Advancing transcriptome analysis in models of disease and ageing

de Jong, Tristan Vincent

DOI:
10.33612/diss.99203371

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Chapter 3

Identification of an RNA polymerase III regulator linked to disease-associated protein aggregation

Olga Sin, Tristan de Jong, Alejandro Mata-Cabana, Michelle Kudron, Mohamad Amr Zaini, Francesco A. Aprile, Renée I. Seinstra, Esther Stroo, Roméo Willinge Prins, Céline Martineau, Hai Hui Wang, Wytse Hogewerf, Anne Steinhof, Erich E. Wanker, Michele Vendruscolo, Cornelis F. Calkhoven, Valerie Reinke, Victor Guryev, Ellen A. A. Nollen

Molecular Cell, 2017 65:1096–1108
SUMMARY

Protein aggregation is associated with age-related neurodegenerative disorders, such as Alzheimer’s and polyglutamine diseases. As a causal relationship between protein aggregation and neurodegeneration remains elusive, understanding the cellular mechanisms regulating protein aggregation will help develop future treatments. To identify such mechanisms, we conducted a forward genetic screen in a C. elegans model of polyglutamine aggregation and identified MOAG-2/LIR-3 as a driver of protein aggregation. In absence of polyglutamine, MOAG-2/LIR-3 regulates the RNA polymerase III-associated transcription of small non-coding RNAs. This regulation is lost in presence of polyglutamine, which mis localizes MOAG-2/LIR-3 from the nucleus to the cytosol. We show biochemically that MOAG-2/LIR-3 can catalyze the aggregation of polyglutamine-expanded huntingtin. These results suggest that polyglutamine induces the aggregation-promoting activity of MOAG-2/LIR-3 in the cytosol. The concept that aggregation-prone proteins convert endogenous proteins to drivers of aggregation and toxicity adds to the understanding of how cellular homeostasis is progressively deteriorated in age-related diseases.
INTRODUCTION

Neurodegenerative disorders such as Alzheimer's, Parkinson's or Huntington's diseases represent a major health problem and demand a better understanding of the molecular mechanisms of pathogenesis in order to make disease-modifying treatments available. The hallmark of many neurodegenerative diseases is the presence of protein aggregates in different brain areas of affected patients (Soto, 2003). These insoluble macromolecular structures are enriched in aggregation-prone proteins which – by exposing certain regions of their amino acid sequence – self-associate in an aberrant manner with other cellular proteins, thereby hampering normal cellular function ((Eisenberg and Jucker, 2012; Olzscha et al., 2011), also reviewed in (Chiti and Dobson, 2009; Sin and Nollen, 2015)). It is not yet clear, however, whether these protein aggregates are actually a cause or a consequence of the disease. The current view is that the soluble precursors of these aggregates – in particular the oligomeric forms – are the cytotoxic species, and that at least in some cases the aggregation process can represent a protective measure to sequester these smaller harmful species (Arrasate et al., 2004; Bolognesi et al., 2010; Kayed et al., 2003; Miller et al., 2011).

These observations focus the attention on the cellular factors that can drive protein aggregation, which are currently poorly understood. In some cases, direct aggregation-promoting factors have been identified, both in humans and in animal models, and include SH3GL3 (Davranche et al., 2011; Sittler et al., 1998), MOAG-4/SERF (Falsone et al., 2012; van Ham et al., 2010) and UNC-30 (Garcia et al., 2007). More studies, however, are needed to acquire a more comprehensive understanding of this phenomenon.

In this context, the roundworm C. elegans is a much used animal model for neurodegenerative diseases and is proving very useful in providing a basic understanding of protein aggregation. If we can use this model organism to identify genetic modifiers of protein aggregation, then we can also characterise the cellular
pathways that are dysregulated in the pathogenesis of human protein misfolding diseases and target them for pharmacological intervention. Indeed, a wide variety of genetic screens have already been performed in *C. elegans* to find genes that regulate protein aggregation and its associated toxicity (Hamamichi et al., 2008; Kuwahara et al., 2008; Lejeune et al., 2012; Nollen et al., 2004; Silva et al., 2011; van der Goot et al., 2012; van Ham et al., 2010). The aim of the current study is to identify genes that drive protein aggregation, to discover the function of such genes and to understand how they are related to protein aggregation. We identified the MOAG-2/LIR-3 as a modifier of aggregation that in the presence of polyglutamine shifts its role from a transcriptional regulator of small ncRNAs to an aggregation and toxicity promoting factor.

**RESULTS**

**Inactivation of MOAG-2/LIR-3 reduces polyglutamine aggregation**

To identify genes whose products are capable of driving protein aggregation, we performed a forward genetic screen in a *C. elegans* model in which the body-wall muscle cells express a transgene consisting of an aggregation-prone polyglutamine stretch of 40 residues fused to YFP (Q40-YFP) (van Ham et al., 2010). We screened for mutants with reduced polyglutamine aggregation, which we named ‘modifiers of aggregation’ (*moag*) (van Ham et al., 2010). At the fourth larval stage (L4), *moag-2(pk2183)* (hereafter designated as *moag-2/lir-3(pk2183)*) showed an about 50% reduction in the number of aggregates relative to the wild type Q40 worms (Figure 1A, B, S1A). Single nucleotide polymorphism (SNP) mapping and genome sequencing allowed us to fine-map the causal mutation and revealed six genes as putative candidates for *moag-2* (Figure S1B). For one of these, which mapped to *lir-3* (*lin-26* related; sequence: F37H8.1; accession number: NC_003280), the causative mutation was in the start codon, replacing the first methionine with an isoleucine (Met1Ile) (Figures 1C and S1B).
Figure 1: Identification of moag-2/lir-3 as an aggregation-promoting factor.
(A) Number of aggregates in Q40 worms and Q40;moag-2(pk2183) worms.
(B) Representative images of Q40 worms and Q40;moag-2(pk2183) worms. Scale bar, 75 μm. See also Figure S1A.
(C) Chromosomal location of lir-3 (F37H8.1, Chromosome II, reverse strand of assembly; http://www.wormbase.org, WS248), showing the point mutation in the start codon (red arrow), the partial deletion (red bar) and the rescue fragment (green). See also Figure S1B and S1C.
(D) Number of aggregates in Q40 worms with either wild type alleles, or heterozygous or homozygous deletion for the lir-3(tm813) allele.
(E) Number of aggregates in Q40 worms and Q40;lir-3(tm813) worms, with and without transgenic overexpression of an injected lir-3 rescue fragment including its endogenous promoter. See also Figure S1D.
(F) Filter retardation assay with five-fold serial dilutions of crude protein extract from Q40 worms and Q40;moag-2/lir-3 mutant worms. The results shown are from a representative experiment of three biological replicates. Q40-YFP and α-tubulin expression were included as controls. See also Figure S1E.
(G) Q40-YFP transcript expression detected by RNA seq in L4-stage worms and protein expression detected in urea-treated L4-stage worms and day 1 adult worms.
In all panels, aggregate counting, representative images and filter retardation assay were performed at the L4 stage and the average of three biological replicates is represented. Data are represented as mean ± SEM and significance was calculated using a one-tailed unpaired Student’s t-test. **p<0.01; ***p<0.001
The *lir*-3 gene encodes a LIN-26-like zinc finger protein of unknown function (http://wormbase.org, June 2016). It is predicted to have two zinc finger domains of the C2H2 type (residues 191-214 and 224-247) at the carboxyl terminus and a nuclear localization signal spanning amino acid residues 132 to 141 (http://nls-mapper.iab.keio.ac.jp/, June 2016) (Figure S1C). LIR-3 shares two non-canonical C2H2 zinc finger motifs with three other *C. elegans* proteins: LIN-26 (31 to 35% similarity), LIR-1 (20 to 31% similarity) and LIR-2 (25% similarity) (Dufourcq et al., 1999).

To assess whether *moag*-2 was *lir*-3, we used a *lir*-3(*tm813*) deletion mutant (hereafter designated as "*moag*-2/*lir*-3(*tm813")") and crossed it with the polyglutamine (Q40) worm model. This strain has a 795 base-pair (bp) deletion spanning residues 276 to 1070 of the F37H8.1 sequence that also causes a premature stop codon (Figures 1C and S1C). This partial deletion of *lir*-3 caused a 35% reduction of aggregates relative to the numbers seen in wild type worms (Figure 1D). Worms heterozygous for *lir*-3 deletion allele had aggregate numbers similar to those seen in the wild type, suggesting that the reduction of aggregation was recessive and due to the loss of function of *lir*-3 (Figure 1D). We next asked whether overexpression of *lir*-3 could restore the aggregation phenotype. To this end, we injected worms with a rescue fragment, consisting of full-length *lir*-3 including a 1.5 kb sequence upstream of the start codon to include its endogenous promoter as well as 330 bp downstream to include the 3' UTR (Figure 1C). Expression of the *lir*-3 rescue fragment in Q40; *lir*-3(*tm813*) worms was able to restore the aggregation phenotype by 2-fold (*p*<0.001), confirming *lir*-3 as a gene responsible for driving aggregation in the polyglutamine model (Figures 1E and S1D). Taken together, these results demonstrate that *lir*-3 is the *moag*-2 gene.

A specific property of protein aggregates in the brains of neurodegenerative disease patients that is also captured by the *C. elegans* polyglutamine model is that they are typically insoluble in strong detergents such as sodium dodecyl sulfate (SDS) (Lee et al., 1999; Shankar et al., 2008; Yanamandra et al., 2015). To establish whether MOAG-2/LIR-3 promoted aggregation of SDS-insoluble polyglutamine aggregates, we performed a filter retardation assay on lysates of wild type and *moag*-2/*lir*-3 mutant.
polyglutamine worms (Scherzinger et al., 1997; Wanker et al., 1999). This assay enables the detection of SDS-insoluble protein aggregates while, soluble species are not captured. Both Q40;moag-2/lir-3(pk2183) and Q40;moag-2/lir-3(tm813) mutants had fewer SDS-insoluble aggregates than their corresponding Q40 controls (Figure 1F). The reduction of SDS-insoluble aggregates was more pronounced in the point mutant (43%; p=0.05) than in the deletion mutant (26%; p=0.18) (Figure S1E). Mutation or partial deletion of moag-2/lir-3 did not cause a discernible reduction in the transcription or the protein expression level of Q40-YFP, indicating that moag-2/lir-3 does not reduce aggregation by reducing expression levels of the Q40-YFP protein (Figure 1F,G). Together, these results indicate that moag-2/lir-3 drives the formation of SDS-insoluble polyglutamine aggregates.

MOAG-2/LIR-3 is a C2H2-domain protein associated with RNA polymerase III promoters

Having established that mutation of moag-2/lir-3 reduces polyglutamine aggregation, we next determined the endogenous function of the protein. C2H2 zinc finger domains are predominantly associated with DNA-binding transcription factors but may also have other functions such as mediating protein-protein interactions or binding to RNA (Brown, 2005; Hall, 2005; Krishna et al., 2003). Bioinformatics analysis combined with manual curation predicted MOAG-2/LIR-3 to be a transcription factor (Reece-Hoyes et al., 2005; 2007). To explore this possibility, we determined the subcellular localization of FLAG-tagged MOAG-2/LIR-3 protein in wild type N2 worms. Indirect immunofluorescence using an anti-FLAG antibody revealed that MOAG/LIR-3 is localized in the nucleus (Figure 2A, S2A).

Next, we performed chromatin immunoprecipitation followed by deep sequencing (ChIP seq) in L4-staged worms that expressed an integrated construct of lir-3 fused to FLAG and GFP (Sarov et al., 2012; Zhong et al., 2010). This analysis yielded a total of 678 unique MOAG-2/LIR-3 binding sites, 404 of which overlapped with 813 C. elegans genes.
Figure 2. MOAG-2/LIR-3 preferentially binds to promoters of small ncRNAs.

(A) Subcellular localization of MOAG-2/LIR-3 in wild type N2 worms. Scale bar, 50 μm. Arrowheads point at nuclei; asterisk is non-specific staining. See also Figure S2A.

(B) Enrichment of binding of MOAG-2/LIR-3 to different gene biotypes (1kb upstream/downstream of TSS) relative to genes distributed genome-wide. See also Figure S2B.

(C) MOAG-2/LIR-3 binding sites within -1000 and +1000 bp of TSS for protein-coding, ncRNA, tRNA and snoRNA genes.

(D) Enriched consensus DNA motifs for MOAG-2/LIR-3 with p value.

(E) Number of MOAG-2/LIR-3 binding sites containing Box A and Box B.
Further analyses of these genes revealed that MOAG-2/LIR-3 binding was enriched in the transcription start sites (TSS) of tRNA genes (35.7%, \(p<0.001\)), snoRNA genes (6.3%, \(p<0.001\)), rRNA genes (2.5%, \(p<0.001\)) and snRNA genes (2.2%, \(p<0.002\)) (Figures 2B and S2B). While MOAG-2/LIR-3 was also found in the vicinity of protein-coding and other ncRNA genes, this binding was not significantly enriched (Figure 2B,C and Figure S2B).

We then asked whether the binding sites were enriched in any consensus sequence motif that could be recognized by MOAG-2/LIR-3. Of the 678 binding sites initially identified in our ChIP seq data, more than half of the sites contained Box A and Box B sequence motifs, 301 of which contained both motifs (Figure 2D,E). Box A and Box B constitute the canonical type 2 promoter site recognized by the RNA polymerase (Pol) III complex (Ikegami and Lieb, 2013; Schramm and Hernandez, 2002). Pol III is responsible for the transcription of structural or catalytic small nuclear RNAs (snRNAs), of tRNAs and of snoRNAs, which mediate chemical modifications of other RNA molecules (Bratkovič and Rogelj, 2014; Guthrie and Patterson, 1988; Schramm and Hernandez, 2002; White, 2011). These findings led us to hypothesize that MOAG-2/LIR-3 may bind to the same target promoters as Pol III. Binding in the proximity of non-coding genes has been shown on occasion for several \(C.\) elegans transcription factors, including PHA-4, PQM-1 and GEI-11 (Niu et al., 2011). These observations prompted us to ask whether the association between MOAG-2/LIR-3 and the Pol III complex resembled the binding of these transcription factors to the promoters of non-coding genes. To answer this question, we collected publicly available ChIP-seq data for known \(C.\) elegans transcription factors (http://www.modencode.org; September 2014) and analysed binding to the promoters of protein-coding genes, snoRNA genes and tRNA genes (Figure 2F). While there is little association between MOAG-2/LIR-3 and the promoters of protein-coding genes, the MOAG-2/LIR-3 binding profile is very similar to that of a group of clustered factors that contain representative components.
of the Pol III complex. These factors include Pol III, TATA binding protein (TBP-1), two subunits of the transcription factor for Pol III C (TFC-1 and TFC-4) and the nuclear pore proteins NPP-3 and NPP-13, which have recently been shown to associate with the Pol III complex to regulate tRNA and snoRNA splicing (Ikegami and Lieb, 2013). In contrast to the majority of the other transcription factors, the binding of these factors is very abundant in the promoters of snoRNAs and of tRNAs (Figure 2F). Therefore, the binding of MOAG-2/LIR-3 to the promoters of small ncRNA genes suggests that MOAG-2/LIR-3 is associated with Pol III transcription.

**MOAG-2/LIR-3 is a positive regulator of Pol III-mediated transcription of small ncRNAs**

Because we found MOAG-2/LIR-3 to bind to the promoters of small ncRNA genes, we next asked what consequences this binding had for the transcription of Pol III downstream targets. We therefore used transcriptome profiling to compare RNA expression in wild type worms with that in moag-2/lir-3 mutant worms. In line with the absence of MOAG-2/LIR-3 at RNA Pol II promoter sites, we did not find any protein-coding genes that were differentially expressed between the mutants and wild type N2 worms, thereby excluding MOAG-2/LIR-3 as a transcriptional regulator of protein-coding genes (Figures 3A and S3A). Mutations in moag-2/lir-3, however, did result in the downregulation of snRNAs (p<0.001), snoRNAs (p<0.001) and tRNAs (p<0.001) in both mutants, demonstrating that MOAG-2/LIR-3 regulates Pol III-mediated transcription of these small ncRNAs (Figures 3A and S3A).

We next asked where in the TSS region MOAG-2/LIR-3 was positioned relative to the Pol III complex. To this end, we compared the positions of the ChIP-seq signals of MOAG-2/LIR-3 with those of the different components of the Pol III complex. For both the tRNA and snoRNA genes, all factors localized to the Box A and Box B-containing promoter region, consistent with previous reports (Figure 3B,C, S3B) (Ikegami and Lieb, 2013). MOAG-2/LIR-3 was also positioned at these same sites (Figure 3B,C).
Figure 3. MOAG-2/LIR-3 regulates transcription of small ncRNAs.

(A) Boxplot showing the relative expression of different gene biotypes in moag-2/lir-3(pk2183) worms relative to the wild type N2 background. TPM: transcripts per kilobase million; Coding: protein-coding genes; ncRNA: non-coding RNA; Pseudo: pseudogenes; snRNA: small nuclear RNA; snlRNA: snRNA-like RNA; snoRNA: small nucleolar RNA; tRNA: transfer RNA. The average of three biological replicates is represented. ***p<0.001. See also Figure S3A.
Next, we tested whether MOAG-2/LIR-3 interacts physically with the Pol III complex by means of immunoprecipitation experiments. We found that FLAG-tagged MOAG-2/LIR-3 protein co-immunoprecipitated with Pol III, but not detectably with Pol II, confirming that MOAG-2/LIR-3 cooperates with the Pol III machinery to drive transcription of small non-coding RNAs (Figure 3D and Figure S3C). The fact that MOAG-2/LIR-3 did not detectably co-immunoprecipitate with Pol II further supports the notion that MOAG-2/LIR-3 is not directly involved in the transcription of protein-coding genes (Figure 3D).

Together these results indicate that MOAG-2/LIR-3 functions as a positive regulator of the Pol III-mediated transcription of small ncRNAs in *C. elegans*.

**Regulation of protein aggregation by MOAG-2/LIR-3 is independent of its role as a transcriptional regulator**

Next, we asked whether MOAG-2/LIR-3 regulated protein aggregation via the Pol III-mediated transcription of small ncRNAs. We therefore used RNAi to knock down, one-by-one, the individual components of the Pol III complex in both wild type worms and *moag-2/lir-3(pk2183)* mutants. To confirm RNAi knockdown, we also looked for RNAi-associated phenotypes other than aggregation. If Pol III-mediated transcription were involved in promoting protein aggregation, this would result in a reduction in the amount of aggregates in the wild type Q40 worms but not in the *moag-2/lir-3* mutants.
Figure 4: Polyglutamine proteins reduce transcription of small ncRNAs.

(A) Number of aggregates measured in Q40 and Q40;moag-2/lir-3(pk2183) worms after RNAi knockdown of individual components of the Pol III complex. As an internal quality control for RNAi, squares indicate penetrance (50% [half-open] and 0% [open]) of all associated visible RNAi phenotypes other than aggregation. See also Figure S4A and S4B.

(B) Number of aggregates measured in Q40 and Q40;moag-2/lir-3(pk2183) worms upon RNAi knockdown of tRNA processing enzymes. See also Figure S4C.

(C-D) Boxplot showing the relative expression of different gene biotypes in Q40 and Q40;moag-2/lir-3(pk2183) worms (C) and in N2 and Q40 wild type worms (Q40 wild type outcrossed from pk2183) (D). TPM: transcripts per kilobase million; Coding: protein-coding genes; ncRNA: non-coding RNA; Pseudo: pseudogenes; snRNA: small nuclear RNA; snlRNA: snRNA-like RNA; snoRNA: small nucleolar RNA; tRNA: transfer RNA. See also Figure S4D, E).

In panels A and B, aggregate counting was performed at the L4 stage and the average of three biological replicates is represented. Data are represented as mean ± SEM. In panels C and D, the average of three biological replicates is represented. ***p<0.001; ns is not significant. See also Figure S4D and S4E.
However, knockdown of Pol III, TBP-1 or TFC-1 did not alter aggregation in the Q40 worms, nor in the Q40;moag-2/lir-3 mutant strains, indicating that – in the absence of moag-2/lir-3 – it is not the resulting lack of Pol III-mediated transcription that is responsible for the reduction in aggregation (Figures 4A and S4A,B).

The nuclear pore protein npp-13 has been shown to interact with the Pol III complex and regulate the processing of small non-coding RNAs (Ikegami and Lieb, 2013). The aggregation phenotype was not altered by knock down of npp-13 or of tRNA-processing enzymes, indicating that neither small RNA processing nor the availability of mature tRNAs is involved in reducing aggregation (Figures 4A,B and S4A-C). These results indicate that the regulation of protein aggregation by MOAG-2/LIR-3 is separate from its involvement in RNA transcription together with the Pol III complex.

**Polyglutamine suppresses transcription of small ncRNAs**

We next addressed why the MOAG-2/LIR-3’s involvement in driving protein aggregation was independent of its role as a transcriptional regulator. We therefore first compared the RNA expression profiles of wild type Q40 worms with those of moag-2/lir-3 mutant Q40 worms. In contrast to wild type N2 worms, there was no longer a reduction in the relative expression levels of the small ncRNAs in the presence of polyglutamine (Figures 4C and S4D). When we then measured the absolute levels of all RNAs, we found that – in contrast to the protein-coding RNAs, pseudogenes and other ncRNAs – the expression of snRNAs, snoRNAs and tRNAs was already strongly reduced in wild type Q40 worms (p<0.001; Figure 4D, S4E). This finding explained why mutations in moag-2/lir-3 could no longer reduce expression in these Q40 worms (Figures 4C and S4D) and indicated that presence of polyglutamine results in downregulation of small ncRNA expression.

**MOAG-2/LIR-3 increases polyglutamine toxicity**

Having identified MOAG-2/LIR-3 as a regulator polyglutamine aggregation we asked whether it regulates the toxicity of polyglutamine as well. For this we compared the motilities of animals with or without overexpression of FLAG-tagged MOAG-2/LIR-3
in absence or presence of polyglutamine over course of 12 days. In absence of polyglutamine, FLAG-tagged MOAG-2/LIR-3 had no effect on the motility in young adults but improved the motility of old adult animals (Figure 5A). In contrast, in worms expressing polyglutamine, overexpression of MOAG-2/LIR-3 accelerated the age-dependent decline in motility (Figure 5B). To note, a deletion in moag-2/lir-3 had no significant effect on the motility of polyglutamine worms up to 13 days of adulthood (Figure S5A). Our results with the MOAG-2/LIR-3 overexpressing worms imply that the effect of a deletion may not have been strong enough to be detected in our assays, because they indicate that the increasing the levels of MOAG-2/LIR-3 increases the toxicity of polyglutamine.

**MOAG-2/LIR-3 can directly drive polyglutamine aggregation**

Our results indicate that MOAG-2/LIR-3 regulates protein aggregation independently of its role in small ncRNA transcription and processing. A possible mechanism by which loss of MOAG-2/LIR-3 could suppress aggregation indirectly could be by increasing the cellular folding capacity. The transcription of protein-coding genes was not affected, however, indicating that increased transcription of folding genes was not involved (Table S3). Alternatively, MOAG-2/LIR-3 may normally sequester folding factors, which then may be released in the Q40;moag-2/lir-3 mutants, a mechanism that has been previously proposed for an increase in protein aggregation by temperature sensitive mutant proteins (Gidalevitz et al., 2006). If such a mechanism would be responsible for the suppression of aggregation in the Q40;moag-2/lir-3 mutants, knockdown of folding factors in these animals would increase aggregation again. To test this possibility, we knocked down by RNAi several folding factors, which have been previously shown to regulate polyglutamine aggregation (Hsu et al., 2003; Nillegoda et al., 2015; Nollen et al., 2004; Rampelt et al., 2012). The depletion of these factors could not revert the aggregation phenotype in the Q40;moag-2/lir-3 mutants (Figure S5B). These results suggest that suppression of protein aggregation by inactivation of MOAG-2/LIR-3 is not caused by an increase in the expression nor a release of folding factors.
Another possibility is that MOAG-2/LIR-3 drives aggregation directly. Precedents for aggregation-promoting factors include MOAG-4/SERF, which can drive aggregation of a variety of disease proteins via a transient direct interaction with early aggregation intermediates (Davranche et al., 2011; Falsone et al., 2012; Sittler et al., 1998; van Ham et al., 2010).

To determine whether MOAG-2/LIR-3 can drive aggregation directly, we incubated GST-tagged huntingtin exon 1 containing 48 CAG repeats (HttQ48) with purified MOAG-2/LIR-3 protein and monitored amyloid formation by measuring thioflavin T (ThT) fluorescence in vitro. Our kinetic data revealed that the presence of an equimolar concentration of MOAG-2/LIR-3 accelerated the aggregation of HttQ48, but not of the non-pathogenic form HttQ23, similarly to what observed in the case of MOAG-4, the C. elegans ortholog of SERF1A (Figure 5C). These results indicate that MOAG-2/LIR-3 can catalyse amyloid formation of aggregation-prone proteins directly.

The presence of polyglutamine mislocalizes MOAG-2/LIR-3 from the nucleus to the cytosol

For lack of an aggregation-catalysing effect to explain the reduction of protein aggregation in Q40;moag-2/lir-3 mutants, MOAG-2/LIR-3 and polyglutamine would, in the wild type animals, have to localize in the same cellular compartment. Thus, to determine their localization, we performed subcellular fractionations with wild type and Q40 worms with or without the expression of FLAG-tagged MOAG-2/LIR-3. We found that in absence of polyglutamine, MOAG-2/LIR-3::FLAG was primarily located in the nucleus (Figure 5D, OE, Figure S5C). Polyglutamine was mostly localised in the cytosol, which was independent of the presence of MOAG-2/LIR-3::FLAG (Figure 5D, Q40 and Q40;OE, Figure S5C). In Q40 worms, however, MOAG-2/LIR-3::FLAG was enriched in the cytosol (Figure 5D, OE and Q40;OE, Figure S5C), which suggests that the presence of polyglutamine altered the localization of MOAG-2/LIR-3. Their joint presence in the cytosol supports a direct interaction between the proteins to drive protein aggregation.
Figure 5. Polyglutamine moves MOAG-2/LIR-3 cytosol where it turns into a positive regulator of protein aggregation.

(A-B) Number of body bends in N2 wild type worms and worms overexpressing MOAG-2/LIR-3 (OE) (A) and Q40 wild type and Q40 worms overexpressing MOAG-2/LIR-3 (Q40;OE) (B). The average of three biological replicates is represented. Data are represented as mean ± SEM and significance was calculated using a two-tailed unpaired Student’s t-test. **p<0.01; ***p<0.001 See also Figure S5A.
(C) Thioflavin T reaction profiles of 4 μM HttQ48 or HttQ23 solutions in the absence or presence of equimolar concentrations of MOAG-2/LIR-3 (black) or the modifier of aggregation MOAG-4 (blue). The concentration of MOAG-2/LIR-3 in whole organism lysates has been determined to be between 0.3 and 8.5 ppm (http://pax-db.org). The average of four replicates is represented and error bars indicate mean ± SD.

(D) Western blot of subcellular fractionation of MOAG-3/LIR-3, which was detected using α-FLAG antibody. Q40-YFP, LMN-1 (nuclear marker) and α-tubulin (cytosolic marker) were used as controls. See Figure S5C.

(E) Western blot analysis of MOAG-2/LIR-3 SDS-solubility from wild type (WT) and MOAG-2/LIR-3-overexpressing (OE) worms in N2 and polyglutamine worms (Q40). Q40-YFP and β-actin expression were included as controls. In panels B and C, the results were generated from L4-staged animals and a representative experiment of three biological replicates is shown.

(F) Model proposing how MOAG-2/LIR-3 drives polyglutamine aggregation. MOAG-2/LIR-3 normally regulates the transcription of small ncRNAs in the nucleus. In the presence of polyglutamine, MOAG-2/LIR-3 is moved to the cytosol where it exposes an ectopic aggregation-promoting activity.

Proteins involved in RNA metabolism have been described to co-aggregate with polyglutamine (Doi et al., 2010; 2008; Schwäb et al., 2008), which could explain the retention of MOAG-2/LIR-3 in the cytosol. We therefore fractionated SDS-soluble and insoluble material to determine whether MOAG-2/LIR-3 co-aggregated with polyglutamine. Q40-YFP was detected as low molecular weight monomers and as SDS-resistant species retained in the stacking gel (Figure 5D). The solubility of MOAG-2/LIR-3::FLAG was not altered in the presence of polyglutamine (Figure 5D), suggesting that MOAG-2/LIR-3 is not recruited to SDS-resistant aggregates, but may be associated with the soluble species.

Together, these data suggest that the presence of polyglutamine moves MOAG-2/LIR-3 from the nucleus to the cytosol where it turns into a positive regulator of protein aggregation, which is at the expense of its function as a transcriptional regulator (Figure 5F).
DISCUSSION

We have used a genetic screen in a C. elegans model of protein aggregation disease to identify MOAG-2/LIR-3 as a regulator of Pol III transcription in the nucleus that – in the presence of polyglutamine – switches into a positive regulator of polyglutamine aggregation in the cytosol.

Our finding that MOAG-2/LIR-3 is a regulator of transcription confirms predictions about to the LIR-3 gene. Its role as a transcriptional regulator has been suggested on basis of its structural similarity to the C2H2 zinc fingers of LIN-26, a fate regulator responsible for the differentiation of non-neuronal ectodermal cells and somatic gonad epithelium (Boer et al., 1998; Dufourcq et al., 1999; Labouesse et al., 1996). We showed that MOAG-2/LIR-3 is required for the transcription of snRNA, snoRNA and tRNA genes and – unlike other transcription factors that have on occasion been found to bind to the promoters of small ncRNAs – that MOAG-2/LIR-3 was associated with the same target genes as the Pol III complex (Niu et al., 2011).

MOAG-2/LIR-3 is known to be expressed in the nuclei of a subset of cell types of C. elegans, namely body wall muscle cells, the vulval muscles, the spermatheca, the head and tail ganglia, and the ventral nerve cord, from embryogenesis throughout adulthood (www.transgeneome.mpi-cbg.de, March 2014 and (Reece-Hoyes et al., 2005; 2007)). A microarray performed in two distinct mechanosensory neurons – the touch receptor neurons and the FLP neurons – revealed that moag-2/lir-3 is upregulated in the FLP sensory neurons, suggesting that moag-2/lir-3 is required for FLP differentiation (Topalidou and Chalfie, 2011). Because the expression of MOAG-2/LIR-3 is limited to a subset of cell types, this suggests that MOAG-2/LIR-3 could be a tissue-specific regulator of transcription rather than a core component of the Pol III machinery. Such cell type-specific regulators of Pol III have been described for human cells and proposed to accommodate cell-specific needs for small non-coding RNAs ((Alla and Cairns, 2014; Oler et al., 2010), also reviewed in (Marshall and White, 2008; White, 2011)). Whether MOAG-2/LIR-3 has a similar role remains to be established,
but it could explain why mutations in the *moag-2/lir-3* gene do not result in any obvious abnormalities in terms of growth or viability.

In *C. elegans*, the nuclear pore protein NPP-13 has been described to associate with the Pol III complex, which regulates the efficient processing of snoRNA and tRNA transcripts (Ikegami and Lieb, 2013). Knockdown of NPP-13 results in abnormally long snoRNA and tRNA transcripts that cannot be processed into their mature form (Ikegami and Lieb, 2013). In *moag-2/lir-3* mutant worms we did not find unprocessed transcripts for snoRNA nor for tRNAs (data not shown), which excludes the possibility that MOAG-2/LIR-3 is required for Pol III transcript processing. Moreover, knockdown of NPP-13 did not alter aggregation, which suggests that mutations in *moag-2/lir-3* do not alter aggregation by interfering with the nuclear pore complex.

Polyglutamine proteins can undergo post-translational modifications, including acetylation, phosphorylation, ubiquitination and sumoylation (reviewed in (Ehrnhoefer et al., 2011; Pennuto et al., 2009)). These post-translational modifications have been shown to modulate polyglutamine toxicity and aggregation (Gu et al., 2009; Jana et al., 2005; Matsumoto et al., 2004; Steffan et al., 2004; Thomas et al., 2004). One possibility that could explain the reduction of aggregation observed in the polyglutamine model is that mutations in *moag-2/lir-3* alter the post-translational status of polyglutamine. Although with our present results we cannot exclude this possibility, our in vitro data show that MOAG-2/LIR-3 is able to drive polyglutamine aggregation directly in conditions where post-translational modifications do not occur, supporting a functional interaction between MOAG-2/LIR-3 and polyglutamine.

In this study, expression of polyglutamine downregulated the levels of snRNAs, snoRNAs and tRNAs, demonstrating that aggregation-prone proteins can affect small non-coding RNA homeostasis. A possible explanation for this downregulation is that the aggregation-prone proteins blocked the nuclear localization of MOAG-2/LIR-3 and perhaps also components of the Pol III complex. Indeed, our data suggest that polyglutamine retains MOAG-2/LIR-3 in the cytosol, which is consistent with the
recent finding that aggregation-prone proteins block trafficking in and out of the nucleus (Woerner et al., 2016). Several intrinsically disordered proteins – which include proteins involved in transcriptional regulation – are known to be sequestered by aggregation-prone proteins (Iakoucheva et al., 2002; Liu et al., 2006; Minezaki et al., 2006; Olzscha et al., 2011; Walther et al., 2015). One example is Sp1, which can no longer bind to its DNA targets due to sequestration by mutant huntingtin (Dunah et al., 2002; Li et al., 2002). We found that MOAG-2/LIR-3 is not sequestered in the polyglutamine aggregates. The concentration of MOAG-2/LIR-3 in whole worm lysates has been determined to be between 0.3 and 8.5 ppm (http://pax-db.org). Although these concentrations do not take into account cell- or stage specific differences in protein expression, its relatively low abundance would support a model in which MOAG-2/LIR-3 drives aggregation via transient interactions with rare SDS soluble polyglutamine species.

In summary, this work has revealed that protein aggregation can affect the non-coding genome by altering the expression of ncRNA genes. This work also opens up another potential perspective on how aggregation-prone proteins can impair cellular homeostasis – these proteins can convert from their normal functions into aggregation-promoting factors by relocation. These results imply a cellular mechanism whereby aggregation-prone disease proteins inactivate, recruit and use endogenous proteins that promote their own aggregation and toxicity, which would resemble viral self-catalysis of pathogenesis. We anticipate that interfering with this class of gene products that promote aggregation-prone interactions can be explored for the development of therapies to slow down progression of age-related protein aggregation diseases.
ACKNOWLEDGEMENTS

We thank the Caenorhabditis Genetics Centre (funded by the National Institutes of Health National Centre for Research Resources) and the Mitani laboratory for the C. elegans strains (funded by the Japan National BioResource Project). We thank Pieter van der Vlies and Jelkje Bergsma for practical assistance with whole-genome sequencing. We thank Pieter Neerincx for help with installing MAQGene. We thank Marianna Bevova and Diana Spierings for assistance with RNA sequencing. We thank Kohta Ikegami for the npp3 RNAi vector and for the protocols on RNA sequencing and co-immunoprecipitation. We thank the modERN project (NIH U41HG007355) for sharing the MOAG-2/LIR-3 ChIP seq data prior to publication. We thank Klaas Sjollema for help with C. elegans imaging (UMIC, UMCG). We thank Sally Hill for editing this manuscript. This study was supported by funding from the Berlin Institute of Health Collaborative Research Grant no. 1.1.2.a.3 funded by the German Federal Ministry for Education and Research (BMBF) (to E.E.W). This project was funded by a European Research Council (ERC) starting grant (to E.A.A.N.), the Fundação para a Ciência e Tecnologia fellowship (SFRH/BPD/51009/2010) (to O.S.), the Alumni chapter Gooische Groningers facilitated by the Ubbo Emmius Fonds and a Groningen University Institute for Drug Exploration (GUIDE) fellowship (O.S.).

REFERENCES


McCarthy DJ, Chen Y and Smyth GK (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Research 40, 4288-4297


Chapter 3


Chapter 3


CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ellen A. A. Nollen (e.a.a.nollen@umcg.nl). Published research materials and reagents from the ERIBA are shared with the academic community under a Material Transfer Agreement (MTA).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

C. elegans

The strains used in this study are listed in Table S1. Standard methods were used for culturing C. elegans at 20°C (Brenner, 1974). The LIR-3 overexpression strain (OP312)
was generated by biolistic transformation to produce an integrated, low-copy transgene of the WRM0637aB05 fosmid, recombined with GFP::3xFLAG in frame at the carboxy terminus of the lir-3 locus (Sarov et al., 2012). To synchronize animals, eggs were collected from gravid hermaphrodites by hypochlorite bleaching and hatched overnight in M9 buffer. The desired numbers of L1 animals were subsequently cultured on nematode growth medium (NGM) agar plates seeded with OP50 bacteria.

**METHODS**

**EMS mutagenesis and mapping**

Mutagenesis was performed using standard *C. elegans* ethyl methanesulfonate (EMS) methodology (Jorgensen and Mango, 2002). 8000 mutagenized genomes were screened for suppressors of aggregation. *moag-2(pk2183)* was identified by single-nucleotide polymorphism mapping to a region between base 9,400,743 and 11,827,697 on linkage group II (Wicks et al., 2001). Next generation sequencing was performed in that region to identify candidate genes for *moag-2*. CLC Bio and MAQGene software were utilized for mapping the mutation in F37H8.1 (listed in Key Resources Table).

**RNA interference**

RNAi experiments were performed on NGM agar plates containing 1 mM IPTG and 50 mg/ml ampicillin and seeded with RNAi bacteria induced with IPTG to produce dsRNA. Worms were synchronized by hypochlorite bleaching; L1 worms were grown on RNAi plates and used for the experiments at L4 stage, unless stated otherwise. Plates were coded so that the experimenter was blind to the genotype of the animal.

**Motility assay**

At day 1, 4, 8 and 12 of adulthood, animals were placed in a drop of M9 and were allowed to recover for 30 seconds after which the number of body bends was counted for 30 seconds. Fifteen animals were counted per experiment and the data from three biological replicates was combined. Plates were coded so that the experimenter was blind to the genotype of the animal. The experiments with the N2 worms and the Q40
animals were performed independently due to growth differences between the strains. In our laboratory conditions, N2 worms take approximately 72 hours to reach L4 stage after hypochlorite treatment, whereas Q40 take approximately 96 hours.

**Generation of transgenic strains**

For the rescue experiment, a genomic construct of *lir*-3 spanning 1500 bp upstream to 330 bp downstream of F37H8.1 was amplified from N2 genomic DNA by nested PCR using primers F1; R1; F2 and R2 (primer sequences listed in the Key Resources Table). The resulting PCR fragment was cloned into the pGEM-T Easy Vector (#A1380, Promega) and sequenced. Transgenic lines were made by injecting ~20 ng/µl of construct along with ~10 ng/µl of pPD136.61 [P(*unc*-54::CFP)] into N2 animals. Two independent lines were obtained for each transgene of interest.

**Quantitative PCR**

Total RNA was extracted from synchronized populations in L4 stage using Trizol (#15596-018, Life Technologies) according to the manufacturer’s description. Total RNA quality and concentration were assessed using a NanoDrop 2000 spectrophotometer (Thermo Scientific). cDNA was made from 2 µg total *C. elegans* RNA with a RevertAid H Minus First Strand cDNA Synthesis kit (#K1632, Life Technologies) using random hexamer primers. Quantitative real-time PCR was performed using a Roche LightCycler 480 Instrument II (Roche Diagnostics) with SYBR green dye (#172-5125, Bio-Rad) to detect DNA amplification. The following cycle conditions were used: 50 °C for 10 minutes, 95 °C for 10 minutes, followed by 40 cycles at 95 °C for 15 seconds and 57 °C for 30 seconds. Relative transcript levels were quantitated using a standard curve of pooled cDNA solutions. Expression levels were normalized against the endogenous reference gene *cdc*-42. The following primers were used: *cdc*-42_F3; *cdc*-42_R3; *lir*-3_F4 and *lir*-3_R4 (primer sequences listed in the Key Resources Table).

To measure the efficiency of Pol III knockdown by RNAi, total RNA from 20 L4 animals was extracted using Trizol (Life Technologies). As a quality control step, total RNA
was measured using Qubit RNA HS Assay Kit (#Q32852, Thermo Fisher). Following DNAse I digestion (#EN0521, Fermentas), total RNA was purified with Trizol (Life Technologies). First strand synthesis was carried out using Superscript II (Invitrogen) using random hexamer primers. Expression levels were normalized to the expression of \textit{rps-21} following the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). The primers used were rpc-1\_F1; rpc1\_R1; rps-21\_F1 and rps-21\_R1 (primer sequences listed in the Key Resources Table).

**Chromatin immunoprecipitation sequencing**

ChIP assays were conducted as previously described (Niu et al., 2011; Zhong et al., 2010). Worm staging was achieved by bleaching and L1 starvation. Arrested L1 worms were plated on peptone-enriched NGM plates seeded with OP50 bacteria and grown for 48 hours for L4 collection at 20°C. Samples were cross-linked with 2\% formaldehyde for 30 minutes at room temperature and then quenched with 1 M Tris pH 7.5. The pelleted worms were subsequently flash frozen in liquid nitrogen and stored at -80°C. Samples were sonicated using a microtip to obtain DNA fragments mostly 200 to 800 bp in length. For each sample, 2.2 or 4.4 mg of cell extract was immunoprecipitated using a goat α-GFP antibody, GoatV (gift from Kevin White).

For library preparation and sequencing we used the enriched DNA fragments and input control (genomic DNA from the same sample) for two biological replicates as previously described (Kasper et al., 2014). Briefly, samples were multiplexed using the Ovation Ultralow DR Multiplex Systems 1-8 and 9-16 (NuGEN Technologies) following the manufacturer’s protocol, with the exception that Qiagen MinElute PCR purification kits were used to isolate the DNA. Library size selection in the 200-800 bp range was achieved using the SPRIselect reagent kit (Beckman Coulter) and sequencing was performed on the Illumina HiSeq 2000 platform. To search for MOAG-2/LIR-3-specific binding sites, we only considered binding sites consistent across both replicates and within a range of -400 bp to +100 bp relative to the TSS. In addition, to avoid false positives in our analysis we excluded highly occupied target regions (Gerstein et al., 2010). The ChIP seq data used was obtained from
modENCODE DCC (http://www.modencode.org; September 2014). All ChIP seq data were deposited in the ENCODE database, under ID code ENCSR408FDZ and in ArrayExpress, under accession number E-MTAB-4174.

**RNA sequencing**

Worms were grown to L4 stage and total RNA was extracted using Trizol (Life Technologies) according to the manufacturer’s description. For polyA RNA sequencing, the TruSeq Sample Preparation V2 Kit was used (Illumina). For RNA sequencing (mRNA, snoRNA, tRNA and large ncRNA), a protocol has been described previously (Ikegami and Lieb, 2013). Briefly, total RNA was treated with DNaseI and depleted from rRNA with the Ribominus Eukaryote Kit (#A10837-08, Invitrogen). Fragmentation of RNA was performed using Fragmentation Buffer (#AM8740, Ambion) and cDNA was generated using the Superscript II Kit (Invitrogen). cDNA libraries were subjected to high-throughput single-end sequencing (50 bp) in an Illumina HiSeq 2500 instrument. RNA sequencing data was mapped to the WS220 genome reference using the TopHat 2.0.9 program (Kim et al., 2013) and gene annotation was derived from Ensembl release 66. All RNA sequencing has been submitted to ArrayExpress, under accession number E-MTAB-4172.

**Immunofluorescence**

Indirect immunofluorescence was performed using the “freeze-cracking” method (Duerr, 2006). Briefly, synchronized worms at L4 stage were fixed in 4% paraformaldehyde in 100 mM sodium phosphate buffer for 1 hour. Samples were washed three times with TBS-T for 10 minutes and blocked in 10% goat serum in antibody buffer (PBS, 0.5% Triton X-100, 1mM EDTA, 0.1% BSA, 0.05% sodium azide, pH 7.2) for 1 hour. Anti-FLAG primary antibody (#F3165, Sigma-Aldrich) was applied at a 1:100 dilution overnight at 4°C. The next day, samples were washed three times in antibody buffer for 20 minutes and incubated with the secondary antibody Cy5-labeled goat anti-mouse IgG (#A10524, Thermo Fisher) at a 1:500 dilution for 4 hours at room temperature. After washing with antibody buffer three times, samples were
mounted in Vectashield mounting medium with DAPI (#H-1200, Vector Labs) and analysed on a TCS SP8 laser scanning confocal microscope with a 40x/1.30 NA objective lens (Leica).

**Filter retardation assay**

The protocol was adapted from Wanker et al (Scherzinger et al., 1997; Wanker et al., 1999). Briefly, crude worm lysates from synchronized L4 animals were resuspended in FTA sample buffer (10 mM Tris-Cl pH 8.0, 150 mM NaCl, 2% SDS) with protease inhibitors (Complete, Roche) and disrupted using a bead beater (FastPrep 24, MP Biomedicals) for 7 cycles of 20 seconds with 5 minutes rest in between cycles. Supernatants were transferred to new 1.5 ml tubes, and protein concentration was determined. To detect SDS-insoluble aggregates, 100 µg of total protein was mixed with 1 M DTT and FTA sample buffer (final concentration 40 µg/100 µl) and heated at 98 ºC for 5 minutes. Samples were filtered through a 0.22 micron cellulose acetate membrane using a 48-well Bio-Dot microfiltration system (Bio-Rad) and 100 µg of total protein was used for the assay plus two five-fold serial dilutions. Proteins were blocked for 30 minutes with 5% milk in TBS-T. Membranes were incubated with the primary antibodies α-GFP (#632381, Clontech Laboratories) or α-tubulin (#T6074, Sigma-Aldrich) at a 1:5000 dilution overnight at 4 ºC. Incubation with the secondary α-mouse antibody was at a 1:10 000 dilution for 1 hour at room temperature. Antibody binding was visualized with an ECL kit (#RPN2232, Amersham).

**Co-immunoprecipitation**

The protocol was adapted from (Ikegami et al., 2010). Briefly, crude worm lysates from synchronized L4 animals were resuspended in FA buffer (50 mM HEPES/KOH pH 7.5, 1 mM EDTA, 1% Triton X-100, 0.1 % sodium deoxycholate, 150 mM NaCl) supplemented with protease inhibitors (Complete, Roche) and 1% sodium lauroyl sarcosinate (sarkosyl). Samples were cross-linked with formaldehyde for 30 minutes and 2.5 M glycine for 5 minutes. Worms were disrupted using a bead beater (FastPrep 24, MP Biomedicals) for 7 cycles of 20 seconds with 5 minutes rest in between cycles.
Supernatants were transferred to new 1.5 ml tubes and protein concentration was
determined. Lysates were pre-cleared for 30 minutes with Protein G-Sepharose beads
(Amersham) and incubated with primary antibodies α-Pol II (#38520002, Novus
Biologicals) or α-Pol III (#53330002, Novus Biologicals) at a dilution of 1:100 overnight
at 4 °C. Lysates were coupled to Protein G-Sepharose beads (Amersham) for 1 hour
and washed five times for 5 minutes with FA buffer supplemented with 1 mM PMSF
and protease inhibitor cocktail (Complete, Roche). Proteins were eluted by boiling the
beads in SDS sample buffer and analysed by western blotting.

**Protein insolubility assay**

Synchronized worms at L4 stage animals were resuspended in FA buffer and disrupted
using a bead beater (FastPrep 24, MP Biomedicals) for 7 cycles of 20 seconds with 5
minutes rest in between cycles, followed by 15 sonication steps of 30 seconds with 30
seconds rest between cycles. Protein concentration was determined and equal
amounts of lysates were incubated with 2% SDS for 1 hour at room temperature. The
lysates were centrifuged at 13 300 rpm for 30 minutes and the supernatant (soluble
fraction) was collected. The pellet (insoluble fraction) was washed two times with FA
buffer and resuspended in urea buffer (8M urea, 2% SDS, 50 mM DTT, 50 mM Tris pH
8) for 1 hour. The samples were subsequently analyzed by western blotting.

**Protein purification**

Recombinant moag-2/lir-3 was expressed and purified fused to the glutathione S-
transferase (GST) from the pGEX-6-P1 vector in E. coli BL21 (DE) gold strain
(Stratagene). Cells were grown in Overnight Express Instant TB Medium (Merck
Millipore) supplemented with ampicillin (100 μg/ml) overnight at 30 °C under
constant shaking at 250 rpm. The cells were harvested by centrifugation, resuspended
in lysis buffer (50 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA and EDTA-Free protease
inhibitor cocktail (Complete, Roche) and lysed by sonication. The cell debris was
removed by centrifugation at 18 000 rpm (JA-25.50 rotor, Beckman Coulter). The
supernatant was loaded onto a column containing Glutathione Sepharose 4 Fast Flow
resin (GE Healthcare LifeSciences) and equilibrated with lysis buffer. The column was washed with 40 CV of lysis buffer and the protein was eluted in 50 mM Tris pH 8, 10 mM reduced glutathione and dialyzed in 50 mM Tris pH 7.4, 150 mM NaCl for the aggregation experiments.

**Aggregation kinetics**

4 μM solutions of GST-HttQ48 or GST-HttQ23 alone or with an equimolar concentration of moag-2/lir-3 or moag-4 were incubated in the presence of 7U of PreScission protease (GE Healthcare LifeSciences) per nmol of Htt protein in 50 mM Tris pH 7.4, 150 mM NaCl, 20 μM Thioflavin T at 37 °C. The aggregation reactions were performed under constant linear shaking at 500 rpm and the ThT fluorescence was monitored in low-binding, clear-bottomed half-area 96-well plates. Emissions at 480 nm were recorded every 300 seconds with excitation at 440 nm, using a CLARIOstar plate reader (BMG Labtech). Fluorescence values were analysed with the following sigmoidal equation:

\[
y = A_0 + \frac{A - A_0}{1 + \exp\left(-\left(t - t_{50\%}\right) \ast k_{agg}\right)}
\]

Where \(A_0\) and \(A\) are the values at the beginning and the end of the aggregation, \(m_{A_0}\) and \(m_A\) are the slopes of the lag phase and the plateau, assuming for them a linear dependency of normalized fluorescence values with the incubation time, \(t_{50\%}\) is the midpoint of aggregation and \(k_{agg}\) is the apparent aggregation rate constant.

**Subcellular fractionation**

The protocol was adapted from (Chen et al., 2000). Briefly, crude worms lysates from synchronized L4 animals were resuspended in hypotonic buffer (15 mM HEPES KOH pH 7.6, 10 mM KCl, 5 mM MgCl₂, 0.1 mM EDTA, 350 mM Sucrose, 1 mM PMSF, 1 mM DTT, supplemented with protease inhibitors cocktail (Complete, Roche). Animals were homogenised using microtube pestle rods and motor (Kontes) by applying 10 strokes of 1 minute followed by 1 minute rest in between strokes. Worm debris was removed by centrifugation at 500 g for 5 minutes at 4 °C. The pellet was discarded and
the supernatant was used as input of the cell fractionation. Nuclei were pelleted at 4000 g for 5 minutes at 4 ºC. The supernatant containing the cytosolic fraction was centrifuged at 17 000 g for 30 minutes at 4 ºC to remove membrane fraction and contaminants. The nuclear fraction was washed two times with hypotonic buffer and resuspended in hypertonic buffer (15 mM Heps KOH pH 7.6, 400 mM KCl, 5 mM MgCl₂, 0.1 mM EDTA, 0.1% Tween 20, 10% Glycerol, 1 mM PMSF, 1 mM DTT, supplemented with protease inhibitors cocktail (Complete, Roche)). The nuclear fraction was homogenized using 29G ½ inch needle insulin syringe (Terumo) and treated with 25 U/µl Benzonase (Millipore) for 30 minutes at 4 ºC. Ten µg of each fraction was loaded on a 12% SDS-PAGE. Gels were transferred to nitrocellulose membranes (Brunschwig Chemie) and incubated with the primary antibodies α-FLAG (#F3165, Sigma-Aldrich) at a 1:1000 dilution; α-LMN-1 (#38530002, Novus Biologicals) at a 1:1000 dilution; α-GFP (#632381, Clontech Laboratories) and α-tubulin (#T6074-200UL, Sigma-Aldrich) at a 1:5000 dilution overnight at 4 ºC. Incubation with the secondary antibody was at a 1:10 000 dilution for 1 hour at room temperature. Antibody binding was visualized with an ECL kit (#RPN2232, Amersham).

**Bioinformatic analysis**

Conserved domains were identified using SMART (Simple Modular Architecture Research Tool) (Letunic et al., 2012; Schultz et al., 1998). Nuclear signal localization was predicted using NLS Mapper (Kosugi et al., 2009). Protein folding was predicted using FoldIndex (Prilusky et al., 2005). The algorithm used for motif discovery was the MEME Suite (Bailey et al., 2009). Orthologues were identified using protein BLAST search (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and aligned with T-Coffee multiple sequence alignment tool (Notredame et al., 2000). Amino acid predictions were performed using ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi).
QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification of aggregates and motility assay

The number of aggregates present in whole worms was counted using a fluorescence dissection microscope (Leica). A minimum of 20 worms was counted and the data from three or four biological replicates were combined. For the motility assay, 15 animals were counted per experiment and the data from three biological replicates was combined.

A biological replicate was defined as an independently grown worm population, before extraction of embryos from gravid adult hermaphrodites to obtain synchronized populations (see EXPERIMENTAL MODEL AND SUBJECT DETAILS). The exact number of biological replicates is indicated in the figure legends. Data are represented as mean ± SEM and significance was calculated using a one-tailed or two-tailed unpaired Student’s t-test. The exact type of Student’s t-test is indicated in the figure legend. *p<0.05 **p<0.01; ***p<0.001

RNA sequencing data analysis

All RNA seq are represented as the mean of n=3 biological replicates, with error bars representing mean ± SEM. A generalized linear model was used to identify differential gene expression between the Q40 wild type and the Q40 mutants with EdgeR (Robinson et al., 2010, McCarthy et al., 2012). Replicates 1-3 were introduced as a technical factor to correct for batch effect. Genes with average expression level below 1 fragment per million (FPM) were excluded from the analysis. The library normalization was left at the standard setting (trimmed mean of M-values, TMM). The resulting p values were corrected for multiple testing using the Benjamini-Hochberg procedure. Per gene expression data was normalized as FPM. Data visualization and statistical tests were conducted using R scripts (available upon request).
Quantitative PCR analysis

All quantitative PCR are represented as the mean of n=3 biological replicates, with error bars representing mean ± SEM. To measure the relative \(lir-3\) mRNA expression, \(lir-3\) transcript levels were quantitated using a standard curve of pooled cDNA solutions and expression levels were normalized against the endogenous reference gene \(cdc-42\). To measure the relative Pol II or Pol III mRNA expression, the corresponding transcript levels were normalized to the expression of \(rps-21\) following the \(2^{-\Delta \Delta CT}\) method (Livak and Schmittgen, 2001) (see METHOD DETAILS).

Filter retardation assay

Immunoblots were quantified by densitometry using ImageJ (listed in Key Resources Table). To quantify the relative amount of SDS-insoluble protein, a ratio (fold change) was calculated by dividing the values of \(Q_{40;moag-2/lir-3}\) mutants by their corresponding wild types (corrected to \(\alpha\)-tubulin as a loading control).

DATA AND SOFTWARE AVAILABILITY

All ChIP seq data generated in this study have been deposited in the ENCODE database, under ID code ENCSR408FDZ and in the ArrayExpress database, under ID code E-MTAB-4174.

All RNA seq data generated in this study have been deposited in the ArrayExpress database, under ID code E-MTAB-4172.

Full images used in this study have been deposited in the Mendeley database under: https://data.mendeley.com/datasets/knzkygbf6x/draft?a=f246576a-2cc7-44c2-b9c4-37de8d6fe98a
SUPPLEMENTARY MATERIAL

In order to reduce the size of this thesis, so it will not take up too much shelf space, all supplemental material of this thesis has been moved online. For additional reading, please see the following link:

https://drive.google.com/open?id=1skLPqE2hToDLGgkJnmCoWcTVeBmiZX1-

If you might encounter a broken link, please contact me:

Tristan_dejong@hotmail.com