Protein homeostasis of a metastable subproteome associated with Alzheimer's disease

Rishika Kundra[a], Prajwal Ciryam[a,b], Richard I. Morimoto[c], Christopher M. Dobson[a], and Michele Vendruscolo[a,1]

[a]Centre for Misfolding Diseases, Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom; [b]Department of Medicine, Columbia University College of Physicans & Surgeons, New York, NY 10032-3784; and [c]Department of Molecular Biosciences, Rice Institute for Biomedical Research, Northwestern University, Evanston, IL 60208

Edited by Ken A. Dill, Stony Brook University, Stony Brook, NY, and approved May 25, 2017 (received for review November 6, 2016)

Alzheimer's disease is the most common cause of dementia. A hallmark of this disease is the presence of aberrant deposits containing by the Aβ peptide (amyloid plaques) and the tau protein (neurofibrillary tangles) in the brains of affected individuals. Increasing evidence suggests that the formation of these deposits is closely associated with the age-related dysregulation of a large set of highly expressed and aggregation-prone proteins, which make up a metastable subproteome. To understand in more detail the origins of such dysregulation, we identify specific components of the protein homeostasis system associated with these metastable proteins by using a gene coexpression analysis. Our results reveal the particular importance of the protein trafficking and clearance mechanisms, including specific branches of the endosomal–lysosomal and ubiquitin–proteasome systems, in maintaining the homeostasis of the metastable subproteome associated with Alzheimer's disease.

Alzheimer's disease | protein aggregation | protein homeostasis | supersaturation

Neurodegenerative diseases are highly complex disorders characterized by extensive neuronal dysfunction, which is typically associated with protein misfolding and aggregation (1–9). A feature common to essentially all of these conditions is the presence of abnormal protein deposits, including amyloid plaques and neurofibrillary tangles in Alzheimer's disease (AD) and Lewy bodies in Parkinson's disease (PD) (1–9). It is increasingly recognized that the formation of such deposits, rather than being an unusual process involving only a small number of proteins, may represent a widespread phenomenon (1), with hundreds of different proteins found to aggregate under stress conditions, in aging, or in disease (5, 10–15).

To rationalize these observations, it has been recently shown that a large number of proteins are inherently supersaturated in the cellular environment (16, 17), as they are expressed at concentrations higher than their solubilities (18, 19) and therefore constitute a metastable subproteome potentially susceptible to aggregation (14–17). It has also been observed that proteins that have been reported to coaggregate with plaques, tangles, and Lewy bodies tend to be supersaturated (16, 17). Therefore, despite their heterogeneous and multifactorial nature, neurodegenerative conditions, including AD, PD, Huntington's disease (HD), and amyotrophic lateral sclerosis (ALS), share the important common attribute of protein supersaturation (16, 17, 20).

Given the intrinsic propensity of proteins to aggregate, it is not surprising that we are endowed with a powerful array of defense mechanisms whose role is to preserve protein homeostasis by helping to maintain proteins in their soluble states and to promote the degradation of those that misfold and aggregate (3–6, 21–25). This protein homeostasis system is comprised of a variety of components, including molecular chaperones, the proteolytic, ubiquitin–proteasome and autophagic degradation pathways, cellular trafficking, and other elements of the cellular stress response (3–6, 21–26). The progressive decline of the efficacy of these regulatory processes upon aging is likely to contribute to the increased susceptibility of the elderly population to age-associated neurodegenerative disorders (3–6, 21–28).

As the proteins within the metastable subproteome that are also transcriptionally down-regulated in AD may be particularly significant for the pathology of this disorder (29), it is important to determine the detailed mechanisms of their regulation by the protein homeostasis system. Our goal here, therefore, is to identify the specific components of this system that control a recently identified metastable subproteome associated with AD (29). To achieve this goal, we adopted the strategy of determining the association between groups of genes by probing their genetic interactions, an approach that is based on the observation that many functionally related genes are coexpressed (30, 31). For example, genes encoding for the various different components of protein complexes tend to have similar expression patterns (30, 32), and if groups of genes are regulated by common mechanisms, then they may be expected to be coexpressed (32).

We have therefore constructed a weighted gene correlation network (33, 34) of this metastable subproteome and of the overall protein homeostasis system (23) to gain a systems-level understanding of the transcriptional relationship between these two sets of proteins. By following this approach, we have identified the protein homeostasis components corresponding to the metastable subproteome specifically associated with AD (29). Our results show that the genes corresponding to this metastable subproteome are tightly coexpressed with specific components of the ubiquitin–proteasome and the endosomal–lysosomal pathways, thereby suggesting that metastable proteins with a high risk of aggregation tend to be closely regulated by the trafficking and degradation machineries.

Results

Protein Homeostasis of a Metastable Subproteome Associated with AD. AD is associated with widespread transcriptional changes (27, 28, 35–38), which can be rationalized in part by the presence

Significance

Alzheimer's disease is a neurodegenerative disorder whose molecular origins have been associated with the dysregulation of a set of metastable proteins prone to aggregation. Under conditions of cellular and organismal health, the protein homeostasis system prevents effectively the misfolding and aggregation of these metastable proteins. Although it is well established that such regulatory mechanisms become progressively impaired with aging, resulting in an accumulation of protein deposits, the specific nature of such impairment has remained incompletely characterized. Through a gene coexpression analysis, here we identify the endosomal–lysosomal and ubiquitin–proteasome systems, and more generally the protein trafficking and clearance mechanisms, as key components of the protein homeostasis system that maintains the metastable proteins in their functional states.

Author contributions: R.K., P.C., R.I.M., C.M.D., and M.V. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. Email: mv245@cam.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1618417114/-/DCSupplemental.

of a set of aggregation-prone proteins in the proteome (29). Many of the proteins involved in this metastable subproteome are components of the mitochondrial respiratory chain, an observation consistent with the well-characterized mitochondrial disruption associated with neurodegenerative disorders and specifically with AD (39). We refer to the proteins expressed by this subset of genes as the "AD metastable subproteome." The primary aim of the present study is to understand the different ways in which these metastable proteins are controlled, as illustrated schematically in Fig. 1. Our goal is thus to identify the specific components of the protein homeostasis system that are most closely involved in the regulation of the AD metastable subproteome.

Coexpression Analysis of the AD Metastable Subproteome and Its Associated Protein Homeostasis Components. Because the proteins in the AD metastable subproteome are intrinsically aggregation-prone, we searched for the specific protein homeostasis components that maintain the solubility and folding of these proteins. We therefore set out to identify an "AD metastable network" as a network of genes that encode for the AD metastable subproteome and its associated protein homeostasis components and that are correlated with the disease status (*Materials and Methods*).

To find the AD metastable network, we first carried out a weighted gene correlation network analysis (WGCNA) (33, 34) of the set of metastable proteins that we previously identified (29) and of the known components of the overall protein homeostasis system (23) (*Materials and Methods*). WGCNA is a robust method of performing gene coexpression analysis that has been shown to be particularly effective when large transcriptional datasets are available (33). As our aim was to study how metastable proteins are regulated across health and disease, we pooled together extensive microarray data obtained from postmortem brain tissues of patients diagnosed with late-onset AD (LOAD) and of matched controls (40) (Dataset S1, Table S1 and *Materials and Methods*). WGCNA uses the Pearson's coefficient of correlation between each pair of genes and their "topological overlap," which is a measure of their connectivity based on their shared neighbors, to identify biologically meaningful groups of coexpressed genes; these groups are called "modules" and labeled by different colors (34) (Dataset S1, Tables S2 and S3 and *Materials and Methods*).

As WGCNA captures the underlying network structure in large-scale gene expression studies, it has been used to study the global changes associated with a range of disease states, with the preservation of groups of coexpressed genes across species, and with the identification of hub genes associated with particular traits (28, 38, 41–44). We observed that the genes encoding for metastable proteins and for certain components of the protein homeostasis network are organized into well-defined modules (Fig. 2*A*, with the genes encoding for proteins in the metastable subproteome shown in pink in each module), where each module consists of tightly coexpressed genes. We found that the majority of metastable proteins belong to four specific modules, and we refer to them as "modules enriched in metastable proteins" (MEMPs)—MEMP-1 (blue), MEMP-2 (turquoise), MEMP-3 (green–yellow), and MEMP-4 (black) (Fig. 2*A*)—which consist of 659, 688, 35, and 74 genes, with 220, 91, 10, and 10 of these genes corresponding to metastable proteins, respectively (Dataset S1, Table S3).

Identification of an AD Metastable Network. As the next step to identify an AD metastable network, we performed an analysis to identify a module eigengene (ME) for each module, which is the first principal component (PC) of the expression values across genes in each module (*Materials and Methods*). The ME therefore provides a representative value for the expression of a group of genes in a particular module (34). This approach offers a significant advantage in correlating pairs of modules, as it eliminates the problem of multiple testing and noise by reducing the number of comparisons to just one instead of several hundreds. The higher the value of the Pearson's coefficient of correlation between two MEs, the more closely the two modules are related (*Materials and Methods*).

We therefore identified which of these modules showed the most significant relationship to the disease status by looking at the correlation between the MEs and disease status (*Materials and Methods*). We found six modules (Fig. 2*B*) to be significantly correlated with disease status, of which three modules (magenta, pink, and red) had very few metastable genes (4, 1, and 1, respectively). We excluded these modules from further analysis, as they mostly contained components of the protein homeostasis machinery whose expression levels do not correlate well with the metastable genes. The other three modules (MEMP-1, MEMP-2, and MEMP-3) were those most enriched for metastable genes that we described in the previous section (Fig. 2*B*).

As these three MEMPs were the only modules both significantly correlated with disease status and significantly enriched in the genes encoding for metastable proteins (Fig. 2*B*), we chose them for further analysis. They are also, in fact, closely related to each other based on the correlation of their MEs with a Pearson's correlation coefficient of 0.78 between MEMP-1 and MEMP-2 and 0.68 between MEMP-2 and MEMP-3 (Dataset S1, Table S4). To control for possible biases of the modules because of the use of a particular dataset, we cross-validated the results of module detection with a hippocampal gene expression dataset, as the hippocampus is among the regions typically most affected in AD (35). We observed that the MEMPs, along with most of the other modules, were well preserved between the two datasets (Fig. S1).

We then performed a gene ontology enrichment analysis to characterize these modules, finding that protein ubiquitination was the most enriched GO term for MEMP-1, MEMP-2, and MEMP-3 (Fig. S2). We also asked if the genes contained in these

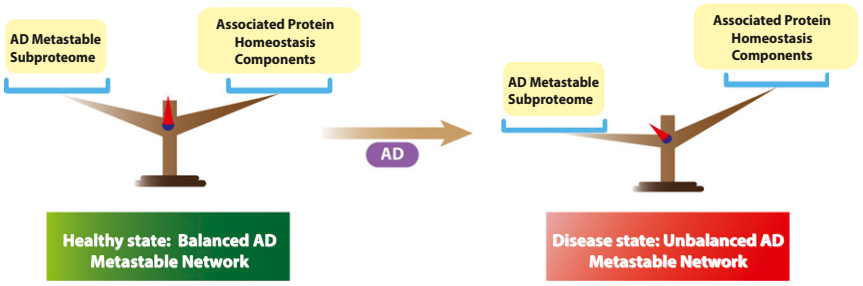

Fig. 1. Protein homeostasis of a metastable subproteome associated with AD. In the healthy state, this metastable subproteome (AD Metastable Subproteome) is effectively regulated by a series of protein homeostasis mechanisms (Associated Protein Homeostasis Components). In a disease state, this balance is compromised, and protein misfolding and aggregation results in the widespread formation of aberrant deposits.

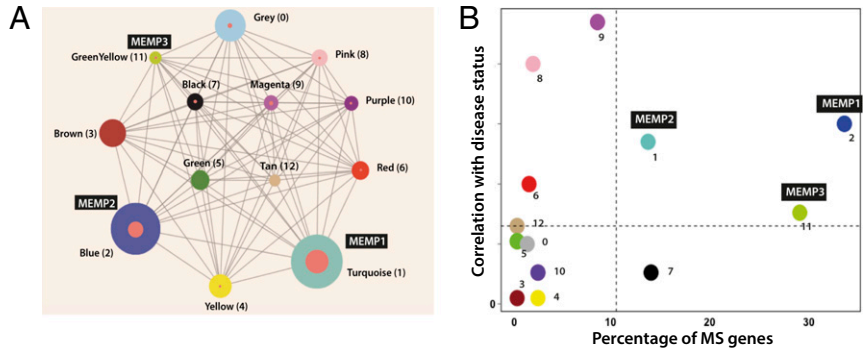

Fig. 2. Identification of the AD metastable network by a coexpression analysis of the AD metastable subproteome and its associated protein homeostasis components. (*A*) By using WGCNA, we carried out a hierarchical clustering of genes on the basis of their topological overlap. Modules of coexpressed genes are labeled by numbers and shown in different colors, with the size of the circles corresponding to the number of genes in the module. The number of genes encoding for the metastable subproteome is highlighted in pink within each module (Dataset S1, Table S3). The vast majority of metastable proteins are found in three specific MEMPs, which are referred to as MEMP-1 (blue, 220 metastable proteins), MEMP-2 (turquoise, 91 metastable proteins), MEMP-3 (green–yellow, 10 metastable proteins), and MEMP-4 (black, 10 metastable proteins); each of the other modules had between 0 and 4 metastable proteins. (*B*) Analysis of MEs. Each module is represented by a circle, which is labeled by number and shown in color, as in *A*. The *x* axis is the percentage of metastable (MS) genes in each module and the *y* axis the negative $\log_{10}$ of the *P* values for the correlation of each module with the disease status (*Materials and Methods*). The horizontal dashed line marks a *P* value of 0.05, and the vertical dashed line marks the 10% value. Hence, we identify the three modules, shown also in *A*—MEMP-1, MEMP-2, and MEMP-3—as the only modules that have a high percentage of metastable proteins and significant correlation to disease status.

modules are overrepresented in any biochemical pathway. To this end, by analyzing the KEGG biochemical pathways (45), we found that they are strongly overrepresented in the pathways associated with the AD metastable subproteome and the ubiquitin–proteasome and endosomal–lysosomal systems (Fig. S3).

Based on these results, we identified the AD metastable network as the set of genes in MEMP-1, MEMP-2, and MEMP-3 (Dataset S1, Table S2).

Identification of the Hub Genes and of Their Roles in the AD Metastable Network. Because any given module is comprised of a large number of genes, it is helpful to identify the most highly connected genes within a particular module, as these "hub" genes are likely to reveal the main processes carried out by the AD metastable network, even if individually they may not be crucial for the disease itself. To achieve this goal, we defined the "module membership" (MM) score by using the intramodular connectivity (kME; *Materials and Methods*), which is a measure of how strongly connected—that is, coexpressed—a given gene is to all of the other genes in a module (34). Hub genes were defined as those genes having an absolute kME value greater than 0.8.

The hub genes in the AD metastable network were found to be highly enriched in the KEGG biochemical pathways of cellular degradation (proteasome and ubiquitin-mediated proteolysis) and trafficking in addition to those previously associated with metastable proteins such as oxidative phosphorylation, AD, PD, and HD (Fig. 3*A*). These results are fully consistent with those reported above for the full list of genes in the AD metastable network (Fig. S3). In any given module, a high mean MM value indicates how tightly coexpressed the genes are within that module. We observed that the genes encoding for the AD metastable subproteome in the AD metastable network have a significantly high mean MM value that is significantly greater than that of other genes in that module (Fig. 3*B*). In addition, more than two thirds of the genes encoding for metastable proteins in these modules are hub genes, indicating their central importance in their respective modules.

Test of Module Generality Using a Consensus Network Analysis with a Visual Cortex Dataset. We next sought to determine whether the modules that we identified are general or instead specific to the dataset or brain region that we analyzed. To check the robustness of the modules identified in this study, we constructed a consensus network (*Materials and Methods*) using WGCNA on another dataset from the visual cortex (VC) of AD patients and healthy controls (40), along with the dataset for the dorsolateral prefrontal cortex (PFC) used previously (*Materials and Methods*), to examine whether or not our network is preserved. To assess the level of preservation, we used the "consensus network" construction, which identifies groups of genes that are tightly coexpressed across multiple studies (46). The consensus MEs (consMEs) represent modules in each of the two sets (46) (*Materials and Methods*). Each gene is assigned to a single consensus module, but there are two sets of consMEs for each module as a given module can have a different expression profile in the two datasets. We found that all of the modules identified in our study have a consensus counterpart in the VC dataset, indicating that the module structure in the two datasets is similar (Fig. S4).

We then constructed the two sets of eigengene dendrograms and eigengene heatmaps based on the consMEs (one for each study), and the results indicate that the overall modular structure in the two sets is quite similar. The preservation heatmap shows the preservation network, defined as one minus the absolute difference of the eigengene networks in the two datasets (Fig. S5*A*). The overall degree of preservation between the two networks is 0.87, and the mean preservation of relationships for each eigengene is consistently high for all of the modules except the "red" one, as shown by the preservation heatmap and bar plot (Fig. S5*B*), thus indicating that the modules identified in the analysis detailed in the study are highly robust.

These results suggest that the difference between a healthy state and a disease state does not involve a reorganization of the modules but rather a variation in the expression levels of specific genes within the modules. In the following, we therefore carried out further investigations to identify such genes.

Protein Homeostasis of the AD Metastable Subproteome. We next asked the central question of this work—How is the AD metastable subproteome regulated? To answer this question, we analyzed which components of the protein homeostasis system are coexpressed with the AD metastable subproteome in the AD metastable network, as we expect that the knowledge of such components could offer insight into the regulation of these metastable proteins (Fig. 4 and Fig. S6). To this end, we identified the most important hub genes, among those described above, by visualizing them within the AD metastable network; we used the Cytoscape software for

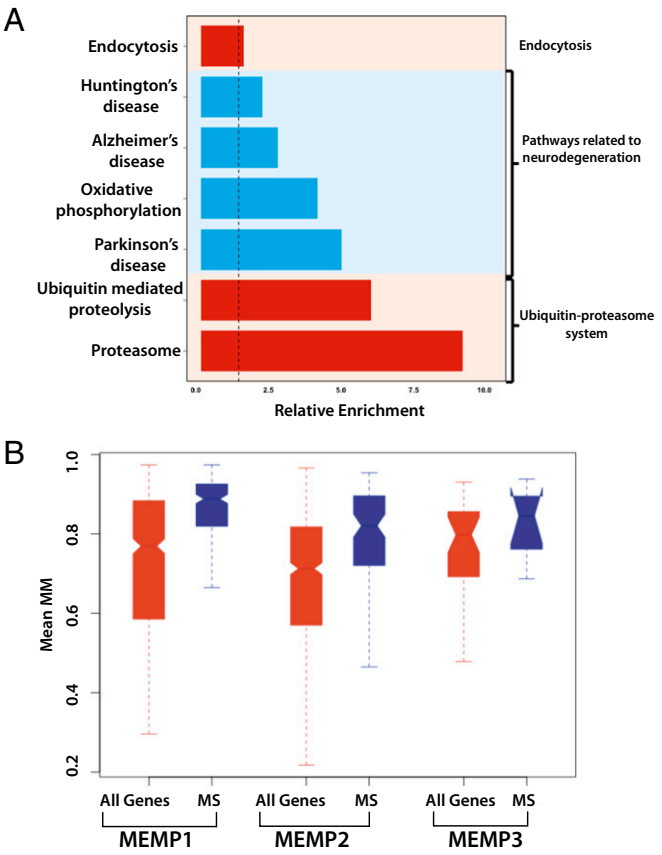

Fig. 3. Identification of KEGG biochemical pathways enriched in hub genes in the AD metastable network. (*A*) KEGG biochemical pathways (45) enriched in hub genes (|kMe| >0.8) in the AD metastable network; the dotted line indicates $P = 0.05$. (*B*) A comparison of the mean MM values of the metastable proteome (MS, blue) and of all genes (red) shows that genes corresponding to metastable proteins are highly coexpressed.

this purpose (47). The top 10% of all hub gene interactions based on their topological overlap were visualized, with those involved in at least 50 of these interactions shown in the center (Fig. 4*A*, inner and middle rings). We observed 10 hub genes related to trafficking (Fig. 4*A*, middle ring, green circles) and 5 hub genes related to the ubiquitin–proteasome pathway (Fig. 4*A*, middle ring, red circles) as the most connected hub genes (Table 1). These results are consistent with experimental evidence that protein trafficking and degradation are components of the protein homeostasis system closely linked with AD (48–54). Our analysis also identifies other components (autophagy, metabolism, signaling, and protein synthesis; Fig. 4*B*), although more extensive data will be needed to clarify their association with the metastable subproteome in greater detail.

Endosomal–Lysosomal System. Although the present analysis of the hub genes is aimed primarily at identifying the main processes within the AD metastable network, it could also be informative to consider the possible specific roles of these genes in AD. Among the hub genes associated with trafficking, we found RAB6A, a small GTPase that helps mediate retrograde transport from the Golgi apparatus to the endoplasmic reticulum (ER), which has an increased expression level in AD brains (55). To explain this finding, it has been suggested that this protein is involved in a regulatory mechanism that responds to increased protein accumulation (55). In addition, overexpression of RAB1, another small GTPase closely related to RAB6A, has been shown to alleviate ER stress in yeast models of PD (56). Hence, RAB6A, which is a central gene in the AD metastable network described in this work, could play an important role in the regulation of the metastable proteins by directing them toward the endosomal–lysosomal degradation machinery, thereby preventing their accumulation in the cytoplasm. Another two genes in the group that we found are *ATP6V1H*, which encodes a protein subunit of a vacuolar ATPase involved in clathrin-mediated endocytosis (57, 58) and whose role in regulating lysosomal pH has been recently been linked to neurodegeneration (59), and ATL1, which is involved in ER trafficking (60, 61). In fact, all 10 genes that we found to be related to trafficking are part of the endosomal–lysosomal system. Specifically, SH3GL2, SLC9A6, and CLTA are localized in the endocytic vesicle membrane (62, 63), and NSF is involved in vesicle-mediated transport and acts as a fusion protein through the SNARE proteins (64). Our results, therefore, indicate the importance of the endosomal–lysosomal system in controlling the metastable subproteome. These findings extend the well-known role of this system in the processing of Aβ (65) to the regulation of a broader range of aggregation-prone proteins.

Ubiquitin–Proteasome System. Among the genes associated with the ubiquitin–proteasome system, we found ENC1, which is an actin binding protein that has been reported to modulate the aggregation of mutant huntingtin under ER stress (66). MYCBP2, FBXL2, and RNF128 are E3 ubiquitin ligases and are essential components of the ubiquitin-dependent degradation of proteins (67–69). These results indicate that metastable proteins are likely to be regulated upstream of the proteasomes at the ubiquitin ligase stage.

Molecular Chaperones. We also found a number of components of molecular chaperone networks coexpressed with the AD metastable subproteome (Dataset S1, Table S5). Such components include co-Hsp70/Hsp90 species, which are known to assist the Hsp70/Hsp90 system to degrade protein aggregates (70, 71). Among such molecular chaperones, we found DNAJC6, a J-domain cochaperone with a role in HSC70-mediated uncoating of the clathrin-coated vesicles in neurons by recruiting HSC70. Also seen as hub genes were TOR1A, with chaperone activity and a member of the AAA family of ATPases, and ERLEC1, which has a role in ER quality control (72, 73).

Taken together, these results indicate that the components of the AD metastable subproteome, which consists of proteins inherently at risk of aggregation, tend to be highly coexpressed with multiple components of the protein homeostasis system. These findings illustrate how during the course of AD, when a dysregulation and collapse of these systems is increasingly likely to occur, these metastable proteins are likely to represent an enhanced risk due to the dysfunction of the regulatory mechanisms associated with their folding, transport, and degradation.

Relationship with Genome-Wide Association Studies (GWASs). To further assess the significance of our analysis, we compared our results with genetic loci identified by GWAS. These studies have reported that several genes associated with the trafficking and degradation systems are closely associated with AD (74, 75). In particular, seven GWAS genes (*PICALM*, *SORL1*, *CD33*, *BIN1*, *CD2AP*, *ABCA7*, and *RIN3*) are associated with the endosomal–lysosomal system, and two GWAS genes (*CLU* and *PTK2B*) are associated with the ubiquitin–proteasome pathway (74, 75). These results are highly consistent with the conclusions of the present study, as 17 GWAS genes (among the 28 that we considered) are present in the AD metastable network identified in this work (Fig. 5). This consistency is remarkable, as the GWAS strategy, where genes are typically associated with disease on the basis of single nucleotide polymorphism (SNP) statistics, is independent from the one that we have used here to associate genes with disease through the combination of their coexpression and the metastability to aggregation of their products. These two approaches

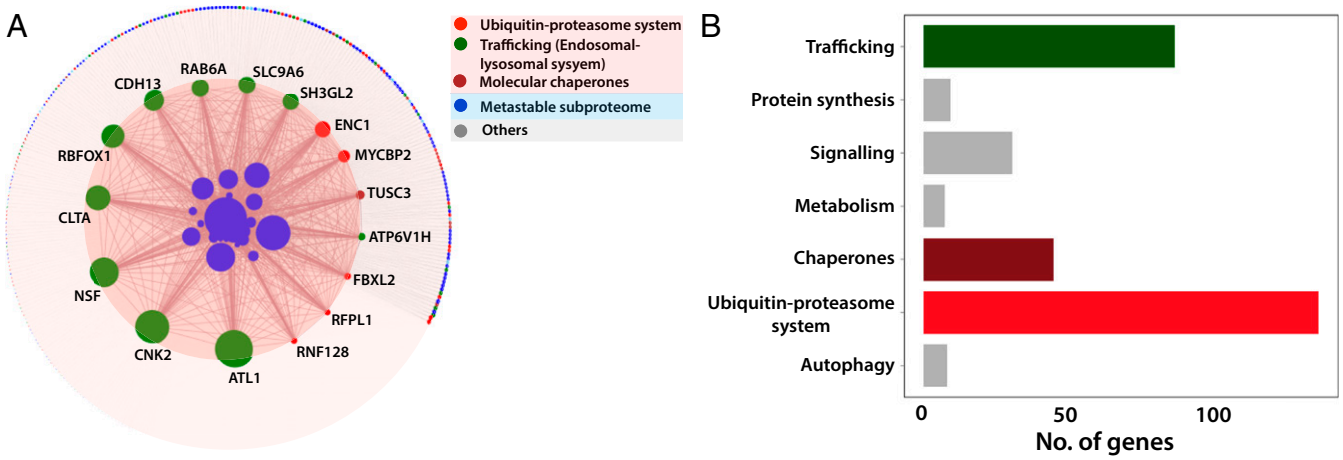

Fig. 4. Identification of the major components of the protein homeostasis system associated with the AD metastable subproteome. (*A*) Network representation of the AD metastable network showing the hub genes of the protein homeostasis system (green, red, and dark red circles, middle ring; Table 1) and the hub genes of the AD metastable subproteome (blue circles, inner ring). This analysis reveals in particular the importance of the ubiquitin–proteasome (red) and trafficking (green) systems in the regulation of aggregation-prone proteins in AD. We visualized the top 10% of the hub gene interactions, with those genes involved in at least 50 interactions shown in the inner and middle rings (see also Fig. S6). The sizes of the nodes correspond to their degrees of connectivity. (*B*) Protein homeostasis components within the hub genes of the AD metastable network. The major components in *A* are shown in the same color code; additional components (autophagy, metabolism, signaling, and protein synthesis) are also shown.

are therefore complementary, as a coexpression analysis can identify a large number of genes and therefore reveal the biochemical pathways involved in the disease and help rationalize the specific roles of the GWAS genes but may not capture important relationships, such as in the present case the role of ADAM10, PSEN1, and PSEN2 in the processing and regulation of APP (Fig. 5).

Consensus Network Analysis of AD, PD, and HD. As noted above, the phenomenon of protein misfolding and aggregation is a common feature of many neurodegenerative disorders, including AD, PD, HD, and ALS. Although these diseases are characterized by a variety of different clinical manifestations and features, there is increasing interest in understanding the extent to which they share common molecular origins (1–9). To address this question in the present context, we investigated whether or not the regulation of the metastable proteins, in terms of their coexpression with specific protein homeostasis components, is similar across AD, PD, and HD.

Because oxidative phosphorylation is the most significantly enriched pathway among the metastable genes, we analyzed the coexpression of genes involved in this specific pathway and in the protein homeostasis components. We built a consensus network for gene expression data (Dataset S1, Table S6) from hippocampal tissue (35), *substantia nigra* (76), and PFC (73), obtained postmortem from patients diagnosed with AD, PD, and HD, respectively, and from age-matched controls. The network heatmaps indicate the correlation of various eigengenes within the AD, PD, and HD networks (Fig. S7) and the preservation heatmaps (Fig. 6) reveal that the overall preservation of the three networks is highly significant (shown in red). The mean preservation of the three networks exceeds 0.7 in all three cases (Fig. S7), indicating that the global structures of the coexpression

Table 1. List of hub genes used to identify the components of the protein homeostasis system associated with the AD metastable subproteome

Hub genes	Proteins	Known functions
ATP6V1H	V-type proton ATPase subunit H	Clathrin coated endocytosis, formation of endosomes
SH3GL2	Endophilin-A1	Synaptic vesicle endocytosis
SLC9A6	Sodium/hydrogen exchanger 6	Exchange of protons across the membrane of early and recycling endosome
RAB6A	Ras-related protein Rab-6A	Retrograde transport from Golgi to ER, transport from endosome to plasma membrane
CDH13	Cadherin-13	Regulation of endocytosis
RBFOX1	RNA binding protein fox-1 homolog 1	RNA binding protein, regulation of alternative splicing events
CLTA	Clathrin light chain A	Major protein of the polyhedral coat of coated pits and vesicles
NSF	Vesicle-fusing ATPase	SNARE binding, regulation of exocytosis
CNK2	Connector enhancer of kinase suppressor of ras 2	Adaptor protein, regulation of signal transduction
ATL1	Atlastin-1	ER to Golgi vesicle transfer
ENC1	Ectoderm-neural cortex protein 1	Proteasomal ubiquitin-independent protein catabolic process
MYCBP2	E3 ubiquitin-protein ligase MYCBP2	Ubiquitin ligase, protein ubiquitination
FBXL2	F-box/LRR-repeat protein 2	Ubiquitin ligase, protein ubiquitination
RFPL1	Ret finger protein-like 1	Zinc ion binding
RNF128	E3 ubiquitin-protein ligase RNF128	Ubiquitin ligase, ubiquitin-dependent protein catabolic process
TUSC3	Tumor suppressor candidate 3	Magnesium transporter

These hub genes are shown in Fig. 4 and are reported here together with their corresponding proteins and their known functions. The list of hub genes corresponding to the AD metastable subproteome is reported in Dataset S1, Table S7.

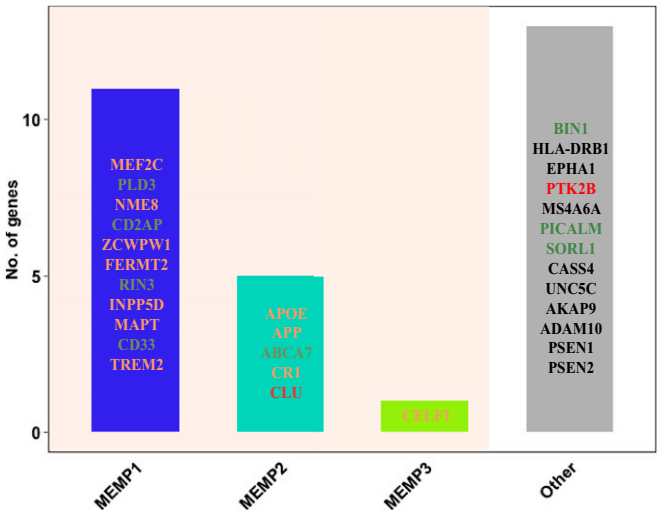

Fig. 5. The majority of GWAS genes are found in the AD metastable network. Shown are the number of genes identified by GWAS (74, 75) that are present in the AD metastable network or in the other modules described in this work. Seventeen out of 28 genes identified by GWAS are present in the AD metastable network. The names of the genes are shown in their respective modules. The genes shown in red belong to the ubiquitin–proteasome system, and those in green belong to the trafficking system.

networks are similar for the three diseases. These results thus suggest that the differences between these diseases may be found in the dysregulation of specific genes within the consensus network (Fig. 6 and Fig. S7).

Discussion

Specific Components of the Protein Homeostasis System That Regulate Protein Aggregation. In this work, we have taken the view that a major hallmark of aging and neurodegeneration is the progressive impairment of the balance between protein aggregation and its control by the protein homeostasis system, which leads to the characteristic accumulation of aberrant protein aggregates (1–17, 21–29) (Fig. 1). In this context, we have previously reported that large numbers of proteins are inherently metastable to aggregation because of their elevated expression levels relative to their solubilities (16, 17). We have also observed a specific transcriptional down-regulation of genes encoding these proteins in AD (29) as well as a tissue-specific vulnerability to AD caused by an imbalance between aggregation-prone proteins and their protein homeostasis regulators (22).

To identify the protein homeostasis mechanisms that control the metastable proteins associated with AD, in this study we have analyzed together a set of proteins inherently prone to aggregation (29) and a set of proteins that make up the overall protein homeostasis system (23). Our analysis started from a metastable subproteome corresponding to the overlap between genes encoding for proteins that are supersaturated and transcriptionally down-regulated in AD but not in aging (29). We then constructed an AD metastable network composed of genes encoding this set of metastable proteins together with the corresponding components of the protein homeostasis system. We have found that this specific AD metastable network consists of well-defined modules of coexpressed genes (Fig. 2), enabling us to identify key players of the ubiquitin–proteasome and endosomal–lysosomal systems, along with some specific molecular chaperones (Fig. 4).

The systems-level approach that we have adopted in this work provides an understanding of the regulation of the AD metastable subproteome as a whole, as opposed to the regulation of individual proteins by specific components of the protein homeostasis system. Our results show that, from a list of about 2,000 components of the protein homeostasis system (23), just a relatively small number of specific proteins in the degradation and trafficking machinery along with specific molecular chaperones are primarily responsible for handling the metastable proteins with a high propensity to misfold and aggregate (Fig. 7).

Endosomal–Lysosomal and Ubiquitin–Proteasome Regulation of the Proteins Involved in Oxidative Phosphorylation. The expression levels of most of the components of the protein homeostasis system identified in this study have been previously seen to decrease with aging (29). Hence, during aging and disease, with the suppression of the protein homeostasis system, these proteins could become particularly vulnerable to aggregation because of their inherent metastability. Because these proteins perform fundamental functions, including in particular energy metabolism through oxidative phosphorylation (29), their aggregation could result in triggering a cascade of events contributing to disease pathology and ultimately to neuronal death. In addition, such dysregulation poses a pronounced threat to neurons due to their postmitotic state and increased dependence on mitochondria for energy production. Indeed, there is a substantial overlap in the genes involved in the pathways associated with oxidative phosphorylation and AD, HD, and PD (Dataset S1, Table S6), indicating again that the proteins encoded by these genes are highly metastable and hence are significant in the context of disease pathology. Mitochondria play a central role in aging and in regulating cell death (39, 77) as well as in the overall maintenance of cellular health. Whether mitochondrial dysfunction is the cause or effect of the disease pathology is still, however, unclear. Mitochondria have been shown to interact with aggregation-prone proteins, including α-synuclein and Aβ. More specifically, Aβ has been shown to be localized on the mitochondrial membrane in a transmembrane arrested form, possibly disrupting protein import into the mitochondria (78, 79).

Our results also point to a possible dependence of the proteins in the respiratory chain complex on the endosomal–lysosomal system, which we identified using the hub genes *RAB6A*, *ATP6V1H*,

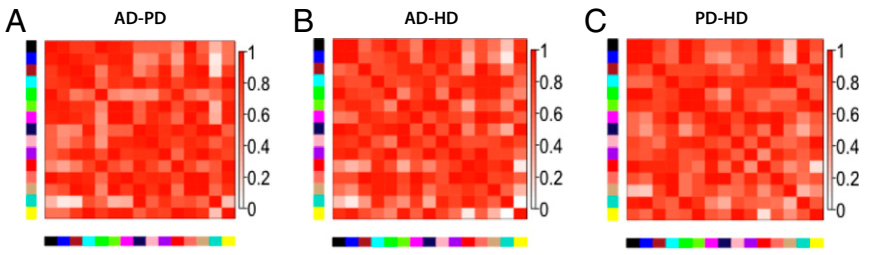

Fig. 6. Consensus network analysis of AD, PD, and HD. Network preservation heatmaps for AD–PD (*A*), AD–HD (*B*), and HD–PD (*C*). The rows and columns represent the ME for different modules in each network. Preservation is defined as one minus the absolute difference of the eigengene networks in the two datasets. Red denotes high preservation, and white denotes low preservation.

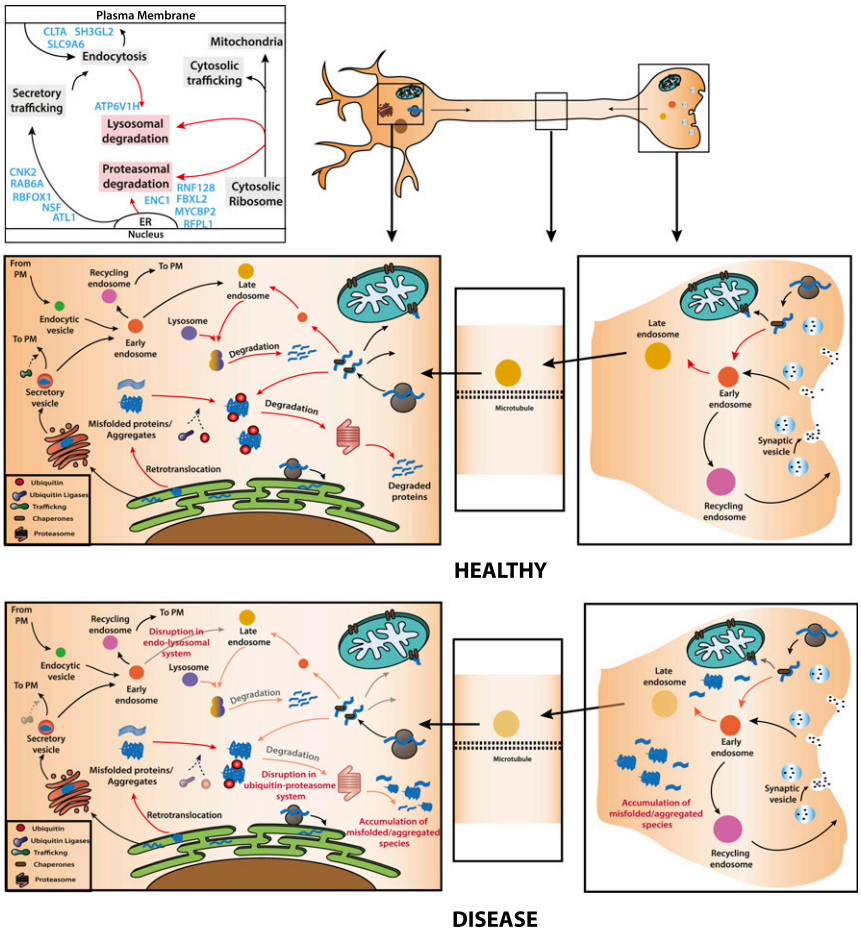

Fig. 7. Schematic representation of the main pathways involved in the homeostasis of the proteins metastable to aggregation in AD. Our results identify the key role of protein degradation, in particular the ubiquitin–proteasome system and the endosomal–lysosomal system, in the homeostasis of the metastable subproteome associated with AD.

ATL1, *SH3GL2*, *SLC9A6*, and *CLTA*, and on the ubiquitin–proteasome system, where the hub genes are the E3 ubiquitin ligases MYCBP2, FBXL2, and RNF128 (Fig. 4). These indications are consistent with the observation in yeast that accumulation of mitochondrial proteins in the cytoplasm leads to activation of the unfolded protein response (80). Furthermore, recent studies have reported the presence of polyubiquitinated mitochondrial proteins, suggesting that they are substrates of the ubiquitin proteasome system (81, 82) and that in yeast the expression of the proteasome is up-regulated upon cytoplasmic accumulation of mitochondrial proteins (80, 83). If mitochondrial import is disrupted and these metastable proteins therefore accumulate in the cytoplasm, the cell responds by clearing them through degradation. When, however, this disruption happens in an environment where protein homeostasis is compromised, these proteins would be particularly at danger of aggregation. We observed, in addition, a similar pattern of coexpression of genes encoding for mitochondrial membrane proteins across AD, PD, and HD, indicating that even though the initial cause of dysregulation might be different, these diseases are likely to share common molecular mechanisms at a later stage of progression, with regulation of mitochondrial membrane proteins playing an important role.

The finding that a metastable subproteome that is specifically associated with AD is primarily regulated by the protein trafficking and degradation systems provides important insights into the control of protein misfolding in this disease. These results suggest that in a setting of compromised protein folding, the maintenance of proteins in their soluble states may move away from regulating conformations and toward regulating concentrations.

Overall, the results that we have reported suggest an extension of the view that AD is associated with an age-related protein homeostasis failure that results in the aggregation of Aβ and tau—thus, neuronal dysfunction appears to be a consequence of a loss of ability of major branches of the protein homeostasis system to regulate a much wider group of aggregation-prone proteins making up a metastable subproteome.

Conclusions

We have described specific components of the protein homeostasis system that regulates a metastable subproteome associated with AD. This analysis has revealed the central roles of the ubiquitin–proteasome and endosomal–lysosomal degradation pathways, whose relevance to AD is well known (48–54), in the maintenance of a pool of metastable proteins prone to aggregation. By identifying a series of regulatory pathways associated with AD, these findings also help to rationalize the roles in the disease of the individual genes resulting from GWASs. We anticipate that an increasingly detailed understanding of the mechanisms of regulation of the metastable subproteome will contribute to the development of therapeutic strategies against neurodegenerative diseases aimed at promoting the maintenance of aggregation-prone proteins in their soluble states.

Materials and Methods

Dataset Acquisition. Microarray data for brain tissues from postmortem AD patients and healthy controls were downloaded from the Gene Expression

Omnibus (GEO) database (84). The following datasets were used for analysis (Dataset S1, Table S1): GSE44770, containing tissues derived through autopsy from the dorsolateral PFC region obtained from LOAD patients and from healthy controls; GSE44771, containing tissues derived through autopsy from the VC region obtained from LOAD patients and from healthy controls; GSE1297, containing hippocampal gene expression data from LOAD patients and from healthy controls; GSE33000, containing dorsolateral PFC tissue from HD patients and from healthy controls obtained from the Harvard Brain Tissue Resource Center (HBTRC); and GSE20292, containing postmortem brain tissue from the *substantia niagra* of PD patients and from healthy controls. Using the GEOquery package, data were downloaded into R and checked for missing values (33).

Sample Clustering. Samples in each dataset were hierarchically clustered within GEOquery to detect outliers. One sample from GSE44771 (GSM1090949) and one sample from GSE20292 (GSM508732) were found to be outliers and hence removed from further analysis.

Generation of a "Weighted Gene Correlation Network". A distance measure commonly used for coexpression analysis is based on the Pearson's coefficient of correlation; in this approach, gene pairs with a coefficient of correlation below a given cutoff value (e.g., 0.8) are considered as not correlated. However, this kind of "hard thresholding" may be insensitive to subtle and yet important expression patterns (85). We therefore used the WGCNA method (33, 34), which uses a "soft thresholding" and the concept of topological overlap or shared neighbors to identify clusters of coexpressed genes. The soft thresholding method assigns a weight to each pair of interacting genes and uses such weights, along with the topological overlap, to identify modules of coexpressed genes in the expression data (33, 34).

The construction of a Weighted Gene Correlation Network was performed using the R package for WGCNA (33). Absolute values of Pearson's coefficient of correlation were calculated for the expression values of each gene pair across all microarray samples. WGCNA uses a power function to transform the coexpression similarities (given by a similarity matrix $S = [s_{ij}]$) into connection strengths (given by an adjacency matrix $A = [a_{ij}]$):

$$a_{ij} = |s_{ij}|^{\beta}, \quad [1]$$

where β is the soft thresholding power. In unweighted networks, the entries a_{ij} of the adjacency matrix are either 1 or 0, indicating whether or not a pair of nodes is connected. In weighted networks, the values are real numbers ranging from 0 to 1. Due to the noise in microarray data and the limited number of samples, we weighted the Pearson's coefficients of correlation by taking their absolute values and raising them to the power β. To choose the value of β, we observed that many biological networks, especially gene expression networks, have been found to exhibit approximate scale-free topology (86)—that is, the connectivity distribution $p(k)$ for each node k follows a power law, $p(k) = k^{-\gamma}$, with exponent γ. This "scale-free" relationship indicates that there are a few nodes that are highly connected, whereas others have much fewer connections. Through these considerations, we chose $\beta = 6$ (34). This procedure results in a weighted network in which the continuous nature of the gene expression values is preserved (as opposed to unweighted networks); the results are robust with respect to the choice of β, as opposed to the high sensitivity to the cutoff value of unweighted networks.

Identification of Modules in the Weighted Gene Correlation Network. Modules were defined as groups of genes having high correlation and high topological overlap (34). The topological overlap of two nodes refers to their interconnectedness, which is measured as the number of shared neighbors between two nodes. It provides a similarity measure that has been shown to be very useful in biological networks (87) and was used here as the basis for average linking hierarchical clustering to identify modules of coexpressed genes.

MEs. The ME, which is defined as the first PC of a given module, can be considered as a representative of the gene expression profiles in a module (33). The connectivity of a gene i with a module k [$MM_k(i)$] is defined as the Pearson's coefficient of correlation of the expression value of that gene with the ME of the module. It is a measure of MM for a particular gene. Specifically:

$$MM_k(i) = \mathrm{cor}(e(i), E_k), \quad [2]$$

where $MM_k(i)$ is a measure of MM for gene i with respect to module k, e(i) is the expression profile of gene I, and E_k is the eigengene of module k. The intramodular connectivity (kME) is defined as the connectivity of a gene within its own module. The ME is also used to calculate the Pearson's coefficient of correlation and the associated student *P* value of each module with disease status; the disease status is encoded as binary information for disease or healthy.

Module Preservation and Consensus Analysis. WGCNA provides various measures of module preservation statistics, which assess whether or not the interconnections among the genes within a module and connectivity patterns of individual modules (for example, intramodular hub gene status) are preserved between two datasets. To assess the preservation of our disease-associated modules found in the PFC dataset (the network that we analyzed) and in a hippocampal gene expression dataset (test network), we used the modulePreservation function in the WGCNA R package (46). In brief, this function provides an average measure of several preservation statistics generated through many permutations of the data, the $Z_{summary}$ value. In general, modules with $Z_{summary}$ scores > 10 are interpreted as strongly preserved (that is, densely connected, distinct, and reproducible modules), $Z_{summary}$ scores between 2 and 10 are weak to moderately preserved, and $Z_{summary}$ scores < 2 are not preserved (46). Another way to look at module preservation is to rank the modules by their overall preservation in the test set, which gives a relative measure of module preservation. Median rank is a measure that relies on observed preservation statistics rather than the permutation Z statistics (46). It is calculated as described previously (46).

Consensus analysis is a way to identify modules present in several independent datasets. Consensus modules group together genes densely connected in all conditions and are defined from the clustering of consensus similarity:

$$\mathrm{Modules}_{consensus} = \min(\mathrm{Network\ 1}, \mathrm{Network\ 2}). \quad [3]$$

Consensus modules are by construction present (i.e., preserved) in all input datasets. If a module identified in a reference dataset is strongly preserved in test datasets, it would also be a consensus module among the reference and test datasets. Each consensus module has one eigengene per dataset. Eigengene correlation helps to visualize the overall network structure and also to compare a given network between different datasets. An eigengene network (A_{ij}) is defined as a signed network with a soft thresholding power of 1. A preservation network ($Pres_{ij}$) measures the correlation of eigengene correlation among different networks (88):

$$Pres_{ij}^{(1,2\ldots)} = 1 - \left[\max\left(A_{ij}^{(1)}, A_{ij}^{(2)}, \ldots\right) - \min\left(A_{ij}^{(1)}, A_{ij}^{(2)}, \ldots\right)\right], \quad [4]$$

where $Pres_{ij}^{(1,2\ldots)}$ is the preservation network for any networks 1 and 2. The overall mean preservation of eigengene networks is given by (88):

$$D^{(1,2\ldots)} = mean_{i<j} P_{i,j}^{(1,2\ldots)}. \quad [5]$$

1. Chiti F, Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* 75:333–366.
2. Holtzman DM, Morris JC, Goate AM (2011) Alzheimer's disease: The challenge of the second century. *Sci Transl Med* 3:77sr1.
3. Balch WE, Morimoto RI, Dillin A, Kelly JW (2008) Adapting proteostasis for disease intervention. *Science* 319:916–919.
4. Hipp MS, Park S-H, Hartl FU (2014) Proteostasis impairment in protein-misfolding and -aggregation diseases. *Trends Cell Biol* 24:506–514.
5. Knowles TP, Vendruscolo M, Dobson CM (2014) The amyloid state and its association with protein misfolding diseases. *Nat Rev Mol Cell Biol* 15:384–396.
6. Labbadia J, Morimoto RI (2015) The biology of proteostasis in aging and disease. *Annu Rev Biochem* 84:435–464.
7. Eisenberg D, Jucker M (2012) The amyloid state of proteins in human diseases. *Cell* 148:1188–1203.
8. De Strooper B, Karran E (2016) The cellular phase of Alzheimer's disease. *Cell* 164:603–615.
9. Selkoe DJ, Hardy J (2016) The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol Med* 8:595–608.
10. Gidalevitz T, Ben-Zvi A, Ho KH, Brignull HR, Morimoto RI (2006) Progressive disruption of cellular protein folding in models of polyglutamine diseases. *Science* 311:1471–1474.
11. Chapman E, et al. (2006) Global aggregation of newly translated proteins in an Escherichia coli strain deficient of the chaperonin GroEL. *Proc Natl Acad Sci USA* 103:15800–15805.
12. David DC, et al. (2010) Widespread protein aggregation as an inherent part of aging in C. elegans. *PLoS Biol* 8:e1000450.
13. Koga H, Kaushik S, Cuervo AM (2011) Protein homeostasis and aging: The importance of exquisite quality control. *Ageing Res Rev* 10:205–215.
14. Olzscha H, et al. (2011) Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell* 144:67–78.
15. Walther DM, et al. (2015) Widespread proteome remodeling and aggregation in aging C. elegans. *Cell* 161:919–932.

16. Ciryam P, Tartaglia GG, Morimoto RI, Dobson CM, Vendruscolo M (2013) Widespread aggregation and neurodegenerative diseases are associated with supersaturated proteins. *Cell Reports* 5:781–790.
17. Ciryam P, Kundra R, Morimoto RI, Dobson CM, Vendruscolo M (2015) Supersaturation is a major driving force for protein aggregation in neurodegenerative diseases. *Trends Pharmacol Sci* 36:72–77.
18. Tartaglia GG, Pechmann S, Dobson CM, Vendruscolo M (2007) Life on the edge: A link between gene expression levels and aggregation rates of human proteins. *Trends Biochem Sci* 32:204–206.
19. Baldwin AJ, et al. (2011) Metastability of native proteins and the phenomenon of amyloid formation. *J Am Chem Soc* 133:14160–14163.
20. Ciryam P, et al. (2017) Spinal motor neuron protein supersaturation patterns are associated with inclusion body formation in ALS. *Proc Natl Acad Sci USA* 114:E3935–E3943.
21. Ben-Zvi A, Miller EA, Morimoto RI (2009) Collapse of proteostasis represents an early molecular event in Caenorhabditis elegans aging. *Proc Natl Acad Sci USA* 106: 14914–14919.
22. Freer R, et al. (2016) A protein homeostasis signature in healthy brains recapitulates tissue vulnerability to Alzheimer's disease. *Sci Adv* 2:e1600947.
23. Brehme M, et al. (2014) A chaperome subnetwork safeguards proteostasis in aging and neurodegenerative disease. *Cell Reports* 9:1135–1150.
24. Hartl FU, Bracher A, Hayer-Hartl M (2011) Molecular chaperones in protein folding and proteostasis. *Nature* 475:324–332.
25. Douglas PM, Dillin A (2010) Protein homeostasis and aging in neurodegeneration. *J Cell Biol* 190:719–729.
26. Khurana V, et al. (2017) Genome-scale networks link neurodegenerative disease genes to α-synuclein through specific molecular pathways. *Cell Syst* 4(2):157–170.e114.
27. Ricciarelli R, et al. (2004) Microarray analysis in Alzheimer's disease and normal aging. *IUBMB Life* 56:349–354.
28. Miller JA, Oldham MC, Geschwind DH (2008) A systems level analysis of transcriptional changes in Alzheimer's disease and normal aging. *J Neurosci* 28:1410–1420.
29. Ciryam P, et al. (2016) A transcriptional signature of Alzheimer's disease is associated with a metastable subproteome at risk for aggregation. *Proc Natl Acad Sci USA* 113: 4753–4758.
30. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 95:14863–14868.
31. Stuart JM, Segal E, Koller D, Kim SK (2003) A gene-coexpression network for global discovery of conserved genetic modules. *Science* 302:249–255.
32. Heyer LJ, Kruglyak S, Yooseph S (1999) Exploring expression data: Identification and analysis of coexpressed genes. *Genome Res* 9:1106–1115.
33. Langfelder P, Horvath S (2008) WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559.
34. Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 4:e17.
35. Blalock EM, et al. (2004) Incipient Alzheimer's disease: Microarray correlation analyses reveal major transcriptional and tumor suppressor responses. *Proc Natl Acad Sci USA* 101:2173–2178.
36. Miller JA, Woltjer RL, Goodenbour JM, Horvath S, Geschwind DH (2013) Genes and pathways underlying regional and cell type changes in Alzheimer's disease. *Genome Med* 5:48.
37. Wu G, et al. (2011) Altered expression of autophagic genes in the peripheral leukocytes of patients with sporadic Parkinson's disease. *Brain Res* 1394:105–111.
38. Miller JA, Horvath S, Geschwind DH (2010) Divergence of human and mouse brain transcriptome highlights Alzheimer disease pathways. *Proc Natl Acad Sci USA* 107: 12698–12703.
39. Lin MT, Beal MF (2006) Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature* 443:787–795.
40. Zhang B, et al. (2013) Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell* 153:707–720.
41. Voineagu I, et al. (2011) Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* 474:380–384.
42. Hawrylycz MJ, et al. (2012) An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* 489:391–399.
43. Oldham MC, et al. (2008) Functional organization of the transcriptome in human brain. *Nat Neurosci* 11:1271–1282.
44. Oldham MC, Horvath S, Geschwind DH (2006) Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proc Natl Acad Sci USA* 103: 17973–17978.
45. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M (2010) KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res* 38:D355–D360.
46. Langfelder P, Luo R, Oldham MC, Horvath S (2011) Is my network module preserved and reproducible? *PLOS Comput Biol* 7:e1001057.
47. Shannon P, et al. (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504.
48. Ihara Y, Morishima-Kawashima M, Nixon R (2012) The ubiquitin-proteasome system and the autophagic-lysosomal system in Alzheimer disease. *Cold Spring Harb Perspect Med* 2:a006361.
49. Treusch S, et al. (2011) Functional links between Aβ toxicity, endocytic trafficking, and Alzheimer's disease risk factors in yeast. *Science* 334:1241–1245.
50. Cirrito JR, et al. (2008) Endocytosis is required for synaptic activity-dependent release of amyloid-β in vivo. *Neuron* 58:42–51.
51. Nixon RA (2013) The role of autophagy in neurodegenerative disease. *Nat Med* 19: 983–997.
52. Vilchez D, Saez I, Dillin A (2014) The role of protein clearance mechanisms in organismal ageing and age-related diseases. *Nat Commun* 5:5659.
53. Kaushik S, Cuervo AM (2015) Proteostasis and aging. *Nat Med* 21:1406–1415.
54. Matus S, Glimcher LH, Hetz C (2011) Protein folding stress in neurodegenerative diseases: A glimpse into the ER. *Curr Opin Cell Biol* 23:239–252.
55. Scheper W, et al. (2007) Rab6 is increased in Alzheimer's disease brain and correlates with endoplasmic reticulum stress. *Neuropathol Appl Neurobiol* 33:523–532.
56. Cooper AA, et al. (2006) Alpha-synuclein blocks ER-Golgi traffic and Rab1 rescues neuron loss in Parkinson's models. *Science* 313:324–328.
57. Molina MF, et al. (2011) Decreased expression of ATP6V1H in type 2 diabetes: A pilot report on the diabetes risk study in Mexican Americans. *Biochem Biophys Res Commun* 412:728–731.
58. Geyer M, Fackler OT, Peterlin BM (2002) Subunit H of the V-ATPase involved in endocytosis shows homology to β-adaptins. *Mol Biol Cell* 13:2045–2056.
59. Colacurcio DJ, Nixon RA (2016) Disorders of lysosomal acidification—The emerging role of v-ATPase in aging and neurodegenerative disease. *Ageing Res Rev* 32:75–88.
60. Namekawa M, et al. (2007) Mutations in the SPG3A gene encoding the GTPase atlastin interfere with vesicle trafficking in the ER/Golgi interface and Golgi morphogenesis. *Mol Cell Neurosci* 35:1–13.
61. Zhu P-P, et al. (2003) Cellular localization, oligomerization, and membrane association of the hereditary spastic paraplegia 3A (SPG3A) protein atlastin. *J Biol Chem* 278: 49063–49071.
62. Brett CL, Wei Y, Donowitz M, Rao R (2002) Human Na(+)/H(+) exchanger isoform 6 is found in recycling endosomes of cells, not in mitochondria. *Am J Physiol Cell Physiol* 282:C1031–C1041.
63. Brodsky FM, Chen C-Y, Knuehl C, Towler MC, Wakeham DE (2001) Biological basket weaving: Formation and function of clathrin-coated vesicles. *Annu Rev Cell Dev Biol* 17:517–568.
64. Rothman JE (1994) Mechanisms of intracellular protein transport. *Nature* 372:55–63.
65. Sannerud R, et al. (2016) Restricted location of psen2/γ-secretase determines substrate specificity and generates an intracellular Aβ pool. *Cell* 166:193–208.
66. Lee H, et al. (2016) Enc1 modulates the aggregation and neurotoxicity of mutant huntingtin through p62 under er stress. *Mol Neurobiol* 53:6620–6634.
67. Han S, et al. (2008) Pam (Protein associated with Myc) functions as an E3 ubiquitin ligase and regulates TSC/mTOR signaling. *Cell Signal* 20:1084–1091.
68. Chen BB, Glasser JR, Coon TA, Mallampalli RK (2012) F-box protein FBXL2 exerts human lung tumor suppressor-like activity by ubiquitin-mediated degradation of cyclin D3 resulting in cell cycle arrest. *Oncogene* 31:2566–2579.
69. Kuchay S, et al. (2013) FBXL2- and PTPL1-mediated degradation of p110-free p85β regulatory subunit controls the PI(3)K signalling cascade. *Nat Cell Biol* 15:472–480.
70. Arndt V, et al. (2010) Chaperone-assisted selective autophagy is essential for muscle maintenance. *Curr Biol* 20:143–148.
71. Yu A, et al. (2014) Protein aggregation inhibits clathrin-mediated endocytosis by chaperone competition. *Proc Natl Acad Sci USA* 111:E1481–E1490.
72. Torres GE, Sweeney AL, Beaulieu J-M, Shashidharan P, Caron MG (2004) Effect of torsinA on membrane proteins reveals a loss of function and a dominant-negative phenotype of the dystonia-associated DeltaE-torsinA mutant. *Proc Natl Acad Sci USA* 101:15650–15655.
73. Hosokawa N, et al. (2008) Human XTP3-B forms an endoplasmic reticulum quality control scaffold with the HRD1-SEL1L ubiquitin ligase complex and BiP. *J Biol Chem* 283:20914–20924.
74. Cuyvers E, Sleegers K (2016) Genetic variations underlying Alzheimer's disease: Evidence from genome-wide association studies and beyond. *Lancet Neurol* 15:857–868.
75. Hinz FI, Geschwind DH (2017) Molecular genetics of neurodegenerative dementias. *Cold Spring Harb Perspect Biol* 9:023705.
76. Zhang Y, James M, Middleton FA, Davis RL (2005) Transcriptional analysis of multiple brain regions in Parkinson's disease supports the involvement of specific protein processing, energy metabolism, and signaling pathways, and suggests novel disease mechanisms. *Am J Med Genet B Neuropsychiatr Genet* 137B:5–16.
77. Danial NN, Korsmeyer SJ (2004) Cell death: Critical control points. *Cell* 116:205–219.
78. Anandatheerthavarada HK, Biswas G, Robin M-A, Avadhani NG (2003) Mitochondrial targeting and a novel transmembrane arrest of Alzheimer's amyloid precursor protein impairs mitochondrial function in neuronal cells. *J Cell Biol* 161:41–54.
79. Devi L, Prabhu BM, Galati DF, Avadhani NG, Anandatheerthavarada HK (2006) Accumulation of amyloid precursor protein in the mitochondrial import channels of human Alzheimer's disease brain is associated with mitochondrial dysfunction. *J Neurosci* 26:9057–9068.
80. Wrobel L, et al. (2015) Mistargeted mitochondrial proteins activate a proteostatic response in the cytosol. *Nature* 524:485–488.
81. Heo J-M, Rutter J (2011) Ubiquitin-dependent mitochondrial protein degradation. *Int J Biochem Cell Biol* 43:1422–1426.
82. Itakura E, et al. (2016) Ubiquilins chaperone and triage mitochondrial membrane proteins for degradation. *Mol Cell* 63:21–33.
83. Wang X, Chen XJ (2015) A cytosolic network suppressing mitochondria-mediated proteostatic stress and cell death. *Nature* 524:481–484.
84. Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30:207–210.
85. Carter SL, Brechbühler CM, Griffin M, Bond AT (2004) Gene co-expression network topology provides a framework for molecular characterization of cellular state. *Bioinformatics* 20:2242–2250.
86. Barabási A-L (2009) Scale-free networks: A decade and beyond. *Science* 325:412–413.
87. Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabási A-L (2002) Hierarchical organization of modularity in metabolic networks. *Science* 297:1551–1555.
88. Langfelder P, Horvath S (2007) Eigengene networks for studying the relationships between co-expression modules. *BMC Syst Biol* 1:54.

Supporting Information

Kundra et al. 10.1073/pnas.1618417114

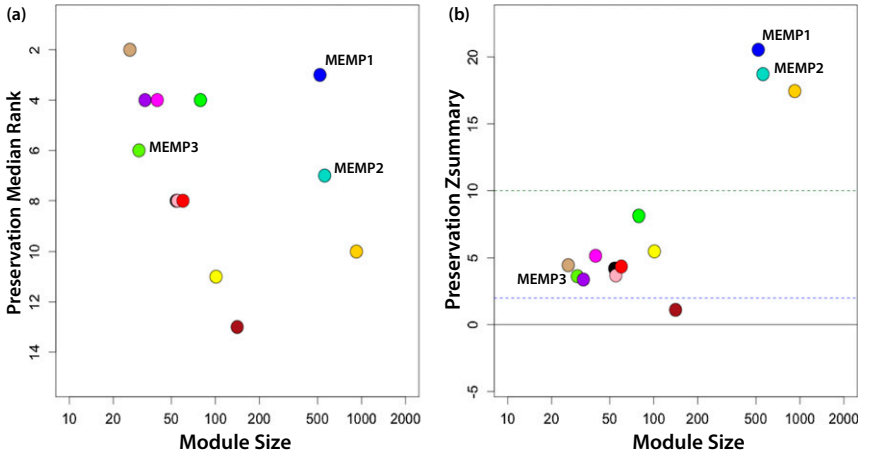

Fig. S1. Module preservation across different datasets. (*A*) Preservation median rank and (*B*) preservation $Z_{summary}$ scores of various modules.

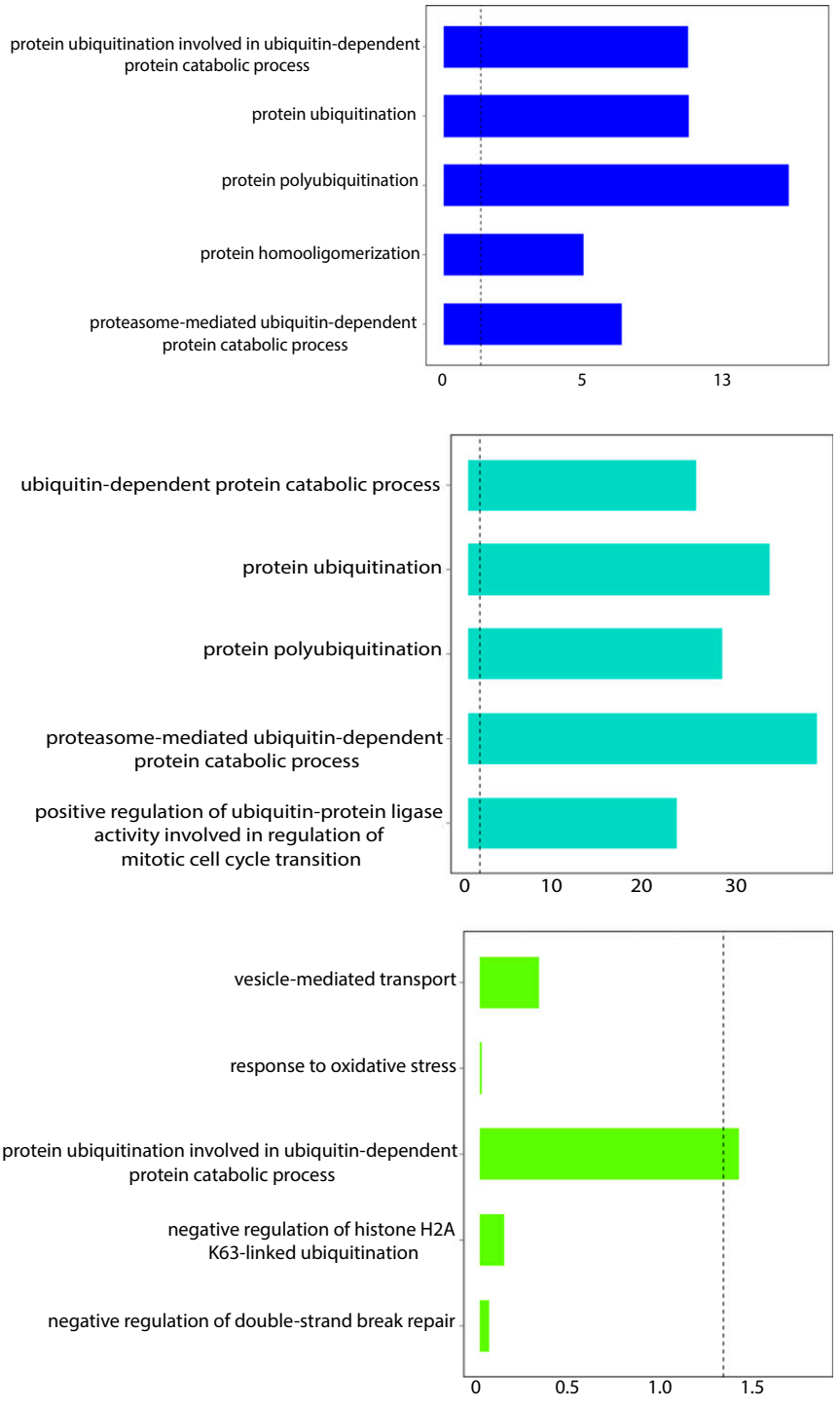

Fig. S2. GO enrichment analysis for the genes in the AD metastable network. This analysis identifies specific components of the ubiquitin–proteasome and endosomal–lysosomal systems in the regulation of the metastable subproteome.

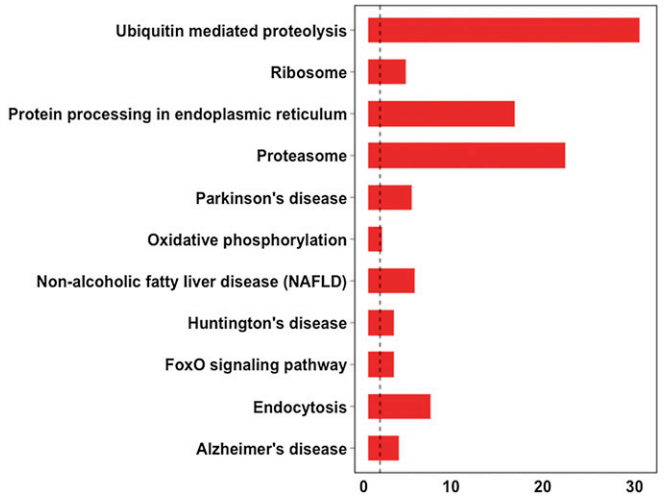

Fig. S3. KEGG pathways enriched in the AD metastable network.

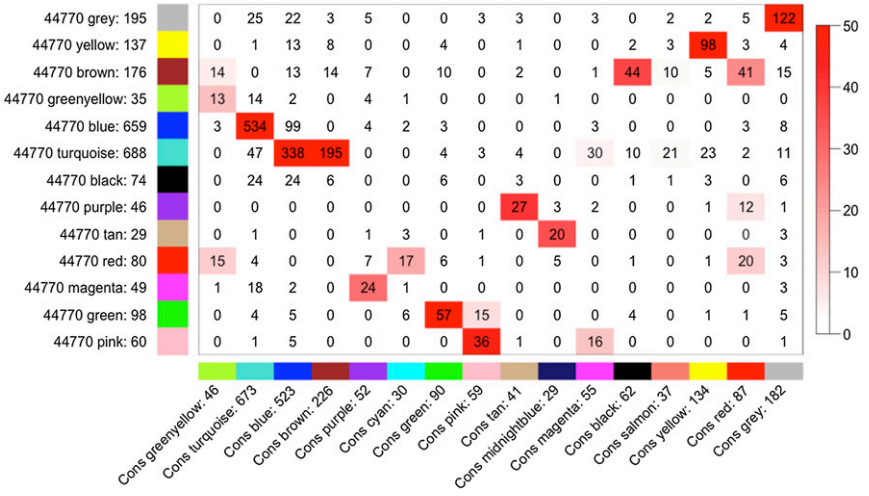

Fig. S4. Correspondence between the modules of the VC dataset (GSE44771) and the consensus modules. Each row of the table corresponds to one of the modules of the VC dataset (labeled by color as well as text), and each column corresponds to one consensus module. Numbers in the table indicate the gene counts in the intersections of the corresponding modules. Coloring of the table encodes −log(p), with p being the Fisher's exact test *P* value for the overlap of the two modules. The table indicates that most of the modules of the VC dataset have a consensus counterpart.

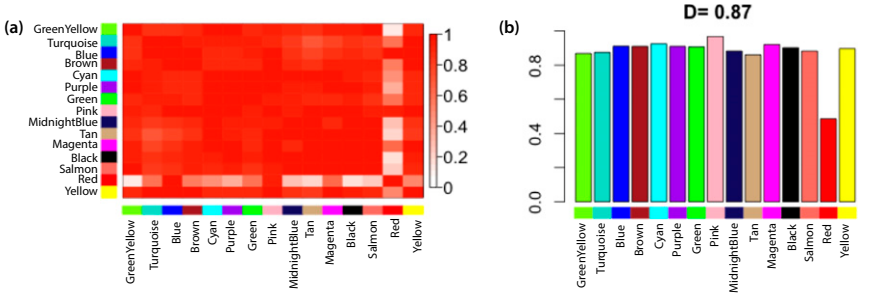

Fig. S5. Consensus eigengene networks for the dorsolateral PFC and the VC. (*A*) Heatmap of the preservation network, defined as one minus the absolute difference of the eigengene networks in the two datasets. (*B*) Mean preservation of adjacency for each of the eigengenes to all other eigengenes. *D* denotes the mean preservation of eigengene networks among the datasets. $D^{(1,2\ldots)} = mean_{i<j} P_{i,j}^{(1,2\ldots)}$ (*Materials and Methods*).

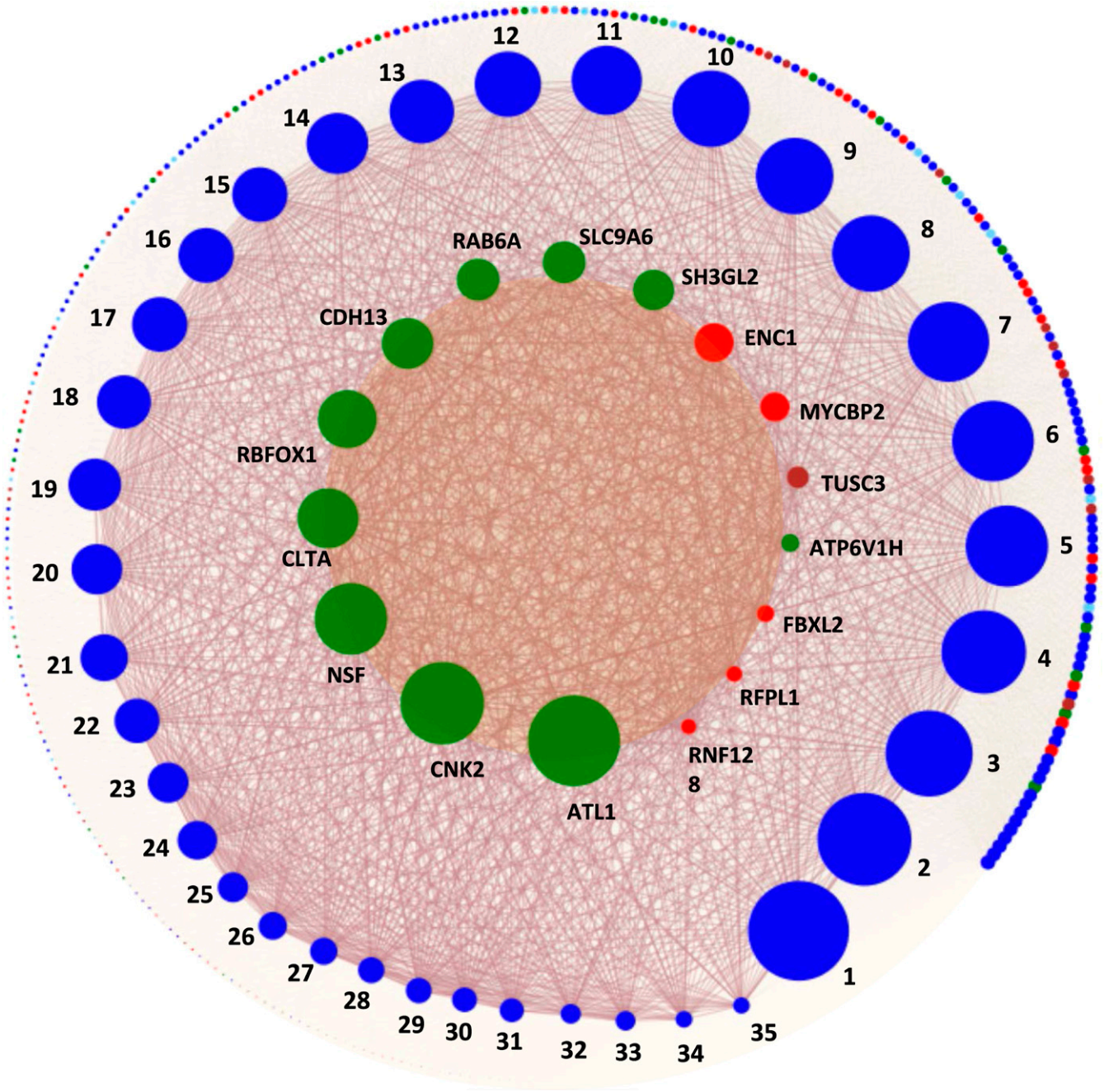

Fig. S6. Network representation of the AD metastable network showing the hub genes and the main components of the protein homeostasis system linked with the AD metastable subproteome. This analysis reveals in particular the importance of the ubiquitin–proteasome (red) and trafficking (green) systems in the regulation of aggregation-prone proteins in AD. The top 10% of the hub gene interactions are visualized, with those genes involved in at least 50 interactions shown in the center. The sizes of the nodes correspond to their degrees of connectivity. The metastable genes are shown in blue. Dataset S1, Table S7 reports the names of the metastable genes according to the numerical labels shown here.

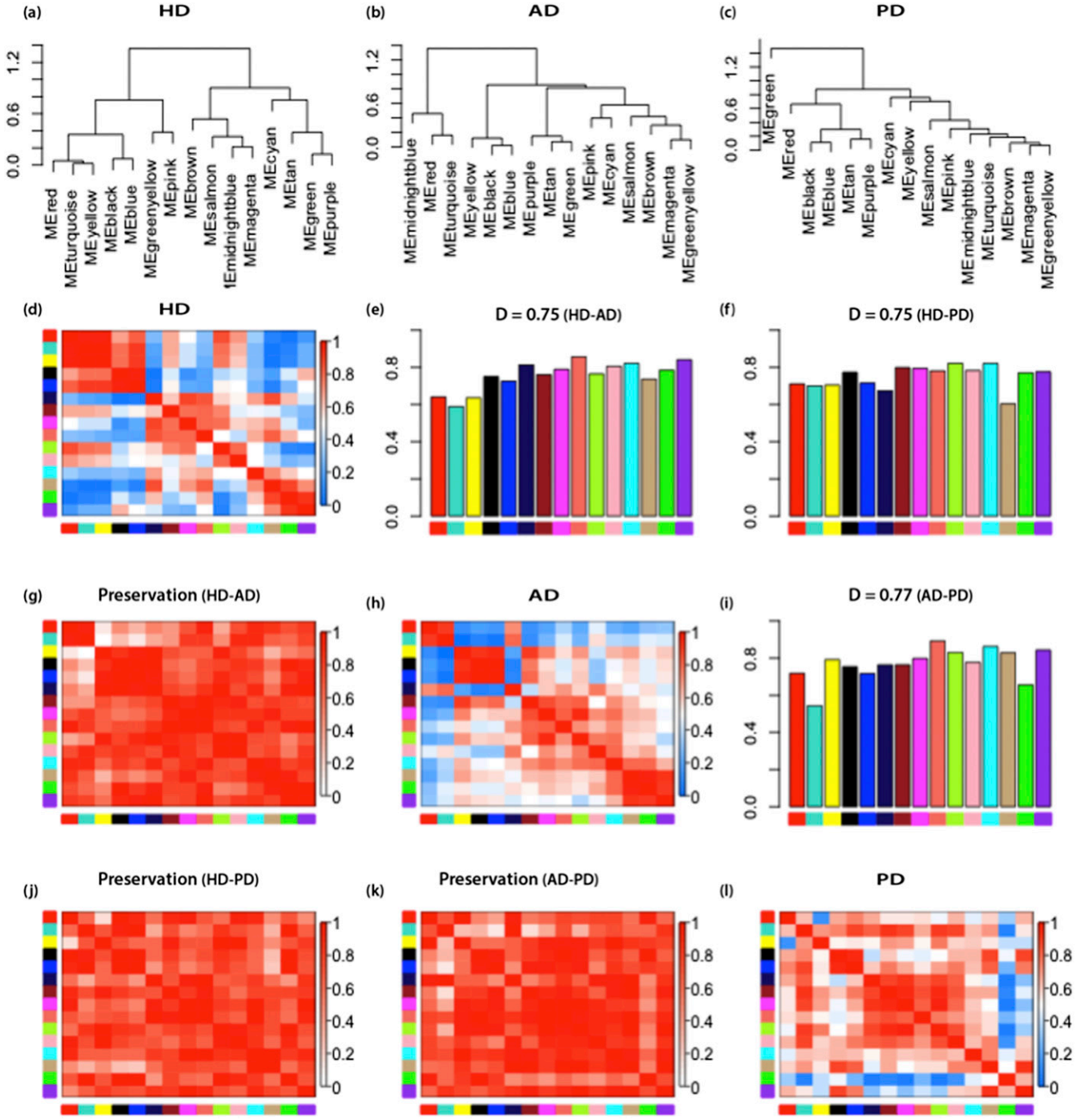

Fig. S7. Consensus eigengene networks and their differential analysis. (*A*–*C*) Dendrograms (clustering trees) of the consensus MEs in the three datasets. (*D*, *H*, and *L*) Eigengene network heatmaps. Red denotes high adjacency (positive correlation), and blue denotes low adjacency. (*G*, *J*, and *K*) Heatmaps of the preservation network, defined as one minus the absolute difference of the eigengene networks in the two datasets. (*E*, *F*, and *I*) Mean preservation of adjacency for each of the eigengenes to all other eigengenes.

Other Supporting Information Files

Dataset S1 (XLSX)