

SCIENTIFIC REPORTS



OPEN

A network biology approach to unraveling inherited axonopathies

Dana M. Bis-Brewer¹, Matt C. Danzi¹ , Stefan Wuchty² & Stephan Züchner¹

Inherited axonopathies represent a spectrum of disorders unified by the common pathological mechanism of length-dependent axonal degeneration. Progressive axonal degeneration can lead to both Charcot-Marie-Tooth type 2 (CMT2) and Hereditary Spastic Paraplegia (HSP) depending on the affected neurons: peripheral motor and sensory nerves or central nervous system axons of the corticospinal tract and dorsal columns, respectively. Inherited axonopathies display an extreme degree of genetic heterogeneity of Mendelian high-penetrance genes. High locus heterogeneity is potentially advantageous to deciphering disease etiology by providing avenues to explore biological pathways in an unbiased fashion. Here, we investigate 'gene modules' in inherited axonopathies through a network-based analysis of the Human Integrated Protein-Protein Interaction rEference (HIPPIE) database. We demonstrate that CMT2 and HSP disease proteins are significantly more connected than randomly expected. We define these connected disease proteins as 'proto-modules' and show the topological relationship of these proto-modules by evaluating their overlap through a shortest-path based measurement. In particular, we observe that the CMT2 and HSP proto-modules significantly overlapped, demonstrating a shared genetic etiology. Comparison of both modules with other diseases revealed an overlapping relationship between HSP and hereditary ataxia and between CMT2 + HSP and hereditary ataxia. We then use the Disease Module Detection (DIAMOND) algorithm to expand the proto-modules into comprehensive disease modules. Analysis of disease modules thus obtained reveals an enrichment of ribosomal proteins and pathways likely central to inherited axonopathy pathogenesis, including protein processing in the endoplasmic reticulum, spliceosome, and mRNA processing. Furthermore, we determine pathways specific to each axonopathy by analyzing the difference of the axonopathy modules. CMT2-specific pathways include glycolysis and gluconeogenesis-related processes, while HSP-specific pathways include processes involved in viral infection response. Unbiased characterization of inherited axonopathy disease modules will provide novel candidate disease genes, improve interpretation of candidate genes identified through patient data, and guide therapy development.

The inherited axonopathies spectrum represents a group of disorders unified by length-dependent axonal degeneration. Progressive axonal degeneration can lead to both Charcot-Marie-Tooth type 2 (CMT2) and hereditary spastic paraplegia (HSP) depending on the affected neurons: peripheral motor and sensory nerves or central nervous system axons of the corticospinal tract and dorsal columns, respectively. Though CMT2 and HSP have been historically classified as two separate disorders, divided by peripheral vs central nerve axons, they are now studied together as a spectrum of inherited axonopathies based on their increasingly apparent clinical and genetic overlap¹⁻³. Charcot-Marie-Tooth (CMT) disease is the most common inherited neurological disorder with an estimated prevalence of 4 to 8 per 10,000⁴⁻⁷. CMT typically causes distal muscle weakness and atrophy, high arched feet, decreased tendon reflexes, and sensory loss resulting from progressive polyneuropathy of the motor and/or sensory nerves^{4,8}. Hereditary spastic paraplegias (HSPs) cause bilateral lower limb spasticity and weakness⁹. HSPs have a prevalence of 1.2–9.6 per 100,000⁹.

Since the advent of next generation sequencing, Mendelian disease gene discovery has made notable progress with over 80 identified CMT2 and over 90 HSP genes¹⁰⁻¹². At least five of these genes have been shown to cause either CMT2, HSP, or a mixed phenotype, depending on the underlying mutation¹. The functional characterization of inherited axonopathy disease genes has improved our understanding of the affected biological

¹Dr. John T. Macdonald Foundation Department of Human Genetics, John P. Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, Florida, USA. ²Department of Computer Science and Department of Biology, University of Miami, Coral Gables, Florida, USA. Correspondence and requests for materials should be addressed to S.Z. (email: szuchner@med.miami.edu)

mechanisms^{2,3,13}. Evidently, the pathogenic molecular mechanisms of CMT2 and HSP overlap, including disruption of axonal transport, mitochondrial dynamics, mitochondrial regulation, membrane trafficking, and organelle shaping¹⁴. These functional similarities cohere as the long axons of the upper and lower motor neurons require adaptable cellular machinery to distribute molecules, maintain neuronal homeostasis, and create a specialized neuronal cytoskeleton³.

Inherited axonopathies display an extreme degree of genetic heterogeneity³. High locus heterogeneity is advantageous to studies of disease etiology by providing avenues to explore biological pathways through a network biology approach. These approaches focus on the complex network of functional interdependencies between cellular components in human disease¹⁵. In extension, network medicine is based on the hypothesis that the impact of a single mutated gene is propagated along the gene's network links¹⁵. The links between a set of genes can be summarized into a disease module. Identification and characterization of such disease modules will help resolve differences between normal and disease pathways, reveal common biological functions in related diseases, and identify novel disease gene candidates and pathways related to disease^{15–18}. As an example, a recent protein interaction network analysis of HSP genes identified three candidate genes, which were subsequently found to be mutated in HSP families¹⁹. Finally, it is unlikely that traditional pharmacological approaches will develop a drug for every single CMT and HSP gene, but rather aim at influencing key gene networks.

Here, we investigate the inherited axonopathy disease modules through an unbiased analysis of protein-protein interactions, provided by the Human Integrated Protein-Protein Interaction rEference (HIPPIE), with the objectives of determining the topological relationship between inherited axonopathies and defining and characterizing their putative disease modules²⁰. We evaluate the overlap between disease modules with a shortest path-based measurement²¹ and observe that the CMT2 and HSP proto-modules significantly overlapped, demonstrating a shared genetic etiology. We expand each disease module by incorporating candidate genes with significant connections to disease-related seed genes¹⁷. Such disease modules indicate an enrichment of ribosomal proteins and pathways, likely involved in inherited axonopathy pathogenesis. Our results suggest that unbiased characterization of inherited axonopathy disease modules provide novel candidate disease genes, improve interpretation of candidate genes identified through patient data, and guide therapy development.

Materials and Methods

Disease Genes. Charcot-Marie-Tooth type 2 (CMT2) and Hereditary Spastic Paraplegia (HSP) input-proteins, also known as “seed protein”, were selected through an extensive literature review by domain experts from the Rare Diseases Clinical Research Network (RDCRN). Specifically, the CMT2 gene list was obtained from the Inherited Neuropathy Consortium and the HSP gene list was obtained from the Alliance for Treatment in HSP and PLS, and the Clinical Research in ALS and Related Disorders for Therapeutic Development consortium. Non-syndromic CMT2 genes ($n = 65$) and both complex and pure HSP genes ($n = 96$) were included as seed proteins for disease module construction. To evaluate the topological distance between CMT2 and HSP in a broader context, we selected seven additional diseases for topological separation analysis. We chose diseases that vary in degree of expected shared etiology with inherited axonopathies. Such considerations are based on clinical symptoms and shared disease genes, with hereditary ataxia expected to be the closest and cancer expected to be the most distance diseases. Seed proteins for comparison of diseases (cancer (CA), deafness (DFN), cardiomyopathy (CM), muscular dystrophy (MD), Parkinson's disease (PD), amyotrophic lateral sclerosis (ALS), and hereditary ataxia (ATX)) were collected from the Online Mendelian Inheritance in Man (OMIM, updated 10-25-2017)²². Supplementary Table 1 contains all seed proteins used in this study, and Supplementary Fig. 1A displays the degree distribution for seed proteins.

Molecular Interactions. The Human Integrated Protein-Protein Interaction rEference (HIPPIE) is a publicly available, scored protein-protein interactions (PPI) dataset²³. HIPPIE integrates data from BioGRID (genetic interactions removed), DIP, HPRD, IntAct, MINT, BIND, and MIPS databases. HIPPIE provides a curated scoring scheme that is based on the experimental technique to identify the PPI, the number of studies that have reported the PPI, and the reproducibility of the PPI across model organisms. The HIPPIE confidence score ranges from 0 to 1 and reflects the quality of experimental evidence supporting each PPI. The HIPPIE authors have predefined confidence levels based on the quartiles of all confidence scores: medium confidence (0.63 – second quartile of the HIPPIE score distribution) or high confidence (0.73 – third quartile) (<http://cbdm-01.zdv.uni-mainz.de/mschaefer/hippie/information.php#sources>). To balance network coverage and reasonable interaction confidence, we accounted for the interactions with at least a medium confidence score of 0.63. The most recent update includes approximately 273,900 experimentally determined PPIs between 17,000 human proteins²⁰. The HIPPIE (v2.0) dataset was pre-filtered to remove interactions from non-human sources and below the predefined medium-confidence score threshold in order to create a more stringent global interactome of 204,215 interactions between 15,690 proteins. The filtered HIPPIE protein-protein interaction network thus obtained is scale-free, as demonstrated by a power-law decay in the corresponding degree distribution (Supplementary Fig. 1B).

Disease Proto-Module Construction. A goal of network analysis is to define and study a disease module. We first define a disease proto-module since our knowledge of both the disease-causing proteins and all human protein interactions is incomplete²¹. Disease proto-modules are considered clusters of highly connected seed proteins that were constructed by retrieving the subgraph of directly connected seed proteins from the filtered HIPPIE network. The connectivity of a proto-module is quantified by the size of the largest connected component (LCC) and shortest distance (SD) between proteins²¹. While LCC indicates the number of directly connected disease proteins, SD measures the localization of disease proteins that do not directly interact. Utilizing publicly available Python scripts, we calculate the LCC, SD, and significance of each disease proto-module²¹. To assess the significance of LCCs, expected null distributions were calculated by randomizing sets of proteins of equal

seed list size. We determined the expected distribution of LCCs after 10,000 randomizations and calculated a proto-module specific z -score²¹. As previously described, a proto-module is obtained if the corresponding z -score was ≥ 1.6 with a significance p -value ≤ 0.05 ²¹. To assess the SDs, the distribution of SDs after 10,000 randomizations was tested using a Mann-Whitney-U test²¹.

Disease Proto-Module Expansion. The Disease Module Detection (DIAMOND) algorithm is one of the latest and most utilized disease-gene prediction methods^{17,24}. We selected the highly cited DIAMOND algorithm for this study based on the previous successful application to a diverse range of diseases^{18,25–28}. Specifically, DIAMOND expands each proto-module to identify the putatively complete disease module¹⁷. Based on the assumption that disease proteins interact more significantly with each other, DIAMOND explores the topological neighborhood of the seed proteins based on the significance of connections to disease proteins¹⁷. Briefly, DIAMOND calculates the connectivity significance for all proteins that interact with the seed proteins; ranks the proteins according to p -values; incorporates the highest ranked protein into the set of seed proteins; and repeats the process to expand the disease seed set¹⁷. Each disease proto-module was expanded using DIAMOND recommended settings, including weighting the seed proteins to 10 and running 200 iterations. The core inherited axonopathy module was constructed from the intersection of the CMT2 and HSP expanded modules.

Disease Proto-Module Separation. The proximity of diseases within a network is useful for understanding their biological and clinical overlap. The network-based separation between two disease proto-modules A and B was assessed by comparing the mean shortest distances of seed proteins within each proto-module (d_{AA} and d_{BB}) to the mean shortest distance between the proteins of each proto-module (d_{AB})²¹. In particular, Menche *et al.* defined the network-based separation as $d_{AB} - \frac{d_{AA} + d_{BB}}{2}$ and calculated a z -score to quantify the difference between the observed separation and random expectation and analytically calculate a corresponding p -value for each z -score²¹. A positive z -score indicates separation of disease pairs while a negative z -score indicates overlap. The $|z\text{-score}|$ must be ≥ 1.6 for a disease pair to be more/less overlapping than randomly expected with a significance p -value ≤ 0.05 ²¹.

Semantic Similarity Within Expanded Module. The biological functions of the expanded modules were assessed with Gene Ontology (GO) semantic similarity measurements using the GOSemSim R package^{29–31}. The semantic similarity was computed using the graph-based Wang method, which computes the semantic similarities between terms based on both their topology within the GO graph structure and their relations with ancestor terms³¹. Semantic similarities of multiple GO terms were combined using the Best-Match Average (BMA) strategy which calculates the average of all maximum similarities on each pair of GO terms³¹. A systematic evaluation of the metrics for GO based semantic similarity found that the BMA is the best combination approach since it yields the highest resolutions and does not show the undesired behaviors of the maximum approach or the average approach³². Before computing the semantic similarities of GO terms, we filtered the GO terms based on annotation evidence codes to produce a high quality dataset. Specifically, we removed Inferred from Electronic Annotation (IEA) evidence codes. Although IEA are the most common evidence codes, they also represent weak associations as they are based on electronic curation without human curation. Since the input for this analysis (DIAMOND proteins) is based on a dataset of physical interactions, we also removed the Inferred from Physical Interaction (IPI) evidence codes to remove bias and circular reasoning from the analysis.

Functional Enrichment Analysis of Disease Module. Functional enrichment analysis was performed for GO Biological Process (BP) terms, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, and Wikipathways^{29,33,34}. Over-Representation Analysis (ORA) of protein clusters, based on GO BP semantic similarity were analyzed using the DAVID Functional Annotation Tool³⁵. ORA of disease modules for KEGG pathways and Wikipathways was performed using WebGestalt and were filtered for FDR ≤ 0.2 ³⁶. Pathway ORA was performed on 6 sets: the expanded CMT2 module, the expanded HSP module, the intersection of the expanded CMT2 and HSP modules (AXONO core module), the union of the expanded CMT2 and HSP modules (AXONO spectrum module), the difference of the expanded CMT2 and HSP modules (CMT2-only module), and the difference of the HSP and CMT2 modules (HSP-only module). Because of gene redundancy in pathway sets, the Cytoscape plugin Enrichment Map was used to visualize the similarity between enriched pathway sets with a Jaccard coefficient greater than 0.1³⁷.

Results

Constructing the inherited axonopathy modules. Disease module construction requires an input seed list of disease proteins and a molecular interaction network. We curated 65 non-syndromic axonal Charcot-Marie-Tooth 2 (CMT2) and 96 Hereditary Spastic Paraplegia (HSP) input seed genes through an extensive literature review (Supplementary Table 1). As a protein-protein interaction data source, we used the Human Integrated Protein-Protein Interaction rEference (HIPPIE, v2.0) that integrates interaction data from high quality databases and provides a high-performing scoring system²⁰. The HIPPIE confidence score is based on the amount and the reliability of supporting experimental evidence for each interaction²³. Applying the predefined medium confidence score (0.63 – second quartile of the HIPPIE score distribution), we created a medium-confidence global interactome consisting of 204,215 interactions between 15,690 human proteins.

In an interactome, disease associated proteins are expected to cluster into specific network neighborhoods through physical interactions^{15,38}. The inherited axonopathy proteins follow this expectation by forming significantly connected disease modules (Fig. 1). The largest connected component (LCC) of a subgraph measures the number of seed proteins in the largest subgraph composed of direct interactions between seed proteins. In our analysis, the CMT2 and HSP LCCs are significantly larger than expected by random chance (CMT2: $n = 35$,

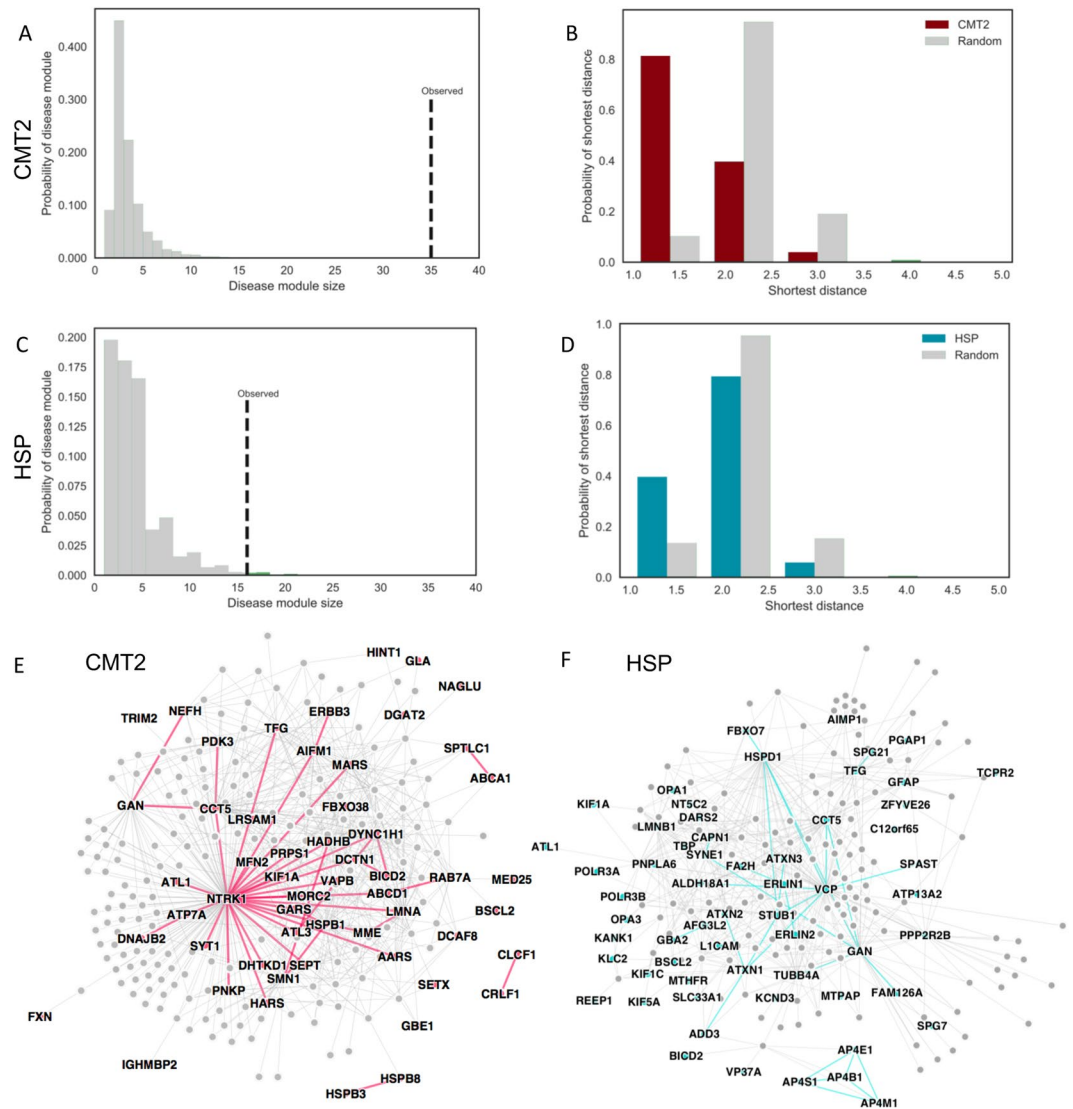


Figure 1. Inherited axonopathy proto-module connectivity (A–D) and expanded module networks (E,F). (A,B) Increased observed largest connected component (LCC) size of CMT2 (A) and HSP (B) disease seed genes in comparison to the expected, random distribution. (C,D) Increased shortest distance between inherited axonopathy disease genes in comparison to the expected, random expectation. (E,F) First degree interactions between known Mendelian disease genes (colored) and DIAMOnD genes (gray) for CMT2 (E) and HSP (F). Disease specific genes are labeled, and direct edges between disease genes are more heavily weighted.

z -score = 17.1; HSP: $n = 16$, z -score = 4.0) (Fig. 1A,C), indicating that the CMT2 and HSP disease proteins form disease proto-modules composed of proteins with high biological and functional similarity. Another measure of connectivity is the shortest distance (SD) between seed proteins. SD also includes fragmented seed proteins that do not connect to proteins in the LCC, providing a more global measure of connectivity than LCC alone. As the knowledge of human protein interactome and disease etiology is incomplete, seed proteins are expected to fragment within the interactome, pointing to larger distances. However, we observed that SD distributions of CMT2 and HSP seed genes are shifted towards 1.0, indicating direct connections between seed proteins. In comparison to the expected distribution based on randomized protein sets (Fig. 1B,D) the observed means are significantly smaller (Mann Whitney U Test; CMT2: p -value = 2.7E-16; HSP: p -value = 5.0E-05). Taken together, the significant connectivity measurements demonstrate that the CMT2 and HSP disease proteins form agglomerations that can be expanded and further studied (Fig. 1).

The significantly connected CMT2 and HSP proto-modules were next used as seed modules to expand into putatively “complete” disease modules using the DISeAse Module Detection (DIAMOnD) method (Supplementary Table 2)¹⁷. The DIAMOnD algorithm identifies the topological neighborhood of the seed proteins based on the significance of connections to seed proteins and has been successfully used to study asthma and cardiovascular endophenotypes^{18,25}. After performing 200 iterations of the DIAMOnD method, an additional 53 CMT2 proteins and 61 HSP proteins were incorporated into each LCC, which were originally fragmented and not

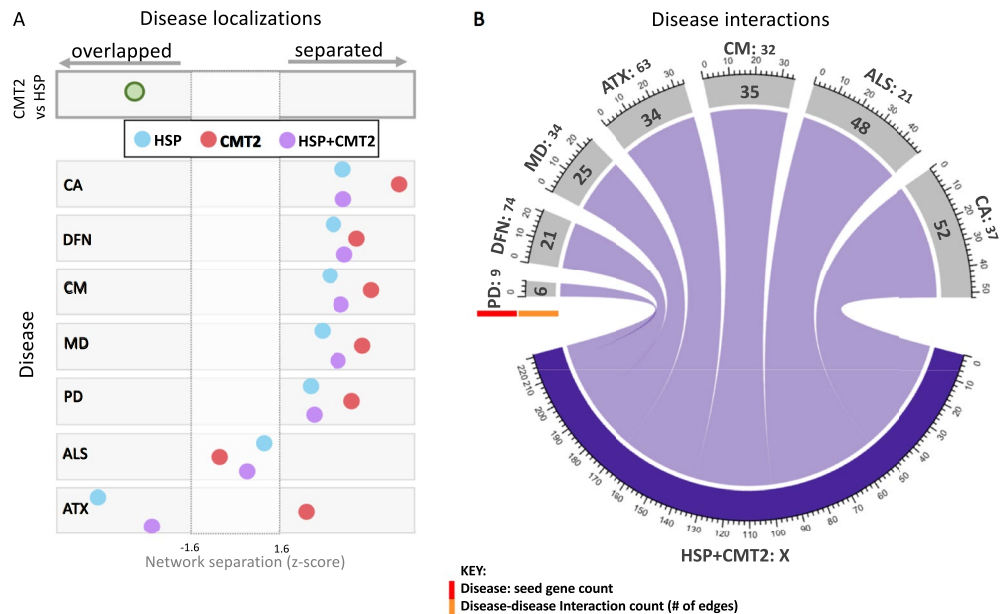


Figure 2. Disease proto-modules separation and connections. **(A)** Significance of topological relationships between axonopathy proto-modules (Charcot-Marie-Tooth type 2 only: CMT2; Hereditary Spastic Paraplegia only: HSP; combined HSP + CMT2) and control diseases (Cancer: CA, Deafness: DFN, Cardiomyopathy: CM, Muscular Dystrophy: MD, Parkinson's Disease: PD, Amyotrophic Lateral Sclerosis: ALS, and Hereditary Ataxia: ATX). Gray dotted lines indicate the threshold for significance ($|Z\text{-scores}| >= 1.6$). Negative (positive) Z-scores indicate overlapped (separated) modules. **(B)** Direct connections between HSP + CMT2 disease proteins and comparison disease proteins. Disease names and input seed protein count (highlighted by red bar) for each disease compared are indicated outside the circle. Direct interactions between disease proteins of HSP + CMT2 are displayed on the outer gray bar (highlight by orange bar).

directly connected (Fig. 1E,F). To limit the incorporation of false positives into each inherited axonopathy module, the DIAMOND proteins were evaluated for biological evidence by comparing the Gene Ontology Biological Process (GO BP) terms of the DIAMOND proteins to the GO BP terms of the known disease proteins. The GO BP terms were compared by hierarchical clustering based on semantic similarity (Supplementary Fig. 2)²⁹, allowing us to infer the biological relationship between DIAMOND proteins and disease proteins. As indicated by the semantic similarity clustering, DIAMOND proteins are functionally related to the disease seed proteins, suggesting that we limited the incorporation of false positives beyond the true biological limit. To understand the underlying biological function, we performed GO BP Over Representation Analysis (ORA) of the proteins within each cluster. In particular, we confirmed known cellular processes including axonal transport, microtubule dysregulation, and mitochondrial involvement (Supplementary Fig. 2). Taken together, these results indicate that the expansion of proto-modules leads to the identification of proteins that are functionally similar to disease proteins and are involved in biological processes that are relevant to the disease. After topologically validating the disease modules with connectivity measures and biologically validating the incorporated proteins through GO BP analysis, we present the final CMT2 and HSP disease modules in Fig. 1E,F. The networks display the connections between disease proteins (in color) and DIAMOND proteins (in gray). The DIAMOND proteins (Supplementary Table 2) will improve our understanding of the disease by adding power to pathway analysis and containing new candidate disease genes and targets.

Topological relationship of disease modules. Disease modules that are located in adjacent network areas will likely share proteins, interactions, and pathways involved in disease pathogenesis²¹. Disruption of overlapping functional modules will likely result in shared clinical characteristics in each disease²¹. The topological relationship between inherited axonopathy proto-modules (i.e. CMT2, HSP) was quantified through the shortest distances between modules (Fig. 2). The close relationship within the spectrum of inherited axonopathies (between CMT2 and HSP) is confirmed in our analysis by significantly overlapping disease proto-modules (network distance = -0.15 ; $z\text{-score} = -2.6$) (Fig. 2A). We next explored the relationships of the inherited axonopathies with cancer (CA), deafness (DFN), cardiomyopathy (CM), muscular dystrophy (MD), Parkinson's disease (PD), amyotrophic lateral sclerosis (ALS), and hereditary ataxia (ATX) (Fig. 2A and Supplementary Fig. 3). Utilizing disease genes curated from the OMIM database (updated 10-25-2017), we compared the corresponding disease-specific proto-modules with CMT2, HSP, and their union (HSP&CMT2). All compared diseases were significantly separated from each axonopathy module except for ALS and ATX. Notably, the ALS proto-module was not significantly separated nor overlapped with the CMT2, HSP, and HSP&CMT2 proto-modules. The ATX proto-module overlapped significantly with both the HSP and HSP&CMT2 proto-modules, yet significantly separated from the CMT2 proto-module. Since the network separation parameter depends on the number of

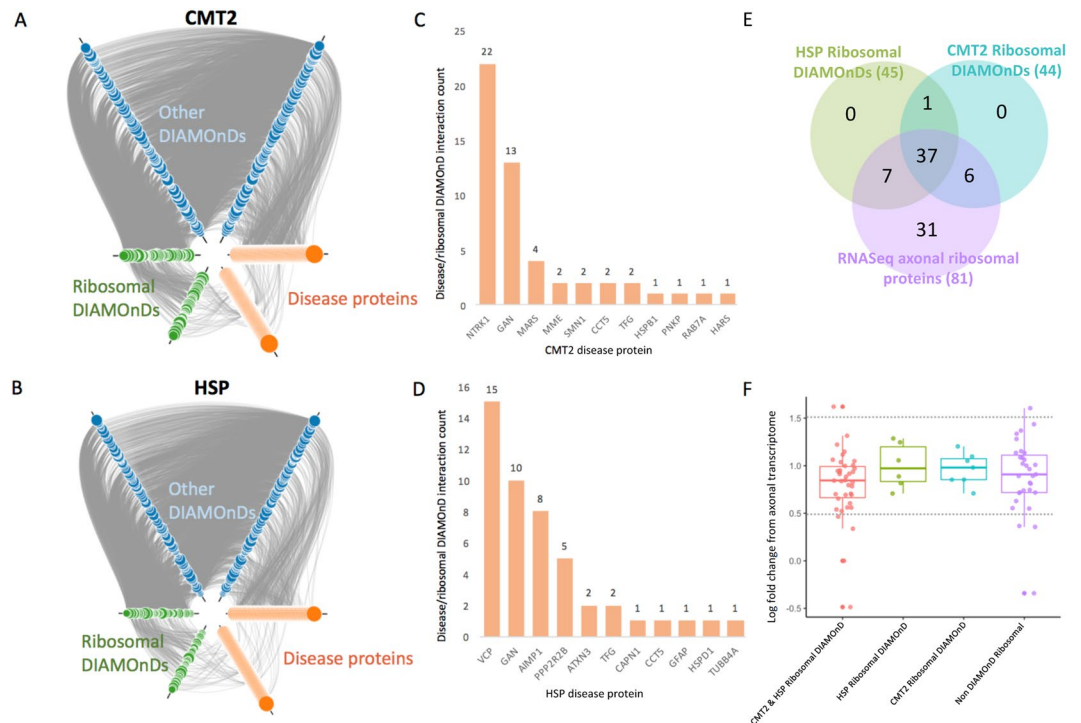


Figure 3. Interactions within the expanded inherited axonopathy modules. **(A,B)** First degree interactions within CMT2 **(A)** and HSP **(B)** expanded modules. DIAMOND proteins are separated into ribosomal proteins (green) and other (blue). Size of DIAMOND nodes are proportional to DIAMOND order of incorporation – smaller nodes were incorporated earlier. **(C,D)** First degree interaction counts between disease protein and ribosomal DIAMOND proteins. **(E)** Relationships between HSP ribosomal DIAMOND proteins, CMT ribosomal DIAMOND proteins, and ribosomal proteins significantly localized to the axons of human motor neurons (q -value ≤ 0.1) **(F)**. Distribution of log fold changes from axonal and soma differential localization.

proteins in each disease and the number of direct interactions between disease proteins, we provided the count of seed proteins (ranging from 9 to 72) and the amount of direct connections between disease proteins (ranging from 6 to 52 connections) (Fig. 2B).

Ribosomal proteins are enriched within the inherited axonopathy modules. The CMT2 and HSP expanded modules both contained an unexpected number of ribosomal proteins incorporated as DIAMOND proteins (CMT2 $n = 47$; HSP $n = 48$; overlap $n = 40$) (Fig. 3). To qualitatively determine if the enriched ribosomal proteins are an artifact of the global network structure, we examined the underlying protein interactions (Fig. 3A,B). The interactions between ribosomal DIAMOND proteins and disease seed proteins (green nodes connecting to orange nodes) are similar to the interactions between disease seed proteins (orange nodes connecting to orange nodes), suggesting true biological interactions instead of a topology-biased artifact. We then quantified the direct interactions between disease seed proteins and ribosomal DIAMOND proteins - CMT2: minimum: 1, maximum: 22, mean: 4.6; HSP: minimum: 1, maximum: 15, mean: 4.2 (Fig. 3C,D). To further verify that ribosomal protein interactions are not an artifact of the HIPPIE network, we expanded each comparison disease proto-module and quantified the first-degree interactions between disease seed proteins and ribosomal DIAMOND proteins (Supplementary Figure 4). We found that neurologically related diseases contained more ribosomal DIAMOND proteins than the non-neurological control disease, demonstrating disease-specific enrichment. Furthermore, in a recent study of the human motor neuron transcriptome, 81 ribosomal proteins showed significant localization (q -value ≤ 0.1) to the axon with modest expression levels (median = 0.89) (Fig. 3E,F)³⁹. In the Venn diagram in Fig. 3E, we observed that 50 of the axonal ribosomal DIAMONDs were significantly localized to the axon, with 37 of these present in both the CMT2 and HSP sets of DIAMONDs. In Fig. 3F, we show that the distributions of expression level log fold changes (LFCs) of the ribosomal DIAMONDs (shared CMT2-HSP: $n = 37$, HSP only: $n = 7$, CMT2 only: $n = 6$) and the ribosomal non-DIAMONDs ($n = 31$) were highly similar, corroborating the biological relevance of the network analysis results. Eight ribosomal proteins were reported as differentially localized to the axons in comparison to the soma (q -value ≤ 0.1 ; LFC ≤ 0.05 or ≥ 1.5)³⁹. RPL26, RPL28, RPL7A, and RPS6KA1 are DIAMOND proteins while RPL31, RPL7L1, RPS6KC1, and RSL24D1 are not (Fig. 3F). Significant axonal localization of 50 ribosomal DIAMOND proteins demonstrates a biological significance, with potential to explain the enrichment of ribosomal proteins within neurologically-related disease network.

Pathway analysis of the inherited axonopathy spectrum modules. Functional analysis of the inherited axonopathy spectrum modules was performed by an Over-Representation Analysis (ORA) of

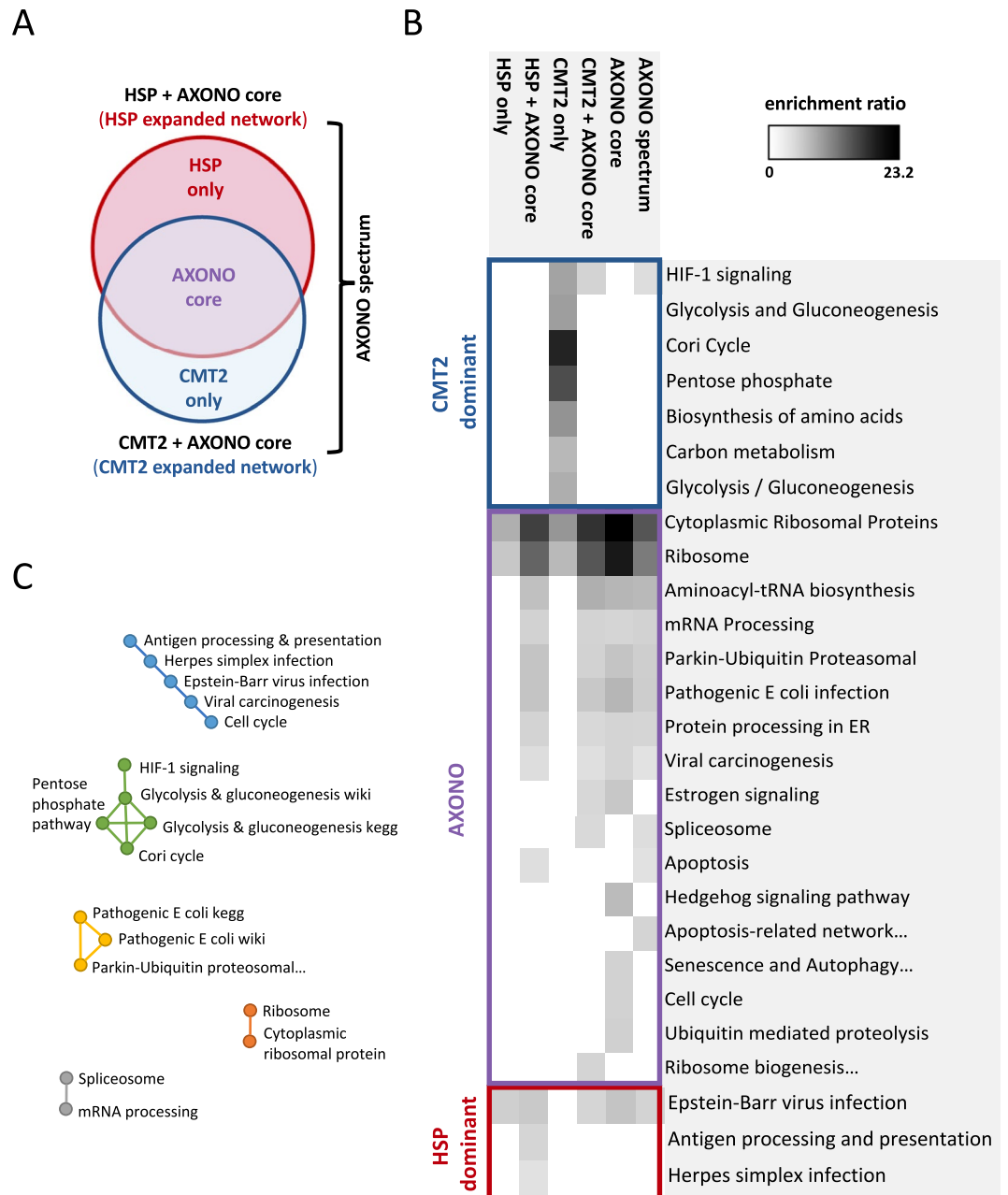


Figure 4. Pathway analysis of inherited axonopathy modules. **(A)** Venn diagram of disease gene sets used for pathway enrichment including the expanded CMT2 module, the expanded HSP module, the intersection of the expanded CMT2 and HSP modules (AXONO core), the union of the expanded CMT2 and HSP modules (AXONO spectrum), the difference of the expanded CMT2 and HSP modules (CMT2-only), and the difference of the HSP and CMT2 modules (HSP-only). **(B)** KEGG and WikiPathways significantly enriched within each disease module set. **(C)** Closely related pathways based on redundancy in proteins involved in each pathway.

WikiPathways and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Supplementary Table 3)^{33,34}. Our results show both unique and shared pathways across inherited axonopathies (Fig. 4). Pathway enrichment analysis was performed on 6 sets of proteins: the expanded CMT2 module ($n = 254$ proteins), the expanded HSP module ($n = 295$ proteins), the intersection of the expanded CMT2 and HSP modules (AXONO core module, $n = 152$ proteins), the union of the expanded CMT2 and HSP modules (AXONO spectrum module, $n = 397$ proteins), the difference of the expanded CMT2 and HSP modules (CMT2-only module, $n = 102$ proteins), and the difference of the HSP and CMT2 modules (HSP-only module, $n = 143$ proteins) (Fig. 4A). We simplified the 6 protein sets to 3 protein sets for pathway classification and visualization purposes and accounted for the CMT2-dominant, HSP-dominant sets and the intersection of CMT2-expanded and HSP-expanded networks (AXONO). We assigned each enriched pathway to the most influenced disease module (CMT2-dominant, HSP-dominant, or AXONO) based on the combination of module sets that was enriched for a pathway (Fig. 4B).

For example, pathways that were enriched in the CMT2 only set, but not in the HSP only set, were categorized as a CMT2-dominant pathway. The pathway results can be ranked by the highest count of observed input proteins within the pathway set or the highest enrichment score within the pathway set, each averaged across the 6 disease module sets. The Cytoplasmic Ribosomal Proteins (wiki) and Ribosome (KEGG) pathways were the most highly enriched with the highest count of input proteins within each module set, likely due to the high amount of ribosomal DIAMOND proteins. Excluding the ribosomal protein pathways, the top 5 pathways (based on the observed count of input proteins) are: **broad axonopathy spectrum**: mRNA Processing, Protein processing in the endoplasmic reticulum, Viral carcinogenesis, Spliceosome, and Apoptosis; **CMT2-dominant**: HIF-1 signaling pathway, Aminoacyl-tRNA biosynthesis, Biosynthesis of amino acids, Carbon metabolism, and Glycolysis and Gluconeogenesis; and **HSP-dominant**: Epstein-Barr virus infection, Herpes simplex infection, and Antigen processing and presentation. The top 5 most enriched pathways, excluding the ribosomal protein pathways, are: **broad axonopathy spectrum**: Aminoacyl-tRNA biosynthesis, Hedgehog signaling pathway, Pathogenic *Escherichia coli* infection, Parkin-Ubiquitin Proteasomal System pathway, and Estrogen signaling pathway; **CMT2-dominant**: Cori Cycle, Pentose phosphate pathway, Biosynthesis of amino acids, Glycolysis and Gluconeogenesis, and Carbon metabolism; and **HSP-dominant**: Epstein-Barr virus infection, Antigen processing and presentation, and Herpes simplex infection. The enriched pathways were simplified for easier understanding by grouping pathways together based on overlapping proteins (Fig. 4C).

Discussion

The goal of this study was to compare the spectrum of inherited axonopathies by identifying and characterizing the disease modules of axonal Charcot-Marie-Tooth disease (CMT2) and Hereditary Spastic Paraplegia (HSP). To our knowledge, this is the first study to apply an unbiased network analysis approach to CMT2 or inherited axonopathies combined.

Inherited axonopathies form significantly connected disease modules. Disease genes are not distributed randomly within biological interaction networks but form local connections with each other as they interact to perform related biological functions^{18,38,40}. We began our analysis by identifying these localized gene neighborhoods of inherited axonopathy proteins. Since the genetic etiology of inherited axonopathies is not completely known, our analysis is limited by incomplete input seed lists. However, we found that CMT2 and HSP genes were significantly more connected to each other than expected by chance (Fig. 1A–D). The identification of these significant disease-specific proto-modules highlights the robustness of network analysis (Fig. 1E,F).

Overlapping disease modules are expected to share clinical characteristics, pointing to physical interaction of proteins, and by extension, common biological pathways²¹. Since CMT2 and HSP have been categorized as inherited axonopathies based on their clinical and genetic overlap, we anticipated and observed a significant overlap between CMT2 and HSP disease proto-modules within the network (Fig. 2A). We found a significant separation of CMT2, HSP, and BOTH proto-modules from cancer, deafness, cardiomyopathy, muscular dystrophy, and Parkinson's disease; an unclear relationship with amyotrophic lateral sclerosis; and discordant distances to ataxia (Fig. 2A). A previous study constructed an “HSPome” from published and candidate HSP proteins, and observed a significant overlap between HSP with Alzheimer's disease, Parkinson's disease (PD), and Amyotrophic Lateral Sclerosis (ALS)¹⁹. However, in our study, HSP was significantly separated from PD and had an unclear relationship with ALS (Fig. 2A). The difference in topological relationships between studies is likely caused by differences in the underlying data sources and input seed lists. The “HSPome” was created from iRefIndex, HumanNet, and STRING with 43 known HSP proteins as input seed list, whereas we used the HIPPIE database and increased our input seed list to 95 proteins. In particular, 40 seeds overlapped between the two inputs^{41–43}. Interestingly, the inherited axonopathy proto-modules appear slightly “less” separated, based on *z*-score, from PD and ALS than the other non-neurological controls (Cancer, Deafness, Cardiomyopathy, and Muscular Dystrophy). Each axonopathy module showed an unclear relationship with ALS (neither significantly overlapped nor separated), which should be clarified in the future with additional seed proteins. Notably, the HSP proto-module is highly overlapping with the ATX proto-module while the CMT2 proto-module is significantly separated. This overlap supports the ataxia-spasticity disease spectrum and warrants further investigation⁴⁴.

Ribosomal proteins in neurological pathogenesis. The CMT2 and HSP disease modules were identified by expanding the seed protein sets with the recommended 200 iterations of the DISeAse Module Detection (DIAMOND) method¹⁷. Exploration of the DIAMOND proteins revealed that close to 25% are ribosomal proteins (Fig. 3). Ribosomal proteins (RPs) surround a catalytic core of ribosomal RNA (rRNA) to form ribosomes^{45,46}. RPs have universal roles in ribosome assembly and function, such as assisting in the proper folding of rRNA^{46,47}. Beyond these core roles, RPs have demonstrated extra-ribosomal functions both within the ribosome system, such as regulating and balancing individual RP synthesis with rRNA, and outside of the ribosome system, such as triggering apoptosis in response to disruption in ribosome synthesis⁴⁷. RPs are responsible for various phenotypes across organisms: the *Minute* phenotype in *Drosophila* (delayed development; short, thin bristles; and impaired fertility), gene-specific abnormalities caused by morpholino-induced knock-downs in zebrafish, developmental defects in mice, and 25% of Diamond-Blackfan anemia cases in humans^{46,48,49}. Additionally, RPs were shown to be major interactors with kinase substrates of LRRK2 - a familial and sporadic PD protein⁵⁰. Specifically, phosphodeficient ribosomal small protein 15 (RPS15) rescued LRRK2 neurotoxicity by affecting translation⁵⁰. Impaired protein translation has been repeatedly implicated in CMT with mutations within five aminoacyl-tRNA synthetases identified to date⁵¹. Furthermore, RPs have tissue-specific expression and have been experimentally shown to localize to nerves *in vivo*, including mature peripheral nerve axons and cortical tract axons^{52,53}. Locally synthesized RPs can dynamically change specific ribosomal compositions which may allow axons to fine-tune mRNA selectivity⁵³. In this study, we show that RPs are appreciable interactors of inherited axonopathy disease proteins.

Interestingly, an enrichment of ribosomal proteins was recently observed within the axonal transcriptome from fibroblast-derived human motor neurons (Fig. 3E,F)³⁹. Forty-six of the DIAMOND RPs show modest localization between axon and cell body in a human motor neuron model and an additional four were significantly differentially localized to the axon³⁹. Furthermore, we observe increased RP involvement in neurological disease controls compared to non-neurological disease (Supplementary Fig. 4). Taken together, our results support the emerging role of translational dysregulation in diverse neurologic diseases. We speculate that RPs may play an important function in neurologic diseases via their extra-ribosomal roles, including alerting the cell to stress⁴⁷.

Pathway analysis provide insights into inherited axonopathy disease mechanisms. Lastly, we compared the 6 inherited axonopathy spectrum module sets through pathway overrepresentation analysis, and assigned each enriched pathway to the most influenced disease module (CMT2-dominant, HSP-dominant, or AXONO) (Fig. 4; Supplementary Table 3). Since each disease module set contained many ribosomal proteins, the most enriched pathways were the Cytoplasmic Ribosomal Proteins (Wikipathway) and Ribosome (KEGG) pathways. Aside from the ribosomal pathways, we identified pathways known to be involved in inherited axonopathies, such as aminoacyl-tRNA biosynthesis and protein processing in the endoplasmic reticulum (ER). Protein processing in the ER was significantly associated with each module involving the “axonopathy core” set, and is important in both CMT2 and HSP pathophysiology. The ER is the primary site for protein/lipid biosynthesis and intracellular calcium storage, and is tightly coordinated with the mitochondria to maintain the energetically demanding mechanisms of long axons⁵⁴. Many affected axonopathy genes localize to the ER or mitochondria and disrupt the organelles’ coordination⁵⁵. For example, *atlastin-1* can cause both CMT2 and HSP through abnormal ER shaping and mitochondrial transport^{3,55}.

Pathways that are not directly implicated in the axonopathy pathogenesis were also identified. The AXONO module was enriched for the Hedgehog signaling pathway and the Estrogen signaling pathway. Sonic hedgehog (Shh) is essential for central nervous system development and is expressed within injured sciatic nerve⁵⁶. Recently, reduced Shh signaling was shown to disturb axon regeneration via misregulated myelin degradation⁵⁶. Estrogen signaling has also been implicated in axon regeneration by accelerating peripheral nerve regeneration and increasing motoneuron participation^{57,58}. Another unexpected pathway enriched within the AXONO module was Pathogenic *Escherichia Coli* infection. The *E. Coli* pathways were significantly overlapping with Parkin-Ubiquitin Proteasomal System pathway through the tubulin proteins in each pathway (Fig. 4C). Alteration of the gut microbiome has been suggested to contribute to neurodegenerative diseases, such as Parkinson’s and Alzheimer’s disease⁵⁹. Furthermore, gut inflammation and microbiome alterations may play a role in the development or progression of motor neuron disease and ALS^{59–61}.

Distinct pathways were revealed between the CMT2-dominant and HSP-dominant modules. The CMT2-dominant module predominantly involved overlapping glucose metabolism pathways: HIF-1 signaling, glycolysis and gluconeogenesis, Cori cycle, and pentose phosphate pathway (Fig. 4B). The disturbance of bioenergetics via mitochondrial dysfunction is a well-documented cause of CMT2^{62–64}. Furthermore, mutations in the gene encoding pyruvate dehydrogenase isoenzyme 3 (PDK3) cause CMT, a protein which links the glycolytic cascade to the Krebs cycle⁶⁵. Based on our results, the involvement of glycolytic metabolism may be underappreciated in CMT2 pathophysiology. We speculate that glucose metabolism could contribute to CMT2 through reduced vesicular transport. Vesicular glycolysis is necessary and sufficient for the energy requirements of fast axonal transport (FAT)⁶⁶. FAT depends on ATP generated by glycolysis, not by the mitochondria, and is reduced when glycolysis is perturbed⁶⁶. FAT alteration is linked to multiple neurodegenerative diseases, and FAT restoration is likely to contribute to neuronal protection⁶⁷.

The HSP-dominant modules were enriched for immune-related pathways: Epstein-Barr virus (EBV) infection, Herpes simplex (HSV) infection, and Antigen processing and presentation. EBV and HSV are members of the *Herpesviridae* family and have large, complex KEGG pathways which involve multiple sub-pathways including, for example, antigen processing and presentation. For a more detailed understanding of which EBV and HSV processes were most involved, we investigated the GO BP terms of the genes involved in the EBV and HSV pathways. The EBV pathway genes were enriched for membrane organization, protein targeting, and protein insertion into the mitochondrial membrane during apoptosis induction. The HSV pathway genes were enriched for transcription from the RNA polymerase II promoter. These results implicate antiviral mechanisms within HSP pathogenesis. Growing evidence documents the important roles of the mitochondria in antiviral immunity, including participation in signaling cascades and inflammation activation^{68,69}. Since mitochondrial dysfunction is known to cause HSP, mitochondrial involvement may link HSP to innate immune responses⁷⁰. Interestingly, EBV and human herpesvirus 6 (HHV-6) have been associated with multiple sclerosis (MS)⁷¹. HSP and MS both cause axonal loss of the corticospinal tract, share clinical similarities, and have concurred within families^{72–74}.

Concluding Remarks

In this study, we extracted insights about the spectrum of inherited axonopathies through the application of network-based tools to publically available protein interaction data. Starting with lists of known CMT2 and HSP disease genes, we determined significantly connected disease proto-modules from the global human interactome. The CMT2 and HSP proto-modules were topologically overlapping, supporting the consideration of these diseases as a related spectrum. We show the inherited axonopathies have an unclear (neither overlapping nor separated) relationship with ALS. These relationships will be elucidated as the molecular etiology of motor neuron diseases is further understood. We found a significant overlap between HSP and inherited ataxias, which is consistent with the ataxia-spasticity disease spectrum, while CMT2 was significantly separated from the ataxias⁴⁴. This demonstrates the utility of network-based approaches in understanding the relationships between spectrums of closely related diseases. Each proto-module was expanded to identify additional proteins likely involved in disease pathogenesis, thus increasing the power of pathway analysis and providing candidate disease genes. We show

that the inherited axonopathies are enriched for interactions with ribosomal proteins, which were recently shown to be preferentially localized to the axonal compartment of motor neurons³⁹. Pathway analysis revealed biological processes enriched for each axonopathy independently, such as glucose metabolism for CMT2 and innate immunity for HSP, as well as for general axonopathies, such as new signaling pathways. Additionally, the identified candidate disease genes may guide exploration into drug target identification^{71,75,76}. Sixty of the identified DIAMOND proteins are present on the updated list of druggable genes: 49 DIAMONDS may be targeted by a small molecule, 18 DIAMONDS may be targeted by a biotherapeutic (monoclonal antibody/enzyme or other protein), and 1 DIAMOND is involved in absorption, distribution, metabolism, and excretion of a compound⁷⁶. Follow-up functional studies will be required to evaluate the potential involvement of DIAMOND proteins in disease and their therapeutic potential. Overall, our study provides useful insights into inherited axonopathies relationships and disease mechanisms, and demonstrates the opportunities available through network-based approaches.

Data Availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

- Fridman, V. & Murphy, S. M. The spectrum of axonopathies: from CMT2 to HSP. *Neurology* **83**, 580–581 (2014).
- Zuchner, S. & Vance, J. M. Emerging pathways for hereditary axonopathies. *J. Mol. Med.* **83**, 935–943 (2005).
- Timmerman, V., Clowes, V. E. & Reid, E. Overlapping molecular pathological themes link Charcot-Marie-Tooth neuropathies and hereditary spastic paraplegias. *Exp. Neurol.* **246**, 14–25 (2013).
- Saporta, M. A. & Shy, M. E. Inherited Peripheral Neuropathies. *Neurol Clin* **31**, 597–+ (2013).
- Dyck, P. J. & Lambert, E. H. Lower motor and primary sensory neuron diseases with peroneal muscular atrophy. II. Neurologic, genetic, and electrophysiologic findings in various neuronal degenerations. *Arch. Neurol.* **18**, 619–625 (1968).
- Skre, H. Genetic and Clinical Aspects of Charcot-Marie-Tooth's Disease. *Clin Genet* **6**, 98–118 (1974).
- Braathen, G. J., Sand, J. C., Lobato, A., Høyer, H. & Russell, M. B. Genetic epidemiology of Charcot-Marie-Tooth in the general population. *European Journal of Neurology* **18**, 39–48 (2010).
- Pagon, R. A. *et al.* Charcot-Marie-Tooth Hereditary Neuropathy Overview (1993).
- Fink, J. K. Hereditary spastic paraplegia: clinico-pathologic features and emerging molecular mechanisms. *Acta Neuropathol.* **126**, 307–328 (2013).
- Timmerman, V., Strickland, A. & Zuchner, S. Genetics of Charcot-Marie-Tooth (CMT) Disease within the Frame of the Human Genome Project Success. *Genes* **5**, 13–32 (2014).
- Pareyson, D., Saveri, P. & Pisciotta, C. New developments in Charcot-Marie-Tooth neuropathy and related diseases. *Curr. Opin. Neurol.* **30**, 471–480 (2017).
- Schule, R. *et al.* Hereditary spastic paraplegia: Clinicogenetic lessons from 608 patients. *Ann Neurol.* **79**, 646–658 (2016).
- Blackstone, C. Cellular Pathways of Hereditary Spastic Paraplegia. *Annu. Rev. Neurosci.* **35**, 25–47 (2012).
- Gonzalez, M. Identifying Novel Genes and Genetic/Phenotypic Spectra for Inherited Neurodegenerative Disorders Using Next-Generation Sequencing. 1–167 (2015).
- Barabási, A.-L., Gulbahce, N. & Loscalzo, J. Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.* **12**, 56–68 (2011).
- Limviphuvadh, V., Tanaka, S., Goto, S., Ueda, K. & Kanehisa, M. The commonality of protein interaction networks determined in neurodegenerative disorders (NDDs). *Bioinformatics* **23**, 2129–2138 (2007).
- Ghiassian, S. D., Menche, J. & Barabási, A.-L. A Disease Module Detection (DIAMOND) algorithm derived from a systematic analysis of connectivity patterns of disease proteins in the human interactome. *PLoS Comput. Biol.* **11**, e1004120 (2015).
- Sharma, A. *et al.* A disease module in the interactome explains disease heterogeneity, drug response and captures novel pathways and genes in asthma. *Hum. Mol. Genet.* **24**, 3005–3020 (2015).
- Novarino, G. *et al.* Exome Sequencing Links Corticospinal Motor Neuron Disease to Common Neurodegenerative Disorders. *Science* **343**, 506–511 (2014).
- Alanis-Lobato, G., Andrade-Navarro, M. A. & Schaefer, M. H. HIPPIEv2.0: enhancing meaningfulness and reliability of protein-protein interaction networks. *Nucleic Acids Res.* gkw985, <https://doi.org/10.1093/nar/gkw985> (2016).
- Menche, J. *et al.* Uncovering disease-disease relationships through the incomplete interactome. *Science* **347**, 1257601–1257601 (2015).
- Hamosh, A., Scott, A. F., Amberger, J., Valle, D. & McKusick, V. A. Online Mendelian Inheritance in Man (OMIM). *Hum. Mutat.* **15**, 57–61 (2000).
- Hwang, Y.-C. *et al.* HIPPIE: a high-throughput identification pipeline for promoter interacting enhancer elements. *Bioinformatics* **31**, 1290–1292 (2015).
- Fiscon, G., Conte, F., Farina, L. & Paci, P. Network-Based Approaches to Explore Complex Biological Systems towards Network Medicine. *Genes* **9**, 437 (2018).
- Ghiassian, S. D. *et al.* Endophenotype Network Models: Common Core of Complex Diseases. *Nature Publishing Group* 1–13, <https://doi.org/10.1038/srep27414> (2016).
- Bezhtentsev, V., Ivanov, S., Kumar, S., Goel, R. & Poroikov, V. Identification of potential drug targets for treatment of refractory epilepsy using network pharmacology. *Journal of Bioinformatics and Computational Biology* **16**, 1840002 (2018).
- Li, J.-W. *et al.* Interactome-transcriptome analysis discovers signatures complementary to GWAS Loci of Type 2 Diabetes. *Sci Rep* **6** (2016).
- Hu, Y. *et al.* Detecting pathway relationship in the context of human protein-protein interaction network and its application to Parkinson's disease. *Methods* **131**, 93–103 (2017).
- Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
- Mistry, M. & Pavlidis, P. Gene Ontology term overlap as a measure of gene functional similarity. *BMC Bioinformatics* **9**, 327–11 (2008).
- Yu, G. *et al.* GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics* **26**, 976–978 (2010).
- Pesquita, C. *et al.* Metrics for GO based protein semantic similarity: a systematic evaluation. *BMC Bioinformatics* **9**(Suppl 5), S4 (2008).
- Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
- Kutmon, M. *et al.* WikiPathways: Capturing the full diversity of pathway knowledge. *Nucleic Acids Res.* **44**, D488–D494 (2016).
- Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).

36. Wang, J., Vasaikar, S., Shi, Z., Greer, M. & Zhang, B. WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res.* **45**, W130–W137 (2017).
37. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G. D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS ONE* **5**, e13984 (2010).
38. Goh, K. I. *et al.* The human disease network. *Proc Natl Acad Sci USA* **104**, 8685–8690 (2007).
39. Maciel, R. *et al.* The human motor neuron axonal transcriptome is enriched for transcripts related to mitochondrial function and microtubule-based axonal transport. *Exp. Neurol.* 1–11, <https://doi.org/10.1016/j.expneurol.2018.06.008> (2018).
40. Oti, M., Snel, B., Huynen, M. A. & Brunner, H. G. Predicting disease genes using protein-protein interactions. *J. Med. Genet.* **43**, 691–698 (2006).
41. Razić, S., Magkharas, G. & Donaldson, I. M. iRefIndex: a consolidated protein interaction database with provenance. *BMC Bioinformatics* **9**, 405 (2008).
42. Lee, I., Blom, U. M., Wang, P. I., Shim, J. E. & Marcotte, E. M. Prioritizing candidate disease genes by network-based boosting of genome-wide association data. *Genome Research* **21**, 1109–1121 (2011).
43. Szklarczyk, D. *et al.* STRINGv10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447–52 (2015).
44. Synofzik, M. & Schule, R. Overcoming the divide between ataxias and spastic paraplegias: Shared phenotypes, genes, and pathways. *Mov Disord.* **32**, 332–345 (2017).
45. Nakao, A., Yoshihama, M. & Kenmochi, N. RPG: The Ribosomal Protein Gene database. *Nucleic Acids Res.* **32** (2004).
46. Marygold, S. J. *et al.* The ribosomal protein genes and Minute loci of *Drosophila melanogaster*. *Genome Biology* **8**, R216 (2007).
47. Warner, J. R. & McIntosh, K. B. How Common Are Extraribosomal Functions of Ribosomal Proteins? *Molecular Cell* **34**, 3–11 (2009).
48. Uechi, T. *et al.* Ribosomal protein gene knockdown causes developmental defects in zebrafish. *PLoS ONE* **1**, e37 (2006).
49. McGowan, K. A. *et al.* Ribosomal mutations cause p53-mediated dark skin and pleiotropic effects. *Nat. Genet.* **40**, 963–970 (2008).
50. Martin, I. *et al.* Ribosomal protein s15 phosphorylation mediates LRRK2 neurodegeneration in Parkinson's disease. *Cell* **157**, 472–485 (2014).
51. Niehues, S. *et al.* Impaired protein translation in *Drosophila* models for Charcot-Marie-Tooth neuropathy caused by mutant tRNA synthetases. *Nat Commun* **6**, 7520 (2015).
52. Kondrashov, N. *et al.* Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell* **145**, 383–397 (2011).
53. Jung, H., Yoon, B. C. & Holt, C. E. Axonal mRNA localization and local protein synthesis in nervous system assembly, maintenance and repair. *Nature Reviews Neuroscience* **13**, 308–324 (2012).
54. Bernard-Marissal, N., Chrast, R. & Schneider, B. L. Endoplasmic reticulum and mitochondria in diseases of motor and sensory neurons: a broken relationship? *Cell Death & Disease* **9**, 292 (2018).
55. Krols, M. *et al.* Mitochondria-associated membranes as hubs for neurodegeneration. *Acta Neuropathol.* **131**, 505–523 (2016).
56. Yamada, Y., Ohazama, A., Maeda, T. & Seo, K. The Sonic Hedgehog signaling pathway regulates inferior alveolar nerve regeneration. *Neuroscience Letters* **671**, 114–119 (2018).
57. Acosta, M. C., Copley, P. A., Harrell, J. R. & Wilhelm, J. C. Estrogen signaling is necessary for exercise-mediated enhancement of motoneuron participation in axon regeneration after peripheral nerve injury in mice. *Dev Neurobiol* **77**, 1133–1143 (2017).
58. McMurray, R., Islamov, R. & Murashov, A. K. Raloxifene analog LY117018 enhances the regeneration of sciatic nerve in ovariectomized female mice. *Brain Res.* **980**, 140–145 (2003).
59. Rowin, J., Xia, Y., Jung, B. & Sun, J. Gut inflammation and dysbiosis in human motor neuron disease. *Physiol Rep* **5**, e13443 (2017).
60. Zhang, Y. G. *et al.* Target Intestinal Microbiota to Alleviate Disease Progression in Amyotrophic Lateral Sclerosis. *Clin Ther* **39**, 322–336 (2017).
61. Wu, S., Yi, J., Zhang, Y. G., Zhou, J. & Sun, J. Leaky intestine and impaired microbiome in an amyotrophic lateral sclerosis mouse model. *Physiol Rep* **3**, e12356 (2015).
62. Zuchner, S. *et al.* Mutations in the mitochondrial GTPase mitofusin 2 cause Charcot-Marie-Tooth neuropathy type 2A. *Nat. Genet.* **36**, 449–451 (2004).
63. Loiseau, D. *et al.* Mitochondrial coupling defect in Charcot-Marie-Tooth type 2A disease. *Ann Neurol.* **61**, 315–323 (2007).
64. Cassereau, J. *et al.* Mitochondrial dysfunction and pathophysiology of Charcot-Marie-Tooth disease involving GDAP1 mutations. *Exp. Neurol.* **227**, 31–41 (2011).
65. Kennerson, M. L. *et al.* A new locus for X-linked dominant Charcot-Marie-Tooth disease (CMTX6) is caused by mutations in the pyruvate dehydrogenase kinase isoenzyme 3 (PDK3) gene. *Hum. Mol. Genet.* **22**, 1404–1416 (2013).
66. Zala, D. *et al.* Vesicular glycolysis provides on-board energy for fast axonal transport. *Cell* **152**, 479–491 (2013).
67. Hinkelmann, M.-V., Zala, D. & Saudou, F. Releasing the brake: restoring fast axonal transport in neurodegenerative disorders. *Trends Cell Biol.* **23**, 634–643 (2013).
68. Pourcelot, M. & Arnoult, D. Mitochondrial dynamics and the innate antiviral immune response. *FEBS J.* **281**, 3791–3802 (2014).
69. West, A. P. Mitochondrial dysfunction as a trigger of innate immune responses and inflammation. *Toxicology* **391**, 54–63 (2017).
70. Giudice, L. T., Lombardi, F., Santorelli, F. M., Kawarai, T. & Orlacchio, A. Hereditary spastic paraplegia: clinical-genetic characteristics and evolving molecular mechanisms. *Exp. Neurol.* **261**, 518–539 (2014).
71. Virtanen, J. O. & Jacobson, S. Viruses and multiple sclerosis. *CNS Neurol Disord Drug Targets* **11**, 528–544 (2012).
72. DeLuca, G. C. *et al.* The role of hereditary spastic paraplegia related genes in multiple sclerosis. *J. Neurol.* **254**, 1221–1226 (2007).
73. Yazıcı, I., Yildirim, N. & Zorlu, Y. The coexistence of multiple sclerosis and hereditary spastic paraparesis in a patient. *Neurol Int* **5**, 17–19 (2013).
74. Mead, S. H. *et al.* A large family with hereditary spastic paraparesis due to a frame shift mutation of the spastin (SPG4) gene: association with multiple sclerosis in two affected siblings and epilepsy in other affected family members. *J. Neurol. Neurosurg. Psychiatr.* **71**, 788–791 (2001).
75. Quan, Y., Wang, Z.-Y., Chu, X.-Y. & Zhang, H.-Y. Evolutionary and genetic features of drug targets. *Med Res Rev* **38**, 1536–1549 (2018).
76. Finan, C. *et al.* The druggable genome and support for target identification and validation in drug development. *Sci Transl Med* **9** (2017).

Acknowledgements

Funded by the Inherited Neuropathy Consortium – Rare Disease Clinical Research Consortium (INC RDCRC - U54NS065712), NINDS/ORDR, NCATS, the Muscular Dystrophy Association (MDA), the Charcot Marie Tooth Association (CMTA), NINDS R01NS075764, and NINDS F31NS106859.

Author Contributions

D.B., M.D., S.W. and S.Z. conceived and planned the experiments. D.B. performed the experiments and prepared the manuscript with support from M.D., S.W. and S.Z.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-37119-z>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019