UNIVERSITY OF CALIFORNIA RIVERSIDE

Robustness of Early Gut Development Influences Metabolism and Physiology in *Caenorhabditis elegans*

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Cell, Molecular and Developmental Biology

by

Ivan Minchev Dimov

December 2021

Dissertation Committee
Dr. Morris Maduro, Chairperson
Dr. Fedor Karginov
Dr. Jeffrey Bachant

The Dissertation of Ivan Minchev Dimov is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENT

The text of this dissertation, in part, is a reprint of the material as it appears in 'The C. elegans intestine: organogenesis, digestion, and physiology' 2019 and 'Decreased robustness of embryonic gut specification causes dyslipidemia in *C. elegans*' a manuscript in preparation. The co-author Dr. Morris Maduro listed in these publications directed and supervised the research that forms the basis for this dissertation.

ABSTRACT OF THE DISSERTATION

Robustness of Early Gut Development Influences Metabolism and Physiology in *Caenorhabditis elegans*

by

Ivan Minchev Dimov

Doctor of Philosophy, Graduate Program in Cell, Molecular and Developmental Biology
University of California, Riverside, December 2021
Dr. Morris Maduro, Chairperson

During development organisms must decide whether to utilize energy or whether to store it. Previous work in Maduro lab has found evidence that early in the intestinal development there is a critical stage in which the timely activation of the transcription factor network, required for the development of the intestine establishes a mode of energy utilization, aimed at maximizing the production of progeny. When gut specification is compromised, by mutating transcription factors required for intestinal development, this changes the metabolism and physiology of those animals and they start to store energy even when the food is abundant. In this thesis, the phenotype associated with the compromised gut specification strains that we call Hypomorphic Gut Specification (HGS) strains is described. Also work was done to look at genes that are differentially expressed in HGS animals and two paralogous, reciprocally expressed genes in wild-type and HGS animals – *fstr-1* and *fstr-2* were characterized. Mutants were generated for *fstr-1* and *fstr-2* genes, and a double *fstr-1,2* mutant. The specific phenotypes associated with those mutations were described. Results show that animals in which *fstr-2* gene is deleted in otherwise normal genetic background, show characteristics associated with caloric

restriction – enlarged intestinal lumen, increased size of lipid droplets, increased amount of gut granules. These phenotypes are similar to those observed in HGS strains. In addition, a reduced brood count, increased lifespan of *fstr-2* mutants and increased resistance to toxins from *Pseudomonas aeruginosa* PA14 were observed. The double *fstr-1,2* mutant phenotype is similar to the one of *fstr-2* mutant. The expression of the *fstr* genes was also examined. *Fstr-1* appears to be a stress response gene, as it has low expression under normal conditions and increased in HGS strains and in caloric restriction. FSTR-2 appears to be a plasma membrane protein expressed mostly in the intestinal cells and based on the conditions in some neurons and muscle cells. These results provide insight about the functions of some of the genes differentially expressed, downstream of improper gut specification, and how they might modulate the metabolism of *C. elegans* and the switch from energy utilization mode to energy storage mode.

Table of contents

List of Figures

List of Tables

Chapter I introduction

C. elegans as a model system

Over 50 years ago Sydney Brenner proposed the free living nematode *Caenorhabditis elegans* as a model system to study questions in developmental biology and neurobiology[1]. Since then, this small worm (the size of adult *C. elegans* is about 1mm) has been studied extensively by hundreds of labs throughout the world and has had great impact in the areas of molecular biology, developmental biology, genetics, and neurobiology, among others. *C. elegans* animals are found around the world, often on rotting fruit, where they likely consume nutrient-rich substrates and bacteria.[2] As a model system, *C. elegans* has several advantages: The size of the animals is small which allows growth on Petri plates and one can grow a large number of worms in relatively short time. The normal diet of the worms is bacteria (as a standard food *E. coli* strain OP50 is used, which is about 55% protein). Other bacteria used in laboratory studies include *E. coli* modified to express double-stranded RNA (dsRNA) for feeding-based RNA interference[3], species that provide a richer food source[4] or pathogenic bacteria. Depending on what strain of *C. elegans* is used, and what kind of food is provided to the strain one can observe profound effects on metabolism, gene expression and growth phenotypes[5,6]. One other characteristic of *C. elegans* is its inability, similar to *Drosophila*, to synthesize cholesterol[7], which has to be provided to the culture media. The fact that the natural food of *C. elegans* are bacteria has been used to study host – pathogen interactions, for example by feeding the worms with a pathogenic bacteria such

as *Pseudomonas aeruginosa*[8]. Another advantage that facilitates experiments with *C. elegans* is that the animal has a transparent body, which allows observation in real time even with a light microscope of many intracellular processes[9]. Animals have a limited number of somatic cells: An adult hermaphrodite has 959 somatic nuclei and makes about 2000 germ cells. *C. elegans* also has a simple anatomy with no respiratory or circulatory systems. The embryonic and post-embryonic cell lineages are known[10].

In terms of genetics the animal exists as hermaphrodites (females) and males. *C elegans* as a small metazoan system is amenable to gene characterization by forward genetics (mutagenesis) and reverse genetics, for example RNAi and CRISPR/Cas9 genome editing[11-14]. Hermaphrodite worms can self-fertilize, which greatly facilitates genetic manipulation since any mutation one introduces can be maintained through self-fertilization[1] of F_1 progeny without mating. Another advantage for genetic experiments is the worm's short life cycle. At 20°C it takes about three days for *C. elegans* to develop from a fertilized egg to adult capable of laying its own fertilized eggs. These advantages can be used to do forward and reverse genetics in a very short time. The genome sequence of *C. elegans* was reported some 20 years ago. Historically, genetic approaches identified small numbers of genes important for the normal function of the animal by mutant phenotype. More recently, systems biology and "omics" work have begun to characterize entire metabolic enzyme networks and determine how the animal responds to changes in its environment[15].

C. elegans intestine

The *C. elegans* digestive tract is essentially a tube (Fig. 1.1)[4,16-19].

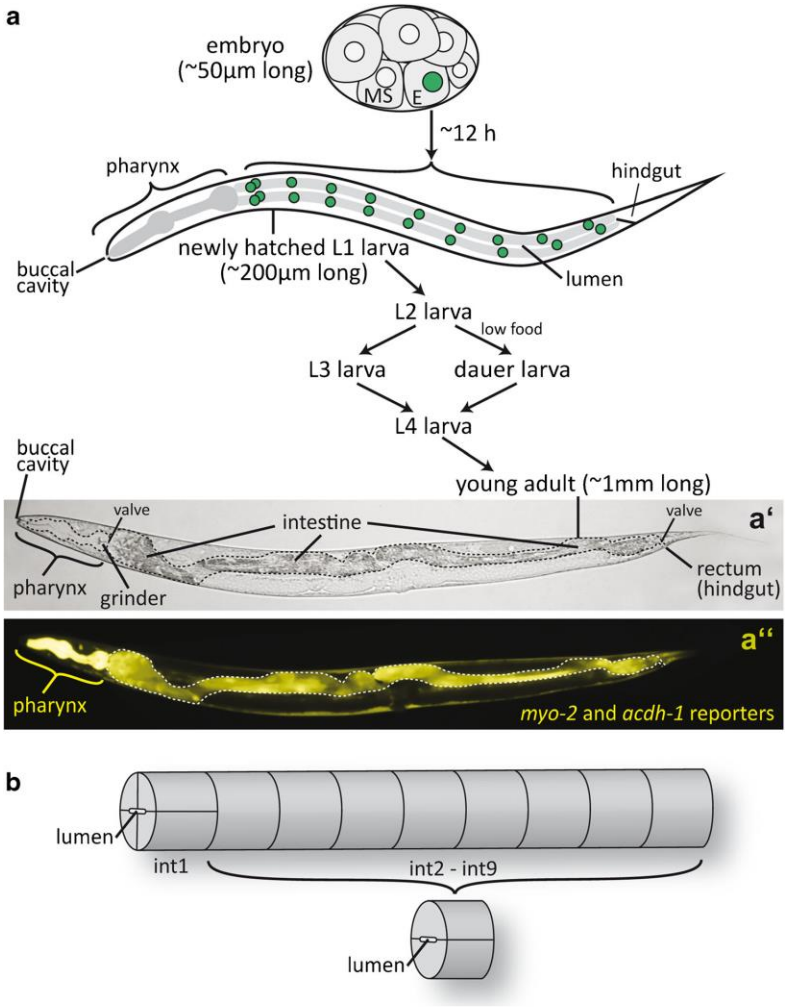


Fig. 1.1 Origin and basic structure of the C. elegans intestine.

a The gut originates as the E blastomere in the early embryo (*Sulston et al.1983*). After approximately 12 h at 25 °C, in the newly hatched first-stage juvenile (L1 larva), the intestine is found between the anterior pharynx and the posterior hindgut. Further development takes approximately 45 h to progress through the L2, L3, and L4 larval stages to adulthood. The alternative Dauer larva stage occurs during conditions of low food availability and/or high population density. In the adult, the intestine (outlined by dashed black lines) is visible by the granular contents of the gut cytoplasm in a light microscope image. In the fluorescence image directly below, expression of myo-2::GFP and acdh-1::GFP transgenes mark pharynx muscle and gut nuclei/cytoplasm, respectively (*MacNeil et al. 2013*; *Okkema et al*. *1993*). The gut is outlined as in the light micrograph.

b The basic structure of the intestine as a set of rings, consisting of four cells in the anterior-most int1 ring and two cells in int2 through int9. The lumen is located apically at the interface between cells in the rings (*Asan et al.* 2016)

The digestive tract consists of three parts: pharynx, intestine and hindgut[1] with valve cells connecting these parts. The pharynx is analogous to the esophagus in vertebrates. It contains 58 cells with its own epithelium, muscle cells and neurons.[17,20] Pumping of the pharynx muscles, and peristalsis through its lumen, directs food to the intestine.[21] At the posterior end of the pharynx is the grinder, a cuticular structure that mechanically breaks up food particles. The intestine is posterior to the grinder, separated from it by the pharyngeal-intestinal valve. The intestine consists of 20 cells arranged as "rings" (Fig. 1.1b). At the anterior, four cells form the ring int1. The remaining 16 cells are arranged in pairs, with each pair forming rings int2 through int9. The outside surface of the intestine is surrounded by a basement membrane that faces the pseudo coelomic space.[22] On the apical side of gut cells is the lumen, elliptical in cross-section, and which is lined with membranous microvilli that form the brush border (Fig. 1.2).

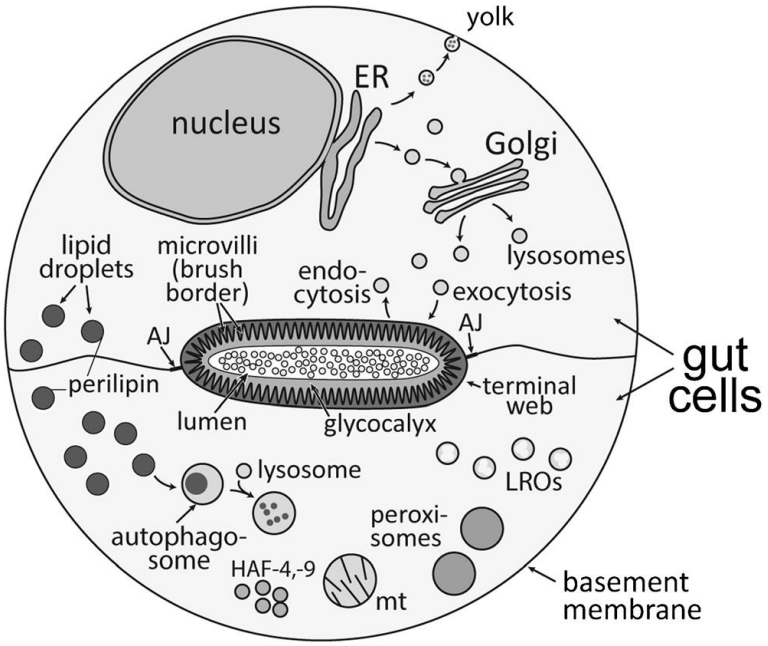


Fig. 1.2 The cytoplasm of the *C. elegans* intestinal cells. A pair of gut cells, part of the same ring is shown as a transverse section. Cellular organelles are shown as examples of the diverse structures and processes that are found in the intestinal cells. Sizes of the structures are approximate. Abbreviations: AJ, adherens junction; ER, endoplasmic reticulum; LRO, lysosome related organelle (gut granules)

The microvillar projections are supported by the terminal web, which consists primarily of an intracellular network of actin. Just outside of the microvilli, facing the luminal space is the glycocalyx, a region rich in glycoproteins that provides a physical barrier from pathogens, serves as an interface between the digestive enzymes and the grinded mass of bacteria coming from the pharynx and likely allows filtering of lumen contents into the intestinal cells.[23] Besides the stem cells, mentioned earlier, the intestine of *C. elegans* lack additional cell types like phagocytes and gland cells that are found in other invertebrates.[24]

At its posterior end, the intestine is connected to the hindgut through the intestinal-rectal valve. The hindgut provides passage of waste to the environment through the anus, regulated by three muscle sets, including two muscles that wrap around the ventral posterior of the intestine.[25]

Being a free-living nematode, the *C. elegans* intestine has significant differences in structure, compared to some parasitic nematodes. The differences in the intestinal structure in the parasitic nematodes are likely reflecting adaptations to their unique environments and way of life. For example, in some plant and animal parasitic nematodes, instead of a cellularized intestine, they have a syncytial one.[26,27] Within gut cells, the terminal web is absent in plant parasitic nematodes, and may only appear in the adult stage of animal parasitic nematodes.[26,27] Another example, whereas the *C. elegans* intestine is connected to a small number of muscles, the intestine of the pinworm *Aspiculuris tetraptera* is surrounded by a muscle fiber network.

Development of the *C. elegans* intestine

The *C. elegans* intestine originates as a single cell called E (endoderm) during early embryogenesis.[17] A series of synchronous mitoses results in 16 E descendants, four of which undergo an additional division to result in 20 cells (Fig. 1.3).[28]

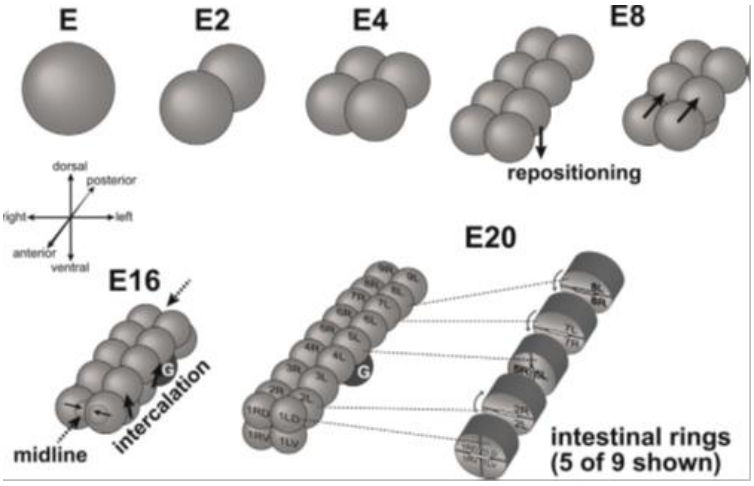


Fig. 1.3 Cell divisions of E cell to form *C. elegans* intestine. *Maduro et al. 2017*

The gut is generally invariant in cell number but occasionally there are 21 or 22 cells instead of 20.[19] Embryogenesis takes about 13 h at 25 °C[17]. Internalization of the gut occurs when the two E daughters migrate from the ventral part of the embryo into the interior. The gut primordium undergoes a set of carefully orchestrated cell movements and polarization to result in longitudinally arranged pairs of cells (in the anterior there are four cells).[19,22] In the latter half of embryogenesis the lumen of the intestine begins to form (Fig. 1.2)[22,28]. The lumen begins to form, first in discontinuous regions at 16E stage, when the intestinal cells start to polarize their contents apicobasally.[22] Lumen formation coincides with formation of adherens junctions around the lumen and between the int rings.[19] Upon hatching the gut is ready to function. After hatching, in the first larval stage, 14 additional nuclear divisions occur without cytokinesis, and a doubling of the nuclear content occurs in each larval molt, resulting in 34 nuclei (32C) in the same 20 cells that were present at hatching.[17]

The reproducible number of cells and nuclei in the *C. elegans* intestine suggests there has been strong selection for its maintenance. However, animals can accommodate slightly fewer numbers of intestinal E descendants or an excess with no apparent consequences to organogenesis and intestine function, at least in the laboratory[29-31]. In mutant backgrounds, that result in more than 30 intestinal cells, rather than the typical 20 cells a functional intestine still forms.[29], thus the morphogenesis of the gut does not require a fixed number of cells. The most commonly observed number of 20 gut cells, likely

represents a trade-off between resource allocation and maintaining enough cells to support integrity and flexibility of the lumen.[19]

Gene regulatory network that controls *C. elegans* gut development

Intestine formation begins with specification of the E blastomere as the gut progenitor. An embryo in which E is not specified develops into an arrested larva, which is shorter than the wild type and is missing all the gut tissue. The resulting larvae are inviable, likely because they have no way of absorbing nutrients.[32] Specification of the E blastomere occurs through activation of a gene regulatory network of maternal and zygotic transcription factors, which are expressed in temporal sequence (Fig. 1.4)[28,33]. The earliest-acting factors that specify E are the paralogous *end-1* and *end-3* genes, which encode structurally similar GATA type transcription factors that are partially redundant.[34,35]

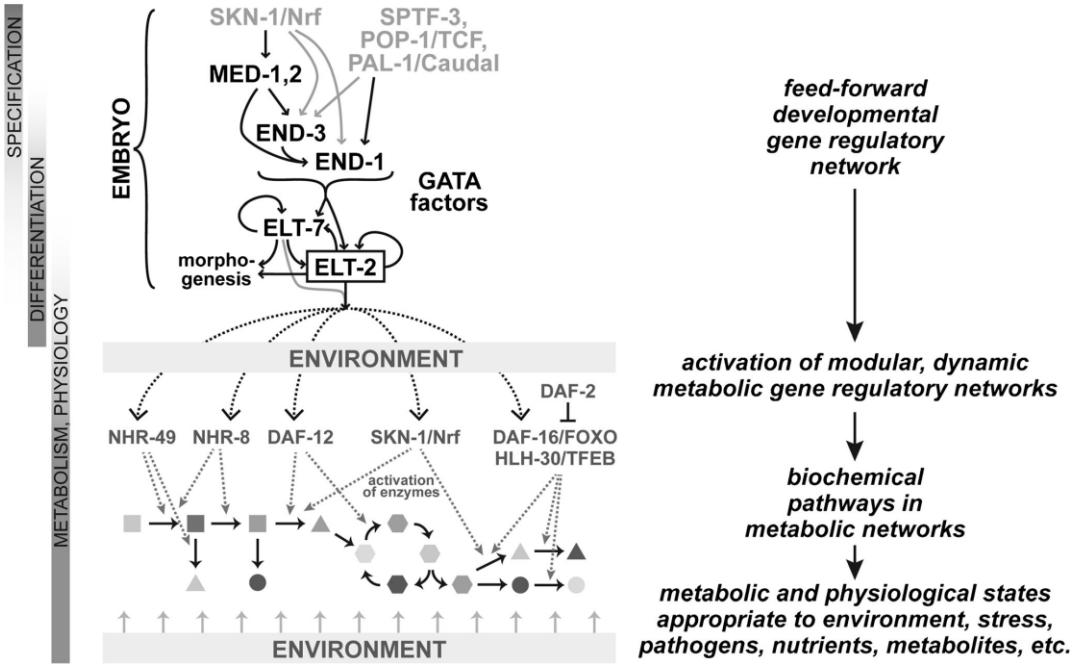


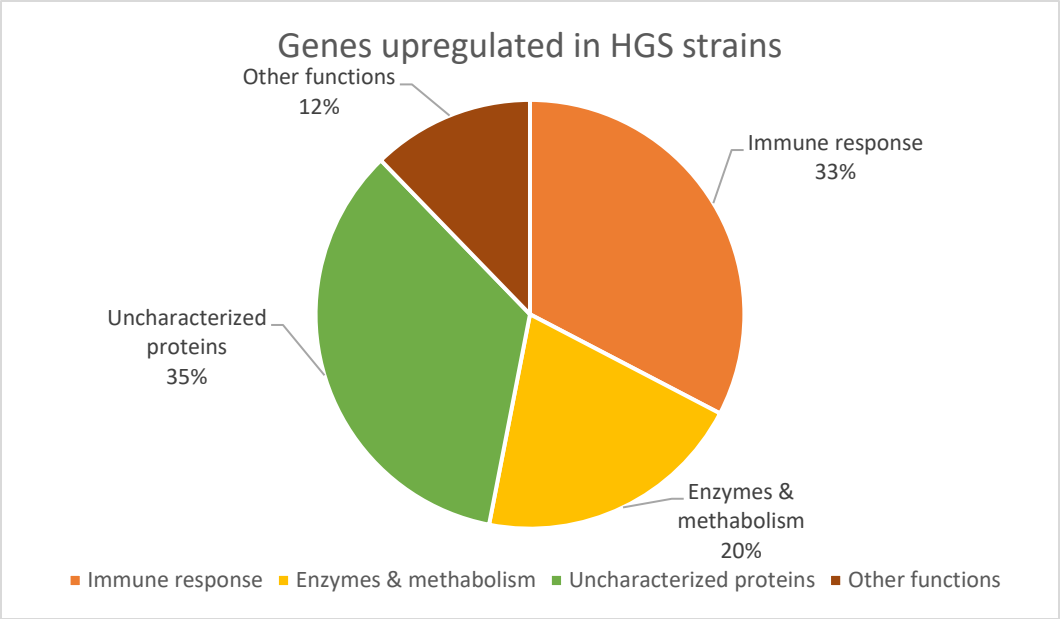
Fig. 1.4 Developmental gene regulatory network (top part) and its integration with metabolic networks (bottom part), which regulate the physiological responses to the changes in the environment. Environmental factors that influence gene expression include temperature, oxygen levels, levels of certain metabolites, presence of pathogens, ascaroside hormones.

Before the activation of the *end* genes, specification of the E cell begins when the four cell stage blastomere P_2 contacts the mother of the E cell, the EMS cell.[36] After EMS divides, the side in contact with P_2 (in normal embryos that's the posterior side) becomes the E cell.[37] Induction of E requires overlapping Wnt/MAPK/Src signaling that produces an endoderm-promoting state of the Wnt nuclear effector TCF/POP-1 in E.[38-41] In this state, POP-1 interacts with the divergent β-catenin SYS-1 to become a weak transcriptional activator of the gene *end-1*. In parallel with the Wnt/β-catenin asymmetry pathway, maternally provided SKN-1, a bZIP/homeodomain-like transcription factor, activates expression of the *med-1* and *med-2* genes.[42] They encode divergent type of GATA-type transcription factors. Together *med-1* and *med-2* genes contribute to the

activation of *end-1* and *end-3* genes.[43] Another factor that participates in the activation of *end* genes is Sp-1 like factor, SPTF3.[44] The result of the inputs of the mentioned above genes is a timely and robust activation of the *elt-2* GATA-factor gene, which is necessary and sufficient to specify and maintain gut cell identity.[45,46] Another gene, *elt-7*, which is paralogous to *elt-2* is also activated in the early E lineage. *Elt-7* is expressed in parallel with *elt-2*, and the two factors have overlapping functions.[44] In *elt-2* mutants, the gut is still specified in the embryo, but after hatching the animals arrest as larvae because they have aberrant intestines.[47] This mutant phenotype is more severe when *elt-7* is also missing, although when *elt-7* alone is missing there's no apparent phenotype.[48] Thus the gut specification in *C. elegans* is a result of the combined input of multiple redundant factors, working together to provide normal gut development.[29]

Prior work in the Maduro lab generated strains that have mutated some of the transcription factors that control gut specification (these strains were labeled, "hypomorphic gut specification" or HGS), leading to a stochastic delay of the activation of *elt*-2 gene. Those mutations are not 100% lethal and the animals that survive are able to form an apparently normal intestine, but they nonetheless exhibit multiple defects, such as increased or decreased number of intestine cells, enlarged lumen of the intestine which is normally a reaction to dietary restriction, and excessive storage of lipids.[29] RNA-seq data from the intestine cells of these strains revealed an altered transcription profile than the one in the control animals. The untimely activation of *elt-2* gene causes not only morphological, and physiological changes in the mutant animals, but also changes the intestinal gene expression profile in a way that resembles the changes

observed in animals under stress. It is known that ELT-2 is a major regulator of intestinal gene expression not only during development but also in the adult animals.[46] In our HGS strains, even though the adult levels of ELT-2 show no significant difference, compared to wild-type animals[29] the delayed activation of *elt-2* during development causes changes in the gene expression profile of the intestinal cells (Fig. 1.5)



From evolution point of view the genes upstream of the *elt-2*, *med* and *end* genes have orthologs in all members of *C. elegans* supergroup, with the *med* genes being much more diverse group than the *end* genes.[49] These are absent in more distant *Caenorhabditis* species such as C. *angaria* and C. *monodelphis*[49] suggesting that in those species another mechanism for activation of *elt-2* exists.[28] It is possible that activation of *elt-2* in E could specify gut directly, as overexpressed ELT-2 can achieve gut specification in *C. elegans* in the absence of *end-1* and *end-3*.[50] Another possibility is that the upstream factors in the

gut specification gene regulatory network may regulate intestine specification differently. That's true even for *Caenorhabditis* species that have *med* and *end* genes. For example, loss of *pop-1* function in C. elegans leads to formation of extra endoderm, the MS cells makes also endoderm, but in *C. briggsae* loss of the *pop-1* ortholog activity results in the loss of endoderm.[51,52]

The use of GATA transcription factors in specification and development of endodermal tissues is evolutionarily shared among protostomes and deuterostomes. In *Drosophila* the gut specification involves the GATA factors Serpent and dGATAe.[53,54] In vertebrates, GATA4/5/6 factors participate in the development of endoderm and other tissues.[55,56]

Assembly of the gut primordium results from the orientation of cell divisions that occur from one round of mitosis to the next within the E lineage, enforced by cell-cell signaling using the Van Gogh/VANG-1 pathway that keeps the primordial cells in a stereotypical alignment.[19,22] VANG-1 is a protein that inhibits Wnt signaling.[57] Between the stages at which there are 8 E descendants and 20 descendants, specific pairs of cells undergo migrations and reorientations within the primordium, guided by signaling of the Ephrin, Notch, and Netrin pathways.[19]

The lumen of the intestine is formed by targeting to the apical surface of the gut cells, components that support the terminal web. The details of how the microvilli get assembled is not very well understood. Two essential components that are known to support the cytoskeleton that support the luminal microvilli are the Ezrin/Radixin/Moesin membrane cytoskeleton linker protein, ERM-1, and the actin ACT-5.[58,59] Another two

factors required for lumen formation are established apicobasal polarity of the intestinal cell membrane and proper endosome trafficking. For example loss of the Clathrin *chc-1* results in mislocalization of ERM-1 and ACT-5 proteins that leads to disruption of the lumen structure.[60] Another example is loss of some of the enzymes for lipid biosynthesis, such as the sphingolipid biosynthetic enzyme LET-767 also disrupts apical localization of lumen components and hence lumen structure.[61]

C. elegans metabolism

Digestion is the first step in metabolizing the nutrients coming from the environment. It begins with mechanical and enzymatic breakdown of the cell walls and membranes of bacteria in the pharynx grinder.[23] From the pharynx the partially digested food proceeds into the intestinal lumen, where it's exposed to the action of enzymes that break down membranes and their lipid constituents. Such enzymes are lysozymes, saponins, lectins. They also enable access to the cellular contents and serve as a defense from pathogenic microbes.[62-65] Throughout the years, the efforts of researchers have been mainly focused on establishing cause-and-effect genetic relationships, by isolating and cloning different digestive enzymes, so experimental details in the literature about subcellular localization of digestion in *C. elegans* are somehow lacking. Another obstacle for studying the localization of certain proteins in the intestine is also the high autofluorescence of the gut.

However, as fundamental digestive mechanisms are conserved, it is likely that in the *C. elegans* gut, macromolecules are partially hydrolyzed in the lumen prior to endocytosis (i.e., extracellular digestion), after which further breakdown occurs when endocytic

vesicles internalize lumen contents and fuse with acidified lysosomes inside gut cells (intracellular digestion).[23,24]

There are many kinds of digestive enzymes expressed in *C. elegans*. The most abundantly expressed class are the proteases.[66] Examples of intestinal peptidases include astacins (NAS proteins). They are metalloproteases and are found in the digestive tract.[67] Another examples are the aspartic protease ASP-1, found both in lumen and intracellular lysosomes[68] and aminopeptidase P, which likely functions in the intracellular hydrolysis of proteins.[69] Enzymes for other macromolecules include an amylase-like enzyme, *C50B6.7*, orthologous of the human amylases, expressed in the intestine[66,70] and the deoxyribonuclease NUC-1.[71] The work done in the past for characterizing the digestion was focused, on working with single genes in many cases because of technical constraints. As the genomic methods became available, the more recent studies are focused to study the metabolome of *C. elegans* by identifying orthologous genes that encode enzymes and integrating information from metabolic pathway databases and published literature.[15,72,73] These types of studies, confirm that both intracellular and extracellular digestion occur in *C. elegans*. In one recent work, over 600 metabolic enzymes in almost 2000 reactions were identified and inferred to occur in the cytosol, mitochondria, or extracellularly based on prior work.[15] The results from this work are emerging metabolic networks, which are now available online, which incorporate the results from metabolomic studies and information from the literature about transport of proteins and metabolites into an online resource that can run simulations about how certain genes or mutations in them could change the metabolome of *C. elegans*.[15] When

discussing digestion in *C. elegans*, one class of biomolecules deserves special attention, that is dsRNA. Unlike DNA, dsRNA is not digested into monomers in *C. elegans*. Double stranded RNA molecules are internalized into the intestinal cells through the lumen, which causes systemic gene silencing by RNA interference.[74] Double stranded RNAs are imported into the intestinal cells by the transporter protein SID-2.[75] This is likely to be an adaptive form of immunity against RNA viruses, as C. elegans strains with defects in the RNAi pathway show increased susceptibility to virus replication.[76,77]

Another interesting aspect of the digestion in *C. elegans* is the dynamic acidification of the lumen. *C. elegans* is a filter feeder. Food travels through the intestine for a few minutes. Such short residence time of the food in the gut appears to be an adaptive mechanism for the short life cycle and the simple diet *C. elegans* has.[78] Enzymatic breakdown of macromolecules in animal digestive tracts usually requires acidic pH. Work in the last decade has elucidated details about the acidification of the lumen in *C. elegans*. The intestinal lumen is weakly acidic with an average pH around 4.4.[79] A region of higher acidity, 1–2 pH units lower, starts in the posterior third of the gut lumen, gets translocated to the very anterior over several seconds, and remains for several seconds more before raising in pH.[80]. During this time the posterior lumen reacidifies over 30s and the cycle repeats every 45s.[79,80] This wave of protons is tied to the defecation motor program (DMP), a set of regular body muscle contractions that propels food through the intestine posteriorly and out the anus.[81] Movement of the acidic region in the opposite direction results from the dynamic activity of proton pumps along the lumen.[80] The DMP itself is regulated by the propagation of a wave of intestinal calcium ions generated in gut

cells.[82,83] The defecation cycle has also a role in the nutrient uptake, the DMP is required for the absorption of fatty acids from the intestinal lumen into the gut cells.[84] Overall, normal *C. elegans* digestion requires, not only the presence of normal lumen structure and expression of different digestive enzymes but also coordinated and rhythmical physiological changes along the digestive tract that produces compartments with increased acidity.

Transcriptional and signaling regulation of *C. elegans* metabolism

The intestinal cells regulated the catabolic and anabolic processes in *C. elegans* through multiple pathways that integrate responses from nutrients and environmental conditions at multiple levels (Fig. 1.4). Transcription factors and regulatory pathways, many of which share conservation with similar pathways in humans, have been identified that modulate changes in biochemical flux and intestinal physiology.[33] One example is SKN-1/Nrf factor. This protein plays an important role early in development for the initial specification of the E cell fate in embryos and in the adult gut cells is a major regulator of the response to oxidative stress[85]. Another example for transcription factors regulating many metabolic pathways is nuclear hormone receptor family (NHR) of proteins. This family of proteins is greatly expanded in *C. elegans* and many of them are known to regulate different metabolic pathways.[33,86] These include, NHR-49/HNF4 a major regulator of stress response and breakdown of fatty acids by beta-oxidation;[87-89] NHR-8, which regulates cholesterol and bile acid metabolism[90] and the NHR DAF-12, which regulates enzymes in catabolism and citric cycle.[33,91,92] Two other proteins that work with

NHR-49 to regulate fat accumulation and the mediator complex protein MDT-15 and the sterol regulatory element-binding protein SBP-1.[93-95]

The environment also modulates metabolic functions during development. *C. elegans* larvae, which are in the first two larval stages, when exposed to conditions of limited food and high population density, enter an alternative third larval stage, known as Dauer larva (Fig. 1.1). In this stage development is paused and metabolism and physiology are altered until food becomes available.[96] A central component of both Dauer formation and metabolism is Insulin/IGF-1 signaling, or IIS.[97] In *C. elegans* the gene that encodes the insulin receptor is *daf-2*. Mutations in this gene result in constitutive Dauer formation at higher temperatures.[98] In adults, mutations in *daf-2* extend the lifespan of the animals, increases their abiotic stress and pathogenic resistance.[99] DAF-2 functions through a conserved phosphatidylinositol-3-kinase/protein kinase B/Target of Rapamycin (PI3K/Akt/TOR) pathway.[100,101] The extension of lifespan in *daf-2* mutants depends on the work of the DAF-16/FOXO transcription factor.[102] DAF-16/FOXO and the helix-loop-helix factor HLH-30/TFEB both translocate to intestinal nuclei to regulate genes that promote longevity and stress resistance.[103] The connection between the intestinal cells and the nervous system that explains how nutritional signal affect metabolism in the gut is mediated through the insulin like peptides (ILP).[104] There are 40 genes for ILPs in the *C. elegans* genome and the ILPs are working with DAF-16/FOXO in a dynamic signaling system to maintain metabolic homeostasis, depending on nutrient availability.[105]

On a cellular level, in the gut cells of *C. elegans*, different types of membrane bound storage organelles have been identified (Fig. 1.2). The most visible of these are gut granules, which contain birefringent material that is visible under polarized light.[106] Gut granules are membrane bound, lysosome related organelles (LROs) enriched in various substances including zinc.[107-109] Under UV light the gut granules fluoresce in blue because of the presence of the glycosylated anthranilic acid.[109] Gut cells also contain conventional acidified lysosomes that are enriched in chloride and calcium ions.[110,111] Lipid droplets are a major site of fat storage that can be detected by different methods, including staining with Oil Red O and dark field microscopy.[112-114] As in other animals including humans, lipid droplets are surrounded by perilipin, a lipid storage regulatory protein that recognizes the phospholipid monolayer surrounding the droplet and regulates lipid metabolism.[115-117] Gut cells also contain peroxisomes, which serve important functions in fatty acid metabolism.[118,119]

C. elegans as model organism has the advantage of being amenable to genetic manipulation and through the years of research many genetic tools have been developed. Another aspect of metabolism, the vesicular trafficking, including lysosomal, and endosomal trafficking have been studied extensively in *C. elegans*. The endosomal trafficking relies on conserved endosomal GTPases such as RAB-5 and RAB-10.[120-122] Maintenance of the lumen over time also relies on proper apical targeting of endosomes, which relies on the conserved adapter proteins, clathrin and AP-1.[60] Enterocytes in hermaphrodites contain vesicles carrying yolk lipoprotein from the endoplasmic reticulum to the basement membrane[123]. The latest addition to the genetic approach of

studying endosomal trafficking by identifying the GTPases, coat proteins that form the different types of vesicles is to also study the protein transporter proteins associated with certain types of vesicles. This leads to identifying novel endosomal types of vesicles, which are different from LROs, lipid droplets, and vesicles carrying yolk baes on the ABC transporters HAF-4 and HAF-9.[124]

C. elegans metabolism. Dietary restriction, lipid storage, autophagy, and life span.

In C. elegans, a complex relationship exists among diet, mobilization or storage of lipids, autophagy, and life span.[125-127] Dietary restriction (DR) in *C. elegans* can be achieved by different ways, such as diluting *E. coli* cultures, growing *C. elegans* in axenic liquid media, or using mutant strains like *eat-2* mutants**,** which have decreased pharyngeal pumping.[128-131] It's been shown that DR during larval stages slows the development and increases the post reproductive life span.[131] In C. elegans the major storage sites for lipid droplets are the intestinal cells and the epidermis and animals under DR conditions show increase in the number of lipid droplets stored.[131,132] Loss of germline stem cells also increases life span and this longevity correlates with increased lipids in the intestine.[133,134] Consistent with this, DR induces fatty acid synthesis gene expression in a manner dependent on DAF-16/FOXO[135], and in response the genes for catabolism of fats are activated through NHR-49 and SKN-1.[136,137] Autophagy also plays role in the process with the endosomal trafficking, which delivers the lipid droplets to lysosomes of the intestinal cells for lipophagy, and in dietary restricted animals, autophagy also increases the life span of the animals.[125,130] Recent studies point to a causative role of yolk

lipoprotein production in the regulation of autophagy and remodeling of lipids, which is proposed to promote long lifespan. In animals undergoing dietary restriction, longevity results from decreased expression of vitellogenin genes, and activation of DAF-16/FOXO.[127] In germline-less animals, the retention of yolk causes activation of SKN-1, which in turn promotes remodeling of lipids to maintain lipid homeostasis[137] As animals age, the intestine undergoes progressive detrimental changes.[138] Ezcurra et al. 2018, proposed that the pathologies observed as the intestinal cells age are due to constant vitellogenesis that "runs off" the intestinal cells by consuming cell contents from within with autophagy. In this model, the conversion of intestinal biomass into yolk into late adulthood results in the reduction of volume of intestinal cells and widening of the lumen (both of which are seen in senescing adults) and extracellular buildup of yolk pools. Autophagy has a role in promoting senescence by converting gut cell biomass into yolk; hence, the role of intestinal autophagy in life span depends on the role it plays in promoting lipoprotein synthesis or lipid breakdown.[130,139]

The HGS strains, mentioned above, generated in our lab exhibit some of the features of animals under DR and aged animals. Grown in normal conditions, with abundance of food and no stress factors they have increased lipid storage a sign of DR and extended lumen of the intestine, consistent with the shrinking of the intestinal cells, which normally happen as the animals age. The only problem they have experienced was the delayed activation of *elt-2* transcription factor due to mutations in *med* and *end* GATA transcription factors that work early in development to specify endoderm. We may say that they have experienced a genetic stress early in development, similar to what happens

in mutants that have disrupted cellular respiration. It was the work with one of those mutants *clk-1* that inspired the work for this thesis. *Clk-1* mutants have disrupted cellular respiration which results in all rhythmic functions being slowed down compared to wild-type animals and significantly changed gene expression profile. The RNA seq data from our HGS strains also showed changed transcriptional profile in a way different of the *clk-1* mutants but the reason is again a disruption of the normal development, and gene expression pattern. Every organism when it completes its development under normal for the organism conditions establishes a certain gene expression profile in each of its tissues which is the expression of the appropriate set of genes and their appropriate expression levels so the cells can function normally. In the case of the *C. elegans* intestinal cells when development has occurred normally the established gene expression profile is one that maximizes utilization of energy, and it provides all the necessary components to produce offspring as it is the major evolution imperative. In the HGS strains the disruption in the development caused by the untimely activation of a key transcription factor results in changed gene expression in the intestinal cells in a way that now maximizes the storage of energy as if the animal is under dietary restriction even though the food is abundant. Our RNA seq data showed changed expression levels in many stress response genes (Fig. 1.5), which supports our hypothesis for this "switch" from "energy utilization" mode to "energy storage" mode. The data also showed differential gene expression for many genes with still unknown function. In the next chapter we will show some of the work we did to characterize the HGS strains and some of the genes with unknown function which were differentially expressed. The focus of this work

became two genes, with unknown function, *fstr-1,* and *fstr-2*. Their expression pattern got our interest because of their reciprocal expression between control and HGS strains: Relative to each other, *fstr-2* has increased expression in control animals, while *fstr-1* has increased expression in HGS animals. The genes got their names from an earlier study of *clk-1* mutants[140] as RNAi on those genes caused alleviation of some of the slowed rhythmical functions in *clk-1* mutants, hence the name *fstr* (faster).

References

1. Corsi, A.K., Wightman, B. & Chalfie, M. A Transparent Window into Biology: A Primer on Caenorhabditis elegans. *Genetics* **200**, 387-407 (2015).

2. Kiontke, K. & Sudhaus, W. Ecology of Caenorhabditis species. *WormBook*, 1-14 (2006).

3. Kamath, R.S., Martinez-Campos, M., Zipperlen, P., Fraser, A.G. & Ahringer, J. Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in Caenorhabditis elegans. *Genome Biol* **2**, RESEARCH0002 (2001).

4. MacNeil, L.T., Watson, E., Arda, H.E., Zhu, L.J. & Walhout, A.J. Diet-induced developmental acceleration independent of TOR and insulin in C. elegans. *Cell* **153**, 240-52 (2013).

5. Celen, I., Doh, J.H. & Sabanayagam, C.R. Effects of liquid cultivation on gene expression and phenotype of C. elegans. *BMC Genomics* **19**, 562 (2018).

6. Reinke, S.N., Hu, X., Sykes, B.D. & Lemire, B.D. Caenorhabditis elegans diet significantly affects metabolic profile, mitochondrial DNA levels, lifespan and brood size. *Mol Genet Metab* **100**, 274-82 (2010).

7. Vinci, G., Xia, X. & Veitia, R.A. Preservation of genes involved in sterol metabolism in cholesterol auxotrophs: facts and hypotheses. *PLoS One* **3**, e2883 (2008).

8. Tan, M.W., Mahajan-Miklos, S. & Ausubel, F.M. Killing of Caenorhabditis elegans by Pseudomonas aeruginosa used to model mammalian bacterial pathogenesis. *Proc Natl Acad Sci U S A* **96**, 715-20 (1999).

9. Hutter, H. Fluorescent protein methods: strategies and applications. *Methods Cell Biol* **107**, 67-92 (2012).

10. Sulston, J.E. & Horvitz, H.R. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Dev Biol* **56**, 110-56 (1977).

11. Ahringer, J. Reverse genetics (April 6, 2006). in *WormBook* (ed. Community, T.C.e.R.) (2006).

12. Jorgensen, E.M. & Mango, S.E. The art and design of genetic screens: caenorhabditis elegans. *Nat Rev Genet* **3**, 356-69 (2002).

13. Waaijers, S. & Boxem, M. Engineering the Caenorhabditis elegans genome with CRISPR/Cas9. *Methods* **68**, 381-8 (2014).

14. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806-11 (1998).

15. Yilmaz, L.S. & Walhout, A.J. A Caenorhabditis elegans Genome-Scale Metabolic Network Model. *Cell Syst* **2**, 297-311 (2016).

16. Dimov, I. & Maduro, M.F. The C. elegans intestine: organogenesis, digestion, and physiology. *Cell Tissue Res* **377**, 383-396 (2019).

17. Sulston, J.E., Schierenberg, E., White, J.G. & Thomson, J.N. The embryonic cell lineage of the nematode Caenorhabditis elegans. *Dev Biol* **100**, 64-119 (1983).

18. Okkema, P.G., Harrison, S.W., Plunger, V., Aryana, A. & Fire, A. Sequence requirements for myosin gene expression and regulation in Caenorhabditis elegans. *Genetics* **135**, 385-404 (1993).

19. Asan, A., Raiders, S.A. & Priess, J.R. Morphogenesis of the C. elegans Intestine Involves Axon Guidance Genes. *PLoS Genet* **12**, e1005950 (2016).

20. Mango, S.E. The C. elegans pharynx: a model for organogenesis. *WormBook*, 1-26 (2007).

21. Song, B.M. & Avery, L. The pharynx of the nematode C. elegans: A model system for the study of motor control. *Worm* **2**, e21833 (2013).

22. Leung, B., Hermann, G.J. & Priess, J.R. Organogenesis of the Caenorhabditis elegans intestine. *Dev Biol* **216**, 114-34 (1999).

23. McGhee, J.D. The C. elegans intestine. *WormBook*, 1-36 (2007).

24. Karasov, W.H. & Douglas, A.E. Comparative digestive physiology. *Compr Physiol* **3**, 741-83 (2013).

25. White, J.G., Southgate, E., Thomson, J.N. & Brenner, S. The structure of the nervous system of the nematode Caenorhabditis elegans. *Philos Trans R Soc Lond B Biol Sci* **314**, 1-340 (1986).

26. Byers, J.R. & Anderson, R.V. Morphology and Ultrastructure of the Intestine in a Plant-Parasitic Nematode, Tylenchorhynchus dubius. *J Nematol* **5**, 28-37 (1973).

27. Colley, F.C. Strongyloides myopotomi: fine structure of the body wall and alimentary tract of the adult and third-stage larva. *Exp Parasitol* **28**, 420-34 (1970).

28. Maduro, M.F. Gut development in C. elegans. *Semin Cell Dev Biol* **66**, 3-11 (2017).

29. Choi, H., Broitman-Maduro, G. & Maduro, M.F. Partially compromised specification causes stochastic effects on gut development in C. elegans. *Dev Biol* **427**, 49-60 (2017).

30. Clucas, C., Cabello, J., Bussing, I., Schnabel, R. & Johnstone, I.L. Oncogenic potential of a C.elegans cdc25 gene is demonstrated by a gain-of-function allele. *Embo J* **21**, 665-74 (2002).

31. Lee, Y.U., Son, M., Kim, J., Shim, Y.H. & Kawasaki, I. CDC-25.2, a C. elegans ortholog of cdc25, is essential for the progression of intestinal divisions. *Cell Cycle* **15**, 654-66 (2016).

32. Owraghi, M., Broitman-Maduro, G., Luu, T., Roberson, H. & Maduro, M.F. Roles of the Wnt effector POP-1/TCF in the C. elegans endomesoderm specification gene network. *Dev Biol* **340**, 209-21 (2010).

33. Watson, E. & Walhout, A.J. Caenorhabditis elegans metabolic gene regulatory networks govern the cellular economy. *Trends Endocrinol Metab* **25**, 502-8 (2014).

34. Zhu, J. *et al.* end-1 encodes an apparent GATA factor that specifies the endoderm precursor in Caenorhabditis elegans embryos. *Genes Dev* **11**, 2883-96 (1997).

35. Maduro, M. *et al.* Genetic redundancy in endoderm specification within the genus Caenorhabditis. *Dev Biol* **284**, 509-522 (2005).

36. Goldstein, B. Induction of gut in Caenorhabditis elegans embryos. *Nature* **357**, 255-7 (1992).

37. Goldstein, B. Establishment of gut fate in the E lineage of C. elegans: the roles of lineage-dependent mechanisms and cell interactions. *Development* **118**, 1267-77 (1993).

38. Rocheleau, C.E. *et al.* Wnt signaling and an APC-related gene specify endoderm in early C. elegans embryos. *Cell* **90**, 707-16 (1997).

39. Shetty, P., Lo, M.C., Robertson, S.M. & Lin, R. C. elegans TCF protein, POP-1, converts from repressor to activator as a result of Wnt-induced lowering of nuclear levels. *Dev Biol* **285**, 584-92 (2005).

40. Shin, T.H. *et al.* MOM-4, a MAP kinase kinase kinase-related protein, activates WRM-1/LIT-1 kinase to transduce anterior/posterior polarity signals in C. elegans. *Mol Cell* **4**, 275-80 (1999).

41. Thorpe, C.J., Schlesinger, A. & Bowerman, B. Wnt signalling in Caenorhabditis elegans: regulating repressors and polarizing the cytoskeleton. *Trends Cell Biol* **10**, 10-7 (2000).

42. Maduro, M.F., Meneghini, M.D., Bowerman, B., Broitman-Maduro, G. & Rothman, J.H. Restriction of mesendoderm to a single blastomere by the combined action of SKN-1 and a GSK-3beta homolog is mediated by MED-1 and -2 in C. elegans. *Mol Cell* **7**, 475-85 (2001).

43. Broitman-Maduro, G., Maduro, M.F. & Rothman, J.H. The noncanonical binding site of the MED-1 GATA factor defines differentially regulated target genes in the C. elegans mesendoderm. *Dev Cell* **8**, 427-33 (2005).

44. Sullivan-Brown, J.L. *et al.* Identifying Regulators of Morphogenesis Common to Vertebrate Neural Tube Closure and Caenorhabditis elegans Gastrulation. *Genetics* **202**, 123-39 (2016).

45. Fukushige, T., Hendzel, M.J., Bazett-Jones, D.P. & McGhee, J.D. Direct visualization of the elt-2 gut-specific GATA factor binding to a target promoter inside the living Caenorhabditis elegans embryo. *Proc Natl Acad Sci U S A* **96**, 11883-8 (1999).

46. McGhee, J.D. *et al.* ELT-2 is the predominant transcription factor controlling differentiation and function of the C. elegans intestine, from embryo to adult. *Dev Biol* **327**, 551-65 (2009).

47. Fukushige, T., Hawkins, M.G. & McGhee, J.D. The GATA-factor elt-2 is essential for formation of the Caenorhabditis elegans intestine. *Dev Biol* **198**, 286-302 (1998).

48. Sommermann, E.M., Strohmaier, K.R., Maduro, M.F. & Rothman, J.H. Endoderm development in Caenorhabditis elegans: the synergistic action of ELT-2 and -7 mediates the specification-->differentiation transition. *Dev Biol* **347**, 154-66 (2010).

49. Maduro, M.F. Evolutionary Dynamics of the SKN-1 --> MED --> END-1,3 Regulatory Gene Cascade in Caenorhabditis Endoderm Specification. *G3 (Bethesda)* **10**, 333-356 (2020).

50. Wiesenfahrt, T. *et al.* The Function and Regulation of the GATA Factor ELT-2 in the C. elegans Endoderm. *Development* (2015).

51. Lin, R., Thompson, S. & Priess, J.R. pop-1 encodes an HMG box protein required for the specification of a mesoderm precursor in early C. elegans embryos. *Cell* **83**, 599-609 (1995).

52. Lin, K.T., Broitman-Maduro, G., Hung, W.W., Cervantes, S. & Maduro, M.F. Knockdown of SKN-1 and the Wnt effector TCF/POP-1 reveals differences in endomesoderm specification in C. briggsae as compared with C. elegans. *Dev Biol* **325**, 296-306 (2009).

53. Murakami, R., Okumura, T. & Uchiyama, H. GATA factors as key regulatory molecules in the development of Drosophila endoderm. *Dev Growth Differ* **47**, 581-9 (2005).

54. Okumura, T., Matsumoto, A., Tanimura, T. & Murakami, R. An endoderm-specific GATA factor gene, dGATAe, is required for the terminal differentiation of the Drosophila endoderm. *Dev Biol* **278**, 576-86 (2005).

55. Lentjes, M.H. *et al.* The emerging role of GATA transcription factors in development and disease. *Expert Rev Mol Med* **18**, e3 (2016).

56. Zorn, A.M. & Wells, J.M. Vertebrate endoderm development and organ formation. *Annu Rev Cell Dev Biol* **25**, 221-51 (2009).

57. Mentink, R.A. *et al.* The planar cell polarity protein VANG-1/Vangl negatively regulates Wnt/beta-catenin signaling through a Dvl dependent mechanism. *PLoS Genet* **14**, e1007840 (2018).

58. Gobel, V., Barrett, P.L., Hall, D.H. & Fleming, J.T. Lumen morphogenesis in C. elegans requires the membrane-cytoskeleton linker erm-1. *Dev Cell* **6**, 865-73 (2004).

59. MacQueen, A.J. *et al.* ACT-5 is an essential Caenorhabditis elegans actin required for intestinal microvilli formation. *Mol Biol Cell* **16**, 3247-59 (2005).

60. Zhang, H. *et al.* Clathrin and AP-1 regulate apical polarity and lumen formation during C. elegans tubulogenesis. *Development* **139**, 2071-83 (2012).

61. Zhang, H. *et al.* Apicobasal domain identities of expanding tubular membranes depend on glycosphingolipid biosynthesis. *Nat Cell Biol* **13**, 1189-201 (2011).

62. Gravato-Nobre, M.J., Vaz, F., Filipe, S., Chalmers, R. & Hodgkin, J. The Invertebrate Lysozyme Effector ILYS-3 Is Systemically Activated in Response to Danger Signals and Confers Antimicrobial Protection in C. elegans. *PLoS Pathog* **12**, e1005826 (2016).

63. Kandasamy, S., Khan, W., Evans, F., Critchley, A.T. & Prithiviraj, B. Tasco(R): a product of Ascophyllum nodosum enhances immune response of Caenorhabditis elegans against Pseudomonas aeruginosa infection. *Mar Drugs* **10**, 84-105 (2012).

64. Mallo, G.V. *et al.* Inducible antibacterial defense system in C. elegans. *Curr Biol* **12**, 1209-14 (2002).

65. Tarr, D.E. Distribution and characteristics of ABFs, cecropins, nemapores, and lysozymes in nematodes. *Dev Comp Immunol* **36**, 502-20 (2012).

66. McGhee, J.D. *et al.* The ELT-2 GATA-factor and the global regulation of transcription in the C. elegans intestine. *Dev Biol* **302**, 627-45 (2007).

67. Mohrlen, F., Hutter, H. & Zwilling, R. The astacin protein family in Caenorhabditis elegans. *Eur J Biochem* **270**, 4909-20 (2003).

68. Tcherepanova, I., Bhattacharyya, L., Rubin, C.S. & Freedman, J.H. Aspartic proteases from the nematode Caenorhabditis elegans. Structural organization and developmental and cell-specific expression of asp-1. *J Biol Chem* **275**, 26359-69 (2000).

69. Laurent, V., Brooks, D.R., Coates, D. & Isaac, R.E. Functional expression and characterization of the cytoplasmic aminopeptidase P of Caenorhabditis elegans. *Eur J Biochem* **268**, 5430-8 (2001).

70. Mulder, N.J. *et al.* The InterPro Database, 2003 brings increased coverage and new features. *Nucleic Acids Res* **31**, 315-8 (2003).

71. Hevelone, J. & Hartman, P.S. An endonuclease from Caenorhabditis elegans: partial purification and characterization. *Biochem Genet* **26**, 447-61 (1988).

72. Gebauer, J. *et al.* A Genome-Scale Database and Reconstruction of Caenorhabditis elegans Metabolism. *Cell Syst* **2**, 312-22 (2016).

73. Witting, M. *et al.* Modeling Meets Metabolomics-The WormJam Consensus Model as Basis for Metabolic Studies in the Model Organism Caenorhabditis elegans. *Front Mol Biosci* **5**, 96 (2018).

74. Timmons, L. & Fire, A. Specific interference by ingested dsRNA. *Nature* **395**, 854 (1998).

75. Winston, W.M., Sutherlin, M., Wright, A.J., Feinberg, E.H. & Hunter, C.P. Caenorhabditis elegans SID-2 is required for environmental RNA interference. *Proc Natl Acad Sci U S A* **104**, 10565-70 (2007).

76. Ashe, A. *et al.* A deletion polymorphism in the Caenorhabditis elegans RIG-I homolog disables viral RNA dicing and antiviral immunity. *Elife* **2**, e00994 (2013).

77. Gammon, D.B. *et al.* The Antiviral RNA Interference Response Provides Resistance to Lethal Arbovirus Infection and Vertical Transmission in Caenorhabditis elegans. *Curr Biol* **27**, 795-806 (2017).

78. Ghafouri, S. & McGhee, J.D. Bacterial residence time in the intestine of Caenorhabditis elegans. *Nematology* **9**, 87-91 (2007).

79. Allman, E., Johnson, D. & Nehrke, K. Loss of the apical V-ATPase a-subunit VHA-6 prevents acidification of the intestinal lumen during a rhythmic behavior in C. elegans. *Am J Physiol Cell Physiol* **297**, C1071-81 (2009).

80. Bender, A. *et al.* Novel acid-activated fluorophores reveal a dynamic wave of protons in the intestine of Caenorhabditis elegans. *ACS Chem Biol* **8**, 636-42 (2013).

81. Thomas, J.H. Genetic analysis of defecation in Caenorhabditis elegans. *Genetics* **124**, 855-72 (1990).

82. Nehrke, K., Denton, J. & Mowrey, W. Intestinal Ca2+ wave dynamics in freely moving C. elegans coordinate execution of a rhythmic motor program. *Am J Physiol Cell Physiol* **294**, C333-44 (2008).

83. Teramoto, T. & Iwasaki, K. Intestinal calcium waves coordinate a behavioral motor program in C. elegans. *Cell Calcium* **40**, 319-27 (2006).

84. Sheng, M. *et al.* Aberrant fat metabolism in Caenorhabditis elegans mutants with defects in the defecation motor program. *PLoS One* **10**, e0124515 (2015).

85. An, J.H. & Blackwell, T.K. SKN-1 links C. elegans mesendodermal specification to a conserved oxidative stress response. *Genes Dev* **17**, 1882-93 (2003).

86. Arda, H.E. *et al.* Functional modularity of nuclear hormone receptors in a Caenorhabditis elegans metabolic gene regulatory network. *Mol Syst Biol* **6**, 367 (2010).

87. Goh, G.Y.S. *et al.* NHR-49/HNF4 integrates regulation of fatty acid metabolism with a protective transcriptional response to oxidative stress and fasting. *Aging Cell* **17**, e12743 (2018).

88. Hu, Q., D'Amora, D.R., MacNeil, L.T., Walhout, A.J.M. & Kubiseski, T.J. The Caenorhabditis elegans Oxidative Stress Response Requires the NHR-49 Transcription Factor. *G3 (Bethesda)* **8**, 3857-3863 (2018).

89. Van Gilst, M.R., Hadjivassiliou, H., Jolly, A. & Yamamoto, K.R. Nuclear hormone receptor NHR-49 controls fat consumption and fatty acid composition in C. elegans. *PLoS Biol* **3**, e53 (2005).

90. Magner, D.B. *et al.* The NHR-8 nuclear receptor regulates cholesterol and bile acid homeostasis in C. elegans. *Cell Metab* **18**, 212-24 (2013).

91. Hochbaum, D. *et al.* DAF-12 regulates a connected network of genes to ensure robust developmental decisions. *PLoS Genet* **7**, e1002179 (2011).

92. McCormick, M., Chen, K., Ramaswamy, P. & Kenyon, C. New genes that extend Caenorhabditis elegans' lifespan in response to reproductive signals. *Aging Cell* (2011).

93. Goh, G.Y. *et al.* The conserved Mediator subunit MDT-15 is required for oxidative stress responses in Caenorhabditis elegans. *Aging Cell* **13**, 70-9 (2014).

94. Taubert, S., Van Gilst, M.R., Hansen, M. & Yamamoto, K.R. A Mediator subunit, MDT-15, integrates regulation of fatty acid metabolism by NHR-49-dependent and -independent pathways in C. elegans. *Genes Dev* **20**, 1137-49 (2006).

95. Moreno-Arriola, E., El Hafidi, M., Ortega-Cuellar, D. & Carvajal, K. AMP-Activated Protein Kinase Regulates Oxidative Metabolism in Caenorhabditis elegans through the NHR-49 and MDT-15 Transcriptional Regulators. *PLoS One* **11**, e0148089 (2016).

96. Androwski, R.J., Flatt, K.M. & Schroeder, N.E. Phenotypic plasticity and remodeling in the stress-induced Caenorhabditis elegans dauer. *Wiley Interdiscip Rev Dev Biol* **6**(2017).

97. Murphy, C.T. & Hu, P.J. Insulin/insulin-like growth factor signaling in C. elegans. *WormBook*, 1-43 (2013).

98. Kimura, K.D., Tissenbaum, H.A., Liu, Y. & Ruvkun, G. daf-2, an insulin receptor-like gene that regulates longevity and diapause in Caenorhabditis elegans. *Science* **277**, 942-6 (1997).

99. Kenyon, C., Chang, J., Gensch, E., Rudner, A. & Tabtiang, R. A C. elegans mutant that lives twice as long as wild type. *Nature* **366**, 461-4 (1993).

100. Chen, D. *et al.* Germline signaling mediates the synergistically prolonged longevity produced by double mutations in daf-2 and rsks-1 in C. elegans. *Cell Rep* **5**, 1600-10 (2013).

101. Paradis, S. & Ruvkun, G. Caenorhabditis elegans Akt/PKB transduces insulin receptor-like signals from AGE-1 PI3 kinase to the DAF-16 transcription factor. *Genes Dev* **12**, 2488-98 (1998).

102. Lin, K., Hsin, H., Libina, N. & Kenyon, C. Regulation of the Caenorhabditis elegans longevity protein DAF-16 by insulin/IGF-1 and germline signaling. *Nat Genet* **28**, 139-45 (2001).

103. Lin, X.X. *et al.* DAF-16/FOXO and HLH-30/TFEB function as combinatorial transcription factors to promote stress resistance and longevity. *Nat Commun* **9**, 4400 (2018).

104. Hung, W.L., Wang, Y., Chitturi, J. & Zhen, M. A Caenorhabditis elegans developmental decision requires insulin signaling-mediated neuron-intestine communication. *Development* **141**, 1767-79 (2014).

105. Kaplan, R.E.W., Maxwell, C.S., Codd, N.K. & Baugh, L.R. Pervasive Positive and Negative Feedback Regulation of Insulin-Like Signaling in Caenorhabditis elegans. *Genetics* **211**, 349-361 (2019).

106. Hermann, G.J. *et al.* Genetic analysis of lysosomal trafficking in Caenorhabditis elegans. *Mol Biol Cell* **16**, 3273-88 (2005).

107. Roh, H.C., Collier, S., Guthrie, J., Robertson, J.D. & Kornfeld, K. Lysosome-related organelles in intestinal cells are a zinc storage site in C. elegans. *Cell Metab* **15**, 88-99 (2012).

108. Hermann, G.J. *et al.* C. elegans BLOC-1 functions in trafficking to lysosome-related gut granules. *PLoS One* **7**, e43043 (2012).

109. Coburn, C. & Gems, D. The mysterious case of the C. elegans gut granule: death fluorescence, anthranilic acid and the kynurenine pathway. *Front Genet* **4**, 151 (2013).

110. Chakraborty, K., Leung, K. & Krishnan, Y. High lumenal chloride in the lysosome is critical for lysosome function. *Elife* **6**(2017).

111. Narayanaswamy, N. *et al.* A pH-correctable, DNA-based fluorescent reporter for organellar calcium. *Nat Methods* **16**, 95-102 (2019).

112. Fouad, A.D. *et al.* Quantitative Assessment of Fat Levels in Caenorhabditis elegans Using Dark Field Microscopy. *G3 (Bethesda)* **7**, 1811-1818 (2017).

113. Lapierre, L.R. *et al.* Autophagy genes are required for normal lipid levels in C. elegans. *Autophagy* **9**, 278-86 (2013).

114. O'Rourke, E.J., Soukas, A.A., Carr, C.E. & Ruvkun, G. C. elegans major fats are stored in vesicles distinct from lysosome-related organelles. *Cell Metab* **10**, 430-5 (2009).

115. Beller, M. *et al.* PERILIPIN-dependent control of lipid droplet structure and fat storage in Drosophila. *Cell Metab* **12**, 521-32 (2010).

116. Chughtai, A.A. *et al.* Perilipin-related protein regulates lipid metabolism in C. elegans. *PeerJ* **3**, e1213 (2015).

117. Copic, A. *et al.* A giant amphipathic helix from a perilipin that is adapted for coating lipid droplets. *Nat Commun* **9**, 1332 (2018).

118. Artyukhin, A.B. *et al.* Metabolomic "Dark Matter" Dependent on Peroxisomal beta-Oxidation in Caenorhabditis elegans. *J Am Chem Soc* **140**, 2841-2852 (2018).

119. Yokota, S. *et al.* Peroxisomes of the nematode Caenorhabditis elegans: distribution and morphological characteristics. *Histochem Cell Biol* **118**, 329-36 (2002).

120. Chen, C.C. *et al.* RAB-10 is required for endocytic recycling in the Caenorhabditis elegans intestine. *Mol Biol Cell* **17**, 1286-97 (2006).

121. Liu, H. *et al.* LET-413/Erbin acts as a RAB-5 effector to promote RAB-10 activation during endocytic recycling. *J Cell Biol* **217**, 299-314 (2018).

122. Sato, K., Norris, A., Sato, M. & Grant, B.D. C. elegans as a model for membrane traffic. *WormBook*, 1-47 (2014).

123. Grant, B. & Hirsh, D. Receptor-mediated endocytosis in the Caenorhabditis elegans oocyte. *Mol Biol Cell* **10**, 4311-26 (1999).

124. Tanji, T. *et al.* Characterization of HAF-4- and HAF-9-localizing organelles as distinct organelles in Caenorhabditis elegans intestinal cells. *BMC Cell Biol* **17**, 4 (2016).

125. Lapierre, L.R., Gelino, S., Melendez, A. & Hansen, M. Autophagy and lipid metabolism coordinately modulate life span in germline-less C. elegans. *Curr Biol* **21**, 1507-14 (2011).

126. Lemieux, G.A. & Ashrafi, K. Neural Regulatory Pathways of Feeding and Fat in Caenorhabditis elegans. *Annu Rev Genet* **49**, 413-38 (2015).

127. Seah, N.E. *et al.* Autophagy-mediated longevity is modulated by lipoprotein biogenesis. *Autophagy* **12**, 261-72 (2016).

128. Cypser, J.R. *et al.* Predicting longevity in C. elegans: fertility, mobility and gene expression. *Mech Ageing Dev* **134**, 291-7 (2013).

129. Lenaerts, I., Walker, G.A., Van Hoorebeke, L., Gems, D. & Vanfleteren, J.R. Dietary restriction of Caenorhabditis elegans by axenic culture reflects nutritional requirement for constituents provided by metabolically active microbes. *J Gerontol A Biol Sci Med Sci* **63**, 242-52 (2008).

130. Gelino, S. *et al.* Intestinal Autophagy Improves Healthspan and Longevity in C. elegans during Dietary Restriction. *PLoS Genet* **12**, e1006135 (2016).

131. Palgunow, D., Klapper, M. & Doring, F. Dietary restriction during development enlarges intestinal and hypodermal lipid droplets in Caenorhabditis elegans. *PLoS One* **7**, e46198 (2012).

132. Zhang, S.O. *et al.* Genetic and dietary regulation of lipid droplet expansion in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **107**, 4640-5 (2010).

133. Hsin, H. & Kenyon, C. Signals from the reproductive system regulate the lifespan of C. elegans. *Nature* **399**, 362-6 (1999).

134. Wang, M.C., O'Rourke, E.J. & Ruvkun, G. Fat metabolism links germline stem cells and longevity in C. elegans. *Science* **322**, 957-60 (2008).

135. Amrit, F.R. *et al.* DAF-16 and TCER-1 Facilitate Adaptation to Germline Loss by Restoring Lipid Homeostasis and Repressing Reproductive Physiology in C. elegans. *PLoS Genet* **12**, e1005788 (2016).

136. Ratnappan, R. *et al.* Germline signals deploy NHR-49 to modulate fatty-acid beta-oxidation and desaturation in somatic tissues of C. elegans. *PLoS Genet* **10**, e1004829 (2014).

137. Blackwell, T.K., Steinbaugh, M.J., Hourihan, J.M., Ewald, C.Y. & Isik, M. SKN-1/Nrf, stress responses, and aging in Caenorhabditis elegans. *Free Radic Biol Med* **88**, 290-301 (2015).

138. McGee, M.D. *et al.* Loss of intestinal nuclei and intestinal integrity in aging C. elegans. *Aging Cell* **10**, 699-710 (2011).

139. Ezcurra, M. *et al.* C. elegans Eats Its Own Intestine to Make Yolk Leading to Multiple Senescent Pathologies. *Curr Biol* **28**, 3352 (2018).

140. Cristina, D., Cary, M., Lunceford, A., Clarke, C. & Kenyon, C. A regulated response to impaired respiration slows behavioral rates and increases lifespan in Caenorhabditis elegans. *PLoS Genet* **5**, e1000450 (2009).

Chapter II Physiological defects caused by compromised gut specification

Introduction

Development of *C. elegans* intestine starts from the E cell, a descendant of the EMS cell. The endodermal fate is induced by the Wnt signaling pathway, whereby the neighboring cell (P_2) contacts the posterior side of EMS giving rise to the E cell in the next cell division. The transcription factors that initiate the gene regulatory network (GRN) for intestinal development are two types of GATA factors *med-1*, and *med-2* and *end-1* and *end-3*. These transcription factors are transiently expressed for a short period of time and in normal development their expression results in timely activation of *elt-2*, another GATA transcription factor which is a key regulator in intestine development. In embryos where the *med* and *end* transcription factors are mutated, a stochastic activation of *elt-2* occurs, leading to importer intestine development[1]. Some animals arrest in early stages of development, before hatching, some at L1 or L2 stage, because the intestine did not develop at all, or with too few cells[1,2]. The adult survivors of this stochastic activation of *elt-2* have managed to form an intestine, in some cases with increased amount of cells (>20)[2]. When these animals reach adulthood, they have several pleiotropic phenotypes including shortened body size, egg laying defects, and some sterility[1]. Many of these phenotypes are associated with reduced fitness and nutrition. In addition, we have recently observed increased fat content and lipid droplet size, as well as distended intestinal lumens size are increased (G. Broitman-Maduro, manuscript in preparation). The strains that exhibit features of this phenotype are called Hypomorphic Gut

Specification (HGS) strains. Here are presented the results from experiments done to characterize and quantify the pleiotropic phenotype of the HGS animals.

Materials and Methods

Worm maintenance and strains used

All strains were grown on *E. coli* OP50 and maintained at 20°C. *C. elegans* animals were grown on *E. coli* OP50 and handled according to standard methods. The reference strain was Bristol N2. Animals were grown in the presence of excess food that was never depleted throughout the experiment. Mutants: *end-3(ok1448), end-3(ir62)*, *med-1(ok804), med-1(ir55), nhr-156(ir67), cdc-25.1(rr31), daf-2(e1370), daf-16(mu86).*

Measurement of intestinal lumen width

To estimate the intestinal lumen size, the fluorescent ERM-1::GFP reporter, was first introduced into N2 animals and then with standard crosses introduced into the HGS strains and controls. ERM-1 is a protein expressed in the apical junctions of the intestinal cells. Fluorescent reporters of this protein were used to outline the apical membrane of the intestinal cells[3]. The intestinal lumen is a space enclosed from apical membranes of the intestinal cells that form the rings of the intestine (see Fig 1.2 in Chapter I). For the experiment adult gravid animals, from N2, *end-3*, *med-1; end-3*, *nhr-156*, were synchronized using bleaching solution[4] placed on 60mm NGM plates seeded with *E.coli* OP50 and left to develop until adulthood. The plates were monitored for contamination

and if detected, animals from that plate were not used for estimation of lumen width. To assess the width of the intestinal lumen day one adults, well fed animals from those three strains were washed with 500 µl of M9 solution and anesthetized with 5 µl of PEPG (phenol ether of propylene glycol). The animals were imaged on an Olympus BX51 microscope, equipped with Canon EOS 77D digital camera under 10X and 60X objectives on DIC and the GFP fluorescence channel. The images were processed with Adobe Photoshop to estimate the approximate lumen width of each animal. The lumen width was measured first as number of pixels from a cross section of the intestinal lumen and then the number of pixels was converted into µm. For every animal, 5 measurements from different parts of the lumen were taken and the numbers were averaged to give an approximate estimate of the lumen size (Table 2 Appendix A).

Generation of graphs and statistical calculations for the significance of the results were done with GraphPad Prism version 7.0.0 for Windows GraphPad Software, San Diego, California USA, www.graphpad.com.

Expression of ACDH-1::GFP

An ACDH-1::GFP reporter transgene was injected into N2 animals and then introduced into the HGS strains and controls through standard crosses. Special attention was paid to avoid any contamination or depletion of food while the animals developed as the *acdh-1* expression is sensitive to stress conditions[5]. When the animals reached adulthood (day one adults) they were imaged the same way described for estimation of intestinal lumen width.

Lifespan of and pathogenic resistance HGS strains

Gravid animals from N2, *age-1*, *end-3*, *med-1; end-3,* and *nhr-156* were synchronized with bleaching solution[4] and the embryos were placed on 60mm NGM agar plates seeded HT115 strain of *E coli* expressing a construct which produces dsRNA for the *egg-5* gene. The *age-1* strain, which is a long living strain was used as positive control and N2 strain as base line. *Egg-5* is a tyrosine phosphatase, which functions in gamete formation and is expressed in the oocytes[6]. RNAi on *egg-5* results in eggs which do not have proper chitin shell and are not viable. The treatment was done to prevent the bagging phenotype as the *med-1; end3* strain is especially prone to holding onto their eggs, which in turn results in their hatching internally and killing the parent prematurely. This treatment is not known to effect longevity studies, as is the case with treatments that disrupt the germline directly. At day one adult animals from all the used strains were transferred into NGM plates seeded with OP50 (starting number between 30-40) and plates were monitored every day, recording the number of dead animals for each strain, and transferring to fresh plates when necessary. The animals that did not die from aging (escaped, herniated, eggs hatched inside the animal) were excluded from the generation of the survival curves. This experiment was done at least 3 times with 30-40 worms.

For the pathogenic resistance experiments, the protocol described by Kirienko et al 2014[7] was followed. Briefly, 10µl of saturated culture of *Pseudomonas aeruginosa* PA14 was spread onto 35mm plates with NGM agar on them. The plates were incubated 24h at 37°C and then another 24h at 25°C. After that, 40µl of 100X FUDR (5-Fluoro-

2'deoxyuridine), 10mg/ml solution in ddH_2O, filter sterilized was applied as small drops close to the edge of the plate[7]. FUDR solution prevents the eggs that the worms will lay from hatching. About 40 L4 stage worms from the strains N2, *end-3*, *med-1; end-3*, *nhr-156*, and *age-1* were transferred on to the plates from normal plates seeded with *E. coli* OP50 to the plates with *Pseudomonas aeruginosa* PA14 in a way that animals were placed on the bare agar and the plates were incubated at 25°C for the duration of the experiment.

Generation of graphs and statistical calculations for the significance of the results were done with GraphPad Prism version 7.0.0

RNA-seq of intestinal cells

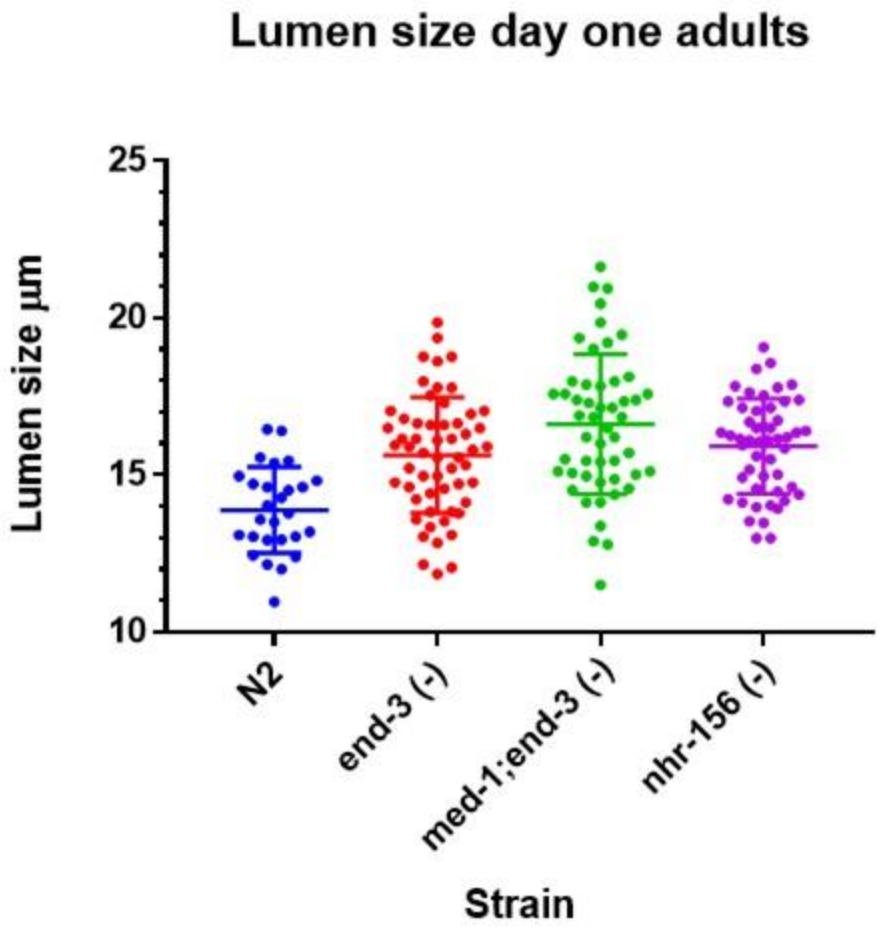
Animals from the N2 and *med-1; end-3* backgrounds were dissected in a watch glass in ice cold 1X PBS, and the intestines of the animals were collected in Eppendorf tubes and frozen in liquid nitrogen until further processing one intestine per tube. For each RNA seq experiment 10 intestines of each strain were pooled and used for RNA extraction and library preparation. RNA seq experiments were repeated three times. Extraction of RNA and subsequent preparation of libraries for RNA seq was done following the protocol published by Hashimshoni et al 2012[8].

All the work for intestine dissection and RNA seq libraries preparation was done by Hailey Choi, Maduro lab.

Results

Measurement of intestinal lumen width

One of the features of the HGS phenotype is increased intestinal lumen size. It was initially observed under DIC in *end-3*, *med-1; end-3* and *nhr-156* mutants. They had increased lumen size compared to wild-type animals. The fluorescent reporter ERM-1::GFP was used to visualize the lumen in living animals (Fig.2.1). The estimated mean for N2 is 13.89 μm ± 0.2636 SEM, for *end-3* 15.64 μm ± 0.2448 SEM, for *med-1; end-3* 16.62 μm ± 0.3117 SEM, for *nhr-156* 15.92 μm ± 0.2114 SEM with p-value of one sample t-test for statistical significance of the differences between lumen widths, $p<0.0001$.



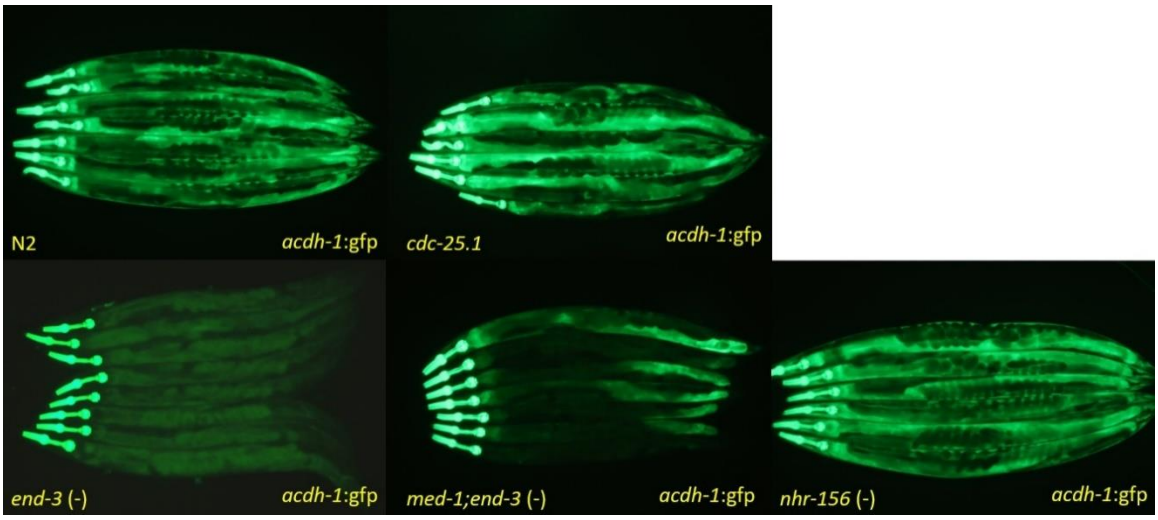
The intestinal lumen increase is usually a result of shrinking of the intestinal cells which is observed when animals experience calorie restriction or starvation, because the contents of the intestinal cells are being consumed to provide energy and nutrients for the animal[7]. Also an increase in the lumen size is observed as the animals age as they "run off" their intestines as the intestinal cells constantly produce products that are required for the eggs[9]. In the case of the HGS strains the cause for the increased lumen size probably comes from the changes in the gene expression profile of the intestinal cells, as

the animals used in the experiment were day one adults, and they were let to develop always in the excess of food.

Expression levels of ACDH-1::GFP

The *acdh-1* gene encodes an acyl-CoA dehydrogenase. It participates in the first step of the beta oxidation cycle and is involved in energy production. The gene is responsive to energy storage state and energy usage state, being upregulated in energy usage state. It has been shown that fasted animals down regulate *acdh-1* expression levels[5,10].

The results from the experiment show that the ACDH-1::GFP expression levels are lower in animals from the HGS strains compared to the wild-type animals (Fig. 2.2), providing further evidence that the HGS strains are in a state of energy storage even when food is abundant. The *cdc-25.1* strain levels of expression are on par with the wild type suggesting the increased number of cells in the *cdc-25.1* mutants is not the reason for down regulation of the reporter in *end-3* and *med-1; end-3* strains.



The images on the figure are an example of the expression levels of ACDH-1::GFP in wild-type and HGS strain. Multiple images for each strain were taken and analyzed to determine quantitatively the expression levels that were mentioned above (G. Broitman-Maduro, personal communication). An adult hermaphrodite is approximately 1mm long.

HGS strains have increased lipid droplet size

Another feature of the HGS phenotype is increased lipid droplet size (Table 2.1) (G. Broitman-Maduro, personal communication).

Strain	Average Lipid droplet size (μm) ± SEM
N2	1.463 ± 0.01811
end-3	1.917 ± 0.02269
med-1; end-3	2.163 ± 0.02275
daf-2	1.864 ± 0.01486
daf-16	1.515 ± 0.01466
end-3; daf-16	1.561 ± 0.01812
med-1; end-3; daf-16	1.75 ± 0.01523
cdc-25.1	1.568 ± 0.01214
nhr-156	2.26 ± 0.02635
nhr-156; daf-16	1.906 ± 0.01343
nhr-156 with rescue	1.441 ± 0.01066
med-1; end-3 with rescue	1.697 ± 0.01533
med-1; end-3; daf-16 with rescue	1.413 ± 0.01011

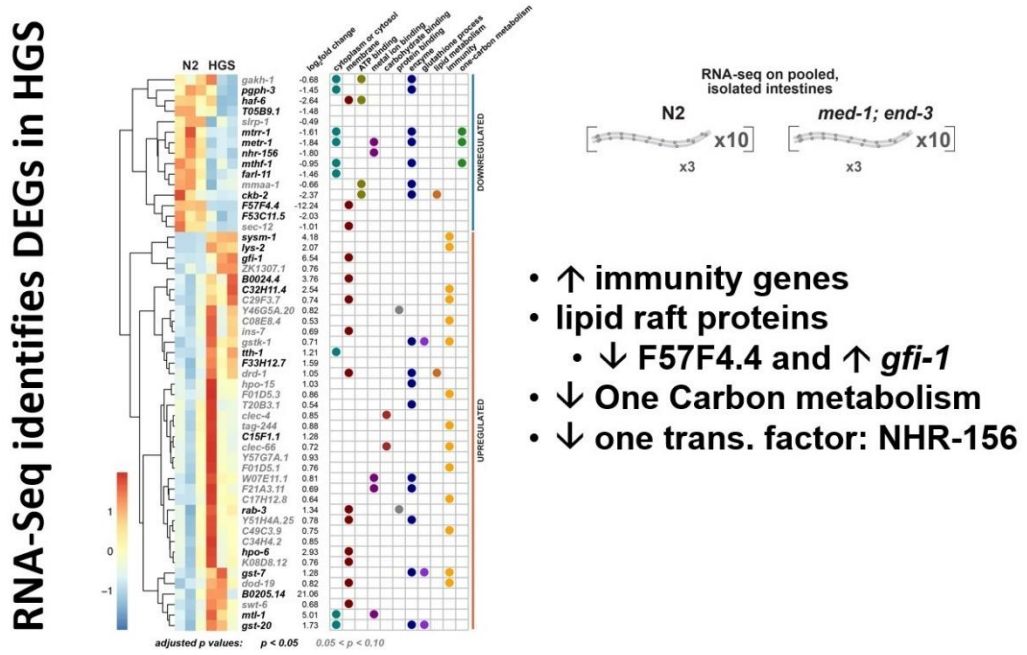
The *daf-2* strain was used as a positive control since previous work has shown that mutations in *daf-2* (the *C. elegans* insulin receptor) lead to constant nuclear localization of the *daf-16*/FOXO transcription factor, which has target genes involved in stress response[11]. The *daf-2* mutants have a significant increase in the average lipid droplet size

because they are storing fats to overcome a stress condition. Increasing the size of lipid droplets makes them less accessible for lipases to digest them. On the other hand, the *daf-16* animals have very little increase since the stress response genes cannot be activated. The *cdc-25.1* strain was used as control. The *cdc-25.1* gene is a phosphatase that positively regulates the proliferation of the E lineage[10]. The strain harbors a dominant mutation and that causes extra intestinal cells to be formed. Since the HGS strain sometimes show increased number of intestinal cells, *cdc-25.1* was used to test whether the increased amount of gut cells could be a reason for increased lipid droplet size. The results show that the HGS strains: *end-3*, *med-1; end-3*and *nhr-156* have significantly increased lipid droplet size compared to wild type (see table). Crossing in the *daf-16* mutation to those strains seems to decrease the average size of lipid droplets. More importantly the experiments shows that when an *nhr-156* rescuing construct was added to the *nhr-156* and *med-1; end-3; daf-16* strains it rescued the lipid droplet size back to wild type average and partially rescued the *med-1; end-3* lipid droplet size. These results suggest that *nhr-156* normally plays a role in regulation of lipolysis with mechanism that is *daf-16* independent.

GO terms analysis of DEG from HGS and wild-type animals

Analysis of the results from the RNA seq experiments with intestinal cells from HGS strain and wild-type animals produced a heatmap showing DEG in HGS and wild-type animals. The map was first generated by Hailey Choi, Maduro lab (updated by MM),

using R. A custom Python script (MM) was used to assign GO terms to the DEG (Fig. 2.3)

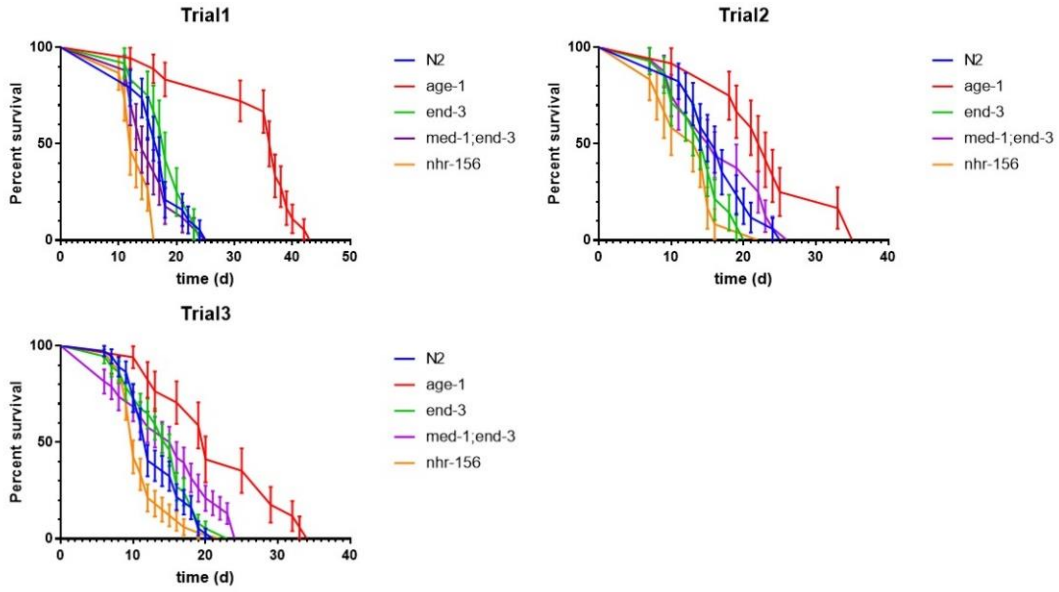


The GO term analysis of the differentially expressed genes showed that in HGS animals, genes that are down regulated are involved in one carbon metabolism, some have enzymatic function, some are lipid raft proteins, and one is a transcription factor *nhr-156*. On the other hand, many of the up regulated genes in the HGS animals were connected with the innate immunity in *C. elegans*. This up regulation of immunity genes prompted us to investigate the lifespan and the pathogenic resistance to infection from *Pseudomonas aeruginosa* PA14.

Lifespan of HGS strains

It's been shown in the literature that when immunity genes are up regulated because of some external factor or mutation in transcription factors or cell signaling cascades that regulate their expression this sometimes results in increased lifespan and increased resistance to certain pathogens[11], therefore experiments were done to characterize the life span and pathogenic resistance of HGS animals.

The results from the lifespan experiments show that there were no significant differences between the lifespan of *end-3*, and *med-1; end-3* strains compared to wild type. The *nhr-156* strain has significantly decreased lifespan, compared to wild type (Fig. 2.3).



The *age-1* strain was used as a positive control. Three trials were done in total. In all trials the *nhr-156* strain had a median lifespan lower than the wild-type animals (Table

2.2). Apparently, the up regulation of immunity genes in HGS animals did not result to significant increase in lifespan.

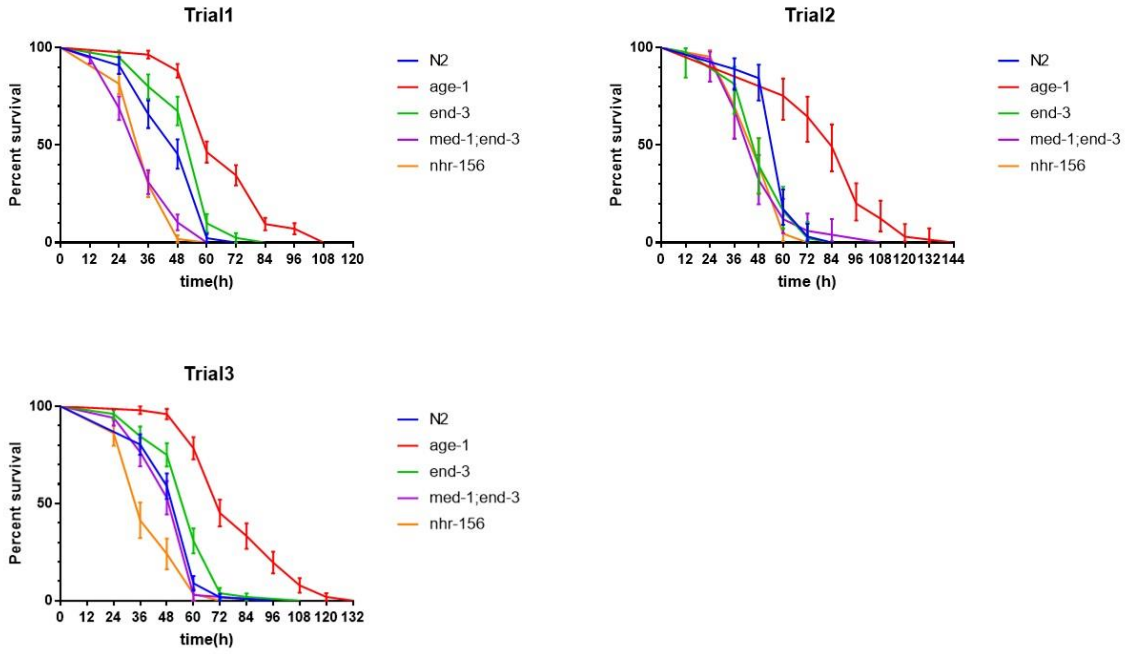
Median lifespan							
Trial 1			Trial 2			Trial 3	
Strain	Median Lifespan (d)		Strain	Median Lifespan (d)		Strain	Median Lifespan (d)
N2	17		N2	16		N2	12
nhr-156 (-)	12		*nhr-156* (-)	13.5		*nhr-156* (-)	10
end-3 (-)	18		*end-3* (-)	14.5		*end-3* (-)	15
med-1;end-3 (-)	14		*med-1;end-3 (-)*	15.5		*med-1;end-3 (-)*	15.5

Comparison between Lifespan of N2 and mutant strains							
Trial 1							
Strains			p value of Log rank (Mantel-Cox) test		p value of Gehan-Breslow-Wilcoxon test		
N2	vs	*nhr-156* (-)	0.0005		0.0012		shorter lifespan
N2	vs	*end-3* (-)	0.6233	ns	0.4306	ns	
N2	vs	*med-1;end-3* (-)	0.409	ns	0.2282	ns	
Trial 2							
N2	vs	*nhr-156* (-)	0.0546		0.035		shorter lifespan
N2	vs	*end-3* (-)	0.0923	ns	0.1413	ns	
N2	vs	*med-1;end-3* (-)	0.5557	ns	0.9712	ns	
Trial 3							
N2	vs	*nhr-156* (-)	0.0534		0.0158		shorter lifespan
N2	vs	*end-3* (-)	0.3269	ns	0.4	ns	
N2	vs	*med-1;end-3* (-)	0.0333		0.3371	ns	

Pathogenic resistance of HGS strains

The "slow" killing mechanism of *Pseudomonas aeruginosa* PA14 was used to test the resistance to a pathogenic bacterium. In this mechanism the bacteria colonize the intestine of the animals, eventually killing them. Walker et al. 2015 reported that animals that are *sams-1* mutants were very sensitive to infection with *Pseudomonas aeruginosa* PA14. *Sams-1*, s-adenosylmethionine synthase is an enzyme that produces s-adenosylmethionine (SAM), a metabolite which is a donor of methyl groups for histone and DNA methylation[12]. The RNA seq data from HSG animals was showing the genes involved in the one carbon metabolism, including *sams-1* were slightly down regulated.

Based on this information three trials were done to test the pathogenic resistance of HGS animal and *nhr-156* animals (Fig. 2.4).



Overall, the *nhr-156* strain appeared to be more sensitive to infection with *Pseudomonas aeruginosa* PA14 compared to the wild type, with median survival of *nhr-156* 36h and 48h, compared to wild type 48h 60h respectively. The data for the two HGS strains shows that in two out of three trials the *med-1;end-3* strain, which has the most severe HGS defects was more sensitive to infection with *Pseudomonas aeruginosa* PA14 compared to the wild type, with the same median survival, compared to wild type like *nhr-156*.The results from the experiments were in concordance with what Walker et al .2015, reported[12]. In conclusion, even though the HGS animals have immunity genes up regulated they did not provide increased resistance to infection with *Pseudomonas aeruginosa* PA14 (Table 2.3).

Median survival on PA14							
Trial 1			Trial 2			Trial 3	
Strain	Median survival (h)		Strain	Median survival (h)		Strain	Median survival (h)
N2	48		N2	60		N2	60
nhr-156 (-)	36		*nhr-156* (-)	48		*nhr-156* (-)	36
end-3 (-)	60		*end-3* (-)	48		*end-3* (-)	60
med-1;end-3 (-)	36		*med-1;end-3 (-)*	48		*med-1;end-3 (-)*	60

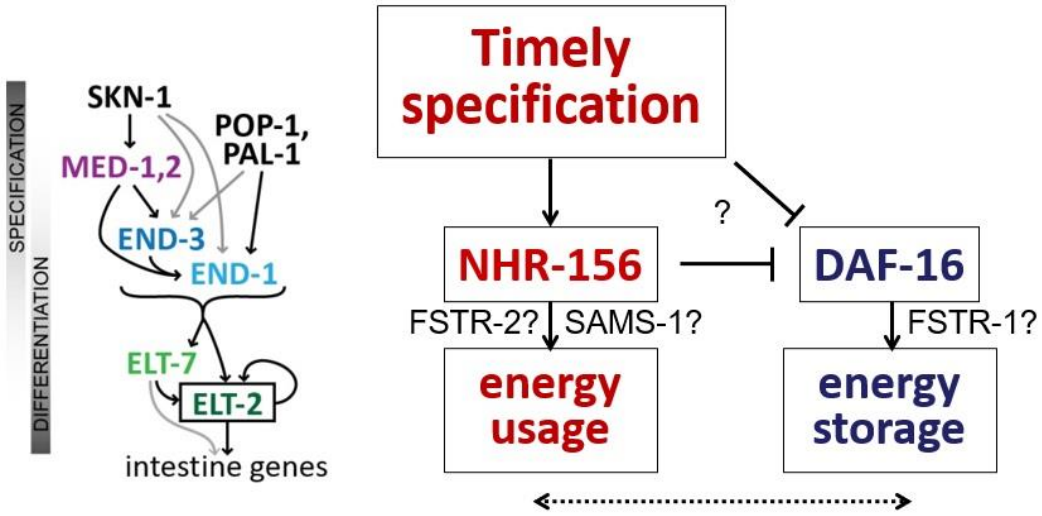
Comparison between survival on PA14, slow kill mechanism of N2 and mutant strains					
Trial 1					
Strains			p value of Log rank (Mantel-Cox) test	p value of Gehan-Breslow-Wilcoxon test	
N2	vs	*nhr-156* (-)	<0.0001	<0.0001	sensitive
N2	vs	*end-3* (-)	0.0215	0.0312	resistant
N2	vs	*med-1;end-3 (-)*	<0.0001	<0.0001	sensitive
Trial 2					
N2	vs	*nhr-156* (-)	<0.0001	<0.0001	sensitive
N2	vs	*end-3* (-)	0.0028	0.0008	sensitive
N2	vs	*med-1;end-3 (-)*	0.0004	<0.0001	sensitive
Trial 3					
N2	vs	*nhr-156* (-)	0.0003	<0.0001	sensitive
N2	vs	*end-3* (-)	0.0111	0.0165	resistant
N2	vs	*med-1;end-3 (-)*	0.453	0.373	ns

Discussion

The experiments shown in this chapter sought to quantify some aspects of the HGS phenotype we observed. The results show that when the early intestine development is compromised, animals that survive, change their development is a way that when they reach adulthood they behave as calorie restricted, even though they feed *ad libitum*. The features of this phenotype are increased lipid content, increased lumen width, increased lipid droplet size, changed gene expression profile in the intestinal cells. To some extent the observed phenotype resembles one described by Palgunow et al 2012[13], where they subjected the animals to 'developmental dietary restriction', which caused a phenotype with increased lipid droplet size, and changes in the gene expression profile of the

affected animals. In the HGS additional defects are observed, such as increased number of intestinal cells[2], and increased lumen width. From the experiments done to characterize the phenotype of the HGS strains it is clear that the untimely specification of the intestine during development results in a change of the gene expression profile and physiology of the animals that survive this developmental defect. The RNA seq experiments showed many differentially expressed genes (Fig. 2.3).

Using the RNA seq results, mentioned above, mutants of some of the downregulated genes were created to additionally investigate the causes of the physiological defects. The mutant described in this chapter was *nhr-156*. The gene is a member of a large family of more than 250 *nhr* genes in *C. elegans*[10]. According to the databases *nhr-156* is with unknown function. The experiments with the *nhr-156* mutants show that it has increased intestinal lumen and lipid droplet size, a shortened lifespan and sensitivity to *Pseudomonas* infection. More importantly the experiments shown here revealed that one function of the *nhr-156* gene is in promoting the lipolysis since the *med-1; end-3* increased lipid droplet size was alleviated to some extent when nhr-156 rescue was introduced into the strain (see results) and in the *med-1; end-3; daf-16* strain, which had less increased lipid droplet size than the *med-1; end-3* the lipid droplet size was rescued to wild-type levels, suggesting that the lack of DAF-16 and presence of NHR-156 contribute in a different way to promotion of lipolysis. Based on the results from the experiments done for characterization of HGS strain a putative model is proposed on how timely specification of the intestine during development affect the metabolism and physiology of adult animals (Fig. 2.5)



Analysis of the differentially expressed genes in HGS and wild-type animals, led to the focus of this thesis, namely the characterization of two reciprocally expressed genes, F57F4.3 (*gfi-1/ fstr-1*) and F57F4.4/*fstr-2*, to elucidate the causes of the physiological defects we observed in the HGS strains.

References

1. Maduro, M.F. *et al.* MED GATA factors promote robust development of the C. elegans endoderm. *Dev Biol* **404**, 66-79 (2015).

2. Choi, H., Broitman-Maduro, G. & Maduro, M.F. Partially compromised specification causes stochastic effects on gut development in C. elegans. *Dev Biol* **427**, 49-60 (2017).

3. Asan, A., Raiders, S.A. & Priess, J.R. Morphogenesis of the C. elegans Intestine Involves Axon Guidance Genes. *PLoS Genet* **12**, e1005950 (2016).

4. Stiernagle, T. Maintenance of C. elegans. *WormBook*, 1-11 (2006).

5. MacNeil, L.T., Watson, E., Arda, H.E., Zhu, L.J. & Walhout, A.J. Diet-induced developmental acceleration independent of TOR and insulin in C. elegans. *Cell* **153**, 240-52 (2013).

6. Parry, J.M. *et al.* EGG-4 and EGG-5 Link Events of the Oocyte-to-Embryo Transition with Meiotic Progression in C. elegans. *Curr Biol* **19**, 1752-7 (2009).

7. Kirienko, N.V., Cezairliyan, B.O., Ausubel, F.M. & Powell, J.R. Pseudomonas aeruginosa PA14 pathogenesis in Caenorhabditis elegans. *Methods Mol Biol* **1149**, 653-69 (2014).

8. Hashimshony, T., Wagner, F., Sher, N. & Yanai, I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep* **2**, 666-73 (2012).

9. Ezcurra, M. *et al.* C. elegans Eats Its Own Intestine to Make Yolk Leading to Multiple Senescent Pathologies. *Curr Biol* **28**, 3352 (2018).

10. Harris, T.W. *et al.* WormBase: a comprehensive resource for nematode research. *Nucleic Acids Res* **38**, D463-7 (2010).

11. Sun, X., Chen, W.D. & Wang, Y.D. DAF-16/FOXO Transcription Factor in Aging and Longevity. *Front Pharmacol* **8**, 548 (2017).

12. Ding, W. *et al.* s-Adenosylmethionine Levels Govern Innate Immunity through Distinct Methylation-Dependent Pathways. *Cell Metab* **22**, 633-45 (2015).

13. Palgunow, D., Klapper, M. & Doring, F. Dietary restriction during development enlarges intestinal and hypodermal lipid droplets in Caenorhabditis elegans. *PLoS One* **7**, e46198 (2012).

Chapter III Characterization of *fstr-1* and *fstr-2* genes

Introduction

We became interested in studying the *fstr-1/gfi-1/* (F57F4.3) and *fstr-2* (F57F4.4) genes, because our RNA-seq data produced from intestinal cells of *med-1; end-3* HGS strain compared to wild-type animals showed that *fstr-1*/F57F4.3 and F57F4.4/*fstr-2* had reciprocal expression. Namely, *gfi-1/fstr-1*was upregulated in the HGS animals, while *fstr-2*/F57F4.4 was upregulated in wild-type animals. The *fstr* genes were named in prior work on *C. elegans clk-1* mutants. CLK-1 is an ortholog of human COQ7 an enzyme that is necessary for ubiquinone biosynthesis, which is part of the electron transport chains.[1] Mutation of *clk-1* results in many pleiotropic effects, including a slowing down of metabolic and rhythmic processes. The F57F4.3/ *fstr-1* and F57F4.4/ *fstr-2* genes were identified in an RNAi screen, for the ability of *fstr-1,2(RNAi)* to restore the speed of rhythmic behaviors in *clk-1* mutants[2]. Also in the same study it was observed that *clk-1* mutants have changed transcription profile with some genes being downregulated, compared to wild-type animals, and RNAi of the F57F4.3 and F57F4.4 genes also caused an increase in the expression levels of some of the downregulated genes in *clk-1* mutants up to levels close to those in wild-type animals[2]. These observations prompted the authors of the study to name the two genes *fstr* (from 'faster') 1 and 2, instead of the cosmid signatures which were assigned to those genes when the *C. elegans* genome was annotated, with *fstr-1* upregulated in *clk-1* mutants (similar to our HGS strains) and *fstr-2* being downregulated in *clk-1* mutants. After generating the HGS stains and obtaining

RNA seq data from their intestines we were interested in genes that may be responsible for modulating the changes we observed in the HGS animals downstream of gut specification. Based on the information in the *clk-1* study and our own data We hypothesized, that FSTR-1 and FSTR-2 play related but differentiable roles in regulating metabolism, and decided to investigate the structure, and function of those two genes.

Materials and methods

Generation of *fstr-1*, *fstr-2* single mutants and *fstr-1,2* double mutant

Deletion alleles of *fstr-1* and *fstr-2* were generated by CRISPR-Cas9 genome editing, by targeting unique regions outside the coding region for each gene. (Fig.3.1). CC1, CC2, CC3, CC4 (shown on the figure below) were the cleavage sites used for the *fstr* mutants. All selected regions were out of the coding sequence of both *fstr-1* and *fstr-2* genes. These regions were chosen specifically because they contain unique sequences, which were used to selectively delete one gene or the other, or both. CC1 and CC2 were used to delete both *fstr-1* and *fstr-2* genes, while CC2, CC4 resulted in a *fstr-1* deletion and CC1, CC3 for *fstr-2* deletion. All constructs were injected into the gonads of N2 day one adult hermaphrodites along with a plasmid construct that carries the Cas9 enzyme. Mutants were first selected on plates that had either co-CRISPR events (visible mutants) or by *rol-6* selection and then confirmed by PCR[3,4].

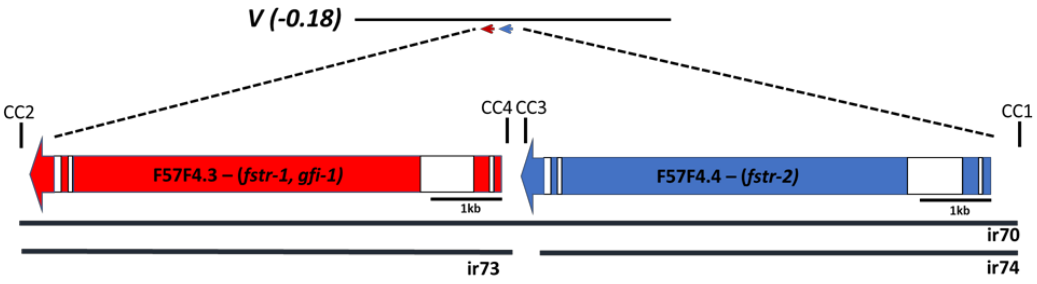


Fig.3.1 Location of *fstr* genes on chromosome V in the *C. elegans* genome. CC1-CC4 are the regions chosen to design gRNAs for CRISPR-Cas9 genome editing.

All strains that are listed as extrachromosomal arrays (Table 1 in Appendix A) were generated by injecting day one adult well fed animals from the appropriate strain, with an injection mixture (10µl), which contained final concentration of approximately 100ng (1µg/1µl) of the fluorescent reporter plasmid and final concentration of 100ng (1µg/µl) of pRF4 plasmid, which contains the *rol-6* gene with a dominant mutation in it as a selectable marker. *Rol-6* encodes a component of the collagen and cuticulin based cuticle[5] and the dominant mutation in the gene causes the animals to corkscrew around in circles.[6] The plasmids are brought to 10µl injection mixture with sterile ddH_2O. After injection in the F_2 of the injected animals, rollers were selected and observed if they express the fluorescent reporter, then grown for a few more generations to establish stable lines.

Assessing the amount of visible gut granules in *fstr* mutants

For this experiment we used animals from the following strains: N2, *fstr-1*, fstr-*2*, *fstr-1,2*, *med-1; end-3*, *clk-1(qm30)* provided by the *Caenorhabditis* Genetics Center (CGC), which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440) and *clk-1; fstr-1,2*. Adult gravid animals from these strains were treated with solution

containing sodium hypochlorite, sodium hydroxide and water to dissolve the animals and gain access to the embryos of the animals. The embryos were placed onto 60mm petri dishes containing NGM agar seeded with *E. coli*, strain OP50 bacteria. The plates were incubated at 20°C until the animals reach adulthood. Plates were inspected daily to monitor how the development of the animals proceeded. Once the animals were 1-day old adults (they had completely developed gonads and were starting laying eggs) some of the animals were washed out of the plate using 500µl of M9 solution, placed in a 1.5ml Eppendorf tube and anesthetized with 5µl PEPG (phenyl ether of propylene glycol). Animals were inspected visually to stop swimming and then spun down and pipetted out of the tube onto a microscope slide with an agar pad on it. Worms were aligned up one to another very closely and observed and photographed under polarized light to assess the quantity and the distribution of the gut granules in their intestines.

All microscope observations and photographs were done on an Olympus BX51 microscope, equipped with Canon EOS 77D digital camera. Observations were made with 10X and 60X objectives of the microscope.

Starvation of L1 larvae

Animals from N2, *fstr-1*, fstr-*2*, *fstr-1,2*, *med-1; end-3*, *clk-1* and *clk-1; fstr-1,2* strains were bleached and the eggs were put in 1ml M9 solution with addition of 0.5µl of 10mg/ml cholesterol. Tubes were next placed on a rotisserie for 8 hours to allow eggs to hatch. After 8h a sample was taken, and larvae were observed under microscope under polarized light to look for presence and amount of gut granules. Photographs of the larvae

were taken at the time of hatching and at 24h, 48h and 72h time points after hatching. For the duration of the experiment the larvae were kept in 1ml M9 solution with 0.5µl of 10mg/ml cholesterol.

Measurement of intestinal lumen width

The estimation of the intestinal lumen width of *fstr-1,2* mutants, and comparison with wild-type and HGS strain animals was done the same way described in the materials and methods of Chapter II. The strains used for the experiment were N2, *med-1; end-3*, and *fstr-1,2*.

Estimating the lipid droplet size of *fstr-1,2(-)* mutants

To estimate the lipid droplet size we created a plasmid construct that expresses the *C. elegans* perilipin gene, *plin-1*[7] fused with a red fluorescent reporter, wrmScarlet, under the control of *opt-2* promoter. *Opt-2* is a gene that encodes for a transmembrane transporter expressed in the intestinal cells[5]. The plasmid construct was injected in animals from the N2, *med-1; end-3*, and *fstr-1,2* animals to assess the size of the lipid droplets. Animals from the three strains were imaged under 60X objective and the images were processed with ImageJ to estimate first the surface of the lipid droplets from the images. After that the diameter of lipid droplets was determined using the approximation that the surface of each lipid droplet is a circle, with the formula: "= (SQRT(measured surface in ImageJ in pixels/3.14)/24)*2 in Microsoft excel. The number 24 is a coefficient that was estimated for the 60X objective we have on the Olympus BX51

microscope, and the digital camera connected with it, which converts the number of pixels measured in ImageJ into micrometers (Table 3 Appendix A).

Brood count of *fstr* mutants

Gravid animals from N2, *fstr-1*, fstr-*2*, *fstr-1,2*, *med-1; end-3*, *clk-1* and *clk-1;fstr-1,2* strains were treated with sodium hypochlorite solution to isolate embryos and the embryos were placed on 60mm NGM agar plates seeded with *E coli* OP50. At L3 stage for each trial 5 hermaphrodite animals were singled onto 35mm plates with NGM agar seeded with *E. coli* OP50, and plates were monitored until the animals transitioned to adulthood and started laying eggs. After that every day the number of hatched progenies was counted and recorded, and the adult animal was transferred onto a new 35mm plate until the adults stopped laying eggs. Results were input in GraphPad Prism version 7.0.0.

Lifespan of *fstr* mutants

Gravid animals from N2, age-*1*, *fstr-1*, *fstr-2*, *fstr-1,2*, *med-1; end-3*, *clk-1* and *clk-1;fstr-1,2* strains were treated with sodium hypochlorite solution to isolate embryos and the embryos were placed on 60mm NGM agar plates seeded HT115 strain of *E coli* expressing a construct which produces dsRNA for the *egg-5* gene. The *age-1* strain, which is a long living strain was used as positive control and N2 strain as base line. *Egg-5* is a tyrosine phosphatase, which functions in gamete formation and is expressed in the oocytes. RNAi against *egg-5* results in eggs which do not have proper chitin shell and are not viable. The treatment was done to prevent bagging phenotype (eggs hatch inside the worm and kill it), as *med-1; end-3* strain is especially prone to that and this treatment

does not make the animals germline less (*cdc-25*; *glp-1 RNAi*), which in its turn changes animals physiology and gene expression significantly. At day one adult animals from all the used strains were transferred into NGM plates seeded with OP50 (starting number between 80-100) and plates were monitored every day, recording the number of dead animals for each strain, and transferring to fresh plates when necessary. The survival curves and statistical calculations were generated with GraphPad Prism version 7.0.0.

Assessing pathogenic resistance to toxin based (fast) killing and colonization of the intestine (slow) killing mechanism from *Pseudomonas aeruginosa* strain PA14

Fast killing mechanism

To setup the experiment we followed the protocol described by Kirienko et al.[8]. Briefly, on 35mm plates with PGS media (peptone, glucose, sorbitol) a 5µl solution from a culture of *Pseudomonas aeruginosa* PA14 grown to high OD, was spread in the middle of the plate using a sterile glass L spreader and then incubated at 37°C for 24h. After that the plates were incubated at room temperature for another 24h. Animals from N2, *fstr-1*, fstr-*2*, *fstr-1,2*, *med-1;end-3*, strains were synchronized and L4 stage animals were put into the PGS plates with PA14 and incubated for the duration of experiment at 25°C. Plates were monitored and the number of dead animals recorded. The results were used to generate survival curves with GraphPad Prism version 7.0.0.

Slow killing mechanism

The preparations for the experiment were done the same way described in Chapter II. About 40 L4 stage worms from the strains N2, *fstr-1*, *fstr-2*, *fstr-1,2*, *med-1; end-3* were transferred on to the plates from normal plates seeded with *E.coli* OP50 to the plates with *Pseudomonas aeruginosa* PA14 in a way that animals were placed on the bare agar and the plates were incubated at 25°C for the duration of the experiment.

Generation of plasmid constructs to explore the expression patterns of *fstr* genes

Transcriptional fusions

The transcriptional fusions generated to probe the expression of the *fstr* genes were designed based on plasmid construct PID3 – p*endu2*:wrmSCR:H2B:*unc-54*UTR (see Appendix A for a full list of the plasmid constructs created for this study), which was originally built for another project. Because PID3 was built modularly, we were able to replace the promoter and the 3'UTR sequences in the vector by using a combination of restriction digest enzymes and ligating the *fstr-1* and *fstr-2* promoters in 3' UTRs. For ligating the promoter regions of *fstr-1* and *fstr-2* genes the PID3 vector was digested with *PstI* and *XbaI* restriction enzymes and the gel purified PCR products of the promoter sequences for both genes which contain the appropriate restriction sites. The restriction sites were embedded in the PCR primers that were used to amplify the promoter sequences from N2 genomic DNA and the promoter sequences were inspected so they should not have any endogenous *PstI* and *XbaI* restriction sites. Once gel purified the PCR products of the promoters were digested with the same combination of restriction

enzymes to produce the appropriate ends so the fragment can be ligated into the vector digested the same way. (See Appendix A for the protocols for PCR, restriction digest, ligation and transformation into *E. coli*) The resulting vectors PID11 – p*fstr-2*:wrmSCR:H2B:*unc-54*_3'UTR and PID12 – p*fstr-1*:wrmSCR:H2B:*unc-54*_3'UTR, were digested with *EcoRI* and *SpeI* to replace the *unc-54*UTR sequence with sequences from the 3'UTRs of *fstr-1* and *fstr-2* genes. This generated the PID14 – p*fstr-2*:wrmSCR:H2B:*fstr-2* 3'UTR and PID15 – p*fstr-1*:wrmSCR:H2B:*fstr-1* 3'UTR constructs, which have the reporter gene wrmScarlet (wrmSCR in the construct's description) and the H2B gene as a nuclear localization signal. The reporter gene wrmScarlet is a codon optimized for expression in *C. elegans*[9] version of mScarlet reporter[10]. Both PID14 and PID15 constructs were injected in *med-1; end-3* (HGS in the figures) animals and then crossed into N2 animals.

Translational fusions

To generate the translational fusions we used PID14 and PID15 constructs in which first the wrmSCR:H2B part was excised and the replaced only with wrmSCR and then the coding sequence of exon1 and exon2 for *fstr-1* and *fstr-2* genes was added resulting in plasmids PID25 – p*fstr-1*:wrmSCR:*fstr-1*ex-1-2:*fstr-1* 3'UTR and PID26 – p *fstr-2*:wrmSCR:*fstr-2*ex-1-2 *fstr-2 3'*UTR. Also, into PID15 and PID14 constructs were introduced by Gibson Assembly[11] the whole *fstr-1* and *fstr-2* genes respectively. Genes were amplified from N2 genomic DNA with hot-start high-fidelity Q5 polymerase (New England Biolabs M0494S). Assembling the whole genes in the

constructs PID14 and PID15 generated PID27 – p*fstr-1*:*fstr-1*:wrmSCR:*fstr-1* 3'UTR and PID28 – p*fstr-2*:*fstr-2*:wrmSCR:*fstr-2* 3'UTR, translational fusions in which the reporter gene is after the end of the coding sequence of *fstr-1* and *fstr-2* genes. Gibson Assembly was used to generate other translational fusions than PID27 and PID28 such as PID52, which was created as 3 piece Gibson Assembly into the vector pBluescript KS-.

Generation of CRISPR integrated chromosome tags of *fstr* genes

We used CRISPR Cas9 genome editing system to integrate the wrmScarlet reporter in different places of the *fstr-2* gene. Briefly, a break is initiated into the specific place in the chromosome where the *fstr-2* gene resides, by injecting into the gonads of wild-type animals of plasmids that contain the guide RNA the Cas-9 enzyme another plasmid that carries the *rol-6* gene harboring a dominant mutation that makes the worms roll as a selectable marker and a specific repair template, which is produced by PCR with a high-fidelity polymerase from a template that has homology with *fstr-2* gene on both sides and also has the full sequence of the wrmScarlet reporter. The sites for the guide RNAs used were CC1 which is the region before the start of the *fstr-2* gene and CC3, which is after the stop codon of the *fstr-2* gene (Fig.3.1)

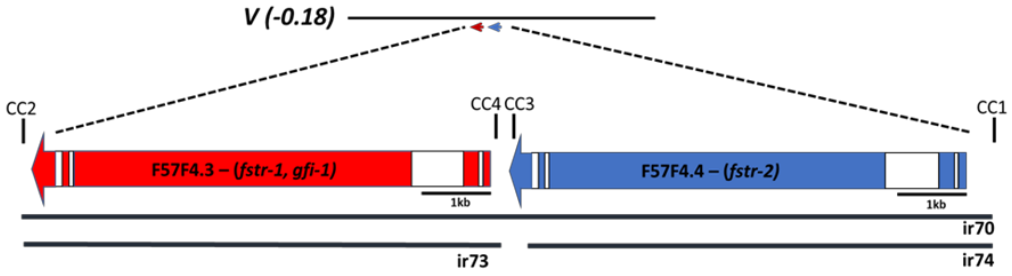


Fig.3.1 Location of *fstr* genes on chromosome V in the *C. elegans* genome. CC1-CC4 are the regions chosen to design gRNAs for CRISPR-Cas9 genome editing.

Using the CC3 and CC1, three integrant lines were generated by delivering a repair template.

For CC1 guide RNA, the repair templates that were successfully integrated were: p*fstr-2*:wrmSCR:exon-1:intron1:exon2 of the *fstr-2* gene (from PID26 – p*fstr-2*:wrmSCR:*fstr-2*exon1:*fstr-2*intron1:*fstr-2*exon2), placing the reporter gene before the start of the coding sequence of *fstr-2*, and *fstr-2* exon-1-intron-1-exon2:wrmSCR:*fstr-2* intron2 (from PID42 – p*fstr-2*:*fstr-2*ex1+2:wrmSCR:*fstr-2*intron2:*fstr-2* 3'UTR), placing the reporter gene after the first two exons in the *fstr-2* gene sequence.

For CC3 guide RNA one repair template was integrated, derived from a version of PID28 that has extended 3'UTR sequence with longer homology region in the 3'UTR to facilitate the process of homology directed repair: *fstr-2* exon-4-intron-4-exon5:wrmSCR: *fstr-2* 3'UTR, placing the reporter gene at the very end of the coding sequence of the *fstr-2* gene.

Results

In silico analysis of gene and protein structure of the *fstr* genes

Fstr-1 (V:6392537..6399931) and *fstr-2* (V:6401379..6408783) are two paralogues genes that have very similar coding sequence, located adjacent to each other on chromosome V in the *C. elegans* genome (Fig.3.2)[5]

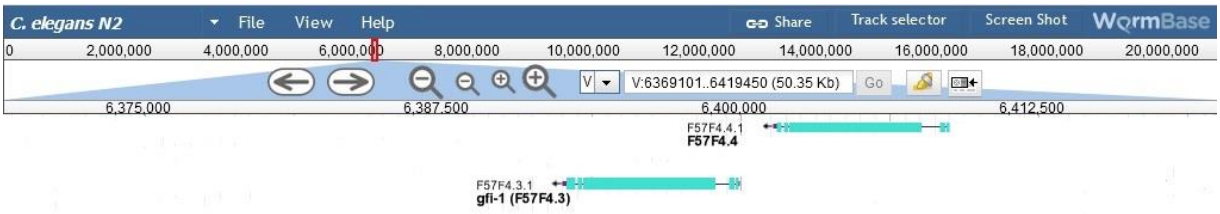


Fig.3.2 A view from the Wormbase genome browser, showing the location of *fstr-1* (F57F4.3) and *fstr-2* (F57F4.4) genes.

Both genes have 5 exons and 4 introns. Exons 1,2,4,5 are with small size, a few hundred base pairs each, while exon 3 carries most of the coding sequence with size above 5000bp. The length of the introns is relatively small, the largest intron is intron 2 with length several hundred base pairs (Fig.3.3)

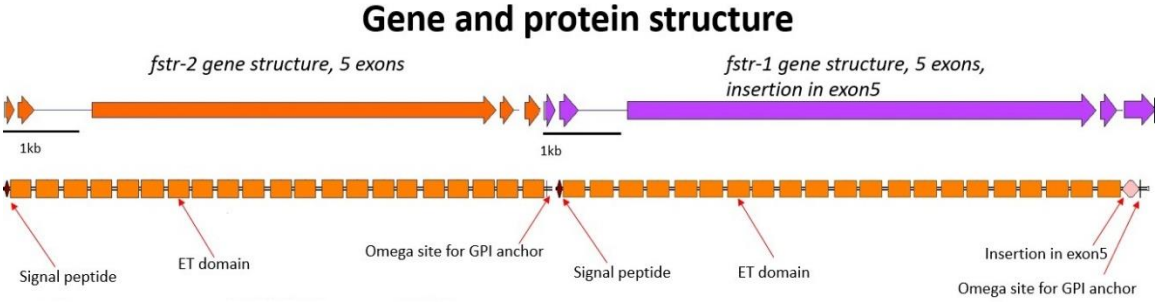


Fig.3.3 Gene and protein structure of *fstr* genes.

The main difference in the coding sequence between *fstr-1* and *fstr-2* genes is a 189bp insertion in the fifth exon of the *fstr-1* gene. We consider this an insertion because when we compared the sequences of potential *fstr* orthologs from all the available *Caenorhabditis* species genomes (www.caenorhabditis.org), the majority of the species have only one ortholog and the sequences do not contain inserted nucleotides in the fifth exon of the *fstr-1* gene, therefore we think that it is likely that *fstr-2* gene is the ancestral one and *fstr-1* arose from a duplication event somewhen in the evolution of *C. elegans*.

We analyzed the promoter regions of *fstr-1* and *fstr-2* genes and the closest orthologs. For that analysis a region of 1.5kb upstream of the translational start was used (sequences were taken from Wormbase)[5]. The MEME suite was used to run the analysis of the promoter sequences[12].

The motifs with lowest *p*-value are located in an area of around 500bp upstream of the translational start (Fig.3.4). For each promoter tested, the motifs are similarly located and in the same order so it is likely that this region corresponds to the core promoter region of the genes. It's been shown in the literature that *elt-2* a GATA transcription factor which plays a critical role in gut development also regulates the transcription of many intestinal genes in adult animals and in all promoters a motif that contains the consensus sequence for a GATA binding site (HGATAR) exists[13]. Interestingly, another transcription factor PQM-1, a protein that has C2H2-type zinc finger and leucine-zipper domains recognizes the exact sequence TGATAAG, which is detected in *fstr* genes promoters. In the literature, this sequence is called DAF-16 Associated Element (DAE)[14]. The DAE motif

is found in the genes that are involved in metabolism, growth, development and reproduction, and the localization of the PQM-1 protein antagonizes the one of DAF-16[14]. Recent studies of the proteins expressed in the germ line showed that expression of the *fstr* genes, there, is PQM-1 dependent[15]. It is possible that PQM-1 regulates the expression of the *fstr* genes in the intestine too. Another thing to note is a motif the contains the *skn-1* transcription factor core recognition sequence (RTCAT). *Skn-1* also regulates the transcription of many intestinal genes and many stress response and immunity genes.[16] Interestingly, in the *fstr-1* promoter, and in the *C. inopinata* promoter, a phylogenetically close species to *C. elegans*, this motif is not detected compared to *fstr-2* promoter. *Fstr-1* promoter also lacks two other motifs, which are otherwise detected in all other tested promoters, this suggests that there may be differences in the transcriptional regulation of *fstr-1* and *fstr-2* genes.

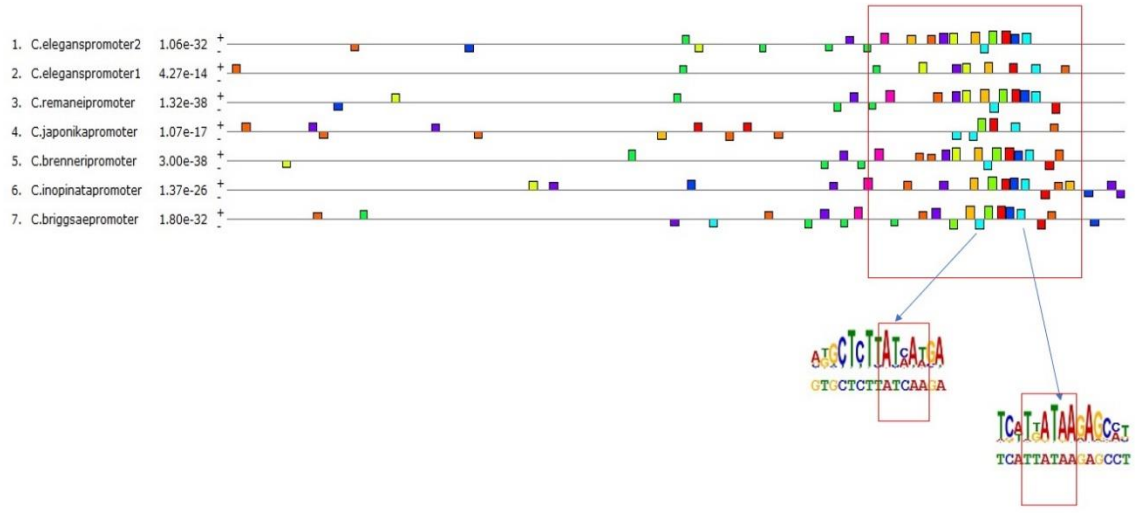


Fig.3.4 Output from the MEME suite for a 1529bp region upstream of the translational start of the *fstr* genes in *C. elegans* and the closest orthologs. The region in the box is localized around 500bp upstream of the translational start. Below the box are two of the detected motifs showing potential binding sites for PQM-1, and SKN-1 transcription factors

The promoter regions of *fstr-1* and *fstr-2* genes that were used to build different plasmid constructs both contain this region around 500bp upstream of the translational start, but overall were shorter than the 1.5kb region that was analyzed with the MEME software. When choosing what promoter region of each gene to amplify, some technical considerations were applied such as presence or absence of certain restriction sites, presence of repetitive sequences, which might make the process of PCR amplification of the promoter region problematic.

Also we've analyzed the 3'UTRs of *fstr-1* and *fstr-2* genes using "TargetScanWorm" software (www.targetscan.org)[17]. The program predicts biological targets of miRNAs by searching for the presence of conserved 8mer, 7mer, and 6mer sites that match the seed region of each miRNA[18]. "TargetScanWorm" uses a database of the seed regions for *C. elegans* microRNA families[19]. The program accepts accession numbers for genes from wormbase, takes a region of 150bp after the stop codon of the genes and subjects it to a scan for potential microRNA binding sites. When comparing the 3'UTRs of fstr-*1* and *fstr-2* genes the program shows that there are binding site for microRNAs common for both genes and sites unique for one of the genes, which are not detected in the other. This observation predicts differences in the post-transcriptional regulation of *fstr-1* and *fstr-2* genes (Fig.3.5)

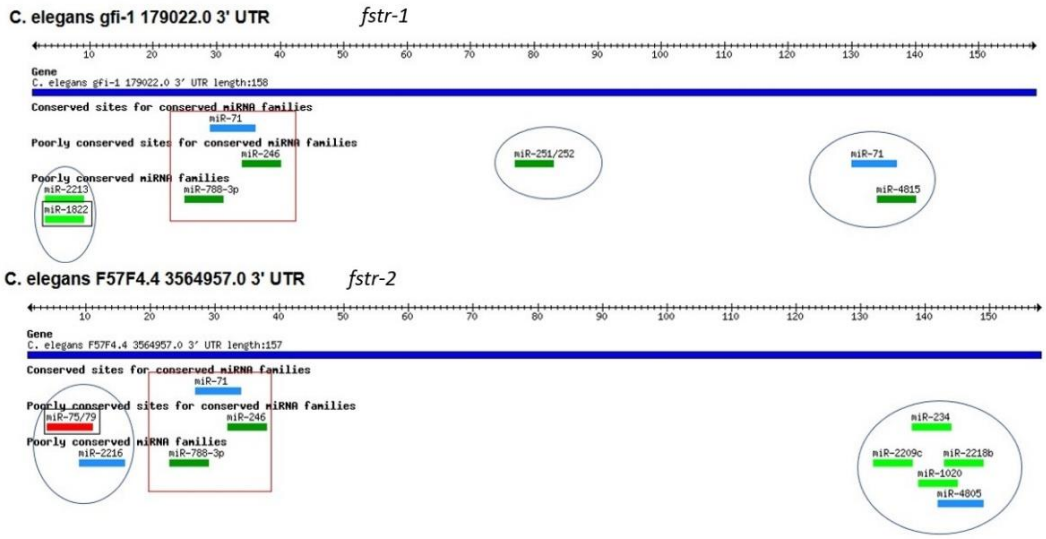


Fig.3.5 Output from "Target Scan Worm" software. In the boxes are the miRNA potential binding sites that are equal for both genes, in ovals are the sites unique for each gene.

Structure of FSTR-1 and FSTR-2 proteins

To analyze the protein structure we used online resources, NCBI CDD (conserved domain database)[20], pfam database[21], InterPro database[22], PredGPI software[23] and MEGA X (molecular evolution genetics analysis) software[24], which a standalone program, was used to generate multiple sequence alignments of protein sequences of FSTR proteins from the available sequences in the *Caenorhabditis* genomes database (www.caenorhabditis.org).

According to the prediction programs used, FSTR-1 and FSTR-2 proteins from *C. elegans* have a signal peptide at the N-terminus of the proteins, 21 ET domains, FSTR-1 has a 63 amino acids insertion after the last ET domain, the sequence is predicted to

resemble those from the endomucin superfamily[20], ω-site for GPI anchor attachment[23,25] and ends with a stretch of hydrophobic amino acids which are recognized, along with the omega site by GPI anchoring enzymes. (Fig.3.3). The signal peptide sequence, to transport to ER, is very well conserved in the *fstr* genes orthologs from the *Caenorhabditis* genus, and has a conserved positively charged amino at the very beginning of the signal peptide followed the by a hydrophobic string of amino acids also conserved, which is likely to orientate the protein to have its N-terminus become a transmembrane part in the ER membrane and the C-terminus of the protein will face the ER lumen (Fig.3.6)

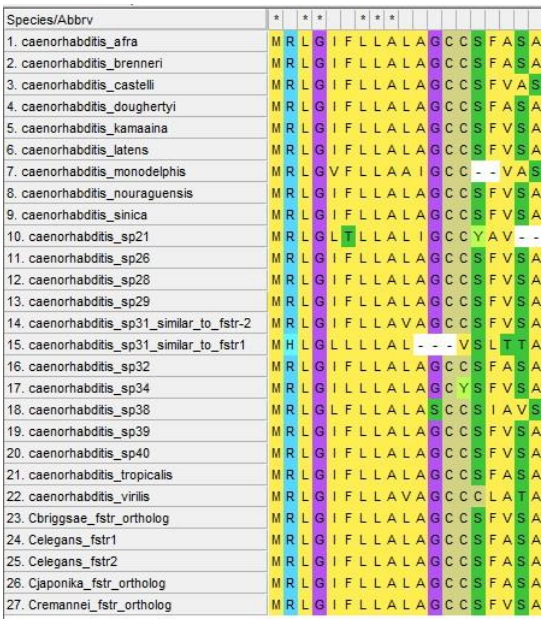


Fig.3.6 Multiple sequence alignment of the signal peptide sequence in the *fstr* genes from Caenorhabditis species

FSTR-1 and FSTR-2 are big proteins with FSTR-2 being 2090aa and FSTR-1 2153aa (63 amino acids insertion at the C-terminus of the protein). Overall, the two proteins are 96%

identical at the amino acid level. Most of the proteins are comprised of ET domains. Both FSTR-1 and FSTR-2 proteins have 21 predicted ET domains[20]. Each ET domain is approximately between 70 and 80 amino acids and has 10 conserved cysteines. According to the pfam database the ET domain appears to be nematode specific and with unknown function[21]. Between the ET domains there is a short stretch of amino acids, between ten and twenty in length, which do not show any significant conservation. In the multiple sequence alignment of the FSTR proteins the number of predicted ET domains varies with the species, but in every ET domain the positions of the 10 cysteines appears to be fully conserved and the conservation of the amino acids flanking them is also high. At the end of each ET domain there's a specific motif of 3 fully conserved cysteines followed by 4-5 other amino acids which also appear to be highly conserved for each ET domain between species but are different for each ET domain in one species, followed by a fully conserved cysteine. A similar motif is observed at the C-terminus of protein domains of similar length found in mammalian proteins (Fig.3.7)

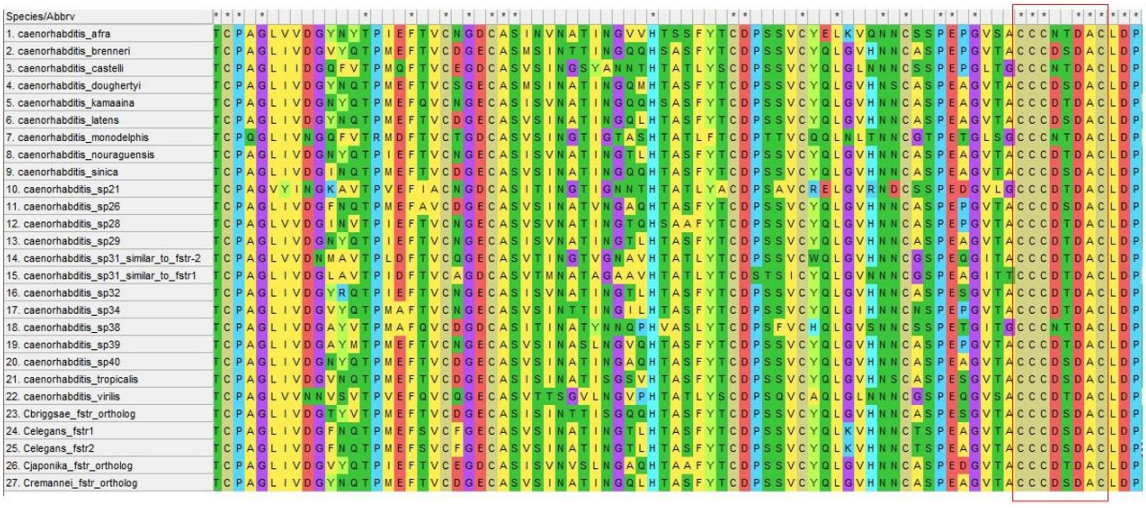


Fig.3.7 Multiple sequence alignment of one ET domain in the *fstr* genes from Caenorhabditis species

After the last ET domain, there is a short stretch of amino acids (10-20) before the predicted ω-site for GPI anchor attachment. The sequence that follows the omega site carries the signatures of the GPI anchor signal found in other animals[26]. FSTR proteins, with so many repeated ET domains, do not have clear orthologs in other species. The closest mammal ortholog mentioned in the literature is STABILIN-2 protein in humans and it shows only 33.7% similarity, based on the presence of similar protein domains in the FSTR proteins and STABILIN-2.[27] Stabilin-2 is a large protein, with several isoforms[28]. The protein has three different types of predicted domains in its structure: fascilin domains, EGF domains, and a C-type lectin-like hyaluronan-binding link module.[20] The similarity with the FSTR proteins comes from the structure of the EGF domains in the stabilin-2 protein. They are tandemly repeated, 2-4 domains with a short stretch of amino acids separating each EGF domain. If we look at two EGF domains together, their length and number of cysteines are similar to the length and number of cysteines in one ET domain (Fig.3.8)

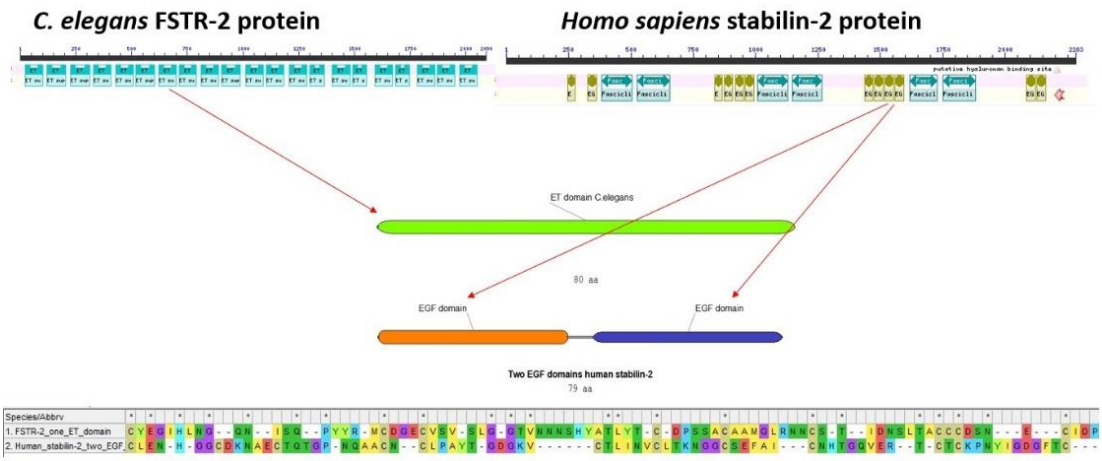


Fig.3.8 Multiple sequence alignment of one ET domain from *C. elegans* FSTR-2 protein and two tandemly repeated EGF domains from the human STABILIN-2 protein

According to the information in the NCBI the stabilin-2 protein is expressed primarily in the endothelial cells of the liver and spleen and in lymph nodes. Its primary function is as a receptor that can bind and endocytose ligands such as hyalurone, low density lipoprotein, Gram-positive and Gram-negative bacteria. If we are to compare the FSTR proteins with stabilin proteins, overall, they have similar amino acid length, the similarity between the two proteins comes from the similarity between 2 tandemly repeated EGF domains in the stabilin and one ET domain in FSTR proteins. According to the GPI anchoring prediction software[23] stabilin-2 protein does not appear to be GPI anchored, whether the FSTR proteins are predicted to be GPI anchored and are found in the lipid rafts from *C. elegans* membranes[25].

Another protein domain that is found in mammals and has similar structure to the ET domain is the LY6/UPAR (lymphocyte antigen-6/urokinase-type plasminogen activator receptor) domain (LU). The LU domain proteins are large family of proteins, which are expressed in different tissue types in mammals.

The LU domain is itself about 80 amino acids long, very similar length to ET domains also has 10 cysteines and the disulfide bridges they are predicted to form create a specific 3 finger structure[29] (Fig.3.9).

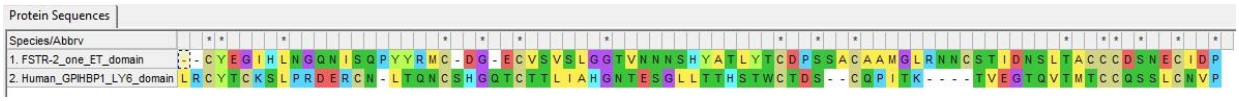


Fig.3.9 Multiple sequence alignment of one ET domain from *C. elegans* FSTR-2 protein and the LU domain of the human GPIHBP1 protein.

There are 35 genes in human and 61 genes in mouse containing this domain, and those genes are involved in very different functions and are expressed in many tissue types with all kinds of origin (ectodermal, mesodermal, endodermal). In human and mouse genes, usually the domain is present only once with some exceptions where it can be tandemly repeated two, sometimes, three times. The similarities with FSTR proteins come from the general protein structure. Both LU proteins and FSTR proteins have a signal peptide at the N-terminus for transport to ER. The LU domain/s is in the middle of the protein, so are the ET domains although they are much more in number, and at the C-terminus of both types of proteins there's a GPI anchor domain (not all members of the LU protein family are GPI anchored). FSTR proteins have been found to participate in lipid rafts isolated from worms and are predicted to be GPI anchored[25]. LU proteins perform many

different functions in mammals and are expressed in different tissue types some of them with endodermal origin. The one member of this family that might have similar function to FSTR proteins is the GPI-anchored high-density lipoprotein-binding protein 1 (GPIHBP1).

It is expressed on the luminal face of capillaries. Human Lipoprotein Lipase (LPL) binds to it. The data in the literature shows that both N-term acidic domain and LU domain play role in LPL binding. In mice deficient of the GPIHBP1 gene the LPL is mis-localized and the animals accumulate fat even when fed with a low fat diet[30,31].

Alignment of an ET domain and a LU domain (Fig.3.9) shows conserved positions of cysteines in both domains. It has been shown in the literature that mutations of the cysteines in the LU lead to reduced or abolished binding of LPL[32]. Aside from the conserved positions of the cysteines in both domains, there's not much conservation in the ET and LU domains. The blast search of an LU domain sequence even with less stringent parameters, against the database for *C. elegans* proteins did not return hits for ET domains in FSTR proteins.

The similarities between the LU domain present in the HGPIHBP1and the ET domains present in the FSTR proteins might be of interest because in *C. elegans* animals that have mutated the *fstr-1* and *fstr-2* genes, also show accumulation of fats.

Establishing the *fstr* mutant phenotype

We observed the *fstr-1,2* mutants under the microscope. Day one adult *fstr-1,2* animals were observed under DIC and polarized light first, and later when fluorescent reporters

were introduced also under fluorescent light. Initial observations with DIC showed that *fstr-1,2* mutants had increased size of lipid droplets, especially at the posterior of the animal and increased intestinal lumen size. Under polarized light a greatly increased number of visible gut granules was observed compared to wild type. The gut granules are lysosome related organelles in the intestinal cells, which contain birefringent substance, which makes them visible under polarized light. Gut granules contain acidified components, which are products of metabolism, serve as storage depo for zinc[33], contain anthranilic acid, which is produced by the kynurenine biochemical pathway from tryptophan. The anthranilic acid in the gut granules makes them fluoresce under UV light in blue and it's been shown in the literature that when the animal dies the anthranilic acid stored in the gut granules is released in a Ca^{2+}-dependent manner leading to a wave of cell necrosis[34]. These three features were similar to what we had observed before in HGS animals, therefore we sought to quantify them and also look for them in the single *fstr-1* and *fstr-2* mutants.

Assessing the amount of visible gut granules in *fstr* mutants

Based on the observations of the different strains of animals listed in the materials and methods section, five different categories were established to describe the amount and the distribution of gut granules in the intestine: Category 0 – no visible or few visible gut granules in one part of the intestine (the intestine of the worm was formally divided into 3 parts: anterior, middle, and posterior). Category 1 – several or many (approximately 20-100) of gut granules in one part of the intestine. Category 2 – numerous gut granules (>100) in one part of the intestine. Category 3 – numerous gut granules in one part of the

intestine and lots or numerous gut granules in another part of the intestine. Category 4 – numerous gut granules in all parts of the intestine (>300). Using the described categories, the majority N2, *fstr-1* and *clk-1* mutants, they have the *fstr-2* gene intact, fall into category 0 some animals into category 1 and only a few into category 2 (between 40 or 50 animals were photographed and many more observed). *Med-1; end-3* animals, they have downregulated *fstr-2* gene and upregulated *fstr-1* gene, are primarily into category 2 with some animals into category 1 and 0 and a few animals in category 3. Animals with *fstr-2* gene deleted, fall into category 3 and 4 with some animals into category 3. *Fstr-1,2* double mutants and *clk-1; fstr-1,2* triple mutant animals fall exclusively into category 4 (Fig.3.10)

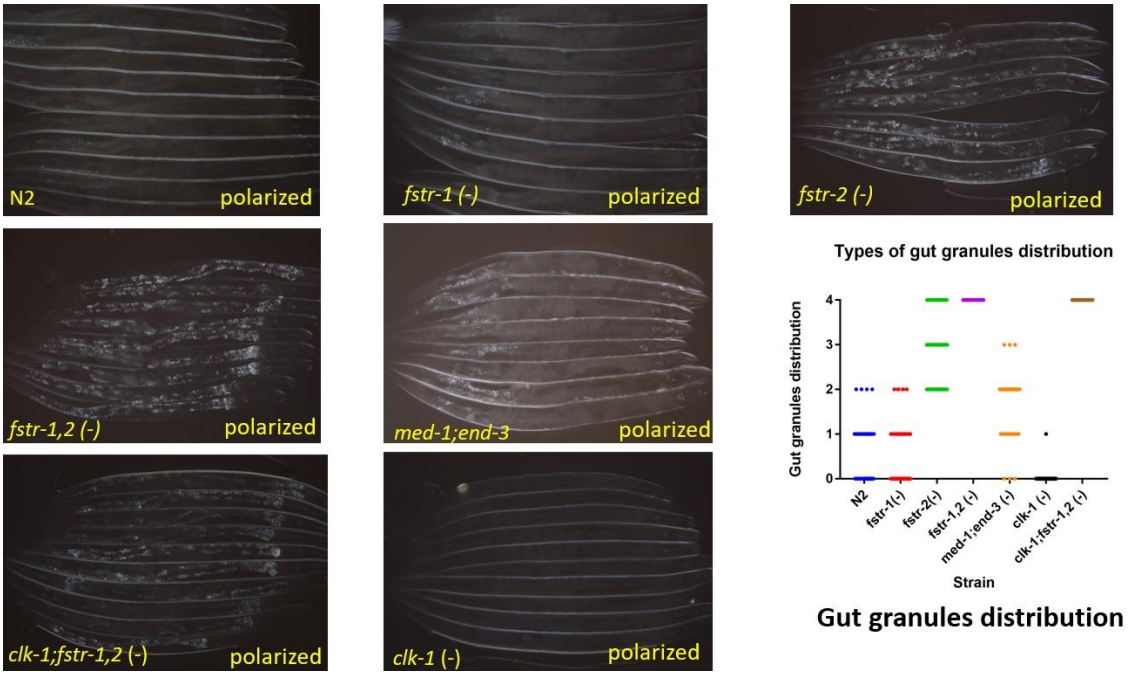


Fig.3.10 Patterns of visible gut granules distribution under polarized light

From prior work, increased amount of gut granules is a sign of caloric restriction as gut granules are lysosome related organelles which store acidified compounds such as fatty acids, zinc, anthranilic acid[32]. We hypothesize that deleting *fstr-2* gene, which is normally expressed in wild-type animals causes the animals to behave as calorie restricted even when the food is abundant and as a result the number of visible gut granules is increased. To test this, three experiments were done.

The first experiment was starvation of L1 larvae to observe the visible gut granules as animals are subject to starvation for 72h after hatching.

At hatching larvae from all the tested strains had numerous gut granules throughout the whole length of the intestine. At 24h of starvation most of the larvae from all the strains still have many gut granules throughout the whole intestine. On some larvae however was observed partial depletion of gut granules always happening from posterior intestine to anterior intestine for all the strains. From N2, *fstr-1*, and *clk-1* strains more animals were observed with partially depleted gut granules (this means no gut granules visible under polarized light) compared to the *fstr-2*, fstr-*1,2* and *clk-1; fstr-1,2* strains, and a few animals with completely depleted gut granules. At 48h of starvation was observed the biggest change in the gut granules content of the larvae. For N2, *fstr-1*, *clk-1* strains at least half of the observed animals were with completely depleted gut granules, whereas for the *fstr-2*, *fstr-1,2* and *clk-1; fstr-1,2* strains the number observed animal with depleted gut granules was less. For the animals that had only partial loss of gut granules due to starvation the same trend was observed: gut granules were depleting posterior to

anterior of the intestine. At 72h starvation only a few animals from N2, *fstr-1* and *clk-1* had visible gut granules usually only at the anterior part of the intestine, for the *fstr-2*, *fstr-1,2* and *clk-1;fstr-1,2* strains the effect of loss of gut granules followed the same pattern just a little more animals as a number were observed to have at least some gut granules left.

Another experiment was performed to test how the amount of visible gut granules changes in a state of calorie restriction. For this experiment, embryos from the N2 strain, which were obtained after hypo-chlorite treatment of adults were placed on 60mm NGM plates seeded with two different strains of *E. coli*: OP50 – the standard lab food for *C. elegans* and HT115 – a strain that is otherwise used for RNAi feeding and for which is known that overall is not very well tolerated from wild-type animals. We observed L4 larvae and one day old adults. On OP50, the L4 larvae had many gut granules throughout the whole intestine similar to the observations with L1 larvae. As the animals transitioned into adulthood in day one adults the gut granules have disappeared leading to the phenotype we have observed before (Fig.3.10) On HT115 however, when L4 larvae transitioned into adults the numerous gut granules did not disappear and day one adult wild-type animals had numerous gut granules throughout the whole intestine, similar to *fstr-1,2* mutant animals (Fig.3.11).

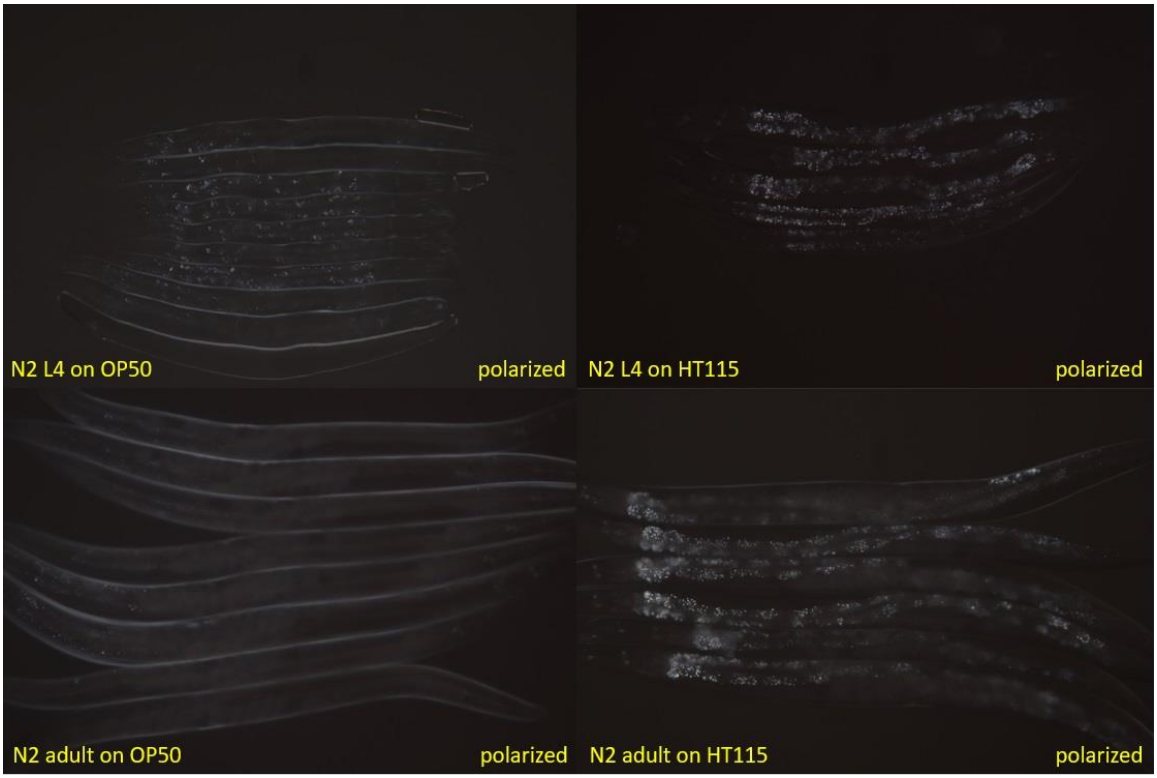


Fig.3.11 N2 animals L4 and day one adults on fed with OP50 *E. coli* (left) and HT115 *E. coli* (right). The HT115 strain is not well tolerated by wild-type *C. elegans*.

To further test the hypothesis that the amount of visible gut granules increases with calorie restriction, we subjected day one adult N2 animals to calorie restriction. We achieved this by putting them onto 60mm NGM plates, without uracil added into the media, seeded with *E. coli* OP50. The OP50 strain is uracil auxotroph so it makes very thin lawn on the NGM agar, creating conditions of calorie restriction. N2 animals were kept for 16h on the plates and at 6h and 16h some of them were anesthetized and photographed (see materials and methods). The results show that in conditions of calorie restriction the amount of visible gut granules increased (Fig.3.12), and the phenotype of the calorie restricted animals is similar to the *med-1; end-3* animals.

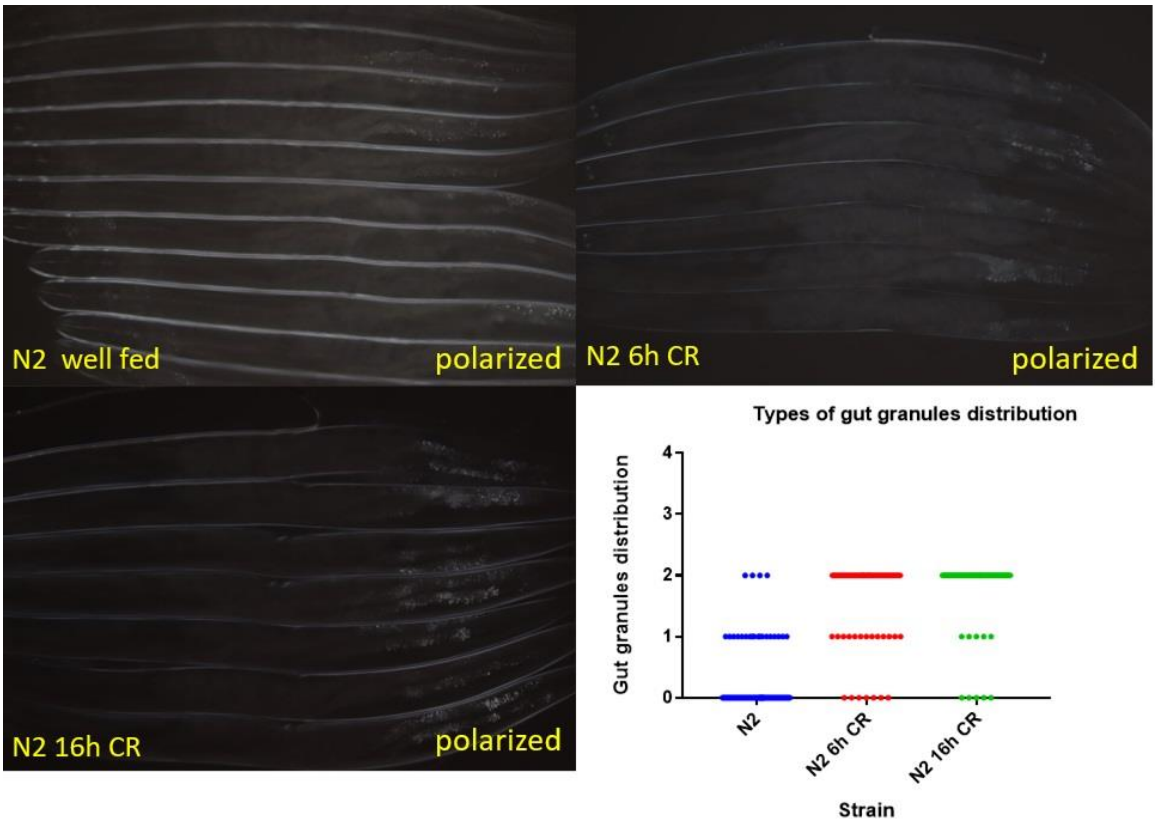


Fig.3.12 Types of gut granules distribution of N2 animals in calorie restriction.

To summarize, the strains that have *fstr-2* gene expression downregulated (*med-1; end-3*) or deleted *fstr-2* and *fstr-1,2* behave as calorie restricted even when the food is abundant compared to the strains that have *fstr-2* gene expressed (N2, *fstr-1*). In event of calorie restriction, the strains with low amount of gut granules (N2) show an increase of gut granules content which suggests that they respond to calorie restriction by starting to store nutrients rather than consume them immediately which happens when food is abundant.

Measurement of the intestinal lumen width

As initial observation of *fstr-1,2* double mutant revealed increased size of the intestinal lumen, we sought to quantify this phenomenon and compare with lumen size of wild-type and HGS animals. (Fig.3.13).

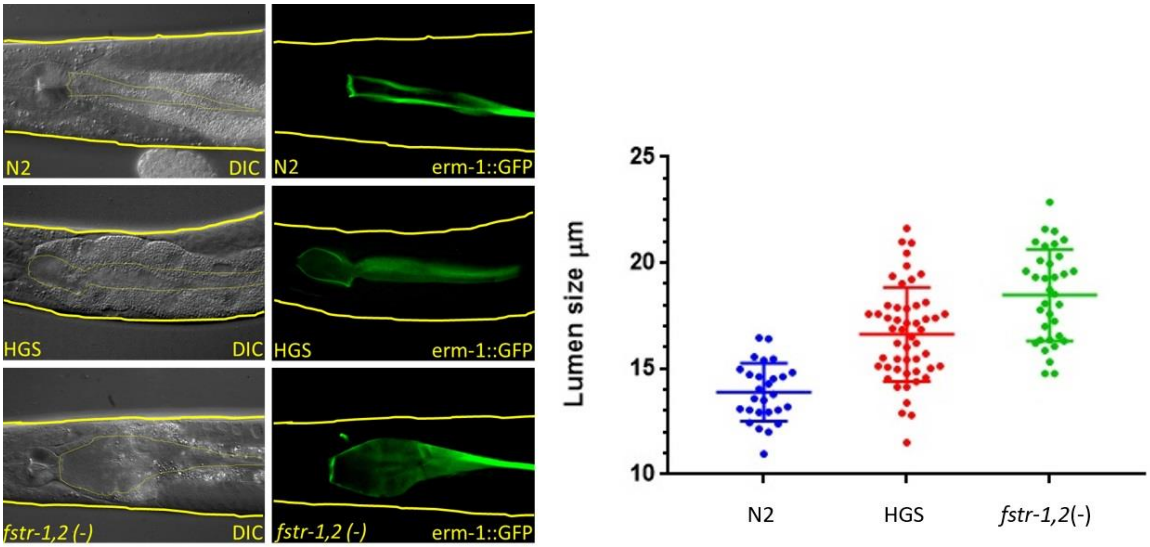


Fig.3.13 Measurement of intestinal lumen width of N2, *med-1; end-3*, and *fstr-1,2* strains

The average size of the lumen of the N2 animals was estimated to be 13.89µm, ± 0.2636 SEM for *med-1;end-3* (HGS on Fig.3.12) 16.62µm ± 0.3117 and for *fstr-1,2* 18.48µm ± 0.3711 SEM with p-value of one sample t-test for statistical significance of the differences between lumen sizes, $p<0.0001$.

To test our hypothesis that *fstr-1,2* mutants experience calorie restriction even when food is abundant we performed additional experiment in which animals from N2 strain, which containing the ERM-1::GFP reporter we starved for 6h and for 16h. To do this, animals

from the same plate that was used to quantify the lumen size of well-fed day one N2 adults were washed several times in M9 solution and placed in two Eppendorf tubes containing 1ml M9 solution and 0.5µl of 10mg/ml cholesterol. The tubes ware placed on rotisserie and after 6h of starvation one of the tubes was used to prepare animals for imaging. The second tube was used after 16h of starvation to estimate the lumen size of the animals. The experiment was stopped after 16h of starvation, because a "bagging" phenotype (eggs hatch inside the mother animal) was observed in some of the animals.

The results from the experiment show that after 6h hours of complete absence of food the lumen size of the intestine of wild-type animal was extended as much as what we've observed in *fstr-1,2* well fed mutants of the same age The median lumen size of the N2 animals subject to 6h starvation was 18.37µm ± 0.3759 SEM and of those with 16h hours starvation was 18.47µm ± 0.2487 SEM. There was no significant difference in the lumen size between 6h starved wild-type animals and 16h starved wild-type animals (Fig.3.14).

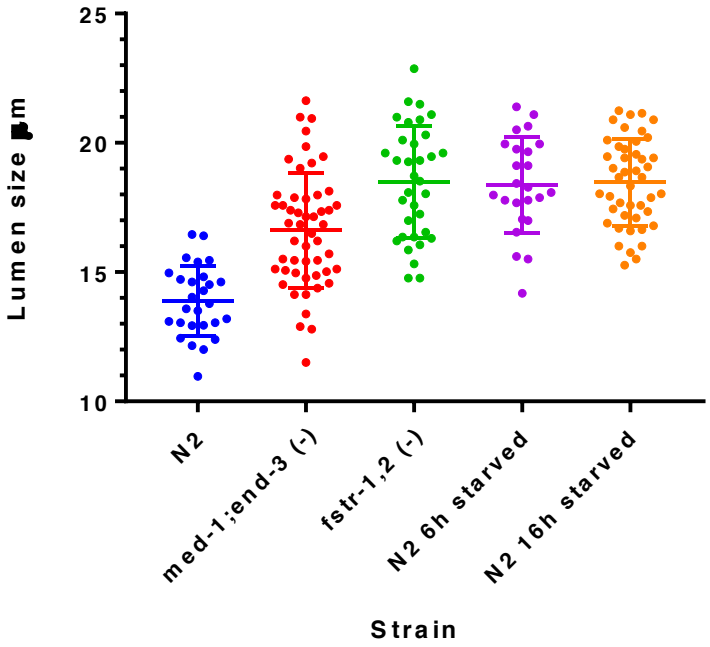


Fig.3.14 Measurement of intestinal lumen width of N2, *med-1; end-3*, *fstr-1,2*, and starved N2 animals.

Estimating the lipid droplet size of *fstr-1,2* mutants

After initial observations under DIC revealed that *fstr1,2* mutants have increased lipid droplets especially at the posterior of the animals we sought to quantify this phenomenon. The results of the experiment showed that the *med-1; end-3* and *fstr-1,2* strains have increased lipid droplet size compared to the control N2 strain under normal lab conditions and abundance of food. The average diameter of the lipid droplets of N2 animals was estimated to be 1.263µm ± 0.03229 SEM. *Med-1; end-3* animals have average diameter of 2.077µm ± 0.03807 SEM and *fstr-1,2* animals – 2.671µm ± 0.1073 SEM with p-value of one sample t-test for statistical significance of the differences between lumen sizes, $p<0.0001$ (Fig.3.15).

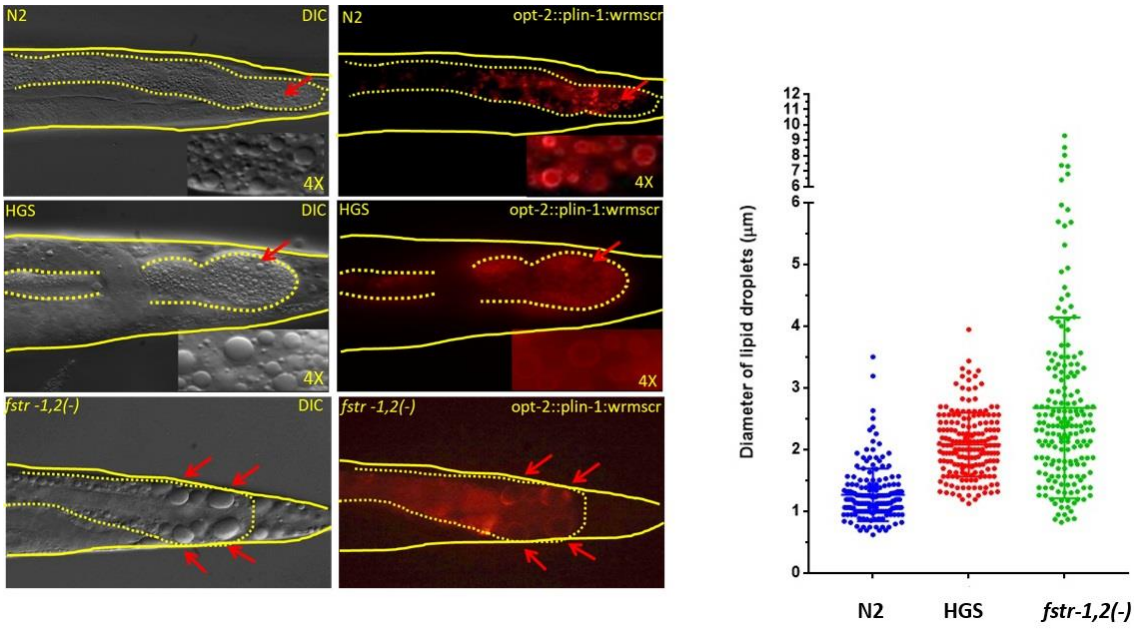


Fig.3.15 Estimation of lipid droplet size of N2, *med-1; end-3*, and *fstr-1,2* strains

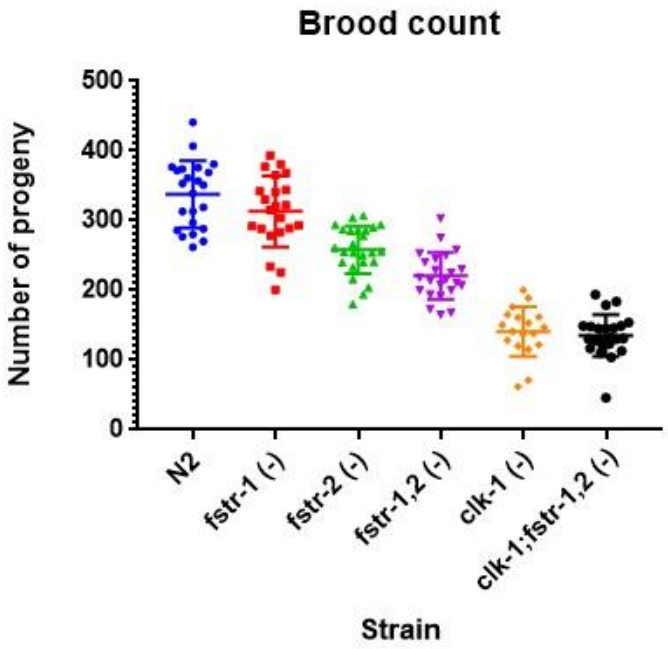
It is important to mention that *fstr-1,2* animals have very characteristic concentration of very large lipid droplets (sometimes around 10µm in diameter) at the posterior part of the intestine (Fig.3.15). Increasing lipid droplet size is another mechanism to store nutrients as it makes lipids less available for processing. This is another sign that *med-1; end-3* and *fstr-1,2* animals, the former have *fstr-2* gene downregulated, the latter have the *fstr-2* gene deleted strains experience calorie restriction even when food is not restricted.

Physiological characteristics of *fstr* mutants

Four experiments were done to assess some of the physiological characteristics of *fstr* mutants: lifespan, brood count, measuring the resistance to pathogenic infection from *Pseudomonas aeruginosa* strain PA14, utilizing two mechanisms of killing, toxin based and colonization of the intestine based.

Brood count of *fstr* mutants

Five trials were done and the results from each trial were combined to generate the brood count graph (Fig.3.16).

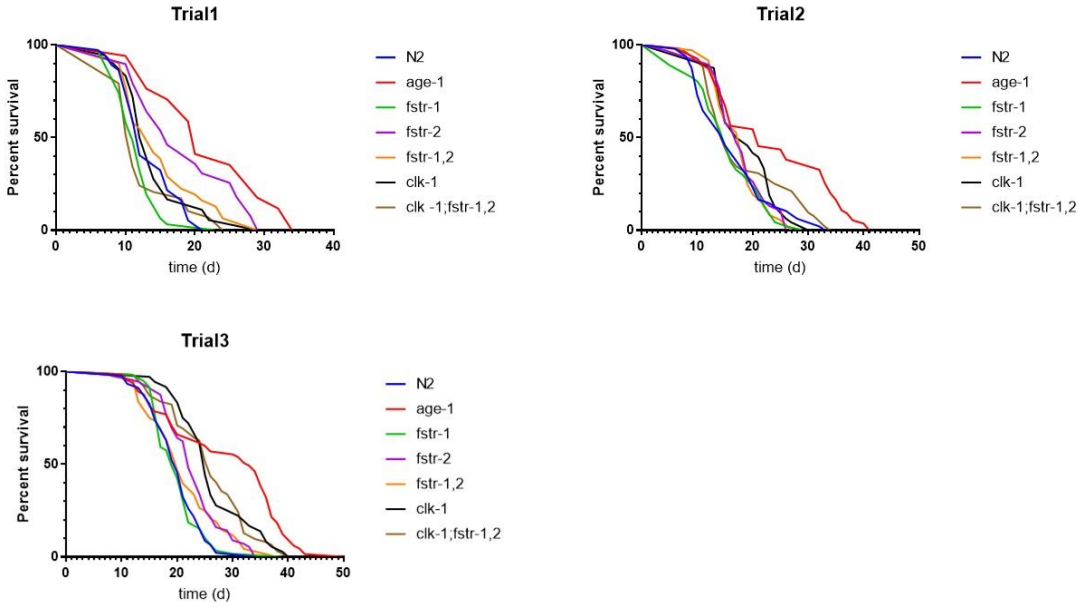


The results from the experiment were: N2 – 337 mean number of progeny ± 10.03 SEM, *fstr-1* – 313 ± 10.86 SEM, *fstr-2* – 258 ± 6.858 SEM, *fstr-1,2* – 220 ± 7.061 SEM, *clk-1* – 140 ± 8.382 SEM, *clk-1; fstr-1,2* – 135 ± 6.264 SEM. *Fstr-2* and *fstr-1,2* mutants had a lot of unfertilized eggs. The results coincide with our predictions, that strains which have the *fstr-2* gene deleted should have reduced brood count as these animals are in a state of calorie restriction even when there's plenty of food. The strains that have *clk-1* mutation have even more reduced brood size as *clk-1* gene encodes a mitochondrial hydroxylase

involved in the production of ubiquinone, Q9 a product required for normal cellular respiration, leading to slowing down of all major functions in the animal.

Lifespan of *fstr* mutants

Three separate trials were done. (Fig.3.17)



The results from the experiments showed that there was no significant difference between the lifespan of N2 and *fstr-1* strains in all 3 trials. In two out of three trial *fstr-2* animals lived significantly longer than the wild type. There was no significant difference between the lifespan of N2 and *fstr-1,2* strains in all 3 trials. In two out of three trials *clk-1* lived significantly longer than N2, which is in concordance with the data available in the literature. In two out of three trials the *clk-1; fstr-1,2* mutants did not live significantly

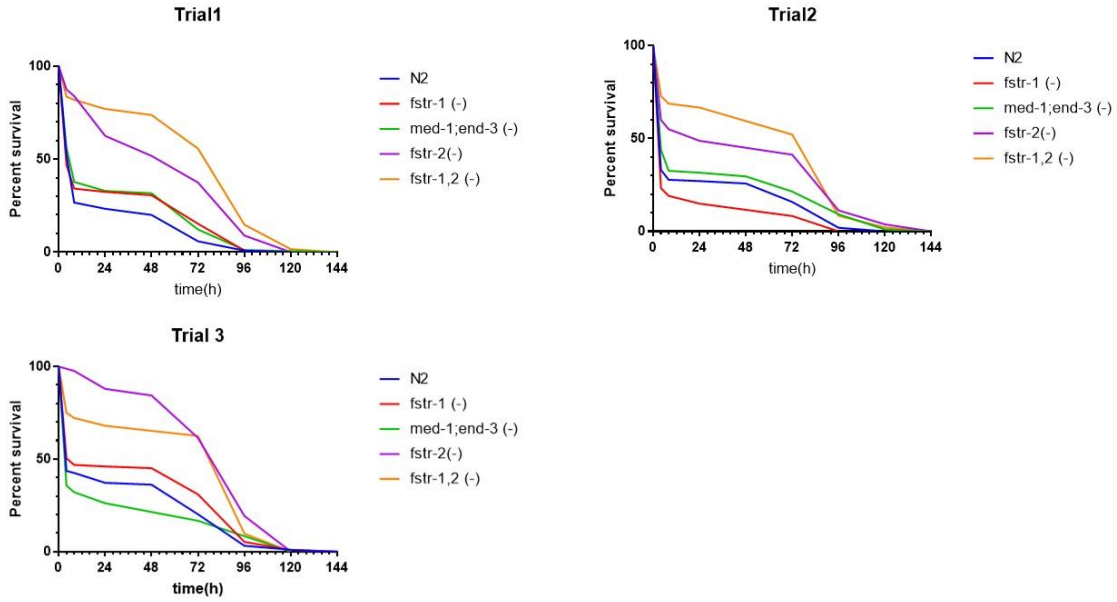
longer than the N2 strain. The deletion of both *fstr* genes apparently abrogated the increase in the lifespan caused by the *clk-1* loss of function mutation. A similar effect was reported before, by Cristina et al. 2009[2], when they RNAi both *fstr* genes. (Table 4 in Appendix A)

Assessing pathogenic resistance to toxin based (fast) killing and colonization of the intestine (slow) killing mechanism from *Pseudomonas aeruginosa* strain PA14

Pseudomonas aeruginosa is a pathogen for *C. elegans* and mammals. In humans it frequently causes an opportunistic infection in the lungs of patients with Cystic fibrosis. Depending on the conditions in which *Pseudomonas aeruginosa* grows, it produces different virulence factors and toxins[35]. *C. elegans* has been used to study host-pathogen interactions to elucidate the mechanisms of infection, and to isolate virulence factors[8]. It has been shown in the literature that *fstr-2* gene gets upregulated in response to infection with *Pseudomonas aeruginosa* PA14 and other pathogens[36,37], therefore we decided to test the resistance of *fstr* mutants to two mechanisms of killing by *Pseudomonas aeruginosa* PA14: the 'fast' killing mechanism, which is toxin based, and the 'slow' killing mechanism in which *Pseudomonas aeruginosa* colonizes the intestine of *C. elegans*. In the fast-killing mechanism *Pseudomonas aeruginosa* PA14 is grown on high osmolarity media, which contains peptone, glucose and sorbitol (PGS media). In this conditions the bacteria produce phenazine toxins, which has been shown to be a major factor in killing the host cells although it has been observed that the process is multifactorial[35]. The mechanism of toxicity of phenazines has been well studied, and

involves their redox capabilities, which can lead to production of reactive oxygen species, and inhibiting antioxidant enzymes in the host cell[38].

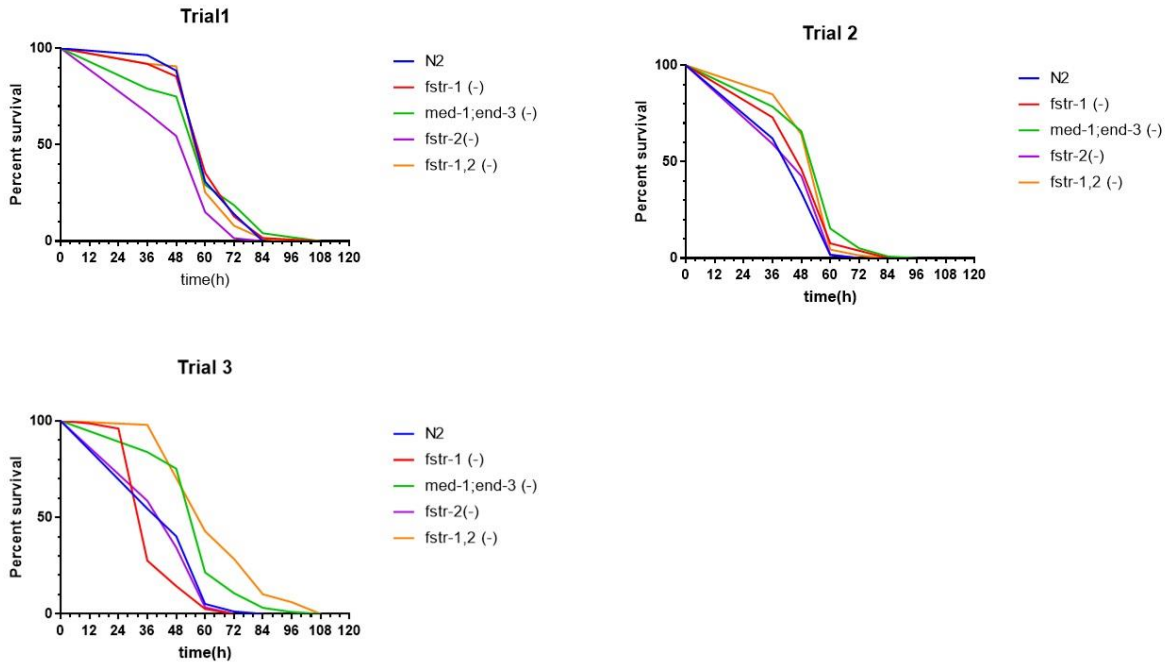
The results from 3 trials showed that *fstr-2* and *fstr-1,2* animals have increased resistance to the toxin based killing mechanism compared to the other strains with median survival of *fstr-2* animal ranging from 24h to 96h compared to median survival of wild-type animals ranging between 4h to 8h, and median survival of *fstr-1,2* animals of 96h. of (Fig.3.18; Table 5 in Appendix A)



When Pseudomonas is grown on NGM media it does not produce phenazine toxins so when *C. elegans* is infected with *Pseudomonas aeruginosa* PA14 under those conditions, the pathogenic bacteria colonize the intestine of the host.

Results showed that in two out of three trials *fstr-1,2* mutants were surviving longer than the wild type with median survival of 60h compared to 48h of the wild-type animals

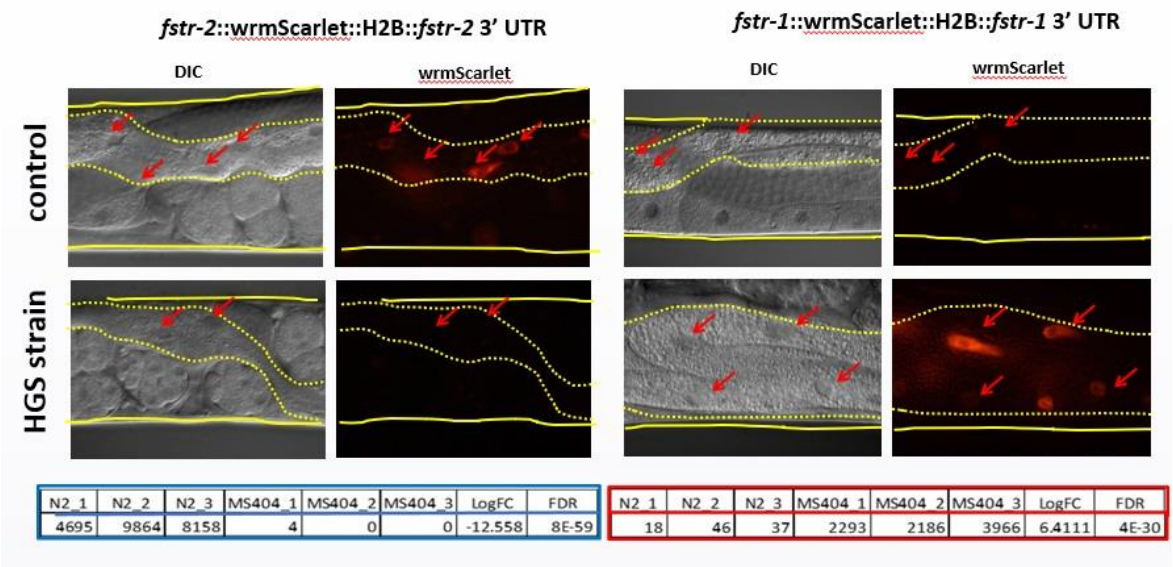
(Table.5 Appendix A). Interestingly *med-1; end-3* strains also survived longer, median survival 60h, than the wild-type animals in two out of three trials. *Med-1; end-3* animals have *fstr-2* gene downregulated (Fig.3.19)



Expression patterns of *fstr-1* and *fstr-2* genes

To analyze the expression pattern of *fstr-1* and *fstr-2* genes we built multiple plasmid constructs, both transcriptional and translational fusions, and constructs that were used to generate a repair template for CRISPR/Cas9 mediated integration of a fluorescent reporter in *fstr-1* and *fstr-2* genes.

(Fig.3.20,3.21,3.22)



N2_1	N2_2	N2_3	MS404_1	MS404_2	MS404_3	LogFC	FDR
4695	9864	8158	4	0	0	-12.558	8E-59

N2_1	N2_2	N2_3	MS404_1	MS404_2	MS404_3	LogFC	FDR
18	46	37	2293	2186	3966	6.4111	4E-30

Fig.3.20 Expression of transcriptional fusions of wrmScarlet reporter gene, driven by *fstr-1* and *fstr-2* promoter and 3'UTR sequences in Wild type and HGS strain (*med-1; end-3* was used). Below each picture are the results from and RNA seq experiment with RNA from intestinal cells showing the expression of *fstr-1* and *fstr-2* genes. (Pictures, Courtesy of G. Broitman-Maduro)

The results from the transcriptional fusions confirmed our results from RNA seq experiment done previously in the lab. *Fstr-2* gene is expressed in wild-type animals under normal conditions in the intestinal cells of the animals, and *fstr-1* gene is expressed in the HGS strain (*med-1; end-3*) under normal conditions (Fig.3.20).

A bioinformatics study from 2014[27], using data from 7 microarray sets identified *fstr-2* gene as down regulated in conditions of dietary restriction. Our observations with PID14 reporter show expression in the intestine and head neurons of wild-type animals, under normal conditions. In HGS strain and in starvation conditions (adult worms were put on plates without any food for two days and then observed), the reporter is downregulated in the intestine, but is expressed in head neurons, body wall muscle and coelomocytes (Fig.3.21)

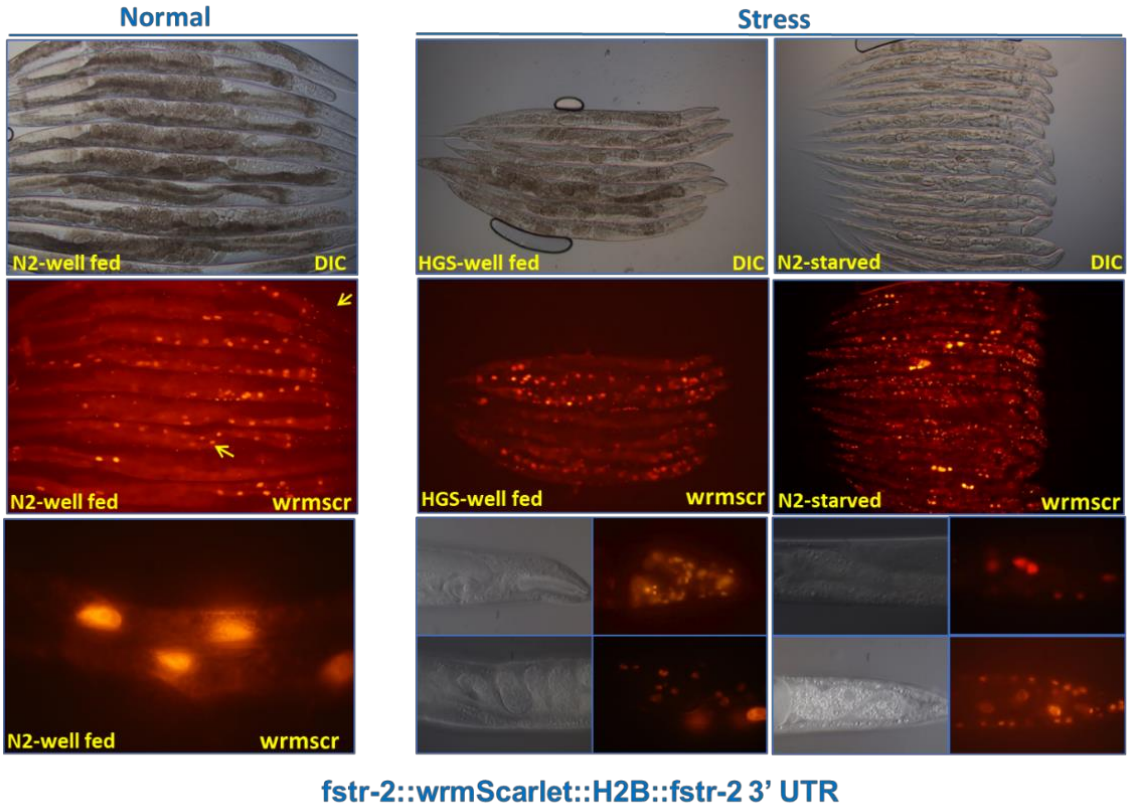


Fig.3.21 Expression of transcriptional fusion of wrmScarlet reporter gene, driven by *fstr-2* promoter and 3'UTR sequences in Wild type and HGS strain (*med-1; end-3* was used). In HGS and in starvation conditions, the reporter is downregulated in the intestine and expressed in head neurons, body wall muscle and coelomocytes. (Pictures, Courtesy of G. Broitman-Maduro)

The *fstr-1* reporter PID15 showed minimal expression in wild-type animals but was up regulated in the HGS strain and under starvation conditions. In all conditions some expression in head neurons was observed (Fig.3.22)

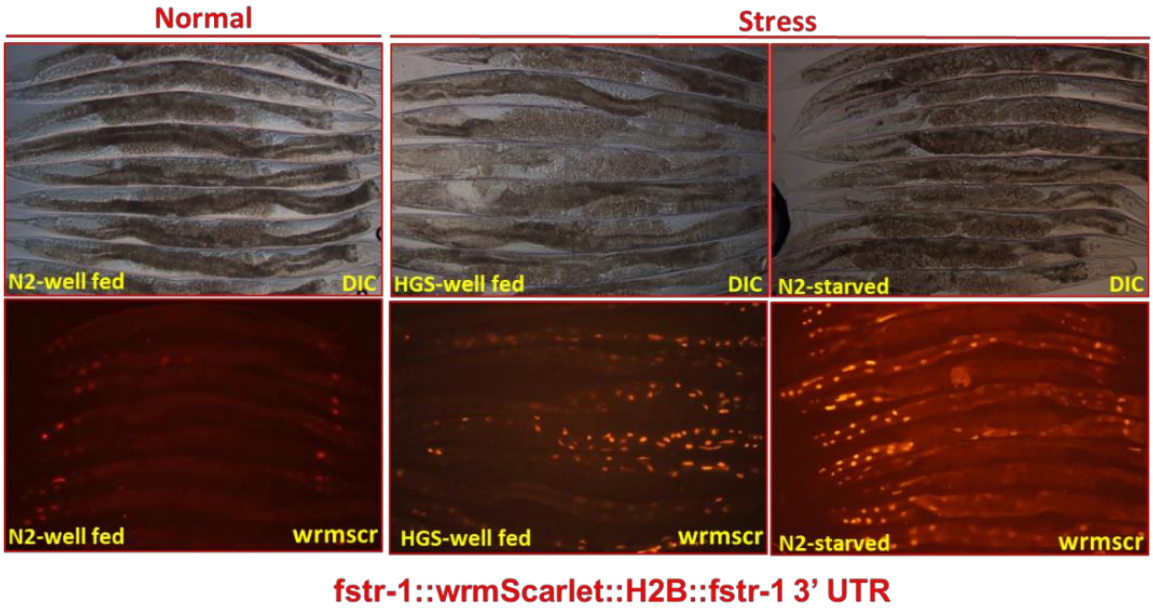


Fig.3.22 Expression of transcriptional fusion of wrmScarlet reporter gene, driven by *fstr-1* promoter and 3'UTR sequences in Wild type and HGS strain (*med-1; end-3* was used). In wild type a minimal expression was observed whether in HGS strains and under starvation conditions an increased expression in the intestinal cells was observed. (Pictures, Courtesy of G. Broitman-Maduro)

We've generated also translational fusions, which contain the whole *fstr-1* and *fstr-2* gene under the control of their own promoter and 3'UTR (plasmid constructs PID27 and PID28)

Both PID27 and PID28 were injected into N2 animals along with another plasmid containing the *rol-6* gene harboring a dominant mutation which makes the animals roll as a selectable marker. Both constructs were introduced as an extrachromosomal array. The results showed expression in the intestine, for FSTR-1 only in certain parts of the intestine with the localization of the protein apparently being cytoplasmic, and for FSTR-2 expression throughout the length of the intestine with apparent cytoplasmic localization (Fig.3.23).

The construct PID28 was also injected along with the *rol-6* selectable marker into *fstr-1,2* mutants to attempt to rescue the *fstr-1,2* mutant phenotype. Animals that were rolling were inspected for presence of visible gut granules, size of lipid droplets at the posterior of the animal and for expression of the construct. Out of 60 animals checked only 1 showed reduction in the visible gut granules and the expression pattern was cytoplasmic similar to the one shown on Fig.3.23. So the construct PID28 - p*fstr-2*:*fstr-2*:wrmSCR:*fstr-2* 3'UTR did not rescue *fstr-1,2* mutants.

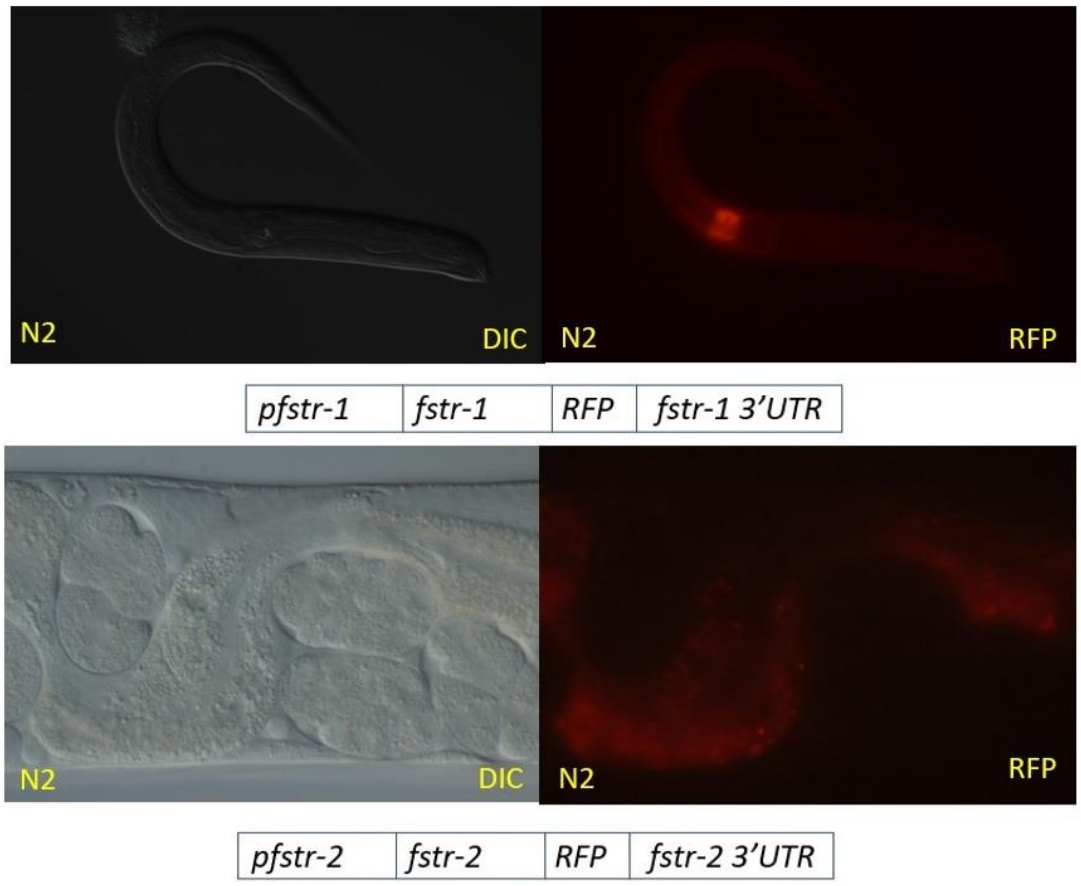


Fig.3.23 Expression of translational fusions of wrmScarlet reporter gene, driven by *fstr-1* and *fstr-2* promoter and 3'UTR sequences. The reporter gene is placed after the end of the coding sequence of the *fstr-1* and *fstr-2* genes.

Thus far, based on the above observations and on the data, we had from prior experiments described, we hypothesize that:

- *Fstr-1* could be a stress response gene
- *Fstr-2* gene is the one that functions in normal genetic background
- *Fstr-1* mutants show no visible phenotype in normal genetic background
- *Fstr-2* and *fstr1,2* mutants show a phenotype consistent with calorie restriction even when food is abundant.

Hence, we next decided to focus our efforts on analyzing the expression of the *fstr-2* gene.

Because the translational fusions injected as extrachromosomal arrays did not show the actual localization of the protein, we decided to integrate the wrmSCR reporter gene into different places of the *fstr-2* gene.

When the reporter gene was integrated at the N-terminus of the FSTR-2 protein this resulted in mis-localization of the protein in multiple vesicles in the intestinal cells. FSTR-2 protein has a signaling peptide sequence for translocation in ER in its N-terminus and is predicted to be a GPI anchored protein, attaching the GPI anchor is a posttranslational modification done also in the ER in which a specific protease cleaves the protein at a specific amino acid called omega site and then a transaminase enzyme attaches the GPI anchor to the protein. For the FSTR-2 protein the omega site is predicted to be a serine amino acid at position 2061[25], which is 11 amino acids after the last predicted ET domain and 29 amino acids before the C-terminus of the protein[20]. We think what we observe is the protein being mis-localized, because the animals with this

integration of the scarlet reporter into the *fstr-2* gene have increased amount of visible gut granules under polarized light the same way *fstr-2* and *fstr-1,2* mutants have, which suggests that the protein function is disrupted (Fig.3.24).

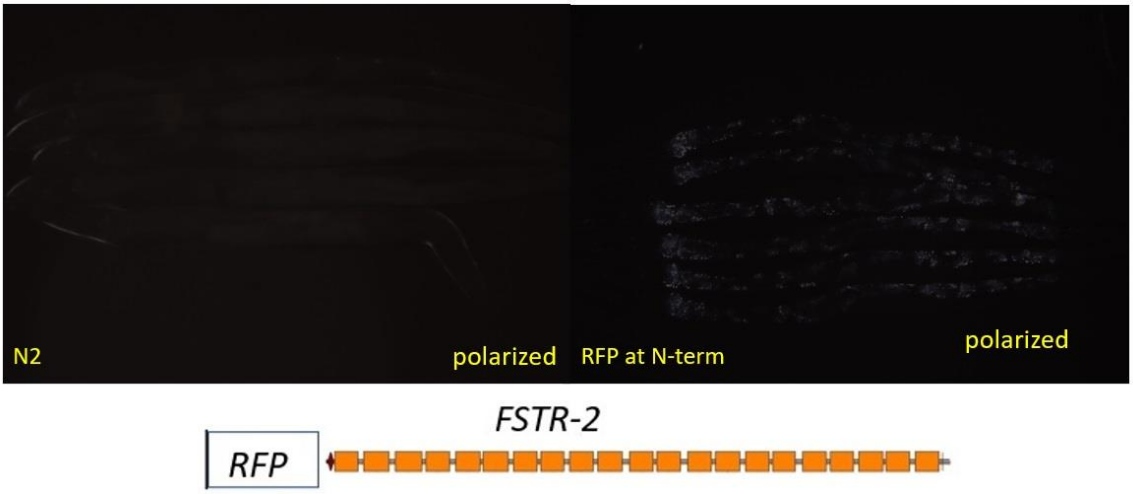


Fig.3.24 Observation of gut granules in N2 animals with integrated wrmScarlet reporter gene (RFP on the picture) at the N-terminus of the FSTR-2. The presence of numerous gut granules shows that the FSTR-2 function is disrupted.

The observation made in this integrant show that this version of the FSTR-2 protein is localized in multiple vesicles in the intestinal cells. Some of the vesicles appear to be gut granules (Fig.3.25)

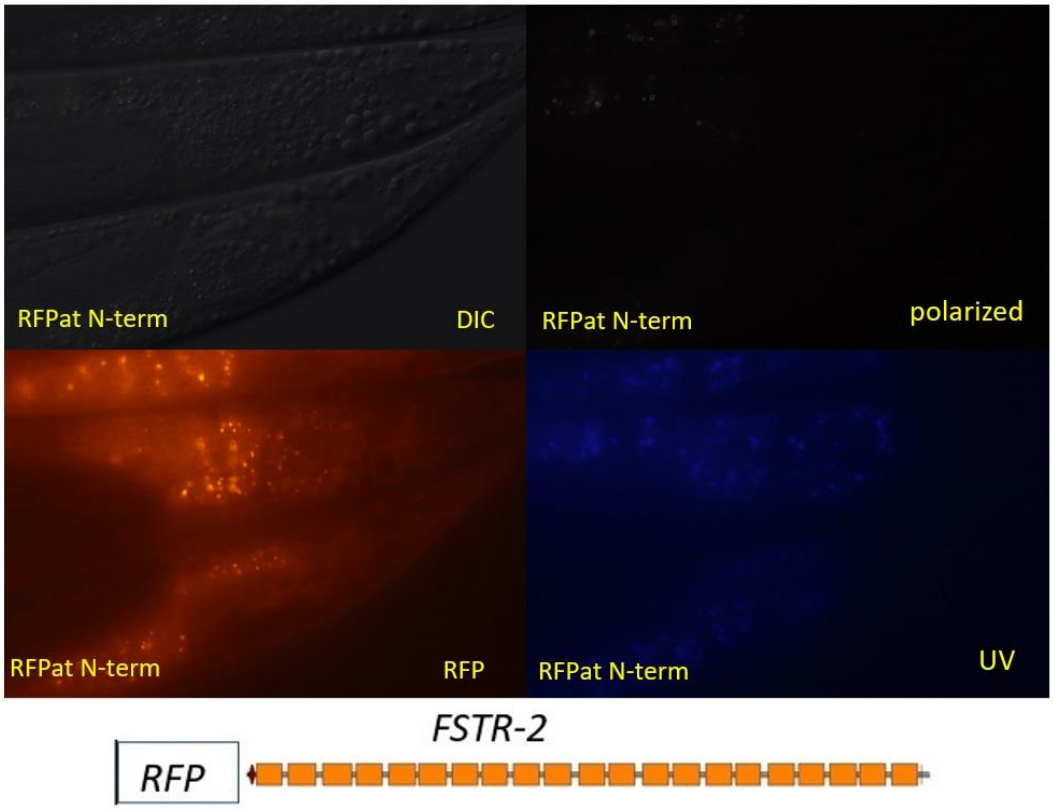


Fig.3.25 Expression pattern of FSTR-2 with integrated wrmScarlet reporter gene (RFP on the picture) at the N-terminus of the FSTR-2.

A possible explanation for the observations is that when the scarlet reporter sequence is translated before the start of FSTR-2 protein it obscures that signal sequence in the N-terminus of the FSTR-2 protein, which is a signal for translocation to ER. Additionally adding more than 200 amino acids to the protein may cause miss-folding which usually leads to marking those proteins for degradation and processing them in organelles such as lysosomes. It is possible that some of the observed fluorescent signal comes from the gut granules, since they are defined as lysosome related organelles and usually contain acidified compounds among other things, which could be products from the degradation of the protein.

We had similar observations when we integrated the scarlet reporter after the first two exons of the *fstr-2* gene (Fig.3.26;3.27). Again, the function of the FSTR-2 appears to be disrupted (increased amount of visible gut granules) and the fluorescent signal is coming from vesicles. In this integrant the signaling sequence for translocation to ER at the N-terminus is intact, but the scarlet reporter was integrated in between the ET domains of the FSTR-2 protein which probably causes mis-folding, and the protein function is disrupted.

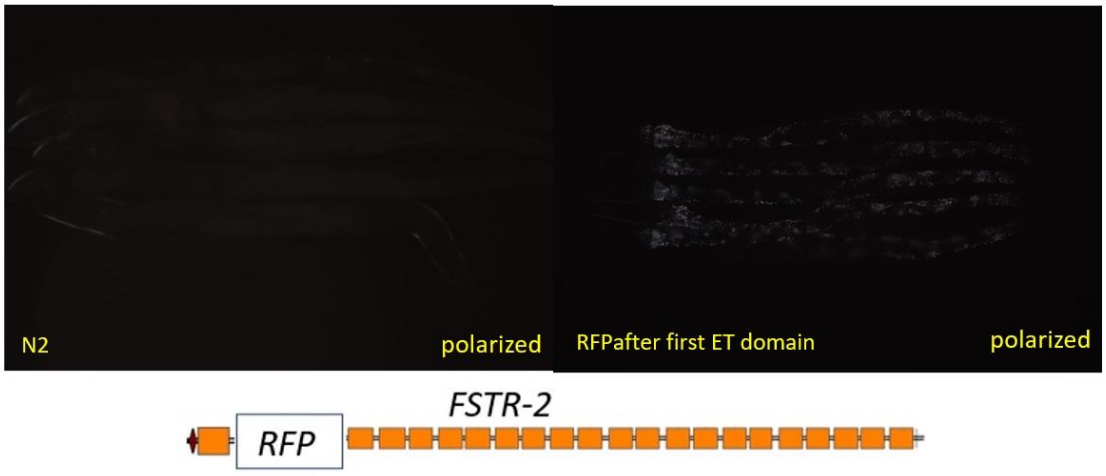


Fig.3.26 Observation of gut granules in N2 animals with integrated wrmScarlet reporter gene (RFP on the picture) after the first ET domain of the FSTR-2. The presence of numerous gut granules shows that the FSTR-2 function is disrupted.

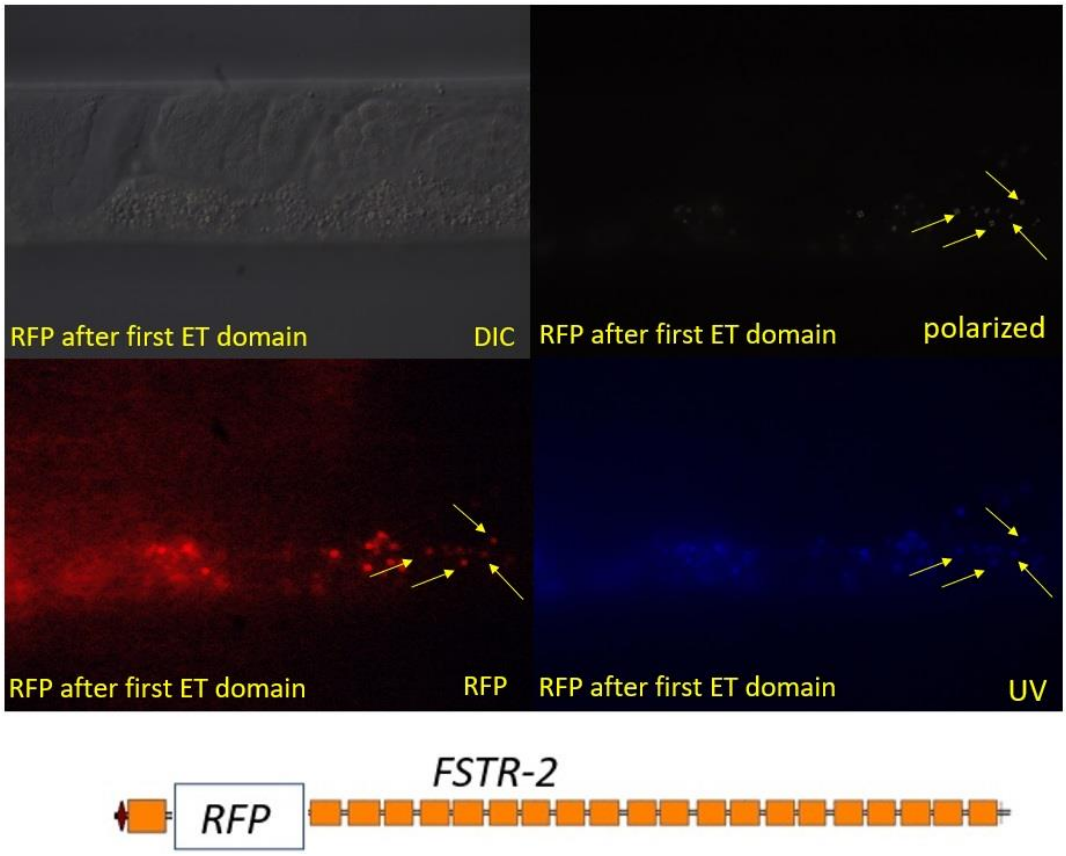


Fig.3.27 Expression pattern of FSTR-2 with integrated wrmScarlet reporter gene (RFP on the picture) after the first ET domain of the FSTR-2.

The third integrant line was established using another guide RNA (CC3) which is situated in the 3'UTR region of the *fstr-2* gene. The goal was to deliver a repair template in which the scarlet reporter is placed in frame right after the coding sequence of the fstr-2 gene. To generate the repair template *fstr-2*exon-4-intron4-exon-5:wrmScr: *fstr-2* 3'UTR a version of the plasmid construct PID28 was used which had extended *fstr-2* 3'UTR sequence.

When observed, most animals which have the wrmSCR reporter integrated that way in the *fstr-2* gene show little to no visible gut granules under polarized light (Fig.3.28). They

also show no increased size of lipid droplets at the posterior of the animal, suggesting that the function of the FSTR-2 protein is at least to some extent preserved.

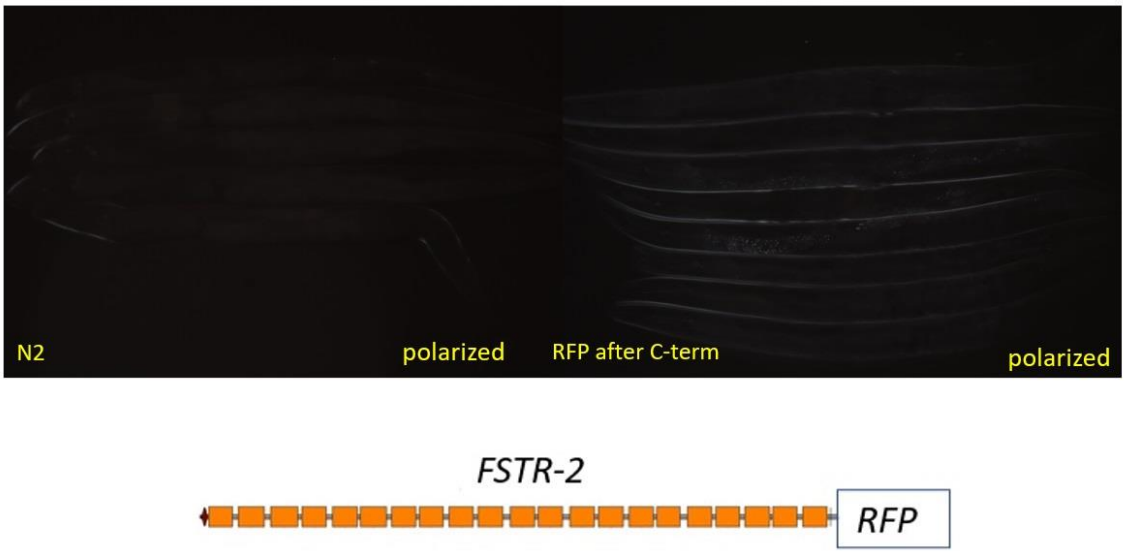


Fig.3.28 Observation of gut granules in N2 animals with integrated wrmScarlet reporter gene (RFP on the picture) at the C-terminus the FSTR-2. The absence of numerous gut granules shows that the FSTR-2 function is intact.

Observation under higher magnification showed that the protein is expressed in all parts of the intestine, with signal coming from certain vesicles and the nuclei of the intestinal cells (Fig.3.29). We were able to confirm with sequencing of genomic DNA from these worms that the expression pattern we observe is not caused by errors introduced in the process of homology directed repair initiated by the CRISPR-Cas9 system nor because sequence was inserted or deleted in the *fstr-2* gene.

A possible explanation for this expression pattern comes from the prediction that FSTR-2 is a GPI anchored protein. The GPI anchor is attached at the ω-site of the protein and the sequence after the ω-site gets cleaved by the GPI anchoring enzyme and the real GPI anchor is placed. Since our integration of wrmScarlet reporter is after the end of the

coding sequence of the *fstr-2* gene it is possible that in the ER lumen when the GPI anchor is attached, the fluorescent reporter got cleaved along with the last twenty something amino acids of the FSTR-2 protein. Because ER is continuous with the nuclear envelope it is possible that what we observe is a result of the fluorescent protein going into the nuclear envelope of the nuclei of the intestinal cells.

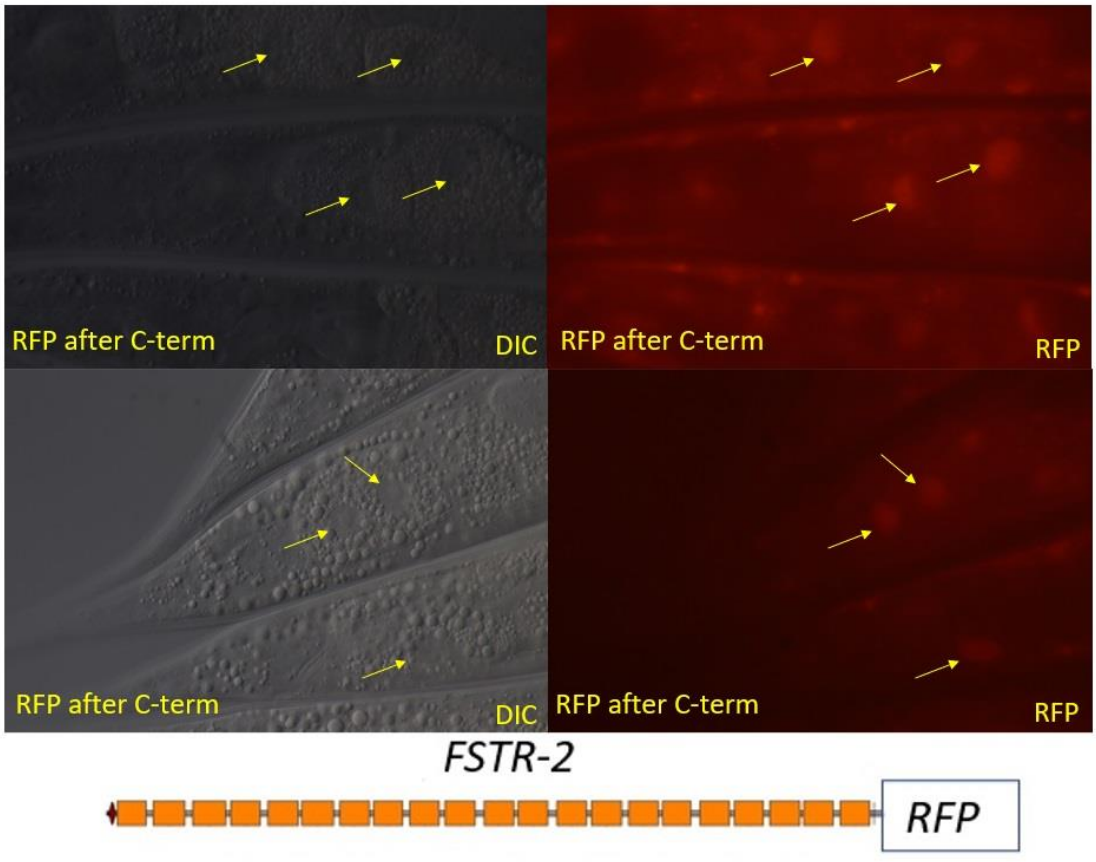


Fig.3.29 Expression pattern of FSTR-2 with integrated wrmScarlet reporter gene (RFP on the picture) at the C-terminus FSTR-2.

The data from the integration experiments, prompted us to design another construct, this time placing the RFP after the last ET domain but before the predicted ω-site for the GPI

anchor. Thus, PID52 construct was built, and first injected in wild-type animals as extra chromosomal array.

The expression pattern shows expression in the intestine with localization of the plasma membrane of the intestinal cells. (Fig.3.30).

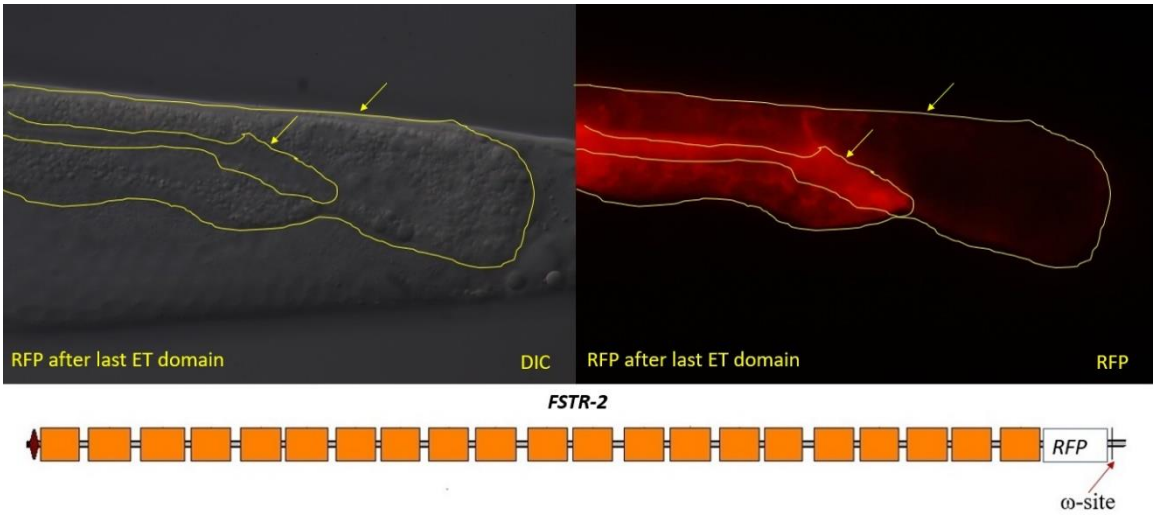


Fig.3.30 Expression pattern of PID52 construct as an extrachromosomal array. Arrows point at the plasma membrane of the intestinal cells, with top arrow pointing the basal plasma membrane and bottom arrow pointing the apical plasma membrane. (Pictures courtesy of G. Broitman-Maduro)

A repair template was generated using the the PID52 construct in an effort to integrate the RFP after the last ET domain in FSTR-2 but with no success so far.

Discussion

Function of the FSTR proteins

The results from our experiments show the FSTR proteins have a role in the normal functioning of *C. elegans* metabolism. Some of the changes we observed in *fstr-2* and *fstr-1,2* mutants are similar to changes we observe in HGS strains, such as increased lumen size; some are more pronounced, such as the increased amount of visible gut granules, while some, like the significant increase of the size of the lipid droplets in the posterior of the animal, are unique for the phenotype. To figure out what could possibly be the function of the protein we have also looked for similar proteins in other organisms. There are no clear orthologs of FSTR proteins in other organisms, mainly because the main protein domain, ET domain, which is the majority of the FSTR proteins, appears to be nematode specific. When we looked for similar protein domains, we found similarities between the ET domain in FSTR proteins and the LY6/UPAR domain in mammals. The proteins that contain LY6/UPAR domain perform many different functions in mammalian cells, and they are expressed in many different types of cells some of them not with endodermal origin, but proteins like the aforementioned GPIHBP1, the LY6/UPAR domain, which is the main functional domain and also have another feature of the FSTR proteins that is likely a GPI-anchored protein. Mutations in GPIHBP1 cause changes in lipid metabolism, in terms of lipids content changes and changes of the expression of genes involved in the lipid metabolism in mammals, so it is possible that the function of FSTR proteins is something similar, we observe a change in the lipid droplet distribution in *fstr-2* and *fstr-1,2* mutants. Another similarity, again based on

similarities in the protein domain structure is with the stabilin-2 protein in humans. Even though stabilin proteins are not predicted to be GPI anchored, their function in metabolism which was mentioned above could be similar to what the function of the FSTR proteins is in *C. elegans*.

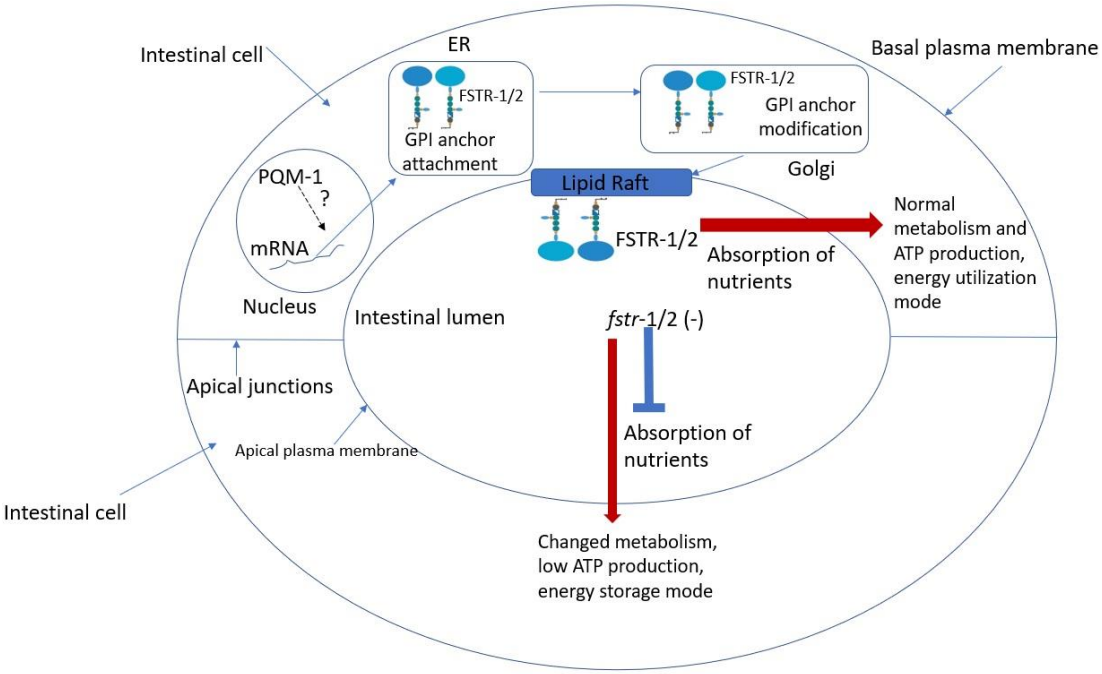
Another observation we made that gives information about the possible function of the FSTR proteins comes from the experiments measuring the resistance to toxic based killing from *Pseudomonas aeruginosa* PA14. The *fstr-2* and *fstr-1,2* mutants were highly resistant, compared to the wild-type and HGS animals. It has been shown in the literature that in mammals toxins from pathogenic microorganisms bind to GPI anchored proteins, which normally reside on the plasma membrane[39]. It is possible that the FSTR proteins function as a structural GPI anchored protein expressed on the plasma membrane surface to which small molecules like toxins or nutrients bind so when the protein is missing in the mutants the normal uptake of the nutrients is disrupted, leading to a behavior of the system like it is under caloric restriction. Another line of evidence for FSTR proteins being structural proteins comes from proteomic studies, in which the main binding partner of FSTR proteins are found to be different types of galectins (a family of β-galactosidase binding proteins)[40,41]. Galectins are expressed in intestinal cells. They are proteins that recognize β-galactosides in different glycoconjugates, and they participate in various cellular processes.[41]

Recent findings about the presence of FSTR proteins in the lipid rafts of the plasma membrane in the germline cells[15] (we had some observations from our translational fusions reporters about germline expression too), show that FSTR proteins appear to be

involved in the proliferation of the germ line cells, as double *fstr* mutants have reduced number of proliferating germline cells, thus a lower brood count[15], which is in concordance with our observations. The exact function of the FSTR proteins in the germ line cells remains unclear, the available data shows that FSTR proteins influence the germ line cell proliferation in a mechanism not related to the Delta-Notch signaling and TGF-β-DAF-4 signaling cascades, for which it is known that regulate proliferation[15]. These findings do not exclude the possibility that FSTR proteins function in the uptake of nutrients from intestinal and germline cells.

Our experiments show that FSTR-2 is one of the necessary factors for normal metabolism in *C. elegans*. In the HGS strains the down regulation of *fstr-2*, probably a result from untimely activation of the *elt-2* transcription factor during development, is likely to contribute to the phenotype, observed in HGS animals. We showed that mutations of the *fstr* genes have a pleiotropic effect on the physiology of wild-type *C. elegans*. *Fstr-1,2* animals were more resistant to infection from *Pseudomonas aeruginosa* PA14 than *fstr-2* animals, showing that deletion of both genes gives better resistance, on the other hand *fstr-1,2* mutants have more gut granules than *med-1; end-3* the HGS strain used in the study, which has *fstr-2* gene down regulated and from the *fstr-2* single mutant, suggesting the state of calorie restriction is more severe. These results suggest that the *fstr-1* is not completely functionally redundant to *fstr-2* since its' deletion has some synergistic input into the observed phenotypes.

Based on the results of the experiments shown in this work the model below is proposed, on what is the function (s) of the FSTR proteins in the intestine of *C. elegans* (Fig. 3.31)


Intestinal cell
ER
FSTR-1/2
GPI anchor attachment
FSTR-1/2
GPI anchor modification
Basal plasma membrane
Golgi
PQM-1
?
mRNA
Nucleus
Lipid Raft
FSTR-1/2
Absorption of nutrients
Normal metabolism and ATP production, energy utilization mode
Intestinal lumen
fstr-1/2 (-)
Apical junctions
Apical plasma membrane
Absorption of nutrients
Intestinal cell
Changed metabolism, low ATP production, energy storage mode

References

1. Liu, J.L., Yee, C., Wang, Y. & Hekimi, S. A single biochemical activity underlies the pleiotropy of the aging-related protein CLK-1. *Sci Rep* **7**, 859 (2017).

2. Cristina, D., Cary, M., Lunceford, A., Clarke, C. & Kenyon, C. A regulated response to impaired respiration slows behavioral rates and increases lifespan in Caenorhabditis elegans. *PLoS Genet* **5**, e1000450 (2009).

3. Kim, H. *et al.* A co-CRISPR strategy for efficient genome editing in Caenorhabditis elegans. *Genetics* **197**, 1069-80 (2014).

4. Friedland, A.E. *et al.* Heritable genome editing in C. elegans via a CRISPR-Cas9 system. *Nat Methods* **10**, 741-3 (2013).

5. Harris, T.W. *et al.* WormBase: a comprehensive resource for nematode research. *Nucleic Acids Res* **38**, D463-7 (2010).

6. Evans, T.C. Transformation and microinjection (April 6, 2006). in *WormBook* (ed. Ambros, V.) (WormBook, 2006).

7. Chughtai, A.A. *et al.* Perilipin-related protein regulates lipid metabolism in C. elegans. *PeerJ* **3**, e1213 (2015).

8. Kirienko, N.V., Cezairliyan, B.O., Ausubel, F.M. & Powell, J.R. Pseudomonas aeruginosa PA14 pathogenesis in Caenorhabditis elegans. *Methods Mol Biol* **1149**, 653-69 (2014).

9. El Mouridi, S. *et al.* Reliable CRISPR/Cas9 Genome Engineering in Caenorhabditis elegans Using a Single Efficient sgRNA and an Easily Recognizable Phenotype. *G3 (Bethesda)* **7**, 1429-1437 (2017).

10. Bindels, D.S. *et al.* mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nat Methods* **14**, 53-56 (2017).

11. Gibson, D.G. Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol* **498**, 349-61 (2011).

12. Bailey, T.L., Johnson, J., Grant, C.E. & Noble, W.S. The MEME Suite. *Nucleic Acids Res* **43**, W39-49 (2015).

13. McGhee, J.D. *et al.* ELT-2 is the predominant transcription factor controlling differentiation and function of the C. elegans intestine, from embryo to adult. *Dev Biol* **327**, 551-65 (2009).

14. Tepper, R.G. *et al.* PQM-1 complements DAF-16 as a key transcriptional regulator of DAF-2-mediated development and longevity. *Cell* **154**, 676-90 (2013).

15. Rikitake, M. *et al.* Analysis of GPI-anchored proteins involved in germline stem cell proliferation in the Caenorhabditis elegans germline stem cell niche. *J Biochem* **168**, 589-602 (2020).

16. An, J.H. & Blackwell, T.K. SKN-1 links C. elegans mesendodermal specification to a conserved oxidative stress response. *Genes Dev* **17**, 1882-93 (2003).

17. Agarwal, V., Bell, G.W., Nam, J.W. & Bartel, D.P. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* **4**(2015).

18. Lewis, B.P., Burge, C.B. & Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15-20 (2005).

19. Jan, C.H., Friedman, R.C., Ruby, J.G. & Bartel, D.P. Formation, regulation and evolution of Caenorhabditis elegans 3'UTRs. *Nature* **469**, 97-101 (2011).

20. Lu, S. *et al.* CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res* **48**, D265-D268 (2020).

21. Sonnhammer, E.L., Eddy, S.R. & Durbin, R. Pfam: a comprehensive database of protein domain families based on seed alignments. *Proteins* **28**, 405-20 (1997).

22. Mulder, N.J. *et al.* The InterPro Database, 2003 brings increased coverage and new features. *Nucleic Acids Res* **31**, 315-8 (2003).

23. Pierleoni, A., Martelli, P.L. & Casadio, R. PredGPI: a GPI-anchor predictor. *BMC Bioinformatics* **9**, 392 (2008).

24. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35**, 1547-1549 (2018).

25. Rao, W., Isaac, R.E. & Keen, J.N. An analysis of the Caenorhabditis elegans lipid raft proteome using geLC-MS/MS. *J Proteomics* **74**, 242-53 (2011).

26. Galian, C., Bjorkholm, P., Bulleid, N. & von Heijne, G. Efficient glycosylphosphatidylinositol (GPI) modification of membrane proteins requires a C-terminal anchoring signal of marginal hydrophobicity. *J Biol Chem* **287**, 16399-409 (2012).

27. Ludewig, A.H., Klapper, M. & Doring, F. Identifying evolutionarily conserved genes in the dietary restriction response using bioinformatics and subsequent testing in Caenorhabditis elegans. *Genes Nutr* **9**, 363 (2014).

28. Pruitt, K.D., Tatusova, T., Brown, G.R. & Maglott, D.R. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res* **40**, D130-5 (2012).

29. Loughner, C.L. *et al.* Organization, evolution and functions of the human and mouse Ly6/uPAR family genes. *Hum Genomics* **10**, 10 (2016).

30. Franssen, R. *et al.* Chylomicronemia with low postheparin lipoprotein lipase levels in the setting of GPIHBP1 defects. *Circ Cardiovasc Genet* **3**, 169-78 (2010).

31. Weinstein, M.M. *et al.* Reciprocal metabolic perturbations in the adipose tissue and liver of GPIHBP1-deficient mice. *Arterioscler Thromb Vasc Biol* **32**, 230-5 (2012).

32. Beigneux, A.P. *et al.* Highly conserved cysteines within the Ly6 domain of GPIHBP1 are crucial for the binding of lipoprotein lipase. *J Biol Chem* **284**, 30240-7 (2009).

33. Roh, H.C., Collier, S., Guthrie, J., Robertson, J.D. & Kornfeld, K. Lysosome-related organelles in intestinal cells are a zinc storage site in C. elegans. *Cell Metab* **15**, 88-99 (2012).

34. Coburn, C. & Gems, D. The mysterious case of the C. elegans gut granule: death fluorescence, anthranilic acid and the kynurenine pathway. *Front Genet* **4**, 151 (2013).

35. Mahajan-Miklos, S., Tan, M.W., Rahme, L.G. & Ausubel, F.M. Molecular mechanisms of bacterial virulence elucidated using a Pseudomonas aeruginosa-Caenorhabditis elegans pathogenesis model. *Cell* **96**, 47-56 (1999).

36. Julien-Gau, I., Schmidt, M. & Kurz, C.L. f57f4.4p::gfp as a fluorescent reporter for analysis of the C. elegans response to bacterial infection. *Dev Comp Immunol* **42**, 132-7 (2014).

37. Hoinville, M.E. & Wollenberg, A.C. Changes in Caenorhabditis elegans gene expression following exposure to Photorhabdus luminescens strain TT01. *Dev Comp Immunol* **82**, 165-176 (2018).

38. Price-Whelan, A., Dietrich, L.E. & Newman, D.K. Rethinking 'secondary' metabolism: physiological roles for phenazine antibiotics. *Nat Chem Biol* **2**, 71-8 (2006).

39. Hong, Y. *et al.* Requirement of N-glycan on GPI-anchored proteins for efficient binding of aerolysin but not Clostridium septicum alpha-toxin. *EMBO J* **21**, 5047-56 (2002).

40. Maduzia, L.L., Yu, E. & Zhang, Y. Caenorhabditis elegans galectins LEC-6 and LEC-10 interact with similar glycoconjugates in the intestine. *J Biol Chem* **286**, 4371-81 (2011).

41. Takeuchi, T., Nemoto-Sasaki, Y., Arata, Y. & Kasai, K. Galectin LEC-6 interacts with glycoprotein F57F4.4 to cooperatively regulate the growth of Caenorhabditis elegans. *Biol Pharm Bull* **34**, 1139-42 (2011).

Chapter IV Conclusions

The experimental work done for this thesis led to the following conclusions:

- Early in the development of the *C. elegans* intestine, there is a period of time in which proper activation of the GRN for intestinal development leads to establishing a gene expression profile in the intestinal cells that favors utilization of energy to maximize the reproduction capacity.
- When the activation of the GRN for intestinal development becomes stochastic, due to mutations in of transcription factors expressed transiently, early in development, a change in the gene expression profile of the intestinal cells causes a pleiotropic phenotype in the surviving animals, named Hypomorphic Gut Specification (HGS).
- The features of the HGS phenotype are increased intestinal lumen size, increased lipid droplet size, increased amount of gut granules, change in expression levels of many genes. The *nhr-156*, *fstr-1* and *fstr-2*, which are down-regulated in HGS strains were additionally investigated.
- *Nhr-156* mutants have some of the features of the HGS phenotype, including increased lumen size, increased lipid droplet size, shortened lifespan, and sensitivity to infection from *Pseudomonas aeruginosa* PA14. Addition of the wild-type *nhr-156* gene as extra chromosomal array rescued the lipid droplet size increase in *nhr-156*, partially rescued it into *med-1; end-3* strain, and rescued the lipid droplet size increase in *med-1;end-3 daf-16*. The function (s) of NHR-156

remain unknown but the evidence presented here suggests that NHR-156 play role in control of lipolysis, possibly by affecting the gene expression of enzymes involved in methionine metabolism, which lead to reduced levels of s-adenosylmethionine (Fig.4.1). Walker et al. 2015 showed that low levels of s-adenosylmethionine lead to accumulation of lipids[1], which suggests that a similar mechanism could contribute to the increase in fat stores in HGS strains.

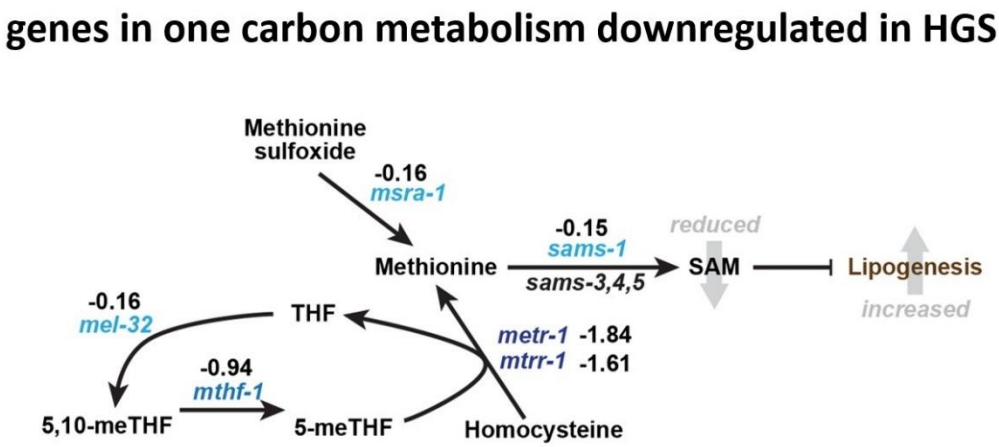


Fig. 4.1 Putative model how down regulation of genes involved in methionine metabolism affect lipogenesis.

- *Fstr-1* and *fstr-2* are two paralogous genes encoding proteins that are 96% identical, identified first for the ability of *fstr-1,2(RNAi)* to alleviate some aspects of the *clk-1* mutant phenotype. Cristina et al. 2009 reported that when RNAi was done on *fstr* genes, some of the slowed rhythmic functions of *clk-1* mutant animals were restored to normal levels[2] (Fig. 4.2). *Fstr-1* is expressed in HGS

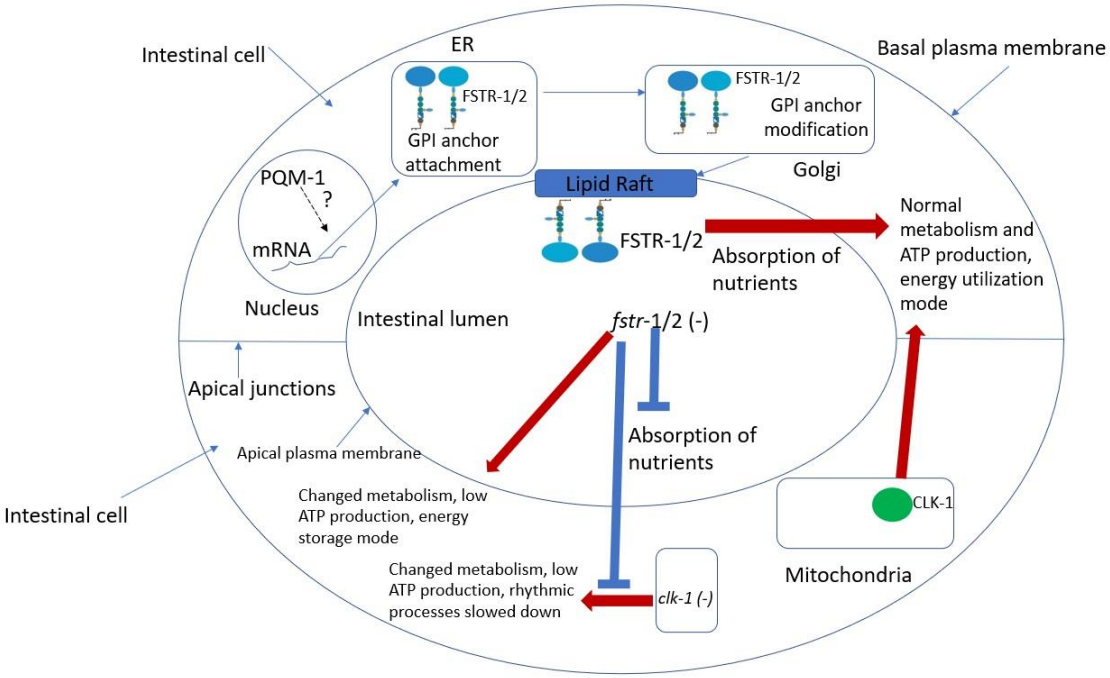
animals and *clk-1* animals and down regulated in wild type under normal conditions and *fstr-2* has reciprocal expression to *fstr-1*.

- *Fstr-1* gene appears to be a stress response gene. *Fstr-2* is the gene that functions in normal genetic background.
- *Fstr-1* mutants show no visible phenotype in wild type genetic background and normal conditions. *Fstr-2* mutant and *fstr-1,2* double mutant in wild type genetic background and normal conditions show a phenotype associated with calorie restriction and similar to the phenotype of the HGS strains even when food is abundant. Also, *fstr-2* mutant and *fstr-1,2* double mutant have reduced brood count, resistance to *Pseudomonas aeruginosa* PA14 infection, both toxin based and colonization mechanisms, *fstr-2* mutants have increased lifespan, whether in *fstr-1,2*, no significant difference was observed.
- In normal genetic background and conditions *fstr-2* is expressed in the intestinal cells. In HGS strain the expression of *fstr-2* is low and in stress conditions, the expression diminishes, and expression was observed in other cell types.
- In normal genetic background and conditions *fstr-1* has low expression in the intestinal cells. In stress conditions and in HGS strain expression *fstr-1* is increased.
- FSTR-2 protein is predicted to be GPI anchored. Translational fusions of FSTR-1 and FSTR-2 injected as extrachromosomal arrays show diffused cytoplasmic localization of the protein in the intestinal cells and in the case of FSTR-2 did not rescue the phenotype of *fstr-1,2*.

- Single copy integration of the scarlet reporter gene in the N-terminus of the FSTR-2 protein and after the first ET domain causes disruption of gene function and possible mis-localization of the protein. Single copy integration of the scarlet reporter after the coding sequence of the *fstr-2* gene does not seem to disrupt the gene function, however the localization of the protein is unexpected, likely because of the attaching of a GPI anchor to the FSTR-2 which removed the fluorescent reporter from the protein, and it diffused into the nuclear envelope.
- Placing a fluorescent reporter after the last ET domain of the FSTR-2 shows a different expression pattern than the others, localization of the protein appears to be in the plasma membrane of the intestinal cells.

The function (s) of FSTR proteins remain unknown, my hypothesis is, based on the experimental results and the analysis of the protein structure is that FSTR proteins are GPI anchored proteins expressed in the plasma membrane of different tissue types. In the intestine their function is probably as a structure protein expressed in the plasma membrane which serves as a scaffold to other types of proteins, normally residing in the ECM of the intestinal cells. Another possible function is involvement of absorption of nutrients and other small molecules such as toxins.

I have proposed a model how FSTR proteins may function and what influence their depletion or deletion may have on the *clk-1* mutant phenotype (Fig. 4.2).



The model proposes that FSTR proteins play role in the absorption of nutrients from the intestinal lumen in the intestinal cells.

In a normal animal where FSTR proteins are present (in wild-type FSTR-2), as the animal reaches adulthood an energy utilization mode is established. For all the processes that cells do a certain level of ATP production is required, so the gene expression profile is set to provide sufficient amount of ATP.

In *fstr-1,2* the absorption of nutrients is compromised, therefore there are not enough resources coming into the intestinal cells to maintain sufficient ATP production for the energy utilization mode so the animals change gene expression and go to energy storage mode – the *fstr* phenotype, (the HGS strains probably do a similar change in gene expression although for different reasons, Jose Soto's data shows the *fstr-1,2* mutants

have decreased ATP production), which is what we observe in HGS strains and in *fstr-1,2*. In energy storage mode the overall ATP requirements of the cells would be lower since we observe increased lipid droplet size, meaning that the process of lipolysis is slowed down, lower brood count, production of eggs requires a lot of energy. Rikitake et al. 2020 show that when *fstr* genes are deleted in the germ line cells, their proliferation is reduced[3], other metabolic processes are probably affected too, another steady state is set.

In *clk-1* mutants the ATP production levels are low, not because of not enough resources absorbed from food but because the electron transport chains are compromised. To compensate, the animals change gene expression in a way that does not go into what we think is an energy storage mode but into another state in which all rhythmic processes are slowed down (the *clk-1* mutants differentially expressed genes are different than the HGS ones). When we RNAi or delete the *fstr* genes and the intestinal cells are not absorbing enough nutrients this is sensed and leads to engaging the energy storage mode, which has a lower overall ATP cost. This, in the case of a *clk-1* mutant which already has lower ATP production levels, I think, makes the cells in the *clk-1* animals better adjusted to the low ATP levels, and what we see on a macroscopic level is alleviation of some aspects of the *clk-1* phenotype, namely restoration to normal or close to normal levels of some of the rhythmic processes, such as pharyngeal pumping and trashing (Reported by Cristina et al. 2009, they also report that levels of some genes which were differentially expressed in *clk-1* mutants were back to close to normal levels), shortening the lifespan to normal[2]. Not all the features on the *clk-1* phenotype were influenced by deleting the *fstr* genes. For example the brood count of *clk-1* and *clk-1; fstr-1,2* is almost the same.

Future work

Some of the experiments that can further test the proposed model for functioning of the FSTR proteins could be:

- RNA seq, from whole animals or dissected intestines on N2 and *fstr-1,2* mutants to looks for changes at gene expression profiles and compare to the HGS strains and *clk-1* mutants.
- Immunostaining of the FSTR-2 protein combined with fluorescent or confocal microscopy to localize more precisely the expression pattern. A construct containing Flag tag epitope sequence at the N-terminus of the FSTR-2 is available.
- Metabolomics of *fstr-1,2* mutants, if there is a compromised absorption of nutrients, compared to wild type a change in the levels of some metabolites could be observed.

References

1. Ding, W. *et al.* s-Adenosylmethionine Levels Govern Innate Immunity through Distinct Methylation-Dependent Pathways. *Cell Metab* **22**, 633-45 (2015).

2. Cristina, D., Cary, M., Lunceford, A., Clarke, C. & Kenyon, C. A regulated response to impaired respiration slows behavioral rates and increases lifespan in Caenorhabditis elegans. *PLoS Genet* **5**, e1000450 (2009).

3. Rikitake, M. *et al.* Analysis of GPI-anchored proteins involved in germline stem cell proliferation in the Caenorhabditis elegans germline stem cell niche. *J Biochem* **168**, 589-602 (2020).

Appendix A

Strain name	Strain genotype
MS2430	N2; irex731 PID14 fstr-2:wrmSCR:H2B:fstr-2 3'UTR + rol-6 D line 1
MS2431	N2; irex732 PID14 fstr-2:wrmSCR:H2B:fstr-2 3'UTR + rol-6 D line 2
MS2452	N2; irex745 PID21 opt2:plin-1:wrmSCR:unc-54 3'UTR + rol-6 D line 1
MS2459	N2; ir 70 (fstr-1,2) fstr-1,2 CC1,2
MS2461	N2;irex 750 PID27 + rol-6 D line 1
MS2462	N2;irex 751 PID27 + rol-6 D line 2
MS2463	ir 70; irex 752 PID21 +rol-6 D in fstr-1,2 (-)
MS2464	ir 70; irex 753 PID21 +rol-6 D in fstr-1,2 (-)
MS2465	ir70 (N2; fstr-1,2) erm-1:GFP in fstr-1,2(-)
MS2466	ir70 erm-1:GFp in fstr-1,2(-)
MS2469	ir70 clk-1(-) (qm30); fstr-1,2(-);clk-1(-)
MS2474	fstr-1,2(-) (ir70); daf-16(-) (mu86) irex752
MS2477	N2; ir73 (fstr-1) fstr-1, CC2,4
MS2478	N2; ir74 (fstr-2) fstr-1, CC1,3
MS2479	ir 74; qm30 (fstr-2(-);clk-1(-))
MS2481	fstr-1(-) (ir73); irex752 PID21 + rol-6 D
MS2482	fstr-2(-) (ir74); irex753 PID21 + rol-6 D
MS2500	unc-119; irex 761 PID31 opt-2:FSTR-1:wrmSCR:fstr-1 3'UTR + unc-119 rescue
MS2501	unc-119; irex 762 PID32 opt-2:FSTR-2:wrmSCR:fstr-2 3'UTR + unc-119 rescue
MS2502	unc-119; irex 763 PID33 opt-2:FSTR-1:wrmSCR:unc-54 3'UTR + unc-119 rescue
MS2503	unc-119; irex 764 same like irex 763 line 2
MS2504	ir74 fstr-2(-); irex 765 PID28 fstr-2: FSTR-2: fstr-2 3 UTR + rol-6 D line 1
MS2505	ir74 fstr-2(-); irex 766 PID28 fstr-2: FSTR-2: fstr-2 3 UTR + rol-6 D line 2
MS2506	ir74 fstr-2(-); irex 767 PID28 fstr-2: FSTR-2: fstr-2 3 UTR + rol-6 D line 3
MS2509	ir75; CC1 fstr-2 wrmScarlet 5' end (chromosome)
MS2510	ir76; CC3 fstr-2 wrmScarlet 3' end (chromosome)
MS2511	unc-119;irex 768 PID 14 fstr-2:wrmSCR:H2B:fstr-2 3'UTR +Myo 3 ch line 1
MS2512	unc-119;irex 769 PID 15 fstr-1:wrmSCR:H2B:fstr-1 3'UTR +Myo 3 ch line 1
MS2513	med-1(-) (de804); end-3(-) (ok1448); irex770 PID15+rol-6 D
MS2514	N2; irex 770 PID15+rol-6 D
MS2515	N2; irex771 PID37 opt2:plin-1:GFP:unc-54 3'UTR line 1
MS2516	N2; irex772 PID37 opt2:plin-1:GFP:unc-54 3'UTR line 2
MS2518	fstr-2 (-) (ir74) ir15129 acdh-1:GFP in fstr-2 (-)
MS2536	fstr-2 (ir78) CC1 fstr-2 Ex1Gly6wrmSCRGly6Ex1 on 5' end; backcrossed into N2
MS2537	fstr-2 (ir76) CC3 fstr-2 wrmScarlet on 3' end; backcrossed into N2
MS2544	N2; irex775 PID52 fstr-2:FSTR-2:wrmSCRafter last ET:fstr-2 3'UTR+rol-6 D line 1
MS2545	N2; irex776 PID52 fstr-2:FSTR-2:wrmSCRafter last ET:fstr-2 3'UTR+rol-6 D line 2
MS2550	fstr-1,2(-) (ir70);irex777 PID56 fstr-2: FSTR-2:fstr-2 3'UTR rescue +rol-6 D
MS2553	fstr-2; irex778 PID52 + PID37 +rol-6 D line 1
MS2555	fstr-2; irex779 PID52 + PID37 +rol-6 D line 2
MS2556	fstr-1,2 irex780 PID 56 + PID37 +rol-6 D line 1

Table.1 List of *C. elegans* strains used in this study.

cdc-25.1 (-)	end-3 (-)	nhr-156 (-)	N26h starved	N2 16h starved	med-1;end-3 (-)	N2	fstr-1,2 (-)
15.9562	13.832	13.9308	17.9816	17.5864	14.9682	10.97386	20.3034
18.278	19.3648	16.1044	18.278	19.4636	15.0176	13.0416	16.3514
15.6598	16.1538	15.8574	15.5116	19.4142	16.0056	15.39633	16.549
18.278	17.784	18.3768	19.1178	17.3394	12.8934	12.4488	17.2406
19.4636	17.9816	17.8828	21.0938	19.4142	16.2032	12.92633	18.7226
16.9442	15.561	16.1538	20.501	20.5998	12.7946	13.50267	14.7706
15.8574	14.4248	17.043	17.043	19.8588	18.1298	12.1524	18.0804
16.2032	18.6238	12.9922	19.76	17.9322	11.5102	13.585	16.302
14.7706	15.561	12.9922	21.3902	17.0924	17.5864	14.9682	17.5864
14.7212	17.043	15.1658	17.6852	16.0056	17.8828	14.82	16.3514
14.326	16.796	14.9188	19.6612	17.6852	15.067	12.3994	18.031
16.9936	12.844	16.549	19.9576	18.031	20.4516	13.091	20.1058
16.9936	19.8588	16.6972	18.4262	19.76	17.3888	12.9428	19.4636
20.1058	17.043	14.3754	14.1778	17.8828	20.9456	13.7826	21.489
17.5864	18.772	14.6224	20.6492	21.1432	19.3648	14.2766	20.8962
14.5236	15.2152	17.3394	19.9576	19.5624	16.8948	14.5236	17.784
16.9442	16.302	16.4008	16.549	20.1058	15.1164	13.1898	15.8574
16.1044	16.4996	17.1418	17.784	18.8708	14.7706	14.7212	14.7706
16.302	14.573	15.0176	19.1178	18.6732	17.9816	13.0416	18.525
15.4622	13.7826	14.1284	17.784	16.6478	14.1284	16.4008	15.314
14.7706	16.6478	16.4996	17.8828	21.0938	17.8334	15.4622	16.055
13.9802	17.784	16.4996	18.0804	18.031	21.6372	14.6224	16.2032
19.3648	14.6224	17.6358	15.6104	19.3648	17.3888	12.0042	19.3154
14.1778	14.7706	17.3394	16.9936	20.0564	15.4622	15.561	21.5878
18.031	16.4996	13.5356		19.019	15.4622	16.4502	16.9936
16.4008	14.1284	15.9562		17.1912	20.995	14.6224	20.7974
16.4008	12.1524	16.3514		17.4382	16.4996	14.0296	19.3154
17.4382	16.1538	15.6104		15.5116	19.4636		20.995
14.7212	13.585	16.2032		20.2046	19.019		19.6118
19.2166	16.5984	13.9802		19.0684	17.1418		19.266
16.8454	12.0536	16.2526		18.6732	14.573		19.9576
16.3514	13.5356	17.8334		20.8962	17.1418		21.0938
17.8334	15.314	16.1538		20.4516	16.8454		19.6118
16.4008	15.9068	13.4862		21.242	17.3394		22.8722
17.2406	17.537	14.1778		16.796	14.1284		
16.055	15.9562	18.5744		15.2646	17.5864		
17.3888	15.7092	16.3514		17.5864	15.4128		
15.4622	17.29	14.2272		18.3274	14.3754		
15.2152	14.7706	16.055		20.8962	15.1164		
17.2406	16.9442	14.4742		16.5984	14.8694		
17.3394	13.091	17.537		16.0056	15.7092		
18.5744	15.2152	16.7466		16.6972	13.3874		
13.7826	16.6478	17.3888		16.8948	16.8454		
16.7466	14.9682	17.1418		15.7586	17.5864		
17.1418	15.9068	14.9682		18.9202	14.5236		
13.8814	13.338	19.0684			19.2166		
17.3394	14.7212	17.784			17.29		
16.4996	13.0416	14.0296			15.5116		
15.37575	16.1538	14.573			16.2032		
18.772	11.856	16.1044			17.9816		
16.9442	13.832	15.5116			19.8588		
16.302	16.5984						
	16.1044						
	14.2272						
	18.772						
	14.9682						
	15.808						

Table.2 Average lumen size of different strains of *C.elegans* mutants used in this study. Each number in the table represents average of 5 measurements in the different parts of the animal's intestine.

N2	med-1; end-3	fstr-1,2
1.14139	2.786509	3.99857
1.474835	3.004491	3.424171
2.389128	2.42751	3.930444
1.14139	2.573079	3.715475
0.955409	1.998743	3.140807
1.433114	1.979136	3.627753
1.069531	2.133654	3.140807
1.004058	2.42751	2.282148
0.930907	1.979136	3.281586
1.238492	1.577057	3.28906
1.526802	1.998743	2.645044
1.526802	2.147822	2.435232
2.073229	2.786509	2.719886
1.470912	2.645044	2.645044
1.243148	1.808294	2.062053
2.860174	2.133654	2.645044
3.00401	2.573079	2.779243
1.526802	2.074622	2.779243
1.741568	3.740657	2.779243
1.470912	2.216003	2.126195
1.715667	2.501354	2.008115
0.968919	3.424592	1.831311
1.238492	2.501354	2.700701
1.977676	2.282148	2.008115
1.328515	1.481674	2.501354
1.119672	1.933361	2.923109
1.433114	2.42751	2.858154
1.976215	1.78579	2.557314
1.57339	2.858154	2.008115
1.160216	2.359931	3.075284
2.216003	2.700701	2.42751
0.859935	2.359931	2.779243
1.998743	1.280914	8.358146
2.573079	1.637268	7.350829
1.211373	1.357551	5.148404
1.979136	1.481674	5.286811

1.78579	2.062053	2.996308
2.133654	2.766743	5.643582
1.777684	2.501354	4.647084
2.133654	2.396371	5.070384
1.637268	2.008115	6.808577
1.919867	2.282148	6.724477
1.715667	1.715667	6.423711
1.338264	1.78579	3.638884
1.577057	2.359931	6.070202
1.137588	2.074622	3.358155
1.137588	1.933361	2.923109
2.435232	1.576141	4.701467
0.786236	1.628423	3.833712
2.133654	2.490939	2.126195
1.357551	2.062053	2.989068
1.438144	2.216003	10.61615
1.280914	1.576141	9.757817
2.28278	1.855598	8.431781
1.338264	1.576141	7.788901
1.777684	2.209476	4.932346
1.610586	2.359931	4.001819
1.637268	1.777684	4.905037
2.147822	1.831311	3.99857
1.628423	2.133654	1.78579
1.637268	1.78579	1.576141
1.357551	1.86259	3.068702
1.429077	3.644436	2.858154
1.577057	3.402166	1.074919
1.481674	1.919867	0.991027
1.211373	3.57158	0.991027
0.786236	3.638884	0.934005
0.844684	2.924097	2.619806
0.704872	2.924097	2.573079
0.786236	1.689368	2.557314
1.074919	1.716509	2.28278
1.357551	1.855598	3.358155
1.137588	2.216003	2.062053
1.86259	2.989068	2.719886
1.19939	1.855598	4.57062

1.280914	2.435232	5.574055
2.282148	2.924097	9.169635
1.074919	1.78579	6.501263
1.637268	2.633004	3.78747
1.933361	3.195965	6.49148
0.844684	2.008115	3.644436
1.637268	2.786509	3.075284
1.211373	2.42751	3.075284
1.438144	2.645044	3.150449
0.934005	2.147822	2.786509
1.429077	1.998743	3.132057
1.004058	2.645044	2.435232
1.074919	2.779243	2.133654
1.280914	3.502984	2.574763
1.502966	2.779243	2.133654
1.211373	2.924097	2.351348
1.637268	2.989068	2.996308
1.136317	2.573079	2.126195
1.137588	2.359931	2.719355
1.136317	2.359931	3.358155
1.280914	2.074622	2.435232
0.905743	2.435232	2.924097
1.429077	3.075284	2.645044
1.502966	1.429077	2.42751
1.357551	1.979136	2.435232
2.216003	1.78579	2.074622
1.280914	2.282148	2.008115
1.137588	2.28278	2.359931
1.429077	3.7806	1.933361
1.715667	2.924097	2.074622
1.291022	2.216003	1.576141
1.357551	3.004491	1.637268
1.280914	2.501354	2.133654
1.136317	3.424171	1.576141
1.074919	2.719886	1.004058
1.19939	2.645044	1.211373
0.844684	2.28278	1.338264
1.291022	1.855598	1.628423
3.644436	2.645044	1.429077

1.715667	2.719886	1.855598
3.99857	1.429077	1.715667
1.137588	2.246427	1.338264
1.136317	1.933361	1.429077
1.502966	2.42751	1.074919
1.211373	2.074622	1.855598
1.502966	2.47465	1.933361
1.405639	2.557314	2.008115
2.074622	2.573079	1.637268
2.700701	3.075284	1.481674
0.991027	2.645044	1.78579
2.645044	2.989068	1.502966
1.576141	2.923109	1.357551
1.280914	3.923824	1.438144
1.577057	1.716509	1.577057
1.280914	2.282148	3.57158
1.405639	2.435232	4.070525
1.137588	2.282148	4.04811
0.991027	2.216003	3.195965
0.844684	2.351348	2.996308
1.074919	2.062053	3.715475
1.004058	1.998743	2.490939
1.280914	2.924097	2.924097
1.280914	3.075284	2.989068
1.137588	1.610586	2.074622
1.438144	1.960807	2.42751
1.78579	2.573079	2.719886
1.19939	1.78579	2.331609
1.357551	2.216003	2.42751
1.211373	2.28278	1.136317
1.136317	2.133654	1.577057
1.211373	2.786509	1.715667
1.280914	2.645044	1.78579
1.280914	2.216003	2.359931
1.429077	1.502966	1.78579
1.357551	4.503129463	1.357551
1.011224	2.786509378	1.78579
1.429077	3.715475417	2.719886
1.502966	3.49885829	1.211373

1.628423	2.923108674	1.502966
1.211373	3.218481518	1.689368
1.481674	2.427509625	4.83654
1.280914	2.924096712	4.716496
1.074919	2.923108674	4.499921
1.429077	1.502966337	4.070525
1.405639	2.216002967	3.78747
1.106697	1.481674044	3.068702
1.280914	1.576141131	3.195965
1.136317	1.429076849	2.859669
1.502966	1.637268344	2.719355
1.074919	1.777683895	2.574763
0.786236	2.216002967	2.719886
1.19939	2.989068287	2.923109
1.074919	3.075283899	2.331609
1.211373	2.996307504	2.786509
1.211373	2.216002967	2.923109
0.844684	2.5340525	2.359931
1.291022	1.357551195	3.860742
1.137588	1.462048356	4.217603
1.429077	2.645044277	4.74154
1.291022	2.85966929	4.212806
1.136317	2.91122607	3.860742
	3.218481518	3.706524
	2.074622153	3.864107
	1.715666885	2.574763
	2.14782224	1.280914
	2.427509625	1.357551
	2.573079374	1.78579
	2.474650136	1.291022
	1.576141131	2.996308
		1.481674

Table.3 Lipid droplets size of N2, *med-1*; *end-3* HGS in text, and *fstr-1,2* strains. Each number represents the diameter of a lipid droplet in μm.

Trial # 1	Median lifespan N2 (d)	Median lifespan *fstr-1* (-) (d)	Median lifespan *fstr-2* (-) (d)	Median lifespan *fstr-1,2* (-) (d)	Median lifespan *clk-1* (-) (d)	Median lifespan *clk-1;fstr-1,2* (-) (d)
1	12	12	15	13	12.5	11
2	15	15	17	19	18	15
3	20	19	22	20	25	26

Highlighted in red is the median lifespan of a strain for which both Log-rank (Mantel-Cox) test and Gehan - Breslow - Wilcoxon test show not significant p-value.
In yellow is highlighted the median lifespan for which both tests produced significant p-value.
In orange is highlighted the median lifespan for which one of the tests produced significant p-value.

Table.4 Median lifespan reported in days of adulthood for the strains used in the experiment. In yellow are highlighted the strains in which was observed statistically significant increase in the lifespan.

Toxin based (fast) killing of *C.elegans* from *Pseudomonas aeroginosa* PA14					
Trial#	Median survival N2 (h)	Median survival *fstr-1* (-) (h)	Median survival *med-1;end-3* (-) (h)	Median survival *fstr-2* (-) (h)	Median survival fstr-1,2 (-) (h)
1	8	4	8	72	96
2	4	4	4	24	96
3	4	8	4	96	96
Colonization based (slow) killing of *C.elegans* from *Pseudomonas aeroginosa* PA14					
Trial #	Median survival N2 (h)	Median survival *fstr-1* (-) (h)	Median survival *med-1;end-3* (-) (h)	Median survival *fstr-2* (-) (h)	Median survival fstr-1,2 (-) (h)
1	60	60	60	60 more sensitive	60
2	48	48	60	48	60
3	48	36	60	48	60

Highlighted in red is the median lifespan of a strain for which both Log-rank (Mantel-Cox) test and Gehan - Breslow - Wilcoxon test show not significant p-value.
In yellow is highlighted the median lifespan for which both tests produced significant p-value.
In orange is highlighted the median lifespan for which one of the tests produced significant p-value.

Table.5 Median survival reported in hours for the strains used in the experiment. In yellow are highlighted the strains in which was observed statistically significant increase in the lifespan.

Primer name	Primer sequence 5'-3'
F57F4.4GAF	CCTGCAGGTGCACTCTAGAGGATCCTCTGACGGTTCACTAGTAAGAACCT
F57F4.4GAB	TCCTTTACTCATTTTTTCTACCGGTGGATGTCTAGTCAGAGCGATATACACAG
F57F4.3GAF	CCTGCAGGTCGACTCTAGAGGATCCATTCAAATTTCAGCTGTATAAC
F57F4.3GAB	TCCTTTACTCATTTTTTCTACCGGTGGATGTCTAGTCAAAGCAATATATAC
fstr-1 F	ACTTGGCCAACTCGTGTGCTTCCATTGACG
FSTR-1 F2	AAAAGAGCTCATGCGCTTGGGCATTTTCCTGCTGGCACTTG
FSTR-1 R2	AAAAGAATTCTTTGGGGCATTAATAACTTCGATTGGCTTTGG
fstr-2opt-2PromF	GTCAGTATCGCCACTTCTAGAATGCGCTTGGGCATTTTCCTGCTGGCACTT
fstr-2opt-2PromB	TCCCTTGCTGACCATTCTAGAATGTCTAGTCAGAGCGATATACACAGCAGCGAAAAGTG
fstr-1opt-2PromF	GTCAGTATCGCCACTTCTAGAATGCGCTTGGGCATTTTCCTGCTGGCACTT
fstr-1opt-2Promb	TCCCTTGCTGACCATTCTAGAATGTCTAGTCAAAGCAATATATACAGCAGCG
fstr-1 -PID15 No scr R	ATTCTCTATAGCTGGCGAATTCCTAATGTCTAGTCAAAGCAATATATACAGCAG
fstr-2 (no scarlet) 1	TTATCTAAAGCTGGCGAATTCCTAATGTCTAGTCAGAGCGATATACACAGCAG
fstr-2 (no scar) 2	GAAATGAGAAAAATTATCTAAAGCTGGCGAATTCCTAATGTCTAGTCAGAGCGA
F57F4.3F	ATTCAAATTTCAGCTGTATAACCGTGGATT
F57F4.3B	ATGTCTAGTCAAAGCAATATATACAGCAGC
f57f4.4f	TCTGACGGTTCACTAGTAAGAACTTTAAAAG
f57f4.4B	ATGTCTAGTCAGAGCGATATACACAGCAGC
F57f4.3AND 4fHALF	CCATCTTACATCTGCAGAATGATGAACTCA
F57f4.3AND 4BHALF	TGAGTTCATCATTCTGCAGATGTAAGATGG
egg-5F	AAAATCTAGAGATGGTGGATTATCCTCAACAAAAGGGAGT
egg-5R	GATTCTGCAGAAGCTGCAATTTTGTAGGCTGCAT
F57F4.3GAF	GCCACTAACACATTCAAATTTCAGTCTAGAATGCGCTTGGGCATTTTCCTGCTGG
F57F4.3GAR	AACTGCCTCTCCCTTGCTGACCATTCTAGAATGTCTAGTC

	AAAGCAATATATACAGCAGC
F57F4.4GAF	TTCACTAACACATTCAAATTTCAGTCTAGAATGCGCTTGGGCATTTTCCTGCTGG
F57F4.4GAR	AACTGCCTCTCCCTTGCTGACCATTCTAGAATGTCTAGTCAGAGCGATATACACAGCAGC
PlinGA-F	CCTCGTCAGTATCGCCACTTCTAGAATGACTGACGTCGAGCAGCCAGTATCAGTT
PlinGA-B	CCTCTCCCTTGCTGACCATTCTAGAGGCACGATTCCTATACAAGTTTCTTGTGCG
h2b sac	GGACGAGCTCTACAAGATGCCACCAAAGCCATCTGCAAAG
h2b eco	GTTGGAATTCTTACTTGCTGGAAGTGTACTTGGTGAC
wrm scarlet cla-I F	AAAAATCGATATGGTCAGCAAGGGAGAGGCAGTTATCAAGGAG
wrm scarlet bgl-II bk	AAAAAGATCTCTTGTAGAGCTCGTCCATTCCTCCGGT
wrmscarF	AAAATCTAGAATGGTCAGCAAGGGAGAGGCAGTTATCAAG
wrmscarB	AAAAGAATTCTTACTTGTAGAGCTCGTCCATTCCTCCGGTG
F4.4 Forward	AAAACTGCAGGACAAAAGACACACACACACAAATACATAG
F4.4 Reverse	AAAATCTAGACTGAAATTTGAATGTGTTAGTGAATAGTAA
F4.3 Forward	AAAACTGCAGTGGAATTTTAAAATATCGATCAATGTTGTT
F4.3 Reverse	AAAATCTAGACTGAAATTTGAATGTGTTAGTGGCAAATTA
FSTR-2 F2	AAAAGAGCTCATGCGCTTGGGCATTTTCCTGCTGGCACTTG
FSTR-2 B2	AAAAGAATTCTTTGGGGCATTAATAACTTCGATTGGCTTTGG
FSTR-1 repF	AAAAGAATTCATGCGCTTGGGCATTTTCCTGCTGGCACTTG
FSTR-1 repbk	AAAAGAGCTCTTTGGGGCATTAATAACTTCGATTGGCTTTGG
FSTR-2 repF	AAAAGAATTCATGCGCTTGGGCATTTTCCTGCTGGCACTTG
FSTR-2 repbk	AAAAGAGCTCTTTGGGGCATTAATAACTTCGATTGGCTTTGG
fstr-2 sequence exon4	AAAAGAGCTCATGATTTTGTGCCGTTCTTCCATCTATCTC
fstr2 4-5	AAAAGAATTCCTAATGTCTAGTCAGAGCGATATACACAG

	CAGC
fstr-1 sequence exon4	AAAAGAGCTCATGATTTTGTGCCGCTCTTCCATCTATCTC
fstr-1 sequence 4-5	AAAAGAATTCCTAATGTCTAGTCAAAGCAATATATACAG CAGC
PID14 sac	ACTCCACCGGAGGAATGGACGAGCTCATGCGCTTGGGCA TTTTCCTG
pID14 Ecori	ATTATCTAAAGCTGGCGAATTCCTAATGTCTAGTCAGAG CGATATACACAG
PID15 sac	ACTCCACCGGAGGAATGGACGAGCTCATGCGCTTGGGCA TTTTCCTG
Pd15 Ecori	ATTCTCTATAGCTGGGCAATTCCTAATGTCTAGTCAAAGC AATATATACAGCAG
perilipinCDS_fwd	CGTCAGTATCGCCACTTATGACTGACGTCGAGCAGCC
perilipinCDS_rev 5	TTGGAGCAGTCATGGCACGATTCCTATACAAGTTTCTTG
GFP_fwd 5	AGGAATCGTGCCATGACTGCTCCAAAGAAGAAG
GFP_rev 5	CCGGCGCTCAGTTGGCTATTTGTATAGTTCATCCATGC
PID15 XbaIF	AACACATTCAAATTTCAGTCTAGAATGCGCTTGGGCATTT TCCTG
PID15 XbaIR	TCTCCCTTGCTGACCATTCTAGATTTGGGGCATTAATAAC TTCGATTGG
Exon1 (RT)F	ATGCGCTTGGGCATTTTCCTGCTGG
Intron2(RT)R	CACGTAACAGTTTGTGGTCAGGAATTGA
Exon4(RT)F	ATGATTTTGTGCCGTTCTTCCATCTATCTC
CC5 (fst-2 e1,2)	GGGCATTAATAACTTCGATGTTTTAGAGCTAGAAATAGC AAGT
Fstr-2Intron2F	ACCGGAGGAATGGACGAGCTCGTGAGTTTTGTTATCACT GTTATTTTTG
Fstr-2Intron2R	AGAAAAATTATCTAAAGCTGGCGAATTCCACGTAACAGT TTGTGGTC
Fstr-2exon1+2fixedR	TCTCCCTTGCTGACCATTCTAGAGTTTGGGGCATTAATAA CTTCGATTGG
Fstr-2 3UTRextF	TTAACTGATATTTGGCTAAACTAGTCTAAGTTTTTTTGTT GATCTGTTGTCAATATTACA
Fstr-2 3UTRextR	GAAAGGGCCCGTACGGCCGACTAGTGCCCACAGGCATGC GGAA
Fstr-2 3UTRextRT	CCACAGGGATGCGGAATTTGTTTTCAGATTT
fstr-2 piece I_fwd	CGGCCGCTCTAGAACTAGTGACAAAAGACACACACACAC
fstr-2 piece I_rev	TGCTGACCATGTTTTGTTTAGCACTAGCG
piece II (wrm-scar)_fwd	TAAACAAAACATGGTCAGCAAGGGAGAG
piece II (wrm-	CGACTAAGCACTTGTAGAGCTCGTCCATTC

scar)_rev	
fstr-2 piece III_fwd	GCTCTACAAGTGCTTAGTCGGAAATCTAATTGTATCG
fstr-2 piece III_rev	CGAGGTCGACGGTATCGATATCCAGCGAGATTGGTGCAG
Fstr-2 piece III_rev (in the 3UTR)	CGAGGTCGACGGTATCGATATTAGCCAAATATCAGTTAAAAACACTAC
WrmScr set1 G-linker F	GGAGGGGGTATGGTCAGCAAGGGAGAGGCAGT
WrmScr set1 G-linker R	CCCTCCGCCCTTGTAGAGCTCGTCCATTCCT
WrmScr set2 G-linker F	TAAACAAAACGGGGGTGGCGGAGGGGGTATGGTCAGCAAGGGAGAGGCAGT
WrmScr set2 G-linker R	CGACTAAGCACCCTCCGCCCCCGCCTCCCTTGTAGAGCTCGTCCATTCCT
Flag Piece I F	CGGCCGCTCTAGAACTAGTGGACAAAAGACACACACAC
Flag Piece I R	CTTTGTAGTCGTTTTGTTTAGCACTAGC
Set 1 Flag Piece II F	GACTACAAAGACGATGACGACAAGTGCTTAGTCGGAAATCTAATTGTATCG
Set 1 Flag Piece II R	TTCCAAGTCGACCTTGTTACACAACTCATC
Flag Piece II F	TAAACAAAACGACTACAAAGACGATGACG
Flag Piece II R	GAATATTTTCTTCCAAGTCGACCTTGTTAC
Flag Piece III F	CGACTTGGAAGAAAATATTCCAAAGCCAATC
Flag Piece III R	CGAGGTCGACGGTATCGATATTAGCCAAATATCAGTTAAAAAC
Piece I pfstr-2:fstr-2_fwd	CGGCCGCTCTAGAACTAGTGGACAAAAGACACACACAC
Piece I pfstr-2:fstr-2_rev	CGCCACCCCCTGGACCAACGCAGTTATTAG
Piece II SCR_fwd	CGTTGGTCCAGGGGGTGGCGGAGGGGGTATG
Piece II SCR_rev	GGCCATTGTTCCCTCCGCCCCCGCCTCC
Piece III fs-2:fs-2UTR_fwd	GGGCGGAGGGAACAATGGCCCAGGAATTGTAACC
Piece III fs-2:fs-2UTR_rev	CGAGGTCGACGGTATCGATAGCCCACAGGGATGCGGAA
pfstr-2:fstr-2_fwd	CGGCCGCTCTAGAACTAGTGGACAAAAGACACACACAC
pfstr-2:fstr-2_rev	TGCTGACCATTGGACCAACGCAGTTATTAG
SCRnolinkers_fwd	CGTTGGTCCAATGGTCAGCAAGGGAGAG
SCRnolinkers_rev	GGCCATTGTTCTTGTAGAGCTCGTCCATTC
fs-2ex5:fs-23'UTRext_fwd	GCTCTACAAGAACAATGGCCCAGGAATTGTAACC
fs-2ex5:fs-23'UTRext_rev	CGAGGTCGACGGTATCGATAGCCCACAGGGATGCGGAA

pfstr-1:fstr-1 lastET_fwd	CGGCCGCTCTAGAACTAGTGTGGAATTTTAAAATATCGATCAATGTTG
pfstr-1:fstr-1 lastET_rev	CTTTACTCATTGGACCAACGCAGTTATTAG
Piece II GFP_fwd	CGTTGGTCCAATGAGTAAAGGAGAAGAAC
Piece II GFP_rev	GCAGCACGTTTTTGTATAGTTCATCCATGC
fstr-1:fstr-1 3UTR_fwd	ACTATACAAAAACGTGCTGCCAAGCGCC
fstr-1:fstr-1 3UTR_rev	CGAGGTCGACGGTATCGATAGGACAAGAGGTTAGATTTAGTAATGTT
Sequencing primers	
pfstr-2CcheckFwd	AGACATGGGCTATCGAATGGCG
fstr-2 Int2CcheckRev	GCCGAGGAGCTAACTAGCCACGT
pFS-2 seq F	GGACCGATTGGCGACCAAGAACCA
scrs seq f	GCGTCTCTACCCAGAGGACGGA
scr seq r	TCCGTCCTCTGGGTAGAGACGC
fs-2 int2 seq r	TCGACGTGTGTTTTTAGTTTCTC

Table.6 List of PCR primers used in this study

List of plasmid constructs

PID001 – p*fstr-2*:FSTR-2:GFP (pPD 95.67)

PID002 – p*fstr-1*:FSTR-1:GFP (pPD 95.67)

PID3 – pEndu2:wrmSCR:H2B:unc-54UTR (PGB561)

PID11 – p*fstr-2*:wrmSCR:H2B:*unc-54*UTR (PID3)

PID12 – p*fstr-1*:wrmSCR:H2B:*unc-54*UTR (PID3)

PID14 – p*fstr-2*:wrmSCR:H2B:*fstr-2* 3'UTR

PID15 – p*fstr-1*:wrmSCR:H2B:*fstr-1* 3'UTR

PID21 – p*opt2*:*plin-1*:wrmSCR:*unc-54*UTR (PGB561) – perilipin marker

PID25 – p*fstr-1*:wrmSCR:*fstr-1*exon1-2:*fstr-1* 3'UTR (fstr-1) (PID15)

PID26 – p*fstr-2*:wrmSCR:*fstr-2*exon1-2:*fstr*-2 3'UTR (fstr-2) (PID14)

PID27 – p*fstr-1*:*fstr-1*:wrmSCR:*fstr*-1 3'UTR (PID25)

PID28 – p*fstr-2*:*fstr-2*:wrmSCR:*fstr-2* 3'UTR (PID26)

PID29 – p*opt2*:wrmSCR:*fstr-1* 3'UTR (PGB561)

PID30 – p*opt2*:wrmSCR:*fstr-2* 3'UTR (PGB561)

PID31 – p*opt2*:*fstr-1*:wrmSCR:*fstr-1* 3'UTR (PID29)

PID32 – p*opt2*:*fstr-2*:wrmSCR:*fstr-2* 3'UTR (PID30)

PID33 – p*opt2*:*fstr-1*:wrmSCR:*unc-54* 3'UTR (PGB561)

PID34 – p*opt2*:*fstr-2*:wrmSCR:*unc-54* 3'UTR (PGB561)

PID35 – p*fstr-1*:*fstr-1*:*fstr-1* 3'UTR (PID15)

PID36 – p*fstr-2*:*fstr-2*:*fstr-2* 3'UTR (PID14)

PID37 – p*opt2*:*plin-1*:GFP:*unc-54*UTR (PID21) – green perilipin marker

PID38 – p*fstr-2*:*fstr-2*ex1+2:wrmSCR:H2B:*fstr-2* 3'UTR (PID14)

PID39 – p*fstr-1*:*fstr-1*ex1+2:wrmSCR:H2B:*fstr-1* 3'UTR (PID15)

PID40 – p*fstr-2*:*fstr-2*ex1+2:wrmSCR:*fstr-2*ex4+5:*fstr-2* 3'UTR (PID38)

PID41 – p*fstr-1*:*fstr-1*ex1+2:wrmSCR:*fstr-1*ex4+5:*fstr-1* 3'UTR (PID39)

PID42 – p*fstr-2*:*fstr-2*ex1+2:wrmSCR:*fstr-2*intron2:*fstr-2* 3'UTR (PID38)

PID44 – p*fstr-2*:*fstr-2*ex1:wrmSCR:*fstr-2*ex-1:*fstr-2* int-1: (the whole gene): *fstr-2* 3'UTR

PID45 – p*fstr-1*:*fstr-1(*lastETdomain):GFP: *fstr-1*ex-4:*fstr-1*int4:*fstr-1*ex-5:*fstr-1* 3'UTR

PID46 – p*fstr-2*:*fstr-2*ex-1:wrmSCR:*fstr-2*ex-1:*fstr-2*int-1:*fstr-2*ex2:*fstr-2*int-2:*fstr-2*ex-3(223bp)

PID48 - p*fstr-2*:*fstr-2*ex1:G6wrmSCRG6:*fstr-2*ex-1:*fstr-2* int-1: (the whole gene): *fstr-2* 3'UTR

PID50 - p*fstr-2*:*fstr-2*ex-1:wrmG6SCRG6:*fstr-2*ex-1:*fstr-2*int-1:*fstr-2*ex2:*fstr-2*int-2:*fstr-2*ex-3(223bp)

PID52 – p*fstr-2*:*fstr-2*(lastETdomain):wrmSCR:*fstr-2*ex-4:*fstr-2*int4:*fstr-2*ex-5:*fstr-2* 3'UTR

PID54 – p*fstr-2*:*fstr-2*(lastETdomain):G6wrmSCRG6:*fstr-2*ex-4:fstr-2int4:*fstr-2*ex-5:*fstr-2* 3'UTR

PID56 – p*fstr-2*:*fstr-2*:*fstr-2* 3'UTR (GA in Bluescript KS-) three different isolates (_1,_2,_3)

PID58 – p*fstr-2*:*fstr-2*ex-1:flagtag:*fstr-2*ex-1:*fstr-2* the rest of the gene:*fstr-2* 3'UTR

PID59 – p*fstr-1*:*fstr-2*(lastETdomain):wrmSCR:*fstr-2*ex-4:*fstr-2*int4:fstr-2ex-5:*fstr-1* 3'UTR

PID60 – p*fstr-1*:*fstr-2*:*fstr-1* 3'UTR (PID15)

Additional methods

PCR

PCR primers were diluted first in TE buffer. The volume of TE buffer used, was calculated based on the nmol concentration of the primer and 4 times of this concentration in microliters of TE buffer were added to the vial containing the lyophilized primer, to produce the stock primer solution. Ten microliters of the stock were diluted ten times in ddH_2O to produce the primer solution used for PCR, the concentration of the primers is approximately 25pmol/µl.

For fragments of 1000bp or less a 2X Taq RED Master Mix (Cat #: 42-138; Apex Bioresearch Products) was used. The total volume of reactions was 20µl. Primers, diluted in the described above way were added in volume of 1µl, and the template DNA volume

was 1µl. The standard PCR cycle used for the most amplifications with this master mix was:

1	95°C 30sec
2	95°C 30sec
3	55°C 30 sec
4	72°C 30 sec
5	72°C 5 min
6	4°C ∞

Steps 2-4 were repeated 30 times. Extension time (step 4) was varied depending of the length of the fragment following the rule described by the manufacturer that at optimal annealing temperature, the rate of polymerization is approximately 1000bp/min.

For fragments with length more than 2kb and in the cases when a repair template for CRISPR integration was generated or fragments were required to be with high quality of amplification for molecular cloning, a Q5 Hot Start High-Fidelity 2X Master Mix (Cat #: M0494S; New England Biolabs inc.) was used.

The total amount of the PCR reactions was 25µl with primers and template in the same volumes described for the previous master mix. The cycle used was

1	98°C 30sec
2	98°C 10sec
3	55°C 30 sec
4	72°C 30 sec
5	72°C 5 min
6	4°C ∞

In step 4 the extension time was varied from 30 sec to 5 min depending on the size of the amplifying fragment. All PCR products were normally resolved onto 0.7% agarose gel, in some cases for better resolution a 1% or 1.2% gel was used. When the PCR products were required for cloning or for sequencing, the bands with the PCR product were cut and the PCR product was gel purified using a GeneJET Gel Extraction Kit (K0692 Thermo Fisher), following the manufacturer's protocol.

Colony PCR

For colony PCR a material from bacterial colony was picked with a pipette tip and placed into 10µl of ddH_2O. The tip was used to streak a LB-agar plate with carbenicillin or kanamycin depending on the vector that was used for transformation. The samples were placed in a thermocycler at 90°C for 15 min. After that a 2µl of the mixture was used as a template for a standard PCR reaction.

Genotyping

For genotyping a few animals from each genotyped strain were place in 10µl of lysis buffer, containing 10µl 10X Taq buffer with $(NH_4)_2SO_4$ and $MgCl_2$, 10µl 25mM $MgCl_2$, 0.5µl Proteinase K (>600 U/ml) and 80µl H_2O. Samples were put into thermocycler at 50°C for 10 minutes or until the animals were dissolved. After that the Proteinase K was inactivated at 70°C for 10 min. and 2µl of that prepared sample was used as a template in a standard PCR reaction.

Restriction digest

Restriction digest with one or two restriction endonucleases was used for molecular cloning or for diagnostic of a certain plasmid constructs. In most cases a 20µl reaction was used. The appropriate 10X reaction buffer was selected using the Promega USA online restriction digest tool (https://www.promega.com/resources/tools/retool/). The amount of restriction enzymes used in each reaction was 1µl, depending on the quality of the enzyme available sometimes a 1:10 dilution was made. Depending on the case, 1µl (usually a precipitated purified plasmid for cloning) or 2µl (usually plasmid miniprep purified plasmids for diagnostic purposes) were digested at 37°C for at least 2h or in many cases overnight.

Molecular cloning

Ligation reaction

A 10 µl ligation reaction was used to ligate a certain fragment, digested with the appropriate combination of restriction enzymes in a vector digested the same way. The reaction contained 1µl 10X ligase buffer, 1µl digested vector, 1µl digested insert and 1µl T4 DNA ligase. The reaction mix was brought to 10µl with ddH_2O. The control for the experiment was the same mixture but without the insert. Ligation reactions were incubated 1h at room temperature.

Gibson Assembly

For Gibson Assembly of 1-4 fragments, we used the protocol described by Mouridi et al. 2017 We've used the reagents described in the paper, briefly, 1.2ml Gibson Assembly master mix contains 320µl 5X ISO buffer, 0.64 µl T5 Exonuclease (10 U/µl), 20µl Phusion DNA polymerase (2 U/µl), 160µl Taq Ligase (40 U/µl), 700 µl H_2O.

The 5X ISO buffer contains in 2 ml, 1ml 1M Tris-HCl pH=7.5, 50µl of 2M $MgCl_2$, 20µl of dCTP, dGTP, dATP, dTTP 100mM each, 100µl of 1M DTT, 200µl 50mM NAD^+, and 0.5g PEG800. Add H_2O up to 2ml.

For Gibson Assembly 15µl of reagents were used, 2 µl of appropriately digested vector and 1µl of each fragment. If the fragments were less than 3, the amount of gel purified fragment was increased accordingly (i.e in a case of one fragment only a 3 µl of the fragment were added) to total volume of 20 µl. In the case of 4-piece assembly the vector

amount was changed to 1µl. The controls for the experiments were the same reaction without the inserts, and when a newly digested vector was used an empty control was made with only the digested vector and H_2O. The reactions were incubated at 50°C for 1h

Bacterial transformation

For transformation an *E. coli* XL Blue 2, strain of chemically competent cells was used. If the transformation was to produce a construct that would be used for RNAi the HT115 strain of *E. coli* was used. If a normal ligation was done a 100µl of cells were transformed with 2µl ligation mixture. If a Gibson Assembly was done a 100µl of cells were transformed with 5µl reaction mixture. In all cases the transformation was done on ice for 20 min, followed by a heat shock at 42°C for 90sec, and then transferring back on ice for 2min. If a Gibson Assembly was done or if a vector with, kanamycin resistance was used to the cells was added 1ml of SOC media (LB-broth also works) and the cells were outgrown for 1h at 37°C at 300rpm. After outgrowing, cells were centrifuged for 1min at 14000rpm and the 1ml of media removed. Cells were plated either on plates with carbenicillin or kanamycin and incubated overnight at 37°C.

Isolation of plasmid constructs and precipitation of DNA

Isolation of the plasmid constructs generated by ligation or Gibson Assembly was carried by using a GeneJET plasmid miniprep kit (Thermo Fisher), following the manufacturer's protocol. After that the plasmids were checked by PCR and restriction digest for appropriate size and presence of specific sequences. After that samples were pooled, and the plasmid DNA precipitated. The precipitation was done by adding 10µl of 3M sodium

acetate for each 100µl of plasmid minipreps pooled, then adding 1ml of 95% ethanol. The samples were centrifuged then at 14000rpm for 10 min. Tubes were inspected to see the precipitated DNA (a white pellet), the ethanol was removed and then another 1ml of 70% ethanol was added and the samples centrifuged again for 1 min at 14000rpm. After centrifugation, all the supernatant was removed carefully, and the samples were vacuum dried for 7 min. The dried DNA pellets were resuspended in same elution buffer, which was used in the plasmid isolation.