A widespread family of heat-resistant obscure (Hero) proteins protect against protein instability and aggregation

Kotaro Tsuboyama[1,2,†], Tatsuya Osaki[3,4], Eriko Suzuki-Matsuura[2,5], Hiroko Kozuka-Hata[6], Yuki Okada[7], Masaaki Oyama[1,6], Yoshiho Ikeuchi[3,4], Shintaro Iwasaki[1,5], and Yukihide Tomari[1,2,†]

[1]Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-0032, Japan.

[2]Laboratory of RNA Function, Institute for Quantitative Biosciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-0032, Japan.

[3]Biomolecular and Cellular Engineering laboratory, Institute of Industrial Science, The University of Tokyo, Meguro-ku, Tokyo 153-8505, Japan

[4]Department of Chemistry and Biotechnology, School of Engineering, The University of Tokyo, Bunkyo-ku, Tokyo 113-0032, Tokyo, Japan

[5]RNA Systems Biochemistry Laboratory, RIKEN Cluster for Pioneering Research, Wako, Saitama 351-0198, Japan.

[6]Medical Proteomics Laboratory, The Institute of Medical Science, The University of Tokyo, Minato-ku, Tokyo 108-8639, Japan.

[7]Laboratory of Pathology and Development, Institute for Quantitative Biosciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-0032, Japan

[†]Corresponding authors. Email: tsuboyama-tky@umin.ac.jp (K.T.) and tomari@iam.u-tokyo.ac.jp (Y.T.)

Abstract

Proteins are typically denatured and aggregated by heat. Exceptions to this principle include highly disordered and heat-resistant proteins found in extremophiles, which help these organisms tolerate extreme conditions such as drying, freezing, and high salinity. In contrast, the functions of heat-soluble proteins in non-extremophilic organisms including humans remain largely unexplored. Here we report that heat-resistant obscure (Hero) proteins, which remain soluble after boiling at 95ºC, are widespread in *Drosophila* and humans. Hero proteins are hydrophilic and highly charged, and function to stabilize various “client” proteins, protecting them from denaturation even under stress conditions such as heat shock, desiccation, and exposure to organic solvents. Hero proteins can also block several different types of pathological protein aggregations in cells and in *Drosophila* strains that model neurodegenerative diseases. Moreover, Hero proteins can extend lifespan of *Drosophila*. Our study reveals that organisms naturally use Hero proteins as molecular shields to stabilize protein functions, highlighting their biotechnological and therapeutic potential.

Introduction

Proteins are polymers composed of 20 different amino acids, whose side chains have various properties such as aliphatic, aromatic, acidic, basic, and sulfur-containing. This diversity allows proteins to fold into three-dimensional structures, which determine their activity and function. On the other hand, proteins are generally vulnerable to heat, which destroys their structures and induces denaturation and aggregation. However, there are some exceptions. For example, Tardigrade Disordered Proteins (TDPs), found in the supernatant of tardigrade lysate after boiling (Yamaguchi et al., 2012), are required for tardigrades to survive desiccation (Boothby et al., 2017). Another example is Late Embryogenic Abundant (LEA) proteins in plants, whose expression is associated with their tolerance to dehydration, freezing, or high salinity (Gai et al., 2011; Galau et al., 1986). LEA proteins are also found in extremophiles including the radiation-resistant bacterium *Deinococcus radiodurans* and desiccation-tolerant animals such as *Artemia*, *C. elegans*, and rotifers (Close et al., 1989; Hand et al., 2011). TDPs and most LEA proteins are extremely hydrophilic and heat-soluble, and can be classified into intrinsically disordered proteins (IDPs). Currently, these heat-resistant proteins are viewed as special cases, required by organisms living in extreme conditions to protect their functional proteins. Although mammals have been reported to produce some heat-soluble proteins (Galea et al., 2009; Kim et al., 2000), their functions remain largely unexplored.

Here we show that a class of heat-soluble proteins, which we call “Hero” (HEat-Resistant Obscure) proteins, are widespread in non-extremophilic animals such as *Drosophila* and humans. Hero proteins act as molecular shields to protect various “client” proteins from denaturation even under stress conditions, such as heat-shock, desiccation, and exposure to organic solvents. Moreover, Hero proteins can suppress several forms of pathogenic protein aggregates in cells, and in vivo models for neurodegenerative diseases, and extend lifespan of *Drosophila*. Our findings not only have significant implications for our fundamental understanding of protein stability and functions but also highlight potential biotechnological and therapeutic applications of Hero proteins.

Results

Boiled supernatants of crude cell lysates improve Ago2 molecular behavior.

Many proteins are naturally unstable even under physiological conditions, probably because they need to prioritize function over structural durability. For example, Argonaute (AGO) proteins, the core of RNA-induced silencing complex (RISC), must load a small RNA duplex, eject the passenger strand from the two strands of the duplex, recognize target RNAs complementary to the remaining guide strand, and endonucleolytically cleave RNA targets or recruit downstream silencing factors (Kobayashi and Tomari, 2016). AGO proteins are highly unstable; in particular, empty AGO is structurally flexible at the single-molecule level (Tsuboyama et al., 2018) and susceptible to proteolytic digestion in vitro (Elkayam et al., 2012). During the course of our studies on AGO, we observed a strange phenomenon: when immunopurifying *Drosophila* Ago2 fused with a TEV-cleavable FLAG-tag using anti-FLAG magnetic beads, we were unable to elute the protein from the beads, even after the FLAG-tag was removed using TEV protease (Figures 1A and B). We concluded that purified, free Ago2 protein is unstable and tends to stick non-specifically to the beads. Strikingly, addition of crude *Drosophila* S2 cell lysate during the TEV protease treatment promoted efficient elution of Ago2 (Figure 1B). To test whether a protein(s) was responsible for this effect, we boiled the crude S2 cell lysate at 95ºC for 15 min and removed the precipitated proteins by centrifugation. Unexpectedly, the supernatant of the heat-denatured lysate, which contained only ~4.3% of the total protein concentration in the original crude lysate, was as competent in eluting Ago2 as the original lysate. Digestion of proteins in the boiled supernatant with Proteinase K, followed by its inactivation by a second round of heating, halved its activity in Ago2 elution (Figure 1B). To explore what materials other than proteins contribute to Ago2 elution, we utilized Benzonase to deplete DNAs and RNAs. Treatment of the boiled lysate by both Proteinase K and Benzonase abolished the elution activity, while the Benzonase treatment alone only modestly decreased the activity (Figure 1B). Elution of *Drosophila* Ago2 was similarly promoted by crude lysate from human HEK293T cells and by its boiled supernatant (~6.0% of the original protein concentration). Again, the efficiency was nearly halved by the Proteinase K treatment, modestly decreased by Benzonase, and abrogated by Proteinase K and Benzonase together (Figure 1B). These results suggest that some heat-soluble proteins as well as nucleic acids contained in the boiled S2 or HEK293T cell supernatants promote elution of free Ago2 from the beads.

We asked whether the boiled supernatants were capable of influencing any of the functional features of Ago2. Assembly of *Drosophila* Ago2-RISC requires not only a small RNA duplex and Ago2 itself but also the Dicer-2/R2D2 heterodimer and the Hsp70/90 chaperone machinery (Iwasaki et al., 2015). All these factors are present in crude S2 cell lysate, and incubation of FLAG-tagged Ago2 immobilized on magnetic beads with a ^{32}P-radiolabeled small RNA duplex and S2 cell lysate efficiently assembled pre-RISC, containing both strands of the duplex, and mature RISC, containing only the single-stranded guide (Figure 1C). Compared to the crude S2 lysate, reconstitution of Ago2-RISC assembly using high concentrations of purified Hsp70/90 chaperone machinery and Dicer-2/R2D2 efficiently formed pre-Ago2-RISC, but its conversion into mature Ago2-RISC was inefficient. Adding the supernatant of heat-denatured S2 or HEK293T cell lysates strongly promoted the formation of both pre- and mature Ago2-RISC, although without the chaperones, the supernatants themselves showed no reconstitution activity (Figures 1C and D). Thus, the boiled supernatants from either fly S2 or human HEK293T cells can promote the assembly of Ago2-RISC by Dicer-2/R2D2 and the chaperone machinery. Together, our data suggest that factors in the boiled supernatants improve the molecular behavior of Ago2, either in its free form (Figures 1A and B) or upon RISC assembly (Figures 1C and D), and that these factors act beyond the species boundary.

Drosophila and human heat-soluble proteins protect LDH activity from desiccation.

We next asked whether heat-soluble proteins in the boiled supernatants could act generally to enhance or protect the activity of other proteins, especially in stress conditions. We first examined the effect of boiled supernatants on protein dehydration. When purified lactate dehydrogenase (LDH) from rabbit muscle was dried up at room temperature overnight, addition of the boiled supernatants from S2 or HEK293T cells efficiently preserved the LDH activity compared to the PBS buffer only condition (Figure 1E). This effect was eliminated by Proteinase K and/or Benzonase treatment of the supernatants (Figure 1E), indicating that heat-soluble *Drosophila* and human proteins as well as nucleic acids can protect the activity of LDH from desiccation. Because it is already known that nucleic acids participate in maintaining protein homeostasis (Docter et al., 2016; Maharana et al., 2018), we decided to focus on heat-soluble proteins in the following analyses.

Hydrophilic and highly charged proteins are enriched in the boiled supernatants.

Heat generally denatures protein structures by exposing hydrophobic regions. Thus, proteins remaining soluble in the boiled supernatant are likely to be hydrophilic and intrinsically disordered. To test this idea, we subjected the crude lysates and their boiled supernatants of fly S2 and human HEK293T cells to LC-MS/MS analysis. In total, we identified 910 (S2) and 980 (HEK293T) proteins with high confidence (FDR [false discovery rate] < 0.01). We then divided the identified proteins into five groups according to the degrees of depletion or enrichment by boiling (Group I [most depleted] to Group V [most enriched]), in such a way that each group contained virtually the same number of proteins (Figures 2A and B; also see Experimental Procedures). To examine the overall characteristics of heat-soluble proteins, we utilized a prediction tool for intrinsically disordered regions (IDRs), called IUPred, whose score above 0.5 indicates that the amino acid residue at a given position is likely to be a part of a disordered region (Dosztányi et al., 2005). For simplicity, we calculated the median IUPred score of all the residues for each protein. As expected, the proteins enriched by boiling showed much higher distributions of the median IUPred score than the depleted ones in both *Drosophila* or humans (Figures 2C and D). This trend was the same when we analyzed the five protein groups by D2P2 (Oates et al., 2013), which integrates nine different algorithms for disorder prediction (data not shown). Moreover, hydrophilic amino acids were overrepresented while hydrophobic amino acids were underrepresented in the boil-enriched proteins (Figures 2E and F). Thus, structurally disordered, hydrophilic proteins were indeed enriched in the boiled supernatant.

Motifs previously associated with LEA proteins were not found among the proteins identified in the boiled supernatants. In fact, no common or conserved motifs were detected among their primary sequences. Moreover, the DIOPT (DRSC Integrative Ortholog Prediction Tool) score, which integrates 15 different ortholog prediction tools for the comparison between *Drosophila* and humans (Hu et al., 2011), were gradually decreased as the enrichment by boiling was increased, indicating that heat-soluble proteins are evolutionarily less conserved in their primary amino-acid sequences than regular heat-vulnerable proteins (Figures 2G and H). These data suggest that heat-soluble proteins, as a whole, cannot be defined by conventional, sequence- and homology-based classifications. Instead, we found that the isoelectric point (pI) values of boil-enriched proteins have a bimodal distribution, concentrating on the acidic (~5) or basic (~10) region but avoiding the neutral (~7–8) region (Figures 2I and J). Thus, heat-resistant proteins tend to be either highly negatively charged or highly positively charged under the physiological pH.

Because the LC-MS/MS analysis has an intrinsic bias that depends on the size range of protease-digested peptide fragments and can only detect proteins expressed in HEK293T cells, we also utilized the IUPred score itself, independently of the MS-identified list, to select proteins from the entire human proteome for further functional analyses. In the human protein atlas (Uhlén et al., 2015), we found ~450 proteins (Figure 2K) that have both high IUPred scores across their entire sequences (median IUPred > 0.6, excluding locally structured proteins with >25 % of residues having an IUPred score of < 0.4) and high expression levels (median RPKM > 25), which also showed a bimodal pI distribution (Figure 2K). Of these, we chose six representative proteins (C9orf16, C11orf58, BEX3, SERBP1, SERF2, and C19orf53; Supplementary Figures 1A and B) whose mRNA levels are as high as that of the most abundant form of Hsp70 (HSPA8) in HEK293T cells, four of which (C9orf16, C11orf58, SERBP1, and SERF2) were detected by the LC-MS/MS analysis of the boiled supernatant. The pI values for the six proteins were 4.13 (C9orf16), 4.57 (C11orf58), 5.31 (BEX3), 8.66 (SERBP1), 10.44 (SERF2), and 11.55 (C19orf53), respectively.

We next assayed the heat solubility of these six candidate proteins, together with GFP and GST as controls (Figure 3A). We expressed them as FLAG-tagged proteins in HEK293T cells, boiled the cell lysate at 95ºC for 15 min, then examined the supernatant by anti-FLAG Western blotting. All of six proteins were readily detectable in the soluble supernatant after boiling, whereas GFP and GST were essentially eliminated from the boiled supernatant (Figure 3A). Based on their unusual heat resistance and unstructured nature, we call this protein class HEat-Resistant Obscure (Hero) proteins (*hero-hero* in Japanese means flimsy, loose, or flexible) and add the molecular weight to distinguish each protein, following the naming convention for heat shock proteins (e.g., Hsp70). Hereafter, we call C9orf16, C11orf58, BEX3, SERBP1, SERF2, and C19orf53 as Hero9, 20, 13, 45, 7, and 11, respectively, for convenience.

Hero proteins protect the activity of various proteins under stress conditions.

To evaluate the function of each of the six representative Hero proteins, we produced recombinant human Hero proteins in *E. coli*. We then mixed 5 µg/mL of LDH with 4 µg/mL of each Hero protein, and subjected the mixture to overnight dehydration. Like the boiled supernatants (Figure 1E), all six Hero proteins protected the LDH activity at ~50%, which was much stronger compared to BSA, GST, or the conventional protein stabilizers arginine or trehalose (Figure 3B). The LDH-protecting activity under desiccation is similar to what was previously observed for TDPs (Boothby et al., 2017),

which share no common sequence motifs with human Hero proteins.

We next explored the protective effect of Hero proteins in organic solvents, a drastically harsh and foreign condition for proteins. When EGFP was exposed to chloroform, its fluorescent intensity was reduced to ~10% of the native state (Figure 3C). In contrast, the addition of four out of six Hero proteins maintained >60% of the EGFP activity (Figure 3C). Strikingly, in the presence of Hero45, EGFP fluorescence remained undiminished even in chloroform (Figure 3C). This is reminiscent of the ability of methacrylate-based random heteropolymers (RHPs) to preserve horseradish peroxidase (HRP) activity in toluene (Panganiban et al., 2018). Thus, naturally occurring Hero proteins act as molecular shields to stabilize proteins, even against harsh conditions.

To investigate the effect of Hero proteins in cells, we expressed firefly luciferase together with each of the six Hero proteins, Hsp70, or Hsp90 in HEK293T cells, then subjected the cells to heat shock at 45ºC for 8 min. Heat shock reduced the luciferase activity to ~10–20% of the original level when co-transfected with an empty or GST-expressing vector. Over-expression of the conventional chaperones, Hsp70 or Hsp90, maintained ~30–50% of luciferase activity (Figure 3D). Remarkably, Hero proteins protected luciferase from the heat shock as well as the chaperones, with Hero9, 7 and 11 preserving ~40% of luciferase activity (Figure 3D). We concluded that Hero proteins have the ability to stabilize various proteins both in vitro and in cells.

Hero proteins prevent pathogenic protein aggregations in vitro and in cells.

Protein instability is often linked to diseases, especially neurodegenerative disorders. For example, aggregation of TAR DNA-binding protein of 43 kDa (TDP-43) is observed in virtually all cases of amyotrophic lateral sclerosis (ALS) and in about half the cases of frontotemporal dementia (FTD) (Neumann et al., 2006). Given the strong activity of Hero proteins to stabilize proteins (Figures 1–3), we tested if Hero proteins could prevent pathogenic aggregation of TDP-43. We constructed TDP-43 lacking the nuclear localization signal (TDP-43ΔNLS), which is more prone to aggregation than the wild type due to its forced cytoplasmic localization. We expressed TDP-43ΔNLS as a GFP-fusion in HEK293T cells, together with each of the six human Hero proteins or GST as a control. Aggregates of GFP-TDP-43ΔNLS and their suppression by co-expression of Hero proteins were apparent at the microscopic level (Figure 4A). To quantitatively analyze the degree of aggregation, the cells were lysed by sonication, and aggregates in the extracts were measured by a filter trap assay using a 0.2 µm cellulose

acetate membrane in the presence of 1% SDS. We found that Hero45, Hero7, and Hero11 strongly suppressed TDP-43ΔNLS aggregation (Figure 4B). In addition to TDP-43, we also analyzed two additional model proteins that use different mechanisms to form pathological aggregates. HTTQ103 consists of a stretch of 103 polyglutamine residues deriving from the abnormal CAG expansion found in Huntington disease-causing huntingtin mutant (Krobitsch and Lindquist, 2000). GA50 (50 Glycine-Alanine repeats) is deriving from the abnormal GGGGCC repeats found in the ALS-causing C9orf72 intron (Mori et al., 2013). Both at the microscopic level and by the filter trap assay, aggregates of HTTQ103 and GA50 were also markedly reduced by Hero9 (HTTQ103), Hero45 (GA50) and several other Hero proteins (Figures 4A and B). Of note, prevention of pathogenic polyglutamine and polyalanine aggregations in the T-REx293 cell line was previously shown for LEA proteins derived from an anhydrobiotic nematode or soya bean (Chakrabortee et al., 2007; Liu et al., 2011). Thus, anti-aggregation may be a common feature for heat-soluble proteins.

Interestingly, among the six Hero proteins tested, there was no “super-Hero” protein able to function for all proteins. Instead, distinct subsets of human Hero proteins were more effective in preventing different types of aggregations (Figure 4B). For example, Hero11 promoted, rather than prevented, aggregation of GA50, while strongly suppressing that of TDP-43ΔNLS and HTTQ103 (Figure 4B). Similarly, Hero7 was previously shown to promote aggregation of HA-tagged HTTQ74 in SK-N-SH cells (van Ham et al., 2010), whereas we found that it strongly suppressed aggregation of GFP-tagged TDP-43ΔNLS in HEK293T cells (Figure 4B). We suggest that different Hero proteins have different preferences for their protein “clients”, which may vary depending on the cellular environments. Nevertheless, it is remarkable that, for a variety of assays with different client proteins and experimental settings (Figures 3–4), there were always multiple Hero proteins (despite looking at only six out of hundreds of putative Hero proteins in humans) that showed strong protein-stabilization and anti-aggregation effects.

Extreme charges of Hero proteins are necessary and sufficient to suppress TDP-43 aggregations.

Apparently, there was no common rule for the effective combination between different Hero proteins and different client proteins. However, we noticed that Hero45, 7, and 11, which were particularly effective in preventing the cellular aggregation of TDP-43ΔNLS (Figure 4B), have unusually high positive charges (pI = 8.66, 10.44, and 11.55,

respectively) with high proportions of basic amino-acid residues (proportion of Arg + Lys = 18.4%, 32.2%, and 27.3%, respectively). This raised a possibility that these positively charged properties may be important for their functions to prevent TDP-43ΔNLS aggregation. To confirm this idea, we constructed mutants of Hero7 and Hero11, in which Lys and Arg (positively charged) were replaced to Gly (non-charged). As shown in Figure 4C, KR-less mutants of Hero7 and Hero11 have largely lost the anti-aggregation activity, suggesting that positive charges of Hero proteins are necessary for this activity, at least for TDP-43 aggregation.

To test the sufficiency of positive charges of Hero proteins, we next randomly scrambled the amino acid sequences of Hero7 and 11, and GST as a control, while keeping their original amino acid composition ratio (i.e., % of each amino acid) and fixing the total amino-acid lengths at 100-aa or 42-aa (Figure 4D and Supplementary Table 1; three different shuffles for 100-aa and two different shuffles for 42-aa). Strikingly, the randomized 100-aa sequences, derived from extremely positively charged Hero7 and 11, prevented the aggregation of TDP-43ΔNLS with comparable or even higher efficiencies than the original ones. Importantly, none of the scrambled 42-aa sequences were able to prevent the TDP-43ΔNLS aggregation, regardless of their amino-acid compositions. These results indicate that their amino-acid composition and length (i.e., their molecular nature as long, hydrophilic and highly charged “polymers”), but not their amino-acid sequence *per se*, are the key for Hero proteins to impart their client-protecting functions, at least for TDP-43ΔNLS. Of note, it was previously shown that short peptide fragments of randomly chosen IDPs in humans (35–45 amino acids) have only modest cryoprotective and lyophilization-protective activities (up to ~1.3–1.5 fold compared to BSA) (Matsuo et al., 2018), suggesting that the amino-acid length is also important for protecting other proteins under stress conditions.

Hero proteins suppress the neurotoxicity caused by protein aggregates in motor neurons.

We next examined the anti-aggregation activity of Hero proteins in human iPS-derived motor neurons. We expressed GFP-tagged TDP-43ΔNLS together with each of the six representative human Hero proteins or the GST negative control, as in HEK293T cells. Aggregates of GFP-tagged TDP-43ΔNLS gave saturated fluorescent signals, which was suppressed by co-expression of Hero7 (Figure 5A). To quantify the degree of aggregation, we defined the ratio of the saturated area as the aggregation index (Figure 5B). Similarly to the analysis in HEK293T cells (Figure 4B), Hero 9, 45, 7 and 11

strongly reduced the aggregation formation in motor neurons (Figure 5C).

The neurite length is widely used as a measurement of the neuronal condition. Indeed, it is known that ALS patient-derived motor neurons with mutations in TDP-43 have shorter neurites compared to those from healthy donors (Fujimori et al., 2018). To evaluate neuronal protective activity of Hero proteins, we measured the total neurite length of the transfected neurons. Except for Hero45, overexpression of Hero proteins themselves had no or little effect on neuronal morphology (Figure 5D). In contrast, overexpression of GFP-TDP-43ΔNLS markedly reduced the neurite outgrowth to ~30% compared to that of control neurons without aggregation (Figure 5F). However, co-expression of Hero9, 7 and 11 almost fully rescued the neurite outgrowth defects by TDP-43 aggregation (Figures 5E and F). Therefore, Hero proteins can efficiently protect human motor neurons from the neurotoxicity of TDP43 aggregation.

Hero proteins suppress the neurotoxicity of protein aggregates in *Drosophila* eyes.

Hero proteins' protective activity can be observed not only in cultured cells but also in *Drosophila* in vivo. The fly eye has been used as a valuable model system for studying neurodegenerative diseases. Indeed, expression of human TDP-43 fused with YFP (TDP-43-YFP) in the differentiating photoreceptor cells using a GMR-Gal4 driver causes retinal degeneration (Elden et al., 2010; Figure 6A, middle). We generated transgenic flies expressing human Hero9, 13, 45 or 11 in the retina and introduced the transgenes into strains expressing TDP-43-YFP. Strikingly, the eye degeneration by TDP-43-YFP was almost completely suppressed by co-expression of Hero9 (Figure 6A, middle). The rescue effect of TDP-43-YFP by Hero45 was moderate, but could be boosted by doubling the copy number of the transgene (Figure 6A, bottom). There were no harmful effects by retina-specific expression of Hero proteins *per se*; neither YFP nor any of the transgenic Hero proteins alone caused a detectable change in the normal eye morphology (Figure 6A, top). These data show that Hero proteins can suppress pathogenic protein aggregations in the living fly eye.

Using this *Drosophila* eye system, we next investigated functions of the endogenous Hero proteins. It is known that Hsc70-4, the most abundant and constitutively expressed Hsp70 chaperone, is required for normal eye development in *Drosophila* (Chang et al., 2002; Hagedorn et al., 2006; Kumar and Tiwari, 2018). Indeed, knockdown of Hsc70-4 in the eyes showed a modest but detectable defect in external eye morphology, whereas mock depletion of germline-specific Piwi caused no

abnormality (Figure 6B). Interestingly, eye-specific knockdown of CG17931, a homolog of human Hero7 (identity = 58%, similarity = 68%), also showed a similar eye degeneration phenotype as that of Hsc70-4. Moreover, irregular arrangements of ommatidia and bristles (Figure 6C) as well as reduction of the retinal depth (Figure 6D) were also caused by knockdown of Hsc70-4 or CG17931. These observations suggest that not only chaperons but Hero proteins also play important roles in fly eye development.

We then evaluated the protective effect of endogenous fly Hero proteins against pathogenic protein aggregates. To this end, we utilized the above-mentioned aggregation-prone TDP-43-YFP (Figure 6E, upper) as well as human Ataxin3 containing a pathogenic expansion of 78 glutamine repeats (MJDtr-Q78; Figure 6E, bottom), which also causes retinal degeneration when overexpressed in *Drosophila* eyes (Warrick et al., 1998; Figure 6E, bottom). Strikingly, defects by these aggregation-prone proteins were markedly exacerbated by simultaneous knockdown of the Hero7 homolog CG17931, compared to the negative control of Piwi knockdown (Figure 6B, middle and bottom). Exacerbated eye phenotypes with MJDtr-Q78 and TDP-43-YFP were also apparent in the eye depleted of other endogenous fly Hero proteins, such as CG14818 and Vig2, which are potential homologs of human Hero9 (identity = 23%, similarity = 40%) and Hero45 (identity = 30%, similarity = 41%), respectively, and CG12384 and CG11444 (Figure 6E). Together, these data show that Hero proteins act to suppress aggregation-associated eye degeneration in living flies, consistent with our in cell data (Figures 4–5).

Hero proteins are essential both in cells and in vivo.

To understand Hero proteins' functions in non-overexpressed, physiological conditions, we performed CRISPR/Cas9-mediated knockout (KO) of six Hero proteins in HEK293T cells. We found that KO of Hero13 and 7 showed defects in cell proliferation under the normal culture condition (Figure 7A). Moreover, whole-body knockdown of endogenous Hero proteins in *Drosophila* using actin-GAL4-driven long hairpin RNAs often caused lethality (Figure 7B). Thus, Hero proteins play essential roles in human cell proliferation and *Drosophila* early development. We note that Frost in *Drosophila*, which is highly upregulated after cold shock (Goto SG., 2001) and required for maintaining female fertility following cold exposure (Newman et al., 2017), is extremely disordered and charged, and presumably represents another example of Hero proteins with physiological functions in flies.

Hero proteins can promote longevity in *Drosophila.*

It is well known that aging is tightly associated to the loss of protein homeostasis (Kaushik and Cuervo, 2015; López-Otín et al., 2013). For example, overexpression of Hsp70 significantly increases life expectancy in flies (Tatar et al., 1997). Such “longevity factors” are not limited to Hsp70 but include other chaperones and proteins related to autophagy and the ubiquitin-proteasome system (Kaushik and Cuervo, 2015). Strikingly, whole-body overexpression of endogenous Hero proteins in *Drosophila* often led to elongation of lifespan by more than ~30%, while YFP overexpression had no significant effect (Figures 7C and D). This was also true for inter-species overexpression of human Hero proteins in the fly whole body (Figure 7D). These observations support the idea that Hero proteins can combat against proteotoxic stress caused by aging.

Discussion

Our study demonstrates that Hero proteins can stabilize various client proteins and protect them from denaturation or aggregation. Thus, not only plants, tardigrades, and other extremophiles, but non-extremophilic organisms such as humans universally utilize heat-resistant, highly hydrophilic and charged proteins as molecular shields against protein instability. The simplest mode of action as molecular shields will be to reduce intermolecular collision/cohesion rates sterically and/or electrostatically, as previously proposed for LEA proteins (Wise and Tunnacliffe, 2004). However, this may not be the only mechanism for the client-protective activities. TDPs, LEA proteins, and Hero proteins belong to a large family of IDPs and can be regarded as their extreme instances with extensive disorders. Currently, IDPs and proteins with IDRs are often linked to their abilities to cause liquid-liquid phase separation (LLPS), which is commonly thought to be a stage prior to aggregation. Indeed, we noticed that purified recombinant Hero7 protein spontaneously undergoes reversible gelation at a low temperature in vitro (data not shown). Moreover, among different conditions tested in this study, there were some cases where Hero proteins promoted, rather than inhibited, aggregations, e.g., Hero20 for GA50 (Figure 4B). However, on the whole, Hero proteins tended to prevent aggregations of various proteins efficiently (Figures 3–7). This client-protective action of Hero proteins is seemingly opposite to the common notion of IDPs/IDRs as autonomous aggregation inducers. Notably, however, it was recently proposed that a prion-like IDR in yeast Sup35 acts to protect Sup35 itself against proteotoxic damage, by extending its soluble/reversible phase space and thereby preventing it from irreversible aggregation (Franzmann and Alberti, 2019; Franzmann et al., 2018). Thus, it is tempting to speculate that Hero proteins can impart at least a part of their protective functions in a similar manner as the prion-like IDR in Sup35, except that they do so in trans on their client proteins.

Trans-acting protective functions of Hero proteins may be reminiscent to those of conventional molecular chaperones. Indeed, overexpression of Hero proteins protected the luciferase activity in HEK293T cells under heat shock, much as Hsp70 and Hsp90 chaperones (Figure 3D). Hero proteins also significantly increased the longevity of *Drosophila* (Figures 7C and D), as previously reported for Hsp70 (Tatar et al., 1997). Moreover, depletion of CG17931 (a homolog of human Hero7) and that of Hsc70-4 caused similar degeneration in the *Drosophila* eye (Figures 6B–D). However, unlike molecular chaperones that often use cycles of ATP binding and hydrolysis, Hero proteins lack any apparent ATPase or any other domains. Moreover, chaperones

usually act on the inactive/denatured state of their client proteins and actively reverse them to the functional state, whereas Hero proteins likely act on the functional state of their client proteins and protect them from lapsing into the inactive/denatured state (Figure 7E). We suggest that cooperation between chaperones and Hero proteins with distinct modes of action is central to the maintenance of the functional proteome.

The fact that the primary sequences of Hero proteins are poorly conserved (Figures 2G and H) suggests that the molecular nature of Hero proteins as hydrophilic, charged and disordered polymers is important for their functions. Supporting this idea, scrambling the amino-acid sequences of Hero7 and Hero11 did not compromise their anti-aggregation activities for TDP-43ΔNLS in cells (Figure 4D). This observation challenges our conventional view of protein functions, in which the amino acid sequence of a protein determines its unique three-dimensional structure and thereby its function. Given the apparent importance of Hero proteins' physical properties, it is not surprising that other charged biopolymers such as nucleic acids are likely to play similar roles as protein stabilizers, as we observed (Figures 1B and E) and as reported recently (Docter et al., 2016; Maharana et al., 2018).

Historically, the goal of biochemistry is to deconstruct biological events, identify enzymes responsible for the reactions of interest, and reconstitute them in a cell-free system. However, the enzymatic activity in reconstituted systems is often much less efficient than in crude cell lysate (Figures 1C and D), even when enzyme concentrations and buffer conditions are optimized. This may be at least in part explained by the absence of Hero proteins in classical reconstitution systems. If so, we might need to revise our approaches of biochemistry to take into consideration not only catalytic enzymes but also Hero proteins (and other environmental biomolecules). In addition, Hero proteins may be useful as stabilizers for protein pharmaceuticals, especially because those from humans are less likely to evoke immune responses than artificial polymers or biomolecules from other organisms. They may be co-expressed, used as additives like polyethylene glycol (PEG), or can be fused to proteins of interest like PEGylation or PASylation (Schlapschy et al., 2013). The finding that some Hero proteins can block pathogenic protein aggregates in vitro, in cells, and in fly models for neurodegenerative diseases may also serve as a starting point for future development of therapeutics or diagnostic methods.

Each organism appears to encode hundreds of Hero proteins with no apparent common motifs, and each Hero protein is likely to deal with many different client

proteins, thus providing a rich but complex source to mine. Indeed, Hero45, which was originally identified as an mRNA-binding protein (Heaton et al., 2001), has been found in the structure of the 80S ribosomes where it occupies the mRNA-binding channel (Habeck et al., 2013), but it is also known to be associated with hundreds of diverse proteins (Oughtred et al., 2019) with dynamic changes in its subcellular localizations in the cytoplasm, stress granules, nucleoli etc. (Lee et al., 2016, 2014; Yeom et al., 2015), highlighting the vast complexity of Hero proteins' actions. Moreover, individual Hero proteins show some preferences toward client proteins (Figures 3–6), and yet they may be at least partially overlapping in their functions. A combination of comprehensive approaches will be needed to decipher the global relationship among Hero proteins and their clients, including genetics to understand the physiological roles of Hero proteins, biochemistry and biophysics to dissect the molecular mechanisms of Hero protein actions, and evolutionary analyses to understand the birth and expansion of these genes.

Figure Legends

Figure 1. Activity of heat-soluble proteins and nucleonic acids on Ago2 molecular behavior and LDH desiccation.

(A) Schematic representation of FLAG-TEV-SNAP-Ago2 immunopurified on magnetic beads.

(B) FLAG-TEV-SNAP-Ago2 was immunopurified onto magnetics beads via anti-FLAG antibody and then the FLAG tag was cleaved off by TEV protease in buffer, crude lysates or their boiled supernatants from fly S2 or human HEK293T cells. PK indicates that the boiled supernatant had been mostly deproteinized by Proteinase K in advance. BZ indicates that DNA and RNA in boiled supernatant had been degraded by Benzonase. BZ/PK indicates that DNA and RNA had been degraded by Benzonase, followed by the deproteinization by Proteinase K. Eluted Ago2 (top) and Ago2 still remaining on the beads (middle) were visualized by a red fluorescent dye covalently attached to the SNAP tag. The values on the top indicate relative amounts of eluted Ago2, normalized to that with S2 crude cell lysate. Proteins remaining in the boiled supernatants have an activity to promote Ago2 elution.

(C) Small RNA pull-down assay for pre-Ago2-RISC and mature Ago2-RISC assembled in the reconstitution system, containing a ^{32}P-radiolabeled small RNA duplex, Ago2, Dicer-2/R2D2, and the Hsp70/Hsp90 chaperone machinery. Supplementation of the boiled supernatants from S2 or HEK293T cells promoted the formation of both pre- and mature Ago2-RISC.

(D) Quantification of (C). Data represent means ± SD from 3 independent experiments.

(E) Heat-soluble proteins in the boiled supernatants protect the LDH activity from desiccation. LDH mixed with the indicated boiled supernatants or buffer was dried up overnight and the remaining LDH activities were measured. Data represent means ± SD from 3 independent experiments.

Figure 2. Hydrophilic and highly charged proteins are enriched in the boiled supernatants.

(A and B) LC-MS/MS analysis of lysates from *Drosophila* S2 (A) and human HEK293T (B) cells before and after boiling. Scatter plots indicate the PSM values (plus 1, to avoid zeros), which correlate with the protein amount, in the original crude lysates on the X-axes and those in the boiled supernatants on the Y-axes. The identified proteins were divided into five groups according to the degrees of depletion or enrichment by boiling (Group I [most depleted] to Group V [most enriched]) and

used for further analysis in C–J.

(C and D) Violin plots of the median IUPred score of the proteins classified into the five groups in *Drosophila* (C) and humans (D). The proteins enriched by boiling are predicted to be much more disordered.

(E and F) Boxplots of the proportion of hydrophilic or hydrophobic amino acids of the proteins classified into the five groups in *Drosophila* (E) and humans (F). The proteins enriched by boiling are more hydrophilic. Boxplot center and box edges indicate median and 25th or 75th percentiles, respectively, while whiskers indicate the median ± 1.5 × IQR or the most extreme observations within these limits, and outliers are plotted individually.

(G and H) Boxplots of the DIOPT score of the proteins classified into the five groups in *Drosophila* (G) and humans (H). The proteins enriched by boiling are predicted to be evolutionally less conserved. Boxplot center and box edges indicate median and 25th or 75th percentiles, respectively, while whiskers indicate the median ± 1.5 × IQR or the most extreme observations within these limits, and outliers are plotted individually.

(I and J) Violin plots of the pI values of the proteins classified into the five groups in *Drosophila* (I) and humans (J). The proteins enriched by boiling show the pI values concentrated in the acidic (~5) or basic (~10) range but avoiding the neutral (~7–8) range.

(K) Violin plots of the median IUPred score of the entire human proteome and that of the proteins selected by their high IUPred scores and expression levels.

(L) Violin plots of the pI values of the entire human proteome and that of the highly disordered proteins selected by their high IUPred scores and expression levels. The selected proteins also show the pI values concentrated in the acidic or basic range but avoiding the neutral range.

Figure 3. Hero proteins protect the activity of various proteins in stress conditions.

(A) Heat solubility of the six Hero proteins selected. Each Hero protein, GFP or GST was fused with a FLAG tag and expressed in HEK293T cells. The cell lysates were prepared with (+) or without (–) boiling and then the proteins in the lysates or their supernatants were detected by anti-FLAG Western blotting.

(B) Hero proteins protect the LDH activity from desiccation. LDH mixed with the indicated Hero protein or controls was dried up overnight and the remaining LDH activities were measured. Note that much higher concentrations of arginine

(500-fold) and trehalose (100-fold) were used, compared to those of the proteins. Data were normalized to the activity of LDH incubated on ice overnight, and represent means ± SD from 3 independent experiments.

(C) Hero proteins protect GFP in organic solvent. GFP mixed with the indicated Hero protein or controls was exposure to chloroform and the remaining GFP fluorescence was measured. Data were normalized to the activity of GFP without the exposure to chloroform and represent means ± SD from 3 independent experiments.

(D) Hero proteins protect the luciferase activity from heat shock. Firefly luciferase was transfected into HEK293T cells, together with each Hero protein or chaperone factor, and then the cells were exposed to heat shock. Remaining luciferase activities after heat shock were measured. Data were normalized to the activity of luciferase in cells without the exposure to heat shock, and represent means ± SD from 5 independent experiments.

Figure 4. Hero proteins prevent pathogenic protein aggregation in vitro and in cells.

(A) Representative microscopic images of GFP signals in the cells transfected with GFP-tagged TDP-43ΔNLS, HTTQ103, or GA50, together with the GST control or each Hero protein. Note that signals from aggregations are saturated.

(B) Filter trap assay of aggregation-prone proteins expressed in cells. TDP-43ΔNLS, HTTQ103 or GA50 was expressed in a GFP-fusion form in HEK293T cells, together with each Hero protein or the GST control. The original or 5-fold diluted samples were loaded on a cellulose acetate membrane in the presence of 1% SDS and the trapped aggregates were probed with anti-GFP antibody.

(C) Filter trap assay of aggregation-prone proteins expressed in cells. GFP-TDP-43ΔNLS was expressed in HEK293T cells, together with each Hero protein, its KR-less mutant, in which positively charged amino acids (Lys and Arg) were substituted with neutral Gly, or the GST negative control. Five-fold serial dilutions of the cell lysates were subjected to the filter trap assay and probed with the anti-GFP antibody. Experiments were simultaneously performed, and the presented data were spliced from the same membrane.

(D) Filter trap assay of aggregation-prone proteins expressed in cells. GFP-TDP-43ΔNLS was expressed in HEK293T cells, together with the GST control, each Hero protein, or their shuffled amino acid sequences with 42 or 100 amino-acid length. Five-fold serial dilutions of the cell lysates were subjected to the filter trap

assay and probed with anti-GFP antibody. Experiments were simultaneously performed, and the presented data were spliced from the same membrane.

Figure 5. Hero proteins suppress the neurotoxicity of TDP-43ΔNLS aggregates in motor neurons.

(A) Representative images of aggregation in human iPS-derived motor neurons transfected with GFP-TDP-43ΔNLS, together with the GST negative control or Hero7. Note that signals from aggregations are saturated.

(B) Scheme for counting the saturated area of TDP-43ΔNLS aggregates.

(C) Aggregation index analysis in motor neurons. GFP-TDP-43ΔNLS was expressed in motor neurons, together with each Hero protein or the GST negative control. Data indicate normalized saturated area (Aggregation index), and represent means ± SD from 3 independent experiments. Hero 9, 45, 7, 11 reduced the aggregate formation.

(D) Neurite length analysis in motor neurons. Each Hero protein or the GST control was individually expressed in motor neurons. Data indicate neurite length, and represent means ± SD from 8 cells. Hero45 showed a modest defect in neurite outgrowth.

(E) Representative images of the neuron marker TUJ1 in motor neurons transfected with GFP-TDP-43ΔNLS, together with Hero7 or the GST control. Note that signals are saturated to clearly visualize the neurites.

(F) Neurite length quantification in motor neurons transfected with GFP-TDP-43ΔNLS together with each Hero protein or the GST negative control. Data indicate neurite length, and represent means ± SD from 8 cells. Hero9, 7 and 11 rescued the neurite outgrowth defect by GFP-TDP-43ΔNLS aggregation.

Figure 6. Hero proteins suppress the neurotoxicity of protein aggregates in *Drosophila* eyes.

(A) Expression of Hero9, 13, 45, or 11 alone showed no phenotype in the external eye morphology, much as the YFP control (upper). Expression of TDP-43-YFP caused the external eye degeneration. Hero9 and double copies of Hero45 strongly suppressed the eye degeneration by TDP-43-YFP (lower).

(B) Knockdown of Hsc70-4 or CG17931 caused detectable external eye degeneration, compared to the mock knockdown of Piwi.

(C) Nail polish imprint of eyes. Knockdown of Hsc70-4 or CG17931 caused irregular arrangements of ommatidia and bristles.

(D) Knockdown of Hsc70-4 or CG17931 decreased the retinal depth. Arrowhead in the

left picture indicates the retinal depth. Data represent means ± SD from 5 individual flies.

(E) Knockdown Hsc70-4, CG17931, CG14818 or Vig2 caused exacerbation of the eye morphology defect by over-expression of TDP-43-YFP (upper). Knockdown of fly Hero proteins, CG17931, CG12384 or CG11444, caused exacerbation of the eye morphology defect by over-expression of aggregation-prone MJDtr-Q78 (lower). Note that female files were used for E, while male files were used for A–D.

Figure 7. Physiological functions of Hero proteins.

(A) Cell growth assay of HEK293T cells acutely knocked out for the indicated genes. 4 days after transfection of gRNA, the numbers of the cells were measured by a coloring assay. Knockout of Hero13 and Hero7 reduced the cell proliferation.

(B) Whole-body knockdown phenotypes in *Drosophila*. Knockdown of endogenous fly Hero proteins were performed by two independent RNAi strains (GD and KK lines). Whole-body depletion of Hero proteins often caused lethality, suggesting that Hero proteins play essential roles in *Drosophila* early development. Note that we used two independent GD lines for CG3760

(C) Survival plots of *Drosophila*. Whole-body overexpression of CG12384 or CG14818 elongated lifespans, compared to the wild type.

(D) Summary of lifespan analysis in *Drosophila*. Whole-body overexpression of Hero proteins except human Hero11 significantly elongated lifespans.

(E) A model for the modes of action of chaperons and Hero proteins. Chaperones usually act on the inactive/denatured state of their client proteins and actively reverse them to the functional state by using ATP, while Hero proteins likely act on the functional state of their client proteins and protect them from lapsing into the inactive/denatured state in an ATP-independent manner.

Supplemental Figure 1. Prediction of disordered regions in Hero proteins

(A) Amino acid sequences of the six representative Hero proteins.

(B) Disorder prediction by IUPred using the default settings (Dosztányi et al., 2005). GFP and GST were used as controls.

Supplemental Table 1. Shuffled amino-acid sequences

Supplemental Table 2. gRNA sequences for acute KO

Acknowledgements

We are grateful to Qinghua Liu for providing the plasmids for Dicer-2 and R2D2 expression, Connie Cepko for pCAGEN vector, Drosophila Genomics Resource Center, supported by NIH grant 2P40OD010949 for pUASg-HA-attB, Vienna Drosophila Resource Center for the RNAi lines, N. Bonini for UAS-SCA3trQ78, UAS-TDP43YFP and UAS-YFP lines, Aaron Gitler for pAG303-Gal-GA50 (Addgene plasmid #84907), Susan Lindquist for p426 103Q GPD (Addgene plasmid #1184), Bob Weinberg for pCMV-VSV-G (Addgene plasmid # 8454), Didier Trono for psPAX2 (Addgene plasmid # 12260), John Doench & William Hahn & David Root for pLX_311-Cas9 (Addgene plasmid # 96924), and Yuko Fukuda for preparing cryosections. We thank Kaori Kiyokawa for experimental assistances, Hiro-oki Iwakawa for proposing the idea of boiling cell lysates, Life Science Editors for editorial assistance, and Phillip Zamore, Olivia Rissland, Shin-ichi Nakagawa, Hisashi Tadakuma, and the members of the Tomari laboratory for critical comments on the manuscript. This work was supported in part by JSPS KAKENHI Grant Number JP16J02134 and 19J30003 (to K.T.), and MEXT KAKENHI Grant Number JP26113007 and JSPS KAKENHI Grant Number 18H05271 (to Y.T.).

Author contribution

K.T. and Y.T. designed the project. K.T. performed most of the experiments. T.O. performed the motor neuron experiments, supervised by Y.I. H. K.-H. performed LC-MS/MS analysis, supervised by M.O. E. S.-M. conceived the anti-aggregation activity of Hero proteins and supervised *Drosophila* experiments. Y.O. directed imaging of *Drosophila* eye cryosections. S.I. found that the boiled supernatants have an activity to promote Ago2 elution and conceived the concept of Hero proteins. Y.T. supervised the research project. K.T. and Y.T. wrote the manuscript.

Data availability

All data supporting the findings of this study are available from the corresponding authors on reasonable request.

Declaration of interests

Y.T., K.T., S.I., E. S.-M., H. K.-H., and M.O. have a patent application related to this work. The other authors declare no competing interests.

References

Arya, R., and Lakhotia, S. (2006). A simple nail polish imprint technique for examination of external morphology of Drosophila eyes. Curr Sci *90*, 1179–1180.

Barinova, Y., Dietzl, G., Dickson, B.J., Fellner, M., Oppel, S., Su, K.-C., Kinsey, K., Gasser, B., Marra, V., Keleman, K., et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. Nature *448*, 151–156.

Bischof, J., Björklund, M., Furger, E., Schertel, C., Taipale, J., and Basler, K. (2012). A versatile platform for creating a comprehensive UAS-ORFeome library in Drosophila. Dev. *140*, 2434–2442.

Boothby, T.C., Tapia, H., Brozena, A.H., Piszkiewicz, S., Smith, A.E., Giovannini, I., Rebecchi, L., Pielak, G.J., Koshland, D., and Goldstein, B. (2017). Tardigrades Use Intrinsically Disordered Proteins to Survive Desiccation. Mol. Cell *65*, 975-984.e5.

Chakrabortee, S., Boschetti, C., Walton, L.J., Sarkar, S., Rubinsztein, D.C., and Tunnacliffe, A. (2007). Hydrophilic protein associated with desiccation tolerance exhibits broad protein stabilization function. Proc. Natl. Acad. Sci. U. S. A. *104*, 18073–18078.

Chambers, S.M., Qi, Y., Mica, Y., Lee, G., Zhang, X.J., Niu, L., Bilsland, J., Cao, L., Stevens, E., Whiting, P., et al. (2012). Combined small-molecule inhibition accelerates developmental timing and converts human pluripotent stem cells into nociceptors. Nat. Biotechnol. *30*, 715–720.

Chang, H.C., Newmyer, S.L., Hull, M.J., Ebersold, M., Schmid, S.L., and Mellman, I. (2002). Hsc70 is required for endocytosis and clathrin function in Drosophila. J. Cell Biol. *159*, 477–487.

Close, T.J., Kortt, A.A., and Chandler, P.M. (1989). A cDNA-based comparison of dehydration-induced proteins (dehydrins) in barley and corn. Plant Mol. Biol. *13*, 95–108.

Docter, B.E., Horowitz, S., Gray, M.J., Jakob, U., and Bardwell, J.C.A. (2016). Do nucleic acids moonlight as molecular chaperones? Nucleic Acids Res. *44*, 4835–4845.

Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., and Root, D.E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat. Biotechnol. *32*, 1262–1267.

Dosztányi, Z., Csizmók, V., Tompa, P., and Simon, I. (2005). The pairwise energy content estimated from amino acid composition discriminates between folded and

intrinsically unstructured proteins. J. Mol. Biol. *347*, 827–839.

Elden, A.C., Kim, H.J., Hart, M.P., Chen-Plotkin, A.S., Johnson, B.S., Fang, X., Armakola, M., Geser, F., Greene, R., Lu, M.M., et al. (2010). Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. Nature *466*, 1069–1075.

Elkayam, E., Kuhn, C.D., Tocilj, A., Haase, A.D., Greene, E.M., Hannon, G.J., and Joshua-Tor, L. (2012). The structure of human argonaute-2 in complex with miR-20a. Cell *150*, 100–110.

Franzmann, T.M., and Alberti, S. (2019). Protein Phase Separation as a Stress Survival Strategy. Cold Spring Harb. Perspect. Biol. a034058.

Franzmann, T.M., Jahnel, M., Pozniakovsky, A., Mahamid, J., Holehouse, A.S., Nüske, E., Richter, D., Baumeister, W., Grill, S.W., Pappu, R. V., et al. (2018). Phase separation of a yeast prion protein promotes cellular fitness. Science (80-.). *359*.

Fujimori, K., Ishikawa, M., Otomo, A., Atsuta, N., Nakamura, R., Akiyama, T., Hadano, S., Aoki, M., Saya, H., Sobue, G., et al. (2018). Modeling sporadic ALS in iPSC-derived motor neurons identifies a potential therapeutic agent. Nat. Med. *24*, 1579–1589.

Gai, Y.-P., Ji, X.-L., Lu, W., Han, X.-J., Yang, G.-D., and Zheng, C.-C. (2011). A novel late embryogenesis abundant like protein associated with chilling stress in Nicotiana tabacum cv. bright yellow-2 cell suspension culture. Mol. Cell. Proteomics *10*.

Galau, G.A., Hughes, D.W., and Dure, L. (1986). Abscisic acid induction of cloned cotton late embryogenesis-abundant (Lea) mRNAs. Plant Mol. Biol. *7*, 155–170.

Galea, C.A., High, A.A., Obenauer, J.C., Mlshra, A., Park, C.G., Punta, M., Schlesslnger, A., Jing, M., Rost, B., Slaughter, C.A., et al. (2009). Large-scale analysis of thermostable, mammalian proteins provides insights into the intrinsically disordered proteome. J. Proteome Res. *8*, 211–226.

Goto SG. (2001). A novel gene that is up-regulated during recovery from cold shock in Drosophila melanogaster. Gene *270*, 259–264.

Guo, L., Kim, H.J., Wang, H., Chook, M., Taylor, J.P., Shorter, J., O 'donovan, K., Fare, C.M., Diaz, Z., Singh, N., et al. (2018). Nuclear-Import Receptors Reverse Aberrant Phase Transitions of RNA-Binding Proteins with Prion-like Domains. Cell *173*, 677–692.

Habeck, M., Beckmann, R., Anger, A.M., Wilson, D.N., Subklewe, M., Armache, J.-P., and Berninghausen, O. (2013). Structures of the human and Drosophila 80S ribosome. Nature *497*, 80–85.

Hagedorn, E.J., Bayraktar, J.L., Kandachar, V.R., Bai, T., Englert, D.M., and Chang,

H.C. (2006). Drosophila melanogaster auxilin regulates the internalization of Delta to control activity of the Notch signaling pathway. J. Cell Biol. *173*, 443–452.

van Ham, T.J., Holmberg, M.A., van der Goot, A.T., Teuling, E., Garcia-Arencibia, M., Kim, H. eui, Du, D., Thijssen, K.L., Wiersma, M., Burggraaff, R., et al. (2010). Identification of MOAG-4/SERF as a regulator of age-related proteotoxicity. Cell *142*, 601–612.

Hand, S.C., Menze, M.A., Toner, M., Boswell, L., and Moore, D. (2011). LEA Proteins During Water Stress: Not Just for Plants Anymore. Annu. Rev. Physiol. *73*, 115–134.

Heaton, J.H., Dlakic, W.M., Dlakic, M., and Gelehrter, T.D. (2001). Identification and cDNA Cloning of a Novel RNA-binding Protein that Interacts with the Cyclic Nucleotide-responsive Sequence in the Type-1 Plasminogen Activator Inhibitor mRNA. J. Biol. Chem. *276*, 3341–3347.

Hu, Y., Flockhart, I., Vinayagam, A., Bergwitz, C., Berger, B., Perrimon, N., and Mohr, S.E. (2011). An integrative approach to ortholog prediction for disease-focused and other functional studies. BMC Bioinformatics *12*.

Iwasaki, S., Sasaki, H.M., Sakaguchi, Y., Suzuki, T., Tadakuma, H., and Tomari, Y. (2015). Defining fundamental steps in the assembly of the Drosophila RNAi enzyme complex. Nature *521*, 533–536.

Jovičič, A., Mertens, J., Boeynaems, S., Bogaert, E., Chai, N., Yamada, S.B., Paul, J.W., Sun, S., Herdy, J.R., Bieri, G., et al. (2015). Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. Nat. Neurosci. *18*, 1226–1229.

Kaushik, S., and Cuervo, A.M. (2015). Proteostasis and aging. Nat. Med. *21*, 1406–1415.

Kawada, J., Kaneda, S., Kirihara, T., Maroof, A., Levi, T., Eggan, K., Fujii, T., and Ikeuchi, Y. (2017). Generation of a Motor Nerve Organoid with Human Stem Cell-Derived Neurons. Stem Cell Reports *9*, 1441–1449.

Kim, T.D., Ryu, H.J., Cho, H. Il, Yang, C.H., and Kim, J. (2000). Thermal behavior of proteins: Heat-resistant proteins and their heat-induced secondary structural changes. Biochemistry *39*, 14839–14846.

Kobayashi, H., and Tomari, Y. (2016). RISC assembly: Coordination between small RNAs and Argonaute proteins. Biochim. Biophys. Acta - Gene Regul. Mech. *1859*, 71–81.

Krobitsch, S., and Lindquist, S. (2000). Aggregation of huntingtin in yeast varies with the length of the polyglutamine expansion and the expression of chaperone proteins. Proc. Natl. Acad. Sci. U. S. A. *97*, 1589–1594.

Kumar, A., and Tiwari, A.K. (2018). Molecular Chaperone Hsp70 and Its Constitutively Active Form Hsc70 Play an Indispensable Role During Eye Development of Drosophila melanogaster. Mol. Neurobiol. *55*, 4345–4361.

Layton, C.J., McMahon, P.L., and Greenleaf, W.J. (2019). Large-Scale, Quantitative Protein Assays on a High-Throughput DNA Sequencing Chip. Mol. Cell *73*, 1075-1082.e4.

Lee, K.H., Zhang, P., Kim, H.J., Mitrea, D.M., Sarkar, M., Freibaum, B.D., Cika, J., Coughlin, M., Messing, J., Molliex, A., et al. (2016). C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. Cell *167*, 774-788.e17.

Lee, Y.J., Wei, H.M., Chen, L.Y., and Li, C. (2014). Localization of SERBP1 in stress granules and nucleoli. FEBS J. *281*, 352–364.

Liang, X., Potter, J., Kumar, S., Zou, Y., Quintanilla, R., Sridharan, M., Carte, J., Chen, W., Roark, N., Ranganathan, S., et al. (2015). Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. J. Biotechnol. *208*, 44–53.

Liu, Y., Chakrabortee, S., Li, R., Zheng, Y., and Tunnacliffe, A. (2011). Both plant and animal LEA proteins act as kinetic stabilisers of polyglutamine-dependent protein aggregation. FEBS Lett. *585*, 630–634.

López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M., and Kroemer, G. (2013). The hallmarks of aging. Cell *153*, 1194.

Maharana, S., Wang, J., Papadopoulos, D.K., Richter, D., Pozniakovsky, A., Poser, I., Bickle, M., Rizk, S., Guillén-Boixet, J., Franzmann, T.M., et al. (2018). RNA buffers the phase separation behavior of prion-like RNA binding proteins. Science (80-.). *360*, 918–921.

Matsuo, N., Goda, N., Shimizu, K., Fukuchi, S., Ota, M., and Hiroaki, H. (2018). Discovery of Cryoprotective Activity in Human Genome-Derived Intrinsically Disordered Proteins. Int. J. Mol. Sci. *19*, 401.

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.

Newman, C.E., Toxopeus, J., Udaka, H., Ahn, S., Martynowicz, D.M., Graether, S.P., Sinclair, B.J., and Percival-Smith, A. (2017). CRISPR-induced null alleles show that Frost protects Drosophila melanogaster reproduction after cold exposure. J. Exp. Biol. *220*, 3344–3354.

Oates, M.E., Romero, P., Ishida, T., Ghalwash, M., Mizianty, M.J., Xue, B., Dosztányi,

Z., Uversky, V.N., Obradovic, Z., Kurgan, L., et al. (2013). D2P2: Database of disordered protein predictions. Nucleic Acids Res. *41*, 508–516.

Oughtred, R., Stark, C., Breitkreutz, B.J., Rust, J., Boucher, L., Chang, C., Kolas, N., O'Donnell, L., Leung, G., McAdam, R., et al. (2019). The BioGRID interaction database: 2019 update. Nucleic Acids Res. *47*, D529–D541.

Panganiban, B., Qiao, B., Jiang, T., DelRe, C., Obadia, M.M., Nguyen, T.D., Smith, A.A.A., Hall, A., Sit, I., Crosby, M.G., et al. (2018). Random heteropolymers preserve protein function in foreign environments. Science *359*, 1239–1243.

Schlapschy, M., Binder, U., Börger, C., Theobald, I., Wachinger, K., Kisling, S., Haller, D., and Skerra, A. (2013). PASylation: A biological alternative to PEGylation for extending the plasma half-life of pharmaceutically active proteins. Protein Eng. Des. Sel. *26*, 489–501.

Stewart, S.A., Dykxhoorn, D.M., Palliser, D., Mizuno, H., Yu, E.Y., An, D.S., Sabatini, D.M., Chen, I.S.Y., Hahn, W.C., Sharp, P.A., et al. (2003). Lentivirus-delivered stable gene silencing by RNAi in primary cells. Rna *9*, 493–501.

Tatar, M., Khazaeli, A.A., and Curtsinger, J.W. (1997). Chaperoning extended life. Nature *390*, 30.

Tsuboyama, K., Tadakuma, H., and Tomari, Y. (2018). Conformational Activation of Argonaute by Distinct yet Coordinated Actions of the Hsp70 and Hsp90 Chaperone Systems. Mol. Cell *70*, 722-729.e4.

Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. Science *347*, 1260419.

Wang, N., Xiang, J., Ma, Z., Quan, J., Chen, J., and Yang, Z. (2008). A concise and diversity-oriented approach to the synthesis of SAG derivatives. J. Comb. Chem. *10*, 825–834.

Warrick, J.M., Paulson, H.L., Gray-Board, G.L., Bui, Q.T., Fischbeck, K.H., Pittman, R.N., and Bonini, N.M. (1998). Expanded polyglutamine protein forms nuclear inclusions and causes neural degeneration in Drosophila. Cell *93*, 939–949.

Wise, M.J., and Tunnacliffe, A. (2004). POPP the question: What do LEA proteins do? Trends Plant Sci. *9*, 13–17.

Yamaguchi, A., Tanaka, S., Yamaguchi, S., Kuwahara, H., Takamura, C., Imajoh-Ohmi, S., Horikawa, D.D., Toyoda, A., Katayama, T., Arakawa, K., et al. (2012). Two Novel Heat-Soluble Protein Families Abundantly Expressed in an Anhydrobiotic Tardigrade. PLoS One *7*, 1–7.

Yeom, J., Kim, J.-H., Kim, S.-Y., Jeong, S., Kwak, H., Ahn, J.-W., Min, K., Baek, S.-J.,

Na, W., Kim, S., et al. (2015). SERBP1 affects homologous recombination-mediated DNA repair by regulation of CtIP translation during S phase. Nucleic Acids Res. *43*, 6321–6333.

Experimental Procedures

General methods

Drosophila S2 cells were cultured at 27°C in Schneider's *Drosophila* Medium (Gibco) supplemented with 10% fetal bovine serum (Gibco) and Antibiotic-Antimycotic (Gibco). HEK293T cells were maintained in Dulbecco's Modified Eagle Medium (Sigma) supplemented with 10% FBS in 5% CO_2 atmosphere at 37°C. Lysis buffer (30 mM HEPES-KOH pH 7.4, 100 mM KOAc, 2 mM $Mg(OAc)_2$), wash buffer (lysis buffer containing 1% Trition X-100 and 800 mM NaCl) and hypotonic lysis buffer (10 mM HEPES-KOH pH 7.4, 10 mM KOAc, 1.5 mM $Mg(OAc)_2$, 1× Complete EDTA-free protease inhibitor cocktail [Roche] and 5 mM DTT) were used. To prepare lysate from S2 cells or HEK293T cells, cells were collected and washed once with PBS. Cell pellets were resuspended with an equal volume of hypotonic lysis buffer, incubated for 15 min, vortexed for 15 sec, and centrifuged at 17,000 × g for 20 min. The supernatant was collected into tubes and shock-frozen in liquid nitrogen and stored at –80 °C. For boiling, lysates were heated at 95°C for 15 min, kept on ice for 1 min, then centrifuged at 20,000 × g for 5 min to remove aggregation.

Plasmid construction

pCold-Hsc70-4, Hsp83, Hop, Droj2 and p23 and pCAGEN-Flag-DEST were described previously (Iwasaki et al., 2015).

pCAGEN-FlagTevSnap-DEST and pCAGEN-GFP-DEST

A DNA fragment containing N-terminal FLAG tag, TEV protease recognition sequence, and SNAP tag or the GFP sequence followed by the Gateway attR sites was inserted into pCAGEN by NEBuilder HiFi DNA Assembly Master Mix (NEB).

pCAGEN-FlagTevSnap-Ago2(WT)

A DNA fragment containing a codon-optimized *Drosophila* Ago2 CDS sequence in pENTR/D-Ago2 (Tsuboyama et al., 2018) was inserted into pCAGEN-*FlagTevSnap*-DEST using Gateway LR Clonase (Invitrogen).

pCAGEN-Flag-GFP, pCAGEN-Flag-GST, and pCAGEN-FLuc

A DNA fragment containing GFP or GST with N-terminal FLAG tag, or Firefly luciferase was inserted into pCAGEN by NEBuilder HiFi DNA Assembly Master Mix (NEB).

pCAGEN-Flag-Hero9, 20, 13, 45, 7 or 11

A DNA fragment containing Hero9 (C9orf16), 20 (C11orf58), Hero13 (BEX3), Hero45 (SERBP1), Hero7 (SERF2), or Hero11 (C19orf53) was amplified from HEK293T cell cDNA by PCR and cloned into pENTR/D-TOPO (Invitrogen), followed by the recombination with pCAGEN-Flag-DEST using Gateway LR Clonase (Invitrogen).

pCAGEN-HTTQ103-GFP

A DNA fragment containing HTT103Q-GFP was amplified from p426 103Q GPD (Addgene #1184; (Krobitsch and Lindquist, 2000)) and inserted into pCAGEN by NEBuilder HiFi DNA Assembly Master Mix (NEB).

pCAGEN-GFP-TDP43ΔNLS

A DNA fragment containing TDP-43 was amplified from HEK293T cell cDNA by PCR and cloned into pENTR/D-TOPO (Invitrogen), followed by the recombination with pCAGEN-GFP-DEST using Gateway LR Clonase (Invitrogen). NLS deletion (78–99) in TDP-43 was performed by PCR-based site-directed mutagenesis.

pCAGEN-GFP-GA50

A DNA fragment containing GA50 was amplified from pAG303-Gal-GA50 (Addgene # 84907; (Jovičič et al., 2015)) and a DNA fragment containing GFP were inserted into pCAGEN by NEBuilder HiFi DNA Assembly Master Mix (NEB).

pCold I-GFP, GST, Hero9, Hero20, 13, 45, 7 or 11-FlagHis

A DNA fragment containing CDS of each Hero protein and C-terminal FLAG and His tags was inserted into pCold I (Takara) by NEBuilder HiFi DNA Assembly Master Mix (NEB). For BEX3 and SERBP1, codon-optimized DNA fragment was synthesized (Eurofins) and utilized as the templates.

pCold I-GstTev-TDP-43-HA

A DNA fragment containing N-terminal GST tag and TEV protease recognition sequence was inserted into pCold I (Takara), followed by the insertion of a DNA fragment containing C-terminal 2×HA tag by NEBuilder HiFi DNA Assembly Master Mix (NEB). Finally, a DNA fragment containing TDP-43 was inserted into pCold I-GstTev-HA by NEBuilder HiFi DNA Assembly Master Mix (NEB).

pCAGEN-Flag-KR-less Hero7 or Hero11

A DNA fragment containing mutants of Hero7 or Hero11, in which positively charged amino acids (Lys and Arg) were substituted with neutral Gly, was inserted into pCAGEN-Flag-DEST vector.

pCAGEN-Flag-Shuffled100[GST, Hero7 or 11]

A DNA fragment containing shuffled 100-aa sequences with the composition of GST, SERBP1, Hero7 or Hero11 was inserted into pCAGEN-Flag-DEST vector. The amino acid sequences are shown in Table S1.

pCAGEN-superFlag-Shuffled42[GST, Hero7, or 1]

A DNA fragment containing N-terminal superFlag (Layton et al., 2019) and shuffled 42-aa sequences with the composition of GST, SERBP1, Hero7 or Hero11 was inserted into pCAGEN vector. The amino acid sequences are shown in Table S1.

Protein purification

Recombinant proteins of Hsc70-4, Hsp83, Hop, Droj2, p23, and Dicer-2/R2D2 heterodimer were expressed and purified as described previously (Iwasaki et al., 2015). The same procedure was used to express and purify recombinant EGFP proteins. Recombinant proteins of GST, Hero9, 20, 13, 45, 7, and 11 were expressed as C-terminal His-tagged proteins in *Escherichia coli* Rosetta 2 (DE3) strain. Typically, the cells were cultivated in 1 L culture to an OD_{600} of 0.4–0.6 at 37°C with ampicillin and then grown at 15°C overnight with 1 mM isopropyl-β-D-thiogalactoside (IPTG). The cell pellets were resuspend in His A buffer (30 mM HEPES-KOH pH 7.4, 200 mM KOAc, 2 mM $Mg(OAc)_2$, 5% glycerol) containing 1×EDTA-free protease inhibitor cocktail (Roche), sonicated and centrifuged at 10,000 × g for 5 min twice. The supernatant was added to a slurry of cOmplete His-Tag Purification Resin (Roche) and incubated for 1 h at room temperature or 2 h at 4°C, then eluted with His B buffer (His A buffer containing 400 mM imidazole). The eluents were collected and buffer exchanged to PBS with PD-10 (GE Healthcare). Recombinant proteins of GST-TEV-TDP-43-HA were expressed in *E. coli* Rosetta 2 (DE3) strain. Typically, the cells in 1 L cultures were cultivated to an OD_{600} of 0.4–0.6 at 37°C with ampicillin and then grown at 15°C overnight with 1 mM isopropyl-β-D-thiogalactoside (IPTG). The cell pellets were resuspend in PBS containing 0.5 M NaCl, 0.1% Triton, 10 mg/L RNaseA (Nacalai tesque), and 1×EDTA-free protease inhibitor cocktail (Roche) (Guo et al., 2018), sonicated and centrifuged at 10,000 × g for 5 min twice. The supernatant was added to a slurry of Glutathione Sepharose 4 Fast Flow (GE Healthcare) and incubated for 2 h at 4°C, then

eluted with GST elution buffer (100 mM Tris-HCl pH 8.0, 20 mM Glutathione, 100 mM NaCl). The eluents were collected and buffer exchanged to PBS with PD-10 (GE Healthcare).

Ago2 elution assay

FLAG-TEV-SNAP-Ago2 expressed in HEK293T cells was labeled with SNAP-Surface Alexa Fluor 647 (NEB) and immunopurified by an anti-FLAG antibody conjugated onto Dynabeads Protein G (Invitrogen). The beads were washed three times by wash buffer, three times by wash buffer containing 2 mM ATP, and finally rinsed by lysis buffer. The beads were incubated with lysis buffer, crude lysates or their boiled supernatants containing 2 U/µL TurboTEV protease (Accelagen) at room temperature for 1 h. To prepare Benzonase and Proteinase K-treated (BZ/PK) supernatants, the supernatants were mixed with 1 U/mL Benzonase (Millipore) and incubated at 37°C for 1 h , then mixed 2 mg/mL Proteinase K and incubated at 37°C for 1 h, and finally reboiled at 95°C for 15 min to deactivate Proteinase K. The Benzonase or Proteinase K treatment was omitted for the PK or BZ condition, respectively.

Small RNA pull-down assay

Small RNA pull-down to detect pre-Ago2-RISC and mature Ago2-RISC formation was performed as described previously (Iwasaki et al., 2015) with slight modifications. FLAG-TEV-Halo-Ago2 expressed in S2 cells were immunopurified by an anti-FLAG antibody conjugated onto Dynabeads Protein G (Invitrogen). The beads were washed three times by wash buffer, three times by wash buffer containing 2 mM ATP, and finally rinsed by lysis buffer. The Ago2-conjugated beads were then incubated with the reconstitution system, ^{32}P-radiolabeled small RNA duplex, ATP regeneration system (1 mM ATP, 25 mM creatine monophosphate [Sigma], 0.03 U/µL creatine kinase [Calbiochem], and 0.1 U/µL RNasin Plus RNase Inhibitor [Promega]) and 3.5 µL of each boiled supernatant in 10 µL lysis buffer at 25°C for 60 min. The reconstitution system contained 20 nM Dicer-2/R2D2, 600 nM Droj2, 1.8 µM Hsc70-4, 1.5 µM Hop, 3 µM Hsp83, and 3 µM p23. In every case, the total protein concentration was adjusted by the supplementation of GST. After incubation, beads were washed four times by wash buffer, then bound RNAs were extracted by Proteinase K treatment and analyzed by native PAGE.

Mass Spectrometry

Proteins in the crude lysates and their boiled supernatants were trypsinized and

subjected to nano liquid chromatography tandem mass spectrometry (nanoLC-MS/MS) analysis. For this analysis, we used LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific) coupled with Dina-2A nanoflow LC system (KYA Technologies). The MS/MS signals then were processed against the UniProt human proteome (Homo sapiens; UP000005640) and fly proteome (*Drosophila melanogaster*; UP000000803) using the Mascot algorithm (version 2.5.1; Matrix Science). Protein identification was based on Mascot scores that exceeded the threshold (a false discovery rate [FDR] < 0.01). One (1) was added to all the peptide spectrum match (PSM) values (to avoid zeros), which correlate with the protein amount, and the ratio between (PSM + 1) in the original crude lysate and that in the boiled supernatant was calculated in order to divide the identified proteins into 5 groups, each of which contained virtually the same number of proteins, according to the degree of enrichment or depletion by boiling.

Prediction of structural disorders

Amino acid sequences were collected from UniProt human proteome (Homo sapiens; UP000005640) and fly proteome (*Drosophila melanogaster*; UP000000803). These sequences were subjected to the IUPred analysis (Dosztányi et al., 2005). The median IUPred score of all amino acid residues was allocated for each protein, and their distributions were plotted on histograms.

Ortholog prediction

Ortholog prediction was performed by DIOPT, DRSC (*Drosophila* RNAi Screening Center) Integrative Ortholog Prediction Tool. The scores of best matched proteins between *Drosophila* and humans were plotted.

Selection of representative Hero proteins

First, proteins with high IUPred scores (median of IUPred scores of all residues > 0.6, and >75 % of the residues have IUPred scores > 0.4) and universally high expression profiles (median RPKM >25 in all tissues and cultured cells) were selected from the human protein atlas (Uhlén et al., 2015). The proteins with clear organelle localization and/or known cellular functions were excluded. Finally, C9orf16 (Uniprot identifier; Q9BUW7), BEX3 (Q00994), C11orf58 (O00193), SERBP1 (Q8NC51), SERF2 (P84101), and C19orf53 (Q9UNZ5) were selected for further analyses.

Western blot analysis

Anti-DDDDK antibody (M185, MBL) was used as the primary antibody at 1:10000.

Chemiluminescence was induced by Luminata Forte Western HRP Substrate (Millipore) and images were acquired by Amersham Imager 600 (GE Healthcare).

Luciferase assay with a heat shock

HEK293T cells in a 96-well plate were transfected with 10 ng pCAGEN-FLuc and 90 ng empty pCAGEN or pCAGEN-Flag-Hero9, 20, 13, 45, 7, or 11 by using Lipofectamine 3000 (Thermo Fisher Scientific). After ~48 h, 4 mM cycloheximide was added to the medium, cells were incubated at 37°C for 10 min and then subjected to a heat shock at 45°C for 8 min. The cells were lysed by Passive Lysis Solution (Perkin Elmer) and incubated for 15 min at room temperature, and the lysates were collected. The luciferase activity of Fluc was measured by using sensilite Enhanced Flash Luminescence (Perkin Elmer).

LDH dehydration assay

LDH from rabbit muscle (Roche) was diluted to 5 µg/mL in 5-fold diluted boiled supernatants (for Figure 1E), or PBS containing 4 µg/mL BSA, GST or each Hero protein, 400 µg/mL Trehalose or 2 mg/mL Arginine (for Figure 2D). Samples were desiccated in SpeedVac vacuum concentrator (Thermo) for 16 h without heating. For normalization, the same amount of LDH in the same conditions was kept on ice for the same duration as desiccation. Dried samples were rehydrated by PBS, and the LDH activity was measured by Cytotoxicity LDH Assay Kit-WST (Dojindo).

Chloroform exposure assay of GFP

10 µg/mL EGFP was mixed with 50 µg/mL BSA, GST or each Hero protein, followed by the addition of the same volume of chloroform. The samples were continuously shaken at room temperate for 30 min and centrifuged at 10,000×g for 1 min. The EGFP intensity in the supernatants (i.e., in the water phase) was measured.

Filter trap assay of aggregates

HEK293T cells in a 12-well plate were transfected with 100 ng pCAGEN-GFP-HTTQ103, 100 ng pCAGEN-GFP-GA50 or 200 ng pCAGEN-GFP-TDP43ΔNLS together with 900 ng (for HTTQ103 and GA50) or 800 ng (for TDP43ΔNLS) of pCAGEN-Flag-Hero9, 20, 13, 45, 7, or 11 by using Lipofectamine 3000 (Thermo Fisher Scientific). After ~48 h, the cells were resuspended in 200 µL PBS containing 1× Complete EDTA-free protease inhibitor cocktail (Roche) and sonicated by Bioruptor II (BMBio). The total protein concentrations were adjusted and then 1% SDS was added. Subsequently, a 5-fold

volume of 1% SDS in PBS was added, then loaded onto a cellulose acetate membrane with a 0.2 μm pore size (GE Healthcare) that had been incubated in 1% SDS in PBS. After washes with 1% SDS, aggregates were detected by using anti-GFP antibody (Santa Cruz; sc-9996; 1:10000).

Microscopy

Cell images were acquired by an inverted type microscopy (IX83, Olympus) with 20× objective lens (UCPLFLN 20×, 0.7 NA, Olympus).

iPS-derived motor neuron differentiation and plasmid transfection

The procedure of motor neuron differentiation has been described previously (Kawada et al., 2017). Briefly, induced pluripotent stem cells (409B2 from Riken cell bank) were seeded to Matigel coated 12-well plate at a density of 70-80% confluency in mTeSR plus medium (STEMCELL technology) with 10 μM of Y-23632 (Rock inhibitor, Wako). After the cells reached to 95% confluent, the differentiation into neural lineage cells was initiated in DMEM/F12 (Sigma) supplemented with 15% knockout serum replacement (KSR, Gibco), 1% GlutaMAX (Gibco), 1% non-essential amino acid (NEAA, Sigma), 10 μM of SB431542 (Wako), and 100 nM of LDN-193189 (Wako) for initial 6 days of culture. For caudalization and ventralization, then, the cells were further supplemented with and 1 μM of retinoic acid (RA) and 1 μM of smoothened against (SAG) and 10 μM of SU-5402 (Sigma), 10 μM of DAPT (Sigma) which accelerate neural differentiation (Chambers et al., 2012) for 4-12 days of culture. SAG was synthesized with previously described protocol and purified by HPLC (Wang et al., 2008). Then, the motor neurons (MN) were maintained in Neurobasal medium (Thermo Fisher Scientific) supplemented with 2% B27 (Gibco), 1% GlutaMAX, 1% of penicillin/streptomycin, and brain-derived neurotrophic factor (BDNF), maintenance medium. Before the Hero plasmid transfection, the MNs were re-seeded into Matrigel coated 96-well plate at a density of 2000 cells/cm2 dissociated by Accutase (Innovative cell technology) for 10-15 min. Then, after 5-7 days from re-seeding, GFP-TDP-43ΔNLS (200 ng/wells) plasmid with and without each Hero plasmid (100 ng/wells) was transfected by Lipofectamine Stem Transfection Reagent in maintenance medium. Then, the aggregation which is observed as GFP saturated area was quantified after 48 hours from the transfection. To quantify the neurite length, the cells were fixed by 4% paraformaldehyde for 15 min and permeabilized by 0.1% Triton-X 100 for 5 min, followed by blocking with 1% BSA. The cells were then incubated with anti-TUJ1 mouse antibody (Biolegend, 1:1000) in PBS at 4°C for overnight and then further incubated with Alexa Fluor 647-conjugated

anti-mouse IgG (Thermo Fisher Scientific, 1:1000) in PBS at 4°C for overnight with Hoechst 33342 for nuclear staining, following three times wash with PBS. The fluorescent images were acquired with a 20x objective by Axio observer (Zeiss).

Image analysis for protein aggregation and quantifying neurite length

The images which show the exogenous expression of transfected GFP-TDP-43 were converted to 16-bit monochrome images and then masked according to saturated pixels (Fig. 5B, red area). The images were binarized and calculated the total number of saturated GFP pixels. Aggregation index is defined by the percentage of the number of GFP saturated pixels divided by total pixels in ROI. All post-image analysis was conducted by using Fiji (http://imagej.nih.gov/ij/). To quantify the neurite length and number of the branches of motor neuron, the neurites which are identified by TUJ1 positive features were traced by using Simple Neurite Tracer plugin (http://imagej.net/Simple_Neurite_Tracer) in Fiji.

***Drosophila* eye degeneration assay**

Fly culture and crosses were carried out at 25°C. All general fly stocks and RNAi lines were obtained from the Bloomington, Kyoto Fly Stock Center, FlyORF (Bischof et al., 2012) or Vienna Drosophila Resource Center (Barinova et al., 2007). Flies carrying pUASg.attB-Hero9, 13, 45, 11 and Vig2 transgenes were generated by standard phiC31 integrase-mediated transgenesis (Bestgene Inc.). Gmr-Gal4 was used to express transgenes and/or induce RNAi by long hairpins in retina. UAS-MJDtr-Q78 (Warrick et al., 1998), UAS-TDP-43-YFP, and UAS-YFP lines (Elden et al., 2010) were kindly provided by Dr. Nancy Bonini. Genotypes in Figure 6 were as follows: In (A), on the upper row, Oregon R is +/+;+/+. Control (YFP), Hero9, 13, 45, or 11 is *gmr-GAL4*; *UAS-YFP*, *UAS-Hero9*, *13*, *45,* or *11*/*SM6a-TM6B*. On the lower row, – is *gmr-GAL4, UAS-TDP-43-YFP*/+; *TM3*/+, Control (YFP) is *gmr-GAL4, UAS-TDP-43-YFP*/+*; UAS-YFP*/+, and Hero9, 13, 45, or 11 is *gmr-GAL4, UAS-TDP-43-YFP*/+; *UAS-Hero9, 13, 45, or 11*/+. Hero45(×2) is *gmr-GAL4, UAS-TDP-43-YFP*/+; *UAS-Hero45*/*UAS-Hero45*. External eyes were imaged in 1d-old male adult flies. In (B), – is *gmr-GAL4/+*, and Piwi, Hsc70-4, CG17931, CG14818 or Vig2 is *gmr-GAL4/UAS-KK(Piwi), KK(Hsc70-4), KK(CG17931), KK(CG14818) or KK(Vig2)*. In (E) – is *gmr-GAL4, UAS-TDP43-YFP/+*, and Piwi, Hsc70-4, CG17931, CG14818 or Vig2 is *gmr-GAL4, UAS-TDP43-YFP /UAS-KK(Piwi), KK(Hsc70-4), KK(CG17931), KK(CG14818) or KK(Vig2)* on the top row. In the bottom row, – is *gmr-GAL4/+; UAS-MJDtr-Q78/+*, and Piwi, Hsc70-4, CG17931, CG14818 or Vig2 is

gmr-GAL4/UAS-KK(Piwi), KK(Hsc70-4), KK(CG17931), KK(CG14818) or KK(Vig2); UAS-MJDtr-Q78/+. External eyes were imaged in 1d-old female adult flies.

Nail polish imprinting of *Drosophila* eyes

Imprints of *Drosophila* eyes by transparent nail polish were taken essentially as previously described (Arya and Lakhotia, 2006; Kumar and Tiwari, 2018), except that the imprints were mounted in 0.1% Tween 20. The preparations were examined using differential interference contrast (DIC) installed in DeltaVision Elite (GE healthcare).

Imaging of *Drosophila* eye cryosections

Heads of wild-type and RNAi flies were fixed by 2.5% glutaraldehyde, 2% paraformaldehyde and 0.1% Tween 20 in PBS at room temperature for overnight. After washing with PBS, fixed fly heads were embedded in OCT compound (SAKURA Finetek Japan) and subjected to cryo-sectioning. The internal retinal depth was quantified using ImageJ from cryosections at the same anatomical position for all animals. Five female heads were quantified from each genotype.

Proliferation assay of acute-KO HEK293T cells

HEK293T cells stably expressing SpCas9 were established by a lenti-viral system, using pLX_311-Cas9 (Addgene #96924) (Doench et al., 2014), psPAX2 (Addgene #12260) and pCMV-VSV-G (Addgene #8454) (Stewart et al., 2003). gRNAs were prepared as previously reported (Liang et al., 2015), and their sequences are listed in Table S2. Cas9-expressing HEK293T cells in a 24-well plate were transfected with 50 ng of two different gRNAs (100 ng in total) for each target gene by using Lipofectamine RNAiMAX Transfection Reagent (Thermo Fisher Scientific). After ~72 h, cell number was counted by using Cell Counting Kit-8 (Dojindo) according to the manufacturer's instruction.

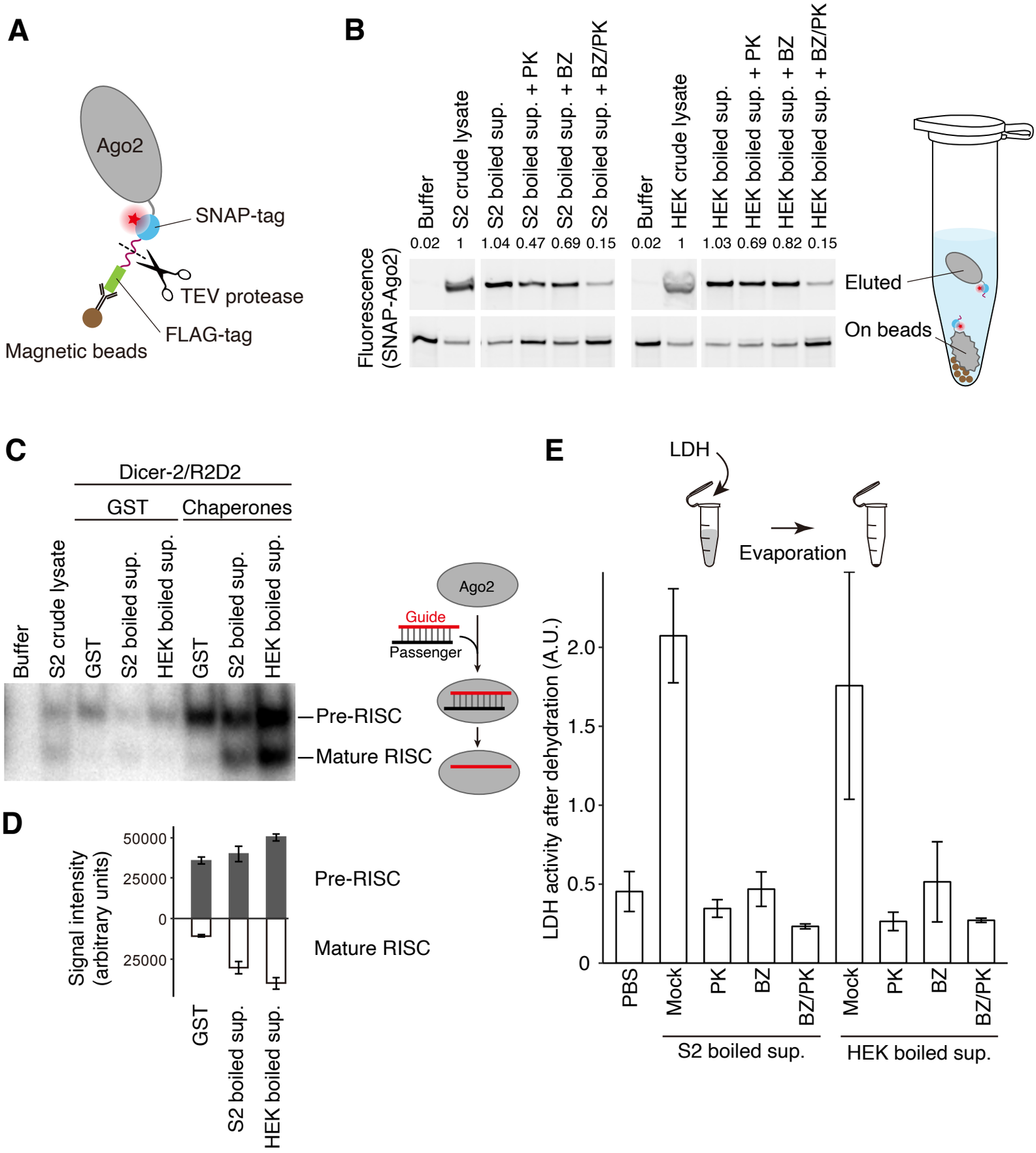
***Drosophila* lethality assay**

Fly culture and crosses were carried out at 25°C, and the viability was checked immediately after emergence. Actin-Gal4 was used to induce RNAi by long hairpins in the whole body. Genotypes in Figure 7B were as follows: *actin-GAL4/UAS-GD(CG14818), actin-GAL4/UAS-KK(CG14818), actin-GAL4/UAS-GD(CG11844; Vig2), actin-GAL4/UAS-KK(CG11844; Vig2), actin-GAL4/UAS-KK(CG17931), actin-GAL4/+; UAS-GD(CG1910)/+, actin-GAL4/UAS-KK(CG1910), actin-GAL4/+; UAS-GD(CG3760-1)/+, actin-GAL4/+; UAS-GD(CG3760-2)/+*.

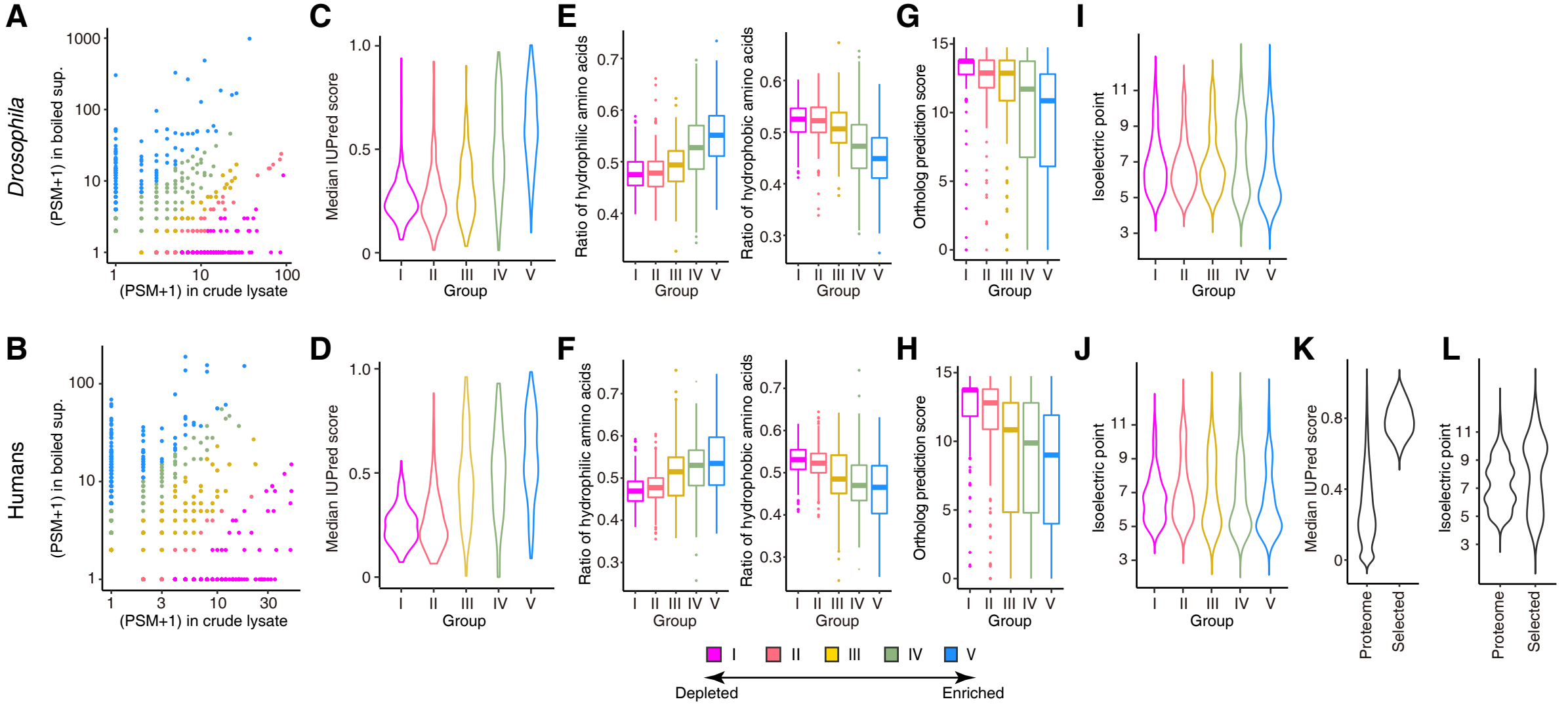
***Drosophila* lifespan assay**

Male files were cultured at 25°C. Every third or fourth day dead flies were scored and the survivors were tipped into a fresh food vial. Kaplan-Meir survival curves were plotted and a log-rank test for trend was performed using R. Actin-Gal4 was used to express trans genes in the whole-body. Genotypes in Figure 7C were as follows: WT is *actin-Gal4/+*, YFP, Hero9, 45, 11, CG14818, or CG12384 is *actin-GAL4/UAS-YFP, Hero9, Hero45, Hero11, CG14818, or CG12384*.

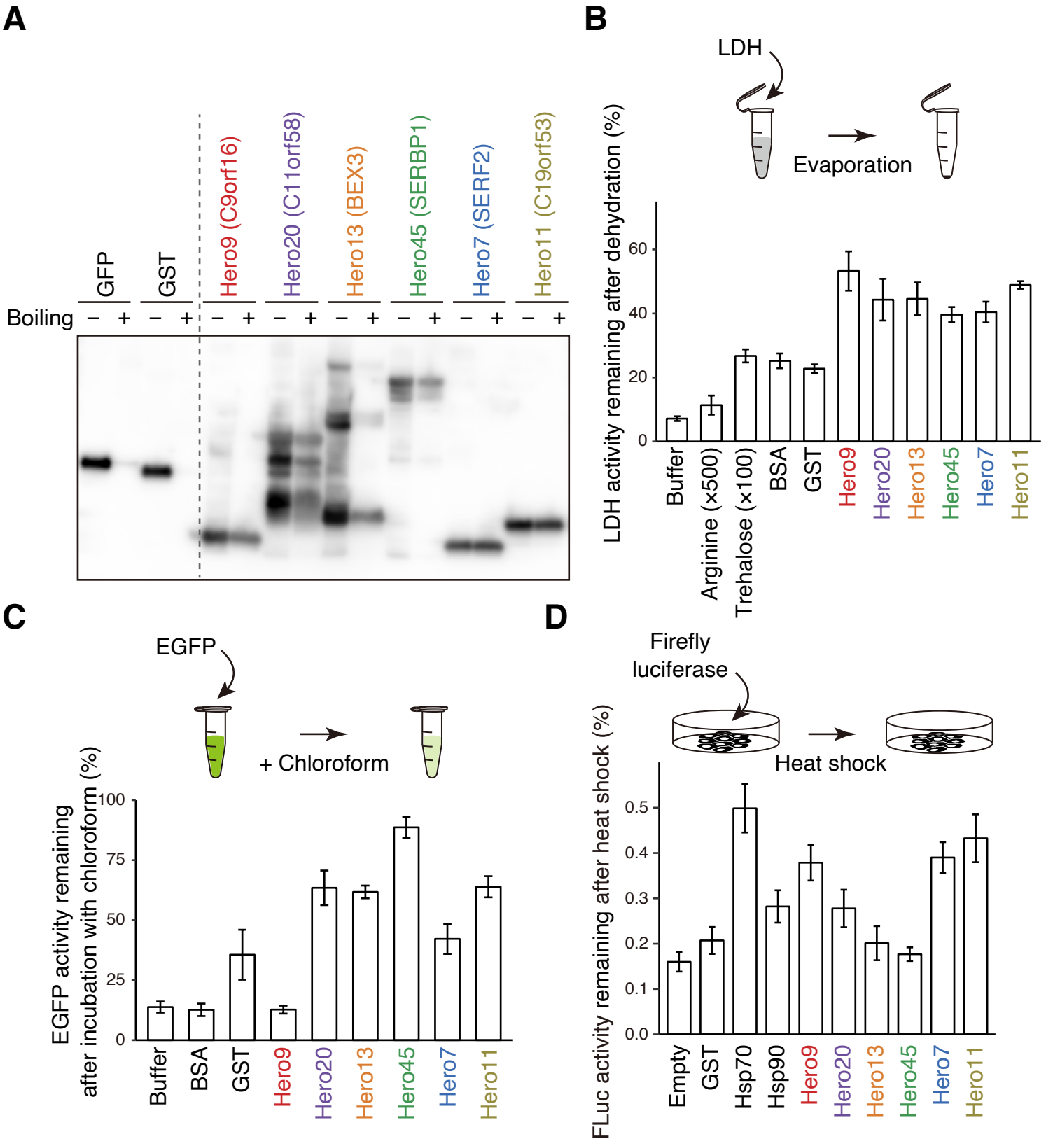
Tsuboyama *et al.* Figure 1



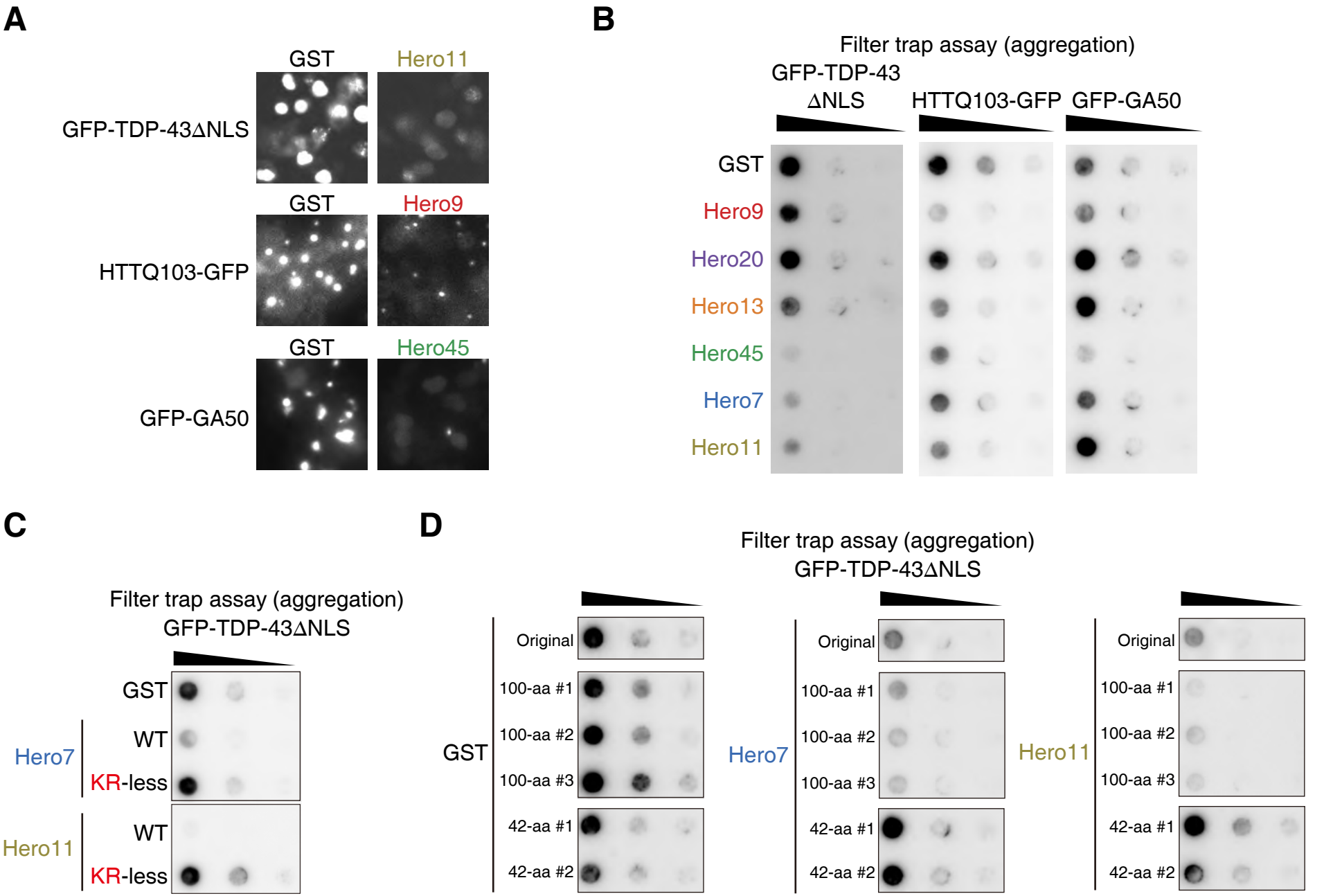
Tsuboyama *et al.* Figure 2



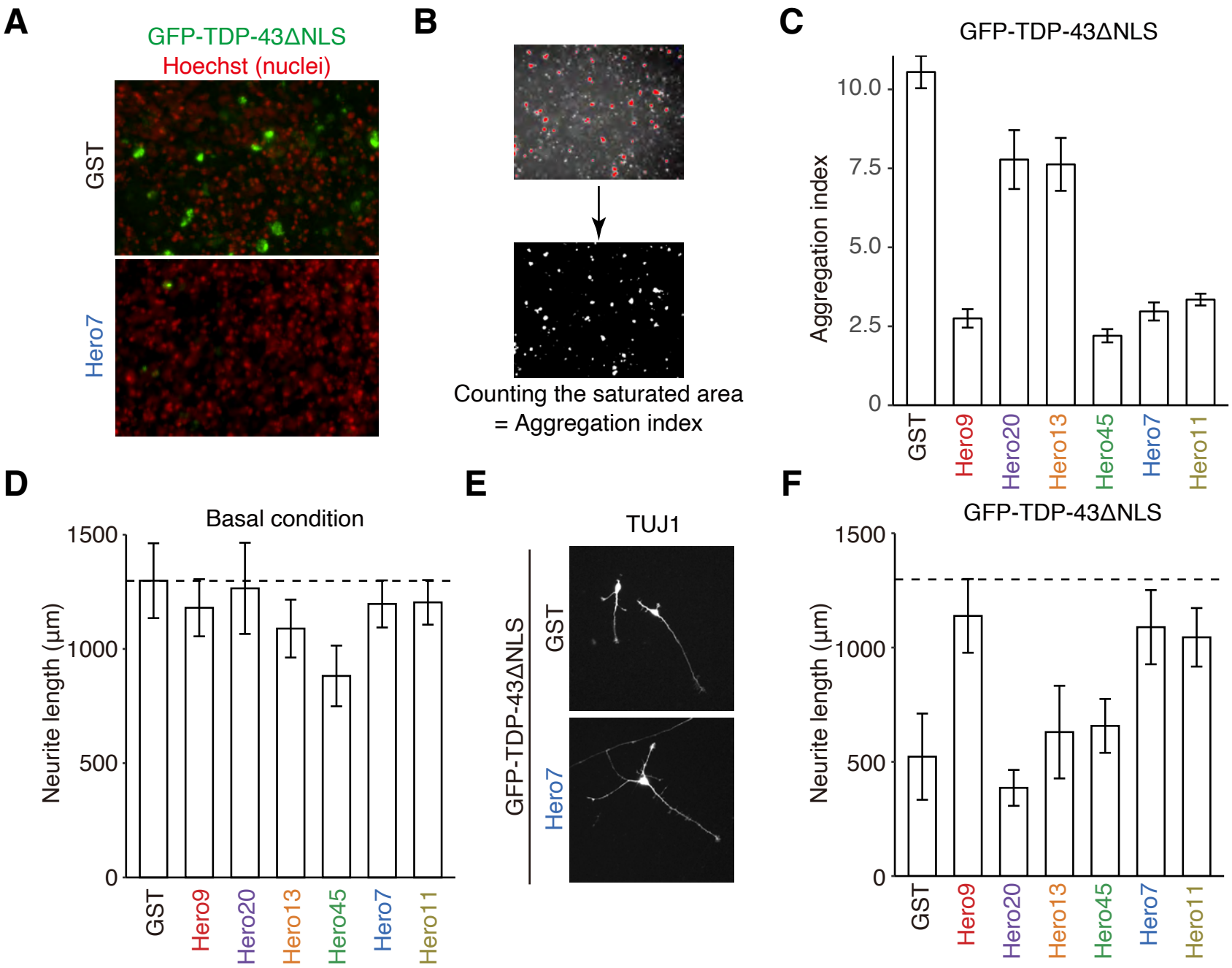
Tsuboyama *et al.* Figure 3



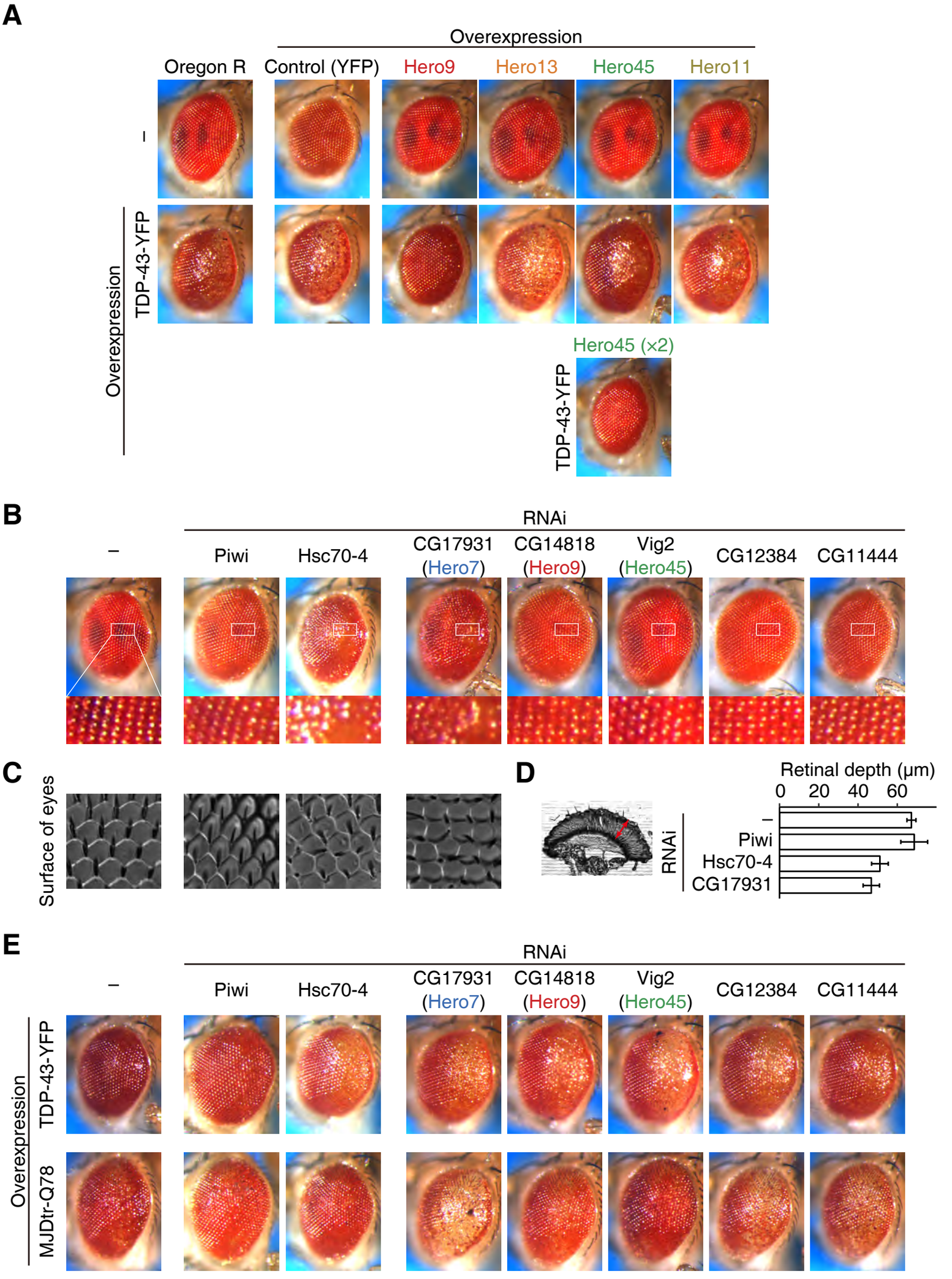
Tsuboyama *et al.* Figure 4



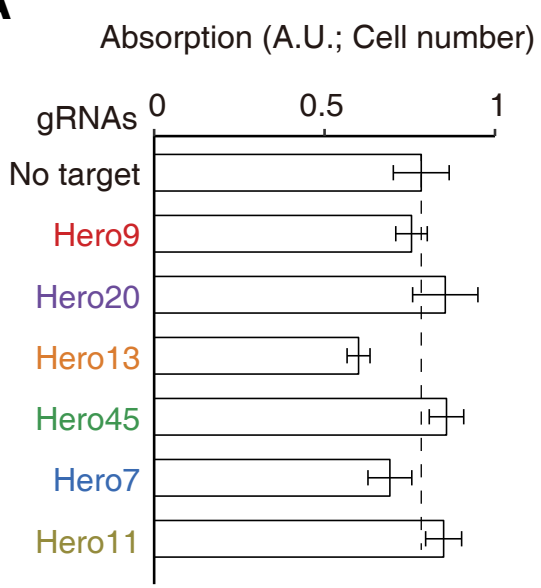
Tsuboyama *et al.* Figure 5



Tsuboyama *et al.* Figure 6



A

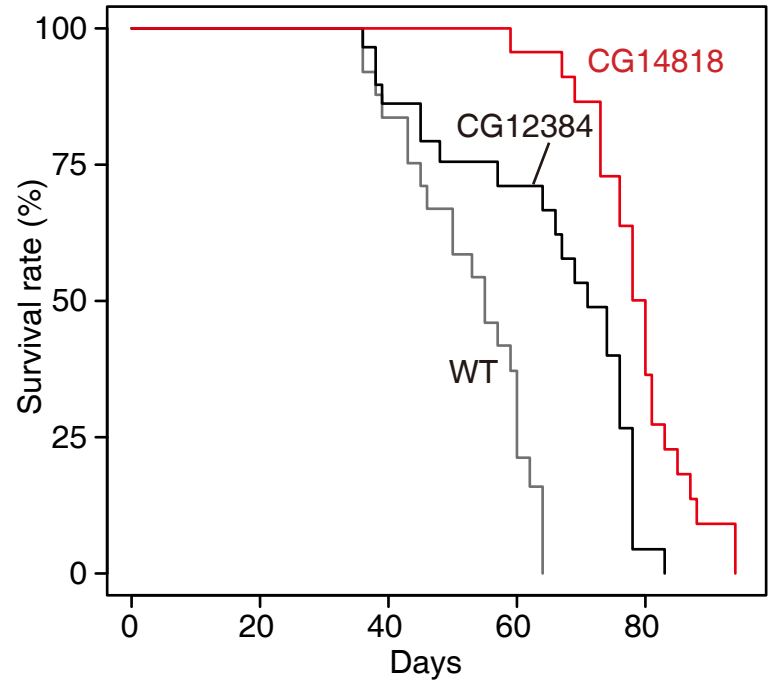


B

Summary of phenotypes in *Drosophila*

Human genes	*Drosophila* genes	GD line	KK line
Hero9	CG14818	Lethal	No phenotype
Hero20	–		
Hero13	–		
Hero45	Vig2	Weak	Lethal
Hero7	CG17931	N.A	Male-specific lethal
Hero11	–		
–	CG1910	Lethal	Lethal
–	CG3760	Lethal/Lethal	N.A

C

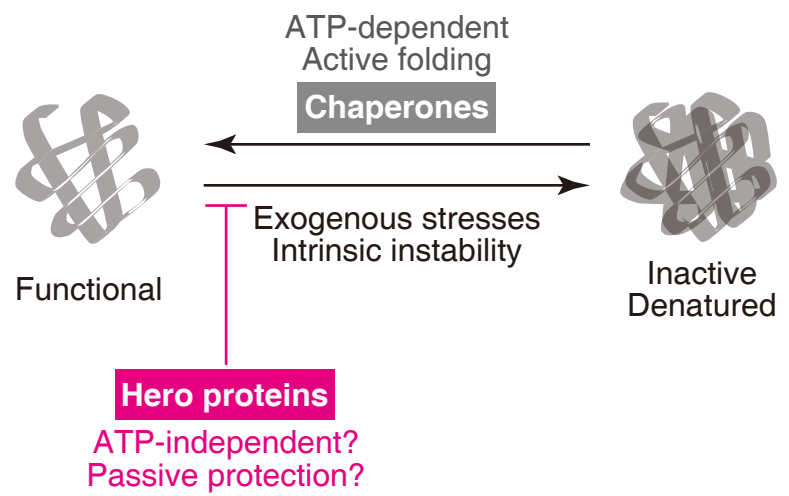


D

Summary of *Drosophila* lifespan

	Genes	Median (95%CI) days	Elongation rate (%)	P value
	WT	55 (50–60)	–	–
	YFP (control)	48 (46–52)	-12	0.09
Drosophila genes	CG14818	78 (76–81)	+42	<0.001
	CG12384	71 (66–76)	+29	<0.001
Human genes	Hero9	78 (76–80)	+42	<0.001
	Hero45	73 (66–74)	+33	<0.001
	Hero11	62 (59–66)	+13	0.02

E



Tsuboyama *et al.* Figure S1

A

Hero9 (C9orf16)

```
MSGPNGDLGM PVEAGAEGEE DGFGEAEYAA INSMLDQINS CLDHLEEKND HLHARLQELL ESNRQTRLEF
QQQLGEAPSD ASP
```

Hero20 (C11orf58)

```
MSAARESHPH GVKRSASPDD DLGSSNWEAA DLGNEERKQK FLRLMGAGKK EHTGRLVIGD HKSTSHFRTG
EEDKKINEEL ESQYQQSMDS KLSGRYRRHC GLGFSEVEDH DGEGDVAGDD DDDDDDSPDP ESPDDSESDS
ESEKEESAEE LQAAEHPDEV EDPKNKKDAK SNYKMMFVKS SGS
```

Hero13 (BEX3)

```
MANIHQENEE MEQPMQNGEE DRPLGGGEGH QPAGNRRGQA RRLAPNFRWA IPNRQINDGM GGDGDDMEIF
MEEMREIRRK LRELQLRNCL RILMGELSNH HDHHDEFCLM P
```

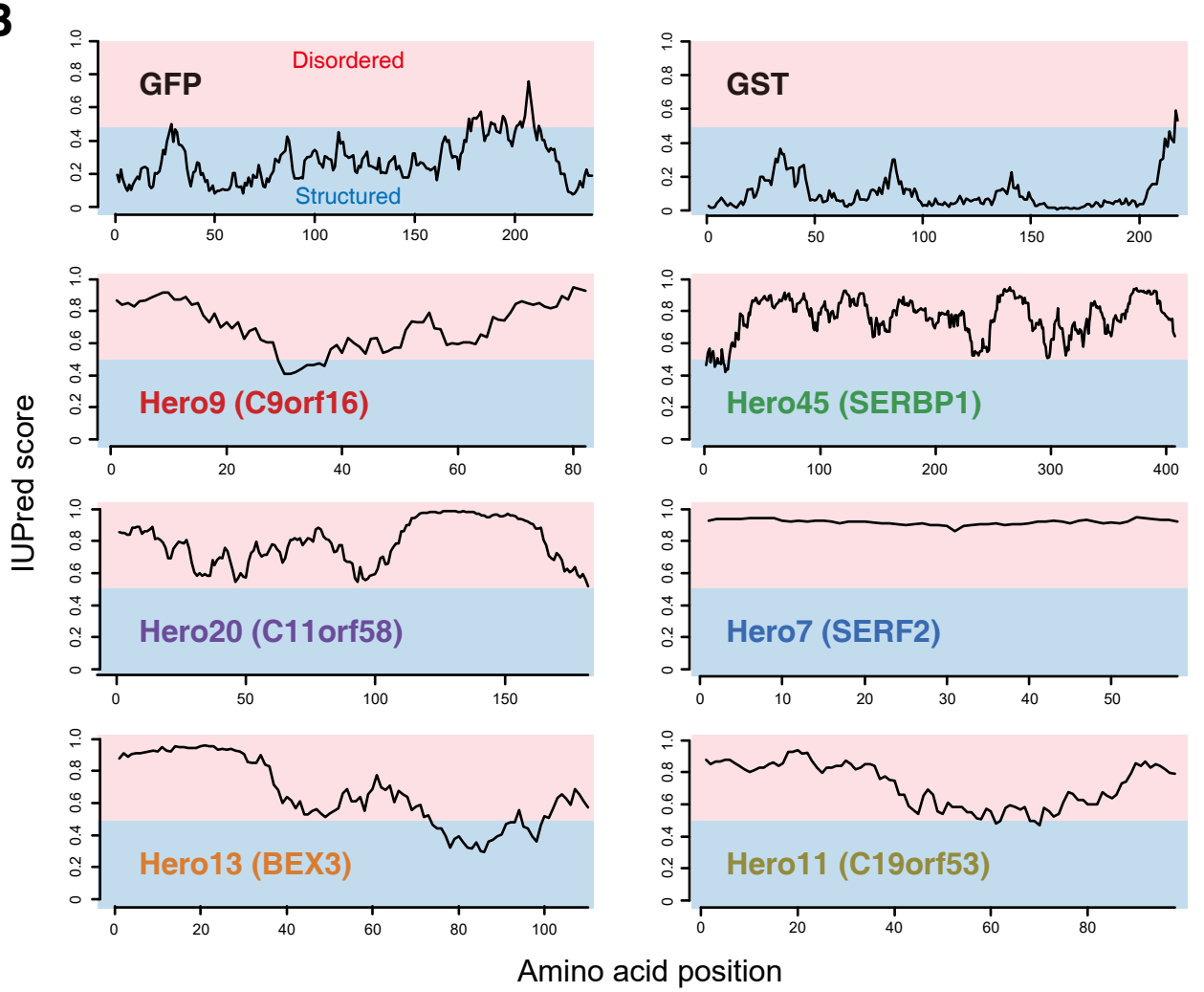
Hero45 (SERBP1)

```
MPGHLQEGFG CVVTNRFDQL FDDESDPFEV LKAAENKKKE AGGGGVGGPG AKSAAQAAAQ TNSNAAGKQL
RKESQKDRKN PLPPSVGVVD KKEETQPPVA LKKEGIRRVG RRPDQQLQGE GKIIDRRPER RPPRERRFEK
PLEEKGEGGE FSVDRPIIDR PIRGRGGLGR GRGGRGRGMG RGDGFDSRGK REFDRHSGSD RSSFSHYSGL
KHEDKRGGSG SHNWGTVKDE LTESPKYIQK QISYNYSDLD QSNVTEETPE GEEHHPVADT ENKENEVEEV
KEEGPKEMTL DEWKAIQNKD RAKVEFNIRK PNEGADGQWK KGFVLHKSKS EEAHAEDSVM DHHFRKPAND
ITSQLEINFG DLGRPGRGGR GGRGGRGRGG RPNRGSRTDK SSASAPDVDD PEAFPALA
```

Hero7 (SERF2)

```
MTRGNQRELA RQKNMKKQSD SVKGKRRDDG LSAAARKQRD SEIMQQKQKK ANEKKEEPK
```

Hero11 (C19orf53)

```
MAQGQRKFQA HKPAKSKTAA AASEKNRGPR KGGRVIAPKK ARVVQQQKLK KNLEVGIRKK IEHDVVMKAS
SSLPKKLALL KAPAKKKGAA AATSSKTPS
```

B