Mutant Huntingtin Protein Interaction Map Implicates Dysregulation of Multiple Cellular Pathways in Neurodegeneration of Huntington’s Disease

Sonia Podvina, Sara Brin Rosenthal, William Poon, Enlin Wei, Kathleen M. Fisch and Vivian Hook

Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA, USA
Center for Computational Biology & Bioinformatics, University of California, San Diego, La Jolla, CA, USA
Department of Obstetrics, Gynecology & Reproductive Sciences, University of California, San Diego, La Jolla, CA, USA
Department of Neuroscience and Dept of Pharmacology, School of Medicine, University of California, San Diego, La Jolla, CA, USA

Pre-press 16 July 2022

Abstract.
Background: Huntington’s disease (HD) is a genetic neurodegenerative disease caused by trinucleotide repeat (CAG) expansions in the human HTT gene encoding the huntingtin protein (Htt) with an expanded polyglutamine tract.

Objective: HD models from yeast to transgenic mice have investigated proteins interacting with mutant Htt that may initiate molecular pathways of cell death. There is a paucity of datasets of published Htt protein interactions that include the criteria of 1) defining fragments or full-length Htt forms, 2) indicating the number of poly-glutamines of the mutant and wild-type Htt forms, and 3) evaluating native Htt interaction complexes. This research evaluated such interactor data to gain understanding of Htt dysregulation of cellular pathways.

Methods: Htt interacting proteins were compiled from the literature that meet our criteria and were subjected to network analysis via clustering, gene ontology, and KEGG pathways using rigorous statistical methods.

Results: The compiled data of Htt interactors found that both mutant and wild-type Htt interact with more than 2,971 proteins. Application of a community detection algorithm to all known Htt interactors identified significant signal transduction, membrane trafficking, chromatin, and mitochondrial clusters, among others. Binomial analyses of a subset of reported protein interactor information determined that chromatin organization, signal transduction and endocytosis were diminished, while mitochondria, translation and membrane trafficking had enriched overall edge effects.

*Correspondence to: Dr. Vivian Hook, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9500 Gilman Dr. MC0657, La Jolla, CA 92037-0657, USA. E-mail: vhook@ucsd.edu.
Conclusion: The data support the hypothesis that mutant Htt disrupts multiple cellular processes causing toxicity. This dataset is an open resource to aid researchers in formulating hypotheses of HD mechanisms of pathogenesis.

Keywords: Huntingtin, mutant, wild-type, protein interactors, networks, translation, signaling, mitochondria, chromatin, trafficking, chromatin

INTRODUCTION

Huntington’s disease (HD) is a progressive neurodegenerative disease caused by CAG triplet repeat expansion mutations in exon 1 of the HTT gene that is autosomal dominant [1–6]. Neurodegeneration in the brain, and especially in striatum, causes a characteristic involuntary movement disorder, chorea, but patients also suffer from deficits in cognitive and psychiatric functions, as well as weight loss and other symptoms resulting from neurodegeneration in brain [1–8]. HD is fatal, and there is a need for drugs that prevent or delay neurodegeneration [9]. HD is a spectrum disorder whereby the CAG repeat length mutation varies among patients. HD develops in adulthood in patients with 40–50 CAG repeats, and longer expansions of more than 60 CAG repeats causes progressively younger age-of-onset and faster HD disease progression [10, 11]. Individuals with expansions between 36 and 39 may or may not develop HD [1–6, 11, 12]. The normal CAG range in humans is 15–21, with a median of 18 [13, 14].

The HTT gene mutation in HD patients was identified in 1993 [1] and has been intensely studied. Yet, gaps remain in our understanding of how the HTT gene product, the mutant huntingtin (Htt) protein, causes neuronal cell toxicity and death leading to characteristic neurodegeneration, disease symptoms, and fatality [1–6, 11, 15]. Htt is a large protein [1, 16] of approximately 3,044 amino acids with varying polyglutamine (polyQ(n)) lengths encoded by the spectrum of CAG repeats. The expanded polyQ(n) peptide sequence domain follows the first 17 amino acids at the amino terminus of Htt. An ongoing challenge in the field has been to determine the cellular functions of the Htt protein and how the expanded polyQ(n) leads to dysfunction in cells to cause repeat-dependent neurodegeneration in HD. The functions of wild-type (wt) Htt are also not fully understood. Htt is processed into multiple protein fragments that are localized to different cellular compartments and may participate in different cellular pathways [17–19].

To address the question of how mutant Htt initiates the HD disease process, researchers have sought to identify proteins that interact with Htt that may potentially reveal the molecular pathways affected by aberrant protein interactions with mutant Htt. Reports over the last 25 years have identified numerous proteins that interact with mutant Htt and wild-type (wt) Htt, involving molecular, cellular, and animal HD model systems using a variety of protein technologies to identify Htt interactors. These extensive data can provide valuable insight into potential cellular functions of the Htt protein. However, a gap in the field is that a complete dataset of all of these potential Htt protein interactors that defines the Htt lengths and polyglutamine expansion lengths of native Htt protein complexes has not yet been fully curated and analyzed. Therefore, the goals of this study were to 1) create a database of all of these Htt protein interactors reported in the literature that meet the criteria of indicating the Htt fragments lengths or full-length Htt, number of glutamines within the Htt form, and analysis of native Htt protein complexes, 2) conduct molecular network analyses to reveal potential cellular pathways of Htt protein interactors that are enriched with expanded poly(Q)n, and 3) provide this dataset and analysis to the public. Results show that both mutant and wt Htt interact with a large number of proteins (2,740 for control wt Htt, and 2,631 for mutant Htt in HD); further, a majority of these protein interactors are shared by wt and mutant Htt forms. Mutant and wt Htt protein interactors cluster into twelve gene ontology (GO) biological functions; several clusters of interactors are significantly enriched or diminished with expanded mutant Htt poly(Q)n which include protein translation, signal transduction, membrane trafficking, and chromatin organization. Assessment of cluster components by KEGG pathway analysis revealed significant inclusion of Htt interactors in functional systems of mitochondria, RNA splicing, and protein modification by the ubiquitin-proteasome pathway. These findings support the hypothesis that Htt has multiple functions in the cell that can potentially be disrupted with CAG expansion mutation of the HTT gene. This new publicly available dataset of wt and mutant Htt protein interactors can enhance
future experimental investigation of the functional consequences of Htt protein interactions in the HD disease process.

**METHODS**

**Acquisition of primary research publications**

PubMed was searched for primary research articles that identified Htt interacting proteins by searching for “huntingtin protein interaction” combined with curation of individual articles for direct identification of proteins interacting with Htt protein forms. Articles were excluded if (i) the study did not evaluate a protein-protein interaction, (ii) the report was a review article and did not report original data, (iii) the article was not about HD or Htt but the article came up in the search, (iv) the article showed that a protein did not interact with Htt, (v) the study used only short synthetic motifs (such as only polyQ) but not a protein segment containing Htt primary sequences. After curation, 193 articles were compiled that investigated proteins interacting with Htt; these articles are provided as references [20–212] (in alphabetical order within the set of citations).

**Dataset recording**

Information from the 193 articles [20–212], meeting the above criteria for studies of Htt interacting proteins, were compiled for analysis with respect to (i) the number of polyQ(n) residues in the Htt protein form used in study, (ii) the species of Htt (human, rodent, or other), (iii) whether the Htt protein utilized was full-length or a fragment, (iv) the known Htt interactor protein used as “bait” in the study if not Htt, (iv) the species of the interactor protein(s) identified, (v) the assays and techniques used to identify the protein interactor, (vi) information about the identified protein interactor with respect to its full name, official gene symbol, and accession number, (vii) how the polyQ(n) expansion affects the interaction between Htt and the interactor protein, if known, ‘+’ indicates more abundant levels interacting with mutant Htt than wt Htt, and ‘-‘ indicates less abundant levels interacting with mutant Htt than wt Htt, and ‘=‘ indicates similar levels, and if blank was not determined), (viii) the age of the model organism used, (ix) article reference information, (x) the protein function assigned based on the authors’ conclusions from the study, or if not indicated (e.g., was from a large dataset) the protein function was based on general cell biological knowledge (verified by genecards [213] or uniprot [214], (xi) if the article tested functional consequences of the interaction between Htt and the protein interactor in a cellular or animal model system and the authors’ conclusion, and finally (xii) if the study tested the interaction in human HD and/or control donor cells or tissue, and the functional consequence of the interaction. This Htt protein interaction data is compiled in Supplementary Table 1 (tab for Master Table).

**Clustering analysis of Htt interactors**

Among reported Htt protein interactors, we sought to identify groups of proteins that shared functional categories, known as “clusters,” of different molecular and cellular pathways. To accomplish this, Htt protein interactors were integrated with the STRING human database [215] to gain information about known connections among interactors, and facilitate analysis of highly connected gene sets, representing the protein sets of this study, to determine the likelihood of an interaction (“edge”) between proteins in the dataset. STRINGdb retrieved information from these types of open source data with respect to (i) genomic/protein context predictions, (ii) high-throughput lab experiments, (iii) gene/protein co-expression, (iv) automated text mining, and (v) previous knowledge in databases. STRING edge confidence assessment selected only edges with high confidence values (>700). An Htt + STRING subnetwork was built from these edges. To allow for reliable clustering, genes were included in the subnetwork if they are in the Htt dataset and have at least 2 neighbors in the background STRING interactome. Next, we quantified structure in Htt + STRING subnetwork by applying Louvain modularity maximization [216] clustering algorithm to break up nodes into highly connected sets related to protein function. These clusters were annotated for biological function by functional enrichment with gene ontology (GO) terms [217, 218] and KEGG [219] pathways, using ToppGene [220]. Clusters were annotated with the most significantly enriched term/pathway. Htt protein interactors in each cluster is provided in Supplementary Table 1.

Select enriched KEGG pathways were examined further. Interactor proteins in significant KEGG pathways were mapped onto KEGG pathway diagrams [219]. HTT interactor proteins were color coded in these diagrams for the effect of the mutant poly(Q) expansion on the interaction of red = more abundant,
Fig. 1. Work-flow for bioinformatics analysis of huntingtin (Htt) interacting proteins reported in the literature. Literature searches for proteins interacting with mutant Htt and wt Htt were conducted with the PubMed resource (a). Information about Htt interactors was organized by experimental parameters of the reported studies, including the polyQ(n) expansion length, Htt fragment length, and whether the interaction is increased or decreased with mutant polyQ(n) Htt fragments (b). Compiled Htt interacting proteins (Supplementary Table 1) were subjected to bioinformatics analysis assessing statistically significant clusters and KEGG pathways using tools that included STRINGdb, gene ontology, and ToppGene (c).

yellow = no reported effect, blue = less abundant, and purple = mixed reports of both more and less abundant interactions.

Data provided in approximately half of the articles tested whether the mutant Htt polyQ(n) expansion in the HD model resulted in higher or lower abundance levels of a particular protein interacting with mutant compared to wt Htt. These data were included in the Htt + STRING dataset as “edge effects.” The significance of these edge effects was evaluated within each cluster by a binomial test of the number of proteins with higher abundance interactions \(n_+\) and lower abundance \(n_-\), with a null hypothesis of 68% positive edges (the rate of positive edges in the full dataset). Significance is defined as \(p < 0.05\) with Benjamini-Hochberg multiple test correction.

RESULTS

Strategy for data acquisition of Htt interacting proteins and bioinformatic evaluation

Data for Htt interacting proteins was compiled from the literature [20–212] obtained by search of PubMed for bioinformatic analyses (Fig. 1). Htt interactors were organized into a dataset that indicates the conditions of identifying interactors with respect to Htt fragment length and species, polyglutamine (polyQ(n)) length of Htt, native Htt protein complexes, model system and its age, assay type for identifying interactors, interacting proteins identified, increased or decreased interaction with mutant Htt, and references (Supplementary Table 1). The Htt interactor dataset was evaluated by bioinformatics tools of STRINGdb, gene ontology, ToppGene, and KEGG pathway mapping with statistical scoring. This unique study compiled and curated articles reporting Htt protein interactors in the literature that meet our criteria that the interactor information includes (1) the lengths of Htt fragments or full-length Htt fragment used in the protein interaction studies, (2) the number of polyglutamines in the mutant and wild-type forms of Htt, and (3) native Htt protein interaction complexes were evaluated without general cross-linking agents, since such cross-linking can include non-specific associated proteins.

Compilation of Htt interacting proteins of mutant Htt and wild-type (wt) Htt

The compiled data of Htt protein interactors in HD animal models shows that a large number of proteins interact with both wild-type (wt) Htt and mutant Htt (Fig. 2). A PubMed search and cura-
Fig. 2. Protein counts for interactors of wt and mutant Htt. The numbers of distinct proteins found to interact with control wt Htt and HD mutant Htt are shown by a Venn diagram. Proteins were found that only interacted with mutant Htt or only with control wt Htt. Many proteins interacted with both wt and mutant Htt.

Table 1

<table>
<thead>
<tr>
<th>Cluster #</th>
<th>Cluster name</th>
<th># Positive edges</th>
<th># Negative edges</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Protein modification</td>
<td>33</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Chromatin organization</td>
<td>84</td>
<td>93</td>
<td>8.26E-04</td>
</tr>
<tr>
<td>3</td>
<td>RNA splicing</td>
<td>88</td>
<td>37</td>
<td>0.1043</td>
</tr>
<tr>
<td>4</td>
<td>Membrane trafficking</td>
<td>43</td>
<td>58</td>
<td>7.86E-04</td>
</tr>
<tr>
<td>5</td>
<td>Signal transduction</td>
<td>82</td>
<td>124</td>
<td>3.64E-09</td>
</tr>
<tr>
<td>6</td>
<td>Mitochondria</td>
<td>63</td>
<td>24</td>
<td>0.1029</td>
</tr>
<tr>
<td>7</td>
<td>Granule membrane</td>
<td>7</td>
<td>8</td>
<td>0.5500</td>
</tr>
<tr>
<td>8</td>
<td>Macrophagocytosis</td>
<td>7</td>
<td>7</td>
<td>0.6452</td>
</tr>
<tr>
<td>9</td>
<td>Cytoplasmic vesicle lumen</td>
<td>30</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>Clathrin-mediated endocytosis</td>
<td>19</td>
<td>30</td>
<td>6.00E-03</td>
</tr>
<tr>
<td>11</td>
<td>Ion channel transport</td>
<td>11</td>
<td>41</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>Translation</td>
<td>122</td>
<td>19</td>
<td>9.52E-10</td>
</tr>
</tbody>
</table>

FDR (false discovery rate) values are corrected for Benjamini-Hochberg. Positive edge (yellow): proteins interact more abundantly with mutant Htt. Negative edge (blue): proteins interact less abundantly with mutant Htt.

Parameters for identification of mutant Htt protein interactors

All studies utilized cellular, animal and in vitro non-human HD model systems to identify potential Htt interactors [20–212]. Some studies used HD tissue or cells in follow-up experiments to confirm interactions. Studies used fragments of Htt or full-length Htt, the majority of which were cloned human or mouse amino-terminal Htt containing the polyQ(n) region (Supplementary Table 1, Master Table). Within the non-pathological human polyQ(n) range of <40Q, the studies identified 2,740 protein interactions. Within the adult HD range of ≥40Q and <60Q, 469 Htt protein interactors were identified (Supplementary Table 1, Table 1). In the
juvenile range of $\geq 60Q$ to $\leq 120Q$, 681 proteins were identified. In the especially long expansion range of $> 120Q$, 1,835 proteins were identified (Supplementary Table 1). Analysis of the distribution of Htt interacting proteins identified by lengths of N-terminal Htt fragments show that Htt of 171 residues was used as the main Htt fragment among the published studies (Supplementary Figure 2).

While there is a clear inverse correlation in human HD age-at onset and polyQ(n) length that spans decades [11], for practical experimental designs, researchers used cellular models (age of < 1 week), or animal models with ages of 1 day (Drosophila larvae) to 24 months (transgenic and knockin mice) (Supplementary Table 1, Master Table). Animal model studies showed neurodegenerative and behavioral phenotypes with expression of mutant Htt, e.g., retinal degeneration in developing Drosophila, motor impairment in C. elegans, and motor and behavioral dysfunction and HD-like neuropathology in transgenic mice.

A variety of approaches were used to test Htt protein interactions using either hypothesis-driven or unbiased identification experimental designs. Hypothesis-driven studies tested approximately 10 or fewer potential Htt protein interactors, while unbiased approaches using Htt protein complex purification and mass spectrometry sequencing identified as many as approximately 1,000 proteins [16, 52, 87, 97, 140, 152]. Yeast 2-hybrid studies identified up to approximately 150 proteins [74, 155]. Approximately half of all potential Htt protein interactors were identified in studies that utilized strategies to determine whether proteins interacted at higher or lower levels with mutant compared to wt Htt (Supplementary Table 1, Master Table). Such results showed whether at protein interactor was present a higher or lower abundance in mutant Htt complexes, e.g., by a darker or lighter band on western blot, or by detection of higher or lower levels of a protein by mass spectrometry. These data were used in our study to determine significant enrichment of gene ontology (GO) clusters in mutant Htt compared to wt Htt.

**Clustering analysis correlates Htt protein interactors with diverse cellular functional roles that may be enriched or diminished in HD-like phenotypes**

We hypothesized that Htt protein interactors could be statistically sorted into functional categories to reveal information about the types of molecular pathways in the cell that are dysfunctional in HD, ultimately leading to cell toxicity and death. Accordingly, the list of Htt protein interactors was integrated with the STRING human database [215] (Search Tool for the Retrieval of Interacting Genes/Proteins) to gain knowledge of potential interactions (“edges”) between proteins. To allow for reliable clustering, genes were included in the subnetwork if they are in the Htt protein interactor dataset and have at least two neighbors in the background STRING interactome.

We then quantified structure in Htt + STRING subnetwork by applying Louvain modularity maximization clustering algorithm [216] to break up nodes into highly connected groups. Louvain modularity maximization is an optimized algorithm that runs on a scale of $-1$ to $1$ to determine the density of connections of possible clusters as the algorithm runs. Modularity is designated as the variable $Q$, the density of connections inside the community compared to links between communities. The resulting clusters are shown in Fig. 3 for mutant Htt and wt Htt. The twelve clusters indicated functions of 1) protein modification, 2) chromatin organization, 3) RNA splicing, 4) membrane trafficking, 5) signal transduction, 6) mitochondria, 7) granule membrane, 8) macroautophagy, 9) cytoplasmic vesicle, 10) clathrin-mediated endocytosis, 11) ion channel transport, and 12) translation. Several of these pathways have been investigated in the field for HD cell death mechanisms including mitochondria, macroautophagy, protein modification, and protein translation which significantly clustered in our analysis.

In the studies that reported differing levels of a protein interacting with mutant Htt compared to wt Htt, we conducted a statistical test for proteins within each cluster in the mutant Htt interactor dataset to determine whether an interactor protein interacted with mutant Htt at significantly greater or lower abundance levels, that is, if a protein was reported to interact with higher or lower abundance with mutant Htt, or the interaction was lost or gained compared to wt Htt protein interactors. We analyzed these data with a binomial test ($N = n_+ + n_-$) for each cluster to determine the probability of observing $n_+$ higher abundance or $n_-$ lower abundance interactions in the mutant Htt interactor dataset out of the total interactions, $N$ in each cluster. Results (Table 1) showed that the function of translation was significantly enriched in mutant Htt interactions, while signal transduction, chromatin organization, membrane trafficking, and
Fig. 3. Clustering analysis of Htt interactors by gene ontology (GO). Htt interactors were analyzed for shared functional categories, “clusters,” in different cellular molecular pathways. The cluster networks are illustrated in color-coded format (see key in figure) for interactions that were lost by mutant Htt (mHtt), weaker with mHtt, stronger with mHtt, gained interaction with mHtt, or represented a neighbor. clathrin-mediated endocytosis clusters were significantly diminished.

Clustered Htt interacting protein identifications achieved by multi-disciplinary techniques

Various assay methods and techniques were used to identify Htt interactor proteins. The number of interactor proteins in each cluster identified by assay type are displayed in a heatmap (Fig. 4). Larger numbers of proteins were identified by high-throughput approaches, particularly methods that purified Htt complexes from tissue or cell lysate, followed by identification of interactors by mass spectrometry (Affinity-MS/MS) and yeast 2-hybrid. Other assays identified interactors by hypothesis-driven approaches; numerous articles utilized immunopurification of Htt complexes, followed by western blot to detect a proposed interactor. Additional studies used co-immunohistochemistry to detect physical proximity between Htt and an interactor protein. Fewer articles utilized in silico or structural analysis approaches, or other assay types. Higher numbers of interactors were identified by affinity-MS/MS, yeast 2-hybrid, and affinity-western blot assays in the chromatin organization, signal transduction, mitochondria, and translation clusters.

Htt poly(Q)n repeat lengths used for identification of clusters of Htt interacting proteins

We next queried the data to determine how many proteins in each cluster were identified by varying polyQ(n) repeat length, assessed in lengths of increasing 10Q (Fig. 5). The largest numbers of proteins were identified in the wt range of < 30Q, and in the very high juvenile range of > 90Q. Notably, fewer interactor proteins were identified with polyQ(n) Htt lengths found in human adult HD of 40–60Q and juvenile HD of 60–120Q, likely due to fewer studies using Htt with these ranges of polyQ(n). HD model systems, particularly genetic mouse models, frequently used especially high polyQ(n) repeat lengths to induce HD pathology and phenotypes within the lifespan of a mouse of less than two years [52, 97, 167]. It will be of interest in future studies to identify Htt interactors within the human Htt polyQ(n) range in HD because proteins that interact with very high polyQ(n) mutant
Htt may or may not interact with mutant Htt present in human HD brain.

It was also found that the majority of studies used N-terminal fragments of mutant Htt and wt Htt containing the polyQ(n) domain for determination of Htt protein interactors (Supplementary Figure 2). Models using N-terminal fragments of 171 amino acids in length were prevalent; the 171 residue fragment corresponds to exons 1–3 [1, 221].

Cellular pathways within mutant Htt interactor clusters reveal potential molecular toxicity mechanisms

The most significant gene ontology (GO) pathways within each cluster were identified by ToppGene [220] with statistical evaluation. All significant pathways for each cluster are provided in Supplementary Table 1 (Master Table, tabs for clusters 1–12).
Kyoto Encyclopedia of Genes and Genomes (KEGG) curates proteins in well-defined cellular pathways [219]. The significant GO terms for each cluster were filtered for KEGG pathways. A summary of the top KEGG pathways for each cluster is provided in Supplementary Table 2. The most significant KEGG pathways for the clusters of translation, signal transduction, chromatin organization, and membrane trafficking are described with respect to Htt interactor components below.

**Protein translation**

Cluster 12 for translation was identified by binomial analysis as the cluster with the most number of positive edges (Table 1). Ribosome was the most significant pathway within the translation cluster (Supplementary Table 2), meaning that protein interactors in this cluster overall had more abundant interactions with mutant compared to wt Htt (Fig. 6). Elongation factors within both the large and small subunits and aminoacyl tRNA transferases interact at higher levels with mutant compared to wt Htt. These results are consistent with a study showing that mutant Htt impedes ribosomal translocation during translation elongation and suppresses protein synthesis [222]. These data together indicate the ribosome as potentially impacted through mutant Htt protein interactions.

**Signal transduction**

Cluster 5 for signal transduction had significantly more negative edges as determined by binomial analysis (Table 1), indicating that protein interactors in this cluster overall had less abundant interactions with mutant Htt compared to wt Htt. The most significant KEGG pathways for cluster 5 (Supplementary Table 2) was the RAP1 signaling pathway (Fig. 7). The RAP1 signaling pathway has not been described as a pathway of cell toxicity and death in HD. Future investigations of this pathway mechanism may reveal new disease mechanisms and potential therapeutic targets.

**Chromatin organization**

Cluster 2 for chromatin organization had significantly more negative edges according to binomial
Fig. 7. Rap1 signaling KEGG pathway of the signal transduction cluster. The Rap1 signaling pathway represented the top significant pathway of the signal transduction cluster (Supplementary Table 2). Preferred interactions of mutant Htt (mHtt) or wt Htt with pathway components are illustrated by the red to blue color color key.

analysis (Table 1). The most significant KEGG pathways within this cluster (Supplementary Table 2) is glycerophospholipid metabolism (Fig. 8). Mutant Htt has aberrant interactions with glycerophospholipids, and biosynthetic and metabolic processes are disrupted [223]. Further investigations into Htt interacting proteins in glycerophospholipid pathways could hold promise for better understanding of mutant Htt associated membrane disruption.

Membrane trafficking

Cluster 4 for membrane trafficking also had significantly more positive edges (Table 1). The most significant KEGG pathway of cluster 3 was SNARE interactions in vesicular transport (Fig. 9 and Supplementary Table 2). Disruption of SNARE proteins in astrocytes by mutant Htt cause impaired gliotransmission associated with behavioral disruptions in HD model mice [224]. Evaluation of Htt protein interactors associated with SNARE vesicular transport may reveal key mechanisms underlying psychiatric and behavioral changes in HD.

Significant KEGG pathways of cluster functions of Htt protein interactors

Mitochondria

Among KEGG pathways assessed, oxidative phosphorylation of mitochondria represents a highly significant cellular system of Htt interacting proteins (Fig. 10 and Supplementary Table 2). Among 71 components of the oxidative phosphorylation KEGG pathway, 70 of these proteins were found as Htt interactors. Mutant Htt forms displayed preferences for interacting with components of oxidative phosphorylation mechanisms of mitochondria. Htt interactors of the mitochondria cluster were also represented by KEGG pathways of metabolism, carbon metabolism, and thermogenesis with oxidative phosphorylation as the top pathways in this cluster.
Fig. 8. Glycerophospholipid metabolism KEGG pathway of the chromatin cluster. The Glycerophospholipid metabolism pathway of the chromatin organization cluster was found to be most significant (Supplementary Table 2). Preferred interactions of mutant Htt (mHtt) or wt Htt with pathway components are illustrated by red to blue color key.
Fig. 9. SNARE interactions in vesicular transport KEGG pathway of the membrane trafficking cluster. The SNARE interactions for vesicular transport of the membrane trafficking cluster was highly significant with respect to interactions with Htt (Supplementary Table 2). Preferred interactions of mutant Htt (mHtt) or wt Htt with pathway components are illustrated by red to blue color key.

(Fig. 11 and Supplementary Table 2). Mitochondrial dysfunctions that compromise energy production are consistent with weight loss that occurs in HD [11].

RNA splicing and protein modification

The RNA splicing cluster (#3) of Htt interactors includes significant KEGG pathways of the spliceosome and RNA transport (Fig. 11 and Supplementary Table 2). Dysregulated RNA splicing occurs in human HD brain and in HD models and variant splicing of human HTT RNA has been observed [225, 226].

The protein modification cluster of Htt interactors (cluster 1 shown in Fig. 11) contains significant KEGG pathways for ubiquitin-mediated proteolysis and the proteasome (Supplementary Table 2). Dysregulation of the ubiquitin-proteasome system in HD participates in the clearance of unwanted proteins, resulting in altered protein homeostasis [227, 228].

Htt interactor dataset available to the public

This dataset of Htt interactor proteins has provided a comprehensive dataset of published reports of wt and mutant Htt protein interactors in non-human HD model systems (provided in Supplementary Table 1) with information on length of Htt fragment or full-length Htt, length of poly-glutamine within Htt, and use of native Htt protein complexes. This Htt interactor data set is publicly available at the website shown at the end of the discussion. A variety of non-human model systems and assays were used to determine potential Htt protein interactors, and the systems used did not affect the associated functions of the proteins identified. Bioinformatics analyses reveal that significant modularity clusters in functional classes of protein translation, signal transduction, chromatin organization, mitochondria, and others that may be enriched or diminished with polyQ(n) expansion in mutant Htt. Access to these data will allow researchers in the field to formulate hypotheses and plan studies to investigate the key roles of Htt pro-
Fig. 10. Oxidative phosphorylation KEGG pathway of the mitochondrial cluster. The oxidative phosphorylation pathway of mitochondria was of high significance for interactions with Htt (Supplementary Table 2). Preferred interactions of mutant Htt (mHtt) or wt Htt with pathway components are illustrated by red to blue color key.
Fig. 11. Top significant KEGG pathways within each cluster of Htt interactor functions. The top five most significant KEGG pathways for each of the 12 clusters of Htt interactor functions are illustrated. The 12 clusters are shown in panels 1–12, respectively, with graphs of the most significant KEGG pathways by -log10(p value). These significant p values are well below the significance level of p < 0.05, which corresponds to -log10(p value) of greater than 1.3.
tein in HD cellular mechanisms of toxicity and cell death.

DISCUSSION

This study achieved collation and statistical bioinformatics analysis of published Htt protein interactors identified in non-human HD model systems that meet the criteria that the studies included 1) defining the lengths of Htt fragments or full-length Htt fragment used in the interaction studies, 2) indicating the number of poly-glutamines in the mutant and wild-type forms of Htt, and 3) assessment of native Htt protein interaction complexes without cross-linking agents, which avoids general cross-linking that may include non-specific associated proteins. A high number of proteins that potentially complex with wt Htt were identified as 2,740 interactors, and with mutant Htt that were identified as 2,631 interactors (Fig. 2). The majority of interactors were shared, 2400 proteins, by the wt Htt and mutant Htt groups. Htt is a large protein of 3044 amino acids whose N-terminal region contains the polyQ(n) domain encoded by the CAG repeat expansion of the human HTT gene. The exact function of Htt is under much investigation, but the data compiled here indicate diverse roles in cellular pathways. In the human brain, Htt is processed into multiple endogenous fragments [17] and each may have distinct cellular processes, although localization to specific cellular compartments and organelles is not known. The presence of the mutant polyQ(n) does not substantially affect the molecular weights of fragments outside of the mutation, although some fragments are present at higher levels in the human brain and are different in motor cortices compared to striatum [17]. Amino terminal regions of mutant Htt, containing the expanded polyQ(n) have been identified as diffuse and aggregated forms of nuclear, cytoplasmic and synaptic distribution [229–233]. Whether the expanded polyQ(n) fragment in particular initiates toxic sequelae or in addition to other endogenous fragments is an area of ongoing investigation.

These data compiled in this study provide a valuable resource to the field for formulating hypothesis related to HD pathogenesis and molecular mechanisms. There is currently a great need for effective drugs to slow or stop the neurodegeneration in HD. Therapeutic targeting of Htt interactor proteins are a potential strategy for drug development.

We note that recent articles recently reported data sets of Htt protein interactors but they did not meet our criteria for including information on Htt fragment length, number of polyglutamine residues in the expansion, and assessment of native Htt complexes [234–236] Articles by Aaronson et al. (2021) [234] and Haenig et al. (2020) [235] did not define Htt fragment lengths or number of Htt polyglutamines for all Htt interactors compiled. Another article by Sap et al. (2021) [236], used a general cross-linking agent formaldehyde to isolate Htt complexes, but formaldehyde is a general cross-linking agent that can include non-specific associated proteins. Therefore, information from these three articles were not included in this study.

Binomial analysis of edge effects found that Htt interacting proteins involved in translation were significantly enhanced, while signal transduction, chromatin organization, membrane trafficking, and clathrin-mediated endocytosis functions were significantly diminished in mutant Htt compared to wt. These results suggest diverse roles of Htt in different cell compartments. Among these clusters, significant KEGG pathways of the ribosome, Rap 1 signaling, glycerophospholipid metabolism, and SNARE interactions in vesicular transport were indicated. Among these pathways, glycerophospholipid synthesis and metabolism dysfunction was found to be dysregulated in HD genetic mouse models [237].

All studies utilized animal or cellular model systems to identify interactors. HD human brains have not been used as a starting material to identify interactors, likely due to the limited availability of such tissue. While some studies utilized Htt with polyQ(n) expansions in the adult human HD range (40Q - 50Q) and juvenile HD range (60Q - 100Q), the majority utilized especially high expansions (>100Q). There is a particular deficit in the reduced penetrance range, 30Q - 40Q. A small subset of articles reported Htt protein-protein interactions in human brain (Supplementary Table 1, Master Table) in follow-up studies subsequent to interactor identification in an HD model. Because some Htt interactors may interact at artificially high polyQ(n) ranges but not in the human HD range (40Q - 100Q), future studies should aim to use the ‘human’ range of polyQ(n).

Disruption or enhancement of key mutant Htt protein interactions is a logical approach to identify and test potential therapeutic targets in experimental systems for reduction of mutant Htt mediated toxicity and phenotypes. HD genetic mouse models expressing polyQ(n) expanded human Htt have HD-
like neuropathology, motor dysfunction and cognitive behavioral phenotypes [238, 239]. Some articles in the compiled dataset tested the consequence of mutant Htt interactors on mouse model HD-like phenotypes (Supplementary Table 1, Master Table). All of these protein interactors were chosen by the authors based on hypothesis driven approaches and demonstrate improvement or exacerbation of HD-like pathology as a consequence of mutant Htt protein interactions. Results of these studies illustrate that aberrant mutant Htt protein interactions cause HD-like phenotypes in mouse genetic models.

The reported Htt interactors in animal model systems can be fruitful for understanding possible mechanisms of mutant Htt in the human disease process. We, therefore, propose the strategy that this list of candidate Htt interactors from animal model systems may be assessed in human HD brain tissues using stringent and rigor in the experimental design to elucidate the different clusters/pathways that may be present in human HD brain. Such future studies will be of value to understand mutant Htt protein interaction mechanisms in the human HD brain.

In summary, this study provides the Htt interacting protein dataset resource to the field that may enhance formulation of hypotheses and designing experiments to evaluate Htt initiation of neurodegenerative cellular pathways via Htt protein interactors. The key unknown mechanistic question of HD pathogenesis is why neurons become toxic and die because of the expanded polyQ(n) mutation. Our clustering and pathway analyses suggest protein interactions in significant GO biological processes for pharmacological or genetic (e.g., siRNA) intervention in model systems. While drugs are available for HD patients that may manage symptoms, there is a great need for development of drugs to delay or prevent fatal neurodegeneration and emaciation. Investigation of the key pathways related to HD-like phenotypes determined by statistical analysis in this study may prove to be effective therapeutic targets.

ACKNOWLEDGMENTS

This research was supported by grants from the Huntington’s Disease Society of America (HDSA) awarded to S. Podvin, C. Mosier, and V. Hook. This project was partially supported by the National Institutes of Health, grant UL1TR001442 of CTSA. The content is solely the responsibility of the authors and does not necessary represent the official view of the NIH. Review by our Institutional Review Board has indicated that for this project use of unidentified human tissue samples is exempt.

CONFLICT OF INTEREST

The authors have no conflict of interest to report.

DATA AVAILABILITY

This compilation of Htt interacting proteins is provided to the public via the location of https://www.ndexbio.org/viewer/networks/8acef3e9-ccdc-11e8-aaa6-0ac135e8bacf

The data is also provided in Supplementary Tables 1 and 2.

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: https://dx.doi.org/10.3233/JHD-220538.

REFERENCES


Turner B, Wilbur JD, Chen CY, Manalo V, Hwang PK, Fletterick RJ, Brodsky FM. Actin binding by Hip1 (huntingtin-


[224] King AC, Wood TE, Rodriguez E, Parpura V, Papura V, Gray M. Differential effects of SNARE-dependent gliotransmis-


