

The microRNA miR-124 controls gene expression in the sensory nervous system of *Caenorhabditis elegans*

Alejandra M. Clark^{1,2}, Leonard D. Goldstein^{3,4,5}, Maya Tevlin⁶, Simon Tavaré^{3,4,5}, Shai Shaham^{6,*} and Eric A. Miska^{1,2,*}

¹Wellcome Trust Cancer Research UK Gurdon Institute, University of Cambridge, The Henry Wellcome Building of Cancer and Developmental Biology, Tennis Court Rd, Cambridge CB2 1QN, ²Department of Biochemistry, University of Cambridge, Tennis Court Rd, Cambridge CB2 1GA, ³Department of Oncology, University of Cambridge, Hills Rd, Cambridge CB2 2XZ, ⁴Cancer Research UK, Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, ⁵Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Rd, Cambridge CB3 0WA, UK and ⁶Laboratory of Developmental Genetics, 1033 Weiss Research Building, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA

Received November 6, 2009; Revised and Accepted January 28, 2010

ABSTRACT

miR-124 is a highly conserved microRNA (miRNA) whose *in vivo* function is poorly understood. Here, we identify miR-124 targets based on the analysis of the first *mir-124* mutant in any organism. We find that miR-124 is expressed in many sensory neurons in *Caenorhabditis elegans* and onset of expression coincides with neuronal morphogenesis. We analyzed the transcriptome of miR-124 expressing and nonexpressing cells from wild-type and *mir-124* mutants. We observe that many targets are co-expressed with and actively repressed by miR-124. These targets are expressed at reduced relative levels in sensory neurons compared to the rest of the animal. Our data from *mir-124* mutant animals show that this effect is due to a large extent to the activity of miR-124. Genes with nonconserved target sites show reduced absolute expression levels in sensory neurons. In contrast, absolute expression levels of genes with conserved sites are comparable to control genes, suggesting a tuning function for many of these targets. We conclude that miR-124 contributes to defining cell-type-specific gene activity by repressing a diverse set of co-expressed genes.

INTRODUCTION

Gene regulation plays a key role in development. While the significance of transcriptional regulation has long been recognized, the potential of post-transcriptional gene regulation mediated by various classes of noncoding small RNAs is beginning to be unravelled (1). microRNAs (miRNAs) are a widespread class of noncoding ~22 nt endogenous RNAs found in animals, plants and algae (2–8). These RNAs modulate gene expression by blocking translation and/or destabilizing target mRNAs (6). The first miRNAs described, *lin-4* and *let-7*, were discovered as entities, mutations in which alter developmental timing in the nematode *Caenorhabditis elegans* (9–11). Since then, a number of approaches, including reverse and forward genetics, have identified functions for miRNAs in animal and plant development, homeostasis and disease (12–20).

Although new sequencing technologies have resulted in a dramatic increase in the number of known miRNAs (21,22), the functions of the majority of miRNAs remain unknown. One approach to get at miRNA function is to identify direct targets. While in plants this task has been facilitated by the high level of complementarity between miRNAs and their targets (23), identification of physiological targets in animals has remained a computational challenge. Animal 3'UTRs often contain short sequence motifs that are complementary to the 5'-region of the

*To whom correspondence should be addressed. Tel: +44 1223 767 220; Fax: +44 1223 334 089; Email: eam29@cam.ac.uk
Correspondence may also be addressed to Shai Shaham. Fax: +1 212 327 7129; Email: shaham@rockefeller.edu

The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors.

miRNA, termed the seed sequence (nucleotides 2–7), which is thought to be the main determinant of miRNA target specificity. These sequence motifs have been conserved during evolution at higher rates than expected by chance (24–27). A 3'UTR match to the miRNA seed sequence can be sufficient for miRNA-mediated repression (24,28). The vast majority of *in vivo* validated miRNA interaction sites are also found in the 3'UTR (9,11,12,29,30). Hence, most existing computational methods for miRNA target prediction are based on the identification of conserved 3'UTR seed matches.

Computational studies suggest that single miRNAs may bind hundreds of targets. Indeed, more than half of all human mRNAs may be under positive selection to maintain miRNA target sites (31). Early experimental confirmation of these hypotheses came from miRNA overexpression studies in cell lines, which demonstrated that hundreds of mRNAs were subtly downregulated in response to ectopic miRNA expression (32). Although initially miRNAs were thought to act predominantly at the translational level, recent proteomics studies suggest that changes in the abundance of transcript and protein are highly correlated and of comparable magnitude (33,34).

Beyond the question of which mRNAs are biologically relevant miRNA targets, general questions about the mode of miRNA function, such as the extent of co-expression of a miRNA and its targets, have remained unanswered. Knowing whether miRNA and target expression is overlapping or not can be useful in elucidating the function of miRNA-dependent target regulation. At the two extremes, overlapping and mutually exclusive expression suggest a tuning and switch-like role for the miRNA, respectively. A study in *Drosophila*, based on mRNA *in situ* data, suggested mutually exclusive expression of miRNAs and their targets (35). In contrast, microarray studies of mRNA expression across different human and mouse tissues demonstrated that miRNA targets are often co-expressed with the miRNA, but at reduced relative levels compared to tissues or developmental time points where the miRNA is absent (36,37). A further computational study concluded that miRNAs and their targets are often positively or negatively co-regulated (38). Together, these studies suggested that miRNAs contribute to tissue-specific mRNA expression. It remained unclear however whether reduced relative expression of miRNA targets in the presence of the miRNA is due to direct miRNA-mediated repression or other regulatory mechanisms acting in concert with the miRNA. More recently, Shkumatava *et al.* (39) and Mishima *et al.* (40) addressed this issue by analyzing mRNA expression in sorted cell populations from wild-type and MZ*dicer* mutant zebrafish.

The miRNA miR-124 provides an excellent opportunity for investigating the mode of action of miRNAs: it is highly conserved and tissue specific, and found in the nervous system of all animals studied to date (41–47). Although several studies have aimed to understand the function of miR-124 in neuronal development, experiments have largely relied on knockdown or overexpression of miR-124 in cell culture. Overexpression of miR-124 in HeLa cells shifts the gene expression profile towards a

brain-like pattern (32) and overexpression of miR-124 in neuroblastoma cell lines and embryonic stem cells can lead to induction of neuronal differentiation (48,49). miR-124 *in vivo* knockdown experiments in the developing chick spinal cord led either to no effect (50) or modest effects on neuronal differentiation (51). In mice, miR-124 knockdown experiments resulted in defects in adult neurogenesis (52).

Here, we examine the mode of action of *C. elegans* miR-124 by studying the impact of miR-124 deletion on the transcriptome of cells expressing the miRNA, which we identify as a subset of sensory neurons. We show that genes upregulated in *mir-124* mutant sensory neurons are enriched for likely direct miR-124 targets. We observe that targets are expressed at reduced relative levels in wild-type sensory neurons compared to the rest of the animal, and that these differences are largely due to direct miRNA-mediated repression. These results suggest that miR-124 contributes to defining gene expression in sensory neurons by regulating a large number of co-expressed genes.

MATERIALS AND METHODS

Strains

Caenorhabditis elegans strains were cultured using standard methods (53) at 20° C. The following strains, CX3716, CX3695, PY2417, CX3553, BZ555, CX3465, DA1262 were provided by the *Caenorhabditis* Genetics Center, which is funded by the NIH National Center for Research Resources (NCRR). For a full list of all strains used in this study see Supplementary Table S1.

Plasmid constructs

Oligos 5'-CGTTAGATTGCTTCTTC-3' and 5'-GGAGAAGAGAGCACTTGAAG-3' were used to amplify ~2 kb of genomic sequence upstream of *mir-124*. GFP was amplified from pPD95.75 and *mir-124* promoter::GFP (*pmir-124::GFP*) generated by PCR fusion (54). *mir-124* promoter::mCherry (*pmir-124::mCherry*) was generated by PCR fusion using the same oligos for *mir-124* promoter indicated above.

Transgenic strains

Germline transformations were carried out as described in ref. (55). *pmir-124::GFP* (5 ng/μl) was coinjected with pEM27 (*plin-15* rescue) into *lin-15* (*n765*) worms, and *pmir-124::mcherry* (5 ng/μl) was injected into N2. Extrachromosomal transgenes [*pmir-124::GFP* + *lin-15*(+)] were integrated by X-ray irradiation and subsequently outcrossed twice into N2 background.

Cell dissociation and cell culture

Embryos were obtained from synchronized populations of SX620 and SX621 worms and dissociated as previously described (56–58). Briefly, 500 000 gravid adults (grown at 20°C in 5 × 15 cm plates) were lysed in hypochlorite solution to release embryos. Embryos were washed in egg buffer containing 118 mM NaCl, 48 mM KCl, 2 mM CaCl₂, 2 mM MgCl₂ and 25 mM HEPES (pH 7.3) and

subsequently treated with Chitinase 1 U/ml (Sigma C-7809) for 45 min to digest egg shells. Embryo pellets were washed in egg buffer and incubated in Trypsin (GIBCO) for 20 mins prior to pipetting several times to facilitate cell dissociation. Dissociated cells were purified by passage through 5 μ m filters (Durapore) using a 10 ml syringe and resuspended in L-15 medium (GIBCO) supplemented with 10% FBS (HYCLONE), 50 U/ml penicillin and 50 μ g/ml streptomycin, plus sucrose for osmolarity. Embryonic cells were plated on poly-L-Lysine (0.01%, Sigma) coated culture dishes at 10^7 cells/ml and cultured for 24 h.

Fluorescence activated cell sorting

Sorting experiments were carried out as previously described in ref. (56) with the following modifications: A BD FACSAriaII (BD Biosciences) equipped with a 50 mw 488 nm laser was used to sort 500 000 cells for each sample directly into extraction buffer (Absolutely RNA microprep kit, Stratagene). Purity of the sorted cells as assessed by resorting GFP+ and GFP- cells after first sort was 74.4% and 91.4%, respectively.

Isolation and amplification of mRNA

mRNA was extracted from 500 000 cells by using the Absolutely RNA microprep kit (Stratagene), yielding ~15 ng mRNA that was subsequently amplified at GeneCore Facility (EMBL). qRT-PCR for miRNAs was carried out as previously described (59,60). For all oligos used, see Supplementary Table 3.

Array hybridization

mRNA isolated and amplified from three biological experiments was hybridized to the Affymetrix GeneChip for *C. elegans*. In total, 12 samples were hybridized: three biological replicates of GFP+ and GFP- cells from both wild-type and *mir-124* mutant animals. Arrays were scanned on an Affymetrix 3000 7G scanner.

Microarray data analysis

Affymetrix CEL files were read into the statistical programming environment R (61). Array quality control was performed using packages available from Bioconductor (62). Arrays were processed using the RMA function (63), performing background correction, quantile normalization and probe set summary, and further analyzed using the *limma* package (64). Since the quality of arrays varied, array quality weights were obtained and used in the linear model fit (65). Probe sets were mapped to Entrez genes using the *celegans.db* package available from Bioconductor. When we used expression values for individual genes, these were obtained from the linear model fit. In the case of multiple probe sets mapping to the same gene, values were summarized by the arithmetic mean (on the \log_2 scale). Array data were submitted to the Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE16050.

3'UTR sequences

A 6-way multiple genome alignment between *C. elegans*, *C. briggsae*, *C. remanei*, *C. brenneri*, *C. japonico* and *Pristionchus pacificus* based on *C. elegans* genome assembly ce6 was downloaded from the UCSC Genome Browser website (<http://genome.ucsc.edu/>) (66,67). Multiple alignments of 3'UTR sequences were extracted based on coordinates in the refGene table (68). In the case of multiple entries for the same RefSeq identifier, the entry with longest sequence was chosen (if there were multiple entries with maximal length, one was chosen at random).

Target prediction

Caenorhabditis elegans 3'UTR sequences were scanned for perfect matches against the miR-124 seed sequence (nucleotides 2–7), and the identity of one upstream and one downstream flanking nucleotide was retained for each seed match. A seed match was considered a putative target site if it was flanked by an adenosine opposite miRNA nucleotide 1 (position t1) or flanked by a perfect match to miRNA nucleotide 8 (position t8), or both. When assessing the expression of miR-124 targets as a class, only those transcripts with a 3'UTR seed match flanked by a t8 match were included in the analysis. Target sites (defined as a seed match and flanking positions t1, t8) were considered conserved if they overlapped target sites in the aligned sequences of *C. briggsae*, *C. remanei* and *C. brenneri* by one or more nucleotides. Predictions are listed in Supplementary Table 4.

Gene annotation

Analysis of the gene expression data was based on Entrez genes. An Entrez gene was considered a (conserved) target if at least one of the associated RefSeq genes contained a (conserved) 3'UTR seed match. For the seed match type analysis in Figure 4J, an Entrez gene was considered to represent a given seed match type if all associated RefSeq genes with available 3'UTR contained a single 3'UTR seed match of that type. For the Sylamer analysis and computation of probabilities for random site occurrences, each Entrez gene was assigned a representative RefSeq gene by choosing the gene with longest annotated 3'UTR.

Sylamer analysis

For each comparison, genes were ranked from most increased to most decreased based on the B-statistic. In the case of multiple probe sets for the same gene, the gene rank was determined by the probe set with largest B-statistic. 3'UTR sequences were purged with default settings as described in ref. (69). Biases in the nucleotide composition of 3'UTR sequences were accounted for using a third-order Markov model. Significance thresholds were adjusted for multiple testing using the Bonferroni correction.

GO analysis

We considered genes represented on the array and with available GO biological process annotation (70). We identified genes with increased expression in

wild-type GFP+ compared to GFP- cells as those with at least one probe set with Benjamini–Hochberg corrected $P < 0.05$ (moderated t -test) and fold-change > 2 . Enrichment and depletion of GO terms was assessed by a two-sided Fisher's exact test. P -values were corrected for multiple testing using the Benjamini–Hochberg method.

Differential expression of miR-124 targets

To assess the differential expression of conserved and nonconserved miR-124 targets, we compared the mean expression level of targets to mean expression levels of 10 000 cohorts of control genes. We computed one-sided P -values for reduced expression as the fraction of cohorts with mean expression levels smaller than or equal to the mean expression level of targets. These P -values were subsequently converted to two-sided P -values. Control genes were chosen to have 3'UTR features comparable to the targets under consideration. More specifically, for a set of targets represented on the array, control genes were obtained by drawing an identical number of genes out of the pool of all genes with available 3'UTR and expression data. The likelihood of a gene being chosen was proportional to the probability of its 3'UTR containing relevant target sites. For an individual gene, we modeled 3'UTR site occurrences as a Poisson process with expected value Np , where p was set to the probability of a given 7-mer in the 3'UTR matching miR-124 nucleotides 2–8 under a second-order Markov model. For conserved and nonconserved sites, N was chosen to be the number of conserved 7-mers (N_c) and nonconserved 7-mers (N_n), respectively. The probability of a gene containing one or more conserved target sites was then estimated as $1 - \exp(-N_c p)$. The probability of a gene containing no conserved target sites and one or more nonconserved target sites was estimated as $\exp(-N_c p) (1 - \exp(-N_n p))$, respectively.

RESULTS

C. elegans miR-124 is expressed in a subset of sensory neurons

miR-124 is a highly conserved miRNA, present in deuterostomes, ecdysozoa and lophotrochozoa (Figure 1A) (Fay Christodoulou and Detlev Arendt, personal communication) (41,71–74). While many miRNAs typically show cross-species sequence conservation in the 5'-region, known as the miRNA seed (nucleotides 2–7, Figure 1A), miR-124 is one of three miRNAs that is conserved throughout its length from *C. elegans* to humans. Unlike vertebrate genomes, the *C. elegans* genome encodes only a single copy of the *mir-124* gene and no other miRNA gene with identical seed sequence. To determine the expression pattern of *C. elegans* miR-124, we generated animals carrying a genomically integrated *mir-124* promoter::*gfp* transgene. *mir-124* promoter::*gfp* transgenic animals show GFP expression in a subset of neurons, detectable from mid-embryogenesis (~350 min post-fertilization), when neuronal differentiation begins, throughout development and in adults (Figure 1B–E). These data are consistent

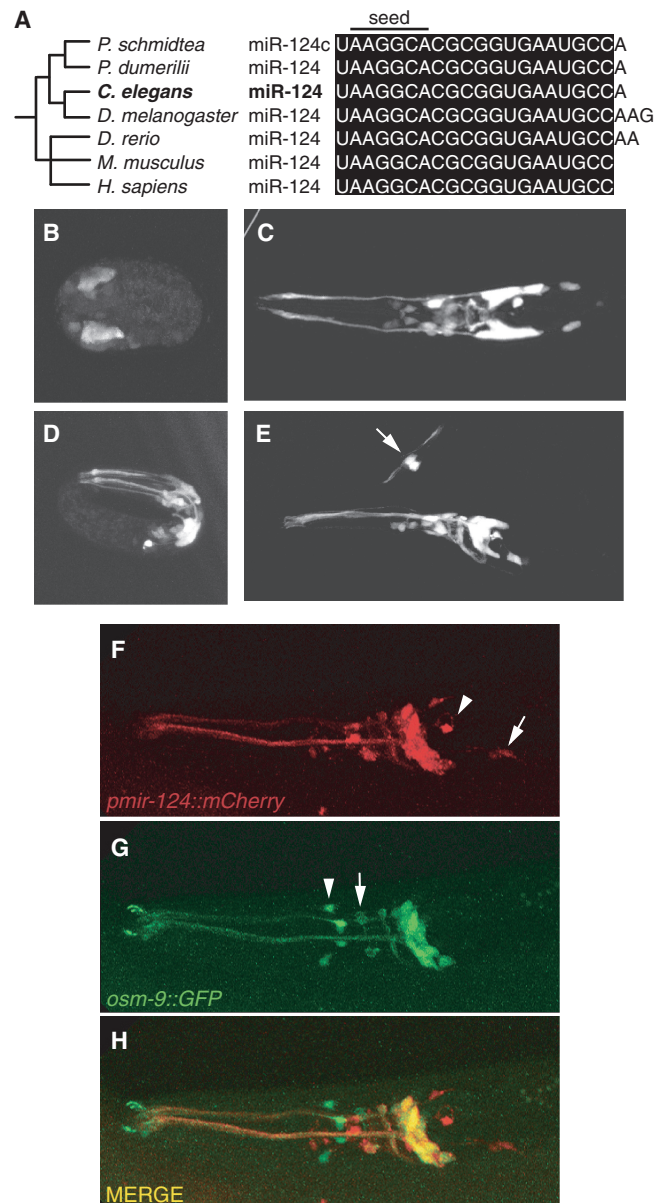


Figure 1. *Caenorhabditis elegans* miR-124 is expressed in ciliated sensory neurons (A) miR-124 primary sequence is highly conserved from *C. elegans* to *Homo sapiens*. Topology of phylogenetic tree is based on ref. (86). (B–E) *mir-124* promoter::*gfp* is expressed in *C. elegans* sensory neurons at different developmental stages. (B and D) Embryos at mid- (350 min post-fertilization) and late- (~600 min post-fertilization) embryogenesis, respectively, (C and E) at early larval stages L2, L1, respectively. This expression persists through adulthood. (E) Arrow indicates *mir-124* promoter::*gfp* expression in the phasmid sensory neurons. (B and C) Dorsal and (D and E) lateral views. (B–E) Anterior is left, (E) ventral is down. (F–H) *mir-124* is expressed in a subset of ciliated neurons as indicated by its overlap with *osm-9::gfp* expression pattern. (F) An adult animal expressing *mir-124* promoter::*mCherry*. Arrowhead and arrow indicate I6 (not ciliated) and ADE neurons, respectively, not overlapping with *osm-9*. (G) *osm-9::gfp*, arrowhead and arrow indicate IL2s and OLQs not overlapping with *mir-124* expression pattern. (H) Merge of *mir-124* promoter::*mCherry* and *osm-9::gfp* shows colocalization in amphid sensory neurons. This colocalization is evident from embryogenesis (data not shown) and persists through adulthood.

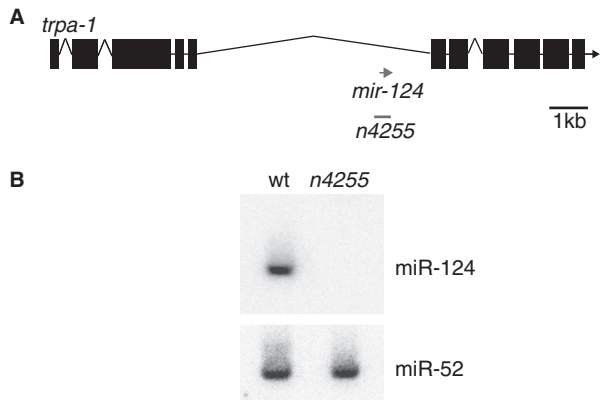


Figure 2. Genomic location of *mir-124* and mutant allele. (A) *mir-124* lies within a ~6-kb intron of host gene *trpa-1*. The *n4255* allele is a 212-bp deletion that spans the entire mature sequence of miR-124. This deletion does not abrogate *trpa-1* expression (see Supplementary Figure S1). (B) Northern blot shows expression of mature miR-124 in wild-type (wt) and absence in *n4255* mutant, miR-52 was used as loading control.

with previous analyses of the temporal pattern of miR-124 expression in *C. elegans* using northern blotting (71) and high-throughput sequencing of small RNA libraries (75). We detected *mir-124* promoter::*gfp* expression in ~40 of the 302 neurons in *C. elegans*. Using direct observation and crosses to strains containing specific neuronal markers (see ‘Materials and Methods’ section), we identified some of these as sensory neurons most of which are ciliated [AWC, AWA, AWB, ASH, ASI, ASK, PVQ (not ciliated), ASE, PHA, PHB, PVD (not ciliated), IL1, ADE, PDE] (Supplementary Table 2). The extensive overlap between *osm-9::gfp*, a reporter expressed in a subset of ciliated neurons (76), and *mir-124* promoter::*mcherry* confirmed that *mir-124* is mainly expressed in ciliated neurons (Figure 1F–H).

To investigate the *in vivo* function of *mir-124*, we characterized a mutant strain carrying a deletion allele, *n4255* (77), that completely abrogates miR-124 expression (Figure 2A and B). As previously described, *mir-124* mutants are viable with no gross abnormalities (77). To determine whether *C. elegans* miR-124 is required for sensory neuron differentiation, we studied the morphology of *mir-124* expressing neurons in wild-type and mutant animals using a number of neuronal cell fate markers. We did not observe any obvious defects in number or differentiation of sensory neurons (Supplementary Table S2 and data not shown).

Isolation of miR-124 expressing cells reveals a distinct sensory expression profile

We took advantage of the fact that mutation of *mir-124* did not cause any gross abnormalities of development of the nervous system. We reasoned that we would be able to detect the direct effect of the miRNA on the transcriptome of sensory neurons. Although, based on our reporter transgene, miR-124 is highly expressed, only ~4% of the cells of the entire animal express the miRNA. Therefore, to study the transcriptional profile of specifically miR-124

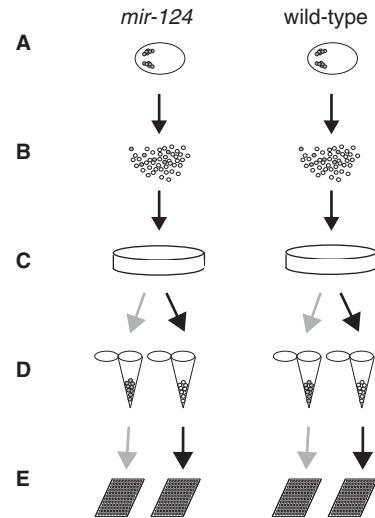


Figure 3. Isolation and expression analysis of miR-124 expressing cells from wild-type and *mir-124* mutants. (A) Embryo isolation, (B) cell dissociation by chitinase and trypsin treatment, (C) embryonic cell culture, (D) enrichment of GFP+ and GFP- cell by FACS, (E) analysis of mRNA expression by Affymetrix arrays.

expressing cells, we used fluorescence-activated cell sorting (FACS) to isolate *mir-124* promoter::*gfp* labeled embryonic cells. We generated four different cell populations: GFP+ (miR-124 expressing) and GFP- cells from both wild-type and mutant animals (Figure 3A–E), and the RNA obtained from these cell populations was subjected to microarray analysis (39,40,56–58).

To assess the purity of the sorted cell populations, we measured the levels of GFP mRNA in GFP+ and GFP- cells by qRT-PCR. As expected, we failed to detect GFP mRNA in GFP- cells (Supplementary Figure S2). To test whether the regulatory regions of the *mir-124* promoter::*gfp* reporter capture the expression domain of endogenous miR-124, we examined the levels of mature miR-124 from GFP+ and GFP- cells by qRT-PCR. Our results show ~5-fold enrichment of miR-124 in GFP+ cells (Supplementary Figure S3), confirming that *mir-124* promoter::*gfp* recapitulates the endogenous miR-124 expression pattern.

Next, we compared the transcriptional profile of miR-124 expressing (GFP+) and nonexpressing (GFP-) cells from wild-type animals. Among genes with increased expression in GFP+ cells, we found an enrichment of genes assigned the GO terms cilium biogenesis and chemotaxis (Table 1). These results are consistent with those of previous studies that examined the profile of ciliated neurons (78–80), thus validating our experimental approach, and confirming our GFP reporter studies suggesting that most miR-124 expressing cells are ciliated sensory neurons.

miR-124 loss of function results in derepression of direct targets

To investigate gene expression changes upon *mir-124* deletion, we first compared *mir-124* mutant versus wild-type GFP+ cells. We ranked genes from most

Table 1. Genes highly expressed in miR-124 expressing cells are enriched for biological processes associated with sensory neurons

ID	Description	Fold Enrichment	<i>P</i> -value
GO:0042384	Cilium biogenesis	14.66	4.10E-07
GO:0046626	Regulation of insulin receptor signaling pathway	12.22	2.40E-04
GO:0017015	Regulation of TGF-beta receptor signaling pathway	10.47	6.70E-04
GO:0008355	Olfactory learning	9.77	2.10E-04
GO:0006972	Hyperosmotic response	8.55	1.20E-04
GO:0007269	Neurotransmitter secretion	8	8.00E-04
GO:0006935	Chemotaxis	7.33	8.30E-06
GO:0009266	Response to temperature stimulus	5.86	6.70E-04
GO:0043053	Dauer entry	5.86	6.70E-04
GO:0043051	Regulation of pharyngeal pumping	5.64	1.20E-04
GO:0009190	Cyclic nucleotide biosynthetic process	5.55	3.90E-06
GO:0007186	G-protein coupled receptor protein signaling pathway	4.48	1.30E-25
GO:0048858	Cell projection morphogenesis	4.43	9.10E-06
GO:0006813	Potassium ion transport	2.71	8.00E-04
GO:0007242	Intracellular signaling cascade	2.31	3.40E-06
GO:0040010	Positive regulation of growth rate	0.5	1.30E-08
GO:0040035	Hermaphrodite genitalia development	0.4	9.70E-05
GO:0009792	Embryonic development ending in birth or egg hatching	0.39	2.00E-23
GO:0010171	Body morphogenesis	0.24	1.90E-06
GO:0002009	Morphogenesis of an epithelium	0.2	3.90E-04
GO:0007049	Cell cycle	0.12	4.30E-04
GO:0006412	Translation	0	1.20E-05

Biological process GO terms were tested for enrichment or depletion among genes with increased expression in miR-124 expressing cells compared to the rest of the animal (see 'Materials and Methods' section). *P*-values are based on a two-sided Fisher's exact test and were corrected for multiple testing using the Benjamini-Hochberg method. Shown are nonredundant terms with corrected $P < 0.001$, ordered by fold enrichment.

increased to most decreased expression and used the Sylamer method (69) to identify nucleotide sequences (words) that are enriched or depleted in the 3'UTR sequences of high-ranking compared to low-ranking genes (Figure 4). Out of all possible 6-, 7- and 8-nt words, the five words with strongest bias in their occurrence (Sylamer $P < 0.05$ after Bonferroni correction) were overrepresented in 3'UTRs of genes with increased expression in mutant GFP+ cells (Figure 4C–E). All of these words matched the seed sequence of mature miR-124 (Figure 4B). Two 6-mers and one 7-mer covered the reverse complement of miR-124 nucleotides 2–8, while one 7-mer and one 8-mer showed perfect Watson–Crick pairing to miR-124 nucleotides 2–7 and a uracil opposite the first nucleotide of the miRNA (position t1 in the mRNA transcript, see Figure 4B). The three words including a match to miRNA nucleotide 8 showed 2-fold or greater enrichment among the 3'UTRs of the most upregulated genes (Figure 4C–E). To test whether this signature was specific to miR-124, we searched for depletion or enrichment of 6-, 7- and 8-nt words in the 3'UTRs of genes with altered expression in cells that do not express miR-124 (GFP– cells, Figure 4F–H). We did not detect any words with biased occurrence in this analysis, confirming that the previously identified signature represents an enrichment of most likely direct miR-124 targets among upregulated genes. We also did not detect any bias in the occurrence of miR-124 associated words when considering the sequences of 5'UTRs or coding sequences (data not shown), consistent with most functional miRNA target sites residing in the 3'UTRs of protein-coding genes (81).

Two of the identified words comprised a perfect seed match flanked by a uracil at position t1. Previous studies in vertebrates suggested that an adenosine at t1 is most beneficial for target repression, irrespective of the identity of the first nucleotide of the miRNA (26). Indeed, when considering the enrichment of seed matches with a t8 match and a t1 A or t1 U, we observed a comparable enrichment in the 3'UTRs of highly ranked genes (Figure 4I). However, the total number of identified seed matches with a t8 match and t1 U ($n = 145$) was almost four times as high as the number of corresponding matches flanked by a t1 A ($n = 38$), explaining the greater statistical significance of the enrichment of matches with a t1 U (Figure 4C). We note that in annotated *C. elegans* 3'UTRs (see 'Materials and Methods' section), the trinucleotide UUU occurs over three times more frequently than the trinucleotide UUA. Therefore, the observed t1 bias can be explained, at least in part, by overall 3'UTR nucleotide composition. Since the increased efficacy in target repression due to a t1 A might be specific to vertebrates, or individual miRNA families only, we repeated the analysis in (26) for individual miRNA families (Supplementary Figures S4 and S5). We confirmed the preferential conservation of an adenosine immediately downstream of conserved 3'UTR seed matches for most miRNA families in vertebrates. In nematodes, however, in most cases the number of identified conserved 3'UTR seed matches for individual miRNA families was too small to conclusively assess biases in nucleotide identity at the t1 position.

To further investigate the effectiveness of distinct seed match types in conferring repression at the mRNA

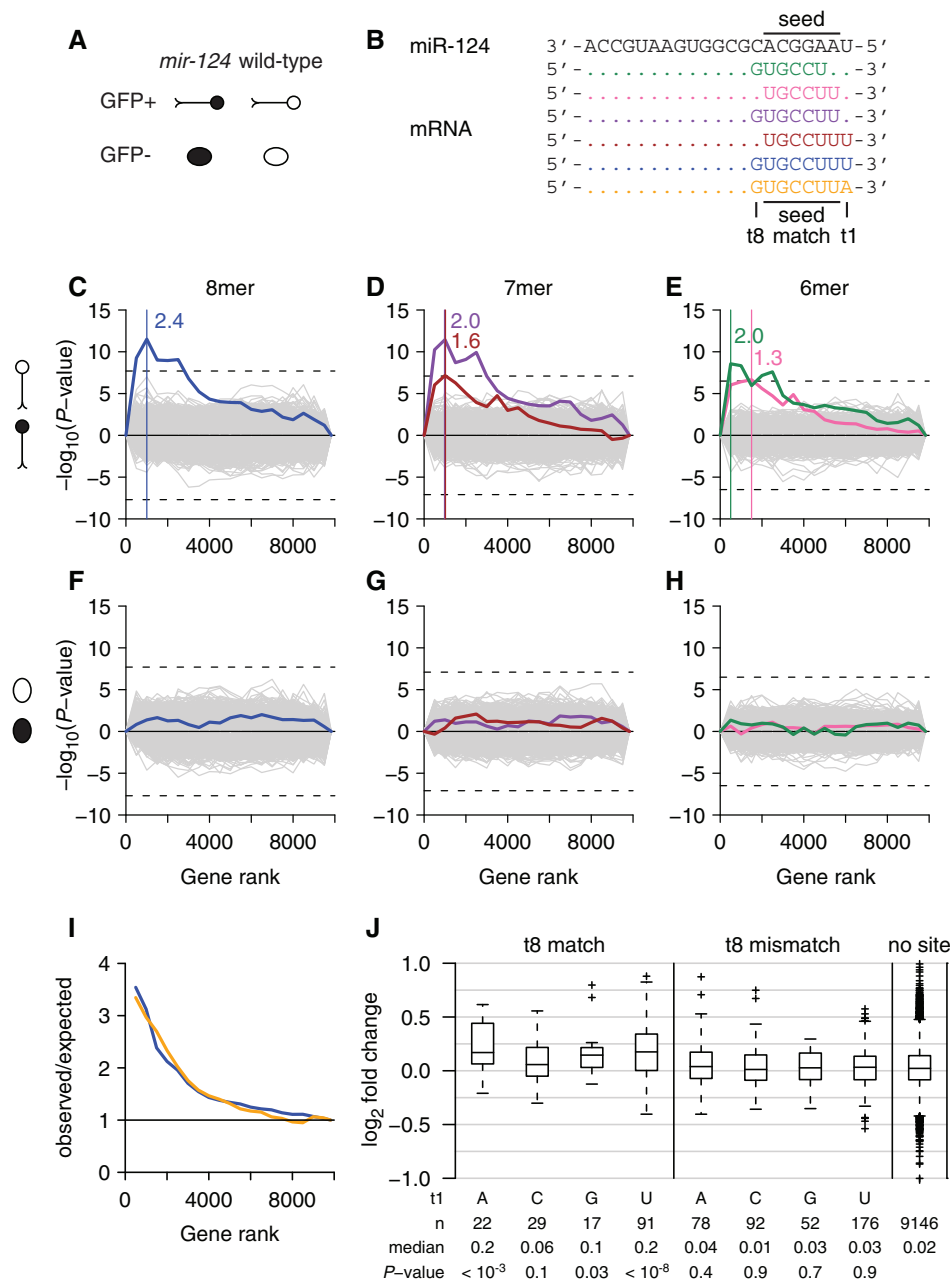


Figure 4. Genes with increased expression in *mir-124* mutant (*n4255*) GFP+ cells are enriched for miR-124 3'UTR seed matches. (A) Schematic representation of sorted cell populations for illustration in subsequent figures. miR-124 expressing cells (GFP+) and other cells of the body (GFP-), are represented by the schematic of a neuron and an oval, respectively. Black indicates *mir-124* mutant cells. (B) Relationship between mature miR-124 and 3'UTR nucleotide sequences (words) associated with gene expression statistical changes in *mir-124* mutant GFP+ cells. miRNA nucleotides 2–7 (seed), the complementary mRNA sequence (seed match) and positions flanking the seed match opposite miRNA nucleotide 1 and 8 (positions t1 and t8, respectively) are indicated. (C–H) Nucleotide words were identified by the Sylamer method. Genes were ranked from most increased to most reduced expression in mutant compared to wild-type GFP+ cells (C–E) and GFP- cells (F–H). We tested for enrichment and depletion of all possible 6-, 7- and 8-nt words among the 3'UTR sequences of highly ranking genes for a range of cut-off values (cut-offs were chosen to be multiples of 500). Left, middle and right panels show results for 8-, 7- and 6-nt words, respectively. Distinct lines correspond to distinct nucleotide words. At a given cut-off (x-axis) the height of a line indicates enrichment (positive y-axis) or depletion (negative y-axis) among sequences to the left of that cut-off. Dashed black lines indicate significance thresholds corresponding to a Bonferroni corrected P-value of 0.05. Words exceeding this threshold were highlighted. Vertical colored lines indicate the cut-off for which highest statistical significance was achieved, colored numbers indicate the enrichment (number of observed/number of expected occurrences) for the corresponding word among sequences to the left of the line. For the same comparison in GFP- cells, nucleotide words did not show any statistically significant bias in their occurrence (F–H). (I) Enrichment of seed matches comprising a t8 match and a t1 A (yellow) or t1 U (blue) among sequences to the left of a given cut-off. (J) Boxplots summarize \log_2 -fold changes between mutant and wild-type GFP+ cells for genes with a 3'UTR seed match of a given type. Only genes with a single seed match were considered. Genes were grouped according to whether the t8 nucleotide formed a Watson-Crick pair with the corresponding base in the miRNA (t8 match/mismatch) and according to the identity of the t1 nucleotide. Boxes indicate interquartile ranges, horizontal solid lines medians, whiskers extend to the most extreme data points that lie within 1.5 times the interquartile range from the quartiles, outliers are indicated by crosses. P-values are based on a two-sided Wilcoxon rank-sum test, comparing fold changes of genes with a given seed match type to those of genes with no seed match.

level, we considered changes in expression of genes with a single 3'UTR seed match (Figure 4J). Changes were modest, with the expression of all but one gene increasing <2-fold in mutant GFP+ cells. Genes harboring a single seed match with a Watson–Crick pair at flanking position t8, augmented by a t1 A or U, showed greatest increase in expression compared to genes with no seed match ($P < 10^{-3}$, $P < 10^{-8}$, respectively, two-sided Wilcoxon rank-sum test). Changes in the expression of genes with a seed match flanked by a t8 mismatch did not achieve statistical significance for any t1 nucleotide. However, we observed a trend for increased expression of seed match harboring transcripts with a t8 mismatch and t1 adenosine.

Taken together, these observations are consistent with miRNA-mediated repression of mRNA targets through perfect matches to the miRNA seed sequence in their 3'UTRs. We also confirmed the hierarchy in the effectiveness of distinct seed match types previously described based on studies in vertebrates (82). For the remainder of the study, we therefore focussed on transcripts with a perfect 3'UTR match to the miR-124 seed sequence (nucleotides 2–7) flanked by a t8 match to the miRNA or with a t1 adenosine (or both). For brevity, we referred to these below as miR-124 targets, and to the corresponding seed matches as target sites, although these are at present only candidate target genes and sites.

miR-124 represses conserved and nonconserved targets in sensory neurons

Different roles have been suggested for genes with conserved and nonconserved miRNA binding sites in their 3'UTR. While conservation is usually interpreted as an indication of biological function, nonconserved targets may constitute biologically important species-specific targets, or non-functional or inconsequential targets. To characterize differences in the behavior of the two target classes, we identified conserved targets by requiring the conservation of at least one 3'UTR site in *C. briggsae*, *C. remanei* and *C. brenneri* (see 'Materials and Methods' section). When assessing the expression of miR-124 targets as a class, we only considered those transcripts with a 3'UTR seed match flanked by a t8 match. We first considered the comparison described in the previous section, analyzing changes in the expression of miR-124 targets upon loss of miR-124. When considering GFP+ cells (Figure 5A), we observed a strong derepression of both conserved and nonconserved targets as compared to control genes ($P < 10^{-4}$, $P < 10^{-4}$, see 'Materials and Methods' section). Approximately 73% and 65% of conserved and nonconserved targets showed increased expression (fold change >1) in the absence of miR-124, respectively (compared to 55% of genes with no site). When considering targets with an increase in expression of at least 20%, the proportions of conserved and nonconserved targets were 38% and 25%, respectively (compared to 11% of genes with no site). Thus, overall conserved targets underwent greater changes in expression upon loss of miR-124 than nonconserved targets. This suggested that site conservation is indeed indicative of

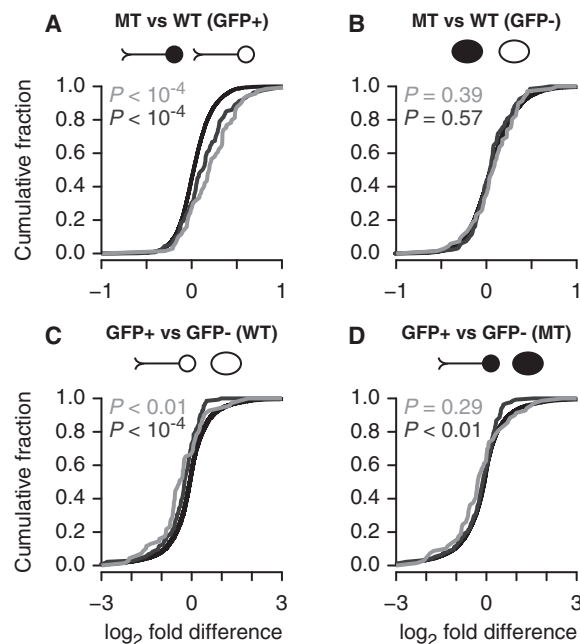


Figure 5. Differential expression of conserved and nonconserved miR-124 targets. Shown are the cumulative distributions of \log_2 -fold differences in the expression of genes with 3'UTRs harboring at least one conserved target site (light grey), exclusively nonconserved target sites (dark grey) or no seed match (black). Only seed matches with a t8 match were considered target sites. Two-sided empirical P -values were obtained by comparing the mean expression level of targets to the mean expression levels of 10000 cohorts of control genes (see 'Materials and Methods' section). Shown are data for the comparison between *mir-124* mutant and wild-type GFP+ (A) and GFP- (B) cells. Observed changes in expression for conserved and nonconserved targets differed from those for control genes in GFP+ cells but not in GFP- cells. (C and D) Shown are the cumulative distributions of \log_2 fold differences between GFP+ and GFP- cells in wild-type (C) and mutant animals (D), otherwise as in previous panels.

functional or highly effective sites or that a greater proportion of conserved targets may be co-expressed with the miRNA. In contrast, in GFP- cells the observed fold changes of targets were similar to control genes (Figure 5B), indicating high fidelity of the *mir-124* promoter::*gfp* reporter and purity of the FACS sorted cell populations.

Reduced relative expression of miR-124 targets in sensory neurons is largely due to miR-124-mediated repression

Previous studies reported reduced relative expression of miRNA targets in tissues where the miRNA is expressed (35–37). However, the mechanisms underlying this phenomenon observed for many tissue-specific miRNAs could only be addressed indirectly, by examining the expression profile of wild-type cells. This observation could be explained by either direct miRNA-mediated repression or a tendency towards mutually exclusive expression of the miRNA and its targets. Our analysis comparing mutant and wild-type GFP+ cells established that the effect could be explained, in part, by direct repression. To further investigate this question, we examined relative expression levels (GFP+/GFP-) in wild-type

and *mir-124* mutant animals (Figure 5C and D). As expected, we observed reduced relative expression of both conserved and nonconserved miR-124 targets in wild-type animals ($P < 0.01$, $P < 10^{-4}$, see 'Materials and Methods' section) (Figure 5C). In the absence of miR-124, this effect was greatly reduced but persisted for nonconserved targets ($P = 0.29$, $P < 0.01$, see 'Materials and Methods' section) (Figure 5D). This suggested that the overall differences in relative expression between GFP+ and GFP- cells observed in wild-type were to a large extent due to direct miRNA-mediated repression. However, in the case of nonconserved targets differences also appeared to be due to other mechanisms.

Conserved and nonconserved targets of miR-124 differ in their absolute expression levels in sensory neurons

To investigate absolute expression levels of miR-124 targets, we compared the mean probe-set intensity of target genes to the mean probe-set intensities of sets of control genes in the same biological sample (see 'Materials and Methods' section). In Figure 6, mean intensities of control gene sets are plotted as histograms, and observed mean intensities of targets are indicated by vertical colored lines. Conserved targets showed mean expression levels similar to control genes in all four cell populations, with a trend for reduced expression in wild-type miR-124 expressing cells (Figure 6A–D). In contrast, absolute expression levels of nonconserved targets were reduced in wild-type GFP+ cells ($P < 0.01$, see Materials and Methods, Figure 6E and F). A trend for reduced expression of nonconserved targets persisted in *mir-124* mutant GFP+ cells, consistent with site avoidance of genes that are highly expressed in sensory neurons and accumulation of nonconserved sites in genes that are highly expressed in other cells of the animal. Nevertheless, we observed a derepression of nonconserved targets upon miR-124 deletion (Figures 5A, 6E and F), suggesting that many nonconserved targets are indeed co-expressed with the miRNA.

Unlike nonconserved targets, conserved targets did not show reduced absolute expression levels in *mir-124* mutant GFP+ cells compared to control genes. Although it is not possible to assess absolute gene expression levels conclusively based on a microarray experiment, these results suggest high expression of many conserved miR-124 targets in cells that express the miRNA.

miR-124 targets show diverse patterns of spatial expression

To further characterize individual target genes, we selected the 50 genes with one or more 3'UTR target sites that showed greatest evidence for differential expression between *mir-124* mutant and wild-type sensory neurons based on the B-statistic (see Materials and Methods). We examined fold-changes between *mir-124* mutant and wild-type GFP+ and GFP- cells (Figure 7A and B). Forty-nine out of the 50 genes showed increased expression in *mir-124* mutant sensory neurons (Figure 7A). Changes between *mir-124* mutant and wild-type GFP- cells were small and less biased towards increased

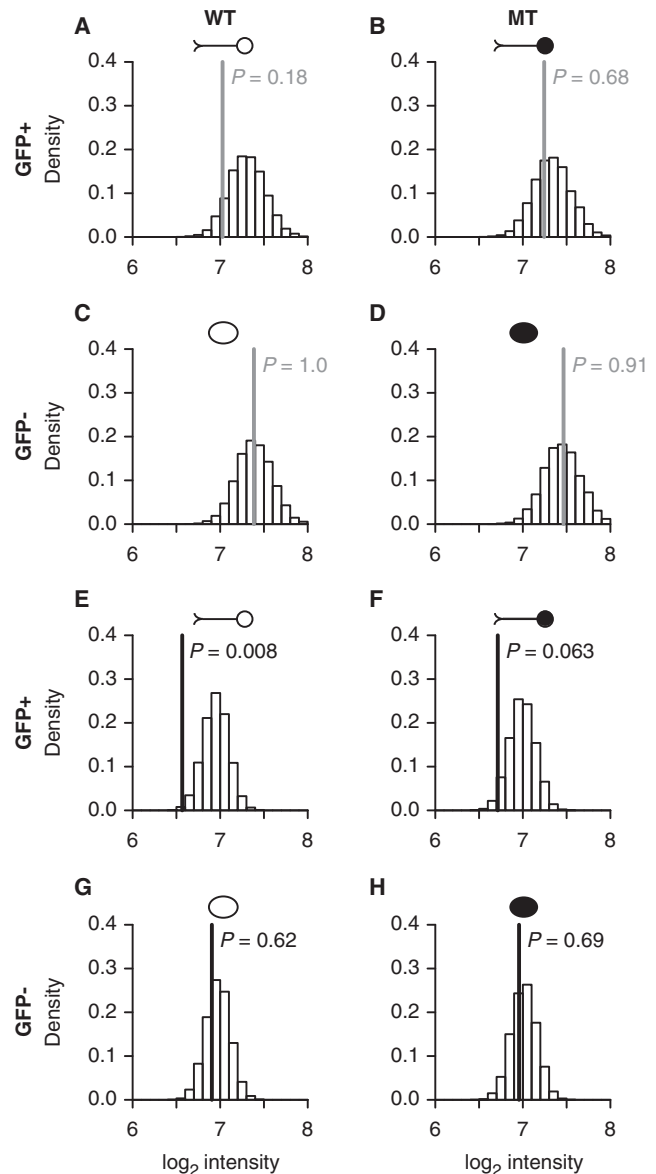


Figure 6. Absolute expression levels of conserved and nonconserved miR-124 targets. Differential expression was assessed by comparing the mean \log_2 intensities of targets to the mean \log_2 intensities of 10000 cohorts of control genes (see 'Materials and Methods' section). Mean \log_2 intensities for the control sets were plotted as histograms. Vertical light grey and dark grey lines indicate the mean \log_2 intensity of conserved (A–D) and nonconserved targets (E–H), respectively.

expression as compared to GFP+ cells (Figure 7B), suggesting that these genes are direct targets of miR-124. In wild-type animals, the relative expression levels of these genes were biased towards reduced expression in GFP+ compared to GFP- cells, as expected. However, some individual genes showed high relative expression in sensory neurons (Figure 7C). In many cases, these were consistent with published expression patterns based on promoter::GFP fusions and immunohistochemistry (Figure 7D).

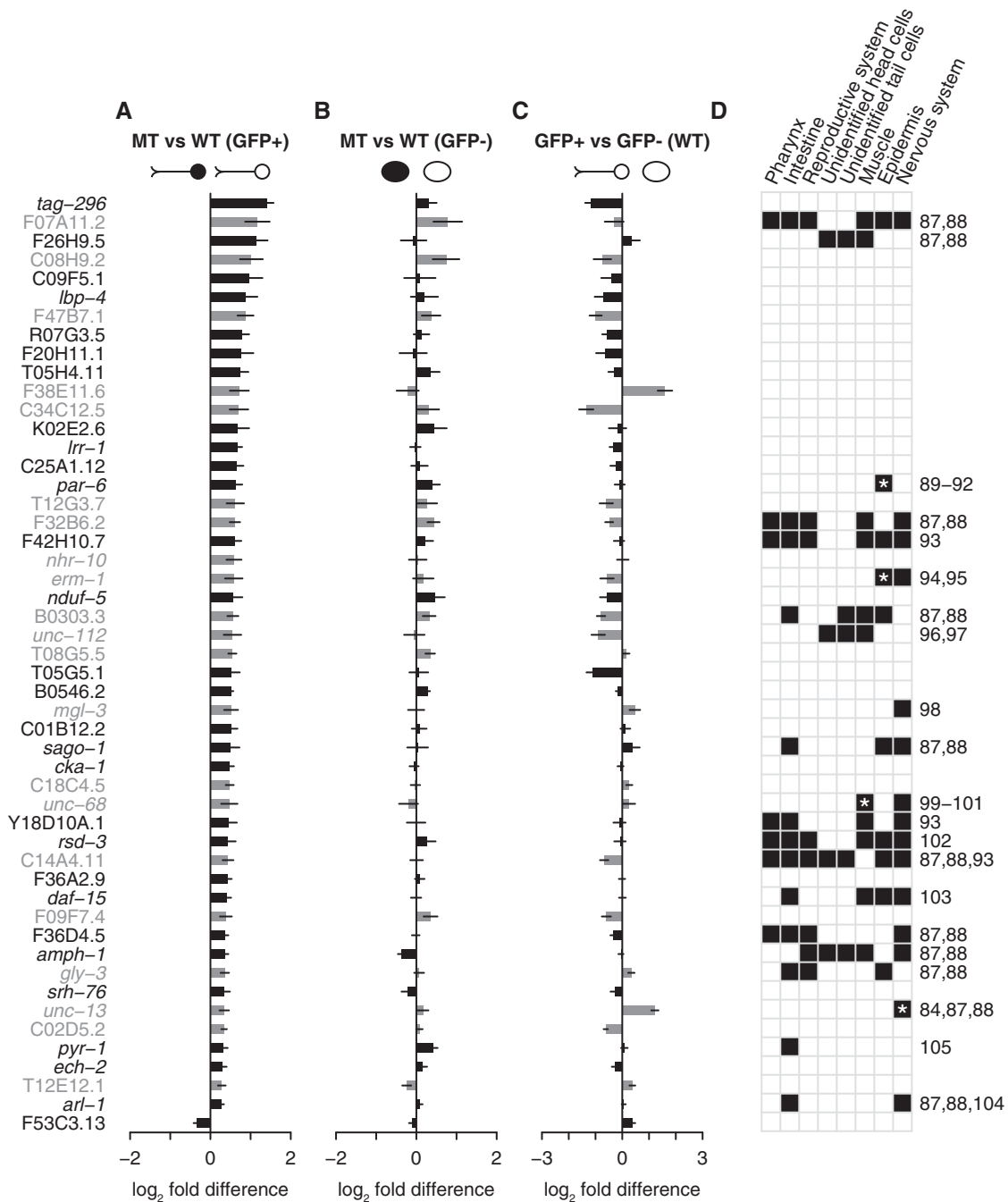


Figure 7. Overview of the 50 miR-124 targets that showed greatest evidence for differential expression between mutant and wild-type GFP+ cells based on the B-statistic. Expression differences are shown for the comparison between mutant and wild-type GFP+ cells (A), mutant and wild-type GFP- cells (B) and *mir-124* expressing and nonexpressing cells in wild-type animals (C). Error bars are standard errors of the mean. Light grey and dark grey indicate conserved and nonconserved targets, respectively. (D) Published expression patterns based on promoter::GFP fusions. A black box indicates reported expression for a specific tissue. Asterisks indicate reported expression based on immunohistochemistry. Numbers in the right margin correspond to relevant publications (84,87–105).

DISCUSSION

Previous studies in vertebrates identified miR-124 as a pan-neuronally expressed miRNA and suggest a role for miR-124 in the differentiation of the nervous system (48,49,51,52). We show that the neuronal expression pattern of miR-124 is conserved in *C. elegans*, suggesting functional conservation from nematodes to humans.

However, in *C. elegans* miR-124 expression is mostly confined to sensory neurons. Interestingly, a recent *in situ* hybridization study in *Aplysia* demonstrated miR-124 expression in sensory but not motor neurons (83), suggesting that in some invertebrates miR-124 might be restricted to the sensory nervous system. The difference between miR-124 expression domains in

vertebrates and invertebrates suggests that miR-124 might have evolved to have different functions and/or targets.

Here, we describe the analysis of the first *mir-124* mutant in any organism. Our characterization of *mir-124* mutant animals did not reveal any obvious defects in the differentiation of the sensory nervous system under standard laboratory conditions. Therefore, we were unable to confirm the observations made for miR-124 in vertebrate systems in *C. elegans in vivo*. As suggested previously, this could be due to differences in miR-124 function among animal species or differences in experimental approaches, as other studies relied on miR-124 overexpression or knockdown by anti-sense probes, in some instances in cell culture.

We investigated the mode of action of *C. elegans* miR-124 by studying its impact on the transcriptome in both miR-124 expressing and nonexpressing cells. Among the genes with increased expression in *mir-124* mutant sensory neurons, we found an enrichment of likely direct miR-124 targets, suggesting that many targets are co-expressed with and actively repressed by miR-124. These data confirm recent studies of mRNA expression in miR-124 expressing cells of wild-type and *MZdicer* mutant zebrafish, which lack the function of all miRNAs (39).

The significance of post-transcriptional regulation by miR-124 was supported by our analysis of absolute expression levels. Expression data from animals lacking miR-124 suggested that many conserved targets were expressed in both sensory neurons and the rest of the animal. In the miR-124 expressing sensory neurons of wild-type animals, we observed a trend for reduced absolute expression compared to control genes, presumably due to direct miRNA-mediated repression. Taken together, these results suggest that miR-124 tunes the expression of many of these genes, rather than acting exclusively as a fail-safe mechanism against spurious transcription.

We observed reduced relative expression of miRNA targets in cells where the miRNA is expressed, a phenomenon that has been described for many tissue-specific miRNAs (35–37). However, it was previously unclear whether this observation could be explained by transcriptional regulation or whether it was due to miRNA-mediated mRNA degradation. Here, we show that for *C. elegans* miR-124 the phenomenon is largely due to a direct effect of the miRNA. However, in the case of nonconserved targets, reduced relative expression in sensory neurons persisted in the absence of the miRNA. This observation is consistent with site avoidance for genes highly expressed in sensory neurons, and accumulation of sites for genes expressed in cells where the miRNA is absent. Similar observations were made in zebrafish for miR-124 (39) and the muscle-specific miRNAs miR-1 and miR-133 (39,40) when analyzing miRNA-expressing cells from wild-type and *MZdicer* animals.

While the biological significance of actively repressed nonconserved targets remains unclear, conserved targets likely represent those key to understanding the function of miR-124. Since miR-124 is expressed from embryogenesis throughout adulthood it is conceivable that various

processes, from development of the nervous system to neuronal function, could be regulated. For example, regulation of the conserved target *unc-13*, a neurotransmitter release regulator localized to most or all synapses (84), might be important for modulating neurotransmission. However, the class of genes with conserved 3'UTR seed matches to miR-124 as a whole, and the subset we found to respond during embryonic development, were not enriched for any particular biological function (not shown), and it remains open why this particular set of genes is experiencing evolutionary pressure to be under post-transcriptional regulation in sensory neurons. The data presented here suggest that miR-124 contributes to the control of numerous biological processes. The phenotypic outcome of this control, however, remains to be elucidated and may only become apparent under more extreme conditions, as recently suggested by a study of miR-7 in *Drosophila* (85).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Svetlana Mazel and Christopher Bare for their help with flow cytometry, Vladimir Benes and Tomi Ivacevic at EMBL for performing the array experiments, Partha Das for help with northern blotting, the *Caenorhabditis* Genetics Center for strains. We thank members of the Miska lab for comments on the article.

FUNDING

Cancer Research UK Programme Grant (C13474 to E.A.M.); core funding to the Wellcome Trust/Cancer Research UK Gurdon Institute provided by the Wellcome Trust (UK); Cancer Research UK; National Institutes of Health (NS064273 to S.S.) in part; EPSRC fellowship (UK); Cancer Research UK (to L.D.G.); Cancer Research UK Programme (C14303 to S.T.); Hutchison Whampoa Limited (to L.D.G., S.T.). Funding for open access charge: Cancer Research UK.

Conflict of interest statement. None declared.

REFERENCES

1. Sharp, P.A. (2009) The centrality of RNA. *Cell*, **136**, 577–580.
2. Arazi, T., Talmor-Neiman, M., Stav, R., Riese, M., Huijser, P. and Baulcombe, D.C. (2005) Cloning and characterization of micro-RNAs from moss. *Plant J.*, **43**, 837–848.
3. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. and Tuschl, T. (2001) Identification of novel genes coding for small expressed RNAs. *Science*, **294**, 853–858.
4. Lau, N.C., Lim, L.P., Weinstein, E.G. and Bartel, D.P. (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science*, **294**, 858–862.
5. Lee, R.C. and Ambros, V. (2001) An extensive class of small RNAs in *Caenorhabditis elegans*. *Science*, **294**, 862–864.
6. Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.

7. Lim, L.P., Glasner, M.E., Yekta, S., Burge, C.B. and Bartel, D.P. (2003) Vertebrate microRNA genes. *Science*, **299**, 1540.
8. Grimson, A., Srivastava, M., Fahey, B., Woodcroft, B.J., Chiang, H.R., King, N., Degnan, B.M., Rokhsar, D.S. and Bartel, D.P. (2008) Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature*, **455**, 1193–1197.
9. Wightman, B., Ha, I. and Ruvkun, G. (1993) Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell*, **75**, 855–862.
10. Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R. and Ruvkun, G. (2000) The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, **403**, 901–906.
11. Lee, R.C., Feinbaum, R.L. and Ambros, V. (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, **75**, 843–854.
12. Johnston, R.J. and Hobert, O. (2003) A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*. *Nature*, **426**, 845–849.
13. Brennecke, J., Hipfner, D.R., Stark, A., Russell, R.B. and Cohen, S.M. (2003) *bantam* encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene *hid* in *Drosophila*. *Cell*, **113**, 25–36.
14. Giraldez, A.J., Cinalli, R.M., Glasner, M.E., Enright, A.J., Thomson, J.M., Baskerville, S., Hammond, S.M., Bartel, D.P. and Schier, A.F. (2005) MicroRNAs regulate brain morphogenesis in zebrafish. *Science*, **308**, 833–838.
15. Rodriguez, A., Vigorito, E., Clare, S., Warren, M.V., Couttet, P., Soond, D.R., van Dongen, S., Grocock, R.J., Das, P.P., Miska, E.A. et al. (2007) Requirement of *bic/microRNA-155* for normal immune function. *Science*, **316**, 608–611.
16. Sokol, N.S. and Ambros, V. (2005) Mesodermally expressed *Drosophila* microRNA-1 is regulated by *Twist* and is required in muscles during larval growth. *Genes Dev.*, **19**, 2343–2354.
17. Juarez, M.T., Kui, J.S., Thomas, J., Heller, B.A. and Timmermans, M.C. (2004) microRNA-mediated repression of rolled *leaf1* specifies maize leaf polarity. *Nature*, **428**, 84–88.
18. Palatnik, J.F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J.C. and Weigel, D. (2003) Control of leaf morphogenesis by microRNAs. *Nature*, **425**, 257–263.
19. Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A. et al. (2005) MicroRNA expression profiles classify human cancers. *Nature*, **435**, 834–838.
20. Teleman, A.A., Maitra, S. and Cohen, S.M. (2006) *Drosophila* lacking microRNA *miR-278* are defective in energy homeostasis. *Genes Dev.*, **20**, 417–422.
21. Griffiths-Jones, S. (2004) The microRNA Registry. *Nucleic Acids Res.*, **32**, D109–D111.
22. Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A. and Enright, A.J. (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.*, **34**, D140–D144.
23. Rhoades, M.W., Reinhart, B.J., Lim, L.P., Burge, C.B., Bartel, B. and Bartel, D.P. (2002) Prediction of plant microRNA targets. *Cell*, **110**, 513–520.
24. Brennecke, J., Stark, A., Russell, R.B. and Cohen, S.M. (2005) Principles of microRNA-target recognition. *PLoS Biol.*, **3**, e85.
25. Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M. et al. (2005) Combinatorial microRNA target predictions. *Nat. Genet.*, **37**, 495–500.
26. Lewis, B.P., Burge, C.B. and Bartel, D.P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, **120**, 15–20.
27. Xie, X., Lu, J., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S. and Kellis, M. (2005) Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature*, **434**, 338–345.
28. Doench, J.G. and Sharp, P.A. (2004) Specificity of microRNA target selection in translational repression. *Genes Dev.*, **18**, 504–511.
29. Abrahante, J.E., Daul, A.L., Li, M., Volk, M.L., Tennessen, J.M., Miller, E.A. and Rougvie, A.E. (2003) The *Caenorhabditis elegans* hunchback-like gene *lin-57/hbl-1* controls developmental time and is regulated by microRNAs. *Dev. Cell*, **4**, 625–637.
30. Lin, S.Y., Johnson, S.M., Abraham, M., Vella, M.C., Pasquinelli, A., Gamberi, C., Gottlieb, E. and Slack, F.J. (2003) The *C. elegans* hunchback homolog, *hbl-1*, controls temporal patterning and is a probable microRNA target. *Dev. Cell*, **4**, 639–650.
31. Friedman, R.C., Farh, K.K., Burge, C.B. and Bartel, D.P. (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.*, **19**, 92–105.
32. Lim, L.P., Lau, N.C., Garrett-Engle, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S. and Johnson, J.M. (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature*, **433**, 769–773.
33. Baek, D., Villen, J., Shin, C., Camargo, F.D., Gygi, S.P. and Bartel, D.P. (2008) The impact of microRNAs on protein output. *Nature*, **455**, 64–71.
34. Selbach, M., Schwanhauss, B., Thierfelder, N., Fang, Z., Khanin, R. and Rajewsky, N. (2008) Widespread changes in protein synthesis induced by microRNAs. *Nature*, **455**, 58–63.
35. Stark, A., Brennecke, J., Bushati, N., Russell, R.B. and Cohen, S.M. (2005) Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell*, **123**, 1133–1146.
36. Farh, K.K., Grimson, A., Jan, C., Lewis, B.P., Johnston, W.K., Lim, L.P., Burge, C.B. and Bartel, D.P. (2005) The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science*, **310**, 1817–1821.
37. Sood, P., Krek, A., Zavolan, M., Macino, G. and Rajewsky, N. (2006) Cell-type-specific signatures of microRNAs on target mRNA expression. *Proc. Natl Acad. Sci. USA*, **103**, 2746–2751.
38. Tsang, J., Zhu, J. and van Oudenaarden, A. (2007) MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals. *Mol. Cell*, **26**, 753–767.
39. Shkumatava, A., Stark, A., Sive, H. and Bartel, D.P. (2009) Coherent but overlapping expression of microRNAs and their targets during vertebrate development. *Genes Dev.*, **23**, 466–481.
40. Mishima, Y., Abreu-Goodger, C., Staton, A.A., Stahlhut, C., Shou, C., Cheng, C., Gerstein, M., Enright, A.J. and Giraldez, A.J. (2009) Zebrafish *miR-1* and *miR-133* shape muscle gene expression and regulate sarcomeric actin organization. *Genes Dev.*, **23**, 619–632.
41. Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W. and Tuschl, T. (2002) Identification of tissue-specific microRNAs from mouse. *Curr. Biol.*, **12**, 735–739.
42. Nelson, P.T., Baldwin, D.A., Kloosterman, W.P., Kauppinen, S., Plasterk, R.H. and Mourelatos, Z. (2006) RAKE and LNA-ISH reveal microRNA expression and localization in archival human brain. *RNA*, **12**, 187–191.
43. Miska, E.A., Alvarez-Saavedra, E., Townsend, M., Yoshii, A., Sestan, N., Rakic, P., Constantine-Paton, M. and Horvitz, H.R. (2004) Microarray analysis of microRNA expression in the developing mammalian brain. *Genome Biol.*, **5**, R68.
44. Sempere, L.F., Freemantle, S., Pitha-Rowe, I., Moss, E., Dmitrov, S. and Ambros, V. (2004) Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. *Genome Biol.*, **5**, R13.
45. Deo, M., Yu, J.Y., Chung, K.H., Tippens, M. and Turner, D.L. (2006) Detection of mammalian microRNA expression by in situ hybridization with RNA oligonucleotides. *Dev. Dyn.*, **235**, 2538–2548.
46. Wienholds, E., Kloosterman, W.P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., Horvitz, H.R., Kauppinen, S. and Plasterk, R.H. (2005) MicroRNA expression in zebrafish embryonic development. *Science*, **309**, 310–311.
47. Aboobaker, A.A., Tomancak, P., Patel, N., Rubin, G.M. and Lai, E.C. (2005) *Drosophila* microRNAs exhibit diverse spatial expression patterns during embryonic development. *Proc. Natl Acad. Sci. USA*, **102**, 18017–18022.
48. Krichevsky, A.M., Sonntag, K.C., Isacson, O. and Kosik, K.S. (2006) Specific microRNAs modulate embryonic stem cell-derived neurogenesis. *Stem Cells*, **24**, 857–864.
49. Makeyev, E.V., Zhang, J., Carrasco, M.A. and Maniatis, T. (2007) The MicroRNA *miR-124* promotes neuronal differentiation by

- triggering brain-specific alternative pre-mRNA splicing. *Mol. Cell*, **27**, 435–448.
50. Cao, X., Pfaff, S.L. and Gage, F.H. (2007) A functional study of miR-124 in the developing neural tube. *Genes Dev.*, **21**, 531–536.
 51. Visvanathan, J., Lee, S., Lee, B., Lee, J.W. and Lee, S.K. (2007) The microRNA miR-124 antagonizes the anti-neural REST/SCP1 pathway during embryonic CNS development. *Genes Dev.*, **21**, 744–749.
 52. Cheng, L.C., Pastrana, E., Tavazoie, M. and Doetsch, F. (2009) miR-124 regulates adult neurogenesis in the subventricular zone stem cell niche. *Nat. Neurosci.*, **12**, 399–408.
 53. Brenner, S. (1974) The genetics of *Caenorhabditis elegans*. *Genetics*, **77**, 71–94.
 54. Hobert, O. (2002) PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic *C. elegans*. *Biotechniques*, **32**, 728–730.
 55. Mello, C. and Fire, A. (1995) DNA transformation. *Methods Cell Biol.*, **48**, 451–482.
 56. Bacaj, T., Tevlin, M., Lu, Y. and Shaham, S. (2008) Glia are essential for sensory organ function in *C. elegans*. *Science*, **322**, 744–747.
 57. Fox, R.M., Von Stetina, S.E., Barlow, S.J., Shaffer, C., Olszewski, K.L., Moore, J.H., Dupuy, D., Vidal, M. and Miller, D.M. III (2005) A gene expression fingerprint of *C. elegans* embryonic motor neurons. *BMC Genomics*, **6**, 42.
 58. Christensen, M., Estevez, A., Yin, X., Fox, R., Morrison, R., McDonnell, M., Gleason, C., Miller, D.M. III and Strange, K. (2002) A primary culture system for functional analysis of *C. elegans* neurons and muscle cells. *Neuron*, **33**, 503–514.
 59. Chen, C., Ridzon, D.A., Broomer, A.J., Zhou, Z., Lee, D.H., Nguyen, J.T., Barbisin, M., Xu, N.L., Mahavakar, V.R., Andersen, M.R. *et al.* (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.*, **33**, e179.
 60. Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R. *et al.* (2008) Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. *Mol. Cell*, **31**, 79–90.
 61. Team, R.D.C. (2009) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
 62. Gentleman, R., Carey, V., Huber, W., Irizarry, R.A. and Dudoit, S. (2005) *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer, New York.
 63. Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B. and Speed, T.P. (2003) Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.*, **31**, e15.
 64. Smyth, G.K. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.*, **3**, Article 3.
 65. Ritchie, M.E., Diyagama, D., Neilson, J., van Laar, R., Dobrovic, A., Holloway, A. and Smyth, G.K. (2006) Empirical array quality weights in the analysis of microarray data. *BMC Bioinformatics*, **7**, 261.
 66. Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M. and Haussler, D. (2002) The human genome browser at UCSC. *Genome Res.*, **12**, 996–1006.
 67. Karolchik, D., Kuhn, R.M., Baertsch, R., Barber, G.P., Clawson, H., Diekhans, M., Giardine, B., Harte, R.A., Hinrichs, A.S., Hsu, F. *et al.* (2008) The UCSC Genome Browser Database: 2008 update. *Nucleic Acids Res.*, **36**, D773–D779.
 68. Pruitt, K.D., Tatusova, T. and Maglott, D.R. (2005) NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.*, **33**, D501–D504.
 69. van Dongen, S., Abreu-Goodger, C. and Enright, A.J. (2008) Detecting microRNA binding and siRNA off-target effects from expression data. *Nat. Methods*, **5**, 1023–1025.
 70. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T. *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.
 71. Lim, L.P., Lau, N.C., Weinstein, E.G., Abdelhakim, A., Yekta, S., Rhoades, M.W., Burge, C.B. and Bartel, D.P. (2003) The microRNAs of *Caenorhabditis elegans*. *Genes Dev.*, **17**, 991–1008.
 72. Palakodeti, D., Smielewska, M. and Graveley, B.R. (2006) MicroRNAs from the Planarian *Schmidtea mediterranea*: a model system for stem cell biology. *RNA*, **12**, 1640–1649.
 73. Aravin, A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J. and Tuschl, T. (2003) The small RNA profile during *Drosophila melanogaster* development. *Dev. Cell*, **5**, 337–350.
 74. Chen, P.Y., Manning, H., Slanchev, K., Chien, M., Russo, J.J., Ju, J., Sheridan, R., John, B., Marks, D.S., Gaidatzis, D. *et al.* (2005) The developmental miRNA profiles of zebrafish as determined by small RNA cloning. *Genes Dev.*, **19**, 1288–1293.
 75. Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S. *et al.* (2008) PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. *Mol. Cell*, **31**, 67–78.
 76. Colbert, H.A., Smith, T.L. and Bargmann, C.I. (1997) OSM-9, a novel protein with structural similarity to channels, is required for olfaction, mechanosensation, and olfactory adaptation in *Caenorhabditis elegans*. *J. Neurosci.*, **17**, 8259–8269.
 77. Miska, E.A., Alvarez-Saavedra, E., Abbott, A.L., Lau, N.C., Hellman, A.B., McGonagle, S.M., Bartel, D.P., Ambros, V.R. and Horvitz, H.R. (2007) Most *Caenorhabditis elegans* microRNAs are individually not essential for development or viability. *PLoS Genet.*, **3**, e215.
 78. Blacque, O.E., Perens, E.A., Boroevich, K.A., Inglis, P.N., Li, C., Warner, A., Khattra, J., Holt, R.A., Ou, G., Mah, A.K. *et al.* (2005) Functional genomics of the cilium, a sensory organelle. *Curr. Biol.*, **15**, 935–941.
 79. Colosimo, M.E., Brown, A., Mukhopadhyay, S., Gabel, C., Lanjuin, A.E., Samuel, A.D. and Sengupta, P. (2004) Identification of thermosensory and olfactory neuron-specific genes via expression profiling of single neuron types. *Curr. Biol.*, **14**, 2245–2251.
 80. Kunitomo, H., Uesugi, H., Kohara, Y. and Iino, Y. (2005) Identification of ciliated sensory neuron-expressed genes in *Caenorhabditis elegans* using targeted pull-down of poly(A) tails. *Genome Biol.*, **6**, R17.
 81. Grimson, A., Farh, K.K., Johnston, W.K., Garrett-Engele, P., Lim, L.P. and Bartel, D.P. (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol. Cell*, **27**, 91–105.
 82. Bartel, D.P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell*, **136**, 215–233.
 83. Rajasethupathy, P., Fiumara, F., Sheridan, R., Betel, D., Puthanveetil, S.V., Russo, J.J., Sander, C., Tuschl, T. and Kandel, E. (2009) Characterization of small RNAs in *aplysia* reveals a role for miR-124 in constraining synaptic plasticity through CREB. *Neuron*, **63**, 803–817.
 84. Kohn, R.E., Duerr, J.S., McManus, J.R., Duke, A., Rakow, T.L., Maruyama, H., Moulder, G., Maruyama, I.N., Barstead, R.J. and Rand, J.B. (2000) Expression of multiple UNC-13 proteins in the *Caenorhabditis elegans* nervous system. *Mol. Biol. Cell*, **11**, 3441–3452.
 85. Li, X., Cassidy, J.J., Reinke, C.A., Fischboeck, S. and Carthew, R.W. (2009) A microRNA imparts robustness against environmental fluctuation during development. *Cell*, **137**, 273–282.
 86. Dunn, C.W., Hejnal, A., Matus, D.Q., Pang, K., Browne, W.E., Smith, S.A., Seaver, E., Rouse, G.W., Obst, M., Edgecombe, G.D. *et al.* (2008) Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature*, **452**, 745–749.
 87. Hunt-Newbury, R., Viveiros, R., Johnsen, R., Mah, A., Anastas, D., Fang, L., Halfnight, E., Lee, D., Lin, J., Lorch, A. *et al.* (2007) High-throughput in vivo analysis of gene expression in *Caenorhabditis elegans*. *PLoS Biol.*, **5**, e237.
 88. McKay, S.J., Johnsen, R., Khattra, J., Asano, J., Baillie, D.L., Chan, S., Dube, N., Fang, L., Goszczynski, B., Ha, E. *et al.* (2004) Gene expression profiling of cells, tissues and developmental stages of the nematode *C. elegans*. *Cold Spring Harb. Symp. Quant. Biol.*, **68**, 159–169.
 89. Totong, R., Achilleos, A. and Nance, J. (2007) PAR-6 is required for junction formation but not apicobasal polarization in *C. elegans* embryonic epithelial cells. *Development*, **134**, 1259–1268.

90. Munro, E., Nance, J. and Priess, J.R. (2004) Cortical flows powered by asymmetrical contraction transport PAR proteins to establish and maintain anterior-posterior polarity in the early *C. elegans* embryo. *Dev. Cell*, **7**, 413–424.
91. Cuenca, A.A., Schetter, A., Aceto, D., Kempfues, K. and Seydoux, G. (2003) Polarization of the *C. elegans* zygote proceeds via distinct establishment and maintenance phases. *Development*, **130**, 1255–1265.
92. Nance, J. and Priess, J.R. (2002) Cell polarity and gastrulation in *C. elegans*. *Development*, **129**, 387–397.
93. Reece-Hoyes, J.S., Shingles, J., Dupuy, D., Grove, C.A., Walhout, A.J., Vidal, M. and Hope, I.A. (2007) Insight into transcription factor gene duplication from *Caenorhabditis elegans* Promoterome-driven expression patterns. *BMC Genomics*, **8**, 27.
94. Gobel, V., Barrett, P.L., Hall, D.H. and Fleming, J.T. (2004) Lumen morphogenesis in *C. elegans* requires the membrane-cytoskeleton linker *erm-1*. *Dev. Cell*, **6**, 865–873.
95. Van Furden, D., Johnson, K., Segbert, C. and Bossinger, O. (2004) The *C. elegans* ezrin-radixin-moesin protein *ERM-1* is necessary for apical junction remodelling and tubulogenesis in the intestine. *Dev. Biol.*, **272**, 262–276.
96. Rogalski, T.M., Mullen, G.P., Gilbert, M.M., Williams, B.D. and Moerman, D.G. (2000) The *UNC-112* gene in *Caenorhabditis elegans* encodes a novel component of cell-matrix adhesion structures required for integrin localization in the muscle cell membrane. *J. Cell Biol.*, **150**, 253–264.
97. Hikita, T., Qadota, H., Tsuboi, D., Taya, S., Moerman, D.G. and Kaibuchi, K. (2005) Identification of a novel Cdc42 GEF that is localized to the PAT-3-mediated adhesive structure. *Biochem. Biophys. Res. Commun.*, **335**, 139–145.
98. Greer, E.R., Perez, C.L., Van Gilst, M.R., Lee, B.H. and Ashrafi, K. (2008) Neural and molecular dissection of a *C. elegans* sensory circuit that regulates fat and feeding. *Cell Metab.*, **8**, 118–131.
99. Maryon, E.B., Saari, B. and Anderson, P. (1998) Muscle-specific functions of ryanodine receptor channels in *Caenorhabditis elegans*. *J. Cell Sci.*, **111** (Pt 19), 2885–2895.
100. Sakube, Y., Ando, H. and Kagawa, H. (1997) An abnormal ketamine response in mutants defective in the ryanodine receptor gene *ryr-1* (*unc-68*) of *Caenorhabditis elegans*. *J. Mol. Biol.*, **267**, 849–864.
101. Hamada, T., Sakube, Y., Ahnn, J., Kim, D.H. and Kagawa, H. (2002) Molecular dissection, tissue localization and Ca²⁺ binding of the ryanodine receptor of *Caenorhabditis elegans*. *J. Mol. Biol.*, **324**, 123–135.
102. Tijsterman, M., May, R.C., Simmer, F., Okihara, K.L. and Plasterk, R.H. (2004) Genes required for systemic RNA interference in *Caenorhabditis elegans*. *Curr. Biol.*, **14**, 111–116.
103. Jia, K., Chen, D. and Riddle, D.L. (2004) The TOR pathway interacts with the insulin signaling pathway to regulate *C. elegans* larval development, metabolism and life span. *Development*, **131**, 3897–3906.
104. Li, Y., Kelly, W.G., Logsdon, J.M. Jr, Schurko, A.M., Harfe, B.D., Hill-Harfe, K.L. and Kahn, R.A. (2004) Functional genomic analysis of the ADP-ribosylation factor family of GTPases: phylogeny among diverse eukaryotes and function in *C. elegans*. *FASEB J.*, **18**, 1834–1850.
105. Franks, D.M., Izumikawa, T., Kitagawa, H., Sugahara, K. and Okkema, P.G. (2006) *C. elegans* pharyngeal morphogenesis requires both de novo synthesis of pyrimidines and synthesis of heparan sulfate proteoglycans. *Dev. Biol.*, **296**, 409–420.