Meta Gene 2 (2014) 819–830

Contents lists available at ScienceDirect

Meta Gene

let-65 is cytoplasmic methionyl tRNA synthetase in *C. elegans*

Maha Z. Alriyami [a,c,*], Martin R. Jones [b], Robert C. Johnsen [a], Yajnavalka Banerjee [c], David L. Baillie [a]

[a] Department of Molecular Biology and Biochemistry, Simon Fraser University Burnaby, 8888 University Drive, Burnaby, BC V5A 1S6 Canada
[b] Genome Sciences Centre, British Columbia Cancer Research Centre, 600 West 10th Avenue, Vancouver, BC V5Z 4E6, Canada
[c] Department of Biochemistry, College of Medicine and Health Sciences, Sultan Qaboos University, 35, Al-Khod 123, Oman

ARTICLE INFO

Article history:
Received 13 September 2013
8 August 2014
Accepted 11 August 2014
Available online 9 November 2014

Keywords:
Caenorhabditis elegans
let-65
Aminoacyl tRNA synthetase, methionyl-tRNA synthetase
Oligonucleotide array comparative genomic hybridization
Essential genes

ABSTRACT

Cytoplasmic methionyl tRNA synthetase (MetRS) is one of more than 20 cytoplasmic aminoacyl tRNA synthetase enzymes (ARS). This family of enzymes catalyzes a process fundamental for protein translation. Using a combination of genetic mapping, oligonucleotide array comparative genomic hybridization, and phenotypic correlation, we show that mutations in the essential gene, *let-65*, reside within the predicted *Caenorhabditis elegans* homologue of MetRS, which we have named *mars-1*. We demonstrate that the lethality associated with alleles of *let-65* is fully rescued by a transgenic array that spans the *mars-1* genomic region. Furthermore, sequence analysis reveals that six *let-65* alleles lead to the alteration of highly conserved amino acids.

Introduction

Analysis of mutations affecting genes that are essential for survival in model organisms, such as the nematode *Caenorhabditis elegans*, have been instrumental in developing our understanding of fundamental biological processes, for example cell division and morphology among others (Green et al., 2011). Many of these processes have a direct relevance to the understanding of human diseases such as metabolic disorders and obesity (Hashmi et al., 2013). The generation of mutations in essential genes and their correlation with a specific genetic locus is, therefore, the first step in understanding the nature of a gene's function in promoting survival. To more easily identify genes that are essential for survival in the model organism *C. elegans*, we created a physical deficiency map that encompasses a portion of Chromosome IV, allowing for the correlation

* Corresponding author at: Department of Biochemistry, College of Medicine and Health Sciences, Sultan Qaboos University, Oman, 35, Al-Khod 123, Oman. Tel.: +968 24141152; fax: +968 2441 3880.
E-mail address: mahazr@squ.edu.om (M.Z. Alriyami)

http://dx.doi.org/10.1016/j.mgene.2014.08.006

of genetically defined mutations with physically defined candidate gene lists. In this study, we describe the use of this physical deficiency map to identify the essential gene (*let-65)*, previously isolated from ethyl methanesulfonate (EMS) genetic screens (Clark and Baillie, 1992; Clark et al., 1988; Moerman and Baillie, 1981; Rogalski and Baillie, 1985; Rogalski et al., 1982). From this analysis we determined that mutations in *let-65* correspond to the methionyl tRNA synthetase gene, *mars-1*.

Organisms utilize four key components to translate genetic information into proteins: ribosomes, messenger RNAs (mRNA), transfer RNAs (tRNA), and aminoacyl-tRNA synthetases (ARS), reviewed in Brown et al. (2010). Ribosomes are two-subunit organelles that direct synthesis of proteins by translating nucleic acid triplets, carried by mRNAs, into amino acids (Crick et al., 1961). mRNAs, which are generated by DNA transcription, act as translation templates, whereas tRNAs act as interfaces between mRNA strands and the amino acids added to a polypeptide chain by transferring the required amino acids to the growing chain (Crick et al., 1961).

The ARS family of proteins encompasses more than 20 multi-domain enzymes. Each of these proteins specifically recognizes one amino acid and its corresponding subset of cognate tRNA molecules to catalyze the aminoacylation of the former to the latter in a two-step reaction. In the first step a specific amino acid is activated by ATP to produce aminoacyl adenylate; in the second step the amino acid is transferred onto the 3′-end of the tRNA (Jakubowski, 2001). ARS are categorized into two classes (I and II) based on their distinct core domains (Wolf et al., 1999). The catalytic domain of Class I synthetases exhibit a typical Rossman fold, which has a three-layer topology (α/β/α) with an inner core of five parallel beta strands. However, Class II synthetases generally consist of only β strands (Berg et al., 2001). Class I and II synthetases recognize different faces of tRNA molecules (Berg et al., 2001). The tRNA CCA terminus (containing the amino acid attachment site) adopts different conformations in tRNA-synthetase complexes for each class. The CCA termini of Class I enzymes have hairpin conformations, whereas, class II enzymes have the same helical confirmations observed in free tRNAs (Berg et al., 2001). An intriguing enzyme that belongs to this family is the Class I synthetase, methionyl tRNA synthetase (MetRS), which recognizes both the initiator tRNA as well as the methionine carrying tRNA that functions in the elongation of peptide chains (Deniziak and Barciszewski, 2001).

In this study, we use oligonucleotide array Comparative Genomic Hybridization (aCGH) to define a physical deficiency map of a region of Chromosome IV (LGIV) in the nematode *C. elegans*. We then use this map to assign precisely defined lists of candidate genes to previously generated and genetically mapped mutations, which confer a lethal phenotype upon the organism. Using this approach, we demonstrate that the lethality associated with alleles of *let-65* is fully rescued by a transgenic array that spans the *mars-1* genomic region.

Materials and methods

Growth and handling of C. elegans

Nematode strains were cultured as previously described (Brenner, 1974). Strains used in this study: N2 Bristol (BC00049), *dpy-5*(*e907*) (CB00907), *let-65*(*s254*); *nT1*(*IV*) (BC00962), *let-65*(*s694*); *nT1*(*IV*) (BC1121), *let-65*(*s1154*); *nT1*(*IV*) (BC2028), *let-65*(*s1222*); *nT1*(*IV*) (BC2116), *let-65*(*s1083*); *nT1*(*IV*) (BC1909), *let-65*(*s1084*); *nT1*(*IV*) (BC1910), *let-65*(*s1730*); *nT1*(*IV*) (BC3258), *let-65*(*s1777*); *nT1*(*IV*) (BC3305), *let-65*(*s1083*), *unc-22*(*s7*)(*IV*) [*WRM0615dH10* + *myo-2*::*GFP* + *pCeh361*] (BC8656), *let-65*(*s1222*), *unc-22*(*s7*) (*IV*) [*WRM0615dH10* + *myo-2*::*GFP* + *pCeh361*] (BC8696).

Oligonucleotide array comparative genomic hybridization (aCGH) data analysis

aCGH experiments and analysis were performed as previously described (Jones et al., 2007).

DNA sequence analysis

Genomic DNA was extracted from *let-65* homozygous animals for eight alleles (*s1730*, *s1083*, *s694*, *s1084*, *s254*, *s1154*, *s1777* and *s1222*). PCR fragments spanning the *mars-1* ORF were amplified from genomic DNA templates with primers designed using Primer 3 (Rozen and Skaletsky, 2000) (primer sequences available upon request). The amplified templates were subsequently sequenced (Macrogen, Korea) with appropriate sequencing primers (Supplementary Table 1).

Transgenic rescue

The fosmid *WRM0615dH10*, which spans the *mars-1* genomic region (11,631,182 to 11,665,626 bp on LGIV), was injected into the syncytial gonads of *dpy-5* (*e907*) animals at a final concentration of 3 ng/μl, along with the pharyngeal targeted green fluorescent protein (GFP) reporter *myo-2*::*GFP* at 10 ng/μl and the plasmid *pCeh361*,which contains wild-type *dpy-5* that rescues the *dpy-5* phenotype, at 100 ng/μl (Thacker et al., 2006). Injected worms were plated (five P_0s per plate) and phenotypically wild-type animals from the F_1s generation that expressed the pharyngeal GFP reporter were individually isolated and propagated. GFP positive wild-type F_2 animals transmitting the transgenic array were selected and a PCR specific to the fosmid backbone performed to confirm the presence of the fosmid (Supplementary Table 1). A single, stably transmitting, animal was selected to create transgenic lines. To conduct transgenic rescue experiments individual hermaphrodites carrying *WRM0615dH10* + *myo-2*::GFP + *pCeh361* were crossed to +/+ N2 males to obtain *dpy-5* (*e907*)/+[*WRM0615dH10* + *myo-2*::GFP + *pCeh361*] males. These GFP-expressing males were then crossed to *let-65* (*s1222*); *unc-22* (*s7*) *and let-65* (*s1803*); *unc-22* (*s7*) heterozygously maintained hermaphrodites. F_1 progeny were screened for the presence of males to ensure the cross was successful and F_1 wild-type hermaphrodites plated individually. F_2 animals were individually plated and screened for unconditional twitchers, which is indicative of viable homozygous animals of the genotype *let-65* (*s1083* or *s1222*) *unc-22* (*s7*)/*let-65* (*s1083* or *s1222*) *unc-22* (*s7*)(*IV*)[*WRM0615dH10* + *myo2*::GFP + *pCeh361*]. Unconditional twitchers were individually plated and incubated at 20 °C to assess rescue.

Microscopy

Screening of transgenic animals was performed using a Zeiss Axioscope (Quorum Technologies) set with a QImaging Camera and appropriate filter sets were used for GFP expression analysis. Animals were immobilized with 100 mM sodium azide (in water) immediately prior to imaging.

Bioinformatics studies

All reference DNA sequences were derived from WormBase release WS208. Analysis of sequence data and image processing were performed using BioEdit (http://www.mbio.ncsu.edu/BioEdit/BioEdit.html), WolFPSORT (Horton et al., 2007) (http://wolfpsort.org/), and BLASTP.

Results

aCGH mapped deficiencies physically define seven zones in the unc-22 region of LGIV

In an effort to characterize genes that are essential for survival in *C. elegans* we previously generated and mapped mutations that confer a lethal phenotype in the *unc-22* region of the *C. elegans* genome (Rogalski et al., 1982). The *unc-22* region (from *unc-43* to *unc-31*) represents approximately two map units (mu) on chromosome IV, which is spanned by 35 well-defined genetic deficiencies (Clark and Baillie, 1992; Clark

Table 1
aCGH defined deficiency breakpoints and gene complement.

Zone	Left breakpoint		Right breakpoint		Genes within zones		Size kb
	Left (bp)	Right (bp)	Left (bp)	Right (bp)	Leftmost	Rightmost	
1	10,942,926	10,947,094	11,070,201	11,076,154	Y69E1A.1	Y5F2A.3	123–134
2	11,070,201	11,076,154	11,629,649	11,629,799	Y5F2A.4	lys-6	553–560
3	11,629,649	11,629,799	11,639,511	11,639,570	F58B3.4	F58B3.7[a]	10–11
4	11,639,511	11,639,570	11,983,621	11,983,655	F58B3.7[a]	*unc-22*[a]	343–344
5	11,983,621	11,983,655	12,243,159	12,251,414	*unc-22*[a]	sre-16	252–253
6	12,243,159	12,251,414	12,551,447	12,553,668	*sre-19*	ZK795.1[a]	315–317
7	12,551,447	12,553,668	12,757,572	12,776,111	ZK795.1[a]	*inx-6*	196–204

Regions are annotated based on the probe sequence falling adjacent to the breakpoint position.
[a] Breakpoints falling within a single gene.

et al., 1988; Marra and Baillie, 1994; Moerman and Baillie, 1981; Rogalski and Baillie, 1985; Rogalski et al., 1982; Schein et al., 1993). Previously, genetic analysis was used to identify and position 31 lethal gene loci within this genetic deficiency map (Clark et al., 1988). However, because the precise molecular extent of the deficiencies was not known the construction of definitive candidate gene lists that corresponded to each mutated loci was not possible.

To physically define the deficiency map in the *unc-22* region we analyzed four of the most informative deficiencies using array comparative genomic hybridization (aCGH). aCGH is a method used to compare the DNA ratio between individual samples from the same organism in order to determine copy number variations on a chromosomal or genome-wide scale (Dhami et al., 2005). This method can be used to determine the precise physical extent of genetic deficiencies in *C. elegans* (Jones et al., 2007). In all four cases, array data was sufficient to position deficiency deletion breakpoints at a single-gene resolution to within approximately 6 kb (Table 1). Deletion breakpoints fell either within single genes or between two genes. Using these data a seven-zone physical deficiency map was constructed spanning approximately 800 kb of the *unc-22* region (Fig. 1A and Table 1). Furthermore, by incorporating previously generated genetic mapping data for the molecularly unidentified lethal loci known to map into the region, we were able to assign each locus to a precisely defined list of candidate genes. The largest number of lethal loci map into zone two of the deficiency map (nine genes). Zone six contains seven lethal loci, while zones four and seven contain six lethal loci each. Zones one and five contain three lethal loci each. Finally, a single lethal locus was positioned into zone three (Fig. 1A).

Sequence analysis of let-65

let-65 was the single lethal loci that mapped into zone three of the deficiency map (Fig. 1A). Based on the physical extent of zone three only four annotated genes, F58B3.4, F58B3.5, F58B3.6 and F58B3.7 were candidates for *let-65*. *let-65* is represented by nine alleles, all of which were isolated in EMS screens for larval lethality (Clark and Baillie, 1992; Clark et al., 1988; Rogalski and Baillie, 1985; Rogalski et al., 1982). Since mutations in *let-65* confer a lethal phenotype we correlated available RNAi data in order to rank the candidate genes based upon the severity of the reported RNAi phenotype (Kamath et al., 2003; Maeda et al., 2001; Piano et al., 2002; Simmer et al., 2003; Sonnichsen et al., 2005) (Fig. 1B). Using this approach we identified F58B3.5 and F58B3.4 as the two strongest candidate genes. To identify mutations in F58B3.5 and F58B3.4 their genomic regions were amplified by PCR from *let-65* (*s254*) homozygous animals and sequenced. A comparison of sequence data to the reference *C. elegans* genome (WormBase release WS208) revealed no mutations that were associated with F58B3.4. However, a point mutation was identified in the coding region of F58B3.5. This mutation resulted in a glutamic acid to lysine substitution (Glu364Lys) in a highly conserved amino acid (Fig. 2 and Table 2). To provide additional confirmation that mutation of F58B3.5 was associated with *let-65* we sequenced the F58B3.5 ORF in the remaining *let-65* alleles, detecting the presence of additional mutations in seven of the eight alleles tested. All mutations identified were G → C to A → T transitions, consistent with the known mutational spectrum for ethane methylsulfonate (EMS) mutagenesis (Fig. 2, Table 2) (Maple and Moller, 2007). In every case, with the exception of *s1777 and s1222*, the identified mutations were substitutions in highly conserved amino acids in the predicted MetRS domain of F58B3.5 (Figs. 1C and 2). The lesion *s1777* resulted in a 508 amino acid truncation, while *s1222* contained a mutation in the first base of the first intron. *s1222* potentially prevents splicing of the first intron, which would result in a premature truncation of the protein at amino acid 28 (Fig. 1C). Together these data demonstrate that *let-65* and F58B3.5 represent the same gene. F58B3.5 encodes a cytoplasmic Class I synthetase, methionyl tRNA synthetase (MetRS) that is evolutionary well conserved, with 58% identity and 71% similarity to its human homologue (BlastP alignment Altschul et al., 1997, 2005) (Fig. 2). We have renamed F58B3.5 *mars-1* (Methionyl Amino-acyl tRNA Synthetase-1).

Rescue of let-65 with a fosmid containing mars-1

We have shown that mutations of *mars-1* are present in *let-65* animals. To provide further evidence that these mutations caused the lethality associated with *let-65*, we performed transgenic rescue experiments. The fosmid clone *WRM0615dH10* spans the entire *mars-1* genomic region, but excludes the F58B3.4 ORF (Fig. 1B). We found that transgenic animals carrying *WRM0615dH10* as an extra-chromosomal array

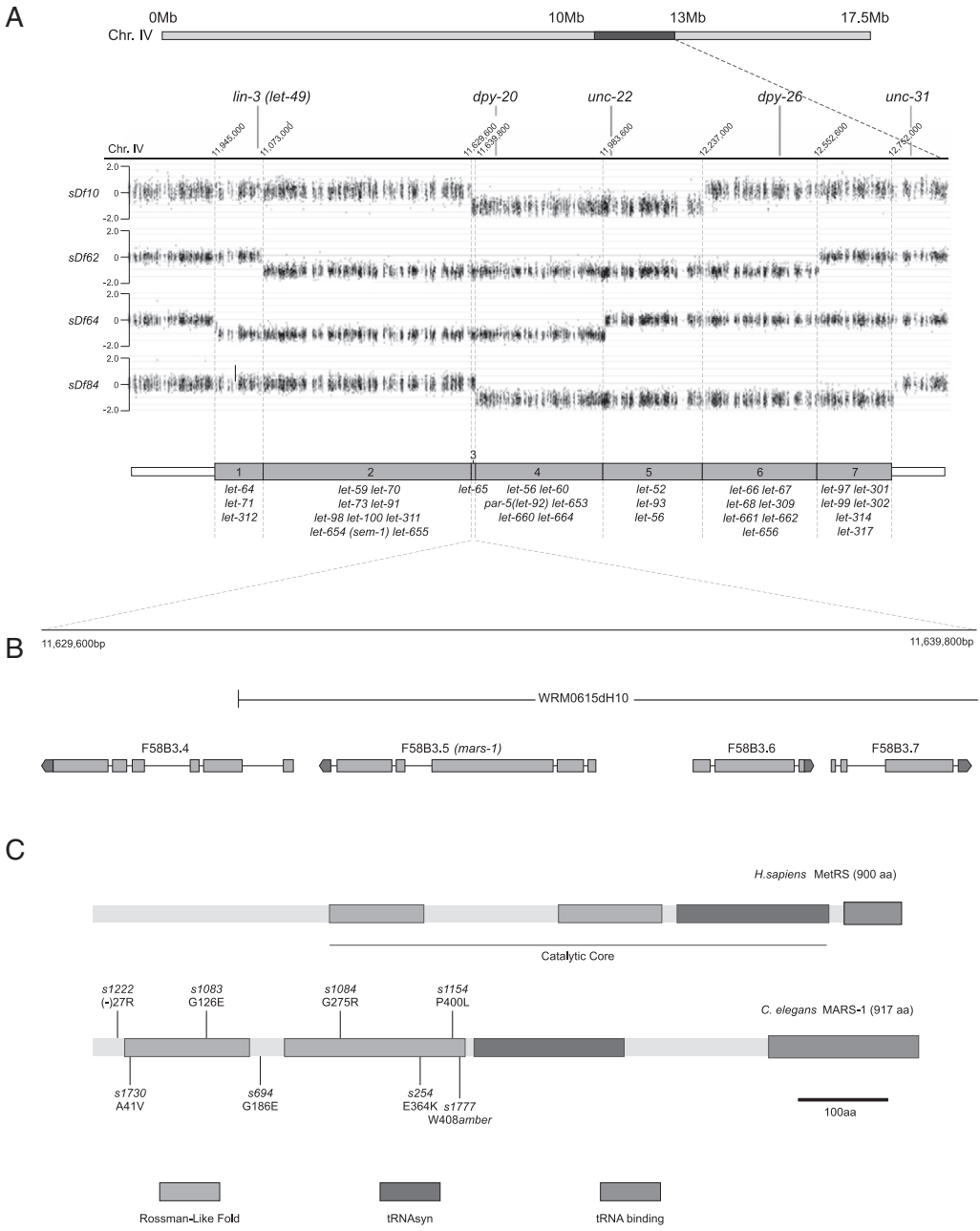

Fig. 1. Physical map of the *unc-22* region of LGIV including our aCGH deficiency mapping data. A. Schematic of LGIV showing the regions uncovered by the deficiencies analyzed in this study is shown at the top of the figure (dark gray). aCGH data for the region around the four analyzed deficiencies is also shown. Commonly used genetic markers and the physical positions of the deficiency breakpoints (estimated to within 0.5 kb) are shown above the array data. A schematic describing the seven zones defined by the breakpoints of the deficiencies is shown below the array data. *let-56* is in either zone five or zone six. B. An expansion of zone three showing the four candidate genes for *let*-65. The region covered by the fosmid used in complementation tests is depicted above the gene models. C. Schematic showing protein structures of human MetRS and *C. elegans* MARS-1. Conserved domains and their locations are indicated. The positions of identified *mars-1* mutations are shown. Figure not to scale.

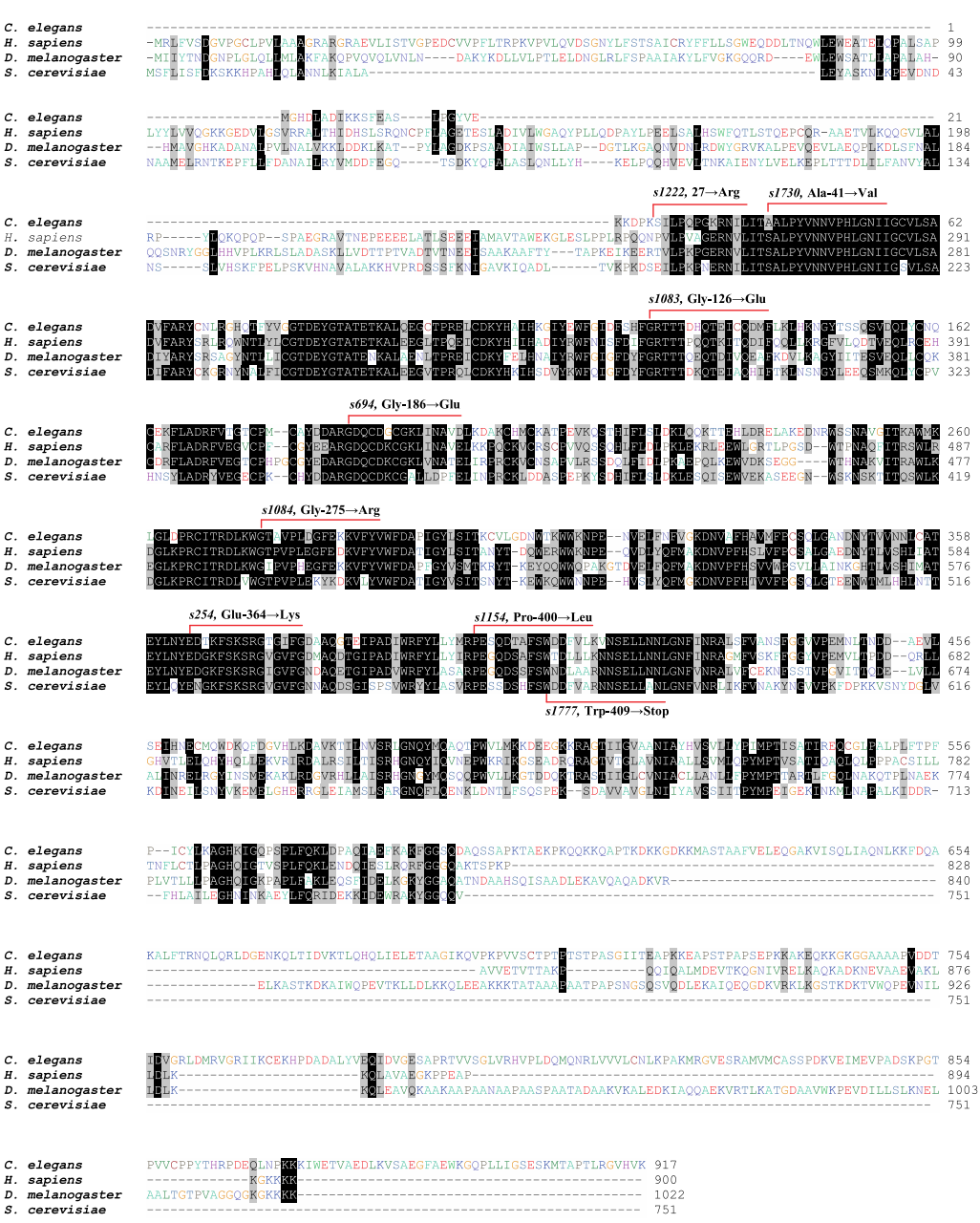

Fig. 2. *C. elegans* MARS-1 alignment with its human (GenBank accession no. NP_004981), *D. melanogaster* (GenBank accession no. NP_611382), and *S. cerevisiae* (GenBank accession no. CAA97293) orthologs showing conserved amino acids and identified *let-65* mutations. MARS-1 is evolutionarily well conserved with 58% identity and 71% similarity to its human ortholog. Red lines indicate the positions of amino acid changes in *let-65* alleles. *s1222* prevents the splicing of the first intron. *s1777* induces a stop codon that prematurely truncates the protein.

Table 2
Mutations identified in *let-65* by sequencing and BioEdit analysis.

Allele	Nucleotide change	Codon mutation	Amino acid change
s174	ND	ND	ND[a]
s254	G1188A	gag → **a**ag	Glu-364 → Lys
s694	G655A	gga → g**a**a	Gly-186 → Glu
s1083	G429A	gga → g**a**a	Gly-126 → Glu
s1084	G921A	gga → **a**ga	Gly-275 → Arg
s1154	C1297T	ccc → c**t**c	Pro-400 → Leu
s1222	G81A	(−) → **a**ga	(−)27 → Arg
s1730	C174T	gct → g**t**t	Ala-41 → Val
s1777	G1324A	tgg → t**a**g	Trp-409 → Stop

[a] Not determined — potentially a complex lesion.

successfully and specifically rescued the lethality associated with the two alleles of *let-65* tested (*s1083* and *s1222*). Furthermore these transgenic animals were superficially wild type in appearance except for the twitching phenotype due to the presence of the *unc-22* (*s7*) homozygous mutation, which was used as a visible phenotypic marker. Individual viable animals were isolated to start the strains BC8656 and BC8696 for alleles *s1083* and *s1222* respectively, and cultured over multiple generations to confirm continued rescue of the lethal phenotype. Together the sequence and rescue data confirm that mutation of *mars-1* gives rise to the lethal phenotype in *let-65* animals.

Mitochondrial methionyl tRNA synthetase candidate

MARS-1 has been predicted to encode the cytoplasmic form of MetRS (Havrylenko et al., 2010). Computational analysis of the predicted MARS-1 protein (WormBase WS209) demonstrated that *mars-1* encodes a protein that localizes subcellularly in the cytoplasm. BLASTP analysis of Y105E8A.20 (Altschul et al., 1997), a second gene predicted to encode methionyl tRNA synthetase in *C. elegans* (WormBase WS211), implies mitochondrial localization for the encoded protein.

Discussion

In this study we used aCGH to define a physical deficiency map for a 2 Mb portion of LGIV in *C. elegans*. Using this deficiency map we correlated physical candidate gene lists with mutations in a number of previously identified essential genes (Clark, 1990; Clark and Lyckegaard, 1988; Moerman, 1980; Rogalski and Baillie, 1985; Rogalski et al., 1982). From our analysis we were able to determine that the genetically identified locus *let-65* corresponds to *mars-1*, which encodes the *C. elegans* homologue of MetRS. *let-65* is one of a number of lethal genes previously identified in EMS screens for essential genes on the right end of LGIV. These screens generated seven alleles for *let-65* six of which resulted in lethality (*s253*, *s1777*, *s1083*, *s1154*, *s1222* and *s1730*), and one additional allele causing sterility (*s694*). Previously it has been reported that inactivation of MARS-1 by RNAi leads to embryonic lethality (Kamath et al., 2003; Sonnichsen et al., 2005). Moreover, this loss of MARS-1 function can also lead to sterility (Havrylenko et al., 2011). Furthermore, correlation of available phenotypic information for other ARS family members deleted by targeted mutations demonstrates a complete correlation of lethal and sterile phenotypes, highlighting the essential nature of this family of enzymes (Table 3). Together these data indicate that *mars-1* is essential for processes that require large scale protein biosynthesis, such as embryogenesis (Havrylenko et al., 2011).

In *C. elegans* cytoplasmic MetRS (MARS-1) is a 917 amino acid protein that has multiple predicted functional domains. The minimal core enzyme, as found in the bacteria *Aquifex aeolicus*, consists of a core catalytic domain (CAT), which is distinguished by a Rossmann fold that is a characteristic domain of Class I aminoacyl tRNA synthetase enzymes (catalyzes the aminoacylation reaction), and an anticodon-binding domain (ABD) (promotes tRNA–protein association) (Nakanishi et al., 2005). Additional domains, found in eukaryotes, are appended in front of or after the core domain of the enzyme and are believed to increase the stability of tRNA–protein complexes (Havrylenko et al., 2010). In humans, MARS carries a C-terminal domain that causes a slow release of aminoacyl-tRNA and provides the aminoacylation reaction with a limiting step (Havrylenko et al., 2010). In *C. elegans*, the MARS-1 protein model contains a t-RNA binding domain (tRBD) of about 170

Table 3
Correlation of known viability phenotypes in ARS-family genes.
Data source: WormBase, http://www.wormbase.org, release WS243, May 2014.

Gene name	RNAi phenotype	Mutation phenotype [allele]
aars-1	ste	
aars-2	emb; lvl; ste	
cars-1	emb; let; lvl; ste	
cars-2	ste	
dars-1	emb; lvl; ste	
dars-2	emb; lvl; ste	let; ste [*tm3799*]
ears-1	emb; lvl	
ears-2	emb; lvl; ste	
fars-1	emb; lvl; ste	
fars-2	ND	
fars-3	emb; lvl; ste	
gars-1	emb; lvl; ste	
hars-1	emb; let; lvl; ste	let; ste [*tm4074*] (Pierce et al., 2011)
iars-1	emb; ste	
iars-1	emb; let	
kars-1	emb; lvl; ste	
lars-1	emb; let; lvl; ste	
lars-2	emb	let; ste [*tm5774*]
mars-1	emb; lvl	
nars-1	emb; let; lvl; ste	let; ste [*tm524*]
nars-2	ND	
pars-1	emb; let; lvl; ste	
pars-2	ND	
qars-1	emb; let; lvl; ste	
rars-1	emb; lvl; ste	emb; lvl; ste [*gc47*]
rars-2	emb; ste	
sars-1	emb; lvl; ste	
sars-2	emb	
tars-1	emb; let; lvl; ste	
vars-1	emb; lvl; ste	
vars-2	emb; let; lvl; ste	let; ste [*tm3947*]
wars-1	emb; lvl; ste	
yars-1	emb; lvl; ste	

emb = embryonic lethality, lvl = larval lethality, let = lethal, ste = sterile, ND = not determined.

residues appended in *cis* at the C-terminus of the enzyme, which is homologous to the tRBD domain appended at the C-terminus of MetRS in *Oriza sativa* (rice) (Havrylenko et al., 2010). However, in *C. elegans*, this domain is separated from the ABD domain by a putative protein-binding domain of approximately 150 residues (Havrylenko et al., 2010). This domain is similar to the human P43 protein, which is a non-synthetase component of the aminoacyl tRNA synthetase complex required for the assembly of the protein within the complex (Havrylenko et al., 2011). The predicted *C. elegans* MARS-1 catalytic core domain comprises amino acids 36 to 407 (Marchler-Bauer et al., 2009). In *let-65* six out of the eight mutations reside in this domain (*s1730*, *s1083*, *s694*, *s1084*, *s254* and *s1154*) suggesting that they disrupt the function of the catalytic core. The EMS induced mutation in *let-65*(*s1777*) creates a truncated protein that lacks the anticodon-binding domain (AA416 to AA545) (Marchler-Bauer et al., 2009) and the tRNA binding domain (AA755 to AA857) (Marchler-Bauer et al., 2009). *let-65*(*s1222*) contains a mutation in the first base of the first intron putatively preventing splicing of the first intron resulting in putative premature truncation at amino acid 28.

mars-1 is highly conserved and encodes a protein with unique significance for translational control. Structures of MARS-1 from *Escherichia coli* (Mechulam et al., 1999) and *Thermus thermophiles* (Sugiura et al., 2000) have been reported. MARS-1 from *E. coli* has greater identity than *T. thermophiles* to the protein in *C. elegans*. Amino acids altered in four of the six *C. elegans* missense mutations are located in conserved regions in both *E. coli* and *T. thermophiles*: *let-65*(*s1083*): βC; *let-65*(*s694*): turn between β2 and β3 (zinc-binding domain); *let-65*(*s254*): turn between β7 and α6; *let-65*(*s1154*): turn between α7 and β8. *let-65*(*s1730*) is in turn between βA and αA and *let-65*(*s1084*) is in turn between β5 and β6. All identified *let-65* changes are located in *E. coli* at the Rossman-like fold.

let-65(*s1777*) contains a nonsense mutation wherein the tryptophan codon at amino acid 408 becomes an amber stop codon. *let-65*(*s1222*) has a mutation that putatively prevents splicing of the first intron resulting in premature truncation, the truncated protein lacks the majority of the synthetase including the catalytic domain, the two binding domains, the anticodon-binding domain, and the tRNA binding domain. *s1777* and *s1222* are, therefore, both putatively null alleles. The phenotypes (mid-larval arrest) (Clark and Baillie, 1992; Clark et al., 1988) of these two mutants are indistinguishable from six other *let-65* alleles. Based on this observation we predict that all six alleles are null for MARS-1 function, or reduce MARS-1 activity below a threshold required for survival. Moreover, the survival of nematodes, containing those alleles to mid-larva, suggests that for the early developmental stages, the mother supplies MARS-1 to eggs during oogenesis.

let-65(*s694*) and *let-65*(*s1083*) contain mutations that change glycine to glutamic acid at amino acids 186 and 126 respectively. These two amino acids are in sequence blocks that are conserved from human to yeast. *let-65*(*s1084*) contains a mutation that changes glycine to arginine at amino acid 275. Generally, a change from a small non-charged non-hydrophilic glycine to a large positively charged hydrophilic arginine could disrupt the structure of a protein enough to render it non-functional, which in the case of *let-65* would cause lethality. *let-65* (*s1730*) contains a mutation that changes alanine to valine at position 41. In humans, *Drosophila melanogaster* and *Saccharomyces cerevisiae*, a serine codon is located at the same position within a highly conserved sequence block. Because a change from a polar amino acid (serine) to a non-polar amino acid (valine) could disrupt the protein's structure more than replacing alanine with valine we analyzed the available sequences to confirm the *C. elegans* wild type reference sequence at position 41. After analyzing the same sequence stretch in the other *let-65* alleles we concluded definitively that an alanine is present at position 41. The mutation in *let-65*(*s254*) replaces a conserved acidic amino acid (glutamic acid) with a basic amino acid (lysine) at position 364. This position is within the catalytic domain of MARS-1. *let-65*(*s1154*) has a polar amino acid (proline) instead of the non-polar amino acid (leucine) at position 400 that is within the catalytic domain. Additionally, since alterations of proline in the flanking regions of active sites are known to change the activity of the protein (Proline bracket hypothesis) (Kini et al., 1998), this mutation may exhibit drastic effects on MARS-1 activity.

Typically, eukaryotic cells have two forms of ARS: cytoplasmic and mitochondrial (Walker et al., 1983). MARS-1 localizes to the cytoplasm (Havrylenko et al., 2010). Therefore, we attempted to find the candidate gene coding for the mitochondrial form of the enzyme. The *C. elegans* genome contains a second ORF, Y105E8A.20 that is predicted to encode a methionyl tRNA synthetase (WormBase WS211). Y105E8A.20 is a 1221 bp gene that encodes a 406 amino acid protein. BlastP was used to search the Swiss-Prot database with MARS-1 amino acid sequences as queries (Altschul et al., 1997, 2005). In the BlastP searches *S. cerevisiae* was used as the canonical genome to identify protein localizations in *C. elegans*. *S. cerevisiae* was used because it's one of the simplest in regard to identifying open reading frames (ORFs), and it is the most highly characterized eukaryotic genome (Fisk et al., 2006). BlastP localization, based on *S. cerevisiae* matches, predicted mitochondrial localization for the protein encoded by Y105E8A.20.

The *C. elegans* million mutations project (MMP) comprises a library of 2007 mutagenized and homozygous *C. elegans* strains that were sequenced to identify the resulting mutations. The aim of this project was to generate mutant alleles in each gene in the *C. elegans* genome. However, a limitation of this approach is that severely detrimental mutations in genes required for viability would be selected against since these strains must be viable and fertile. Of the 14 MMP mutations that are recorded for *mars-1*, all are missense mutations with the single exception of a mutation that is predicted to affect splicing of an intron in the C-terminus of the protein. This intron is not conserved in other species (Table 4). Out of the remaining 13 alleles 10 affect residues that are also not conserved. Interestingly three mutations reside within the catalytic core of the enzyme. Since these strains are viable it is, however, unlikely that these lesions severely affect the function of the protein. Alternatively secondary mutations within these strains might be present that could compensate for any detrimental effect on enzyme function. Further studies would be required to ascertain if this is the case.

In our study we isolated and identified eight *mars-1* alleles all of which lead to a loss-of-function phenotype. Furthermore, the number of alleles isolated in *mars-1* is greater than the average number of alleles isolated for most genes in screens of this type (Berg et al., 2001; Brenner, 1974; Deniziak and Barciszewski, 2001; Horton et al., 2007; Jones et al., 2007; Rozen and Skaletsky, 2000; Thacker et al., 2006). This indicates that *let-65* (*mars-1*) is a large mutagenic target for EMS. The availability of the loss of function alleles allows for further exploration of the gene's structure and function in a tractable model system. Moreover, it will

Table 4
mars-1 MMP mutations.

Allele	Mutation	Effect	Domain affected	Type	Conserved residue
gk213073	C → T	A887T	None	Missense	No
gk410589	G → A	P865S	None	Missense	No
gk445673	C → T	Affects splicing	None	Intron, splicing	No
gk485164	C → T	A453T	Anticodon_1	Missense	No
gk509072	C → T	D190N	tRNA-synt_1g	Missense	Yes
gk515000	C → T	G668E	coiled_coil_region	Missense	No
gk541772	G → A	P265L	tRNA-synt_1g	Missense	Yes
gk550543	C → T	A613T	Anticodon_1	Missense	No
gk636296	G → A	S837L	tRNA_bind	Missense	No
gk656240	T → C	M844V	tRNA_bind	Missense	No
gk719138	C → T	A431T	tRNA-synt_1g	Missense	Yes
gk780068	C → T	D752N	None	Missense	No
gk793812	A → T	F412L	tRNA-synt_1g	Missense	No
gk889082	G → A	A520V	Anticodon_1	Missense	No

facilitate a better understanding of ARS, which will lead to a deeper understanding of the complete protein synthesis machinery in all organisms including humans.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.mgene.2014.08.006.

Abbreviations

mRNA	messenger RNAs
tRNA	transfer RNAs
ARS	aminoacyl-tRNA synthetases
METRS	methionyl tRNA synthetase
aCGH	oligonucleotide array comparative genomic hybridization
EMS	ethyl methanesulfonate
mars-1	*C. elegans* methionyl tRNA synthetase encoding gene
GFP	green fluorescent protein
NLS	nuclear localization sequence
MARS-1	*C. elegans* methionyl tRNA synthetase

Funding

This project was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) [RGPIN9975] Canada. Maha Z. AlRiyami was supported by a fellowship from Sultan Qaboos University, Sultanate of Oman.

Acknowledgments

We thank Domena Tu for her help and for injecting all constructs made in this study. We also thank Shu Yi Chua for designing the fosmid PCR primers.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.
Altschul, S.F., Wootton, J.C., Gertz, E.M., Agarwala, R., Morgulis, A., Schaffer, A.A., Yu, Y.K., 2005. Protein database searches using compositionally adjusted substitution matrices. FEBS J. 272, 5101–5109.
Berg, J.M., Tymoczko, J.L., Stryer, L., 2001. Biochemistry. Biochemistry. W.H. Freeman and Company, New York, p. 822.
Brenner, S., 1974. The genetics of *Caenorhabditis elegans*. Genetics 77, 71–94.
Brown, M.V., Reader, J.S., Tzima, E., 2010. Mammalian aminoacyl-tRNA synthetases: cell signaling functions of the protein translation machinery. Vasc. Pharmacol. 52, 21–26.

Clark, D.V., 1990. The unc-22 (IV) region of *Caenorhabditis elegans*: genetic analysis and molecular mapping(PhD Thesis) Simon Fraser University.
Clark, D.V., Baillie, D.L., 1992. Genetic analysis and complementation by germ-line transformation of lethal mutations in the unc-22 IV region of *Caenorhabditis elegans*. Mol. Gen. Genet. 232, 97–105.
Clark, A.G., Lyckegaard, E.M., 1988. Natural selection with nuclear and cytoplasmic transmission. III. Joint analysis of segregation and mtDNA in *Drosophila melanogaster*. Genetics 118, 471–481.
Clark, D.V., Rogalski, T.M., Donati, L.M., Baillie, D.L., 1988. The unc-22(IV) region of *Caenorhabditis elegans*: genetic analysis of lethal mutations. Genetics 119, 345–353.
Crick, F.H., Barnett, L., Brenner, S., Watts-Tobin, R.J., 1961. General nature of the genetic code for proteins. Nature 192, 1227–1232.
Deniziak, M.A., Barciszewski, J., 2001. Methionyl-tRNA synthetase. Acta Biochim. Pol. 48, 337–350.
Dhami, P., Coffey, A.J., Abbs, S., Vermeesch, J.R., Dumanski, J.P., Woodward, K.J., Andrews, R.M., Langford, C., Vetrie, D., 2005. Exon array CGH: detection of copy-number changes at the resolution of individual exons in the human genome. Am. J. Hum. Genet. 76, 750–762.
Fisk, D.G., Ball, C.A., Dolinski, K., Engel, S.R., Hong, E.L., Issel-Tarver, L., Schwartz, K., Sethuraman, A., Botstein, D., Cherry, J.M., 2006. *Saccharomyces cerevisiae* S288C genome annotation: a working hypothesis. Yeast 23, 857–865.
Green, R.A., Kao, H.L., Audhya, A., Arur, S., Mayers, J.R., Fridolfsson, H.N., Schulman, M., Schloissnig, S., Niessen, S., Laband, K., Wang, S., Starr, D.A., Hyman, A.A., Schedl, T., Desai, A., Piano, F., Gunsalus, K.C., Oegema, K., 2011. A high-resolution *C. elegans* essential gene network based on phenotypic profiling of a complex tissue. Cell 145, 470–482.
Hashmi, S., Wang, Y., Parhar, R.S., Collison, K.S., Conca, W., Al-Mohanna, F., Gaugler, R., 2013. A *C. elegans* model to study human metabolic regulation. Nutr. Metab. 10, 31.
Havrylenko, S., Legouis, R., Negrutskii, B., Mirande, M., 2010. Methionyl-tRNA synthetase from *Caenorhabditis elegans*: a specific multidomain organization for convergent functional evolution. Protein Sci. 19, 2475–2484.
Havrylenko, S., Legouis, R., Negrutskii, B., Mirande, M., 2011. *Caenorhabditis elegans* evolves a new architecture for the multi-aminoacyl-tRNA synthetase complex. J. Biol. Chem. 286, 28476–28487.
Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., Nakai, K., 2007. WoLF PSORT: protein localization predictor. Nucleic Acids Res. 35, W585–W587.
Jakubowski, H., 2001. Translational accuracy of aminoacyl-tRNA synthetases: implications for atherosclerosis. J. Nutr. 131, 2983S–2987S.
Jones, M.R., Maydan, J.S., Flibotte, S., Moerman, D.G., Baillie, D.L., 2007. Oligonucleotide array comparative genomic hybridization (oaCGH) based characterization of genetic deficiencies as an aid to gene mapping in *Caenorhabditis elegans*. BMC Genomics 8, 402.
Kamath, R.S., Fraser, A.G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D.P., Zipperlen, P., Ahringer, J., 2003. Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. Nature 421, 231–237.
Kini, R.M., Caldwell, R.A., Wu, Q.Y., Baumgarten, C.M., Feher, J.J., Evans, H.J., 1998. Flanking proline residues identify the L-type Ca2 + channel binding site of calciseptine and FS2. Biochemistry 37, 9058–9063.
Maeda, I., Kohara, Y., Yamamoto, M., Sugimoto, A., 2001. Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi. Curr. Biol. 11, 171–176.
Maple, J., Moller, S.G., 2007. Mutagenesis in *Arabidopsis*. Methods Mol. Biol. (Clifton, N.J.) 362, 197–206.
Marchler-Bauer, A., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Lu, S., Marchler, G.H., Mullokandov, M., Song, J.S., Tasneem, A., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Bryant, S.H., 2009. CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Res. 37, D205–D210.
Marra, M.A., Baillie, D.L., 1994. Recovery of duplications by drug resistance selection in *Caenorhabditis elegans*. Genome 37, 701–705.
Mechulam, Y., Schmitt, E., Maveyraud, L., Zelwer, C., Nureki, O., Yokoyama, S., Konno, M., Blanquet, S., 1999. Crystal structure of *Escherichia coli* methionyl-tRNA synthetase highlights species-specific features. J. Mol. Biol. 294, 1287–1297.
Moerman, D.G., 1980. A genetic analysis of the unc-22 region in *Caenorhabditis elegans*(PhD Thesis) Simon Fraser University.
Moerman, D.G., Baillie, D.L., 1981. Formaldehyde mutagenesis in the nematode *Caenorhabditis elegans*. Mutat. Res. 80, 273–279.
Nakanishi, K., Ogiso, Y., Nakama, T., Fukai, S., Nureki, O., 2005. Structural basis for anticodon recognition by methionyl-tRNA synthetase. Nat. Struct. Mol. Biol. 12, 931–932.
Piano, F., Schetter, A.J., Morton, D.G., Gunsalus, K.C., Reinke, V., Kim, S.K., Kemphues, K.J., 2002. Gene clustering based on RNAi phenotypes of ovary-enriched genes in *C. elegans*. Curr. Biol. 12, 1959–1964.
Pierce, S.B., Chisholm, K.M., Lynch, E.D., Lee, M.K., Walsh, T., Opitz, J.M., Li, W., Klevit, R.E., King, M.C., 2011. Mutations in mitochondrial histidyl tRNA synthetase HARS2 cause ovarian dysgenesis and sensorineural hearing loss of Perrault syndrome. Proc. Natl. Acad. Sci. U. S. A. 108, 6543–6548.
Rogalski, T.M., Baillie, D.L., 1985. Genetic organization of the unc-22 IV gene and the adjacent region in *Caenorhabditis elegans*. Mol. Gen. Genet. 201, 409–414.
Rogalski, T.M., Moerman, D.G., Baillie, D.L., 1982. Essential genes and deficiencies in the unc-22 IV region of *Caenorhabditis elegans*. Genetics 102, 725–736.
Rozen, S., Skaletsky, H.J., 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz, S., M.S.e (Eds.), Methods in Molecular Biology. Humana Press, Totowa, NJ, pp. 365–386.
Schein, J.E., Marra, M.A., Benian, G.M., Fields, C., Baillie, D.L., 1993. The use of deficiencies to determine essential gene content in the let-56-unc-22 region of *Caenorhabditis elegans*. Genome 36, 1148–1156.
Simmer, F., Moorman, C., van der Linden, A.M., Kuijk, E., van den Berghe, P.V., Kamath, R.S., Fraser, A.G., Ahringer, J., Plasterk, R.H., 2003. Genome-wide RNAi of *C. elegans* using the hypersensitive rrf-3 strain reveals novel gene functions. PLoS Biol. 1, E12.
Sonnichsen, B., Koski, L.B., Walsh, A., Marschall, P., Neumann, B., Brehm, M., Alleaume, A.M., Artelt, J., Bettencourt, P., Cassin, E., Hewitson, M., Holz, C., Khan, M., Lazik, S., Martin, C., Nitzsche, B., Ruer, M., Stamford, J., Winzi, M., Heinkel, R., Roder, M., Finell, J., Hantsch, H., Jones, S.J., Jones, M., Piano, F., Gunsalus, K.C., Oegema, K., Gonczy, P., Coulson, A., Hyman, A.A., Echeverri, C.J., 2005. Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. Nature 434, 462–469.
Sugiura, I., Nureki, O., Ugaji-Yoshikawa, Y., Kuwabara, S., Shimada, A., Tateno, M., Lorber, B., Giege, R., Moras, D., Yokoyama, S., Konno, M., 2000. The 2.0 Å crystal structure of *Thermus thermophilus* methionyl-tRNA synthetase reveals two RNA-binding modules. Structure (London, England: 1993) 8, 197–208.

Thacker, C., Sheps, J.A., Rose, A.M., 2006. *Caenorhabditis elegans* dpy-5 is a cuticle procollagen processed by a proprotein convertase. Cell. Mol. Life Sci. 63, 1193–1204.
Walker, E.J., Treacy, G.B., Jeffrey, P.D., 1983. Molecular weights of mitochondrial and cytoplasmic aminoacyl-tRNA synthetases of beef liver and their complexes. Biochemistry 22, 1934–1941.
Wolf, Y.I., Aravind, L., Grishin, N.V., Koonin, E.V., 1999. Evolution of aminoacyl-tRNA synthetases—analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. Genome Res. 9, 689–710.