Development 137, 1799-1805 (2010) doi:10.1242/dev.046219

The Groucho ortholog UNC-37 interacts with the short Groucho-like protein LSY-22 to control developmental decisions in *C. elegans*

Eileen B. Flowers[1,*], Richard J. Poole[1,*], Baris Tursun[1], Enkelejda Bashllari[1], Itsik Pe'er[2] and Oliver Hobert[1,†]

SUMMARY

Transcriptional co-repressors of the Groucho/TLE family are important regulators of development in many species. A subset of Groucho/TLE family members that lack the C-terminal WD40 domains have been proposed to act as dominant-negative regulators of Groucho/TLE proteins, yet such a role has not been conclusively proven. Through a mutant screen for genes controlling a left/right asymmetric cell fate decision in the nervous system of the nematode *C. elegans*, we have retrieved loss-of-function alleles in two distinct loci that display identical phenotypes in neuronal fate specification and in other developmental contexts. Using the novel technology of whole-genome sequencing, we find that these loci encode the *C. elegans* ortholog of Groucho, UNC-37, and, surprisingly, a short Groucho-like protein, LSY-22, that is similar to truncated Groucho proteins in other species. Besides their phenotypic similarities, *unc-37* and *lsy-22* show genetic interactions and UNC-37 and LSY-22 proteins also physically bind to each other in vivo. Our findings suggest that rather than acting as negative regulators of Groucho, small Groucho-like proteins may promote Groucho function. We propose that Groucho-mediated gene regulatory events involve heteromeric complexes of distinct Groucho-like proteins.

KEY WORDS: *C. elegans*, Left/right asymmetry, Neuronal development, Transcriptional control

INTRODUCTION

Drosophila Groucho, *C. elegans* UNC-37 and their four vertebrate orthologs Tle1-4 (Grg1-4) are WD40 domain-containing co-repressors of gene expression (Buscarlet and Stifani, 2007; Chen and Courey, 2000; Gasperowicz and Otto, 2005). Co-repressors, which are recruited to DNA by a variety of transcription factors, recruit other proteins, such as histone deacetylases (HDACs), that modify chromatin and thereby control gene expression (Privalsky, 2001). Groucho/UNC-37 and their vertebrate orthologs can interact with a plethora of different DNA-binding transcription factors and have been implicated in a number of diverse biological processes, including neuronal pattern formation, hematopoiesis, osteopoiesis, muscle, gland and eye development and tumorigenesis (Buscarlet and Stifani, 2007; Chen and Courey, 2000; Gasperowicz and Otto, 2005).

Aside from the canonical, WD40 domain-containing family members, there are several unusual and little-studied variants of Groucho/TLE proteins. These proteins, which are sometimes referred to as short Groucho proteins, are either short splice isoforms of longer Groucho/TLE family members (e.g. Grg1-S or Grg3b) or are encoded by entirely distinct genetic loci [e.g. mouse or human AES (Grg5)] (Gasperowicz and Otto, 2005; Miyasaka et al., 1993). In either case, these short Groucho proteins lack the C-terminal WD40 domains that are involved in transcription factor interactions, but contain the so-called Q-rich and GP (glycine/proline-rich) N-terminal domains. Within canonical Groucho/TLE proteins, the Q-rich domain mediates oligomerization of Groucho/TLE family members, as well as interaction with other DNA-binding transcription factors, whereas the GP domain recruits other proteins, such as HDACs (Gasperowicz and Otto, 2005; Miyasaka et al., 1993). In contrast to full-length Groucho/TLE proteins, the biological and molecular function of short Groucho proteins is not well understood. In vitro evidence suggests that they function either as antagonists or agonists of normal Groucho proteins (Buscarlet and Stifani, 2007; Gasperowicz and Otto, 2005); this contradiction illustrates the need for genetic loss-of-function studies that address the precise biological role of short Groucho proteins and their relation to normal Groucho function. We report here the isolation and functional characterization of a short Groucho protein, LSY-22, in the nematode *C. elegans*.

MATERIALS AND METHODS

Strains and transgenes

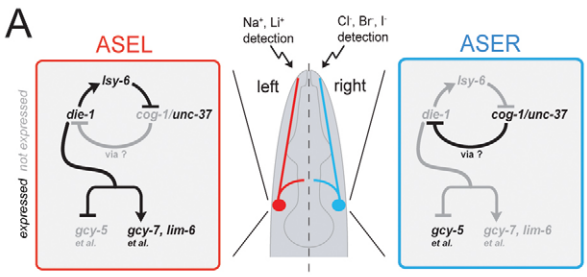
The following were used: N2 Bristol wild-type (Brenner, 1974); OH7410, *lsy-5(ot37) otIs114/hT2[bli-4(e937) let-?(q782) qIs48]* (I;III); OH9305, *lsy-5(ot240) otIs114*; OH7116, *lsy-22(ot114)/hT2[bli-4(e937) let-?(q782) qIs48]* (I;III); *otIs3*; OH7115, *lsy-22(ot244) otIs114/hT2[bli-4(e937) let-?(q782) qIs48]* (I;III) (Sarin et al., 2007); CB262, *unc-37(e262)*; OH9242, *unc-37(wd17wd22)/hT2[bli-4(e937) let-?(q782) qIs48]* (I;III) (Miller et al., 1993); OH9289, *lsy-22(tm3750)/hT2[bli-4(e937) let-?(q782) qIs48]* (I;III); OH1997, *otIs114; him-8; dpy-11 lsy-6(ot71)*; OH351, *otIs114; die-1(ot26)*; OH1583, *otIs114; cog-1(ot28),* DR293, *dpy-5(e61) unc-101(m1)*.

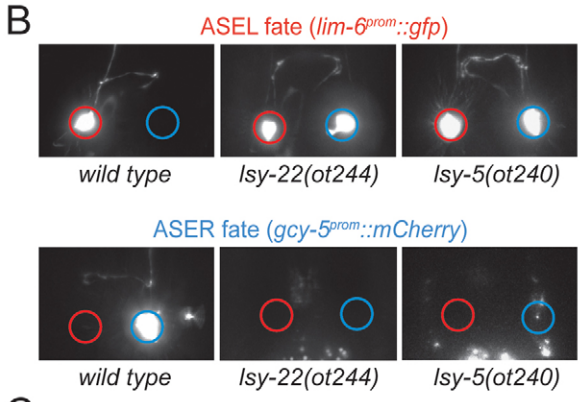
Transgenes to label ASEL and ASER fates included: *otIs3* V and *otIs4* X, Is[*gcy-7^{prom}::gfp; lin-15* (+)]; *otIs114* I, Is[*lim-6^{prom}::gfp; rol-6(d)*]; *otIs220* IV, Is[*gcy-5^{prom}::mCherry; rol-6(d)*]; *ntIs1* V, Is[*gcy-5^{prom}::gfp; lin-15 (+)*]; and the ASEL/ASER markers *otIs151* V, Is[*ceh-36^{prom}::DsRed2; rol-6(d)*] and *otIs232* V, Is[*che-1^{prom}::mCherry; rol-6(d)*]. Reporters for

[1]Department of Biochemistry and Molecular Biophysics, Howard Hughes Medical Institute, Columbia University Medical Center and [2]Department of Computer Science, Columbia University, New York, NY 10032, USA.

*These authors contributed equally to this work
†Author for correspondence (or38@columbia.edu)

Accepted 25 March 2010





C

Genotype	ASEL fate [1]	L	R	ASER fate [2]	L	R
Wild type	100% (n>100) *,**	●	○	100% (n>100) #,##	○	●
unc-37(ot240)	63% (n=256) *	●	●	69% (n=49) ##	○	○
lunc-37(ot37) [3]	49% (n=231) *	●	●	74% (n=112) #,5	○	○
unc-37(e262)	23% (n=52) *	●	●			
unc-37(wd17wd22) [3]	74% (n=50) **	●	●			
lsy-22(ot244) [3]	95% (n=59) *	●	●	73% (n=30) ##	○	○
lsy-22(ot114) [3]	70% (n=50) *	●	●			
lsy-22(tm3750) [3]	72% (n=50) **	●	●	44% (n=50) ##	○	○
lsy-6(ot71)	100% (n=160) *	○	○			
unc-37(ot37); lsy-6(ot71)	49% (n=49) *	●	●			
lsy-22(ot244); lsy-6(ot71) [3]	87% (n=52) *	●	●			
cog-1(ot28)	73% (n=75) *	●	●			
cog-1(ot28)/+	0% (n>50) *	●	●			
+/unc-37(e262)	0% (n>50) *	●	●			
+/lsy-22(ot244)	0% (n>50) *	●	●			
cog-1(ot28)/unc-37(e262)	40% (n=30) *	●	●			
cog-1(ot28)/lsy-22(ot244)	77% (n=13) *	●	●			
lsy-22(ot244)/unc-37(e262)	0% (n=81)	●	●			
cog-1(OE) [4]	56% (n=31) *	○	○			
unc-37(ot37); cog-1(OE)	66% (n=27) *	●	●			
lsy-22(ot244); cog-1(OE)	71% (n=34) *	●	●			
die-1(ot26)	100% (n=40) *	○	○			
unc-37(ot37); die-1(ot26) [3]	96% (n=26) *	○	○			
lsy-22(ot244); die-1(ot26) [3]	82% (n=51) *	○	○			

Fig. 1. Left/right asymmetry of the ASE neurons is controlled by a complex network of regulatory factors, including *lsy-5* and *lsy-22*. (**A**) Summary of the previously known factors controlling the laterality of the ASEL/ASER neurons in *C. elegans*. Red indicates ASEL fate and blue indicates ASER fate. A bistable feedback loop controls expression of the downstream terminal differentiation genes, of which only a selected subset is shown. The negative regulation of *die-1* expression by *cog-1* involves 3′UTR regulation of *die-1* (Chang et al., 2004) as well as uncharacterized regulatory interactions (Didiano et al., 2010). (**B**) Loss of *lsy-22* and *lsy-5/unc-37* show a similar effect on downstream *gfp* reporters, in which expression of the ASER fate marker *gcy-5::mCherry* is lost and the ASEL fate marker *lim-6::gfp* is ectopically expressed in ASER (the 'two ASEL' phenotype). See C for quantification of data. Red and blue circles indicate ASEL and ASER neurons, respectively. (**C**) Quantification of laterality defects in *lsy-22* and *lsy-5/unc-37* mutants. We refer to *lsy-5* as *unc-37* (see text). Shown is the percentage of animals within the population with a given phenotype. Circles represent ASEL and ASER and green shading indicates whether and where the respective fate marker is expressed. *lim-6* marker gene expression for alleles *ot244*, *ot114*, *ot37*, *ot240* and *e262* was rescored and found to be similar to that previously reported (Chang et al., 2003; Sarin et al., 2007). [1]Scored with *lim-6::gfp (otIs114)* (*) and/or *gcy-7::gfp (otIs3* or *otIs4)* (**). If scored with both, the *lim-6::gfp* data are shown. [2]Scored with *gcy-5::gfp (ntIs1)* (#) or *gcy-5::mCherry (otIs220)* (##). [3]Homozygous *unc-37* or *lsy-22* worms were scored from the heterozygous balanced strain *lsy-22* or *unc-37/hT2[bli-4(e937)let-?(q782)qIs48]*. [4]*cog-1(OE)* is an extrachromosomal array, *otEx4124* Ex[*ceh-36::cog-1, elt-2::gfp*], that ectopically expresses *cog-1* in ASEL and ASER. [5]Data from Sarin et al. (Sarin et al., 2007) and shown for comparison purposes only.

other cell types included: *bxIs14*, Is[*pkd-2::gfp*, *pha-1*] (Lints et al., 2004); *wdIs3*, Is[*del-1::gfp*] (Winnier et al., 1999); and *qIs19*, Is[*lag-2::gfp*] (Blelloch et al., 1999).

Transgenic arrays included: OH9296, *otEx4124*, Ex[*ceh-36::cog-1,elt-2*]; OH9297, *otEx4125*, Ex[*lsy-22^{FOS}::2×FLAG::venus, elt-2::rfp*]; *otIs232*; OH9557, *lsy-22(ot244) otIs114*, Ex[lsy-22^{FOS}, lsy-22$^{5'FOS}$, *elt-2::rfp*]; and OH9556, *otIs288* Is*[unc-37^{FOS}::yfp, elt-2::rfp] otIs232.*

Generation of reporter genes

unc-37 and *lsy-22* reporter genes were created utilizing λ-Red-mediated recombineering in bacteria as described (Tursun et al., 2009). Briefly, the *lsy-22*-containing fosmid (WRM0628cC04) and the *unc-37*-containing fosmid (WRM0612cG05) were electroporated into *E. coli* strain SW105 (Warming et al., 2005). Using FLP recombinase-removable *galK*-based cassettes, we inserted *yfp* immediately preceding the stop codon at the C-terminus of *lsy-22* or *unc-37*, resulting in a translational fusion. Recombineered fosmids were sequenced at their recombineered junctions and correct clones were maintained in *E. coli* strain EPI-300 T1R (Epicentre). Fosmids were digested with *Sbf*I and injected at 10 ng/μl, together with 2 ng/μl *Sca*I-digested *rol-6(d)* (pRF4; 2 ng/μl) or *Hin*dIII-digested *elt-2::NLS-dsRed* and *Pvu*II-digested bacterial genomic DNA (150 ng/μl) to generate a complex array. The DNA was injected into wild-type N2. The resulting array is called *otEx4125* (*lsy-22^{FOS}::2×FLAG::venus*) and *otEx4126* (*unc-37^{FOS}::yfp*). *otEx4126* subsequently spontaneously integrated to generate *otIs288*.

Mapping and whole-genome sequencing

lsy-22(ot114) was mapped by standard three-factor mapping between *dpy-5* (0 cM) and *unc-101* (13.3 cM) using strain DR293 *dpy-5(e61) unc-101(m1)*. Briefly, upon picking ~100 F2s, Lsy Dpy Unc animals were never obtained, whereas both Dpy Lsy non-Unc and Unc Lsy non-Dpy recombinants were found, suggesting that *lsy-22* lies between *dpy-5* and *unc-101*. This interval was analyzed for sequence variants by whole-genome sequencing (see below).

lsy-5(ot37) and *lsy-5(ot240)* were placed on chromosome I because they showed linkage to the integrated transgene *otIs114*, with which we scored ASEL fate. Further mapping of both alleles, as well as the *otIs114* reporter, using single-nucleotide polymorphisms in the CB4856 Hawaiian wild-type isolate revealed two regions of possible linkage on chromosome I, from –8.23 cM to –5.78 cM and from –1.50 cM to +3.01 cM. *lsy-5(ot240)* was chosen for whole-genome sequencing and both possible linkage regions were analyzed for the presence of sequence variants.

Genomic DNA from homozygous *lsy-22(ot114)* animals was prepared by picking 1000 progeny of heterozygous *ot114* animals with a Lsy phenotype and using the Gentra GenePure DNA extraction kit. Genomic DNA was prepared from *lsy-5(ot240)* from a homozygous population of viable animals. Genomic DNA (5 μg) prepared from either *lsy-5(ot240)* or *lsy-22(ot114)* was sequenced as previously described (Sarin et al., 2008).

ot240 was sequenced on an in-house Illumina Genome Analyzer II according to the manufacturer's specification and the resultant sequence data were analyzed with MAQGene (Bigelow et al., 2009). The following stringency criteria were used to filter variants: quality score ≥3, loci multiplicity ≤1, sequencing depth ≥3 (both orientations required). *ot114* data were generated by Illumina's sequencing service and data were analyzed using MAQ (as MAQGene was not available at that time). In both cases, all detected sequence variants were analyzed for whether they were present in any one of 14 additional genomes of different genotypes that we sequenced in the laboratory; if so, they were considered background and eliminated from further analysis.

Yeast two-hybrid analysis

Yeast two-hybrid experiments were performed according to the Matchmaker Gold Yeast Two-Hybrid System protocol (Clontech). cDNA was amplified from a total RNA population using primers for the *lsy-22*, *unc-37* or *cog-1* locus and cloned into the yeast vectors pGBKT7 and pGADT7. Vectors were co-transformed into the yeast strain AH109 and transformations were plated on SC-Leu-Trp media. Five transformants were randomly selected from each transformation and diluted in 100 µl of water to a uniform optical density. Then 10 µl of each dilution was spotted onto the following plates: SC-Leu-Trp, SC-Leu-Trp-His + 3-amino-1,2,4-triazole (3-AT), and SC-Ade-Leu-Trp-His + X-α-Gal. Plates were incubated for 4-5 days at 30°C and scored for growth and on X-α-Gal-containing plates for blue staining of the spotted yeast culture dilutions.

Co-immunoprecipitation

Packed worms [1.5 ml; either *otEx4125* (*lsy-22::FLAG::venus*) worms that express FLAG-tagged *lsy-22* or wild-type N2] were washed with M9 buffer and resuspended in 2 ml IP buffer [50 mM Hepes-KOH pH 7.6, 100 mM NaCl, 1 mM EDTA, 0.25 M LiCl, 1% sodium deoxycholate, 0.5% NP40, 10% glycerol and protease inhibitors (Sigma S8820)] and homogenized in a glass homogenizer on ice. After a 30-minute incubation on ice, lysates were centrifuged for 20 minutes at 15,000 ***g*** at 4°C to clear of debris. Each cleared lysate was incubated at 4°C for 4 hours with either 100 µl of a 50% slurry of anti-FLAG M2-agarose beads (Sigma A2220) in IP buffer or anti-HA HA-7-agarose beads (Sigma A2095) as negative control. By volume, 2.5% of each lysate was kept as input. The beads were washed three times with IP buffer and precipitates and lysates were subjected to immunoblotting using anti-UNC-37 [R-4677, a gift from D. Miller (Winnier et al., 1999)] or anti-FLAG (Sigma A8592) antibody.

RESULTS AND DISCUSSION

lsy-5 and *lsy-22* are required for left/right asymmetric cell fate specification in the nervous system

The ASE neurons of the nematode *C. elegans* are a pair of morphologically symmetric gustatory neurons that are functionally lateralized and express distinct chemoreceptors in the left (ASEL) versus right (ASER) neuron (Fig. 1A) (Hobert et al., 2002; Ortiz et al., 2009). Genetic screens have revealed a complex gene regulatory network composed of transcription factors and miRNAs that control ASE laterality (Chang et al., 2004; Chang et al., 2003; Hobert, 2006; Johnston and Hobert, 2003; Johnston et al., 2005; Sarin et al., 2007) (Fig. 1A). Two distinct, previously uncloned loci retrieved by these screens, *lsy-5* and *lsy-22*, displayed identical defects in their ability to execute the ASER fate, which instead converted to ASEL fate (Fig. 1B,C; note that *lsy-5* is referred to as *unc-37* in Fig. 1C, for reasons explained below).

The ASER fate-inducing activity of the *lsy-5* and *lsy-22* genes appears intimately linked as they not only share similar ASE phenotypes but also show similar epistatic relationships to previously described members of the regulatory network that control ASE laterality. That is, both mutants suppressed the phenotypic consequence (i.e. ASEL to ASER conversion) of the

A

	ot114	*ot240*
Average coverage	16.1x	21.4x
Size of genetically defined interval	7,076,277 bp (13 cM)	4,270,868 bp (6 cM)
Total variants in interval compared to wild-type reference genome	54	37
Total variants minus strain background variants *	33	27
Intergenic/intronic/silent	26/33	24/27
Splice junction/protein coding	7/33	3/27
Splice junction/protein coding confirmed by manual resequencing	7/7**	3/3

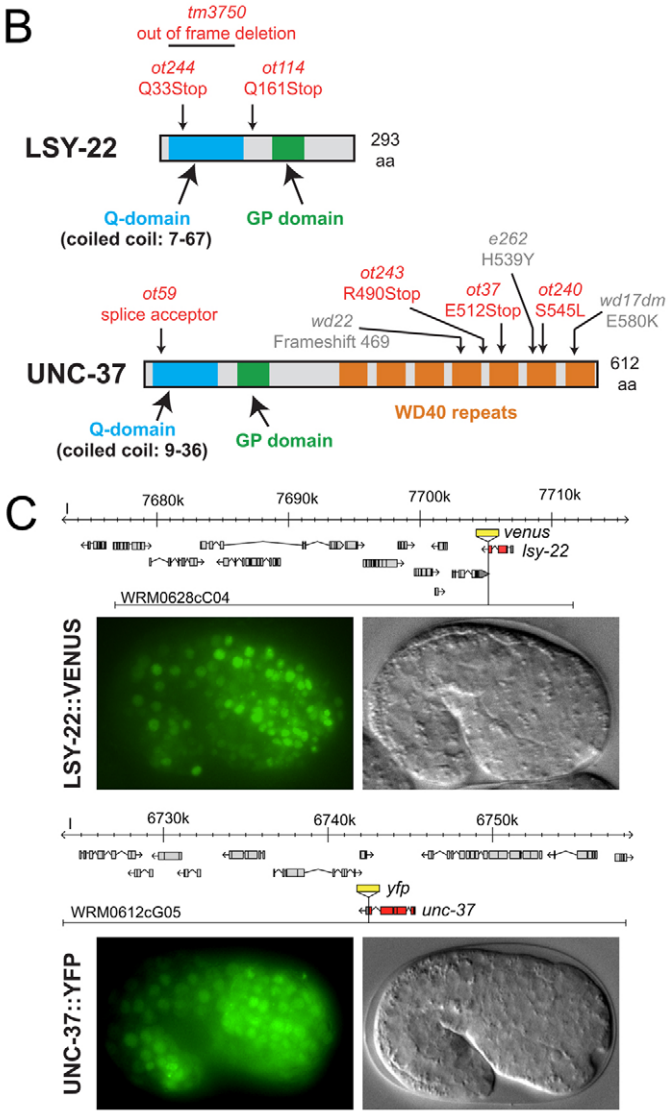


Fig. 2. Molecular identity of *lsy-22* and *lsy-5*. (**A**) Whole-genome sequencing (WGS) data for the *lsy-22(ot114)* and *lsy-5(ot240)* alleles. *, Variants were considered as background if they were also found in other WGS datasets from the laboratory (see Materials and methods). **, 1/7 was found to be a heterozygous variant. (**B**) LSY-22 and UNC-37, showing the location of the various mutant alleles. Some previously isolated *unc-37* alleles (Pflugrad et al., 1997) that we used here are indicated in gray text. Detailed sequence comparisons, including coiled-coil domains, are shown in Fig. S1 in the supplementary material. (**C**) *lsy-22::venus* and *unc-37::yfp* reporters generated by fosmid recombineering are broadly expressed during embryogenesis and larval development. Representative ~1.5-fold stage *C. elegans* embryos are shown.

loss of the ASEL fate-inducing miRNA encoded by *lsy-6* (Fig. 1C). Both genes also genetically interact in a similar manner with *cog-1*, which encodes an Nkx6.1-like, EH1 domain-containing homeodomain protein that is thought to directly bind UNC-37 (Chang et al., 2003). First, both mutants failed to complement *cog-1* (Fig. 1C); such failure to complement is often an indication that the encoded gene products act in the same complex (e.g. Hays et al., 1989). Second, *cog-1* requires both *lsy-5* and *lsy-22* to specify ASER fate. Forced bilateral expression of *cog-1* in both ASE cells resulted in ectopic ASER fate induction in ASEL (Fig. 1C). This ability of *cog-1* to induce ASER fate was abrogated in *lsy-5* and *lsy-22* mutants (Fig. 1C), demonstrating that COG-1 requires both genes to fulfil its function as an ASER fate inducer. Lastly, both *lsy-5* and *lsy-22* act upstream of the ASEL fate inducer *die-1*, as the 'two ASEL fate' mutant phenotype observed in *lsy-5* and *lsy-22* mutants was suppressed in *die-1* mutants (Fig. 1C). We conclude that *lsy-5* and *lsy-22* affect the bistable ASE cell fate decision in a similar manner.

Molecular identity of *lsy-5* and *lsy-22*

We mapped *lsy-5* and *lsy-22*, each represented by two independently isolated alleles, to genetic intervals on chromosome I and then used the recently introduced 'gene cloning' technology of whole-genome sequencing (Sarin et al., 2008) for one mutant allele of each to identify the molecular nature of the affected genes. The genome sequence data are summarized in Fig. 2A. This approach identified a small number of protein-coding sequence variants in the respective genetic intervals, including a premature stop codon in *F27D4.2* of *lsy-22(ot114)* animals and a missense mutation in *W02D3.9* of *lsy-5(ot240)* animals (Fig. 2A,B). These are the phenotype-causing variants based on the following evidence: (1) the second allele of each locus showed a mutation in the same, respective gene (Fig. 2B); (2) additional, independently isolated alleles of *F27D4.2* and *W02D3.9* resulted in similar ASE laterality defects (Fig. 1C); (3) RNAi against *F27D4.2* and *W02D3.9* reproduced both the lethality and ASE phenotypes of the strongest alleles of each locus (see Table S1 in the supplementary material); and (4) mutant alleles of either locus could be rescued by wild-type genomic copies of the respective genes (see Table S1 in the supplementary material). We conclude that *F27D4.2* is *lsy-22* and *W02D3.9* is *lsy-5*.

W02D3.9/lsy-5 encodes the previously described Groucho ortholog UNC-37 (from here on we refer to *lsy-5* as *unc-37*) (Fig. 2B). *unc-37* was previously implicated in controlling ASE laterality through its interaction with the EH1 domain-containing homeobox factor COG-1 (Chang et al., 2003); we had not considered *unc-37* as a candidate for *lsy-5* as the *ot37* allele that we isolated complemented the canonical *unc-37* allele *e262* (see Table S2 in

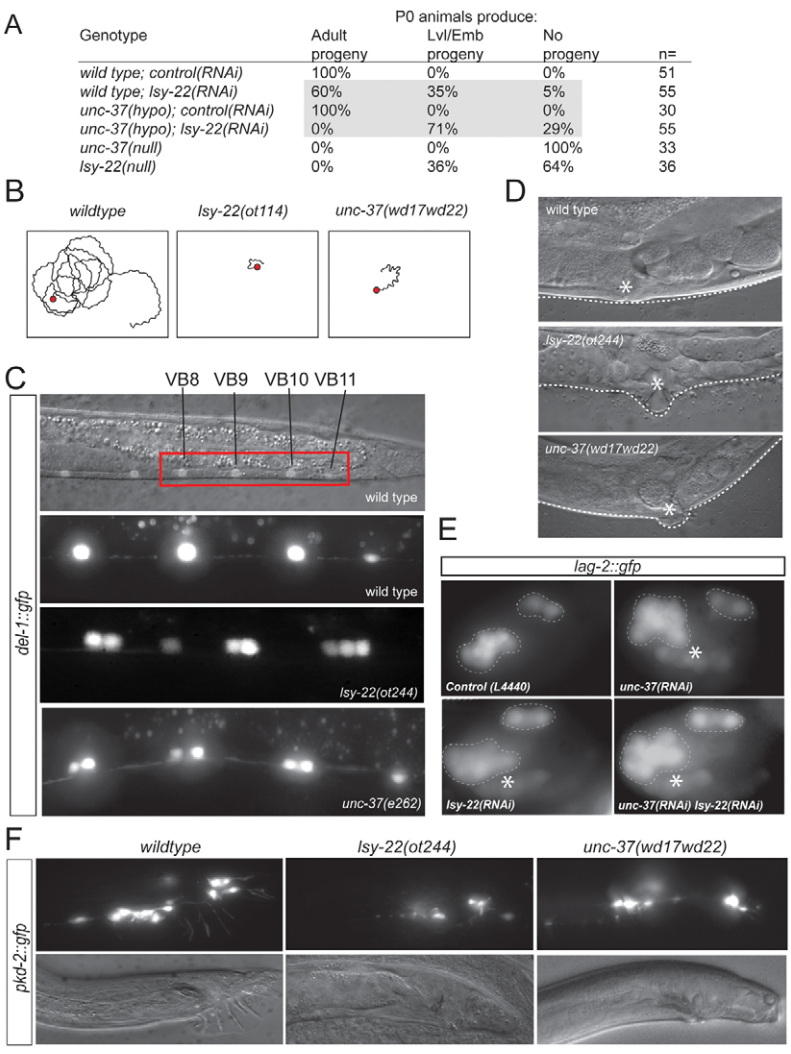


Fig. 3. *lsy-22* and *unc-37* control similar developmental processes. (**A**) *lsy-22* synergizes with *unc-37* for sterility/lethality. *unc-37(hypo)* is the hypomorphic mutant allele *e262*, *unc-37(null)* is the null allele *wd17wd22* and *lsy-22(null)* is the null allele *ot244*. Null mutant animals that were maternally rescued for viability were picked based on their Lsy phenotype and their progeny scored for viability. For more information on RNAi, see Table S1 in the supplementary material. Control RNAi refers to the L4440 vector clone. (**B**) Unc phenotypes of *lsy-22* and *lsy-5/unc-37* (observed in 50/50 animals of each genotype; 0/50 wild-type animals show comparable locomotory tracks). Single homozygous adults were placed in the middle of an *E. coli* lawn and observed after 5 minutes. The red circle is where the animal was originally placed and tracks were drawn over a photo of the animal's tracks as an estimation of its path. One representative example is shown. (**C**) *lsy-22* and *unc-37* mutants show ectopic expression of the *wdIs3* (*del-1::gfp*) reporter in the presumptive VA motoneurons [21/21 animals for *lsy-22(ot244)* and 37/40 animals for *unc-37(e262)*]. The reporter is expressed only in the VB motoneurons in wild-type animals at the L2 stage. (**D**) *unc-37* and *lsy-22* animals display protruding vulva (Pvl) phenotypes (observed in 50/50 animals of each mutant genotype). The ventral surface is indicated by a dotted line and the position of the vulva is indicated with an asterisk. (**E**) Removal of maternal *lsy-22* and *unc-37* by RNAi results in ectopic expression of the *qIs19(lag-2::gfp)* reporter in embryos. An asterisk indicates ectopic expression in certain AB descendants; dotted lines encircle endogenous expression in AB and MS descendants. Images are lateral views, anterior to the left. Percentages of ectopic expression observed were: *control (L4440)*, 4% (*n*=28); *unc-37(RNAi)*, 97% (*n*=28); *lsy-22(RNAi)*, 81% (*n*=26); *unc-37(RNAi) lsy-22(RNAi)*, 95% (*n*=21). (**F**) *pkd-2::gfp* (*bxIs14*) reveals missing rays in both *unc-37* and *lsy-22* maternally rescued null mutant males [20/20 and 9/9 for *lsy-22(ot244)* and *unc-37(wd17wd22)*, respectively].

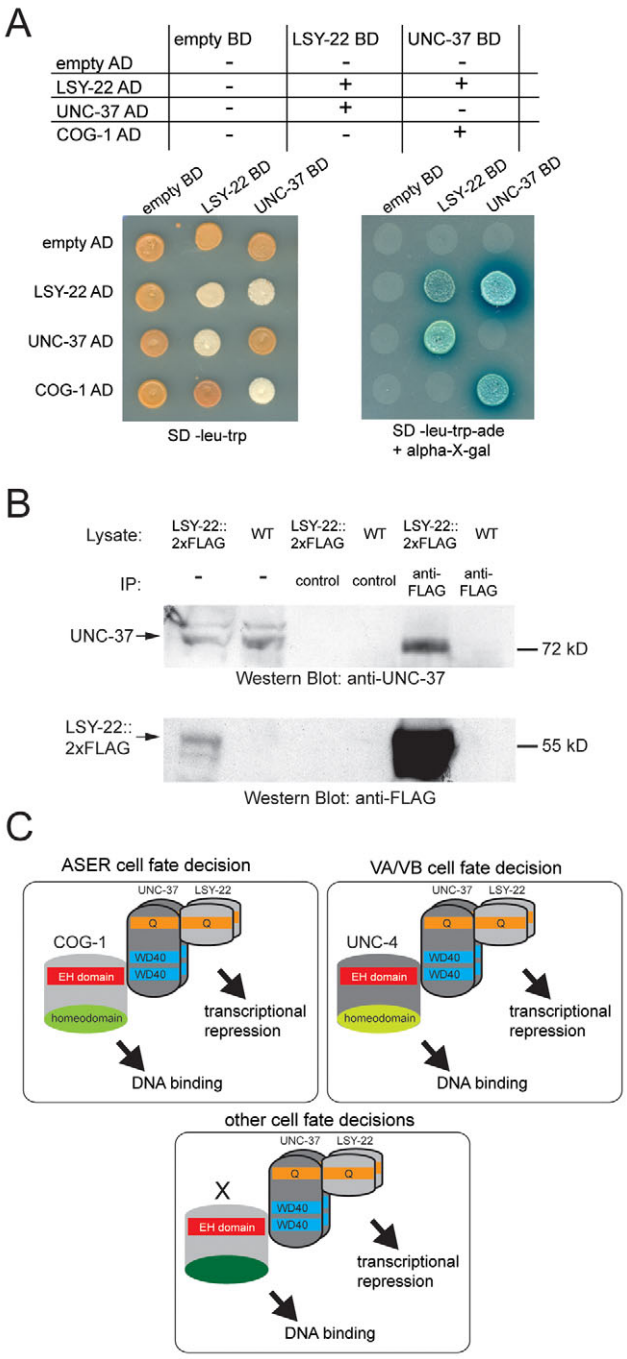


Fig. 4. LSY-22 physically interacts with UNC-37. (**A**) Yeast two-hybrid assay with UNC-37 and LSY-22 fused to the GAL4-activating domain (AD) or DNA-binding domain (BD) shows interaction of these two proteins. Additionally, the assay indicates that LSY-22 is able to interact with itself, and COG-1 is able to interact with UNC-37. (**B**) Co-immunoprecipitation of UNC-37 and LSY-22 from transgenic *C. elegans* expressing 2×FLAG-VENUS epitope-tagged LSY-22 (*otEx4125*). The lysate column indicates whether lysates were prepared from wild-type animals (WT) or from *otEx4125* animals. Control refers to beads with non-specific antibody. Endogenous UNC-37 was detected using an anti-UNC-37 antibody (Winnier et al., 1999). (**C**) Model for UNC-37–LSY-22 interaction. Given the reported Q-rich domain-mediated oligomerization of UNC-37/Groucho in other species, we suggest that the interaction occurs via their respective Q-rich domains, but interaction via other domains, such as the GP domain, is conceivable as well. This might occur in a tetrameric configuration, as previously reported for Groucho (Chen et al., 1998). Our data indicating a lack of interaction between UNC-37 and itself suggest that LSY-22 may be required for UNC-37 co-repressor complex formation. The interaction of UNC-37 and COG-1 might occur via a WD40/EH1 domain interaction, as suggested by crystallographic studies of such domains (Pickles et al., 2002). We propose that a broadly expressed LSY-22–UNC-37 complex interacts in a tissue-specific manner with specific DNA-binding transcription factors to control cell fate decisions; in ASE, this transcription factor is COG-1, in the VA motoneurons it might be the EH1 domain-containing UNC-4 protein (Winnier et al., 1999), and in other cellular contexts it might be any one of the dozens of EH1 domain-containing proteins (marked X) present in *C. elegans* (Copley, 2005) (or any other transcription factor interacting with UNC-37 in a non-EH1 domain-dependent manner).

the supplementary material). Allelic complementation is a relatively rare phenomenon, but is known to occur in multimeric multidomain proteins (e.g. Griffin and Chan, 2006).

F27D4.2/lsy-22 encodes a small, previously uncharacterized protein. Close sequence inspection reveals similarity to the N-terminus of UNC-37 (see Fig. S1 in the supplementary material); both proteins contain a Q-rich domain and a GP domain, yet only UNC-37, not LSY-22, contains C-terminal WD40 domains (Fig. 2B). Like other Groucho-related proteins, LSY-22 is predicted to contain two amphipathic alpha-helices within the Q-rich domain, the first of which contains a coiled-coil region (see Fig. S1 in the supplementary material). The Q-rich domain of long and short Groucho proteins has been shown to mediate their multimerization, whereas the GP domain is thought to mediate interactions with HDACs (Chen et al., 1998; Pinto and Lobe, 1996). Overall, LSY-22 is therefore similar to a small number of invertebrate and vertebrate Groucho-related proteins that also share similarity to the N-terminus of Groucho, but lack the C-terminal WD40 domains (AES proteins and shorter splice forms of Groucho/TLE family members) (Buscarlet and Stifani, 2007). No other obvious AES-related proteins were identified by sequence similarity in the *C. elegans* genome.

lsy-22 and *unc-37* are expressed in many cell types

To analyze their expression, we recombineered a fluorescent reporter gene into ~40 kb fosmids that contain the *lsy-22* and *unc-37* genomic loci (Fig. 2C). Each fosmid reporter rescued several of the mutant phenotypes of *lsy-22* and *unc-37* (data not shown). *lsy-22* was very broadly expressed in the nucleus of most, if not all, cells tested, starting in early embryogenesis and persisting throughout larval and adult life (Fig. 2C). Consistent with previously described antibody staining (Pflugrad et al., 1997), a fosmid-recombineered *unc-37* reporter showed a similar, broad expression pattern in apparently all cells at all times (Fig. 2C). Such broad expression is characteristic of Groucho/TLE family members in other organisms as well, including short Groucho proteins such as AES (Mallo et al., 1993).

Phenotypic similarities of *unc-37* and *lsy-22*

In addition to their similar ASE laterality defect and indistinguishable epistatic relationship to other ASE fate regulators (Fig. 1C), *unc-37* and *lsy-22* null mutants showed a striking set of phenotypic similarities. First, *lsy-22* and *unc-37* null mutants

displayed a maternally rescued sterility (Ste) and embryonic lethality (Emb) phenotype and partial removal of both genes together revealed a genetic interaction (Fig. 3A). Second, viable homozygous offspring of *lsy-22* and *unc-37* heterozygous parents not only displayed a Lsy phenotype, but also uncoordinated locomotion (Fig. 3B). As in *unc-37* mutants (Miller et al., 1993; Pflugrad et al., 1997), the *lsy-22* phenotype is caused by the misspecification of a specific class of ventral cord motoneurons (Fig. 3C). Third, both *unc-37* and *lsy-22* mutants display morphological defects in the vulva (Pvl phenotype) (Fig. 3D). Fourth, removal of maternal and zygotic *unc-37* through RNAi results in ectopic expression of the Notch ligand *lag-2* during embryogenesis (Neves and Priess, 2005); the same phenotype was observed upon removal of *lsy-22* function (Fig. 3E). Finally, in both *unc-37* and *lsy-22* mutants, sensory ray structures in the male tail were transformed into hypodermal cells (Pal phenotype) (Zhang and Emmons, 2002) (Fig. 3F). In summary, in addition to their shared ASE laterality defects and broad nuclear co-expression, *unc-37* and *lsy-22* null mutants share a significant series of phenotypes and therefore function in similar biological processes.

UNC-37 and LSY-22 proteins physically interact

The spectrum of phenotypic similarities of *unc-37* and *lsy-22*, their genetic interactions and their nuclear co-expression in many cell types suggest that both proteins might physically interact. Groucho/TLE family proteins are thought to homomultimerize via the N-terminal Q-rich domain and it has been proposed that the short Groucho family members might engage in heteromeric interactions with full-length Groucho/TLE proteins (Pinto and Lobe, 1996; Ren et al., 1999). We tested for an UNC-37–LSY-22 interaction in two different settings. First, we undertook a yeast two-hybrid analysis and indeed found that the two proteins interact with one another (Fig. 4A). In addition to interacting with UNC-37, our yeast two-hybrid experiment indicated that LSY-22 is able to interact with itself. The two-hybrid data also revealed a COG-1–UNC-37 interaction (Fig. 4A), corroborating their genetic interaction data (Fig. 1C) (Chang et al., 2003). However, we did not find two-hybrid evidence for UNC-37 homodimers; this is not because the UNC-37 prey or bait was not functional, as both interacted with LSY-22. This raises the intriguing possibility that *C. elegans* UNC-37 does not homomultimerize, but rather requires LSY-22 to form a functional heteromeric complex. Second, we investigated the interaction between LSY-22 and UNC-37 in vivo by co-immunoprecipitation from worm total lysates and found that endogenous UNC-37 protein is in a complex with LSY-22 (Fig. 4B). Taken together, our data suggest that COG-1, UNC-37 and LSY-22 exist in a complex that controls ASE cell fate (Fig. 4C).

Conclusions

Our genetic and physical interaction tests indicate that LSY-22 and UNC-37 act together to control a number of distinct developmental processes. We propose that UNC-37 does not exist in a homomultimeric state, but rather in a heteromultimeric complex with LSY-22, possibly, although not necessarily, mediated through Q-rich domain interactions. Our study therefore provides a novel perspective on the small Groucho proteins that were previously thought to antagonize canonical Groucho proteins. Such antagonistic relationships were exclusively inferred from overexpression studies that often required significant amounts of exogenously added short Groucho proteins (Buscarlet and Stifani, 2007). We also find that overexpression of *lsy-22* can mimic certain loss-of-function phenotypes of *unc-37*. *lsy-22*-overexpressing animals display a mild ASE laterality defect, similar to that observed in *lsy*-22 mutants (see Table S1 in the supplementary material). Yet, as our loss-of-function analysis clearly shows, this mild overexpression effect is not indicative of the endogenous function of *lsy-22*. Rather, LSY-22 appears to promote, not antagonize, UNC-37 function, and it might do so in what we propose to be a novel co-repressor complex. The stoichiometry of this complex might be tightly regulated, such that overexpression of LSY-22 disrupts its formation. Taken together, these findings should prompt a careful examination of the composition of vertebrate Groucho/TLE-containing co-repressor complexes to better understand the important roles that they play in development and disease.

Acknowledgements

We thank Q. Chen for expert DNA injection; A. Boyanov for expert handling of the Illumina Genome Analyzer; Illumina for providing the *ot114* genome sequence data; S. Mitani at Tokyo Women's Medical University School of Medicine for the *lsy-22(tm3750)* allele; D. Miller for advice and reagents; and members of the O.H. laboratory for discussion and comments on the manuscript. We acknowledge funding by the NIH (R01NS039996-05; R01NS050266-03 to O.H.) and the NSF (NSF-CISE 0829882 to I.P.). B.T. is funded by a Francis Goelet Postdoctoral Fellowship. R.J.P. is a Charles H. Revson Senior Fellow in Biomedical Science. O.H. is an Investigator of the HHMI. Deposited in PMC for release after 6 months.

Competing interests statement

The authors declare no competing financial interests.

Supplementary material

Supplementary material for this article is available at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.046219/-/DC1

References

Bigelow, H., Doitsidou, M., Sarin, S. and Hobert, O. (2009). MAQGene: software to facilitate C. elegans mutant genome sequence analysis. *Nat. Methods* **6**, 549.

Blelloch, R., Anna-Arriola, S. S., Gao, D., Li, Y., Hodgkin, J. and Kimble, J. (1999). The gon-1 gene is required for gonadal morphogenesis in Caenorhabditis elegans. *Dev. Biol.* **216**, 382-393.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. *Genetics* **77**, 71-94.

Buscarlet, M. and Stifani, S. (2007). The 'Marx' of Groucho on development and disease. *Trends Cell Biol.* **17**, 353-361.

Chang, S., Johnston, R. J., Jr and Hobert, O. (2003). A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of C. elegans. *Genes Dev.* **17**, 2123-2137.

Chang, S., Johnston, R. J., Frokjaer-Jensen, C., Lockery, S. and Hobert, O. (2004). MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode. *Nature* **430**, 785-789.

Chen, G. and Courey, A. J. (2000). Groucho/TLE family proteins and transcriptional repression. *Gene* **249**, 1-16.

Chen, G., Nguyen, P. H. and Courey, A. J. (1998). A role for Groucho tetramerization in transcriptional repression. *Mol. Cell. Biol.* **18**, 7259-7268.

Copley, R. R. (2005). The EH1 motif in metazoan transcription factors. *BMC Genomics* **6**, 169.

Didiano, D., Cochella, L., Tursun, B. and Hobert, O. (2010). Neuron-type specific regulation of a 3'UTR through redundant and combinatorially acting cis-regulatory elements. *RNA* **16**, 349-363.

Gasperowicz, M. and Otto, F. (2005). Mammalian Groucho homologs: redundancy or specificity? *J. Cell. Biochem.* **95**, 670-687.

Griffin, E. E. and Chan, D. C. (2006). Domain interactions within Fzo1 oligomers are essential for mitochondrial fusion. *J. Biol. Chem.* **281**, 16599-16606.

Hays, T. S., Deuring, R., Robertson, B., Prout, M. and Fuller, M. T. (1989). Interacting proteins identified by genetic interactions: a missense mutation in alpha-tubulin fails to complement alleles of the testis-specific beta-tubulin gene of Drosophila melanogaster. *Mol. Cell. Biol.* **9**, 875-884.

Hobert, O. (2006). Architecture of a MicroRNA-controlled gene regulatory network that diversifies neuronal cell fates. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 181-188.

Hobert, O., Johnston, R. J., Jr and Chang, S. (2002). Left-right asymmetry in the nervous system: the Caenorhabditis elegans model. *Nat. Rev. Neurosci.* **3**, 629-640.

Johnston, R. J. and Hobert, O. (2003). A microRNA controlling left/right neuronal asymmetry in Caenorhabditis elegans. *Nature* **426**, 845-849.

Johnston, R. J., Jr, Chang, S., Etchberger, J. F., Ortiz, C. O. and Hobert, O. (2005). MicroRNAs acting in a double-negative feedback loop to control a neuronal cell fate decision. *Proc. Natl. Acad. Sci. USA* **102**, 12449-12454.
Lints, R., Jia, L., Kim, K., Li, C. and Emmons, S. W. (2004). Axial patterning of C. elegans male sensilla identities by selector genes. *Dev. Biol.* **269**, 137-151.
Mallo, M., Franco del Amo, F. and Gridley, T. (1993). Cloning and developmental expression of Grg, a mouse gene related to the groucho transcript of the Drosophila Enhancer of split complex. *Mech. Dev.* **42**, 67-76.
Miller, D. M., 3rd, Niemeyer, C. J. and Chitkara, P. (1993). Dominant unc-37 mutations suppress the movement defect of a homeodomain mutation in unc-4, a neural specificity gene in Caenorhabditis elegans. *Genetics* **135**, 741-753.
Miyasaka, H., Choudhury, B. K., Hou, E. W. and Li, S. S. (1993). Molecular cloning and expression of mouse and human cDNA encoding AES and ESG proteins with strong similarity to Drosophila enhancer of split groucho protein. *Eur. J. Biochem.* **216**, 343-352.
Neves, A. and Priess, J. R. (2005). The REF-1 family of bHLH transcription factors pattern C. elegans embryos through Notch-dependent and Notch-independent pathways. *Dev. Cell* **8**, 867-879.
Ortiz, C. O., Faumont, S., Takayama, J., Ahmed, H. K., Goldsmith, A. D., Pocock, R., McCormick, K. E., Kunimoto, H., Iino, Y., Lockery, S. et al. (2009). Lateralized gustatory behavior of C. elegans is controlled by specific receptor-type guanylyl cyclases. *Curr. Biol.* **19**, 996-1004.
Pflugrad, A., Meir, J. Y., Barnes, T. M. and Miller, D. M., 3rd (1997). The Groucho-like transcription factor UNC-37 functions with the neural specificity gene unc-4 to govern motor neuron identity in C. elegans. *Development* **124**, 1699-1709.
Pickles, L. M., Roe, S. M., Hemingway, E. J., Stifani, S. and Pearl, L. H. (2002). Crystal structure of the C-terminal WD40 repeat domain of the human Groucho/TLE1 transcriptional corepressor. *Structure* **10**, 751-761.
Pinto, M. and Lobe, C. G. (1996). Products of the grg (Groucho-related gene) family can dimerize through the amino-terminal Q domain. *J. Biol. Chem.* **271**, 33026-33031.
Privalsky, M. L. (2001). *Transcriptional Corepressors: Mediators of Eukaryotic Gene Repression.* Heidelberg: Springer.
Ren, B., Chee, K. J., Kim, T. H. and Maniatis, T. (1999). PRDI-BF1/Blimp-1 repression is mediated by corepressors of the Groucho family of proteins. *Genes Dev.* **13**, 125-137.
Sarin, S., O'Meara, M, M., Flowers, E. B., Antonio, C., Poole, R. J., Didiano, D., Johnston, R. J., Jr, Chang, S., Narula, S. and Hobert, O. (2007). Genetic screens for Caenorhabditis elegans mutants defective in left/right asymmetric neuronal fate specification. *Genetics* **176**, 2109-2130.
Sarin, S., Prabhu, S., O'Meara, M. M., Pe'er, I. and Hobert, O. (2008). Caenorhabditis elegans mutant allele identification by whole-genome sequencing. *Nat. Methods* **5**, 865-867.
Tursun, B., Cochella, L., Carrera, I. and Hobert, O. (2009). A toolkit and robust pipeline for the generation of fosmid-based reporter genes in C. elegans. *PLoS ONE* **4**, e4625.
Warming, S., Costantino, N., Court, D. L., Jenkins, N. A. and Copeland, N. G. (2005). Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res.* **33**, e36.
Winnier, A. R., Meir, J. Y., Ross, J. M., Tavernarakis, N., Driscoll, M., Ishihara, T., Katsura, I. and Miller, D. M., 3rd (1999). UNC-4/UNC-37-dependent repression of motor neuron-specific genes controls synaptic choice in Caenorhabditis elegans. *Genes Dev.* **13**, 2774-2786.
Zhang, H. and Emmons, S. W. (2002). Caenorhabditis elegans unc-37/groucho interacts genetically with components of the transcriptional mediator complex. *Genetics* **160**, 799-803.

Table S1. Rescue, RNAi and overexpression of *lsy-5* and *lsy-22*

Genotype/strain	Viable progeny	Two ASEL phenotype*
Wild type	Yes	0 (>100)
Rescue		
unc-37(ot240)	Yes	63 (256)[†]
unc-37(wd17wd22)	No	74 (50)[†]
unc-37(ot240) +Ex[unc-37^{FOS}][¶,**]	Yes	2.8 (143)
lsy-22(ot244)	No	95 (59)[†]
lsy-22(ot244) +Ex[F27D4.2^{FOS}][¶,††]	Yes	40 (72)
lsy-22(ot244) +Ex[F27D4.2^{FOS}][¶,‡‡]	Yes	62 (143)
lsy-22(ot244) +Ex[F27D4.2^{FOS}][¶,‡‡]	Yes	50 (21)
RNAi phenocopy		
RNAi(empty vector)	Yes	0 (50)
RNAi(W02D3.9)	No	52 (65)[‡]
RNAi(F27D4.2)	Yes	51 (97)
nre-1 lin-15b; RNAi(empty vector)	Yes	0 (>100)
nre-1 lin-15b; RNAi(W03D3.9)	No	n.s.[§]
nre-1 lin-15b; RNAi(F27D4.2)	No	n.s.[§]
Overexpression		
+Ex[hsp16-2::lsy-22][¶,§§]	n.d.	3 (92)
–Ex[hsp16-2::lsy-22][¶,§§]	n.d.	0 (106)
+Ex[hsp16-2::che-1][¶,¶¶]	n.d.	100 (>100)***

For RNAi experiments, animals were fed on the respective dsRNA-producing bacteria and assessed for their ability to produce viable, fertile brood. In all experiments, animals were scored at the L3 larval-young adult stage. RNAi in a non-sensitized background allows scoring of the 'late' (Lsy) phenotype, whereas RNAi in the *nre-1; lin-15* sensitized background shows the 'early' (Emb) phenotype.

*Values indicate the percentage, with *n* in parentheses. Examined using the *gfp* reporter for ASEL *otIs114 (lim-6::gfp)* or *otIs4 (gcy-7::gfp)*.

[†]Data from Fig. 1C, this paper.

[‡]Scored from Lvl escapers.

[§]Not scorable owing to lethality.

[¶]'+Ex' refers to array-carrying worms and '–Ex' refers to non-array-carrying worms.

**The array contains the *unc-37(W03D3.9)* locus in the context of fosmid WRM0612cG05. We were unable to produce homozygous *unc-37(ot240)* in this strain owing to lethality and therefore scored the complete brood of an *ot240*/+ heterozygous parent. Only 1/4 of this brood is homozygous for *ot240*, so the maximum expected phenotype is 25%; we observed a 14.3% two ASEL phenotype in non-array-carrying worms. Z-test analysis indicates that the reduction to 2.8% in array-carrying worms is statistically significant ($P<0.005$).

[††]The array contains the *lsy-22(F27D4.2)* locus in the context of two fosmids, WRM0628cC04 and WRM0628aF08.

[‡‡]These two independent arrays each contain the *lsy-22(F27D4.2)* locus in the context of fosmid WRM0619c09.

[§§]The array contains *hsp16-2* driving a *lsy-22* cDNA.

[¶¶]The array contains *hsp16-2* driving a *che-1* genomic fragment known to induce ASE cell fate and serves as a control for activity of the heat-shock promoter in the ASE neurons.

***Expression is also induced in many other cells besides ASEL/ASER

Table S2. Allelic complementation of *unc-37* alleles

Genotype	ASEL fate marker (*lim-6*) ectopic in ASER	ASER fate marker (*gcy-5*) lost in ASER
Wild type	0 (100)	0 (>100)
ot240	63 (256)	69 (49)
*ot37**	62 (229)	74 (112)
e262	23 (52)	32 (122)
ot240/+	0 (>50)	n.d.
ot37/+	0 (>50)	n.d.
ot262/+	0 (>50)	n.d.
ot240/*e262*†	n.d.	22 (46)
ot37/*e262*†	0 (24)	n.d.
ot37/*ot240*	26 (31)	n.d.

This table explains why we did not initially consider *ot37* and *ot240* [previously called *lsy-5* (Sarin et al., 2007)] to be alleles of *unc-37*. We found that *ot37* and *ot240* failed to complement one another [see data in table; taken from Sarin et al. (Sarin et al., 2007)] and that *ot37* complements the canonical *unc-37(e262)* allele (see data in table; this complementation test was previously performed by Sarin et al. and was repeated with the new numbers and reporter shown here). After we identified *ot37* and *ot240* as *unc-37* alleles, we performed a complementation test of *ot240* with *e262*; these alleles fail to complement (as shown in table). Values show percentage (with *n* in parentheses) of animals in the population with a given phenotype. The alleles *ot37* and *ot240* were previously reported for the *lim-6* marker (Sarin et al., 2007) and are the same as shown in Fig. 1. Fate markers for ASEL are *otIs114 (lim-6::gfp)*, *otIs4* and *otIs3 (gcy-7::gfp),* and fate markers for ASER are *ntIs1 (gcy-5::gfp)* and *otIs220 (gcy-5::mCherry).* The data for the ASE phenotype of the three *unc-37* alleles alone are the same as shown in Fig. 1C, with the exception of the ASER fate marker for *unc-37(e262)*, which was previously reported by Chang et al. (Chang et al., 2003).

*Homozygous *unc-37* worms were scored from the heterozygous balanced strain *unc-37/hT2[bli-4(e937)let-?(q782)qIs48].*

†Scored from a cross in which males were heterozygous for one of the mutations.