also, there was a correlation between this mislocalization and the amount of tau. Since RNA polymerase is such a vitally important protein, mislocalization of its subunit could have very damaging results [DYFH21]. This mislocalization might also relate to the localization process that was enriched in the clusters.

The results highlight only a few of the many processes that appear to be enriched in the overlapping networks. Most BiNGO graphs of the network clusters were quite large in width, such as in Figure 9. Also, only the clusters that contained the proteins APP, PSEN1, PSEN2, APOE, MAPT, or HTT were used for enrichment analysis. This is probably why the clusters and processes for the different networks were quite similar. Still, relating the processes found in the results back to the literature gives a nice overlap, and implies that enrichment of the clusters could give quite accurate indications of the important overlapping pathways between AD and HD.

7 Conclusion and Further Research

In this research, an Alzheimer's disease protein-protein consensus network was created using the Cytoscape software. The goal was to compare this network with a Huntington's disease consensus network and observe how they overlap.

One of the main results was, that about two-thirds of the proteins in these networks overlapped. Further analysis into the clusters and enrichment of these networks gave an abundance of results. Filtering and looking at significantly enriched nodes in the BiNGO graph showed that the processes of autophagy, synaptic transmission, regulation of transcription, and apoptosis were significantly enriched in multiple clusters. It was also confirmed by a literature analysis that these processes played a role in both diseases.

What can be seen from this research is that network analyses such as these generate a lot of data and that the hardest part is finding the right information within that data. Still, identifying overlapping processes is interesting because it could give new insights into shared drug targets. Also, there might be even more overlap than is currently known, which suggests further research would be useful.

This research also gives rise to a lot of new questions and reasons for further exploration. Firstly, since these bigger overlapping processes were correctly identified by the network, there might be much smaller processes that also overlap for which more research is needed. The MCODE analysis is biased toward finding biological complexes. So, this might explain why there was a lot of evidence found for the processes that overlap in the networks. Secondly, the results focused mostly on the overlap of networks, but it would be also interesting to see how the ADCN and HDCN differ from each other. This could give insights into disease-specific characteristics.

Also, AD is a tauopathy while HD is a polyQ disease. So, they belong to different categories of neurodegenerative diseases but still, there was quite some overlap between the networks. What could also be done is comparing AD with another tauopathy such as Pick's disease, and HD with another polyQ disease such as SBMA (Spinal and bulbar muscular atrophy). Another disease that was often discussed in the literature in combination with AD and HD was Parkinson's disease. So, a comparison between these three diseases could lead to more comprehension of the disease patterns.

To summarize, this research gave more insight into what kind of information can be gained from PPINs and future research should be focused on discovering new overlapping processes and developing methods for identifying shared drug targets.

References

[ABR$^+$17] Avraham Ashkenazi, Carla F Bento, Thomas Ricketts, Mariella Vicinanza, Farah Siddiqi, Mariana Pavel, Ferdinando Squitieri, Maarten C Hardenberg, Sara Imarisio, Fiona M Menzies, et al. Polyglutamine tracts regulate beclin 1-dependent autophagy. *Nature*, 545(7652):108–111, 2017.

[BH03] Gary D Bader and Christopher WV Hogue. An automated method for finding molecular complexes in large protein interaction networks. *BMC bioinformatics*, 4(1):1–27, 2003.

[BK17] Jae Ryul Bae and Sung Hyun Kim. Synapses in neurodegenerative diseases. *BMB reports*, 50(5):237, 2017.

[BO04] Albert-Laszlo Barabasi and Zoltan N Oltvai. Network biology: understanding the cell's functional organization. *Nature reviews genetics*, 5(2):101–113, 2004.

[CMM22] Gaetano Calabrese, Cristen Molzahn, and Thibault Mayor. Protein interaction networks in neurodegenerative diseases: From physiological function to aggregation. *Journal of Biological Chemistry*, 298(7), 2022.

[Con04] Gene Ontology Consortium. The gene ontology (go) database and informatics resource. *Nucleic acids research*, 32(suppl_1):D258–D261, 2004.

[DH07] Jun Dong and Steve Horvath. Understanding network concepts in modules. *BMC systems biology*, 1(1):1–20, 2007.

[DKKK18] Albena T Dinkova-Kostova, Rumen V Kostov, and Aleksey G Kazantsev. The role of nrf2 signaling in counteracting neurodegenerative diseases. *The FEBS journal*, 285(19):3576–3590, 2018.

[DMGJ18] Nadezhda T Doncheva, John H Morris, Jan Gorodkin, and Lars J Jensen. Cytoscape stringapp: network analysis and visualization of proteomics data. *Journal of proteome research*, 18(2):623–632, 2018.

[DYFH21] John R Dickson, Hyejin Yoon, Matthew P Frosch, and Bradley T Hyman. Cytoplasmic mislocalization of rna polymerase ii subunit rpb1 in alzheimer disease is linked to pathologic tau. *Journal of Neuropathology & Experimental Neurology*, 80(6):530–540, 2021.

[EWH11] Dagmar E Ehrnhoefer, Bibiana KY Wong, and Michael R Hayden. Convergent pathogenic pathways in alzheimer's and huntington's diseases: shared targets for drug development. *Nature reviews Drug discovery*, 10(11):853–867, 2011.

[GLCL18] Fang Guo, Xinyao Liu, Huaibin Cai, and Weidong Le. Autophagy in neurodegenerative diseases: pathogenesis and therapy. *Brain pathology*, 28(1):3–13, 2018.

[Hen23] Nina Anna Maria Henninger. Extending consensus knowledge in Huntington's Disease Protein Interaction Networks. *LIACS Thesis Repository*, 2023.

[HHB+21] Harald Hampel, John Hardy, Kaj Blennow, Christopher Chen, George Perry, Seung Hyun Kim, Victor L Villemagne, Paul Aisen, Michele Vendruscolo, Takeshi Iwatsubo, et al. The amyloid-β pathway in alzheimer's disease. *Molecular psychiatry*, 26(10):5481–5503, 2021.

[Jia22] Chen Ji Rong Jiang. Finding consensus knowledge in the Huntington's Disease pathway. *LIACS Thesis Repository*, 2022.

[KART+18] Reza Karbalaei, Marzieh Allahyari, Mostafa Rezaei-Tavirani, Hamid Asadzadeh-Aghdaei, and Mohammad Reza Zali. Protein-protein interaction analysis of alzheimers disease and nafld based on systems biology methods unhide common ancestor pathways. *Gastroenterology and Hepatology from bed to bench*, 11(1):27, 2018.

[KFT+17] Minoru Kanehisa, Miho Furumichi, Mao Tanabe, Yoko Sato, and Kanae Morishima. Kegg: new perspectives on genomes, pathways, diseases and drugs. *Nucleic acids research*, 45(D1):D353–D361, 2017.

[KPA+12] Gavin CKW Koh, Pablo Porras, Bruno Aranda, Henning Hermjakob, and Sandra E Orchard. Analyzing protein–protein interaction networks. *Journal of proteome research*, 11(4):2014–2031, 2012.

[LCC03] Ruth Luthi-Carter and Jang-Ho J Cha. Mechanisms of transcriptional dysregulation in huntington's disease. *Clinical neuroscience research*, 3(3):165–177, 2003.

[LCM18] Adam Labadorf, Seung H Choi, and Richard H Myers. Evidence for a pan-neurodegenerative disease response in huntington's and parkinson's disease expression profiles. *Frontiers in molecular neuroscience*, 10:430, 2018.

[Mat03] Mark P Mattson. Excitotoxic and excitoprotective mechanisms: abundant targets for the prevention and treatment of neurodegenerative disorders. *Neuromolecular medicine*, 3:65–94, 2003.

[MCY+11] Herman Moreno, Soonwook Choi, Eunah Yu, Janaina Brusco, Jesus Avila, Jorge E Moreira, Mutsuyuki Sugimori, and Rodolfo R Llinás. Blocking effects of human tau on squid giant synapse transmission and its prevention by t-817 ma. *Frontiers in synaptic neuroscience*, 3:3, 2011.

[MHK05] Steven Maere, Karel Heymans, and Martin Kuiper. Bingo: a cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, 21(16):3448–3449, 2005.

[MLEH15] Dale DO Martin, Safia Ladha, Dagmar E Ehrnhoefer, and Michael R Hayden. Autophagy in huntington disease and huntingtin in autophagy. *Trends in neurosciences*, 38(1):26–35, 2015.

[MV07] Oliver Mason and Mark Verwoerd. Graph theory and networks in biology. *IET systems biology*, 1(2):89–119, 2007.

[OBND] Keiichiro Ono, Jorge Bouças, Kozo Nishida, and Barry Demchak. py4cytoscape.

[OMK+15] Keiichiro Ono, Tanja Muetze, Georgi Kolishovski, Paul Shannon, and Barry Demchak. Cyrest: turbocharging cytoscape access for external tools via a restful api. *F1000Research*, 4(478):478, 2015.

[PKW+19] Xiongfeng Pan, Atipatsa C Kaminga, Shi Wu Wen, Xinyin Wu, Kwabena Acheampong, and Aizhong Liu. Dopamine and dopamine receptors in alzheimer's disease: a systematic review and network meta-analysis. *Frontiers in aging neuroscience*, 11:175, 2019.

[Ram10] Karthik Raman. Construction and analysis of protein–protein interaction networks. *Automated experimentation*, 2:1–11, 2010.

[RDW+15] T Rothe, M Deliano, AM Wójtowicz, A Dvorzhak, D Harnack, S Paul, T Vagner, I Melnick, H Stark, and R Grantyn. Pathological gamma oscillations, impaired dopamine release, synapse loss and reduced dynamic range of unitary glutamatergic synaptic transmission in the striatum of hypokinetic q175 huntington mice. *Neuroscience*, 311:519–538, 2015.

[RO06] Brigit E Riley and Harry T Orr. Polyglutamine neurodegenerative diseases and regulation of transcription: assembling the puzzle. *Genes & development*, 20(16):2183–2192, 2006.

[Roo10] Raymund AC Roos. Huntington's disease: a clinical review. *Orphanet journal of rare diseases*, 5:1–8, 2010.

[SDSK+21] Philip Scheltens, Bart De Strooper, Miia Kivipelto, Henne Holstege, Gael Chételat, Charlotte E Teunissen, Jeffrey Cummings, and Wiesje M van der Flier. Alzheimer's disease. *The Lancet*, 397(10284):1577–1590, 2021.

[SGN+21] Damian Szklarczyk, Annika L Gable, Katerina C Nastou, David Lyon, Rebecca Kirsch, Sampo Pyysalo, Nadezhda T Doncheva, Marc Legeay, Tao Fang, Peer Bork, et al. The string database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic acids research*, 49(D1):D605–D612, 2021.

[SH16] Frédéric Saudou and Sandrine Humbert. The biology of huntingtin. *Neuron*, 89(5):910–926, 2016.

[SKH+18] Denise N Slenter, Martina Kutmon, Kristina Hanspers, Anders Riutta, Jacob Windsor, Nuno Nunes, Jonathan Mélius, Elisa Cirillo, Susan L Coort, Daniela Digles, et al. Wikipathways: a multifaceted pathway database bridging metabolomics to other omics research. *Nucleic acids research*, 46(D1):D661–D667, 2018.

[SLM+20] Annerieke Sierksma, Ashley Lu, Renzo Mancuso, Nicola Fattorelli, Nicola Thrupp, Evgenia Salta, Jesus Zoco, David Blum, Luc Buée, Bart De Strooper, et al. Novel alzheimer risk genes determine the microglia response to amyloid-β but not to tau pathology. *EMBO molecular medicine*, 12(3):e10606, 2020.

[SLZL+11] Montserrat Soler-López, Andreas Zanzoni, Ricart Lluís, Ulrich Stelzl, and Patrick Aloy. Interactome mapping suggests new mechanistic details underlying alzheimer's disease. *Genome research*, 21(3):364–376, 2011.

[SSO+12] Rintaro Saito, Michael E Smoot, Keiichiro Ono, Johannes Ruscheinski, Peng-Liang Wang, Samad Lotia, Alexander R Pico, Gary D Bader, and Trey Ideker. A travel guide to cytoscape plugins. *Nature methods*, 9(11):1069–1076, 2012.

[STW+98] SK Singhrao, P Thomas, JD Wood, JC MacMillan, JW Neal, PS Harper, and AL Jones. Huntingtin protein colocalizes with lesions of neurodegenerative diseases: An investigation in huntington's, alzheimer's, and pick's diseases. *Experimental neurology*, 150(2):213–222, 1998.

[TGL19] Sarah J Tabrizi, Rhia Ghosh, and Blair R Leavitt. Huntingtin lowering strategies for disease modification in huntington's disease. *Neuron*, 101(5):801–819, 2019.

[TMA+10] Andrey S Tsvetkov, Jason Miller, Montserrat Arrasate, Jinny S Wong, Michael A Pleiss, and Steven Finkbeiner. A small-molecule scaffold induces autophagy in primary neurons and protects against toxicity in a huntington disease model. *Proceedings of the National Academy of Sciences*, 107(39):16982–16987, 2010.

[TMP03] Piero Andrea Temussi, Laura Masino, and Annalisa Pastore. From alzheimer to huntington: why is a structural understanding so difficult? *The EMBO journal*, 22(3):355–361, 2003.

[VD98] Jean Paul G Vonsattel and Marian DiFiglia. Huntington disease. *Journal of neuropathology and experimental neurology*, 57(5):369, 1998.

[VSLvL19] Sofie Voet, Sahana Srinivasan, Mohamed Lamkanfi, and Geert van Loo. Inflammasomes in neuroinflammatory and neurodegenerative diseases. *EMBO molecular medicine*, 11(6):e10248, 2019.

[Xin21] Yihang Xin. Github - cytoscape - cytoscape-automation - wiki, 2021.

[ZK23] Benedikt Zott and Arthur Konnerth. Impairments of glutamatergic synaptic transmission in alzheimer's disease. In *Seminars in Cell & Developmental Biology*, volume 139, pages 24–34. Elsevier, 2023.

A Figures

A.1 Workflow

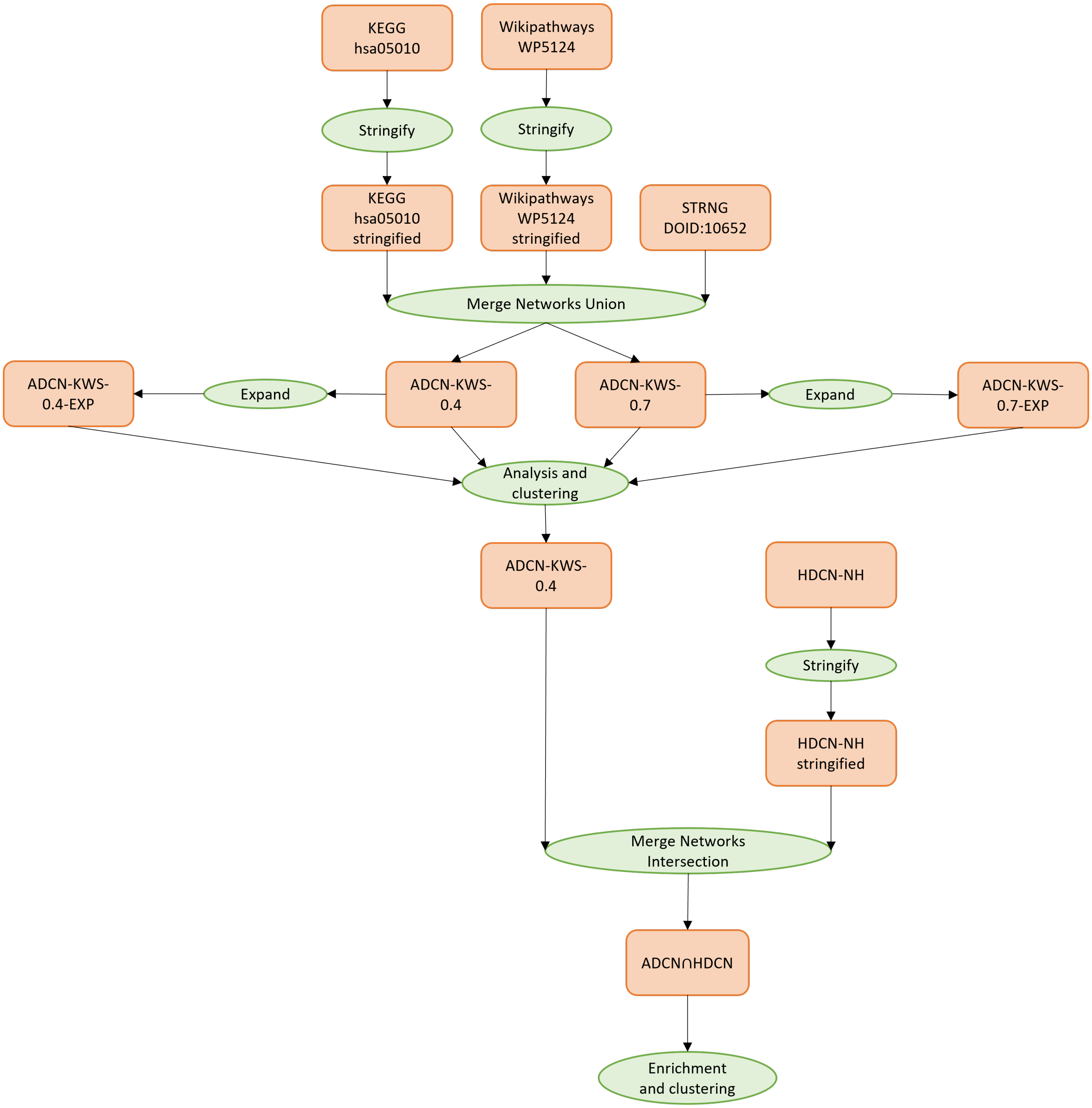


Figure 1: Overview of the full workflow for creating the Alzheimer's disease consensus networks and merging with the Huntington's disease consensus network. Orange nodes are networks in Cytoscape and green nodes indicate Cytoscape network operations.

A.2 Networks

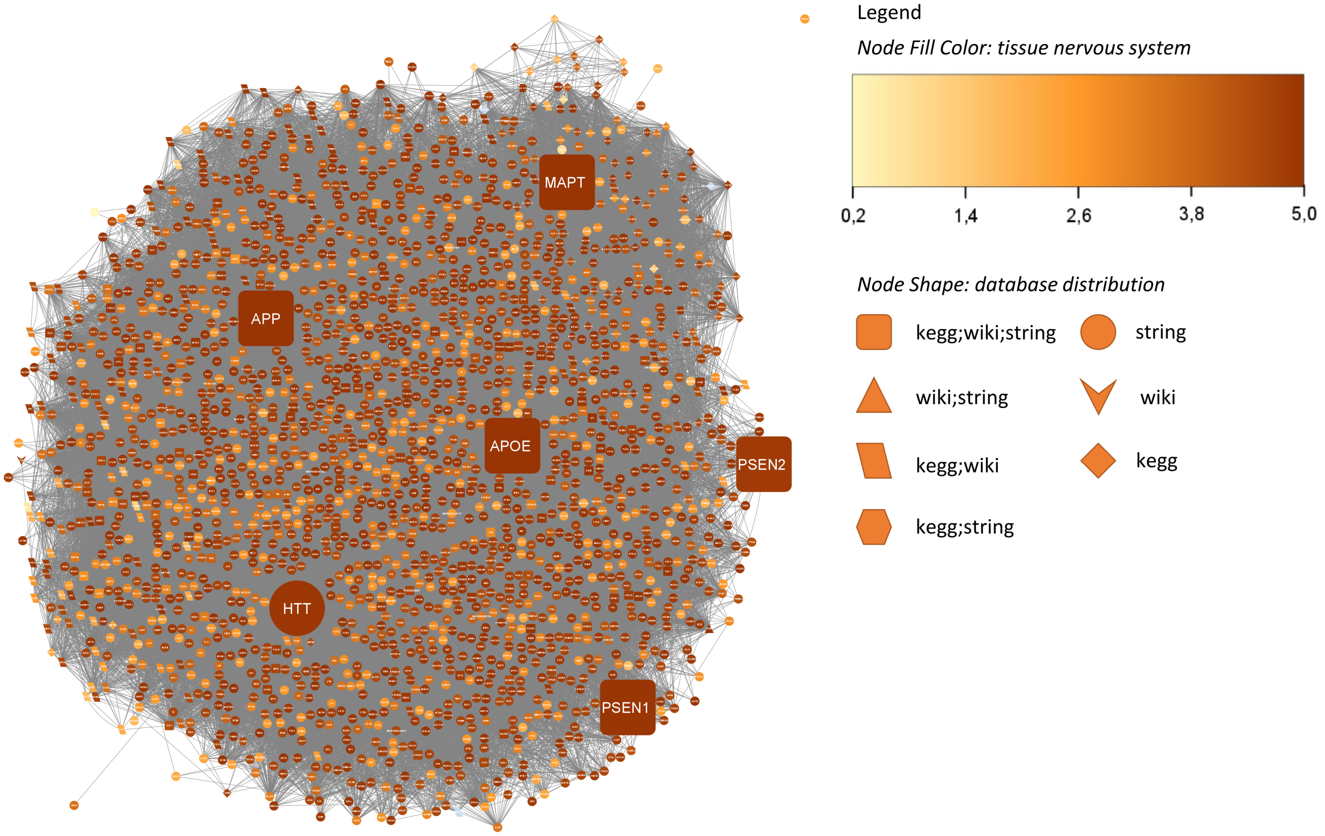


Figure 2: **ADCN-KWS-0.4**, the dark orange nodes indicate a higher certainty that the proteins are active in the nervous system. The database distributions are indicated in the shapes and the APP, MAPT, PSEN1, PSEN2, APOE and HTT proteins are enlarged. The important proteins for AD were obtained from KEGG, Wikipathways and STRING. The HTT protein is also present, but came from only the STRING interactions database.

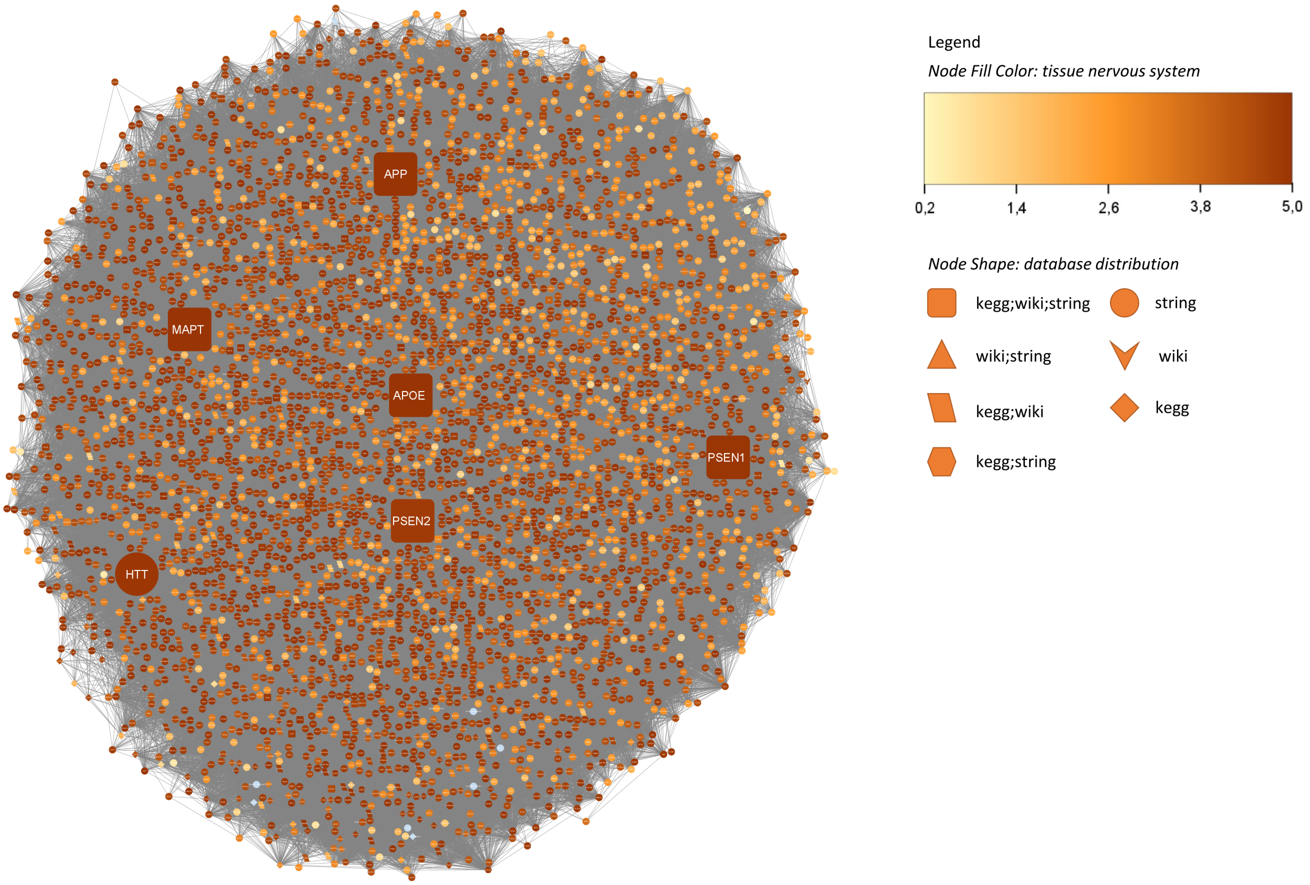


Figure 3: **ADCN-KWS-0.4-EXP**, the dark orange nodes indicate a higher certainty that the proteins are active in the nervous system. The database distributions are indicated in the shapes and the APP, MAPT, PSEN1, PSEN2, APOE and HTT proteins are enlarged. This network was expanded with an additional 2000 proteins from STRING.

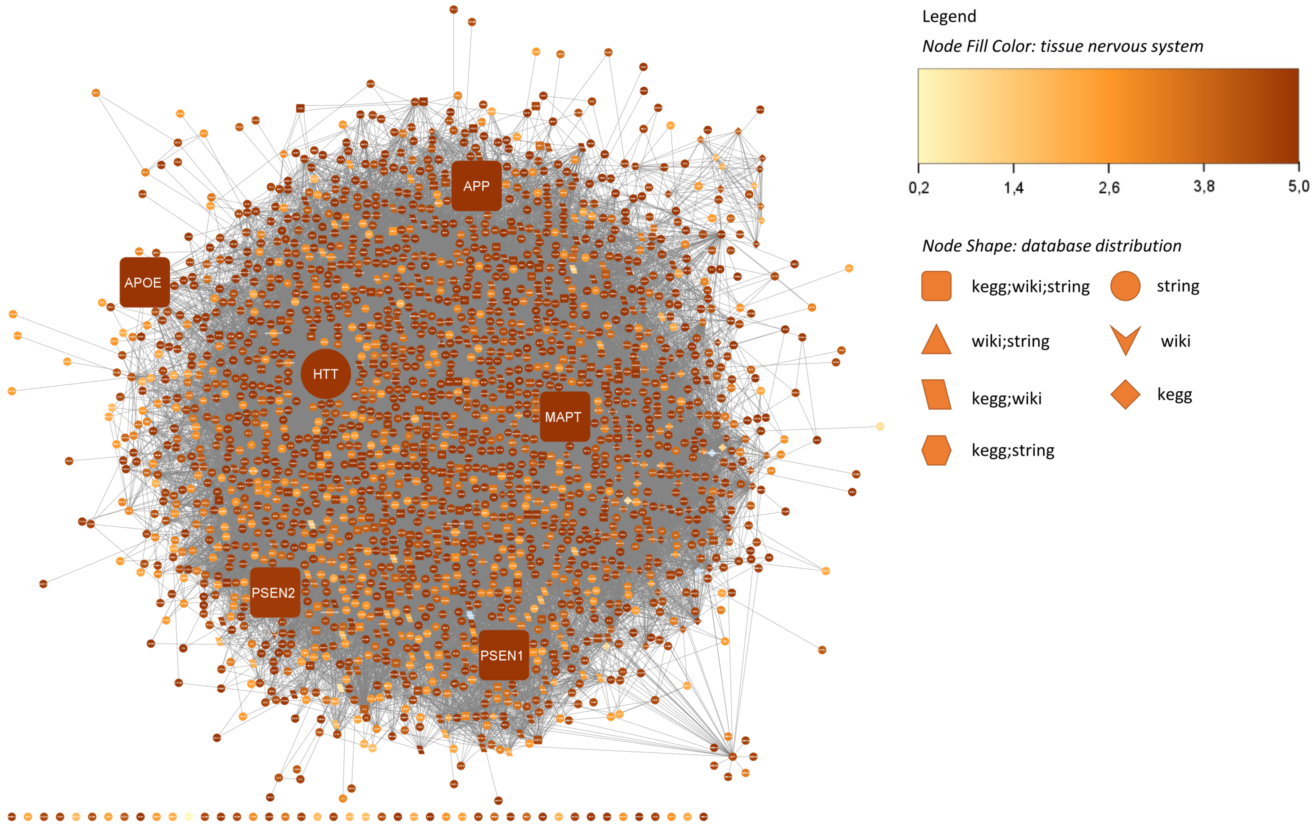


Figure 4: **ADCN-KWS-0.7**, the dark orange nodes indicate a higher certainty that the proteins are active in the nervous system. The database distributions are indicated in the shapes and the APP, MAPT, PSEN1, PSEN2, APOE and HTT proteins are enlarged. The STRING confidence for this network was higher, which causes it to have fewer edges than the networks with a confidence of 0.4. Therefore, there are more isolated nodes.

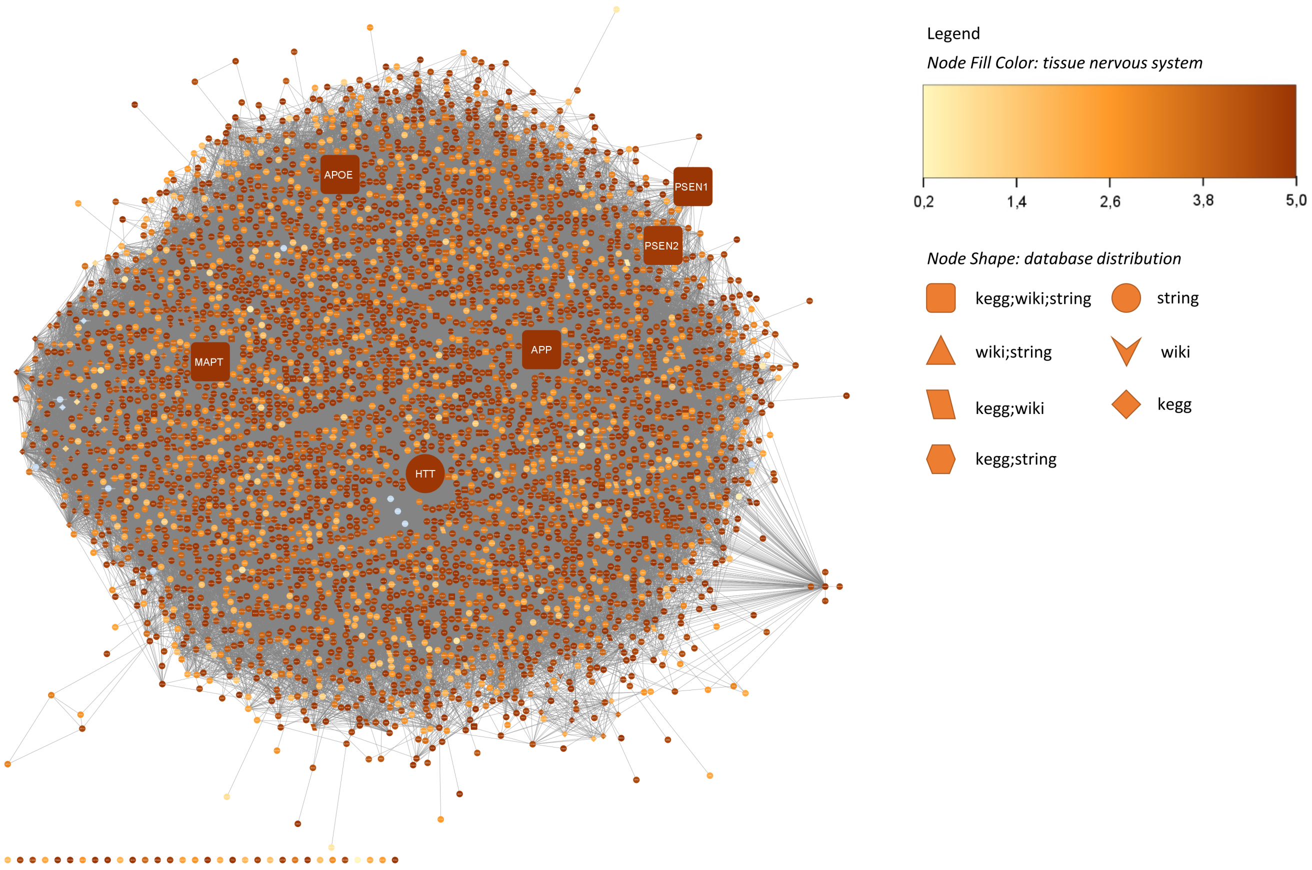


Figure 5: **ADCN-KWS-0.7-EXP**, the dark orange nodes indicate a higher certainty that the proteins are active in the nervous system. The database distributions are indicated in the shapes and the APP, MAPT, PSEN1, PSEN2, APOE and HTT proteins are enlarged. This network was expanded with an additional 2000 nodes from STRING, but there are still isolated nodes that do not connect to any of the others.

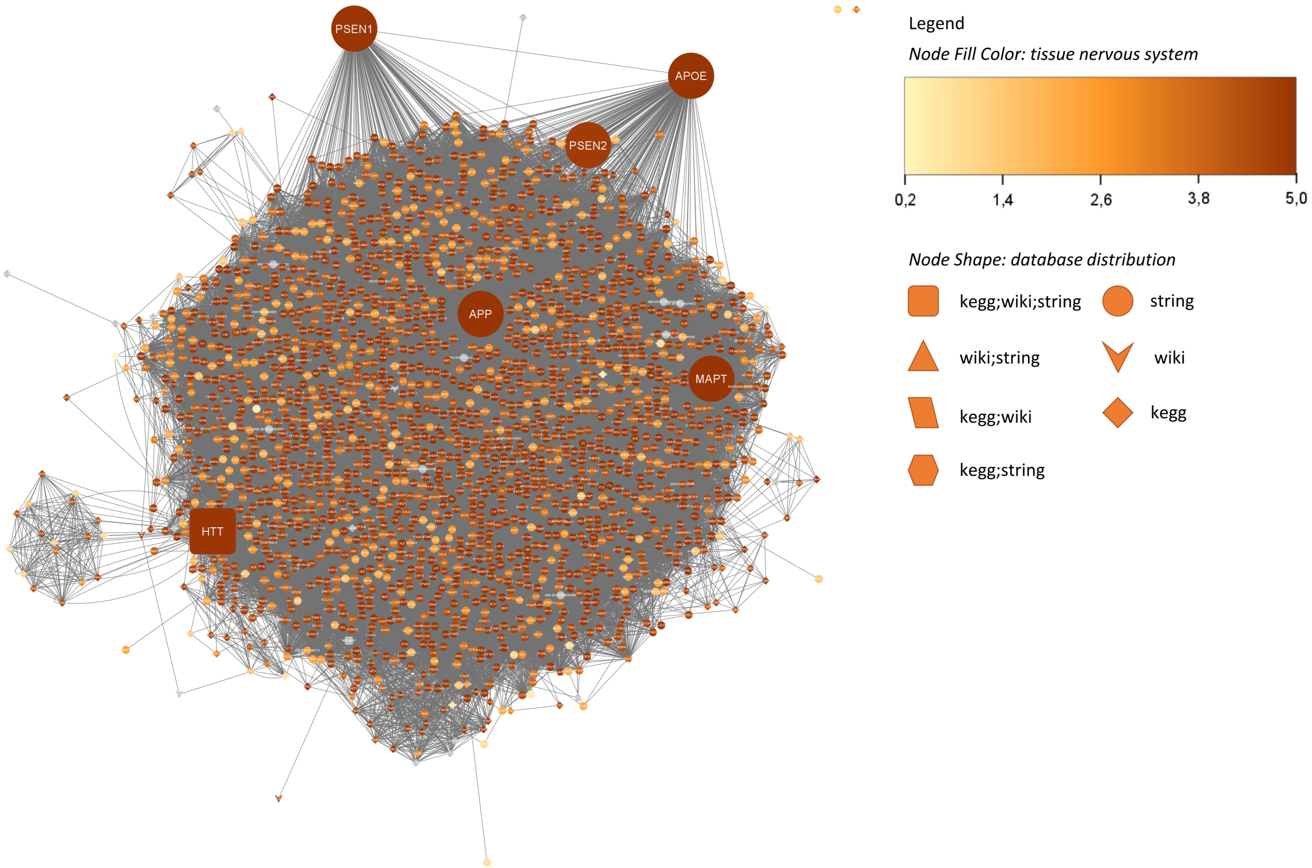


Figure 6: **HDCN-NH**, the dark orange nodes indicate a higher certainty that the proteins are active in the nervous system. The database distributions are indicated in the shapes and the APP, MAPT, PSEN1, PSEN2, APOE and HTT proteins are enlarged. This network was created by Nina Henninger.

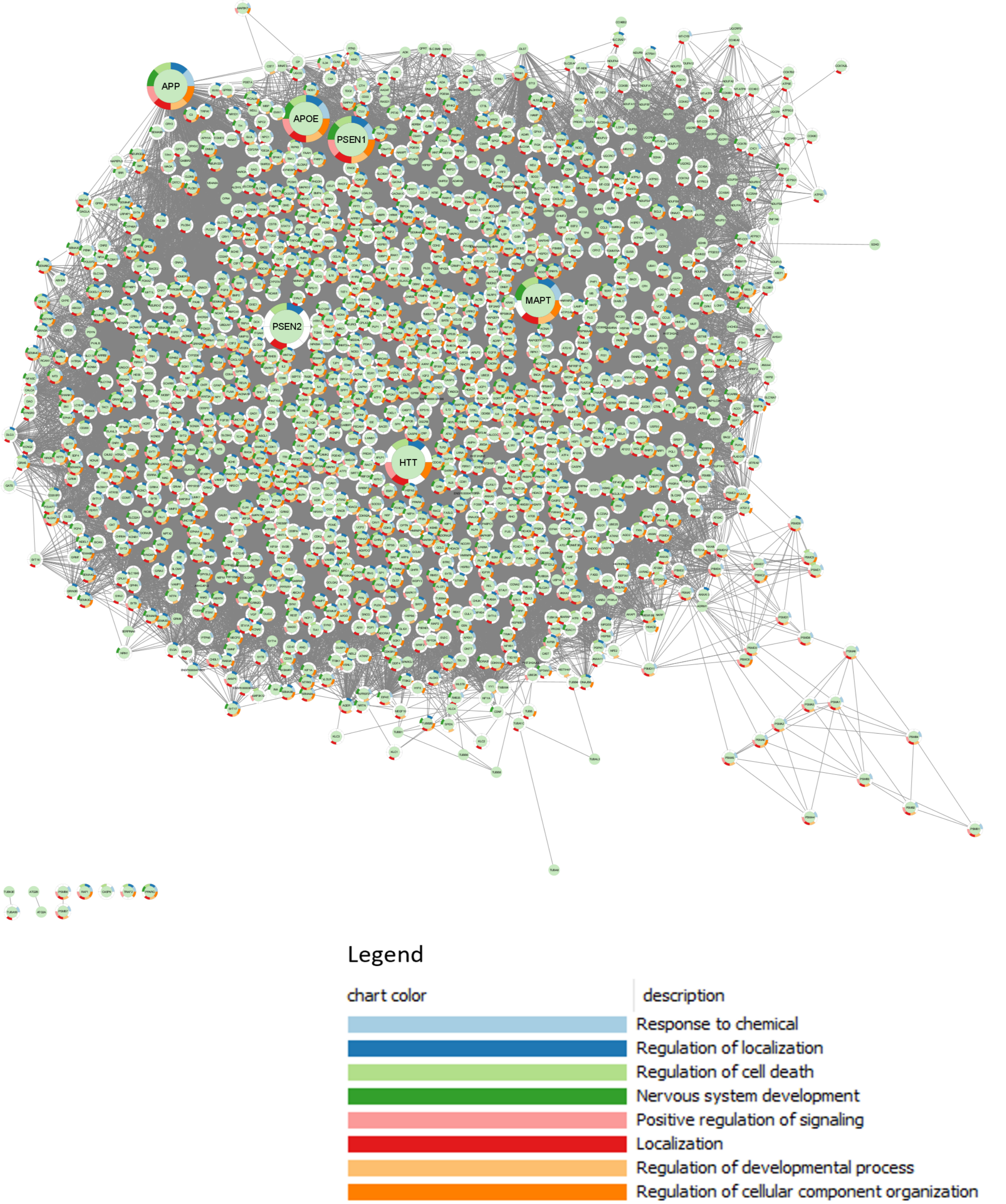
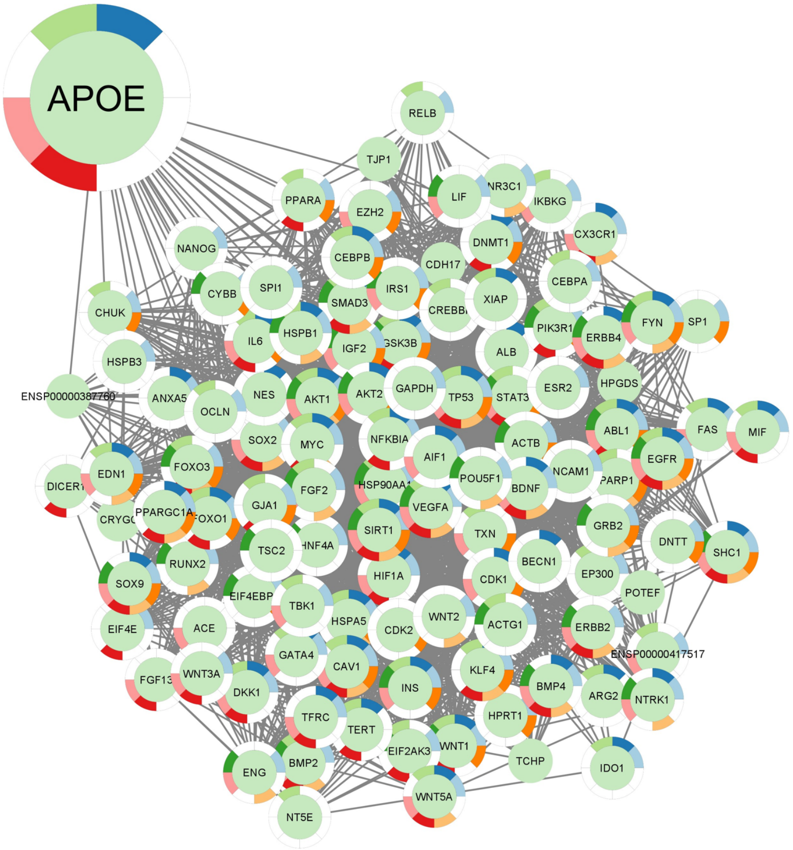
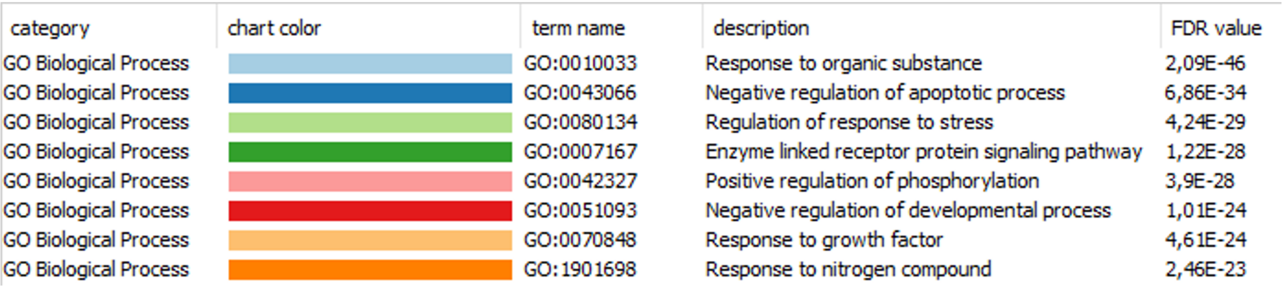

APP
APOE
PSEN1
PSEN2
MAPT
HTT
Legend
chart color
description
Response to chemical
Regulation of localization
Regulation of cell death
Nervous system development
Positive regulation of signaling
Localization
Regulation of developmental process
Regulation of cellular component organization

Figure 7: **ADCN∩HDCN**, the eight processes with the lowest FDR value from the STRING enrichment analysis can be seen. The APP, APOE, PSEN1, PSEN2, MAPT and HTT proteins are enlarged.

A.3 Clusters





category	chart color	term name	description	FDR value
GO Biological Process	🟦	GO:0010033	Response to organic substance	2,09E-46
GO Biological Process	🟦	GO:0043066	Negative regulation of apoptotic process	6,86E-34
GO Biological Process	🟩	GO:0080134	Regulation of response to stress	4,24E-29
GO Biological Process	🟩	GO:0007167	Enzyme linked receptor protein signaling pathway	1,22E-28
GO Biological Process	🟥	GO:0042327	Positive regulation of phosphorylation	3,9E-28
GO Biological Process	🟥	GO:0051093	Negative regulation of developmental process	1,01E-24
GO Biological Process	🟧	GO:0070848	Response to growth factor	4,61E-24
GO Biological Process	🟧	GO:1901698	Response to nitrogen compound	2,46E-23

Figure 8: **Cluster 2 ADCN∩HDCN**, this cluster contained the protein APOE. The eight processes with the lowest FDR value from the STRING enrichment analysis can be seen.

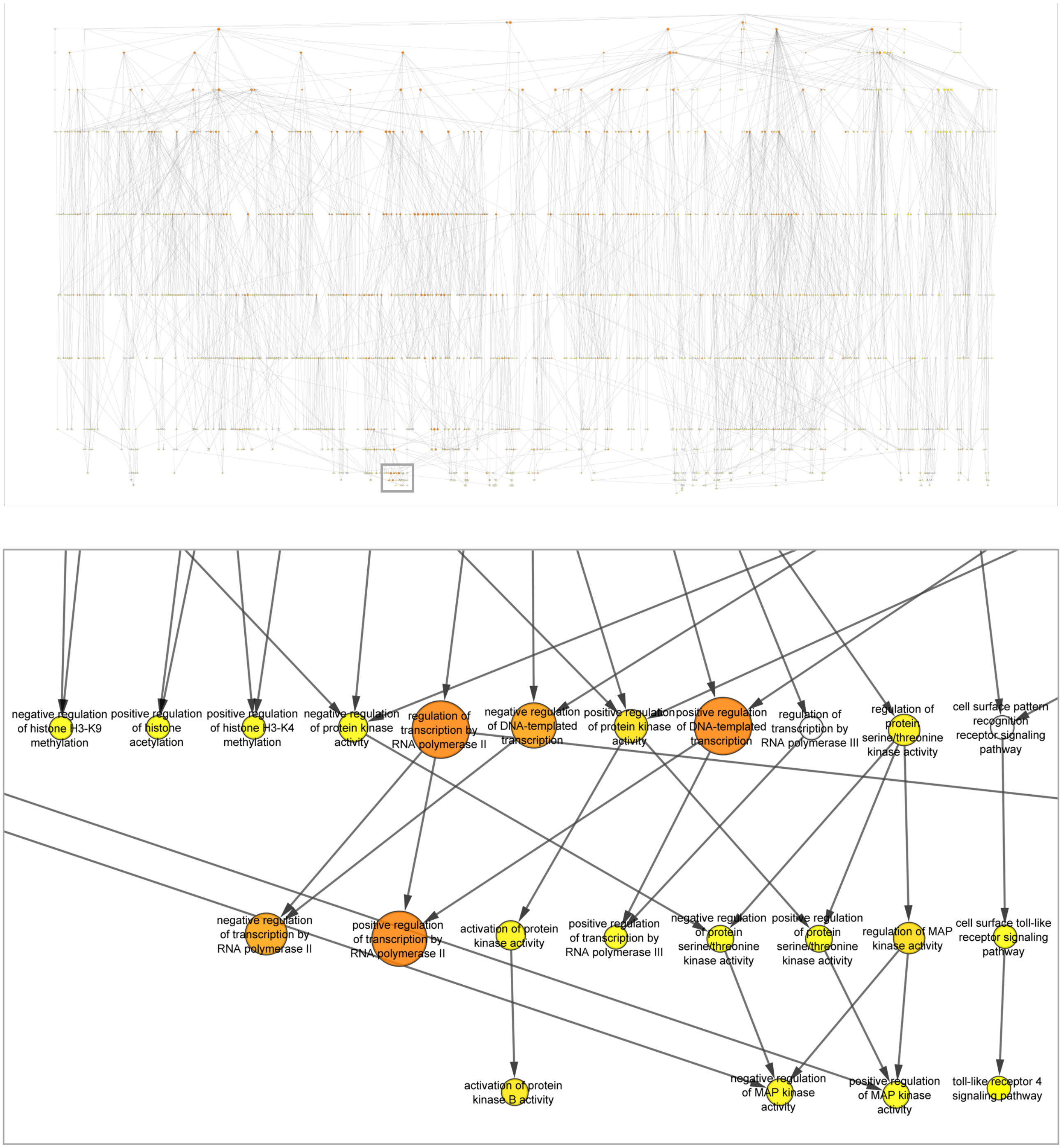
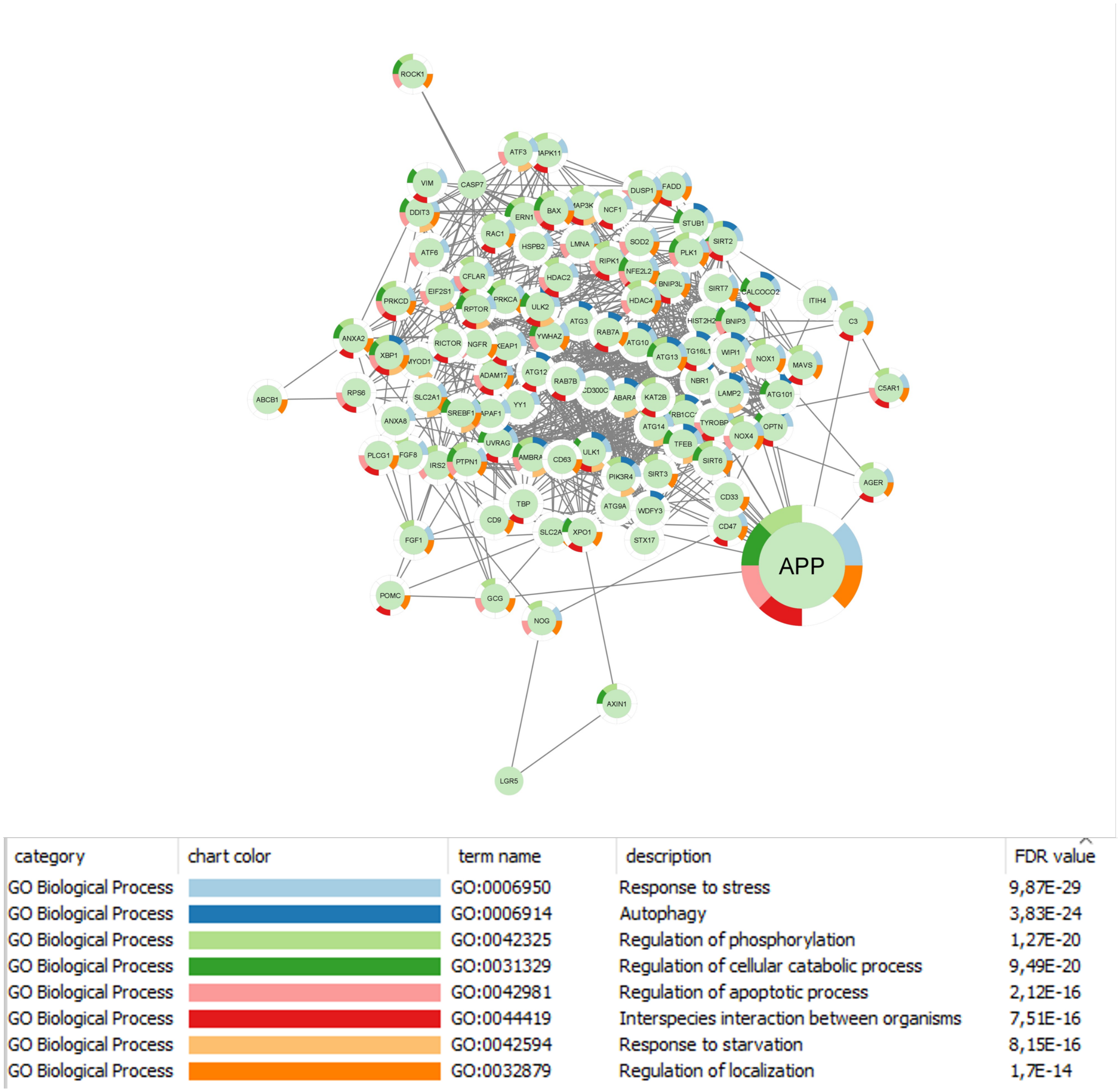

negative regulation of histone H3-K9 methylation
positive regulation of histone acetylation
positive regulation of histone H3-K4 methylation
negative regulation of protein kinase activity
regulation of transcription by RNA polymerase II
negative regulation of DNA-templated transcription
positive regulation of protein kinase activity
positive regulation of DNA-templated transcription
regulation of transcription by RNA polymerase III
regulation of protein serine/threonine kinase activity
cell surface pattern recognition receptor signaling pathway
negative regulation of transcription by RNA polymerase II
positive regulation of transcription by RNA polymerase II
activation of protein kinase activity
positive regulation of transcription by RNA polymerase III
negative regulation of protein serine/threonine kinase activity
positive regulation of protein serine/threonine kinase activity
regulation of MAP kinase activity
cell surface toll-like receptor signaling pathway
activation of protein kinase B activity
negative regulation of MAP kinase activity
positive regulation of MAP kinase activity
toll-like receptor 4 signaling pathway

Figure 9: **BiNGO graph Cluster 2 ADCN∩HDCN**, part of the BiNGO graph is enlarged, where it can be seen that the GO Biological Process *positive regulation of transcription by RNA polymerase II* is significantly enriched.



category	chart color	term name	description	FDR value
GO Biological Process	🟦	GO:0006950	Response to stress	9,87E-29
GO Biological Process	🟦	GO:0006914	Autophagy	3,83E-24
GO Biological Process	🟩	GO:0042325	Regulation of phosphorylation	1,27E-20
GO Biological Process	🟩	GO:0031329	Regulation of cellular catabolic process	9,49E-20
GO Biological Process	🟥	GO:0042981	Regulation of apoptotic process	2,12E-16
GO Biological Process	🟥	GO:0044419	Interspecies interaction between organisms	7,51E-16
GO Biological Process	🟧	GO:0042594	Response to starvation	8,15E-16
GO Biological Process	🟧	GO:0032879	Regulation of localization	1,7E-14

Figure 10: **Cluster 4 ADCN∩HDCN**, this cluster contained the protein APP. The eight processes with the lowest FDR value from the STRING enrichment analysis can be seen

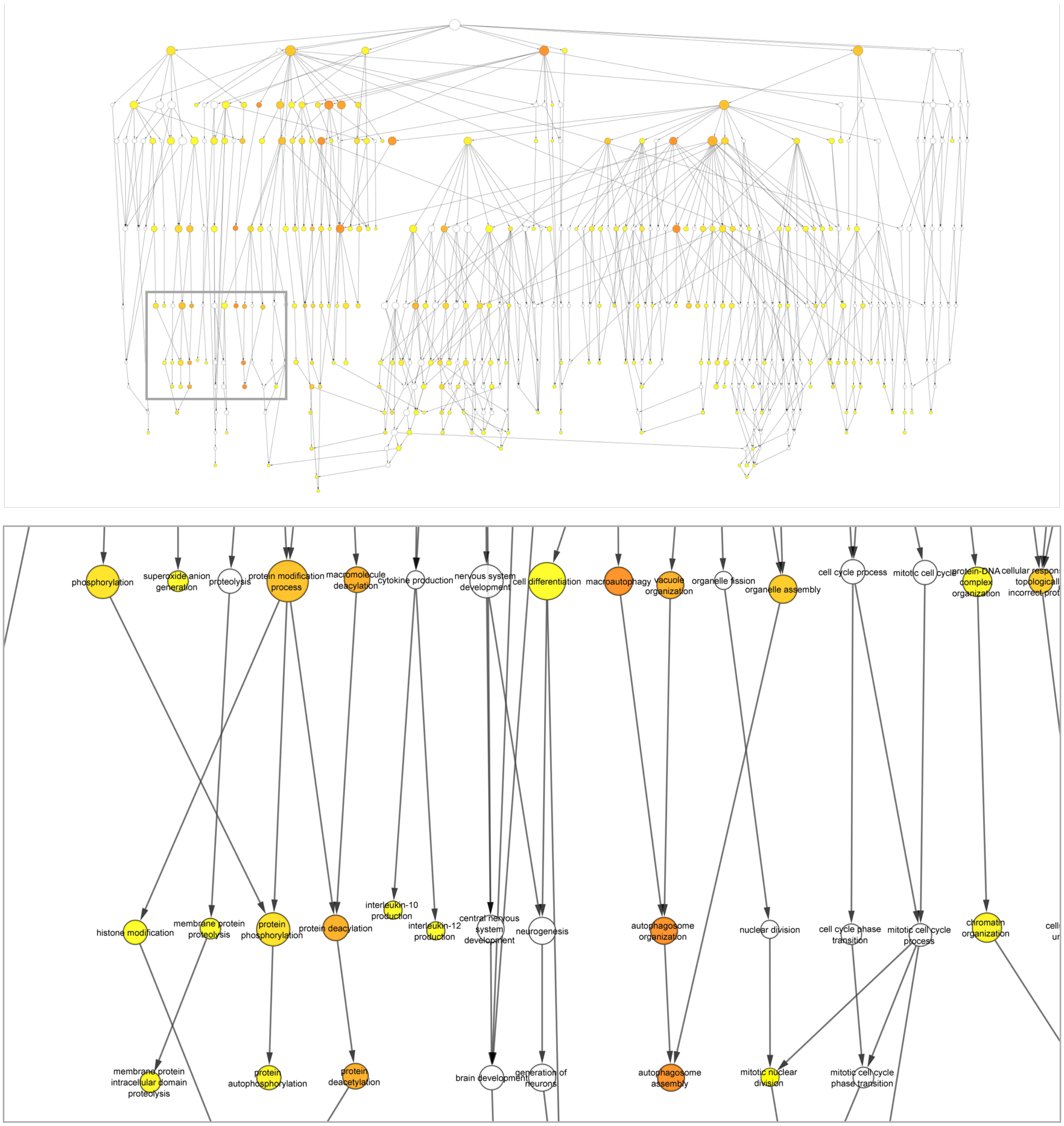
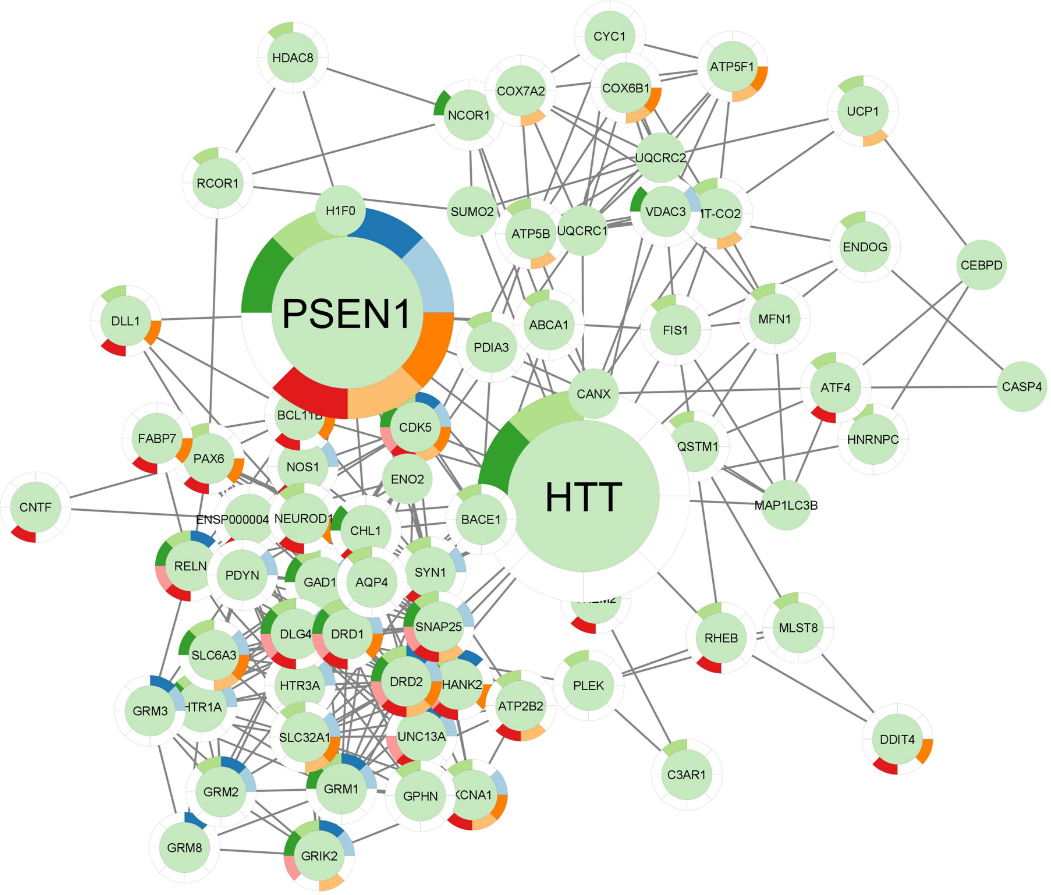
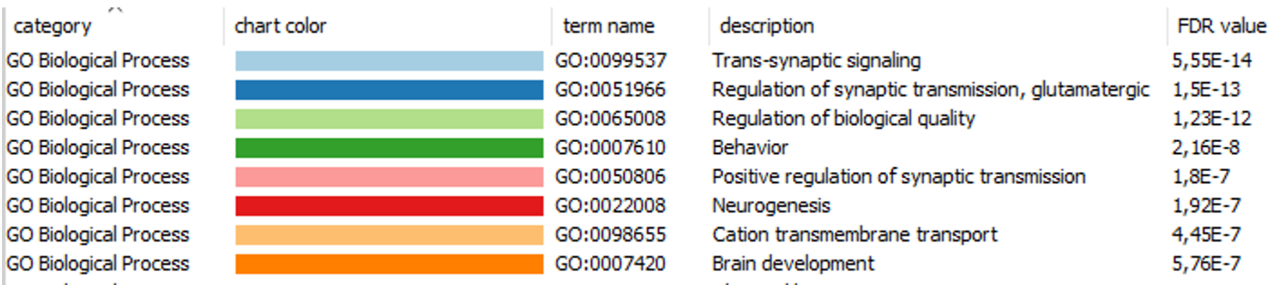

phosphorylation
superoxide anion generation
proteolysis
protein modification process
macromolecule deacylation
cytokine production
nervous system development
cell differentiation
macroautophagy
vacuole organization
organelle fission
organelle assembly
cell cycle process
mitotic cell cycle
protein-DNA complex organization
cellular respons
topologicall
incorrect prot
interleukin-10 production
histone modification
membrane protein proteolysis
protein phosphorylation
protein deacylation
interleukin-12 production
central nervous system development
neurogenesis
autophagosome organization
nuclear division
cell cycle phase transition
mitotic cell cycle process
chromatin organization
cell
ur
membrane protein intracellular domain proteolysis
protein autophosphorylation
protein deacetylation
brain development
generation of neurons
autophagosome assembly
mitotic nuclear division
mitotic cell cycle phase transition

Figure 11: **BiNGO graph Cluster 4 ADCN∩HDCN**, part of the BiNGO graph is enlarged, where it can be seen that the GO Biological Process *autophagosome assembly* is significantly enriched





category	chart color	term name	description	FDR value
GO Biological Process	🟦	GO:0099537	Trans-synaptic signaling	5,55E-14
GO Biological Process	🟦	GO:0051966	Regulation of synaptic transmission, glutamatergic	1,5E-13
GO Biological Process	🟩	GO:0065008	Regulation of biological quality	1,23E-12
GO Biological Process	🟩	GO:0007610	Behavior	2,16E-8
GO Biological Process	🟥	GO:0050806	Positive regulation of synaptic transmission	1,8E-7
GO Biological Process	🟥	GO:0022008	Neurogenesis	1,92E-7
GO Biological Process	🟧	GO:0098655	Cation transmembrane transport	4,45E-7
GO Biological Process	🟧	GO:0007420	Brain development	5,76E-7

Figure 12: **Cluster 6 ADCN∩HDCN**, this cluster contained the proteins PSEN1 and HTT. The eight processes with the lowest FDR value from the STRING enrichment analysis can be seen

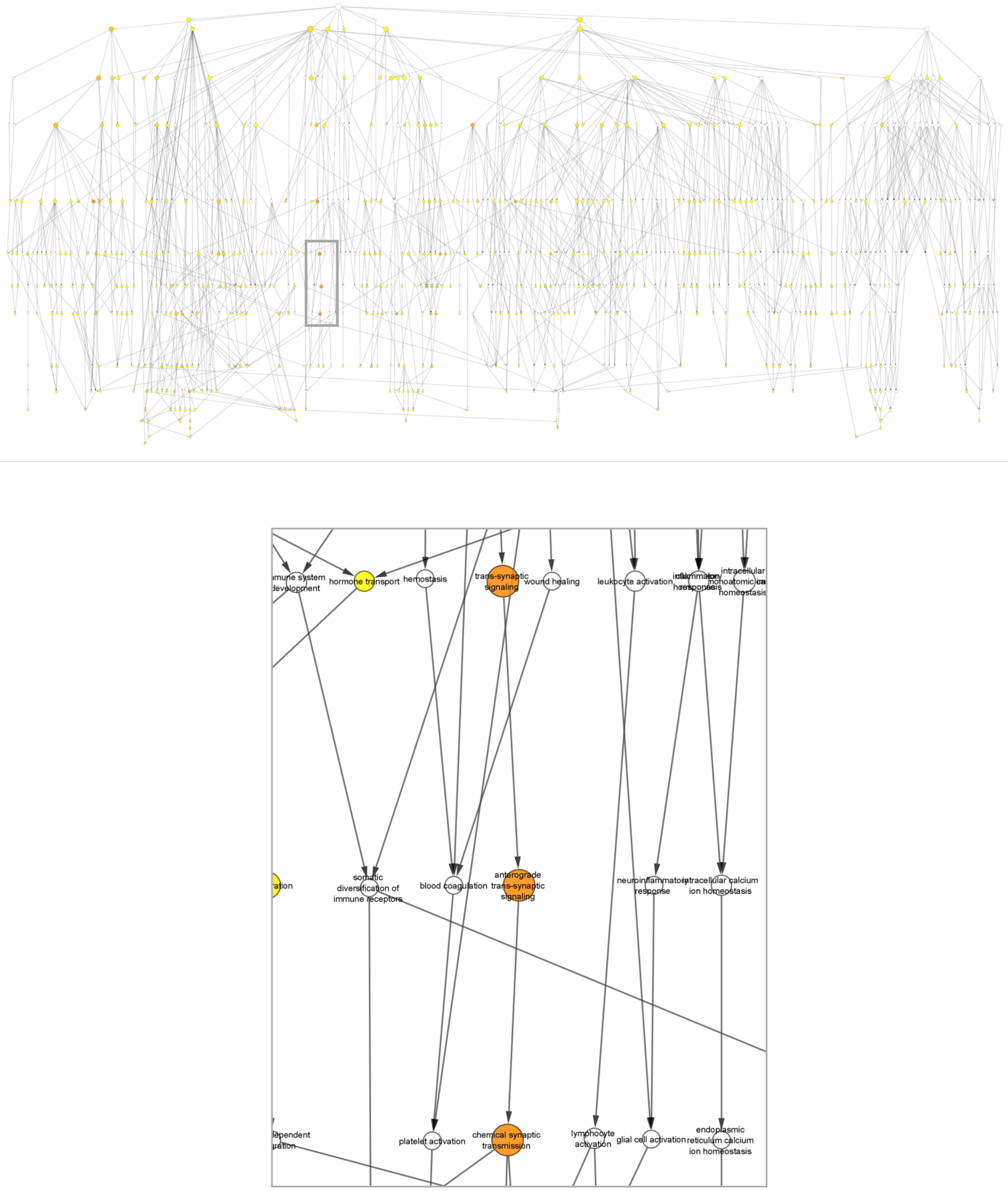
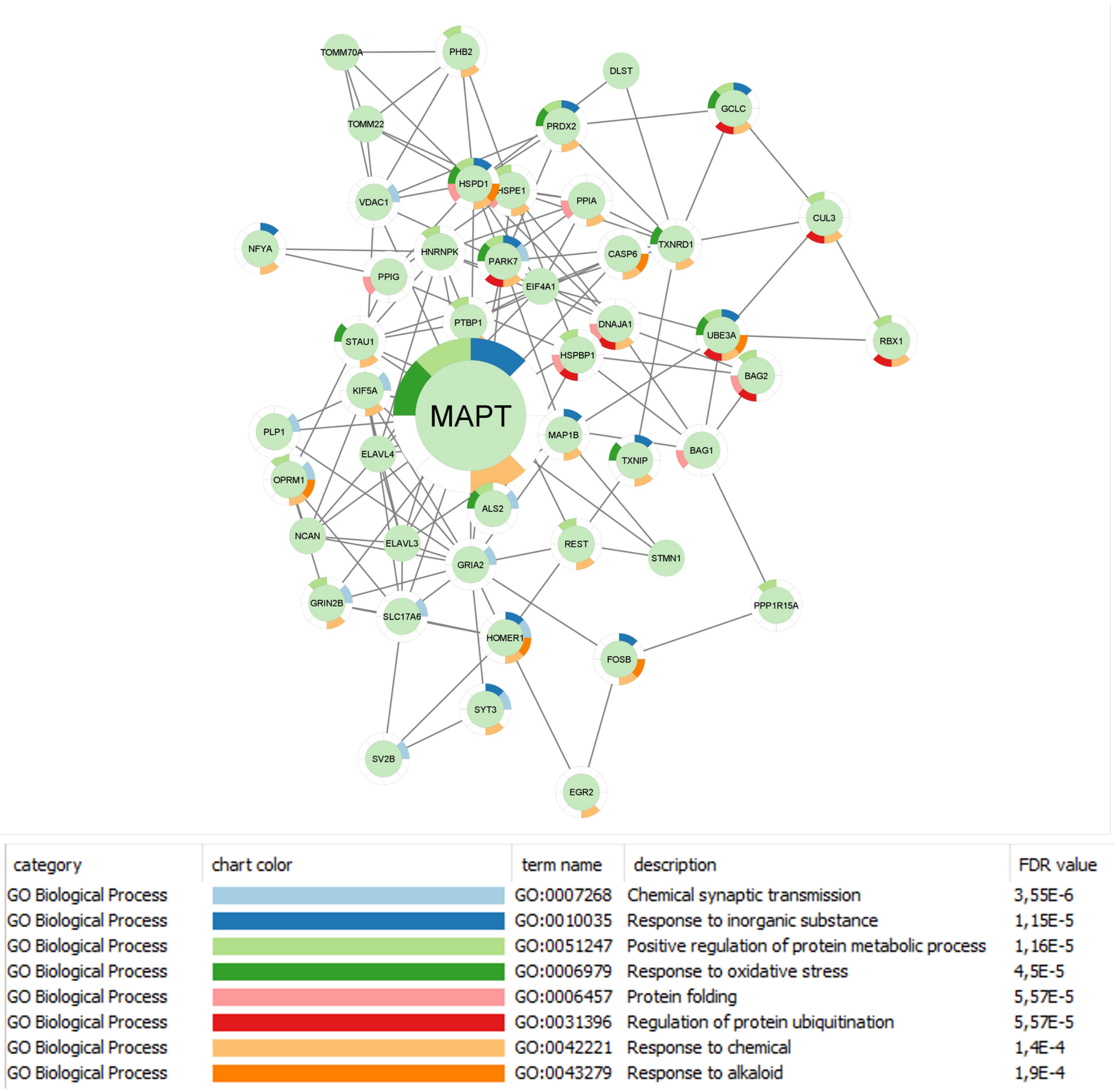

mune system development
hormone transport
hemostasis
trans-synaptic signaling
wound healing
leukocyte activation
inflammatory response
ion homeostasis
intracellular monoatomic ca homeostasis
ation
somatic diversification of immune receptors
blood coagulation
anterograde trans-synaptic signaling
neuroinflammatory response
intracellular calcium ion homeostasis
ependent ration
platelet activation
chemical synaptic transmission
lymphocyte activation
glial cell activation
endoplasmic reticulum calcium ion homeostasis

Figure 13: **BiNGO graph Cluster 6 ADCN∩HDCN**, part of the BiNGO graph is enlarged, where it can be seen that the GO Biological Process *chemical synaptic transmission* is significantly enriched.



category	chart color	term name	description	FDR value
GO Biological Process	🟦	GO:0007268	Chemical synaptic transmission	3,55E-6
GO Biological Process	🟦	GO:0010035	Response to inorganic substance	1,15E-5
GO Biological Process	🟩	GO:0051247	Positive regulation of protein metabolic process	1,16E-5
GO Biological Process	🟩	GO:0006979	Response to oxidative stress	4,5E-5
GO Biological Process	🟥	GO:0006457	Protein folding	5,57E-5
GO Biological Process	🟥	GO:0031396	Regulation of protein ubiquitination	5,57E-5
GO Biological Process	🟧	GO:0042221	Response to chemical	1,4E-4
GO Biological Process	🟧	GO:0043279	Response to alkaloid	1,9E-4

Figure 14: **Cluster 8 ADCN∩HDCN**, this cluster contained the protein MAPT. The eight processes with the lowest FDR value from the STRING enrichment analysis can be seen.

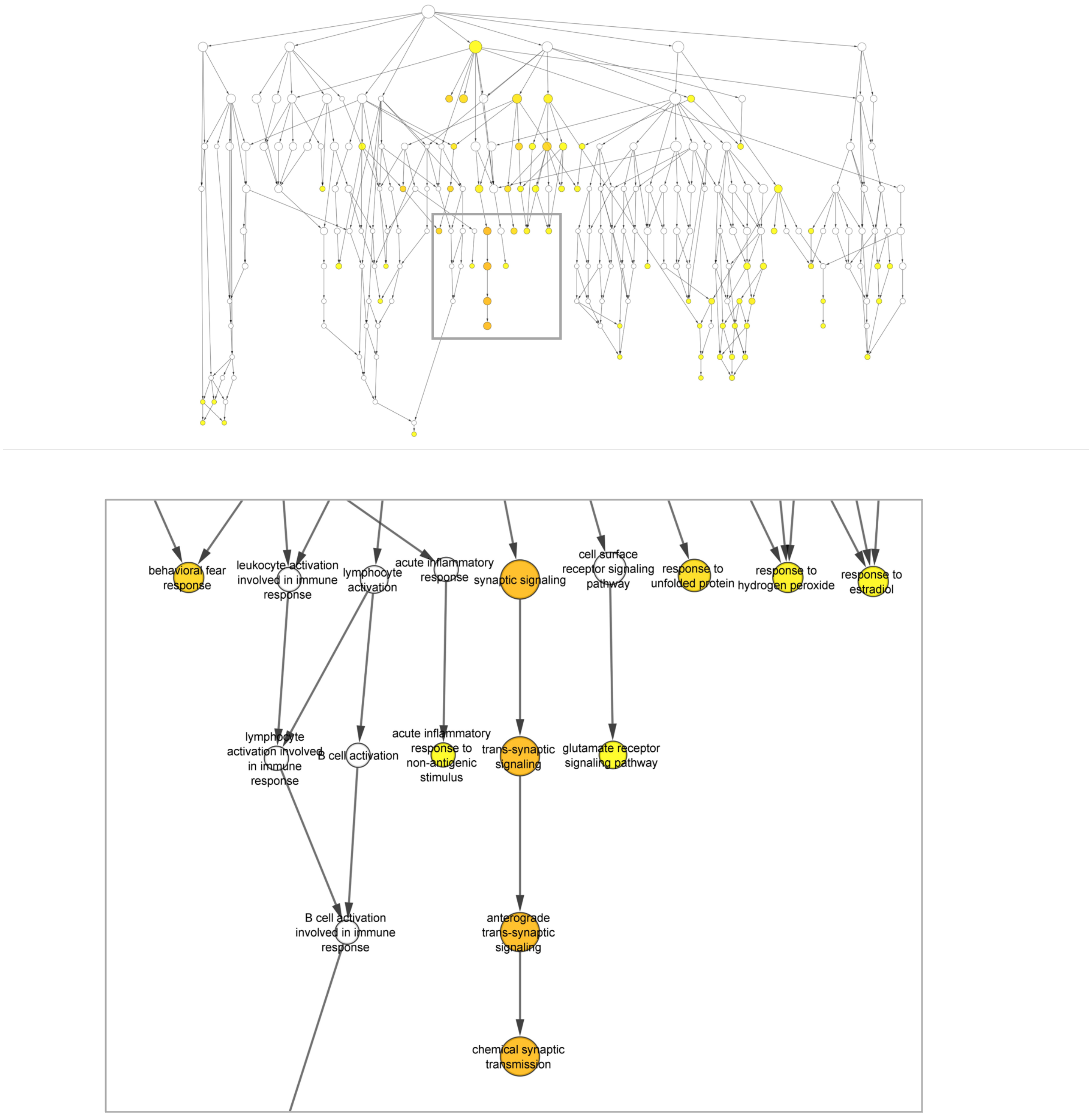

behavioral fear response
leukocyte activation involved in immune response
lymphocyte activation
acute inflammatory response
synaptic signaling
cell surface receptor signaling pathway
response to unfolded protein
response to hydrogen peroxide
response to estradiol
lymphocyte activation involved in immune response
B cell activation
acute inflammatory response to non-antigenic stimulus
trans-synaptic signaling
glutamate receptor signaling pathway
B cell activation involved in immune response
anterograde trans-synaptic signaling
chemical synaptic transmission

Figure 15: **BiNGO graph Cluster 8 ADCN∩HDCN**, part of the BiNGO graph is enlarged, where it can be seen that the GO Biological Process *chemical synaptic transmission* is significantly enriched