Identifying modifiers of age-dependent protein aggregation in *C. elegans*

Dissertation

zur Erlangung des Grades eines
Doktors der Naturwissenschaften

der Mathematisch-Naturwissenschaftlichen Fakultät
und
der Medizinischen Fakultät
der Eberhard-Karls-Universität Tübingen

vorgelegt
von

Marie Lechler
aus Gräfelfing, Deutschland

Dezember – 2017

Tag der mündlichen Prüfung: 19.07.2018

Dekan der Math.-Nat. Fakultät: Prof. Dr. W. Rosenstiel
Dekan der Medizinischen Fakultät: Prof. Dr. I. B. Autenrieth

1. Berichterstatter: Prof. Dr. Peter Heutink

2. Berichterstatter: Prof. Dr. Ralf Sommer

Prüfungskommission: Prof. Dr. Thomas Gasser

Prof. Dr. Peter Heutink

Prof. Dr. Ralf Sommer

Dr. Della David

Erklärung / Declaration:

Ich erkläre, dass ich die zur Promotion eingereichte Arbeit mit dem Titel:

„Identifying modifiers of age-dependent protein aggregation in *C. elegans*“

selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen als solche gekennzeichnet habe. Ich versichere an Eides statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Mir ist bekannt, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

I hereby declare that I have produced the work entitled

*“*Identifying modifiers of age-dependent protein aggregation in *C. elegans”,*

submitted for the award of a doctorate, on my own (without external help), have used only the sources and aids indicated and have marked passages included from other works, whether verbatim or in content, as such. I swear upon oath that these statements are true and that I have not concealed anything. I am aware that making a false declaration under oath is punishable by a term of imprisonment of up to three years or by a fine.

Tübingen, den ....

Datum / Date Unterschrift / Signature

Abstract

The misfolding of specific proteins and their accumulation in insoluble aggregates has long been recognized as a pathological hallmark of several neurodegenerative diseases. In recent years, widespread protein aggregation occurring during healthy aging has become a hot topic of research. However, to this date little is known about the regulation of this aggregation, the tissue-specificity and the consequences in a disease context.

This thesis answers several questions about different aspects of protein aggregation with aging and in disease. Notably, we analysed the solubility of RNA-binding proteins that are important for the formation of stress granules (sgRBPs) in the nematode *Caenorhabditis elegans* (*C. elegans*). We showed the impact of sgRBP insolubility on organismal health and the importance of maintaining their solubility in long-lived animals. We identified regulators of sgRBP aggregation. In addition, we showed that aggregation-prone sgRBPs are highly prone to interact with other proteins and that this co-localization can influence aggregation patterns or protein localization. Furthermore, we analysed the tissue-specificity of the regulation of age-related protein aggregation. Disruption of the protein-quality control network has contrasting effects on protein aggregation in different tissues, surprisingly reducing age-related protein aggregation in the pharyngeal muscle of *C. elegans*. Specifically, we showed that impaired protein-quality control prevented the accumulation of newly synthesized aggregation-prone proteins. Additionally we demonstrated how screening approaches identifying mutations that influence disease-associated phenotypes, like protein aggregation in *C. elegans*, can help prioritise variants found by whole exome sequencing in large cohorts of patients with Parkinson's disease. To validate promising candidates found to be influencing protein aggregation in *C. elegans*, we have established a cell culture model of age-related protein aggregation.

In conclusion, these findings give important insights into mechanism and regulation of age-related protein insolubility and highlight the importance of age-related protein aggregation for neurodegenerative diseases.

Table of Contents

1 Synopsis

1.1 Introduction

There are many different definitions for the term aging, but the most common one is the age-dependent decline of the intrinsic physiological function increasing the likelihood for death and a decrease in age-specific reproductive rate [1]. Common features of human aging, like tissue deterioration, are evolutionally conserved from humans to smaller organisms like the nematode *Caenorhabditis elegans* (*C. elegans*) [2]. Furthermore, several genetic pathways and biochemical processes are known to be controlling aging [3]. With the increase of human life expectancy and aging being the main risk factor for neurodegenerative diseases, the understanding of the aging process in connection to these diseases is a priority [4, 5].

Most neurodegenerative diseases are characterized by the deposition of specific insoluble proteins in aggregates. Examples for this are Alzheimer's disease (AD), with deposits containing the protein Amyloid-β (Aβ), Parkinson's disease (PD), characterized by the deposition of α-synuclein, Huntington's disease (HD), with inclusions containing Huntingtin, and the prion-diseases (spongiform encephalopathies, Creutzfeldt–Jacob disease) [6]. These proteins associated with diseases differ in size, amino acid composition and sequence. Also the localization of the protein aggregates is diverse: they can be found in the nucleus (Huntingtin in HD), in the cytoplasm (α-synuclein in PD and tau in AD) or extracellularly (in prion diseases and Aβ in AD) [7]. However, a common denominator in each of these neurodegenerative diseases is the amyloid conformation of the aggregates characterized by unbranched fibrils positive for Congo Red staining and producing a cross-beta fiber diffraction pattern [8, 9]. Those amyloid fibrils originate from native proteins that misfold and form oligomers, which develop into the mature fibrils [10]. Recently, the smaller soluble oligomers have been demonstrated to be a main source of toxicity in several neurodegenerative diseases [11, 12].

In addition to the aggregation of specific proteins in a disease-context, is has recently been shown that a part of the proteome is also aggregating with age in the absence of disease. Studies in *C. elegans*, yeast, *Drosophila* and mouse show an age-dependent drastic increase in insolubility of a subset of proteins [13-20]. To a lesser extent, changes in protein insolubility with age were also observed in the hippocampal region from rat brains [21]. Among the proteins aggregating with age proteins of the functional categories of the proteasome, chaperones and ribosomal proteins were overrepresented [13]. Also in structure and amino acid composition the proteins aggregating with age differ from the whole proteome, as proteins with enrichment in aliphatic amino acid residues or high beta-sheet content were prone to aggregate with age [13].

C. elegans is model of choice for studying the genetic basis of aging and age-related diseases, as wild type animals have a short lifespan of approximately two weeks, which is furthermore associated with characteristic features observed in mammalian aging, like tissue deterioration [22]. Many modifiers of lifespan and models for aggregation-related diseases have been discovered in *C. elegans* [23]. Age-related protein aggregation seems to be impairing the health of the organism, as it has been shown in *C. elegans* that reduction of the protein level of aggregation prone proteins leads in half of the tested candidates to an extended lifespan [16]. Furthermore, slowing aging through reduced Insulin/IGF-1 like signalling by impairing the function of the Insulin/IGF-1 receptor DAF-2, decreased protein insolubility of a certain subset of proteins, which could be specifically very important for the longevity of these worms [13].

One important question is why proteins aggregate with age. In young healthy cells protein homeostasis is balanced by the concerted action of protein folding and refolding by molecular chaperones, protein degradation by autophagy or the proteasome, detoxification and adaptive stress responses, like the heat shock response or the unfolded protein response [24]. Aging is accompanied by a decline in these proteostasis control mechanisms and an increase in oxidative stress and protein modifications. Furthermore, the fidelity of protein synthesis is decreasing with aging, leading to increasing numbers of misfolded proteins [24], which together with impaired proteostasis control could lead to the increase of protein aggregation with aging.

Genetic mutations or polymorphisms causing variations in the amino acid sequence of a protein can influence the stability of the native state of a protein [24], and can be risk factors for certain neurodegenerative diseases, as discussed below.

Factors regulating proteostasis are known to be expressed to varying degrees in different tissues [25, 26], indicating tissues specificity for the occurrence and extend of protein aggregation. Furthermore, selective vulnerability of certain brain regions to protein aggregation is well established in a disease context [27, 28]. However, little is known about the susceptibility of different tissues to age-related protein aggregation and tissue-specific modifiers thereof. In Chapter 1 of this thesis, we demonstrate that age-related protein aggregation is differentially regulated in *C. elegans* in a tissue-specific manner. Furthermore, we bring new insights into the susceptibility of newly synthesized and older proteins to aggregate in the aging organism.

The mechanism of age-related protein aggregation influencing disease-related aggregation is not yet fully understood. One possibility is that age-related protein aggregation is stressing the cell and using up the anti-aggregation factors of the cell, thereby indirectly accelerating disease-related aggregation. Another possibility is the direct interaction and coaggregation of age-related and disease-related proteins [14]. Apart from homologous seeding as a cause for disease propagation [28] heterologous cross-seeding between disease-associated proteins has been shown [29]. Recently, it has been shown that insoluble proteins from aged *C. elegans* or aged mouse brains directly interact and initiate Aβ aggregation in vivo, showing that age-related aggregation can cross-seed Aβ aggregation conserved from *C. elegans* to mammals [14]. Furthermore, proteins known to be minor components in the disease-associated aggregates, like Aβ plaques and neurofibrillary tangles, are found to be aggregating with age [13, 14]. The large fraction of minor components aggregating with age themselves could induce heterologous seeding of other disease-associated proteins [14].

There is accumulating evidence, that RNA-binding proteins (RBPs) are linked to pathogenesis in several age-related neurodegenerative diseases. Pathological inclusions of RBPs are found associated with several familiar and sporadic forms of neurodegenerative diseases and these hallmark inclusions often contain RNA. Several pathogenic mutations have been discovered in genes encoding RBPs, among them TDP-

43 and FUS, which if mutated cause amyotrophic lateral sclerosis (ALS) and related diseases and frontotemporal lobar degeneration (FTLD), respectively [30-33]. Additionally, further RBPs, namely EWSR1, TAF15, hnRNPA2B1 and hnRNPA1, are linked to neurodegenerative diseases [34, 35]. Genetic mutations in FUS and hnRNPA1 were shown to increase the aggregation propensity of these proteins from a liquid-like to an aggregated state [36-38]. Many of these RBPs, including FUS, TDP-43, TIA-1 and hnRNPA1, contain a prion-like low complexity (LC) domain. These sequences of low complexity contain a subset of polar amino acids with repetitive nature and were initially identified in prion proteins in budding yeast [39-41]. The formation of membrane-less compartments by liquid-liquid phase separation (LLPS) is a process in which super-saturation of an initially homologous solution leads to demixing and formation of liquid droplets [42-44]. Recent findings show that the LC domain is required for this process [36, 38]. Among the compartments formed by LLPS are the nucleolus and RNA granules like stress granules (SGs).

In times of cellular stress, synthesis of proteins related to stress adaptation is prioritized and synthesis of housekeeping genes is inhibited [45]. The mRNA and RBPs released from the polyribosome after translation inhibition can be stored and protected in SGs [46]. Those highly dynamic structures are formed rapidly by self-assembly of RBPs with LC domains and their RNA targets. Apart from additional factors like 40S ribosomal subunits, SGs contain proteins involved in mRNA stabilization, processing and translation, such as the Poly(A)-binding protein PABP-1, T cell internal antigen-1 (TIA-1) and TIA-1-related (TIAR) [47-50]. Those proteins can be used as SG markers, as they are found specifically in SGs, but not in other cytoplasmic RNA granules [50]. Reversible post-translational modifications regulate SG assembly, composition and also their rapid disassembly after subsiding of stress conditions [51-59].

Recent findings distinguish between physiological SGs and aberrant SGs. Aberrant SGs lose their dynamic properties, accumulate misfolded proteins and can sequester factors of protein-quality control and the autophagy machinery [60, 61]. Also many disease-associated RBPs, like FUS and TDP-43, which are predominantly nuclear proteins, are recruited into cytoplasmic SGs and found to be aggregating in aberrant SGs [50, 60].

The presence of misfolded SG proteins and the accumulation of other misfolded proteins inside the aberrant SGs lead to a decrease in dynamics and resistance to RNase treatment, suggesting that the SGs have reached a stabile conformation independent of RNA content [61]. It has been shown that disease associated mutations in genes encoding SG components increase their propensity to aggregate and thereby disturb SG dynamics [35, 36, 62-64]. In Chapter 2 and Chapter 3 we show that several SG marker RBPs (sgRBPs) are aggregating with age in *C. elegans* and that their age-related aggregation is reduced by activation of longevity pathways, indicating the importance of their functionality for longevity. We further demonstrate the impact of sgRBP aggregation on health, its modifiers and its interactors.

Identification of pathogenic mutations in genes encoding RBPs [30-33] brought these proteins in the focus of research and therefore led to the advanced insights into the pathogenesis of several neurodegenerative diseases. These advances show the importance of identifying risk mutations associated with neurodegenerative diseases to help solve the underlying disease processes. One of the diseases in which identification of genes and variants associated with pathology have furthered understanding of pathogenesis is PD. PD is the second most common neurodegenerative movement disorder with a prevalence of 2-3% in individuals over 65 years [65]. It is characterized symptomatically by a wide spectrum, including motor symptoms such a tremor, slow movement, rigidity and gait and postural impairment as well as non-motor features including cognitive impairment, depression and sleep disorders [66-68]. The pathological hallmark of PD is the loss of predominantly dopaminergic neurons in the substantia nigra pars compacta associated with intracellular insoluble proteinaceous inclusions called Lewy bodies, which include aggregated α-synuclein [67]. While the molecular basis of the disease is not completely understood, there have been huge advances in the establishment of the genetic background [69]. Only 5-10 % of the disease cases are familial, the majority of PD cases are sporadic and originate probably from a combination of genetic and environmental risk factors [69]. Current treatments of PD focus on the cardinal motor symptoms but identification of genes associated with pathogenesis of PD aims at enhancing knowledge on the underlying molecular basis and thereby improving disease modifying treatments [70]. In Chapter 4 we demonstrate the

power of a new approach which is combining whole exome sequencing of large cohorts of PD patients and controls with a series of functional screenings regarding protein aggregation and other disease phenotypes to identify promising new PD risk gene candidates.

Collectively, this thesis is bringing light into diverse aspects of protein aggregation with age and in disease. We demonstrate tissue-specificity of age-related protein aggregation and show the specific reduction of newly synthesized aggregation prone protein as a tissue-specific response to proteostasis stress. We highlight the impact of age-related sgRBP insolubility on organismal health, its connection to neurodegenerative diseases and determine regulators and interactors of sgRBP aggregation. Additionally, we show how screening approaches identifying mutations with influence on disease associated phenotypes like protein aggregation, can help prioritise variants found by whole exome sequencing in large cohorts of PD patients. Furthermore, we have established a cell culture model of age-related protein aggregation to validate promising candidates found to be influencing protein aggregation in *C. elegans* in cell culture.

1.2 Chapter 1

Tissue-specific reduction of aggregation-prone protein levels following disruption of the quality control network as regulator of age-related protein aggregation in *C. elegans*

Lechler et al, in preparation

The focus of this project was to understand more about the regulation of age-related protein aggregation. Furthermore, we were interested in the susceptibility of different tissues to age-related protein aggregation.

In a disease context it is well established that different tissues and cell types vary in their propensity to accumulate misfolded disease-associated proteins [27, 28]. Increase in protein insolubility with age has been shown in *C. elegans* using the whole animal and reproduced with individual proteins in several tissues [13, 14, 16, 17]. We were wondering whether age-related aggregation of a protein is occurring to a varying extend and with differing regulation in different tissues. We chose to investigate differences in endogenous protein aggregation in two types of muscle tissues – the striated body wall muscle and the non-striated pharyngeal muscle. To analyse this we used two proteins found in our previous study [13] to be aggregating with age, RHO-1 and KIN-19, each with a fluorescent tag.

We found strong differences in the aggregation propensity in the two different tissues. While KIN-19 formed highly insoluble puncta in both tissues with age, but slower in the body wall muscle, RHO-1 aggregated strongly in the pharynx but not in the body wall muscle. Importantly, we found that aggregation is differentially regulated in the two muscle types: Disruption of proteostasis led to the expected increase of insoluble aggregate formation in the body wall muscle. Surprisingly, we found that disruption of proteostasis in the pharyngeal muscle, either by inhibiting parts of the protein degradation machinery or by disrupting chaperone expression, leads to a reduction of protein aggregation.

What are the possible reasons behind the opposing influence of proteostasis disruption on protein insolubility in the pharyngeal and the body wall muscle? Factors known to influence aggregation behaviour of different proteins are pH, temperature, amino acid sequence of the protein, ion strength, expression level of molecular chaperones and other proteins preventing aggregation and the concentration of the protein and cosolutes [6]. From our data we can conclude that the cellular environment of both cell types is favouring protein aggregation, as KIN-19 aggregates in both tissues with age. However, differences in the mentioned factors could be responsible to the differing responses to proteostasis stress between the tissues. The reduced aggregation propensity of both RHO-1, which is expressed under the control of a strong promoter, and KIN-19, which is expressed at much lower levels than RHO-1, makes it unlikely that protein concentration of the aggregating protein is the deciding factor for the different response to proteostasis disruption. Furthermore, differential expression of chaperones seems not to be the critical factor, as knockdown of the key transcription factor for chaperones in *C. elegans*, HSF-1, also leads to reduced aggregation in the pharyngeal muscle while increasing aggregation in the body wall muscle.

Knockdown of several negative regulators of the heat shock response (HSR), among them proteasomal subunits and chaperone machineries, have previously been shown to upregulate the HSR tissue-specifically, likely through activation of HSF-1 [25, 26, 71]. This demonstrates that upregulation of one part of the proteostasis machinery can compensate for another. We tested the influence of double knockdown of autophagy and HSF-1 and found that it did not restore aggregation in the pharynx, indicating that there is no compensatory upregulation of degradation or the HSR after proteostasis disruption.

It would be very interesting for future research to find the deciding factors or cellular pathways in the different tissues which are responsible for the switch from proteostasis disruption leading to increased or reduced aggregation levels. A promising tactic for solving this puzzle would be the application of tissue-specific proteomics or RNA sequencing approaches to identify the exact conditions leading to improved proteostasis conditions in the cell.

We next focused our research on the reason behind protein aggregation occurring with age. Is protein aggregation due to the slow accumulation of older misfolded proteins during the aging process, leading to the formation of visible aggregates in the old animal or is it due to a collapse of proteostasis in the aged animal because of the decrease in degradation efficiency and chaperone expression with age, causing proteins newly synthesized in the aged animals to aggregate? To investigate these possibilities we used the photoconvertible tag mEOS2, allowing to differentiate between proteins and aggregates present before photoconversion and those newly formed after conversion.

We found that mainly proteins synthesized during later life stages are responsible for the strong increase of protein aggregation with age. While older protein aggregates were removed over time, the amount of aggregates consisting of newly synthesized proteins is increasing continuously, causing the evident increased levels of aggregation in aged animals. It is known that a fraction of newly synthesized proteins, so called defective ribosomal products (DRIPs), are directly targeted for degradation and accumulate in the insoluble fraction of proteins following proteasome inhibition. Those DRIPs are unable to achieve a stable conformation due to errors in protein folding or translation, e.g. by truncation [72-74]. Our findings of newly synthesized proteins aggregating with age show that mechanisms for degrading DRIPs are also likely to fail during aging, causing those newly synthesized proteins to accumulate and aggregate. Our results are strengthened by recent finding showing that the vulnerability of newly synthesized proteins to thermal stress leads to protein insolubility of these proteins, indicating that proteostasis disruption has the greatest impact on newly synthesized proteins [75]. We propose that with age, as the protein-quality control system decreases, the newly synthesized proteins that are directly degraded in young animals, accumulate and aggregate. DRIPs could therefore be making a large contribution to the age-related increase in protein aggregation. Furthermore, aging is accompanied by an increase of free radicals and reactive oxygen and nitrogen species leading to proteostasis stress, which has been shown to impact aggregation of newly synthesized proteins [24, 75].

We find that proteostasis disruption is preventing the aggregation of newly synthesized proteins in the pharynx, but not in the body wall muscle. Importantly, this is achieved by

a reduction in protein levels specifically of aggregation-prone proteins. These results emphasize the importance of newly synthesized proteins in the age-related aggregation process and regulation thereof. Newly synthesized proteins that are especially vulnerable to proteostasis stress [75] are aggregating with age and disruption of proteostasis can impact this aggregation in a tissue-specific manner: While enhancing age-related aggregation of newly synthesized proteins in the body wall muscle as expected, proteostasis disruption leads to reduced levels of these newly synthesized aggregation-prone proteins in the pharyngeal muscle.

An interesting topic for future research will be to investigate the mechanism leading to reduced levels of newly synthesized proteins following proteostasis disruption. Possibilities for this effect are either reduced protein synthesis or direct removal of the newly synthesized proteins. It will be important to analyse the basis for the specificity leading to reduced levels of only aggregation-prone proteins. Furthermore, to determine the tissue-specific factors leading to this protective response in one tissue but not the other could be a significant step to developing powerful tools for preventing protein aggregation. Overall, our study provides interesting new insights into tissue-specificity of proteostasis regulation and highlights the vulnerability of newly synthesized proteins to age-related protein aggregation. It shows the importance of studying protein aggregation in a tissue-specific manner, as mechanisms of regulation are not ubiquitous.

Age-related protein aggregation is widely considered to be a significant factor in several neurodegenerative diseases. Therefore finding factors leading to improved proteostasis especially in aged organism could be an essential step in preventing those diseases.

1.3 Chapter 2 and Chapter 3

Reduced Insulin/IGF-1 Signalling Restores the Dynamic Properties of Key Stress Granule Proteins during Aging

Lechler et al, Cell Reports, 2017

More stressed out with age? Check your RNA granule aggregation

Lechler et al, Prion, 2017

In recent years mutations in several known RBPs have been shown to be associated with insoluble aggregates in neurodegenerative diseases. Those disease-associated RBPs contain a LC domain and an RNA-recognition motive (RRM) [39, 40]. The LC domain has been shown in a non-disease context to be required for LLPS, a process leading to the formation of membrane-less organelles, like the nucleolus or SGs [36, 38]. Computational predictions count ~250 proteins with LC domains in the human proteome [40], including 12% containing RRM [76]. As a substantial part of the proteome has been shown to be aggregating with age alone, we were wondering how disposed to aggregate with age RBPs with LC domains are. Furthermore, as aging is the most important risk factors for neurodegenerative diseases [5], we asked how aging alone is impacting SG dynamics and solubility.

In our proteomic study of the insoluble proteome of wildtype and long-lived *daf-2(-)* worms we found that a several RNA granule components aggregate with age and that reduced *daf-2* signalling preferentially abrogates the aggregation of RNA granule components. Four proteins containing LC domains were found to be no longer aggregating in *daf-2(-)* conditions: PAB-1, FIB-1, CAR-1 and HRP-1. As PAB-1, CAR-1 and FIB-1 are known components of different kinds of RNA granules (stress granules, P-bodies and the nucleolus, respectively) it is likely that the dynamic nature of RNA granules is negatively influenced by aging. Apart from the importance of reduced stress granule dynamics discussed below, this would have important consequences for the

removal of aberrant mRNAs by nonsense-mediated decay in P-bodies and the formation of the nucleolus.

The fact that HRP-1, a nuclear localized RBP with LC domain, is aggregating with age is particularly interesting, as HRP-1 is associated with multisystem proteinopathy and ALS. Mutations in hnRNPA1 and hnRNPA3, the human homologs of HRP-1, have been shown to cause the aberrant cytoplasmic accumulations in those diseases [35, 77]. We overexpressed fluorescently tagged HRP-1 and could confirm the nuclear localization in young animals. However, aging the animals with elevated temperatures as mild stress leads to mislocalization of HRP-1 in cytoplasmic accumulations. This indicates that without disease associated mutations aging together with additional stressors alone can lead to the mislocalization of RBPs with LC domains and therefore could be the first step in the development of neurodegenerative diseases associated with cytoplasmic accumulation of those proteins.

The overrepresentation of RNA granule components being protected from aggregation in *daf-2* conditions indicates that the functionality and solubility of those proteins is of specific importance for longevity. We therefore decided to further analyse RNA granule solubility and focus on the dynamics of sgRBPs with aging in vivo. We chose to investigate this topic based on the proteins PAB-1 and TIAR-2, the homologues of the known human stress granule markers PABP-1 and TIA-1. Both RBPs contain a LC domain and have also been shown to be minor components of pathological inclusions in ALS and FTLD [39, 78]. Furthermore, PABP-1 and TIA-1 have been found to accumulate in pathological aggregates causing two rare diseases [63, 79]. Overexpressing PAB-1 and TIAR-2 with a fluorescent tag under the control of the pharyngeal muscle promotor allowed us to visualize dynamic formation and disassembly of SGs after heat shock.

Strikingly, we found that both SG markers form solid aggregates with age in *C. elegans*. These results clearly demonstrate that even in the absence of disease and mutations, aging alone leads to insolubility of certain important SG markers. Recent work in human cell culture shows that aberrant SGs accumulating misfolded proteins are less dynamic and exhibit less fusion and fission events. They are stabile after RNase treatment and contain other SG components that are impaired in their motility, indicating that the presence of misfolded proteins in these SGs impairs their function [61]. Our findings

suggest that with aging key sgRBPs become insoluble themselves. It has yet to be demonstrated that age-related SG protein aggregation directly impairs SG dynamics and function. However, the recruitment of other misfolded proteins into aberrant SGs strongly suggests that age-dependent misfolded sgRBPs also accumulate in SGs and disturb the dynamic nature and fusion and fission events [61].

Recent in vitro studies on purified sgRBPs capable of LLPS show their slow maturation into a more solid state with time, a process enhanced by repeated cycles of liquid droplet formation and dissolution [37, 38]. As cells are subjected to various stresses during the lifetime of an organism, we asked whether repeated LLPS of SGs likely to transpire in an aging organism leads to formation of aberrant SGs. We show that sustained SG formation in vivo by mild heat stress strongly increased SG protein aggregation. These results indicate that continuous LLPS during aging is enhancing aggregation and therefore dynamic impairment of sgRBPs also in vivo. Collectively, these results show that SG proteins become insoluble with age, possibly enhanced by repeated exposure to stresses or by a general decrease in proteostasis control.

Another important question is whether aggregation of sgRBPs in the absence of disease influences the organism. We show in *C. elegans* that aggregation of sgRBPs impairs the health of the animal, demonstrated by smaller body size, slower movement and less stress resistance of worms with higher aggregation levels. Furthermore, our results reveal that three independent longevity pathways in *C. elegans*, reduced Insulin-IGF-1-like signalling, mitochondrial inhibition or dietary restriction, prevent aggregation of sgRBPs. This implies that maintaining solubility of sgRBPs and thereby SG dynamics is a common strategy to prolong lifespan and increase organismal fitness.

Research up to this point has put great importance on the role of mutations in disease associated aggregation-prone proteins and their influence on the dynamics and aggregation formation of SGs [38, 61]. Our results now reveal the impact of the aging process on sgRBPs in the absence of disease. But how is sgRBP aggregation occurring with age connected with disease pathology? As multiple RBPs also found to be recruited into SGs, have been shown to be causing cytotoxic aggregates if mutated or to be components of the hallmark aggregates occurring in disease, age-dependent insolubility of those components is very likely to have an impact on disease development [50].

One possibility of how age-related sgRBP aggregation can influence disease development is the recruitment of further proteins into aberrant SGs. Various in vitro and in vivo experiments from yeast and cell culture show that SGs can interact with and accumulate misfolded proteins [14, 61, 64, 80-82]. Several of our results demonstrate the propensity of age-related sgRBP aggregates to interact with other proteins in *C. elegans*: We first show that the two SG markers PAB-1 and TIAR-2 if co-expressed co-localize and change their aggregation pattern with age. Notably, aggregation of PAB-1 pattern changes from formation of large aggregates to smaller puncta reminiscent of SGs. PAB-1 also co-localized with KIN-19, homologue of casein kinase isoform alpha (CK1alpha), a non-SG component (no RRM or LC domain) found in our study to be aggregating with age. Co-aggregation with KIN-19 leads to a strong increase in PAB-1 aggregation. Furthermore, overexpression of FUST-1, the *C. elegans* homolog of disease-associated human FUS, led to formation of aberrant cytoplasmic FUST-1 aggregates which recruited PAB-1 if co-expressed. These results all demonstrate the strong propensity of sgRBPs aggregating with age to interact with other proteins with and without LC domain in *C. elegans* and influence their aggregation pattern. As the inherently aggregating protein KIN-19 is enhancing sgRBP aggregation and misfolded proteins are known to accumulate in SGs [61], cross-seeding of age-related protein misfolding is likely contributing to sgRBP aggregation with age.

One factor enhancing SG protein aggregation with age is the decrease in proteostasis regulation occurring during aging [83-87]. We were especially interested in finding factors able to modify age-related aggregation of sgRBPs. As we found in our proteomic study key sgRBPs to be prevented from aggregation with age in long-lived animals with reduced insulin/IGF-1-like signalling, we first confirmed reduced age-related aggregation of PAB-1 and TIAR-2 in *daf-2* mutant animals. Furthermore, we show that this reduction in aggregation is dependent on the two transcription factors upregulated in animals with reduced *daf-2* signalling, heat shock factor HSF-1 and DAF-16/FOXO. Interestingly, HSF-1 activity is not required for reduction of pharyngeal KIN-19 aggregation by *daf-2* inhibition, indicating that sgRBP aggregation is differently regulated than other age-related aggregation.

Recent findings show the importance of the chaperone system in regulating SG disassembly and thereby preventing the formation of aberrant SGs. While some aberrant SGs can be degraded by autophagy, most SGs disassemble in a chaperone-mediated manner [60]. In yeast and cell culture, the chaperones HSP-70, HSP-40 and HSP-110 have been shown to co-localize with SG and modulate their disassembly [80, 81, 88, 89]. The yeast HSP40s *sis1* and *ydj1* are essential for the the disassembly and/or clearance of SGs [90] and in mammalian cells several chaperones are recruited into SGs [49, 60, 91, 92]. Recent findings show that the small heat shock protein (HSP) HSPB8 is recruited into SG to prevent interactions between SG components and DRIPs accumulated in SGs. Together with the nucleotide exchange factor BAG3 and HSP70 SG disassembly is mediated [60]. Inhibition of HSP70 has been shown to lead to an accumulation of DRIPs in aberrant SGs [60].

We investigated the influence of the chaperone system on age-related aggregation of sgRBPs in *C. elegans*. Knockdown of the transcription factor HSF-1, which controls the expression of chaperones in *C. elegans*, led to a strong increase in sgRBP aggregation already in young animals, but also with aging. We found that HSF-1 activity specifically during development is important for sgRBP solubility. This is in accordance with findings showing the dependence of *daf-2* mediated longevity on HSF-1 activity during development [93]. Although DNJ-13 and DNJ-19, the *C. elegans* homologs of yeast HSP40s *sis1* and *ydj1*, co-localized with sgRBP aggregates, their overexpression did not modulate sgRBP aggregation. Furthermore, knockdown of HSP70 did not alter aggregation of PAB-1, indicating that those HSPs are not modulating sgRBP aggregation. These results show that sgRBP aggregation with age may not be dependent on the same chaperones influencing SG dynamics. However, upregulation or activation of other chaperones could compensate for the inhibited chaperones, showing functional redundancy.

In the last years mutations in the components of the chaperone system VCP/p97, HSPB1, HSPB8 and BAG3 have been associated with multiple neurological diseases [94-98]. These findings highlight the importance of the chaperone system on SG solubility and dynamics in the context of disease. But how is impairment of the chaperone system connected to diseases associated with aggregating SG components? The following

explanation concerning the impact of diseases-causing mutations of RNPs and mutations in chaperone system components has been proposed: Mutations in RNPs found to be recruited into SGs increase their aggregation propensity and lead to formation of aberrant SG, which accumulate further misfolded proteins and DRIPs. The thus formed aberrant SG containing misfolded proteins are impaired in their functionality. Additionally, deficiency of the chaperone system responsible for SG disassembly and removal of misfolded proteins from SGs lead to retention of aberrant SGs and thereby promote the formation of fibrillary aggregated SG components [47]. Our findings about age-related sgRBP aggregation now broaden this explanation of pathogenesis: We find that sgRBPs aggregate with age alone in the absence of disease associated mutations. This indicates that with aging the functionality of SGs will be impaired, as the SGs accumulate misfolded proteins and become aberrant. The decrease of the functionality of the chaperone system with age is advancing this process by limiting the removal of misfolded proteins and aberrant SGs. Furthermore, the general increase in protein insolubility with age that we previously demonstrated [13] will lead to enhanced localization of misfolded proteins in aberrant SGs with age. Misfolded SG proteins and aberrant SGs are building up during normal aging and are impairing health. However, disease associated mutations accelerate this process, leading to earlier formation of aberrant SGs and toxic aggregates.

Previous findings show that DRIPs accumulate in aberrant SGs [60]. Furthermore, we have shown in Chapter 1 that protein aggregation with age is mostly due to the misfolding of proteins newly synthesized in the aging animal, in contrast to a steady accumulation of older misfolded proteins during the aging process. From this we can conclude that with aging the amount of newly synthesized misfolded proteins leads to an enhanced accumulation of these proteins in aberrant SGs.

Together these results demonstrate the importance of the functionality of sgRBPs during the aging process and in disease. Our results and other advances in the field [47] connect the functionality of sgRBPs with the formation of toxic aggregates in various diseases, showing that understanding the mechanism leading to the formation of aberrant SGs is an important step in solving the pathogenesis of these diseases.

1.4 Chapter 4

Discovery and functional prioritization of Parkinson's disease candidate genes from large-scale whole exome sequencing

Jansen et al, Genome Biology, 2017

Identifying genes and variants responsible for pathogenesis is an important step towards a more complete understanding of the molecular and genetic basis of neurodegenerative diseases. Research on genes associated with PD has mainly been focused on familiar forms of disease, which make up 5-10% of disease cases. To this date this led to the identification of six genes (*SNCA*, *LRRK2*, *VPS35*, *Parkin*, *PINK1* and *DJ-1*) conclusively associated with PD and inherited in an autosomal dominant or recessive manner [69]. As most forms of PD are sporadic, understanding the genetic involvement in pathogenesis is of vital importance.

The application of new technologies in next generation sequencing, such as exome sequencing and the advent of genome-wide association studies (GWAS) have accelerated the progress in the field of PD genetics in the last years [99, 100]. Whole exome sequencing (WES) is a very cost-effective way to sequence the whole genome by only considering the coding regions of the genome. These coding regions only make up 1% of the genome but contain 85% of mutations associated with Mendelian diseases [69, 100, 101]. As WES does yield over 20000 exonic single nucleotide variants per individual, this number has to be narrowed down to the important variants [102]. In sporadic PD cases filtering the variants using segregation analysis is not possible due to the cohort of unrelated individuals, leading to an increased number of candidate variants [103]. Using WES in a large cohort of unrelated PD case/control samples would be a valuable method to discover rare variants. With the absence of DNA samples from family members to perform segregation analysis and the difficulties in statistical analysis in rare variants that only occur in few cases a new approach of this work is to cover this gap using functional screens in mammalian cell culture and the model organisms *C. elegans* and *Drosophila*. This integrated approach helps to prioritize the variants found

by WES analyses for further study by using medium- to high-throughput screening assays.

In this study, our collaborators performed WES in 1148 unrelated young-onset PD cases and 503 control individuals with European ancestry, the largest PD WES dataset studied to date. As young onset of disease is often associated with a recessive inheritance model [104-106], they concentrated on genes with homozygous or putative compound heterozygous loss of function (LoF) variants for the analysis. 920,896 variants were found in this WES dataset and were narrowed down to those found rarely in the control group. 27 LoF variants were identified, 18 of which are homozygous and nine putative compound heterozygous, with nearly all not recurring in more than one PD individual.

Several independent datasets were used for human genetic validation, as no sufficient exome data set was available, showing further genetic evidence for seven genes (*GPATCH2L*, *FAM83A*, *CD36*, *UHRF1BP1L*, *PTPRH*, *ARSB*, and *VPS13C*). Furthermore, the tolerability of the recessive LoF variants was tested by analysing the occurrence of our variants in recently predicted complete gene knock-outs in the Icelandic population [107] and carrying out systematic LoF analysis in *C. elegans*. We performed high-throughput RNAi gene knockdown assays for 66 candidate genes with available orthologs in *C. elegans* and tested effects on survival, development, fertility and the influence on polyglutamine aggregation. We show that knockdown results in developmental arrest and/or reduced lifespan for three of the candidates (*DIS3* (*dis-3*), *KALRN* (*unc-73*), and *PTCHD3* (*ptr-10*)).

The human genetic studies were then coupled with functional screening to strengthen the functional involvement of the 27 candidate genes in well-known PD mechanisms. It has been proposed that mitochondrial dysfunction is involved in PD pathogenesis as some proteins associated with PD, like Parkin (*PARK2*), *DJ-1*, and *PINK1*, play roles in protein-quality control and dynamics of mitochondria [108] while the PD hallmark protein alpha-synuclein has been associated to mitochondrial injury [109]. The influence of LoF candidate genes found in this WES study on mitochondrial function therefore might give insights on the importance of those genes in PD pathology. BE(2)-M17 neuroblastoma cells were used to test the influence of gene knockdown on mitochondrial number, axial length ratio and roundness to determine abnormalities in mitochondrial morphology [110]. 13 candidate genes showing a significant effect on at

least one of these parameters were found. An assay using cells expressing Parkin-GFP was used to test Parkin translocation [111-115] and showed significant modification of Parkin translocation by six candidate genes.

Furthermore, our collaborators tested whether the candidate genes modify alpha-synuclein toxicity in a well-established model of α-synuclein toxicity in *Drosophila* [116-118]. In flies expressing α-synuclein in the retina knockdown of the 13 well-conserved homologs of the candidate genes found by WES was performed and screened for retinal tissue integrity. Knockdown of four of these genes (*ARSB*, *TMEM134*, *PTPRH*, and *VPS13C*) led to robust enhancement of α-synuclein-mediated neurodegeneration in the retina.

This integrated approach of combining genetic data with functional screening in cell culture and animal models identifies five strong candidates for PD susceptibility genes (*GPATCH2L*, *UHRF1BP1L*, *PTPRH*, *ARSB* and *VPS13C*), which are supported by both functional data and the additional genetic analysis consistent with replication. The results demonstrate the power of this method to determine and prioritize rare variants in PD and potentially other neurodegenerative diseases.

There are several problems facing the discovery of rare risk variants in population based PD samples: Due to the absence of available genetic data from large family pedigrees segregation analysis is not possible in the sample of unrelated individuals, leading to a high number of potential candidate variants. This leads to the use of stringent filtering criteria, in this case the focus on only strongly damaging LoF variants and the restriction to a recessive inheritance model to narrow down the variants which concomitantly leads to a loss of other important variants, among them dominantly acting alleles. An important task for future studies will be the analysis of dominantly acting alleles. PD is strongly age-dependent and has a long pre-symptomatic phase, leading to the assumption that certain pathogenic variants might be missed because of their presence in the control cohort, as pre-symptomatic cases of PD in this cohort are not unlikely. In future studies, careful neurological testing to exclude mild early-stage forms of PD from the control cohort and increased sample sizes will be an important improvement. The combination of genetic WES data with functional screening in cell culture and model organisms in this new approach allows reliable prioritization of discovered gene variants for future investigation. This application of functional follow up studies demonstrates an

important advantage of WES over GWAS for this approach, as this is only possible with location of the variants in gene-coding areas.

Our approach in combination with other related studies also allows the grouping of variants into common pathways, leading to a broader understanding of PD pathogenesis. For example in addition to our study *VPS13C* has recently been found in a study in families with autosomal recessive early onset PD and dementia [119]. Our findings of *VPS13C* disrupting mitochondrial morphology are in accordance with this study reporting localization of *VPS13C* to the outer mitochondrial membrane and *VPS13C* LoF association with reduced mitochondrial membrane potential, mitochondrial fragmentation and increased Parkin-dependent mitophagy [119].

These results demonstrate the strength of this study as a powerful and integrative strategy to discover variants in large-scale genomic studies of PD. Analysis in the largest PD WES dataset to date was performed and replication in additional available cohorts was established. This data and the and the additional data from the multiple functional assays with relevance to PD pathology in cell culture and model organisms will be a valuable resource for future studies of PD pathogenesis.

2 Method chapter: Establishing a cell culture model for age-related protein insolubility

2.1 Introduction

The preservation of the functional state of the proteome of individual cells, but also in tissues and organs, is very important for the health of an organism. The tight control of protein homeostasis requires a complex network of cellular pathways. Several layers of protein-quality control, including regulation of transcriptional and translational rates, molecular chaperones, as well as protein degradation by autophagy or the ubiquitin-proteasome system are necessary to keep proteins in a functional state and prevent aggregation [120-122].

However, as the organism ages, the carefully controlled protein-quality control system gets out of balance. Aging has been associated with a decrease in chaperone expression, stress responses and decline in the degradation activity of the proteasomal or autophagic pathways in different model organisms [83-87]. This disruption of proteostasis in the aging organism leads to the formation of misfolded and aggregated proteins. Indeed, aging is the most important risk factor for neurodegenerative diseases associated with deposition of insoluble proteins in aggregates [5]. Furthermore, an increase of widespread non-disease associated protein aggregation with aging has been shown in *C. elegans*, and also in several other model organisms [13-21].

We have been studying age-related protein aggregation and modifiers thereof in the model organism *C. elegans*. These nematodes have long been a useful tool to study age related questions, as their medium lifespan is very short and genetic modifications are relatively easily achievable. However, in this project we were looking for a model more closely related to the human body to be able in the future to verify the influence of modifiers of age-related protein aggregation found in *C. elegans*.

An established method to study cellular aging in cell culture is the serial cultivation of somatic cells and calculation of the number of times the cell number has doubled (population doublings). The proliferation potential of normal cells in culture is limited

due to multiple environmental and genetic mechanisms [123-125] and this process called cellular senescence mimics the cellular and molecular changes associated with development and aging [126]. Interestingly, the replicative lifespan of fibroblasts in vitro has been shown to correlate directly with the potential of maximum life-span of the donor species [127] and the proliferative lifespan of normal human cells in culture is negatively correlated with the age of the donor [128-131].

The cultivation of fetal lung fibroblasts has widely been used as a cell culture model for to study cellular senescence and aging, as fibroblasts exhibit hallmarks of aging over time, and become senescent after a finite number of divisions in culture [132]. Recently, Ayyadevara et al have demonstrated an increase in endogenous protein insolubility in the heart tissue of older compared to younger mice. Furthermore they showed an increase of protein aggregation with passaging of heart fibroblasts from young mice direct to replicative senescence [18].

To establish a model for age-related protein aggregation in our laboratory, we therefore decided to extract insoluble proteins from lung and heart fibroblasts of young and aged mice. Furthermore, we analysed the protein aggregation in lung fibroblasts of young and aged mice harvested at early- and at late-passage stage.

2.2 Material and Methods

Organ extraction

All mice were bred and maintained by Chiara Valori at the German Center for Neurodegenerative Diseases, Tübingen. Mice used were non-transgenic males from the hFUS R495X-6 strain at different ages (2,1 months (Pretrial), 1 month, 18 months).

Organ extraction was performed by Chiara Valori at the German Center for Neurodegenerative Diseases, Tübingen. The extracted lungs and heart were directly put into PBS and used for cell extraction.

Cell extraction

Extraction of lung fibroblasts

Extraction of lung fibroblasts was performed following a modified version of a previously published protocol [133]. The freshly extracted lung tissue was transferred to a clean tissue culture dish under a tissue culture hood and cut into small pieces using sterile scalpels. Tissue fragments were thereafter digested by 0,13 Wunsch units/ml of Liberase TM Research Grade enzyme mix (Sigma-Aldrich) in 10 ml DMEM/F-12 media (life technologies) containing Penicillin/Streptomycin (Invitrogen), stirring slowly for 60 minutes at 37°C.

The solution was further disrupted by pipetting up and down and transferred into warm DMEM/F12 media containing 15% fetal bovine serum (FBS) (Invitrogen), 1x Penicillin/Streptomycin. The enzyme mix was removed by washing three times with warm DMEM/F12 media containing 15% FBS, 1x Penicillin/Streptomycin (centrifugation at 524 g for 1 minute). Afterwards tissue fragments were resuspended in warm DMEM/F12 media containing 15% FBS, 1x Penicillin/Streptomycin, transferred into a 75T tissue culture flask and placed into a tissue culture incubator at 37°C, 5% CO_2. After one day incubation, tissue fragments were transferred to a new flask with warm DMEM/F12 media containing 15% FBS, 1x Penicillin/Streptomycin and new medium was added to the old flask (1x wash with PBS). Once the flasks reach confluency, cells were trypsinated (0,25% Trypsin (Invitrogen), 2 minutes 37°C), tissue fragments were removed by sedimentation and the cells were split 1:2 into new T75 flasks using Minimum Essential Medium Eagle (EMEM) medium (Sigma-Aldrich) containing 15% FBS, non-essential amino acids (Invitrogen) and 1x Penicillin/Streptomycin for culturing.

Extraction of heart fibroblasts

For extraction of heart fibroblasts two different protocols were used:

Trypsin digestion protocol:

Extraction of heart fibroblasts with trypsin was performed following a modified version of a previously published protocol [18]. The freshly extracted heart tissue was

transferred to a clean tissue culture dish under a tissue culture hood and cut into small pieces using sterile scalpels. Tissue fragments were thereafter digested by 0,25% Trypsin at 37°C for 1 hour. The sedimented tissue fragments were again Trypsin-digested and this cycle repeated 5 times. The collected cells were pooled and resuspended and cultured in DMEM/F12 containing 5% FBS and 1x Penicillin/Streptomycin at 37°C, 5% CO_2.

Liberase TM digestion protocol:

For extraction of heart fibroblasts using the Liberase TM Research Grade enzyme mix, the tissue was treated as described for lung tissue in "Extraction of lung fibroblasts". Heart fibroblasts were extracted and cultivated in the same media as used for lung fibroblasts.

Culturing of cells

Cells were cultured in T75 flasks using EMEM medium containing 15% FBS, non-essential amino acids and 1x Penicillin/Streptomycin in a tissue culture incubator at 37°C, 5% CO_2. Once confluency was reached the cells were detached using 0,25% Trypsin (2 minutes 37°C), washed with medium and split 1:2 into new T75 flasks.

Freezing cells for protein extraction

For protein extraction cells were detached from confluent T75 flasks using 0,25% Trypsin and washed with medium. For counting the trypsinated cells were resuspended in medium and 10 µl of this suspension loaded onto an hemacytometer with coverslip. Cells on the grids were counted under the microscope (100x magnification) and cell number per millilitre medium was estimated as 10^4 times the average number of cells per grid. After washing the cells with warm PBS, the cells were centrifuged (centrifugation at 524 g for 3 minutes) and the supernatant removed. The resulting cell pellet was frozen directly in liquid nitrogen.

Protein extraction

For extraction of detergent-soluble and -insoluble proteins 150 µl RIPA buffer (50 mM Tris pH 8, 150 mM NaCl, 5 mM EDTA, 0,5% SDS, 0,5% SDO, 1% NP40 (Applichem), 1 mM PMSF, 1x Complete Protease Inhibitor Cocktail (Roche)) and 50µl RAB buffer (0,1 M MES, 1 mM EGTA, 0,1 mM EDTA, 0,5 mM $MgSO_4$, 0,75 M NaCl, 0,02 M NaF, 2x Complete Protease Inhibitor Cocktail (Roche)) were added to the frozen cell pellet. The cell pellet was solubilized on ice by vortexing and pipetting and the solution was homogenized using a 1ml syringe and needle (27 G x ½ ", 0.4 mm x 13 mm) by drawing the solution up and down ten times on ice. After a centrifugation step at 4°C (14000 rpm for 20 minutes) the supernatant was collected in aliquots (=soluble fraction) and frozen at -80°C, while the pellet was washed once with 100 µl RIPA buffer. The pellet containing the insoluble proteins was resolubilized in 75 µl Urea/SDS buffer (8 M Urea, 2% SDS, 50 mM DTT, 50 mM Tris pH 8) at room temperature. After incubation for ten minutes the sample was frozen in aliquots at -80°C (= insoluble fraction).

Gel electrophoresis and whole protein staining

Detergent-insoluble and total (= detergent insoluble fraction mixed 1:2 with soluble fraction) protein fractions were analysed using gel electrophoresis and whole protein staining. For sample preparation 4x NuPAGE LDS sample buffer (Thermo Fischer) and 10x sample reducing agent (Thermo Fischer) were added to the samples. The samples were heated 10 minutes at 70°C and thereafter centrifuged at 14000 rpm for 10 minutes at room temperature. As protein marker 5µl Novex Sharp Pre-Stained Protein Standard (Thermo Fisher) was used. The samples were loaded on a 4-12% gradient gel (NuPAGE 4-12% BisTris protein gels, Thermo Fisher) and run with NuPAGE MES SDS running buffer (Thermo Fisher). The gel was stained with Sypro Ruby protein gel stain following the manufacturer's instructions (Thermo Fisher) and detected with the Stella 3200 imaging system (Raytest) or used for western blotting.

For western blotting proteins were transferred from the SDS gel onto Immobilon-FL Polyvinylidene difluoride membrane (Merck Millipore) as following the manufacturer's instructions of the NuPAGE system (Thermo Fisher). Blotting was carried out with 35 mA overnight at 4°C. The blots were stained with Sypro Ruby protein blot stain following the

manufacturer's instructions (Thermo Fisher) and analysed using the Odyssey CLx imaging system (LiCor). For quantification of protein amounts the Fiji [134, 135] program was used.

2.3 Results

To be able to analyse modifiers of age-related protein aggregation found in *C. elegans* in a cell culture model, the aim of this project was to establish a model for inherent protein aggregation with age in cell culture. We chose to extract lung and heart fibroblasts from young and aged mice and compare the insoluble protein content in their early- and late-passage stages.

Cell extraction, passaging and collection of cells for protein extraction

Pre-Trial

In a pre-trial to set up the experiments, mouse lung fibroblasts were extracted from a 2,1 month old mouse using a previously published protocol [133]. The first confluent T75 flask after medium change to culturing medium, containing $23x10^5$ cells, was harvested for protein extraction and analysis of the total and insoluble protein content.

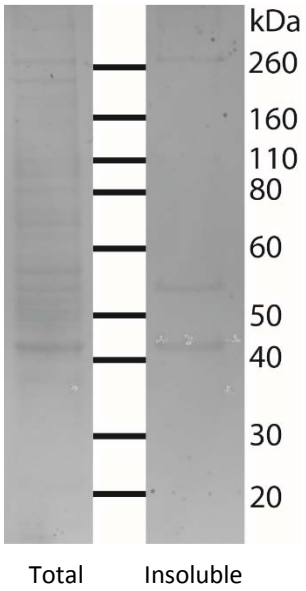


Figure 1: Whole protein staining of the total and insoluble fraction from lung fibroblasts extracted from a young mouse.

The whole protein staining shows the number of cells used for extraction was sufficient to analyse insoluble protein content and that protein extraction was successful.

After the successful pre-trial, fibroblasts were extracted in parallel from a young (1 month) and an aged (18 months) mouse.

Heart fibroblasts

For the extraction of heart fibroblasts two different protocols were used. In a first trial using the heart tissue of a young mouse, the previously published protocol of Ayyadevara et al [18] was replicated. However, no fibroblast culture grew out of the thus prepared tissue fragments.

Thereafter the protocol for the extraction of lung fibroblasts was also used for the heart tissue of the aged mouse. Using this protocol a live population of heart fibroblasts could be obtained and proliferated. After three population doublings cells could be harvested for protein extraction. However, after the third population doubling the cells stopped proliferating and no late-passage cell culture could be obtained.

Lung fibroblasts

For the extraction of lung fibroblasts a modified version of a previously published protocol was used (see methods) [133]. From both the extraction from the young and the aged mouse a growing culture of lung fibroblasts could be obtained. After three population doublings, a portion of the cells were harvested for protein extraction (early passage). The rest of the cultures were maintained to population doubling level (PDL) 7, at which time point the interdivision time of the fibroblasts extracted from the young mouse had more than doubled and cells were harvested for protein extraction (late passage). Cells kept from this culture after PDL 7 did not grow to a confluent PDL 8. Lung fibroblasts obtained from the aged mice were harvested at PDL 7 for comparison with those obtained from the young mouse. This culture grew to a confluent PDL 8 but stopped proliferating afterwards.

Cells harvested for protein extraction were counted before freezing. We find that although the same number of flasks were harvested per frozen sample, the number of

cells per sample vary greatly. Samples of lung fibroblasts with PDL 3, obtained from the young and the aged mouse, contained approximately factor 5 more cells than those with PDL 7 (see Table 1).

Table 1: Lung and mouse fibroblasts extracted from young (1 month) and aged (18 months) mice at PDL 3 or PDL 7

Tissue	**Mouse age**	**Population doubling level**	**Days in culture**	**Flasks harvested**	**Number of samples**	**Cells /sample**
Lung	1 month	PDL 3	13	15	3	$41{,}8x10^5$
		PDL 7	42	5	1	$7{,}4x10^5$
	18 months	PDL 3	11	15	3	$51{,}6x10^5$
		PDL 7	23	15	3	$9{,}8x10^5$
Heart	1 month	Did not grow				
	18 months	PDL 3 (Growth stopped at P3)	15	15	3	$4{,}7x10^5$

Analysis of insoluble protein content

To analyse the insoluble protein content of early- and late-passage fibroblasts from young and aged mice, we performed protein extraction from the harvested cells.

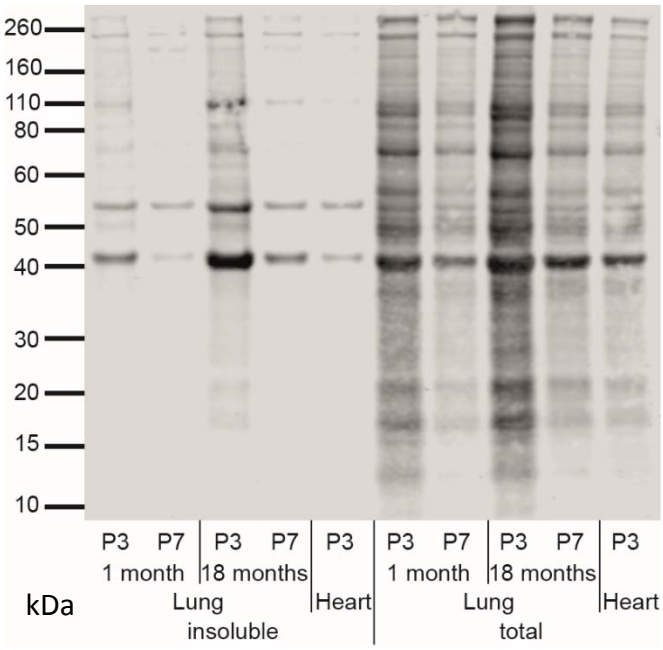


Figure 2: Whole protein staining of insoluble and total extracts of lung fibroblasts at PDL 3 (P3) and PDL 7 (P7) from young (1 month) and aged (18 months) mice and heart fibroblasts at PDL 3 from aged mice.

We used whole protein staining to determine protein amounts in the insoluble and total fractions of our samples. As we have varying amounts of cells per sample, the insoluble fraction had to be normalized to the total protein level per sample. Figure 3 shows the relative increase of insoluble content compared to that of lung fibroblasts at PDL 3 extracted from the young mouse.

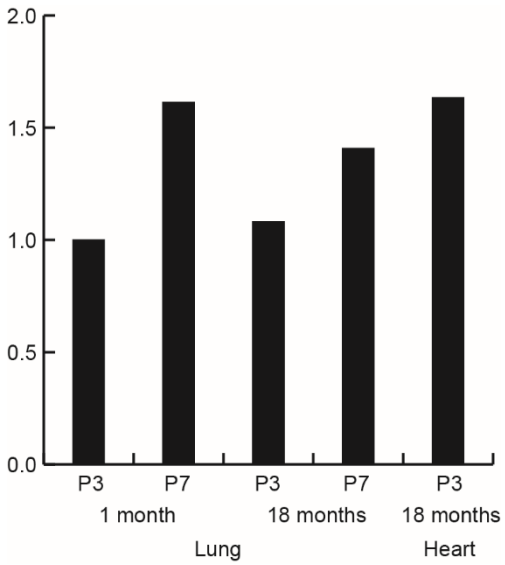


Figure 3: Quantification of insoluble protein content from whole protein stain Figure 2. Protein amounts of insoluble fraction were corrected for total protein amount and normalized to Lung, 1 month, P3.

While we do not find an increase of protein insolubility in fibroblasts extracted from aged mice compared to young mice, we see an increase of insoluble protein content from early- to late-passaged lung fibroblasts extracted from young as well as from aged mice. Heart fibroblasts extracted from aged mice already exhibit a relative high insoluble protein content compared to lung fibroblasts. In the absence of heart fibroblasts cultures obtained from young mice and late-passaged heart fibroblasts we cannot conclude whether this insoluble protein content would increase further with age of the donor mouse and longer passaging.

The increase of insoluble protein content of late- compared to early-passaged lung fibroblasts allows us to use this model to study age-related protein insolubility in cell culture.

2.4 Discussion

The goal of this project was to establish a cell culture model for age related protein insolubility. Our results show that aging of lung fibroblasts, extracted from young as well as aged mice, leads to a significant increase of insoluble protein content (Figure 3). However, it is important to consider that the in vitro culturing conditions may stress the

cells unnaturally. Keeping cells in culture requires them to continuously proliferate, which is not necessary in an organism, and cultured cells have different needs in regard to metabolic supplements, growth conditions and further factors [136-138]. Most importantly, the physiological levels of oxygen in most mammalian tissues are with 3% much lower than the standard culture conditions at approximately 20% [137]. This exposure to a hypertoxic environment is sufficient to induce growth arrest due to DNA damage by oxidation after 10-15 population doublings in mouse cells, which are sensitive to ambient oxygen [137]. Culturing normal human diploid cells in 2% oxygen has been shown to significantly extend their lifespan [139] while normal mouse cells can divide nearly indefinitely under physiologic oxygen conditions without reaching senescence [137]. This shows that in using standard conditions for our cell culture model we add high oxidative stress to the cellular aging process. Our findings therefor demonstrate that passaging of lung fibroblasts under stressful conditions leads to an increase in protein aggregation and can be used as a model to study age-related protein aggregation in cell culture.

Several issues concerning this model should be addressed before it can be used to study factors modifying age-related protein aggregation:

Previous publications propose that with increasing age of the donors, extracted cells need more time to double the cell number (population doubling time) and reach senescence after fewer population doublings [129]. However, this view has been challenged later on [140]. In our hands cells extracted from aged mice had an initial population doubling time of 3 days, while that of cells extracted from young mice was 2 days. Furthermore, the lung fibroblasts of young mice reached senescence after 42 days at PDL 7 and stopped proliferation. In contrast, lung fibroblasts from aged mice reached senescence after 30 days at PDL 8. Cell extractions from young and aged mice were not performed at the same time, allowing the possibility of different experimental conditions to influence the experiment. Levels of familiarity with the protocol can influence incubation times and cell handling, therefore possibly stressing the cells to a different extend. We can speculate that the cells extracted from young mice were stressed more during the extraction process as this experiment was performed first, leading slower growth, thereby limiting the PDL before senescence. To be able to

compare insoluble protein content of cells extracted from young and aged mice cell extractions and culture should be performed in parallel in future. However, as the population doubling time of cells from young and aged animals should not be equal, it will most likely not be possible to perform the following steps of harvesting and freezing cells at the same time point.

Another issue that should be addressed is the strongly varying number of cells per harvested confluent flask. Our results show that by harvesting the same number of confluent flasks of different PDLs, the number of obtained cells is approximately factor 5 higher at PDL 3 than at PDL 7. This can easily be explained by the increase of cell size with PDL. Therefore a smaller number of cells is needed to cover the ground of the flask completely. In future experiments the number of flasks harvested for the higher PDL should be increased, allowing the collection of an approximately equal number of cells per sample. Cells should be counted and thereafter distributed at a defined cell number in the different samples. Use of the same cell number for protein extraction would allow an easier comparison of insoluble protein content. However, as our results were corrected for total protein amount per sample, the influence of different cell numbers on relative insoluble protein amount should be negligible.

To reduce the influence of variability between different individual mice, lung fibroblasts from several same-aged mice should be harvested in one experiment. Combining the cut lung tissue from several mice during the extraction does thereafter not increase the work load of the experiment while helping to remove the effect of variability on the insoluble protein content.

Our results indicate that aging lung fibroblasts from aged or from young mice under stressful conditions by passaging to senescence at 20% oxygen can be used as a model for age-related protein insolubility. This model can hereafter be used to analyse the impact of modifiers of protein aggregation found in *C. elegans* on mammalian cells. Using specific antibodies the aggregation behaviour of proteins used in our *C. elegans* study can be tested in cell culture. Knockdown of proteins found to be relevant for age-related protein aggregation in *C. elegans* can be re-evaluated in this model using shRNA. The impact of these knockdowns on total protein insolubility or on specific proteins

using antibody staining could be determined. These experiments could give valuable insights on the transferability of our finding into a mammalian model.

The test of modifiers on protein aggregation on several model systems is a very important step to verify the conserved role of these modifiers. Proving their impact on mammalian cells in cell culture leads a step closer to determining their role in protein aggregation with age in human and ultimately in human diseases.

3 References

1. Flatt, T., *A new definition of aging?* Front Genet, 2012. **3**: p. 148.
2. Lopez-Otin, C., et al., *The hallmarks of aging.* Cell, 2013. **153**(6): p. 1194-217.
3. Kenyon, C.J., *The genetics of ageing.* Nature, 2010. **464**(7288): p. 504-12.
4. Trojanowski, J.Q., *PENN neurodegenerative disease research - in the spirit of Benjamin Franklin.* Neurosignals, 2008. **16**(1): p. 5-10.
5. Niccoli, T. and L. Partridge, *Ageing as a risk factor for disease.* Curr Biol, 2012. **22**(17): p. R741-52.
6. Fink, A.L., *Protein aggregation: folding aggregates, inclusion bodies and amyloid.* Fold Des, 1998. **3**(1): p. R9-23.
7. Aguzzi, A. and T. O'Connor, *Protein aggregation diseases: pathogenicity and therapeutic perspectives.* Nat Rev Drug Discov, 2010. **9**(3): p. 237-48.
8. Eisenberg, D. and M. Jucker, *The amyloid state of proteins in human diseases.* Cell, 2012. **148**(6): p. 1188-203.
9. Sipe, J.D., et al., *Amyloid fibril protein nomenclature: 2010 recommendations from the nomenclature committee of the International Society of Amyloidosis.* Amyloid, 2010. **17**(3-4): p. 101-4.
10. Verma, M., A. Vats, and V. Taneja, *Toxic species in amyloid disorders: Oligomers or mature fibrils.* Ann Indian Acad Neurol, 2015. **18**(2): p. 138-45.
11. Sengupta, U., et al., *Pathological interface between oligomeric alpha-synuclein and tau in synucleinopathies.* Biol Psychiatry, 2015. **78**(10): p. 672-83.
12. Sengupta, U., A.N. Nilson, and R. Kayed, *The Role of Amyloid-beta Oligomers in Toxicity, Propagation, and Immunotherapy.* EBioMedicine, 2016. **6**: p. 42-49.
13. David, D.C., et al., *Widespread protein aggregation as an inherent part of aging in C. elegans.* PLoS Biol, 2010. **8**(8): p. e1000450.
14. Lechler, M.C., et al., *Reduced Insulin/IGF-1 Signaling Restores the Dynamic Properties of Key Stress Granule Proteins during Aging.* Cell Rep, 2017. **18**(2): p. 454-467.
15. Demontis, F. and N. Perrimon, *FOXO/4E-BP signaling in Drosophila muscles regulates organism-wide proteostasis during aging.* Cell, 2010. **143**(5): p. 813-25.
16. Reis-Rodrigues, P., et al., *Proteomic analysis of age-dependent changes in protein solubility identifies genes that modulate lifespan.* Aging Cell, 2012. **11**(1): p. 120-7.
17. Walther, D.M., et al., *Widespread Proteome Remodeling and Aggregation in Aging C. elegans.* Cell, 2015. **161**(4): p. 919-32.
18. Ayyadevara, S., et al., *Age- and Hypertension-Associated Protein Aggregates in Mouse Heart Have Similar Proteomic Profiles.* Hypertension, 2016. **67**(5): p. 1006-13.
19. Tanase, M., et al., *Role of Carbonyl Modifications on Aging-Associated Protein Aggregation.* Sci Rep, 2016. **6**: p. 19311.
20. Peters, T.W., et al., *Tor1 regulates protein solubility in Saccharomyces cerevisiae.* Mol Biol Cell, 2012. **23**(24): p. 4679-88.
21. Ottis, P., et al., *Aging-induced proteostatic changes in the rat hippocampus identify ARP3, NEB2 and BRAG2 as a molecular circuitry for cognitive impairment.* PLoS One, 2013. **8**(9): p. e75112.
22. Garigan, D., et al., *Genetic analysis of tissue aging in Caenorhabditis elegans: a role for heat-shock factor and bacterial proliferation.* Genetics, 2002. **161**(3): p. 1101-12.
23. Markaki, M. and N. Tavernarakis, *Modeling human diseases in Caenorhabditis elegans.* Biotechnol J, 2010. **5**(12): p. 1261-76.

24. Kikis, E.A., T. Gidalevitz, and R.I. Morimoto, *Protein homeostasis in models of aging and age-related conformational disease.* Adv Exp Med Biol, 2010. **694**: p. 138-59.
25. Guisbert, E., et al., *Identification of a tissue-selective heat shock response regulatory network.* PLoS Genet, 2013. **9**(4): p. e1003466.
26. Ma, J., et al., *Cellular Proteomes Drive Tissue-Specific Regulation of the Heat Shock Response.* G3 (Bethesda), 2017. **7**(3): p. 1011-1018.
27. Jackson, W.S., *Selective vulnerability to neurodegenerative disease: the curious case of Prion Protein.* Dis Model Mech, 2014. **7**(1): p. 21-9.
28. Jucker, M. and L.C. Walker, *Self-propagation of pathogenic protein aggregates in neurodegenerative diseases.* Nature, 2013. **501**(7465): p. 45-51.
29. Morales, R., I. Moreno-Gonzalez, and C. Soto, *Cross-seeding of misfolded proteins: implications for etiology and pathogenesis of protein misfolding diseases.* PLoS Pathog, 2013. **9**(9): p. e1003537.
30. Arai, T., et al., *TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis.* Biochem Biophys Res Commun, 2006. **351**(3): p. 602-11.
31. Neumann, M., et al., *Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis.* Science, 2006. **314**(5796): p. 130-3.
32. Kwiatkowski, T.J., Jr., et al., *Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis.* Science, 2009. **323**(5918): p. 1205-8.
33. Vance, C., et al., *Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6.* Science, 2009. **323**(5918): p. 1208-1211.
34. Neumann, M., et al., *FET proteins TAF15 and EWS are selective markers that distinguish FTLD with FUS pathology from amyotrophic lateral sclerosis with FUS mutations.* Brain, 2011. **134**(Pt 9): p. 2595-609.
35. Kim, H.J., et al., *Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS.* Nature, 2013. **495**(7442): p. 467-73.
36. Molliex, A., et al., *Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization.* Cell, 2015. **163**(1): p. 123-33.
37. Murakami, T., et al., *ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function.* Neuron, 2015. **88**(4): p. 678-90.
38. Patel, A., et al., *A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation.* Cell, 2015. **162**(5): p. 1066-77.
39. Alberti, S., et al., *A systematic survey identifies prions and illuminates sequence features of prionogenic proteins.* Cell, 2009. **137**(1): p. 146-58.
40. King, O.D., A.D. Gitler, and J. Shorter, *The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease.* Brain Res, 2012. **1462**: p. 61-80.
41. Malinovska, L., S. Kroschwald, and S. Alberti, *Protein disorder, prion propensities, and self-organizing macromolecular collectives.* Biochim Biophys Acta, 2013. **1834**(5): p. 918-31.
42. Brangwynne, C.P., et al., *Germline P granules are liquid droplets that localize by controlled dissolution/condensation.* Science, 2009. **324**(5935): p. 1729-32.
43. Li, P., et al., *Phase transitions in the assembly of multivalent signalling proteins.* Nature, 2012. **483**(7389): p. 336-40.
44. Hyman, A.A., C.A. Weber, and F. Julicher, *Liquid-liquid phase separation in biology.* Annu Rev Cell Dev Biol, 2014. **30**: p. 39-58.
45. Lindquist, S., *Regulation of protein synthesis during heat shock.* Nature, 1981. **293**(5830): p. 311-4.
46. Anderson, P. and N. Kedersha, *Stressful initiations.* J Cell Sci, 2002. **115**(Pt 16): p. 3227-34.

47. Alberti, S., et al., *Granulostasis: Protein Quality Control of RNP Granules.* Front Mol Neurosci, 2017. **10**: p. 84.
48. Kedersha, N., et al., *Stress granules and processing bodies are dynamically linked sites of mRNP remodeling.* J Cell Biol, 2005. **169**(6): p. 871-84.
49. Kedersha, N.L., et al., *RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules.* J Cell Biol, 1999. **147**(7): p. 1431-42.
50. Bentmann, E., C. Haass, and D. Dormann, *Stress granules in neurodegeneration--lessons learnt from TAR DNA binding protein of 43 kDa and fused in sarcoma.* FEBS J, 2013. **280**(18): p. 4348-70.
51. Gallouzi, I.E., et al., *A novel phosphorylation-dependent RNase activity of GAP-SH3 binding protein: a potential link between signal transduction and RNA stability.* Mol Cell Biol, 1998. **18**(7): p. 3956-65.
52. Tourriere, H., et al., *RasGAP-associated endoribonuclease G3Bp: selective RNA degradation and phosphorylation-dependent localization.* Mol Cell Biol, 2001. **21**(22): p. 7747-60.
53. Tourriere, H., et al., *The RasGAP-associated endoribonuclease G3BP assembles stress granules.* J Cell Biol, 2003. **160**(6): p. 823-31.
54. Schmidlin, M., et al., *The ARE-dependent mRNA-destabilizing activity of BRF1 is regulated by protein kinase B.* EMBO J, 2004. **23**(24): p. 4760-9.
55. Stoecklin, G., et al., *MK2-induced tristetraprolin:14-3-3 complexes prevent stress granule association and ARE-mRNA decay.* EMBO J, 2004. **23**(6): p. 1313-24.
56. De Leeuw, F., et al., *The cold-inducible RNA-binding protein migrates from the nucleus to cytoplasmic stress granules by a methylation-dependent mechanism and acts as a translational repressor.* Exp Cell Res, 2007. **313**(20): p. 4130-44.
57. Kwon, S., Y. Zhang, and P. Matthias, *The deacetylase HDAC6 is a novel critical component of stress granules involved in the stress response.* Genes Dev, 2007. **21**(24): p. 3381-94.
58. Goulet, I., et al., *TDRD3, a novel Tudor domain-containing protein, localizes to cytoplasmic stress granules.* Hum Mol Genet, 2008. **17**(19): p. 3055-74.
59. Ohn, T., et al., *A functional RNAi screen links O-GlcNAc modification of ribosomal proteins to stress granule and processing body assembly.* Nat Cell Biol, 2008. **10**(10): p. 1224-31.
60. Ganassi, M., et al., *A Surveillance Function of the HSPB8-BAG3-HSP70 Chaperone Complex Ensures Stress Granule Integrity and Dynamism.* Mol Cell, 2016. **63**(5): p. 796-810.
61. Mateju, D., et al., *An aberrant phase transition of stress granules triggered by misfolded protein and prevented by chaperone function.* EMBO J, 2017. **36**(12): p. 1669-1687.
62. Hackman, P., et al., *Welander distal myopathy is caused by a mutation in the RNA-binding protein TIA1.* Ann Neurol, 2013. **73**(4): p. 500-9.
63. Klar, J., et al., *Welander distal myopathy caused by an ancient founder mutation in TIA1 associated with perturbed splicing.* Hum Mutat, 2013. **34**(4): p. 572-7.
64. Lin, Y., et al., *Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins.* Mol Cell, 2015. **60**(2): p. 208-19.
65. Poewe, W., et al., *Parkinson disease.* Nat Rev Dis Primers, 2017. **3**: p. 17013.
66. Postuma, R.B., et al., *MDS clinical diagnostic criteria for Parkinson's disease.* Mov Disord, 2015. **30**(12): p. 1591-601.
67. Tolosa, E., G. Wenning, and W. Poewe, *The diagnosis of Parkinson's disease.* Lancet Neurol, 2006. **5**(1): p. 75-86.
68. Bostantjopoulou, S., et al., *Evaluation of non-motor symptoms in Parkinson's Disease: An underestimated necessity.* Hippokratia, 2013. **17**(3): p. 214-9.

69. Kalinderi, K., S. Bostantjopoulou, and L. Fidani, *The genetic background of Parkinson's disease: current progress and future prospects.* Acta Neurol Scand, 2016. **134**(5): p. 314-326.
70. Rousseaux, M.W.C., J.M. Shulman, and J. Jankovic, *Progress toward an integrated understanding of Parkinson's disease.* F1000Res, 2017. **6**: p. 1121.
71. Neef, D.W., et al., *A direct regulatory interaction between chaperonin TRiC and stress-responsive transcription factor HSF1.* Cell Rep, 2014. **9**(3): p. 955-66.
72. Vabulas, R.M. and F.U. Hartl, *Protein synthesis upon acute nutrient restriction relies on proteasome function.* Science, 2005. **310**(5756): p. 1960-3.
73. Schubert, U., et al., *Rapid degradation of a large fraction of newly synthesized proteins by proteasomes.* Nature, 2000. **404**(6779): p. 770-4.
74. Cardinaud, S., et al., *The synthesis of truncated polypeptides for immune surveillance and viral evasion.* PLoS One, 2010. **5**(1): p. e8692.
75. Xu, G., et al., *Vulnerability of newly synthesized proteins to proteostasis stress.* J Cell Sci, 2016. **129**(9): p. 1892-901.
76. Li, Y.R., et al., *Stress granules as crucibles of ALS pathogenesis.* J Cell Biol, 2013. **201**(3): p. 361-72.
77. Mori, K., et al., *hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with C9orf72 mutations.* Acta Neuropathol, 2013. **125**(3): p. 413-23.
78. Bentmann, E., et al., *Requirements for stress granule recruitment of fused in sarcoma (FUS) and TAR DNA-binding protein of 43 kDa (TDP-43).* J Biol Chem, 2012. **287**(27): p. 23079-94.
79. Brais, B., et al., *Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy.* Nat Genet, 1998. **18**(2): p. 164-7.
80. Kroschwald, S., et al., *Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules.* Elife, 2015. **4**: p. e06807.
81. Cherkasov, V., et al., *Coordination of translational control and protein homeostasis during severe heat stress.* Curr Biol, 2013. **23**(24): p. 2452-62.
82. Kato, M., et al., *Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels.* Cell, 2012. **149**(4): p. 753-67.
83. Ben-Zvi, A., E.A. Miller, and R.I. Morimoto, *Collapse of proteostasis represents an early molecular event in Caenorhabditis elegans aging.* Proc Natl Acad Sci U S A, 2009. **106**(35): p. 14914-9.
84. Brown, M.K. and N. Naidoo, *The endoplasmic reticulum stress response in aging and age-related diseases.* Front Physiol, 2012. **3**: p. 263.
85. Anselmi, B., et al., *Dietary self-selection can compensate an age-related decrease of rat liver 20 S proteasome activity observed with standard diet.* J Gerontol A Biol Sci Med Sci, 1998. **53**(3): p. B173-9.
86. Keller, J.N., K.B. Hanni, and W.R. Markesbery, *Possible involvement of proteasome inhibition in aging: implications for oxidative stress.* Mech Ageing Dev, 2000. **113**(1): p. 61-70.
87. Cuervo, A.M. and J.F. Dice, *Age-related decline in chaperone-mediated autophagy.* J Biol Chem, 2000. **275**(40): p. 31505-13.
88. Gilks, N., et al., *Stress granule assembly is mediated by prion-like aggregation of TIA-1.* Mol Biol Cell, 2004. **15**(12): p. 5383-98.
89. Walters, R.W. and R. Parker, *Coupling of Ribostasis and Proteostasis: Hsp70 Proteins in mRNA Metabolism.* Trends Biochem Sci, 2015. **40**(10): p. 552-9.
90. Walters, R.W., et al., *Differential effects of Ydj1 and Sis1 on Hsp70-mediated clearance of stress granules in Saccharomyces cerevisiae.* RNA, 2015. **21**(9): p. 1660-71.

91. Collier, N.C. and M.J. Schlesinger, *The dynamic state of heat shock proteins in chicken embryo fibroblasts.* J Cell Biol, 1986. **103**(4): p. 1495-507.
92. Scharf, K.D., et al., *The tomato Hsf system: HsfA2 needs interaction with HsfA1 for efficient nuclear import and may be localized in cytoplasmic heat stress granules.* Mol Cell Biol, 1998. **18**(4): p. 2240-51.
93. Volovik, Y., et al., *Temporal requirements of heat shock factor-1 for longevity assurance.* Aging Cell, 2012. **11**(3): p. 491-9.
94. Evgrafov, O.V., et al., *Mutant small heat-shock protein 27 causes axonal Charcot-Marie-Tooth disease and distal hereditary motor neuropathy.* Nat Genet, 2004. **36**(6): p. 602-6.
95. Irobi, J., et al., *Hot-spot residue in small heat-shock protein 22 causes distal motor neuropathy.* Nat Genet, 2004. **36**(6): p. 597-601.
96. Ghaoui, R., et al., *Mutations in HSPB8 causing a new phenotype of distal myopathy and motor neuropathy.* Neurology, 2016. **86**(4): p. 391-8.
97. Selcen, D., et al., *Mutation in BAG3 causes severe dominant childhood muscular dystrophy.* Ann Neurol, 2009. **65**(1): p. 83-9.
98. Johnson, J.O., et al., *Exome sequencing reveals VCP mutations as a cause of familial ALS.* Neuron, 2010. **68**(5): p. 857-64.
99. Krebs, C.E. and C. Paisan-Ruiz, *The use of next-generation sequencing in movement disorders.* Front Genet, 2012. **3**: p. 75.
100. Wang, Z., et al., *The role and challenges of exome sequencing in studies of human diseases.* Front Genet, 2013. **4**: p. 160.
101. Choi, M., et al., *Genetic diagnosis by whole exome capture and massively parallel DNA sequencing.* Proc Natl Acad Sci U S A, 2009. **106**(45): p. 19096-101.
102. Bamshad, M.J., et al., *Exome sequencing as a tool for Mendelian disease gene discovery.* Nat Rev Genet, 2011. **12**(11): p. 745-55.
103. Farlow, J.L., et al., *Whole-Exome Sequencing in Familial Parkinson Disease.* JAMA Neurol, 2016. **73**(1): p. 68-75.
104. Kitada, T., et al., *Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism.* Nature, 1998. **392**(6676): p. 605-8.
105. Valente, E.M., et al., *Hereditary early-onset Parkinson's disease caused by mutations in PINK1.* Science, 2004. **304**(5674): p. 1158-60.
106. Bonifati, V., et al., *DJ-1(PARK7), a novel gene for autosomal recessive, early onset parkinsonism.* Neurol Sci, 2003. **24**(3): p. 159-60.
107. Sulem, P., et al., *Identification of a large set of rare complete human knockouts.* Nat Genet, 2015. **47**(5): p. 448-52.
108. Cookson, M.R., *Parkinsonism due to mutations in PINK1, parkin, and DJ-1 and oxidative stress and mitochondrial pathways.* Cold Spring Harb Perspect Med, 2012. **2**(9): p. a009415.
109. Kamp, F., et al., *Inhibition of mitochondrial fusion by alpha-synuclein is rescued by PINK1, Parkin and DJ-1.* EMBO J, 2010. **29**(20): p. 3571-89.
110. Koopman, W.J., et al., *Simultaneous quantitative measurement and automated analysis of mitochondrial morphology, mass, potential, and motility in living human skin fibroblasts.* Cytometry A, 2006. **69**(1): p. 1-12.
111. Narendra, D., et al., *Parkin is recruited selectively to impaired mitochondria and promotes their autophagy.* J Cell Biol, 2008. **183**(5): p. 795-803.
112. Narendra, D.P., et al., *PINK1 is selectively stabilized on impaired mitochondria to activate Parkin.* PLoS Biol, 2010. **8**(1): p. e1000298.
113. Vives-Bauza, C., et al., *PINK1-dependent recruitment of Parkin to mitochondria in mitophagy.* Proc Natl Acad Sci U S A, 2010. **107**(1): p. 378-83.
114. Geisler, S., et al., *PINK1/Parkin-mediated mitophagy is dependent on VDAC1 and p62/SQSTM1.* Nat Cell Biol, 2010. **12**(2): p. 119-31.

115. Vincow, E.S., et al., *The PINK1-Parkin pathway promotes both mitophagy and selective respiratory chain turnover in vivo.* Proc Natl Acad Sci U S A, 2013. **110**(16): p. 6400-5.
116. Feany, M.B. and W.W. Bender, *A Drosophila model of Parkinson's disease.* Nature, 2000. **404**(6776): p. 394-8.
117. Chen, L. and M.B. Feany, *Alpha-synuclein phosphorylation controls neurotoxicity and inclusion formation in a Drosophila model of Parkinson disease.* Nat Neurosci, 2005. **8**(5): p. 657-63.
118. Cullen, V., et al., *Cathepsin D expression level affects alpha-synuclein processing, aggregation, and toxicity in vivo.* Mol Brain, 2009. **2**: p. 5.
119. Lesage, S., et al., *Loss of VPS13C Function in Autosomal-Recessive Parkinsonism Causes Mitochondrial Dysfunction and Increases PINK1/Parkin-Dependent Mitophagy.* Am J Hum Genet, 2016. **98**(3): p. 500-13.
120. Balch, W.E., et al., *Adapting proteostasis for disease intervention.* Science, 2008. **319**(5865): p. 916-9.
121. Hartl, F.U., A. Bracher, and M. Hayer-Hartl, *Molecular chaperones in protein folding and proteostasis.* Nature, 2011. **475**(7356): p. 324-32.
122. Kettern, N., et al., *Chaperone-assisted degradation: multiple paths to destruction.* Biol Chem, 2010. **391**(5): p. 481-9.
123. Hayflick, L. and P.S. Moorhead, *The serial cultivation of human diploid cell strains.* Exp Cell Res, 1961. **25**: p. 585-621.
124. Swim, H.E. and R.F. Parker, *Culture characteristics of human fibroblasts propagated serially.* Am J Hyg, 1957. **66**(2): p. 235-43.
125. Cristofalo, V.J. and R.J. Pignolo, *Replicative senescence of human fibroblast-like cells in culture.* Physiol Rev, 1993. **73**(3): p. 617-38.
126. Phipps, S.M., et al., *Aging cell culture: methods and observations.* Methods Mol Biol, 2007. **371**: p. 9-19.
127. Rohme, D., *Evidence for a relationship between longevity of mammalian species and life spans of normal fibroblasts in vitro and erythrocytes in vivo.* Proc Natl Acad Sci U S A, 1981. **78**(8): p. 5009-13.
128. Bierman, E.L., *The effect of donor age on the in vitro life span of cultured human arterial smooth-muscle cells.* In Vitro, 1978. **14**(11): p. 951-5.
129. Schneider, E.L. and Y. Mitsui, *The relationship between in vitro cellular aging and in vivo human age.* Proc Natl Acad Sci U S A, 1976. **73**(10): p. 3584-8.
130. Martin, G.M., C.A. Sprague, and C.J. Epstein, *Replicative life-span of cultivated human cells. Effects of donor's age, tissue, and genotype.* Lab Invest, 1970. **23**(1): p. 86-92.
131. Cristofalo, V.J., C. Volker, and R.G. Allen, *Use of the fibroblast model in the study of cellular senescence.* Methods Mol Med, 2000. **38**: p. 23-52.
132. Chen, H., Y. Li, and T.O. Tollefsbol, *Cell senescence culturing methods.* Methods Mol Biol, 2013. **1048**: p. 1-10.
133. Seluanov, A., A. Vaidya, and V. Gorbunova, *Establishing primary adult fibroblast cultures from rodents.* J Vis Exp, 2010(44).
134. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis.* Nat Methods, 2012. **9**(7): p. 671-5.
135. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nat Methods, 2012. **9**(7): p. 676-82.
136. Wright, W.E. and J.W. Shay, *Telomere dynamics in cancer progression and prevention: fundamental differences in human and mouse telomere biology.* Nat Med, 2000. **6**(8): p. 849-51.
137. Parrinello, S., et al., *Oxygen sensitivity severely limits the replicative lifespan of murine fibroblasts.* Nat Cell Biol, 2003. **5**(8): p. 741-7.

138. Wright, W.E. and J.W. Shay, *Inexpensive low-oxygen incubators.* Nat Protoc, 2006. **1**(4): p. 2088-90.
139. Packer, L. and K. Fuehr, *Low oxygen concentration extends the lifespan of cultured human diploid cells.* Nature, 1977. **267**(5610): p. 423-5.
140. Cristofalo, V.J., et al., *Relationship between donor age and the replicative lifespan of human cells in culture: a reevaluation.* Proc Natl Acad Sci U S A, 1998. **95**(18): p. 10614-9.

4 Publications

4.1 List of publications/manuscripts appended

- Lechler et al, in preparation

  "Tissue-specific reduction of aggregation-prone protein levels following disruption of the quality control network as regulator of age-related protein aggregation in C. elegans"

  Marie C. Lechler, Raimund Jung, Della C. David

- Lechler et al., Cell Reports, 2017

  "Reduced Insulin/IGF-1 Signalling Restores the Dynamic Properties of Key Stress Granule Proteins during Aging"

  Marie C. Lechler, Emily D. Crawford, Nicole Groh, Katja Widmaier, Raimund Jung, Janine Kirstein, Jonathan C. Trinidad, Alma L. Burlingame, and Della C. David

- Lechler et al., Prion, 2017

  "More stressed out with age? Check your RNA granule aggregation"

  Marie C. Lechler and Della C. David

- Jansen et al., Genome Biology, 2017

  "Discovery and functional prioritization of Parkinson's disease candidate genes from large-scale whole exome sequencing"

  Iris E. Jansen, Hui Ye, Sasja Heetveld, Marie C. Lechler, Helen Michels, Renée I. Seinstra, Steven J. Lubbe, Valérie Drouet, Suzanne Lesage, Elisa Majounie, J. Raphael Gibbs, Mike A. Nalls, Mina Ryten, Juan A. Botia, Jana Vandrovcova, Javier Simon-Sanchez, Melissa Castillo-Lizardo, Patrizia Rizzu, Cornelis Blauwendraat, Amit K. Chouhan, Yarong Li, Puja Yogi, Najaf Amin, Cornelia M. van Duijn, International Parkinson's Disease Genetics Consortium (IPGDC), Huw R. Morris, Alexis Brice8,16, Andrew B. Singleton, Della C. David, Ellen A. Nollen, Shushant Jain, Joshua M. Shulman and Peter Heutink

4.2 Statement of contributions

- *„Tissue-specific reduction of aggregation-prone protein levels following disruption of the quality control network as regulator of age-related protein aggregation in C. elegans“*

  Marie C. Lechler, Raimund Jung, Della C. David

  MCL, JR, and DCD designed and performed experiments. MCL performed all RHO-1 aggregation counting and imaging experiments and the fluorescence quantification of mEOS2-tag. MCL and DCD performed data analysis and wrote the paper.

- "*Reduced Insulin/IGF-1 Signalling Restores the Dynamic Properties of Key Stress Granule Proteins during Aging*"

  Marie C. Lechler, Emily D. Crawford, Nicole Groh, Katja Widmaier, Raimund Jung, Janine Kirstein, Jonathan C. Trinidad, Alma L. Burlingame, and Della C. David

  MCL, NG, JCT, and DCD designed and performed experiments. MCL performed all PAB-1 and TIAR-2 aggregation counting experiments, confocal imaging, immunostainings (except DNJ-13, DNJ-19), size measurements and lifespan analysis. EDC performed bioinformatics analysis. KW, RJ, and JK generated reagents. JK performed DNJ-13 and DNJ-19 immunostainings. ALB provided analytical tools, reagents, and materials. MCL, JCT, and DCD wrote the paper.

- "*More stressed out with age? Check your RNA granule aggregation"*

  Marie C. Lechler and Della C. David

  MCL and DCD designed and performed experiments. MCL performed all PAB-1 aggregation counting experiments and confocal imaging.

  MCL and DCD wrote the paper.

- *"Discovery and functional prioritization of Parkinson's disease candidate genes from large-scale whole exome sequencing"*

  Iris E. Jansen, Hui Ye, Sasja Heetveld, Marie C. Lechler, Helen Michels, Renée I. Seinstra, Steven J. Lubbe, Valérie Drouet, Suzanne Lesage,

Elisa Majounie, J. Raphael Gibbs, Mike A. Nalls, Mina Ryten, Juan A. Botia, Jana Vandrovcova, Javier Simon-Sanchez, Melissa Castillo-Lizardo, Patrizia Rizzu, Cornelis Blauwendraat, Amit K. Chouhan, Yarong Li, Puja Yogi, Najaf Amin, Cornelia M. van Duijn, International Parkinson's Disease Genetics Consortium (IPGDC), Huw R. Morris, Alexis Brice8,16, Andrew B. Singleton, Della C. David, Ellen A. Nollen, Shushant Jain, Joshua M. Shulman and Peter Heutink

Conception/design: IEJ, HY, AKC, JMS, PH;

Data analysis: IEJ, HY, SH, MCL, EM, JRG, MAN, MR, JAB, JV, JSS, MCL, PR, CB, SJ;

Data acquisition: IEJ, HY, SH, MCL, HM, RIS, SJL, VD, SL, EM, JRG, MAN, MR, JAB, JV, JSS, MCL, PR, AKC, YL, PY, IPGDC, HRM, AB, ABS, DCD, EAN, SJ;

Data interpretation: IEJ, HY, SH, MCL, SJ, JMS, PH;

Writing manuscript: IEJ, HY, SH, MCL, JAB, SJ, JMS, PH;

Revising manuscript: HM, RIS, SJL, VD, SL, EM, JRG, MAN, MR, JAB, JV, JSS, MCL, PR, CB, AKC, YL, PY, IPGDC, HRM, AB, ABS, DCD, EAN;

MCL performed high throughput knockdown screens for survival, developmental effects, fertility and influence on polyglutamine aggregation with 66 candidate genes in *C. elegans*. MCL participated in planning of experiments, data analysis and writing of the manuscript.

5 Appendix

5.1 Abbreviations

Aβ: Amyloid-β

AD: Alzheimer's Disease

ALS: amyotrophic lateral sclerosis

C. elegans: *Caenorhabditis elegans*

CK1alpha: casein kinase isoform alpha

DRIPs: defective ribosomal products

EMEM: Minimum Essential Medium Eagle

FBS: fetal bovine serum

FTLD: frontotemporal lobar degeneration

GWAS: genome-wide association study

HD: Huntington's disease

HSP: heat shock protein

HSR: heat shock response

LC: low complexity

LLPS: liquid-liquid phase separation

LoF: loss of function

PAB-1: poly(A)-binding protein

PD: Parkinson's disease

PDL: population doubling level

RBP: RNA-binding protein

RRM: RNA-recognition motive

SG: stress granules

sgRBP: stress granule RNA-binding protein

TIA-1: T-cell-restricted intracellular antigen-1

TIAR: TIA-1-related

UPR: unfolded protein response

WES: whole exome sequencing

5.2 Acknowledgements

With this I want to sincerely thank my supervisor Dr. Della David for her continuous encouragement and support during my PhD thesis. I also want to thank Prof. Heutink and Prof. Sommer for their helpful feedback and advice as members of my advisory board.

To my fellow labmates I am very grateful for many stimulating discussions and for creating a productive and fun work environment over these past years.

Last but not least, I would like to express my gratitude to my family: to my parents and to my brothers and sisters for all the loving encouragement and motivation during my studies and my PhD work.

5.3 Appended publications/manuscripts

- Lechler et al, in preparation

  "Tissue-specific reduction of aggregation-prone protein levels following disruption of the quality control network as regulator of age-related protein aggregation in C. elegans"

- Lechler et al., Cell Reports, 2017

  "Reduced Insulin/IGF-1 Signalling Restores the Dynamic Properties of Key Stress Granule Proteins during Aging"

- Lechler et al., Prion, 2017

  "More stressed out with age? Check your RNA granule aggregation"

- Jansen et al., Genome Biology, 2017

  "Discovery and functional prioritization of Parkinson's disease candidate genes from large-scale whole exome sequencing"

Tissue-specific reduction of aggregation-prone protein levels following disruption of the quality control network as regulator of age-related protein aggregation in *C. elegans*

Marie C. Lechler[1, 2,]*, Raimund Jung[1,]*, Della C. David[1]

[1] German Center for Neurodegenerative Diseases, 72076 Tübingen, Germany
[2] Graduate Training Centre of Neuroscience, 72074 Tübingen, Germany
*equally contributing

Abstract

The role of the complex quality control network (PQC) is to maintain protein homeostasis in resting and stress conditions, as the functionality of the proteome is vital for organismal health. However, as the organism ages, a decline of the PQC and general increase in protein insolubility have been observed. As the proteome varies between different cell types, tissue-specific regulation by the PQC is necessary. To this point, little is known about the regulation of age-related protein aggregation in different tissues.

Here, we demonstrate tissue-specificity in the regulation of age-related protein aggregation in *C. elegans*. Impairment of the PQC either by inhibition of proteasomal or autophagic degradation or by knockdown of the chaperone system, affects age-related protein aggregation in different tissues differently, leading to increased aggregation in the body-wall muscle, but unexpectedly decreasing aggregation in the pharyngeal muscle. We show that age-related protein aggregation originates from misfolding of newly synthesized proteins in aging animals. Thereby, disruption of the PQC triggers a compensatory response, protecting the pharyngeal muscle from protein aggregation with age by reducing the levels of newly synthesized aggregation-prone proteins.

Overall, our study offers exciting new insights into tissue-specificity of proteostasis regulation and highlights the vulnerability of newly synthesized proteins to age-related protein aggregation. This shows the significance of studying protein aggregation in a tissue-specific manner, as regulation of age-related protein aggregation is not ubiquitous.

Introduction

Maintenance of the proteome in a functional state is critical for cellular and organismal health, with defects in protein homeostasis (proteostasis) being linked to many diseases, including neurodegenerative diseases, immunological and metabolic diseases and cancer [1]. Multiple studies have shown that even in the absence of disease several hundred proteins become insoluble with age [2-8]. To prevent protein aggregation and maintain proteostasis, cells have evolved a complex quality control network. The conventional protein-quality control (PQC) network consists of molecular chaperone pathways, mediating protein folding and refolding, and proteolytic clearance pathways, the ubiquitin-proteasome system (UPS) and autophagy, which remove irreversibly misfolded proteins [9-13]. During aging, the ability of the quality control network to maintain proteostasis under resting and stress conditions gradually declines [2, 14]. While the induction of chaperones is impaired with age [15] and proteasome activity gradually decreases [16], the altered gene expression of certain co-chaperones leads to an enhanced autophagy induction with age [17].

Due to the variation in the proteomes in different cell types and tissues, the need for a tissue-specific proteostasis regulation arises. Components and regulators of the heat shock response (HSR) vary in different tissues [18, 19] and it has been shown that knockdown of certain chaperones or other negative regulators involved in gene expression, protein folding, trafficking or clearance induce HSR in some tissues, but not in others [20]. Due to a tissue-specific hierarchy in the induction of HSF-1, the neuronal heat shock response is delayed compared to muscle cells, leading to neurons being more sensitive to protein aggregation during heat stress. However, neurons have been found to have a higher potential for refolding and disaggregation than muscle cells [21]. Furthermore, muscle and neuronal cells differ in their rate of UPS degradation, with aging influencing the degradation efficiency in a tissue-specific manner. These effects could be explained by different proteasome subunit composition in the two cell types [22]. However, little is known at this stage about how age-related protein aggregation is regulated in different tissues following proteostasis stress.

Here we analyse differences in regulation of age-related protein aggregation in different tissues of the nematode *C. elegans*. We find that PQC impairment, either by inhibition of proteasomal or autophagic degradation or by knockdown of the chaperone system, affect the pharyngeal and the body wall muscle differently. While impaired PQC leads to an increase of protein aggregation in the body wall muscle, we demonstrate an unexpected reduction of aggregation in the pharyngeal muscle. We can show that age-related aggregation is mostly due to the misfolding of proteins newly synthesized in the aging animal and that proteostasis disruption inhibits this aggregation by selectively reducing levels of newly synthesized aggregation-prone proteins.

Results

Tissue-specific reduction of inherent protein aggregation is caused by proteostasis disruption.

As the induction of pathways responsible for proteostasis varies in different tissues [18-20], we were wondering whether regulation of age-related protein aggregation is tissue-specific.

For analysis of inherent protein aggregation with age in different tissues we chose to evaluate aggregation of KIN-19, the mammalian homolog of casein kinase 1 isoform alpha (CK1a), tagged with the fluorescent tags tagRFP or mEOS2, in the body wall muscle and the pharyngeal muscle. We previously showed by mass spectrometry that endogenous KIN-19 becomes 2-fold or more insoluble with age [4]. KIN-19 overexpressed under its endogenous promotor forms immobile puncta with age in the cytoplasm of the pharyngeal cells (Figure S1B) [4]. KIN-19 overexpressed in the body wall muscle accumulates in immobile puncta in young worms and continues to aggregate increasingly with age (Figure 1A and S1A) [4]. The early aggregation of KIN-19 in the body-wall muscle is likely due to high protein expression driven by the strong *myo-3* promotor [4].

To examine how inherent protein aggregation is regulated in the different tissues we chose to disrupt proteostasis by inhibiting the proteasome, the autophagy machinery or the chaperone network. As these systems are important for the correct folding and degradation of proteins we would assume their inhibition to enhance inherent protein aggregation. As expected, downregulation of the proteasome by RNAi against the core subunit PBS-3 or of the key autophagy component ATG-18 leads to an increase of KIN-19 aggregation in the body-wall muscle (Figure 1A). Furthermore, knockdown of the heat shock factor 1 (HSF-1), the key transcription factor regulating chaperone levels, significantly increases KIN-19 aggregation with age in the body wall muscle (Figure 1A).

Surprisingly however, disruption of proteostasis by these factors has a beneficial effect on KIN-19 aggregation in the pharyngeal muscle. Downregulating the proteasome by PBS-3/PAS-6/RPT-6 RNAi reduces KIN-19 aggregation at day 7, while loss-of-function ATG-18 mutation reduces KIN-19 aggregation at all ages (Figure 1B and S1C). Also reduced chaperone levels caused by a loss-of-function HSF-1 mutation leads to a decrease in the rate of KIN-19 aggregation at all ages examined (Figure 1B and S1C).

As reduction of aggregation as a consequence of proteostasis disruption in the pharynx is unexpected, we confirmed those results with another aggregation-prone protein expressed in the pharyngeal muscle. Under the control of the strong promotor *myo-2*, Ras-like GTP-binding protein RHO-1 forms aggregates already in young worms, and shows an increase of aggregation with age (Figure 1C). Inhibiting HSF-1 or protein degradation via defective proteasome or autophagy

significantly reduces RHO-1 aggregation compared to controls. Notably, as we obtained similar results in the pharynx with both aggregation reporters which aggregate to varying extends and on a different timescale, we can rule out that differences in the outcome upon proteostasis disruption between the pharynx and the body-wall muscle are simply a consequence of different levels of aggregation-prone proteins.

Next, we asked whether the conventional protein-quality control systems could be potentially promoting directly or indirectly through their activity the process of inherent protein aggregation in the pharynx. For example, this could be by degradation of an important anti-aggregation factor by the proteasome or by HSF-1 controlling the expression of a prominent pro-aggregation factor involved in inherent protein aggregation. However, we found that increased proteasome activity by RPN-6 overexpression [23] reduces RHO-1 aggregation with age (Figure 1D). Similarly, overexpression of HSF-1 and thereby upregulation of chaperones leads to a clear reduction of RHO-1 and KIN-19 aggregation (Figure 1D and S1C). This demonstrates that both the proteasome and chaperone systems normally prevent inherent protein aggregation.

To determine the overall effect of the disruption of conventional PQC for the organism as a whole, we extracted the highly-insoluble proteins from animals lacking gonads treated with RNAi targeting different proteasomal subunits. Analysis of the total insoluble protein fraction from all somatic tissues clearly demonstrates that disrupted proteostasis leads to enhanced protein aggregation (Figure S2A). Analysis by antibody detection shows a similar trend for KIN-19 insolubility (Figure S2B). These findings suggest that a reversal of protein aggregation upon proteostasis impairment is likely restricted to the pharynx system and possibly a few less prominent tissues.

Together, these results reveal that disruption of cytoprotective proteostasis mechanisms usually responsible for preventing protein aggregation induces a compensatory effect that protects against age-related protein aggregation specifically in the pharyngeal muscles.

Reduced protein aggregation is not due to compensatory upregulation of autophagy or folding capacity.

From these results the question arises how impaired PQC can so effectively reduce the amount of protein aggregates in the pharynx. We wondered whether knockdown of select PQC components reduces inherent protein aggregation by a compensatory upregulation of other conventional PQC components. For example, autophagy activity is up-regulated upon proteasome inhibition [24-26]. In that case we would expect rescue of protein aggregation after knockdown of both degradation pathways. However, we found that reduced proteasome levels did not rescue KIN-19 aggregation in

animals with abrogated autophagy (Figure 2A). Next, we asked whether HSF-1 inhibition reduces aggregation by upregulating degradation of misfolded proteins. Similarly, knockdown of autophagy in HSF-1 mutant animals did not restore RHO-1 aggregation in these animals (Figure 2B).

These results indicate that reduction of protein aggregation triggered by disruption of the proteostasis network in the pharynx is not due to compensatory upregulation of the other conventional network components.

Newly formed proteins are prone to aggregate with age.

To understand better how PQC impairment has beneficial effects on protein aggregation in the pharynx we studied the aggregation kinetics of age-related protein aggregation. Aggregation formation is possible at several stages of the proteins' life cycle. It is yet unclear whether age-related protein aggregation happens with "new" proteins shortly after translation or rather with "aged" proteins after they have accumulated damage. Aggregation of newly synthesized proteins with age could be due to translational mistakes combined with reduced folding and degradation capacity. Notably, a recent study demonstrated codon-specific ribosome pausing induces widespread protein aggregation [27]. Furthermore, a large fraction (estimated between 6% and 30%) of newly synthesized proteins are not correctly folded and are degraded by the proteasome directly after synthesis [28, 29]. In addition, newly synthesized proteins are prone to aggregate after thermal stress [30]. Conversely, abundant evidence exists for the role of post-translational modifications in promoting misfolding and protein aggregation in disease [31]. Importantly, the aggregating proteome is increasingly carbonylated with age and oxidative stress enhances aggregation [7]. To analyse whether age-related protein aggregation is due to accumulation of older proteins with age or to misfolding of newly synthesized protein, we used the photoconvertible fluorescent tag mEOS2 [32] fused to our proteins of interest. Irreversible photoconversion from a green to a red fluorescent state using blue light enables to distinguish between proteins present before the photoconversion (red) and proteins newly synthesized after photoconversion (green). We examined to which extent newly synthesized protein and older proteins are contributing to the age-related protein aggregation observed in the different tissues by quantifying the number of green and red aggregates. Of note, conversion transformed the fluorescence of all of the aggregates already present into the red spectrum. However the conversion was not complete and the core region of some larger aggregates continued to emit green fluorescence. Already 24 hours after conversion, we observed a doubling of the number of animals with green-labelled aggregates in both the pharynx and body-wall muscle.

Conversely, analysis of the red-labelled aggregates demonstrates the capacity of both tissues to remove a portion of aggregates over time (Figure 3 A, B and S3A).

Confocal imaging shows that in less than 24 hours, newly translated proteins aggregate over pre-existing aggregates suggesting a seed-induced aggregation. To a lesser extent, we also observed aggregation independently of pre-existing aggregates (Figure 3C and S3B). Together these results demonstrate that aggregation-prone proteins aggregate shortly after their translation.

Protein-quality control impairment reduces levels of newly synthesized aggregating protein to prevent aggregation.

Next, we asked which step of the aggregation process is modulated by PQC impairment. For this, we quantified green and red KIN-19 aggregates after photoconversion and in combination with proteasome inhibition. Reduced proteasome levels did not influence the slow rate of old aggregate removal in the body wall muscle or in the pharynx. However, proteasome impairment strongly prevented new aggregate formation in the pharynx but not in the body wall muscle (Figure 4A and 4B).

How does PQC disruption prevent the generation of aggregates? Possibilities for intervening with the misfolding and aggregation of newly synthesized proteins are increasing protein refolding or reducing of the amount of newly synthesized proteins. The latter could be achieved by a decrease in protein synthesis or an increase in protein removal in form of degradation or sequestration. Alternatively PQC impairment could be delaying the assembly of large aggregates and causing an accumulation of misfolded soluble proteins.

To assess these alternative possibilities, we analysed the total protein amount of mEOS2-labeled protein by quantifying the fluorescent intensity in the pharynx. For this we measured green fluorescence intensity following photoconversion in control conditions and after PQC impairment. The conversion was performed in day 2 young adults before the appearance of large aggregates. In control conditions, KIN-19::mEOS2 translated after conversion accumulated over 48h due to increased transcriptional activity of the promoter (Figure 4C and Figure S4). Upon proteasome inhibition the accumulation of newly synthesized KIN-19::mEOS2 is significantly reduced (Figure 4C). This compensatory mechanism is only triggered by aggregation-prone proteins as proteasomal knockdown did not significantly change levels of newly synthesized mEOS2-tag alone (Figure S4A). We observed similar findings when inhibiting autophagy (Figure 4C and Figure S4B).

As reduced aggregation by an increase in protein refolding would not change total KIN-19 protein levels, we can exclude that increased folding capacity is responsible for the reduction in protein aggregation after proteasomal inhibition. These results show that disruption of proteostasis in the pharynx selectively reduces the levels of newly synthesized aggregation-prone proteins in the cell and thereby prevents the strong age-dependent increase of aggregation of these proteins.

Discussion

In this study we show that the regulation of age-related protein aggregation is highly tissue-specific. We find that impaired PQC impacts age-related protein aggregation in different tissues in opposing directions by increasing protein aggregation in the body wall muscle but leading to an unexpected reduction of aggregation with age in the pharyngeal muscle. We show that age-related aggregation is generated from proteins newly synthesized in the aging body-wall muscles and pharyngeal muscles. Disruptions in proteasome, autophagy or chaperone systems triggered a protective compensatory response that selectively reduced levels of newly synthesized aggregation-prone proteins with age.

Our studies examining the dynamics of protein aggregation demonstrate that proteins do not need to age by accumulating damage over time to localize into aggregates. Instead our data reveal that aggregation-prone proteins assemble into aggregates shortly after translation. Recent findings show the vulnerability specifically of newly synthesized proteins to thermal stress and proteostasis disruption [30]. Furthermore, newly synthesized proteins are known to be especially vulnerable to misfolding through errors during protein synthesis and folding, for example codon pausing [27], lack of cofactors or inadequate chaperone availability [33, 34]. Studies show that a significant part of all newly synthesized proteins (Defective ribosomal products (DRIPS)) are directly targeted for degradation because they are unable to achieve a stable conformation due to folding and translation errors and accumulate following proteasome inhibition [28, 35]. Our results now connect these findings with aging, as newly synthesized proteins are responsible for the increase of protein aggregation with age, indicating that the control mechanisms to prevent newly synthesized proteins from aggregating are decreasing in efficiency with age, leading to accumulation and aggregation of these proteins.

An important question is how the contrasting response to proteostasis stress is triggered in the different tissues. Variance in the expression level of molecular chaperones and other proteins preventing aggregation [36], or in other factors influencing protein aggregation, like pH, temperature, ion strength, concentration of cosolutes and the concentration of the protein itself [36], could influence the resilience of cells and could trigger compensatory mechanisms specifically in

one tissue and not the other. The cellular environments of both the body wall and pharyngeal muscles seem to favour protein aggregation, as we find KIN-19 aggregating with age in both tissues. Our results strongly suggest that concentration of aggregation-prone proteins is unlikely to be the main factor influencing this response. Furthermore, we can exclude the compensatory upregulation of the HSR as the reason for the reduction of protein aggregation following proteostasis disruption in the pharynx, as double knockdown of HSF-1 and autophagy did not rescue the reduced aggregation phenotype.

Further research will be needed to determine how this compensatory reduction of levels of newly synthesized specifically aggregation prone proteins can be achieved following PQC disruption. Some possibilities are reduced synthesis of specifically aggregation-prone proteins or direct removal of these proteins by an unconventional degradation pathway, like microautophagy, chaperone mediated autophagy [37] or exopher formation [38].

Our results provide interesting new insights into tissue-specificity of proteostasis regulation, showing the ability of certain tissues to compensate in situations of proteostasis stress. Furthermore we show the vulnerability of newly synthesized proteins to age-related protein aggregation. As age-related protein aggregation is considered to be an important factor in several diseases, finding ways of improving proteostasis in aged organisms is an important step in preventing those diseases.

Figure 1

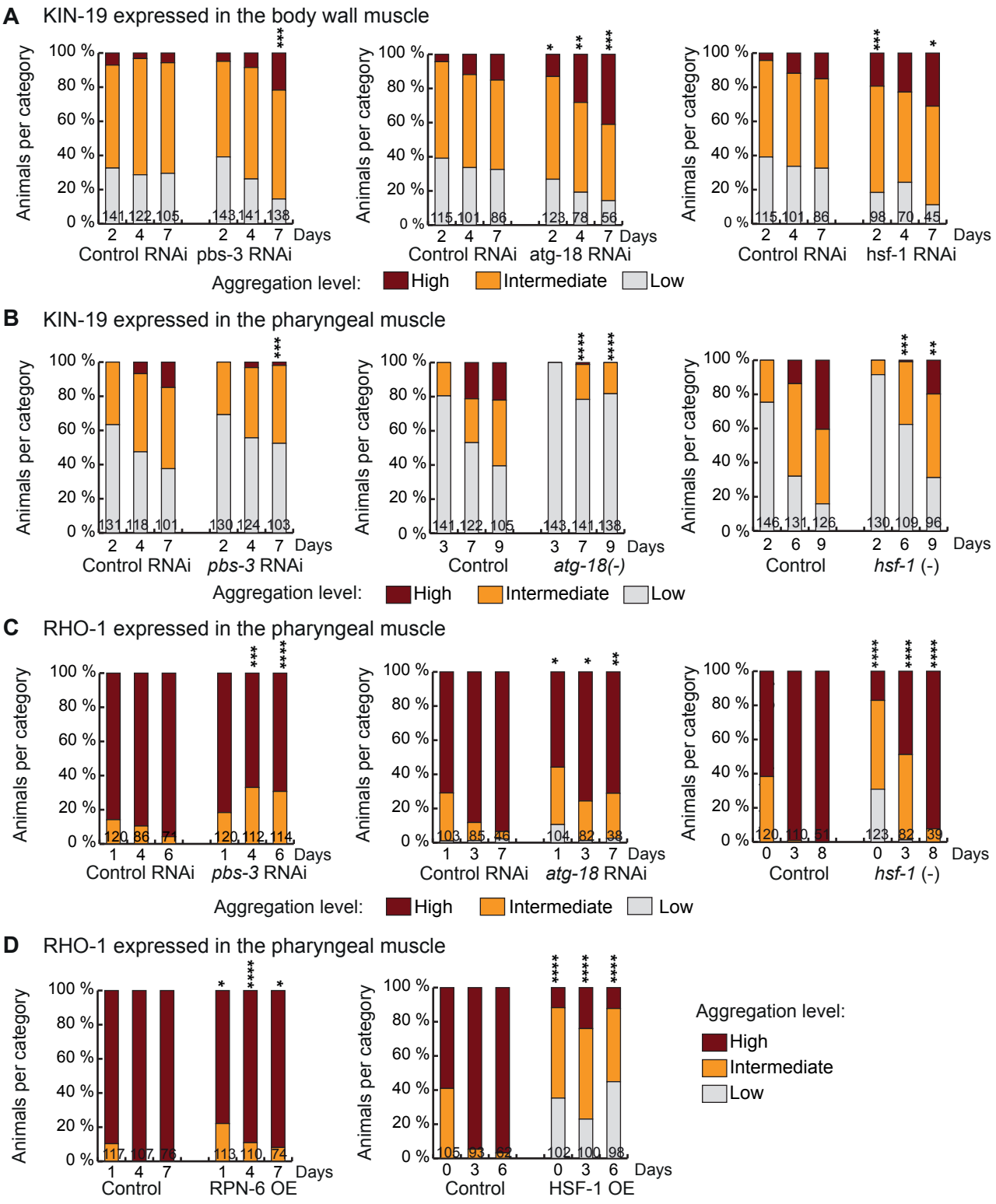


Figure 1. Proteostasis disruption differentially regulates age-related protein aggregation in the pharynx and the body wall muscle

(A) Age-related aggregation of KIN-19::tagRFP expressed in the body wall muscle is increased by inhibition of the proteasome (*pbs-3*), autophagy (*atg-18*) or the chaperone system (*hsf-1*).
(B) Age-related aggregation of KIN-19::tagRFP expressed in the pharyngeal muscle is reduced by inhibition of the proteasome (*pbs-3*), autophagy (*atg-18*) or the chaperone system (*hsf-1*).
(C) Proteostasis disruption by inhibition of the proteasome (*pbs-3*), autophagy (*atg-18*) or the chaperone system (*hsf-1*) reduces age-related aggregation of RHO-1::tagRFP expressed in the pharyngeal muscle.
(D) Activation of the proteasome (RPN-6 overexpression) or the chaperone system (HSF-1 overexpression) delays age-related RHO-1::tagRFP aggregation in the pharynx.

Percentage of worms exhibiting low, intermediate or high aggregation levels in a population of *C. elegans*. Numbers of worms indicated in the bars. Significance calculated low+medium versus high aggregation levels; *p<0.05, **p<0.01, ***p<0.001, ****p < 0.0001.
See also Figure S1, Figure S2

Figure 2

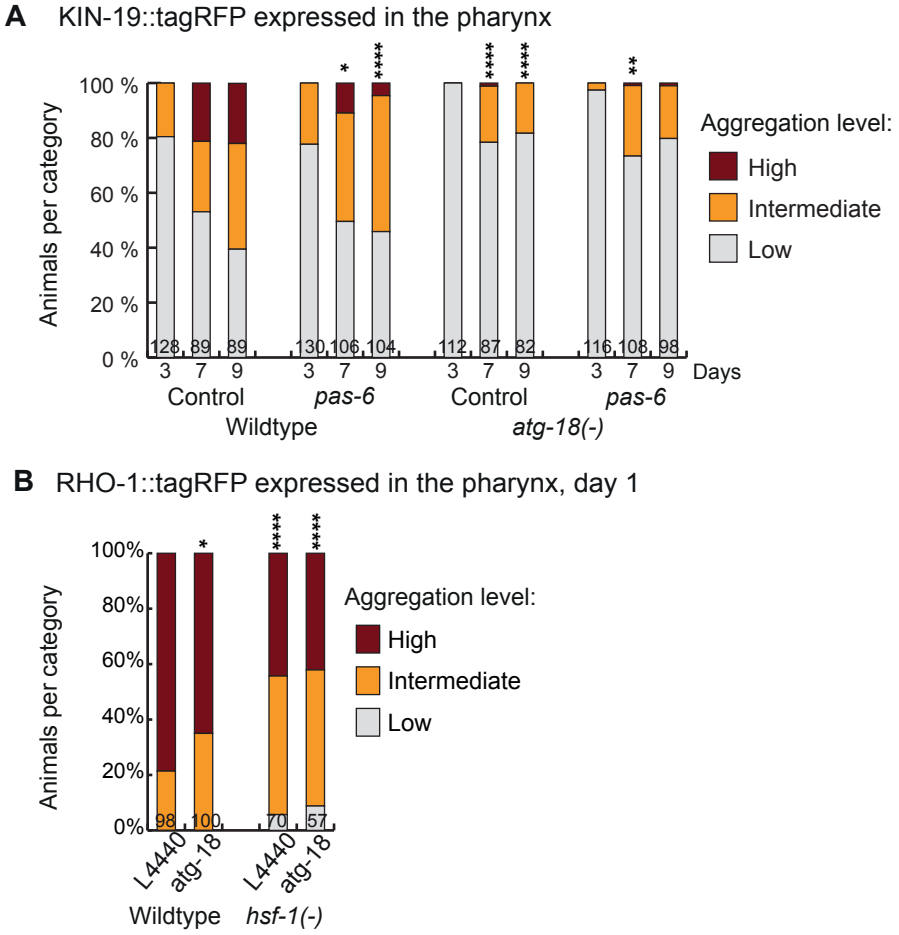


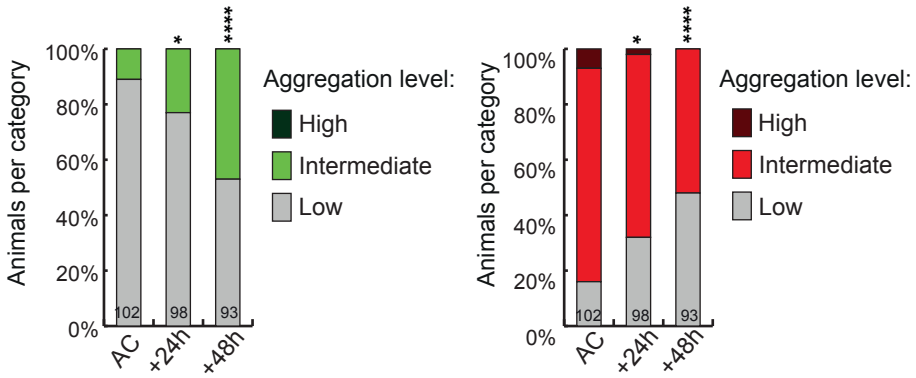
Figure 2. Reduced protein aggregation following proteostasis disruption is not due to compensatory upregulation of the proteasome, autophagy or HSF-1.

(A) Inhibition of both the proteasome and autophagy systems does not restore KIN-19::tagRFP aggregation in the pharynx.
(B) Inhibition of both the chaperone and autophagy systems does not restore KIN-19::tagRFP aggregation.

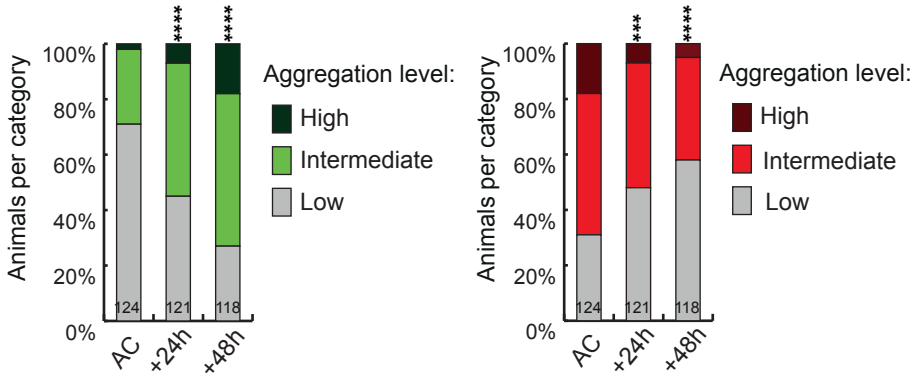
Percentage of worms exhibiting low, intermediate or high aggregation levels in a population of *C. elegans*. Numbers of worms indicated in the bars. Significance calculated low+medium versus high aggregation levels; *p < 0.05, **p < 0.01,****p < 0.0001.

Figure 3

A KIN-19::mEOS2 aggregation in the body wall muscle



B KIN-19::mEOS2 aggregation in the pharynx



C KIN-19::mEOS2 expressed in the pharynx, day 7, 24h after conversion

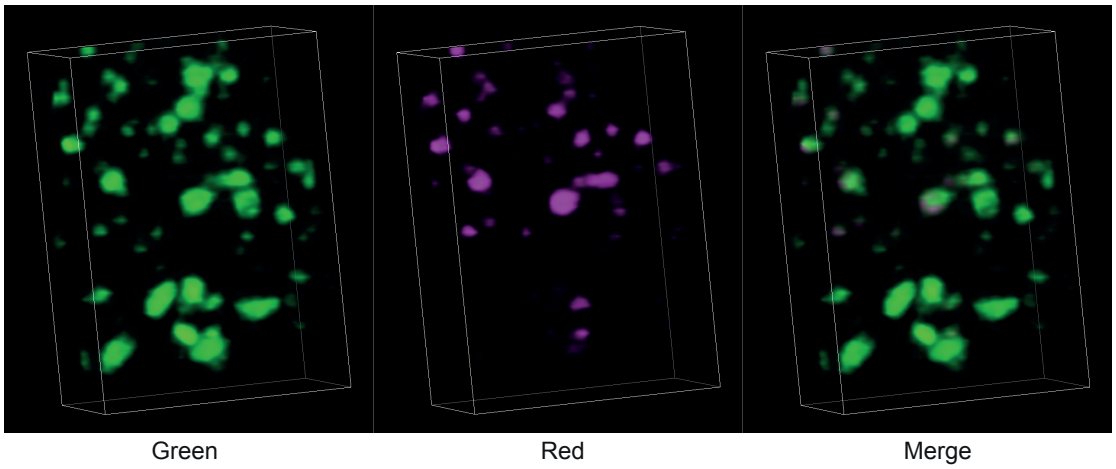


Figure 3. Age-related protein aggregation is due to the accumulation of newly synthesized proteins in aging animals.

(A) In the body wall muscle aggregation of newly synthesized (green) KIN-19::mEOS2 rapidly increases after photoconversion at day 2 whereas levels of red-labelled aggregates decline.
(B) Following photoconversion at day 5 in the pharynx, the number of animals with green-labelled KIN-19::mEOS2 aggregates doubles over 24 hours. Conversely old aggregates labelled in red are slowly removed.

Percentage of worms exhibiting low, intermediate or high aggregation levels in a population of *C. elegans* directly (AC), 24h or 48h after photoconversion. Numbers of worms indicated in the bars. Significance calculated low versus medium+high aggregation levels; *$p < 0.05$, ***$p < 0.001$,****$p < 0.0001$.

(C) 24 hours after photoconversion at day 7, newly synthesized KIN-19::mEOS2 (green) forms new aggregates and associates around older aggregates (red). 3D reconstruction of a confocal stack in the pharyngeal anterior bulb region.

See also Figure S3.

Figure 4

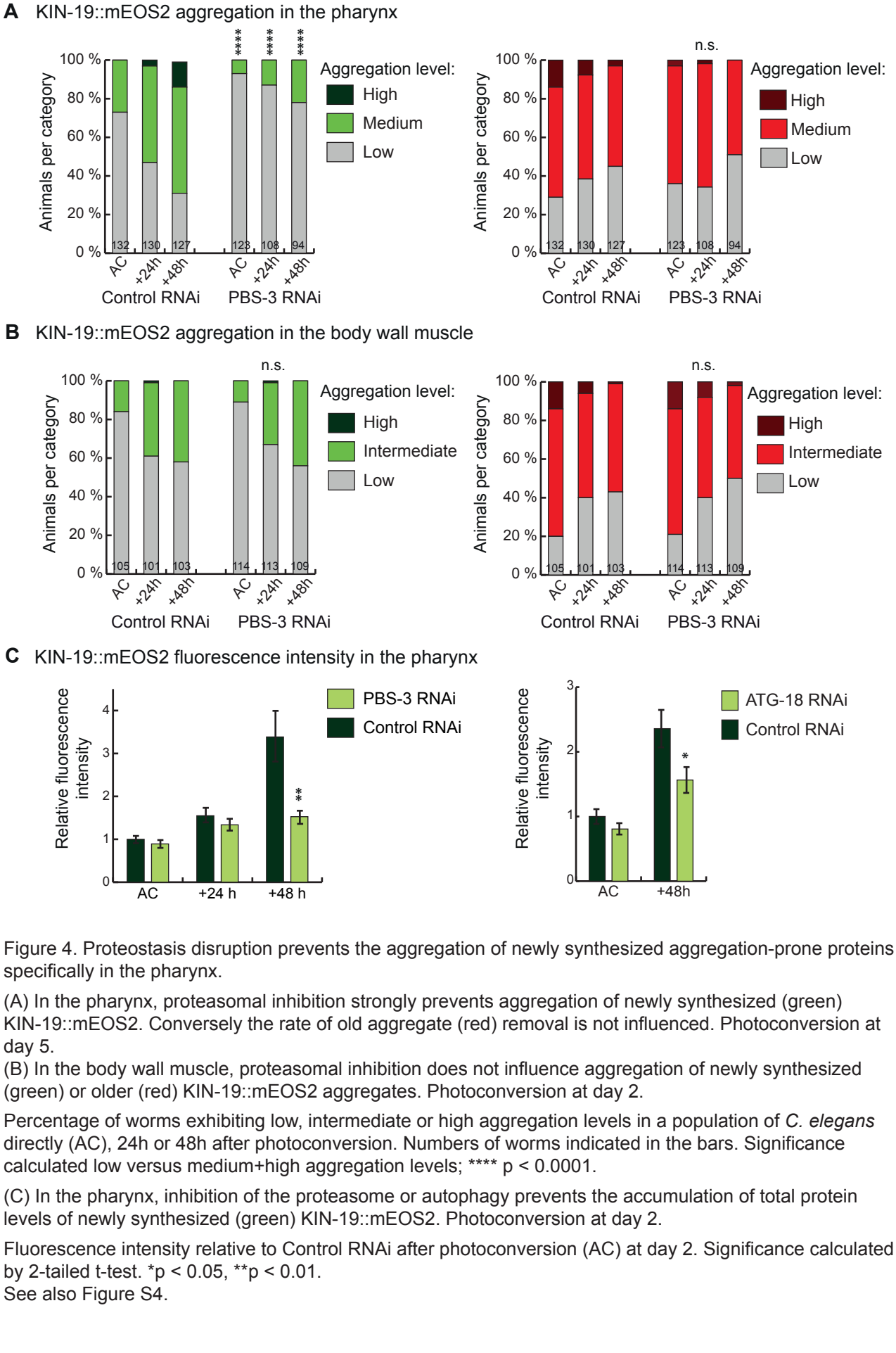


Figure 4. Proteostasis disruption prevents the aggregation of newly synthesized aggregation-prone proteins specifically in the pharynx.

(A) In the pharynx, proteasomal inhibition strongly prevents aggregation of newly synthesized (green) KIN-19::mEOS2. Conversely the rate of old aggregate (red) removal is not influenced. Photoconversion at day 5.
(B) In the body wall muscle, proteasomal inhibition does not influence aggregation of newly synthesized (green) or older (red) KIN-19::mEOS2 aggregates. Photoconversion at day 2.

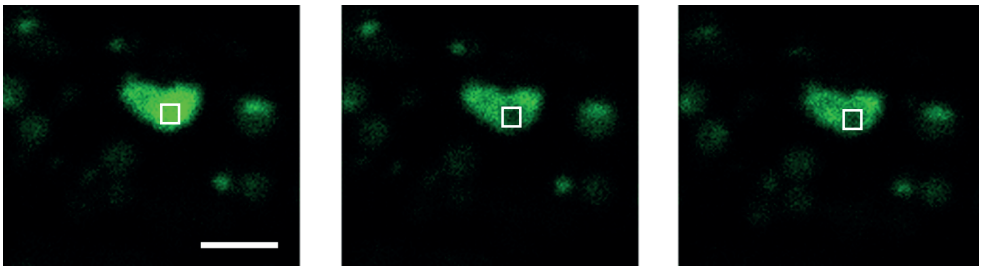
Percentage of worms exhibiting low, intermediate or high aggregation levels in a population of *C. elegans* directly (AC), 24h or 48h after photoconversion. Numbers of worms indicated in the bars. Significance calculated low versus medium+high aggregation levels; **** $p < 0.0001$.

(C) In the pharynx, inhibition of the proteasome or autophagy prevents the accumulation of total protein levels of newly synthesized (green) KIN-19::mEOS2. Photoconversion at day 2.

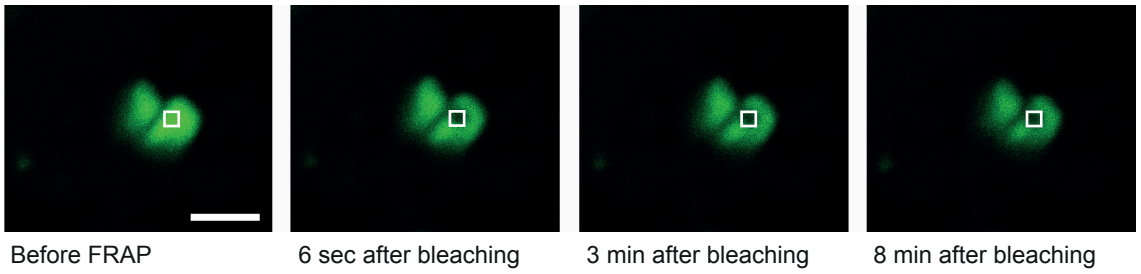
Fluorescence intensity relative to Control RNAi after photoconversion (AC) at day 2. Significance calculated by 2-tailed t-test. *$p < 0.05$, **$p < 0.01$.
See also Figure S4.

Figure S1

A KIN-19::mEOS2 expressed in the body wall muscle



B KIN-19::mEOS2 expressed in the pharynx



C KIN-19 expressed in the pharyngeal muscle

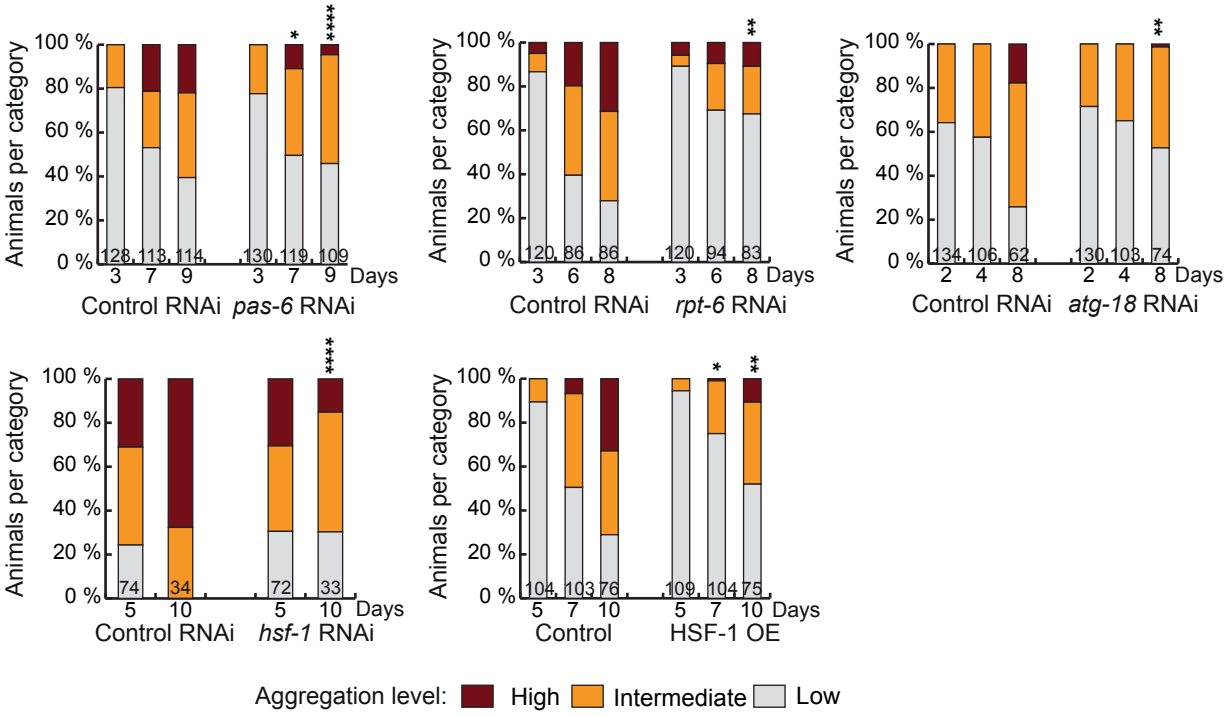


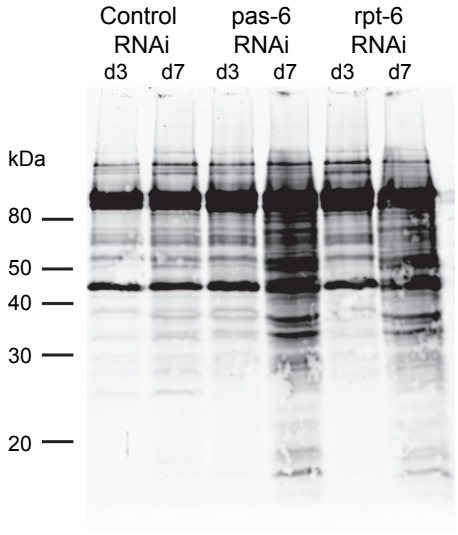
Figure S1. Proteostasis disruption reduces age-related KIN-19 aggregation in the pharynx.

(A) KIN-19::mEOS2 puncta visible in the body wall muscle contain immobile protein, demonstrated by lack of fluorescent recovery after photobleaching (FRAP) after 5 minutes. Scale bar 2µm.
(B) KIN-19::mEOS2 puncta visible in the pharynx contain immobile protein, demonstrated by FRAP after 8 minutes. Scale bar 2µm.
(C) Proteostasis disruption by inhibition of the proteasome (*pbs-6, rpt-6*), autophagy (*atg-18* mutation) or the chaperone system (*hsf-1* mutation) reduces age-related aggregation of KIN-19::tagRFP expressed in the pharyngeal muscle. Improving proteostasis by HSF-1 overexpression also reduces KIN-19::tagRFP aggregation. Percentage of worms exhibiting low, intermediate or high aggregation levels in a population of *C. elegans*. Numbers of worms indicated in the bars. Significance calculated low+medium versus high aggregation levels; *p < 0.05, **p < 0.01, ***p < 0.001,****p < 0.0001.

Corresponding to Figure 1.

Figure S2

A Insoluble fraction, total protein staining



B Insoluble fraction, KIN-19 antibody

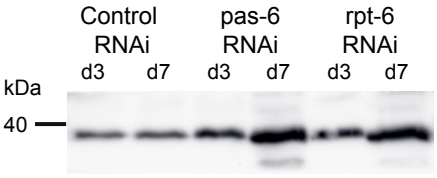


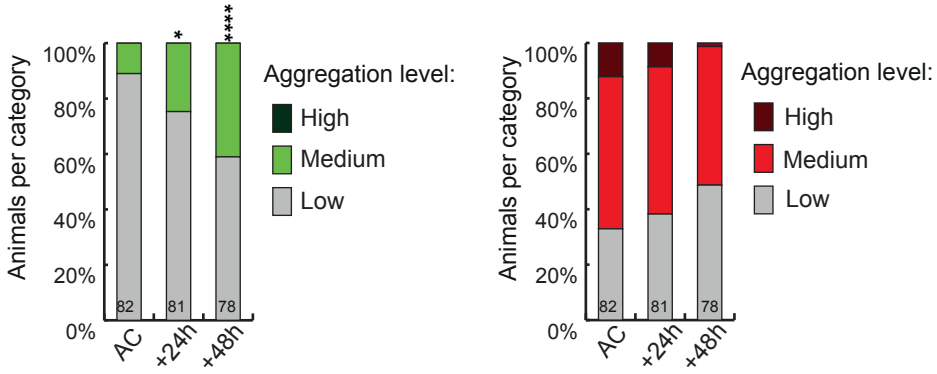
Figure S2. Proteostasis disruption leads to enhanced protein aggregation in whole worm extracts.

(A) Insoluble protein fractions of *gon-2(-)* worms grown on control, *pas-6* or *rpt-6* RNAi till day 3 or day 7 stained with total protein staining demonstrate increase of total protein insolubility following proteasome inhibition.

(B) Insoluble protein fractions of *gon-2(-)* worms grown on control, *pas-6* or *rpt-6* RNAi till day 3 or day 7 stained with KIN-19 antibody show increased levels of insoluble KIN-19 after proteasome inhibition.

Figure S3

A KIN-19::mEOS2 aggregation in the body wall muscle after conversion at day 5



B RHO-1::mEOS2 expressed in the pharynx at day 1, 24h after photoconversion

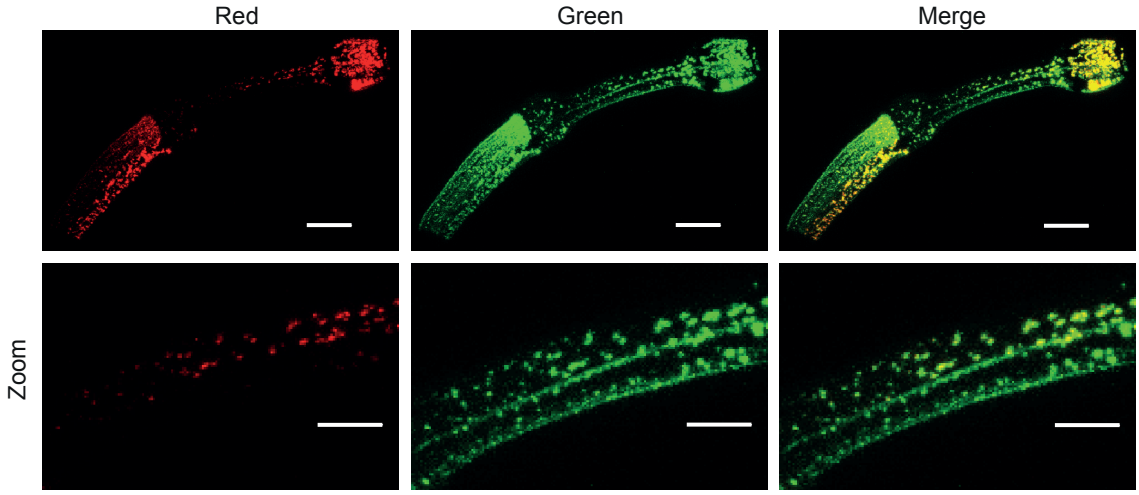
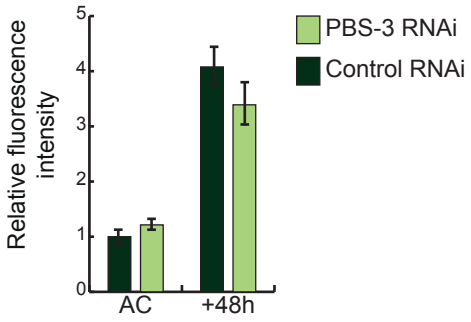


Figure S3. Age-related protein aggregates form from newly synthesized protein in aged animals without seeding and around older aggregates.

(A) In the body wall muscle, aggregation of newly synthesized (green) KIN-19::mEOS2 rapidly increases after photoconversion at day 5 whereas levels of red-labelled aggregates decline. Percentage of worms exhibiting low, intermediate or high aggregation levels in a population of *C. elegans* directly (AC), 24h or 48h after photoconversion. Numbers of worms indicated in the bars. Significance calculated low versus medium+high aggregation levels; *$p < 0.05$,****$p < 0.0001$.
(B) RHO-1::mEOS2 aggregates strongly already at day 1 and forms new aggregates (green) 24 h after photoconversion and associates with older aggregates (red). Scale bar 15µm, 7µm in zoom.
Corresponding to Figure 3.

Figure S4

A mEOS2 transcriptional reporter fluorescence intensity in the pharynx after proteasomal inhibition



B mEOS2 transcriptional reporter fluorescence intensity in the pharynx after autophagy inhibition

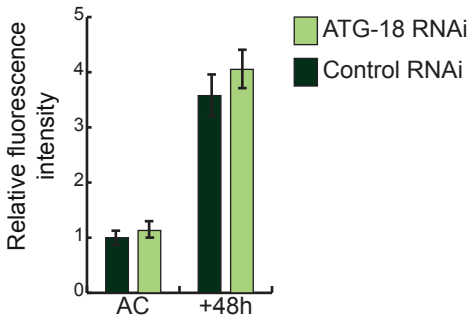


Figure S4. Proteostasis disruption does not generally influence levels of newly synthesized proteins.

(A) In the pharynx, proteasomal inhibition does not influence total protein levels of newly synthesized (green) or older (red) mEOS2 control protein.
(B) In the pharynx, inhibition of autophagy does not influence total protein levels of newly synthesized (green) or older (red) mEOS2 control protein.

Fluorescence intensity relative to Control RNAi after photoconversion (AC) at day 2. Significance calculated by 2-tailed t-test.
Corresponding to Figure 4.

Experimental Procedures

Strains:

Wild type: N2

CF2253: *gon-2(q388)I*

Transgenics:

CF3166: *muEx473[pC03C10.1::C03C10.1::tagrfp + Ptph-1::gfp]*

CF3649: N2; *muIs209[Pmyo-3::kin-19::tagrfp + Ptph-1::gfp]*

CF3706: N2; *muEx587[Pkin-19::kin-19::meos2 + Punc-122::gfp]*

DCD13: N2; *uqIs9[Pmyo-2::rho-1::tagrfp + Ptph-1::gfp]*

DCD69: N2; *uqEx4[Pmyo-3::kin-19::meos2]*

DCD83: *ttTi5605II; unc-119(ed3)III; uqEx[Pmyo-2::rho-1::meos2 + Punc-122::gfp + cb-unc-119(+)]*

DCD88: *muIs115[Phsfp::hsf-1 + Pmyo-3p::gfp]; uqIs9[Pmyo-2::rho-1::tagrfp + Ptph-1::gfp]*

DCD173: *hsf-1(sy441)I; muEx587[Pkin-19::kin-19::meos2 + Punc-122::gfp]*

DCD174: *atg-18(gk378)V; muEx587[Pkin-19::kin-19::meos2 + Punc-122::gfp]*

DCD190: N2; *uqIs9[Pmyo-2::rho-1::tagrfp + Ptph-1::gfp]; uthEx557[Sur-5p::rpn-6 + Myo-3p::gfp]*

DCD245: N2; *uqEx49[Pkin-19::meos2]*

DCD258: *hsf-1(sy441)I; muIs209[Pmyo-3::kin-19::tagrfp + Ptph-1::gfp]*

Maintenance:

All strains were kept at 15°C on NGM plates inoculated with OP50 using standard techniques. Age-synchronization was achieved by transferring adults of the desired strain to 20°C and selecting their progeny at L4 stage. From L4 stage on experiments were performed at 20°C.

RNAi treatment:

RNA treatment was performed by feeding as previously published [39]. The RNAi clones were acquired from the Marc Vidal or the Julie Ahringer RNAi feeding library (Source BioScience, UK) and sequenced. HT115 containing the empty vector L4440 was used as control.

Imaging:

For confocal analysis using a Leica SP8 confocal microscope with the HC PL APO CS2 63x1.40 oil objective, worms were mounted onto slides with 2% agarose pads using 2 µM levamisole for anaesthesia. Worms were examined using the Leica HyD hybrid detector. The tag mEOS2 was detected using 506nm as excitation and an emission range from 508-525nm for green fluorescence and 571nm as excitation and an emission range from 573-602nm for red fluorescence. 3D reconstructions were performed using the Leica Application Suite (LAS X).

FRAP analysis:

FRAP analysis was performed as previously described [4] using the Leica SP8 confocal microscope with the HC PL APO CS2 63x 1.30 glycerol objective and PMT detector. Relative fluorescence intensity (RFI) was analysed as described previously following the equation RFI = (Tt/Ct)/(T0/C0), where T0 is the intensity in the region of interest (ROI) before photobleaching; Tt, the intensity in the ROI at a defined time after photobleaching; C0, the intensity in the non-bleached part of the puncta before photobleaching; and Ct, the intensity in the non-bleached part of the puncta after bleaching [40].

Protein extraction and western blotting analysis:

Isolation of insoluble proteins for western blot analysis and western blotting was performed as previously published [41]. Bands were detected by immunoblot using KIN-19 antibody (Cell Signalling, 1:500) or whole blot staining by Sypro Ruby blot staining (Thermo Scientific) for total protein levels.

Photoconversion of mEOS2-tag and quantification of fluorescence levels:

For photoconversion worms were placed onto small (diameter 35 mm) NGM plates without food and exposed to blue light five times for six minutes, with 2 minutes breaks. For quantification of fluorescence levels worms were mounted onto slides with 2% agarose pads using 2 µM levamisole for anaesthesia. Using an Axio Observer.Z1 (Zeiss), levels of green (eGFP set 38HE, excitation 470 ±40nm, emission 525 ±50nm) and red (dsRED set 43HE, excitation 550 ±25nm, emission 605 ±70nm) fluorescence were detected and intensities determined using ImageJ [42, 43].

Aggregation quantification in vivo:

Aggregation levels were determined using Leica fluorescence microscope M165 FC with a Planapo 2.0x objective. Animals overexpressing *Pmyo-2::Rho-1::TagRFP* were divided into three categories according to the aggregation visible in the Isthmus: animals with no aggregation (low aggregation), animals with aggregation in up to 50% (intermediate aggregation) and animals with aggregation in more than 50% (high aggregation) of the isthmus. Animals overexpressing *Pmyo-3::KIN-19::TagRFP* were divided into three categories: animals with more than 15 puncta in the head or the middle body region (low aggregation), animals with more than 15 puncta in the head and the middle body region (intermediate aggregation) and animals with puncta in head, middle body and tail region (high aggregation). Animals expressing *Pkin-19::KIN-19::mEOS2* were divided into less than 10 puncta (low aggregation), between 10 and 100 puncta (intermediate aggregation) and over 100 puncta in the anterior bulb (high aggregation).

Counting was done in a blind fashion. Two-tailed Fisher's exact test using an online tool (http://www.socscistatistics.com/tests/fisher/default2.aspx) was performed for statistical analysis.

Authors contributions:

MCL, RJ and DCD designed and performed experiments. MCL and DCD wrote the paper.

Acknowledgements:

This work was supported by funding from the DZNE and a Marie Curie International Reintegration Grant (322120 to DCD). Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440).

References:

1. Balchin, D., M. Hayer-Hartl, and F.U. Hartl, *In vivo aspects of protein folding and quality control.* Science, 2016. **353**(6294): p. aac4354.
2. David, D.C., *Aging and the aggregating proteome.* Front Genet, 2012. **3**: p. 247.
3. Ayyadevara, S., et al., *Age- and Hypertension-Associated Protein Aggregates in Mouse Heart Have Similar Proteomic Profiles.* Hypertension, 2016. **67**(5): p. 1006-13.
4. David, D.C., et al., *Widespread protein aggregation as an inherent part of aging in C. elegans.* PLoS Biol, 2010. **8**(8): p. e1000450.
5. Demontis, F. and N. Perrimon, *FOXO/4E-BP signaling in Drosophila muscles regulates organism-wide proteostasis during aging.* Cell, 2010. **143**(5): p. 813-25.
6. Reis-Rodrigues, P., et al., *Proteomic analysis of age-dependent changes in protein solubility identifies genes that modulate lifespan.* Aging Cell, 2012. **11**(1): p. 120-7.
7. Tanase, M., et al., *Role of Carbonyl Modifications on Aging-Associated Protein Aggregation.* Sci Rep, 2016. **6**: p. 19311.
8. Walther, D.M., et al., *Widespread Proteome Remodeling and Aggregation in Aging C. elegans.* Cell, 2015. **161**(4): p. 919-32.
9. Sontag, E.M., R.S. Samant, and J. Frydman, *Mechanisms and Functions of Spatial Protein Quality Control.* Annu Rev Biochem, 2017. **86**: p. 97-122.
10. Houck, S.A., S. Singh, and D.M. Cyr, *Cellular responses to misfolded proteins and protein aggregates.* Methods Mol Biol, 2012. **832**: p. 455-61.
11. Jeng, W., et al., *Molecular chaperones: guardians of the proteome in normal and disease states.* F1000Res, 2015. **4**.
12. Lindberg, I., et al., *Chaperones in Neurodegeneration.* J Neurosci, 2015. **35**(41): p. 13853-9.
13. Chen, B., et al., *Cellular strategies of protein quality control.* Cold Spring Harb Perspect Biol, 2011. **3**(8): p. a004374.
14. Labbadia, J. and R.I. Morimoto, *The biology of proteostasis in aging and disease.* Annu Rev Biochem, 2015. **84**: p. 435-64.
15. Calderwood, S.K., A. Murshid, and T. Prince, *The shock of aging: molecular chaperones and the heat shock response in longevity and aging--a mini-review.* Gerontology, 2009. **55**(5): p. 550-8.
16. Saez, I. and D. Vilchez, *The Mechanistic Links Between Proteasome Activity, Aging and Age-related Diseases.* Curr Genomics, 2014. **15**(1): p. 38-51.
17. Gamerdinger, M., et al., *Protein quality control during aging involves recruitment of the macroautophagy pathway by BAG3.* EMBO J, 2009. **28**(7): p. 889-901.
18. Ma, J., et al., *Cellular Proteomes Drive Tissue-Specific Regulation of the Heat Shock Response.* G3 (Bethesda), 2017. **7**(3): p. 1011-1018.
19. Sala, A.J., L.C. Bott, and R.I. Morimoto, *Shaping proteostasis at the cellular, tissue, and organismal level.* J Cell Biol, 2017. **216**(5): p. 1231-1241.
20. Guisbert, E., et al., *Identification of a tissue-selective heat shock response regulatory network.* PLoS Genet, 2013. **9**(4): p. e1003466.
21. Kern, A., et al., *HSF1-controlled and age-associated chaperone capacity in neurons and muscle cells of C. elegans.* PLoS One, 2010. **5**(1): p. e8568.
22. Hamer, G., O. Matilainen, and C.I. Holmberg, *A photoconvertible reporter of the ubiquitin-proteasome system in vivo.* Nat Methods, 2010. **7**(6): p. 473-8.
23. Vilchez, D., et al., *RPN-6 determines C. elegans longevity under proteotoxic stress conditions.* Nature, 2012. **489**(7415): p. 263-8.
24. Iwata, A., et al., *HDAC6 and microtubules are required for autophagic degradation of aggregated huntingtin.* J Biol Chem, 2005. **280**(48): p. 40282-92.

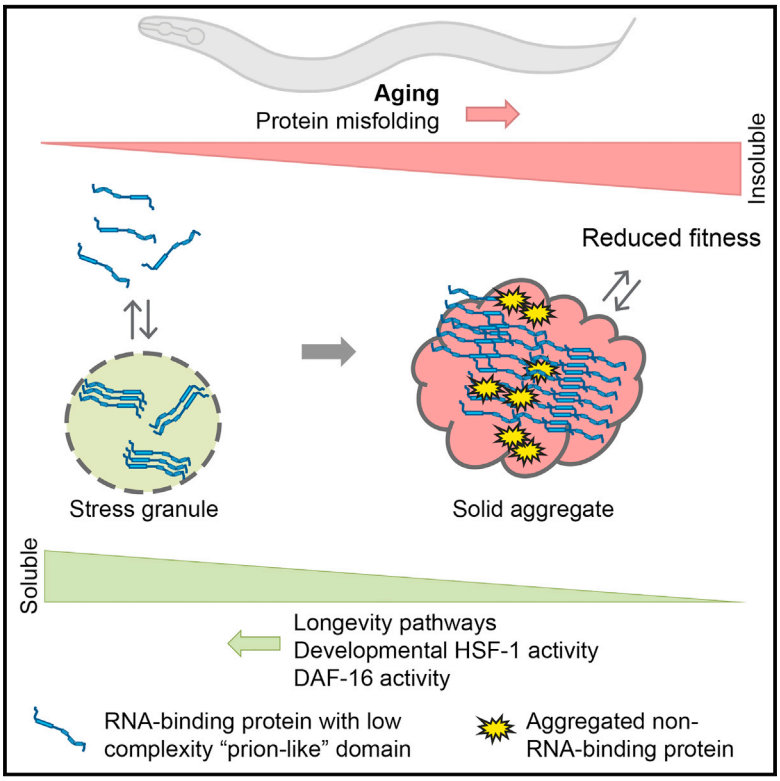
25. Ding, W.X., et al., *Linking of autophagy to ubiquitin-proteasome system is important for the regulation of endoplasmic reticulum stress and cell viability.* Am J Pathol, 2007. **171**(2): p. 513-24.
26. Pandey, U.B., et al., *HDAC6 rescues neurodegeneration and provides an essential link between autophagy and the UPS.* Nature, 2007. **447**(7146): p. 859-63.
27. Nedialkova, D.D. and S.A. Leidel, *Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity.* Cell, 2015. **161**(7): p. 1606-18.
28. Schubert, U., et al., *Rapid degradation of a large fraction of newly synthesized proteins by proteasomes.* Nature, 2000. **404**(6779): p. 770-4.
29. Qian, S.B., et al., *CHIP-mediated stress recovery by sequential ubiquitination of substrates and Hsp70.* Nature, 2006. **440**(7083): p. 551-5.
30. Xu, G., et al., *Vulnerability of newly synthesized proteins to proteostasis stress.* J Cell Sci, 2016. **129**(9): p. 1892-901.
31. Del Monte, F. and G. Agnetti, *Protein post-translational modifications and misfolding: new concepts in heart failure.* Proteomics Clin Appl, 2014. **8**(7-8): p. 534-42.
32. McKinney, S.A., et al., *A bright and photostable photoconvertible fluorescent protein.* Nat Methods, 2009. **6**(2): p. 131-3.
33. Hartl, F.U. and M. Hayer-Hartl, *Molecular chaperones in the cytosol: from nascent chain to folded protein.* Science, 2002. **295**(5561): p. 1852-8.
34. Frydman, J., *Folding of newly translated proteins in vivo: the role of molecular chaperones.* Annu Rev Biochem, 2001. **70**: p. 603-47.
35. Vabulas, R.M. and F.U. Hartl, *Protein synthesis upon acute nutrient restriction relies on proteasome function.* Science, 2005. **310**(5756): p. 1960-3.
36. Fink, A.L., *Protein aggregation: folding aggregates, inclusion bodies and amyloid.* Fold Des, 1998. **3**(1): p. R9-23.
37. Arias, E. and A.M. Cuervo, *Chaperone-mediated autophagy in protein quality control.* Curr Opin Cell Biol, 2011. **23**(2): p. 184-9.
38. Melentijevic, I., et al., *C. elegans neurons jettison protein aggregates and mitochondria under neurotoxic stress.* Nature, 2017. **542**(7641): p. 367-371.
39. Hansen, M., et al., *New genes tied to endocrine, metabolic, and dietary regulation of lifespan from a Caenorhabditis elegans genomic RNAi screen.* PLoS Genet, 2005. **1**(1): p. 119-28.
40. Brignull, H.R., et al., *Modeling polyglutamine pathogenesis in C. elegans.* Methods Enzymol, 2006. **412**: p. 256-82.
41. Lechler, M.C., et al., *Reduced Insulin/IGF-1 Signaling Restores the Dynamic Properties of Key Stress Granule Proteins during Aging.* Cell Rep, 2017. **18**(2): p. 454-467.
42. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nat Methods, 2012. **9**(7): p. 676-82.
43. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis.* Nat Methods, 2012. **9**(7): p. 671-5.

Article

Cell Reports

Reduced Insulin/IGF-1 Signaling Restores the Dynamic Properties of Key Stress Granule Proteins during Aging

Graphical Abstract



Authors

Marie C. Lechler, Emily D. Crawford, Nicole Groh, ..., Jonathan C. Trinidad, Alma L. Burlingame, Della C. David

Correspondence

della.david@dzne.de

In Brief

Lechler et al. show that RNA-binding proteins (RBPs) including stress granule proteins are prone to aggregate with age in *C. elegans*. Aggregation of stress granule RBPs with "prion-like" domains is associated with reduced fitness. Their aggregation is prevented by longevity pathways and promoted by the aggregation of other misfolded proteins.

Highlights

- RNA-binding proteins (RBPs) with "prion-like" domains form solid aggregates with age
- Reduced *daf-2* signaling preferentially prevents insolubility of RNA granule proteins
- Co-aggregation with other misfolded proteins promotes stress granule RBP aggregation
- Aggregation of key stress-granule-related RBPs is associated with impaired health

Accession Numbers

PXD003451

Lechler et al., 2017, Cell Reports *18*, 454–467
January 10, 2017
http://dx.doi.org/10.1016/j.celrep.2016.12.033

OPEN ACCESS
CellPress

Cell Reports
Article

Reduced Insulin/IGF-1 Signaling Restores the Dynamic Properties of Key Stress Granule Proteins during Aging

Marie C. Lechler,[1,2] Emily D. Crawford,[1,5] Nicole Groh,[1,2] Katja Widmaier,[1] Raimund Jung,[1] Janine Kirstein,[3] Jonathan C. Trinidad,[4,6] Alma L. Burlingame,[4] and Della C. David[1,7,*]

[1]German Center for Neurodegenerative Diseases, 72076 Tübingen, Germany
[2]Graduate Training Centre of Neuroscience, 72074 Tübingen, Germany
[3]Leibniz-Institut für Molekulare Pharmakologie im Forschungsverbund Berlin, 13125 Berlin, Germany
[4]Mass Spectrometry Facility, Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94158, USA
[5]Present address: Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158, USA
[6]Present address: Department of Chemistry, Indiana University, Bloomington, IN 47405, USA
[7]Lead Contact
*Correspondence: della.david@dzne.de
http://dx.doi.org/10.1016/j.celrep.2016.12.033

SUMMARY

Low-complexity "prion-like" domains in key RNA-binding proteins (RBPs) mediate the reversible assembly of RNA granules. Individual RBPs harboring these domains have been linked to specific neurodegenerative diseases. Although their aggregation in neurodegeneration has been extensively characterized, it remains unknown how the process of aging disturbs RBP dynamics. We show that a wide variety of RNA granule components, including stress granule proteins, become highly insoluble with age in *C. elegans* and that reduced insulin/insulin-like growth factor 1 (IGF-1) *daf-2* receptor signaling efficiently prevents their aggregation. Importantly, stress-granule-related RBP aggregates are associated with reduced fitness. We show that heat shock transcription factor 1 (HSF-1) is a main regulator of stress-granule-related RBP aggregation in both young and aged animals. During aging, increasing DAF-16 activity restores dynamic stress-granule-related RBPs, partly by decreasing the buildup of other misfolded proteins that seed RBP aggregation. Longevity-associated mechanisms found to maintain dynamic RBPs during aging could be relevant for neurodegenerative diseases.

INTRODUCTION

Young, healthy organisms strive to maintain their proteome in a functional state through the tight control of rates of protein synthesis, folding, and degradation. Extensive quality-control systems are set up throughout the cell to prevent and manage protein damage. As the organism ages, these control mechanisms become less efficient, leading to a disruption in protein homeostasis (Balch et al., 2008; David, 2012). Aging is the main risk factor for a variety of neurodegenerative diseases where specific proteins accumulate as pathological aggregates. Recently, there has been considerable interest in investigating widespread protein aggregation in the absence of disease. Multiple studies have demonstrated that several hundred proteins become highly detergent-insoluble in aged animals (Ayyadevara et al., 2016; David, 2012; David et al., 2010; Demontis and Perrimon, 2010; Reis-Rodrigues et al., 2012; Tanase et al., 2016; Walther et al., 2015). Computational analysis of the insoluble proteome indicates an overrepresentation of proteins with functional and structural similarities (David et al., 2010). The examination of some of these proteins in vivo reveals their assembly into large "solid" aggregates with age similar to those formed in the context of disease. The discovery of endogenous age-dependent protein aggregation in model organisms gives us the unprecedented opportunity to dissect the intrinsic cellular machineries responsible for preventing protein aggregation without using ectopically expressed human disease-associated proteins. At this time, very little is known concerning the regulation of widespread protein insolubility with age and its consequences for the health of the organism. Interestingly, several studies show that protein insolubility is modified in long-lived animals with reduced insulin/insulin growth factor (IGF)-1 *daf-2* signaling, but it remains unclear to which extent (David et al., 2010; Demontis and Perrimon, 2010; Walther et al., 2015).

A growing number of familial and sporadic forms of neurodegenerative diseases show pathological inclusions caused by abnormal aggregation of RNA-binding proteins (RBPs). The first RBPs identified in these inclusions were TAR DNA binding protein of 43 kDa (TDP-43) and fused in sarcoma (FUS), associated with amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD) (Arai et al., 2006; Neumann et al., 2009; Neumann et al., 2006). Since then additional RBPs such as TAF15, EWSR1, hnRNPA2B1, hnRNPA1,

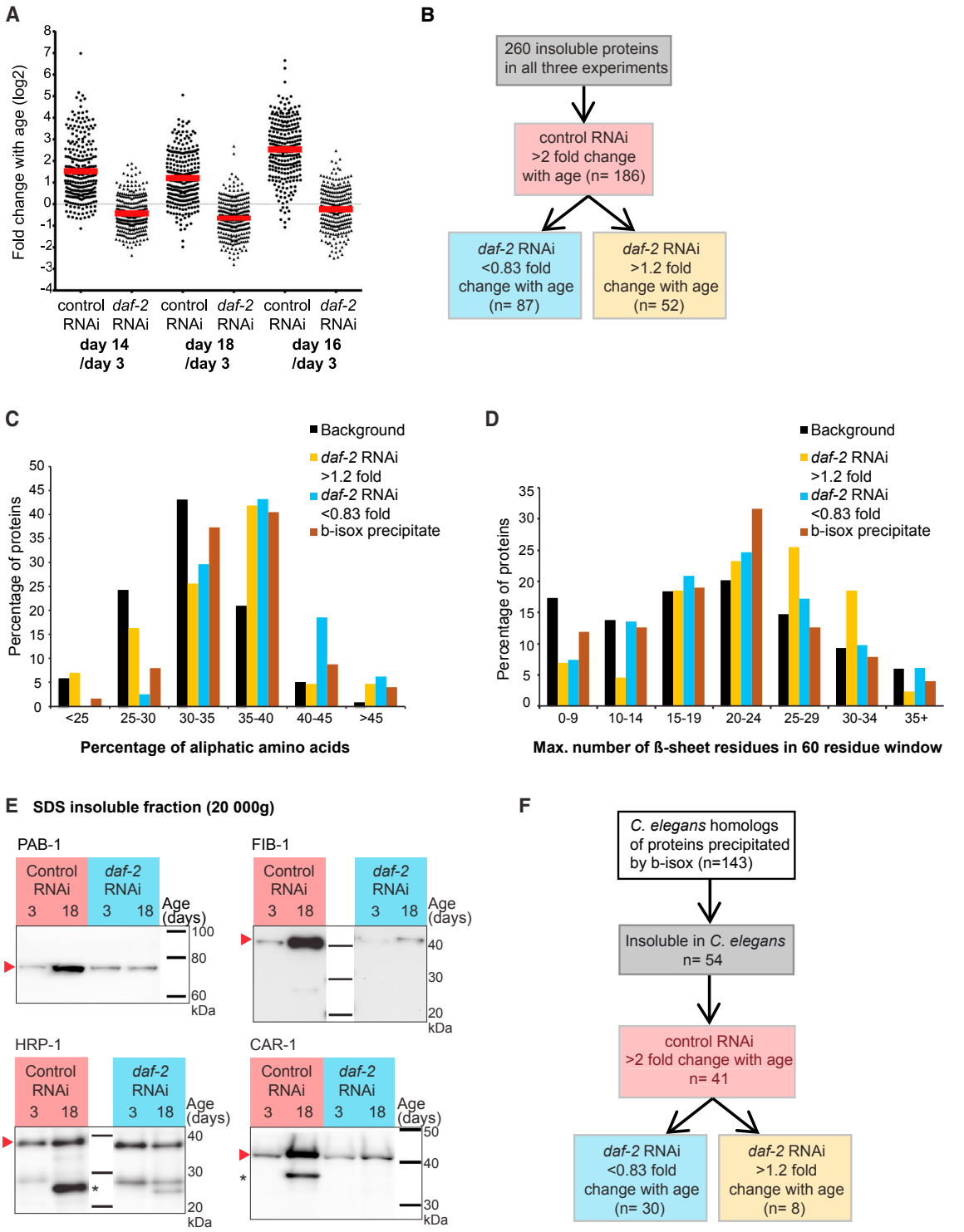


Figure 1. RNA Granule Components No Longer Aggregate in Long-Lived Animals with Reduced *daf-2* Signaling

(A) Distribution of fold changes with age in insolubility for 260 proteins (n = 3, biological replicates). Fold changes are measured by iTRAQ quantification.

(B) Flowchart describing the segregation of insolubility fold changes with age into different groups for analysis.

(C) Proteins with reduced aggregation in *daf-2* RNAi conditions and U2OS proteins precipitated by b-isox are enriched in aliphatic amino acids. Unequal variance t test: proteins aggregating more with age in long-lived animals (n = 43, >1.2-fold), p = 0.04; proteins aggregating less with age in long-lived animals (n = 81, <0.83-fold), p = 6.1E-15; proteins precipitated by b-isox (n = 126), p = 1.3E-9.

(legend continued on next page)

and hnRNPA3 have been associated with neurodegenerative diseases (Kim et al., 2013; Neumann et al., 2011). All of the known RBPs associated with dementia contain a low-complexity (LC) "prion-like" domain enriched in glycines and uncharged polar amino acids, and similar to the sequences driving yeast prion aggregation (Alberti et al., 2009; King et al., 2012). Mutations in this domain enhance pathology by accelerating aggregation (Johnson et al., 2009; Kim et al., 2013). LC prion-like domains are also present in key RBPs that mediate the assembly of RNA granules by liquid-liquid phase separation (Lin et al., 2015; Molliex et al., 2015; Murakami et al., 2015; Patel et al., 2015). Significantly, a small proportion of liquid droplets made by RBPs transform into solid aggregates over time in vitro (Lin et al., 2015; Molliex et al., 2015; Murakami et al., 2015; Patel et al., 2015). For clarity, we will use the term *aggregation* only when referring to the formation of non-dynamic RBP aggregates. An important question is whether the special assembly properties of RBPs puts them at risk of aggregating during aging in a multicellular organism and not just in the context of disease. Interestingly, several RBPs with LC prion-like domains were identified in the insoluble proteome of aged animals (David et al., 2010). Overall, it is imperative to know the causes and consequences of wild-type RBP aggregation during aging in order to fully understand RBP aggregation in neurodegenerative diseases. Furthermore, it is likely that the organism has evolved specific mechanisms to control liquid droplet protein aggregation.

In the current study, we chose to focus on key RBPs responsible for stress granule formation. Stress granules are a specific type of RNA granule that protect the cell by sequestering mRNA from the translational machinery during periods of stress. Importantly, stress granule proteins are often found to co-localize with pathological inclusions of TDP-43 and FUS (Bentmann et al., 2013; Li et al., 2013). Whether these stress granule proteins are innocent bystanders transiently interacting with TDP-43 and FUS or whether they co-aggregate and accelerate disease-associated RBP aggregation remains intensely debated (Bentmann et al., 2013; Li et al., 2013).

We show that key stress-granule-related RBPs (sgRBPs) accumulate in aberrant stress granule-like puncta and in large solid aggregates in aged *C. elegans*. Proteomic analysis revealed that long-lived animals with reduced *daf-2* signaling preferentially abrogate the insolubility of RNA granule components. Importantly, sgRBP aggregates are associated with reduced animal size, motility, and lifespan. We show that sgRBP aggregation is triggered at an earlier age by their co-aggregation with other misfolded proteins, a process that is prevented by DAF-16 in *daf-2* mutants. In addition, the proteostasis network established by heat shock transcription factor 1 (HSF-1) during development is required to maintain dynamic stress granule proteins throughout the animal's life.

RESULTS

Long-Lived Animals with Reduced *daf-2* Signaling Prevent Widespread Protein Insolubility with Age

To identify and quantify changes in aggregation-prone proteins in animals with reduced *daf-2* signaling, we performed an in-depth proteomic analysis of the insoluble proteome from both control and long-lived animals (Figure 1A; Table S1). Because protein misfolding and aggregation is highly abundant in aged *C. elegans* gonads and masks changes in other somatic tissues (David et al., 2010; Goudeau and Aguilaniu, 2010; Zimmerman et al., 2015), we used a gonad-less mutant to focus our analysis on protein insolubility in non-reproductive tissues. We isolated large aggregates that are pelleted by low centrifugal forces (20,000 × *g*) and insoluble in 0.5% SDS. In three biological replicates, we identified 260 insoluble proteins, of which 186 were highly prone to aggregate with age in control animals (Figure 1B). The strong correlations between the control replicates ($r = 0.77$, $r = 0.85$, and $r = 0.67$) and long-lived replicates ($r = 0.87$, $r = 0.86$. and $r = 0.82$) attest to the quality of the quantification and experimental reproducibility (Figures S1A and S1B). None of these insoluble proteins was more prone to aggregate with age in the long-lived animals as compared with controls. These results are consistent with previous observations (David et al., 2010; Demontis and Perrimon, 2010). Surprisingly, a recent proteomic study showed that endogenous protein insolubility is higher in *daf-2* mutants than in wild-type animals (Walther et al., 2015). To account for procedural differences, we performed the extraction following the less stringent extraction protocol from Walther et al. (2015). By omitting SDS and using ultracentrifugation at 500,000 × *g*, Walther et al. analyzed highly insoluble large aggregates (as in this study), as well as smaller and more soluble aggregates. However, we did not observe a general change in the action of *daf-2* signaling on protein insolubility with age after using the less stringent extraction protocol (Figure S1C). Next, we asked whether the inconsistencies between the studies could be related to protein aggregation in the gonad and indeed, we found that long-lived animals with gonads have proportionally more insoluble proteins compared with wild-type animals with gonads (Figure S1D). These results suggest that aggregation in the gonad masks the protective effect of reduced *daf-2* signaling in somatic tissues. Importantly, we confirmed this protective action of reduced *daf-2* signaling with several candidates (see below).

Taken together, these data demonstrate that reduced *daf-2* signaling promotes protein solubility in most tissues with age.

(D) Proteins that aggregate in both control and *daf-2* RNAi conditions are enriched in extended stretches of β-sheet propensity. Unequal variance t test: proteins aggregating more with age in long-lived animals (n = 43, >1.2-fold), p = 0.002; proteins aggregating less with age in long-lived animals (n = 81, <0.83-fold), p = 0.09; proteins precipitated by b-isox (n = 126), p = 0.6.

(E) Large SDS-insoluble aggregates precipitated by 20,000 × *g*. Immunoblots detecting RBPs PAB-1, FIB-1, HRP-1 (unspecific band noted by asterisk), and CAR-1 (truncated band marked by asterisk). Relevant protein bands are indicated by red arrow.

(F) Flowchart showing a high overlap between proteins precipitated by b-isox and aggregation-prone proteins in *C. elegans*, in particular those no longer aggregating in *daf-2(–)* conditions.

See also Figures S1, S2, and S3 and Tables S1, S2, S3, and S4.

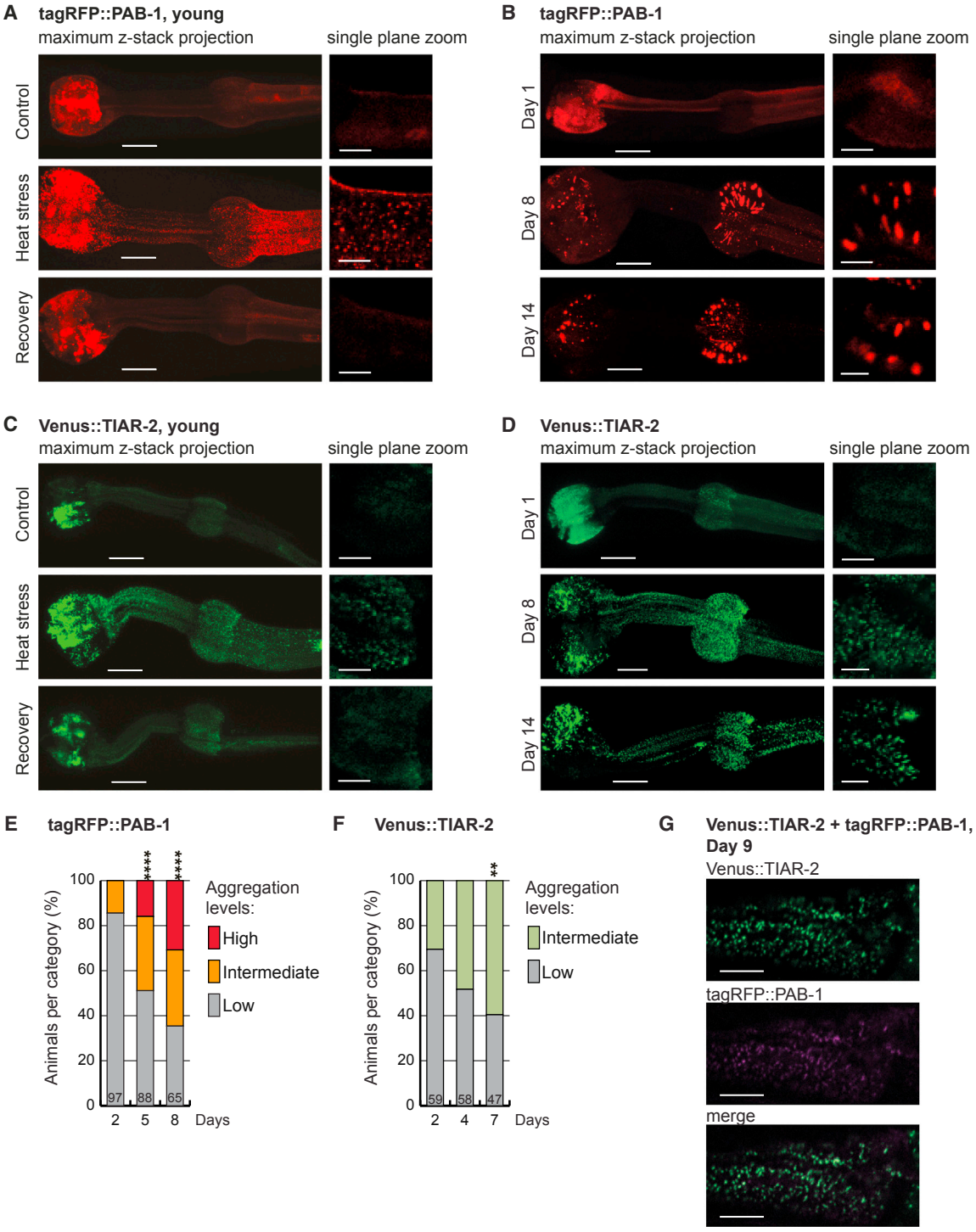


Figure 2. Stress-Granule-Related RBPs PAB-1 and TIAR-2 Aggregate with Age in *C. elegans*

(A) tagRFP::PAB-1 expressed in the pharyngeal muscles forms stress granules upon heat stress (2 hr, 32°C) on day 1 of adulthood. No stress granules are visible after recovery (+24 hr). Scale bars: z stack projection, 15 μm; single-plane insets, 5 μm.

(B) tagRFP::PAB-1 distribution changes from a diffuse pattern in young animals to a punctate pattern in aged worms. Scale bars: z stack projection, 15 μm; single-plane insets, 5 μm.

(C) Venus::TIAR-2 expressed in the pharyngeal muscles forms stress granules upon heat stress (2 hr, 32°C) on day 1 of adulthood. No stress granules are visible after recovery (+24 hr). Scale bars: z stack projection, 15 μm; single-plane insets, 5 μm.

(D) Venus::TIAR-2 accumulates mainly in stress-granule-like puncta with age. Scale bars: z stack projection, 15 μm; single-plane insets, 5 μm.

(legend continued on next page)

Reduced *daf-2* Signaling Preferentially Abrogates the Insolubility of RNA Granule Components

To investigate specifically changes in insolubility regulated by reduced *daf-2* signaling, we sorted proteins prone to aggregate with age in control animals (>2-fold) by their fold change in insolubility with age in the long-lived conditions. We restricted our analysis to 87 proteins that aggregated less with age (<0.83-fold) and 52 proteins that continued to aggregate with age (>1.2-fold) in the long-lived animals (Figure 1B; Table S1). Previous bioinformatics analysis of aggregation-prone proteins revealed an enrichment in both β-sheet propensity and aliphatic amino acids (David et al., 2010). A similar enrichment in the whole insoluble proteome was identified in this study (Figures S1E and S1F). Intriguingly, when examining the two groups of insoluble proteins that were differentially regulated by reduced *daf-2* signaling, we found a segregation of the two properties. Proteins with abrogated aggregation with the *daf-2* RNAi treatment were highly enriched in aliphatic amino acids, in particular alanine, glycine, and valines, but not in β sheets (Figures 1C, 1D, and S1G). Conversely, proteins that still aggregate in the *daf-2(−)* condition had a significant propensity to form β sheets but were only modestly enriched in aliphatic amino acids (Figures 1C and 1D).

Next, we searched for functional differences between the two groups differentially regulated by reduced *daf-2* signaling. Strikingly, ribosomal proteins and RNA granule components, including stress granule and P-granule RBPs, were highly overrepresented among the proteins that were prevented from aggregating with age in the long-lived animals (Tables S2A and S3). Conversely, chaperones and vitellogenin yolk proteins were overrepresented among proteins that were still prone to aggregate with age in *daf-2(−)* conditions (Table S2B). Among the RBPs, four are predicted to have LC prion-like domains (Alberti et al., 2009): PAB-1, FIB-1, HRP-1, and CAR-1 (Figure S2A). By western blot, we confirmed that reduced *daf-2* signaling abrogated their aggregation with age (Figures 1E and S2B). In addition, we evaluated two proteins without RNA-binding or LC prion-like domains, PAR-5 and DAF-21, quantified by mass spectrometry as more insoluble with age in both control and long-lived animals. We confirmed that PAR-5 and DAF-21 continued to aggregate with age in long-lived animals (by >7-fold and 10-fold, respectively), albeit to a reduced extent compared with controls (Figure S2C). Of note, changes in aggregation were not correlated with changes in total protein levels (Figure S3).

A previous study found that the chemical b-isox exclusively causes RNA granules to precipitate out from whole-cell lysates by inducing their assembly into a hydrogel (Han et al., 2012; Kato et al., 2012). Proteins precipitated by b-isox were also enriched in aliphatic amino acids (in particular glycine and to a lesser extent alanine and valine) and not in the propensity to form β sheets (Figures 1C, 1D, and S1G). We checked for proteins in common with our study and found a very significant overlap between proteins precipitated by b-isox and proteins no longer aggregating in the long-lived conditions (Figure 1F; Table S4).

Together, these results indicate that different types of RNA granule components including key RBPs responsible for their assembly become insoluble with age, and long-lived animals with reduced *daf-2* signaling are highly successful in preventing their aggregation. Interestingly, higher levels of aliphatic amino acids in RNA granule components could help their assembly and/or drive their age-dependent aggregation.

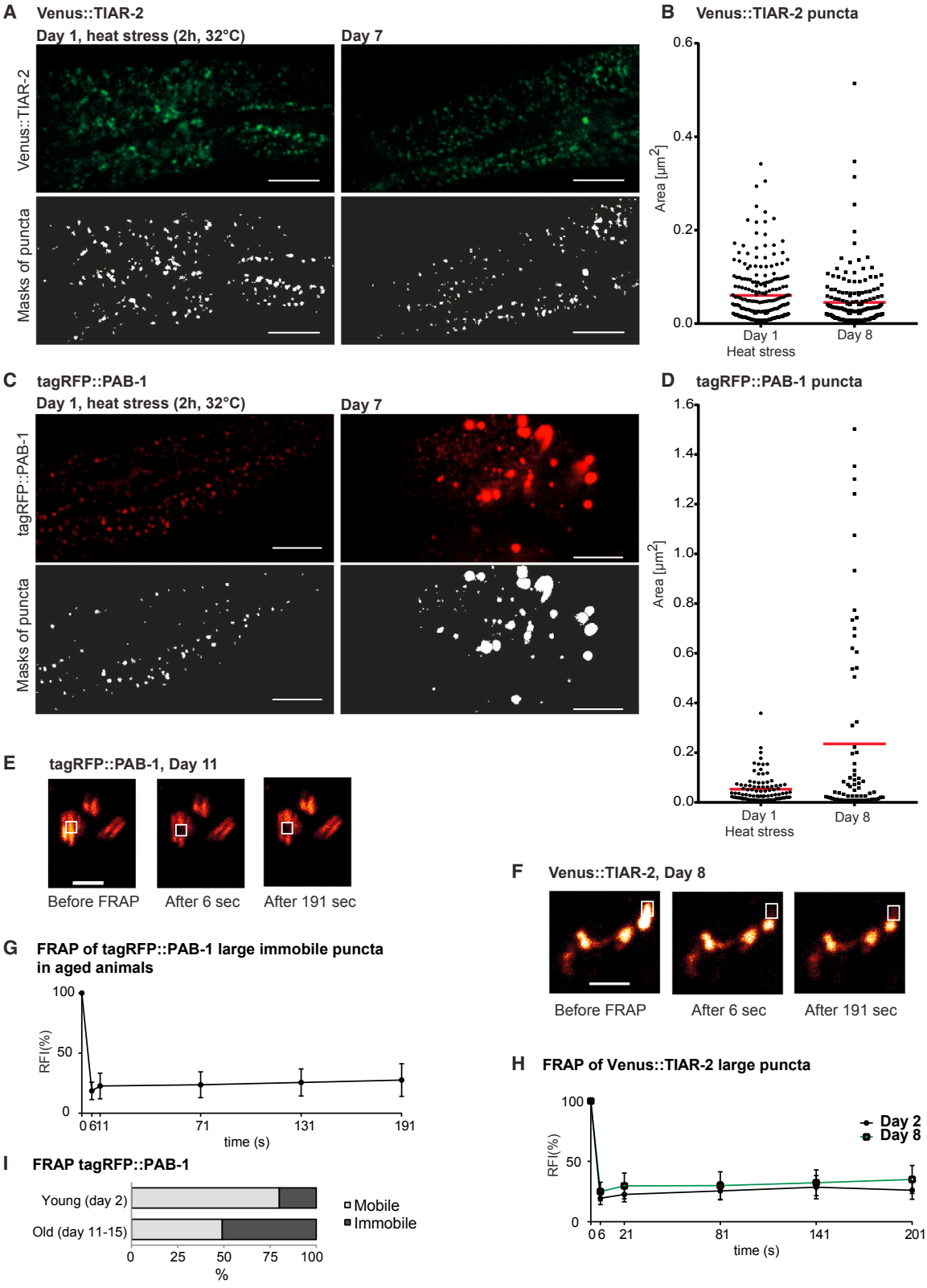
Key Stress Granule Proteins PAB-1 and TIAR-2 Form Solid Aggregates in Aged *C. elegans*

To investigate further the aggregation of key RBPs with LC prion-like domains and to understand the mechanisms involved, we generated *C. elegans* strains expressing PAB-1 and TIAR-2 in the pharyngeal muscles, fused to fluorescent tags. PAB-1 and TIAR-2 are the *C. elegans* homologs of human polyadenylate-binding protein 1 (PABP-1) and T-cell-restricted intracellular antigen-1 (TIA-1), two prominent RBPs that localize to stress granules and are also minor components of pathological inclusions that occur in ALS and FLTD (Bentmann et al., 2012). Both of these RBPs harbor LC prion-like domains (Figure S2A) (Alberti et al., 2009).

Exposing cells to stressors such as heat induces RBPs with LC prion-like domains to form liquid droplets (Molliex et al., 2015; Patel et al., 2015). Similarly, heat shock in *C. elegans* caused PAB-1 and TIAR-2 to form stress granules (Figures 2A and 2C; Figure S4A) (Murakami et al., 2012; Rousakis et al., 2014). When co-expressed, PAB-1 and TIAR-2 localized to the same stress granules (Figure S4B). Consistent with the dynamic nature of stress granules, these puncta were no longer observed 24 hr after the heat shock (Figures 2A and 2C). Using antibodies, we observed a similar change in pattern with endogenous PAB-1, indicating the effect is not merely due to overexpression or the fluorescent tag (Figure S4C). As an additional control, we showed that kinase KIN-19, which is not an RBP and lacks a LC prion-like domain, does not localize to stress granules upon heat shock (Figure S4D).

With age, we observed a striking change in the distribution pattern of these key stress granule RBPs. Whereas the majority of PAB-1 and TIAR-2 proteins were diffusely localized in non-stressed young animals, we found that both stress granule components accumulated in distinct puncta in aged animals (Figures 2B and 2D; Figure S4F). This change was specific for PAB-1 and TIAR-2 as the fluorescent tags alone (Venus or tagRFP) remained diffuse with age (Figure S4E) (David et al., 2010). The results were not simply caused by overexpression because endogenous PAB-1 formed similar puncta with age, in the different head regions where it is natively expressed (Figure S4G). For both PAB-1 and TIAR-2, we observed a significant increase with

(E) Increased tagRFP::PAB-1 aggregation with age in a population of *C. elegans*. Day 5 or 8 versus day 2: ****$p < 0.0001$.
(F) Increased Venus::TIAR-2 aggregation with age in a population of *C. elegans*. Day 7 versus day 2: **$p < 0.01$.
(G) tagRFP::PAB-1 co-localizes with stress-granule-like Venus::TIAR-2 puncta in double-transgenic animals. Representative single-plane images. Scale bar, 7 μm.
See also Figure S4.



(legend on next page)

age in the number of worms with large puncta that are visible by low-magnification microscopy (Figures 2E and 2F; Figure S4H). Interestingly, imaging at high magnification revealed that TIAR-2 in aged animals was localized predominantly in small puncta highly reminiscent of stress granules assembled during heat shock, whereas PAB-1 accumulated to a greater extent in large puncta that were not normally observed upon heat shock in young animals (Figures 2B, 2D, single-plane insets, 3A, 3C, and 3D). Quantification confirmed that stress granules induced during heat shock and aberrant TIAR-2 stress granule-like puncta formed during aging had similar sizes (Figure 3B). Remarkably, when co-expressed, PAB-1 preferentially co-localized within TIAR-2-positive age-dependent stress-granule-like puncta (Figure 2G), suggesting that interactions between RBPs change their aggregation patterns.

A hallmark of protein aggregation associated with disease is the immobility of proteins within the aggregates. To evaluate this aspect, we monitored fluorescence recovery after photobleaching (FRAP) in large PAB-1 and TIAR-2 puncta (Figures 3E–3I). All large TIAR-2 puncta and half of the large PAB-1 puncta in aged animals showed no fluorescence recovery, demonstrating that these are solid aggregates (Figures 3E–3I). Because of the small size of the age-dependent stress-granule-like puncta, it was not possible to assess the mobility of TIAR-2 or PAB-1 in these structures. Consequently, we used the large puncta as a readout for sgRBP aggregation in subsequent experiments.

Aggregation of Key Stress Granule Component PAB-1 Is Associated with Reduced Fitness

The consequences of age-dependent protein aggregation for the animal's health are poorly understood. Here, we evaluated the impact of the aggregation of a key stress granule component during aging. Surprisingly, PAB-1 overexpression was protective as animals, grown at 15°C, 20°C, or 25°C all their life, lived longer than non-transgenic animals (Table S5). However, this effect is likely due to higher levels of functional PAB-1 and not related to protein aggregation because PAB-1 does not aggregate in animals at 15°C (Figure S5A). In order to distinguish effects specifically related to PAB-1 aggregation, we separated PAB-1 transgenics at day 7 into three groups depending on their aggregation levels. We found that animals with PAB-1 aggregation were significantly smaller in size compared with animals without PAB-1 aggregation (Figure 4A; Figure S5B). In addition, smaller animals were visibly less motile (Movie S1). Importantly, mildly stressed animals with the highest levels of aggregation died earlier (Figure 4B; Table S5). Overall, these results demonstrate that PAB-1 aggregation is associated with impaired health.

HSF-1 Activity during Development Protects against sgRBP Aggregation

Our proteomic study revealed that reduced *daf-2* signaling efficiently prevents the insolubility of stress granule components with age. We confirmed that long-lived *daf-2* mutants greatly delay the formation of both stress-granule-like structures and large aggregates of PAB-1 and TIAR-2 (Figures 5A and 5B; Figures S6A and S6B). To gain insight into the mechanisms controlling protein aggregation, we investigated the role of the transcription factor HSF-1 activated by reduced *daf-2* signaling (Hsu et al., 2003; Volovik et al., 2012). Chaperones HSP110, HSP70, and HSP40 modulate stress granule dynamics in *Saccharomyces cerevisiae* and/or in cell culture (Cherkasov et al., 2013; Gilks et al., 2004; Kroschwald et al., 2015; Walters et al., 2015). Because chaperone expression is controlled by HSF-1 in *C. elegans*, we speculated that HSF-1 may regulate sgRBP aggregation. Indeed, impairing HSF-1 activity in both *daf-2(−)* and wild-type backgrounds caused severe PAB-1 aggregation already in young adults as well as in aged individuals, an effect that was reversed by overexpressing HSF-1 (Figures 5C–5E; Figures S6C and S6D). Interestingly, HSF-1 in *daf-2* mutants did not control the aggregation of the kinase KIN-19, which was previously shown to misfold and form solid aggregates with age (Figure 5G) (David et al., 2010). These results suggest that HSF-1 regulates different types of endogenous protein aggregation with age to different extents. It was previously shown that to assure *daf-2(−)* longevity, HSF-1 is most highly expressed and acts mainly during development (Volovik et al., 2012). We observed that impairing HSF-1 activity by RNAi during adulthood had no effect on PAB-1 aggregation (Figure 5F; Figure S6E). Conversely, reducing HSF-1 activity by RNAi during development caused PAB-1 aggregation in young adults, albeit mainly in the anterior bulb (Figure 5F).

Figure 3. TIAR-2 and PAB-1 Accumulate in Stress-Granule-like Puncta and Large Immobile Puncta in Aged *C. elegans*

(A) Small Venus::TIAR-2 puncta formed with age are similar to stress granules formed during heat stress. Representative single-plane images and masks of puncta for size quantification. Scale bars, 5 μm.

(B) Size quantification of Venus::TIAR-2 puncta from masks of representative single-plane images in (A).

(C) Representative single-plane images and masks of puncta for size quantification showing large tagRFP::PAB-1 puncta formed with age compared with stress granules assembled during heat stress. Scale bars, 5 μm.

(D) Size quantification of tagRFP::PAB-1 puncta from masks of representative single-plane images in (C). Puncta larger than 0.5 μm^2 were considered as "large" puncta.

(E) Representative immobile tagRFP::PAB-1 puncta at day 11 assayed by FRAP. Bleached area is marked by white box. Scale bar, 4 μm.

(F) Representative immobile Venus::TIAR-2 puncta at day 8 assayed by FRAP. Bleached area is marked by white box. Scale bar, 2 μm.

(G) FRAP analysis of immobile tagRFP::PAB-1 puncta present in aged worms (days 11–12). Quantification of relative fluorescence intensity (RFI) over time. Number of animals = 6, puncta evaluated = 6, mean ± SD is represented.

(H) Venus::TIAR-2 puncta monitored by FRAP were highly immobile both in young (day 2) and in aged (day 8) animals. In both young and aged animals: animals = 5, puncta evaluated = 5, mean ± SD is represented.

(I) Quantification of FRAP results shows increased immobility of tagRFP::PAB-1 puncta with age. Twenty percent of tagRFP::PAB-1 puncta present in young worms (day 2) were immobile (number of animals = 18, puncta evaluated = 57) compared with 51% in aged worms (days 11–15) (number of animals = 18, puncta evaluated = 55).

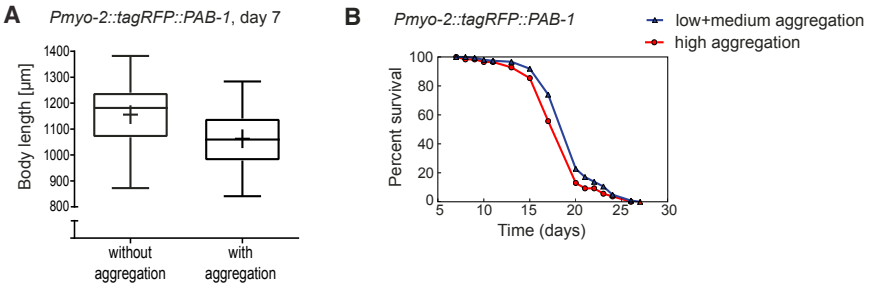


Figure 4. Aggregation of Stress Granule Component PAB-1 with Age Is Associated with Reduced Fitness

(A) Animals with tagRFP::PAB-1 aggregation are significantly smaller than animals without aggregation (day 7, $p < 0.0001$). Data are represented with Tukey-style box plots and mean indicated by + (animals without aggregation n = 99, with aggregation n = 125). See also Figure S5B.

(B) High levels of tagRFP::PAB-1 aggregation are associated with reduced survival. Survival curve of *Pmyo-2::tagRFP::PAB-1* animals grown at 20°C until day 7, sorted by their aggregation levels at day 7, and then transferred to 25°C (repeat 1: $p = 0.029$; see Table S5).

See also Figure S5, Table S5, and Movie S1.

The same chaperones discovered to regulate stress granule dynamics in other model systems could also play a role in preventing sgRBP aggregation. In yeast the Hsp40 proteins Sis1 and Ydj1 were shown to co-localize with stress granules and to play a role in stress granule disassembly (Walters et al., 2015). We evaluated worm strains overexpressing yellow fluorescent protein (YFP)-tagged DNJ-13 and DNJ-19, the worm orthologs of Sis1 and Ydj1, together with tagRFP::PAB-1. We observed occasional co-localization of both chaperones with both heat-induced PAB-1 stress granules and age-dependent large and stress-granule-like PAB-1 puncta (Figures S6G and S6H). These findings were confirmed in single tagRFP::PAB-1 transgenics using antibodies against DNJ-13 and DNJ-19 (data not shown). However, overexpression of DNJ-19 and DNJ-13 did not significantly reduce PAB-1 aggregation with age (Figures S6F and S6I), indicating that these chaperones may not modulate sgRBP aggregation. In addition, we performed immunostaining for HSP110. However, we observed no co-staining with either PAB-1 stress granules induced by heat shock or PAB-1 puncta formed during aging. Therefore, it is possible that HSP110 does not modulate stress granule dynamics and sgRBP aggregation in *C. elegans*. Finally, we investigated the role of HSP70 on sgRBP aggregation and found that inhibition of HSP70 by RNAi did not alter PAB-1 aggregation (Table S6). Of note, the later results could be caused by redundancy between chaperone functions.

Collectively, these results reveal that HSF-1 is an important regulator of sgRBP aggregation throughout life, and it contributes to maintaining dynamic stress granule proteins in long-lived *daf-2* mutants. The exact chaperones responsible for preventing sgRBP aggregation remain to be determined.

daf-2 Mutants Avoid sgRBP Aggregation in Part by Eliminating Putative Cross-Seeding

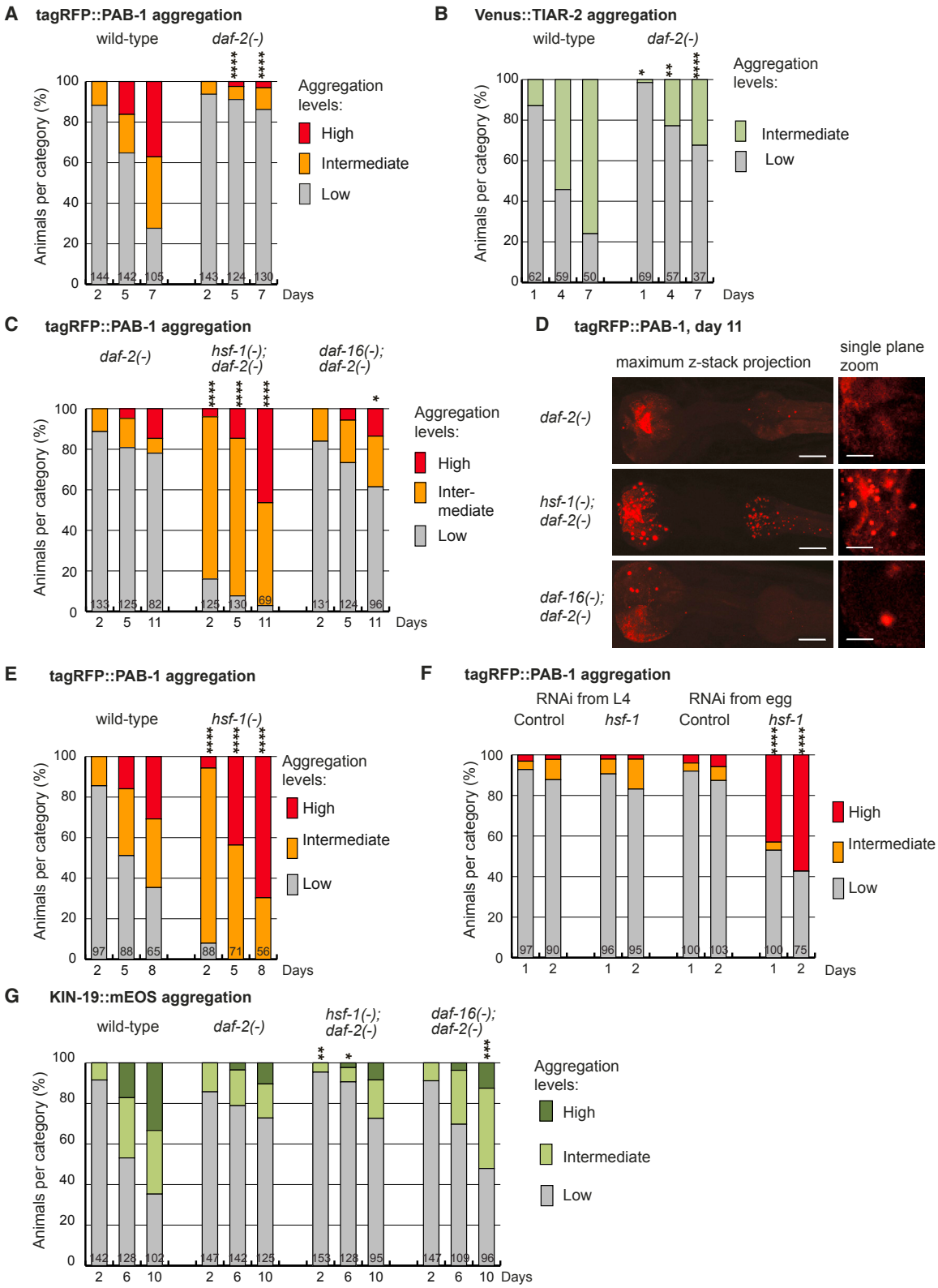
Next, we examined the role of the transcription factor DAF-16 activated by reduced *daf-2* signaling (Lin et al., 1997; Ogg et al., 1997). In *daf-2* mutants, DAF-16 protected against PAB-1 aggregation in aged animals (Figures 5C and 5D). Importantly, *daf-2* mutants use DAF-16 to control different types of age-dependent protein aggregation because delayed aggregation of the kinase KIN-19 was also dependent on DAF-16 (Figure 5G).

Recent work performed in *S. cerevisiae* and *Drosophila* reveals that stress granules dynamically interact with misfolded proteins (Cherkasov et al., 2013; Kroschwald et al., 2015). We speculated that widespread protein misfolding and aggregation occurring with age could promote sgRBP aggregation. In support of this hypothesis, we observed the co-localization of PAB-1 and KIN-19 in large immobile aggregates in double transgenics (Figures 6A and 6B). Next, we evaluated the rate of PAB-1 and KIN-19 aggregation in the double transgenics relative to their aggregation rates in single transgenics (Figures 6C and 6D). In the single transgenics, KIN-19 aggregated faster and to a greater extent than PAB-1. Importantly, when both PAB-1 and KIN-19 were co-expressed, the presence of misfolded KIN-19 triggered more abundant PAB-1 aggregation at an earlier age (Figure 6C). Conversely, KIN-19 aggregation was slightly impeded by PAB-1 overexpression (Figure 6D), indicating that sgRBP aggregation does not cross-seed KIN-19 aggregation. Notably, PAB-1 aggregation was not accelerated by the co-expression of a fluorescent tag alone (Figure 6E). Furthermore, we did not observe a general induction of PAB-1 stress granules in young double-transgenic animals expressing both KIN-19 and PAB-1 (Figure S6J). These data strongly suggest that PAB-1 aggregation is not simply the consequence of generalized cellular stress induced by KIN-19 overexpression. Rather, the co-localization of KIN-19 and PAB-1 in the same aggregates as well as the earlier and accelerated aggregation of PAB-1 is consistent with a seeding mechanism related to KIN-19 aggregation.

Taken together, we interpret these findings to imply that the accumulation of misfolded proteins with age acts as a seed for sgRBP aggregation. Therefore, the overall reduction in widespread protein aggregation in long-lived *daf-2(–)* conditions as evidenced by our proteomic analysis, at least in part through increased DAF-16 activity, could be an effective strategy to prevent sgRBP aggregation.

Other Longevity Pathways Prevent Age-Dependent sgRBP Aggregation

Several experimental manipulations have been shown to extend the lifespan of *C. elegans* and protect against proteotoxicity (Kenyon, 2010; Taylor and Dillin, 2011). We wondered whether other pathways extending lifespan could also protect against sgRBP aggregation. We found that both dietary restriction mimicked in *eat-2* mutant animals as well as inhibition of mitochondrial function achieved by targeting *cyc-1* with RNAi strongly limited PAB-1 aggregation with age (Figures 7A and



(legend on next page)

7B). Therefore, maintaining dynamic RBPs could be a common strategy associated with longevity.

DISCUSSION

We show that a wide variety of RNA granule components become highly insoluble with age in *C. elegans*. Together with previous in vitro and cell culture results, our findings demonstrate that the capacity of RBPs to cycle between assembled and disassembled states can become a liability in aging organisms. Already in young animals, maintaining sgRBP dynamics necessitates an active control system established by HSF-1. The accumulation of other misfolded proteins during age acts as a seed for the aggregation of key sgRBPs. Significantly, one of the main outcomes of the longevity program initiated by reduced *daf-2* signaling while responding to widespread protein aggregation is the preservation of RBP solubility with age.

In this study, we have examined in detail the aggregation pattern of PAB-1 and TIAR-2, two key RBPs with LC prion-like domains that are important for the formation of stress granules. During the aging process, both proteins spontaneously assembled into small puncta similar to liquid droplets induced during stress and into larger aggregates. Significantly, upon co-expression, both PAB-1 and TIAR-2 co-localized in these age-related stress-granule-like structures. Because our proteomic analysis revealed a number of stress granule components in the insoluble proteome, it is likely that secondary stress granule proteins are also incorporated. Therefore, an attractive hypothesis is that these small puncta represent stress granules formed as a response to stress related to aging. The inherent aggregation propensity of sgRBPs would induce at least some of these droplets to undergo the irreversible transition into a solid state. These stabilized stress granules could then grow into large aggregates as we observed with PAB-1, or simply accumulate with age as seen with TIAR-2.

An important question remains how the inherent propensity of RNA granule components to aggregate with age could influence pathogenesis in neurodegenerative diseases.

One possibility is that inherent RBP aggregation impacts cellular health and thereby indirectly accelerates pathology. We observed reduced lifespan and a striking decrease in size and mobility of animals with higher levels of PAB-1 aggregation. As yet, it remains unclear whether reduced fitness is a cause or consequence of PAB-1 aggregation. In support of a gain of function related to sgRBP aggregation, two rare diseases are caused by mutated PABPN1 and TIA-1, the human homologs of PAB-1 and TIAR-2, which accumulate in pathological aggregates (Brais et al., 1998; Klar et al., 2013). Our proteomic analysis of aging *C. elegans* highlighted three other RBPs with LC prion-like domains that are highly prone to aggregate with age: HRP-1, FIB-1, and CAR-1. The aggregation of the human homologs of HRP-1, hnRNP-A1, and hnRNP-A3 was recently discovered to cause multisystem proteinopathy and ALS/FTLD (Kim et al., 2013; Mori et al., 2013). The aggregation of both FIB-1 and CAR-1 could also be detrimental. Indeed, a loss in nucleolar protein FIB-1 function caused by its aggregation with age could impair ribosomal biogenesis (Tollervey et al., 1991). The mammalian LSM14B and LSM14A homologs of CAR-1 localize to P bodies (Eulalio et al., 2007), and aggregation of key P-body components could impair non-sense-mediated decay. Therefore, it will be important to investigate the consequences of FIB-1 and CAR-1 aggregation on cellular health.

Apart from accelerating pathology indirectly by reducing cellular health, aggregating RBPs could directly influence pathological protein aggregation. The presence of stress granule proteins in pathological protein aggregates is emerging as a common denominator in different types of neurodegenerative diseases including ALS, FTLD, Alzheimer's disease, and Huntington's disease (Aulas and Vande Velde, 2015; Bentmann et al., 2013). The inherent aggregation propensity of stress granule proteins demonstrates that they are unlikely to be transient interacting partners in pathological aggregates. It remains to be determined whether age-related sgRBP aggregation acts as a seed for disease-associated protein aggregation. Overall, the role of stress granules in neurodegenerative diseases is clearly highly complex because there is evidence supporting the recruitment of disease-associated proteins to stress granules and vice versa. Interestingly, recent cell culture data show that the assembly of stress granules caused by disease-associated protein aggregation in turn promotes pathological aggregation (Vanderweyde et al., 2016).

The age-dependent aggregation of sgRBPs and prevalence of stress granule components in neurodegenerative diseases underline their relevance as therapeutic targets. One successful strategy would be to prevent the initial assembly of stress granules (Kim et al., 2014). Our work suggests another possibility,

Figure 5. HSF-1 Activity during Development Protects against PAB-1 Aggregation in Adulthood in *daf-2* Mutant and Wild-Type Adults

(A) Delayed tagRFP::PAB-1 aggregation with age in *daf-2* mutant background. Days 5 and 7, *daf-2(–)* versus wild-type background: ****p < 0.0001.

(B) Delayed Venus::TIAR-2 aggregation with age in *daf-2* mutant background. Days 1, 4 and 7, *daf-2(–)* versus wild-type background: *p < 0.05, **p < 0.01, and ****p < 0.0001, respectively.

(C) Levels of tagRFP::PAB-1 aggregation are highly increased at all ages examined in *hsf-1(–); daf-2(–)* animals compared with *daf-2(–)* animals. DAF-16 moderately protects against tagRFP::PAB-1 aggregation at day 11. *daf-2(–)* compared with *hsf-1(–);daf-2(–)*: ****p < 0.0001; *daf-2(–)* compared with *daf-16(–);daf-2(–)*: *p = 0.02.

(D) Head regions of representative animals expressing *Pmyo-2::tagRFP::PAB-1* in *daf-2(–), hsf-1(–); daf-2 (–)* and *daf-16(–); daf-2(–)* mutants at day 11. Scale bars: z stack projection, 15 μm; single-plane zoom, 5 μm.

(E) *hsf-1* mutation alone increases tagRFP::PAB-1 aggregation dramatically, even at day 2. ****p < 0.0001.

(F) HSF-1 activity during development is essential in order to delay tagRFP::PAB-1 aggregation (control: L4440 empty vector). Days 1 and 2 with RNAi treatment from egg, ****p < 0.0001.

(G) Delayed KIN-19::mEOS (monomeric EOS) aggregation with age in *daf-2* mutants is dependent on DAF-16, but not on HSF-1. *daf-2(–)* compared with *hsf-1(–); daf-2(–)*: day 2, **p = 0.0049; day 6, *p = 0.01. *daf-2(–)* compared with *daf-16(–);daf-2(–)*: ***p = 0.0003, day 10.

See also Figure S6 and Table S6.

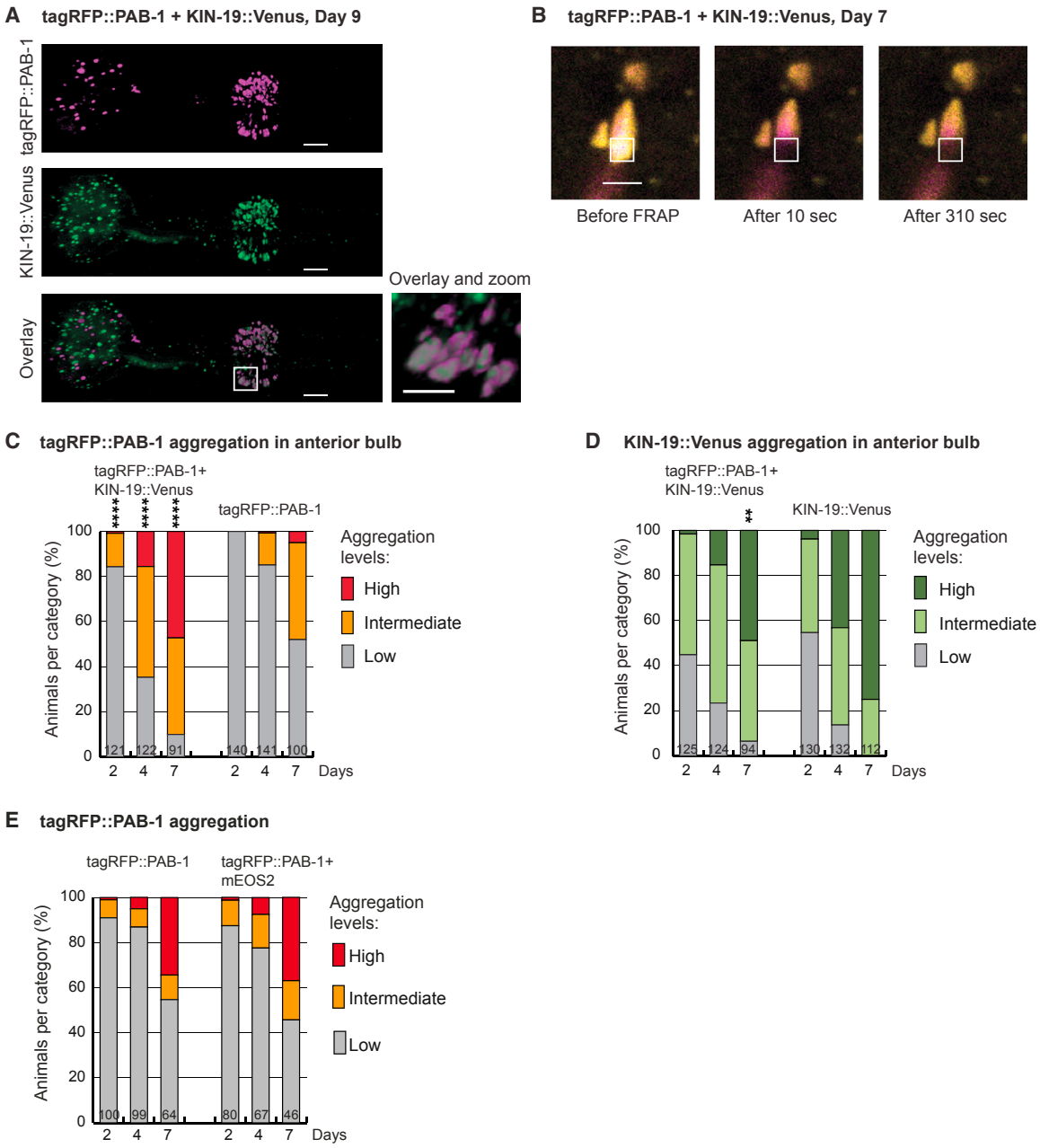


Figure 6. PAB-1 Aggregation Is Accelerated by KIN-19

(A) tagRFP::PAB-1 co-localizes with KIN-19::Venus in large aggregates in double-transgenic animals. Representative head region displayed in 3D. Scale bars, 10 μm; overlay and zoom scale bar, 5 μm.

(B) Representative immobile mixed tagRFP::PAB-1 (magenta) and KIN-19::Venus (yellow) puncta at day 7 assayed by FRAP. Bleached area is marked by white box. Scale bar, 2 μm.

(C) Accelerated tagRFP::PAB-1 aggregation in the anterior pharyngeal bulb in double transgenics compared with single transgenics. Days 2, 4, and 7: ****$p < 0.0001$.

(D) Moderately reduced KIN-19::Venus aggregation in the anterior pharyngeal bulb in double transgenics compared with single transgenics. Day 7: **$p = 0.0083$.

(E) No significant increase of tagRFP::PAB-1 aggregation in the presence of mEOS2 overexpression. At all ages, $p > 0.05$.

See also Figure S6.

namely abrogating sgRBP aggregation. In future work, it will be important to understand how longevity pathways relying on dietary restriction or defective mitochondrial respiration efficiently prevent sgRBP aggregation. In the case of reduced *daf-2* signaling, sgRBP aggregation is suppressed by at least two mechanisms: DAF-16 activation prevents cross-seeding by

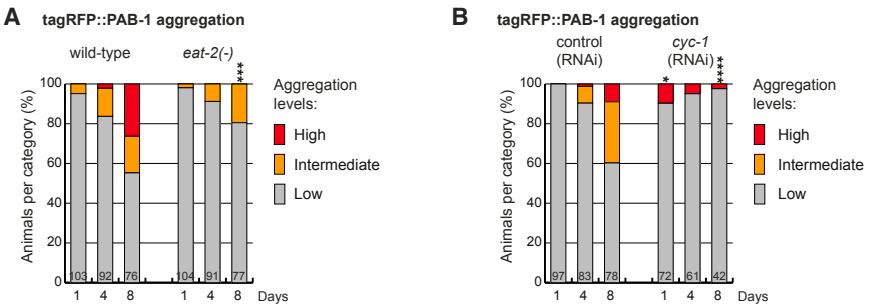


Figure 7. Other Longevity Pathways Prevent PAB-1 Aggregation with Age

(A) Dietary restriction delays tagRFP::PAB-1 aggregation with age. Day 8, *eat-2(–)* versus wild-type background: ***p < 0.001.

(B) Inhibition of mitochondrial function by *cyc-1* RNAi halts tagRFP::PAB-1 aggregation. Days 1 and 8, *cyc-1* RNAi versus control: *p < 0.05 and ****p < 0.0001, respectively.

delaying the accumulation of misfolded aggregation-prone proteins, and increased activity of HSF-1 during development assures enhanced sgRBP proteostasis throughout adulthood. To date, cross-seeding of RBP aggregation has been observed only with the Huntingtin protein containing an expanded polyglutamine repeat region (Furukawa et al., 2009). It will be important to investigate by which means endogenous aggregation-prone proteins lacking motifs similar to the LC prion-like domain would cross-seed RBPs aggregation. Overall, both of these strategies used to restore the dynamic nature of stress granule proteins could also directly prevent disease-associated RBP aggregation.

EXPERIMENTAL PROCEDURES

A list of strains, strain maintenance, RNAi treatment, and lifespan assays is described in the Supplemental Experimental Procedures.

Heat Shock

Nematodes were heat shocked in M9 medium with OP50 on a nutator. Control worms were treated the same at 20°C. Worms were either fixed for imaging analysis directly after heat shock or allowed to recover on normal growth (NG) plates kept at 20°C before fixation.

Size Measurements

Images were taken of synchronized live *C. elegans* using a Leica fluorescence microscope M165 FC with a Planapo 2.0× objective and Leica DFC310 FX camera. Body length was determined for each worm using Fiji software (Schindelin et al., 2012). Significance was evaluated by unpaired, two-tailed t test using GraphPad Prism 6.

Imaging and Immunofluorescence Staining

Worms were examined with a Leica SP8 confocal microscope. Fixation and immunostaining protocol including imaging parameters are described in the Supplemental Experimental Procedures. Representative confocal images are displayed as maximum z stack projection unless mentioned otherwise.

Aggregation Quantification In Vivo

In a population of synchronized live *C. elegans*, aggregation levels were determined using a Leica fluorescence microscope M165 FC with a Planapo 2.0× objective. Animals overexpressing *Pmyo-2::Venus::TIAR-2* were divided into two categories: animals with up to 10 (low aggregation) or more than 10 Venus::TIAR-2 puncta (intermediate aggregation) in the anterior and posterior pharyngeal bulb. Animals overexpressing *Pmyo-2:tagRFP:PAB-1* were divided into three categories: animals with up to 10 (low aggregation) or more than 10 (intermediate aggregation) tagRFP::PAB-1 puncta in the posterior bulb, or more than 10 (high aggregation) tagRFP:PAB-1 puncta in the anterior bulb. The latter mostly had more than 10 puncta in the posterior bulb. Animals expressing *Pkin-19::KIN-19::mEOS* or *Pkin-19::KIN-19::Venus* were divided into less than 10 puncta (low aggregation), between 10 and 100 puncta (intermediate aggregation), and more than 100 puncta in the anterior bulb (high aggregation). To compare aggregation levels between KIN-19::Venus and tagRFP::PAB-1, we evaluated tagRFP::PAB-1 puncta formation only in the anterior bulb and with the same counting scheme as for KIN-19::Venus. When comparing different strains, counting was done in a blind fashion. For statistics, two-tailed Fisher's exact test was performed using an online tool (http://www.socscistatistics.com/tests/fisher/default2.aspx). Significance of high + intermediate against low aggregation levels was calculated unless indicated otherwise. Numbers of animals per time point are indicated in the graphic bars.

FRAP Analysis

FRAP analysis was performed as previously described (David et al., 2010) using the Leica SP8 confocal microscope with the harmonic compound, plan, apochromatic (HC PL APO) CS2 63× 1.30 glycerol objective and photomultiplier tube (PMT) detector. Further experimental details are described in the Supplemental Experimental Procedures.

Insoluble Protein Extraction, Quantification, and Mass Spectrometry Analysis

To obtain large synchronized populations of aged animals and quantify protein aggregation only in the somatic tissues, we used temperature-induced sterile *gon-2* mutants as previously described (David et al., 2010). To induce longevity, we subjected animals to *daf-2* RNAi and control animals to *gfp* RNAi from the last larval stage L4 onward. Animals aged at 25°C were collected at day 3 of adulthood (young) and when half the control animals had died between days 14 and 18 (aged). No significant death was observed in long-lived animals on *daf-2* RNAi. Isolation of large SDS insoluble aggregates for immunoblot and mass spectrometry analysis were performed as previously described (David et al., 2010). See Supplemental Experimental Procedures for further details and antibodies used for immunoblots. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (Vizcaíno et al., 2014) with the dataset identifier PXD003451.

Bioinformatics Analysis

Aliphatic amino acid residues were defined as A, G, I, L, and V. Secondary structure content was predicted using PSIPRED v2.6 (Jones, 1999). The p values were calculated using the unequal variance t test compared with the background set of all proteins detected by mass spectrometry (n = 5,637). Additional details and functional analysis are described in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for the mass spectrometry proteomics data reported in this paper is PRIDE: PXD003451.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, six tables, and one movie and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2016.12.033.

AUTHOR CONTRIBUTIONS

M.C.L., N.G., J.C.T., and D.C.D. designed and performed experiments. E.D.C. performed bioinformatics analysis. K.W., R.J., and J.K. generated reagents. J.K. performed chaperone immunostainings. A.L.B. provided analytical tools, reagents, and materials. M.C.L., J.C.T., and D.C.D. wrote the paper.

ACKNOWLEDGMENTS

We are grateful to Cynthia Kenyon for help in the early stages of this project. We thank Aimee Kao and Sivan Henis-Korenblit for critical input, David Maltby for mass spectrometry support, Aenoch Lynn for help with the proteomic data analysis, Brian Lee for basic gateway constructs, and Kristin Arnsburg for integrating DNJ-13 and DNJ-19 transgenics. We are grateful to Simon Alberti for sharing the *C. elegans* list of proteins with prion-like domains. For the generous donation of antibodies, we would like to thank Rafal Ciosk (anti-PAB-1), Junho Lee (anti-HRP-1), and Keith Blackwell (anti-CAR-1). The *gon-2* mutants were kindly provided by Eric Lambie. This work was initiated in Cynthia Kenyon's lab and was funded by an Ellison/AFAR postdoctoral fellowship (to D.C.D.) and by the NIH (NIGMS grant 8P41GM103481 to A.L.B. and grant P50 GM081879 to J.C.T. and A.L.B.). Subsequently, this work was supported by funding from the DZNE and a Marie Curie International Reintegration Grant (322120 to D.C.D.).

Received: August 11, 2016
Revised: October 28, 2016
Accepted: December 12, 2016
Published: January 10, 2017

REFERENCES

Alberti, S., Halfmann, R., King, O., Kapila, A., and Lindquist, S. (2009). A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. Cell *137*, 146–158.

Arai, T., Hasegawa, M., Akiyama, H., Ikeda, K., Nonaka, T., Mori, H., Mann, D., Tsuchiya, K., Yoshida, M., Hashizume, Y., and Oda, T. (2006). TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Biochem. Biophys. Res. Commun. *351*, 602–611.

Aulas, A., and Vande Velde, C. (2015). Alterations in stress granule dynamics driven by TDP-43 and FUS: a link to pathological inclusions in ALS? Front. Cell. Neurosci. *9*, 423.

Ayyadevara, S., Mercanti, F., Wang, X., Mackintosh, S.G., Tackett, A.J., Prayaga, S.V., Romeo, F., Shmookler Reis, R.J., and Mehta, J.L. (2016). Age- and hypertension-associated protein aggregates in mouse heart have similar proteomic profiles. Hypertension *67*, 1006–1013.

Balch, W.E., Morimoto, R.I., Dillin, A., and Kelly, J.W. (2008). Adapting proteostasis for disease intervention. Science *319*, 916–919.

Bentmann, E., Neumann, M., Tahirovic, S., Rodde, R., Dormann, D., and Haass, C. (2012). Requirements for stress granule recruitment of fused in sarcoma (FUS) and TAR DNA-binding protein of 43 kDa (TDP-43). J. Biol. Chem. *287*, 23079–23094.

Bentmann, E., Haass, C., and Dormann, D. (2013). Stress granules in neurodegeneration–lessons learnt from TAR DNA binding protein of 43 kDa and fused in sarcoma. FEBS J. *280*, 4348–4370.

Brais, B., Bouchard, J.P., Xie, Y.G., Rochefort, D.L., Chrétien, N., Tomé, F.M., Lafrenière, R.G., Rommens, J.M., Uyama, E., Nohira, O., et al. (1998). Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. Nat. Genet. *18*, 164–167.

Cherkasov, V., Hofmann, S., Druffel-Augustin, S., Mogk, A., Tyedmers, J., Stoecklin, G., and Bukau, B. (2013). Coordination of translational control and protein homeostasis during severe heat stress. Curr. Biol. *23*, 2452–2462.

David, D.C. (2012). Aging and the aggregating proteome. Front. Genet. *3*, 247.

David, D.C., Ollikainen, N., Trinidad, J.C., Cary, M.P., Burlingame, A.L., and Kenyon, C. (2010). Widespread protein aggregation as an inherent part of aging in C. elegans. PLoS Biol. *8*, e1000450.

Demontis, F., and Perrimon, N. (2010). FOXO/4E-BP signaling in Drosophila muscles regulates organism-wide proteostasis during aging. Cell *143*, 813–825.

Eulalio, A., Behm-Ansmant, I., and Izaurralde, E. (2007). P bodies: at the crossroads of post-transcriptional pathways. Nat. Rev. Mol. Cell Biol. *8*, 9–22.

Furukawa, Y., Kaneko, K., Matsumoto, G., Kurosawa, M., and Nukina, N. (2009). Cross-seeding fibrillation of Q/N-rich proteins offers new pathomechanism of polyglutamine diseases. J. Neurosci. *29*, 5153–5162.

Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L.M., and Anderson, P. (2004). Stress granule assembly is mediated by prion-like aggregation of TIA-1. Mol. Biol. Cell *15*, 5383–5398.

Goudeau, J., and Aguilaniu, H. (2010). Carbonylated proteins are eliminated during reproduction in C. elegans. Aging Cell *9*, 991–1003.

Han, T.W., Kato, M., Xie, S., Wu, L.C., Mirzaei, H., Pei, J., Chen, M., Xie, Y., Allen, J., Xiao, G., and McKnight, S.L. (2012). Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. Cell *149*, 768–779.

Hsu, A.L., Murphy, C.T., and Kenyon, C. (2003). Regulation of aging and age-related disease by DAF-16 and heat-shock factor. Science *300*, 1142–1145.

Johnson, B.S., Snead, D., Lee, J.J., McCaffery, J.M., Shorter, J., and Gitler, A.D. (2009). TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. J. Biol. Chem. *284*, 20329–20339.

Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. J. Mol. Biol. *292*, 195–202.

Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J., et al. (2012). Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell *149*, 753–767.

Kenyon, C.J. (2010). The genetics of ageing. Nature *464*, 504–512.

Kim, H.J., Kim, N.C., Wang, Y.D., Scarborough, E.A., Moore, J., Diaz, Z., MacLea, K.S., Freibaum, B., Li, S., Molliex, A., et al. (2013). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. Nature *495*, 467–473.

Kim, H.J., Raphael, A.R., LaDow, E.S., McGurk, L., Weber, R.A., Trojanowski, J.Q., Lee, V.M., Finkbeiner, S., Gitler, A.D., and Bonini, N.M. (2014). Therapeutic modulation of eIF2α phosphorylation rescues TDP-43 toxicity in amyotrophic lateral sclerosis disease models. Nat. Genet. *46*, 152–160.

King, O.D., Gitler, A.D., and Shorter, J. (2012). The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. Brain Res. *1462*, 61–80.

Klar, J., Sobol, M., Melberg, A., Mäbert, K., Ameur, A., Johansson, A.C., Feuk, L., Entesarian, M., Orlén, H., Casar-Borota, O., and Dahl, N. (2013). Welander distal myopathy caused by an ancient founder mutation in TIA1 associated with perturbed splicing. Hum. Mutat. *34*, 572–577.

Kroschwald, S., Maharana, S., Mateju, D., Malinovska, L., Nüske, E., Poser, I., Richter, D., and Alberti, S. (2015). Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. eLife *4*, e06807.

Li, Y.R., King, O.D., Shorter, J., and Gitler, A.D. (2013). Stress granules as crucibles of ALS pathogenesis. J. Cell Biol. *201*, 361–372.

Lin, K., Dorman, J.B., Rodan, A., and Kenyon, C. (1997). daf-16: an HNF-3/forkhead family member that can function to double the life-span of Caenorhabditis elegans. Science *278*, 1319–1322.

Lin, Y., Protter, D.S., Rosen, M.K., and Parker, R. (2015). Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. Mol. Cell *60*, 208–219.

Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J., Mittag, T., and Taylor, J.P. (2015). Phase separation by low complexity domains

promotes stress granule assembly and drives pathological fibrillization. Cell *163*, 123–133.

Mori, K., Lammich, S., Mackenzie, I.R., Forné, I., Zilow, S., Kretzschmar, H., Edbauer, D., Janssens, J., Kleinberger, G., Cruts, M., et al. (2013). hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with C9orf72 mutations. Acta Neuropathol. *125*, 413–423.

Murakami, T., Yang, S.P., Xie, L., Kawano, T., Fu, D., Mukai, A., Bohm, C., Chen, F., Robertson, J., Suzuki, H., et al. (2012). ALS mutations in FUS cause neuronal dysfunction and death in Caenorhabditis elegans by a dominant gain-of-function mechanism. Hum. Mol. Genet. *21*, 1–9.

Murakami, T., Qamar, S., Lin, J.Q., Schierle, G.S., Rees, E., Miyashita, A., Costa, A.R., Dodd, R.B., Chan, F.T., Michel, C.H., et al. (2015). ALS/FTD mutation-induced phase transition of FUS liquid droplets and reversible hydrogels into irreversible hydrogels impairs RNP granule function. Neuron *88*, 678–690.

Neumann, M., Sampathu, D.M., Kwong, L.K., Truax, A.C., Micsenyi, M.C., Chou, T.T., Bruce, J., Schuck, T., Grossman, M., Clark, C.M., et al. (2006). Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science *314*, 130–133.

Neumann, M., Rademakers, R., Roeber, S., Baker, M., Kretzschmar, H.A., and Mackenzie, I.R. (2009). A new subtype of frontotemporal lobar degeneration with FUS pathology. Brain *132*, 2922–2931.

Neumann, M., Bentmann, E., Dormann, D., Jawaid, A., DeJesus-Hernandez, M., Ansorge, O., Roeber, S., Kretzschmar, H.A., Munoz, D.G., Kusaka, H., et al. (2011). FET proteins TAF15 and EWS are selective markers that distinguish FTLD with FUS pathology from amyotrophic lateral sclerosis with FUS mutations. Brain *134*, 2595–2609.

Ogg, S., Paradis, S., Gottlieb, S., Patterson, G.I., Lee, L., Tissenbaum, H.A., and Ruvkun, G. (1997). The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in C. elegans. Nature *389*, 994–999.

Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T.M., et al. (2015). A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. Cell *162*, 1066–1077.

Reis-Rodrigues, P., Czerwieniec, G., Peters, T.W., Evani, U.S., Alavez, S., Gaman, E.A., Vantipalli, M., Mooney, S.D., Gibson, B.W., Lithgow, G.J., and Hughes, R.E. (2012). Proteomic analysis of age-dependent changes in protein solubility identifies genes that modulate lifespan. Aging Cell *11*, 120–127.

Rousakis, A., Vlanti, A., Borbolis, F., Roumelioti, F., Kapetanou, M., and Syntichaki, P. (2014). Diverse functions of mRNA metabolism factors in stress defense and aging of Caenorhabditis elegans. PLoS ONE *9*, e103365.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Tanase, M., Urbanska, A.M., Zolla, V., Clement, C.C., Huang, L., Morozova, K., Follo, C., Goldberg, M., Roda, B., Reschiglian, P., and Santambrogio, L. (2016). Role of carbonyl modifications on aging-associated protein aggregation. Sci. Rep. *6*, 19311.

Taylor, R.C., and Dillin, A. (2011). Aging as an event of proteostasis collapse. Cold Spring Harb. Perspect. Biol. *3*, a004440.

Tollervey, D., Lehtonen, H., Carmo-Fonseca, M., and Hurt, E.C. (1991). The small nucleolar RNP protein NOP1 (fibrillarin) is required for pre-rRNA processing in yeast. EMBO J. *10*, 573–583.

Vanderweyde, T., Apicco, D.J., Youmans-Kidder, K., Ash, P.E., Cook, C., Lummertz da Rocha, E., Jansen-West, K., Frame, A.A., Citro, A., Leszyk, J.D., et al. (2016). Interaction of tau with the RNA-binding protein TIA1 regulates tau pathophysiology and toxicity. Cell Rep. *15*, 1455–1466.

Vizcaíno, J.A., Deutsch, E.W., Wang, R., Csordas, A., Reisinger, F., Ríos, D., Dianes, J.A., Sun, Z., Farrah, T., Bandeira, N., et al. (2014). ProteomeXchange provides globally coordinated proteomics data submission and dissemination. Nat. Biotechnol. *32*, 223–226.

Volovik, Y., Maman, M., Dubnikov, T., Bejerano-Sagie, M., Joyce, D., Kapernick, E.A., Cohen, E., and Dillin, A. (2012). Temporal requirements of heat shock factor-1 for longevity assurance. Aging Cell *11*, 491–499.

Walters, R.W., Muhlrad, D., Garcia, J., and Parker, R. (2015). Differential effects of Ydj1 and Sis1 on Hsp70-mediated clearance of stress granules in Saccharomyces cerevisiae. RNA *21*, 1660–1671.

Walther, D.M., Kasturi, P., Zheng, M., Pinkert, S., Vecchi, G., Ciryam, P., Morimoto, R.I., Dobson, C.M., Vendruscolo, M., Mann, M., and Hartl, F.U. (2015). Widespread proteome remodeling and aggregation in aging C. elegans. Cell *161*, 919–932.

Zimmerman, S.M., Hinkson, I.V., Elias, J.E., and Kim, S.K. (2015). Reproductive aging drives protein accumulation in the uterus and limits lifespan in C. elegans. PLoS Genet. *11*, e1005725.

Prion, 11:313–322, 2017
Published with license by Taylor & Francis
ISSN: 1933-68961933-690X online
DOI: 10.1080/19336896.2017.1356559

EXTRA VIEWS

More stressed out with age? Check your RNA granule aggregation

Marie C. Lechler[a,b] and Della C. David[a]

[a]German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany;
[b]Graduate School of Cellular & Molecular Neuroscience, Tübingen, Germany

ABSTRACT. Low complexity (LC) prion-like domains are over-represented among RNA-binding proteins (RBPs) and contribute to the dynamic nature of RNA granules. Importantly, several neurodegenerative diseases are characterized by cytoplasmic "solid" aggregates formed by mainly nuclear RBPs harboring LC prion-like domains. Although RBP aggregation in disease has been extensively characterized, it remains unknown how the process of aging disturbs RBP dynamics. Our recent study revealed that RNA granule components including 2 key stress granule RBPs with LC prion-like domains, PAB-1 and TIAR-2, aggregate in aged *Caenorhabditis elegans* in the absence of disease. Here we present new evidence showing that sustained stress granule formation triggers RBP aggregation. In addition, we demonstrate that mild chronic stress during aging promotes mislocalization of nuclear RBPs. We discuss the consequences of aberrant interactions between age-related RBP aggregation and disease-associated RBP aggregation. In particular, we show that FUST-1 and PAB-1 co-localize in aberrant cytoplasmic accumulations. Significantly, long-lived animals with reduced insulin/IGF-1 signaling abrogate stress granule RBP aggregation through activation of the transcription factors HSF-1 and DAF-16. We evaluate the different mechanisms that could maintain dynamic

Correspondence to: Della C. David; German Center for Neurodegenerative Diseases (DZNE), Otfried-Müller-Str. 23, D-72076 Tübingen, Germany; Email: della.david@dzne.de

Received June 7, 2017; Revised July 6, 2017; Accepted July 10, 2017

Extra View to: Lechler MC, Crawford ED, Groh N, Widmaier K, Jung R, Kirstein J, Trinidad JC, Burlingame AL, David DC. Reduced insulin/IGF-1 signaling restores the dynamic properties of key stress granule proteins during aging. Cell Rep 2017; 10;18(2):454-467; https://doi.org/10.1016/j.celrep.2016.12.033

stress granules. Together these findings highlight how changes with age could contribute to pathogenesis in neurodegenerative diseases and disruption of RNA homeostasis.

KEYWORDS. RNA-binding proteins, protein aggregation, prion-like domains, aging, stress granules

INTRODUCTION

A number of diseases including certain neurodegenerative disorders are characterized by the presence of pathological highly intractable or "solid" protein aggregates formed by one or several distinct proteins. In the last decade, the list of proteins identified in aggregates associated with disease has been considerably extended with the addition of several RNA-binding proteins (RBPs). These include RBPs such as TDP-43 and FUS observed in aggregates of patients with amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD).[1-3] RBPs associated with disease contain a low complexity (LC) "prion-like" domain similar to sequences identified in yeast prion proteins.[4,5] Highlighting the importance of this domain, familial cases of these diseases are frequently related to aggregation-promoting mutations in the LC prion-like domain.[6,7] In a non-disease context, RBPs with LC prion-like domains are key components of RNA granules. Depending on their composition, RNA granules such as stress granules, P-bodies, P-granules and neuronal granules perform different functions in the cell. RNA granules are highly dynamic membrane-less organelles and their assembly is mediated by association of RBPs through their LC prion-like domains and subsequent recruitment of RNA and associated proteins by RNA-binding domains.[8-10] In particular, weak interactions built between LC prion-like domains in RBPs promote a liquid-liquid phase separation in vitro consistent with the observed liquid-like proprieties of RNA granules in vivo.[11-14] Considering the special nature of the interactions between RBPs with LC prion-like domains and the growing number of RBPs forming hallmark aggregates in different neurodegenerative disorders, we hypothesized that aberrant aggregation of RBPs and RNA granules could be an important problem that the organism needs to actively avoid especially during aging.

RBPs with LC Prion-Like Domains and RNA Granule-Associated Proteins are Highly Prone to Aggregate with Age in C. elegans

Recently it has become clear that protein aggregation is not restricted to specific proteins in a disease context. Notably, we and others have shown that several hundred proteins are highly prone to aggregate during normal aging in different model organisms.[15-19] In our recent study,[20] the proteomic analysis of the aggregating proteome from long-lived *C. elegans* with reduced *daf-2*/insulin/IGF-1 receptor signaling highlighted a specific group of proteins namely RNA granule components. In particular we identified several stress granule components as well as 4 RBPs with LC prion-like domains that became highly insoluble with age in control animals but not in long-lived animals. These results imply that RNA granules or at least their core structures lose their dynamic characteristics with age. Significantly, preserving dynamic RBPs is a mechanism associated with longevity. We focused our main study on stress granule proteins as they are frequently found to co-localize with pathological TDP-43 and FUS inclusions and have been implied to play a role in pathogenesis.[21,22] We overexpressed 2 key stress-granule-related RBPs (sgRBPs) with LC prion-like domains, PAB-1, homolog of human polyadenylate-binding protein 1 (PABP-1) and TIAR-2, homolog of human T-cell-restricted intracellular antigen-1 (TIA-1) in *C. elegans* pharyngeal muscles.[20] Upon heat stress, both fluorescently-labeled proteins assembled into stress granules as expected. Importantly without exposing animals to additional stress, PAB-1 and TIAR-2 accumulated in stress-granule-like structures as well as large puncta in aged *C. elegans*. The absence of fluorescence recovery after photobleaching confirmed that the large PAB-1 and TIAR-2 puncta contained highly immobile protein demonstrating that these are solid aggregates.

Importantly, we found that animals with higher levels of PAB-1 aggregation were smaller, less motile and shorter lived than animals without aggregation.[20] These results demonstrate that aggregation of sgRBPs is potentially toxic and could accelerate the aging process.

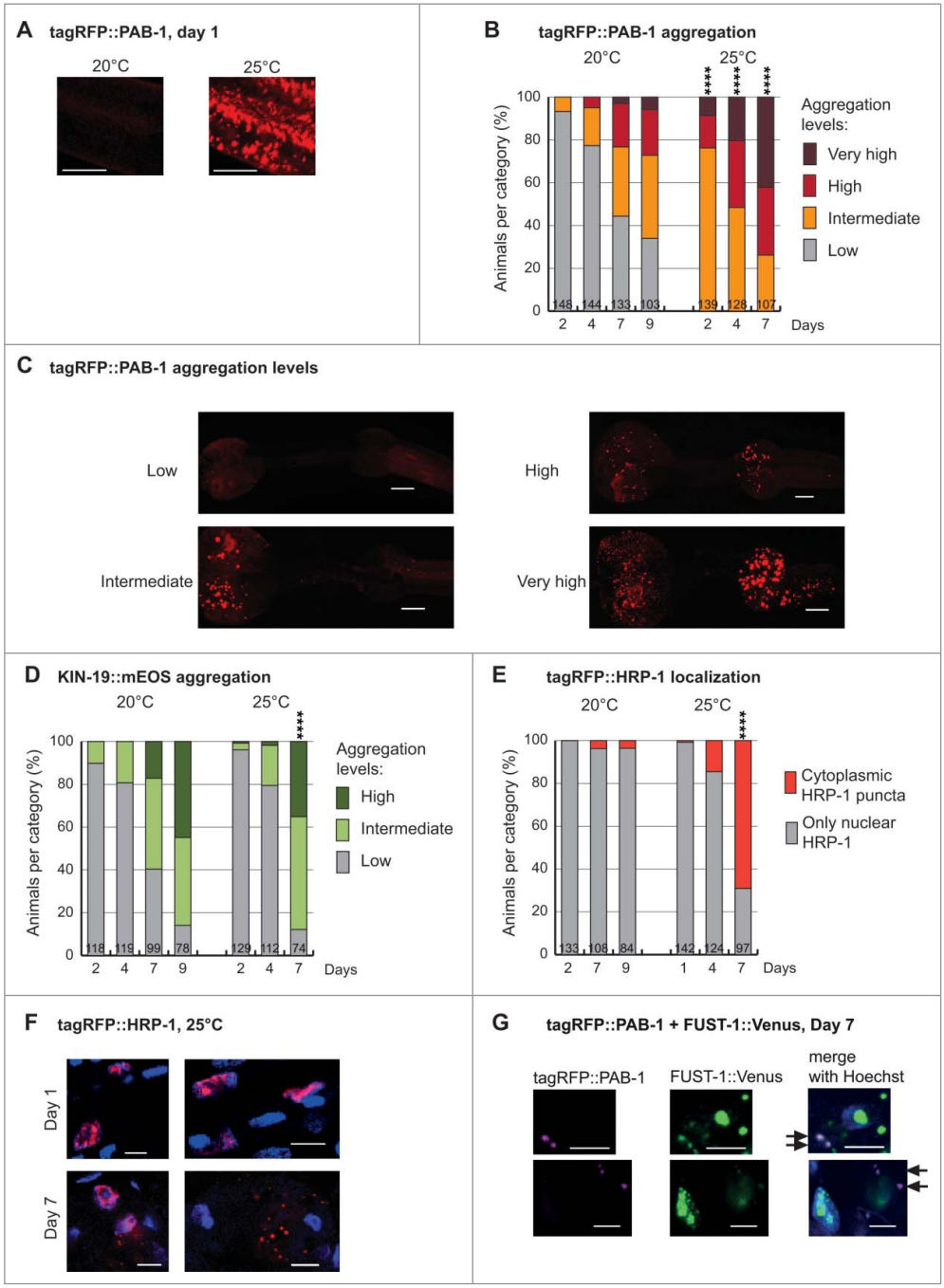
Sustained Stress Granule Formation Triggers sgRBP Aggregation

In vitro studies have shown that liquid droplets formed by purified RBPs with LC prion-like domains will eventually transition into a more solid state over time, thus impairing their disassembly.[11-14] Solidification of liquid droplets and concomitantly the formation of fibrils was enhanced by disease-related mutations promoting aggregation and repeated cycles of phase separation.[11,13] These findings lead us to the hypothesis that sustained stress granule formation could enhance their aggregation in vivo. 24h after continuous mild stress at 25°C, confocal analysis revealed the formation of abundant stress granules in day-1-old adults (Fig. 1A). Already on day 2, these animals started to form large PAB-1 aggregates visible at low-magnification. Overall, PAB-1 aggregation in a population of *C. elegans* grown at 25°C versus 20°C was greatly enhanced at all ages evaluated (Fig. 1B, C). As control, we examined the rate of aggregate formation by KIN-19, an age-dependent aggregation-prone protein without a LC prion-like domain or RNA-binding domain. In this case, we observed only a minor enhancement of KIN-19 aggregation rate at 25°C that is probably due to the accelerated rate of *C. elegans* aging at this higher temperature (Fig. 1D). These results demonstrate that continuous liquid-liquid phase separation over time plays an important role in initiating the aggregation of RBPs, potentially by allowing stress granules to stabilize and act as a nucleation site for aggregation.

Mild Stress Combined with Aging Promotes Mislocalization of RBPs with LC Prion-Like Domains

Several neurodegenerative diseases such as ALS and FTLD are characterized by cytoplasmic

FIGURE 1 (next page). Aging in combination with mild chronic stress enhances abnormal distribution of RNA granule components in *C. elegans* pharyngeal muscles. (A) Mild stress by elevation of growth temperature from 20°C to 25°C causes abundant tagRFP::PAB-1 stress granule formation in *Pmyo-2::tagRFP::PAB-1* transgenics at day 1. Representative confocal images displayed as single plane. Scale bar 7 μm. (B) Mild chronic stress highly enhances tagRFP::PAB-1 aggregation with age in *Pmyo-2::tagRFP::PAB-1* transgenics (Number of animals per group and time point is indicated in the columns; Fisher's exact test 2-tailed comparing low aggregation vs. all other categories at day 2, 4 and 7, 20°C vs. 25°C, ****$p < 0.0001$). (C) Representative images of tagRFP::PAB-1 aggregation levels in pharyngeal muscles (maximal z-stack projection). Low, up to 10 puncta in posterior bulb; intermediate, more than 10 puncta in posterior bulb; high, more than 10 puncta in anterior bulb; very high, more than 50 puncta in anterior bulb. Scale bar 10 μm. (D) Modest increase in KIN-19::mEOS aggregation in *Pkin-19::KIN-19::mEOS* animals aged at 25°C compared with 20°C. (Aggregation determined in anterior pharyngeal bulb: low, up to 10 puncta; intermediate, between 10 and 100 puncta; high, over 100 puncta. Fisher's exact test 2-tailed comparing high + intermediate against low aggregation levels at day 7, 20°C vs. 25°C $p < 0.0001$). (E) Mild chronic stress in combination with aging causes abnormal cytoplasmic tagRFP::HRP-1 localization in aged *Pmyo-2::tagRFP::HRP-1* transgenics (Fisher's exact test 2-tailed at day 7, 20°C vs. 25°C, ****$p < 0.0001$). (F) High-magnification representative confocal images with nuclear Hoechst33342 staining showing cytoplasmic tagRFP::HRP-1 puncta at day 7 but not day 1 of adulthood, in *Pmyo-2::tagRFP::HRP-1* transgenics aged at 25°C. Scale bar 5 μm. (G) Representative confocal images showing tagRFP::PAB-1 co-localization with small cytoplasmic FUST-1::Venus puncta in double transgenic animals expressing *Pmyo-2::tagRFP::PAB-1; Pkin-19::FUST-1::Venus* at day 7. Arrows mark co-localization. Scale bar: 4 μm.



inclusions of mainly nuclear-localized RBPs. Therefore RBP mislocalization from the nucleus to the cytoplasm is a key step toward pathogenesis. We investigated whether the aging process triggers mislocalization of nuclear RBPs. For this, we selected HRP-1, a nuclear RBP with a LC prion-like domain, which we identified in our proteomic analysis to become highly insoluble with age. Significantly, the human homologs of HRP-1, hnRNPA1 and hnRNPA3 form aberrant cytoplasmic inclusions in multisystem proteinopathy and they are also found to co-aggregate in inclusions in C9orf72 mutation-associated ALS/FTLD.[6,23] Using a fluorescent-tagged HRP-1, we confirmed its primary nuclear localization (Fig. 1E and 1F) and we observed that HRP-1 is

not a normal constituent of cytoplasmic stress granules as it remained in the nucleus upon acute heat stress (2 hour heat shock at 32°C, data not shown). Next we tested if aging modulates HRP-1 localization. In *C. elegans* maintained in standard conditions at 20°C, HRP-1 remained in the nucleus in the majority of aged animals (Fig. 1E). However, this was strikingly different when animals were exposed to an additional mild stress by aging them at 25°C. Whereas HRP-1 was localized in the nucleus in young animals, we observed the formation of distinct cytoplasmic puncta in aged animals (Fig. 1F). Our quantification reveals that 69% of day-7-old animals grown at 25°C developed cytoplasmic HRP-1 puncta (Fig. 1E). Therefore, the combination of changes related to aging and chronic exposure to mild environmental stress drives the aberrant cytoplasmic localization of a nuclear RBP. It remains to be investigated whether the aberrant cytoplasmic localization of RBPs that we observed in *C. elegans* is due to a disruption in nuclear integrity with age and/or a specific impairment in the nuclear import machinery.[24,25]

Aberrant Interactions Between RBPs and Other Aggregation-Prone Proteins

An important question is whether RBP aggregation with age influences the aggregation of other RBPs. In vitro phase separated RBPs with LC prion-like domains recruit others into the same assembly.[10,12] Our data revealed that PAB-1 and TIAR-2 have very different aggregation patterns with age.[20] Significantly, TIAR-2 distribution in aged animals was remarkably reminiscent of stress granule assembly caused by heat shock in young animals whereas PAB-1 tended to form mostly large aggregates with age. Notably in double transgenics, TIAR-2 recruited PAB-1 preferentially into stress granule-like structures with age.[20] These results reveal that interactions between 2 RBPs with different aggregation propensities can change their aggregation patterns.

We also investigated the *C. elegans* ortholog of FUS, FUST-1. Like the human FUS, nematode FUST-1 has a LC prion-like domain, an RNA recognition motif, a zinc finger domain and a PY-NLS motif. In healthy cells, human FUS is predominately nuclear while in the disease state, FUS forms cytoplasmic inclusions. Similarly in *C. elegans*, FUST-1 was mostly located in the nucleus. However, its overexpression alone was sufficient to cause the formation of small cytoplasmic puncta visible at higher magnification in young and aged animals (data not shown). As heat shock did not result in a clear increase in cytoplasmic FUST-1 puncta (data not shown), these aberrant puncta are unlikely to be related to functional stress granules. This is similar to observations made with wild-type human FUS in mammalian cells and in *C. elegans* showing that FUS does not normally localize to stress granules upon exposure to heat stress.[26-28] Importantly when co-expressed, FUST-1 recruited PAB-1 into these aberrant cytoplasmic puncta in both aged as well as young animals (Fig. 1G).

Several studies have highlighted an aberrant interaction between stress granules and misfolded proteins in yeast and in cell culture.[29-32] Significantly, recent work showed that stress granules containing misfolded proteins tend to be more stable.[31,32] In *C. elegans* we found that sgRBP PAB-1 forms large solid aggregates together with a non-RBP and age-dependent aggregation-prone protein KIN-19, when both proteins were overexpressed.[20] As KIN-19 greatly accelerated PAB-1 aggregation, cross-seeding mechanisms due to protein misfolding with age is likely to be a main cause of sgRBP aggregation. Important questions remain and in particular related to how this aberrant interaction is initiated: Are non-RBP aggregation-prone proteins first recruited into age-induced stress granules and then act as seeds for their solidification into large aggregates? Or do small age-dependent aggregates of non-RBPs trigger stress granule formation and co-aggregation? One argument in favor of the first possibility is the fact that stress granules induced by heat stress in HeLa cells subsequently recruit ALS-linked mutant SOD-1.[31]

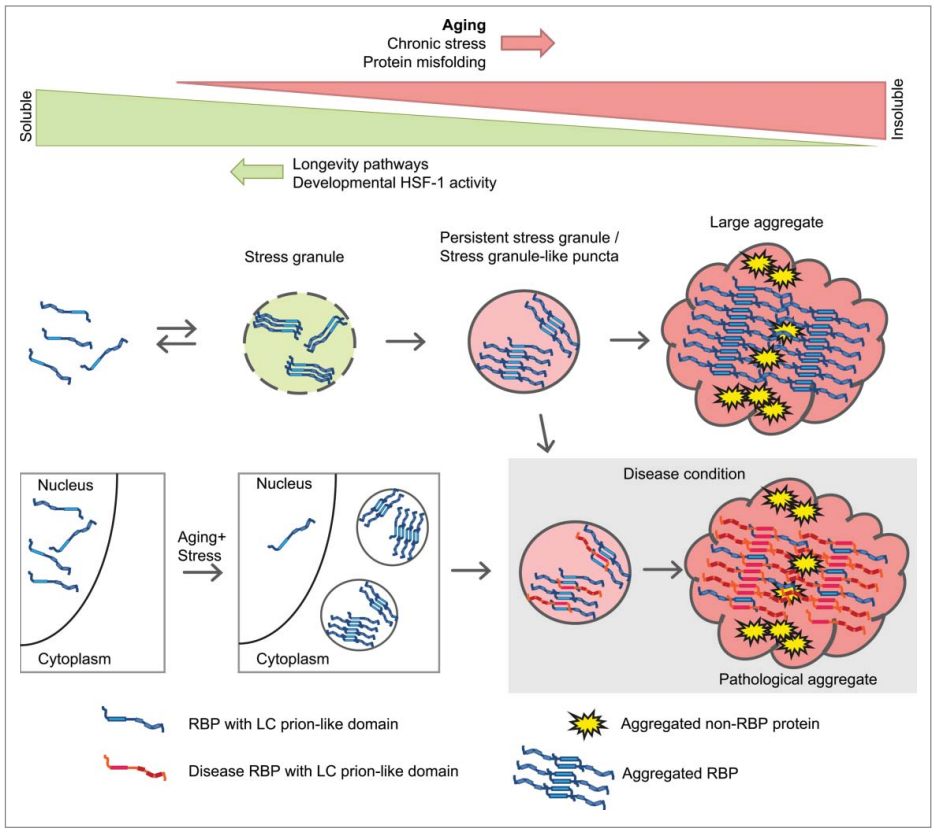
Overall, these findings show that RBPs with LC prion-like domains are prone to recruit other RBPs into aberrant cytoplasmic accumulations and to interact with other aggregation-prone proteins. This could explain why stress granule components are frequently observed not only in disease-associated RBP aggregates

but also in other pathological aggregates identified in Alzheimer disease and Huntington's disease.[21] Moreover our results strongly suggest that sgRBPs are not passive players in neurodegenerative disease.

Strategies to Prevent Stress Granule Aggregation

We found that sgRBP aggregation is efficiently prevented by at least 3 independent longevity pathways in *C. elegans*, namely reduced *daf-2*/insulin/IGF-1-like receptor signaling, dietary restriction and reduced mitochondrial function.[20] Accordingly, maintaining dynamic stress granules is likely to be a common strategy to delay the course of aging. This concept is further supported by the reduced fitness of aged animals with sgRBP aggregation.[20] Two transcription factors, heat shock factor HSF-1 and DAF-16/FOXO, are strongly activated in animals with reduced *daf-2* signaling and play an essential role in mediating longevity and improving proteostasis. We demonstrated that both factors are involved in promoting dynamic RBPs.[20] Notably HSF-1 activity during development was essential to set-up an active quality-control system to avoid sgRBP aggregation both in wild-type and in long-lived *daf-2* mutants. As HSF-1 is the master regulator of the chaperone system, we sought to determine which specific chaperones may be involved in abrogating sgRBP aggregation. Interestingly, we did not find a connection between sgRBP aggregation and the activity of the *C. elegans* homologs of HSP40, HSP70 and HSP110, which have been previously implicated in stress granule dynamics in *S. cerevisiae* and/or cell culture.[9,29,30,33] In addition, HSP40, HSP70

FIGURE 2. Model depicting aggregation of sgRBPs during aging and the relationship with disease RBPs.



and HSP110 chaperones act together to promote protein disaggregation.[34,35] However, we did not observe the involvement of the nematode disaggregase machinery in the modulation of sgRBP aggregation (in particular DNJ-13, HSP-1 and HSP-110, [20] and data not shown). In a small RNA interference screen targeting 15 *C. elegans* chaperones individually, only the knock-down of small heat shock protein HSP-16.11 accelerated PAB-1 aggregation (statistically significant in 3 out of 4 repeats, data not shown). In future work, it will be important to test the chaperone complex HSPB8-BAG3-HSP70 which is required to avert defective ribosomal products from accumulating in stress granules [32] and thus might also play a role in sgRPB aggregation regulation. Furthermore a recent study demonstrated that mini-chromosome maintenance MCM proteins stabilize stress granules.[36] As several *mcm* genes are upregulated in *hsf-1* mutants in *C. elegans*,[37] these could be promising candidates as pro-aggregation factors contributing to sgRBP aggregation with age. In addition to strategies enhancing stress granule disassembly, our results implicate that preventing cross-seeding by the aggregating proteome will be an effective way to reduce RBP aggregation. Significantly, increasing DAF-16 activity could be used as a general strategy to reduce widespread protein aggregation with age.[20]

CONCLUSIONS

We propose the model depicted in Fig. 2 to explain our results and their possible implications for neurodegenerative diseases. First, RBPs with LC prion-like domains form stress granules during aging and undergo a change in structure allowing them to stabilize the whole stress granule and avoid its disassembly. The presence of other age-dependent aggregation-prone proteins helps accelerate a further change in structure and leads to the formation of large aggregates. Second, the occurrence of a mild chronic stress during aging triggers the mislocalization of certain nuclear RBPs in cytoplasmic accumulations. Third, we hypothesize that aberrant interactions between age-dependent aggregating RBPs and disease-associated RBPs also affects disease protein aggregation. As stress granule proteins are found not only in disease-associated RBP aggregates but also in tau aggregates and Huntingtin aggregates,[21] their inherent propensity to aggregate with age could impact pathogenesis in a range of neurodegenerative diseases.

Furthermore, aggregation of RNA granule components with age could have dramatic consequences for the normal function of RNA granules. For example, the inability of stress granules to disassemble after stress subsides and the permanent sequestration of stress granule components including mRNAs would stunt cellular adaptations and recovery from stress. As such, stress granule aggregation could explain reduced stress resistance with age. In the age-dependent aggregating proteome we also identified CAR-1, an RBP with LC prion-like domain with homology to the P-body component LSM14B.[20] This could indicate that P-bodies may solidify with age. This would have important consequences for the removal of aberrant mRNAs and could contribute to the decline of nonsense-mediated decay observed with age.[38] Significantly the function of other membrane-less structures such as the nucleolus are likely to be negatively affected by the aggregation of RBPs with LC prion-like domains. For instance, we identified that the nucleolus RBP Fibrillarin (FIB-1) was highly prone to aggregate with age.[20] Previously FIB-1 was shown to undergo liquid-liquid phase separation and became less dynamic with age.[39,40]

The conservation of LC prion-like domains throughout evolution from yeast to mammals underlines the importance of this domain in the cellular biology. In the last 10 years, an increasing number of dementias have been associated with the aberrant aggregation of specific RBPs. Overall our study highlights the vulnerability of membrane-less organelles organized by RBPs with LC prion-like domains to transition into a solid state during normal aging in a multicellular organism. This has important implications for our understanding of the role of aging in pathogenesis and toxicity in neurodegenerative diseases.

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

No potential conflicts of interest were disclosed.

FUNDING

This work was supported by funding from the DZNE and a Marie Curie International Reintegration Grant (322120 to DCD)

REFERENCES

[1] Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, Mann D, Tsuchiya K, Yoshida M, Hashizume Y, et al. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Biochem Biophys Res Commun. 2006;351(3):602-11. doi:10.1016/j.bbrc.2006.10.093. PMID:17084815

[2] Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM, et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science. 2006;314 (5796):130-3. doi:10.1126/science.1134108. PMID:17023659

[3] Neumann M, Rademakers R, Roeber S, Baker M, Kretzschmar HA, Mackenzie IR. A new subtype of frontotemporal lobar degeneration with FUS pathology. Brain. 2009;132(Pt 11):2922-31. doi:10.1093/brain/awp214. PMID:19674978

[4] King OD, Gitler AD and Shorter J. The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. Brain Res. 2012;1462:61-80. doi:10.1016/j.brainres.2012.01.016. PMID:22445064

[5] Alberti S, Halfmann R, King O, Kapila A and Lindquist S. A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. Cell. 2009;137(1):146-58. doi:10.1016/j.cell.2009.02.044. PMID:19345193

[6] Kim HJ, Kim NC, Wang YD, Scarborough EA, Moore J, Diaz Z, MacLea KS, Freibaum B, Li S, Molliex A, et al. Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. Nature. 2013;495(7442):467-73. doi:10.1038/nature11922. PMID:23455423

[7] Johnson BS, Snead D, Lee JJ, McCaffery JM, Shorter J and Gitler AD. TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. J Biol Chem. 2009;284(30):20329-39. doi:10.1074/jbc.M109.010264. PMID:19465477

[8] Couthouis J, Hart MP, Shorter J, DeJesus-Hernandez M, Erion R, Oristano R, Liu AX, Ramos D, Jethava N, Hosangadi D, et al. A yeast functional screen predicts new candidate ALS disease genes. Proc Natl Acad Sci U S A. 2011;108(52):20881-90. doi:10.1073/pnas.1109434108. PMID:22065782

[9] Gilks N, Kedersha N, Ayodele M, Shen L, Stoecklin G, Dember LM and Anderson P. Stress granule assembly is mediated by prion-like aggregation of TIA-1. Mol Biol Cell. 2004;15(12):5383-98. doi:10.1091/mbc.E04-08-0715. PMID:15371533

[10] Kato M, Han TW, Xie S, Shi K, Du X, Wu LC, Mirzaei H, Goldsmith EJ, Longgood J, Pei J, et al. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell. 2012;149(4):753-67. doi:10.1016/j.cell.2012.04.017. PMID:22579281

[11] Murakami T, Qamar S, Lin JQ, Schierle GS, Rees E, Miyashita A, Costa AR, Dodd RB, Chan FT, Michel CH, et al. ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. Neuron. 2015;88(4):678-90. doi:10.1016/j.neuron.2015.10.030. PMID:26526393

[12] Lin Y, Protter DS, Rosen MK, Parker R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. Mol Cell. 2015;60 (2):208-19. doi:10.1016/j.molcel.2015.08.018. PMID:26412307

[13] Molliex A, Temirov J, Lee J, Coughlin M, Kanagaraj AP, Kim HJ, Mittag T and Taylor JP. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. Cell. 2015;163(1):123-33. doi:10.1016/j.cell.2015.09.015. PMID:26406374

[14] Patel A, Lee HO, Jawerth L, Maharana S, Jahnel M, Hein MY, Stoynov S, Mahamid J, Saha S, Franzmann TM, et al. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. Cell. 2015;162(5):1066-77. doi:10.1016/j.cell.2015.07.047. PMID:26317470

[15] David DC, Ollikainen N, Trinidad JC, Cary MP, Burlingame AL and Kenyon C. Widespread protein aggregation as an inherent part of aging in C. elegans. PLoS Biol. 2010;8(8):e1000450. doi:10.1371/journal.pbio.1000450. PMID:20711477

[16] Reis-Rodrigues P, Czerwieniec G, Peters TW, Evani US, Alavez S, Gaman EA, Vantipalli M, Mooney SD, Gibson BW, Lithgow GJ et al. Proteomic analysis of age-dependent changes in protein solubility identifies genes that modulate lifespan. Aging Cell. 2012;11(1):120-7. doi:10.1111/j.1474-9726.2011.00765.x. PMID:22103665

[17] Demontis F, Perrimon N. FOXO/4E-BP signaling in Drosophila muscles regulates organism-wide proteostasis during aging. Cell. 2010;143(5):813-25. doi:10.1016/j.cell.2010.10.007. PMID:21111239

[18] Tanase M, Urbanska AM, Zolla V, Clement CC, Huang L, Morozova K, Follo C, Goldberg M, Roda B, Reschiglian P, Santambrogio L. Role of Carbonyl Modifications on Aging-Associated Protein Aggregation. Sci Rep. 2016;6:19311. doi:10.1038/srep19311. PMID:26776680

[19] Ayyadevara S, Mercanti F, Wang X, Mackintosh SG, Tackett AJ, Prayaga SV, Romeo F, Shmookler Reis RJ and Mehta JL. Age- and Hypertension-Associated Protein Aggregates in Mouse Heart Have Similar Proteomic Profiles. Hypertension. 2016; 67(5):1006-13. doi:10.1161/HYPERTENSIONAHA.115.06849. PMID:26975704

[20] Lechler MC, Crawford ED, Groh N, Widmaier K, Jung R, Kirstein J, Trinidad JC, Burlingame AL and David DC. Reduced Insulin/IGF-1 Signaling Restores the Dynamic Properties of Key Stress Granule Proteins during Aging. Cell Rep. 2017;18(2):454-467. doi:10.1016/j.celrep.2016.12.033. PMID:28076789

[21] Bentmann E, Haass C and Dormann D. Stress granules in neurodegeneration–lessons learnt from TAR DNA binding protein of 43 kDa and fused in sarcoma. FEBS J. 2013;280(18):4348-70. doi:10.1111/febs.12287. PMID:23587065

[22] Li YR, King OD, Shorter J, Gitler AD. Stress granules as crucibles of ALS pathogenesis. J Cell Biol. 2013;201(3):361-72. doi:10.1083/jcb.201302044. PMID:23629963

[23] Mori K, Lammich S, Mackenzie IR, Forne I, Zilow S, Kretzschmar H, Edbauer D, Janssens J, Kleinberger G, Cruts M, et al. hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with C9orf72 mutations. Acta Neuropathol. 2013;125(3):413-23. doi:10.1007/s00401-013-1088-7. PMID:23381195

[24] D'Angelo MA, Raices M, Panowski SH and Hetzer MW. Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells. Cell. 2009;136(2):284-95. doi:10.1016/j.cell.2008.11.037. PMID:19167330

[25] Dormann D, Haass C. TDP-43 and FUS: a nuclear affair. Trends Neurosci. 2011;34(7), 339-48. doi:10.1016/j.tins.2011.05.002. PMID:21700347

[26] Sama RR, Ward CL, Kaushansky LJ, Lemay N, Ishigaki S, Urano F and Bosco DA. FUS/TLS assembles into stress granules and is a prosurvival factor during hyperosmolar stress. J Cell Physiol. 2013;228(11):2222-31. doi:10.1002/jcp.24395. PMID:23625794

[27] Bosco DA, Lemay N, Ko HK, Zhou H, Burke C, Kwiatkowski TJ, Jr., Sapp P, McKenna-Yasek D, Brown RH, Jr., Hayward LJ. Mutant FUS proteins that cause amyotrophic lateral sclerosis incorporate into stress granules. Hum Mol Genet. 2010;19(21):4160-75. doi:10.1093/hmg/ddq335. PMID:20699327

[28] Murakami T, Yang SP, Xie L, Kawano T, Fu D, Mukai A, Bohm C, Chen F, Robertson J, Suzuki H, et al. ALS mutations in FUS cause neuronal dysfunction and death in Caenorhabditis elegans by a dominant gain-of-function mechanism. Hum Mol Genet. 2012;21(1):1-9. doi:10.1093/hmg/ddr417. PMID:21949354

[29] Cherkasov V, Hofmann S, Druffel-Augustin S, Mogk A, Tyedmers J, Stoecklin G, Bukau B. Coordination of translational control and protein homeostasis during severe heat stress. Curr Biol. 2013;23(24):2452-62. doi:10.1016/j.cub.2013.09.058. PMID:24291094

[30] Kroschwald S, Maharana S, Mateju D, Malinovska L, Nuske E, Poser I, Richter D and Alberti S. Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. Elife. 2015; 4:e06807. doi:10.7554/eLife.06807. PMID:26238190

[31] Mateju D, Franzmann TM, Patel A, Kopach A, Boczek EE, Maharana S, Lee HO, Carra S, Hyman AA, Alberti S. An aberrant phase transition of stress granules triggered by misfolded protein and prevented by chaperone function. EMBO J. 2017;36(12):1669-1687. doi:10.15252/embj.201695957. PMID:28377462

[32] Ganassi M, Mateju D, Bigi I, Mediani L, Poser I, Lee HO, Seguin SJ, Morelli FF, Vinet J, Leo G, et al. A Surveillance Function of the HSPB8-BAG3-HSP70 Chaperone Complex Ensures Stress Granule Integrity and Dynamism. Mol Cell. 2016;63(5):796-810. doi:10.1016/j.molcel.2016.07.021. PMID:27570075

[33] Walters RW, Muhlrad D, Garcia J and Parker R. Differential effects of Ydj1 and Sis1 on Hsp70-mediated clearance of stress granules in Saccharomyces cerevisiae. RNA. 2015;21(9):1660-71. doi:10.1261/rna.053116.115. PMID:26199455

[34] Shorter J. The mammalian disaggregase machinery: Hsp110 synergizes with Hsp70 and Hsp40 to catalyze protein disaggregation and reactivation in a cell-free system. PLoS One. 2011;6(10):e26319. doi:10.1371/journal.pone.0026319. PMID:22022600

[35] Nillegoda NB, Kirstein J, Szlachcic A, Berynskyy M, Stank A, Stengel F, Arnsburg K, Gao X, Scior A, Aebersold R, Guilbride DL, Wade RC, Morimoto RI, Mayer MP and Bukau B. Crucial HSP70 co-chaperone complex unlocks metazoan protein disaggregation. Nature. 2015;524(7564):247-51. doi:10.1038/nature14884. PMID:26245380

[36] Jain S, Wheeler JR, Walters RW, Agrawal A, Barsic A and Parker R. ATPase-Modulated Stress

Granules Contain a Diverse Proteome and Substructure. Cell. 2016;164(3):487-98. doi:10.1016/j.cell.2015.12.038. PMID:26777405

[37] Brunquell J, Morris S, Lu Y, Cheng F and Westerheide SD. The genome-wide role of HSF-1 in the regulation of gene expression in Caenorhabditis elegans. BMC Genomics. 2016;17:559. doi:10.1186/s12864-016-2837-5. PMID:27496166

[38] Son HG, Seo M, Ham S, Hwang W, Lee D, An SW, Artan M, Seo K, Kaletsky R, Arey RN, et al. RNA surveillance via nonsense-mediated mRNA decay is crucial for longevity in daf-2/insulin/IGF-1 mutant C. elegans. Nat Commun. 2017;8:14749. doi:10.1038/ncomms14749. PMID:28276441

[39] Berry J, Weber SC, Vaidya N, Haataja M and Brangwynne CP. RNA transcription modulates phase transition-driven nuclear body assembly. Proc Natl Acad Sci U S A. 2015;112(38):E5237-45. doi:10.1073/pnas.1509317112. PMID:26351690

[40] Feric M, Vaidya N, Harmon TS, Mitrea DM, Zhu L, Richardson TM, Kriwacki RW, Pappu RV, Brangwynne CP.. Coexisting Liquid Phases Underlie Nucleolar Subcompartments. Cell.2016;165(7):1686-97. doi:10.1016/j.cell.2016.04.047. PMID:27212236

Jansen et al. Genome Biology (2017) 18:22
DOI 10.1186/s13059-017-1147-9

Genome Biology

RESEARCH Open Access

Discovery and functional prioritization of Parkinson's disease candidate genes from large-scale whole exome sequencing

Iris E. Jansen[1,2†], Hui Ye[3†], Sasja Heetveld[1†], Marie C. Lechler[1,4†], Helen Michels[5], Renée I. Seinstra[5], Steven J. Lubbe[6,7], Valérie Drouet[8], Suzanne Lesage[8], Elisa Majounie[9], J. Raphael Gibbs[10], Mike A. Nalls[10], Mina Ryten[11,12], Juan A. Botia[11], Jana Vandrovcova[11], Javier Simon-Sanchez[1,13], Melissa Castillo-Lizardo[1], Patrizia Rizzu[1], Cornelis Blauwendraat[1], Amit K. Chouhan[14], Yarong Li[14], Puja Yogi[14], Najaf Amin[15], Cornelia M. van Duijn[15], International Parkinson's Disease Genetics Consortium (IPGDC), Huw R. Morris[6], Alexis Brice[8,16], Andrew B. Singleton[10], Della C. David[1], Ellen A. Nollen[5], Shushant Jain[1], Joshua M. Shulman[3,14,17,18*†] and Peter Heutink[1,2,13*†]

Abstract

Background: Whole-exome sequencing (WES) has been successful in identifying genes that cause familial Parkinson's disease (PD). However, until now this approach has not been deployed to study large cohorts of unrelated participants. To discover rare PD susceptibility variants, we performed WES in 1148 unrelated cases and 503 control participants. Candidate genes were subsequently validated for functions relevant to PD based on parallel RNA-interference (RNAi) screens in human cell culture and *Drosophila* and *C. elegans* models.

Results: Assuming autosomal recessive inheritance, we identify 27 genes that have homozygous or compound heterozygous loss-of-function variants in PD cases. Definitive replication and confirmation of these findings were hindered by potential heterogeneity and by the rarity of the implicated alleles. We therefore looked for potential genetic interactions with established PD mechanisms. Following RNAi-mediated knockdown, 15 of the genes modulated mitochondrial dynamics in human neuronal cultures and four candidates enhanced α-synuclein-induced neurodegeneration in *Drosophila*. Based on complementary analyses in independent human datasets, five functionally validated genes—*GPATCH2L*, *UHRF1BP1L*, *PTPRH*, *ARSB*, and *VPS13C*—also showed evidence consistent with genetic replication.

Conclusions: By integrating human genetic and functional evidence, we identify several PD susceptibility gene candidates for further investigation. Our approach highlights a powerful experimental strategy with broad applicability for future studies of disorders with complex genetic etiologies.

Keywords: Parkinson's disease, Genomics, Whole-exome sequencing, Loss-of-function, Rare variants, Functional screening, Mitochondria, Parkin, α-synuclein, Animal model

* Correspondence: Joshua.Shulman@bcm.edu; Peter.Heutink@dzne.de
Joshua M. Shulman and Peter Heutink are co-senior authors.
Iris E. Jansen, Hui Ye, Sasja Heetveld and Marie C. Lechler are co-first authors.
†Equal contributors
[3]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA
[1]German Center for Neurodegenerative Diseases (DZNE), Otfried-Müller-Str. 23, Tübingen 72076, Germany
Full list of author information is available at the end of the article

Background

Next-generation sequencing (NGS) approaches have recently accelerated the identification of variants responsible for familial Parkinson's disease (PD) [1–4]. While a positive family history is common in PD, large, multigenerational pedigrees, especially with available DNA and clinical evaluations, remain exceptional, hindering progress in unraveling the genetic underpinnings. Importantly, several genes initially discovered to cause PD in families, such as *LRRK2*, *GBA*, and *PARK2/parkin*, were subsequently discovered with surprisingly high frequency in "sporadic" PD cohorts [5, 6]. To date, large population samples of individuals with PD have primarily contributed to the discovery of common variant susceptibility loci, based on genome-wide association studies (GWAS) of case/control cohorts [7]. The variants identified by GWAS have modest effect sizes and collectively fail to account for current estimates of PD heritability [8, 9]. Considering the above, it seems likely that additional less common alleles, with larger effect sizes, contribute to PD risk in the population and NGS is one promising approach to identify such alleles. Despite recent successes in other neurodegenerative disease with complex genetic etiologies, including Alzheimer's disease [10–12] and amyotrophic lateral sclerosis [13, 14], sequencing has yet to be deployed in large, unrelated PD case/control samples for rare variant discovery.

The successful discovery of rare variant risk alleles in population-based PD samples faces a number of potential challenges. Perhaps most importantly, analyses of rare variants in large family pedigrees is greatly facilitated by segregation analysis which is not possible in cohorts of unrelated individuals, leading to an increased number of candidate variants to consider. Assumptions of a recessive inheritance model and the application of stringent filters, such as consideration of only strongly damaging, loss-of-function (LoF) variants, is one potential solution, but this is likely to miss many important variants, including dominantly acting alleles. Further, PD is characterized by extensive genic and allelic heterogeneity and extremely large cohorts may be required to document sufficient numbers of cases to facilitate meaningful statistical comparisons [15]. Lastly, as PD is: (1) common (~1–3% prevalence); (2) strongly age-dependent; and (3) often preceded by a prolonged presymptomatic or minimally symptomatic phase, we may expect to find truly pathogenic rare variants, including those with large effect sizes, in "control" cohorts of adults (due to unrecognized or early disease stages with minimal symptoms). Therefore, given the occurrence of rare variants, including potentially damaging variants, in most genomes of presumably healthy individuals [16], it may be difficult to identify genes/variants that truly cause disease. Importantly, recent advances in cellular and animal models, along with improved understanding of PD pathogenesis, enable an integrated approach, in which variant discovery is coupled with a functional screening pipeline for prioritization of those genes worthy of more intensive study.

In this collaborative study of the International Parkinson's Disease Genomics Consortium (IPDGC), we report the results of whole-exome sequencing (WES) in 1148 PD cases, the largest such cohort examined to date. Consistent with the younger age of PD onset in this cohort, which is often associated with a recessive inheritance [17–19], and to prioritize candidate genes/variants for initial investigation, our analysis focuses on genes with homozygous or compound heterozygous LoF variants. We further couple the human genetic studies with functional screening in mammalian cell culture and invertebrate animal models, successfully identifying those candidate genes showing interactions with established PD mechanisms, including mitochondrial dynamics and α-synuclein-mediated neurodegeneration. Although no sufficiently powered exome dataset was available for definitive replication, human genetic validation was undertaken in several independent datasets. Our integrated approach identifies five strong candidate PD susceptibility genes worthy of further investigation, and exemplifies a powerful strategy with potential broad applicability to the follow-up of future rare variant studies in PD and other neurologic disorders with complex genetic etiologies.

Results

Discovery of recessive LoF variants from PD exomes

A total of 920,896 variants (93.2% single nucleotide variants and 6.8% insertions and deletions) were called in a WES dataset of 1651 participants, including 1148 young-onset PD cases (average age of onset, 40.6 years; range, 5–56 years) and 503 control participants with European ancestry. As our cohort has an average age at onset of less than 45 years, we focused our search on homozygous and putative compound heterozygous variants, consistent with a recessive inheritance model. Although most PD cases were prescreened for mutations in established PD genes, we identified two participants with homozygous exonic variants in *parkin* and *PINK1* (Additional file 1: Table S1). In order to identify novel PD gene candidates, we focused on variants that are rare in control populations. Considering the worldwide prevalence for PD (0.041% in individuals aged 40–49 years) [20], we used a minor allele frequency (MAF) threshold of 1% and only considered LoF variants causing a premature stop codon or splicing site mutations (see "Methods"). When co-occurring with a heterozygous LoF variant, we also considered rare, heterozygous amino-acid changing missense alleles that were predicted to be

deleterious (CADD > 20), consistent with a compound heterozygous recessive genotype.

Figure 1 displays each variant filtering step along with the corresponding numbers of implicated variants. Following Sanger sequencing confirmation, we identified a total of 27 candidate genes—18 genes encompassing homozygous variants and nine genes harboring putative compound heterozygous variants—all predicted to cause a loss of gene function (Table 1). Approximately 17% of the variants are absent in public allele frequency databases (1000 Genomes Project (1000G), Exome Sequencing Project v. 6500 (ESP6500), or Exome Aggregation Consortium (ExAC)) and therefore implicated to be novel. Except in the case of *ARSB*, the other 26 genes harbor LoF variants in only a single case, consistent with the hypothesis that novel recessive PD alleles may consist of many rare, "private" mutations. Four PD cases in our cohort were identified with a LoF variant in the *ARSB* gene, in which mutations have previously been linked with the recessive lysosomal storage disorder, MPS VI (also called Maroteaux-Lamy syndrome). All four individual cases, along with one control participant, were homozygous for a variant (rs138279020) predicted to disrupt splicing. Although this variant is neither reported in ExAC nor was frequency information available from dbSNP, the MAF was 0.065 in our cohort ($MAF_{CASES} = 0.073$, $MAF_{CONTROLS} = 0.052$, $p = 0.054$). Although relatively frequent in our control dataset (MAF > 1%), we have retained it among our results, based on three considerations. First, information was not present in dbSNP, ExAC, or ESP6500, which was the basis for applying this frequency filter in all other cases. Second, at least one of the homozygous individuals had clinical manifestations consistent with MPS VI, supporting potential pathogenicity of this allele (see "Discussion"). Lastly, as detailed below, our functional

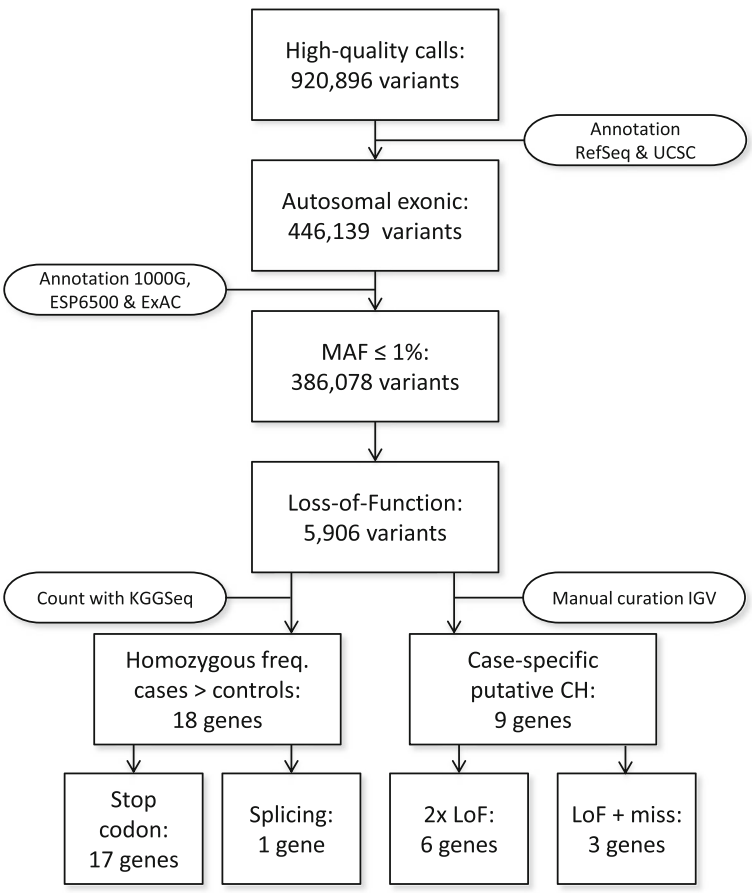


Fig. 1 *Flowchart* explaining multiple filtering steps to select LoF variants with assumed recessive inheritance pattern. Functional annotation was performed with transcripts of RefSeq and UCSC databases. MAF annotations were based on 1000 Genomes project, Exome variant Server, and the ExAC database. Seventeen genes harbored homozygous variants causing stopgain or loss and one gene contained a homozygous splicing variant. For the putative compound heterozygous genes, six genes were selected based on the presence of two LoF variants, and three genes were based on the presence of one LoF variant and one missense variant (predicted to belong to the 1% most harmful variants of the genome)

Table 1 Recessive LoF variants

WES results		Variant characteristics					Functional information				Validation	
					ExAC (EU)							
Gene	Type	Chr:bp (hg19)	Ref/Alt all	dbSNP137	MAF	*n* HZ (freq.)	Variant type	AA change	AA length	CADD	Functional	Genetic
ANKRD30A	CH	10:37438759	G/T	x	x	x	stopgain	NM_052997:p.E487X	1341	35.0		
		10:37508648	A/AT	x	0.0031%	0	fs insertion	NM_052997:p.Q1280fs	1341	24.2		
ARSB	HZ	5:78281383	T/TA	rs138279020	x	x	splicing	NM_198709:c.190-2insT	533	2.2	S	neuroX
c11orf21	HZ	11:2321829	C/T	rs74048215	0.037%	0	stopgain	NM_001142946:p.W69X	178	35.0		
CALML4	HZ	15:68497597	G/A	rs11071990	0.75%	1 (0.00003%)	stopgain	NM_001031733:p.R40X	120	25.9		
CAPS2	HZ	12:75687045	C/CTATTTGAT	x	0.68%	4 (0.00012%)	stopgain	NM_001286547:p.L321X	382	33.0		
CD36	CH	7:80285955	C/T	x	0.0075%	0	stopgain	NM_001127444:p.Q74X	472	26.5		GRIP
		7:80302705	C/T	x	0.011%	0	missense	NM_001127444:p.P412S	472	25.9		
COL6A5	CH	3:130132401	C/T	rs190283135	0.12%	0	stopgain	NM_001278298:p.Q1559X	2611	38.0		
		3:130107827	G/A	x	0.32%	0	missense	NM_001278298:p.V756M	2611	23.0		
DIS3	CH	13:73355096	G/A	rs201493043	0.0045%	0	stopgain	NM_014953:p.Q92X	958	37.0	Mm	
		13:73333935	A/G	rs141067458	0.22%	0	stoploss	NM_014953:p.X959Q	958	14.1		
FAM71A	HZ	1:212799882	A/T	rs143861665	0.80%	2 (0.00006%)	stopgain	NM_153606:p.K555X	594	36.0	Mm	
FAM83A	CH	8:124195352	G/T	rs148011353	0.23%	2 (0.00006%)	stopgain	NM_032899:p.G86X	434	26.2		PPMI
		8:124195506	T/G	x	x	x	missense	NM_032899:p.V137G	434	25.6		
GH2	HZ	17:61957946	C/T	rs150668018	0.33%	1 (0.00003%)	stopgain	NM_022557:p.W214X	256	26.8		
GPATCH2L	HZ	14:76644266	C/T	rs117516637	0.53%	1 (0.00003%)	stopgain	NM_017972:p.R362X	482	18.8	Mp,Mm	PPMI
KALRN	HZ	3:124303696	C/T	rs56407180	0.39%	0	stopgain	NM_007064:p.R10X	1289	37.0	Mp	
KCNK16	HZ	6:39282816	G/A	rs138573996	0.44%	2 (0.00006%)	stopgain	NM_001135107:p.Q251X	322	9.5	Mm	
MNS1	CH	15:56736045	G/A	x	x	x	stopgain	NM_018365:p.Q232X	495	38.0	Mm	
		15:56748680	G/A	x	0.003%	0	stopgain	NM_018365:c.C265T:p.Q89X	495	37.0		
OR7G3	HZ	19:9236969	G/A	rs61751875	0.15%	1 (0.00003%)	stopgain	NM_001001958:p.R220X	312	35.0	Mm	
PCDHA9	HZ	5:140229907	C/G	x	0.0046%	0	stopgain	NM_014005:p.Y609X	950	34.0	Mm	
PTCHD3	HZ	10:27688101	G/A	rs142646098	0.97%	6 (0.00018%)	stopgain	NM_001034842:p.R476X	767	34.0	Mp,Mm	
PTPRH	CH	19:55697712	G/A	rs147881000	0.39%	1 (0.00003%)	stopgain	NM_002842:p.Q887X	1115	36.0	S,Mm	neuroX
		19:55716715	C/A	rs201517965	0.0030%	0	stopgain	NM_002842:p.E200X	1115	26.0		
PZP	CH	12:9321534	G/A	rs145240281	0.85%	4 (0.00012%)	stopgain	NM_002864:p.R680X	1482	32.0		
		12:9333626	G/A	rs117889746	0.021%	0	stopgain	NM_002864:p.Q598X	1482	36.0		
SSPO	HZ	7:149493519	C/T	rs57595625	0.011%	0	stopgain	NM_198455:p.Q2199X	5150	10.4		
SVOPL	HZ	7:138341219	G/A	rs117871806	0.35%	0	stopgain	NM_174959:p.R18X	492	38.0	Mp,Mm	

Table 1 Recessive LoF variants *(Continued)*

TCHHL1	HZ	1:152058192	G/A	rs150014958	0.037%	0	stopgain	NM_001008536:p.Q656X	904	35.0	Mp	
TMEM134	HZ	11:67235051	G/A	rs143199541	0.70%	3 (0.00009%)	stopgain	NM_001078650:p.R84X	195	34.0	S	
UHRF1BP1L	HZ	12:100433523	T/A	x	x	x	stopgain	NM_015054:p.K1376X	1464	35.0	Mm	neuroX
VPS13C	CH	15:62174851	C/A	x	x	x	stopgain	NM_017684:p.E3147X	3628	55.0	S,Mm	neuroX,GWAS
		15:62305257	C/CTCTG	x	x	x	fs insertion	NM_017684:p.R226fs	3628	27.4		
ZNF543	HZ	19:57838058	G/A	rs150392165	0.60%	1 (0.00003%)	stopgain	NM_213598:p.W68X	600	25.0	Mp,Mm	

The current designation of rs138279020 is rs11424557

Type type of affecting 2 alleles, *CH* putative compound heterozygote, *HZ* homozygote, *Chr:bp* chromosome and base-pair position, *hg19* reference build, *Ref* reference, *Alt* alternative, *ExAC* Exome Aggregation consortium, *EU* European, *fs* frameshift, *AA change* amino acid change for specified RefSeq transcript, *MAF* minor allele frequency, *n HZ* number of individuals with homozygous variant, *AA length* length of specified transcript in amino acids, *PPMI* Parkinson Progression Markers Initiative, *CADD* functional algorithm prediction (>20 = belongs to 1% most damaging variants of total genome), *Mp* mitochondrial Parkin translocation assay, *Mm* mitochondrial morphology assay, *S* α-synuclein assay

studies identify links between manipulation of ARSB and cellular/organismal phenotypes consistent with a potential role in PD.

Of note, while the analyses of the IPDGC WES dataset and subsequent work described here were in progress, an independent family-based sequencing study identified *VPS13C* as a cause of autosomal recessive parkinsonism [21]. Although the single IPDGC subject with compound heterozygous *VPS13C* LoF alleles was published as a replicate case in that work, we retained it among the 27 candidates described here, since it was independently carried forward for all analyses detailed below.

Tolerability of gene LoF in humans and animal models

The "tolerability" of recessive LoF genotypes has important implications for understanding the genetic basis of adult-onset, age-influenced disorders such as PD. As most of the identified homozygous and putative compound heterozygous LoF genotypes are based on a single individual, we also examined for their occurrence in a large, recently published study [16] of predicted complete gene knock-outs in the Icelandic population, including 104,220 participants with imputed genotypes, based on whole genome sequencing from a subset of 2363 individuals. The Icelandic population is enriched for rare disease-causing mutations with a recessive inheritance pattern, given a strong founder effect and non-random mating patterns. Twelve of the variants that we identified are also present in the Icelandic study (Additional file 1: Table S2); however, the observed homozygote frequencies are not sufficiently high to confidently exclude them as possible PD genes and importantly, detailed phenotypic data are not publicly available for these participants. For example, 29 Icelandic participants are reported homozygous for the identical *PTCHD3* stopgain variant (c.C1426T, p.R476X) as the single PD case in our WES study. However, this is only 0.028% of the total sample set and below the reported prevalence of young-onset PD (0.041%).

We additionally examined for the presence of other LoF variants with a recessive inheritance pattern in our implicated candidate genes (Additional file 1: Table S2). For a subset of genes, we indeed identified several variants with particularly high homozygote frequencies including *OR7G3* (9.16%), *SSPO* (9.38%), and *PTCHD3* (16.55%). This is consistent with prior reports describing a homozygous deletion covering *PTCHD3* in apparently healthy individuals, consistent with a non-essential role [22]. Assuming that the variants in *OR7G3*, *SSPO*, and *PTCHD3* confer similar LoF to the alleles identified in our PD WES data, their high variant frequency makes these genes unlikely to be highly penetrant PD-risk loci.

Human genes harboring homozygous LoF variants—especially those observed recurrently in large population-based datasets—potentially identify genes that are dispensable for fetal and subsequent child development. Given the limited human phenotypic information available, we further investigated the potential tolerability for the implicated genes using a cross-species approach, performing systematic LoF analysis in the nematode, *C. elegans*. Out of the 27 candidate genes identified in our WES analysis, ten were well conserved in the *C. elegans* genome and nine had readily available RNA-interference (RNAi) reagents for LoF screening (see "Methods"). Each gene was targeted for knockdown using RNAi and we assessed for developmental lethality and survival. The results of these studies, along with other LoF data from public databases, are available in Additional file 1: Table S3. Knockdown of homologs of *DIS3* (*dis-3*), *KALRN* (*unc-73*), and *PTCHD3* (*ptr-10*) resulted in developmental arrest and/or reduced survival in *C. elegans*. Notably, homologs of *KALRN* and *DIS3* are also associated with reduced viability following genetic disruption in both *Drosophila* [23, 24] and mice [25, 26]. Thus, these results are potentially consistent with conserved, early, and/or essential developmental roles for these genes and the absence of individuals harboring homozygous LoF variants in the Icelandic cohort [16].

Since the human genome contains multiple gene paralogs for *KALRN* and *PTCHD3*, genetic redundancy might account for how LoF might be tolerated in humans but not in simple animal models. Alternatively, it is possible that the allelic variants implicated in our PD WES cohort and Icelandic study might not cause a complete LoF (i.e. genetic null) despite the algorithmic predictions, instead causing only a partial LoF. Nevertheless, these cross-species comparisons suggest essential and early developmental roles for homologs of *PTCHD3*, *DIS3*, and *KALRN*, and informing our consideration of potential contribution to adult-onset disorders, such as PD.

Variant aggregation analyses

For the 27 genes implicated based on our primary analyses of homozygous or compound heterozygous LoF variants, we additionally considered evidence for the presence of other allelic variants conferring risk for PD in our cohort. We therefore performed burden analyses leveraging our IPDGC WES data, testing two nested classes of variants: (1) a subset predicted to be deleterious (CADD > 20); and (2) all amino-acid changing missense alleles. Rare variants (MAF < 0.018) were considered either selectively or in joint models with common variants (MAF > 0.018). As detailed in Additional file 1: Table S4, the rare variant aggregation association analyses provided further evidence in support of four candidate genes: *GH2*, *PTPRH*, *UHRF1BP1L*, and *ZNF453*. Interestingly, the burden association at the *PTPRH* gene is further enhanced

when common and rare variants are simultaneously modeled.

Our analyses of LoF variants in PD exomes identify a number of promising candidate genes. However, even though a positive family history was observed for almost 40% of the cases, segregation analysis of the variants in families is not feasible, as DNA samples are not available from additional family members. Further, since most of the genes implicated contribute to single or few cases, we are unable to perform meaningful statistical comparisons, based on the limited numbers of LoF variants identified by WES in cases versus controls. As an alternative strategy, we therefore deployed a combination of cell-based and model organism functional screens to define potential links between the 27 candidate genes (Table 1) and well-established mechanisms of PD susceptibility and pathogenesis, including (1) mitochondrial health and (2) α-synuclein-mediated toxicity.

Functional prioritization: mitochondrial health

Although the mechanism of neurodegeneration in PD remains incompletely defined and may be heterogeneous, mitochondrial dysfunction has been proposed to play an important role, particularly in young onset PD [27–29]. Notably, *parkin (PARK2)*, *DJ-1*, and *PINK1*, associated with autosomal recessive, juvenile-onset Parkinsonism, have roles in mitochondrial dynamics and quality control [30]. Specifically, Parkin is an E3 ubiquitin ligase and recruited selectively to dysfunctional mitochondria with a low membrane potential [31]. Further, the neurotoxicity of α-synuclein, the primary constituent of Lewy body inclusions in PD, has also been linked to mitochondrial injury [32]. We therefore hypothesized that LoF in candidate genes identified from our analyses of WES, might similarly impact mitochondria, consistent with roles in PD susceptibility.

Therefore we quantified mitochondrial morphology after gene knockdown in BE(2)-M17 neuroblastoma cells by examining three parameters commonly used for quantification of mitochondrial morphology: mitochondrial number, axial length ratio, and roundness [33]. Cells transduced with the short hairpin RNA (shRNA) encoding a scrambled sequence were used for normalization and positive controls for mitochondrial morphology were included in each experiment. For example, knockdown of the mitochondrial fission gene dynamin 1-like (*DNM1L*), a positive control, results in elongated mitochondria and therefore decreases mitochondrial axial length ratio and roundness (Fig. 2a, b) [34]. Knockdown of 13 genes show a significant effect on at least one of the three parameters (Additional file 1: Table S5 and Table S6 and Additional file 2: Figure S1). *GPATCH2L* shows the largest increase in mitochondrial roundness, while *UHRF1BP1L* displays the largest decrease (Fig. 2c, d).

We also took advantage of a well-established Parkin translocation assay [31, 35–38] based on BE(2)-M17 human neuroblastoma cells stably expressing Parkin-GFP. As expected, upon exposure to the mitochondrial toxin and electron transport chain uncoupling reagent, CCCP, we observed robust translocation of Parkin-GFP from the cytoplasm (Fig. 3a, untreated) to the mitochondria (Fig. 3a, CCCP-SCR transduced) and this was PINK1-dependent (Fig. 3a, CCCP-PINK1 shRNA), which provides an internal, positive control in our assay. CCCP-induced Parkin accumulation was assessed by high-content microscopy and automated image analysis following systematic shRNA-knockdown of our 27 candidate genes (Fig. 3b). Based on stringent criteria (see "Methods"), six genes significantly modified Parkin translocation (Fig. 3c and d; Additional file 2: Figure S2; Additional file 1: Table S5 and Table S6), including four genes (*GPATCH2L*, *PTCHD3*, *SVOPL*, and *ZNF543*) with consistent activities in both the mitochondrial morphology and Parkin translocation assays.

Functional prioritization: α-synuclein-mediated toxicity

A wealth of evidence also supports a central role for α-synuclein-mediated toxicity in PD pathogenesis. α-synuclein aggregates, termed Lewy bodies, are the defining disease pathology and α-synuclein gene (*SNCA*) mutations, locus multiplication, and promoter polymorphisms are associated with PD susceptibility [5]. Further, expression of α-synuclein in numerous animal models including in the fruit fly [39–41], *Drosophila melanogaster*, recapitulates features of PD-related neurodegenerative pathology. Transgenic expression of α-synuclein in the fly retina leads to neurotoxic changes [39] and is amenable for detection of genetic modifiers [42, 43]. Genetic manipulation of established PD susceptibility genes, including *PARK2* [44, 45] and *VPS35* [46], modulate α-synuclein toxicity in transgenic flies, similar to findings in mammalian models [44, 47]. We therefore hypothesized that LoF in homologs of novel PD genes may similarly enhance α-synuclein-induced retinal degeneration.

Out of the 27 candidate genes implicated by our WES analyses, 13 were well-conserved in *Drosophila* (Additional file 1: Table S7). Available RNAi stocks targeting each of the 18 fly homologs (some genes had multiple conserved paralogs) were crossed to flies in which the human α-synuclein transgene was directed to adult photoreceptors using the *Rhodopsin1-GAL4 (Rh1)* driver (*Rh1 > α-synuclein*) [48]. For rapid screening, retinal neurodegeneration was monitored using the optical neutralization technique which allows assessment of retinal tissue integrity in intact, unfixed heads. In *Rh1 > α-synuclein* animals, the retina appears

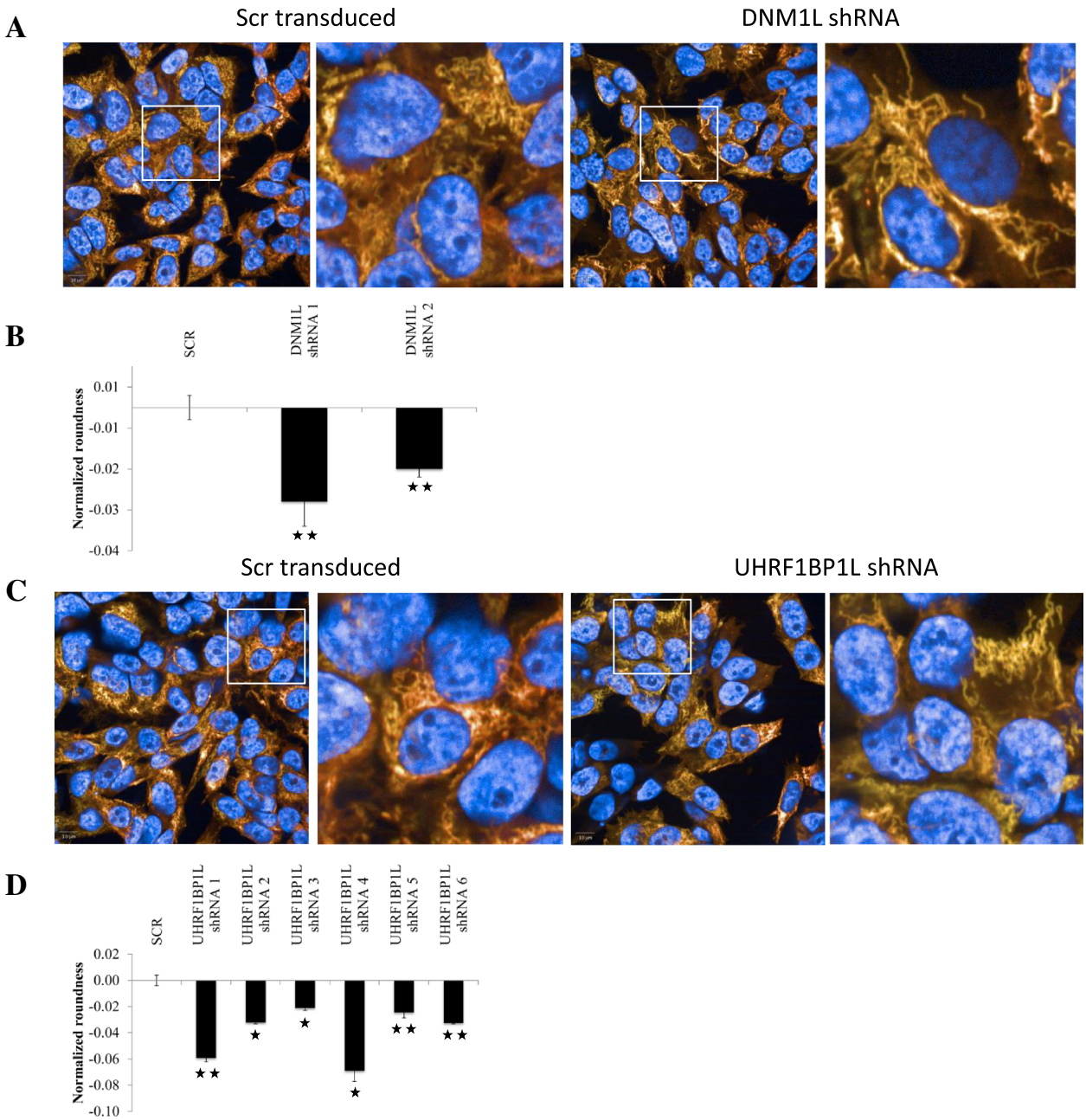


Fig. 2 High-content assay for mitochondrial morphology. Effect of *DNM1L* shRNA (**a**, **b**) and *UHRF1BP1L* shRNA (**c**, **d**). BE(2)M17 cells stained with Hoechst (*blue*; nuclei), MitoTracker CMXros, and MitoTracker Deepred (*yellow*; mitochondria). **a** Cells infected with shRNA encoding a scrambled sequence (SCR, *left panel*) and decrease in mitochondrial axial length ratio and roundness for *DNM1L* (positive control, *right panel*). **b** The graph displays normalized mitochondrial roundness. **c** Cells infected with shRNA encoding a SCR sequence (*left panel*) and decrease in number of mitochondria per cell, mitochondrial axial length ratio, and roundness for *UHRF1BP1L* (*right panel*). **d** The graph displays normalized mitochondrial roundness. Data are median values ± median absolute deviation (MAD) of N = 6 measurements. *$p < 0.05$ and **$p < 0.01$, Mann–Whitney *U* test (see "Methods"). All values were normalized to the negative control (infected with SCR shRNA) and all shRNA clones that meet the cutoff criteria are shown (**b**, **d**)

morphologically normal at 1 day (Fig. 4), but demonstrates age-dependent degeneration leading to progressive vacuolar changes, rhabdomere loss, and culminating with extensive tissue destruction by 30 days. At the 15-day time point selected for screening, only mild, if any, retinal pathology is detectable on most histologic sections, consistent with a weakly penetrant degenerative phenotype following optical neutralization (mean penetrance ~25%) (Fig. 4). However, co-expression of RNAi targeting fly homologs of four candidate genes (*ARSB*, *TMEM134*,

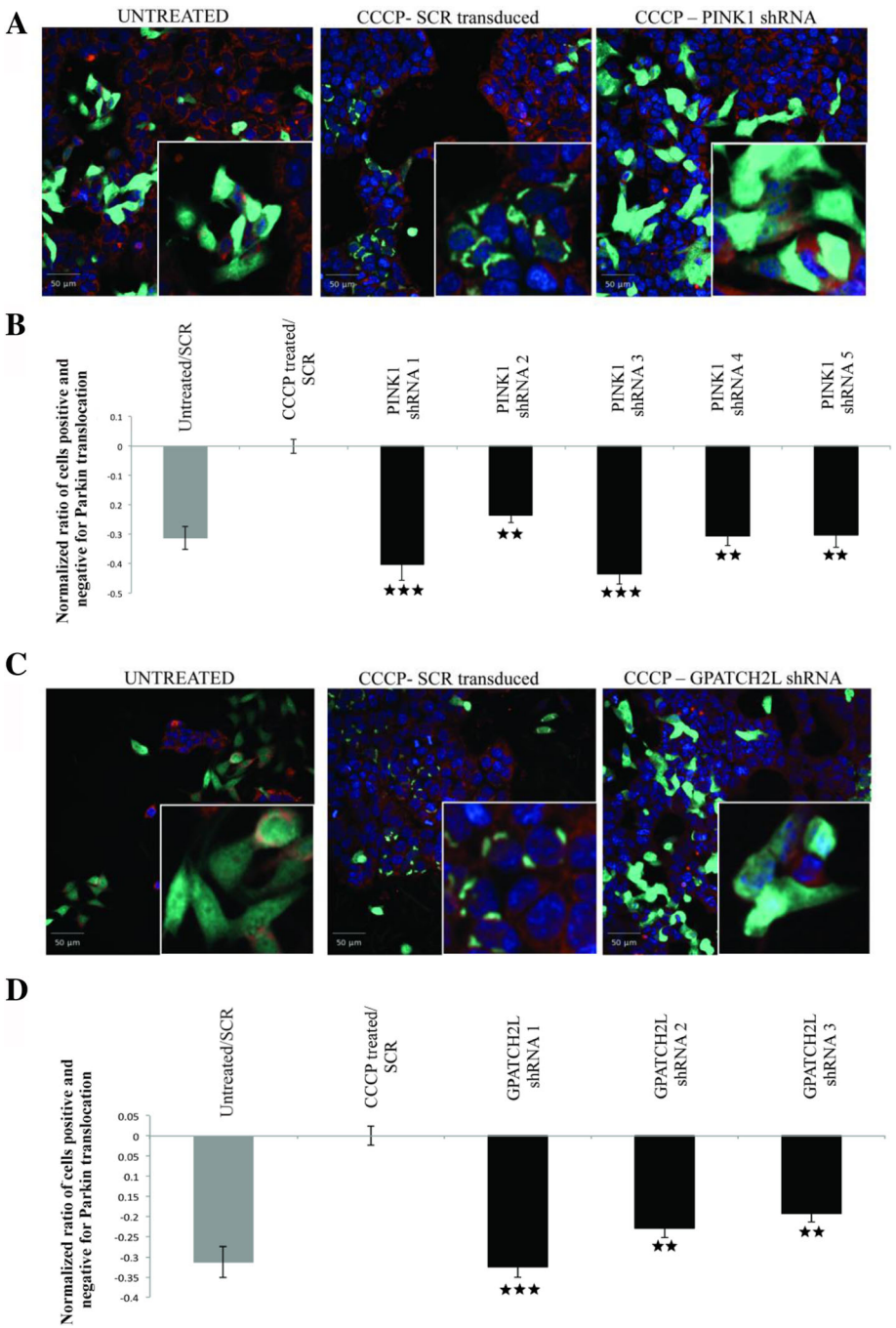


Fig. 3 High content assay for Parkin translocation. Effect of *PINK1* shRNA (**a**, **b**) and *GPATCH2L* shRNA (**c**, **d**). **a**, **c** Cells are labeled for nuclei (*blue*; Hoechst), Parkin-GFP (*green*), mitochondria (*red*, Mitotracker Deepred). Untreated cells infected with shRNA encoding a scrambled sequence show absence of puncta (*left panel*). Cells infected with a scrambled sequence but treated with CCCP show a significant increase in puncta formation (*middle panel*). Infection of cells with shRNA targeting *PINK1* or *GPATCH2L* prevents the accumulation of Parkin on mitochondrial (*right panel*). **b**, **d** The *graph* displays the normalized ratio of cells positive for translocation and cells negative for parkin translocation. All values were normalized to the negative control (CCCP treated infected with shRNA encoding a scrambled sequence). Data are median values ± median absolute deviation (MAD) of N = 6 measurements. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$, Mann–Whitney *U* test (see "Methods"). All shRNA clones that meet the cutoff criteria (see "Methods") are shown

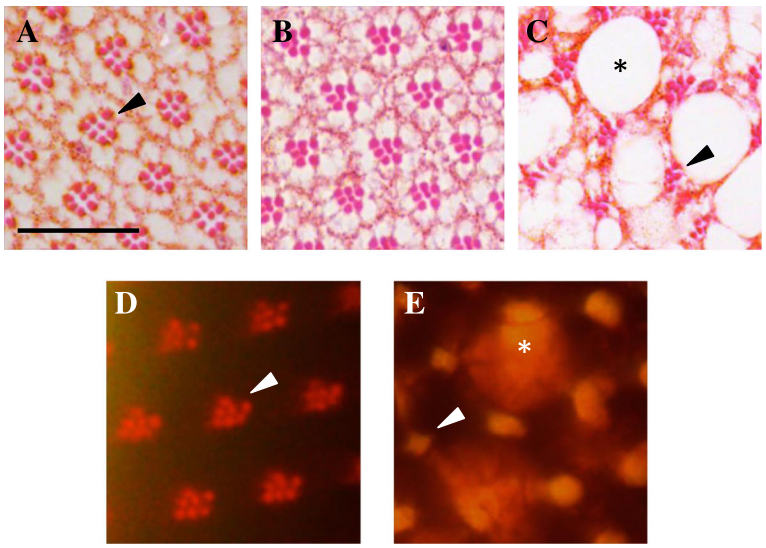


Fig. 4 α-synuclein-induced retinal degeneration and screening assays in *Drosophila* transgenic animals. Tangential sections through the fly retina stained with hematoxylin and eosin reveal the ordered ommatidial array in control animals (**a** *Rh1-GAL4 / +*). Each ommatidia consists of a cluster of eight photoreceptive neurons (seven visible at the level examined). The photoreceptors each contain a single rhabdomere, the specialized organelle subserving phototransduction, giving the ommatidia cluster its characteristic appearance (arrowhead). Expression of *α-synuclein* in adult photoreceptors (**b**, **c** *Rh1-GAL4 / +; UAS-α-synuclein / +*) causes age-dependent, progressive retinal degeneration. Compared to one-day-old *Rh1 > α-synuclein* flies (**b**), histologic sections in 30-day-old animals (**c**) demonstrate rhabdomere/cell loss and substantial vacuolar changes (*asterisk*). The pseudopupil preparation allows visualization of rhabdomeres (*arrowhead*) in intact, unfixed intact fly heads, permitting medium-throughput screening for progression of *α-synuclein*-induced retinal pathology. Compared to controls (**d** *Rh1-GAL4 / +*), in 30-day-old *α-synuclein* transgenic animals (**e** *Rh1-GAL4 / +; UAS-α-synuclein / +*) rhabodomeres frequently appear indistinct (*arrowhead*) and vacuolar changes disrupt light refraction (*asterisk*). Representative control histology (**a**) and pseudopupil images (**d**) are shown for 15-day-old animals, the timepoint used for screening, in order to facilitate comparison with Fig. 5. Scale bar: 20 μm

PTPRH, and *VPS13C*) was observed to robustly enhance α-synuclein-mediated neurodegeneration in the retina (mean penetrance ~ 75%; Additional file 1: Table S8).

All candidate enhancers of α-synuclein identified using the screening assay were further confirmed based on retinal histology, demonstrating accelerated pathologic changes with a significantly increased overall extent and severity of degeneration compared to *Rh1 > α-synuclein* controls without RNAi transgenes present (Fig. 5). Importantly, when each of these genes were targeted under similar experimental conditions (*Rh1 > RNAi*), but independent of α-synuclein expression, we did not observe any significant retinal pathology in 15-day-old animals (Fig. 5). Therefore, within the *Drosophila* α-synuclein transgenic model system, the implicated LoF enhancers appear consistent with synergistic (non-additive) effects on α-synuclein-mediated retinal degeneration. Since increased α-synuclein expression levels are one important mechanism of PD susceptibility [5], western blot analyses were performed to determine whether any of the identified genetic enhancers alter α-synuclein protein levels. However, following RNAi-mediated knockdown, none led to significant changes (Additional file 2: Figure S3). Thus, we hypothesize potential interactions with more downstream mechanisms of α-synuclein neurotoxicity.

For 3 out of 4 candidate enhancers (*ARSB*, *VPS13C*, *PTPRH*), available siRNAs permitted additional testing of gene homologs as candidate modifiers in an established *C. elegans* model of α-synuclein toxicity [49]. However, no significant differences were detected in the α-synuclein-induced locomotor phenotype observed in one-week-old worms following knockdown of these genes (Additional file 2: Figure S4). We speculate that these contradictory results might stem from differences in assay sensitivity and/or tissue-specific toxic mechanisms as the fly and worm models are based on α-synuclein expression in the retina versus muscle, respectively.

Of the four genes discovered to interact with α-synuclein toxicity in *Drosophila*, we were able to obtain additional genetic reagents, including classical LoF alleles, for the two homologs of *PTPRH*: *Ptp10D* and *Ptp4E*. In our screen, two independent RNAi lines targeting *Ptp10D* robustly enhanced α-synuclein toxicity, but only one of the two available lines for *Ptp4E* met our threshold criteria (Additional file 1: Table S8). Interestingly, prior studies in *Drosophila* suggest that *Ptp10D* and *Ptp4E* are the result of a gene duplication event and these genes show evidence of partial functional redundancy, including for nervous system phenotypes [50]. Consistent with this, we found that transheterozygosity for

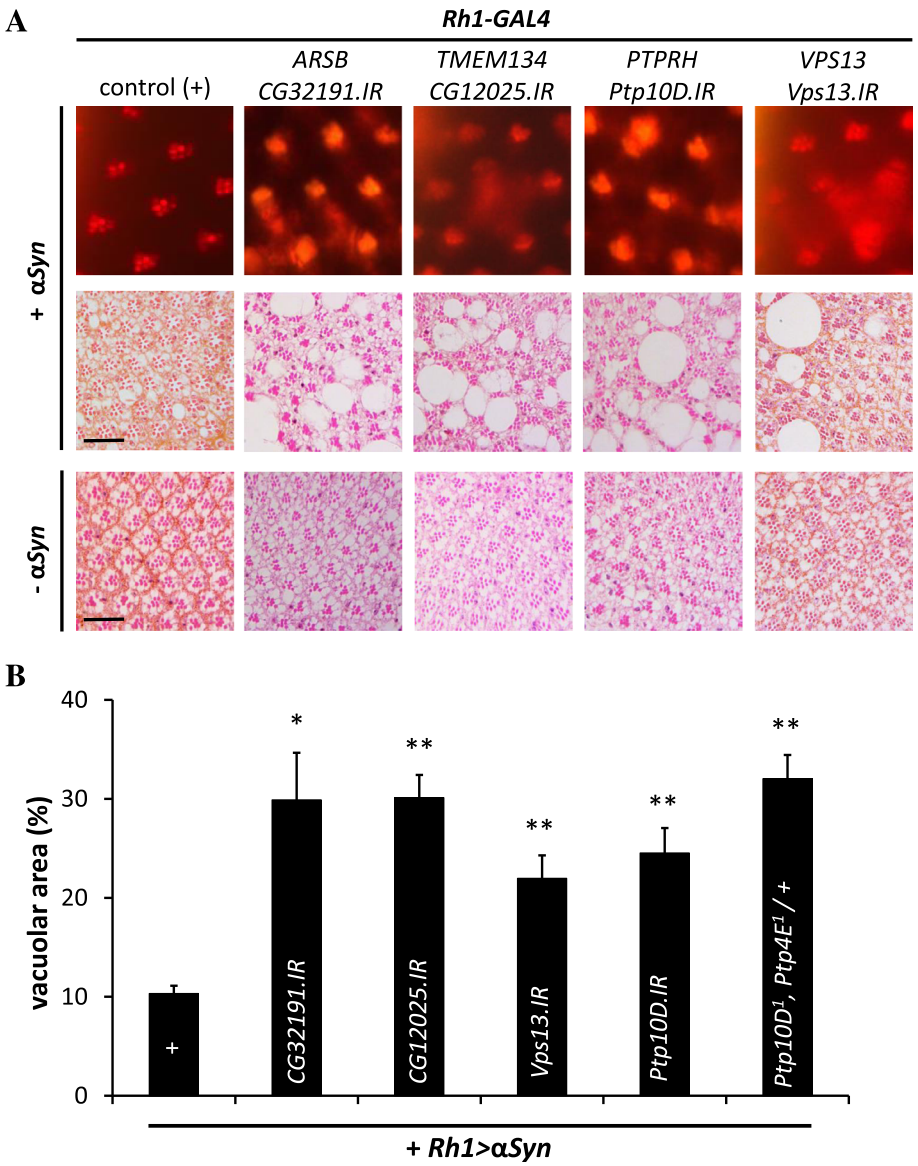


Fig. 5 PD gene candidates harboring LoF variants enhance α-synuclein toxicity in *Drosophila*. Conserved fly orthologs of human genes discovered from WES analysis were targeted with RNAi (IR) and screened for enhancement of α-synuclein pathology using the pseudopupil assay (**a** *top row*). For each line evaluated, the severity of retinal degeneration was scored based on penetrance of the α-synuclein pseudopupil phenotype and enhancers required consistent results for at least two independent RNAi lines (see Additional file 1: Table S8). Representative results from the primary screen are shown for controls (*Rh1-GAL4 / +*; *UAS-α-synuclein / +*) and one IR line each for the implicated enhancers [Human Gene-Fly Ortholog (experimental genotype shown)]: *ARSB-CG32191* (*Rh1-GAL4 / +*; *UAS-α-synuclein / UAS-CG32191.IR.v14294*), *TMEM134-CG12025* (*Rh1-GAL4 / UAS-CG12025.IR.v104336*; *UAS-α-synuclein / +*), *PTPRH-Ptp10D* (*Rh1-GAL4 / UAS-Ptp10D.IR.v1102*; *UAS-α-synuclein / +*), and *VPS13-Vps13* (*Rh1-GAL4 / UAS-Vps13.IR.HMS02460*; *UAS-α-synuclein / +*). At the 15-day-old time point, *Rh1 > α-synuclein* causes a weakly-penetrant pseudopupil phenotype and mild histopathologic changes which are amenable to modifier screening (compare with Fig. 4, panels c and e). Enhancers identified in the primary screen were confirmed based on retinal histology (**a** *middle row*) and demonstrated increased tissue destruction and disorganization. Activation of RNAi was not associated with any significant retinal degeneration in the absence of α-synuclein co-expression (**a** *bottom row*, *Rh1-GAL4 / IR transgene*). Scale bars: 20 μm. **b** Enhancement of α-synuclein-induced retinal degeneration was quantified based on the extent of vacuolar changes (area occupied by vacuoles / total retinal area). For quantification, three animals were examined per genotype. For *PTPRH*, additional confirmation was obtained by evaluating flies doubly heterozygous for strong alleles of the paralogs *Ptp10D* and *Ptp4e* (see also Additional file 2: Figure S5). Statistical comparisons were made using unpaired *t*-tests. *Error bars* are based on Standard Error of the Mean. *$p < 0.05$; **$p < 0.01$

strong (null) alleles of both genes enhanced α-synuclein-induced retinal degeneration (*Ptp4E*^{1}, *Ptp10D*^{1} / +; *Rh1-Gal4* / +; *UAS-α-synuclein* / +); whereas heterozygosity for either allele in isolation showed no significant enhancement (Fig. 5b and Additional file 2: Figure S5).

Genetic replication of candidate PD genes from WES

We next evaluated our 27 gene candidates in additional available genetic datasets including: (1) an independent exome sequencing dataset from the Parkinson Progression Markers Initiative (PPMI) project [51]; (2) a whole-genome sequencing dataset including PD index cases of a Dutch genetic isolate belonging to the Genetic Research in Isolated Population (GRIP) program [52]; (3) an independent NeuroX exome array dataset [7, 53]; and (4) a large PD GWAS dataset [53]. Within the PPMI exome dataset, including 462 PD cases and 183 controls, evidence supporting replication was discovered for two genes, in which we identified the identical variants from the IPDGC discovery exome dataset (Additional file 1: Table S9). A PD case from PPMI carries the same homozygous stopgain variant (p.R362X) in *GPATCH2L* as observed for an IPDGC case. Although the age of onset differs 20 years between these two PD cases (47 and 68 years for the IPDGC and PPMI patients, respectively), they share similar asymmetric clinical symptoms at onset, which are characterized by resting tremor, bradykinesia, and rigidity. Furthermore, both PD cases have a father diagnosed with PD, implying the variant to be highly penetrant. We excluded the possibility that these two PD cases might be related by computing pairwise genetic relationships [54] from common SNPs (MAF ≥ 0.01). No evidence of relatedness was observed ($A_{jk} = -0.0018$). Based on ExAC, only one (0.003%) out of 32,647 European individuals has this same homozygous variant. The observation of two PD cases (0.12%) of our 1610 studied PD patients (1148 IPDGC WES plus 462 PPMI WES) with this *GPATCH2L* mutation is consistent with a 40-fold enrichment in our PD cohort. The second gene harboring an identical LoF variant is *FAM83A*. The p.G86X variant in *FAM83A*, detected within an IPDGC participant with sporadic PD diagnosed at the age of 28 years, was also observed in a single sporadic PD case from PPMI with an age of onset of 62 years. These *FAM83A* carriers presented with similar symptoms, including bradykinesa, rigidity, and resting tremor. In both datasets, the p.G86X allele is predicted to be in *trans* with another variant: p.R347X or p.V137G in PPMI and IPDGC, respectively.

The second genetic independent dataset that was investigated included a whole-genome sequencing study (39 PD index cases and 19 controls) of a genetic GRIP isolate from the Netherlands, focusing on variants within our candidate genes that were present in at least two PD index cases and absent in controls. We identified a heterozygous missense variant (NM_001127444:c.1176G > T:p.L392F) in *CD36* for three PD index cases. Although not consistent with a recessive inheritance model, this variant has not been observed in the 60,706 unrelated individuals of the ExAC database, suggesting potential enrichment in PD cases. These heterozygote variant carriers have a substantial higher age of onset (range, 61–79 years) in comparison to the PD patient (age of onset, 38 years) with the putative compound heterozygous variant within the discovery WES dataset. This observation supports an additive model of pathogenicity, implying more severe disease onset when two alleles are affected. Further, *CD36* (p.L392F) is predicted to represent the top 1% most harmful variants within the genome (CADD score = 23.3). In the IPDGC discovery dataset, the discovered compound heterozygous variants, p.Q74X and p.P412S (Table 1), are also predicted to be strongly deleterious (CADD scores of 26.5 and 25.9, respectively).

We next interrogated the independent IPDGC NeuroX dataset, including genotypes from 6801 individuals with PD and 5970 neurologically healthy controls. NeuroX is a genotyping array that includes pre-selected exonic variants and is therefore not suitable to search for the identical recessive LoF variants implicated by our WES analyses. Instead, we examined the burden of multiple variant classes within the 27 candidate genes, following the same variant categories as for the original IPDGC WES dataset (Additional file 1: Table S10). When only considering variants predicted to be deleterious (CADD > 20), an association is detected for *UHRF1BP1L* with PD risk ($p = 0.005$). This gene also shows an association with PD in the IPDGC WES dataset when performing a similar burden analysis considering missense variants (see above, $p = 0.016$). Using the NeuroX dataset, we additionally confirmed the enrichment of rare *PTPRH* variants in participants with PD (WES: $p = 0.034$, NeuroX: $p = 0.045$). Furthermore, *VPS13C* and *ARSB* show significant associations to PD when considering the joint effect of all variants, both common and rare (Additional file 1: Table S10).

Leveraging available IPDGC GWAS data (13,708 cases/95,282 controls), we next assessed for potential common variant association signals ($p < 1 \times 10^{-4}$) using a 1-Mb genomic window centered on each of the 27 candidate genes. Three loci (*VPS13C*, *PCDHA9*, and *TCHHL1*) showed evidence consistent with an association peak (Additional file 2: Figure S6). A genome-wide significant association at the *VPS13C* locus, was in fact recently reported [7]; the best SNP (rs2414739, $p = 3.59 \times 10^{-12}$) maps ~150 kb distal to *VPS13C*. Based on local patterns of linkage disequilibrium defined by Hapmap (Additional file 2: Figure S6), it is unlikely that rs2414739 is a proxy for

p.E3147X or similar LoF variants in *VPS13C*; however, it might be possible that the SNP influences *VPS13C* expression by affecting the long non-coding RNA lnc-VPS13C-1 [55] in which the SNP is located. The other two candidate association peaks, adjacent to *PCDHA9* and *TCHHL1*, are considerably weaker signals (rs349129 = 1.40×10^{-5} and rs7529535 = 7.66×10^{-5}, respectively) and given the distances (~500 kb) many other candidate genes are potentially implicated.

In sum, we identify additional genetic evidence consistent with replication for seven genes (*GPATCH2L*, *FAM83A*, *CD36*, *UHRF1BP1L*, *PTPRH*, *ARSB*, and *VPS13C*) that were implicated by our WES analysis, of which five (*GPATCH2L*, *UHRF1BP1L*, *PTPRH*, *ARSB*, and *VPS13C*) are further validated based on functional evidence from PD-relevant experimental models.

Transcriptomics-based functional exploration

Lastly, we examined each candidate gene from our WES analysis for co-expression with established PD susceptibility gene in expression networks derived from human substantia nigra, leveraging available data from the United Kingdom Brain Expression Consortium (UKBEC) and the Genotype-Tissue Expression project [56]. Of the 27 candidate genes, seven were not sufficiently expressed in substantia nigra on the basis of UKBEC. Except for *DIS3*, these genes were also expressed poorly in publicly available data of the Genotype-Tissue Expression (GTEx) project [56]. Consequently, expression values for these genes were not used for construction of the UKBEC gene co-expression network (GCN). The remaining 20 genes were assessed for co-expression with known Mendelian PD genes (*ATP13A2*, *FBXO7*, *LRRK2*, *PARK2*, *PARK7*, *PINK1*, *RAB39B*, *SNCA*, and *VPS35*) using the UKBEC GCN (Additional file 1: Table S11 and Additional file 2: Figure S7). This approach highlighted three genes (*UHRF1BP1L*, *GPATCH2L*, and *PTPRH*) and the implicated networks were further interrogated based on gene set enrichment analysis using gene ontology (GO) terms to denote potential functions. *UHRF1BP1L* was co-expressed with *SNCA*, *PINK1*, *GBA*, and *ATP13A2* in a network significantly enriched for genes with roles in synaptic transmission ($p = 2.27 \times 10^{-11}$) as well as astrocytic ($p = 8.18 \times 10^{-8}$) and dopaminergic neuronal markers ($p = 3.98 \times 10^{-46}$). *GPATCH2L* was co-expressed with *PARK7* in a network enriched for other neuronal genes ($p = 3.41 \times 10^{-12}$) with cellular roles in metabolism of macromolecules ($p = 3.82 \times 10^{-15}$). Lastly, *PTPRH* was assigned to a co-expression module including *FBX07* and enriched for oligodendrocyte markers ($p = 8.69 \times 10^{-22}$). Importantly, the implicated modules were preserved (Z.summary > =10) in the independent GTEx dataset.

Discussion

We report the results from WES analysis in the largest PD cohort studied to date. Assuming a recessive inheritance model, we identified 27 candidate genes harboring rare homozygous or compound heterozygous LoF variants. With the exception of *ARSB*, we did not identify recurrent recessive alleles in more than a single PD case. This result—potentially consistent with a highly heterogeneous genetic etiology for PD—creates significant barriers for statistical confirmation and genetic replication of novel PD susceptibility loci. Additional genetic samples were not available for segregation analysis and given the rarity and heterogeneity of the implicated alleles, definitive human genetic replication would likely require very large sample sizes, including many thousands of PD cases with either WES or gene resequencing. We therefore coupled our WES analyses with functional studies in both mammalian cells and experimental animal models, including *Drosophila* and *C. elegans*, in order to prioritize genes for future study. Our results highlight 15 out of the 27 gene candidates that interact with mitochondrial dynamics and five loci that enhance α-synuclein-mediated neurodegeneration. As discussed below, while these results highlight a promising subset of genes with potential links to PD-relevant mechanisms, we cannot exclude contributions from other implicated genes/variants. All of these data, including promising variants from the human genetic analyses and results of functional studies, will be a valuable resource for future investigations of PD genomics. Analyses of several other WES and complementary large-scale, genetic datasets provide additional evidence supporting replication for 7 out of 27 genes. Evidence from human genetics and functional studies converge to most strongly implicate five gene candidates discussed below; however, further investigation will be required to definitively link each of these loci to PD susceptibility and elucidate the relevant mechanisms. Nearly all of these genes are robustly expressed in brain [56], including the substantia nigra, thereby consistent with their implication in PD. A subset (*GPATCH2L*, *UHRF1BP1L*, and *PTPRH*) are co-expressed with established Mendelian PD genes in the substantia nigra based on analyses of UKBEC and GTEx expression data. In sum, our results define several promising new susceptibility loci candidates for further investigation and illustrate a powerful, integrative discovery strategy for future, large-scale PD genomic studies.

Mitochondrial mechanisms have been strongly implicated in PD risk and pathogenesis [28, 30]. Following shRNA-mediated knockdown, 15 candidate recessive loci identified in our WES dataset showed effects on mitochondrial morphology and Parkin translocation to mitochondria in cell culture. We focus our initial

discussion on three genes, *GPATCH2L*, *UHRF1BP1L*, and *VPS13C*, for which we discovered additional genetic evidence consistent with replication in independent cohorts. In the IPDGC cohort, a single PD case was identified with a homozygous stopgain variant (p.R362X) in *GPATCH2L* and a second individual with the identical, rare genotype was discovered in PPMI. This variant is reported with a low frequency of 0.003% in ExAC. Although minimal clinical or demographic information is available within ExAC, this finding is compatible with population prevalence estimates for PD [20]. Nevertheless, genotyping of p.R362X in additional large PD case and control cohorts will be required to definitively establish an association with PD susceptibility. *GPATCH2L* knockdown both increased mitochondrial roundness and impaired Parkin translocation. The encoded protein, GPATCH2L, which has not previously been studied, contains a glycine-rich RNA-binding motif, the "G-patch" domain [57]. *GPATCH2*, a paralog of *GPATCH2L*, is upregulated in cancer cells, localizes to the nucleus where it interacts with RNA-processing machinery, and manipulation in culture alters cell proliferation [58, 59]. Notably, *GPATCH2L* is non-conserved in either the *C. elegans* or *Drosophila* genomes, precluding study of this candidate in these models. While our results using cellular assays implicate *GPATCH2L* in mitochondrial quality control mechanisms, further follow-up studies in mammalian model systems will be needed to confirm a role in PD pathogenesis.

Another promising gene, *UHRF1BP1L*, harbored a homozygous stopgain variant (p.K1376X) in a single IPDGC case. This is a novel variant, based on its absence from the ExAC cohort. Additional support for *UHRF1BP1L* as a *bona fide* PD locus comes from complementary analyses in both the IPDGC WES and NeuroX datasets, documenting a burden of rare missense and LoF variants in association with disease risk. In the UKBEC, *UHRF1BP1L* was associated with a substantia nigra co-expression module including both *SNCA* and *PINK1*, reinforcing potential links with established PD genetic mechanisms. Indeed, *UHRF1BP1L* knockdown cause sharply reduced mitochondrial numbers and altered morphology. Interestingly, *UHRF1BP1L* encodes a protein bearing an amino terminal homologous to yeast VPS13 and studies in cell culture provide support for a role in retrograde transport from the endosome to the trans-Golgi network [60].

Notably, LoF in human *VPS13C* was also implicated by our analyses of IPDGC WES data and knockdown disrupted mitochondrial morphology. Besides the single IPDGC case, several families with autosomal recessive early onset Parkinsonism and dementia due to *VPS13C* were recently reported [21] and this locus also harbors common PD susceptibility variants based on GWAS [7]. Our findings of a potential mitochondrial role for *VPS13C* agree with those of Lesage et al. who additionally reported that VPS13C localizes to the outer membrane of mitochondria and LoF was associated with reduced mitochondrial membrane potential, fragmentation, and increased Parkin-dependent mitophagy. Importantly, *VPS35*, which causes autosomal dominant, late-onset PD, is similarly involved in endosomal trafficking [61] and has also recently been implicated in mitochondrial dynamics [62], including interactions with Parkin [63]. Like *UHRF1BP1L*, *VPS13C* and *GPATCH2L* are expressed in the brain, including within the substantia nigra; however, additional work will be needed to define their functions, including potential interactions with other established disease genes (e.g. *VPS35*, *parkin*) and requirements for mitochondrial maintenance.

Based on functional screening in *Drosophila*, four candidate genes from our WES analyses were implicated as LoF enhancers of α-synuclein neurotoxicity, which also has a central role in PD pathogenesis. We discuss the three genes (*VPS13C*, *PTPRH*, and *ARSB*) where additional human genetic evidence supports replication. Interestingly, besides its requirement for mitochondrial maintenance, RNAi-mediated knockdown of *Drosophila Vps13* enhanced α-synuclein toxicity. In the single reported *VPS13C* PD case with a completed autopsy, neuropathological findings included abundant α-synuclein aggregates in both the brainstem and cortex [21]. Thus, *VPS13C* and associated endosomal sorting pathways (including *VPS35*) may represent a point of convergence for mitochondrial and α-synuclein-mediated PD mechanisms. Consistent with this, evidence for the impact of α-synuclein toxicity on mitochondria has recently emerged [28], including from studies in mammals [64].

In the IPDGC WES cohort, a single PD case was discovered with compound heterozygous LoF variants in *PTPRH* (p.Q887X and p.E200X). Both variants were also observed at low frequencies in the ExAC database (0.039% and 0.003%, respectively); however, they each met our prespecified threshold of < 1% based on the population prevalence of PD. Encoding a receptor protein tyrosine phosphatase, *PTPRH* (also called *SAP-1*) was first discovered for its potential association with gastrointestinal cancers [65, 66] and remains poorly studied in the nervous system context. In studies of both vertebrates and invertebrates, receptor protein tyrosine phosphatases have been strongly implicated as key neural cell adhesion receptors, with roles in neurodevelopment and synaptic function, and other members of this family have been implicated in numerous neuropsychiatric disorders [67]. In *Drosophila*, RNAi-mediated knockdown of the conserved *PTPRH* ortholog, *Ptp10D*, enhanced α-synuclein-triggered retinal degeneration, but was not

associated with substantial neurotoxicity independent of α-synuclein expression. *Ptp10D* mutant flies are also viable and fertile but demonstrate long-term memory deficits in behavioral assays [68]. More recent studies further implicated Ptp10D in neural-glial interactions during development of the central nervous system [69], potentially consistent with our findings that human *PTPRH* participates in a substantia nigra gene co-expression network strongly enriched for oligodendrocyte markers. Besides our discovery of homozygous LoF in *PTPRH*, further analyses of the IPDGC WES dataset, and the substantially larger, independent NeuroX cohort, implicate a burden of rare variants at this locus in association with PD susceptibility.

α-synuclein-induced neurodegeneration was also enhanced by knockdown of *CG32191*, a *Drosophila* homolog of *ARSB*. RNAi transgenic lines targeting three other conserved fly *ARSB* homologs showed consistent interactions with α-synuclein (Additional file 1: Table S7 and Table S8). In the IPDGC cohort, we discovered four PD cases homozygous for a variant predicted to disrupt splicing of exons 1 and 2 in *ARSB*. Although the identified variant has not previously been documented in ExAC, we identified a single IPDGC control homozygote. Additional evidence supporting association of the *ARSB* gene with PD susceptibility comes from burden analysis in the independent NeuroX cohort. The surprisingly common *ARSB* splicing variant (rs138279020, MAF = 0.065 in IPDGC) is a single nucleotide insertion allele within a poly-A repeat, which we speculate might lead to inefficient capture in prior WES and possibly explain the absence of this variant from ExAC and the 1000 Genomes project reference. All four PD cases in our data with the homozygous *ARSB* splicing variant were confirmed by Sanger sequencing. Intriguingly, mutations in *ARSB*, encoding the lysosomal enzyme Arylsulfatase B, are associated with the recessive lysosome disorder, Mucopolysaccharidosis type VI (MPS VI, also called Maroteaux-Lamy syndrome), in which the glycosaminoglycan, dermatan sulfate, accumulates causing skeletal dysplasia and other heterogeneous manifestations [70]. Substrate accumulation and associated cellular stress has been reported to induce markers of impaired autophagy and mitochondrial dysfunction in *ARSB* deficient fibroblasts from MPSVI patients, as in other lysosomal disorders [71, 72]. Importantly, Maroteaux-Lamy can be characterized by minimal or even absent clinical signs, leading to incidental discovery or diagnosis in adulthood, and such mild phenotypes have been suggested to accompany partial LoF with preserved low-level ARSB enzymatic activity [70, 73, 74]. Similar genotype–phenotype relationships have been documented for other lysosomal-storage disorders, including Gaucher's disease, which has established links with PD risk [75, 76]. While a full accounting is outside the scope of this study, at least one of the three IPDGC cases for which records were available revealed clinical features potentially overlapping with MPS VI.

The strengths of our study include the largest PD WES discovery dataset assembled to date, complementary analyses in independent available cohorts to establish replication, and integration of promising human genetic findings with multiple functional assays relevant to PD mechanisms. Nevertheless, we also make note of several inherent limitations. In order to prioritize candidate genes for initial investigation, assumptions were made concerning the specific inheritance model (recessive) and stringent criteria were employed for variant filtering. In the future, it will be important to also consider the possibility of dominantly acting alleles; however, this substantially increases the number of variants to consider and also potentially complicates functional studies (i.e. compared with LoF screening using RNAi). Our study design excluded consideration of many non-synonymous variants that could potentially cause loss (or gain) of gene function, along with certain non-truncating, frameshifting alleles (see "Methods"). Even with fairly stringent criteria for variant filtering and the assumption of recessive inheritance, we found evidence for substantial etiologic heterogeneity. Improved confidence for the discovery of PD causal variants will likely come from PD WES cohorts with significantly enhanced sample sizes, as well as increased numbers of adult controls, including those with careful neurological assessments to exclude mild PD symptoms. Indeed, most of the variants implicated by the IPDGC WES cohort were represented at low frequencies within the largest available public database, ExAC [77, 78]; however, we have no information about potential PD manifestations in such individuals or even participant age.

Since no single cellular or animal experimental model is expected to universally recapitulate all potential facets of disease biology, we note that the employed functional screening assays are potentially liable to false-negative or false-positive findings. Importantly, experimental evidence of a genetic interaction with either mitochondrial dynamics or α-synuclein-mediated neuronal injury in our screening assays cannot in isolation confirm a role in disease causation, but rather serves to prioritize genes for future investigation. Out of the 27 candidate genes implicated in the IPDGC WES discovery analysis, 14 were insufficiently conserved for follow-up in α-synuclein transgenic flies. While simple animal models, including *Drosophila* or *C. elegans*, have made important contributions to our understanding of PD pathogenesis, selected mechanisms, such as the potential role of adaptive immunity or basal ganglia circuit dysfunction, cannot be addressed in invertebrates [79, 80]. We were

unable to confirm our findings from *Drosophila* in a published *C. elegans* model of α-synuclein toxicity. In the future, it will also be important to examine potential genetic interactions in other PD models, including *LRRK2* transgenic flies or those containing mutations in other PD loci, such as *VPS35* or *parkin*. While neuroblastoma cells offer the convenience of robust mitochondrial readouts, they are limited by their undifferentiated, transformed state distinct from that of postmitotic neurons. In the future, human-induced pluripotent stem cells, including those derived from individuals with PD, can be differentiated into dopaminergic or other neuronal types and potentially deployed for functional screening strategies. Additionally, genome-editing technologies may facilitate systematic functional evaluation of candidate disease-associated variants of unknown significance.

Conclusions

We have identified five excellent PD gene candidates (*GPATCH2L*, *UHRF1BP1L*, *PTPRH*, *ARSB*, and *VPS13C*), harboring homozygous or compound heterozygous LoF variants in PD exomes, demonstrating functional interactions with mitochondrial and/or α-synuclein-mediated mechanisms, and supported by evidence of replication in independent human datasets. The recent report [21] of additional PD families segregating LoF mutations in *VPS13C* along with other experiments supporting a role in mitochondrial mechanisms significantly strengthens the evidence in support of this gene in PD and validates our overall approach. These loci are well-suited for future efforts directed at human genetic replication and in-depth functional dissection. We also make available results, including findings from human genetic analyses and functional studies in most cases, on 22 other promising loci. These data will serve as a valuable reference for ongoing and future PD genetic studies. More broadly, our approach of integrating high-throughput sequencing in PD case/control cohorts with parallel systematic screening in cells and model organisms for functional prioritization exemplifies a powerful experimental strategy with great promise for future genomic studies of PD and other human disorders.

Methods

Genetic analyses

Whole-exome sequencing

WES was performed on 1148 PD cases and 503 neurologically healthy controls of European descent. All participants provided written informed consent. Relevant local ethical committees for medical research approved participation in genetic studies. If PD patients were prescreened for known pathogenic mutations, they were excluded for exome sequencing when having such a variant. The cases were diagnosed with PD at a relatively young average age of 40.6 years (range, 6–56 years), of which approximately 37% reported a positive family history. The neurologically healthy controls are on average 48.2 years of age (range, 10–97 years). A more extensive overview of demographic information is reported in Additional file 2: Figure S8.

Due to improvements of the exome sequencing protocol over time, the exome sample libraries were prepared with different capture kits. For this study, three different capture kits were used: Illumina TruSeq (San Diego, CA, USA) (62 Mb target); Roche (Basel, Switzerland) Nimblegen SeqCap (44.1 Mb target); and Agilent (Santa Clara, CA, USA) SureSelect (37.6 Mb target), which captured 96%, 81%, and 71% of the targeted exome at least ten times, respectively (Additional file 1: Table S12). Exome libraries were sequenced on a HiSeq 2000 (Illumina, San Diego, CA, USA). The Burrows Wheeler Aligner MEM v0.7.9.a [81] was used to align the 100-bp paired-end reads to the human reference genome build hg19. We called the single nucleotide variants (SNVs) and insertions/deletions (indels) for all samples simultaneously using Genome Analysis Toolkit (GATK) 3.x [82], followed by the exclusion of low-quality variant calls not passing the default GATK filters. Individual genotypes were removed with genotype quality Phred-scores below 40. ANNOVAR [83] was applied to annotate the variants with information concerning variant type (valid annotations when Refseq in concordance with UCSC), MAF in the general population, and predictions of the variant's effect on gene function, implementing CADD [84].

Variant identification in IPDGC WES dataset

Considering the worldwide prevalence of 0.041% for PD in the age range of 40–49 years [20], we selected rare variants with a MAF < 1% (corresponding to a homozygous frequency of 0.01%) in the European population. Because the specified 0.041% of the population with young-onset Parkinson's disease (YOPD) is not caused by one shared genetic factor, we expect a homozygous frequency of 0.01% to be an adequate cutoff, which would be able to determine variants present in approximately 25% of the YOPD population. As a comparison to the most common genetic cause of YOPD, *parkin* [85], the most frequent mutation is an exon 3 deletion, which has been identified in 16.4% of YOPD patients [86]. Using ANNOVAR [83], all variants were annotated with MAF information of ESP6500si (European American population) [87], 1000 Genomes Project (European population of April 2012 version) [88], and the ExAC browser (non-Finish European population) [77, 78]. When no public allele frequency was available for homozygous variants, the in-house control dataset of 503

individuals was used as a reference for the general population. Homozygous variants were excluded when being common (>1%) in controls or having a relative higher frequency in controls than in cases. KGGseq [89] was used to count the number of homozygous variants for the cases versus controls.

In addition to the population allele frequency filters, we only selected SNVs and indels affecting the position of the stop codon or located at a splice site (within 2 bp of splicing junction), which are variants expected to result in a loss of gene function. As the aim of this study was to validate our approach to identify high promising PD candidate genes, rather than discovering all putative PD genes present within our WES dataset, we set a conservative selection criteria by only including frameshifts that caused an immediate stopcodon at the position of the indel. Splice-site variants were only considered when being adjacently located to an exon that is coding for amino acids. As a final filter for the homozygous variants, we manually excluded variants that failed GATKVQSR and hard filtering. Quality predictions based on the ExAC database are more adequate, as it includes ~37× more samples than our dataset.

For the putative compound heterozygous mutations, both variants should be located within the same transcript and at least one allele should contain a LoF variant. The second variant could be: (1) a LoF variant; or (2) a missense variant that is absent in dbSNP137 [90] database and with a CADD score > 20 (predicted to belong to the 1% most deleterious variants of the total genome), indicating a pathogenic effect. The latter two filter criteria should decrease the chance of including benign missense variants. The putative compound heterozygous variants were identified by scoring the number of variants per sample per gene with PSEQ (https://atgu.mgh.harvard.edu/plinkseq/pseq.shtml). The reads of variants located within approximately 200 base pairs were visualized in IGV [91] to judge the authenticity of the compound heterozygous variant. When the different variants are located on distinct alleles, the combination of variants was considered a true compound heterozygous mutation.

All recessive variants that remained after the filtering procedures were Sanger sequenced to confirm the variant calls generated by the exome pipeline.

Variant aggregation analyses in the IPDGC WES dataset

SKAT-c [92] was used to analyze the burden of coding variants for each identified gene. Both rare variants only and the joint effect of common and rare variants were tested. Because variant aggregation tests are prone to coverage differences, capture usage and population stratification, we performed a more stringent individual and variant QC, resulting in a reduced dataset of 1540 samples (1062 cases and 478 controls) covering 268,038 variants. Individuals were excluded when failing gender test, showing evidence of relatedness, having dubious heterozygosity/genotype calls, or being a population outlier. Variants were removed when having a genotype missingness > 5%, a Hardy–Weinberg equilibrium p value < $1e^{-6}$ or a p value for non-random missingness by phenotype < $1e^{-5}$. Variants were only considered for association analyses if located in a region targeted by all different capture kits.

Benign variants have the potential to dilute a true association signal of the combined effect of functional variants in a gene. We therefore annotated variants with ANNOVAR [83] to group variants according to their type or predicted pathogenicity. Two subsets of variants were examined: (1) predicted pathogenic variants, including LoF variants and missense mutations that are predicted to be pathogenic by the CADD framework; and (2) missense variants, including amino-acid changing and LoF variants.

As suggested by SKAT, we selected a MAF cutoff of 0.018, which is based on the total sample size and separates rare and common variants. Common variants (MAF > 0.018) were pruned using PLINK [93] (indep settings 50 5 1.5). Due to confounding factors (usage different capture kits and multiple CEU populations), 20 principle components, 10× coverage, and gender were taken into account as covariates. Both a traditional one-sided burden (assuming all variants to have a harmful effect) and a two-sided SKAT test (allowing variants to be either damaging or protective) were performed. Empirical p values were calculated by comparison of the nominal p value to 10,000 permutations of affection status. Genes with an empirical p value < 0.05 were considered to be significantly associated to PD.

Genetic replication 1: variant identification in PPMI WES dataset

We obtained permission to access WES data generated by the PPMI [51]. After standard variant and individual QC, the dataset includes 477,512 variants for 462 PD cases and 183 neurologically healthy controls. A similar search for homozygous and putative compound heterozygous LoF variants, as described for the original IPDGC WES dataset, was applied for this second independent PPMI WES dataset by using ANNOVAR [83] and KGGSeq [89].

Genetic replication 2: GRIP genetic isolate

The southwest of the Netherlands contains a recently isolated population which is part of the GRIP program [52]. A total of 39 PD index cases and 19 controls of this isolate were subjected to whole-genome sequencing to explore the genetic factors underlying PD within this

geographic region. Missense and LoF variants which were present in at least two index cases and a MAF < 0.1% in public databases (ExAC, 1000G dbSNP138, and ESP6500) were considered as potential PD variants. Genes harboring such variants were surveyed for overlap with our list of candidate genes.

Genetic replication 3: variant aggregation analyses in NeuroX

We investigated the genetic burden of common and rare variants in these genes by using the independent NeuroX dataset, which is generated by a custom-made genotype array [53] using a backbone of ~240,000 standard Illumina Exome content as a basis with an additional ~24,000 variants that are suggested to be involved neurological diseases. The same procedures as described for the burden test in the IPDGC WES dataset were applied. After QC, a total of 6801 PD cases and 5970 neurologically healthy controls remained with high-quality genotype data for 178,779 variants. Based on the sample size, the MAF cutoff was 0.0063.

Genetic replication 4: overlap PD risk loci

Approximately 70% of the participants included in this study have also been included in previous published GWAS [7, 94, 95]. To explore the possibility that our candidate genes might also contain common risk variants increasing the risk to develop PD, next to the identified LoF variants with assumed high penetrance, we searched for GWAS loci within 1 Mb upstream and downstream of the gene of interest using the recent PD meta-analysis through pdgene.org [7]. Significant associations and suggestive p values < 1e-4 were considered. To understand the underlying linkage disequilibrium structure, LocusZoom [96] was applied to visualize the European 1000G recombination events for the candidate genes that were closely located to a GWAS locus.

Gene co-expression analyses

We constructed gene co-expression networks (GCN) from two different substantia nigra datasets using the R software package, WGCNA (weighted gene co-expression network analysis) [97]. This was followed by the same post-processing of WGCNA gene modules based on k-means: a heuristic to rearrange misplaced genes between modules using the number of modules detected by the standard WGCNA as k and the eigengenes as centroids. The first GCN is based on 19,152 genes from 65 substantia nigra control brains from the UKBEC consortium. The gene expression profiles are based on Affymetrix Exon 1.0 ST Arrays [98]. The second GCN is based on 63 samples from the same tissue, GTEx [56] V6 gene RPKM values. Genes were filtered with a RPKM based cutoff of 0.2 and missingness < 30% resulting in the analysis of 18,363 Ensembl genes. We corrected this gene expression dataset for the principal components significantly correlated with GTEx samples covariates using the Swamp R package. WGCNA gene modules were functionally annotated with gProfileR [99] R software package using GO database, accounting for multiple testing with gSCS's gProfiler test. Background genes used were all genes in the substantia nigra GCN. Cell type enrichment analysis was performed with the userListEnrichment function with brain specific enrichment, implemented in the WGCNA R package. Preservation analysis of UKBEC GCN in GTEx's substantia nigra profiles was performed with WGCNA's preservation analysis. Results are reported with the Z.summary statistic [100]. Graphical representation of the GCN subnetworks were constructed by using the 27 candidate genes and known PD genes (ATP13A2, FBXO7, LRRK2, PARK2, PARK7, PINK1, RAB39B, SNCA, and VPS35) as seed genes. For each of these genes sequentially, in a round robin fashion, we added the gene with highest adjacency, based on TOM values, and the links this gene has with all the seed genes. We used Cytoscape 3.3 for display with a Kamada-kawai layout algorithm [101].

Human cellular screen

shRNA virus production

Bacterial glycerol stocks containing the shRNA vectors (Sigma, St. Louis, MO, USA; TRC1 and 1.5) were grown overnight in Luria-Bertani media containing 100 μg/mL of ampicillin (Sigma-Aldrich, St. Louis, MO, USA). We selected at least five shRNA clones per gene. Endotoxin-free shRNA plasmids were extracted according to the manufacturer's protocol (Zymo, Irvine, CA, USA; ZR Plasmid Miniprep Classic kit). Lentivirus was produced as follows: HEK293T packaging cells were seeded at a density of $4 \times 10^{\Delta 5}$/mL (100 μL per well) in cell culture media, Optimem (Invitrogen, Carlsbad, CA, USA) containing 10% fetal bovine serum (FBS) in 96-well tissue culture plates. Cells were incubated for 24 h (37 °C, 5% CO_2). Each well was subsequently transfected with 100 ng of shRNA plasmid, 90 ng of packaging plasmid (pCMV-dr8.74psPAX2), and 10 ng of envelope plasmid (VSV-G/pMD2.G) combined with 0.6 μL of FugeneHD (Promega, Madison, WI, USA) in a total volume of 10 μL. Transfection efficiency was monitored using the pKLO.1 GFP plasmid (Sigma, St. Louis, MO, USA) and had to be greater than 90%. Sixteen hours after transfection, media was refreshed and supernatant harvested after a further 24 h. Virus was stored at −80 °C.

To ensure successful lentivirus production, HEK293T cells were plated out at a density of $2 \times 10^{\Delta 5}$/mL (100 μL per well) in Optimem containing 10% FBS and 15 μg/mL of protamine sulfate (Sigma, St. Louis, MO, USA). Cells were infected with 10 μL, 25 μL, and 50 μL of

lentivirus. The following day, media was refreshed with media containing 2.5 μg/mL of puromycin. After a further three days, plates were manually inspected to determine cell viability of each well. If more than 10% of the wells contained dead cells, lentiviral production for that plate was repeated.

Neuroblastoma cell culture

BE(2)-M17 (ATCC® CRL-2267™) and HEK 293 T (ATCC® CRL-3216™) cell lines were obtained from the American Type Culture Collection (Manassas, VA, USA). BE(2)-M17 cell lines were cultured in Dulbecco's Modified Eagle/Nutrient Mixture F-12 Medium (DMEM/F-12) with GlutaMAX (Invitrogen, Carlsbad, CA, USA) supplemented with 10% FBS, 1× non-essential amino acids (NEAA), and 1% Penicillin/Streptomycin. HEK 293 T cells were cultured in Opti-MEM (Invitrogen, Carlsbad, CA, USA) containing 10% FBS and 1× NEAA. All cell lines were routinely tested for mycoplasma contamination. For lentivirus infection, 25 μL of the lentivirus was added to each well of a 96-well plates and protamine sulfate was added at a final concentration of 1 μg/mL in each well of the 96-well plate. Specific wells on each lentiviral plate contained GFP expressing virus to ensure efficient transduction.

Cell-based screening assays

Four phenotypes were studied in two different assays:

Mitochondrial morphology [33] was examined in a single assay with BE(2)-M17 cells, which were expanded and plated at a density of $5 \times 10^{\Delta 4}$/mL (100 μL per well) in 96-well black CellCarrier plates (PerkinElmer, Waltham, MA, USA) pre-pipetted with 25 μL of the lentivirus. On day 2, media was refreshed with DMEM/F12 (with 10% FBS) supplemented with 2 μg/mL puromycin. On day 4, the cells were incubated with 100 nM MitoTracker Red CMXros, 100 nM MitoTracker DeepRed (Molecular Probes), and 1 μg/mL Hoechst for 20 min at room temperature. Media was refreshed and the cells were incubated for a further 2 h before fixation with 4% paraformaldehyde (pH 7.3).We examined three parameters commonly used for quantification of mitochondrial morphology: mitochondrial number, axial length ratio, and roundness.

For the Parkin translocation assay BE(2)-M17 cells were also utilized. The PLVX inducible vector (Clontech, Mountain View, CA, USA) overexpressing C-terminally tagged Parkin-GFP was used to make polyclonal stable BE(2)-M17 cells. Stable cell lines were cultured in DMEM/F12 supplemented with 10% FBS, 1% NEAA, 1% P/S, 250 ng/mL Puromycin, 200 μg/mL G418, and 1 μg/mL of doxycycline. BE(2)-M17 cells were expanded and plated at a density of 7.5 × 10^4/mL (100 μL per well) in 96-well black CellCarrier plates (PerkinElmer, Waltham, MA, USA) pre-pipetted with 25 μL of the lentivirus. The following day, media was exchanged with media without doxycycline to induce the expression of Parkin-GFP. On day 5, the cells were incubated with 100 nM MitoTracker DeepRed (Molecular Probes, Eugene, OR, USA) and 1 μg/mL Hoechst. After 20 min, media was refreshed with media containing 15 μM Carbonyl cyanide m-chlorophenyl hydrazone (CCCP). Cells were incubated for 2 h before fixation in 4% paraformaldehyde (pH 7.3).

Image acquisition and analysis

Image acquisition was carried out using the automated confocal imaging system, Cell Voyager CV7000 (Yokogawa, Tokyo, Japan). The mitochondrial morphology assay involved a total of 60 fields per well using a 60× water immersion objective lens for improved resolution. Nuclei were imaged utilizing the 405 nm laser, Mitotracker CMXros utilizing the 561 nm laser, and mitotracker DeepRed utilizing the 640nM laser. For the translocation assay, a total of 60 fields per well were taken using a 20× objective lens. Nuclei were imaged utilizing the 405 nm laser, Parkin-GFP utilizing the 488 nm laser, and mitotracker DeepRed utilizing the 640 nm laser.

Images were stored and analyzed by the Columbus Image Data storage (PerkinElmer, Waltham, MA, USA). Image quality control: only well-segmented interphase cells were included. Mitotic, apoptotic badly segmented, and out-of-focus cells were excluded. Cells touching the border of the image were removed to avoid analysis of artificially cropped cells. All wells where the perturbation strongly decreased cell number were disregarded. Morphological characteristics and signal intensities were quantified and results exported to R package CellHTS2. To quantify mitochondrial morphology, the median mitochondrial number per object, roundness, axial length ratio, and intensity of mitorackerCMXros (mitochondrial potential) were calculated.

To differentiate between CCCP-treated Parkin stable cell lines and untreated cells, the number of spots formed on mitochondria was calculated. Cells containing more than two spots were considered positive for Parkin translocation. The ratio of cells positive for translocation versus the number of cells negative for translocation was calculated per well to give a cell number independent measure of Parkin translocation. CCCP-treated cells transduced with a scrambled shRNA and CCCP-treated cells transduced with shRNA targeting *PINK1* were included on each plate. An average Z' of 0.61 was calculated for the entire screen, with a minimum Spearman's Rank correlation between replicates of 0.8.

Data from high content imaging assays were analyzed using the BioConductor CellHTS2 package for the R

software environment (R version 2.11.1, BioConductor version 2.6). Data were normalized to negative controls on a per-plate basis to minimize plate-to-plate variation. For the Parkin-translocation screen, negative controls were considered as wells which had been transduced with lentivirus encoding a scrambled sequence and had been treated with CCCP. For the remaining screens, negative controls were considered as wells that had been transduced with lentivirus encoding a scrambled sequence.

Statistical analysis

For each of the shRNA screens, each assay plate was completed with six replicates to enable the detection of subtle effects and minimize false negatives. For each shRNA, Mann–Whitney U tests with false discovery rate (FDR) correction were performed and the robust strictly standardized median difference (SSMD*) was calculated [102]. Effects were considered significant when the SSMD* normalized effect of shRNA treatment was greater than or less than 4 or –4 and at least two independent clones per gene showed a significant effect. Seed sequences were manually inspected to ensure no common sequence.

For each assay, a positive control plate containing known modifiers of the phenotype in question was run in parallel to ensure the assay worked optimally. The robust Z-factor was calculated as previously described [103], using the normalized values for the controls from all plates. For the mitochondrial assay, known regulators of mitochondrial fission or fusion were included. For the Parkin translocation assay, TOMM7 and PINK1 were used as positive controls.

shRNA knockdown validation

Cell culture and shRNA mediated knockdown were performed as described above. Cells were harvested for RNA isolation using the SV 96 Total RNA Isolation System (Promega, Madison, WI, USA) according to the manufacturer's protocol. Total RNA primed with oligo dT (Qiagen, Hilden, Germany) was used for cDNA synthesis with Superscript III RT (Life Technologies, Carlsbad, CA, USA) according to the manufacturer's specifications. Quantitative polymerase chain reaction (PCR) was carried out in triplicates on a ViiA7 real-time PCR system using SYBR Green PCR master mix (Life Technologies, Carlsbad, CA, USA) and 0.04 μM specific primer pairs for all targets. For multiple exons, gene primers were designed to span exon-exon junctions or to be separated by one intron on the corresponding genomic DNA. Normalized relative quantities were calculated with HMBS as housekeeping gene by using the qbasePLUS software (Biogazelle, Gent, Belgium) and knockdown efficiencies per clone were calculated using scrambled control wells ($n = 3$) as a reference.

Animal models

Orthologue selection

The function of the candidate genes and their involvement in neurodegeneration was tested in two animal models; *C. elegans* and *Drosophila*. The DRSC Integrated Ortholog Prediction Tool (DIOPT) [104] was used to identify the conserved homologs of human genes in the nematode or fly genomes. Orthologues were defined based on a minimum unweighted DIOPT score of 2, such that two independent bioinformatics algorithms were in agreement concerning the orthologue pairing. In cases where multiple genes were identified as potential orthologues for a given human gene, we carried forward all candidates with DIOPT scores greater than 3.

Fly stocks and husbandry

The human α-synuclein transgenic flies with codon-optimization for *Drosophila* (*UAS-α-synuclein* line #7), were recently described [48] and are available from the Bloomington Stock Center (Bloomington, IN, USA). RNAi transgenic lines were obtained from the Vienna *Drosophila* RNAi Centre (Vienna, Austria) or from Bloomington for the Harvard Transgenic RNAi Project. All RNAi lines used for this study are detailed in Additional file 1: Table S8. The GAL4-UAS system [105] was used for ectopic co-expression of both the α-synuclein and RNAi transgene. The *Rh1-Gal4* driver line (second-chromosome insertion) has been previously described [48, 106]. For screening, individual RNAi (IR) lines or Canton S (as a control) were crossed to animals of the genotype: *Rh1-Gal4/CyO*; *UAS-Syn/TM6B*. All crosses were established at 18 °C and F1 experimental animals (*Rh1-Gal4 / UAS-IR*; *UAS-Syn / +* or *Rh1-Gal4 / +*; *UAS-Syn / UAS-IR*) were shifted to 25 °C within 24 h of eclosion and aged 15 days. To examine for potential α-synuclein independent retinal degeneration, each *UAS-IR* transgenic line was separately crossed to *Rh1-Gal4*, using identical conditions. Based on the results of the primary RNAi screen, we also obtained from Bloomington available mutant alleles for the fly orthologues of *PTPRH*: *Ptp10D* and *Ptp4E.* The following additional stocks were used: (1) *w, $Ptp4E^1$*; (2) *w, $Ptp10D^1$*; (3) *yw, $Ptp4E^1$, $Ptp10D^1$ / FM7C*. All experimental results were quantified and photographed in female animals.

Characterization of retinal degeneration in Drosophila

For optical neutralization (also known as the pseudopupil preparation), fly heads of 15-day-old animals were immersed in mineral oil and transilluminated using a 40× objective on a Leica (Wetzlar, Germany) DM6000B

light microscope. Eyes from at least four animals were examined per genotype (at least eight retinae). All candidate modifier lines and controls were scored blinded by three independent examiners. The penetrance of degeneration caused by each RNAi line was calculated by dividing the number of abnormal retinae, showing evidence of either reduced rhabodomere numbers or altered refraction of light indicative of vacuolar changes, by the total number of retinae examined. For identification of genetic enhancers, we required two independent RNAi lines targeting non-overlapping sequences with 50% or greater degenerate retinaes observed using the pseudopupil assay. Following our initial screen of two RNAi lines targeting each of 18 fly gene homologs, additional RNAi lines and mutant strains were evaluated, where possible, for the most promising candidates. For each enhancer gene, the strongest RNAi line was independently re-tested for consistency using the pseudopupil assay and retinal histologic sections were also performed for further confirmation. To examine for potential α-synuclein-independent retinal degeneration, the strongest RNAi modifier for each gene was separately crossed to *Rh1-Gal4* and histologic sections were examined for 15-day-old animals. For histology, fly heads from 15-day-old animals were fixed in 8% glutaraldehyde and embedded in paraffin. Tangential (3 μm) retinal sections were cut using a Leica Microtome (RM2245) and stained with hematoxylin and eosin. Retinae from at least three animals were examined and quantified per genotype. Enhancement of α-synuclein-induced retinal degeneration was quantified based on the severity of retinal vacuolar changes seen in stained histologic sections. We examined representative photographs taken with a 40× objective from well-oriented, intact tangential sections at a depth in which the retina achieves maximal diameter. Using ImageJ software [107], we recorded the area occupied by all vacuoles with a diameter greater than 4 μm and divided by the total retinal area to compute a percentage. Statistical comparisons were implemented using a two-tailed student's t-test. α-synuclein expression levels were determined by immunoblot (clone 42, 1:1000, BD Transduction Laboratories, San Diego, CA, USA).

C. elegans media and strains

All strains were maintained as described previously [108]. For this study, the worm strains N2 (wild-type), CF512 (*fer-15(b26)II; fem-1(hc17)III*), and OW40 (zgIs15[P(unc-54)::α-synuclein::YFP]IV) were used. Strains were grown at 20 °C on Nematode Growth medium (NGM) seeded with *Escherichia coli* stain OP50. For each orthologue, one RNAi clone was selected to target the corresponding gene.

Phenotype assays for basal phenotypes in C. elegans

The systematic RNAi screen was carried out as described [109]. RNAi clones targeting the genes of interest (9/27; Additional file 1: Table S3) were obtained from the Vidal cDNA RNAi library or the Ahringer RNAi library. Bacteria expressing the empty vector L4440 were used as negative control. For the survival assay, we employed a sterile strain, CF512 (*fer-15(b26); fem-1(hc17*)) [110]. To induce sterility, eggs were collected and kept in M9 medium at 25 °C overnight until they reached L1 arrest. Approximately 25 L1 worms were added to plates seeded with RNAi clones of interest and empty vector control and allowed to develop to adults at 25 °C. At day 9 of adulthood at 25 °C, when approximately half of the worms grown on control plates were dead, the survival of worms on RNAi plates was determined.

The offspring and developmental phenotypes were tested in a single assay. N2 worms were grown at 20 °C until L4 stage on OP50 bacteria and then transferred to plates seeded with RNAi clones of interest and empty vector control. At day 2 of adulthood, ten worms were put onto a new plate seeded with the same RNAi clone for 1 h to produce progeny. The plates containing the progeny were kept at 20 °C until the F1 generation of the control worms reached L4 stage. The number and developmental phenotypes of the offspring were scored at the last time point using a dissecting microscope. A one-sided student's t-test was used to determine the significant changes compared to controls. All counting was done in a blind fashion in which the identity of the samples was concealed and each experiment was performed in three biological replicates.

Motility assay for α-synuclein toxicity model in C. elegans

Animals were age-synchronized by hypochlorite treatment, hatched overnight in M9 buffer, and subsequently cultured on NGM containing isopropylthio-β-D-galactoside (IPTG, 15 mg/L) and 50 μg/mL ampicillin (plates for RNAi treatment). Plates were seeded with RNAi bacteria. Prior to the experiment, the plates were kept at room temperature for two days to allow the production of dsRNA by the bacteria. On day 1 of adulthood (one day after larval stage L4), animals were transferred to RNAi plates containing 5-fluoro-2'deoxy-uridine (FUDR) to prevent the offspring from growing. RNAi clones targeting C54D2.4 (*ARSB*), T08G11.1 (*VPS13C*), and F44G4.8 (*PTPRH*) were used from the Ahringer *C. elegans* RNAi library. All clones were verified by sequencing. RNAi clones for the *C. elegans* orthologue F21F3.7 (*TMEM134*) was not available.

Animals were scored at day 4 and day 8 of adulthood. Animals were placed in a drop of M9 and allowed to adjust for 30 s, after which the number of body bends was

counted for another 30 s. Fifteen animals were scored per condition. Relative body bends were calculated by normalizing to control values. Error bars are showing the standard error of mean. Assays were repeated in three independent experiments and the relative body bends of one representative experiment is shown.

Additional files

Additional file 1: Includes all 12 additional tables with every table on a separate spreadsheet. The file format is an Excel spreadsheet. (XLSX 183 kb)

Additional file 2: Includes all eight additional figures with every figure on a separate slide with the exception of Additional file 1: Figure S1 (divided over two slides). The file formats are PowerPoint (.pptx) and pdf. (PDF 2504 kb)

Abbreviations

ESP6500: Exome Sequencing Project v. 6500; ExAC: Exome Aggregation Consortium; FBS: Fetal bovine serum; GCN: Gene co-expression network; GO: Gene ontology; GRIP: Genetic Research in Isolated Population; GTEx: The Genotype-Tissue Expression; GWAS: Genome-wide association study; indels: Insertions/deletions; IPDGC: International Parkinson's Disease Genomics Consortium; IR: Interfering RNA; LoF: Loss of function; MAF: Minor allele frequency; MPS VI: Mucopolysaccharidosis type VI; NGM: Nematode Growth medium; PD: Parkinson's disease; PPMI: Parkinson Progression Markers Initiative; RNAi: RNA-interference; SNVs: Single nucleotide variants; SSMD: Strictly standardized median difference; UKBEC: United Kingdom Brain Expression Consortium; WES: Whole-exome sequencing; WGCNA: Weighted gene co-expression network analysis; YOPD: Young-onset Parkinson's disease

Acknowledgements

We would like to thank all the participants who donated their time and biological samples to be a part of this study. This study was supported by the UK Brain Expression Consortium (UKBEC), the French Parkinson's Disease Genetics Study (PDG), and the Drug Interaction with Genes in Parkinson's Disease (DIGPD) study. Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/data). For up-to-date information on the study, visit www.ppmi-info.org. We also thank the Bloomington Drosophila stock center, the Vienna Drosophila RNAi Center, and the TRiP at Harvard Medical School for providing fly strains.

IPDGC consortium members and affiliations:

Mike A Nalls (Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA), Vincent Plagnol (UCL Genetics Institute, London, UK), Dena G Hernandez (Laboratory of Neurogenetics, National Institute on Aging; and Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Manu Sharma (Centre for Genetic Epidemiology, Institute for Clinical Epidemiology and Applied Biometry and Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen Germany), Una-Marie Sheerin (Department of Molecular Neuroscience, UCL Institute of Neurology), Mohamad Saad (INSERM U563, CPTP, Toulouse, France; and Paul Sabatier University, Toulouse, France), Javier Simón-Sánchez (Genetics and Epigenetics of Neurodegeneration, German Center for Neurodegenerative Diseases (DZNE)- Tübingen and Hertie Institute for Clinical Brain Research (HIH)), Claudia Schulte (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research), Suzanne Lesage (Sorbonne Université, UPMC Univ Paris 06, UM 1127, ICM; Inserm, U 1127, ICM; Cnrs, UMR 7225, ICM; ICM, Paris), Sigurlaug Sveinbjörnsdóttir (Department of Neurology, Landspítali University Hospital, Reykjavík, Iceland; Department of Neurology, MEHT Broomfield Hospital, Chelmsford, Essex, UK; and Queen Mary College, University of London, London, UK), Sampath Arepalli (Laboratory of Neurogenetics, National Institute on Aging), Roger Barker (Department of Neurology, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK), Yoav Ben-Shlomo (School of Social and Community Medicine, University of Bristol), Henk W Berendse (Department of Neurology and Alzheimer Center, VU University Medical Center), Daniela Berg (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research and DZNE, German Center for Neurodegenerative diseases), Kailash Bhatia (Department of Motor Neuroscience, UCL Institute of Neurology), Rob M A de Bie (Department of Neurology, Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands), Alessandro Biffi (Center for Human Genetic Research and Department of Neurology, Massachusetts General Hospital, Boston, MA, USA; and Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA), Bas Bloem (Department of Neurology, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands), Zoltan Bochdanovits (Department of Clinical Genetics, Section of Medical Genomics, VU University Medical Centre), Michael Bonin (Department of Medical Genetics, Institute of Human Genetics, University of Tübingen, Tübingen, Germany), Jose M Bras (Department of Molecular Neuroscience, UCL Institute of Neurology), Kathrin Brockmann (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research and DZNE, German Center for Neurodegenerative diseases), Janet Brooks (Laboratory of Neurogenetics, National Institute on Aging), David J Burn (Newcastle University Clinical Ageing Research Unit, Campus for Ageing and Vitality, Newcastle upon Tyne, UK), Elisa Majounie (Laboratory of Neurogenetics, National Institute on Aging), Steven Lubbe (Department of Clinical Neuroscience, UCL Institute of Neurology, London, UK), Iris E Jansen (Department of Clinical Genetics, VU University Medical Center, Amsterdam, The Netherlands, and German Center for Neurodegenerative diseases (DZNE), Tübingen, Germany), Ryan Price (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA); Aude Nicolas (Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA); Gavin Charlesworth (Department of Molecular Neuroscience, UCL Institute of Neurology), Codrin Lungu (National Institutes of Health Parkinson Clinic, NINDS, National Institutes of Health), Honglei Chen (Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, NC, USA), Patrick F Chinnery (Neurology M4104, The Medical School, Framlington Place, Newcastle upon Tyne, UK), Sean Chong (Laboratory of Neurogenetics, National Institute on Aging), Carl E Clarke (School of Clinical and Experimental Medicine, University of Birmingham, Birmingham, UK; and Department of Neurology, City Hospital, Sandwell and West Birmingham Hospitals NHS Trust, Birmingham, UK), Mark R Cookson (Laboratory of Neurogenetics, National Institute on Aging), J Mark Cooper (Department of Clinical Neurosciences, UCL Institute of Neurology), Jean Christophe Corvol (Sorbonne Université, UPMC Univ Paris 06, UM 1127, ICM; Inserm, U 1127, ICM; Cnrs, UMR 7225, ICM; ICM, Paris; and INSERM CIC-9503, Hôpital Pitié-Salpêtrière, Paris, France), Carl Counsell (University of Aberdeen, Division of Applied Health Sciences, Population Health Section, Aberdeen, UK), Philippe Damier (CHU Nantes, CIC0004, Service de Neurologie, Nantes, France), Jean-François Dartigues (INSERM U897, Université Victor Segalen, Bordeaux, France), Panos Deloukas (Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge, UK), Günther Deuschl (Klinik für Neurologie, Universitätsklinikum Schleswig-Holstein, Campus Kiel, ChristianAlbrechts-Universität Kiel, Kiel, Germany), David T Dexter (Parkinson's Disease Research Group, Faculty of Medicine, Imperial College London, London, UK), Karin D van Dijk (Department of Neurology and Alzheimer Center, VU University Medical Center), Allissa Dillman (Laboratory of Neurogenetics, National Institute on Aging), Frank Durif (Service de Neurologie, Hôpital Gabriel Montpied, Clermont-Ferrand, France), Alexandra Dürr (Sorbonne Université, UPMC Univ Paris 06, UM 1127, ICM; Inserm, U 1127, ICM; Cnrs, UMR 7225, ICM; ICM, Paris; and AP-HP, Pitié-Salpêtrière Hospital, Département de Génétique et Cytogénétique), Sarah Edkins (Wellcome Trust Sanger Institute), Jonathan R Evans (Cambridge Centre for Brain Repair, Cambridge, UK), Thomas Foltynie (UCL Institute of Neurology), Jing Dong (Epidemiology Branch, National Institute of Environmental Health Sciences), Michelle Gardner (Department of Molecular Neuroscience, UCL Institute of Neurology), J Raphael Gibbs (Laboratory of Neurogenetics, National Institute on Aging; and Department of Molecular Neuroscience, UCL Institute of Neurology), Alison Goate (Department of Psychiatry, Department of Neurology, Washington University School of Medicine, MI, USA), Emma Gray (Wellcome Trust Sanger Institute), Rita Guerreiro (Department of Molecular Neuroscience, UCL Institute of Neurology), Clare Harris (University of Aberdeen), Jacobus J van Hilten (Department of Neurology, Leiden University Medical Center, Leiden, Netherlands), Albert Hofman (Department of Epidemiology, Erasmus University Medical Center, Rotterdam, Netherlands), Albert Hollenbeck (AARP, Washington DC, USA), Janice Holton (Queen Square Brain Bank for Neurological Disorders, UCL Institute of Neurology), Michele Hu (Department of Clinical Neurology, John Radcliffe

Hospital, Oxford, UK), Xuemei Huang (Departments of Neurology, Radiology, Neurosurgery, Pharmacology, Kinesiology, and Bioengineering, Pennsylvania State University– Milton S Hershey Medical Center, Hershey, PA, USA), Isabel Wurster (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research and German Center for Neurodegenerative diseases), Walter Mätzler (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research and German Center for Neurodegenerative diseases), Gavin Hudson (Neurology M4104, The Medical School, Newcastle upon Tyne, UK), Sarah E Hunt (Wellcome Trust Sanger Institute), Johanna Huttenlocher (deCODE genetics), Thomas Illig (Institute of Epidemiology, Helmholtz Zentrum München, German Research Centre for Environmental Health, Neuherberg, Germany), Pálmi V Jónsson (Department of Geriatrics, Landspítali University Hospital, Reykjavík, Iceland), Jean-Charles Lambert (INSERM U744, Lille, France; and Institut Pasteur de Lille, Université de Lille Nord, Lille, France), Cordelia Langford (Cambridge Centre for Brain Repair), Andrew Lees (Queen Square Brain Bank for Neurological Disorders), Peter Lichtner (Institute of Human Genetics, Helmholtz Zentrum München, German Research Centre for Environmental Health, Neuherberg, Germany), Patricia Limousin (Institute of Neurology, Sobell Department, Unit of Functional Neurosurgery, London, UK), Grisel Lopez (Section on Molecular Neurogenetics, Medical Genetics Branch, NHGRI, National Institutes of Health), Delia Lorenz (Klinik für Neurologie, Universitätsklinikum Schleswig-Holstein), Codrin Lungu (National Institutes of Health Parkinson Clinic, NINDS, National Institutes of Health), Alisdair McNeill (Department of Clinical Neurosciences, UCL Institute of Neurology), Catriona Moorby (School of Clinical and Experimental Medicine, University of Birmingham), Matthew Moore (Laboratory of Neurogenetics, National Institute on Aging), Huw R Morris (National Hospital for Neurology and Neurosurgery, University College London, London, UK), Karen E Morrison (School of Clinical and Experimental Medicine, University of Birmingham; and Neurosciences Department, Queen Elizabeth Hospital, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK), Valentina Escott-Price (MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Cardiff, UK), Ese Mudanohwo (Neurogenetics Unit, UCL Institute of Neurology and National Hospital for Neurology and Neurosurgery), Sean S O'Sullivan (Queen Square Brain Bank for Neurological Disorders), Justin Pearson (MRC Centre for Neuropsychiatric Genetics and Genomics), Joel S Perlmutter (Department of Neurology, Radiology, and Neurobiology at Washington University, St Louis), Hjörvar Pétursson (deCODE genetics; and Department of Medical Genetics, Institute of Human Genetics, University of Tübingen), Pierre Pollak (Service de Neurologie, CHU de Grenoble, Grenoble, France), Bart Post (Department of Neurology, Radboud University Nijmegen Medical Centre), Simon Potter (Wellcome Trust Sanger Institute), Bernard Ravina Translational Neurology, Biogen Idec, MA, USA), Tamas Revesz (Queen Square Brain Bank for Neurological Disorders), Olaf Riess (Department of Medical Genetics, Institute of Human Genetics, University of Tübingen), Fernando Rivadeneira (Departments of Epidemiology and Internal Medicine, Erasmus University Medical Center), Patrizia Rizzu (Department of Clinical Genetics, Section of Medical Genomics, VU University Medical Centre), Mina Ryten (Department of Molecular Neuroscience, UCL Institute of Neurology), Stephen Sawcer (University of Cambridge, Department of Clinical Neurosciences, Addenbrooke's hospital, Cambridge, UK), Anthony Schapira (Department of Clinical Neurosciences, UCL Institute of Neurology), Hans Scheffer (Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands), Karen Shaw (Queen Square Brain Bank for Neurological Disorders), Ira Shoulson (Department of Neurology, University of Rochester, Rochester, NY, USA), Joshua Shulman (Departments of Neurology, Molecular and Human Genetics, and Neuroscience, Baylor College of Medicine and The Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital), Ellen Sidransky (Section on Molecular Neurogenetics, Medical Genetics Branch, NHGRI), Colin Smith (Department of Pathology, University of Edinburgh, Edinburgh, UK), Chris C A Spencer (Wellcome Trust Centre for Human Genetics, Oxford, UK), Hreinn Stefánsson (deCODE genetics), Francesco Bettella (deCODE genetics), Joanna D Stockton (School of Clinical and Experimental Medicine), Amy Strange (Wellcome Trust Centre for Human Genetics), Kevin Talbot (University of Oxford, Department of Clinical Neurology, John Radcliffe Hospital, Oxford, UK), Carlie M Tanner (Clinical Research Department, The Parkinson's Institute and Clinical Center, Sunnyvale, CA, USA), Avazeh Tashakkori-Ghanbaria (Wellcome Trust Sanger Institute), François Tison (Service de Neurologie, Hôpital HautLévêque, Pessac, France), Daniah Trabzuni (Department of Molecular Neuroscience, UCL Institute of Neurology), Bryan J Traynor (Laboratory of Neurogenetics, National Institute on Aging), André G Uitterlinden (Departments of Epidemiology and Internal Medicine, Erasmus University Medical Center), Daan Velseboer (Department of Neurology, Academic Medical Center), Marie Vidailhet (Sorbonne Université, UPMC Univ Paris 06, UM 1127, ICM; Inserm, U 1127, ICM; Cnrs, UMR 7225, ICM; ICM, Paris), Robert Walker (Department of Pathology, University of Edinburgh), Bart van de Warrenburg Department of Neurology, Radboud University Nijmegen Medical Centre), Mirdhu Wickremaratchi (Department of Neurology, Cardiff University, Cardiff, UK), Nigel Williams (MRC Centre for Neuropsychiatric Genetics and Genomics), Caroline H Williams-Gray (Department of Neurology, Addenbrooke's Hospital), Sophie Winder-Rhodes (Department of Psychiatry and Medical Research Council and Wellcome Trust Behavioural and Clinical Neurosciences Institute, University of Cambridge), Kári Stefánsson (deCODE genetics), Maria Martinez (INSERM UMR 1043; and Paul Sabatier University), Nicholas W Wood (UCL Genetics Institute; and Department of Molecular Neuroscience, UCL Institute of Neurology), John Hardy (Department of Molecular Neuroscience, UCL Institute of Neurology), Peter Heutink (Department of Clinical Genetics, Section of Medical Genomics, VU University Medical Centre), Alexis Brice (Sorbonne Université, UPMC Univ Paris 06, UM 1127, ICM; Inserm, U 1127, ICM; Cnrs, UMR 7225, ICM; ICM, Paris; AP-HP, Hôpital de la Salpêtrière, Département de Génétique et Cytogénétique), Thomas Gasser (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, and DZNE, German Center for Neurodegenerative Diseases), Andrew B Singleton (Laboratory of Neurogenetics, National Institute on Aging).

Funding

Funding for this study was provided by the Prinses Beatrix Spierfonds (IEJ, PH); the EU joint Program-Neurodegenerative Diseases (JPND): COURAGE-PD (PH, VD, SL, AB); the Federal Ministry of Education and Research Germany (BMBF): MitoPD (PH); NIH grants (K08AG034290, R21NS089854, R01AG033193, U01AG046161, C06RR029965, R01NS037167, R01CA141668, P50NS071674) (JMS, ABS); the Alzheimer's Association (JMS); the American Federation for Aging Research (JMS); Huffington Foundation (JMS); the Robert and Renee Belfer Family Foundation (JMS); the Jan and Dan Duncan Neurological Research Institute at Texas Children's Hospital (JMS); a Career Award for Medical Scientists from the Burroughs Wellcome Fund (JMS); the Wellcome Trust under awards 076113, 085475, and 090355 (SL, HM), Parkinson's UK (grants 8047, J-0804, F1002, and F-1201) (SL, HM); the Medical Research Council (G0700943 and G1100643) (SL, HM); the France-Parkinson Association (VD, SL, AB); the Roger de Spoelberch Foundation (R12123DD) (VD, SL, AB); the French Academy of Sciences (VD, SL, AB); the French program "Investissements d'avenir" (ANR-10-IAIHU-06) (VD, SL, AB); the Intramural Research Program of the National Institute on Aging, National Institutes of Health, Department of Health and Human Services (ZO1 AG000957) (MN, JRG, ABS), the National Institute of Neurological Disorders and Stroke (NINDS) (Z01-AG000949-02), and the National Institute of Environmental Health Sciences (Z01-ES101986). Department of Defense (award W81XWH-09-2-0128); The Michael J Fox Foundation for Parkinson's Research; American Parkinson Disease Association (APDA); Barnes Jewish Hospital Foundation; Greater St Louis Chapter of the APDA; Hersenstichting Nederland; the German National Genome Network (NGFNplus number 01GS08134, German Ministry for Education and Research); the German Federal Ministry of Education and Research (NGFN 01GR0468, PopGen); 01EW0908 in the frame of ERA-NET NEURON and Helmholtz Alliance Mental Health in an Ageing Society (HA-215); European Community Framework Programme 7, People Programme; IAPP on novel genetic and phenotypic markers of Parkinson's disease, and Essential Tremor (MarkMD), contract number PIAP-GA-2008-230596 MarkMD.
PPMI – a public–private partnership – is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, including Abbvie, Avid, Biogen, Bristol-Myers Squibb, Covance, GE Healthcare, Genentech, GlaxoSmithKline, Lilly, Lundbeek, Merck, Meso Scale Discovery, Pfizer, Piramal, Roche, Servier, Teva, UCB, and Golub Capital.
The Pathology and Histology Core at Baylor College of Medicine is supported by NIH grant P30CA125123. The DIGPD cohort was sponsored by the Assistance Publique Hôpitaux de Paris and founded by the French clinical research hospital program-PHRC (code AOR08010). This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD, USA (https://hpc.nih.gov/systems/) and DNA panels, samples, and clinical data from the National Institute of

Neurological Disorders and Stroke Human Genetics Resource Center DNA and Cell Line Repository. The TRiP at Harvard Medical School, which provided fly stocks, is supported by R01GM084947.
The relevant funding bodies (above) did not participate in the design of the study; collection, analysis, or interpretation of data; or in drafting the manuscript.

Availability of data and materials

Based on individual participant consents and institutional ethics committee approval (above), we are permitted to release WES data for 766 cases and 395 controls (70.3% of the IPDGC discovery cohort) into the public domain. These data are deposited in the National Institutes of Health database of Genotypes and Phenotypes (dbGaP; accession phs001103.v1.p2 (National Institute of Neurological Disorders and Stroke (NINDS) Exome Sequencing in Parkinson's Disease)) and the European Genome-phenome Archive (EGA; accession EGAS00001002103 (Whole-exome sequencing of Dutch Parkinson's disease patients), EGAS00001002110 (North American Brain Expression Consortium (NABEC) Exome Sequencing), EGAS00001002113 (Exome sequencing of United Kingdom Brain Expression Consortium samples), EGAS00001002156 (Whole-exome sequencing of Parkinson's disease patients from the United Kingdom)). For the remaining 382 cases and 108 controls within the IPDGC WES dataset, consent does not allow for public release of the data; investigators wishing to access these datasets should apply through the IPDGC webpage (http://pdgenetics.org/resources). Additional file 1: Table S13 provides a complete overview of the public availability of the IPDGC discovery cohort. The NeuroXdata have been deposited in dbGaP (accession phs000918.v1.p1 (International Parkinson's Disease Genomics Consortium (IPDGC), NeuroX Dataset)). PD GWAS summary statistics are accessible through http://www.pdgene.org, which includes a straightforward search on chromosomal region, gene name, or SNP identifier. We used the summary statistics of the pdgene.org website. For example, focusing on VPS13C (one of the genes we describe in the context of PD GWAS loci), we used the 1 MB region around VPS13C (chr15:61144588–63352664) as input for the search field, and copied and pasted the variants with *p* values below 1e-4 from the website in a text file. Then we converted the layout in a locus zoom format, so we were able to plot the GWAS variants together with our WES variants. Data from the PPMI WES cohort are available at http://www.ppmi-info.org/access-data-specimens/download-data/. Subsequent to approval of data access by PPMI, the PPMI WES dataset can be downloaded through an account login (DOWNLOAD > Genetic Data > Exome Sequencing > Data). All *Drosophila* strains used in this work are already available from public stock centers (Bloomington Drosophila Stock Center and Vienna Drosophila RNAi Center).

Authors' contributions

IEJ: Conception/design, data analysis/acquisition, data interpretation, and writing manuscript. HY: Conception/design, data analysis/acquisition, data interpretation, and writing manuscript. SH: Data analysis/acquisition, data interpretation, and writing manuscript. ML: Data analysis/acquisition, data interpretation, and writing manuscript. HM: Data acquisition and revising manuscript. RIS: Data acquisition and revising manuscript. SJL: Data acquisition and revising manuscript. VD: Data acquisition and revising manuscript. SL: Data acquisition and revising manuscript. EM: Data analysis/acquisition and revising manuscript. JRG: Data analysis/acquisition and revising manuscript. MAN: Data analysis/acquisition and revising manuscript. MR: Data analysis/acquisition and revising manuscript. JAB: Data analysis/acquisition and writing/revising manuscript. JV: Data analysis/acquisition and revising manuscript. JSS: Data analysis/acquisition and revising manuscript. MCL: Data analysis/acquisition and revising manuscript. PR: Data analysis/acquisition and revising manuscript. CB: Data analysis and revising manuscript. AKC: Conception/design, data acquisition and revising manuscript. YL: Data acquisition and revising manuscript. PY: Data acquisition and revising manuscript. International Parkinson's Disease Genetics Consortium: Data acquisition and revising manuscript. HRM: Data acquisition and revising manuscript. AB: Data acquisition and revising manuscript. ABS: Data acquisition and revising manuscript. DCD: Data acquisition and revising manuscript. EAN: Data acquisition and revising manuscript. SJ: Data analysis/acquisition, data interpretation, and writing manuscript. JMS: Conception/design, data interpretation, and writing manuscript. PH: Conception/design, data interpretation, and writing manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Ethics approval and consent to participate

All human participants contributing clinical data and genetic samples to this study provided written informed consent, subject to oversight by the relevant institutional review boards and ethical committees at each study site, including the NIH Office of Human Subjects Research (protocols: 4813; 12250), NIA IRB (protocol: 03-AG-N39), National Hospital for Neurology and Neurosurgery ethics committee (protocol: 10/H07166/3), Wales Research Ethics Committee 3 (protocols: 14/WA/1179; 05MRE09/58), INSERM Comité de Protection des Personnes (protocols: 00001072; RBM 03–48), CMO Regio Arnhem-Nijmegen (protocol: 2005/190) and LUMC Commissie Medische Ethiek (protocol: P03.140). All experimental methods pertaining to human participants are in accordance with the Helsinki Declaration.

Author details

[1]German Center for Neurodegenerative Diseases (DZNE), Otfried-Müller-Str. 23, Tübingen 72076, Germany. [2]Department of Clinical Genetics, VU University Medical Center, Amsterdam 1081HZ, The Netherlands. [3]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA. [4]Graduate School of Cellular & Molecular Neuroscience, Tübingen 72074, Germany. [5]European Research Institute for the Biology of Aging, University of Groningen, University Medical Centre Groningen, Groningen 9700AD, The Netherlands. [6]Department of Clinical Neuroscience, UCL Institute of Neurology, London, UK. [7]Northwestern University Feinberg School of Medicine, Ken and Ruth Davee Department of Neurology, Chicago, IL, USA. [8]Inserm U1127, CNRS UMR7225, Sorbonne Universités, UPMC Univ Paris 06, UMR_S1127, Institut du Cerveau et de la Moelle épinière, Paris, France. [9]Institute of Psychological Medicine and Clinical Neurosciences, MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, UK. [10]Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA. [11]Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK. [12]Department of Medical & Molecular Genetics, King's College London, London, UK. [13]Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany. [14]Department of Neurology, Baylor College of Medicine, Houston, TX, USA. [15]Genetic Epidemiology Unit, Department of Epidemiology, Erasmus MC, Rotterdam, The Netherlands. [16]Assistance Publique Hôpitaux de Paris, Hôpital de la Salpêtrière, Département de Génétique et Cytogénétique, Paris, France. [17]Department of Neuroscience and Program in Developmental Biology, Baylor College of Medicine, Houston, TX, USA. [18]Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, 1250 Moursund St., N.1150, Houston, TX 77030, USA.

Received: 17 June 2016 Accepted: 3 January 2017
Published online: 30 January 2017

References

1. Zimprich A, Benet-Pages A, Struhal W, Graf E, Eck SH, Offman MN, et al. A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease. Am J Hum Genet. 2011;89:168–75.
2. Vilarino-Guell C, Wider C, Ross OA, Dachsel JC, Kachergus JM, Lincoln SJ, et al. VPS35 mutations in Parkinson disease. Am J Hum Genet. 2011;89:162–7.
3. Funayama M, Ohe K, Amo T, Furuya N, Yamaguchi J, Saiki S, et al. CHCHD2 mutations in autosomal dominant late-onset Parkinson's disease: a genome-wide linkage and sequencing study. Lancet Neurol. 2015;14:274–82.
4. Farlow JL, Robak LA, Hetrick K, Bowling K, Boerwinkle E, Coban-Akdemir ZH, et al. Whole-exome sequencing in familial Parkinson disease. JAMA Neurol. 2016;73:68–75.
5. Shulman JM, De Jager PL, Feany MB. Parkinson's disease: genetics and pathogenesis. Annu Rev Pathol. 2011;6:193–222.
6. Trinh J, Farrer M. Advances in the genetics of Parkinson disease. Nat Rev Neurol. 2013;9:445–54.
7. Nalls MA, Pankratz N, Lill CM, Do CB, Hernandez DG, Saad M, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. Nat Genet. 2014;46:989–93.
8. Hamza TH, Payami H. The heritability of risk and age at onset of Parkinson's disease after accounting for known genetic risk factors. J Hum Genet. 2010;55:241–3.

9. Keller MF, Saad M, Bras J, Bettella F, Nicolaou N, Simon-Sanchez J, et al. Using genome-wide complex trait analysis to quantify 'missing heritability' in Parkinson's disease. Hum Mol Genet. 2012;21:4996–5009.
10. Jonsson T, Atwal JK, Steinberg S, Snaedal J, Jonsson PV, Bjornsson S, et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. Nature. 2012;488:96–9.
11. Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson PV, Snaedal J, et al. Variant of TREM2 associated with the risk of Alzheimer's disease. N Engl J Med. 2013;368:107–16.
12. Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, et al. TREM2 variants in Alzheimer's disease. N Engl J Med. 2013;368:117–27.
13. Smith BN, Ticozzi N, Fallini C, Gkazi AS, Topp S, Kenna KP, et al. Exome-wide rare variant analysis identifies TUBA4A mutations associated with familial ALS. Neuron. 2014;84:324–31.
14. Cirulli ET, Lasseigne BN, Petrovski S, Sapp PC, Dion PA, Leblond CS, et al. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. Science. 2015;347:1436–41.
15. Moutsianas L, Agarwala V, Fuchsberger C, Flannick J, Rivas MA, Gaulton KJ, et al. The power of gene-based rare variant methods to detect disease-associated variation and test hypotheses about complex disease. PLoS Genet. 2015;11:e1005165.
16. Sulem P, Helgason H, Oddson A, Stefansson H, Gudjonsson SA, Zink F, et al. Identification of a large set of rare complete human knockouts. Nat Genet. 2015;47:448–52.
17. Kitada T, Asakawa S, Hattori N, Matsumine H, Yamamura Y, Minoshima S, et al. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. Nature. 1998;392:605–8.
18. Bonifati V, Rizzu P, Squitieri F, Krieger E, Vanacore N, van Swieten JC, et al. DJ-1(PARK7), a novel gene for autosomal recessive, early onset parkinsonism. Neurol Sci. 2003;24:159–60.
19. Valente EM, Abou-Sleiman PM, Caputo V, Muqit MM, Harvey K, Gispert S, et al. Hereditary early-onset Parkinson's disease caused by mutations in PINK1. Science. 2004;304:1158–60.
20. Pringsheim T, Jette N, Frolkis A, Steeves TD. The prevalence of Parkinson's disease: a systematic review and meta-analysis. Mov Disord. 2014;29:1583–90.
21. Lesage S, Drouet V, Majounie E, Deramecourt V, Jacoupy M, Nicolas A, et al. Loss of VPS13C function in autosomal-recessive Parkinsonism causes mitochondrial dysfunction and increases PINK1/Parkin-dependent mitophagy. Am J Hum Genet. 2016;98:500–13.
22. Ghahramani Seno MM, Kwan BY, Lee-Ng KK, Moessner R, Lionel AC, Marshall CR, et al. Human PTCHD3 nulls: rare copy number and sequence variants suggest a non-essential gene. BMC Med Genet. 2011;12:45.
23. Newsome TP, Schmidt S, Dietzl G, Keleman K, Asling B, Debant A, et al. Trio combines with dock to regulate Pak activity during photoreceptor axon pathfinding in Drosophila. Cell. 2000;101:283–94.
24. Neumuller RA, Richter C, Fischer A, Novatchkova M, Neumuller KG, Knoblich JA. Genome-wide analysis of self-renewal in Drosophila neural stem cells by transgenic RNAi. Cell Stem Cell. 2011;8:580–93.
25. Ma XM, Kiraly DD, Gaier ED, Wang Y, Kim EJ, Levine ES, et al. Kalirin-7 is required for synaptic structure and function. J Neurosci. 2008;28:12368–82.
26. Mandela P, Yankova M, Conti LH, Ma XM, Grady J, Eipper BA, et al. Kalrn plays key roles within and outside of the nervous system. BMC Neurosci. 2012;13:136.
27. Greenamyre JT, Hastings TG. Biomedicine. Parkinson's–divergent causes, convergent mechanisms. Science. 2004;304:1120–2.
28. Haelterman NA, Yoon WH, Sandoval H, Jaiswal M, Shulman JM, Bellen HJ. A mitocentric view of Parkinson's disease. Annu Rev Neurosci. 2014;37:137–59.
29. Pickrell AM, Youle RJ. The roles of PINK1, parkin, and mitochondrial fidelity in Parkinson's disease. Neuron. 2015;85:257–73.
30. Cookson MR. Parkinsonism due to mutations in PINK1, parkin, and DJ-1 and oxidative stress and mitochondrial pathways. Cold Spring Harb Perspect Med. 2012;2:a009415.
31. Narendra D, Tanaka A, Suen DF, Youle RJ. Parkin is recruited selectively to impaired mitochondria and promotes their autophagy. J Cell Biol. 2008;183:795–803.
32. Kamp F, Exner N, Lutz AK, Wender N, Hegermann J, Brunner B, et al. Inhibition of mitochondrial fusion by alpha-synuclein is rescued by PINK1, Parkin and DJ-1. EMBO J. 2010;29:3571–89.
33. Koopman WJ, Visch HJ, Smeitink JA, Willems PH. Simultaneous quantitative measurement and automated analysis of mitochondrial morphology, mass, potential, and motility in living human skin fibroblasts. Cytometry A. 2006;69:1–12.
34. Chang CR, Blackstone C. Dynamic regulation of mitochondrial fission through modification of the dynamin-related protein Drp1. Ann N Y Acad Sci. 2010;1201:34–9.
35. Narendra DP, Jin SM, Tanaka A, Suen DF, Gautier CA, Shen J, et al. PINK1 is selectively stabilized on impaired mitochondria to activate Parkin. PLoS Biol. 2010;8:e1000298.
36. Vives-Bauza C, Zhou C, Huang Y, Cui M, de Vries RL, Kim J, et al. PINK1-dependent recruitment of Parkin to mitochondria in mitophagy. Proc Natl Acad Sci U S A. 2010;107:378–83.
37. Geisler S, Holmstrom KM, Skujat D, Fiesel FC, Rothfuss OC, Kahle PJ, et al. PINK1/Parkin-mediated mitophagy is dependent on VDAC1 and p62/SQSTM1. Nat Cell Biol. 2010;12:119–31.
38. Vincow ES, Merrihew G, Thomas RE, Shulman NJ, Beyer RP, MacCoss MJ, et al. The PINK1-Parkin pathway promotes both mitophagy and selective respiratory chain turnover in vivo. Proc Natl Acad Sci U S A. 2013;110:6400–5.
39. Feany MB, Bender WW. A Drosophila model of Parkinson's disease. Nature. 2000;404:394–8.
40. Auluck PK, Chan HY, Trojanowski JQ, Lee VM, Bonini NM. Chaperone suppression of alpha-synuclein toxicity in a Drosophila model for Parkinson's disease. Science. 2002;295:865–8.
41. MacLeod DA, Rhinn H, Kuwahara T, Zolin A, Di Paolo G, McCabe BD, et al. RAB7L1 interacts with LRRK2 to modify intraneuronal protein sorting and Parkinson's disease risk. Neuron. 2013;77:425–39.
42. Chen L, Feany MB. Alpha-synuclein phosphorylation controls neurotoxicity and inclusion formation in a Drosophila model of Parkinson disease. Nat Neurosci. 2005;8:657–63.
43. Cullen V, Lindfors M, Ng J, Paetau A, Swinton E, Kolodziej P, et al. Cathepsin D expression level affects alpha-synuclein processing, aggregation, and toxicity in vivo. Mol Brain. 2009;2:5.
44. Petrucelli L, O'Farrell C, Lockhart PJ, Baptista M, Kehoe K, Vink L, et al. Parkin protects against the toxicity associated with mutant alpha-synuclein: proteasome dysfunction selectively affects catecholaminergic neurons. Neuron. 2002;36:1007–19.
45. Yang Y, Nishimura I, Imai Y, Takahashi R, Lu B. Parkin suppresses dopaminergic neuron-selective neurotoxicity induced by Pael-R in Drosophila. Neuron. 2003;37:911–24.
46. Miura E, Hasegawa T, Konno M, Suzuki M, Sugeno N, Fujikake N, et al. VPS35 dysfunction impairs lysosomal degradation of alpha-synuclein and exacerbates neurotoxicity in a Drosophila model of Parkinson's disease. Neurobiol Dis. 2014;71:1–13.
47. Dhungel N, Eleuteri S, Li LB, Kramer NJ, Chartron JW, Spencer B, et al. Parkinson's disease genes VPS35 and EIF4G1 interact genetically and converge on alpha-synuclein. Neuron. 2015;85:76–87.
48. Chouhan AK, Guo C, Hsieh Y-C, Ye H, Senturk M, Zuo Z, et al. Uncoupling neuronal death and dysfunction in Drosophila models of neurodegenerative disease. Acta Neuropathol Comm. 2016;4:62.
49. van Ham TJ, Holmberg MA, van der Goot AT, Teuling E, Garcia-Arencibia M, Kim HE, et al. Identification of MOAG-4/SERF as a regulator of age-related proteotoxicity. Cell. 2010;142:601–12.
50. Jeon M, Nguyen H, Bahri S, Zinn K. Redundancy and compensation in axon guidance: genetic analysis of the Drosophila Ptp10D/Ptp4E receptor tyrosine phosphatase subfamily. Neural Dev. 2008;3:3.
51. Parkinson Progression Marker Initiative. The Parkinson Progression Marker Initiative (PPMI). Prog Neurobiol. 2011;95:629–35
52. Pardo LM, MacKay I, Oostra B, van Duijn CM, Aulchenko YS. The effect of genetic drift in a young genetically isolated population. Ann Hum Genet. 2005;69:288–95.
53. Nalls MA, Bras J, Hernandez DG, Keller MF, Majounie E, Renton AE, et al. NeuroX, a fast and efficient genotyping platform for investigation of neurodegenerative diseases. Neurobiol Aging. 2015;36(1605):e7–e12.
54. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. Nat Genet. 2010;42:565–9.
55. Xu J, Bai J, Zhang X, Lv Y, Gong Y, Liu L, et al. A comprehensive overview of lncRNA annotation resources. Brief Bioinform. 2016. doi:10.1093/bib/bbw015.
56. GTEx-Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science. 2015;348:648–60.
57. Aravind L, Koonin EV. G-patch: a new conserved domain in eukaryotic RNA-processing proteins and type D retroviral polyproteins. Trends Biochem Sci. 1999;24:342–4.

58. Lin ML, Fukukawa C, Park JH, Naito K, Kijima K, Shimo A, et al. Involvement of G-patch domain containing 2 overexpression in breast carcinogenesis. Cancer Sci. 2009;100:1443–50.
59. Hu F, Gou L, Liu Q, Zhang W, Luo M, Zhang X. G-patch domain containing 2, a gene highly expressed in testes, inhibits nuclear factor-kappaB and cell proliferation. Mol Med Rep. 2015;11:1252–7.
60. Otto GP, Razi M, Morvan J, Stenner F, Tooze SA. A novel syntaxin 6-interacting protein, SHIP164, regulates syntaxin 6-dependent sorting from early endosomes. Traffic. 2010;11:688–705.
61. Wang S, Bellen HJ. The retromer complex in development and disease. Development. 2015;142:2392–6.
62. Wang W, Wang X, Fujioka H, Hoppel C, Whone AL, Caldwell MA, et al. Parkinson's disease-associated mutant VPS35 causes mitochondrial dysfunction by recycling DLP1 complexes. Nat Med. 2016;22:54–63.
63. Song P, Trajkovic K, Tsunemi T, Krainc D. Parkin modulates endosomal organization and function of the endo-lysosomal pathway. J Neurosci. 2016;36:2425–37.
64. Chen L, Xie Z, Turkson S, Zhuang X. A53T human alpha-synuclein overexpression in transgenic mice induces pervasive mitochondria macroautophagy defects preceding dopamine neuron degeneration. J Neurosci. 2015;35:890–905.
65. Matozaki T, Suzuki T, Uchida T, Inazawa J, Ariyama T, Matsuda K, et al. Molecular cloning of a human transmembrane-type protein tyrosine phosphatase and its expression in gastrointestinal cancers. J Biol Chem. 1994;269:2075–81.
66. Matozaki T, Murata Y, Mori M, Kotani T, Okazawa H, Ohnishi H. Expression, localization, and biological function of the R3 subtype of receptor-type protein tyrosine phosphatases in mammals. Cell Signal. 2010;22:1811–7.
67. Takahashi H, Craig AM. Protein tyrosine phosphatases PTPdelta, PTPsigma, and LAR: presynaptic hubs for synapse organization. Trends Neurosci. 2013; 36:522–34.
68. Qian M, Pan G, Sun L, Feng C, Xie Z, Tully T, et al. Receptor-like tyrosine phosphatase PTP10D is required for long-term memory in Drosophila. J Neurosci. 2007;27:4396–402.
69. Lee HK, Cording A, Vielmetter J, Zinn K. Interactions between a receptor tyrosine phosphatase and a cell surface ligand regulate axon guidance and glial-neuronal communication. Neuron. 2013;78:813–26.
70. Valayannopoulos V, Nicely H, Harmatz P, Turbeville S. Mucopolysaccharidosis VI. Orphanet J Rare Dis. 2010;5:5.
71. Tessitore A, Pirozzi M, Auricchio A. Abnormal autophagy, ubiquitination, inflammation and apoptosis are dependent upon lysosomal storage and are useful biomarkers of mucopolysaccharidosis VI. Pathogenetics. 2009;2:4.
72. Lieberman AP, Puertollano R, Raben N, Slaugenhaupt S, Walkley SU, Ballabio A. Autophagy in lysosomal storage disorders. Autophagy. 2012;8:719–30.
73. Brooks DA, Gibson GJ, Karageorgos L, Hein LK, Robertson EF, Hopwood JJ. An index case for the attenuated end of the mucopolysaccharidosis type VI clinical spectrum. Mol Genet Metab. 2005;85:236–8.
74. Karageorgos L, Brooks DA, Pollard A, Melville EL, Hein LK, Clements PR, et al. Mutational analysis of 105 mucopolysaccharidosis type VI patients. Hum Mutat. 2007;28:897–903.
75. Sidransky E, Nalls MA, Aasly JO, Aharon-Peretz J, Annesi G, Barbosa ER, et al. Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. N Engl J Med. 2009;361:1651–61.
76. Sidransky E, Lopez G. The link between the GBA gene and parkinsonism. Lancet Neurol. 2012;11:986–98.
77. Exome Aggregation Consortium (ExAC) C, MA. http://exac.broadinstitute.org [April 2015].
78. Lek M, Karczewski K, Minikel E, Samocha K, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536:285–91.
79. Dawson TM, Ko HS, Dawson VL. Genetic animal models of Parkinson's disease. Neuron. 2010;66:646–61.
80. Shulman JM. Drosophila and experimental neurology in the post-genomic era. Exp Neurol. 2015;274:4–13.
81. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–60.
82. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43:491–8.
83. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38:e164.
84. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46:310–5.
85. Klein C, Westenberger A. Genetics of Parkinson's disease. Cold Spring Harb Perspect Med. 2012;2:a008888.
86. Grunewald A, Kasten M, Ziegler A, Klein C. Next-generation phenotyping using the parkin example: time to catch up with genetics. JAMA Neurol. 2013;70:1186–91.
87. Exome Variant Server NGESPE, Seattle, WA. http://evs.gs.washington.edu/EVS/ [September 2013 and April 2015].
88. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012;491:56–65.
89. Li MX, Gui HS, Kwan JS, Bao SY, Sham PC. A comprehensive framework for prioritizing variants in exome sequencing studies of Mendelian diseases. Nucleic Acids Res. 2012;40:e53.
90. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 2001;29:308–11.
91. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29:24–6.
92. Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X. Sequence kernel association tests for the combined effect of rare and common variants. Am J Hum Genet. 2013;92:841–53.
93. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75.
94. Simon-Sanchez J, Schulte C, Bras JM, Sharma M, Gibbs JR, Berg D, et al. Genome-wide association study reveals genetic risk underlying Parkinson's disease. Nat Genet. 2009;41:1308–12.
95. International Parkinson's Disease Genomics Consortium, Wellcome Trust Case Control Consortium. A two-stage meta-analysis identifies several new loci for Parkinson's disease. PLoS Genet. 2011;7:e1002142.
96. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010;26:2336–7.
97. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.
98. Forabosco P, Ramasamy A, Trabzuni D, Walker R, Smith C, Bras J, et al. Insights into TREM2 biology by network analysis of human brain gene expression data. Neurobiol Aging. 2013;34:2699–714.
99. Reimand J, Kull M, Peterson H, Hansen J, Vilo J. g:Profiler–a web-based toolset for functional profiling of gene lists from large-scale experiments. Nucleic Acids Res. 2007;35:W193–200.
100. Langfelder P, Luo R, Oldham MC, Horvath S. Is my network module preserved and reproducible? PLoS Comput Biol. 2011;7:e1001057.
101. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504.
102. Zhang XD. Illustration of SSMD, z score, SSMD*, z* score, and t statistic for hit selection in RNAi high-throughput screens. J Biomol Screen. 2011;16:775–85.
103. Zhang JH, Chung TD, Oldenburg KR. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. J Biomol Screen. 1999;4:67–73.
104. Hu Y, Flockhart I, Vinayagam A, Bergwitz C, Berger B, Perrimon N, et al. An integrative approach to ortholog prediction for disease-focused and other functional studies. BMC Bioinformatics. 2011;12:357.
105. Brand AH, Perrimon N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development. 1993;118:401–15.
106. Xiong B, Bayat V, Jaiswal M, Zhang K, Sandoval H, Charng WL, et al. Crag is a GEF for Rab11 required for rhodopsin trafficking and maintenance of adult photoreceptor cells. PLoS Biol. 2012;10:e1001438.
107. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9:671–5.
108. Brenner S. The genetics of Caenorhabditis elegans. Genetics. 1974;77:71–94.
109. Hansen M, Hsu AL, Dillin A, Kenyon C. New genes tied to endocrine, metabolic, and dietary regulation of lifespan from a Caenorhabditis elegans genomic RNAi screen. PLoS Genet. 2005;1:119–28.
110. Garigan D, Hsu AL, Fraser AG, Kamath RS, Ahringer J, Kenyon C. Genetic analysis of tissue aging in Caenorhabditis elegans: a role for heat-shock factor and bacterial proliferation. Genetics. 2002;161:1101–12.