A functional study of the conserved LSM proteins in *C. elegans* reveals their involvement in the stress response of metazoans

Eric Cornes Maragliano

DOCTORAL THESIS UPF / 2015

THESIS CO DIRECTED BY

- Dr. Julián Cerón

Cancer and Human Molecular Genetics, Bellvitge Biomedical Research Institute-IDIBELL, L'Hospitalet de Llobregat, Barcelona, Spain

- Dr. Denis Dupuy

Univ. Bordeaux, IECB, Laboratoire ARNA, F-33600 Pessac, France

À mon ami Canardo,

un des co-fondateurs du *coin chercheur*.

Acknowledgements

Quisiera aprovechar este espacio para dar las gracias especialmente a un tridente sin el que no hubiera sido posible realizar esta tesis. Tesis que empezó como “un salto a la piscina” y la cual resume perfectamente la siguiente frase: “esto es como una maratón y nosotros empezamos con 1 hora de retraso”. Esta frase es de Julián y a él tengo que agradecerle su “salto a la piscina” gratuito (además de tantos otros) y toda la atención y dedicación que me ha prestado desde el primer día hasta el último tanto dentro como fuera del lab. Ha sido un gran supervisor y un amigo a lo largo de estos 4 años, donde siempre se ha preocupado por nuestros proyectos y ha sabido balancear la entrada y salida de charcos (dentro y fuera del lab) en los momentos clave. Pese a las dificultades y los barrizales siempre ha buscado la manera de hacernos participar en congresos, estancias, etc. Así que directa, o indirectamente, la base de TODO lo que ha ocurrido estos últimos 4 años es gracias a él ☺ A pesar de los “golpes de riñón”, “partido a partido” y momentos “amarrategui” a lo largo de la tesis, he disfrutado siempre con el balón. Estoy siempre agradecido y feliz de haber compartido y aprendido tantas cosas juntos a lo largo de esta tesis.

J’aimerais aussi remercier à mon autre boss, Denis. J’ai eu la chance de tomber par hasard à la Californie française, à Pessac, au milieu de Hollywood et j’ai eu l’opportunité de voir comment les choses marchent ailleurs. Quand je regarde en arrière je vois comment le temps passe vite et je comprends toutes les choses que j’ai appris avec lui maintenant. Et si je continue en recherche j’aimerais garder un peu cette partie de Denis qui m’a fait voir les choses différemment. Il a été toujours ouvert à discuter sur les projets, à perdre le temps à essayer de me faire comprendre mes erreurs avec un esprit critique que j’ai trouvé très cool. Grace à lui cette thèse à été possible aussi. Je remercie toute cette attention et préocupation pour qu’on puisse arriver à la fin, toutes les corrections et commentaires sur les papiers, la thèse etc. J’ai

beaucoup apprécié tout mon temps passé à Bordeaux où j'ai maintenant de très bons amis et je considère ma ville, Denis inclus, thanks a lot.

Y no se puede llegar al final de una carrera tan larga y cuesta arriba sin la energía de Alberto Villanueva al que tengo que agradecerle siempre su disposición y su ayuda esenciales. Siempre se ha preocupado por como iban las cosas y ha sido un amigo más a lo largo de esta tesis. Por cierto, te debo una botella de vino de Burdeos, la tengo en casa, mañana te la traigo. Con su buen rollo y hiperactividad, junto a Julián y Denis forman el tridente perfecto, gracias a los tres. De ellos quisiera dejar aquí algunos grandes highlights que resumen una tesis.

JC "No dejes que el arbol te impida ver el bosque"
"Barrizal" "Charco" "Focus!"
DD "Hola cabrón" "Donde está la biblioteca" "Life is short"
AV "Venga. Hasta luego"

También quiero agradecer a los miembros del Cerón Lab y de Genética Molecular, pasados y presentes, por los buenos momentos que hemos compartido todos los días y el *support* en los momentos de hundimiento global. Con ellos la esquizofrenia del investigador se hecho más llevadera. Especial agradecimiento también a *los del LRT* que han aguantado mis chapas y mis robos de comida con una sonrisa y que han formado parte de todos los buenos momentos que he pasado en este lab. Con especial carño y amor incondicional a la "Comunidad de la Amistad".

Si hablo de personas o cosas esenciales para la realización de esta tesis, no puedo acabar sin citar "al portátil de Rashid" y "los pendrives de Rashid" sin los que – pese a los robos- esta tesis no estaría en este papel. A ellos y a su dueño Gracias! A lo largo de esta tesis he tenido la oportunidad de encontrarme con mucha gente especial para mí, en lugares distintos y he pasado tan buenos momentos que no tengo espacio para recordarlos a todos, pero a

toda esa gente especial y a mi familia, gracias por estar cerca siempre ☺

Quand je parle de famille, je n'oublie pas de remercier à toute ma famille bordelaise, ma maison Boivin et tous mes collègues de Hollywood /Carrère. Je garde un très bon souvenir et le plus important j'ai connu des personnes très speciales dans ma vie. Merci à tous mes *gabachos*.

Summary

Lsm proteins regulate RNA metabolism and are conserved in the three domains of life, typically functioning as RNA-binding complexes involved in a wide range of post-transcriptional mechanisms. Generally, their functions have been explored in unicellular models using biochemical approaches; however their physiological roles in multicellular organisms remain unknown. This gap in knowledge is biomedically relevant since alterations of individual LSM proteins functions have been related to cancer development.

We performed a functional study of the eleven LSM proteins encoded in the *C. elegans* genome. We found that although *lsm-1* and *lsm-3* genes are not essential for the viability of the organism, they are required for wild type healthspan. In addition LSM-1 and LSM-3 proteins function in stress responses by promoting cytoplasmic LSM foci formation and influencing Insulin/IGF-like signaling pathway, a major regulator of development and stress response in metazoans. This study uncovers a physiological role for the LSM proteins in multicellular organisms as essential players for healthspan maintenance and stress adaptation.

Resumen

Las proteínas de la familia Lsm están conservadas desde bacterias a humanos y participan en el metabolismo de ARN. Aunque el estudio de sus funciones ha sido generalmente abordado mediante aproximaciones bioquímicas en modelos unicelulares, sus funciones en organismos multicelulares son desconocidas. Su estudio en organismos modelo es de especial relevancia biomédica ya que la alteración específica de ciertas proteínas Lsm ha sido relacionada con el desarrollo del cáncer.

Mediante el estudio funcional de las once proteínas Lsm presentes en *C. elegans*, mostramos cómo los genes *lsm-1* y *lsm-3*, pese a no ser esenciales para la viabilidad del organismo, son necesarios para el mantenimiento de su salud. Además, LSM-1 y LSM-3 funcionan durante la respuesta a estrés promoviendo la agregación citoplasmática de proteínas LSM y contribuyendo a la correcta señalización a través de la vía de la Insulina/IGF-like, importante en la regulación del metabolismo y la respuesta a estrés en metazoos. Este estudio destaca el importante papel fisiológico de las proteínas LSM en el desarrollo y la adaptación al estrés de un organismo multicelular.

Preface

The multifunctional role of splicing factors

The conceptual framework of this thesis started in a large scale RNAi screen in *C. elegans* searching for genetic interactors of *lin-35/*Rb (Ceron et al. 2007), the worm homolog of the human retinoblastoma gene. Interestingly, this study identified specific splicing factors among the top *lin-35/*Rb interactor candidates. Since *lin-35*/Rb generally functions as a transcriptional repressor, the authors raised the hypothesis about additional regulatory possibilities for specific splicing components. A particular example of such interactors is RSR-2, the *C. elegans* ortholog of the human spliceosomal protein SRm300/SRRM2. Its functional characterization in worms, showed that rather than having a direct role in splicing, RSR-2 influences transcription during animals development (Fontrodona et al. 2013).

Over the last years, the regulatory potential of the splicing mechanism has become an exciting research field. An example of that is the growing number of publications investigating co-transcriptional splicing and suggesting the possibility that transcription and splicing machineries are mechanistically coupled (reviewed in Bentley 2014). Moreover, the systematic analysis of the knockdown of individual splicing components has been recently applied to investigate specific functional interactions among core splicing factors. By this approach it has been shown how specific subsets of splicing factors can have distinct functional implications

on different cellular processes such as alternative splicing or apoptosis (Papasaikas et al. 2014). All together these examples show how splicing is mechanistically and dynamically coupled to other cellular processes rather than being a functionally static and isolated process in time and space. A perception that can be extended to other essential cellular RNA processing events.

The functional study of the splicing-related LSM family of proteins

In this context, this thesis started as a proof-of-principle functional study of a splicing-related family of proteins in *C. elegans.* Our approach relied on the well developed genetic toolbox available in this model organism to study gene functions. The generation and analysis of high-quality functional information related to splicing components would be helpful to reveal their possible multiple functions beyond splicing and their impact on organisms' physiology.

We started working with components of the Sm-like (LSM) family of proteins. Members of this family are conserved from bacteria to humans and typically function as RNA-binding complexes (Tharun 2009). To date, the components of two canonical heteroheptamers (the cytoplasmic LSM1-7 and nuclear LSM2-8) are the best studied LSM proteins in eukaryotes and function in many steps of RNA metabolism, including splicing (LSM2-8 complex) and messenger

RNA decay (LSM1-7 complex). However, their precise function in these cellular mechanisms still remains obscure.

There are different observations pointing to the possible additional functions of LSM proteins beyond splicing (i) their presence in prokaryotes, where splicing has not been described, (ii) protein components are shared between a cytoplasmic complex functioning in mRNA degradation and a nuclear complex involved in splicing. Additionally, the human homolog of LSM-1 (a cytoplasmic-specific LSM protein) has been found overexpressed in tumors and has been characterized as an oncogene (Schweinfest et al. 1997; Watson et al. 2008) suggesting also specific functions for individual components of the LSM complexes.

Since the loss-of-function analysis of splicing components is difficult because of their essential functions, we took advantage of RNA interference approaches in *C. elegans* to modulate the inactivation of genes. The results presented in this thesis offer the first complete description of LSM proteins functions in a metazoan organism. It also suggests the existence of a functional heterogeneity among components of the canonical LSM complexes both at phenotypic and transcriptional regulatory levels.

By genetic characterizations, we describe LSM-1 and LSM-3 as two non-essential LSM proteins required for a correct organism development and stress responses. We also show that LSM-1 and LSM-3 are involved in the stress response of the organism by

mechanisms affecting the correct functioning of the Insulin/IGF-like signaling pathway, a central pathway coupling metabolism to development and stress response in metazoans. We propose a model for the functions of specific LSM proteins during development and stress conditions.

The functional link between cytoplasmic LSM proteins and the IIS pathway could be an argument to justify the high level of conservation of LSM proteins by having an ancestral role linking stress adaptation to RNA metabolism as it is discussed in this thesis.

Table of contents

List of abbreviations

Emb :embryonic lethality (*C. elegans* phenotype)

Gro : growth variant (*C. elegans* phenotype)

Let : lethal (*C. elegans* phenotype)

Lva : larval arrest (*C. elegans* phenotype)

IGF : Insulin-like growth factor

IIS : Insulin/IGF-like signaling

mRNA : messenger RNA

ncRNA : non coding RNA

dsRNA : double stranded RNA

PB : P body

RBP : RNA-binding protein

Rbs : reduced brood size

snRNA : small nuclear RNA

snRNP : small nuclear ribonucleoprotein

RNP : ribonucleoprotein

RNAi : RNA interference

SG : stress granule

Ste : sterile (*C. elegans* phenotyp

1. INTRODUCTION

1.1. The complex life of RNA

The view of RNA as a central molecule in Biology is reflected in what is commonly known as the sequence conservation hypothesis. In this concept enunciated by Francis Crick in 1958, RNA appears as a transitory template in the middle of the sequential transfer of information between DNA and proteins, considered at that time the final functional units of the cells (Figure 1).

Figure 1. RNA as the messenger molecule. Diagram of the sequence hypothesis in which the information encoded in the sequence of nucleotides specifies the protein aminoacid sequence.

Active research in the RNA field over the last 50 years has contributed to include additional layers of complexity to this information flow. In fact, many functions have been attributed to RNA apart from its role as messenger. The identification of thousands of individual new non coding RNAs with functions in the regulation of genome organization and gene expression - in addition to the previously known transfer and ribosomal RNAs with similar functions in the regulation of protein translation - drive the actual recognized view of RNA as a regulatory paradigm (Morris and Mattick 2014). Indeed, nowadays RNA is considered to perform essential cellular and developmental functions either directly, as

non-coding RNAs, or indirectly, as a protein coding messenger RNA (mRNA) molecule.

Independently from their nature, most cellular RNAs do not act in isolation as *naked* molecules but rather as RNA-protein (RNP) complexes. In fact, each step in the life of RNA, from transcription to processing to function, involves the dynamic organization of RNP complexes (Glisovic et al. 2008) (Figure 2).

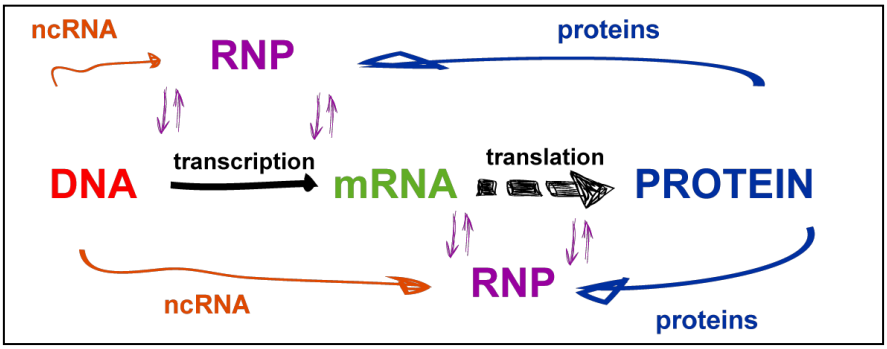


Figure 2. Regulatory and messenger roles of RNA. Individually or in combination with proteins as ribonucleoproteins (RNP), non coding RNAs (ncRNA) can perform structural, catalytic and regulatory functions and influence the expression and fate of messenger RNAs (mRNAs).

By binding to RNA, proteins can participate in the coupling of different metabolic processes or influence its localization. This, for example, allows the spatiotemporal confinement of RNA-related processes to specific cellular regions meaning that the composition and dynamics of RNPs constituents dictate the fate of an mRNA (Moore 2005, Müller-McNicoll and Neugebauer 2013). Consequently, gene expression is vastly influenced by post

transcriptional mechanisms including RNA-RNA and RNA-Protein interactions and miRNA-mediated regulation. An example of the importance of the post transcriptional regulation affecting RNAs is highlighted by imprecise correlation existing between the steady state levels of mRNA versus the corresponding proteins *in vivo* (Ghaemmaghami et al. 2003; Greenbaum et al. 2003; Schwanhäusser et al. 2011; Grün et al. 2014).

1.2. mRNA-binding proteins and messenger ribonucleoproteins (mRNPs): at the interface of gene expression

As for other RNA molecules, the control of mRNA function is modulated by complex interactions between mRNA, proteins and non coding RNAs in what is known as messenger ribonucleoproteins (mRNPs).

The life of an mRNA can be summarized in five essential steps of mRNP formation and remodeling where mRNAs are:

1- Transcribed,

2- Spliced,

3- Exported to the cytoplasm,

4- Translated and

5- Degraded.

The mRNA is led along this cellular trip, through interactions with different mRNP machineries, functioning in essential steps of mRNA biogenesis and decay (Figure 3) (Mitchell and Parker 2014).

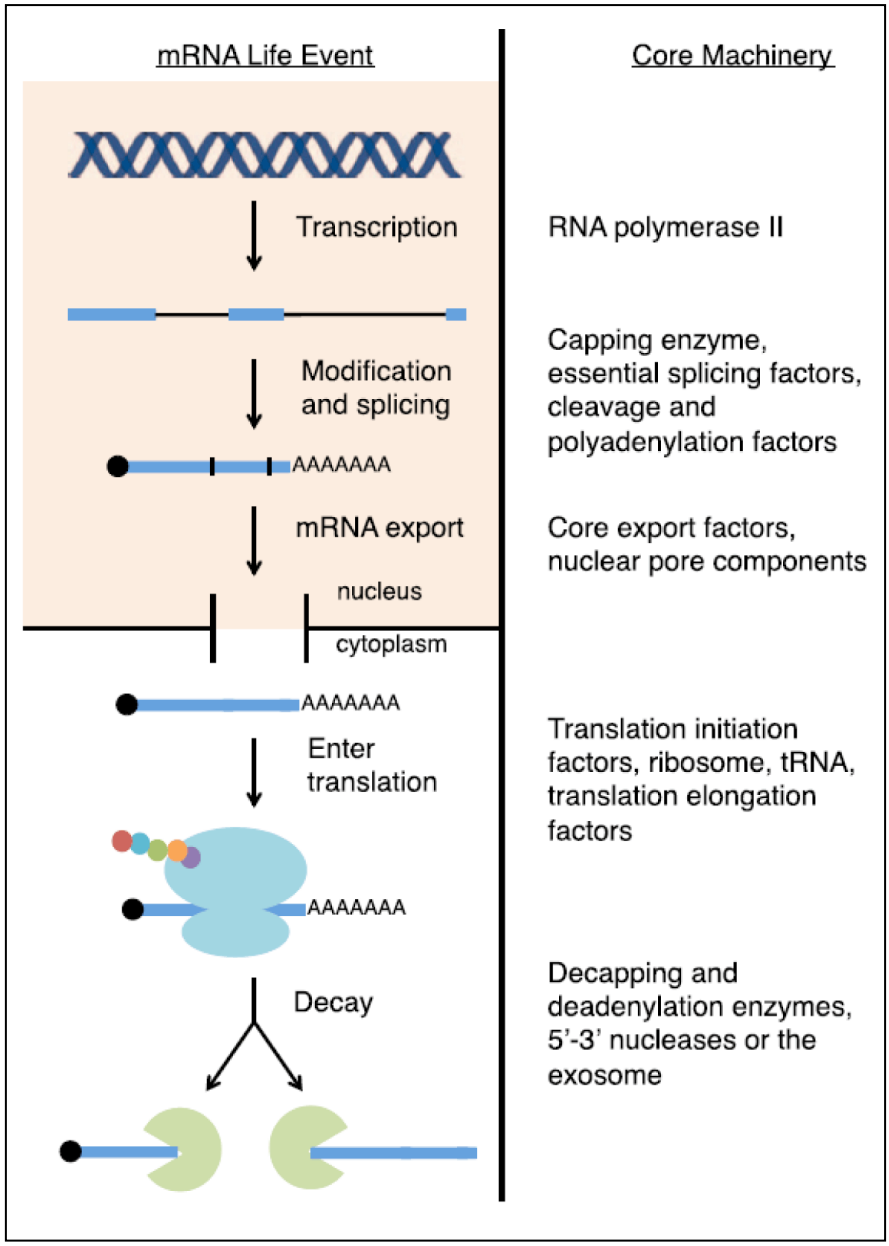


Figure 3. The central steps of mRNA metabolism. Most mRNA molecules pass through the metabolic steps schematized on the left part of the figure. Every step is performed by mRNPs referred as "core machineries" on the right (extracted from Mitchell & Parker 2014).

Major components of mRNPs are mRNA-binding proteins (RBPs), which commonly contain specific protein domains with RNA-binding capacities. To date, hundreds of such conserved domains have been characterized and can be organized in different families (Baltz et al 2012, Castello et al 2012, Kwon et al 2013, Mitchell et al 2013). A protein with a single RNA binding domain can bind many different RNA targets and a single RNA binding protein can contain combinations of multiple RNA binding domains. These two simple features highlight the relevance, versatility and specificity of RNA-Protein interaction networks in post transcriptional regulation.

As regulators of mRNA life, alterations in RBPs expression and function have profound effects in mRNP organization and cellular physiology and are at the origin of many human diseases (Lukong et al. 2008; Castello et al. 2013).

1.2.1. The spliceosome as a prototypical RNP machine

A clear example of a complex and dynamic eukaryotic RNP is the spliceosome, a multi megadalton complex of proteins and ncRNAs that interact dynamically over time to perform the removal of introns from transcribed pre-mRNAs (Wahl, Will, and Lührmann 2009; Will and Lührmann 2011). The building blocks of the spliceosome are the U1, U2, U5, and U4/U6 small nuclear RNPs (snRNPs) and a variable number of non-snRNP proteins (Figure 4A). Probably one of the most striking features of the spliceosome is the complex structural rearrangements and the massive

remodelling of spliceosomal subunits that take place during the splicing reaction (Figure 4B). During the assembly and activation of spliceosomal subunits the protein inventory associated to them is extensively remodelled over time.

The spliceosome composition, dynamics and function is a clear example of the intrincated and tightly regulated network of Protein-Protein, Protein-RNA and RNA-RNA interactions taking place around a pre-mRNA to ensure its correct maturation.

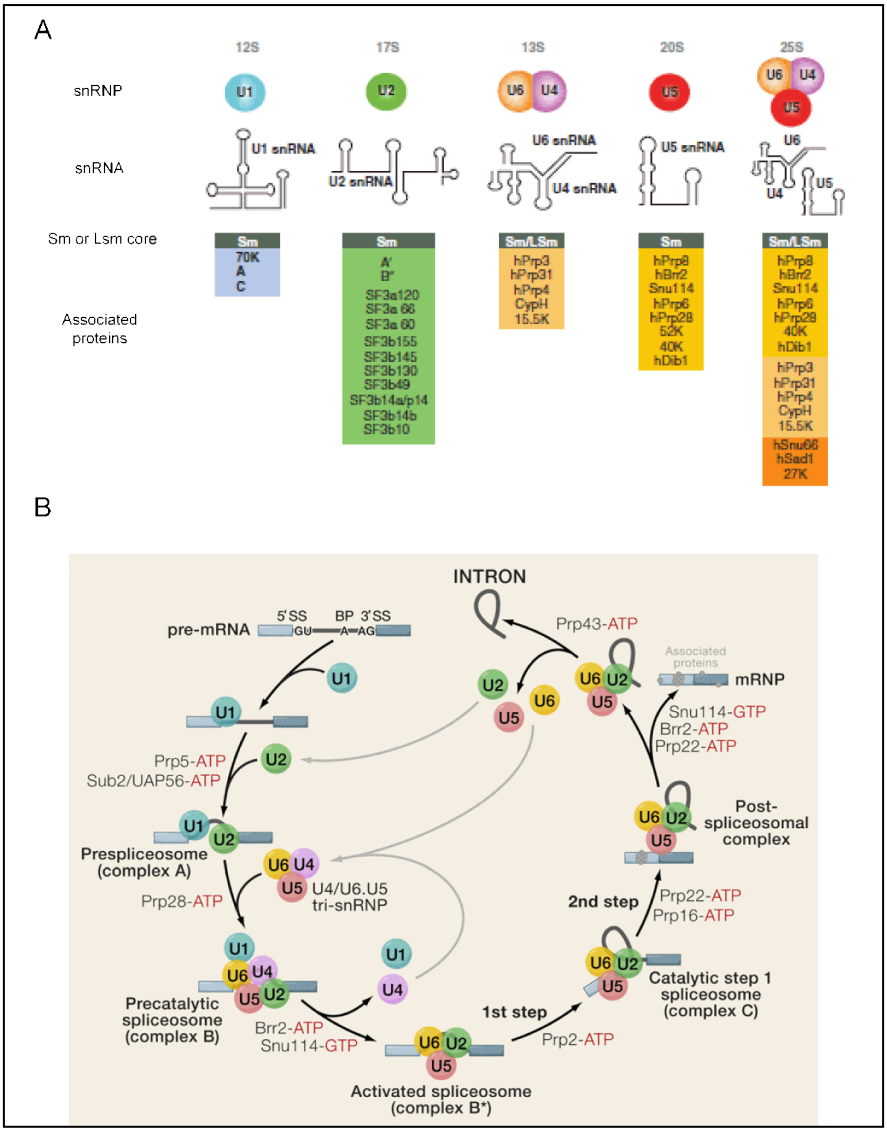


Figure 4. Protein components of the human spliceosomal snRNPs and spliceosome dynamics. A). Each snRNP is composed by a small nuclear RNA (snRNA) (or two in the case of U4/U6) associated to a protein complex of seven Sm (in U1, U2, U4, U5 snRNPs) or Lsm (in U6 snRNP) proteins and a variable number of associated proteins B) Representation of the interactions of the spliceosomal snRNP (colored circles) during the splicing of an intron from a pre-mRNA containing two exons. Additional proteins that participate as ATPases and helicases involved in RNA-RNA and RNP remodeling events are also shown (modified from Will & Lührmann 2011)

1.3. The conserved LSM family of proteins: multitask RNA-binding complexes

The presence of the domain forming a tertiary structure known as "Sm-fold" is the common signature of the large Sm/Lsm(Sm-like) family of proteins which exist in Archaea, Bacteria and Eukaryotes (Weichenrieder 2014; Mura et al. 2013) and play essential roles in RNA metabolism.

1.3.1 Common features of Sm-like proteins

Usually, members of the Sm/Lsm family are small proteins (~10-25 kDa) that share at least four structural and functional properties:

- The presence of a Sm domain that occupies most of the proteins length. It is a bipartite domain containing two Sm motifs, referred as Sm1 and Sm2, separated by a variable short-length (~20 amino acids) linker sequence (Séraphin 1995; Achsel et al. 1999; Salgado-Garrido et al. 1999). The integrity of the Sm domain is important for the RNA-binding activity of the Lsm proteins (Chowdhury et al. 2012) (Figure 5A).

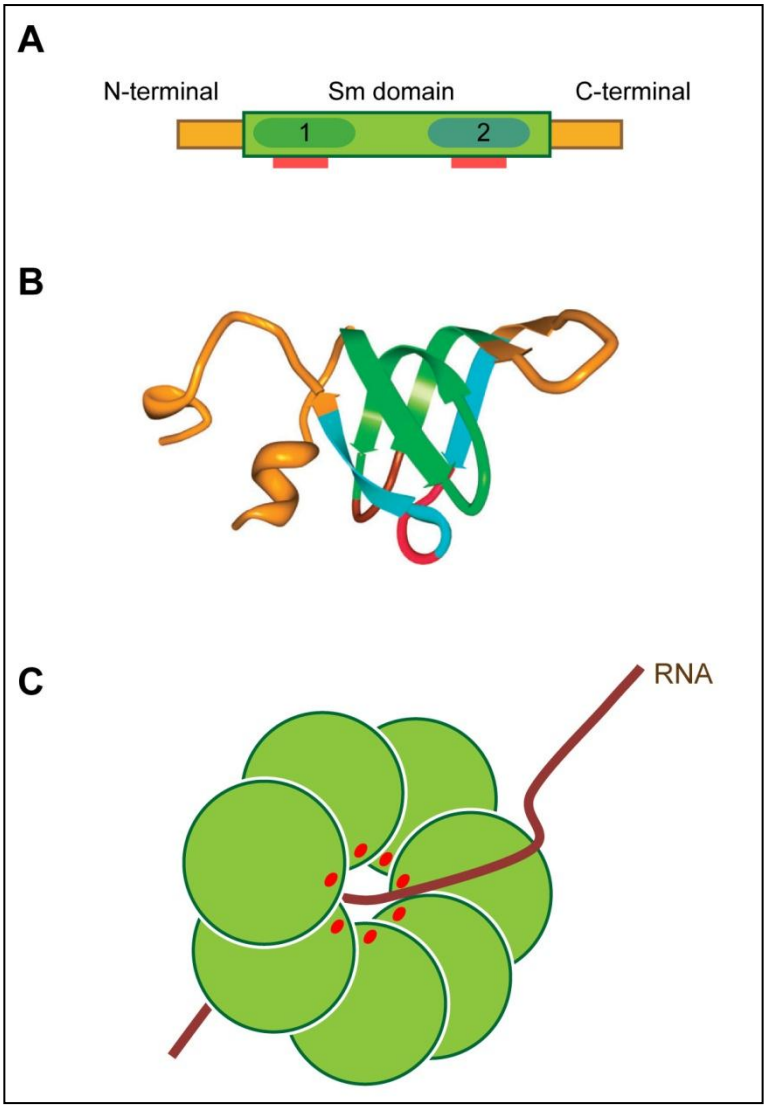


Figure 5. Sm-like protein features. A) Scheme of a typical Lsm protein. Green box represents the region encoding the Sm domain containing the two Sm motifs 1 and 2; red rectangles represent RNA-binding sites. B) Human SmD1 protein structural model. When a single Sm protein is folded, both RNA-binding sites (in red) are found within loops located in close proximity on one side of the protein. C) Scheme of a heteroheptameric Sm/Lsm complex interacting with an RNA molecule. When an entire Sm/Lsm ring is assembled, RNA-binding sites (in red) are oriented toward the central pore of the ring. RNA molecules can either pass through the central pore or interact with one of its faces (extracted from Scofield & Lynch 2008)

- A conserved tertiary structure encoded by the Sm domain and known as the "*Sm fold"* consisting in a N-terminal α-helix followed by a five stranded antiparallel β-sheet *(*Figure 5B).
- A conserved quaternary structure: The Sm-fold mediates the interaction between adjacent Sm/Lsm proteins in order to make multimeric complexes in vivo. (Figure 5C).
- Lsm complexes are RNA-binding complexes with functions in different steps of RNA metabolism (from RNA splicing to mRNA cytoplasmic decay in eukaryotes). RNAs can pass through the pore formed by the Lsm complex or interact with one of their faces. It is suggested that this complexes participate in facilitating mRNP re-arrangements around RNA, essential for RNA related processes (Tharun 2009).

1.3.2. Canonical Sm and Lsm proteins

The founding members of the family are the seven conserved eukaryotic Sm proteins named B/B', D1, D2, D3, E, F and G. These proteins associate forming nuclear heteroheptameric rings (Figure 6A) in the presence of snRNAs and are core components of different snRNPs in the spliceosome (U1, U2, U4, U5, U11, U12). They are required for the stability of their associated snRNAs and in consequence essential for the biogenesis of snRNPs and splicing integrity (Will and Lührmann 2011). Every Sm protein (D1, D2, D3, E, F and G) has an Lsm homolog (Lsm2, Lsm3, Lsm4, Lsm5, Lsm6, Lsm7) except Sm-B/B' which is weakly related to Lsm1 and Lsm8. These Lsm proteins can also function forming heteromeric

complexes (to date the best characterized are Lsm1-7 and Lsm2-8) (Figure 6B) but have also other specialized RNA-related functions including splicing, nuclear RNA processing and messenger RNA decay (Tharun 2009; Veretnik et al. 2009).

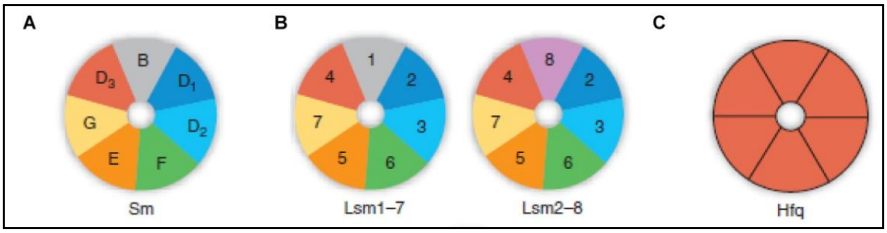


Figure 6. Organization of the eukaryotic and prokaryotic Lsm complexes. The Sm-like proteins often exists as ring-shaped hexameric or heteromeric complexes. Schematics of the eukaryotic A) Sm heteroheptamer B) Lsm heteroheptamers and C) the prokaryotic Hfq homohexamer (from Wilusz & Wilusz 2005).

1.3.3. Divergent Lsm proteins

Additionally, other individual Lsm proteins exist and are conserved in eukaryotes, such as Lsm9, Lsm10, Lsm11. Although Lsm9 has a more divergent Sm domain and has not been related with any RNA-related function, the small Lsm10 (human protein ~14kDa) and large Lsm11 (human protein ~50kDa) can associate in complexes with several Sm proteins and function in histone pre-mRNA processing (Azzouz and Schumperli 2003; Pillai et al. 2001; Pillai et al. 2003)

1.3.4. Lsm proteins with additional domains

Based on bioinformatic searches, three different classes (named Lsm12, Lsm16 and Lsm13-15) of larger proteins containing other domains in addition to the Sm domain have been described in eukaryotes (Albrecht and Lengauer 2004). Classification into these three classes are defined by the presence of specific domains but their functions remain unknown (Tharun 2009). Although many of these non canonical Lsm proteins remain uncharacterized, the yeast protein Lsm16-class protein Edc3p has been shown to have RNA-related functions.

1.3.5. Prokaryotic Lsm proteins

In prokaryotes, usually one or two Lsm homologs exist. The Hfq protein is the best studied Lsm protein in bacteria species. They differ from their eukaryotic counterparts since they only contain a single Sm1 motif (Moller et al. 2002) and organize in homomeric ring structures, usually as hexamers (Figure 6C).

In summary, Lsm proteins are a versatile and widespread family of proteins with the ability to assemble into complexes with diverse functions and RNA binding specificities (Beggs 2005; Wilusz and Wilusz 2005). The presence of proteins containing Sm domains across the tree of life indicates the relevance of their functions.

1.3.6. Lsm conserved functions

The study of Lsm functions is complicated due to the complex organization of Lsm proteins in stochiometric complexes. Although the *in vitro* reconstruction of human Lsm1-7 and Lsm2-8 complexes has been achieved (Zaric et al. 2005) and the crystal structures of yeast complexes obtained (Sharif and Conti 2013; Zhou et al. 2014), the interpretation of genetic mutations in individual Lsm proteins is difficult by the possibility of alternative complex formation in the absence of a single component.

However, the compilation of information regarding protein-protein interactions in several systems including *E. coli* (Butland et al. 2005), *S.cerevisiae* (Uetz et al. 2000; Ito et al. 2001; Gavin et al. 2002; Ho et al. 2002; Krogan et al. 2004), *D.melanogaster* (Giot et al. 2003) and human cells (Lehner and Sanderson 2004) shows a high degree of conservation for functional interaction partners between prokaryotic and eukaryotic Lsm complexes (Figure 7). These observations suggest that the functions of Lsm complexes in processes involving, among others, the decay and processing, chaperoning and splicing/transcription of RNAs are conserved. Although the majority of these interactions has not been fully validated ore demonstrated to be direct, the high level of conservation and functional diversity of interaction partners along evolution highlights the versatility of Lsm complexes in the participation on RNA-related mechanisms.

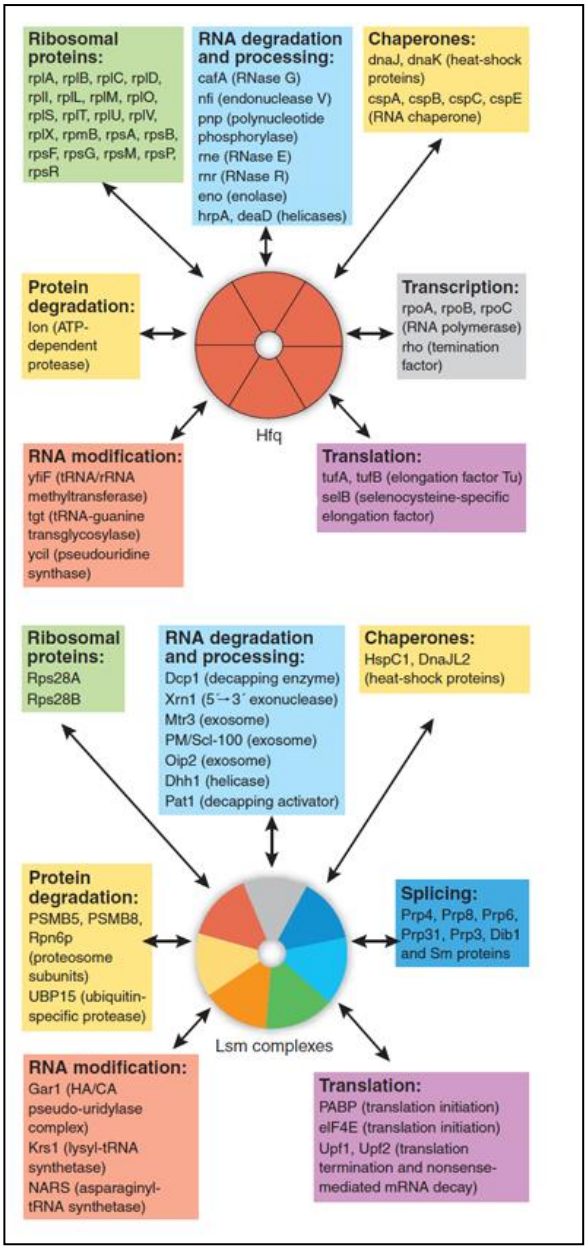


Figure 7. Protein interactors of the Hfq and Lsm complexes. Protein interactors identified from different co-purification experiments are grouped in conserved functional categories (from Wilusz & Wilusz 2005).

1.3.7. Two canonical LSM complexes function in nuclear splicing and mRNA degradation

Biochemical and genetic characterizations in yeast have been the main approaches to study Lsm functions in eukaryotes leading to the characterization of two *canonical Lsm complexes* with different biological functions, the nuclear Lsm2-8 which associates to U6 snRNA and the cytoplasmic Lsm1-7 which functions in mRNA degradation (Tharun 2009). Therefore, Lsm2 to Lsm7 are common subunits of the two complexes while Lsm8 and Lsm1 are specific for the nuclear and cytoplasmic compartments respectively (Figure 8).

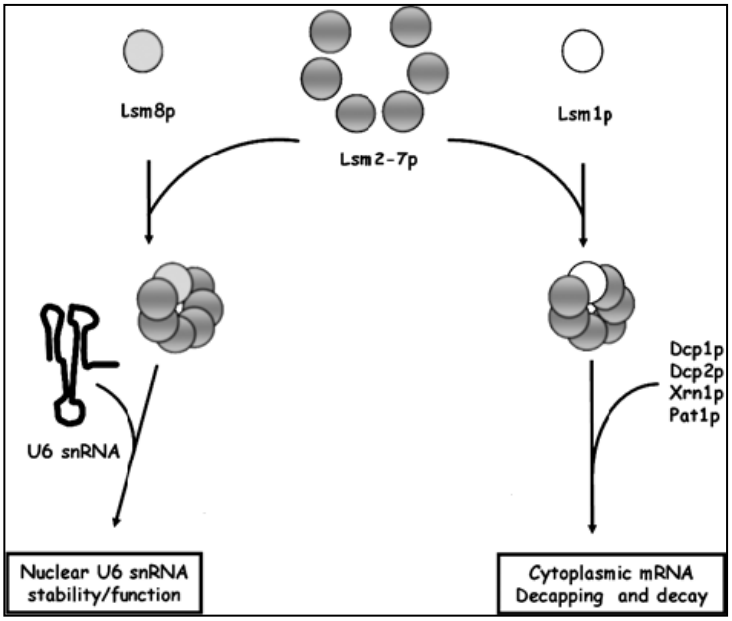


Figure 8. Lsm canonical complexes function in mRNA splicing and decay. Schematic representation of the Lsm interactions forming two different complexes in *S. cerevisiae*. Additional factors interacting with Lsm complexes to perform specific cellular functions are also shown (modified from Beggs 2005)**.**

This model comes from the observations that only seven Lsm proteins (Lsm2 to Lsm8) can associate with free U6 snRNA in the nucleus (Achsel et al. 1999). Consistently, only these Lsm proteins interact with previously known splicing factors (Fromont-Racine, Rain, and Legrain 1997) and lsm2 to lsm8 mutants affect U6 snRNA levels and pre-mRNA splicing (Mayes et al. 1999; Beggs 2005b). A suggested function for the nuclear Lsm2-8 complex is to bind specifically to U6 snRNA in the U6 snRNP and participate as a chaperone in facilitating specific rearangements between spliceosomal snRNPs during and after the splicing reaction (Achsel et al. 1999; Verdone et al. 2004) (Figure 9).

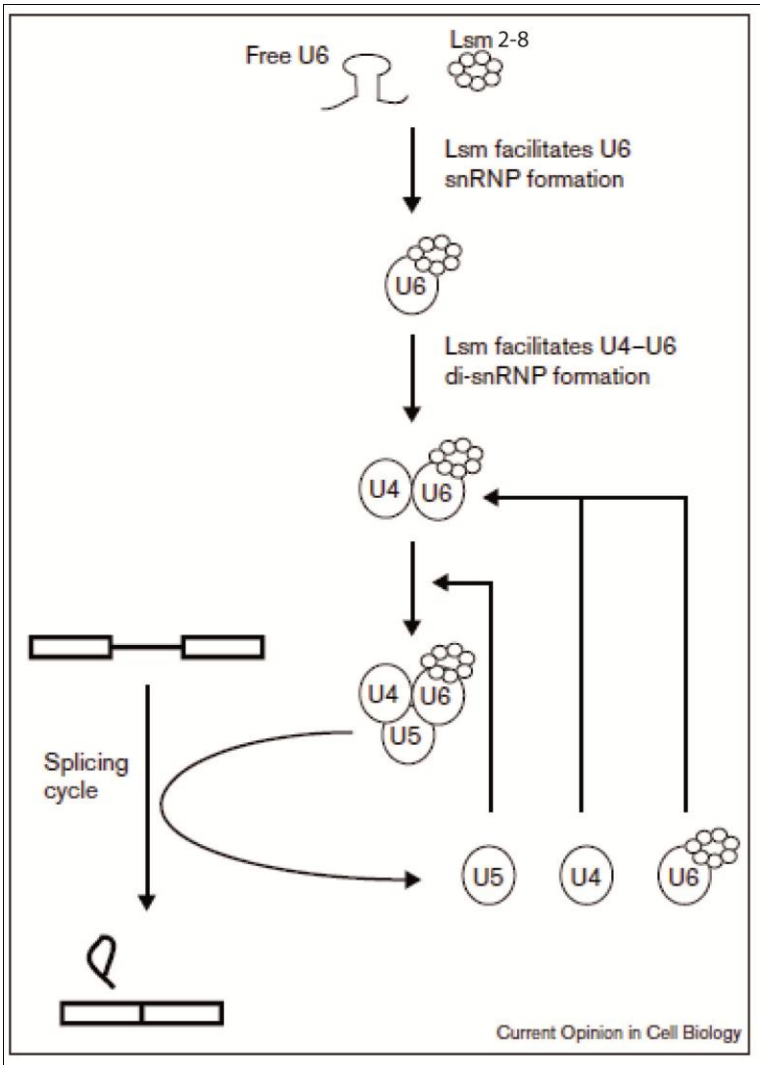


Figure 9. Lsm2-8 complex functions in U6 snRNA metabolism. A proposed role for the nuclear Lsm2-8 complex in facilitating U6 snRNP biogenesis and U4/U6 di-snRNP formation as a mediator in the rearrangement of RNA-protein complexes (extracted from He & Parker 2000).

Interestingly, Lsm1 has never been found co-precipitating with U6 snRNA species and mutations in this gene do not cause splicing-related phenotypes (Mayes et al. 1999).

Indeed, Lsm1 to Lsm7 mutant strains, but not Lsm8, are affected in the rates of cytoplasmic mRNA degradation, leading to the stabilization of capped and oligoadenylated RNAs in the cytoplasm (Tharun et al. 2000, Parker and Sheth 2007). Additionally, Lsm1 to Lsm7 proteins physically interact with mRNA decay factors such as the decapping enzyme Dcp1p, the 5'-3' exonuclease Xrn1p and the decapping activators Pat1p and Dcp2p (Tharun et al. 2000). Altogether these observations indicate that Lsm1 to Lsm7 proteins function in cytoplasmic mRNA degradation forming a complex that interacts with the decapping and decay machinery in the cytoplasm (Figure 10).

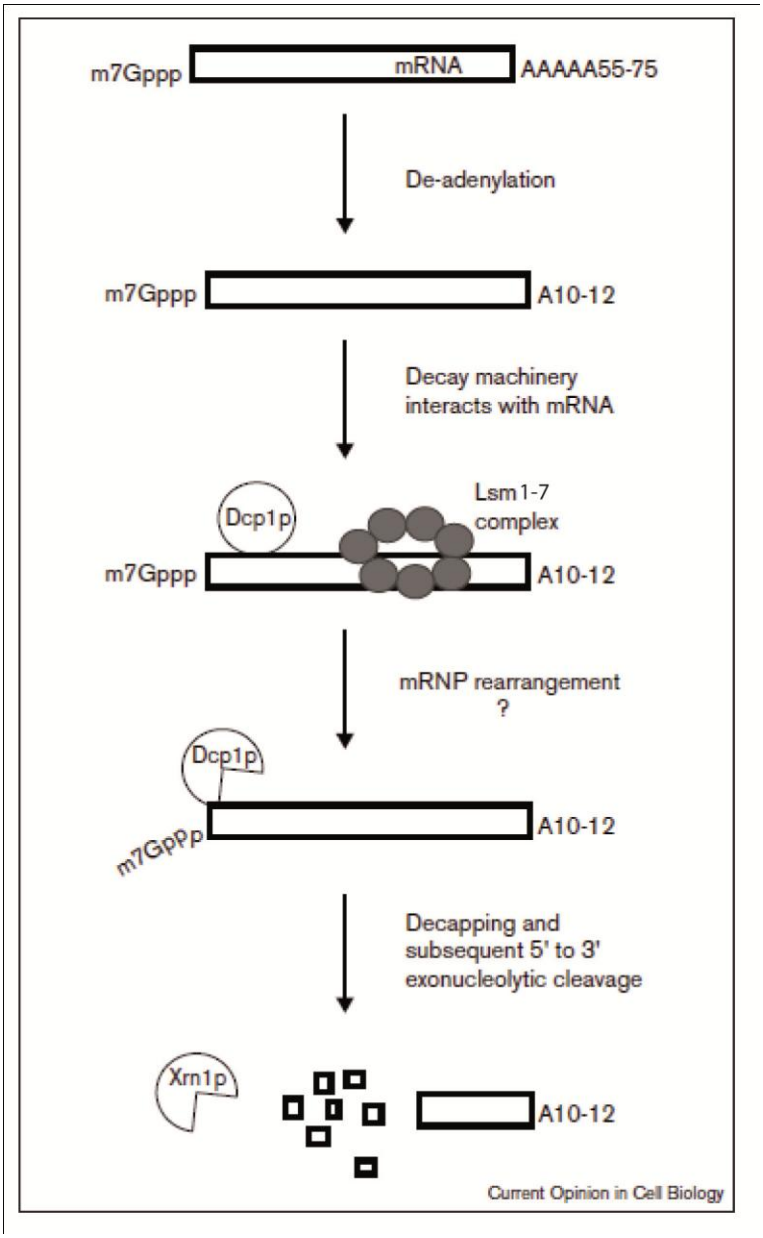


Figure 10. Lsm1-7 is an activator of decapping in the major 5'-3' mRNA degradation pathway. Lsm1-7 complex binds preferentially to oligo-adenylated mRNAs and facilitates its decapping probably by interacting with decapping enzymes and promoting Protein-mRNA rearrangements (extracted from He & Parker 2000).

Additional characterizations in human cells and plants support this model and suggest that the two complexes organization and functions might be conserved in multicellular eukaryotes (Ingelfinger et al. 2002; Perea-Resa et al. 2012). Finally, recent work has defined the structures of the nuclear and cytoplasmic eukaryotic Lsm complexes, showing that they share the same topology "Lsm-(1 or 8)-2-3-6-5-7-4" and highlighting the relevance of individual subunits in the establishment of specific interactions with additional factors involved in splicing and mRNA degradation (Sharif and Conti 2013; Zhou et al. 2014).

Besides the above-mentioned canonical LSM complexes, other heteromeric rings have been characterized in yeast and vertebrate cells pointing towards an expanded catalog of LSM functions in the modulation of RNA-RNA and RNA-protein interactions (Tomasevic and Peculis 2002; Pillai et al. 2003; Fernandez et al. 2004). A clear example the the U7 snRNP involved in histone RNA 3′ end processing. In vertebrates and flies, its protein core component is an heteromeric ring analogous to the splicing-related Sm nuclear heterocomplex. In this case, Sm D1 and D2 subunits are replaced by specific Sm-like proteins Lsm10 and Lsm11 whose function is important for histone RNA processing.(Azzouz & Schumperli 2003; Pillai et al. 2001).

1.3.8. LSM1-7, RNA turnover and cellular bodies

Many proteins involved in post transcriptional processes such as mRNA silencing, translational repression and decay localize into diverse types of cytoplasmic foci such as P bodies (PBs) and Stress Granules (SGs) (Sheth and Parker 2006; Parker and Sheth 2007;). Both are RNA-protein granules in which non-translating RNAs are localized for degradation or storage. P bodies usually contain translationally inactive mRNAs and proteins involved in mRNA degradation and translation inhibitors. On the other hand, SG contains mRNAs and translation initiation factors, indicating that these granules components associate with mRNAs in which translation has been stalled (Buchan et al. 2010). According to their known function as components of the mRNA decapping associated machinery, members of the LSM1-7 complex localize in PBs in yeast and human cells (Kedersha and Anderson 2007; Tharun 2009).

However PBs and SGs are dynamic structures that share some protein components and can physically co-localize and exchange protein components under specific conditions (Kedersha and Anderson 2007; Buchan, Nissan, and Parker 2010). The physiological roles of these aggregates are not completely understood and the classification of proteins within one or another is mostly based on protein co-localization experiments with different specific components (Figure 11).

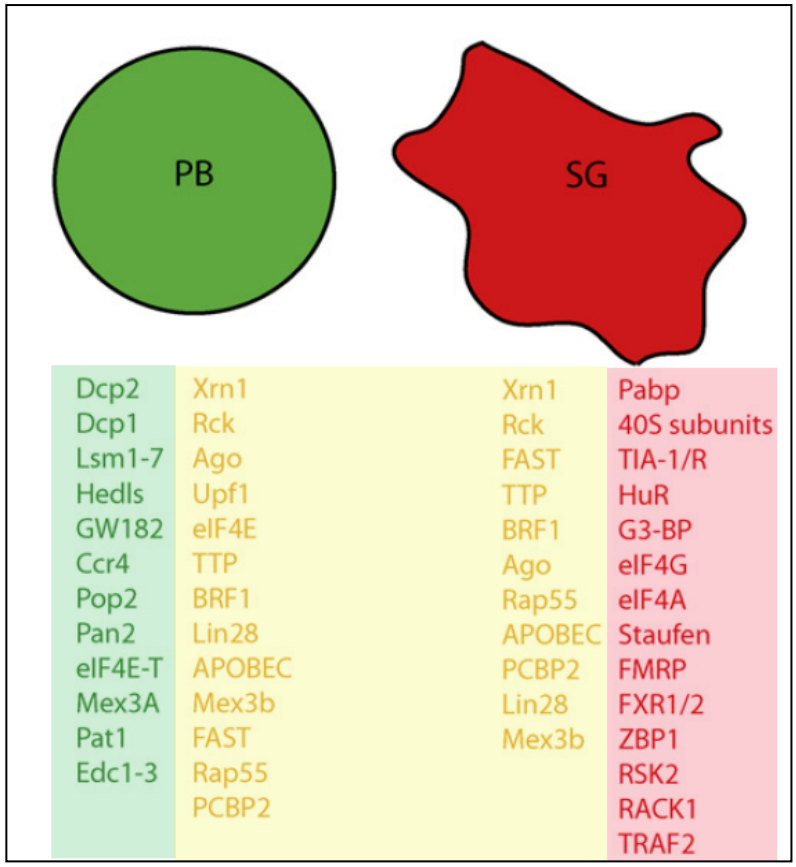


Figure 11. P bodies and Stress Granules common and specific components. Green and Red structures represent P Bodies (PB) and Stress Granules (SG) respectively. Some protein components have been characterized as PB (in green) or SG (red) specific. Others (in yellow) are found co-localizing within PB or SG (Extracted from Buchan & Parker 2009)

1.3.9. LSM proteins and disease

There is a strong correlation between the overexpression of the human LSM1 protein and the malignant development in diverse types of human cancers and cancer-derived cell lines (Schweinfest et al. 1997; Streicher et al. 2007; Watson et al. 2008). For that reason the human Lsm1 homolog gene is also referred as CaSm

(Cancer-associated Sm-like). Lsm1 overexpression is often associated to a specific chromosomal amplification found in many of these cancer types (8p11-12 region, which contains the region encoding Lsm1). The downregulation of hLsm1 expression in tumorous cell lines by using antisense and siRNA is sufficient to rescue the malignant phenotypes. suggesting that Lsm1 acts as an oncogene and can be a target for cancer therapy (Kelley et al. 2001; Kelley et al. 2003; Fraser et al. 2005; Streicher et al. 2007). All these observations underscore the importance of the control of Lsm1 cellular levels for the maintenance of a correct cellular physiology. However, although many hypotheses exist nothing is known regarding the causal oncogenic mechanism of Lsm1.

The *C. elegans* Lsm proteins are highly conserved (human and *C. elegans* Lsm1 share 90% of sequence identity) (Table 1). The powerful genetic and genomic tools available in *C. elegans* research offer the possibility to explore the functions of the Lsm proteins in a living complex system and therefore can provide insights regarding their roles in the physiology and development of multicellular organisms.

ORF	C. elegans Gene	Human gene	Blast e-value	Length
F40F8.9	lsm-1	LSm1	2,30E-22	89,60%
T10G3.6	lsm-2	LSm2	3,70E-40	95,90%
Y62E10A.12	lsm-3	LSm3	3,10E-36	97,10%
F32A5.7	lsm-4	LSm4	1,10E-38	95,90%
F28F8.3	lsm-5	LSm5	2,50E-34	98,90%
Y71G12B.14	lsm-6	LSm6	2,90E-31	96,10%
ZK593.7	lsm-7	LSm7	8,90E-30	95,20%
Y73B6BL.32	lsm-8	LSm8	4,00E-27	91,80%

Table 1. Conservation of *C. elegans* Lsm proteins.
ORF: open reading frame associated to a *C.elegans lsm* gene. Blast e-value: statistical significance for the alignment between nematode and human proteins. Length: percentage of the polypeptide length that is conserved between species.

1.3.10. LSM proteins in *C. elegans*

A phylogenetic analysis based on protein sequences shows that the *C. elegans* genome contains the eight genes (*lsm-1*, *lsm-2*, *lsm-3*, *lsm-4*, *lsm-5*, *lsm-6*, *lsm-7* and *lsm-8*) coding for small proteins (77 to 125 aminoacids) of the eukaryotic canonical complexes, and three other genes (Y48G1C.9, K07A1.15, and C49H3.4) coding also for small proteins with Sm domains only. In addition to this set of Lsm genes, other well-characterized and conserved Sm-like genes such as *edc-3* and *car-1* encode larger proteins containing other functional domains (Squirrell et al. 2006; Tritschler et al. 2007) (Figure 12).

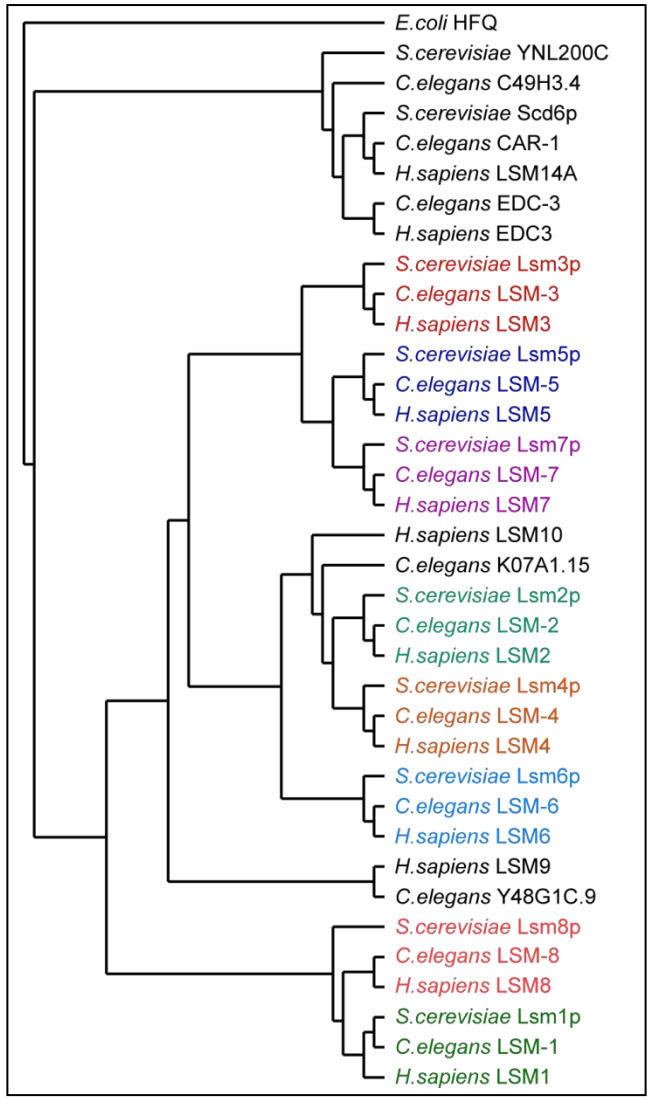


Figure 12. Phylogenetic tree of the *C. elegans* LSM proteins. Multiple protein sequence alignment was done using CLUSTAL X V2.0 and the result is represented as a cladogram where the *E. coli* LSM ortholog Hfq has been used as outgroup.

1.3.11. LSM functions in *C. elegans*

Despite the conservation of Lsm functions along evolution, little is known regarding the role of LSM proteins in metazoans.

In the context of the *C. elegans* embryo, it has been shown that LSM-1 and LSM-3 co-localizes with other PB components in somatic blastomeres, but the presence of visible LSM-1 granules is not required for mRNA degradation (Gallo et al. 2008). This is in accordance with other studies in *Drosophila* showing that processes such as mRNA silencing and decay are still functioning in cells lacking visible P body structures (Eulalio et al. 2007).

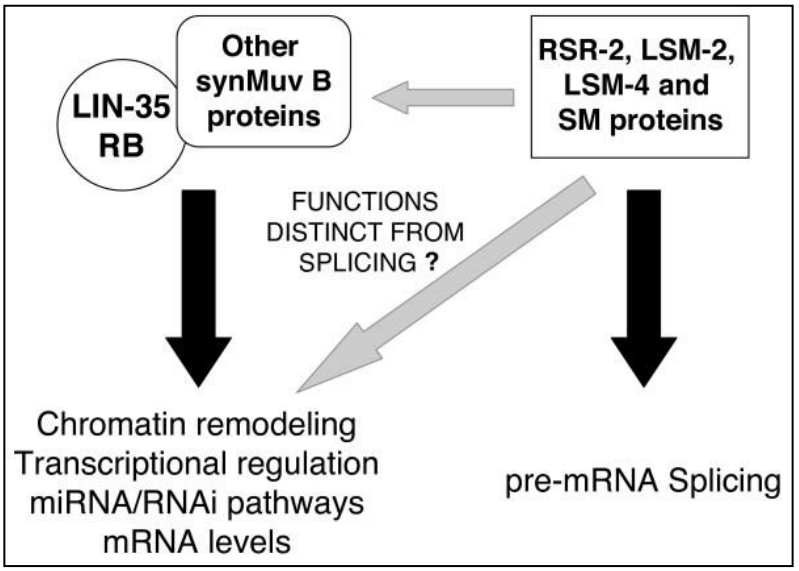


Figure 13. Proposed roles for multifunctional splicing components (extracted from Ceron et al. 2007).

On the basis of a large scale RNAi screening looking for genetic interactors of *lin-35/Rb* gene, the human Retinoblastoma homolog in *C. elegans*, two *lsm* genes (*lsm-2* and *lsm-4)* appeared as positive hits in addition to other splicing components (Ceron et al. 2007). Since *lin-35/Rb* participates in transcriptional regulation and chromatin remodelling it was proposed that specific splicing components (such as *lsm-2* and *lsm-4)* could cooperate with *lin-35* related complexes in those additional functions (Figure 13).

Together, these observations show that some LSM proteins can function as P body components in *C. elegans* and raise the possibility of multiple individual functions among core components of the eukaryotic Lsm complexes.

1.4. *C. elegans:* a top model to perform functional studies

The roundworm *Caenorhabditis elegans* was raised to the status of model organism in biologic research for the first time in the early seventies by Sydney Brenner (Brenner 1974) in the search for a simple model to unravel the genetic basis of the development of metazoan systems. In his 1974 paper, Sydney Brenner showed how it was possible to induce gene mutations in worms and observe the subsequent phenotypes at the whole organism level.

As the first animal to have its full genome sequenced in 1998 (*C. elegans* sequencing consortium 1998), *C. elegans* research rapidly benefited from the combination of powerful genetics and modern functional genomics approaches to study gene function. The model organism encyclopedia of DNA elements (modENCODE) project database is an example of the generation of a comprehensive annotation of functional elements in the *C. elegans* genome from high throughput studies (Celniker 2009, Gerstein 2010).

Extensive knowledge is now available regarding its detailed anatomy and cell lineages (Sulston & Horvitz 1977; Altun et al. 2007) or its behavioral, developmental, physiological, cellular and molecular biology (Girard et al. 2007). Additionally the WormBase database (www.wormbase.org) is used as a repository of genomic information and curation, providing data regarding the biology of *C. elegans* and related nematodes: genomes, gene models, allelic variations, mutant phenotypes, expression patterns or human disease relevance (Harris et al. 2014).

Indeed, many essential human genes and pathways are highly conserved between humans and worms making possible to model and study the molecular and cellular basis of human diseases in worms (Kaletta and Hengartner 2006). It is not casual that many of the most relevant findings in biomedicine in the last twenty years have been discovered in worms. The mechanism of programmed cell death (Yuan et al. 1993) and RNA mediated silencing (Fire et al. 1998) are just two famous examples showing how looking at worms can bring light in the understanding of mechanisms causing human diseases and the development of therapies against them.

1.4.1. *C. elegans* biology

C. elegans is a nematode specie found worldwide, who originally feeds on different bacteria present in decomposing plant material (Rézal and Félix, 2015). However, as a model organism it is commonly maintained in agar plates between 15ºC and 25ºC, feeding on *E. coli*. Among the most relevant biological features that make the worm a powerful model organism to perform genetic and functional studies are:

- A dual mode of reproduction: *C. elegans* is a diploid animal and has two sexes, hermaphrodites (XX) and males (X0). Hermaphrodites can produce both sperm and oocytes and can reproduce by auto-fertilization laying around three hundred eggs per generation. Males appear spontaneously in hermaphrodite populations but at very low frequency rates

(0.1%). They can only produce sperm and thus, to reproduce they need to mate with hermaphrodites.

Two separated sexes offer the possibility of conveniently obtaining and maintaining homozygous strains by either self-fertilization or combine genotypes by cross-fertilization. More specifically, cross-fertilization strategies can be designed to generate new strains containing different combinations of mutations and/or markers, to uncover meaningful genetic interactions.

- A short life cycle consisting in an embryonic development followed by four main larval stages (L1 to L4) and adulthood (Figure 14). This life cycle is temperature dependent and can be achieved in three days at 20°C. After the reproductive period worms can live up to 20-25 days at 20°C.

- An invariant cell lineage: The development and fate of somatic cells has been characterized at the single cell resolution (Sulston and Horvitz 1977; Sulston 1983). These cell lineages are invariant among individuals (959 cells in hermaphrodites, 1031 in males) facilitating the detection of phenotypes caused by mutations affecting cell divisions and developmental processes (Horvitz and Sulston 1980).

- A transparent body that allows the direct visualization of internal structures (Figure 2B), the cellular or subcellular phenotypes caused by specific mutations during

development and the in vivo visualization of fluorescent reporters.

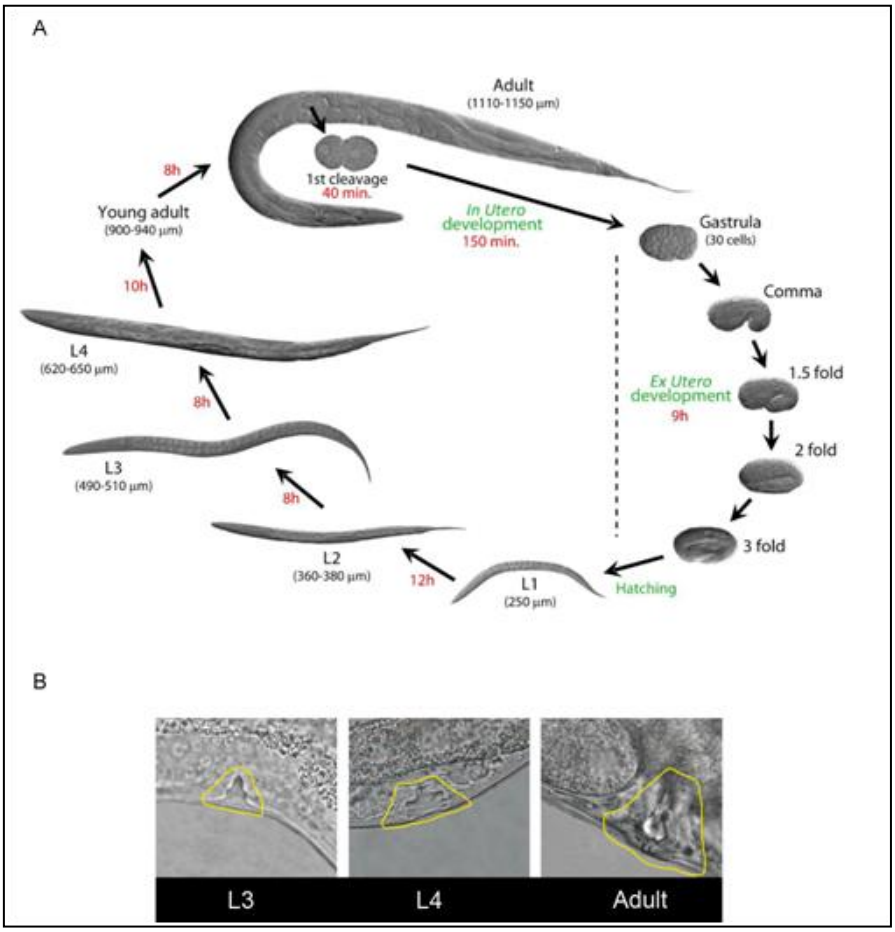


Figure 14. *C. elegans* life cycle and development. (A) Simplification of an unperturbated worm life cycle at 22°C. Numbers in red indicate the approximate length of time spent by the animals at each specific stage. (B) Pictures showing the vulval development along the specified developmental stages. Yellow line highlights the vulval structure (extracted from Porta-de-la-Riva et al. 2012).

- Generation and maintenance of mutant strains: mutations in the *C. elegans* genome can be induced by exposing worms to different chemical compounds or ionizing radiations. The abovementioned short life cycle and dual mode of reproduction facilitates the propagation and isolation of detrimental mutations as homozygous strains along genetic screens (Jorgensen and Mango 2002). Organizations such as the japanese National Bioresource Project (NBP) generate and distribute new mutant strains.

- Easy maintenance: *C. elegans* strains are cheap to maintain and easy to keep as frozen backup stocks at -80ºC for long-term storage. Organizations such as the Caenorhabditis Genetics Center (CGC) collect, maintain, and distribute published mutant and transgenic strains.

- Gene silencing by RNA-mediated interference (RNAi) is straitghtforward: the administration of gene specific dsRNA into worms by different approaches is sufficient to induce the silencing of the corresponding gene. This effect is systemic and transgenerational depending on the gene targeted. Two RNAi libraries are available and together they cover around 94% of the *C. elegans* genome allowing the study of gene function both at the single gene or genome wide level.

1.4.2. RNA-mediated interference (RNAi)

The mechanism of gene silencing by RNAi was first discovered in *C. elegans* (Fire et al. 1998) and nowadays is used as a powerful approach to specifically modulate gene function in worms and other systems. It is a sequence-specific gene regulation mechanism conserved in the majority of eukaryotes. Briefly, once the dsRNA is introduced inside the cells it is recognized by the Dicer-RDE complex who cleaves the dsRNA template in small dsRNA molecules (~20 nucleotides) also known as siRNAs. As double stranded RNA molecules, they contain a sense (homologous to the targeted gene, also known as passenger strand) and an antisense strand (complementary to the targeted gene, known as guide strand). After unwinding of the small dsRNA, the passenger strand is degraded and a complex of proteins called RNA-induced silencing complex (RISC) is guided by the antisense strand to the mRNA from the targeted gene. The silencing of the target mRNA can take place through diverse mechanisms including mRNA degradation, and transcriptional or translational blocking (Grishok 2005)

There are at least three different ways to administrate dsRNA to worms:

by feeding on genetically modified bacteria synthetising the dsRNA of interest or by dsRNA microinjection or soaking (Figure 15).

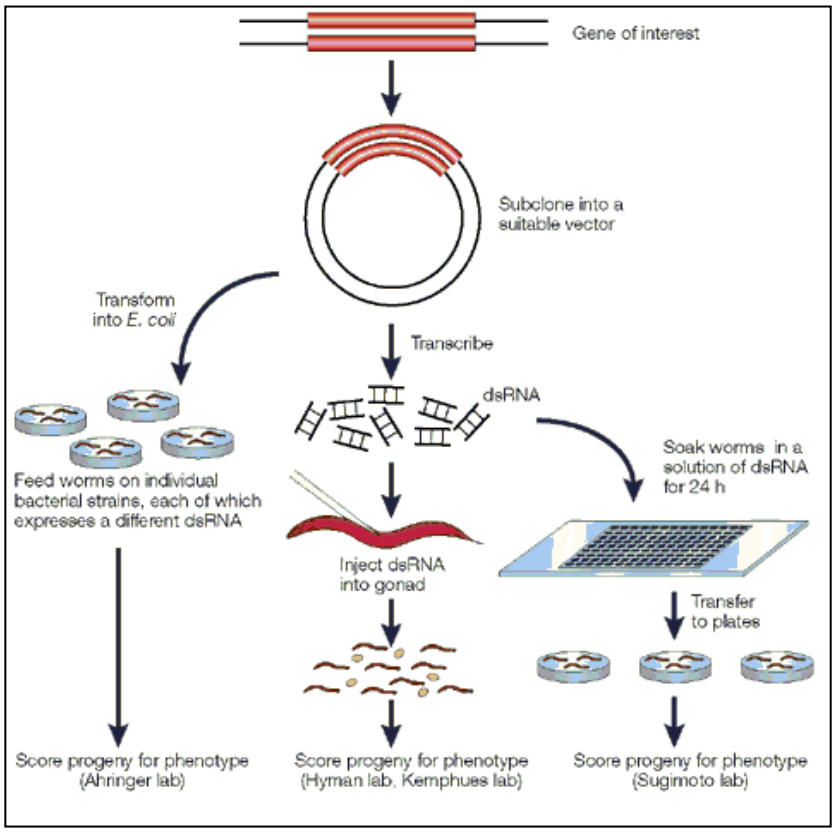


Figure 15. Different RNAi approaches used to inactívate gene function in *C. elegans*. RNAi by feeding (left), by dsRNA microinjection (middle) and by soaking (right), (extracted from Kim 2001)

a) RNAi clone libraries

Two different RNAi clone libraries exist and together they cover around 94% of the *C. elegans* genome. Basically the principle of the two libraries construction relies on the cloning of gene-specific fragments between two inverted T7 RNA polymerase promoters in the L4440 vector and transformation in the HT115 bacteria strain (A. G. Fraser et al. 2000; Kamath et al. 2003); HT115 bacteria has

IPTG inducible T7 polymerase and a non functional dsRNAse gene allowing the IPTG inducible dsRNA accumulation inside the cells (Timmons, Court, and Fire 2001). One library was made by Julie Ahringer's lab and contains 19,762 clones made by cloning genomic fragments into the L4440 vector (Fraser et al. 2000; Kamath et al. 2003). The other library is from Marc Vidal's lab and contains around 12.000 clones made by cloning full length open reading frames (ORF) from cDNA sequences (Rual et al. 2004). When a specific gene is not available in any of these libraries, a RNAi vector can be constructed by amplifying a gene-specific sequence and cloning it into a L4440 vector by classical molecular biology methods.

These libraries have been used successfully to perform large-scale RNAi screenings to find genes involved in specific processes (Gönczy et al. 2000; Piano et al. 2000; Maeda et al. 2001; Piano et al. 2002; Kamath et al. 2003; Simmer et al. 2003; Rual et al. 2004). Moreover, RNAi can be combined to target multiple genes and this approach has been used to uncover genetic interactions (Tischler, Lehner, and Fraser 2008).

b) RNAi by feeding

Bacterial clones from any of the two available libraries are fed to worms and embryonic or postembryonic phenotypes can be scored depending on the developmental stage chosen to start the experiment. Different parameters, such as temperature or the developmental stage at which the assay is performed can influence

the strength of a given phenotype. Although RNAi by feeding can give more variability compared to RNAi by microinjection, it is a good cost-effective way to treat a larger number of animals at once and perform large-scale screenings on agar plates or liquid medium.

c) RNAi by injection

dsRNA can be synthesized *in vitro* using any of the Ahringer or Vidal RNAi clone as a template (see MM section) and the dsRNA solution can be injected directly into young adult worms gonad or intestine. Usually injection of dsRNA has a reliable and potent effect observable on the F1 generation and has been used as an efficient approach to study the effect of essential and non-essential gene inactivation during embryonic development (Sönnichsen et al. 2005).

Although RNAi is a straightforward method to study gene function, genetics is a robust method to validate the biological functions of a single gene. Thus, the combination of RNAi and mutant phenotypes on a gene of interest gives solid information about the gene function.

1.4.3. *C. elegans* transgenesis

In addition to RNAi and genetic approaches, transgenesis combined with fluorescent markers is another powerful way to study in vivo gene function in *C. elegans* (Chalfie et al. 1994). DNA transformation is commonly used to study gene expression and protein localization, ectopically express genes or fluorescent reporters and rescue mutant phenotypes (Giordano-Santini and Dupuy 2011).

The principle of DNA transformation in *C. elegans* relies on the introduction of exogenous DNA (plasmid or PCR product) directly into the syncytial gonad of the self-fertilizing hermaphrodite, either by DNA microinjection or by particle bombardment.

DNA microinjection is a relatively easy technique, which results in the formation of extrachromosomal arrays, consisting of multiple copies (80–300) of the exogenous DNA arranged as concatemers (Mello et al. 1991; Stinchcomb et al. 1985). These arrays behave as artificial chromosomes, as they are efficiently replicated and segregated to the progeny, producing stable transgenic lines (Mello et al. 1991; Stinchcomb et al. 1985).

During gene bombardment, DNA-coated beads are used as vectors to introduce DNA into the animals (Wilm et al. 1999). This method also produces extrachromosomal arrays, but in addition, random integration of several copies of the transgene into the genome is observed in 1/4 to 1/8 of the obtained strains (Praitis et al. 2001; Reece-Hoyes et al. 2007). Nowadays, the most established

bombardment protocol uses *unc-119(ed3)* mutant animals as a recipient strain. These animals display an easy observable locomotor defect and are unable to enter the diapause state Dauer upon starvation (Maduro and Pilgrim 1995). During bombardment, the *unc-119* gene is used as a co-transformation marker, allowing the selection of transformed animals based on their wild-type locomotion and their ability to survive starvation (Maduro & Pilgrim 1995). The majority of transformation markers used for *C. elegans* transgenesis are easily scorable under a dissecting scope and are mostly based on the rescue of nonlethal mutations or the use of dominant or fluorescent markers, allowing visual identification of specific traits. Additionally, antiobiotic selection has been successfully applied to *C. elegans* for a more efficient and effective transformation, strain maintenance and genetic studies (Giordano-Santini and Dupuy 2011; Cornes et al. 2014)

With the complete sequencing of the *C. elegans* genome in 1998, platforms such as the Transgeneome Project (Sarov et al. 2012) or the *C. elegans* Promoterome (Dupuy et al. 2004) offer a genome wide resource of green fluorescence protein (GFP) tagged genes under cis regulatory sequences and promoters respectively to study in vivo gene expression patterns throughout animal development.

1.4.4. The Insulin/IGF-like signaling pathway in worms

Since the discovery in the 80's and early 90's of single gene mutations (*age-1* and *daf-2*) extending considerably the worms' lifespan (Friedman and Johnson 1988; Kenyon et al. 1993), *C. elegans* has become a powerful model to study biological pathways related with aging because of its short lifespan and the possibility to screen for mutations in genes that can shorten or increase it. One of the major pathways involved in the regulation of lifespan is the Insulin/IGF-like signaling (IIS) pathway whose major components are conserved in metazoans from worms to humans.

The gene *daf-2* encodes the only *C. elegans* homolog of the family of insulin-like growth factor receptors (IGFs) (Baugh & Sternberg, 2006). In the presence of food, DAF-2 is activated by binding to its ligands (Insulin-Like Peptides), inducing an AKT signaling cascade that results in the phosphorylation of the FOXO transcription factor DAF-16 (Figure 16A). As a consequence, DAF-16 is inactive and located in the cytoplasm. Upon reduced insulin pathway signalling, DAF-16 translocates to the nucleus and mediates the specific activation and downregulation of a subset of genes referred as Class I and Class II respectively (Murphy et al. 2003; Tepper et al. 2013). This transcriptional regulation dependent on DAF-16 mediates stress resistance, innate immunity and metabolic adaptation, processes that influence lifespan (Singh & Aballay 2009; Coleen T. Murphy and Patrick J. Hu. 2013).

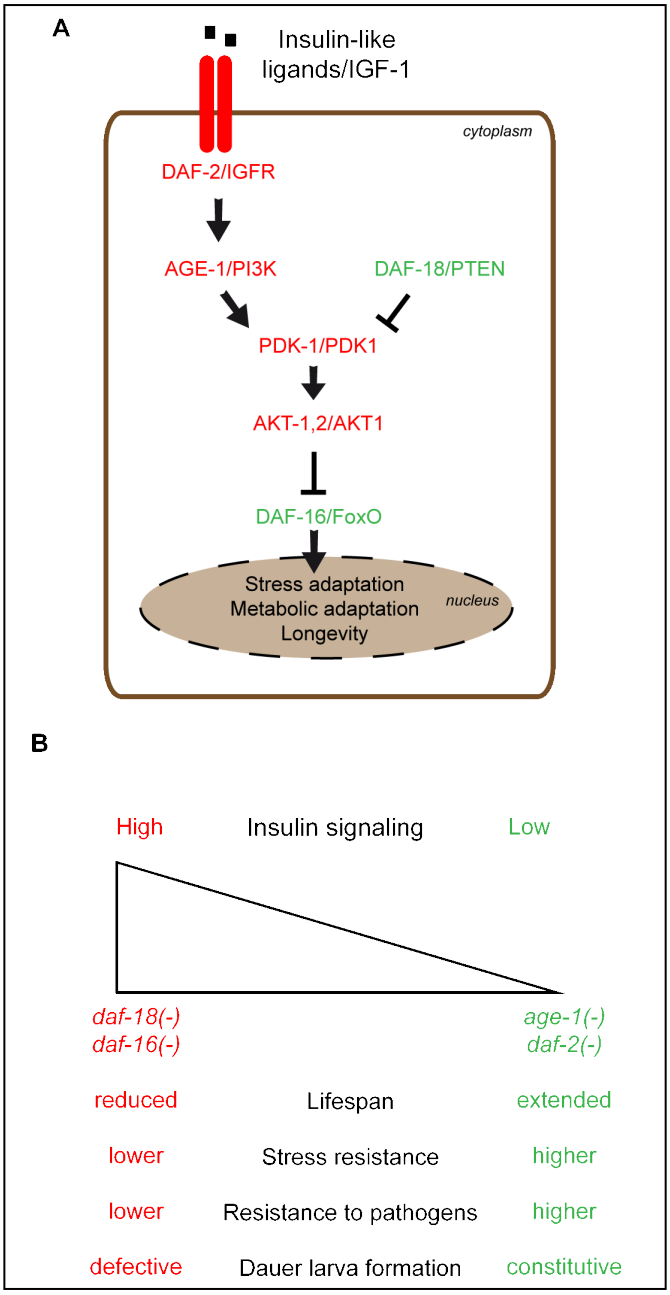


Figure 16. Insulin pathway components are conserved between worms and humans. A) Schematic of the main IIS pathway components. *C. elegans* and human names are shown. Genes that contribute to IIS are colored in red. Genes whose activities are associated with a reduced insulin signaling are colored in green. B) The IIS levels can be modulated using specific mutants producing pleiotropic phenotypes.

Many alleles have been isolated for *daf-2*, *daf-16* and other components of the pathway (Gems et al. 1998). These mutants display pleiotropic phenotypes affecting larval development, stress resistance and longevity depending on how they affect insulin signaling levels (Patel et al. 2008) (Figure 16B). The identification of the major components of the IIS pathway in worms allows using forward and reversing genetics to look for direct or indirect modulators of the IIS. For example, large-scale RNAi screenings in combination with epistatic analyses have been used to define new players and pathways contributing directly or indirectly to stress resistance and lifespan regulation (Hamilton et al. 2005; Lee et al. 2003; Hansen et al. 2007).

2. OBJECTIVES

1- A functional characterization of the LSM family in *C. elegans*.

2- To investigate the mechanism of action of *lsm-1* gene in a multicellular organism.

3. RESULTS

3.1. *lsm-1* AND *lsm-3* are non-essential genes for *C. elegans* viability

As a first step to investigate the functions of the individual *lsm* genes, we generated information regarding both phenotypes and expression. Comparing *Phenome* and *Localizome* data we intend to set the fundaments about individual and common LSM functions.

3.1.1. Phenome analysis by RNAi classifies *lsm* genes in distinct functional categories

To evaluate the relevance of every individual *lsm* gene at the organism level we took advantage of the versatility of the RNAi in the modulation of gene silencing.

We performed RNAi by feeding for all the members of the *lsm* family in L1 synchronized populations of wild type animals (Table 2). The individual knockdown of *lsm* genes has a wide diversity of effects depending on the gene targeted, ranging from the absence of any obvious phenotype after C39H3.4*(RNAi)* to the highly penetrant adult sterility caused by *lsm-2(RNAi)* in adulthood. The reduction of mRNA levels by RNAi is sufficient to cause reproducible phenotypes among *lsm* genes already on the first generation (Figure 17) using diverse RNAi conditions (ie: bacterial culture concentration, IPTG concentration between 1 to 6 mM or temperature range 15ºC – 25ºC).

Gene targeted	Canonical complex	RNAi by feeding		RNAi by injection
		P0	F1	F1
lsm-1[a]	Cytoplasmic	Rbs	Rbs,Gro	Rbs, Gro, low % Emb
lsm-2	Cytoplasmic/Nuclear	Rbs,Ste	-	100 % Emb
lsm-3[a]	Cytoplasmic/Nuclear	Rbs	Rbs,Gro	Rbs, Gro, low %Emb
lsm-4	Cytoplasmic/Nuclear	Rbs	Ste,Lva	100 % Emb
lsm-5	Cytoplasmic/Nuclear	Rbs	Ste	100 % Emb
lsm-6	Cytoplasmic/Nuclear	Rbs	Ste,Lva	100 % Emb
lsm-7	Cytoplasmic/Nuclear	Rbs	Ste,Lva	100 % Emb
lsm-8	Nuclear	Rbs	Ste,Lva	100 % Emb
Y48G1C.9[a]	Not applicable	WT	WT	WT
K07A1.15[a]	Not applicable	WT	WT	WT
C49H3.4[a]	Not applicable	WT	WT	WT

Table 2. RNAi phenotypes scored for *lsm* genes
The table summarizes the RNAi phenotypes observed at 25ºC for the eleven *lsm* genes tested relative to control animals (L4440 vector). Phenotypes are abbreviated as follows: Rbs (Reduced brood size), Gro (Slow growth), Lva (Larval arrest), Ste (Sterilie), Emb (Embryonic lethal), and WT (Wild type).
[a]*lsm* genes whose RNAi knockdown produces viable phenotypes.

We observed that *lsm* gene knockdown by RNAi during two subsequent generations produces a general enhancement of the phenotypes compared to a single generation treatment, suggesting a strongest gene inactivation after two generations (Table 2).

Considering the resulting phenotypes after exposure to RNAi for two generations, *lsm* family members can be classified in two different categories: essential and non-essential genes.

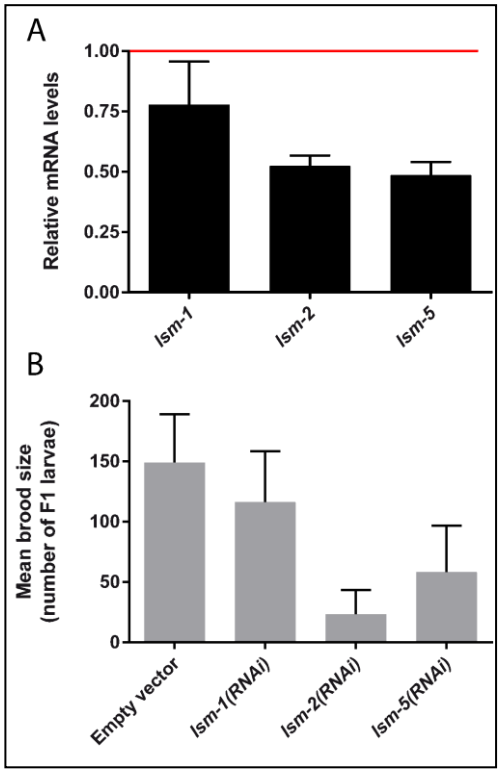


Figure 17. RNAi by feeding knockdown of *lsm* genes causes a Rbs phenotype. A) Measure of the mRNA levels of the indicated *lsm* genes upon their corresponding RNAi by quantitative RT-PCR. Synchronized L1 worms were fed with the corresponding RNAi clone for 36h at 25ºC. The fold change is expressed relative to gene expression in control worms fed with the L4440 vector (red line). *act-1* gene was used as internal control. Error bars represent the standard deviation (SD) between three samples from a single RNAi experiment. B) Quantification of the Rbs after RNAi of *lsm* genes at 25ºC (progeny scored from n=15 worms). Error bars represent the SD among individual progenies.

However, heterogeneity in the phenotypes among *lsm* genes can be due to gene-specific differences in the effectiveness of the RNAi treatment. RNAi by microinjection is a method to further enhance the effectiveness of the RNAi. We synthesized dsRNA for every gene of the family that were injected in the germline of young adult N2 worms. The microinjection of *lsm-2, lsm-4, lsm-5, lsm-6, lsm-7* and *lsm-8* dsRNA produced 100% embryonic lethality while the injection of *lsm-1* and *lsm-3* dsRNA did not, confirming an heterogeneous phenotypic signature among genes coding for the LSM canonical components. Additionally, the three genes encoding LSM proteins that are not present in the canonical complexes did not show any obvious phenotype (Table 2).

3.1.2. Lack of functional redundancies between non essential *lsm* genes

In *Arabidopsis thaliana*, functional redundancies between some Lsm paralogues exist (Perea-Resa et al. 2012). Combinatorial RNAi has been proven an effective way to identify synthetic interactions in *C. elegans* and thus identify paralogues sharing a common function (Tischler et al. 2006).

We microinjected different combinations of dsRNAs in young adult worms (Table 3) and scored the phenotypes in the F1 generation.

Strain	Injected dsRNA	Synthetic phenotype?
N2	*lsm-1* + *lsm-3*	No
	lsm-3 + C49H3.4	No
	C49H3.4 + K07A1.15 + Y48G1C.9	No
	lsm-1 + C49H3.4 + K07A1.15 + Y48G1C.9	No
lsm-3(-)	C49H3.4	No

Table 3. Diverse combinations of dsRNA injections from non essential *lsm* genes do not cause synthetic lethal phenotypes. More than 10 animals were injected for each experimental condition. No obvious synthetic phenotypes were observed for different combinations of dsRNA compared to the phenotypes caused by the respective individual dsRNA.

We did not observe any phenotypic differences when compared to injections of single dsRNAs (Table 2). Our combinatorial dsRNA injection experiments suggest that: (i) *lsm-1* do not present redundant functions with *lsm-3,* (ii) the uncharacterized genes C49H3.4, K07A1.15 and Y48G1C.9 do not seem to be redundant in their functions and (iii) none of the uncharacterized *C. elegans lsm* genes act redundantly with *lsm-1*.

Alignments of protein sequences suggest that C49H3.4 is the non essential *lsm* gene with higher similarity to a canonical gene, particularly *lsm-3* (Figure 18). To test whether the viable phenotypes observed after *lsm-3(RNAi)* could be due to a functional redundancy with C49H3.4 we injected C49H3.4 dsRNA in *lsm-3(tm5166)* mutants (see below on section 3.1.3) (Table 3). However

we did not observe an enhanced phenotype suggesting that *lsm-3* and C49H3.4 are not functionally redundant.

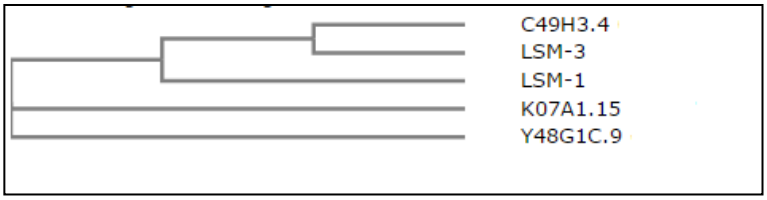


Figure 18. Phylogenetic relationship among non essential LSM proteins. Multiple protein sequence alignment was done using CLUSTAL w2 available at (**http://www.ebi.ac.uk/Tools**) and the result is represented as a cladogram.

3.1.3. Validation of RNAi results using mutant alleles

A robust way to validate any observation generated from an RNAi experiment is the use of mutants. Two homozygous deletion mutants for *lsm-1* and *lsm-3* were available at the Japanese Consortium "National Bioresource Project for the Experimental Animal Nematode *C. elegans* (NBP)" (http://www.shigen.nig.ac.jp/c.elegans/index.jsp). This consortium generates mutations by random mutagenesis using TMP/UV (Trimethyl Psoralen/Ultraviolet) and uses gene-specific primers to identify deletion alleles in the resulting deletion library (Gengyo-Ando and Mitani, 2000).

lsm-1(tm3585) and *lsm-3(tm5166)* strains harbor deletion alleles that affect almost half of the predicted transcripts, disrupting the regions encoding the Sm domain of the LSM-1 and LSM-3 proteins

(Figure 19). Both strains are catalogued as Sterile/Lethal on the NBRP database. After request and genotyping, we observed that these homozygous mutants were viable in our hands, although after three back-crosses with wild-type animals, they displayed obvious phenotypes.

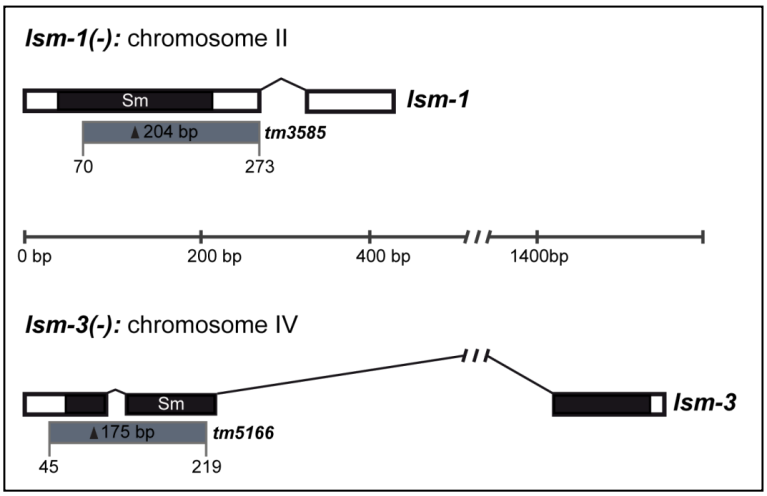


Figure 19. Gene structures of *lsm-1* and *lsm-3.* White boxes represent exons. Connecting lines represent introns. Black areas indicate the regions encoding the conserved Sm domain. Grey boxes represent regions deleted by *lsm-1(tm3585)* and *lsm-3(tm5166)* alleles. The horizontal black bar indicates the base pair scale.

To verify the molecular nature of *lsm-1* and *lsm-3* mutations, we analyzed the transcripts corresponding to *lsm-1(tm3585)* and *lsm-3(tm5166)* by PCR using cDNA as a template (Figure 19). The transcripts corresponding to *lsm-1* and *lsm-3* cDNA were shorter in their respective mutant backgrounds compared to wild type worms. These differences in sequence length correspond to the predicted

deletion sizes (see Figure 3) suggesting that *lsm-1(tm3585)* and *lsm-3(tm5166)* are in frame deletions that produce truncated transcripts.

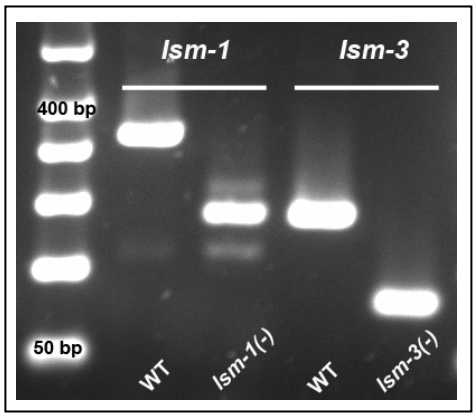


Figure 19. *lsm-1(tm3585)* and *lsm-3(tm5166)* deletion alleles produce truncated transcripts. PCR analysis of *lsm-1* and *lsm-3* specific cDNA sequences synthesized from total RNA extracted from mixed populations of the indicated strains.

3.2. Characterization of *lsm-1* and *lsm-3* mutants

3.2.1 *lsm-1* and *lsm-3* are required for proper development, reproduction and motility

lsm-1(tm3585) and *lsm-3(tm5166)* mutant strains are viable but display pleiotropic phenotypes whose penetrance is influenced by temperature. Both mutations affect embryonic and larval development causing low incidence of embryonic lethality and larval arrest. We also observed a low percentage of lethality in adult worms (Table 4).

Strains	%Emb	%Lva	%Let
lsm-1(-)	19.1 ± 19.8 (n=511)	9.4 ± 1.5 (n=2035)	8.1 ± 0.9 (n=1340)
lsm-3(-)	8.4 ± 5.8 (n=845)	12.5 ± 0 (n=1912)	5.7 ± 1.6 (n=1911)

Table 4. Phenotypes caused by mutations in *lsm-1* and *lsm-3*
Synchronized L1 worms were seeded on NGM plates and phenotypes scored after 48 hours at 20ºC. Emb: Embryonic lethality, Lva: Larval arrest and Let: Adult lethality. Mean values and ± standard deviations are shown. n represent the total number of eggs or individuals scored for each phenotype.

Moreover *lsm-1(tm3585)* and *lsm-3(tm5166)* exhibit a reduced locomotor activity (Figure 20A), are smaller than wild type worms (Figure 20B) and present a reduced brood size phenotype at 15ºC that increases its severity with temperature (Figure 20C).

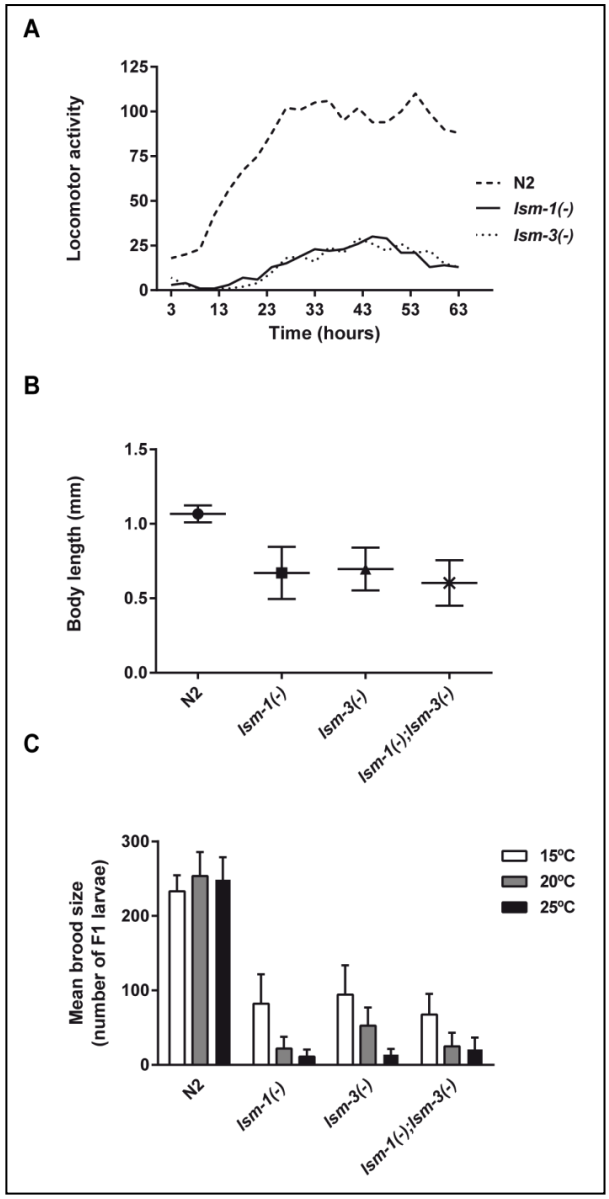


Figure 20. Characterization of *lsm-1* and *lsm-3* mutants. A) Measure of animal motility using an automated locomotor tracking system (Simonetta & Golombek 2007). Locomotor activity plots showing the mean activity of L4 worms at 20ºC (measured in bins/hour of three hours intervals). B) Graph representing the body length of young adult worms (after 48 hours at 25ºC from L1 stage). Mean and standard deviations are plotted (n≥100 worms per genotype tested). C) Brood sizes of wild type (N2) and the indicated mutant strains. Error bars represent the standard deviation among individuals (n≥20).

Given that (i) *lsm-1(tm3585)* phenocopies *lsm-3(tm5166)* (suggesting that they might participate in the same cellular process) and ii- *lsm-1* and *lsm-3* are individually dispensable for the viability of the worms, we made the double homozygous *lsm-1(tm3585);lsm-3(tm5166)* mutant strain to evaluate a possible functional redundancy between these two genes. The lack of a synthetic phenotype (Figure 20B-C), as it was observed by double dsRNA injection (see Table 3), suggests that these genes are in the same genetic pathway during reproduction and development.

3.2.2. *lsm-1(tm3585)* and *lsm-3(tm5166)* are null alleles

Although the extent of the deletions in *lsm-1* and *lsm-3*, including a disruption of the Sm domain, suggest that these deletion alleles are null alleles, we further investigated this point. We microinjected dsRNA corresponding to *lsm-1* and *lsm-3* in their respective mutant backgrounds as an approach to evaluate if these alleles keep some functionality.

These RNAi assays did not enhance the phenotype of the mutants, indicating that *lsm-1(tm3585)* and *lsm-3(tm5166)* are probably null alleles (Figure 21).

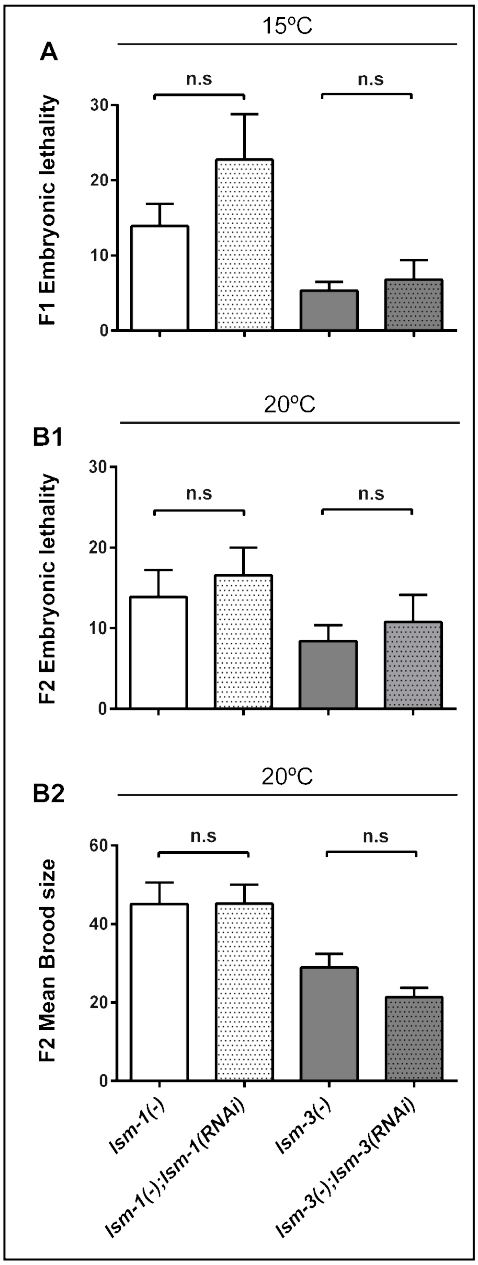


Figure 21. *lsm-1(tm3585)* and *lsm-3(tm5166)* are probably null alleles. A) No significant differences in the mean percentage of F1 embryonic lethality at 15ºC were observed after microinjection (n>10 worms injected, embryos scored n>490, p value > 0,05 in all cases). B) Phenotypic scoring was monitored in the F1 progeny of injected worms ($n \geq 30$) at 20ºC. No significant differences were observed between control and microinjected mutant animals for (B1) embryonic lethality (total embryos scored 1100>n>530), and (B2) brood size (p value > 0,05 in all cases). All the graphs represent the mean value and the standard error of the mean (SEM). *p* values were calculated using a Student-t test.

Given that *lsm-1* and *lsm-3* are involved in RNA metabolism, and the implication of P bodies in RNAi mechanisms (Jagannath & Wood 2009) we tested whether *lsm-1* and *lsm-3* mutants are resistant to RNAi. Thus, we scored the RNAi effect of several genes known to produce somatic and developmental phenotypes. As a positive control we included rrf-*1(pk1417)* mutant worms in this experiment. *rrf-1* mutants are resistant to RNAi in somatic tissues because RRF-1 is implicated in the production of secondary siRNA, a necessary step for a proper function of the RNAi mechanism (Sijen et al., 2001). *prp-8(RNAi)* causes 100% larval arrest in L3 stage in N2, *lsm-1(tm3585)* and *lsm-3(tm5166),* but not in *rrf-1(pk1417)* animals (Figure 22). Additionally after *pos-1(RNAi)* we observed 100% of embryonic lethality in N2, *lsm-1(tm3585)* and *lsm-3(tm5166),* suggesting that *lsm-1* and *lsm-3* mutants are not RNAi resistant in somatic and reproductive tissues.

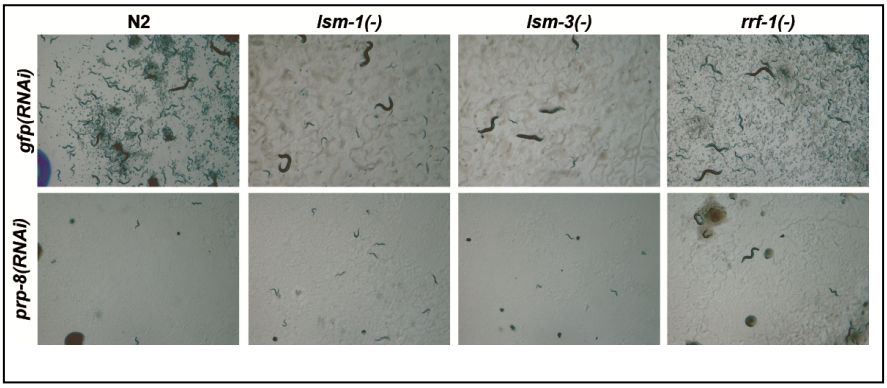


Figure 22. *lsm-1* and *lsm-3* mutants are not RNAi resistant. Synchronized L1 populations from the indicated strains were fed on the respective RNAi plates and phenotypes scored after 72h at 20ºC. Notice that in control conditions *lsm-1* and *lsm-3* mutants present a strong Rbs phenotype. *prp-8(RNAi)* cause a 100% L3 arrest in all strains except in *rrf-1(-)* in which worms can reach adult stages and present protruding vulvas.

Since the delivery of *lsm-1* and *lsm-3* dsRNA into their respective mutant backgrounds does not significantly modify the phenotype of the mutants, we conclude that *lsm-1(tm3585)* and *lsm-3(tm5166)* can be considered as null alleles.

3.2.3. *lsm-1* mutations affect germline development and fertility

One of the most dramatic phenotypes of *lsm-1(tm3585)* mutants is a reduced brood size at the permissive temperature of 15ºC. To get insights on the origins of this phenotype, we first studied the expression pattern of LSM-1 in transgenic worms expressing LSM-1 under the control of a germline specific promoter. Since the GFP expression of this transgenic line is very low at the regular microscope, we amplified the GFP signal by immunofluorescence. Immunostaining of germlines shows a cytoplasmic LSM-1 expression all along the germline, from the proliferative region to oocites (Figure 23). After confirming the expression of LSM-1 in the *C. elegans* germline, we studied the germline proliferation in *lsm-1* mutants and we observed a reduction of germ cells compared to wild type worms during development and in adulthood (Figure 24).

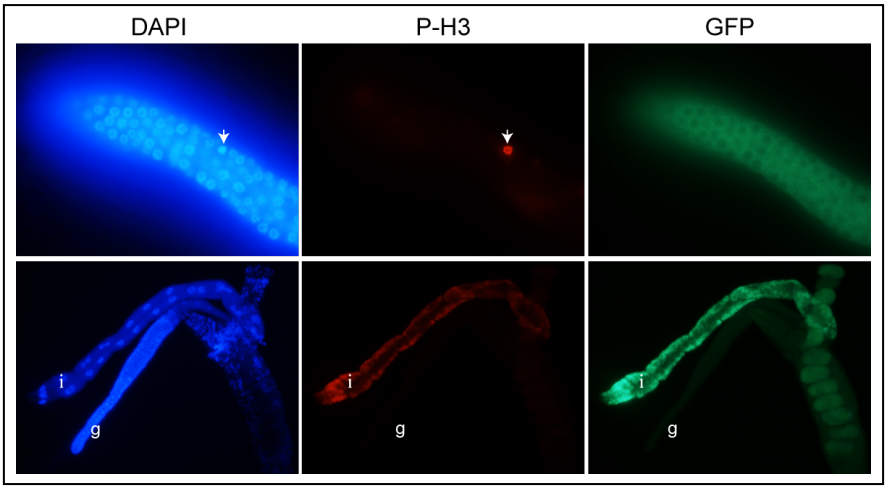


Figure 23. *lsm-1* is expressed in the syncytial gonad of JH1986 worms. Upper panels: Proliferative regions of the germline of JH1986 [pie-1::LSM-1::GFP] adult hermaphrodites immunostained using an antibody against GFP and counterstained with DAPI. As a positive control for the immunostaining, an antibody against phosphorylated histone 3 at serine 10 residue (P-H3) was used as a marker of mitotic cells (white arrow). Lower panels: control for unespecific signal in which no primary antibody was used. "i" indicates the intestine, a tissue typically giving autofluorescence."g" indicates the tissue corresponding to the germline.

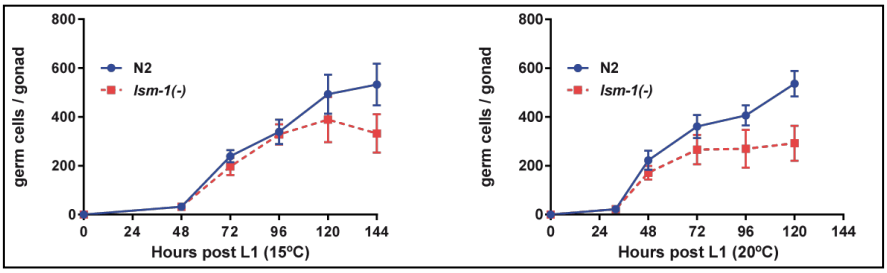


Figure 24. The *lsm-1(tm3585)* allele affects germline development. Graphs represent the mean number of germ cells per gonad at different developmental timepoints of N2 and *lsm-1* mutant worms at two different temperatures. Error bars represent ± standard deviation (SD) among different gonads (n>15).

Finally, we performed a rescue experiment to confirm the causal link between the phenotypes observed and the *lsm-1(tm3585)* mutation. We generated a strain containing integrated copies of a LSM-1::GFP reporter in combination with the *lsm-1(tm3585)* allele. We observed a total rescue of the developmental phenotypes (Larval arrest, Reduced brood size, Adult lethality, Embryonic lethality) of *lsm-1* mutants (Figure 25), indicating that the deletion in *lsm-1* is directly responsible for the observed phenotypes and demonstrating that our integrated LSM-1::GFP reporter is functional.

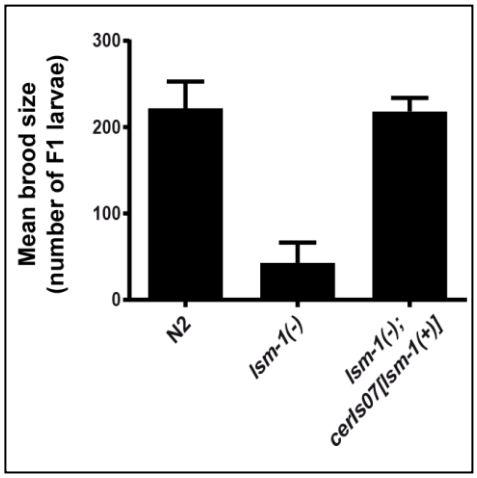


Figure 25. LSM-1::GFP rescues the brood size defect of *lsm-1* mutants. Histograms represent the mean brood sizes at 20°C of the indicated strains. Error bars represent SD among individuals (n>5).

3.2.4. *lsm-3* is not required for constitutive splicing

An interesting observation arising from the analysis of phenotypes of *lsm* family members in *C. elegans* is the viability of *lsm-3* mutants. In yeast, Lsm3 is an essential gene necessary for correct splicing (Mayes et al. 1999). Since spliceosome core components are expected to be essential in *C. elegans* (Kamath et al. 2003a; Kerins et al. 2010; J. F. Rual et al. 2004), the viability of the *lsm-3* mutant argues against a central role of LSM-3 in splicing.

To study the relevance of *lsm-3* in the splicing process, we used a reporter strain for constitutive splicing, expressing an intron-containing GFP transgene in intestinal cells (Figure 26) (MacMorris et al. 2003).

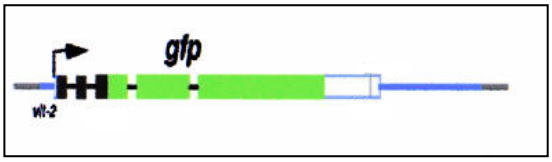


Figure 26. Schematics of the inIs173 [pNvitgfp] transgene. BL3466 worms express integrated copies of a GFP-intron-GFP transgene under the transcriptional control of an intestinal promoter (*vit-2p*). Extracted from (MacMorris et al. 2003).

In this splicing reporter, GFP expression levels are reduced when worms are fed with a *hrp-1(RNAi)* (*hrp-1* encodes a heterogeneous nuclear ribonucleoprotein) (Figure 27). The same effect is also observed in worms treated with *lsm-2(RNAi)*, the *lsm* gene presenting a strongest RNAi phenotype by feeding (see Table 2). A reduction of GFP levels is also observed after *lsm-6(RNAi)* and *lsm-8(RNAi)* treatment for two generations (Figure 27).

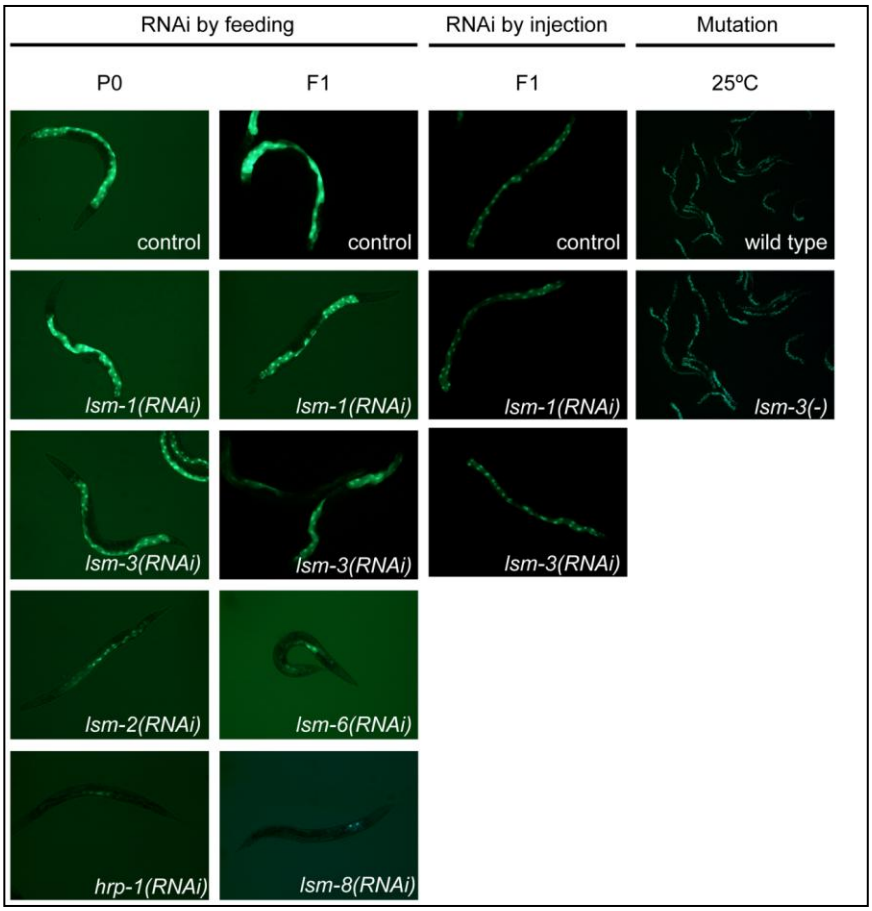


Figure 27. *lsm-3* is not necessary for constitutive splicing in a BL3466 reporter strain. GFP expression in BL3466 adult animals fed or injected with the indicated RNAi, or in the *lsm-3(tm5166)* mutant background.

Interestingly, no alterations in GFP expression were observed in worms fed or injected with *lsm-3* dsRNA or *lsm-1* dsRNA (used as an additional negative control since it is expected to function exclusively in the cytoplasm). To further force the inactivation of *lsm-3,* we generated a strain combining the constitutive splicing reporter and the *lsm-3(tm5166)* deletion allele. These mutant animals display normal GFP expression in the intestinal nuclei (Figure 27), indicating that LSM-3 is not essential for the constitutive splicing of this intron-containing GFP reporter.

Altogether, *lsm-1* and *lsm-3* seem to have similar functions. Since LSM-1 is the only described cytoplasm-specific canonical LSM protein, we decided to study the subcellular localization of LSM-3. We made a construct containing the full LSM-3 ORF under the control of the specific intestinal promoter P*elt-2,* and the *lsm-3* 3'UTR by Gateway cloning. Animals carrying this construct as an extrachromosomal array, express LSM-3 both in the nucleus and the cytoplasm of intestinal cells, as it occurs for transgenic worms expressing fosmid-derived translational reporters for other *lsm* core components such as LSM-4 and LSM-7 (Figure 28).

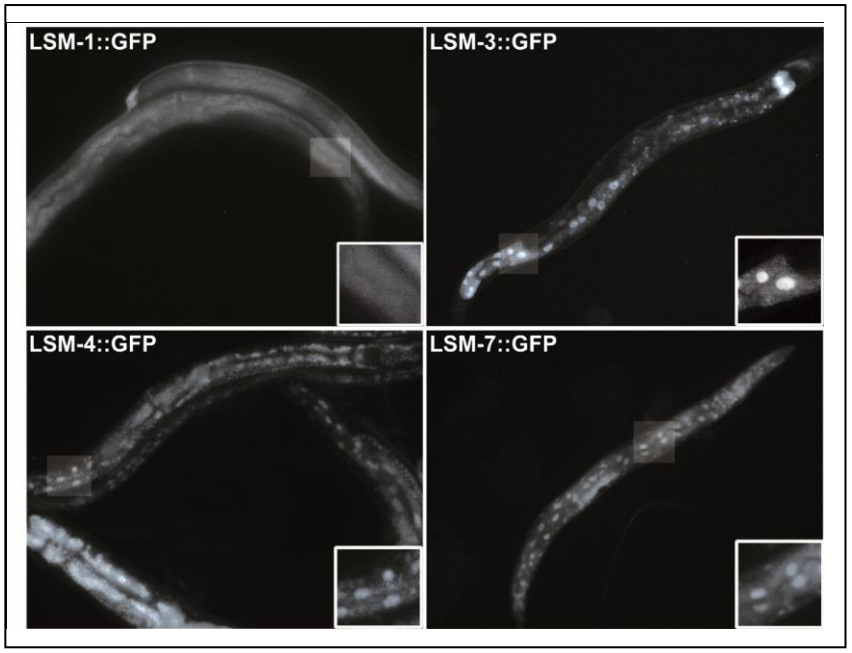


Figure 28. *Pelt*-2::GFP::LSM-3 is expressed in the nucleus and cytoplasm of intestinal cells. Representative images of transgenic worms expressing translational reporters for the indicated genes. Small panels show a magnified view of the highlighted area.

3.3. Expression of *lsm* genes in *C. elegans*

Protein components of the nuclear (LSM2-8) and cytoplasmic (LSM1-7) heteroheptameric rings are suposed to assemble in a 1:1 ratio (Tharun, 2009). However, we observe functional differences among individual members of these complexes and other uncharacterized *lsm* proteins pointing towards a functional versatility among members of the family. In addition to the apparent functional heterogeneity of the *lsm* genes, we wanted to know if differences also exist at the expression level.

The analysis of *lsm* gene structures indicates heterogeneity among family members in terms of regulatory regions because *lsm* genes present distinct UTR sequences and four out of the eleven genes are located in predicted operons (Figure 29). For these genes, only *lsm-1* is annotated as first gene on an operon.

3.3.1. Localizome analysis uncovers differences in expression patterns among *lsm* genes

We produced reporter constructs by PCR fusion (Hobert 2002) to generate promoter::GFP::*unc-54*_3'UTR transgenic lines for all the *lsm* genes (see Annex4 for a detailed annotation of transgenic strains). All *lsm* promoters reported expression except for the K07A1.15 gene. We analyzed fluorescence intensities for several identifiers of the *C. elegans* anatomy and observed overlapping but distinct gene expression

patterns across family members (Figure 30), indicating only a partial co-regulation in the expression of lsm genes. Expression pattern scoring was performed based on GFP fluorescence intensities in different tissues as defined by Wormbase IDs. Tissue expression intensity was classified in color-coded categories ranging from "no expression" to "highest expression". Young adult animals from two independent extrachromosomal transgenic lines were analyzed for each gene (with the exception of *lsm-6*, for which only one transgenic line was studied).

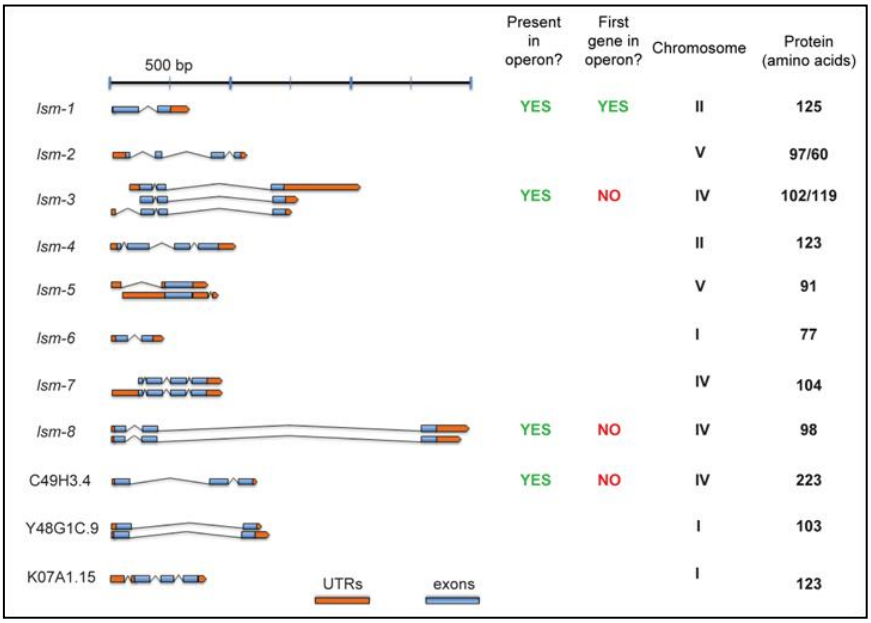


Figure 29. Differences in regulatory sequences among *lsm* genes. *lsm* gene structures. Blue boxes represent exonic sequences. Connecting lines represent introns. Orange boxes represent untranslated regions (UTRs). *lsm* genes have different gene structures and regulatory sequences. They do not form genomic clusters since they are present in different chromosomes and some *lsm* genes are present in independent operons. Association of *lsm* genes to operons was determined in base to a recent revision of *C. elegans* operon annotations (Allen et al. 2011).

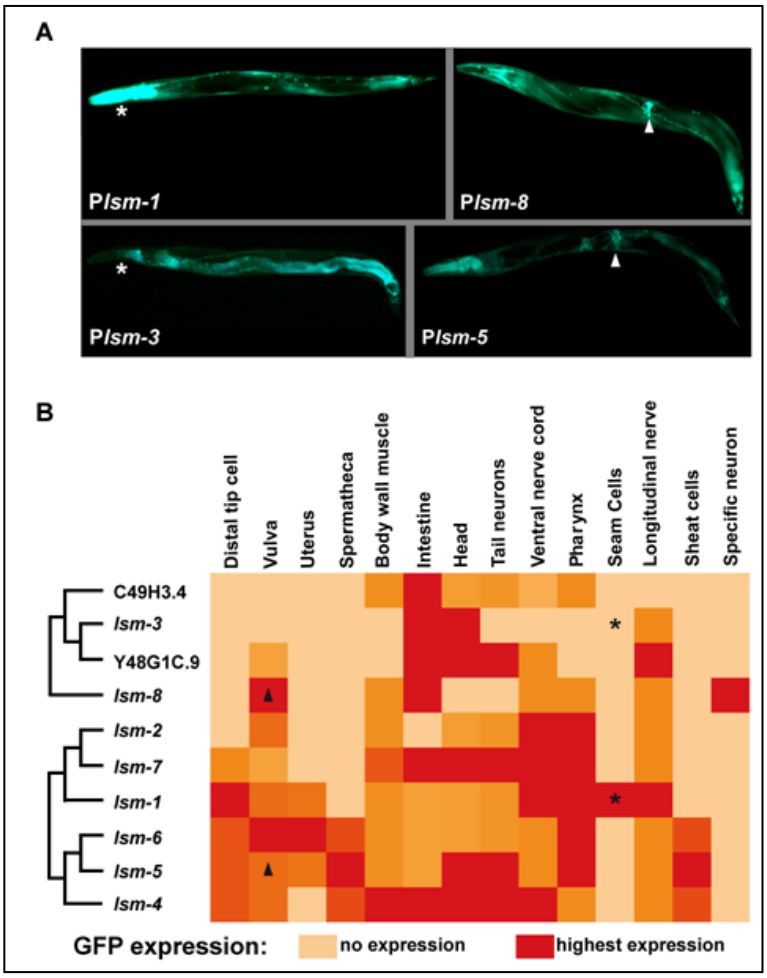


Figure 30. Transcriptional reporters of *lsm* genes display different expression patterns. (A) Representative images showing whole animal fluorescence micrographs for *lsm* transcriptional reporters. (B) Heat map showing the clustering of *lsm* genes based on the expression of the transcriptional reporters. Heat map and clusters were generated using Spotfire software. Δ and * refer to the images in the A and B panels to illustrate the scoring criteria in the indicated tissues (n≥10 worms analyzed for each transgenic line).

This localizome study shows that the expression of *lsm* genes located within different purported operons (as *lsm-3* and *lsm-8*) can be driven by internal promoters. Since both SL1 and SL2 splice leaders have been detected upstream of the messenger RNA (Allen et al. 2011), the activity of these internal promoters may not fully capture the endogenous expression pattern that might resemble the more ubiquitous expression exhibited by the rest of the *lsm* family promoters. For example, we observed that a translational reporter for *lsm-4* in a fosmid context (including UTRs, introns and other regulatory regions) is ubiquitously expressed although with varying levels depending on the cell type (Figure 31).

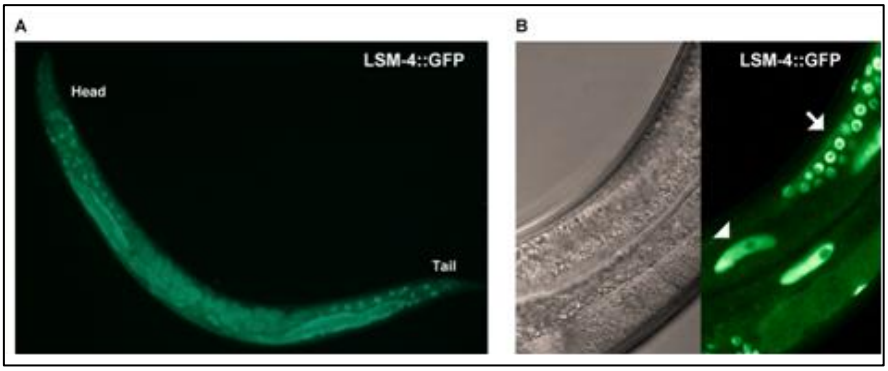


Figure 31. Ubiquitous expression of an *lsm-4* translational reporter. A) Representative image showing a whole animal fluorescence micrograph for *LSM*-4::GFP. B) Confocal image showing in detail the expression of LSM-4 in intestinal (white arrowhead) and germ cells (white arrow).

3.4. *lsm-1* and *lsm-3* promote stress responses through the Insulin/IGF-like signaling (IIS) pathway

The human Lsm1 gene is biomedically relevant because it has been found overexpressed in many cancer types and this altered expression is directly involved in cancer development (Schweinfest et al. 1997; Watson et al. 2008). To get insights on LSM-1 functions in worms, we compared the transcriptomes of synchronized *lsm-1(tm3585)* mutants and wild type animals at the L3 stage (where the germline is not yet fully developed).

3.4.1. Transcriptomic analysis of *lsm-1* mutants uncovers a missregulation of IIS pathway targets

The analysis of RNA-seq data shows a significant gene deregulation in *lsm-1* mutants compared to wild type animals. We found a total of 210 significantly ($p<0.05$) up-regulated and 108 down-regulated genes in *lsm-1(tm3585)* animals compared to wild type (Figure 32).

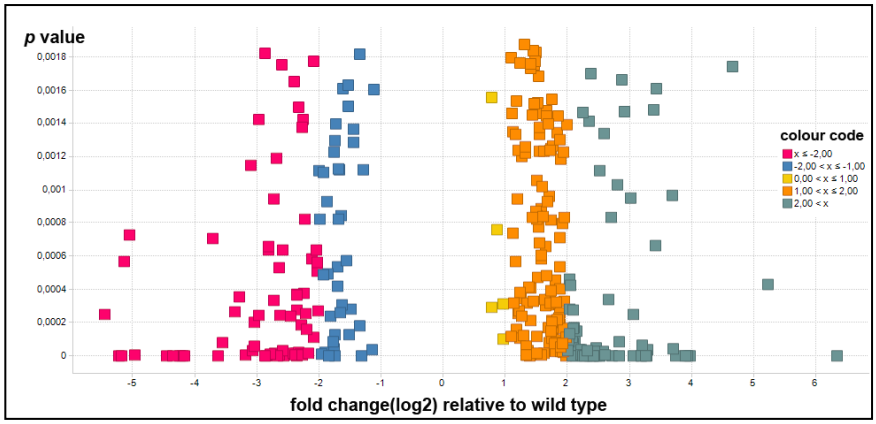


Figure 32. Transcriptomic analysis of *lsm-1* mutants. Volcano plot showing genes significantly deregulated ($p<0.05$) in *lsm-1(tm3585)* compared to wild type (N2). Each dot represents a single gene. Color code represents the indicated fold change scale (pink dots x<-2, grey dots x>2. x represents de log2 fold change in gene expression relative to wild type worms).

We used the DAVID functional annotation tool (Huang, Brad T Sherman, et al. 2009; Huang, Brad T. Sherman, et al. 2009b) to identify significantly enriched annotation terms associated with our gene lists. We used a score p-value ($p\leq0.001$) to analyze protein domains and Gene Ontology (GO) terms statistically more enriched than random among the up regulated (Table 5) and down regulated (Table 6) genes in *lsm-1(tm3585)* mutants. Among the up regulated genes, we found an enrichment of protein domains related with *C. elegans* stress responses as evidenced by the GO terms associated (Table 5). For example, CUB-like domains are found in secreted proteins with induced expression during pathogen infections (Shapira et al. 2006), heat shock proteins are involved in a wide variety of stress responses (Morimoto 1998) and glutathione-S-transferases and thioredoxins are involved in the detoxification of oxidative stress products (van Rossum et al. 2001).

Protein domain	Gene count	P-value
CUB-like region	9	1.2E-7
Alpha crystallin/Heat shock protein	6	5.5E-7
Glutathione S-transferase	8	2.2E-6
Thioredoxin fold	8	6.2E-4
Nematode cuticle collagen	8	1.7E-3
GO term	Gene count	P-value
Response to heat	7	1.2E-6
Aging	15	4.1E-6
Cytoskeleton	10	4.3.E-4
Endoplasmic reticulum unfolded protein response	3	7.6E-3
Response to abiotic stimulus	7	1.1E-3
Response to stress	14	7,7E-5

Table 5. Protein domains and Gene Ontology terms significantly enriched in the dataset for up regulated genes in *lsm-1* mutants.

The interpretation of the analysis of protein domains and GO terms enriched in the down regulated gene list is more complicated due to the unknown functions associated to DUF19 enriched protein domains (Table 6).

Protein domain	Gene count	P-value
DUF19	5	1.2E-6
ABC transporter	3	4.1E-6
GO term	Gene count	P-value
Oxidoreductase activity	10	8.8E-4
Catalytic activity	28	7.8E-3
Membrane	29	8.8.E-2

Table 6. Protein domains and Gene Ontology terms significantly enriched in the dataset for down regulated genes in *lsm-1* mutants.

In addition to functional annotation analysis, one of the most interesting observations in our dataset regards the apparent missregulation of Insulin/IGF-like signaling (IIS) targets in *lsm-1* mutant worms. More in detail, 34 out of 210 significantly up-regulated genes in *lsm-1* mutants have been previously described as IIS pathway targets in different studies (Murphy et al. 2003; McElwee et al. 2003; Liu et al. 2004; Halaschek-Wiener et al. 2005; Oh et al. 2006; Pinkston-Gosse and Kenyon 2007; Lee et al. 2009; Schuster et al. 2010).

Moreover, among the top 24 upregulated genes (fold change >3, electronic Table S3) in *lsm-1(tm3585)* mutants we found 4 genes that are normally downregulated by the IIS pathway effector DAF-16 (C32H11.4, C32H11.9, *dod-21* and *dod-24*) (Murphy, S. a McCarroll, et al. 2003). On the contrary, *pud-2.1* and *pud-2.2*, which are upregulated in *daf-2* mutants (Dong et al. 2007), are

among the top 17 downregulated genes (fold change <-3, electronic Table S3) in *lsm-1(tm3585)* mutants.

Therefore, our transcriptomic data points towards a deregulation of stress-related transcripts and a DAF-16 deficiency in *lsm-1(tm3585)* mutants in normal growth conditions.

3.4.2. *lsm-1* and *lsm-3* promote the nuclear translocation of the DAF-16/FOXO transcription factor during heat stress

To further investigate the functional relationship between LSM-1 and the IIS pathway, we studied the stress-induced nuclear translocation dynamics of DAF-16 in the *lsm-1(tm3585)* mutant background. In control conditions, DAF-16 is ubiquitously expressed in the cytoplasm in both wild type and *lsm-1* mutant animals. We observed that upon thermal stress the translocation of DAF-16 to the nucleus is impaired in *lsm-1(tm3585)* mutants (Figure 33A). More precisely, we observed a delay in DAF-16 nuclear translocation dynamics (Figure 33B), although most of DAF-16 protein is eventually relocated to the nuclei after 60 min at 35ºC. Such effect has previously been observed in mutants of the *C. elegans* sirtuin SIR-2.4-homolog of mammalian SIRT6 and SIRT7 proteins, affecting the correct function of the IIS pathway (Chiang et al. 2012).

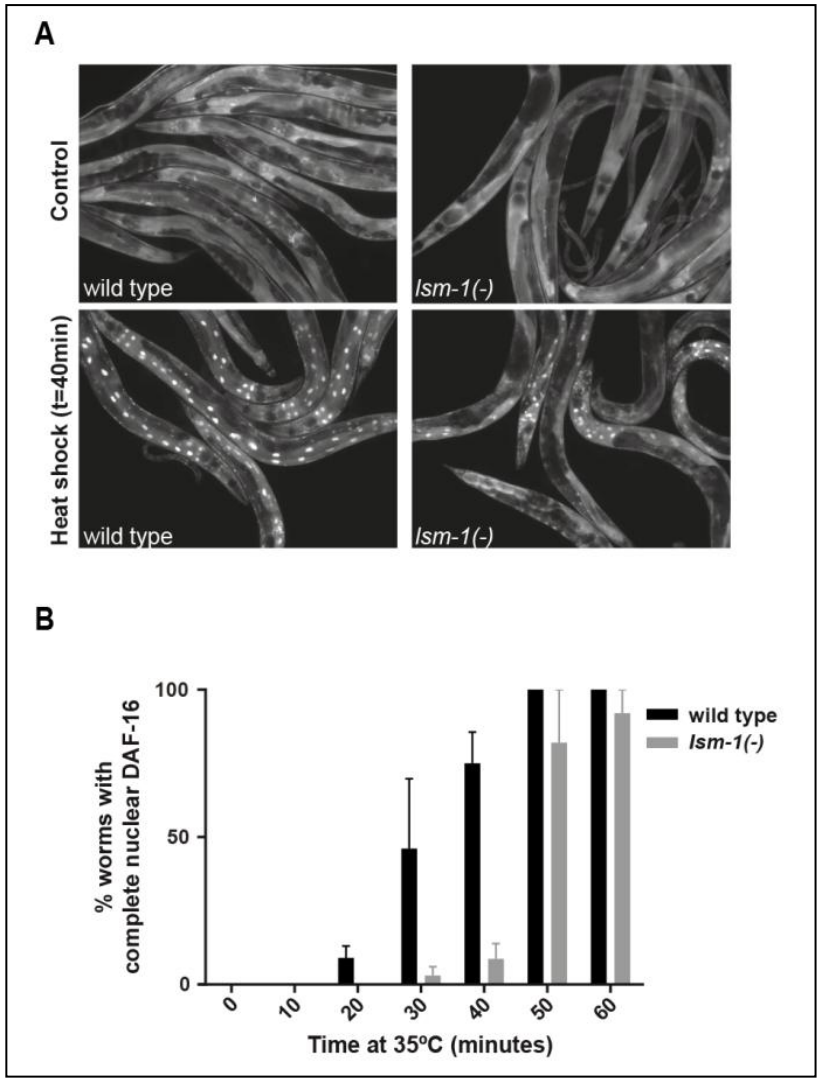


Figure 33. IIS pathway is affected in *lsm-1* mutants. A) Fluorescence images of DAF-16::GFP in wild-type (N2) and *lsm-1* mutant animals before and after a 40 minutes heat-shock at 35°C. B) *lsm-1* affects the nuclear translocation dynamics of DAF-16 during heat-stress. Time-course analyses of DAF-16::GFP nuclear accumulation in response to thermal stress. The histograms represent the average percentages of worms with full nuclear DAF-16::GFP in three independent experiments ($n \geq 50$ per time point). Error bars indicate the standard error of the mean (SEM).

Since *lsm-3* and *lsm-1* mutants are phenocopies, we wanted to assess the effect of mutations in *lsm-3* on the correct signaling through the IIS pathway. However, the DAF-16::GFP reporter used for this experiment and *lsm-3* are located in the same chromosome, making the generation of a genetic hybrid impractical. Thus, to test the impact of *lsm-3* on DAF-16 nuclear re-localization kinetics we injected *lsm-3* dsRNA (using *lsm-1* dsRNA as positive control) in a DAF-16::GFP strain and we scored the phenotype in F1 adult animals. In both RNAi assays we observed a defect in DAF-16 nuclear translocation signal after 15 minutes heat shock at 35ºC (Figure 34).

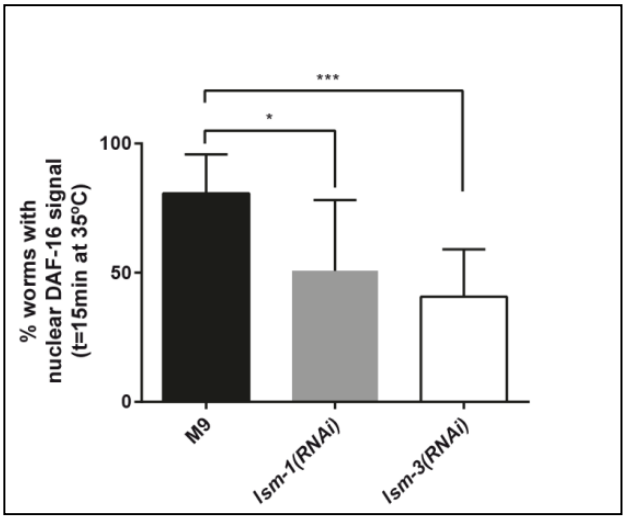


Figure 34. Injection of *lsm-1* and *lsm-3* dsRNA produces a defect in DAF-16::GFP nuclear localization during thermal stress. Error bars represent the SD from seven heat shock experiments (total n=140 for every treatment). *p* values were calculated using a student t- test (**p* value < 0.05, *** *p* value < 0.001).

2.4.3. Mutations in *lsm-1* and *lsm-3* affect stress responses mediated by the IIS pathway

A reduction in DAF-16 nuclear localization is a defect that has been previously associated with a diminished stress resistance (Lin et al. 2001). Accordingly, we found that *lsm-1(tm3585)* and *lsm-3(tm5166)* mutants were sensitive to thermal stress (Figure 35A) while strains overexpressing *lsm-1* displayed a significant thermoresistance, similar to that reported for *daf-2(e1370)* mutants where DAF-16 is constitutively nuclear and transcriptionally active (McColl et al. 2010) (Figure 35B). Furthermore, we also observed that *lsm-1(tm3585)* and *lsm-3(tm5166)* mutant L1 larvae were more sensitive to starvation than wild-type animals, whereas larvae with additional copies of *lsm-1* presented higher survival rate than wild type worms (Figure 35C). These observations suggest that *lsm-1* and *lsm-3* affect the outcome of stress responses involving the IIS pathway, and that *lsm-1* levels influence the robustness of these responses.

The IIS pathway is also involved in host responses to bacterial infection and longevity (Evans et al. 2008). We observed that mutations in *lsm-1* and *lsm-3* cause heightened sensitivity to infection by diverse bacterial pathogens (Figure 36) and have negative effects on lifespan (Figure 37). However, in these cases ectopic *lsm-1* expression does not significantly protect from pathogens or extend lifespan (Figures 36 and 37).

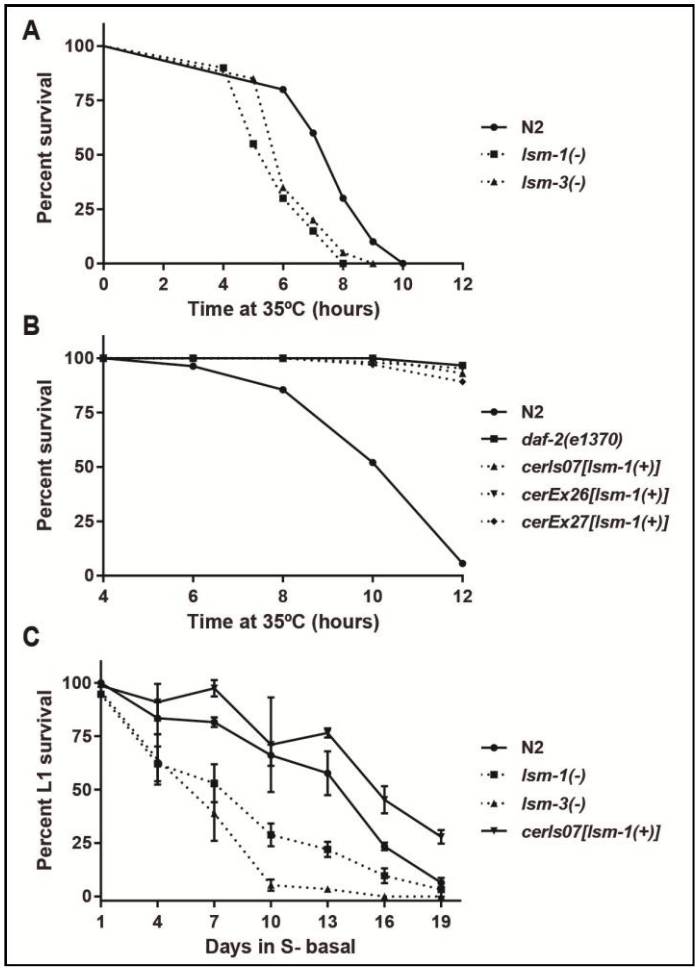


Figure 35. *lsm-1* and *lsm-3* mutants are sensitive to heat stress and starvation. A) Representative survival curves (see Annex 1A for the complete set of experiments) showing the significant reduced resistance to thermal stress caused by mutations in *lsm-1* and *lsm-3* (*p* value<0,001). Young adult worms were incubated at 35ºC and survival was scored every one or two hours. B) Representative survival curves (see Annex 1A for the complete set of experiments) showing the significant resistance to thermal stress caused by three distinct strains overexpressing *lsm-1* (*p* value<0,001). *daf-2(e1370)* mutants were used as positive control for heat stress resistance. C) *lsm-1* and *lsm-3* mutations reduce the survival of L1 larvae during starvation at 20^{o}C. Ectopic expression of *lsm-1* through the integrated reporter LSM-1::GFP displayed a significant survival to starvation compared to wild type (N2) animals after 19 days (*p* value<0,001). Graph represents the mean percentage of survival, and error bars the standard deviations from three different experimental replicates (n ≥100 for each replicate and timepoint). *p* values were calculated by a log rank Mantel-Cox test.

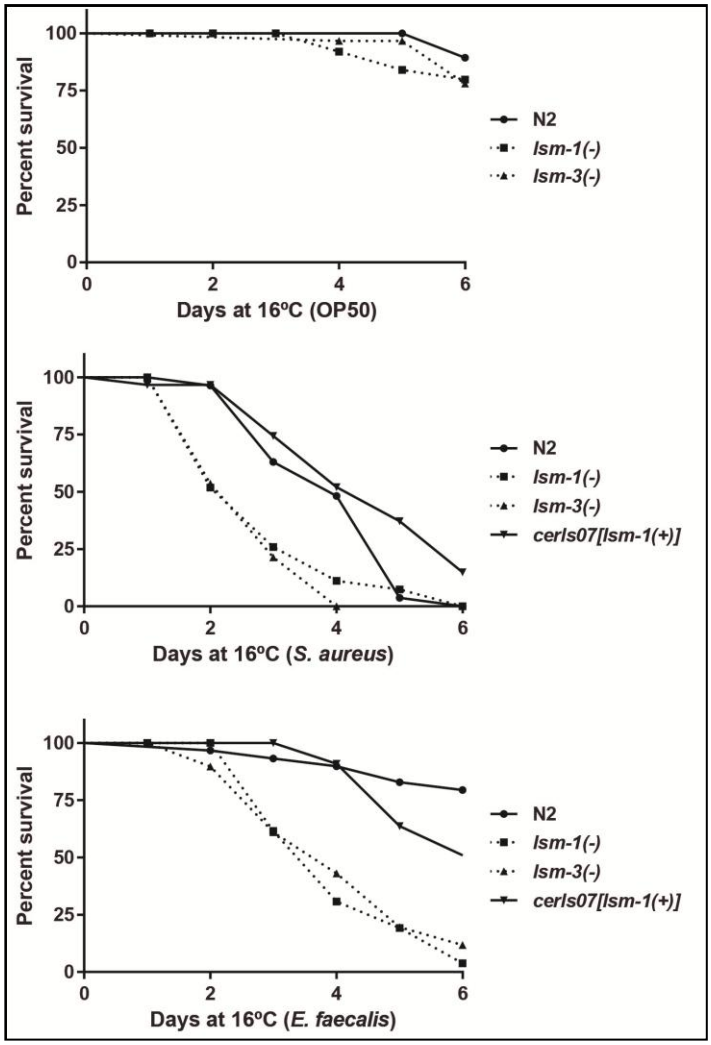


Figure 36. Mutations in *lsm-1* and *lsm-3* cause hypersensitivity to pathogen infection. A) Representative survival curves (see Annex 1B for the complete set of experiments) of the indicated strains on *E. coli* OP50, *S. aureus* and *E. faecalis* (n=30). No significant differences were observed after 6 days in *E.coli* OP50 strain between *lsm-1* and *lsm-3* mutants compared to wild type (N2) animals. Differences between *lsm-1* and *lsm-3* mutants were significant compared to wild type (N2) after pathogen infection ($0{,}01 < p$ value $< 0{,}001$). No significant differences were observed between the survival of wild type and *cerIS07* animals on different pathogens after 12 days.

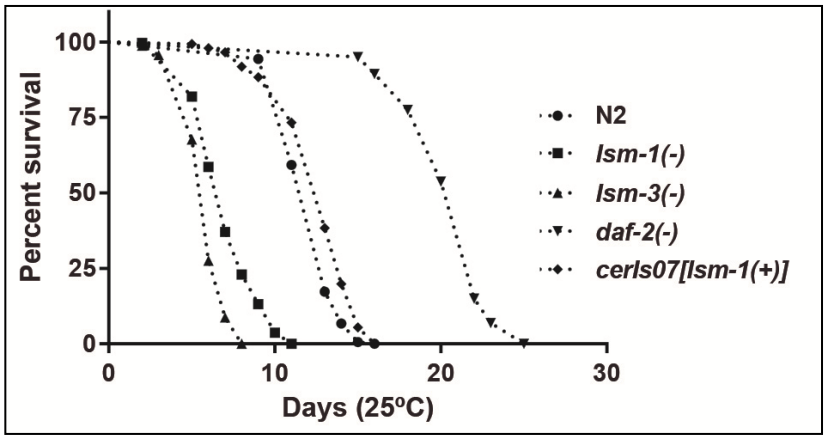


Figure 37. Mutations in *lsm-1* and *lsm-3* affect lifespan. Representative graph showing lifespan analysis of the indicated strains (see Annex 1C for the complete set of experiments). *daf-2(m577)* mutant was used as a control for extended lifespan.

Genetic screens for mutants hypersensitive to pathogen infection have shown that the insulin pathway acts in association with other stress related pathways in the control of innate immune responses in *C. elegans* (Troemel et al. 2006; Singh & Aballay 2009; Coleen T. Murphy and Patrick J. Hu. 2013). One of these pathways is the p38 mitogen-activated protein kinase (MAPK) pathway which has conserved roles in the cellular adaptation and response to a wide variety of stresses form yeast to humans (Sheikh-Hamad & Gustin 2004). In *C. elegans*, the p38 homolog is PMK-1, whose activity is critical for immune protection during development (Kim et al. 2002; Shivers et al. 2010).

In order to assess the integrity of the p38 MAPK signaling pathway, we measured the levels of active (phosphorilated) form of PMK-1

in *lsm-1* and *lsm-3* mutants and we observed a decrease in abundance of PMK-1 activation in larval development compared to wild type animals (Figure 38).

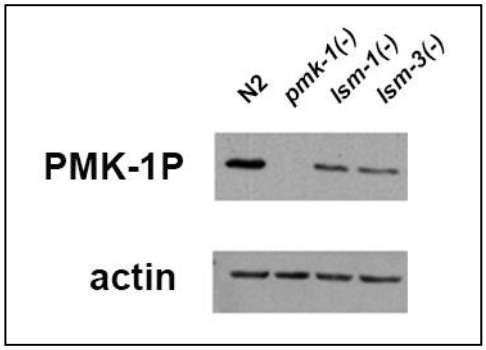


Figure 38. Mutations in *lsm-1* and *lsm-3* affect PMK-1 activity. Immunoblot of active PMK-1. Total protein isolated from L3 synchronized worms of the indicated strains grown at 25ºC was separated by SDS-PAGE and immunoblots were performed to detect the doubly phosphorilated form of PMK-1 (PMK-1P) and actin.

A general decline in the PMK-1 activation is associated with a decrease in the capacity of the organisms to face pathogen infections, a defect also observed in aging worms (Youngman et al. 2011).

In summary, these experiments show the requirements of *lsm-1* and *lsm-3* for heat stress, starvation survival, pathogen infection resistance and normal lifespan. Additionally, the ectopic expression of *lsm-1* only seems to confer enhanced resistance to heat stress and starvation but not to pathogen infection and does not increase lifespan in normal growth conditions.

Regarding pathogen infection, it has been proposed that although IIS pathway regulates the transcriptional response through the DAF-2/DAF-16 axis as part of a basal stress response, other pathways such as PMK-1 are specific and indispensable for immune responses (Troemel et al. 2006). In this context, we observed a constitutive reduction of active PMK-1 in *lsm-1* and *lsm-3* mutants.

Mutations in components of the insulin pathway can affect the overall lifespan of the organism by influencing IIS levels. Regarding the effect of *lsm-1* ectopic expression on lifespan, the assays were performed at 25ºC a temperature considered physiological for wild-type animals.

Altogether our results suggest that mutations in *lsm-1* and *lsm-3* affect stress responses known to be mediated in part by the IIS pathway and can also influence signaling through other stress pathways. Additionally the ectopic expression of *lsm-1* only seems to confer enhanced under acute stress conditions.

3.5. LSM proteins accumulate in cytoplasmic foci at specific stages and environmental conditions

Due to the apparent dual requirement of *lsm-1* and *lsm-3* during development and stress conditions, we decided to study more in detail the subcellular localization of the LSM proteins in different developmental and environmental conditions.

3.5.1. Distinct distribution of cytoplasmic LSM proteins in embryonic and postembryonic cells

To study the subcellular localization of LSM proteins we used two integrated transgenic reporter lines for LSM-1 and LSM-4, allowing the tracking of cytoplasmic and nuclear LSM complexes in the germline and the early embryo. In embryos, according to previous localization studies in other systems (Ingelfinger et al. 2002; Spiller et al. 2007), LSM-1 localizes exclusively in the cytoplasm (Figure 39-40A) whereas LSM-4 is expressed both in nucleus and cytoplasm (Figure 40A).

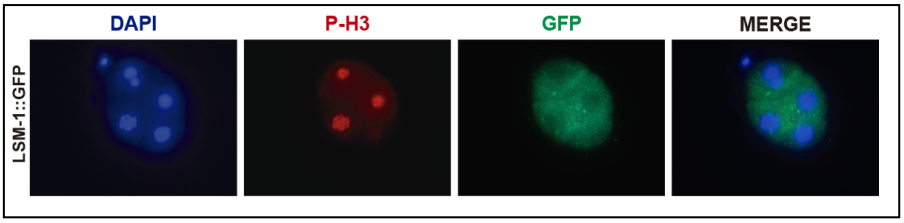


Figure 39. LSM-1 is expressed in the cytoplasm of embryonic cells. Microscope images of LSM-1::GFP transgenic animals immunostained using an anti GFP antibody and counterstained with DAPI. An antibody against phoshphorilated histone 3 at serine 10 residue was used as a positive control for immunostaining.

As previously described in the context of the *C. elegans* embryo, we also observed the constitutive accumulation of LSM proteins in cytoplasmic foci in somatic blastomeres (Gallo et al. 2008) (Figure 40A). When we look at post embryonic tissues, in young adult animals, the distribution of cytoplasmic LSM proteins is diffuse rather than located in aggregates (Fig 40B).

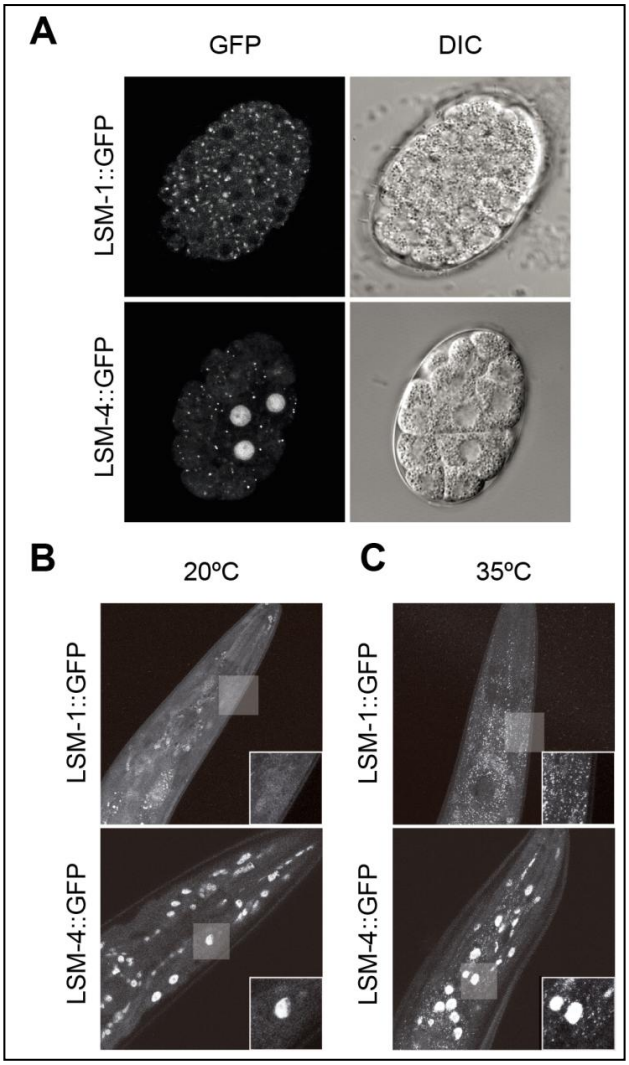


Figure 40. LSM proteins accumulate in cytoplasmic foci under specific developmental and environmental conditions. A) Confocal images of embryos at 20ºC showing LSM-1::GFP and LSM-4::GFP in cytoplasmic foci. Right panels show the same embryos visualized under Differential Interference Contrast (DIC). B) Confocal images (z hyperstack) of the head of young adult animals expressing LSM-1::GFP and LSM-4::GFP at 20ºC. C) Heat stress (1 hour at 35ºC) in young adult animals causes the cytoplasmic accumulation of LSM-1::GFP and LSM-4::GFP. Small panels show a magnified view of the highlighted area.

3.5.2. Redistribution of cytoplasmic LSM proteins during heat stress

In yeast and humans, the LSM1-7 complex has been associated with PBs (Kedersha & Anderson 2007). Interestingly when we applied a thermal stress, adult animals expressing LSM-1::GFP or LSM-4::GFP showed cytoplasmic GFP aggregates (Figure 40C). These stress-dependent LSM cytoplasmic foci appeared at all cell types and were reproducibly visible at the regular and confocal microscope (see Annex 2). Our previous observations show that *lsm-1* and *lsm-3* mutants are required for the correct thermal stress response. Microscopic analysis of the LSM-4::GFP reporter in *lsm-1* and *lsm-3* mutants indicate that the formation of cytoplasmic LSM-4::GFP granules during heat-stress is also dependent on the presence of LSM-1 and LSM-3 (Figure 41). Moreover, the expression of LSM-1::GFP in a *lsm-1* mutant background, in addition to rescue all the phenotypes caused by *lsm-1* (see Figure 11) also recapitulates the LSM-1 cytoplasmic aggregation after heat stress (Figure 42).

These observations suggest that *lsm-1* and *lsm-3* are required for correct stress related responses, in part, through the accumulation in cytoplasmic LSM granules.

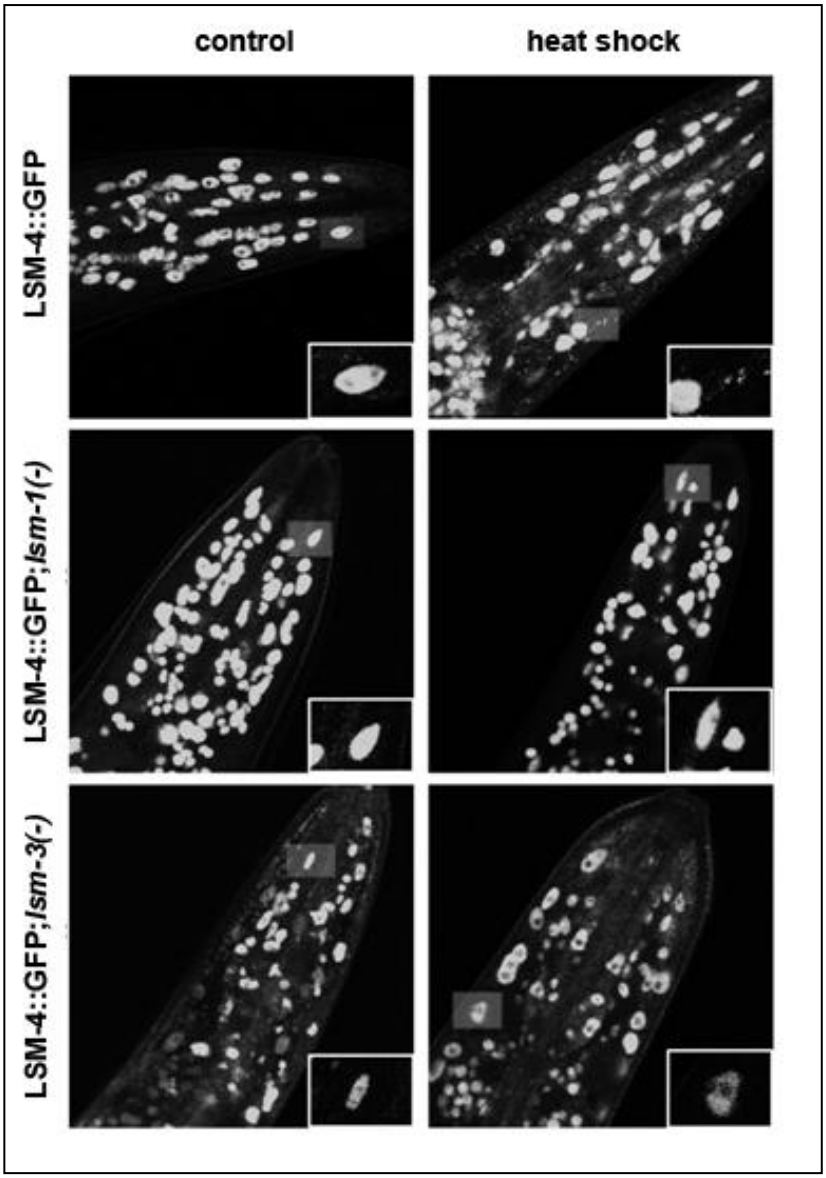


Figure 41. Mutations in *lsm-1* and *lsm-3* abbrogate LSM-4::GFP cytoplasmic foci formation upon heat stress. Confocal images of LSM-4::GFP in different genetic backgrounds at 20ºC (control), or after 1 hour at 35ºC (heat-shock). Small panels show a magnified view of the highlighted area.

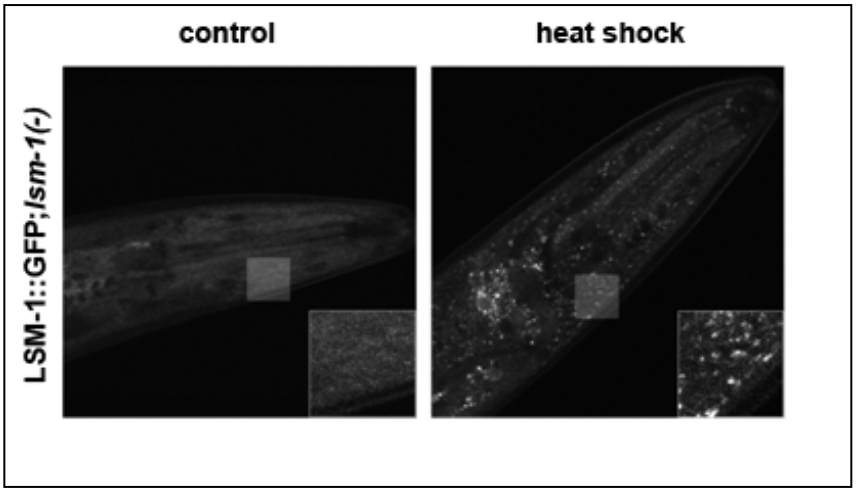


Figure 42. Figure 27. LSM-1::GFP expression recapitulates LSM-1 foci formation after heat stress in a *lsm-1* mutant background. Confocal images of LSM-1::GFP in a *lsm-1(tm3585)* background at 20ºC (control), or after 1 hour at 35ºC (heat-shock). Small panels show a magnified view of the highlighted area.

3.5.3. LSM cytoplasmic proteins co-localize with Stress Granules during heat stress

It has been described that distinct cytoplasmic aggregates can share many RNA binding proteins and, depending on cellular conditions, these factors can relocate from one type of granule to another (Buchan 2014). Since we observed LSM-1::GFP aggregates only under specific stress conditions in post embryonic tissues, LSM-1 may shuttle between PBs and other cytoplasmic granules, such as Stress Granules (SG), depending on physiological or environmental conditions. To confirm this hypothesis, we transformed LSM-1::GFP transgenic worms with an RFP-tagged TIAR-1 (TIA-1-related) reporter, which is a RNA-binding protein described as a SG component (Kedersha et al. 1999). We observed a partial co-

localization of LSM-1::GFP and TIAR-1::RFP aggregates under heat stress (Figure 43) in transgenic worms expressing red fluorescent TIAR-1 as an extrachromosomal array, suggesting that LSM-1 can also be present in SGs under specific conditions.

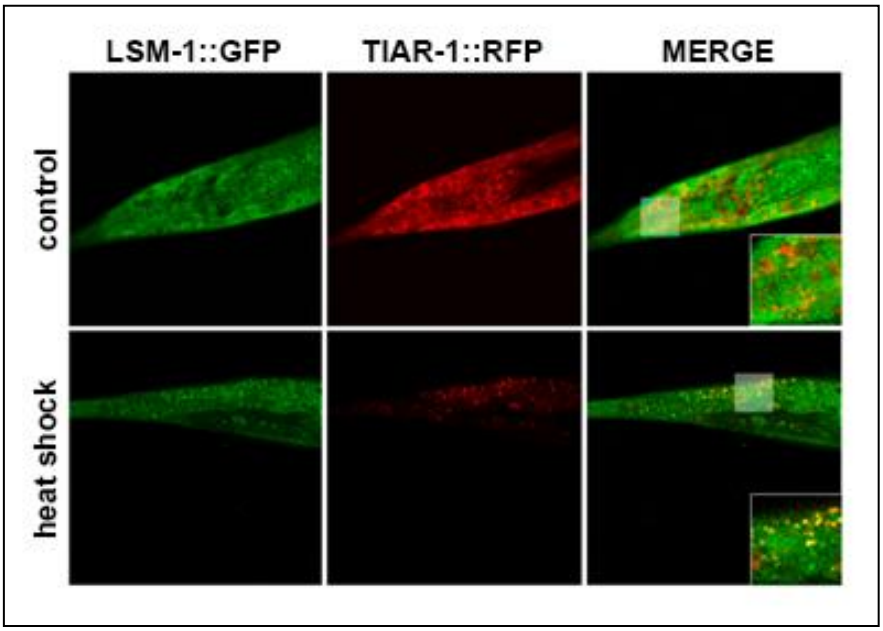


Figure 43. LSM-1 granules co-localize with TIAR-1 stress granules under heat-stress conditions. Confocal images of the tail of L4 worms co-expressing LSM-1::GFP and RFP::TIAR-1 at control conditions (20ºC), or after a heat-shock (1 hour at 35ºC). Small panels show a magnified view of the highlighted area.

3.6. Functional interactions between LSM and IIS pathway components

Our studies support the idea that *lsm-1* and *lsm-3* could contribute to stress responses through the IIS pathway in *C. elegans* by affecting the DAF-16 nuclear localization and the cytoplasmic accumulation of LSM proteins. To further study the possible functional link existing between LSM proteins and the IIS pathway we performed genetic experiments with different LSM and IIS pathway components.

3.6.1. *lsm-1* is required for the *daf-2(m577)* dauer constitutive and extended lifespan phenotypes

Previous work has shown that mutations in the *C. elegans* ortholog of the human insulin receptor *daf-2* gene result in a variety of phenotypes related with alterations in metabolism and lifespan extension. The temperature sensitive *daf-2* mutants have been classified in two overlapping classes (class I and class II) based on the strength and pleiotropy of their phenotypes; class II alleles generally presenting more pleiotropic and stronger phenotypes than class I (Gems et al. 1998).

In order to study the interaction between *lsm* genes and the insulin pathway, we combined *lsm-1* and *lsm-3* mutations with two different classes of *daf-2* alleles: the class II *daf-2(e1370)* and the

class I *daf-2(m577)*. Double mutants were only viable when *lsm-1* and *lsm-3* mutants were crossed with a class I *daf-2* allele (Table 7).

Genotype	Phenotype
daf-2(e1370);lsm-1(tm3585)	Synthetic lethal
daf-2(e1370);lsm-3(tm5166)	Synthetic lethal
daf-2(m577);lsm-1(tm3585)	Viable
daf-2(m577);lsm-3(tm5166)	Viable

Table 7. *lsm-1* and *lsm-3* mutations combined with *daf-2(m577)* class I and *daf-2(e1370)* class II alleles.

It has been shown a correlation between the reduction of IIS pathway activity caused by class I and class II *daf-2* alleles and the degree of stress resistance and lifespan extension phenotypes. Thus, whereas class II alleles display lower insulin signaling levels and constitutive DAF-16 nuclear activity at 25ºC and 15ºC, class I alleles generally do not (Patel et al. 2008).

One of the best characterized *daf-2* associated phenotypes is the constitutive formation of dauer larvae (Daf-c phenoptype) at the restrictive temperature (25ºC). Dauer larvae are nonfeeding, developmentally arrested larvae, which form in response to reduced food supply, and are long lived relative to the adult (Hu 2007). Another well characterized *daf-2* phenotype is the lifespan extension. All *daf-2* mutant phenotypes, including the Daf-c phenotype have been found to be suppressed by mutations in *daf-16* (Gems et al. 1998).

We observed that *lsm-1(tm3585)* and *lsm-3(tm5166)* are sufficient to strongly suppress the lifespan extension phenotype caused by

daf-2(m577) allele at 25°C (Figure 44A), suggesting that the *daf-2(m577)* extended lifespan requires *lsm-1* and *lsm-3*. Moreover, the short lifespan caused by the null *daf-16(mu86)* mutation is significantly shortened by the *lsm-1(tm3585)* mutation (Figure 44B), suggesting that the harmful effect of the *lsm-1* mutation on longevity is independent on the regulation of DAF-16 activity.

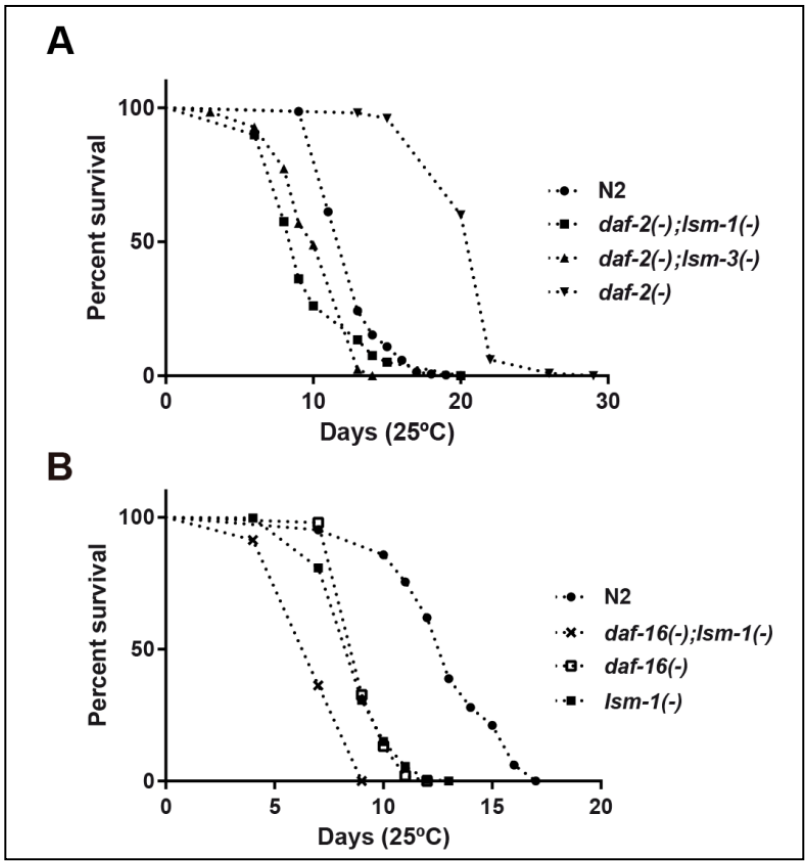


Figure 44. The extended lifespan of *daf-2(m577)* requires functional *lsm-1* and *lsm-3*. A) Representative graphs showing lifespan analysis (see Annex 1C for additional lifespan experiments) of wild type (N2), and the combination of *lsm-1(tm3585)* and *lsm-3(tm5166)* mutations with a *daf-2(m577)* allele. B) Increased effect on lifespan shortening by combining *lsm-1(tm3585)* and *daf-16(mu86)* alleles ((see Annex 1C for additional lifespan experiments)). *daf-2(m577)* and *daf-16(mu86)* mutants were used as control for extended and shortened lifespan respectively.

We also observed that a mutation in *lsm-1* but not in *lsm-3* is sufficient to strongly reduce the dauer constitutive formation caused by *daf-2(m577)* allele at 25°C (Figure 45).

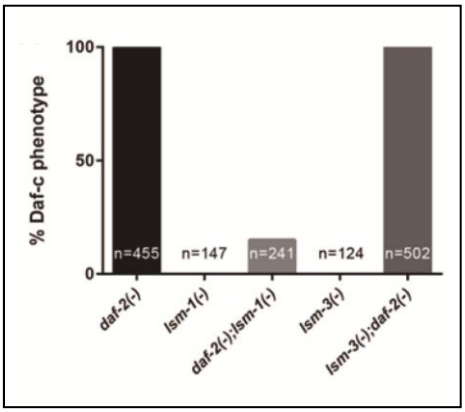


Figure 45. Mutations in *lsm-1* but not *lsm-3* rescue the *daf-2(m577)* dauer constitutive phenotype. Percentage of dauer larva scored for the indicated mutant strains after 55 hours at 25ºC post L1 stage. n represents the total number of individuals scored.

In addition, *lsm-1(tm3585);daf-2(m577)* double mutant larvae present a proportion of adult lethality similar to that observed for individual *lsm-1* mutants at 25°C (Figure 46).

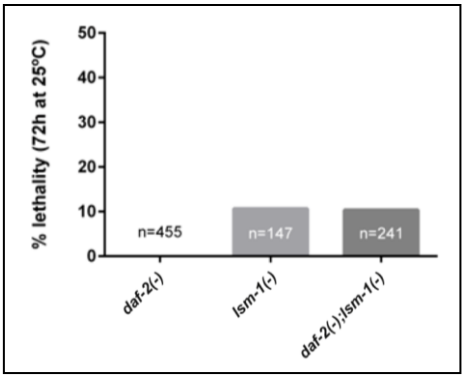


Figure 46. *daf-2(m577)* does not rescue the *lsm-1(tm3585)* adult lethality. Lethality phenotype scoring for the indicated mutant strains after 55 hours at 25ºC post L1 stage. n represents the total number of individuals scored.

The genetic analysis combining *lsm-1* and *lsm-3* with different *daf-2* alleles shows that *lsm-1(tm3585)* and *lsm-3(tm5166)* deletions are not compatible with a class II *daf-2(e1370)* background in which

insulin signaling levels are lower compared to a *daf-2(m577)*. In addition, *lsm-1* and *lsm-3* are required for the extended lifespan of a *daf-2(m577)* classI allele, suggesting that *lsm-1* and *lsm-3* are required at any point in the processes downstream *daf-2* inactivation. Surprisingly, only *lsm-1(tm3585)* but no *lsm-3(tm3585)* can rescue the dauer constitutive phenotype caused by *daf-2(m577)* at 25ºC. This observation can be an argument favoring a stronger interaction between *lsm-1* and the downstream processes derived from *daf-2* inactivation. In addition, it can also been interpreted as a specific interaction between *lsm-1* and additional pathways participating in the formation of dauer larvae.

Regarding the possible interaction of lsm-1 with other pathways contributing to longevity control, the increased lifespan reduction in *daf-16(mu86);lsm-1(tm3585)* double mutants suggests that *lsm-1* and *daf-16* participate in independent pathways influencing longevity.

3.6.2. Effect of IIS levels on the formation of LSM cytoplasmic foci

Since LSM proteins accumulate in cytoplasmic aggregates in a stress-specific manner, and *lsm-1* and *lsm-3* contribute to IIS pathway-related phenotypes, we wanted to study whether the cytoplasmic LSM aggregation mechanism is stress-specific or can be required for the IIS pathway in normal conditions. As already commented in previous sections, the levels of IIS can be

modulated environmentally (i.e. heat-stress) or genetically (i.e mutations in *daf-2),* and the degree of DAF-16 nuclear localization or its transcriptional activity can be used a robust readout of the overall insulin signaling rates (Patel et al. 2008).

To evaluate the influence of the insulin signaling levels on the cytoplasmic LSM aggregation we crossed LSM-1::GFP and LSM-4::GFP transgenic strains with a *daf-2(e1370)* allele. In this *daf-2(e1370)* mutant background, a DAF-16::GFP reporter appears constitutively nuclear at 25ºC compared to a *daf-16* rescued background (Figure 47).

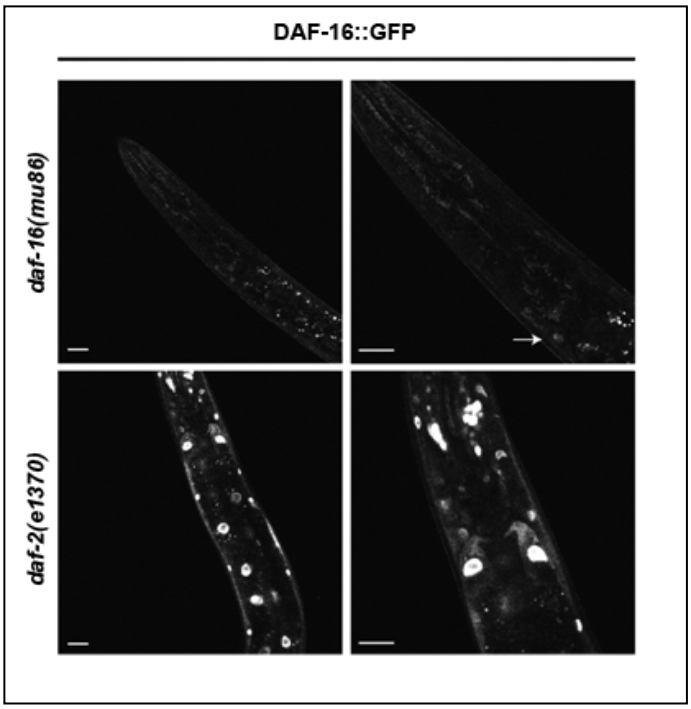


Figure 47. Genetically-induced low IIS causes constitutive DAF-16 nuclear localization. Confocal images of transgenic young adult worms expressing DAF-16::GFP in the indicated mutant backgrounds at 25ºC. Right panels show a higher magnification. White arrow indicates a basal nuclear DAF-16 signal in the rescued strain.

In *daf-2(e1370)* animals, we observed LSM-1 and LSM-4 accumulation in cytoplasmic foci (Figure 48). Although the accurate quantification of LSM cytoplasmic foci is difficult, the qualitative observation of these aggregates suggests that downstream signals coming from the DAF-2 receptor can trigger the onset of cytoplasmic LSM foci.

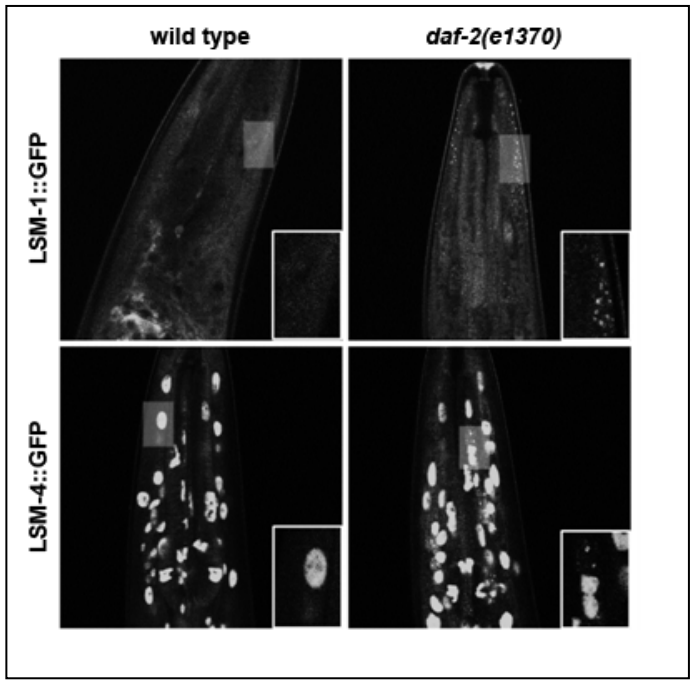


Figure 48. Genetically-induced low IIS causes LSM foci formation. Confocal images of transgenic young adult worms expressing LSM-1::GFP and LSM-4::GFP in the indicated mutant background at 25ºC. Small panels show a magnified view of the highlighted area.

To test whether this foci formation is dependent on DAF-16, we combined LSM-4::GFP transgenic animals with the *daf-16(mu86)* null allele and analyzed the LSM-4 foci formation upon heat stress,

an environmental stress capable of induce DAF-16 nuclear translocation (See Figure 49). We observed the formation of cytoplasmic LSM-4 foci in the absence of DAF-16 activity (Figure 34) suggesting that LSM protein aggregation is stress-dependent and seems to be mediated through DAF-16 independent signals downstream of DAF-2.

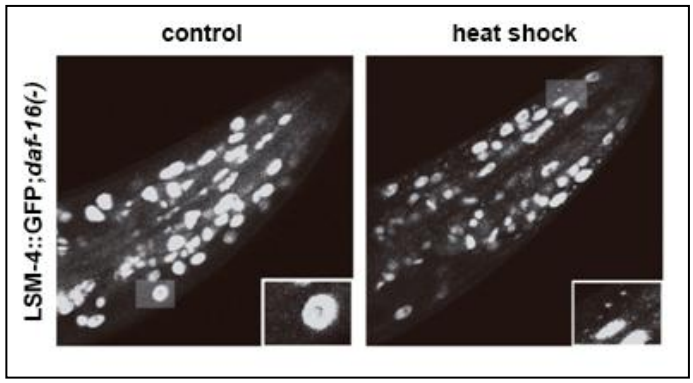


Figure 49. Stress-induced cytoplasmic LSM foci formation does not require DAF-16 activity. Confocal images of transgenic young adult worms expressing LSM-4::GFP in the indicated mutant background at 20ºC (control) or after 1 hour at 35ºC (Heat-shock). Small panels show a magnified view of the highlighted area.

4. DISCUSSION

4.1. Functional diversity among components of heteromeric LSM complexes

Our phenome and localizome studies of the LSM protein family members introduce interesting observations regarding the functional peculiarities of individual components of the two canonical LSM complexes in a multicellular organism. By using two RNAi methods that cause gene silencing at two different levels, we described a heterogeneous phenotypic signature among members of the family. These experiments resulted in the classification of *lsm* genes in two classes, essential and non essential genes. Our observations were further validated by the use of mutants. Moreover, the expression analysis using transcriptional reporters for all the *C. elegans lsm* genes supports the idea that individual LSM proteins may present a wider functional landscape than expected for proteins supposed to function as stoechiometric heterocomplexes.

4.1.1. Functional heterogeneity among proteins that are expected to assemble and function as stoichiometric heterocomplexes

The functional heterogeneity among components of the LSM1-7 and LSM2-8 canonical complexes can be hypothesized looking at public datasets based on high throughput RNAi screenings in *C. elegans*, available at Wormbase (electronic Table S1). One example is the case of *lsm-2* and *lsm-4*, which are the only *lsm* genes found to be genetically interacting with *lin-35*/Rb, a

chromatin remodeler involved in cell cycle control (Ceron et al. 2007). Another example is *lsm-3*, identified as a splicing-related component involved in pathways regulating small RNAs (Tabach et al. 2013).

Gene	Canonical complex	Phenobank dsRNA injection phenotype	This study dsRNA injection phenotype
lsm-1	Cytoplasmic	Viable	Viable
lsm-2	Nuclear/Cytoplasmic	Embryonic lethal	Embryonic lethal
lsm-3	Nuclear/Cytoplasmic	Viable	Viable
lsm-4	Nuclear/Cytoplasmic	Embryonic lethal	Embryonic lethal
lsm-5	Nuclear/Cytoplasmic	Embryonic lethal	Embryonic lethal
lsm-6	Nuclear/Cytoplasmic	Viable	Embryonic lethal
lsm-7	Nuclear/Cytoplasmic	Embryonic lethal	Embryonic lethal
lsm-8	Nuclear/Cytoplasmic	Viable	Embryonic lethal
Y48G1C.9	Not applicable	Viable	Viable
K07A1.15	Not applicable	Not tested	Viable
C49H3.4	Not applicable	Not tested	Viable

Table 8. RNAi by injection phenotypes from two independent studies (data extracted from http://www.worm.mpi-cbg.de/phenobank/)

We performed RNAi by feeding and by injection, and both approaches underscored the distinct phenotypes of *lsm-1* and *lsm-3*. Moreover, our study consolidates the phenotypic information existing for the *lsm* family in *C. elegans*. For example, our work completes the Phenobank data set. The Phenobank database contains information regarding early embryonic and postembryonic

phenotypes caused by dsRNA injection of ~900 essential *C. elegans* genes (~5% of the genome)(Sönnichsen et al. 2005). Moreover, we made new RNAi clones, added phenotypic descriptions for the uncharacterized *lsm* genes, and reported new phenotypes for *lsm-6* and *lsm-8* (Table 8). Our RNAi results are consistent with genetic data obtained in yeast with the exceptions of lsm6 and lsm7 that are dispensable in yeast, and *lsm-3* that is non essential in worms (Mayes et al. 1999; Salgado-Garrido et al. 1999) (electronic Table S2). Therefore, lsm1/*lsm-1* is the only family member that is non-essential in both organisms. Thus, both unicellular and multicellular eukaryotes have members of the heptameric complexes that are more essential than others.

The implications of LSM components in P body formation have been studied in *Drosophila* S2 cell lines, in which the efficient knockdown of Lsm1 or Lsm3 by siRNA prevent the formation of P bodies (Eulalio et al. 2007). Although maybe casual, we wonder if the absence of information for the rest of canonical LSM proteins is because their essential functions (i.e splicing) and therefore because Lsm1 and Lsm3 are the only Lsm genes whose siRNA knockdown produced viable phenotypes in their hands.

Few years ago, the statements about LSM functions were established taking into account biochemical experiments. However, we and others are demonstrating that the combination of biochemistry and genetics should provide a more realistic view of the functional diversity among *lsm* family members.

4.1.2. LSM-3, a spliceosome component out of the splicing party?

As a core component of both the nuclear LSM2-8 and the cytoplasmic LSM1-7 complexes, LSM-3 has been associated to nuclear pre-mRNA splicing and cytoplasmic mRNA decay in yeast (Tharun 2009). Lsm3 protein immunoprecipitates with U6 snRNAs in yeast (Salgado-Garrido et al. 1999) and localizes in the nucleus and cytoplasm in human cells (Ingelfinger et al. 2002). Although yeast lsm3 mutants are non viable (Séraphin 1995), genetic analyses of conditional alleles show defects affecting the correct nuclear splicing and the cytoplasmic decay rates of many RNAs (Mayes et al. 1999; Tharun et al. 2000). Similar results have been observed in a multicellular system such as *Arabidopsis thaliana* (Perea-Resa et al. 2012), suggesting that lsm3 nuclear and cytoplasmic functions are evolutionary conserved.

Our confirmation regarding the dispensability of *lsm-3* for *C. elegans* viability was obtained from the characterization of *lsm-3* deletion mutants lacking a functional Sm domain. No other alleles for canonical *lsm* components were viable in homozygosis except *lsm-1(tm3585)* and *lsm-3(tm5166)*. This fact can be considered as a validation of our RNAi results but still there are questions regarding the viability of *lsm-3* mutants. Since *lsm-3(tm5166)* phenocopies *lsm-1(tm3585)*, a mutation for a cytoplasm-specific LSM protein, we constructed a P*elt-2*::LSM-3::*lsm-3* 3'UTR reporter to study the subcellular localization of LSM-3 in intestinal cells. This transgenic

reported nuclear and cytoplasmic GFP expression of the LSM-3 translational reporter in intestinal cells.

The reasons to use P*elt-2* as a promoter were: (i) intestinal cells are the biggest somatic cells in *C. elegans*, allowing an easier characterization of subcellular localizations and, (ii) we did not observe defects in splicing of an intron-containing GFP reporter in intestinal cells of *lsm-3(tm5166)* mutants. However, the use of a non endogenous promoter could be causing an abnormal nuclear accumulation of LSM-3, since it does not reproduce the basal transcriptional levels of an endogenous *lsm-3* promoter.

In a recent study in *C. elegans*, LSM-3 has been identified along with LSM-2 to LSM-8, and other U4-U6 components, in co-immunoprecipitation experiments with SART-3, an essential nuclear protein involved in U6 snRNA recycling (Rüegger et al. 2015) demonstrating that, to some extent, LSM-3 is present in the nucleus in wild type worms.

The crystal structures of the yeast nuclear and cytoplasmic LSM2-8 and LSM1-7 have been recently described (Sharif & Conti 2013; Zhou et al. 2014) confirming the presence of LSM-3 in both complexes as expected from previous genetic and biochemical characterizations. A recent study in yeast has shown the relevance of Lsm2 and Lsm3 for the interaction of the complex with Pat1 a component required for the decapping function of the LSM1-7 complex in vivo (Wu et al. 2014). This observation demonstrates that interactions occur among individual members of the complex

and specific protein partners, underscoring the distinct "functional weight" of certain LSM proteins in the heptameric complex.

Still, we cannot explain why *lsm-3* is not essential in *C. elegans* after *lsm-3* dsRNA injection in *lsm-3(tm5166)* mutants. One possibility is that a residual functionality of *lsm-3* can be sufficient for animal viability. Recently, CRISPR/Cas9 technology has been successfully applied in *C. elegans* to manipulate the genome and generate mutations *"à la carte"* at specific sites (Waaijers et al. 2013). A definitive step to prove the requirement of *lsm-3* for animal viability would be the application of CRISPR/Cas9 approaches to delete the full *lsm-3* coding sequence.

4.1.3. Specificities in transcription Factors binding to *lsm* promoters

The apparent functional heterogeneity of LSM proteins inferred from the analysis of RNAi and mutant phenotypes is supported by the localizome analysis of promoter-driven GFP reporters. The modENCODE database contains experimental data regarding binding sites for 300 transcription factors (TFs) in *C. elegans* (Gerstein et al. 2010) under specific developmental and environmental conditions. We extracted the information existing for TF binding to *lsm* promoters to draw a functional network based on TF-binding signatures (Figure 50 and Annex 3). We observed that *lsm* genes evidence specificity in binding TFs. This observation is consistent with our phenome and localizome analyses. Some functions of LSM complexes components may require a common subset of TFs. However, specific expression of certain *lsm* genes

influenced by the tissue, the developmental stage, or environmental conditions could exist in a multicellular animal.

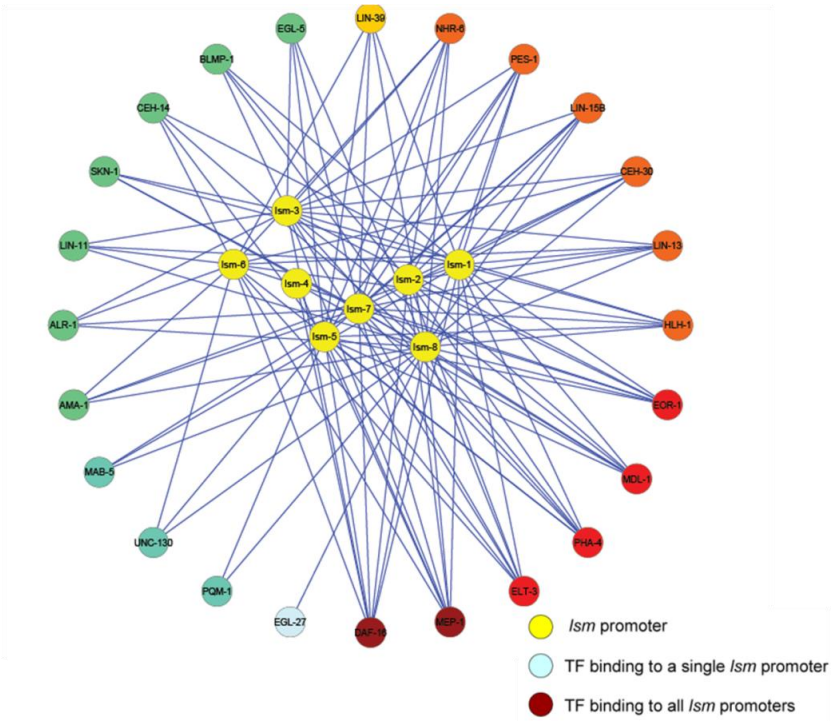


Figure 50. Network based on functional annotations reveal differences on *lsm* genes regulation. Circles represent nodes. Yellow circles represent *lsm* genes and colored circles represent transcription factors (TF). A color scale was assigned to TFs based on the number of interactions. Blue lines represent interactions ($p<10^{-5}$) between nodes, in this case: binding to the promoter region of the gene.

4.1.4. Uncharacterized *lsm* genes

The analysis of LSM protein sequences encoded in the *C. elegans* genome revealed the existence of three uncharacterized *lsm* genes K07A1.15, Y48G1C.9 and C49H3.4. The phylogenetic analysis related K07A1.15 and Y48G1C.9 to human Lsm10 and Lsm9 respectively (Figure R12). C49H3.4 felt in another branch related to human Lsm proteins containing additional domains such as Edc3. None of these human homologs have been described as components of the splicesome. Accordingly, we did not observe any obvious phenotype after RNAi knockdown for these genes.

While Lsm10 protein is a component of the splicing-related U7 snRNP with roles histone 3' end pre-mRNA processing conserved in eukaryotes (Azzouz & Schumperli 2003), yeast Lsm9/Mak31 has been characterized as a subunit of a N^{α}-terminal acetyltransferase (NatC) with roles in protein acetylation (Polevoda & Sherman 2001). We were not able to identify any functional redundancy among these uncharacterized *lsm* genes by doing diverse RNAi by injection experiments. Another explanation for the absence of phenotypes associated to these three genes could be the existence of functional redundancies with other components of unknown processes.

We explored the gene expression of these three genes different at developmental stages in the modENCODE database (Gerstein et al. 2010). In young adult animals, K07A1.15 expression levels appear very low compared to the surrounding genes. Interestingly, we were not able to detect GFP expression in a transcriptional reporter for this gene. Moreover, this gene is located

in a region corresponding to the 3'UTR of *ile-1* a fact that can interfere with its expression.

C49H3.4 is a member of the CEOP4264 operon formed by *ntl-4* and C49H3.4. While *ntl-4* (the first gene of the operon) shows expression in young adult worms, C49H3.4 levels are very low but still visible in our transcriptional reporters.

Finally Y48G1C.9 seems to be predominantly expressed in young adult worms. Thus, the residual expression of K07A1.7 and C49H3.4 in normal growth conditions can be another argument for the absence of phenotypes observed for these proteins.

These observations are also consistent with modENCODE data regarding TF binding to *lsm* promoters (Annex 3). C49H3.4, K07A1.15 and Y48G1C.9 present a low occupancy of TF binding to their promoter regions compared to the rest of the *lsm* family members.

However, since our study highlights a dual role of the canonical LSM proteins in normal growth conditions and stress, an obvious step to dig into the functions of these uncharacterized LSM proteins is the characterization of null mutants to study their functions in diverse enviromental conditions.

4.1.5. LSM complexes evolve towards heteromerization

Members of the Lsm family of proteins have been related with RNA metabolism in all domains of life. They are present in Archaea, Prokaryotes and Eukaryotes. While archaeal and prokaryotic organisms typically contain one or two Lsm homologs that assemble forming homomeric rings, more than 16 Lsm proteins have been described in eukaryotes usually assembling in heteromeric rings. The phylogenetic analysis of Lsm protein sequences show that eukaryotic Lsm proteins derive from prokaryotic ones following a diversification-duplication process (Salgado-Garrido et al. 1999).

One of the main differences between prokaryotic and eukaryotic Lsm proteins resides in their quaternary structure. It has been suggested that compared to the homomeric prokaryotic Lsm ring, a heteromeric complex can provide diverse spatial binding sites along its rotational axis ensuring a precise conformational specificity in RNA and protein binding (Figure 51).

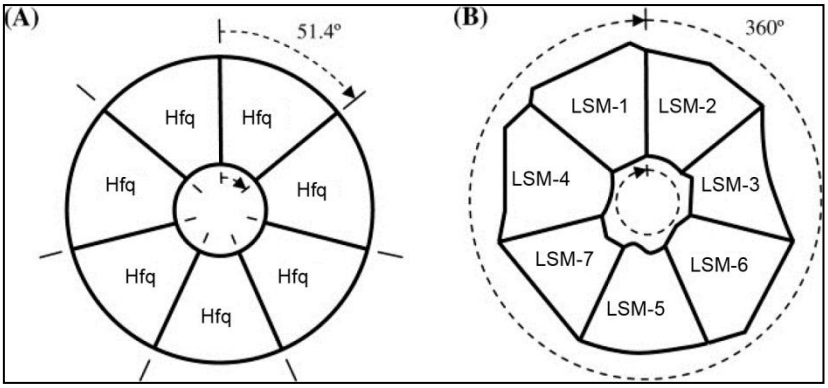


Figure 51. Steric specificity for seven membered Lsm rings. (*A*) Homomorphic ring typical of prokaryotes. (*B*) Heteromorphic ring typical of eukaryotes (modified from Scotfield & Lynch 2008).

In heteromeric rings, different Lsm proteins can evolve and develop specific and functionally divergent side chains around the ring, increasing the combinatiorial possibilities of Lsm proteins in different complexes. As a consequence of the functional versatility offered by the heteromerization, homomeric and heteromeric complexes have been named as *fixed* and *flexible* Lsm complexes respectively (Scofield & Lynch 2008).

4.2. Are cytoplasmic LSM proteins P Bodies or Stress Granules?

Post-transcriptional processes such as translational repression and mRNA degradation play key roles in the regulation of gene expression regulation the functional activity of genes (Wilusz & Wilusz 2004). Proteins involved in these RNA-related processes are often localized in cytoplasmic compartments of eukaryotic cells.

Two of the best characterized cytoplasmic Ribonucleoprotein (RNP) aggregates in eukaryotic cells are Processing Bodies (PBs) and Stress Granules (SGs). Both aggregates contain a overlapping set of protein constituents (Anderson & Kedersha 2009; Balagopal & Parker 2009) but the functional and physical interactions between these aggregates are still unclear. Additionally, distinct cytoplasmic aggregates can share many RNA binding proteins and depending on cellular conditions, these factors can relocalize from one type of granule to another (Buchan 2014). One model regarding the role of RNP granules presents them as centers for cytoplasmic mRNA turnover, but different studies have demonstrated that P Body formation is not necessary for a functional mRNA decay (Eulalio et al. 2007). As a consequence, to date the biological functions of PB and SGs require further investigation.

In yeast and humans, the LSM1-7 complex has been associated with P Bodies (PBs) (Kedersha & Anderson 2007), however many examples exist in the literature showing differences between LSM proteins and other characterized PB components.

In the context of the *C. elegans* embryo, it has been shown that LSM-1::GFP co-localizes with other PB components, but the presence of visible LSM-1::GFP granules is not required for mRNA degradation (Gallo et al. 2008). Compared to a well-known PB component such as DCAP-2, the catalytic enzyme of the 5'-3' decapping step, in developing *C. elegans* embryos, LSM-1 is present only in somatic blastomeres whereas DCAP-2 localizes in granules (P granules) in germline blastomeres (Lall et al. 2005). Another PB component, the decapping enzyme DCAP-1, localizes in granules that are only weakly reduced upon *lsm-1* RNAi inactivation (Sun et al. 2011). Moreover, in adult animals, though DCAP-1-positive granules have been described to increase in size with age (Sun et al. 2011), we did not observe this age-associated effect in LSM-1::GFP worms (not shown). These observations raise the possibility of (i) differences in the cellular functions between LSM-1 and other canonical PB components in *C. elegans,* and (ii) an imprecise labeling of the LSM1-7 complex components as exclusive PB proteins.

Our observations of a partial co-localization between LSM-1::GFP and TIAR-1::RFP foci under heat stress in *C. elegans* suggest that LSM-1 can also accumulate in SG under specific conditions. In mammalian cells LSM1 is mostly P Body-specific in the absence of stress, although it has been observed associated with

some type of stress granules (Kedersha & Anderson 2007). Regarding the absence of constitutive LSM cytoplasmic foci under normal growth conditions in *C. elegans* adult cells, many examples exist demonstrating that the majority of P body components can also be detected diffusely throughout the cytoplasm (Eulalio et al. 2007). In summary, in terms of subcellular localization, LSM-1 has been observed associated to PBs and SGs.

In functional terms, mutations in the *C. elegans lsm-1* and *lsm-3* genes cause a Rbs (Reduced Brood Size) phenotype that is also observed in mutants of other P body genes such as homologues of the decapping enzyme DCAP-2/Dcp2 or the translation repression related protein CGH-1/Dhh1 (Figure 52).

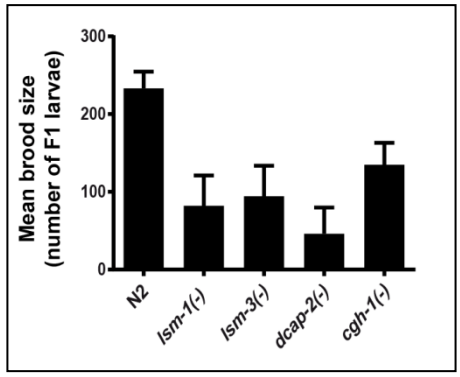


Figure 52. Reduced brood size of P body mutants compared to *lsm-1(-)* and *lsm-3(-)*. Histograms represent the mean brood sizes of wild type and the indicated mutants. Error bars represent Standard Deviations among individuals (n>15).

We show how in *C. elegans, lsm-1* and *lsm-3* are required for the accumulation of LSM-4::GFP in cytoplasmic foci under stress conditions. Similarly to the LSM-4 subunit of the LSM1-7 cytoplasmic complex that play important roles in the assembly of

yeast P-bodies through a Q/N-rich domain (Decker et al. 2007), mutations in *lsm-1* and *lsm-3* would also potentially affect the global aggregation of P bodies. These observations are relevant because in yeast, stress granule assembly is dependent on P-body formation, whereas P-body formation does not require the SG formation (Buchan et al. 2008), suggesting that for SGs to form, a previous P-body related mRNP is needed. Related to this point, although we do not observe a constitutive aggregation of cytoplasmic LSM proteins in adult cells in *C. elegans*, and we only observe this aggregation under stress conditions, we cannot exclude the possibility about LSM proteins being P body components (either diffusely distributed in the cytoplasm or forming aggregates) that can co-localize with SGs in specific stress conditions and being necessary for SG formation. To solve this question, SG formation under stress conditions should be evaluated in *lsm-1* and *lsm-3* mutant backgrounds.

Finally, a good approach to study the nature of the stress induced cytoplasmic LSM aggregates would need the combination of immunoprecipitation of a specific LSM granule component and mass spectrometry in order to determine the components associated to LSM granules in different environmental conditions and/or genetic backgrounds. For example, along with this study, we generated transgenic strains expressing fosmid vectors that contain GFP-FLAG tagged proteins for LSM-1 and LSM-4 (Sarov et al. 2012), enabling the use of α-GFP or α-FLAG antibodies in proteomic approaches.4.3. LSM-1 and LSM-3 are promoting healtshpan from the cytoplasm

4.3 LSM-1 and LSM-3 are promoting healthspan from the cytoplasm

4.3.1. A stress-related transcriptional landscape in *lsm-1(tm3585)* animals

To study the functions affected by LSM-1 during animal development, we analyzed the transcriptome of *lsm-1(tm3585)* mutants. We observed that under normal growth conditions, a subset of genes encoding proteins related to a wide variety of stress responses were enriched in *lsm-1* mutants.

One of the most interesting observations arising from our transcriptomic data is the apparent missregulation of DAF-16 target genes in *lsm-1(tm3585)* animals. The identification of DAF-16 dependent transcription targets has been an active area of research over the last years, leading to the finding of many genes with important roles in aging and stress responses. These genes have been classified as class I or class II, which are positively or negatively regulated by DAF-16 respectively (Murphy, S. A. McCarroll, et al. 2003; Tepper et al. 2013). Under normal growth conditions, the transcription factor PQM-1 antagonizes DAF-16 and activates class II genes promoting growth and development, whereas an activation of class I genes by DAF-16 allows mild stresses to be faced during development. Under acute stresses, DAF-16 activity stops development and activates a stress-dependent transcriptional response. PQM-1 and DAF-16 are mutually

exclusive in the nucleus and their activities depend on the IIS pathway. The balance between PQM-1 and DAF-16 has been proposed as responsable to trigger development versus stress programs during *C. elegans* development (Tepper et al. 2013).

We further confirmed the incorrect functioning of DAF-16 signaling by using a DAF-16::GFP reporter in an *lsm-1* mutant background and observed a delay in the total DAF-16 nuclear localization upon heat stress. These observations suggest that LSM-1 is required for the correct stress response by influencing DAF-16 activity. A similar delay in the stress-related DAF-16 nuclear localization kinetics has been observed in *C. elegans sir-2.4* mutants (Chiang et al. 2012). (Chiang et al. 2012). Sirtuins are members of a histone deacetylase family that have been shown to play key roles in cell survival and stress, regulating lifespan in worms and flies (Tissenbaum & Guarente 2001; Rogina & Helfand 2004; Vaquero 2009). Through its interaction with DAF-16 activity, SIR-2.4 has been found to be a new factor in stress resistance and DAF-16 regulation in worms.

4.3.2. LSM-1 functions in two distinct scenarios: with and without stress

Our results show that LSM-1 is not only required for the proper animal development, but also for the rapid adaptation to adverse environmental conditions. Under favorable growth conditions, LSM-1 is expressed diffusely in the cytoplasm of animal cells and in parallel, signaling through the IIS pathway keeps DAF-

16 transcription factor inactive in the cytoplasm. In these conditions, the function of LSM-1 is required for the basal DAF-16 transcriptional activity, as it can be deduced from the missregulation of IIS pathway related class I and class II genes observed in the transcriptomes of *lsm-1* mutants. A basal DAF-16 activity would be necessary to face environmental fluctuations potentially causing mild or sublethal stresses during development. As a consequence *lsm-1* mutants display reduced lifespan and developmental defects. On the other hand, the additive effect on lifespan reduction observed in *lsm-1(tm3585);daf-16(mu86)* double mutants along with the inability of a *daf-2(m577)* allele to rescue the adult lethality caused by *lsm-1(tm3585)*, suggests that the contribution of LSM-1 to the correct development of the organism is through different pathways (Figure 53A).

During acute stress conditions such as heat stress, we observed a cytoplasmic LSM-1 aggregation that was partly dependent on reduced IIS (see Figure 48). In conditions of reduced IIS, inhibitory signaling coming from DAF-2 receptor are reduced, enabling the nuclear translocation of DAF-16 and the activation of a transcriptional response for stress adaptation. As it also seem to happen in normal growth conditions, LSM-1 is required for the proper nuclear localization of DAF-16 and consequently, *lsm-1* mutants display a heightened sensitivity to all the different stresses tested. The rescue of the *daf-2(m577)* extended lifespan and dauer constitutive phenotypes by *lsm-1(tm3585)* suggest that LSM-1 is required for DAF-16 mediated phenotypes. However, since LSM-4

aggregates are observed in *daf-16(mu86)* null mutants, the formation of LSM cytoplasmic granules is not a consequence of the transcriptional response to stress activated by DAF-16.

Related to the requirements for LSM-1 and LSM-3 in conditions causing reduced IIS, especially interesting is the synthetic lethality observed when combining *lsm-1(tm3585)* or *lsm-3(tm5166)* with *daf-2(e1370),* a class II *daf-2* allele with a strong reduction of IIS.

Taking together all these observations together, we hypothesize multiple roles for LSM-1 also during stress conditions, at least by promoting DAF-16 signaling and forming cytoplasmic aggregates (Figure 53B).

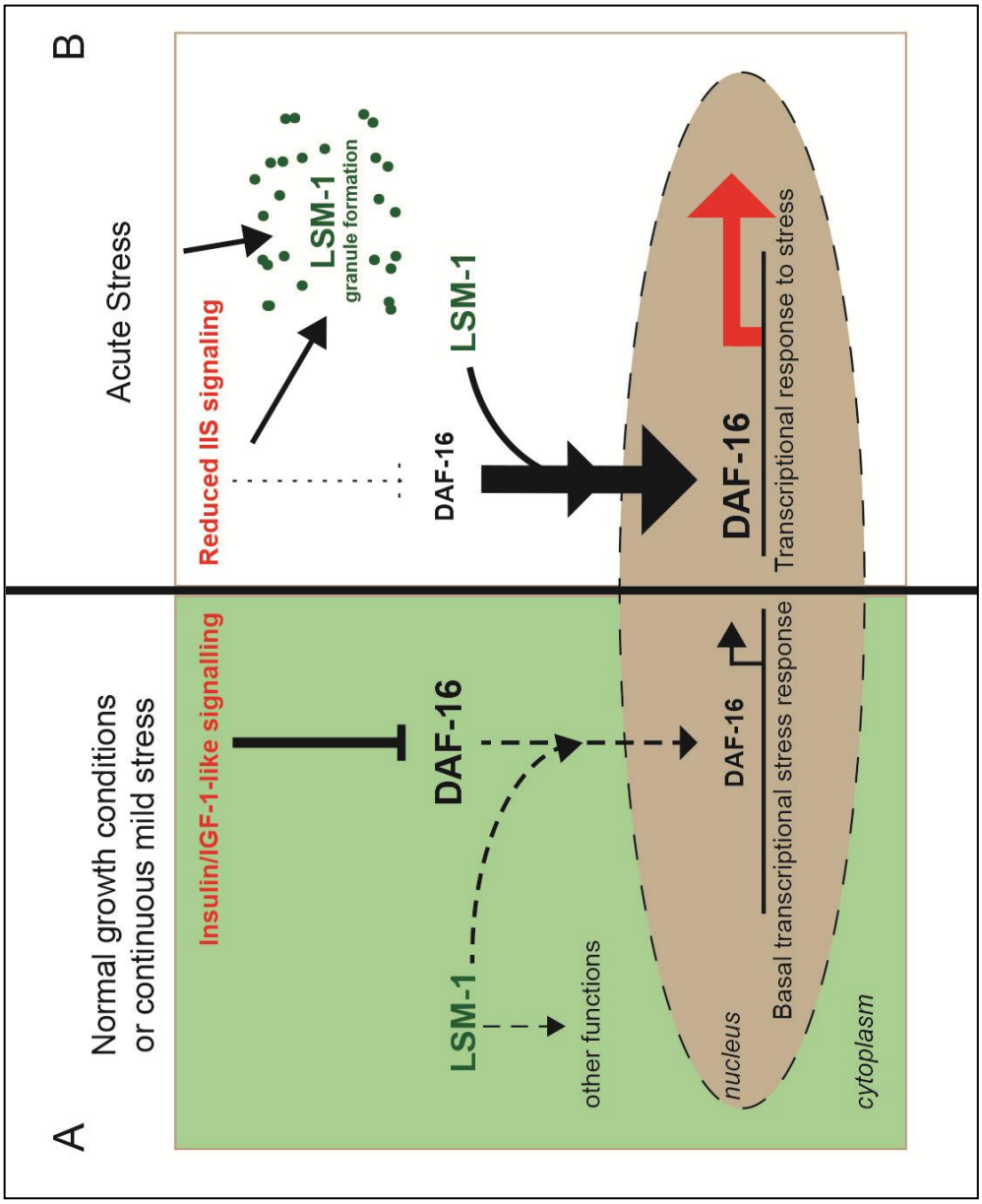


Figure 53. A model for the somatic roles of LSM-1 in normal and stress conditions. A) Under normal growth conditions, LSM-1 appears diffuse in the cytoplasm and promotes basal DAF-16 activity in the nucleus, ensuring a basal transcriptional stress response. B) Under conditions of acute stress, a reduced IIS allows DAF-16 to translocate to the nucleus. As it occurs during normal growth, LSM-1 is promotes the correct nuclear localization of DAF-16. In parallel, a reduced IIS also promotes LSM-1 cytoplasmic aggregation. These mechanisms are required for stress adaptation in multicellular animals.

The exception to the model described above is the *C. elegans* embryo, in which LSM-1 cytoplasmic foci are constitutively observed in somatic blastomeres. Regarding the link between Lsm1 and cancer, it is tempting to establish a parallelism between cancer cells and the rapid cell division occurring in the embryo.

In this context, although we do not provide information regarding the regulation of LSM-3 expression in embryos or under acute stress conditions, our phenotypic characterization suggest that LSM-1 and LSM-3 function in the same genetic pathway during development and stress adaptation. Interestingly, elevated LSM-3 expression has been correlated with metastasic phenotypes in advanced cervical carcinomas (Lyng et al. 2006) suggesting that LSM-1 and LSM-3 may influence cancer progression through similar mechanisms. In this context, it would be interesting to study Lsm1 expression in Lsm1-positive tumor cell lines and investigate if a dynamic subcellular localization influences the malignant phenotypes.

4.3.4. Does LSM-1 have a specific role in the heptameric complex?

Differently from LSM-1 overexpression, preliminary analysis of the response of LSM-4::GFP transgenic animals to heat stress suggest that LSM-4 ectopic expression does not confer enhanced resistance to heat stress (results not shown). Although we do not know if the *lsm-1* and *lsm-4* expression levels are comparable in LSM-1::GFP and LSM-4::GFP reporter strains, we

observed enhanced thermotolerance in all the obtained transgenic lines expressing LSM-1::GFP as extrachromosomal or integrated reporter suggesting a specific resistant effect associated to LSM-1 overexpression. Interestingly LSM-1::GFP seems to confer enhanced protection under acute stresses only.. We hypothesize that LSM-1 ectopic expression, only confers enhanced stress resistance, to conditions causing cytoplasmic LSM aggregation such as heat stress. Pathogen infection or aging are two conditions that cannot be considered as acute stresses and in which LSM-1 overexpression does not confer enhanced protection. It would be interesting to explore is if an additional LSM-1 expression exacerbates the long-lived and pathogen resistant phenotypes of *daf-2* mutants for example. That would mean that LSM-1 ectopic expression protects from a mild continuous effect under reduced IIS conditions.

Going back to the main question, "*Why only Lsm-1?*". One possibility is that Lsm1 has an independent function out of the complex. The other explanation is that it is essential for a specific function associated to the cytoplasmic LSM1-7 complex.
Several studies in yeast and human cells suggest that a balance in the levels of cytoplasmic LSM1-7 versus nuclear LSM2-8 exists. Lsm1 overexpression in yeast causes a depletion of U6 snRNA (Luhtala & Parker 2009). By contrast, Lsm8 expression levels inversely correlate with cytoplasmic LSM foci formation in human cells (Novotny et al. 2012) suggesting that the balance between complexes can specifically mediated by Lsm1 and Lsm8 levels.

In order to understand the mechanisms by which Lsm1 levels influence stress responses, a future question to be investigated is how the balance between cytoplasmic/nuclear complexes affects their associated functions (ie: splicing and mRNA degradation) and how this changes influence development and stress resistance.

Concluding remarks

It seems evident from our characterization of LSM proteins in *C. elegans*, is that by interfering with the correct functioning of the IIS, *lsm-1* and *lsm-3* affect animals' healthspan. Healthspan is defined as the lifetime in which an individual is active and free from age associated diseases, a concept of healthy aging for which main features have been defined in worms (Keith et al. 2014), and which is starting to become the focus of aging research (Tissenbaum 2012).

Our findings regarding the functional link existing between the IIS pathway and the LSM proteins offer an argument to explain the conservation of *lsm-1* and *lsm-3* despite being non essential genes. In fact, *lsm-1* and *lsm-3* are required for wild type fitness and probably they participate in the organism adaptability to environmental changes.

Prokaryotic species have a Lsm homolog gene known as *Hfq*. Deletion of the complete Hfq sequence in prokaryotic species is viable (Fischer et al. 2010), but causes severe growth defects in different growth conditions and affect the response to different cellular stresses (Chambers & Bender 2011; Christiansen et al. 2004; Tsui et al. 1994; Fischer et al. 2010). Although prokaryotic organisms do not contain IIS pathway homologs, the functions of Hfq/Lsm1 regarding RNA metabolism seem to be conserved. Hfq has been described functioning in the post-transcriptional regulation of mRNA stability and translation (Vogel & Luisi 2011). By this means *Hfq* participates in a sort of primitive stress response and

adaptation as a post-transcriptional regulator of prokaryotic gene expression (Muffler et al. 1996; Brown & Elliott 1996; Guisbert et al. 2007). It is then plausible, as we show along this study, that in addition to the RNA metabolism-related functions, the role in stress response adaptation is conserved in *Hfq*/LSM-1 from prokaryotic to multicellular organisms. In *C. elegans* this function is linked to the correct functioning of the IIS pathway, a central regulator of metabolism, stress response and longevity.

5. CONCLUSIONS

1- We found a functional heterogeneity among members of the LSM family in *C. elegans.* Some *lsm* genes are essential for animals viability and others not.

2- *lsm-1* and *lsm-3* are non-essential genes but are required for wild-type animal healthspan.

3- A subset of DAF-16/FOXO transcriptional targets is missregulated in *lsm-1* mutants compared to wild-type animals.

4- *lsm-1* and *lsm-3* promote stress responses through the Insulin/IGF1-like signaling (IIS) pathway by affecting DAF-16/FOXO activity.

5- Under stress conditions, cytoplasmic LSM proteins aggregate in cytoplasmic foci, and this aggregation is dependent on *lsm-1* and *lsm-3.*

6- Cytoplasmic LSM foci co-localize with stress granules markers under stress conditions.

7- Reduced IIS can trigger cytoplasmic LSM aggregation in a DAF-16 independent manner.

8- LSM-1 functions both in normal and under stress conditions.

6. MATERIALS AND METHODS

6.1. Worm strains and general methods

Caenorhabditis elegans strains were cultured and maintained between 15ºC and 25ºC on standard NGM (Nematode Growth Medium) agar plates seeded with an overgrown OP50 bacterial culture (Stiernagle 2006; Porta-de-la-Riva et al. 2012). Bristol N2 was used as wild type strain; alleles/transgenes used in this study are listed in Annex 4 Table A1. Double mutants and transgenic mutants were generated by crossing transgenic or mutant hermaphrodites with males of the desired mutant or transgenic background.

Synchronization at L1 stage was carried out by hypochlorite treatment (or *bleaching*). A mixed population of worms enriched in gravid adult worms and embryos was treated with a bleaching solution (2% sodium hypochlorite, 1M NaOH, diluted in sterile H20). After three washes with M9 buffer, the embryo preparation was filtered using a 40 μm nylon cell strainer (BD Falcon) in order to remove debris and larger clumps of unhatched eggs derived from the bleaching . Embryos resuspended in M9 were allowed to hatch in a 15mL falcon tube for at least 24 hours at 15ºC on agitation to obtain a synchronized population of L1 larvae.

6.2. RNA-mediated interference (RNAi)

RNAi libraries contain DNA sequences corresponding to almost every single gene of the *C. elegans* genome cloned into an L4440 vector. Depending on the library, this DNA fragment corresponds to gene-specific genomic DNA sequences (Ahringer library) or cDNA sequences (ORFeome library). In both cases, this gene-specific sequence is cloned in a multi cloning site of the vector flanked by two inverted T7 RNA polymerase specific promoters whose transcriptional activity is inducible by IPTG (isopropyl β-D-1-thiogalactopyranoside) (Figure 54).

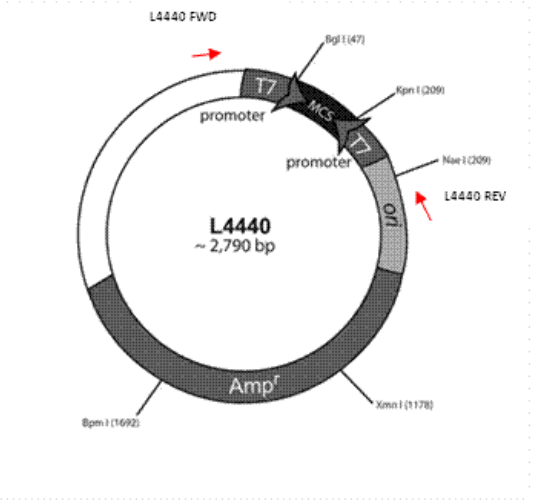


Figure 54. Schematics of a L4440 vector. A multi cloning site is flanked by two inverted T7 promoters. Red arrows represent the areas where L4440 primers anneal.

The RNAi clones used in this study were obtained from the ORFeome library (J.-F. Rual et al. 2004) (*lsm-1*, *lsm-3*, *lsm-5*, *lsm-7*, C49H6.4) or the Ahringer library (RS Kamath et al. 2003) (*lsm-2*, *lsm-4*, *lsm-6*). We completed the *lsm* family RNAi clone collection

by amplifying the cDNAs of *lsm-8,* Y48G1C.9 and K07A1.15, and cloning them into a L4440 vector by ligation after digestion with restriction enzymes. All RNAi clones were verified by sequencing using L4440 forward and reverse primers (Figure 54).

6.2.1. RNAi by feeding

To induce RNAi by feeding, standard NGM plates were supplemented with 50µg/µL ampicillin, 12,5µg/µL tetracycline and 3mM IPTG. For each experiment, the corresponding RNAi clones were picked from the corresponding RNAi library or from a previously validated glycerol stock. Colonies were grown in LB medium supplemented with 50µg/µL ampicillin and 12,5µg/µL tetracycline for at least 16 hours at 37ºC. After seeding bacteria on the RNAi plates, plates were let dry overnight at RT to allow dsRNA production induced by IPTG. RNAi plates were used to feed synchronized L1 worms, unless otherwise stated. Phenotype scoring was carried out at 25ºC unless otherwise stated. An empty L4440 vector was used as negative control for RNAi.

6.2.2. RNAi by microinjection

To induce RNAi by microinjection, each specific dsRNA was synthesized from its corresponding RNAi clone by using the MEGAscript® T7 Kit (Ambion). This kit allows the synthesis of dsRNA directly from PCR products containing an RNA polymerase promoter. L4440 vector system contains two inverted RNA polymerase T7 promoters flanking the multicloning site, allowing the synthesis of the transcribed *sense* and *antisense* strands of the

cloned fragment. We generated the templates for in vitro dsRNA synthesis for every *lsm* gene by PCR amplification from their corresponding clone by using L4440 forward and reverse primers.

Young adult animals were injected with 1µg/µl of the dsRNA of interest diluted in M9 buffer. M9 buffer alone was used as a control for microinjection experiments and phenotypes scored at 20ºC or 25ºC depending on the experiment.

6.3. Quantitative RT-PCR

Synchronized worms were washed from the plates with M9 buffer and incubated in rotation at the experimental temperature for 30 minutes to allow the maximum elimination of the bacteria present in the intestine of the worms. Then, samples were washed three more times with M9 buffer, and the pellets resuspended in 7 volumes of TRI Reagent (TR-118, MRC) and total RNA extracted using manufacturers instructions. RNA integrity was assesses by agarose electrophoresis and fluorometric methods. For cDNA synthesis using oligodT, a RevertAid H Minus First Strand cDNA Synthesis Kit (Fermentas. Cat.No. K1632) was used following manufacturers instructions. For the quantification of gene expression levels, the Roche LightCycler 480 Instrument I and the LightCycler 480 SYBR green I Master Kit (Cat. 04707516001) were used according to the manufacturer's instructions. The fold change expression of the corresponding genes was based on the ddCT method and

normalized relative to the amplification obtained using act-1 (actin) primers. Primer sequences are available upon request.

6.4. RNA-Seq Analyses

N2 wild-type and *lsm-1(tm3585)* L1 synchronized worms were grown and harvested at L3 stage (26 hours at 25 ° C). Animals were washed with M9 buffer to remove bacteria, and pellet frozen in Trizol. Total RNA purification was performed using the *mirVana* miRNA isolation kit (Ambion) followed by Ribosomal RNA depletion with the RiboMinus Eukaryote Kit (Invitrogen). RNA quality was verified in the *Experion Bioanalyzer* (Bio-Rad). We used an *Illumina* kit to make libraries that were run through a Genome Analyzer IIX Ultrasequencer (*Illumina*). Each sample produced about 10 million reads that were processed using TopHat (Trapnell et al. 2009) to be mapped against the *C. elegans* genome (WS225). BAM files were analyzed with the SeqSolve NGS software (Integromics, S.L., Spain) using a False Discovery Rate of 0.05, and filtering reads displaying multiple mapping sites. SeqSolve uses Cufflinks (Trapnell et al. 2010) and Cuffdiff (Trapnell et al. 2013) programs to perform differential gene expression analyses between samples (p value <0.05). Expression values were normalized in FPKM (fragments per kilobase of exon per million fragments mapped).

6.5. Stress Assays

For thermotolerance assays, L4 animals grown at 16ºC were transferred to plates seeded with OP50 bacteria and grown to day 1 of adulthood. Next, worms (n=40) were transferred to two 6 cm plates without any food and incubated at 35°C. Viability was scored at several time points; death was determined by the lack of movement and/or pharyngeal pumping after prodding.

To quantify DAF-16::GFP subcellular localization, L4 worms grown at 16ºC were shifted to 25ºC until they reached day one adult stage. GFP was analyzed using an Axio Imager Z1 Zeiss microscope at 40X magnification before and after heat shock at 35ºC. For heat shock time-course analyses, worms were scored for the presence or absence of GFP accumulation in the nuclei of somatic cells along the body every 10 minutes at 35ºC. Animals were scored as having full nuclear GFP if DAF-16::GFP was observed in the nucleus of somatic cells homogeneously from head to tail.

L1 starvation assay was performed as previously described by Zhang et al (Zhang et al. 2011). Briefly, adult worms were bleached and the resulting eggs were resuspended in 4-6mL S- basal without cholesterol in 15-mL tubes. Egg prep was filtered using a 40 μm nylon cell strainer (BD Falcon) in order to remove debris and larger clumps of unhatched eggs derived from the bleaching. Tubes were incubated rotating at 20ºC. To determine larval viability, 20-μL aliquots (~100worms) were placed every 3 days onto three 6-cm

nematode growth medium (NGM) plates and survival rates were calculated.

For bacterial Pathogen assays L4 worms grown at 16ºC (n=30 for each genotype) were cultured in the presence of pathogenic bacteria as the sole food source, and the number of survivors was counted every day. *Enterococcus faecalis* strain OG1RF was grown in BHI with 40 μg/ml of gentamycin. *Staphylococcus aureus* strain NTCT8325 was grown in TSB with 10 μg/ml of nalidixic acid.
For all the stress assays, graphical representation of survival curves and Kaplan-Meier statistical analyses were made using GraphPad Prism 4.0 software (GraphPad Software Inc.) *p* values were obtained applying the Mantel-Cox logrank test.

6.6. Lifespan experiments

A synchronized population of L1 larvae was grown at 15ºC for 72 hours until they reached L4 stage. At that point, worms were washed from the plate using sterile M9 buffer and were transferred (at least 100 animals per experiment) onto NGM plates incubated at 25ºC and seeded with OP50 and 0.1mg/mL 5-fluoro-2′-deoxyuridine (FUDR). FUDR was used to chemically inhibit reproduction and avoid lethality in mutant worms caused by reproductive problems. The viability of the worms, cultured at 25°C, was scored every two or three days. Animals that failed to respond to stimulation by touch were referred as dead. Day 0 of adulthood was defined as the day that mid-to-late L4s were

transferred to NGM-FUDR plates and maintained at 25°C. Survival curves were made with GraphPad Prism 4.0 software.

6.7. Nematode transgenesis

6.7.1. Generation of *lsm* translational reporters

For the generation of *lsm-1* and *lsm-4* translational reporters, fosmid vectors containing a GFP tagged version of these genes were requested from the Transgeneome resource (Sarov, John I Murray, et al. 2012) and transformation was performed by bombardment with gold particles (Biolistic Helium Gun, Caenotec). *unc-119(ed3)* young adults were shot with 16μg of the purified DNA of interest (Praitis et al. 2001). Fosmid constructs were verified by difestion with restriction enzymes.

For the co-localization experiments using *rfp::tiar-1* marker, CER157 animals expressing integrated copies of *lsm-1::gfp* were injected with a mix containing 80ng/μL of a *tiar*-1::RFP construct (Rousakis et al. 2014) and 20ng/μL of the linearized roller marker pRF4 [*rol-6(su1006)*].

For the generation of P*elt-2*::GFP::LSM-3 translational reporters, the MultiSite Gateway Three-Fragment Vector Construction Kit (Invitrogen) was used. We combined three pEntry vectors, a pEntry1 entry vector from the *C. elegans* Promoterome platform (Dupuy et al. 2004) containing the *Pelt-2* sequence, a pEntry2 vector contining a *gfp* open reading frame and a pEntry3 vector

containing the *lsm-3* open reading frame and its own 3'UTR. The final expression vector obtained was transformed in Library Efficiency DH5 competent bacteria (Invitrogen), purified and verified by sequencing using M13 primers. Transgenic animals were generated by microinjecting 5ng/uL of the purified DNA together with a mix containing 5 ng/µL of the red muscular marker pCFJ104 [P*myo-3::mCherry::unc-54*_3'UTR] and 10 ng/µL of the red neuronal marker pGH8 [P*rab-3::mCherry::unc54*_3'UTR]

6.7.2. Generation of *lsm* transcriptional reporters

To generate transcriptional reporters, molecular constructs were made by PCR-fusion (Hobert 2002). We amplified promoter sequences from wild type N2 genomic DNA and the pPD95.75 vector was used to amplify a GFP sequence fused to *unc-54* 3'UTR. Transgenic strains were generated by microinjection into N2 animals and/or in *dpy-5(e907)* mutant animals. In N2 worms, 70 ng/µL of the purified PCR product was microinjected together with a mix containing 20 ng/µL of the linearized roller marker pRF4 [*rol-6(su1006)*], 5 ng/µL of the red muscular marker pCFJ104 [P*myo-3::mCherry::unc-54*_3'UTR] and 10 ng/µL of the red neuronal marker pGH8 [P*rab-3::mCherry::unc54*_3'UTR] (list of transcriptional reporters generated in Annex 4 Table A2). In *dpy-5(e907)* animals, promoter::GFP constructs were co-injected with the *dpy-5*(+) plasmid (pCeh-361), (Hunt-Newbury et al. 2007), (list of transcriptional reporters generated in Annex 4 Table A3).

6.8. Gonad dissection and germ line quantification

For germ line quantification, adult worms were immobilized in 0.3mM Levamisole. Next, gonads were dissected and fixed (formaldehyde 3%, methanol 75%, K2HPO4 6.2 mM) and stained with DAPI (0.6 mg/mL) at the indicated timepoints. The stained gonads were photographed using a Leica DM5000B microscope. Digital pictures were used for germ cell quantification. For the germline time-course experiment, germ nuclei from the distal part to the bend of the gonad were scored in a single Z stack. At least 15 germlines were quantified for each experiment.

6.9. Gonad and embryo immunostaining

For gonad immunostaining, adult worms were immobilized in 0.3mM Levamisole. Next, gonads were dissected and fixed in Formaldehyde 4% for 20 min (Formaldehyde stabilized with MetOH, Panreac, Cat. 252931). Incubations with primary and secondary antibodies were performed overnight at 4ºC and for 2 hours at room temperature, respectively. Antibody dilutions were done in PBS 0.05% Tween 1% BSA. After fixation and antibody incubations, gonads were washed three times with PBS 0.05% Tween. All samples were counterstained with DAPI (0.6 μg/mL) in order to visualize the nuclei.

For embryo immunostaining, an alternative protocol based on Duer J 2006 (Duerr 2006) was used. This protocol includes practical modifications allowing the removal of the embryo cuticle for a better penetration of the antibodies prior to immunostaining.

Antibody	Epitope	Host	Nature	Dilution	Source
A11120	GFP	mouse	primary	1:200	Molecular Probes
	H3P (Ser 10)	rabbit	primary	1:800	Millipore
A11001 (Alexa fluor 488 IgG)	Anti-mouse	goat	secondary	1:400	Molecular Probes
A11001 (Alexa fluor 568 IgG)	Anti-rabbit	goat	secondary	1:400	Molecular Probes

Table 9. List of antibodies used for germline and embryo immunostaining.

6.10. Protein extraction and Western blotting

Synchronized worms were washed off the plates using M9 buffer and let in rotation for 30 minutes. Samples were washed three more times with M9 buffer and pellets were mixed with lysis buffer 2X (4% SDS, 100 mM Tris HCl ph 6.8, 20% glycerol, 1X protease and phosphatase inhibitors (Roche, Cat. 1187350001). Total protein quantification was done using the Bio-Rad DC protein assay (Bio-Rad, Cat. 500-0112). Protein samples were electrophoresed on SDS 10% polyacrylamide gels and electroblotted onto nitrocellulose membranes. 5% Non-fat milk diluted in TBS-Tween buffer (25mM Tris-HCl, pH 7.5, 137mM NaCl, 0.1% Tween) was used to block the membranes. Blotting was carried out using a primary antibody

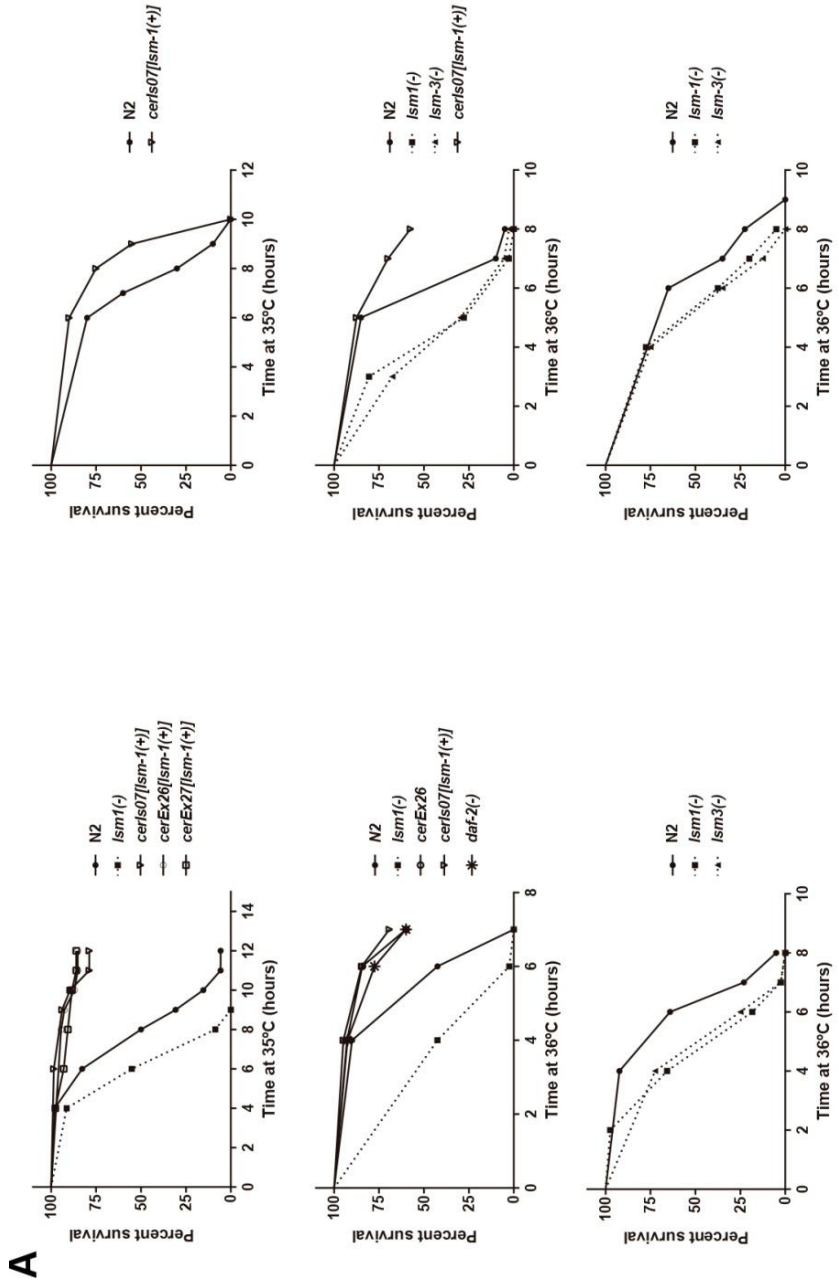
recognizing PMK-1/p38 dually phosphorylated at Threonine 180 and Tyrosine 182 (Cell Signaling) and actin (MP Biomedicals). Primary antibodies were incubated overnight at 4ºC. Secondary antibodies, HRP-conjugated anti-rabbit and anti-mouse (Dako) were incubated for 1 hour at room temperature. Membranes were developed using luminal, a substrate for HRP, and autoradiographic films (CL-xposure films, Cultek S.A.) were used to detect protein signal.

Antibody	Epitope	Host	Nature	Dilution	Source
#9212S	Phospho PMK-1/p38 (T180/Y182)	rabbit	primary	1:1000	Cell Signaling
C4 (69100)	ACT-1	mouse	primary	1:500	MP biomedicals
P0260 (HRP conjugated)	Anti-mouse	rabbit	secondary	1:2000	Dako
P0448 (HRP conjugated)	Anti-rabbit	goat	secondary	1:2000	Dako

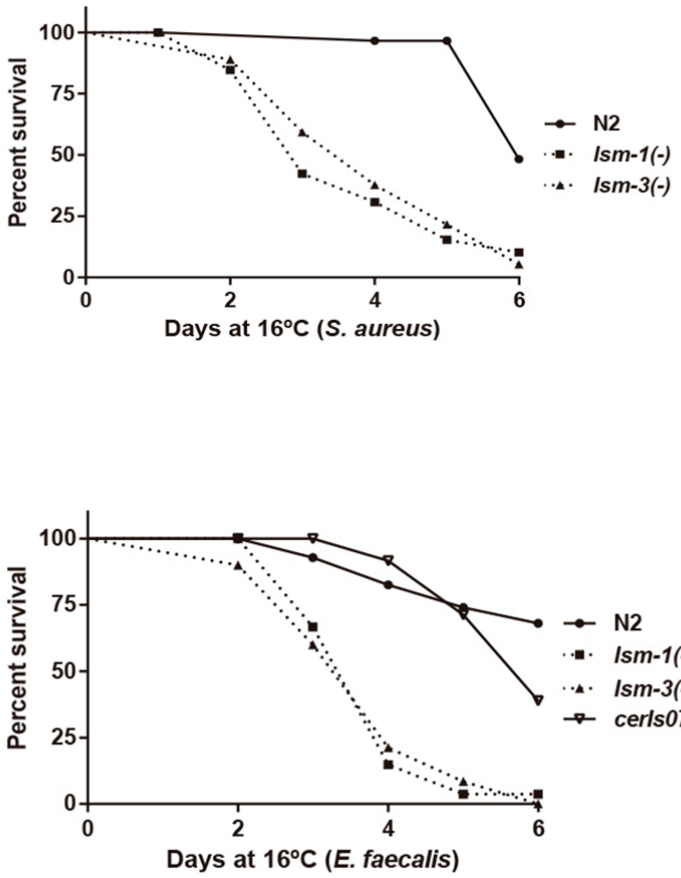
Table 10. List of antibodies used for western blot analysis.

7. ANNEX

Annex 1 – Stress and lifespan assays



B

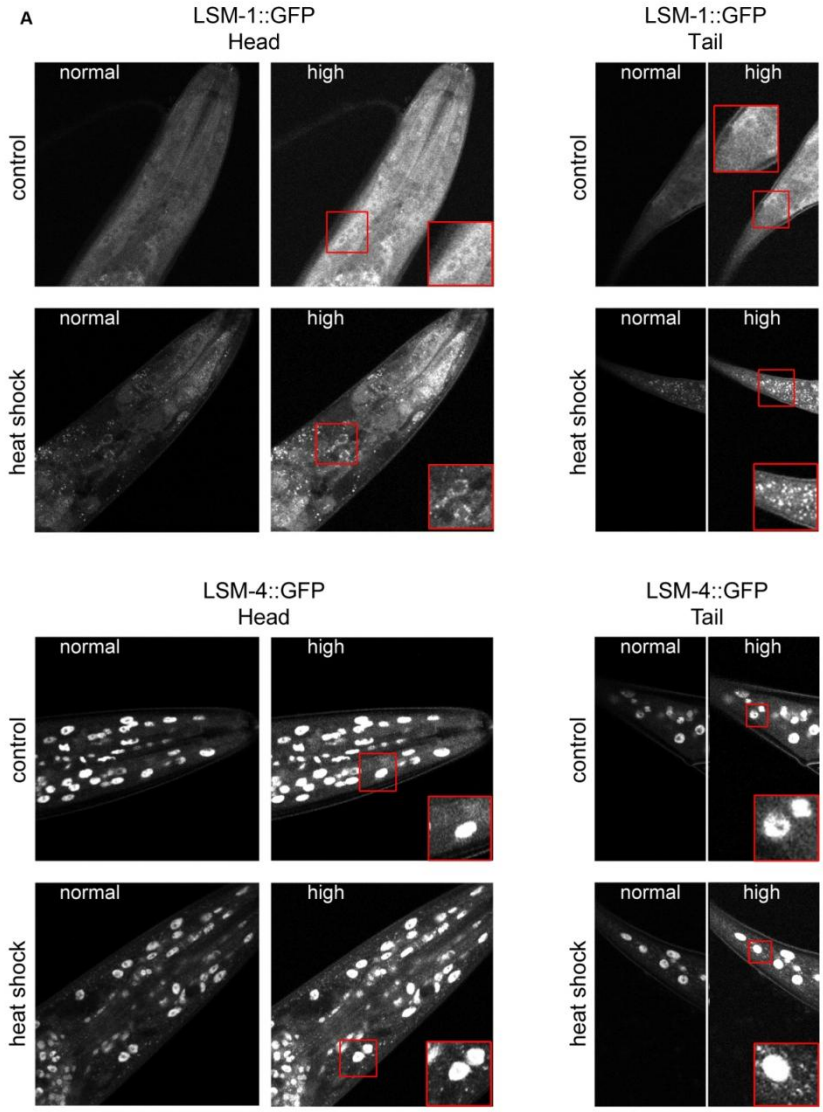

Percent survival
100
75
50
25
0
0
2
4
6
Days at 16°C (S. aureus)
N2
lsm-1(-)
lsm-3(-)
Percent survival
100
75
50
25
0
0
2
4
6
Days at 16°C (E. faecalis)
N2
lsm-1(-)
lsm-3(-)
cerIs07[lsm-1(+)]

C

Genotype	Mean lifespan (days)	Median lifespan (days)	Maximum lifespan (days)	n
Replicate 1				
N2[a]	12.3 ± 1.5	13	16	167
lsm-1(tm3585)[a]	7.1 ± 1.8	7	11	306
lsm-3(tm5166)[a]	5.9 ± 1.1	6	8	319
daf-2(m577);lsm-1(tm3585)	7.7 ± 1.8	8	11	170
cerIs07[lsm-1(+)][a]	12.7 ± 2.3	13	16	146
daf-2(m577)[a]	21.0 ± 2.9	22	27	172
Replicate 2				
N2[c]	13,0 ± 2.4	13	17	146
daf-16(mu86);lsm-1(tm3585)[c]	7.4 ± 1.4	7	9	185
daf-16(mu86)[c]	9.4 ± 0.8	9	12	196
lsm-1(tm3585)[c]	9.1 ± 1.3	9	12	287
lsm-3(tm5166)	9.1 ± 1.7	9	13	162
daf-2(m577);lsm-1(tm3585)	10.0 ± 2.2	10	15	127
daf-2(m577);lsm-3(tm5166)	11.7 ± 2.2	11	14	155
Replicate 3				
N2[b]	12.8 ± 1.9	13	20	276
daf-16(mu86);lsm-1(tm3585)	7.0 ± 1.2	8	9	202
daf-16(mu86)	7.9 ± 1.7	8	10	229
lsm-1(tm3585)	9.0 ± 1.9	9	13	200
lsm-3(tm5166)	8.8 ± 2.2	9	13	193
daf-2(m577);lsm-1(tm3585)[b]	10.6 ± 4.2	9	22	252
daf-2(m577);lsm-3(tm5166)[b]	10.6 ± 2.6	10	14	123
cerIs07[lsm-1(+)]	12.8 ± 2.3	13	20	186
daf-2(m577)[b]	21.3 ± 2.3	22	29	101

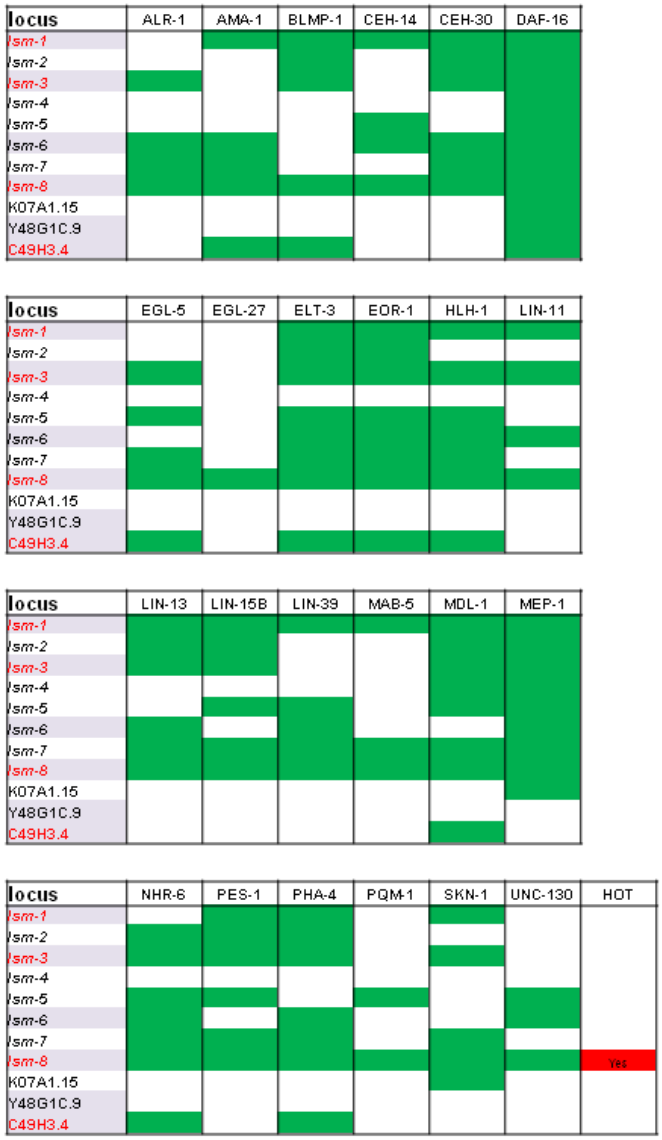
Annex 1. Additional data related to stress assays A) Additional thermal stress assays. Survival curves showing the significant sensitivity of *lsm-1* and *lsm-3* mutants to thermal stress at 35ºC and/or 36ºC ($p<0.001$ in all experiments), and the significant thermo resistance of transgenic strains over expressing *lsm-1* ($p<0.001$ in all experiments) **B) Additional data related to pathogen infection assays.** Survival curves showing the significant sensitivity of *lsm-1* and *lsm-3* mutants to infection with *S. aureus* and *E. faecalis* ($p<0.001$ in all cases) **C) Additional data related to lifespan assays.** Table showing the lifespan parameters observed for different genotypes in three different biological replicates. Strains represented in [a]Figure R22, [b]Figure R29A and [c]Figure R29B respectively.

Annex 2 – Cytoplasmic LSM foci



Annex 2. Additional confocal images showing specific cytoplasmic accumulation of LSM proteins under heat stress. Confocal images showing Z projections of three 1µm z slides. Images from the head and the tail of worms expressing LSM-1::GFP (strain CER157) or LSM-4::GFP (strain CER41) upon normal conditions and heat stress (1hour at 35ºC). Two different exposures are shown (normal or high) to confirm that the presence of foci is observed only under heat stress.

Annex 3 – Transcription factors binding to LSM promoters



Annex 3. Table summarizing transcription factor binding patterns for *lsm* promoter regions. Green cases represent a positive interaction ($p<10^{-5}$) found between a transcription factor (TF) and an *lsm* promoter region. White cases represent no interaction. This organization allows the comparison between TF binding patterns to different *lsm* promoter regions. *lsm* genes highlighted in red are present in operons. HOT (Hot occupancy of target) regions have higher density of binding than would be expected by chance. Based on modENCODE

data for TF binding sites, only *lsm-8* promoter is considered a HOT region (red case) among the rest of *lsm* genes. Consequently, all the TF tested bind to *lsm* promoter region. K07A1.15, C49H3.4 and Y48G1C.9 present in general a more distinct TF binding signature compared to the rest of the members of the family. TF binding sites were defined as the 500bp upstream of the coding sequence of the gene. For genes present in operons, the promoter region of the first gene was analyzed.

Annex 4 – List of alleles and transgenic strains

Strain	Genotype	Source
CER 60	*lsm-1(tm3585)* II	NBP
CER59	*lsm-3(tm5166)*IV	NBP
CER36	*lsm-1(tm3585)*II;*lsm-3(tm5166)*IV	This study
CB1370	*daf-2(e1370)*III	CGC
CF1038	*daf-16(mu86)*I	CGC
BL3466	*inIs173[*P*Nvitgfp]*	T. Blumenthal lab
CER129	*inIs173[*P*Nvitgfp];lsm-3(tm5166)*IV	This study
TJ356	*zIs356[Pdaf-16::daf-16::gfp,rol-6(su1006)]*IV	CGC
CER155	*zIs356*IV;*lsm-1(tm3585)*II	This study
DR1567	*daf-2(m577)*III	CGC
CER154	*lsm-1(tm3585)*II;*daf-2(m577)*III	This study
CER158	*lsm-1(tm3585)*II;*daf-16(mu86)*I	This study
CER159	*lsm-3(tm5166)*IV;*daf-2(m577)*III	This study

DG1701	*cgh-1(tn691)*III	CGC
RB1641	*dcap-2(ok2023)*IV	CGC
CF1139	*daf-16(mu86) I; muIs61[daf16::GFP(pKL78) + rol-6(pRF4)]*	CGC
CF1724	*daf-16(mu86) I; daf-2(e1370)III; muIs105[*daf-16p::GFP::daf-16 + rol-6]	This study
CER26	*unc-119(ed3)*III; *cerEx26[*P*lsm-1*::*lsm-1::gfp*::*lsm*-1_3'UTR, *unc-119(+)]*	This study
CER27	*unc-119(ed3)*III; *cerEx27*[P*lsm-1*::*lsm-1::gfp*::*lsm*-1_3'UTR, *unc-119(+)]*	This study
CER41	*unc-119(ed3)*III; *cerIs02*[P*lsm*-4::*lsm-4::gfp*::*lsm*-4_3'UTR, *unc-119(+)]*	This study
CER157	*unc-119(ed3)*III; *cerIs07*[P*lsm-1*::*lsm-1::gfp*::*lsm*-1_3'UTR, *unc-119(+)]*	This study
CER152	*cerIs02*; *lsm-1(tm3585)*II	This study
CER153	*cerIs02; lsm-3(tm5166)IV*	This study
CER160	*cerIs02;daf-16(mu86)I*	This study
CER161	*cerIs02;daf-2(e1370)III*	This study
CER162	*cerIs07;lsm-1(tm3585)II*	This study

CER166	*cerIs07;cerEx42[Ptiar-1::RFP::tiar-1ORF::tiar-1 3'UTR+rol-6(+)]*	This study
CER42	*N2;cerEx26; cerEx26[Plsm-7::lsm-7ORF::gfpORF::lsm-73'UTR+unc-119(+)]*	This study
CER35	*N2; cerEx20; cerEx20 [Pelt-2::gfp::lsm-3ORF-3'UTRlsm-3 + myo-3::mcherry + rab-3::mcherry]*	This study

Table A1. List of mutant and transgenic strains used in this study. CGC: *Caenorhabditis* Genetics Center (Minnessota, USA) , NBP: National Bioresource Project (Tokyo, Japan).

STRAIN	GENOTYPE	ADDITIONAL INFORMATION
CER15	*cerEx14[plsm-8::gfp + rol-6(+)+ pmyo-3::mcherry + prab-3::mcherry]*	*lsm-8* promoter
CER16	*cerEx14[plsm-5::gfp + rol-6(+)+ pmyo-3::mcherry + prab-3::mcherry]*	*lsm-5* promoter
CER17	*cerEx10[plsm-3::gfp + rol-6(+) + pmyo-3::mcherry + prab-3::mcherry]*	*lsm-3* promoter
CER18	*cerEx14[pC49H3.4::gfp + rol-6(+)+ pmyo-3::mcherry + prab-3::mcherry]*	C49H3.4 promoter
CER19	*cerEx14[plsm-7::gfp + rol-6(+)+ pmyo-3::mcherry + prab-3::mcherry]*	*lsm-7* promoter
CER20	*cerEx14[plsm-6::gfp + rol-6(+)+ pmyo-3::mcherry + prab-3::mcherry]*	*lsm-6* promoter
CER21	*cerEx14[plsm-1::gfp + rol-6(+)+ pmyo-3::mcherry + prab-3::mcherry]*	*lsm-1* promoter
CER22	*cerEx14[pY48G1C.9::gfp + rol-6(+)+ pmyo-3::mcherry + prab-3::mcherry]*	Y48G1C.9 promoter no roller marker
CER23	*cerEx14[plsm-2::gfp + rol-6(+)+ pmyo-3::mcherry + prab-3::mcherry]*	*lsm-2* promoter low roller marker
CER24	*cerEx14[plsm-4::gfp + rol-6(+)+ pmyo-3::mcherry + prab-3::mcherry]*	*lsm-3* promoter low roller marker
-	No construct available for K07A1.15	K07A1.15

Table Annex A2. List of promoter::GFP strains for *lsm* genes generated in our lab.

STRAIN	GENOTYPE	ADDITIONAL INFORMATION
BC20377	*sEx20377[rCesC49H3.4::GFP + pCeh361];dpy-5(e907)I*	C49H3.4 promoter
BC20380	*sEx20380[rCesF28F8.3::GFP + pCeh361];dpy-5(e907)I*	*lsm-5* promoter
BC20381	*sEx20381[rCesF28F8.3::GFP + pCeh361];dpy-5(e907)I*	*lsm-5* promoter
BC20382	*sEx20382[rCesF32A5.7::GFP + pCeh361];dpy-5(e907)I*	*lsm-4* promoter
BC20383	*sEx20383[rCesF32A5.7::GFP + pCeh361];dpy-5(e907)I*	*lsm-4* promoter
BC20384	*sEx20384[rCesF40F8.9::GFP + pCeh361];dpy-5(e907)I*	*lsm-1* promoter
BC20385	*sEx20385[rCesF40F8.9::GFP + pCeh361];dpy-5(e907)I*	*lsm-1* promoter
BC20386	*sEx20386[rCesT10G3.6::GFP + pCeh361];dpy-5(e907)I*	*lsm-2* promoter
BC20388	*sEx20388[rCesY48G1C.9::GFP + pCeh361];dpy-5(e907)I*	Y48G1C.9 promoter
BC20391	*sEx20391[rCesK07A1.15::GFP + pCeh361];dpy-5(e907)I*	k07A1.15 promoter
BC20393	*sEx20393[rCesY73B6BL.32::GFP + pCeh361];dpy-5(e907)I*	*lsm-8* promoter
BC20395	*sEx20395[rCesY62E10A.12::GFP + pCeh361];dpy-5(e907)I*	*lsm-3* promoter
BC20396	*sEx20396[rCesZK593.7::GFP + pCeh361];dpy-5(e907)I*	*lsm-7* promoter
-	No construct available for *lsm-6*	*lsm-6* promoter

Table A3. Table Annex A2. List of promoter::GFP strains for *lsm* genes generated in Robert Johnsen lab (Simon Fraser University, Canada).

Annex 5- Publications during this thesis

1- Cornes, E. Quéré, CAL. Giordano-Santini, R. Dupuy, D. **Applying antibiotic selection markers for nematode genetics**. *Methods* (68):403-408 (2014).

2- Cornes, E. Porta-de-la-Riva, M. Aristizábal-Corrales, D. Brokate-Llanos, AM. García-Rodríguez, FJ. Ertl, I. Díaz, M. Fontrodona, L. Reis, K. Johnsen, R. Baillie, D. Muñoz, MJ. Sarov, M. Dupuy, D. Cerón, J. **Cytoplasmic LSM-1 protein regulates stress responses through the insulin/IGF-1 signaling pathway in *C. elegans.*** *RNA (In press)* (2015).

3- Rubio, K. Fontrodona, L. Aristizábal-Corrales, D. Torres, S. Cornes, E. García-Rodríguez, FJ. Serrat, X. **Modeling of autosomal dominant Retinitis Pigmentosa in *Caenorhabditis elegans* uncovers a nexus between global impaired functioning of certain splicing factors and cell-type specific apoptosis.** *RNA* (Under review) (2014).

4- Ertl,I. Porta-de-la-Riva, M. Gómez-Orte, E. Aristizábal-Corrales, D. Cornes, E. Fontrodona, L. Askjaer, P. Cabello, J. Cerón, J. **Functional interplay of two paralogs encoding SWI/SNF chromatin-remodeling accessory subunits during *C. elegans* development.**. *Genetics.* (Under Review) (2015).

REFERENCES

REFERENCES

Altun, Z.F. et al., 2007. WormAtlas and WormImage: The Websites and the Book. *International Worm Meeting*.

Allen, M.A. et al., 2011. A global analysis of *C. elegans* trans-splicing. *Genome Research*, 21(2), pp.255–264.

Anderson, P. & Kedersha, N., 2009. RNA granules: post-transcriptional and epigenetic modulators of gene expression. *Nature reviews. Molecular cell biology*, 10(6), pp.430–436.

Azzouz, T.N. & Schumperli, D., 2003. Evolutionary conservation of the U7 small nuclear ribonucleoprotein in Drosophila melanogaster. *RNA (New York, N.Y.)*, 9(12), pp.1532–1541.

Balagopal, V. & Parker, R., 2009. Polysomes, P bodies and stress granules: states and fates of eukaryotic mRNAs. *Current Opinion in Cell Biology*, 21(3), pp.403–408.

Beggs, J.D., 2005. Lsm proteins and RNA processing. *Biochemical Society Transactions*, 33(Pt 3), pp.433–438.

Brown, L. & Elliott, T., 1996. Efficient translation of the RpoS sigma factor in Salmonella typhimurium requires host factor I, an RNA-binding protein encoded by the hfq gene. *Journal of Bacteriology*, 178(13), pp.3763–3770.

Buchan, J.R., 2014. mRNP granules: Assembly, function, and connections with disease. *RNA Biology*, 11(8), p.11.

Buchan, J.R., Muhlrad, D. & Parker, R., 2008. P bodies promote stress granule assembly in Saccharomyces cerevisiae. *Journal of Cell Biology*, 183(3), pp.441–455.

Buchan, J.R. & Parker, R., 2009. Eukaryotic Stress Granules: The Ins and Outs of Translation. *Molecular Cell*, 36(6), pp.932–941.

Ceron, J. et al., 2007. Large-scale RNAi screens identify novel genes that interact with the *C. elegans* retinoblastoma pathway as well as splicing-related components with synMuv B activity. *BMC developmental biology*, 7, p.30.

Coleen T. Murphy and Patrick J. Hu., 2013. Insulin/insulin-like growth factor signaling in *C. elegans*. In ed. The *C. elegans* Research Community, ed. *WormBook*.

Chambers, J.R. & Bender, K.S., 2011. The RNA chaperone Hfq is important for growth and stress tolerance in Francisella novicida. *PloS one*, 6(5), p.e19797.

Chiang, W.-C. et al., 2012. *C. elegans* SIRT6/7 homolog SIR-2.4 promotes DAF-16 relocalization and function during stress. *PLoS genetics*, 8(9), p.e1002948.

Christiansen, J.K. et al., 2004. The RNA-Binding Protein Hfq of Listeria monocytogenes: Role in Stress Tolerance and Virulence. *Journal of Bacteriology*, 186(11), pp.3355–3362.

Decker, C.J., Teixeira, D. & Parker, R., 2007. Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in Saccharomyces cerevisiae. *Journal of Cell Biology*, 179(3), pp.437–449.

Dong, M.-Q. et al., 2007. Quantitative mass spectrometry identifies insulin signaling targets in *C. elegans*. *Science (New York, N.Y.)*, 317(5838), pp.660–3.

Duerr, J.S., 2006. Immunohistochemistry. *WormBook : the online review of C. elegans biology*, pp.1–61.

Dupuy, D. et al., 2004. A first version of the Caenorhabditis elegans promoterome. *Genome Research*, 14(10 B), pp.2169–2175.

Eulalio, A. et al., 2007a. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Molecular and cellular biology*, 27(11), pp.3970–3981.

Eulalio, A. et al., 2007b. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Molecular and cellular biology*, 27(11), pp.3970–81.

Evans, E. a, Chen, W.C. & Tan, M.-W., 2008. The DAF-2 insulin-like signaling pathway independently regulates aging and immunity in *C. elegans*. *Aging cell*, 7(6), pp.879–93.

Fischer, S. et al., 2010. The archaeal lsm protein binds to small RNAs. *Journal of Biological Chemistry*, 285(45), pp.34429–34438.

Fontrodona, L. et al., 2013. RSR-2, the Caenorhabditis elegans Ortholog of Human Spliceosomal Component SRm300/SRRM2, Regulates Development by Influencing the Transcriptional Machinery. *PLoS Genetics*, 9(6).

Gallo, C.M. et al., 2008. Processing bodies and germ granules are distinct RNA granules that interact in *C. elegans* embryos. *Developmental biology*, 323(1), pp.76–87.

Gems, D. et al., 1998. Two pleiotropic classes of daf-2 mutation affect larval arrest, adult behavior, reproduction and longevity in Caenorhabditis elegans. *Genetics*, 150(1), pp.129–155.

Gerstein, M.B. et al., 2010. Integrative Analysis of the Caenorhabditis elegans Genome by the modENCODE Project. *Science*, 330(6012), pp.1775–1787.

Grishok, A., 2005. RNAi mechanisms in Caenorhabditis elegans. *FEBS Letters*, 579(26), pp.5932–5939.

Grün, D. et al., 2014. Conservation of mRNA and protein expression during development of *C. elegans*. *Cell reports*, 6(3), pp.565–77.

Guisbert, E. et al., 2007. Hfq modulates the sigmaE-mediated envelope stress response and the sigma32-mediated cytoplasmic stress response in Escherichia coli. *Journal of bacteriology*, 189(5), pp.1963–73.

He, W. & Parker, R., 2000. Functions of Lsm proteins in mRNA degradation and splicing. *Current Opinion in Cell Biology*, 12(3), pp.346–350.

Hobert, O., 2002. BioTechniques - PCR Fusion-Based Approach to Create Reporter Gene Constructs for Expression Analysis in Transgenic *C. elegans*. *Biotechniques*, 32(4), pp.728–730.

Hu, P.J., 2007. Dauer. *WormBook : the online review of C. elegans biology*, pp.1–19.

Huang, D.W., Sherman, B.T. & Lempicki, R.A., 2009. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, 37(1), pp.1–13.

Huang, D.W., Sherman, B.T. & Lempicki, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*, 4(1), pp.44–57.

Hunt-Newbury, R. et al., 2007. High-throughput in vivo analysis of gene expression in Caenorhabditis elegans. *PLoS Biology*, 5(9), pp.1981–1997.

Ingelfinger, D. et al., 2002. The human LSm1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrnl in distinct cytoplasmic foci. *RNA (New York, N.Y.)*, 8(12), pp.1489–1501.

Jagannath, A. & Wood, M.J.A., 2009. Localization of double-stranded small interfering RNA to cytoplasmic processing bodies is Ago2 dependent and results in up-regulation of GW182 and Argonaute-2. *Molecular biology of the cell*, 20(1), pp.521–529.

Kamath, R.S. et al., 2003a. Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. *Nature*, 421(6920), pp.231–237.

Kamath, R.S. et al., 2003b. Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. *Nature*, 421(6920), pp.231–7.

Kedersha, N. & Anderson, P., 2007. Mammalian stress granules and processing bodies. *Methods in enzymology*, 431, pp.61–81.

Kedersha, N.L. et al., 1999. RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. *The Journal of cell biology*, 147(7), pp.1431–42.

Keith, S.A. et al., 2014. The *C. elegans* healthspan and stress-resistance assay toolkit. *Methods*, 68(3), pp.476–486.

Kerins, J.A. et al., 2010. PRP-17 and the pre-mRNA splicing pathway are preferentially required for the proliferation versus meiotic development decision and germline sex determination in Caenorhabditis elegans. *Developmental Dynamics*, 239(5), pp.1555–1572.

Kim, D.H. et al., 2002. A conserved p38 MAP kinase pathway in Caenorhabditis elegans innate immunity. *Science*, 297(5581), pp.623–626.

Kim, S.K., 2001. Http://*C. elegans*: mining the functional genomic landscape. *Nature reviews. Genetics*, 2(9), pp.681–689.

Lall, S., Piano, F. & Davis, R.E., 2005. Caenorhabditis elegans Decapping Proteins : Localization and Functional Analysis of Dcp1 , Dcp2 , and DcpS during Embryogenesis. , 16(December), pp.5880–5890.

Lin, K. et al., 2001. Regulation of the Caenorhabditis elegans longevity protein DAF-16 by insulin/IGF-1 and germline signaling. *Nature genetics*, 28(2), pp.139–45.

Luhtala, N. & Parker, R., 2009. LSM1 over-expression in Saccharomyces cerevisiae depletes U6 snRNA levels. *Nucleic Acids Research*, 37(16), pp.5529–5536.

Lyng, H. et al., 2006. Gene expressions and copy numbers associated with metastatic phenotypes of uterine cervical cancer. *BMC genomics*, 7, p.268.

MacMorris, M., Brocker, C. & Blumenthal, T., 2003. UAP56 levels affect viability and mRNA export in Caenorhabditis elegans. *RNA (New York, N.Y.)*, 9(7), pp.847–857.

Mayes, a E. et al., 1999. Characterization of Sm-like proteins in yeast and their association with U6 snRNA. *The EMBO journal*, 18(15), pp.4321–31.

McColl, G. et al., 2010. Insulin-like signaling determines survival during stress via posttranscriptional mechanisms in *C. elegans*. *Cell metabolism*, 12(3), pp.260–72.

Mitchell, S.F. & Parker, R., 2014. Principles and Properties of Eukaryotic mRNPs. *Molecular Cell*, 54(4), pp.547–588.

Morimoto, R.I., 1998. Regulation of the heat shock transcriptional response: cross talk between a family of heat shock factors, molecular chaperones, and negative regulators. *Genes & development*, 12(24), pp.3788–96.

Muffler, A., Fischer, D. & Hengge-Aronis, R., 1996. The RNA-binding protein HF-I, known as a host factor for phage Q?? RNA replication, is essential for rpoS translation in

Escherichia coli. *Genes and Development*, 10(9), pp.1143–1151.

Murphy, C.T., McCarroll, S. a, et al., 2003. Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. *Nature*, 424(6946), pp.277–83.

Murphy, C.T., McCarroll, S.A., et al., 2003. Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. *Nature*, 424(6946), pp.277–283.

Novotny, I. et al., 2012. Nuclear LSm8 affects number of cytoplasmic processing bodies via controlling cellular distribution of Like-Sm proteins. *Molecular biology of the cell*, 23(19), pp.3776–85.

Papasaikas, P. et al., 2014. Functional Splicing Network Reveals Extensive Regulatory Potential of the Core Spliceosomal Machinery. *Molecular Cell*, 57(1), pp.7–22.

Patel, D.S. et al., 2008. Clustering of genetically defined allele classes in the Caenorhabditis elegans DAF-2 insulin/IGF-1 receptor. *Genetics*, 178(2), pp.931–946.

Perea-Resa, C. et al., 2012. LSM proteins provide accurate splicing and decay of selected transcripts to ensure normal Arabidopsis development. *The Plant cell*, 24(12), pp.4930–4947.

Pillai, R.S. et al., 2001. Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *EMBO Journal*, 20(19), pp.5470–5479.

Polevoda, B. & Sherman, F., 2001. NatC N??-terminal Acetyltransferase of Yeast Contains Three Subunits, Mak3p, Mak10p, and Mak31p. *Journal of Biological Chemistry*, 276(23), pp.20154–20159.

Porta-de-la-Riva, M. et al., 2012. Basic Caenorhabditis elegans methods: synchronization and observation. *Journal of visualized experiments : JoVE*, (64), p.e4019.

Praitis, V. et al., 2001. Creation of low-copy integrated transgenic lines in Caenorhabditis elegans. *Genetics*, 157(3), pp.1217–1226.

Rogina, B. & Helfand, S.L., 2004. Sir2 mediates longevity in the fly through a pathway related to calorie restriction. *Proceedings of the National Academy of Sciences of the United States of America*, 101(45), pp.15998–16003.

Van Rossum, A.J. et al., 2001. Proteomic identification of glutathione S-transferases from the model nematode Caenorhabditis elegans. *Proteomics*, 1(11), pp.1463–1468.

Rousakis, A. et al., 2014. Diverse functions of mRNA metabolism factors in stress defense and aging of Caenorhabditis elegans. *PloS one*, 9(7), p.e103365.

Rual, J.F. et al., 2004. Toward improving Caenorhabditis elegans phenome mapping with an ORFeome-based RNAi library. *Genome Research*, 14(10 B), pp.2162–2168.

Rual, J.-F. et al., 2004. Toward improving Caenorhabditis elegans phenome mapping with an ORFeome-based RNAi library. *Genome research*, 14(10B), pp.2162–8.

Rüegger, S. et al., 2015. The ribonucleotidyl transferase USIP-1 acts with SART3 to promote U6 snRNA recycling. *Nucleic acids research*, 43(6), pp.3344–57.

Salgado-Garrido, J. et al., 1999. Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin. *EMBO Journal*, 18(12), pp.3451–3462.

Sarov, M., Murray, J.I., et al., 2012. A genome-scale resource for in vivo tag-based protein function exploration in *C. elegans*. *Cell*, 150(4), pp.855–866.

Sarov, M., Murray, J.I., et al., 2012. A genome-scale resource for in vivo tag-based protein function exploration in *C. elegans*. *Cell*, 150(4), pp.855–66.

Scofield, D.G. & Lynch, M., 2008. Evolutionary diversification of the Sm family of RNA-associated proteins. *Molecular Biology and Evolution*, 25(11), pp.2255–2267.

Schweinfest, C.W. et al., 1997. CaSm: An Sm-like protein that contributes to the transformed state in cancer cells. *Cancer Research*, 57(14), pp.2961–2965.

Séraphin, B., 1995. Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs. *The EMBO journal*, 14(9), pp.2089–2098.

Shapira, M. et al., 2006. A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. *Proceedings of the National Academy of Sciences of the United States of America*, 103(38), pp.14086–14091.

Sharif, H. & Conti, E., 2013. Architecture of the Lsm1-7-Pat1 Complex: A Conserved Assembly in Eukaryotic mRNA Turnover. *Cell Reports*, 5(2), pp.283–291.

Sheikh-Hamad, D. & Gustin, M.C., 2004. MAP kinases and the adaptive response to hypertonicity: functional preservation from yeast to mammals. *American journal of physiology. Renal physiology*, 287(6), pp.F1102–F1110.

Shivers, R.P. et al., 2010. Phosphorylation of the conserved transcription factor ATF-7 by PMK-1 p38 MAPK regulates innate immunity in Caenorhabditis elegans. *PLoS Genetics*, 6(4).

Simonetta, S.H. & Golombek, D.A., 2007. An automated tracking system for Caenorhabditis elegans locomotor behavior and circadian studies application. *Journal of neuroscience methods*, 161(2), pp.273–80.

Singh, V. & Aballay, A., 2009. Regulation of DAF-16-mediated Innate Immunity in Caenorhabditis elegans * □. , 284(51), pp.35580–35587.

Sönnichsen, B. et al., 2005. Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. *Nature*, 434(7032), pp.462–469.

Spiller, M.P., Reijns, M.A.M. & Beggs, J.D., 2007. Requirements for nuclear localization of the Lsm2-8p complex and competition between nuclear and cytoplasmic Lsm complexes. *Journal of Cell Science*, 120(Pt 24), pp.4310–4320.

Sulston, J.E. & Horvitz, H.R., 1977. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Developmental Biology*, 56(1), pp.110–156.

Sun, Y. et al., 2011. A genome-wide RNAi screen identifies genes regulating the formation of P bodies in *C. elegans* and their functions in NMD and RNAi. *Protein & cell*, 2(11), pp.918–39.

Tabach, Y. et al., 2013. Identification of small RNA pathway genes using patterns of phylogenetic conservation and divergence. *Nature*, 493(7434), pp.694–698. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23364702.

Tepper, R.G. et al., 2013. PQM-1 complements DAF-16 as a key transcriptional regulator of DAF-2-mediated development and longevity. *Cell*, 154(3), pp.676–690.

Tharun, S., 2009. Roles of eukaryotic Lsm proteins in the regulation of mRNA function. *International review of cell and molecular biology*, 272, pp.149–189.

Tharun, S. et al., 2000. Yeast Sm-like proteins function in mRNA decapping and decay. *Nature*, 404(6777), pp.515–8.

Tissenbaum, H.A. & Guarente, L., 2001. Increased dosage of a sir-2 gene extends lifespan in Caenorhabditis elegans. *Nature*, 410(6825), pp.227–30.

Trapnell, C. et al., 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature biotechnology*, 31(1), pp.46–53.

Trapnell, C. et al., 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology*, 28(5), pp.511–5.

Trapnell, C., Pachter, L. & Salzberg, S.L., 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics (Oxford, England)*, 25(9), pp.1105–11.

Troemel, E.R. et al., 2006. p38 MAPK Regulates Expression of Immune Response Genes and Contributes to Longevity in *C. elegans* S. Kim, ed. *PLoS Genetics*, 2(11), p.e183.

Tsui, H.C., Leung, H.C. & Winkler, M.E., 1994. Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in Escherichia coli K-12. *Molecular microbiology*, 13(1), pp.35–49.

Vaquero, A., 2009. The conserved role of sirtuins in chromatin regulation. *International Journal of Developmental Biology*, 53(2-3), pp.303–322.

Vogel, J. & Luisi, B.F., 2011. Hfq and its constellation of RNA. *Nature reviews. Microbiology*, 9(8), pp.578–589.

Waaijers, S. et al., 2013. CRISPR/Cas9-targeted mutagenesis in Caenorhabditis elegans. *Genetics*, 195(3), pp.1187–1191.

Watson, P.M. et al., 2008. CaSm (LSm-1) Overexpression in Lung Cancer and Mesothelioma Is Required for Transformed Phenotypes.

Wilusz, C.J. & Wilusz, J., 2004. Bringing the role of mRNA decay in the control of gene expression into focus. *Trends in Genetics*, 20(10), pp.491–497.

Wilusz, C.J. & Wilusz, J., 2005. Eukaryotic Lsm proteins: lessons from bacteria. *Nature structural & molecular biology*, 12(12), pp.1031–6.

Will, C.L. & Lührmann, R., 2011. Spliceosome structure and function. *Cold Spring Harbor perspectives in biology*, 3(7).

Wu, D. et al., 2014. Lsm2 and Lsm3 bridge the interaction of the Lsm1-7 complex with Pat1 for decapping activation. *Cell research*, 24(2), pp.233–46.

Youngman, M.J., Rogers, Z.N. & Kim, D.H., 2011. A decline in p38 MAPK signaling underlies immunosenescence in Caenorhabditis elegans. *PLoS genetics*, 7(5), p.e1002082.

Zhang, X. et al., 2011. microRNAs play critical roles in the survival and recovery of Caenorhabditis elegans from starvation-induced L1 diapause. *Proceedings of the National Academy of Sciences of the United States of America*, 108(44), pp.17997–8002.

Zhou, L. et al., 2014. Crystal structures of the Lsm complex bound to the 3' end sequence of U6 small nuclear RNA. *Nature*, 506(7486), pp.116–20.