


RESEARCH

Open Access



Systematic evaluation of *C. elegans* lincRNAs with CRISPR knockout mutants

Shuai Wei^{1†}, He Chen^{1†}, Emmanuel Enoch Dzakah^{1,2†}, Bin Yu^{1,4}, Xiaolin Wang¹, Tao Fu¹, Jingxin Li¹, Lei Liu¹, Shucheng Fang¹, Weihong Liu^{3,5} and Ge Shan^{1,6*} 

Abstract

Background: Long intergenic RNAs (lincRNAs) play critical roles in eukaryotic cells, but systematic analyses of the lincRNAs of an animal for phenotypes are lacking. We generate CRISPR knockout strains for *Caenorhabditis elegans* lincRNAs and evaluate their phenotypes.

Results: *C. elegans* lincRNAs demonstrate global features such as shorter length and fewer exons than mRNAs. For the systematic evaluation of *C. elegans* lincRNAs, we produce CRISPR knockout strains for 155 of the total 170 *C. elegans* lincRNAs. Mutants of 23 lincRNAs show phenotypes in 6 analyzed traits. We investigate these lincRNAs by phenotype for their gene expression patterns and potential functional mechanisms. Some *C. elegans* lincRNAs play cis roles to modulate the expression of their neighboring genes, and several lincRNAs play trans roles as ceRNAs against microRNAs. We also examine the regulation of lincRNA expression by transcription factors, and we dissect the pathway by which two transcription factors, UNC-30 and UNC-55, together control the expression of *linc-73*. Furthermore, *linc-73* possesses a cis function to modulate the expression of its neighboring kinesin gene *unc-104* and thus plays roles in *C. elegans* locomotion.

Conclusions: By using CRISPR/cas9 technology, we generate knockout strains of 155 *C. elegans* lincRNAs as valuable resources for studies in noncoding RNAs, and we provide biological insights for 23 lincRNAs with the phenotypes identified in this study.

Keywords: *C. elegans*, lincRNA, MicroRNA, CRISPR, Phenotype, Transcription factor

Background

Long intergenic RNAs (lincRNAs) are a specific class of long noncoding RNAs (lncRNAs) that are encoded by genomic sequences without overlap with genomic sequences of known coding genes [1, 2]. LincRNAs were identified first in mammalian cells, and they are key regulators of diverse biological processes such as transcription and chromatin epigenetics [3, 4]. Mutations in lincRNAs have been shown to promote the development of many complex diseases, such as inflammation,

viral infection, and carcinogenesis [3, 5, 6]. For example, one extensively studied lincRNA, *hotair*, regulates epidermal differentiation and associates with cancer metastasis by interacting with epigenetic factors such as Polycomb repressive complex 2 (PRC2) [7, 8]. LincRNA-p21 has been shown to play crucial roles in hypoxia-enhanced glycolysis by forming a positive feedback loop between HIF-1 α and lincRNA-p21 to enhance glycolysis under hypoxia [9]. These roles have been characterized mostly with cultured cells, tumor xenografts, tissues, and only recently and for a very limited number of lincRNAs, also at the whole organismal level [10, 11]. For example, *linc1405* has recently been found to modulate the Eomes/WDR5/GCN5 complex in mouse ESCs, and at the whole animal level, depletion of *linc1405* impedes heart development in mice [10]. In another study, *lincRNA-EPS* was found to play a trans role in recruiting the heterochromatin binding protein hnRNP L to control nucleosome positioning

* Correspondence: shange@ustc.edu.cn

[†]Shuai Wei, He Chen and Emmanuel Enoch Dzakah contributed equally to this work.

¹Division of Molecular Medicine, Hefei National Laboratory for Physical Sciences at Microscale, the CAS Key Laboratory of Innate Immunity and Chronic Disease, School of Life Sciences, University of Science and Technology of China, Hefei 230027, China

⁶CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, CAS, Shanghai 200031, China

Full list of author information is available at the end of the article



and inhibit the transcription of immune response genes, and lincRNA-EPS traditional knockout mice demonstrate enhanced inflammation [11].

Hundreds of lincRNAs have also been identified in other metazoans such as *Caenorhabditis elegans*, *Drosophila*, and zebrafish [12–14]. There are 170 lincRNAs encoded in the current annotated *C. elegans* genome [15, 16]. Thus far, little is known about the functions and phenotypes associated with these *C. elegans* lincRNAs. Furthermore, there has been essentially no systematic analysis of all lincRNAs with knockout strains for any given animal.

CRISPR technology enables efficient production of *C. elegans* knockout and insertion strains [17–23]. In this study, we generated knockout strains using CRISPR for 155 of the 170 *C. elegans* lincRNAs. Among the 6 traits we analyzed, mutants of 23 lincRNAs exhibited phenotypes. We also provided mechanistic insights for these lincRNAs.

Results

Genome-wide characteristics of *C. elegans* lincRNAs

We performed H3K4me3 and H3K9me3 ChIP-seq and sequenced the expression profiles of embryos, L1 stage, L2 stage, dauer stage, L3 stage, L4 stage, young adults, males (*him-5* worms), and mixed stages of worms under starvation and then analyzed the 170 *C. elegans* lincRNAs for their global features (Fig. 1a, b). Several lincRNAs showed stage-specific expression (Fig. 1b, Additional file 1: Table S1). For example, linc-28, linc-131, and linc-155 were expressed only in embryos; linc-148 was expressed exclusively in L2 worms; linc-52 was expressed in young adults only; linc-141 and linc-168 were expressed only in dauer; and linc-23 was expressed in males only (Additional file 1: Table S1). There were 12 lincRNAs expressed at all stages examined, and their expression levels showed low variations (the ratio of the highest to the lowest levels of each lincRNA respectively was within tenfold) (Additional file 1: Table S1). These results indicated that the expression of some lincRNAs was under tight control for stage-specific expression and functions, while some other lincRNAs might play ubiquitous roles with expression at all stages. H3K4me3 is generally an activation marker, and H3K9me3 is a suppressive marker. We noticed that in L4 worms, H3K4me3 bound to genomic regions of the majority of lincRNAs, although H3K9me3 only bound to genomic regions of 12 lincRNAs (Fig. 1c). These results suggested a dynamic and regulated expression of *C. elegans* lincRNAs, and further investigations are necessary to dissect the relevant mechanisms and factors such as transcription factors and histone modifications.

Compared to mRNAs, lincRNAs were less conserved in 26 nematode species (Fig. 1c). When there were conserved sequences, the length of these sequences was also shorter in lincRNAs than in mRNAs (Fig. 1c). The

exon numbers of lincRNAs were significantly fewer than of mRNAs (Fig. 1d). lincRNAs were also significantly shorter than mRNAs (Fig. 1e). These features of exon numbers and sequence length were also true for lincRNAs in several other organisms [1, 12].

Phenotypes of lincRNA CRISPR knockout strains

To investigate the roles of these lincRNAs, we generated CRISPR knockout (KO) strains of 155 *C. elegans* lincRNAs (Additional file 2: Figure S1, Additional file 3: Table S2). None of the 155 lincRNA mutants showed obvious abnormality in morphology, and they did not have a severe lethal phenotype. Actually, for the 15 lincRNAs we failed to obtain CRISPR knockouts, the failure might be technical and was not due to lethality of mutants as we did not even get heterozygotes. We then examined the locomotion, defecation, pharyngeal pumping, egg retention, developmental delay, and offspring numbers of these KO strains. Twenty-three lincRNA KO strains showed defects in these 6 traits (Fig. 2, Additional file 4: Table S3); 6 lincRNAs (linc-37, linc-60, linc-73, linc-107, linc-150, and linc-159) showed uncoordination (Fig. 2a, b); 6 lincRNAs (linc-27, linc-60, linc-67, linc-72, linc-107, and linc-126) had defects in defecation (Fig. 2a, c); 5 lincRNAs (linc-2, linc-5, linc-22, linc-109, and linc-140) showed defects in pharyngeal pumping (Fig. 2a, d); 2 lincRNAs (linc-4 and linc-92) showed egg retention (Fig. 2a, e); and 2 lincRNAs, linc-10 and linc-155 had decreased numbers of progeny (Fig. 2a, f). linc-10 and linc-155 mutants actually laid fewer eggs, although essentially all eggs hatched. Four lincRNAs (linc-17, linc-18, linc-36, and linc-74) demonstrated a delay in development (Fig. 2a, g). Two lincRNAs, linc-60 and linc-107, showed pleiotropic effects in locomotion and defecation (Fig. 2a, b, d).

Expression patterns of lincRNAs with a mutant phenotype

Next, we examined the expression of lincRNAs with phenotypes using transcriptional reporter (Fig. 3). For the six lincRNAs with the uncoordination phenotype, we noticed that five (excluding linc-107) were expressed in neurons and/or muscles (Fig. 3a). For the five lincRNAs with defects in pharyngeal pumping, four (excluding linc-140) showed expression in pharyngeal muscles and neurons (Fig. 3b). The expression patterns indicated that these nine lincRNAs were expressed in cells in association with their specific phenotypes, and thus, they might play cell-autonomous roles. The other four phenotypes, defecation, egg retention, developmental delay, and offspring numbers, were relatively more complex and might be related to multiple cell types; thus, a direct link between the expression patterns of the lincRNA and the corresponding phenotype was difficult to establish (Fig. 3d–f). Additionally, 14 lincRNAs (*Is* strains) had integrated reporters and 9

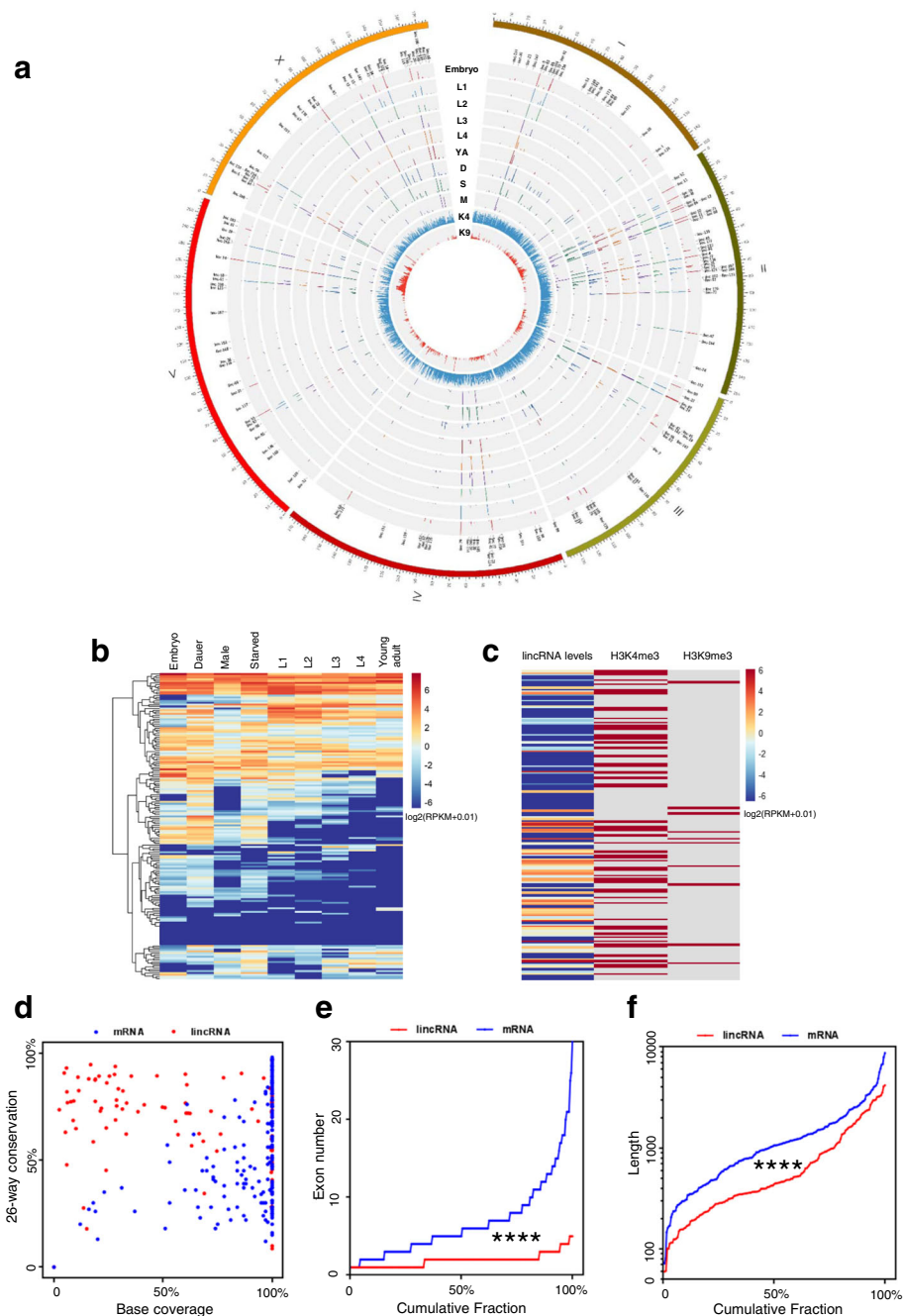


Fig. 1 Genomic characterization of *C. elegans* lincRNAs. **a** Circos plot of the 170 lincRNAs in the *C. elegans* genome. The expression levels of 170 lincRNAs in nine developmental stages and populations: embryo, L1, L2, L3, and L4, YA (young adult), D (dauer), S (mixed stages of worms under starvation), and M (male, him-5 mix worms) are shown in the inner tracks. The two innermost tracks represent distributions of H3K4me3 (K4) and H3K9me3 (K9) ChIP-seq signals (L4 worms), at the whole genome (not just for lincRNA genes). **b** Hierarchical clustering of the relative expression levels of the 170 lincRNAs. RNA-seq data from 9 developmental stages were normalized to $\log_2(\text{RPKM}+0.01)$. **c** lincRNA expression levels (heatmap of RPKM) along with H3K4me3 and H3K9me3 binding (binary map with binding in red) on lincRNA genes. **d** Conservation score of lincRNAs and mRNAs ($n = 200$, randomly chosen). “Base coverage” refers to the percentage of annotated bases. Conservation phastCons scores of 26 nematodes were interrogated from the UCSC genome browser [61], and the degree of conservation along with the portion of conserved sequences to the full-length (base coverage) lincRNAs and mRNAs were compared. **e** Cumulative plot of exon numbers of randomly selected lincRNAs and mRNAs ($n = 200$, randomly chosen). **f** Length distribution of lincRNAs and mRNAs ($n = 200$, randomly chosen). For the analysis of sequence conservation, 26 nematode conservation phastCons scores were interrogated from UCSC [61] for each base of an individual *C. elegans* lincRNA or mRNA, and the scores of each transcript were averaged. For **d** and **e**, ****, $p < 0.0001$ by the two-sided Mann-Whitney U test

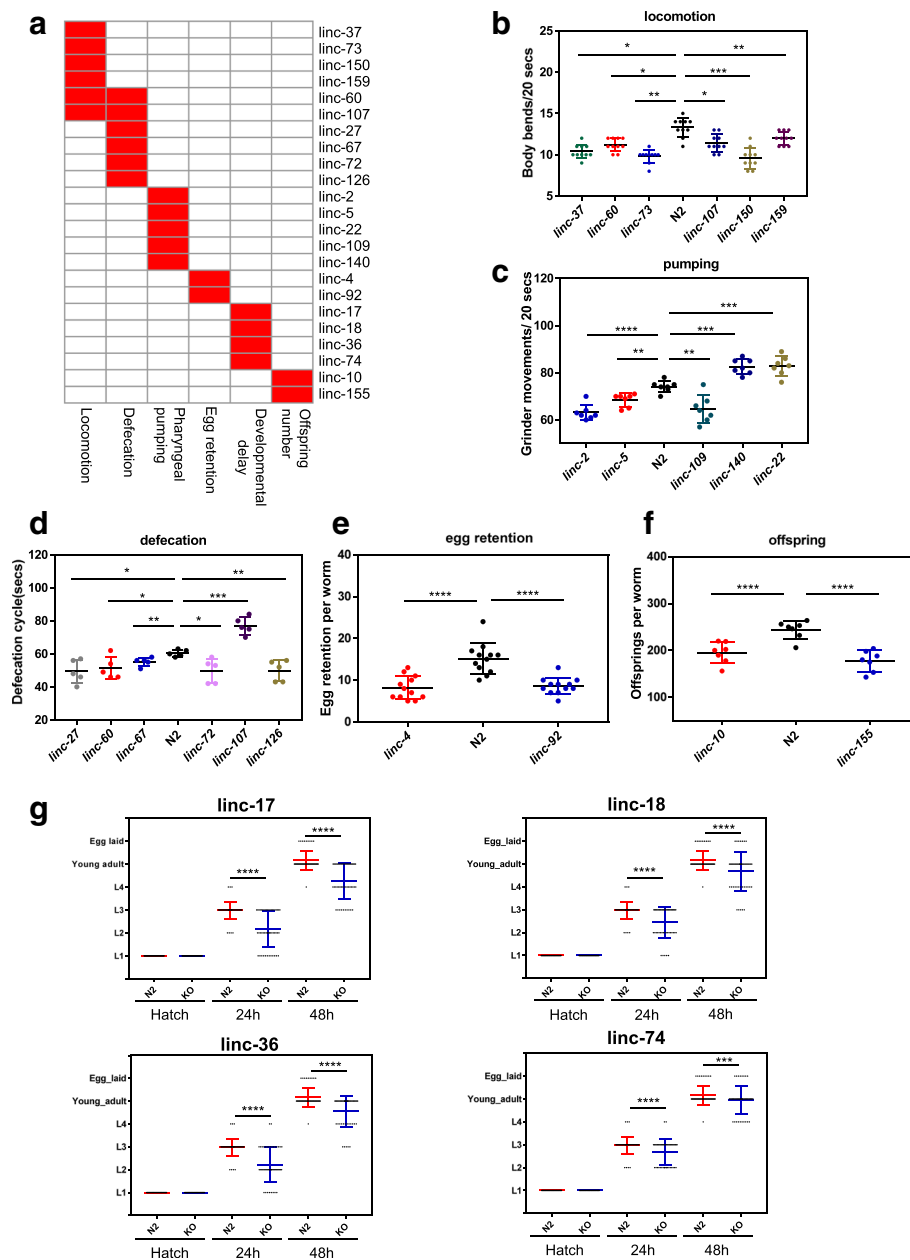


Fig. 2 Phenotypic analysis of lincRNA mutants. **a** Summary of the phenotypic characteristics of lincRNA mutants. Six phenotypic traits (locomotion, defecation, pharyngeal pumping, egg retention, rate of development, and number of progeny) were examined in 155 lincRNA mutants. The red cell represents phenotypic data of the corresponding lincRNA mutant that were with statistically significantly different compared with the wild-type data. **b** Six lincRNA mutants showed uncoordination. **c** Six lincRNA mutants had defects in defecation. **d** Five lincRNA mutants showed defects in pharyngeal pumping. **e** Two lincRNA mutants showed egg retention defects. **f** Two lincRNA mutants had decreased numbers of progeny. **g** Four lincRNA mutants demonstrated a delay in development. $n = 50$. For **g**, data for N2 worms were reused in the figure for comparison to lincRNA mutants. For **b-f**, * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; p values were calculated by the unpaired Student's t test; for **g**, *** $p < 0.001$, **** $p < 0.0001$, p values were calculated by the chi-square test

lincRNAs (*Ex* strains) had non-integrated extrachromosomal reporters (Fig. 3, Additional file 3: Table S2).

Correlations between lincRNAs and mRNAs

For the lincRNAs with a mutant phenotype, we examined whether they affected the expression of their neighboring

genes (Fig. 4a, b). For certain lincRNAs such as linc-67, linc-5, and linc-74, there were no substantial changes in the expression levels of their adjacent genes once the lincRNAs were knocked out (Fig. 4a). For lincRNAs such as linc-17 and linc-18, there were significant increases in the expression levels of their adjacent genes in the

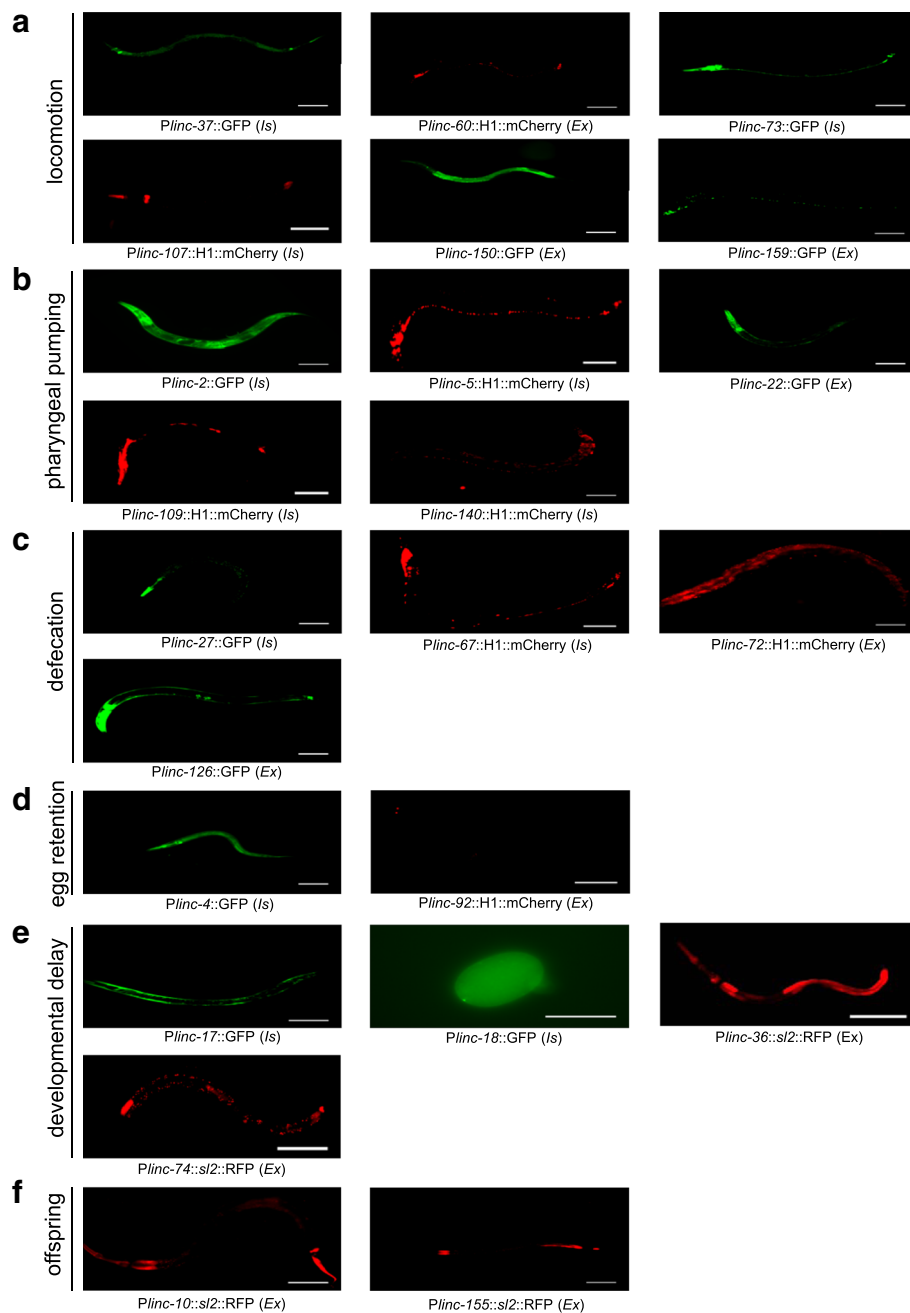


Fig. 3 Transcriptional reporters of lincRNAs. **a** The expression of transcriptional reporters of lincRNAs with locomotion defects. **b** The expression of transcriptional reporters of lincRNAs with defects in pharyngeal pumping. **c** The expression of transcriptional reporters of lincRNAs with defecation defects. **d** The expression of transcriptional reporters of lincRNAs with defects in egg retention. **e** The expression of transcriptional reporters of lincRNAs with developmental delay. **f** The expression of transcriptional reporters of lincRNAs with a decreased number of progeny. Is, integrated strain; Ex, extrachromosomal strain. Scale bar, 50 μ m

corresponding knockouts (Fig. 4a). Interestingly, the majority of these 23 lincRNAs showed complex effects on the expression of neighboring genes, with some adjacent genes demonstrating increased expression levels and some other adjacent genes simultaneously exhibiting decreased expression levels in the knockouts

(Fig. 4a). When considered as a whole, the positions of neighboring genes from the lincRNA locus showed no specific trend in how lincRNAs affected their neighboring genes (Fig. 4b). These results indicated that some of these lincRNAs had *cis* effects on the expression of their neighboring genes, and they could either

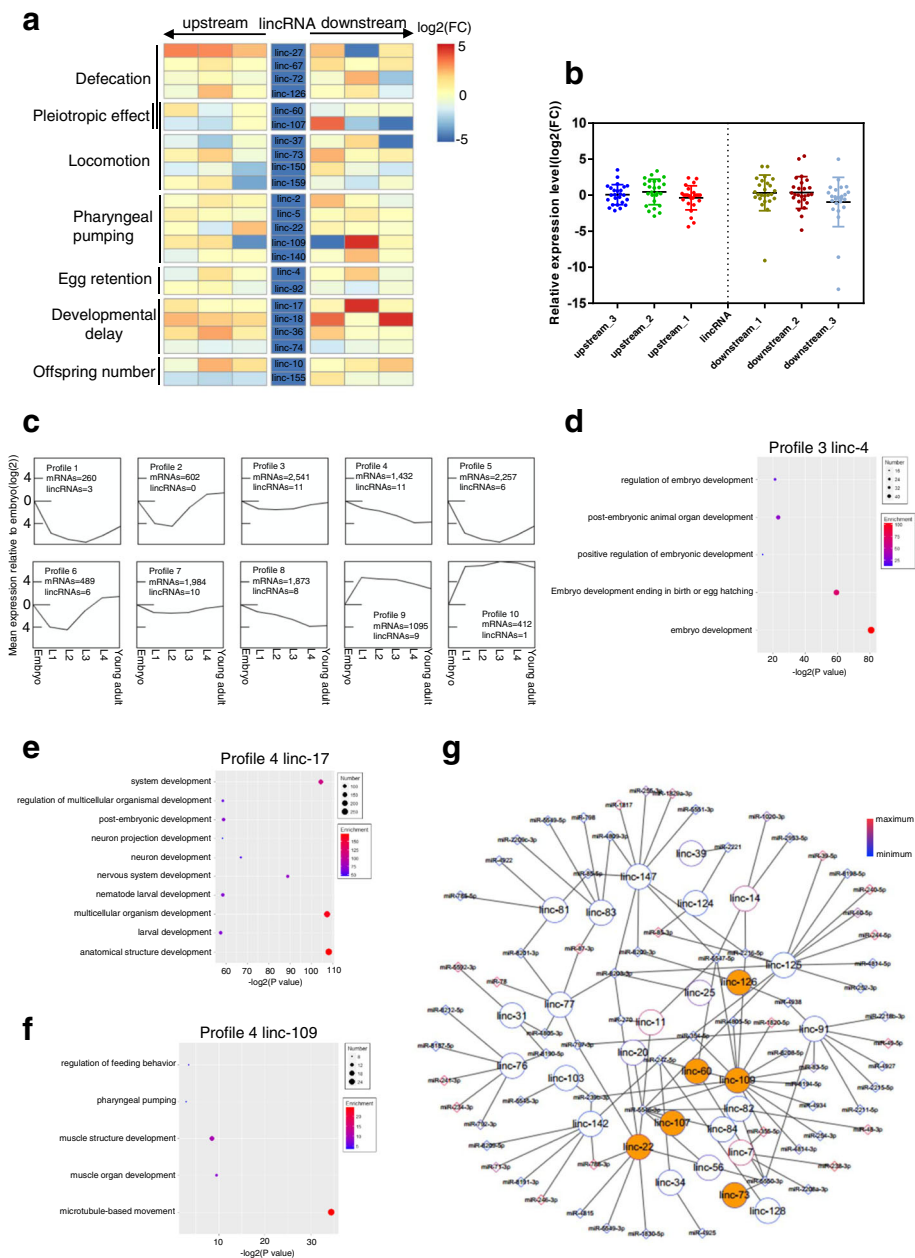


Fig. 4 Connections of lincRNAs to mRNAs and microRNAs. **a** Heatmap of the expression levels of lincRNA-neighboring genes. The expression level of each gene was assessed by qRT-PCR, and $\log_2(\text{FC})$ compared to wild type in gene expression was converted to heatmap (FC, fold change). Forward and backward arrows indicate the downstream and upstream genes. $\log_2(\text{FC})$ were set between -5 and 5 mandatorily to draw the heat map. **b** Relative expression levels of neighboring genes of the 23 lincRNAs with mutant phenotypes; data are the same as in **a**, except that the $\log_2(\text{FC})$ are actual values. **c** Mean expression profiles of mRNAs and lincRNAs using our RNA-seq data from six developmental stages (embryo, L1, L2, L3, L4, young adult). Data were analyzed by Short Time-series Expression Miner (STEM) [24] using k-means clustering. Signals for each profile cluster were normalized to signals of the embryonic stage. **d** GO analysis of coding genes in profile 3 for ontology matching the *linc-4* phenotype of egg retention. **e** GO analysis of protein coding genes in profile 4 for ontology matching the *linc-17* phenotype of developmental delay. **f** GO analysis of protein-coding genes in profile 4 for ontology matching the *linc-109* phenotype of pharyngeal pumping. **g** Global network of the lincRNA-miRNA interaction constructed with our RNA-seq data for long RNAs and microRNAs from nine developmental stages and worm populations. LincRNAs in golden brown-filled circles represent lincRNAs with mutant phenotypes in this study. The line colors of the circle (for lincRNA) and diamond (for miRNA) represent relative expression levels (scale shown to the right)

activate and/or suppress gene expression. For each individual lincRNA, however, further experiments are necessary to validate the potential *cis* role.

We also analyzed the expression correlations between the lincRNAs and the corresponding coding genes within the 100 kb upstream and downstream genomic regions (Additional file 5: Figure S2a, b); for either all the 170 lincRNAs or the 23 lincRNAs with phenotypes, the correlation between the expression of lincRNAs and mRNAs seemed to have no relevance to the position of the mRNA from the lincRNA locus. We further examined the relationship between the mean expression profiles of mRNAs and lincRNAs based on RNA-seq data for embryos, L1, L2, L3, and L4, and young adults generated by our group using Short Time-series Expression Miner (STEM) [24]. Ten expression profile patterns were obtained after normalizing the mean expression of both lincRNAs and mRNAs in L1, L2, L3, and L4, and young adults to the mean expression in the embryo (Fig. 4c). Nine of the 10 expression profiles (missing the expression profile pattern 2) contained lincRNAs that showed a correlated expression similar to the mRNAs. In these 10 expression profile patterns, profile patterns 3 and 4 showed an enrichment for the largest number of lincRNAs (11 lincRNAs in each pattern) (Fig. 4c). Gene ontology (GO) analysis of coding genes in profile 3 revealed enrichment for genes involved in the regulation of embryonic development and embryo development ending in birth or egg hatching, among others (Fig. 4d). Among the 11 lincRNAs in profile 3, only one lincRNA, *linc-4*, had a phenotype (egg retention) (Figs. 2a and 4d). Among the 11 lincRNAs in profile 4, two lincRNAs, *linc-17* (developmental delay) and *linc-109* (pharyngeal pumping), had phenotypes (Fig. 2a). GO terms in profile 4 showed enrichment for genes in system development, larval development, and pharyngeal pumping (Fig. 4e, f).

Interactions between lincRNAs and microRNAs

Thus far, it has been known that some lincRNAs play *cis* regulatory roles, and we were interested in whether some lincRNAs might have *trans* roles. Many lincRNAs play *trans* roles as competing endogenous RNAs (ceRNAs) to block the inhibitory regulation of microRNA (miRNAs) on mRNA targets [25–27].

To illustrate the interaction of lincRNAs and microRNAs, we also sequenced the microRNA expression profiles of *C. elegans* in the nine different stages and populations. A functional interaction network between lincRNAs and miRNAs was then built (Fig. 4g). We observed that of the 170 lincRNAs, 28 of them contained at least two miRNA seed regions in their sequences and showed a negative correlation with the corresponding microRNA at expression levels (Fig. 4g, Additional file 6: Table S4). Among these 28 lincRNAs, six, *linc-22*, *linc-60*,

linc-73, *linc-107*, *linc-109*, and *linc-126*, showed phenotypes in this study (Figs. 2a and 4g). In fact, *linc-109* was the lincRNA with the most microRNA interactions in this network.

A dual-color system was used to determine the interaction of lincRNA-miRNA pairs *in vivo*, in which the 3' UTR region of a GFP reporter was replaced with the complete sequences of the lincRNA of interest, and the corresponding lincRNA harboring the mutated microRNA binding sites was used as a negative control (Fig. 5). The relative GFP intensity of *Plinc-60::GFP::linc-60* was stronger in N2 worms than *mir-5550* overexpressing worms (Fig. 5a). *linc-109* was predicted to be regulated by 11 miRNAs (miR-5547-5p, miR-4805-5p, miR-1820-5p, miR-6208-5p, miR-8194-5p, miR-4934, miR-254-3p, miR-4814-3p, miR-355-5p, miR-5546-3p, and miR-239b-3p), and we examined 4 of the 11 miRNAs. *Plinc-109::GFP::linc-109* showed weaker GFP expression in worms overexpressing *mir-355*, *mir-254*, or *mir-4934* (Fig. 5b–d). However, another tested microRNA, miR-5546, had no effect on the expression of *Plinc-109::GFP::linc-109* (Additional file 7: Figure S3a). Another predicted lincRNA and microRNA pair, *linc-126* and *mir-4938*, also did not show an interaction in the dual-color *in vivo* assay (Additional file 7: Figure S3b). These results strongly indicated that certain lincRNAs could play *trans* roles as ceRNAs in *C. elegans*.

Rescuing lincRNA phenotypes

Rescue experiments can provide further insights into molecular mechanisms, and thus, we expressed the corresponding lincRNA with its own promoter in the 23 lincRNA mutants. Among these 23 lincRNA mutants, the phenotypes of 9 lincRNA mutants were fully rescued, those of 7 lincRNA mutants were partially rescued, and those of 9 lincRNA mutants were not rescued (Fig. 6a, Additional file 8: Table S5). Here, partial rescue meant that the rescuing line showed a statistically significant difference from the lincRNA mutants, although the defect was not fully recovered as data from the rescuing line were still significantly different from those of the wild-type worms. For locomotion defects, three lincRNA mutants, such as *linc-37*, could be fully rescued, two including *linc-73* could be partially rescued, and *linc-159* mutant was not rescued (Fig. 6b). For the other phenotypes in pharyngeal pumping, defecation, egg retention, offspring number, and developmental delay, we observed that two lincRNA mutants with defects in number of progenies could not be rescued with overexpression, and lincRNA mutants with one of the other four phenotypes could either be fully rescued, partially rescued, or not rescued (Fig. 6c–g). LincRNAs (e.g., *linc-109*) with phenotypes that could be fully rescued by overexpressing the corresponding

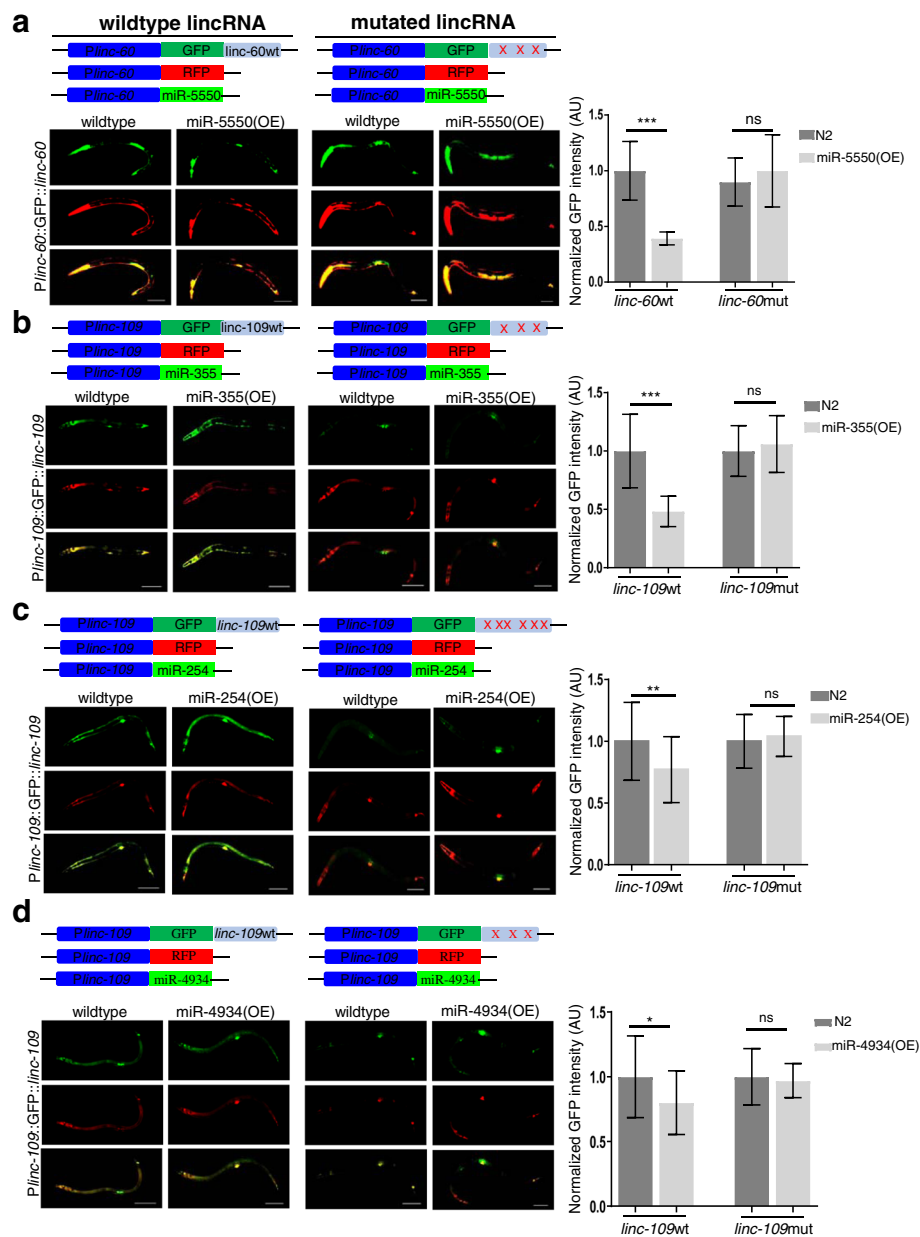


Fig. 5 Regulation of lincRNAs by miRNAs. **a** Relative GFP expression level of *linc-60* in N2 worms with or without the overexpression of *mir-5550* ($n = 20$). **b** Relative GFP expression level of *linc-109* in N2 worms with or without the overexpression of *mir-355* ($n = 20$). **c** Relative GFP expression level of *linc-109* in N2 worms with or without the overexpression of *mir-254* ($n = 20$). **d** Relative GFP expression level of *linc-109* in N2 worms with or without the overexpression of *mir-4934* ($n = 20$). Constructs with mutations in the miRNA binding site of lincRNA were used as negative controls, and the positions of mutations are represented by red crosses. ns, no significance; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; Student's t test. Data are the means \pm SD. Images shown are representative of the control and experimental groups. Scale bar, 20 μ m

lincRNA might mainly play *trans* roles, while those with phenotypes that could not be rescued by overexpressing the corresponding lincRNA (e.g., *linc-27*) might mainly play *cis* roles. LincRNAs (*linc-73*) with phenotype that could be partially rescued might possess both *trans* and *cis* roles. For phenotypes that likely link to germline expression (e.g., *linc-10* and *linc-155*), failure to rescue might be due to silencing of the

overexpressing extrachromosomal constructs. Of course, links between the rescuing result and the molecular mechanism might be more complex, and we were comparing it with other results.

Transcriptional regulation of lincRNAs

The transcriptional regulation of noncoding RNAs has not been clearly understood because most studies have

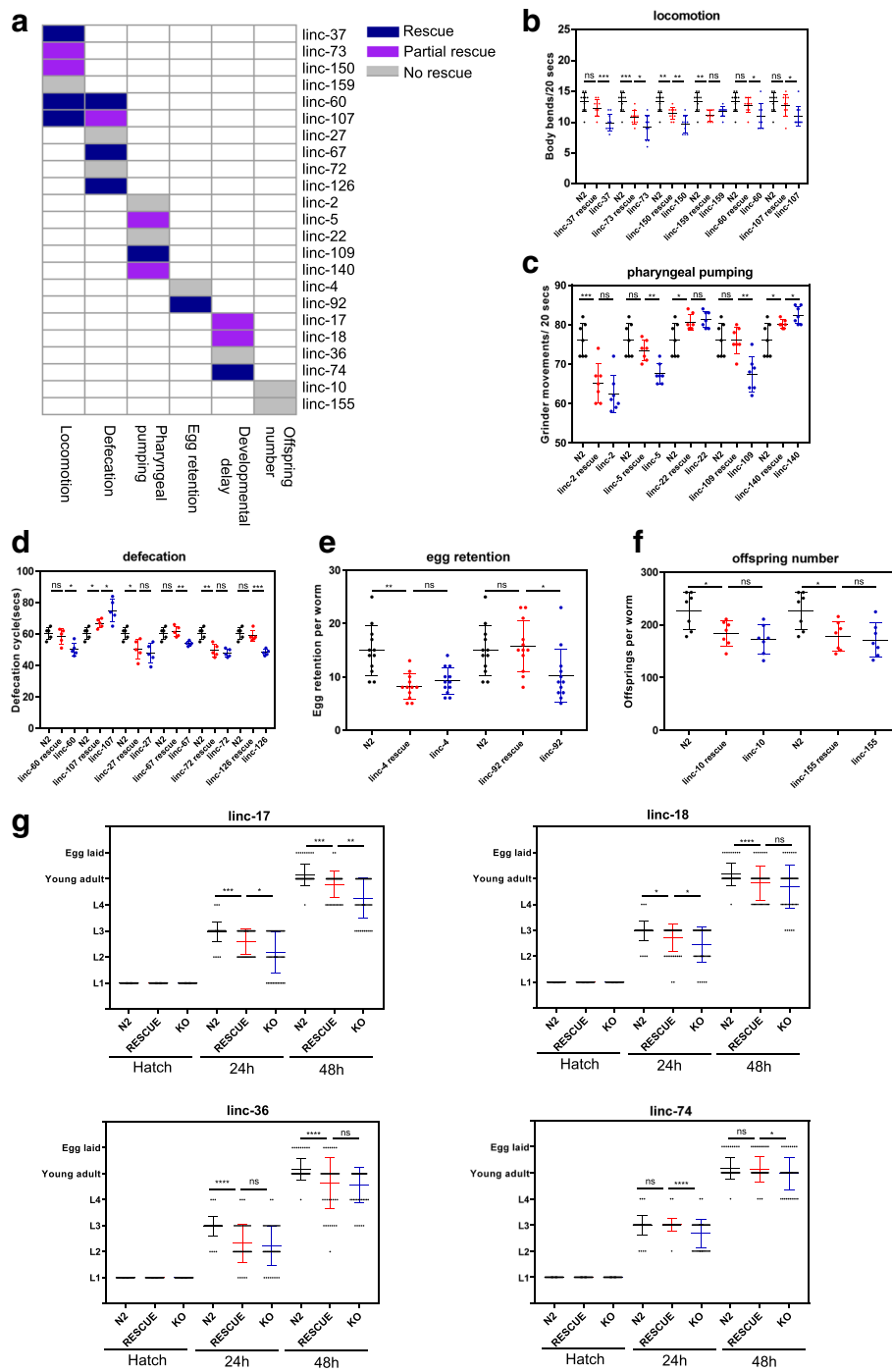


Fig. 6 Rescuing lincRNA phenotypes. **a** Summary of the rescuing experiment results in all 23 lincRNA mutants. Rescue, the mutant phenotype was fully rescued. Partial rescue, the mutant phenotype was rescued, although data from the rescuing line were still significantly different from those of wild-type worms. **b** Rescuing data for the locomotion phenotypes in 6 lincRNA mutants. **c** Rescuing data for the pharyngeal pumping defects in 5 lincRNA mutants. **d** Rescuing data for the defecation defects in 6 lincRNA mutants. **e** Rescuing data for the decreased egg retention in 2 lincRNA mutant worms. **f** Rescuing data for the reduced number of progeny phenotype in 2 lincRNA mutants. **g** Rescuing data for the developmental delay in 4 lincRNA mutants. For b-g, the data for N2 worms were reused in the figure for comparison to the lincRNA mutants and rescue lines. For b-f, ns, no significance; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; unpaired Student's t-test. Data are the means \pm SD. For **g**, ns, no significance; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; p values were calculated by the chi-square test

focused on protein-coding genes [28, 29]. We analyzed the chromatin immunoprecipitation sequencing (ChIP-seq) data of ~300 transcription factors in *C. elegans* downloaded from modENCODE to examine their binding sites on the genomic sequences of lincRNAs in 6 different stages [30, 31]. According to our re-analyzed data, 60 of 79 transcription factors were found to regulate a total of 136 lincRNAs in the embryo (Fig. 7a); 96 of the 116 transcription factors showed binding to the genomic region of 130 lincRNAs in L1 stage (Fig. 7b); 99 of 107 transcription factors regulated 131 lincRNAs in L2 stage (Fig. 7c); 85 of 108 transcription factor genes at L3 stage regulated the transcriptional expression of 143 lincRNAs (Fig. 7d); 93 of 110 transcription factors might regulate the expression of 129 lincRNAs at L4 stage (Fig. 7e); and 37 of 39 the transcriptional factors showed binding to 109 lincRNA genes in young adults (Fig. 7f). Interestingly, the 23 lincRNAs with a phenotype in this study were regulated by more transcription factors than the other 147 lincRNAs in L1, L2, and L3 worms (Fig. 7g–i), while there was no significant difference in the number of transcription factors regulating these two groups of lincRNAs in embryos, L4 worms, and young adults (Additional file 9: Figure S4).

Previous studies by our group and others have shown that two transcription factors, UNC-30 and UNC-55, work together to specify GABAergic DD and VD motor neurons (mns) in *C. elegans* [32–34]. Therefore, we analyzed the ChIP-seq data from endogenously expressed UNC-30 and UNC-55 for their lincRNA targets [32]. UNC-30 regulated 10 lincRNAs, and UNC-55 regulated 9 lincRNAs (Fig. 7j). UNC-30 and UNC-55 shared 6 lincRNA target genes (*linc-5*, *linc-58*, *linc-73*, *linc-146*, *linc-149*, and *linc-152*) (Fig. 7j, k, Additional file 10: Figure S5). The 6 shared lincRNA targets showed a higher relative enrichment in ChIP-seq compared with lincRNA targets that were regulated by either UNC-30 or UNC-55 alone (Fig. 7k). Among the shared lincRNA targets of UNC-30 and UNC-55, *linc-5* and *linc-73* had phenotypes of pharyngeal pumping and locomotion, respectively (Fig. 2a). Promoter reporters of *linc-5* and *linc-73* demonstrated that both lincRNAs were expressed in the head region and the D mns (Fig. 3a, b).

Molecular mechanism of *linc-73* in locomotion

The *linc-73* CRISPR KO strain showed uncoordinated backward movement resulting in the formation of a ventral coil, which resembled the phenotype of the *unc-55* mutant (Fig. 8a). *linc-73* was expressed in GABAergic D mns and other cells (Fig. 8b), and its expression levels were decreased in either *unc-55(e1170)* or *unc-30(e191)* mutants (Fig. 8c). The decrease of *linc-73* expression in *unc-55(e1170)* or *unc-30(e191)* was mild, which could be explained by the expression

of *linc-73* in cells without *unc-55* or *unc-30* expression. These results indicated that both UNC-30 and UNC-55 activated *linc-73* expression. The immediate downstream gene of *linc-73* was *unc-104*, a *C. elegans* kinesin gene [35–37], and the expression levels of *unc-104* were significantly increased in *linc-73* KO worms (Fig. 8d). We noticed that this change in expression levels was inconsistent with changes in H3K4me3 (activation marker) and H3K9me3 (suppressive marker) at the promoter region of *unc-104* when comparing the *linc-73* mutant to wild-type worms (Fig. 8e). When transcription terminal sites were inserted into the *linc-73* genomic region, the expression levels of *unc-104* were increased (Fig. 8f, Additional file 11: Figure S6). When the binding site of UNC-30 or UNC-55 in the promoter of *linc-73* was mutated, the expression levels of *unc-104* were also increased (Fig. 8f, Additional file 11: Figure S6). These results supported a model in which both UNC-30 and UNC-55 could activate the expression of *linc-73* RNA, which played a *cis* role to modulate the histone epigenetic status of the *unc-104* promoter and therefore inhibit the expression of *unc-104*.

It is well known that *unc-104* plays essential roles in the transportation of presynaptic proteins [35–37]. There was a slight decrease in the dorsal presynaptic puncta for DD mns in *linc-73* mutants compared to a more dramatic decrease in the number of ventral presynaptic VD mn puncta (Fig. 8g). The detailed mechanism about how increased levels of UNC-104 in D mns resulted in asymmetric presynaptic punctum distribution remained for further investigation. These changes in DD and VD mns in *linc-73* mutants would result in relatively weaker inhibition of ventral vs dorsal body wall muscles in *linc-73* mutants and thus to a ventral coil phenotype. Taken together, these data suggested a model in which two transcription factors, UNC-30 and UNC-55, co-regulated the expression of *linc-73*, which then regulated the expression of *unc-104* *in cis* by affecting histone modifications to modulate the formation of presynapses in the D mns and further to play roles in *C. elegans* locomotion (Fig. 8h).

Discussion

LincRNAs are now recognized as critical players in eukaryotic cells [1–4]. Studies at the cellular level have uncovered a myriad of functions and functional mechanisms for many mammalian lincRNAs [7, 9, 38]. These lincRNAs can play roles either in the nucleus or in the cytoplasm with an array of *trans* and *cis* mechanisms [39, 40].

CRISPR enables fast and efficient genetic engineering, thus providing an opportunity to generate KO strains for nearly all the lincRNAs of an animal, *C. elegans*. Systematic analyses of these strains for just six traits

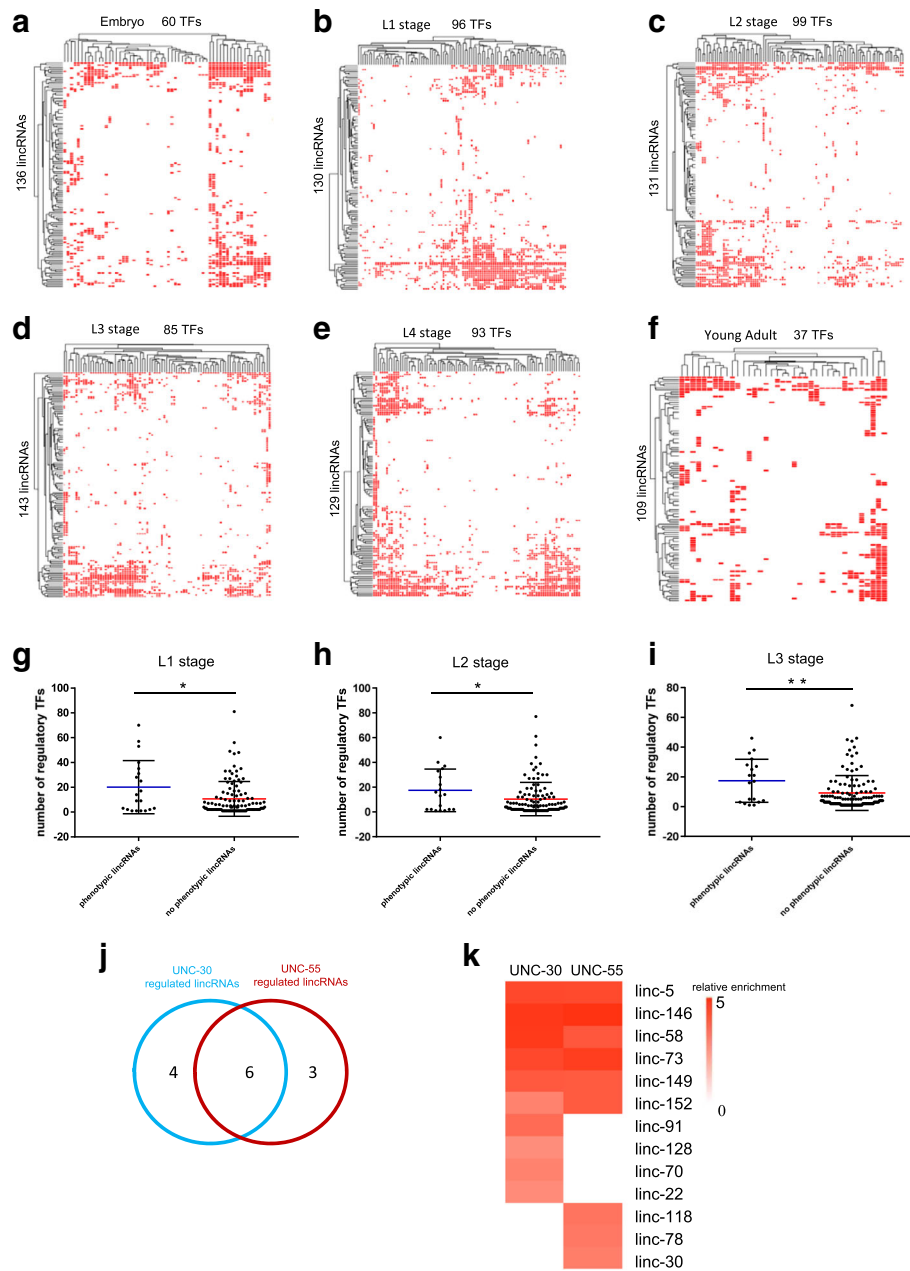


Fig. 7 Regulation of lincRNAs by transcription factors. **a** Clustering map illustrating the binding of transcription factors to lincRNA genes in embryos. **b** Clustering map illustrating the binding of transcription factors to lincRNA genes in L1 stage. **c** Clustering map illustrating the binding of transcription factors to lincRNA genes in L2 stage. **d** Clustering map illustrating the binding of transcription factors to lincRNA genes in L3. **e** Clustering map illustrating the binding of transcription factors to lincRNA genes in L4 stage. **f** Clustering map illustrating the binding of transcription factors to lincRNA genes in young adult worms. **g–i** Number of transcription factors regulating the 23 lincRNAs with phenotypes in this study and the other 147 lincRNAs in L1 worms (**g**), L2 worms (**h**), and L3 worms (**i**). **j** UNC-30 and UNC-55 regulated lincRNA targets. **k** Heatmap of the relative ChIP-seq enrichment of UNC-30 and UNC-55 lincRNA targets. All data illustrated in a–f were downloaded from modENCODE. * $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; p values were calculated by two-sided Mann-Whitney U test

identified 23 phenotypic lincRNAs; it would be reasonable to speculate that many lincRNAs or even most of them may be phenotypic lincRNAs given the analysis of more (or more complex) traits, such as chemosensory, longevity, and male mating. Researchers have just

started to explore the roles of lincRNAs and other lincRNAs systematically with CRISPR screening in mammalian cell cultures [41–44]. LincRNAs do not have overlapping sequences with other genes, which makes them relatively more adaptive to perturbation,

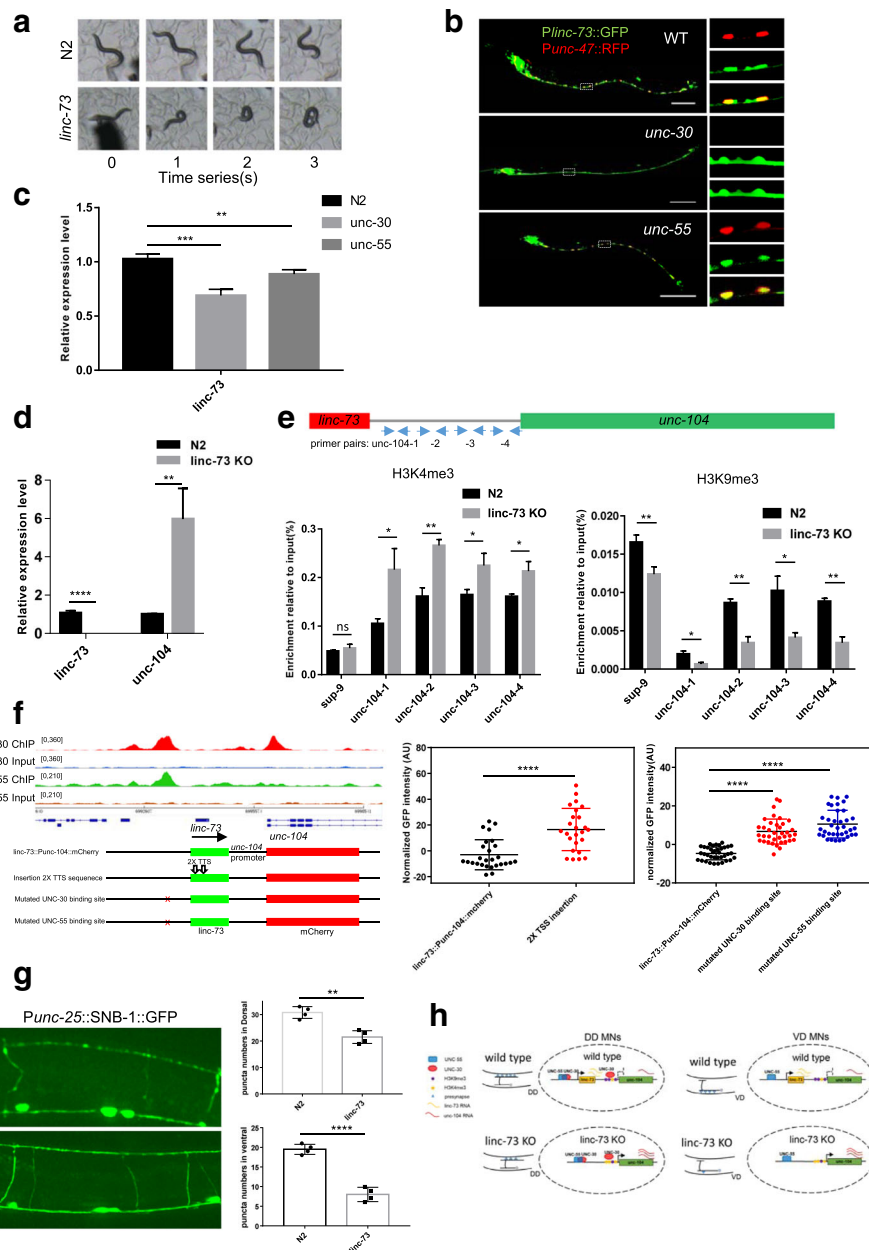


Fig. 8 *cis* effect of *linc-73* on the neighboring gene *unc-104*. **a** Time-lapse observation of the uncoordinated backward movement of the *linc-73* CRISPR KO strain. **b** Expression pattern of *linc-73* in the wild-type, *unc-30*, and *unc-55* mutant backgrounds. Areas inside the dashed boxes are enlarged at the sides. *Punc-47::RFP* is a GABAergic marker. **c** qRT-PCR of *linc-73* RNA levels in L2 worms of N2, *unc-55(e1170)*, and *unc-30(e191)*. **d** qRT-PCR of *unc-104* mRNA levels in N2 and *linc-73* KO worms (L2). **e** H3K4me3 (activation marker) and H3K9me3 (suppressive marker) at the promoter region of *unc-104* in N2 and *linc-73* mutants (L2 worms). The positions of the primer pairs used are indicated in the diagram. **f** Quantification of the relative expression levels of *unc-104::mCherry* in the cell body of D mns. Positions of the mutated UNC-30 (Δ UNC-30) & UNC-55 (Δ UNC-55) and insertion of the transcription terminal site (TTS) are shown along with the UNC-30 and UNC-55 ChIP-seq peaks. **g** Quantification of the dorsal and ventral presynaptic puncta (SNB-1::GFP) of DD mns in N2 and *linc-73* mutants (L2). Representative images are shown. **h** a working model for the regulation of UNC-30 & UNC-55 on *linc-73*, which then regulates the expression of *unc-104* to modulate *C. elegans* locomotion. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.0001$ by the Student's *t* test. Scale bar, 50 μ m

and the results from the manipulations are relatively easier to explain. Our understandings of lincRNAs could also be true for other lincRNAs, as lincRNAs have multiple features that are shared by many other

lincRNAs. The study of lincRNAs and lncRNAs in *C. elegans* is relatively lagging behind that in mammalian cells. The *C. elegans* KO strains of lincRNAs from this study would be valuable resources for future studies, as

this animal is a supreme model organism with powerful genetic and cell biology tools.

Critical roles of lincRNAs at the cellular level sometimes do not justify their physiological significance at the whole organismal level. For example, studies at the cellular level have demonstrated that MALAT1 plays major roles in nuclear speckles for mRNA processing, splicing, and export [45, 46]. However, there is no obvious phenotype in MALAT1 KO mice [47, 48]. Additionally, some recent arguments have been raised about the physiological roles of hotair, as some researchers believe that hotair KO mice do not show an apparent phenotype [49, 50]. Therefore, it is of great value to study lincRNAs both at the cellular level and with animals. Our lincRNA KO strains would facilitate studies at the whole organismal level. A pilot study using traditional method has generated KO strains for 18 murine lincRNAs, and essentially, all these mutants have phenotypes of embryonic lethal or severe defects in development leading to early death [51]. It is somewhat surprising that none of the 155 *C. elegans* lincRNA mutants have a lethal phenotype. It could be that the mammalian development is much more complicated, and the previous study also selected for lincRNAs with expression patterns of greater association with neural development [51].

To analyze the connections of *C. elegans* lincRNAs to other transcripts and epigenetic markers, we performed ChIP-seq of H3K4me3 and H3K9me3 for L4 worms and RNA-seq for both long RNAs (e.g., lincRNAs, mRNAs, and circular RNAs) and small RNAs (e.g., microRNAs) in nine worm developmental stages and populations (GSE115324). These are also valuable resources for future studies. Network construction and expression profile association can provide mechanism insights in the roles of lincRNAs. For example, co-expression analysis revealed that linc-109 was associated with muscle development and pharyngeal pumping, as well as microtubule-based movement (Fig. 4f), and the phenotype of *linc-109* mutant was a pharyngeal pumping defect. The lincRNA-microRNA co-expression and bioinformatic analyses revealed that linc-109 might be regulated by multiple microRNAs (Fig. 4g), and indeed, some of these regulatory effects were experimentally confirmed (Fig. 5). These points and the complete rescue of the *linc-109* phenotype by overexpressing this lincRNA (Fig. 6a, c) strongly suggested a *trans* regulatory role of linc-109, making it highly plausible that it serves as a ceRNA against microRNAs. lincRNAs can play *trans* roles other than ceRNA [39, 52, 53], and other potential *trans* roles of *C. elegans* lincRNAs require further investigations.

For the 8 lincRNAs that were expressed exclusively at one particular stage, only the linc-155 mutant had a phenotype, and the phenotype of a decreased number

of progenies seemed to match its exclusive expression in early embryo (Figs. 1a, b and 2a, f). For the 12 lincRNAs that were ubiquitously expressed, only the *linc-4* mutant demonstrated a phenotype, egg retention (Figs. 1a, b and 2a, e), and it was difficult to speculate on any direct link between the ubiquitous expression of *linc-4* with the mutant phenotype. For the remaining 150 lincRNAs that were not expressed either ubiquitously or exclusively, the mutants of 21 lincRNAs showed phenotypes in the six traits examined (Figs. 1a, b and 2). For locomotion, defecation, pharyngeal pumping, egg retention, and offspring number, young adults were examined. Therefore, it was difficult to identify links between the corresponding expression pattern and the phenotype. For the four lincRNAs (*linc-17*, *linc-18*, *linc-36*, and *linc-74*) with a developmental delay, their mutants already did display retardation in early development within 24 h of hatching (Figs. 1a, b and 2a, g). All four of them showed relatively high expression levels in the embryo (Fig. 1a, b, Additional file 1: Table S1).

The expression of lincRNAs is under the control of transcription factors, and we noticed that a small portion (8 of ~300) of transcription factors (LIN-39, EOR-1, BLMP-1, NHR-77, HLH-1, DAF-16, W03F9.2, and NHR-237) regulated the expression of ≥ 50 lincRNAs (Fig. 7a–f). It would be interesting to further investigate the biological relevance underlying this regulatory phenomenon. A lincRNA can be transcriptionally regulated by multiple transcription factors together (Fig. 7). For example, lincRNA-73 is regulated by 48 transcription factors, including UNC-30 and UNC-55, two transcription factors that converge to control the differentiation and plasticity of GABAergic D mns [32–34]. Six lincRNAs are co-regulated by UNC-30 and UNC-55 (Fig. 7j) [23]. It was surprising that CRISPR knockout of only one of the six lincRNAs, *linc-73*, gave rise to uncoordination (Figs. 2a, b and 8). It is known how *linc-73* plays a cell-autonomous role in D mns to regulate the expression of *unc-104* (Fig. 8), but the roles of the other 5 lincRNAs that are also commonly regulated by UNC-30 and UNC-55, and why KO strains of these lincRNAs do not show a locomotion defect, remain to be elucidated. The 23 lincRNAs with mutant phenotypes in this study tended to be regulated by more transcription factors in L1, L2, and L3 worms (Fig. 7g–i). It is possible that these lincRNAs are related to greater physiological regulation, and thus, their perturbation may be more likely to cause defects. As for the regulation by histone modifications, our results show that both H3K4me3 and H3K9me3 regulate *linc-73* at L2 stage, although only H3K4me3 but not H3K9me3 binds to *linc-73* at L4 stage (Figs. 1c and 8e). H3K9me3 does not have that many genomic binding peaks as compared to H3K4me3 in our study and also in data from others (NCBI BioProject: PRJEB20485).

We have presented data to support that *linc-73* plays a *cis* role to regulate the expression of *unc-104* (Figs. 4a and 8), although it is possible that *linc-73* also has a *trans* role because the overexpression of *linc-73* via an extrachromosomal construct could partially rescue the *linc-73* phenotype (Fig. 6a, b). However, *linc-109* has been shown to function with *trans* roles (Figs. 4g, 5b–d, and 6a, c), although the expression of neighboring genes is altered in *linc-109* KO, which may be an indication of a *cis* role (Fig. 4a). The effects of *linc-109* KO on the expression of its neighboring genes may not contribute to the mutant phenotype, as the extrachromosomal construct could fully rescue the *linc-109* phenotype (Fig. 6a, c). The application of CRISPR actually deletes the DNA sequences of lincRNAs, which may harbor DNA elements that regulate the expression of neighboring genes. Thus, for each individual lincRNA, an array of experiments must be performed to elucidate the potential *cis* and/or *trans* role.

Conclusions

By using CRISPR, we have generated knockout strains of 155 *C. elegans* lincRNAs as valuable resources for studies in ncRNAs. Systematic analyses of these strains for just six traits identified phenotypes in 23 lincRNA mutants. We have characterized some aspects of the expression patterns, molecular mechanisms, and other regulatory relevance of these lincRNAs.

Methods

Animal cultures and strains

Unless otherwise stated, all *C. elegans* strains used in this study were maintained on standard nematode growth medium (NGM) at 20 °C or 25 °C [54]. N2 Bristol was obtained from the Caenorhabditis Genetic Center (CGC). Eight strains including XIL0375, XIL0389, XIL1172, XIL0354, XIL1177, XIL0386, XIL0411, and XIL1237 were gifts from Dr. Xiao Liu. All worm strains generated or used in this study are listed in Additional file 3: Table S2.

Worm synchronization

Gravid adult worms were washed three times with M9 and collected into 1.5 ml tubes, after which the tubes were centrifuged at 600 g. Animals were then treated with hypochlorite. Synchronized embryos were cultured at 20 °C on NGM plates with seeded OP50.

Plasmid construction

pDD162, expressing Cas9 II protein, was a kind gift from Dr. Guangshuo Ou. For long lincRNAs (> 2 kb), 3–6 sgRNAs were designed to target the 5' ends of the lincRNA. In the case of short lincRNAs (< 2 kb), 2–3 sgRNAs targeting the 5' and 3' ends of each lincRNA

were used. In order to enhance the efficiency of the sgRNA, we specifically selected sgRNAs containing two NGG PAM motif in the 3' ends of sgRNA sequence. All the sgRNA sequence used in this work were assessed at <http://crispor.tefor.net/>. The 20 nt sgRNA sequence was inserted behind the U6 promoter of pPD162 plasmid between the EcoRI and HindIII restriction endonuclease sites. Homology recombination plasmids were generated by cloning the 1.5 kb DNA sequence upstream of the site of interest, 2 kb lincRNA promoter sequence, GFP sequence, and 1.5 kb DNA sequence downstream of the site of interest between Sph I and Apa I of the pPD117.01. For lincRNA transcriptional reporters, approximately 2.5 kb promoter sequence was cloned from genomic DNA. The corresponding product was fused with sl2 sequence and was inserted between the Sph I and Age I of pPD117.0 expressing GFP or between Pst I and Age I of pPD95.67 expressing RFP (Andrew Fire collection, Addgene). For rescue plasmids, 2 kb promoter sequence was cloned from the genomic DNA, and lincRNA full-length sequence was cloned from cDNA. All those products were inserted into the pPD117.01 between the Sph I and Apa I double-digested sites. In the dual-color system for the in vivo analysis of miRNA-lincRNA interaction, we constructed GFP reporters for the selected lincRNA by replacing the 3' UTR region of pPD117.01 with the complete wild-type sequence of the lincRNA of interest. As a control, the mutated versions of each lincRNA, in which the respective miRNA binding sites within the lincRNAs were mutated, was also cloned into pPD117.01. miRNA overexpression plasmids were constructed cloning the pri-miRNA sequence of the miRNA into pPD95.67 driven by promoter of the corresponding lincRNA. For *linc-73::Punc-104::mCherry* plasmid, *linc-73* (TTS insertion)::*Punc-104::mCherry* plasmid and *linc-73* (mutated UNC-30 binding site)::*Punc-104::mCherry* plasmid construction, *linc-73* promoter and gene body sequence, *unc-104* promoter sequence, mCherry sequence were cloned separately and inserted into the pPD117.01. The UNC-30 or UNC-55 binding site in *linc-73* promoter was mutated from GATTA to CTCAG (for UNC-30) or from ATCG ATCCAT to CGATCGAACG (for UNC-55). 2X transcriptional terminal site (2X TTS, AAATAAAAT TTTCAGAAATAAAATTTTACA) was inserted into the 5' portion of *linc-73*. A list of primers used is provided in Additional file 12: Table S6.

Injection of CRISPR/Cas9 knockout and knock-in and other plasmids

CRISPR/Cas9 system was carried out as previously described with modifications [23]. For the knockout system, we mixed 3–6 Pu6::lincRNA sgRNA plasmids

(30 ng/ μ l of each) and pPD162 expressing Cas9 II protein (30 ng/ μ l), as well as co-injection marker *Pmyo-2::mCherry* (PCFJ90) (10 ng/ μ l) together. The mixture was injected into about 30 N2 adults (adulthood day 1). For the CRISPR knock-in, the upstream locus of *linc-1* in chromosome I was selected as the knock-in site due to the presence of fewer genes located in the *linc-1* neighborhood. PU6::sgRNA plasmids (30 ng/ μ l), PPD162 plasmid (30 ng/ μ l), co-marker plasmid (10 ng/ μ l), and homologous recombination plasmid (40 ng/ μ l) were injected into the 30 gravid worms, and transgenes were selected as described above. All the knockout or knock-in mutant worms were transferred to new plates and outcross for at least three generations to eliminate off targets. In the dual-color system, wild-type or mutated lincRNA reporters (20 ng/ μ l) were mixed with miRNA overexpression plasmids (20 ng/ μ l), control plasmids (20 ng/ μ l), and a 1-kb DNA ladder (Invitrogen) standard. For rescue experiment, overexpression plasmid of lincRNAs (20 ng/ μ l) was mixed with co-maker plasmid (PCFJ90 20 ng/ μ l) as well as DNA ladder (Invitrogen). *linc-73::Punc-104::mCherry* plasmid (20 ng/ μ l), *linc-73* (TTS insertion)::*Punc-104::mCherry* plasmid (20 ng/ μ l), *linc-73* (mutated UNC-30 binding site)::*Punc-104::mCherry* plasmid (20 ng/ μ l) and *linc-73* (mutated UNC-55 binding site)::*Punc-104::mCherry* plasmid (20 ng/ μ l) was mixed with *myo-2::GFP* separately, and injected into gravid young adults. Standard micro-injection techniques were used.

Screening for CRISPR deletion and knock-in strains

Approximately, 200 F1 worms were singled after the injection and cultured at 25 °C. Genomic DNA of the F3 generation was extracted and examined by PCR. Worms were harvested and transferred to 100 μ l lysis buffer (20 μ g/ml Proteinase K, 100 mM KCl, 10 mM PH8.3 Tris-HCl, 1.5 mM MgCl₂), and then placed at –80 °C for 10 mins, thawed at 65 °C for at least 2 h. Worms were then placed at 95 °C for 15 mins to inactivate proteinase K, and 2 μ l each worm lysate was used as DNA template for PCR amplification with primers spanning sgRNA-targeted regions. For the verification of the knock-in strains, we amplified genomic regions spanning the point of insertion. Worms with the corrected PCR products were singled to NGM plates and further confirmed by DNA sequencing of the genomic PCR products. CRISPR worms were outcrossed at least three times before being used in experiments. The primers used for PCR screening are listed in Additional file 12: Table S6.

Locomotion

To examine locomotion of worms, young adult worms were removed from the bacterial lawn of an agar culture

plate to bacteria-free plates at room temperature, and allowed to crawl away from any food remains for about 10–20 s. Complete body bends per 20 s were then counted under a dissecting microscope after animals were gently touched at the tail end (*n*, number of worms = 5; *N*, number of replicates = 3) [55].

Defecation assay

Defecation cycles were performed according to previous report [56]. Data was presented by recording the time between defecation cycles of young adult worms (*n*, number of worms = 5; *N*, number of replicates = 3).

Pharyngeal pumping

Pharyngeal pumping behavior was assayed as previously described [55, 57]. Pharyngeal pumping was examined by counting grinder movements for 20 s at 20 °C (*n*, number of worms = 7; *N*, number of replicates = 3).

Egg retention

Egg retention assay was carried out as described earlier with some modifications [58]. One day (post the last molt) old adult worms were singled out and lysed in hypochlorite solution for 6 mins in 96-well plate, and the number of eggs was counted (*n*, number of worms = 12; *N*, number of replicates = 3).

Examination of development stages

To examine the development stages of worms, synchronized eggs were allowed to hatch at 20 °C and allowed to grow at NGM plates with adequate food and their developmental stages were examined after 24 h and 48 h (*n*, number of worms = 30; *N*, number of replicates = 3) [59].

Number of progenies

L4 worms were singled on NGM plates and allowed to lay eggs at 20 °C [60]. Individual worms were transferred daily from the start of egg laying until egg laying stopped. The number of live offspring (L1) were counted (*n*, number of worms = 7; *N*, number of replicates = 3). All experiments were performed under a dissecting microscope.

Quantitative RT-PCR (qRT-PCR) and quantitative PCR (qPCR)

RNAs were extracted from worms in TRIzol L/S solution (Invitrogen) after three cycles of freezing at –80 °C and thawing at room temperature. Five hundred nanogram total RNAs were reverse transcribed into cDNA by cDNA synthesis kit (Goscript™ Reverse Transcription System, Promega). qRT-PCR (with cDNA template) and qPCR (with genomic DNA template) were performed using a GoTaq qPCR Master Mix kit (Promega) on a PikoReal 96 real-time PCR system (Thermo Scientific) according to standard procedures. 18S RNA

was used for normalization. All PCR products were sequenced for confirmation. All primers used are listed in Additional file 12: Table S6.

Microscopy and calculating the relative fluorescence intensity

For all the lincRNAs reporter worms, Axio Scope A1 compound microscope (Zeiss, Oberkochen, Germany) was used for the examination of fluorescence. L4 stage worms were anesthetized in 10 mM sodium azide, and images were taken using the $\times 20$ objective. All the images were analyzed by the ImageJ (an open-source image processing software). Confocal imaging was carried out as previously reported with some modification [32]. Imaging of anesthetized worms were carried out on Andor Revolution XD laser confocal microscope system (Andor Technology PLC) based on a spinning-disk confocal scanning head CSU-X1 (Yokogawa Electric Corporation) under control of Andor IQ 10.1 software or two-photon confocal laser scanning microscopy FV1200MPE (Olympus) with GaAsP-NDD detector. Z-stack images were obtained on Olympus IX-71 inverted microscope (Olympus Corporation) with $\times 60$ 1.45 NA oil-immersion objective. An Andor iXonEM+ DV897K EM CCD camera was used for capturing the 14-bit digital images with Andor LC-401A Laser Combiner with diode-pumped solid state (DPSS) lasers, emissions at 458 nm, 488 nm, 515 nm, and 561 nm.

Counting the presynaptic puncta in ventral and dorsal

Dorsal nerve cord and ventral nerve cord images were obtained and counted between VD9 and VD11. ImageJ plot profile tool was used to plot nerve cords, and the number of SNB-1::GFP (Punc-25::snb-1::gfp) puncta was calculated by counting the number of crests of the plot file ($n = 4$).

RNA sequencing

For next-generation RNA sequencing, total RNAs were isolated from nine different stages of worms (embryos, L1, L2, dauer, L3, L4, young adult, male, and mix stage with starvation). Sequencing libraries were carried out as previously described with modifications [53]. Whole transcriptome libraries were constructed by the TruSeq Ribo Profile Library Prep Kit (Illumina, USA), according to the manufacturer's instructions. In brief, 10 μ g total RNA was depleted rRNA with an Illumina Ribo-Zero Gold kit and purified for end repair and 5'-adaptor ligation. Then, reverse transcription was performed with random primers containing 3' adaptor sequences and randomized hexamers. The cDNAs were purified and amplified, and PCR products of 200–500 bp were purified, quantified, and stored at -80°C until sequencing. For RNA sequencing of long RNAs, the

libraries were prepared according to the manufacturer's instructions and subjected to 150 nt paired-end sequencing with an Illumina HiSeq 2500 system (Novogene, China). We sequenced each library to a depth of 10–50 million read pairs, and the reads were mapped to the *C. elegans* genome (ce11). For small RNA (sRNA) sequencing, nine sRNA libraries were generated with TruSeq small RNA (Illumina, USA) according to the manufacturer's instructions. Then, the prepared libraries were sequenced with an Illumina Nextseq 500 system (Novogene, China). After filtering out the reads shorter than 15 nt, the remaining reads were mapped to the *C. elegans* genome (ce11) and the miRNA database in miR-Base with bowtie (-v 1).

Conservation, length, and exon number analysis of lincRNAs

For the genome-wide feature analysis of lincRNAs, the control was 200 mRNAs randomly picked from *C. elegans* transcriptome. The information of length and exon number for lincRNAs and mRNAs was extracted from the annotation of *C. elegans* [61]. For the analysis of sequence conservation, we interrogated 26 nematode conservation phastCons scores from UCSC [61] for each base of individual *C. elegans* lincRNA or mRNA and averaged the scores of each transcript. The distribution of lincRNA and mRNA was compared by two-sided Mann-Whitney *U* test.

Construction of lincRNA-miRNA co-expression network

Functional networks of miRNA and lincRNA pairs were illustrated with cytoscape v3.5.1 [62]. For the one-to-one connection, the expression of lincRNA with at least two 7-mers matches of particular miRNA was negatively correlated to the expression of miRNA (Pearson $R < -0.1$) across nine stages.

Chromatin immunoprecipitation (ChIP)

ChIP assays were performed as described in our previous report with modifications [23]. N2 and linc-73 mutant worms were bleached with hypochlorite solution, and the eggs were incubated at 20°C on NGM plates seeded with OP50 to be synchronized to L2 (for ChIP-qPCR experiments) or L4 stage (for ChIP-seq experiments). Synchronized worms were then washed with three changes of M9 buffer and fixed with 2% formaldehyde for 35 min followed by stopping with 100 mM Tris pH 7.5 for 2 min. Worm pellets were washed with FA buffer supplemented with 10 μ l 1 M DTT, 50 μ l 0.1 M PMSE, 100 μ l 10% SDS, 500 μ l 20% N-Lavroyl sarcosine sodium, and 2 tablets protease inhibitors in 10 ml FA buffer. Worms were sonicated on ice for 15 min with the setting of high power, 4°C , and 15 cycles, 30 s on, 30 s off. The tubes were then spun at 14,000 g for 10 min at 4°C . The supernatant was carefully

removed into new tubes, and an aliquot (5% of each sample) was taken as input. Prewashed salmon sperm Protein G beads were added to the supernatant for 1 h for pre-cleaning. Beads were discarded, and 2 μ g anti-H3K4me3 or anti-H3K9me3 (Abcam) were added to each tube overnight at 4 °C. The beads were washed twice with 150 mM NaCl FA buffer for 5 min each, washed once with 1 M NaCl FA buffer for 5 min, twice with 500 mM NaCl FA buffer for 10 min, once with TEL buffer (0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl, pH 8.0) for 10 min, and finally with three changes of 1X TE buffer (1 M Tris-HCl, 0.5 M EDTA). DNA-protein complexes were eluted in 200 μ l of ChIP elution buffer (1% SDS in TE with 250 mM NaCl) and incubated at 65 °C for 20 min with regular shaking every 5–10 min. Both samples and inputs were treated with RNase A (2 μ g/ μ l) and proteinase K (2 μ g/ μ l) for 2 h at 55 °C for 1 h and then reverse cross-linked at 65 °C overnight. DNA was purified by phenol/chloroform/isoamyl extraction and then used for ChIP-qPCR or ChIP-seq. For ChIP-seq, DNA from ChIP (along with the input) was iron fragmented at 95 °C followed by end repair and 5' adaptor ligation, then purified and amplified. PCR products corresponding to 200–500 bps were purified for sequencing. Illumina Nextseq 500 system for 150 nt pair-end sequencing was then performed (Novogene).

Analysis of ChIP-seq data of transcription factors and H3K4me3 and H3K9me3

A total of 774 ChIP-seq raw fastq data and 561 computed gff3-file data were downloaded from modENCODE (<ftp://data.modencode.org/C.elegans/Transcriptional-Factor/ChIP-seq/>) [30, 31], and the regulation patterns of all transcriptional factor genes by lincRNAs in *C. elegans* were analyzed. The quality of all these 774 raw fastq data was verified using bowtie2 to map the reads to the *C. elegans* genome (ce11). We then re-analyzed the calculated peaks in order to investigate the regulation of transcriptional factors by the various lincRNAs. Considering the shorter length of the lincRNAs transcripts as compared to the mRNAs, we used the scale within 1 kb upstream or 200 bp downstream of the transcription start site of the lincRNAs. ChIP-seq data of UNC-30::GFP and UNC-55::GFP from our previous study using endogenous GFP knock-in *unc-30* and *unc-55* mutant worms were also analyzed (GEO: GSE102213) [23]. Reads were first filtered from genomic repeats, and the unique reads were then mapped to the *C. elegans* genome (ce11) with bowtie2. Peaks of UNC-30 and UNC-55 were assigned by the cis-Genome with default parameters (cutoff > 3 and *p* value < 10⁻⁵). H3K4me3 and H3K9me3 ChIP-seq data of L4 were mapped to the *C. elegans* genome (ce11) with bowtie2 using the default parameters. Samtools were used to filter

*sam files and remove duplicated reads. Macs2 was used for peak calling (*q* parameter was set as 0.001).

Short Time-series Expression Miner analysis (STEM)

The co-expression patterns of lincRNAs and mRNAs were calculated by STEM [2 (a software program designed for clustering comparing, and visualizing gene expression data from short time series experiments) using RNA-seq data from nine different developmental stages. RNA-seq data from the embryonic stage was set as 0 point, and the other developmental stages were normalized to the embryonic stage data. *K*-means method was used to cluster the genes into specific profile according to their expression pattern. In all, nearly 20,000 genes with reads per kilo million (RPKM) greater than 1 were clustered into 10 profiles according changes in their expression patterns at different stages of development. The function of genes in specific clusters with similar expression patterns was analyzed by gene ontology analysis.

GO analysis

The significant enriched genes were analyzed with Gorilla web-server [63]. *P* values were calculated with default parameters.

Statistical analysis

For Student's *t* tests, the values reported in the graphs represent averages of independent experiments, with error bars showing s.e.m. in all figures, except for Fig. 5 and Additional file 7: Figure S3, in which error bars show S.D. Statistical methods are also indicated in the figure legends. All statistical significances were determined using GraphPad Prism software (version 7). Two-sided Mann-Whitney *U* test was used in Figs. 1d, e and 7g–i and Additional file 7: Figure S3. Unpaired Student's *t* test was used in Figs. 2b–f, 5, 6b–f, and 8c–g and Additional file 7: Figure S3. Chi-square test was used in Figs. 2g and 6g.

Additional files

Additional file 1: Table S1. Expression profiles of lincRNAs. (XLSX 68 kb)

Additional file 2: Figure S1. Schematic diagram of the CRISPR/cas9 knockout system used in this study. (DOCX 117 kb)

Additional file 3: Table S2. List of worm strains. (XLSX 20 kb)

Additional file 4: Table S3. Data for the lincRNA mutant phenotypes. (XLSX 66 kb)

Additional file 5: Figure S2. Correlation of the expression levels of lincRNAs and their neighboring protein-coding genes. a Relative expression levels of 170 *C. elegans* lincRNAs and their neighboring mRNAs (100 kb upstream and downstream of the lincRNA locus). b Relative expression levels of 23 lincRNAs with distinct phenotypic characteristics and their neighboring mRNAs (100 kb upstream and downstream of the lincRNA locus) (DOCX 111 kb)

Additional file 6: Table S4. lincRNA-microRNA interactions. (XLSX 10 kb)

Additional file 7: Figure S3. miRNA regulation of lincRNAs. **a** Relative GFP expression levels of *linc-126* in N2 worms with or without overexpression of miR-4938 ($n = 20$). **b** Relative GFP expression levels of *linc-109* in N2 worms with or without overexpression of miR-5546 ($n = 20$). Data are the means \pm SD. ns, not significant by the Student's *t*-test. Images shown are representative of the control and experimental groups. Scale bar, 20 μ m (DOCX 84 kb)

Additional file 8: Table S5. Data for the rescue experiments. (XLSX 12 kb)

Additional file 9: Figure S4. Numbers of transcription factors regulating the 23 lincRNAs with phenotypes in this study and the other 147 lincRNAs in embryos (**a**), L4 worms (**b**), and young adults (**c**). ns, no significant difference, by the two-sided Mann-Whitney U test (DOCX 46 kb)

Additional file 10: Figure S5. Shared UNC-30 and UNC-55 lincRNA targets with peak patterns in both UNC-30 and UNC-55 ChIP-seq. **a** *linc-5*, **b** *linc-58*, **c** *linc-146*, **d** *linc-149*, **e** *linc-152*. (DOCX 75 kb)

Additional file 11: Figure S6. Expression of *Punc-104::mCherry* in the cell body of D mns. Representative images of data presented in Fig. 8f. Scale bar, 25 μ m (DOCX 100 kb)

Additional file 12: Table S6. Primers used in this study. (XLSX 39 kb)

Abbreviations

C. elegans: *Caenorhabditis elegans*; ChIP-seq: Chromatin immunoprecipitation sequencing; CRISPR: Clustered regularly interspaced short palindromic repeats; D mns: D motor neurons; GO: Gene ontology; KO: Knockout; lincRNA: Long intergenic noncoding RNA; PRC2: Polycomb repressive complex2; RPKM: Reads per Kilobase per Million mapped reads

Acknowledgements

We thank Dr. Guangshuo Ou for providing plasmids, Dr. Xiao Liu for providing strains used in this work, and Dr. Shouhong Guang for providing experimental facilities. We thank the Bioinformatics Center of the USTC, School of Life Sciences, for providing supercomputing resources.

Funding

This work was supported by the National Basic Research Program of China (2015CB943000), the National Key R&D Program of China (2018YFC1004500), the National Natural Science Foundation of China (31725016 and 31471225), the Major/Innovative Program of Development Foundation of Hefei Center for Physical Science and Technology (2016FXCX006), and the Open Project of the CAS Key Laboratory of Innate Immunity and Chronic Disease (KLIICD-201603), and the Strategic Priority Research Program (Pilot study) "Biological basis of aging and therapeutic strategies" of the Chinese Academy of Sciences (XDPB10).

Availability of data and materials

The 9 different stages of lincRNAs and microRNAs sequencing data as well as H3K4me3 and H3K9me3 ChIP-seq data generated in this study have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO accession number GSE115324. Accession number for ChIP-seq data of UNC-30 and UNC-55 is GSE102213. ChIP-seq data of transcription factors at different stages of development are publicly available at <http://www.modencode.org/> [30, 31].

Authors' contributions

GS designed and initiated this project, provided the major funding, and supervised the experiments. SW, HC, EED, TF, BY, XW, JL, LL, SF, and WL performed experiments. SW, HC, EED, and GS analyzed the data. GS, SW, EED, and HC wrote the manuscript. All authors have discussed the results and made comments on the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Division of Molecular Medicine, Hefei National Laboratory for Physical Sciences at Microscale, the CAS Key Laboratory of Innate Immunity and Chronic Disease, School of Life Sciences, University of Science and Technology of China, Hefei 230027, China. ²Department of Molecular Biology and Biotechnology, School of Biological Sciences, College of Agriculture and Natural Sciences, University of Cape Coast, Cape Coast, Ghana. ³MOE Key Laboratory of Bioinformatics, Center for Synthetic and Systems Biology, School of Life Sciences, Tsinghua University, Beijing 100084, China. ⁴Present address: Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309, USA. ⁵Present address: Hanwang Technology Co., Ltd., Haidian District, Beijing 100193, China. ⁶CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, CAS, Shanghai 200031, China.

Received: 22 November 2018 Accepted: 27 December 2018

Published online: 08 January 2019

References

- Ransohoff JD, Wei Y, Khavari PA. The functions and unique features of long intergenic non-coding RNA. *Nat Rev Mol Cell Biol.* 2018;19(3):143–57.
- Popadin K, Gutierrez-Arcelus M, Dermitzakis ET, Antonarakis SE. Genetic and epigenetic regulation of human lincRNA gene expression. *Am J Hum Genet.* 2013;93(6):1015–26.
- Kopp F, Mendell JT. Functional classification and experimental dissection of long noncoding RNAs. *Cell.* 2018;172(3):393–407.
- Batista PJ, Chang HY. Long noncoding RNAs: cellular address codes in development and disease. *Cell.* 2013;152(6):1298–307.
- Gomez JA, Wapinski OL, Yang YW, Bureau JF, Gopinath S, Monack DM, Chang HY, Brahic M, Kirkegaard K. The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon- γ locus. *Cell.* 2013; 152(4):743–54.
- Rutenberg-Schoenberg M, Sexton AN, Simon MD. The properties of long noncoding RNAs that regulate chromatin. *Annu Rev Genomics Hum Genet.* 2016;17:69–94.
- Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai M-C, Hung T, Argani P, Rinn JL. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature.* 2010;464(7291):1071.
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Bruggmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell.* 2007;129(7):1311–23.
- Yang F, Zhang H, Mei Y, Wu M. Reciprocal regulation of HIF-1 α and lincRNA-p21 modulates the Warburg effect. *Mol Cell.* 2014;53(1):88–100.
- Guo X, Xu Y, Wang Z, Wu Y, Chen J, Wang G, Lu C, Jia W, Xi J, Zhu S. A linc1405/Eomes complex promotes cardiac mesoderm specification and cardiogenesis. *Cell Stem Cell.* 2018;22(6):893–908.
- Atianand MK, Hu W, Satpathy AT, Shen Y, Ricci EP, Alvarez-Dominguez JR, Bhatta A, Schattgen SA, McGowan JD, Blin J, et al. A long noncoding RNA lincRNA-EP5 acts as a transcriptional brake to restrain inflammation. *Cell.* 2016;165(7):1672–85.
- Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. *Cell.* 2013;154(1):26–46.
- Young RS, Marques AC, Tibbit C, Haerty W, Bassett AR, Liu J-L, Ponting CP. Identification and properties of 1,119 candidate lincRNA loci in the *Drosophila melanogaster* genome. *Genome Biol Evol.* 2012;4(4):427–42.
- Pauli A, Valen E, Lin MF, Garber M, Vastenhouw NL, Levin JZ, Fan L, Sandelin A, Rinn JL, Regev A. Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome Res.* 2012;22(3):577–91.
- Nam J-W, Bartel DP. Long noncoding RNAs in *C. elegans*. *Genome Res.* 2012;22(12):2529–40.
- Wormbase Version:264 [<https://wormbase.org/>] Accessed 10 Oct 2016.
- Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc.* 2013;8(11):2281.

18. Burstein D, Harrington LB, Strutt SC, Probst AJ, Anantharaman K, Thomas BC, Doudna JA, Banfield JF. New CRISPR–Cas systems from uncultivated microbes. *Nature*. 2017;542(7640):237.
19. Zhao P, Zhang Z, Lv X, Zhao X, Suehiro Y, Jiang Y, Wang X, Mitani S, Gong H, Xue D. One-step homozygosity in precise gene editing by an improved CRISPR/Cas9 system. *Cell Res*. 2016;26(5):633.
20. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini L. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339(6121):819–23.
21. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR–Cas9 for genome engineering. *Cell*. 2014;157(6):1262–78.
22. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. 2013;152(5):1173–83.
23. Yu B, Wang X, Wei S, Fu T, Dzakah EE, Waqas A, Walthall WW, Shan G. Convergent transcriptional programs regulate cAMP levels in *C. elegans* GABAergic motor neurons. *Dev Cell*. 2017;43(2):212–26 e217. <https://doi.org/10.1016/j.devcel.2017.09.013>.
24. Ernst J, Bar-Joseph Z. STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinformatics*. 2006;7:191.
25. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta stone of a hidden RNA language? *Cell*. 2011;146(3):353–8.
26. Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. *Nature*. 2014;505(7483):344–52.
27. Kumar MS, Armenteros-Monterroso E, East P, Chakravorty P, Matthews N, Winslow MM, Downward J. HMG2A functions as a competing endogenous RNA to promote lung cancer progression. *Nature*. 2014;505(7482):212–7.
28. Fujimoto A, Furuta M, Totoki Y, Tsunoda T, Kato M, Shiraishi Y, Tanaka H, Taniguchi H, Kawakami Y, Ueno M, et al. Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer. *Nat Genet*. 2016;48(5):500–9.
29. Hajarnis SS, Patel V, Aboudehen K, Attanasio M, Cobo-Stark P, Pontoglio M, Igarashi P. Transcription factor hepatocyte nuclear factor-1beta (HNF-1beta) regulates microRNA-200 expression through a long noncoding RNA. *J Biol Chem*. 2015;290(41):24793–805.
30. The National Human Genome Research Institute model organism Encyclopedia Of DNA Elements [<http://www.modencode.org/>] Accessed 10 May 2018.
31. Niu W, Lu ZJ, Zhong M, Sarov M, Murray JI, Brdlik CM, Janette J, Chen C, Alves P, Preston E, Slightam C, Jiang L, Hyman AA, Kim SK, Waterston RH, Gerstein M, Snyder M, Reinke V. Diverse transcription factor binding features revealed by genome-wide ChIP-seq in *C. elegans*. *Genome Res*. 2011;21(2):245–54 <https://doi.org/10.1101/gr.114587.110>.
32. Shan G, Kim K, Li C, Walthall W. Convergent genetic programs regulate similarities and differences between related motor neuron classes in *Caenorhabditis elegans*. *Dev Biol*. 2005;280(2):494–503.
33. Westmoreland JJ, McEwen J, Moore BA, Jin Y, Condie BG. Conserved function of *Caenorhabditis elegans* UNC-30 and mouse Pitx2 in controlling GABAergic neuron differentiation. *J Neurosci*. 2001;21(17):6810–9.
34. Jin Y, Hoskins R, Horvitz HR. Control of type-D GABAergic neuron differentiation by *C. elegans* UNC-30 homeodomain protein. *Nature*. 1994;372(6508):780–3.
35. Niwa S, Lipton DM, Morikawa M, Zhao C, Hirokawa N, Lu H, Shen K. Autoinhibition of a neuronal kinesin UNC-104/KIF1A regulates the size and density of synapses. *Cell Rep*. 2016;16(8):2129–41.
36. Otsuka AJ, Jeyaprakash A, Garcia-Anoveros J, Tang LZ, Fisk G, Hartshorne T, Franco R, Born T. The *C. elegans* unc-104 gene encodes a putative kinesin heavy chain-like protein. *Neuron*. 1991;6(1):113–22.
37. Hall DH, Hedgecock EM. Kinesin-related gene unc-104 is required for axonal transport of synaptic vesicles in *C. elegans*. *Cell*. 1991;65(5):837–47.
38. Perry RB, Hezroni H, Goldrich MJ, Ulitsky I. Regulation of neuroregeneration by long noncoding RNAs. *Mol Cell*. 2018;72(3):553–67.
39. Yu B, Shan G. Functions of long noncoding RNAs in the nucleus. *Nucleus*. 2016;7(2):155–66.
40. Rashid F, Shah A, Shan G. Long non-coding RNAs in the cytoplasm. *Genomics, Proteomics Bioinformatics*. 2016;14(2):73–80.
41. Liu SJ, Nowakowski TJ, Pollen AA, Lui JH, Horlbeck MA, Attenello FJ, He D, Weissman JS, Kriegstein AR, Diaz AA, et al. Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biol*. 2016;17:67.
42. Liu SJ, Horlbeck MA, Cho SW, Birk HS, Malatesta M, He D, Attenello FJ, Villalta JE, Cho MY, Chen Y et al: CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science*. 2017; 355(6320):aah7111.
43. Cho SW, Xu J, Sun R, Mumbach MR, Carter AC, Chen YG, Yost KE, Kim J, He J, Nevins SA, et al. Promoter of lncRNA gene PVT1 is a tumor-suppressor DNA boundary element. *Cell*. 2018;173(6):1398–412.
44. Goyal A, Fiškin E, Gutschner T, Polycarpou-Schwarz M, Groß M, Neugebauer J, Gandhi M, Caudron-Herger M, Benes V, Diederichs S. A cautionary tale of sense-antisense gene pairs: independent regulation despite inverse correlation of expression. *Nucleic Acids Res*. 2017;45(21):12496–508.
45. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, Bennett CF, Sharma A, Bubulya PA, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell*. 2010;39(6):925–38.
46. Miyagawa R, Tano K, Mizuno R, Nakamura Y, Ijiri K, Rakwal R, Shibato J, Masuo Y, Mayeda A, Hirose T, et al. Identification of cis- and trans-acting factors involved in the localization of MALAT-1 noncoding RNA to nuclear speckles. *RNA*. 2012;18(4):738–51.
47. Zhang B, Arun G, Mao YS, Lazar Z, Hung G, Bhattacharjee G, Xiao X, Booth CJ, Wu J, Zhang C, et al. The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Rep*. 2012;2(1):111–23.
48. Nakagawa S, Ip JY, Shioi G, Tripathi V, Zong X, Hirose T, Prasanth KV. Malat1 is not an essential component of nuclear speckles in mice. *RNA*. 2012;18(8):1487–99.
49. Amandio AR, Necsculea A, Joye E, Mascrez B, Duboule D. Hotair is dispensable for mouse development. *PLoS Genet*. 2016;12(12):e1006232.
50. Li L, Liu B, Wapinski OL, Tsai MC, Qu K, Zhang J, Carlson JC, Lin M, Fang F, Gupta RA, et al. Targeted disruption of Hotair leads to homeotic transformation and gene derepression. *Cell Rep*. 2013;5(1):3–12.
51. Sauvageau M, Goff LA, Lodato S, Bonev B, Groff AF, Gerhardinger C, Sanchez-Gomez DB, Hacisuleyman E, Li E, Spence M, et al. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. *Elife*. 2013;2:e01749.
52. Bassett AR, Akhtar A, Barlow DP, Bird AP, Brockdorff N, Duboule D, Ephrussi A, Ferguson-Smith AC, Gingeras TR, Haerty W, Higgs DR, Miska EA, Ponting CP. Considerations when investigating lncRNA function in vivo. *Elife*. 2014;3: e03058.
53. Hu S, Wang X, Shan G. Insertion of an Alu element in an lncRNA leads to primate specific modulation of alternative splicing. *Nat Struct Mol Biol*. 2016; 23(11):1011–9.
54. Brenner S. The genetics of *Caenorhabditis elegans*. *Genetics*. 1974;77(1):71–94.
55. Miska EA, Alvarez-Saavedra E, Abbott AL, Lau NC, Hellman AB, McGonagle SM, Bartel DP, Ambros VR, Horvitz HR. Most *Caenorhabditis elegans* microRNAs are individually not essential for development or viability. *PLoS Genet*. 2007;3(12):e215.
56. Thomas JH. Genetic analysis of defecation in *Caenorhabditis elegans*. *Genetics*. 1990;124(4):855–72.
57. Raizen DM, Lee RY, Avery L. Interacting genes required for pharyngeal excitation by motor neuron MC in *Caenorhabditis elegans*. *Genetics*. 1995; 141(4):1365–82.
58. Ringstad N, Horvitz HR. FMRFamide neuropeptides and acetylcholine synergistically inhibit egg-laying by *C. elegans*. *Nat Neurosci*. 2008;11(10): 1168–76.
59. Zeller P, Padeken J, van Schendel R, Kalck W, Tijsterman M, Gasser SM. Histone H3K9 methylation is dispensable for *Caenorhabditis elegans* development but suppresses RNA:DNA hybrid-associated repeat instability. *Nat Genet*. 2016;48(11):1385–95.
60. Boag PR, Nakamura A, Blackwell TK. A conserved RNA-protein complex component involved in physiological germline apoptosis regulation in *C. elegans*. *Development*. 2005;132(22):4975–86.
61. UCSC [<http://hgdownload.soe.ucsc.edu/goldenPath/ce11/>]. Accessed 28 Mar 2018.
62. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–504.
63. Eran E, Roy N, Israel S, Doron L, Zohar Y. Gorilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*. 2009;10:48.