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

Eileen B. Flowers1,*, Richard J. Poole1,*, Baris Tursun1, Enkeleida Bashllari1, Itsik Pe‘er2 and Oliver Hobert1,†

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, lssy-5(ot37) otls114/hT2[bli-4(e937) let-? (q782) qIs48] (I;III); OH9305, lssy-5(ot240) otls114; OH7116, lssy-22(ot11) otls144/hT2[bli-4(e937) let-? (q782) qls48] (I;III); otls3; OH7115, lssy-22(ot244) otls114/hT2[bli-4(e937) let-? (q782) qls48] (I;III) (Sarin et al., 2007); CB262, unc-37(e262); OH9242, unc-37(evd122)/hT2[bli-4(e937) let-? (q782) qls48] (I;III); OH9289, lssy-22(ot37) otls114/hT2[bli-4(e937) let-? (q782) qls48] (I;III); OH9997, otls114; him-8; dpy-11 lssy-6(ot71); OH351, otls114; die-1(ot26); OH1583, otls114; cog-1(ot28); DR293, dpy-5(e61) unc-101(m1).

Transgenes to label ASEL and ASER markers included: otls3 V and otls4 X, lsgcy-7prom::gfp; lin-15 (+); otls114 I, lsgcy-7prom::gfp; rol-6(d); otls220 IV, lsgcy-7prom::mCherry; rol-6(d); nds1 V, lsgcy-7prom::gfp; lin-15 (+); and the ASEL/ASER markers otls151 V, lsgcy-36prom::DsRed2; rol-6(d)] and otls232 V, lsgcy-7prom::mCherry; rol-6(d)]. Reporters for

1Department of Biochemistry and Molecular Biophysics, Howard Hughes Medical Institute, Columbia University Medical Center and 2Department 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

© 2010. Published by The Company of Biologists Ltd

Development 137 (11) 1800

**RESEARCH REPORT**

**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 ASE/ASER neurons in *C. elegans*. Red indicates ASE 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 ASE fate marker *lim-6::gfp* is ectopically expressed in ASER (the ‘two ASE’ phenotype). See C for quantification of data. Red and blue circles indicate ASE 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 ASE and ASER and green shading indicates whether and where the respective fate marker is expressed. *lim-6* marker gene expression for alleles *ot244, ot114, ot240* and *e262* was rescored and found to be similar to that previously reported (Chang et al., 2003; Sarin et al., 2007).

**Genotype** | **ASEL fate** | **ASER fate**
--- | --- | ---
Wild type | 100% (n=100) | 100% (n=100)
unc-37(ot240) | 0% (n=356) | 65% (n=46) 65% (n=112)
lim-37(ot37) | 0% (n=331) | 74% (n=112) 74% (n=331)
unc-37(ot262) | 23% (n=52) | 74% (n=56) 74% (n=56)
unc-37(red17) | 74% (n=56) | 74% (n=56)
lsy-22(ot244) | 96% (n=56) | 73% (n=30) 44% (n=50)
lsy-22(ot114) | 70% (n=56) | 73% (n=30) 44% (n=50)
lsy-5(ot3750) | 72% (n=56) | 73% (n=30) 44% (n=50)
lsy-8(ot71) | 0% (n=160) | 0% (n=160) 0% (n=160)
unc-37(ot37); lsy-8(ot71) | 0% (n=48) | 0% (n=48) 0% (n=48)
lsy-22(ot244); lsy-8(ot71) | 87% (n=52) | 87% (n=52) 87% (n=52)
cog(1-28) | 73% (n=75) | 73% (n=75) 73% (n=75)
cog(1-28)+ | 0% (n=50) | 0% (n=50) 0% (n=50)
+unc-37(ot262) | 0% (n=50) | 0% (n=50) 0% (n=50)
+lsy-22(ot244) | 0% (n=50) | 0% (n=50) 0% (n=50)
cog(1-28);unc-37(ot262) | 0% (n=50) | 0% (n=50) 0% (n=50)
+lsy-22(ot244) | 0% (n=50) | 0% (n=50) 0% (n=50)
cog(1-2OE) | 0% (n=81) | 0% (n=81) 0% (n=81)
+unc-37(ot37); cog(1-2OE) | 0% (n=81) | 0% (n=81) 0% (n=81)
+lsy-22(ot244); cog(1-2OE) | 0% (n=81) | 0% (n=81) 0% (n=81)
die-1(ot206) | 0% (n=40) | 0% (n=40) 0% (n=40)
unc-37(ot26); die-1(ot206) | 0% (n=40) | 0% (n=40) 0% (n=40)
lsy-22(ot244); die-1(ot206) | 0% (n=40) | 0% (n=40) 0% (n=40)

**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 scored with 2 ng/µl Scal-digested rol-6(d) (pRF4; 2 ng/µl) or HindIII-digested elt-2::NLS:dsRed and PmlI-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*2004;2*::FLAG::venus*) and *otEx4126* (*unc-37*2004;2*::yfp*). *otEx4126* subsequently spontaneously integrated to generate *otls288*.

**Generation of reporter genes**

*unc-37* and *lsy-22* reporter genes were created utilizing λ-Red-mediated recombinase in bacteria as described (Tursun et al., 2009). Briefly, the *lsy-2*-containing fosmid (WRM06283004) and the *unc-37*-containing fosmid (WRM0612G050) 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. Recombinered fosmids were sequenced at their recombinerend junctions and correct clones were maintained in *E. coli* and *hsy-22* subsequently spontaneously integrated to generate *otls288*.
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 systems 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 to clear 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) or with 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
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 fulfill its function as an ASER fate inducer. Lastly, both lsy-5 and lsy-22 act upstream of the ASE fate inducer die-1, as the ‘two ASE 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.](image-url)
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 recombinereed 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), 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).
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; Pfuiigrad 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 fag-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 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 tcf14 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.). It 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**


Function of a Groucho-like regulatory factor


