

A Protein–Protein Interaction Network for Human Inherited Ataxias and Disorders of Purkinje Cell Degeneration

Janghoo Lim,¹ Tong Hao,⁴ Chad Shaw,¹ Akash J. Patel,¹ Gábor Szabó,^{4,5} Jean-François Rual,⁴ C. Joseph Fisk,¹ Ning Li,⁴ Alex Smolyar,⁴ David E. Hill,⁴ Albert-László Barabási,^{4,5} Marc Vidal,⁴ and Huda Y. Zoghbi^{1,2,3,*}

¹Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA

²Departments of Pediatrics and Neuroscience, Baylor College of Medicine, Houston, TX 77030, USA

³Howard Hughes Medical Institute, USA

⁴Center for Cancer Systems Biology and Department of Cancer Biology, Dana-Farber Cancer Institute and Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

⁵Center for Complex Network Research and Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA

*Contact: hzoghbi@bcm.tmc.edu

DOI 10.1016/j.cell.2006.03.032

SUMMARY

Many human inherited neurodegenerative disorders are characterized by loss of balance due to cerebellar Purkinje cell (PC) degeneration. Although the disease-causing mutations have been identified for a number of these disorders, the normal functions of the proteins involved remain, in many cases, unknown. To gain insight into the function of proteins involved in PC degeneration, we developed an interaction network for 54 proteins involved in 23 inherited ataxias and expanded the network by incorporating literature-curated and evolutionarily conserved interactions. We identified 770 mostly novel protein–protein interactions using a stringent yeast two-hybrid screen; of 75 pairs tested, 83% of the interactions were verified in mammalian cells. Many ataxia-causing proteins share interacting partners, a subset of which have been found to modify neurodegeneration in animal models. This interactome thus provides a tool for understanding pathogenic mechanisms common for this class of neurodegenerative disorders and for identifying candidate genes for inherited ataxias.

INTRODUCTION

Many human inherited neurodegenerative disorders, such as the spinocerebellar ataxias (SCAs), are characterized by cerebellar Purkinje cell (PC) degeneration that causes ataxia, or loss of balance and coordination (Sakaguchi et al., 1996; Engert et al., 2000; Zoghbi and Orr, 2000; Moore et al., 2001; Sun et al., 2001; Matsuda et al., 2004; Taroni and DiDonato, 2004). Similarly, several ataxic

mouse mutants also display PC degeneration phenotypes (Hamilton et al., 1996; Fernandez-Gonzalez et al., 2002; Klein et al., 2002; Isaacs et al., 2003; Noveroske et al., 2005). The inherited ataxias are caused by either gain-of-function or loss-of-function mutations in seemingly unrelated genes (Zoghbi and Orr, 2000; Taroni and DiDonato, 2004). To date, genes have been identified for over 23 inherited ataxias, but the normal function of the majority of ataxia-causing proteins is poorly understood. Furthermore, although several human and mouse ataxias share similar motor dysfunction and cell-specific neuropathology, the molecular mechanism(s) mediating such overlapping features are largely unknown. Independent disease-based studies have revealed the importance of protein–protein interactions in understanding the normal function of the disease-causing protein and are beginning to identify pathways that could be targeted therapeutically (Stefan et al., 2001; Yoshida et al., 2002; Chen et al., 2003; Goehler et al., 2004; Ravikumar et al., 2004; Kaytor et al., 2005; Tsuda et al., 2005).

Because of the overlap in phenotypes and the prominence of PC pathology in human and mouse ataxias, we hypothesized that the gene products involved in this class of neurodegenerative diseases might play roles in common molecular pathways that are essential for PC function and survival. Furthermore, we reasoned that insight into the cellular functions of ataxia-causing proteins would deepen our understanding of the molecular events underlying the PC pathology. One approach to characterizing unknown proteins is to identify interacting protein partners and link the interacting pairs to known cellular pathways. A first-draft human protein–protein interaction network has recently been generated (Rual et al., 2005; Stelzl et al., 2005). Undoubtedly, this network will be a valuable resource for many biological studies. Nevertheless, to uncover pathogenic mechanisms requires characterizing in greater detail the network around specific groups of proteins that are implicated in a particular class of disorders. We therefore set out to develop a proteome-scale

protein–protein interaction network for ataxia-causing proteins. Such a network will have more depth around proteins involved in a specific biological problem and will be more useful to disease-oriented researchers than the larger human interaction networks. Furthermore, phenotype-based interaction networks are more likely to bring out some of the key interactions mediating pathogenesis than the broader and less deeply developed entire proteome networks.

In this study, we developed a stringent protein–protein interaction network for the 20 or more different inherited cerebellar ataxias characterized by PC degeneration. This phenotype-based interactome network revealed several previously unsuspected interactions between the various ataxia-causing proteins. Bioinformatic analysis showed high connectivity between different ataxia-causing proteins and revealed common cellular pathways that might lead to PC dysfunction and degeneration. Importantly, many genetic modifier proteins identified by studies using mouse and *Drosophila* disease models showed direct physical interactions with the corresponding disease protein, linking known genetic pathways within the protein interaction network. Finally, this protein interaction network provides candidate genes for cerebellar ataxias whose genetic defect has not yet been identified.

RESULTS

Three Yeast Two-Hybrid Screens for Proteins Involved in Inherited Ataxias

To generate a proteome-scale interactome network for inherited ataxias, we selected 23 human cDNAs encoding proteins directly involved in 23 different types of inherited ataxias (Table 1). We refer to these as “ataxia-causing” proteins to indicate that mutations in the genes encoding them cause ataxia in humans or mice. This list includes 11 recessive and 12 dominant ataxias encompassing the polyglutamine-mediated SCAs. We also selected 31 additional proteins that either interact with the ataxia-causing proteins or are paralogs of such proteins (Table S1). We refer to this extended group as “ataxia-associated proteins.” A total of 122 full-length or partial open reading frames (ORFs) encoding 54 different proteins were cloned into Gateway-compatible yeast two-hybrid Y2H destination vectors to generate GAL4 DNA binding domain (DB)- or activation domain (AD)-fusion proteins, referred to as DB-ataxia or AD-ataxia constructs, respectively (Table S2). Using these clones, we screened both the human ORFeome v1.1 (hORFeome) (Rual et al., 2004; Rual et al., 2005) and adult brain cDNA libraries (Figure 1A).

To screen the hORFeome, we performed matrix-based mating type screens in 96-well format. The DB-ataxia transformed MaV203 yeasts were individually mated with MaV103 AD-188 ORFs minilibrary pools (Rual et al., 2005). As a second round of screening, we tested reciprocal pair-wise interactions between AD-ataxia preys and DB-hORFeome baits. We chose to screen both the AD- and DB-hORFeome clones for two reasons. First, we recognized that the folding of the fusion protein in Y2H vectors

might vary depending on the vector backbone. Second, by using AD-ataxia clones we could include autoactivating baits in our screen. From these reciprocal hORFeome screens, we identified 269 potential protein–protein interactions involving 36 different ataxia-associated proteins (Figure 1B; Table S3). We found 14 interactions in common between the AD- and DB-hORFeome screens. The overlap comprises 5.2% (14/269) of the observed interactions, which is typical for reciprocal ORFeome studies (Rual et al., 2005).

Because the inherited ataxias affect the central nervous system, we reasoned that it is important to screen a library that represents the affected tissue. We performed a third screen using adult human brain cDNA library and identified 530 potential protein–protein interactions (Figure 1B; Table S3).

In total, the hORFeome and the brain cDNA library screens revealed 770 protein–protein interactions involving 42 ataxia-associated baits and 561 interacting prey proteins (Figure 1B; Table S3). Interestingly, 29 pairs of protein–protein interactions were identified as common to both screens, revealing the advantage of performing hORFeome and brain cDNA library screens in parallel. The vast majority (741/770 = 96%) of the Y2H interactions we identified are novel, with only 29 interactions having been reported previously (Table S4). We subdivided the Y2H interactions into two classes based on the number of clones identified for each interacting protein (Table S3). Core-1 interacting proteins were those for which more than three independent clones were found, whereas Core-2 had less than three clones.

Validation of Y2H Interactions by Coaffinity Purification and Computational Analyses

To validate the Y2H interactions, we randomly sampled interacting pairs and tested them using coaffinity purification co-AP glutathione-S-transferase (GST) pull-down assays in human HEK293T cells (Li et al., 2004; Rual et al., 2005). Proteins that have more interactions than others appeared more frequently in the sampled set. Of 75 interactions (about 10% of all interactions) assayed, 62 (or ~83%) were confirmed by the co-AP assays (Figures 2 and S1 and Table S3). The co-AP results showed a similar success rate between Core-1 (80%) and Core-2 (86%) datasets (Table S3). The similarly high success rate of co-AP validation suggests that the quality of the Y2H screens is very good and that both Core-1 and Core-2 Y2H interactions are reliable.

In addition to the experimental validation, we performed bioinformatic analyses to confirm the experimental Y2H interactions (see Supplemental Data). We implemented a coannotation analysis using the cellular component branch of Gene Ontology (GO). We found that 72% of the Y2H interactions that have known compartment annotations involve proteins that colocalize to the same compartment. This is a high colocalization rate, given that GO annotations are not comprehensive and that biological reagents to detect all possible subcellular localizations are not available. We extended this analysis to examine the

biological process and molecular function branches of the GO. We found that 98% of the Y2H interactions are coannotated in at least one GO branch.

Protein–Protein Interaction Network for Inherited Cerebellar Ataxias

Analysis of the Y2H interaction data revealed one large interconnected network consisting of 752 protein–protein interactions between 36 ataxia-associated proteins and 541 prey proteins (Figure 3). The finding of a single dominant connected component in our network suggests that proteins involved in inherited cerebellar ataxias are functionally linked. Thirteen inherited ataxia-causing bait proteins were linked either directly or through common interacting proteins within the large component. In contrast, ten ataxia-causing proteins had only a few or no Y2H interactors and were isolated from the main component (Figure 3). These ten proteins were PSAP, RORA, TBP, AGTPBP1, CACNA1A, ATXN10, PPP2R2B, ATM, PRND, and AF4. These proteins may have had a limited number of partners or been isolated from the main component because they were represented each by a single full-length clone (Table S2). This raises the possibility that the baits might not have been ideal for screening and emphasizes the importance of using multiple overlapping clones per protein to increase the likelihood of generating functional or properly folded Y2H fusion proteins. Another possible explanation is that the pathogenesis of some of these inherited ataxias may be less directly related to that of other ataxias.

Identification of Literature-Curated and Evolutionarily Conserved Potential Interactions

In order to expand the interaction dataset, we added relevant direct protein–protein interactions from currently available human protein–protein interaction networks (Rual et al., 2005; Stelzl et al., 2005). We also searched public databases, including BIND (Bader et al., 2003), DIP (Xenarios et al., 2002), HPRD (Peri et al., 2003), MINT (Zanzoni et al., 2002), and MIPS (Pagel et al., 2005), to identify literature-based binary interactions involving the 54 ataxia-associated baits and the 561 interacting prey proteins. We identified 4796 binary protein–protein interactions for our Y2H baits and prey proteins (Table S4) and incorporated them in the Y2H protein–protein interaction map (Figures 4A–4C).

In addition, we searched for potentially conserved interactions, or “interologs,” whose ortholog pairs are known to interact in other organisms (Walhout et al., 2000). Using the InParanoid database (<http://inparanoid.cgb.ki.se>) with confidence score 1.0, we examined the 615 human proteins (54 baits and 561 preys identified from Y2H screen) against the yeast, worm, fly, and mouse predicted proteomes. We identified 1527 potential human interologs from one or more species, and 92/1527 interologs were known to interact with each other in human (Figures 4A–4C and Table S4).

Coannotation analysis revealed that 68% and 63% of the literature-curated and interolog interactions, respectively,

are annotated to similar GO compartments (see [Supplemental Data](#)). This suggests that the literature-curated and interolog interactions are of a quality similar to interactions identified in our Y2H screens.

In total, we identified 6972 pairs of potential protein–protein interactions between 3607 proteins, including 18 proteins out of 23 ataxia-causing proteins used for the initial Y2H screens (Figures 4A–4C). These 6972 interactions include 119 homodimerization interactions. All protein–protein interactions are first-, second-, or third-order interactions with ataxia-causing proteins. This ataxia interaction network comprises two integrated components that we call the “direct” ataxia network and the “expanded” ataxia network. The direct ataxia network contains direct, first-order protein interactions with ataxia-causing proteins, whereas the expanded network contains additional second- or third-order interactions.

Properties of the Ataxia-Based Network

We performed a number of analyses to explore the properties of our ataxia network. The controls for comparison with our results are networks constructed from a list of proteins associated with a phenotypically diverse group of disorders. In addition, we performed a large-scale simulation study. These control networks, constructed using an approach similar to the one employed in our study (see [Supplemental Data](#)), determine a space of protein interaction networks that are nucleated around disease proteins but where no bias for a particular disease phenotype should influence the character of the network. Importantly, the control networks have similar numbers of proteins, interactors, and connections to the ataxia network. Such a control permits us to analyze the extent to which the phenotype-based ataxia network has unusual topological properties.

To assess our ataxia network against this control, we calculated the mean shortest-path length between the ataxia-causing proteins in our network to be 3.11. This path length is significantly smaller ($p < .0004$) than the path length (3.67) between disease proteins in the network nucleated around 30 disease proteins sampled independent of phenotypes. The simulated networks showed the same property (Figure 4D).

In addition to the mean shortest path length, we searched for proteins linking pairs of ataxia-causing proteins in relations we call disease triples. Central proteins in a triple directly interact with two different ataxia-causing proteins in the pattern: Ataxia protein1–interactor–Ataxia protein2 (Table S5) as well as quadruples (two interacting nodes that also bind to two different ataxia-causing proteins: Ataxia protein1–interactor1–interactor2–Ataxia protein2) (Table S6). There were 63 triples and 608 quadruples in the ataxia network. In contrast, there were only 42 triples in the manually constructed, varied phenotype disease network, and the simulation-based networks had a mean of 31 triples. The number of triples in the ataxia interactome was significantly higher ($p < .005$).

Table 1. Inherited Ataxias and Ataxic Mouse Mutants

Symbol	Entrez Gene ID	Proteins	Inherited Ataxias or Ataxic Mutant Mice	Disease and Mutational Basis	OMIM No.
ATXN1	6310	Ataxin 1	Spinocerebellar ataxia type 1 (SCA1)	Autosomal dominant cerebellar ataxia, CAG repeats/coding	164400
ATXN2	6311	Ataxin 2	Spinocerebellar ataxia type 2 (SCA2)	Autosomal dominant cerebellar ataxia, CAG repeats/coding	183090
ATXN3	4287	Ataxin 3	Spinocerebellar ataxia type 3 (SCA3) / Machado-Joseph disease	Autosomal dominant cerebellar ataxia, CAG repeats/coding	109150
CACNA1A	773	Calcium channel, voltage-dependent, P/Q type, alpha 1A subunit	Spinocerebellar ataxia type 6 (SCA6)	Autosomal dominant cerebellar ataxia, CAG repeats/coding	183086
ATXN7	6314	Ataxin 7	Spinocerebellar ataxia type 7 (SCA7)	Autosomal dominant cerebellar ataxia, CAG repeats/coding	164500
TBP	6908	TATA box binding protein	Spinocerebellar ataxia type 17 (SCA17)	Autosomal dominant cerebellar ataxia, CAG repeats/coding	607136
ATN1	1822	Atrophin 1	Dentatorubropallidoluysian atrophy (DRPLA)	Autosomal dominant cerebellar ataxia, CAG repeats/coding	125370
ATXN10	25814	Ataxin 10	Spinocerebellar ataxia type 10 (SCA10)	Autosomal dominant cerebellar ataxia, ATTCT repeats/Intronic	603516
PPP2R2B	5521	Protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), beta isoform	Spinocerebellar ataxia type 12 (SCA12)	Autosomal dominant cerebellar ataxia, CAG expansion/5'-UTR	604326
PRKCG	5582	Protein kinase C, gamma	Spinocerebellar ataxia type 14 (SCA14)	Autosomal dominant cerebellar ataxia, Missense mutations	605361
FXN	2395	Frataxin	Friedreich ataxia (FRDA)	Autosomal recessive ataxia, GAA expansion/Intronic	229300
ATM	472	Ataxia telangiectasia mutated	Ataxia telangiectasia (A-T)	Autosomal recessive cerebellar ataxia, Point mutations, deletions	208900
AGTPBP1	23287	ATP/GTP binding protein 1	Purkinje cell degeneration (Nna1 pcd, pcd) mutant mouse	Autosomal recessive cerebellar ataxia	606830
APTX	54840	Aprataxin	Ataxia with oculomotor apraxia (AOA1)	Autosomal recessive cerebellar ataxia, Insertion, deletion, missense mutations	208920
SACS	26278	Saccsin	Spastic ataxia of Charlevoix-Saguenay	Autosomal recessive spastic ataxia, Deletion	270550

Table 1. Continued

Symbol	Entrez Gene ID	Proteins	Inherited Ataxias or Ataxic Mutant Mice	Disease and Mutational Basis	OMIM No.
AIF (PDCD8)	9131	Programmed cell death 8 (apoptosis-inducing factor)	Harlequin (Hq) mutant mouse	X-linked, Progressive neuronal degeneration in cerebellum, Proviral insertion	300169
RORA	6095	RAR-related orphan receptor A	Staggerer (sg) mouse	Autosomal recessive cerebellar ataxia, Deletion	600825
PSAP	5660	Prosaposin	Purkinje cell degeneration	Autosomal recessive Purkinje cell degeneration by Saposin D knockout	176801
PRND	23627	Prion protein 2 (dublet)	Purkinje cell degeneration	Autosomal dominant Purkinje cell degeneration by increased expression of Doppel	604263
PRNP	5621	Prion protein (p27-30) (Creutzfeld-Jakob disease, Gerstmann-Strausler-Scheinker syndrome, fatal familial insomnia)	Purkinje cell degeneration	Autosomal recessive Purkinje cell degeneration	176640
AF4	4299	Myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 2	Robotic mouse	Autosomal dominant cerebellar ataxia, Missense mutations	159557
ERCC5	2073	Excision repair cross-complementing rodent repair deficiency, complementation group 5 (xeroderma pigmentosum, complementation group G [Cockayne syndrome])	Purkinje cell degeneration	Autosomal recessive ataxia	133530
QKI	9444	Quaking homolog, KH domain RNA binding (mouse)	Quaking mouse	Autosomal recessive mouse mutation, ataxia	609590

To determine whether the properties of the ataxia network are compatible with the known topological characteristics of protein interaction networks (Barabasi and Oltvai, 2004; Albert, 2005), we investigated the degree distribution $P(k)$ of the participating proteins and found that our disease network shows a scale-free degree distribution similar to that of other interactome studies of model

organisms. Yet as Figure 4E shows, for the combined two-hybrid and literature-curated human interactome (Rual et al., 2005), the value of the degree exponent is 2.7 ± 0.2 , considerably higher than the value 2.2 ± 0.1 obtained for the ataxia network. This is somewhat unexpected, as a topologically unbiased subset of a scale-free network should display the same exponent as the

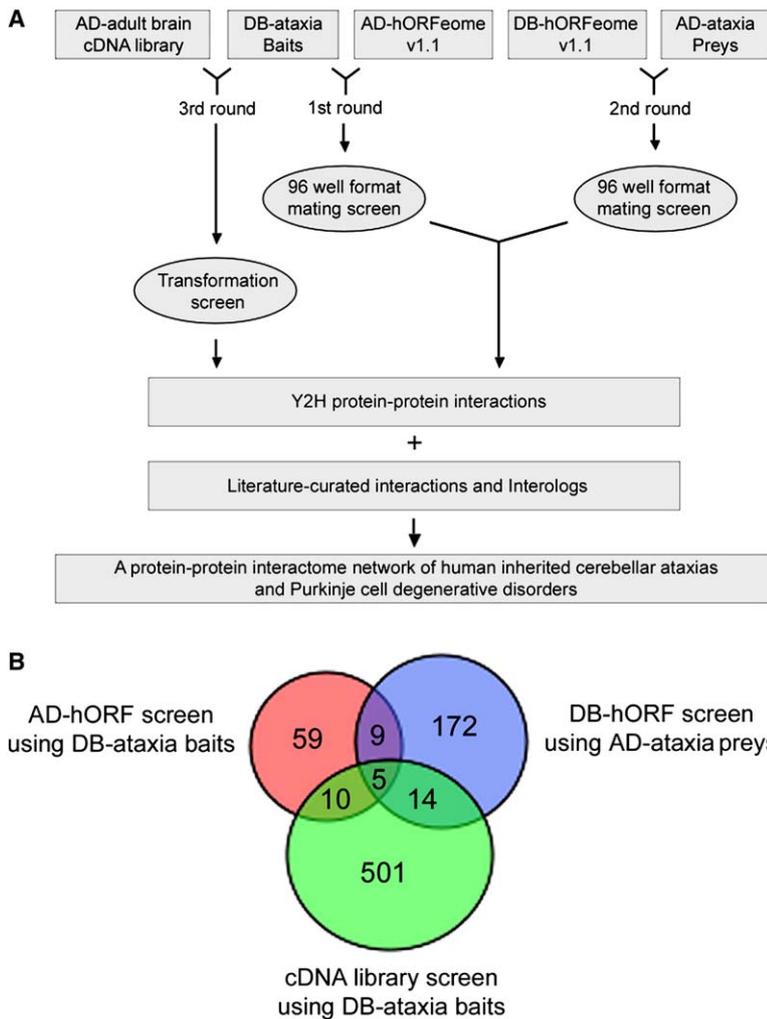


Figure 1. Generation of a Protein Interaction Network for Inherited Cerebellar Ataxias

(A) Schematic diagram illustrates two major steps: (1) identification of Y2H interactions from three Y2H screens of the hORFeome and the adult human brain cDNA library; (2) searching for and adding literature-curated and evolutionarily conserved interactions, Interologs.

(B) Experimental Y2H interaction data for 54 ataxia-associated proteins.

full network. As we show in the Supplementary Data, such a difference emerges each time the network is constructed around a selected set of proteins. Indeed, high-degree proteins will be overrepresented in the sample, as the probability that they will interact with the ataxia proteins is proportional to their degree k . Therefore, the exponent of the ataxia subnetwork should be one less than the exponent of the full network. Given, however, that the protein interaction network has a hierarchical structure (Yook et al., 2004) and displays degree-degree correlations (Maslov and Sneppen, 2002), the difference between the two exponents will be less than one. Based on these data, we predict that future studies that focus on the neighborhood of a selected group of proteins will likely encounter a similar effect, resulting in an exponent that is smaller than the exponent obtained for the whole network. From a biological perspective, this means that in all studies focused on the neighborhood of selected disease proteins, the hubs will be overrepresented in comparison to a random sample of a similar number of proteins.

Biological Functions of Proteins in the Ataxia Network

To ascertain which biological functions are associated with cerebellar ataxias and PC degeneration, we applied GO term enrichment analysis to the ataxia network (Young et al., 2005). GO term analysis was applied to the network as a whole as well as to 44 cohorts. The cohorts were defined as those protein sets that interact directly with each ataxia-causing protein as well as each collection of first- and second-order interactors (the union of the proteins that directly interact with ataxia-causing proteins as well as their interactors). GO analysis was also applied to the proteins linking any pair of ataxia-causing proteins through a ternary or quaternary sequence of interactions. The full collection of results can be viewed at <http://franklin.imgen.bcm.tmc.edu/ppi/>.

Applying the most stringent criteria, we identified 56 GO biological process categories, 13 GO cellular component categories, and 13 molecular functions as overrepresented in the ataxia network compared to what would be expected from a random collection of proteins

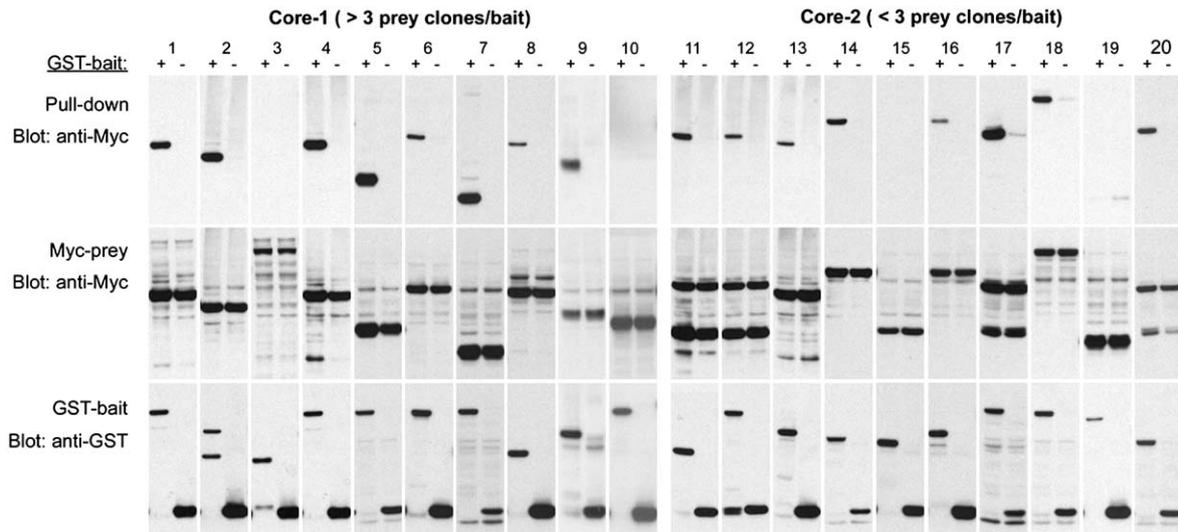


Figure 2. Verification of Y2H Interactions by co-AP Assays in Human Cells

Ten representative examples of co-AP assays from Core-1 and Core-2 Y2H datasets are shown. Top panels show expression of Myc-tagged prey proteins after affinity purification on glutathione-Sepharose 4B bead, demonstrating binding to GST-tagged bait proteins (+) or to GST alone (–). Middle and bottom panels show expression levels of Myc-prey and GST-bait proteins, respectively. Interacting pairs are shown in Table S3 with the lane number. All figures of co-AP assays have similar configurations unless otherwise indicated.

(Figure 5). The biological process GO categories with significantly enhanced representation in the expanded ataxia network include transcriptional regulation, RNA splicing, ubiquitination, cell cycle, and a few others. The cellular component categories enriched in the GO term analysis include the nucleus, proteasome complex, spliceosome complex, and nuclear membrane. The molecular function enriched GO categories include transcription factor binding, transcription cofactor activity, DNA binding, and kinase inhibitor activity. The GO term analysis data are consistent with previous genetic studies illustrating that many genetic modifiers belong to GO categories such as transcriptional regulation, RNA splicing, and protein degradation (Fernandez-Funez et al., 2000; Cummings et al., 2001; Zoghbi and Botas, 2002; Ghosh and Feany, 2004).

Comparing the Inherited Ataxia Network with Human Diseases in the OMIM Database

To see if any newly identified interacting proteins in the ataxia network are associated with human diseases, we searched the OMIM disease database. Interestingly, we uncovered Puratrophin-1 (Puratrophin-1; also called DKFZP434I216 or PLEKHG4), a pleckstrin homology domain containing guanine-nucleotide exchange factor (GEF) for Rho GTPases, as a Coilin (COIL)-interacting protein in our Y2H screen (Table S3). Puratrophin-1 was recently implicated in the etiology of the 16q22.1-linked autosomal dominant cerebellar ataxia (OMIM #117210), and the protein was found in aggregates in PCs (Ishikawa et al., 2005). Interestingly, our data link Puratrophin-1 to Ataxin-1 (ATXN1), which causes spinocerebellar ataxia type 1 (SCA1), through interactions with

COIL (Figure 6A and Table S4). This suggests that SCA1 and the 16q22.1-linked autosomal dominant cerebellar ataxia might share certain disease processes. This finding also suggests that our network could contain candidate genes for ataxias whose genes have not yet been identified. Amongst the primary interactors of the ataxia-causing proteins, 57 proteins are associated with human diseases listed in OMIM. The expanded network, which includes secondary interactions and beyond, contains 496 of the disease proteins listed in OMIM (Table S4). Thus, the network should be valuable for investigators studying these diseases.

Linking Genetic Modifiers with Protein–Protein Interaction Networks

To verify the *in vivo* relevance of the ataxia network, we turned our attention to animal models of inherited ataxias to see if any of the identified interacting proteins affect the neurodegenerative phenotypes. Genetic modifier screens using *Drosophila* and mouse disease models have identified factors that modulate the toxicity of disease-causing proteins, although their biochemical and molecular relationship to the disease proteins is not known even in the model organisms (Fernandez-Funez et al., 2000; Kazemi-Esfarjani and Benzer, 2000; Zoghbi and Botas, 2002; Ghosh and Feany, 2004). These modifiers belong to several different classes that include protein folding and degradation pathways, RNA metabolism, and transcription. Many human orthologs of the genetic modifiers previously described in *Drosophila* and mouse SCA1 and SCA3 models (Fernandez-Funez et al., 2000; Ghosh and Feany, 2004) were identified in our Y2H screen as potential protein

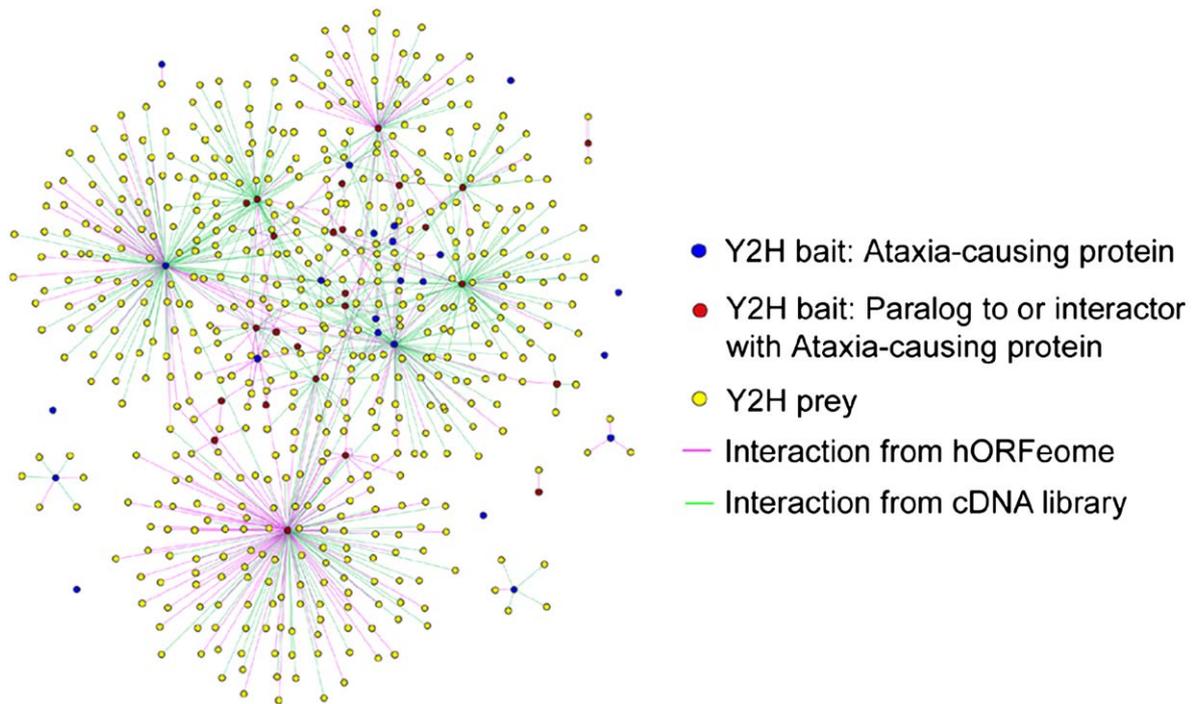


Figure 3. Network Graph of the Y2H Interactions

Blue nodes depict proteins corresponding to genes mutated in ataxias; red nodes depict proteins that are paralogs or interact with ataxia-causing proteins; and yellow nodes depict Y2H interactors from this screen. Purple and green edges show protein interactions from the hORFeome and brain cDNA library screens, respectively. All edges are either first- or second-order interactions to ataxia-causing proteins.

interactors of the disease-causing proteins (Figure 6A). We verified some of the new Y2H interactions for one disease-causing protein, ATXN1, to further support the finding that the genetic interaction is due to direct physical interactions. Two SCA1 genetic modifiers, Ataxin-2 (ATXN2) and RBPMS (ortholog of *Drosophila* “Couch potato”), showed a strong physical interaction with ATXN1 by co-AP assays (Figure 6B). We further determined the domains of ATXN1 required for the interaction with RBPMS using the Y2H system (Figure 6C) and co-AP assays (Figure 6D). Our results suggest that the N-terminal portion of ATXN1 is required for the interaction with RBPMS. The finding that a subset of the interacting proteins are established genetic modifiers of disease-related phenotypes suggests that the ataxia network will most likely reveal novel genetic modifiers for the ataxias.

Interestingly, we found that some of the ATXN1 genetic modifiers also physically interact with other ataxia-causing proteins. For example, RBPMS is one of the main hubs in the network interacting with many other proteins, including two other cerebellar ataxia-causing proteins, Atrophin-1 (ATN1) and Quaking (QKI) (Figure 6A). This suggests that some genetic modifiers for a specific disease could also modify other diseases of the same class and might play a key role in a common molecular pathway. Taken together, data from the protein interaction network can be combined with data from the genetic modifier network to

study biological functions and pathogenic mechanisms of disease-causing proteins.

DISCUSSION

Several human inherited cerebellar ataxias and many ataxic mouse mutants share clinical and pathological features such as loss of balance and coordination due to Purkinje cell degeneration. To determine whether shared molecular pathways can explain the phenotypic similarities, we developed an interaction network for proteins involved in inherited cerebellar ataxias. The ataxia network was built first using Y2H screens and then expanded by incorporating literature-curated and evolutionarily conserved interactions. This network revealed that several ataxia proteins interact and that indeed there are common processes and pathways involved in this class of neurodegenerative diseases. Furthermore, the network should provide insight into the normal functions of each individual ataxia protein and reveal novel candidate genes for diseases with overlapping phenotypes.

Generation and Characterization of the Ataxia Protein Network

Y2H screens provide an efficient way to develop protein–protein interaction networks. The limitations of Y2H-based assays, however, include high false-positive and

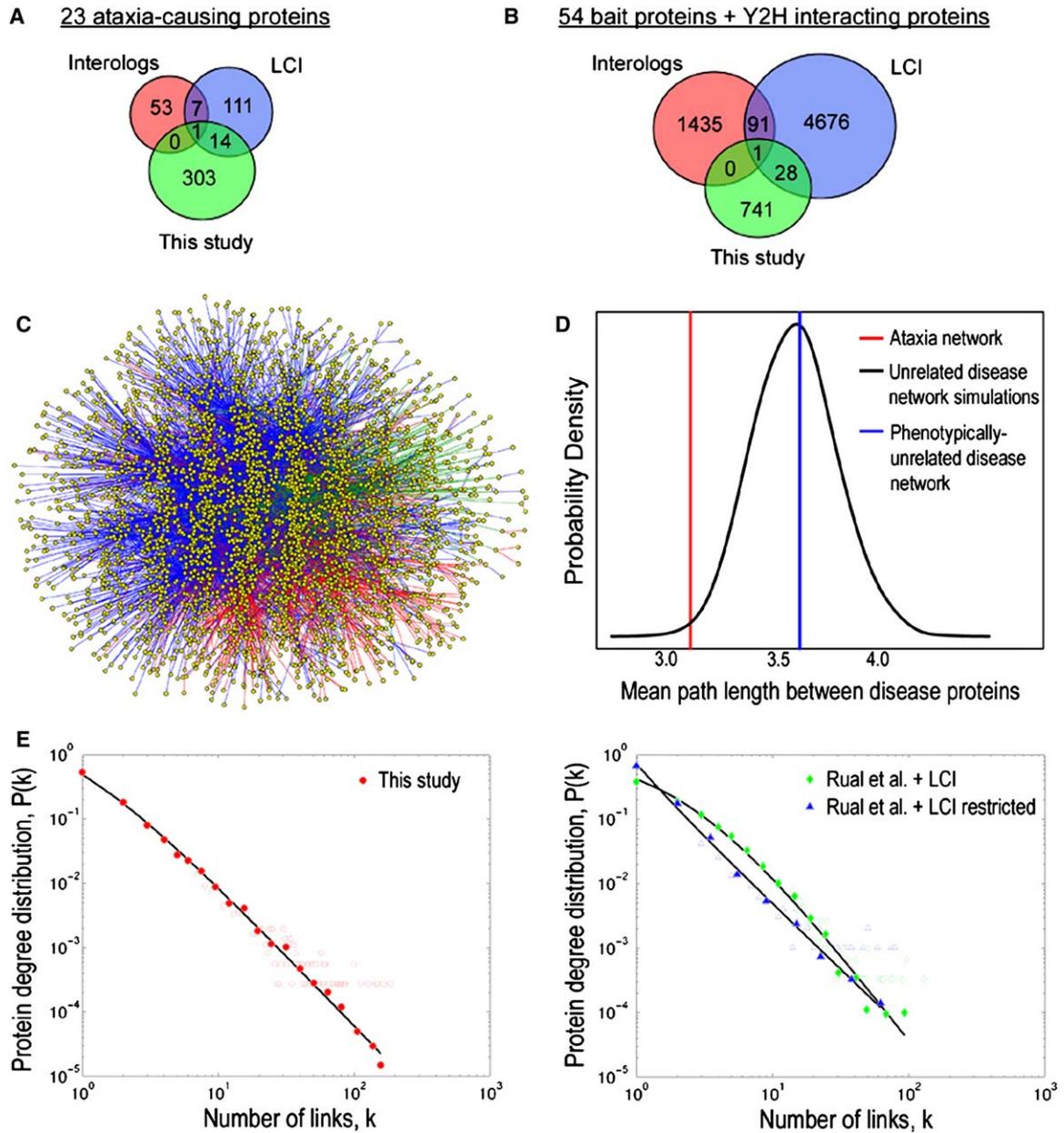


Figure 4. Generation and Analysis of the Ataxia Network

(A and B) Literature-curated interactions (LCIs) and Interologs are searched (A) for only 23 ataxia-causing proteins or (B) for all 54 ataxia-associated proteins and 561 Y2H-interacting proteins.

(C) Expanded ataxia network graph. Proteins are shown as yellow nodes. Interactions are shown as three different edge colors: green (Y2H interactions in this study), blue (LCI), and red (Interolog). All edges are first-, second-, or third-order interactions with ataxia-causing proteins.

(D) Mean path length. Any two ataxia-causing proteins in our network are connected via 3.11 links on average (red line), which is significantly ($p < .0004$) shorter than a mean path length between two proteins in a network based on phenotypically *unrelated* diseases (3.67, blue line) or in a simulated network of *unrelated* random diseases (3.65, black line).

(E) The degree distribution (normalized histogram) of the ataxia network in the left panel. Unfilled circles denote the probability that a protein will have exactly k links, and filled symbols correspond to the logarithmically binned version of the degree distribution, providing similar statistical weights in each bin. Both representations follow power laws $P(k) \sim (k + k_0)^{-\gamma}$, where the exponent is $\gamma = 2.2 \pm 0.1$ (the black line indicates the fit). The right panel shows with diamonds the degree distribution of the human interactome network expanded with literature-curated data, and the fit to the logarithmically binned points is $P(k) \sim (k + 2.3)^{-2.7 \pm 0.2}$. In contrast, for the subnetwork generated by taking interactions of at most third order with the ataxia-causing proteins (triangles), the best fit to the degree distribution changes to $P(k) \sim k^{-2.0 \pm 0.1}$.

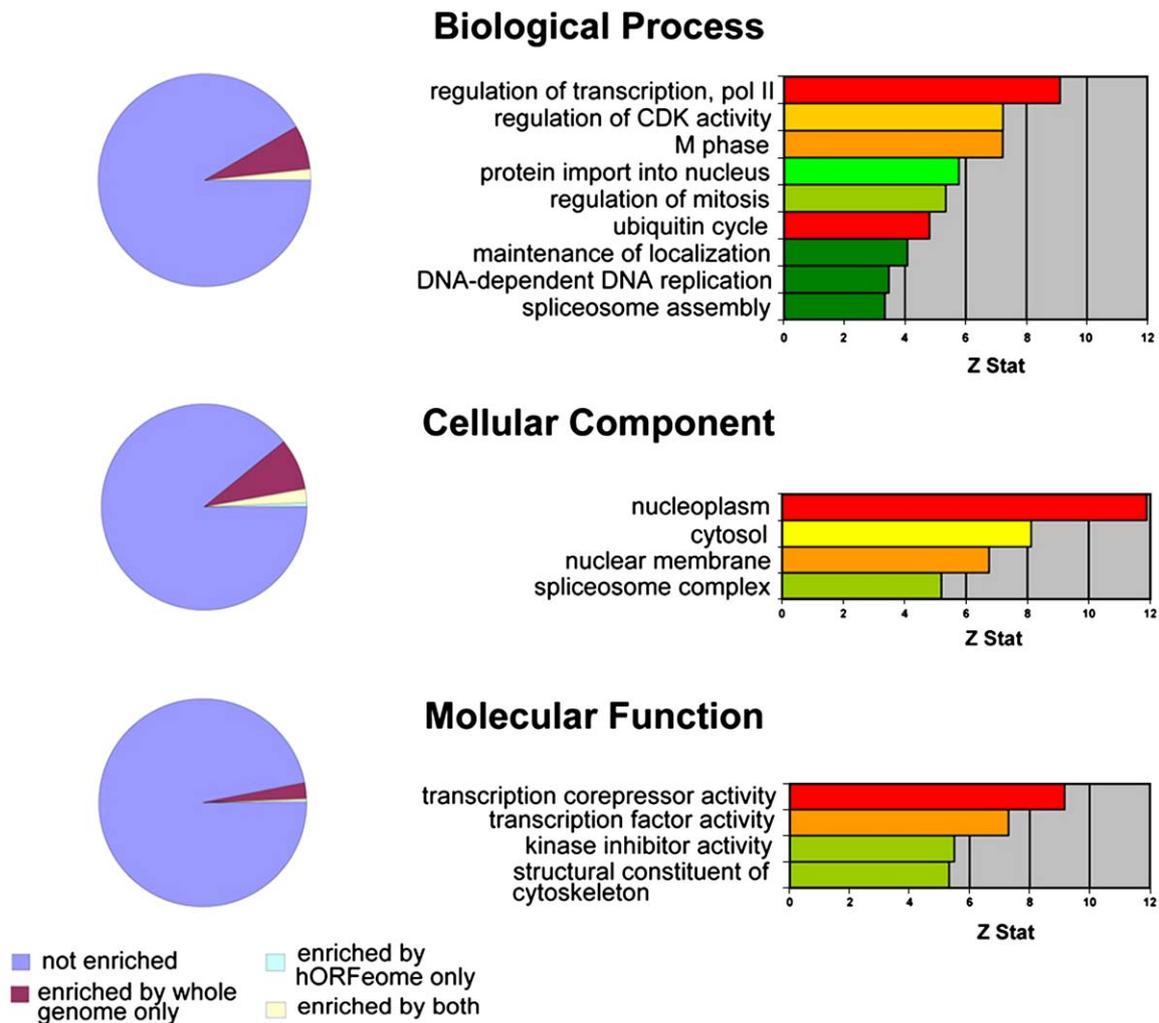


Figure 5. The Enriched GO Categories of the Ataxia Network

The pie charts depict the relative number of GO terms enriched in each of the three branches of the GO structure when using the whole genome and/or the hORFeome as reference distributions. Enriched categories are identified as those significant ($p < 0.05$) after adjusting for multiple testing. The bar charts depict the enriched GO terms from the common set (whole genome and hORFeome). The number of bars shown is less than the total number of common enriched categories because bars represent only the terminal GO terms. The length of the bar is the standardized enrichment score.

false-negative discovery rates. Therefore, when we set out to develop the ataxia-based network, we implemented several strategies to reduce these limitations. We used a stringent version of the Y2H system designed to express fusion proteins at low levels, permitting the identification of protein-protein interactions at physiological concentrations and minimizing the likelihood of false-positive interactions (Vidalain et al., 2004; Rual et al., 2005). The high validation rate of our Y2H interaction datasets by independent co-AP assays in mammalian cells (83%) as well as the biological information gathered from informatic-based analyses support the notion that our network has a low rate of false-positives and is of high quality.

To decrease the false-negative discovery rate, we screened hORFeome and adult human brain cDNA libraries. The hORFeome has the advantage of being nearly

90% normalized (i.e., >88% of the clones are equally represented in the library). In addition, the hORFeome allowed us to screen both AD- and DB-hORFeome reciprocally using both DB- and AD-ataxia fusion proteins, respectively. This approach also made it possible to screen some ataxia baits that were autoactivating at the DB position by placing them in the AD-position (Du et al., 1996). Using an adult brain cDNA library, we had the added advantage that multiple AD-cDNAs in that library encode fragments with different AD-fusion junctions representing each protein and its isoforms. Our results show that screening both libraries was indeed a worthwhile effort, given that only 29 out of 770 interaction pairs were identified as common.

There are several possible reasons that these two Y2H screens yielded largely different subsets of interacting proteins. The most prominent contributing factor is the

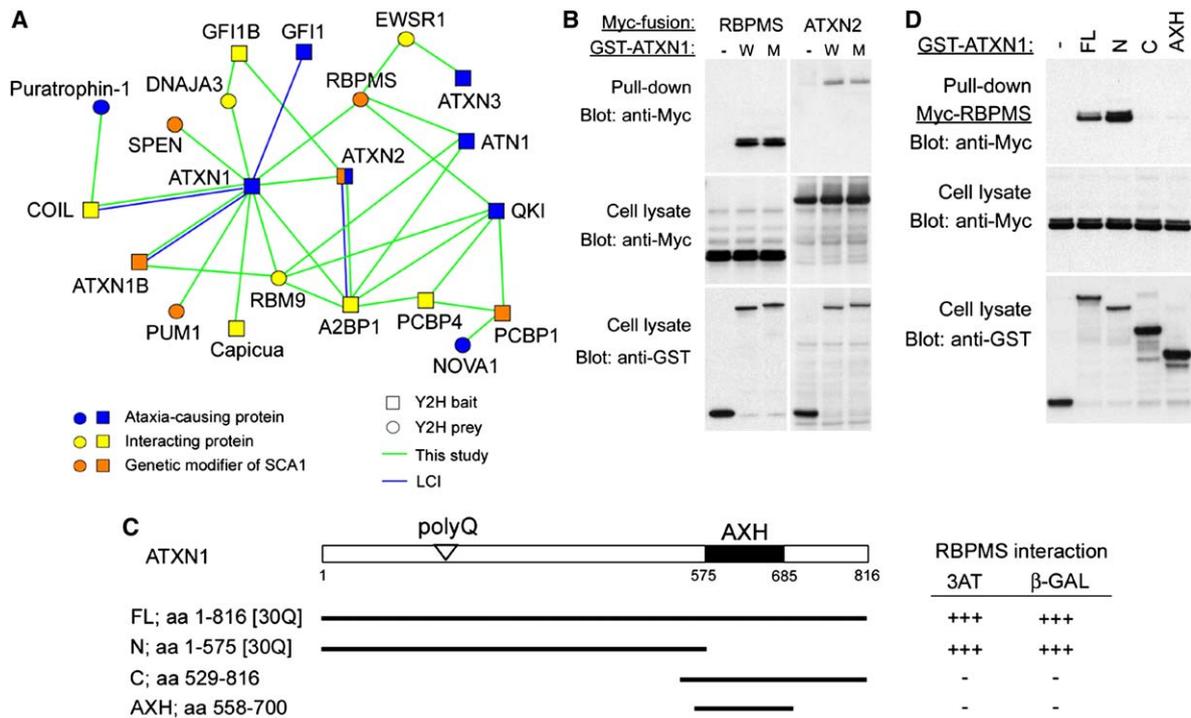


Figure 6. Genetic and Protein Interaction Data Reveal Functional Relevance of the Ataxia Network

(A) Graphic representation of an ATXN1 subnetwork; several interactors of ATXN1 proved to be genetic modifiers as well. Nodes depict interacting proteins in yellow, ataxia-causing proteins in blue, and genetic modifiers of SCA1 pathology in orange. Squares and circles indicate Y2H baits or preys, respectively. Edge colors define interactions: green (this study) and blue (LCI).

(B) Co-AP validation of interactions. GST-tagged proteins are GST alone (–), GST-ATXN1 [30Q] (W), and GST-ATXN1 [84Q] (M).

(C and D) Mapping of ATXN1's domain that binds RBPMS. Deletion constructs of ATXN1 were tested for binding to RBPMS by Y2H assays (C) and by co-AP assays in HEK293T cells (D).

difference between the libraries themselves: The hORFeome has about 8000 clones, whereas the brain cDNA library has at least two to three times that representation of clones. The representation of specific proteins in each library is also vastly different (full-length clone versus full-length and fragments for hORFeome and brain cDNA libraries, respectively). Previous work has shown that many biologically relevant protein–protein interactions cannot be detected in the Y2H assay when using full-length proteins (Flajolet et al., 2000; Legrain and Selig, 2000). In fact, our own data show that baits containing full-length as well as portions of the coding region of a specific gene are more successful at identifying protein partners than the full-length bait by itself. For example, the hORFeome screen yielded on average 8.34 interactions per gene for baits represented by both full-length and fragment clones in contrast with 1.19 interactions per gene represented only by full-length bait. In addition, full-length baits may fail to identify interacting proteins because they do not generate properly folded Y2H fusion proteins. The use of multiple overlapping protein fragments may increase the likelihood of exposing protein interfaces required for successful interactions in Y2H screening. Another factor contributing to discordances between the two Y2H screens is that each screen is done using totally

different methodology (mating versus cotransformation screen). Thus, it is not surprising that their sensitivity might differ. Lastly, it is possible that one method might yield a higher false-positive rate than the other. Our data, however, exclude this as a major factor, given the similar validation rate for hORFeome and cDNA library based interactions. In summary, the findings of this study suggest that screening both libraries increases the chances of identifying interacting proteins.

Combining a stringent Y2H screen with literature-curated and evolutionarily conserved interaction data, we generated a large protein–protein interaction network focused on approximately 20 proteins associated with inherited ataxias. After developing the ataxia network, we sought to determine whether the ataxia-causing proteins were more directly or indirectly connected to each other than would be expected in a similar network developed around proteins involved in diseases with *unrelated* phenotypes. The data clearly show that the ataxia network is significantly different from a disease network based on a similar number of proteins with nonoverlapping disease phenotypes. These differences were apparent in both the path length between disease-causing proteins (shorter in the ataxia network) and the number of triples connecting two disease proteins (higher in the ataxia network).

To determine whether the protein–protein interactions in our network are biologically relevant (Jansen et al., 2003; Jansen and Gerstein, 2004; Rual et al., 2005), we used informatic approaches. Gene Ontology (GO) analysis revealed that interacting proteins colocalize in subcellular compartments and are similarly annotated. Overall, the nucleus and nuclear subdomains are the dominant subcellular compartments implicated in the ataxia network. GO term analysis of the network revealed enrichment for specific biological functions.

Biological Implications of the Ataxia Network

A surprising and interesting outcome of this study is that the majority (18/23) of the ataxia-causing proteins interact either directly or indirectly. This was certainly not expected for such a large proportion of proteins; indeed, none of the 23 ataxia-causing proteins were known to interact with each other. Given the interconnectivity and short path length between the ataxia gene products, this network is likely to contain candidate genes for the 20 or more inherited ataxias whose causative genes have not yet been identified (Table S7). A case in point is the example of Puratrophin-1. When we started this study, we did not know that Puratrophin-1 is associated with an inherited ataxia, so it was not included among the baits. After completing the Y2H screen, Puratrophin-1 was identified as an interactor of COIL (an interactor of ATXN1). In the meantime, Ishikawa et al. (2005) reported that Puratrophin-1 is implicated in 16q22.1-linked cerebellar ataxia.

We demonstrated that many human orthologs of genetic modifiers identified in *Drosophila* and murine disease models interact directly with the ataxia-causing proteins. This is also quite interesting given that the majority of such modifiers were not known to be physical interactors and were not selected as baits in our study. Having shown that several genetic modifiers are physical interactors of ataxia-causing proteins, we would predict that some of the physical interactors are modifiers of ataxia phenotypes. Identifying such modifiers might expand the list of potential targets for therapeutic interventions. The finding that many protein interactions involve orthologous modifiers studied in model organisms supports the fact that many biologically relevant protein–protein interactions are conserved throughout evolution. Therefore, searching for evolutionarily conserved interactions in other organisms as we did in this study should provide insight into the functions of a protein and its role in pathogenesis. Protein–protein interaction networks generated in model organisms such as yeast (Uetz et al., 2000; Ito et al., 2001), worm (Walhout et al., 2000; Li et al., 2004), and fruit fly (Giot et al., 2003) will be a valuable resource for identifying functionally conserved interactions in the human and their relevance to different diseases or biological processes.

Human protein–protein interaction maps have been generated (Rual et al., 2005; Stelzl et al., 2005). Although such maps are still far from being complete, they will definitely provide a framework upon which to expand and deepen the network for many biological studies. In this

study we constructed an expanded network around the ataxias to deepen the network around specific disease processes and to provide a rich source of proteins that will permit in-depth functional studies of each specific ataxia protein.

Interestingly, some components in the network interact with only a single ataxia-causing protein, but others called “hubs” interconnect many (Barabasi and Albert, 1999). For example, three proteins (RBM9, A2BP1, and RBPMS) represent one of the main hubs in the ataxia network and interact with several different ataxia-causing proteins (Figure 6A). These proteins are involved in RNA binding or splicing (Jin et al., 2003; Nakahata and Kawamoto, 2005; Underwood et al., 2005) and genetically modify neurodegeneration in animal models (Fernandez-Funez et al., 2000). These data reveal one potential common mechanism of pathogenicity in this class of diseases and suggest that a subset of inherited ataxias might represent disorders of RNA splicing. This result is interesting in light of the discovery that alternative splicing defects are at the root of pathogenesis of mice lacking the protein NOVA, which is implicated in paraneoplastic opsoclonus myoclonus ataxia (POMA) syndrome (Ule et al., 2003; Ule et al., 2005). Identifying such hubs and modifiers might focus pathogenesis research efforts on key proteins and interactors relevant to the broader class of ataxias rather than a single disorder.

In summary, the data in this study show that a protein–protein interaction network for seemingly unrelated gene products involved in a group of different diseases that share clinical and pathological phenotypes (inherited cerebellar ataxias) has a high rate of verifiable interactions, is highly connected, and has the potential to reveal key mediators of pathogenesis. The finding that many genetic modifier proteins show direct physical interactions with the corresponding disease protein validates and underscores the power of both approaches. Lastly, we propose that phenotype-based protein–protein interaction studies can be applied to many human diseases, particularly common disorders that are sporadic in the majority of cases but do result from single gene defects in a small subset of patients. Included among these are Parkinson disease, diabetes, and hypertension. Such phenotype or disease-based interaction networks might reveal candidate disease genes, novel modifiers, and key pathogenic mechanisms that could be targeted therapeutically.

EXPERIMENTAL PROCEDURES

Yeast Two-Hybrid Screens

Starting from cDNAs of genes mutated in inherited ataxias, full-length ORFs or various fragments thereof were amplified by PCR and cloned into the pDONR223 “Donor” vector using the Gateway system as described (Rual et al., 2004). From the resulting “Entry” clones, ORFs or fragments were transferred individually into pDB-dest and pAD-dest-CYH Gateway-compatible “Destination” vectors to generate DB-ataxia and AD-ataxia fusion constructs, respectively.

To screen the hORFeome, we performed two independent mating Y2H screens in a 96-well format as described (Rual et al., 2005). The

DB-ataxia baits were used against the AD-hORFeome library in the first screen, whereas the AD-ataxia constructs were used in the second screen. DB-ataxia baits that did not show autoactivation (Table S2) were used to screen an adult human brain cDNA library (ProQuest, Invitrogen). See Supplemental Data for details about vectors and screens.

Coaffinity Purification Experiments

The co-AP experiments were performed in human HEK293T cells as described (Rual et al., 2005). See Supplemental Data for details.

Bioinformatics Analyses

Three types of analyses were performed on the network: validation of the protein interactions based on GO annotation data, analysis of the network topology, and GO content analysis. The calculations on the network were largely conducted in the R statistical programming environment (Gentleman and Ihaka, 1996) using the RBGL and graph package extensions. See Supplemental Data for more details.

Supplemental Data

Supplemental Data include one figure, seven tables, Experimental Procedures, and References and can be found with this article online at <http://www.cell.com/cgi/content/full/125/4/801/DC1/>.

ACKNOWLEDGMENTS

We thank Gabriela David, Candace Mainor, Matija Dreze, and Christophe Simon for technical assistance; Drs. Juan Botas, Hugo Bellen, Hamed Jafar-Nejad, and members of the Zoghbi and Vidal labs for helpful discussions and comments on the manuscript; and our colleagues who provided clones: Dr. Stefan-M. Pulst (Ataxin-2), Dr. Bert W. O'Malley (TBP), Dr. Roger D. Everett (USP), Dr. Harry T. Orr (Ataxin-7, RORA, and A1Up), Dr. Hitoshi Okazawa (PQBP-1), and Drs. Stephen J. Elledge and Yossi Shiloh (ATM). This research was supported by NIH grants NS27699 and HD24064 to H.Y.Z., an AMA seed grant to A.J.P., and an interactome mapping grant from NHGRI and NIGMS to M.V. H.Y.Z. is an investigator with the Howard Hughes Medical Institute.

Received: November 9, 2005

Revised: February 8, 2006

Accepted: March 13, 2006

Published: May 18, 2006

REFERENCES

- Albert, R. (2005). Scale-free networks in cell biology. *J. Cell Sci.* *118*, 4947–4957.
- Bader, G.D., Betel, D., and Hogue, C.W. (2003). BIND: the Biomolecular Interaction Network Database. *Nucleic Acids Res.* *31*, 248–250.
- Barabasi, A.L., and Albert, R. (1999). Emergence of scaling in random networks. *Science* *286*, 509–512.
- Barabasi, A.L., and Oltvai, Z.N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* *5*, 101–113.
- Chen, H.K., Fernandez-Funez, P., Acevedo, S.F., Lam, Y.C., Kaytor, M.D., Fernandez, M.H., Aitken, A., Skoulakis, E.M., Orr, H.T., Botas, J., and Zoghbi, H.Y. (2003). Interaction of Akt-phosphorylated ataxin-1 with 14-3-3 mediates neurodegeneration in spinocerebellar ataxia type 1. *Cell* *113*, 457–468.
- Cummings, C.J., Sun, Y., Opal, P., Antalfy, B., Mestrlil, R., Orr, H.T., Dillmann, W.H., and Zoghbi, H.Y. (2001). Over-expression of inducible HSP70 chaperone suppresses neuropathology and improves motor function in SCA1 mice. *Hum. Mol. Genet.* *10*, 1511–1518.
- Du, W., Vidal, M., Xie, J.E., and Dyson, N. (1996). RBF, a novel RB-related gene that regulates E2F activity and interacts with cyclin E in *Drosophila*. *Genes Dev.* *10*, 1206–1218.
- Engert, J.C., Berube, P., Mercier, J., Dore, C., Lepage, P., Ge, B., Bouchard, J.P., Mathieu, J., Melancon, S.B., Schalling, M., et al. (2000). ARSACS, a spastic ataxia common in northeastern Quebec, is caused by mutations in a new gene encoding an 11.5-kb ORF. *Nat. Genet.* *24*, 120–125.
- Fernandez-Funez, P., Nino-Rosales, M.L., de Gouyon, B., She, W.C., Luchak, J.M., Martinez, P., Turiegano, E., Benito, J., Capovilla, M., Skinner, P.J., et al. (2000). Identification of genes that modify ataxin-1-induced neurodegeneration. *Nature* *408*, 101–106.
- Fernandez-Gonzalez, A., La Spada, A.R., Treadaway, J., Higdon, J.C., Harris, B.S., Sidman, R.L., Morgan, J.I., and Zuo, J. (2002). Purkinje cell degeneration (pcd) phenotypes caused by mutations in the axotomy-induced gene, *Nna1*. *Science* *295*, 1904–1906.
- Flajolet, M., Rotondo, G., Daviet, L., Bergametti, F., Inchauspe, G., Tiollais, P., Transy, C., and Legrain, P. (2000). A genomic approach of the hepatitis C virus generates a protein interaction map. *Gene* *242*, 369–379.
- Gentleman, R., and Ihaka, R. (1996). R: A language for data analysis and graphics. *J. Comp. and Stat. Graph.* *5*, 299–314.
- Ghosh, S., and Feany, M.B. (2004). Comparison of pathways controlling toxicity in the eye and brain in *Drosophila* models of human neurodegenerative diseases. *Hum. Mol. Genet.* *13*, 2011–2018.
- Giot, L., Bader, J.S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y.L., Ooi, C.E., Godwin, B., Vitols, E., et al. (2003). A protein interaction map of *Drosophila melanogaster*. *Science* *302*, 1727–1736.
- Goehler, H., Lalowski, M., Stelzl, U., Waelter, S., Stroedicke, M., Worm, U., Droege, A., Lindenberg, K.S., Knoblich, M., Haenig, C., et al. (2004). A protein interaction network links Git1, an enhancer of huntingtin aggregation, to Huntington's disease. *Mol. Cell* *15*, 853–865.
- Hamilton, B.A., Frankel, W.N., Kerrebrock, A.W., Hawkins, T.L., Fitz-Hugh, W., Kusumi, K., Russell, L.B., Mueller, K.L., van Berkel, V., Birren, B.W., et al. (1996). Disruption of the nuclear hormone receptor RORalpha in staggerer mice. *Nature* *379*, 736–739.
- Isaacs, A.M., Oliver, P.L., Jones, E.L., Jeans, A., Potter, A., Hovik, B.H., Nolan, P.M., Vizor, L., Glenister, P., Simon, P.M., et al. (2003). A mutation in *Af4* is predicted to cause cerebellar ataxia and cataracts in the robotic mouse. *J. Neurosci.* *23*, 1631–1637.
- Ishikawa, K., Toru, S., Tsunemi, T., Li, M., Kobayashi, K., Yokota, T., Amino, T., Owada, K., Fujigasaki, H., Sakamoto, M., et al. (2005). An autosomal dominant cerebellar ataxia linked to chromosome 16q22.1 is associated with a single-nucleotide substitution in the 5' untranslated region of the gene encoding a protein with spectrin repeat and Rho guanine-nucleotide exchange-factor domains. *Am. J. Hum. Genet.* *77*, 280–296.
- Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. USA* *98*, 4569–4574.
- Jansen, R., and Gerstein, M. (2004). Analyzing protein function on a genomic scale: the importance of gold-standard positives and negatives for network prediction. *Curr. Opin. Microbiol.* *7*, 535–545.
- Jansen, R., Yu, H., Greenbaum, D., Kluger, Y., Krogan, N.J., Chung, S., Emili, A., Snyder, M., Greenblatt, J.F., and Gerstein, M. (2003). A Bayesian networks approach for predicting protein-protein interactions from genomic data. *Science* *302*, 449–453.
- Jin, Y., Suzuki, H., Maegawa, S., Endo, H., Sugano, S., Hashimoto, K., Yasuda, K., and Inoue, K. (2003). A vertebrate RNA-binding protein Fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG. *EMBO J.* *22*, 905–912.

- Kaytor, M.D., Byam, C.E., Tousey, S.K., Stevens, S.D., Zoghbi, H.Y., and Orr, H.T. (2005). A cell-based screen for modulators of ataxin-1 phosphorylation. *Hum. Mol. Genet.* *14*, 1095–1105.
- Kazemi-Esfarjani, P., and Benzer, S. (2000). Genetic suppression of polyglutamine toxicity in *Drosophila*. *Science* *287*, 1837–1840.
- Klein, J.A., Longo-Guess, C.M., Rossmann, M.P., Seburn, K.L., Hurd, R.E., Frankel, W.N., Bronson, R.T., and Ackerman, S.L. (2002). The harlequin mouse mutation downregulates apoptosis-inducing factor. *Nature* *419*, 367–374.
- Legrain, P., and Selig, L. (2000). Genome-wide protein interaction maps using two-hybrid systems. *FEBS Lett.* *480*, 32–36.
- Li, S., Armstrong, C.M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P.O., Han, J.D., Chesneau, A., Hao, T., et al. (2004). A map of the interactome network of the metazoan *C. elegans*. *Science* *303*, 540–543.
- Maslov, S., and Sneppen, K. (2002). Specificity and stability in topology of protein networks. *Science* *296*, 910–913.
- Matsuda, J., Kido, M., Tadano-Aritomi, K., Ishizuka, I., Tominaga, K., Toida, K., Takeda, E., Suzuki, K., and Kuroda, Y. (2004). Mutation in saposin D domain of sphingolipid activator protein gene causes urinary system defects and cerebellar Purkinje cell degeneration with accumulation of hydroxy fatty acid-containing ceramide in mouse. *Hum. Mol. Genet.* *13*, 2709–2723.
- Moore, R.C., Mastrangelo, P., Bouzamondo, E., Heinrich, C., Legname, G., Prusiner, S.B., Hood, L., Westaway, D., DeArmond, S.J., and Tremblay, P. (2001). Doppel-induced cerebellar degeneration in transgenic mice. *Proc. Natl. Acad. Sci. USA* *98*, 15288–15293.
- Nakahata, S., and Kawamoto, S. (2005). Tissue-dependent isoforms of mammalian Fox-1 homologs are associated with tissue-specific splicing activities. *Nucleic Acids Res.* *33*, 2078–2089.
- Noveroske, J.K., Hardy, R., Dapper, J.D., Vogel, H., and Justice, M.J. (2005). A new ENU-induced allele of mouse quaking causes severe CNS dysmyelination. *Mamm. Genome* *16*, 672–682.
- Pagel, P., Kovac, S., Oesterheld, M., Brauner, B., Dunger-Kaltenbach, I., Frishman, G., Montrone, C., Mark, P., Stumpflen, V., Mewes, H.W., et al. (2005). The MIPS mammalian protein-protein interaction database. *Bioinformatics* *21*, 832–834.
- Peri, S., Navarro, J.D., Amanchy, R., Kristiansen, T.Z., Jonnalagadda, C.K., Surendranath, V., Niranjani, V., Muthusamy, B., Gandhi, T.K., Gronborg, M., et al. (2003). Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Res.* *13*, 2363–2371.
- Ravikumar, B., Vacher, C., Berger, Z., Davies, J.E., Luo, S., Oroz, L.G., Scaravilli, F., Easton, D.F., Duden, R., O’Kane, C.J., and Rubinsztein, D.C. (2004). Inhibition of mTOR induces autophagy and reduces toxicity of polyglutamine expansions in fly and mouse models of Huntington disease. *Nat. Genet.* *36*, 585–595.
- Rual, J.F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S., Dricot, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P.O., et al. (2004). Human ORFeome version 1.1: a platform for reverse proteomics. *Genome Res.* *14*, 2128–2135.
- Rual, J.F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G.F., Gibbons, F.D., Dreze, M., Ayivi-Guedehoussou, N., et al. (2005). Towards a proteome-scale map of the human protein-protein interaction network. *Nature* *437*, 1173–1178.
- Sakaguchi, S., Katamine, S., Nishida, N., Moriuchi, R., Shigematsu, K., Sugimoto, T., Nakatani, A., Kataoka, Y., Houtani, T., Shirabe, S., et al. (1996). Loss of cerebellar Purkinje cells in aged mice homozygous for a disrupted PrP gene. *Nature* *380*, 528–531.
- Steffan, J.S., Bodai, L., Pallos, J., Poelman, M., McCampbell, A., Apostol, B.L., Kazantsev, A., Schmidt, E., Zhu, Y.Z., Greenwald, M., et al. (2001). Histone deacetylase inhibitors arrest polyglutamine-dependent neurodegeneration in *Drosophila*. *Nature* *413*, 739–743.
- Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F.H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., et al. (2005). A human protein-protein interaction network: a resource for annotating the proteome. *Cell* *122*, 957–968.
- Sun, X.Z., Harada, Y.N., Takahashi, S., Shiomi, N., and Shiomi, T. (2001). Purkinje cell degeneration in mice lacking the xeroderma pigmentosum group G gene. *J. Neurosci. Res.* *64*, 348–354.
- Taroni, F., and DiDonato, S. (2004). Pathways to motor incoordination: the inherited ataxias. *Nat. Rev. Neurosci.* *5*, 641–655.
- Tsuda, H., Jafar-Nejad, H., Patel, A.J., Sun, Y., Chen, H.K., Rose, M.F., Venken, K.J., Botas, J., Orr, H.T., Bellen, H.J., and Zoghbi, H.Y. (2005). The AXH Domain of Ataxin-1 Mediates Neurodegeneration through Its Interaction with Gfi-1/Senseless Proteins. *Cell* *122*, 633–644.
- Uetz, P., Giot, L., Cagney, G., Mansfield, T.A., Judson, R.S., Knight, J.R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., et al. (2000). A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* *403*, 623–627.
- Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. (2003). CLIP identifies Nova-regulated RNA networks in the brain. *Science* *302*, 1212–1215.
- Ule, J., Ule, A., Spencer, J., Williams, A., Hu, J.S., Cline, M., Wang, H., Clark, T., Fraser, C., Ruggiu, M., et al. (2005). Nova regulates brain-specific splicing to shape the synapse. *Nat. Genet.* *37*, 844–852.
- Underwood, J.G., Boutz, P.L., Dougherty, J.D., Stoilov, P., and Black, D.L. (2005). Homologues of the *Caenorhabditis elegans* Fox-1 protein are neuronal splicing regulators in mammals. *Mol. Cell. Biol.* *25*, 10005–10016.
- Vidalain, P.O., Boxem, M., Ge, H., Li, S., and Vidal, M. (2004). Increasing specificity in high-throughput yeast two-hybrid experiments. *Methods* *32*, 363–370.
- Walhout, A.J., Sordella, R., Lu, X., Hartley, J.L., Temple, G.F., Brasch, M.A., Thierry-Mieg, N., and Vidal, M. (2000). Protein interaction mapping in *C. elegans* using proteins involved in vulval development. *Science* *287*, 116–122.
- Xenarios, I., Salwinski, L., Duan, X.J., Higney, P., Kim, S.M., and Eisenberg, D. (2002). DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions. *Nucleic Acids Res.* *30*, 303–305.
- Yook, S.H., Oltvai, Z.N., and Barabasi, A.L. (2004). Functional and topological characterization of protein interaction networks. *Proteomics* *4*, 928–942.
- Yoshida, H., Yoshizawa, T., Shibasaki, F., Shoji, S., and Kanazawa, I. (2002). Chemical chaperones reduce aggregate formation and cell death caused by the truncated Machado-Joseph disease gene product with an expanded polyglutamine stretch. *Neurobiol. Dis.* *10*, 88–99.
- Young, A., Whitehouse, N., Cho, J., and Shaw, C. (2005). Ontology-Traverser: an R package for GO analysis. *Bioinformatics* *21*, 275–276.
- Zanzoni, A., Montecchi-Palazzi, L., Quondam, M., Ausiello, G., Helmer-Citterich, M., and Cesareni, G. (2002). MINT: a Molecular Interaction database. *FEBS Lett.* *513*, 135–140.
- Zoghbi, H.Y., and Botas, J. (2002). Mouse and fly models of neurodegeneration. *Trends Genet.* *18*, 463–471.
- Zoghbi, H.Y., and Orr, H.T. (2000). Glutamine repeats and neurodegeneration. *Annu. Rev. Neurosci.* *23*, 217–247.