Caenorhabditis elegans aristaless/Arx gene alr-1 restricts variable gene expression

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Caenorhabditis elegans aristless/Arx gene alr-1 restricts variable gene expression

Irini Topalidou, Alexander van Oudenaarden, and Martin Chalfie

*Department of Biological Sciences, Columbia University, New York, NY 10027; ‡Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139; and †Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by Martin Chalfie, January 25, 2011 (sent for review December 11, 2010)

Variable expressivity of mutant phenotypes in genetically identical individuals is a phenomenon widely reported but poorly understood. For example, mutations in the gene encoding the transcription factor ALR-1 in Caenorhabditis elegans result in variable touch receptor neuron (TRN) function. Using single-molecule in situ hybridization, we demonstrate that this phenotypic variability reflects enhanced variability in the expression of the selector gene mec-3, which is needed, together with unc-86, for the differentiation of the TRNs. In a yeast expression system, ALR-1 enhances MEC-3/UNC-86–dependent transcription from the mec-3 promoter, showing that ALR-1 can enhance bulk mec-3 expression. We show that, due to stochastic fluctuations, autoregulation of mec-3 is not sufficient for TRN differentiation; ALR-1 provides a second positive feedback loop that increases mec-3 expression, by restricting variability, and thus ensures TRN differentiation. Our results link fluctuations in gene expression to phenotypic variability, which is seen in many mutant strains, and provide an explicit demonstration of how variable gene expression can be curtailed in developing cells to ensure their differentiation. Because ALR-1 and similar proteins (Drosophila Aristaless and human ARX) are needed for the expression of other transcription factors, we propose that proteins in this family may act to ensure differentiation more generally.

Aristaless family | LIM homeodomain proteins | stochastic gene expression | terminal differentiation

Phenotypic variability in genetically identical mutant organisms is widely reported (1–3) but poorly understood. Gene expression studies in unicellular organisms have shown that stochastic events can lead to molecular and phenotypic variability (4). In multicellular organisms, where cells should acquire particular cell fates, such noise in gene expression could have dramatic developmental consequences, implying that tight regulatory control is needed for robust development. Little is currently known about the mechanisms by which this noise is prevented in differentiating cells.

In general, cells are thought to acquire specific fates during development because regulatory transcription factors (encoded by selector or terminal selector genes) activate the expression of cell-type–specific realizer or effector genes that define cell identity (5). Hobert (7) and Colman and Dresesen (8) have proposed that the stability and maintenance of a newly acquired state is mediated by autoregulatory loops whereby selector genes initiate and maintain the cell-specific gene program by autoregulating their own expression throughout the life of the cell.

The differentiation of the six touch receptor neurons (TRNs) in the nematode Caenorhabditis elegans uses such an autoregulatory loop. These cells sense gentle touch to the body. Sreenos for touch-insensitive mutants identified two genes, unc-86 and mec-3, that are needed for the generation and specification of the TRNs (9, 10). unc-86 encodes a homeodomain protein (UNC-86), which is expressed in the TRNs and their immediate precursors (11), whereas mec-3 encodes a LIM-type homeodomain protein that is expressed only in the terminal cells (12). Mutants lacking unc-86 activity do not generate the cells that become TRNs. In contrast, these cells are produced in mec-3 mutants, but they do not differentiate as TRNs (13). Therefore, mec-3 acts as a typical terminal selector gene. UNC-86 is needed not only to produce touch cell lineages but also to initiate transcription from the mec-3 gene (12). The maintained expression of mec-3 and the subsequent expression of TRN characteristics require the combinatorial action of MEC-3 and UNC-86; MEC-3 acts (as a coactivator with UNC-86) in an autoregulatory loop to maintain its own expression and to induce the expression of TRN-specific genes (13–18).

In this study, we investigate the function of the transcription factor ALR-1 in the regulation of TRN cell fate. Tucker et al. (19) found that alr-1, the C. elegans homolog of the Drosophila paired-type homeobox gene aristless (al), is expressed in these cells. Aristaleless-like proteins have important developmental roles in several organisms. alr-1 regulates the development of sensory neurons in the head and GABAergic motor neurons (19, 20). Drosophila aristless is needed for the development of several organs, including the aristae and wings (21, 22). The mouse ortholog Arx controls neuroblast proliferation and migration of GABAergic neurons (23, 24). Finally, defects in human Arx cause mental retardation (23, 25–28).

Here we report that, in contrast to mec-3, alr-1 does not act as a selector gene to determine the fate of the TRNs. Instead, ALR-1 ensures TRN differentiation by providing a second positive feedback circuit to maintain mec-3 expression. We show that loss of alr-1 produces variable touch sensation that correlates with increased variability in mec-3 expression. These results demonstrate that mec-3 autoregulation cannot maintain mec-3 mRNA levels. These levels are maintained in wild-type animals because ALR-1 restricts mec-3 expression to the high end of its expression range, a process that we call refinement. Our results demonstrate that stochastic fluctuations in gene expression can account for phenotypic variability and provide a mechanistic explanation for these genetic phenomena. In addition, we provide an example of how transcription factors like ALR-1 may function to reduce the variability of gene expression and ensure differentiation. Given that alr-1 and its homologs (aristaless and Arx) are required for the expression of other transcription factors (20, 23, 29, 30), our data support the hypothesis that proteins in the Aristaleless/ARX family may function as stabilizers of gene expression that ensure differentiation.

Results

alr-1 Expression in the TRNs Depends on MEC-3. Because MEC-3 controls TRN differentiation, we tested whether alr-1 expression required mec-3. Using a transcriptional fusion in which the alr-1 upstream sequences regulate gfp expression (P_{alr-1} gfp) (19), we confirmed that alr-1 is expressed in the ALM, PLM, and AVM

Author contributions: I.T. and M.C. designed research; I.T. performed research; A.v.O. contributed new reagents/analytic tools; I.T. and M.C. analyzed data; and I.T. and M.C. wrote the paper.

The authors declare no conflict of interest.

1To whom correspondence should be addressed. E-mail: mc21@columbia.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1101329108/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1101329108
TRNs, but not in the TRN PVM (Fig. S1A). This expression, but not that in other cells, was completely abolished in mec-3(e1338) animals (Fig. S1B). This mec-3–dependent expression is consistent with the finding of the heptanucleotide AATGCAT (the consensus core for MEC-3 targets) (31) in the alr-1 promoter. An ALR-1::GFP protein fusion was localized exclusively to the nucleus of all of the expressing TRNs throughout all larval development (Fig. S1C and Fig. S2), a result consistent with a role for ALR-1 in transcriptional control.

**alr-1 Mutants Are Variably Touch-Insensitive.** The TRNs sense gentle touch to the body in *C. elegans* (9). As expected from a role in the differentiation or function of the cells, mutations of alr-1 (including the null mutation oy42) (20) cause touch insensitivity (Fig. L4). The touch insensitivity of alr-1(oy42) mutants could be rescued with a genomic fragment of alr-1 using a mec-3 promoter (Fig. L4), indicating that the touch-insensitive phenotype was a consequence of loss of alr-1 function cell-autonomously in the TRNs.

The alr-1 touch-insensitive phenotype was variably expressed, ranging from animals that were indistinguishable from wild type (responding to 7–8 of 10 touches) to those that were completely insensitive to touch (responding to 1–2 touches) (Fig. L4). This great variability is not a characteristic of all strains with partial touch insensitivity because the loss of mec-18 results in partial, but more restricted, touch insensitivity (Fig. L4). Such incomplete penetrance and variable expressivity have also been observed for phenotypes associated with the AWA and ASG cells in alr-1 mutants (20). Individuals with a particular touch response showed similar touch responses during the course of development, suggesting that this phenotypic variability was set early during their differentiation and remained stable (Fig. S3).

The variability of the alr-1 phenotype appeared to be cell-specific. We observed that individual alr-1 mutants respond to anterior and posterior touch to different degrees, implying a cell-specific requirement for ALR-1 in ALM/AVM (sensing anterior touch) and PLM (sensing posterior touch) neurons, respectively. We also found that the ability of alr-1 mutants to be attracted to diacetyl, which shows a similar variability in alr-1 mutants (20), was independent of their ability to respond to touch, a result that is consistent with cell-autonomous activity.

**alr-1 Is Required Throughout Development for TRN Function.** alr-1 is expressed in the TRNs at all larval stages (Fig. S2). To determine whether alr-1 was needed for touch sensitivity throughout this time, we tested the touch sensitivity of adults that expressed alr-1 for different amounts of times in larvae. We generated a strain that expressed alr-1 in a temperature-dependent manner by inserting a mec-8–dependent intron before the coding region of alr-1 and using a temperature-sensitive allele of mec-8 (32). The extensive temperature-sensitive period (region of crossover between the downshift and upshift curves) (Fig. L1B) seen for touch sensitivity in this strain suggests that alr-1 activity is needed throughout larval development, which is consistent with the ALR-1::GFP expression pattern.

**ALR-1 Functions in TRN Differentiation as a Transcriptional Activator.** Aristalless-like proteins are thought to act as either transcriptional activators or transcriptional repressors (29, 30, 33–37). In *C. elegans*, Melkman and Sengupta (20) found that alr-1 blocked expression of the DD motor neuron marker f1p-13 in LD motor neurons and both down-regulated and up-regulated lin-11 expression in the AWA and ASG neurons, respectively. These changes, however, could be indirect.

To test how ALR-1 acts in the TRNs, we determined whether we could restore ALR-1 function with fusions of the ALR-1 homeodomain (ALR-1HD) with either the activation domain of VP16 (VP16AD::ALR-1HD) or the repressor domain of Engrailed (EnRD::ALR-1HD). The VP16AD::ALR-1HD fusion partially rescued the touch-insensitive phenotype of alr-1(oy42) mutants (three of five stable lines), but the EnRD::ALR-1HD fusion did not (six stable lines) (Fig. S4). These results suggest that ALR-1 affects TRN fate by acting mainly as a transcriptional activator.

These observations are consistent with the lack of the octapeptide domain and polyalanine tract 4 domain in ALR-1 (28), which McKenzie et al. (38) showed were needed for ARX repression. We cannot, however, exclude the possibility that ALR-1 can act as a repressor because our VP16 fusion only partially rescued the alr-1 mutant phenotype in the TRNs and did not rescue the chemosensory phenotype of alr-1 mutants (although the Engrailed fusion also failed to rescue this phenotype).

**alr-1 Is Required Selectively for TRN Gene Expression.** Consistent with the loss of touch sensitivity and ALR-1 being a homeodomain transcription factor, we found that the expression of genes required for TRN function was reduced in alr-1 mutants. All three available mutant alleles of alr-1, including the null allele oy42, greatly reduced MEC-18 protein (Fig. S5A) and mRNA (Fig. S5B). The decreased expression was more dramatic at the early larval stages of the animals although it was still obvious at later stages (Table S1).

The expression of mec-2, mec-7, and mec-17, detected using antibodies or GFP fusions, was similarly reduced by the oy42
and ALR-1 is required to maintain expression. (mec-3 is not sufficient for wild-type expression of unstable GFP from the mec-3 promoter (P_mec-3 praja::gfp) (Fig. 3A) together (Table S1) and the amount of mec-3 mRNA (Fig. S6B) was reduced in the TRNs (except PVM) (Fig. S6C) of young larvae and adult alr-1 (oy42) animals. The reduction of mec-3 expression observed with P_mec-3 praja::gfp displayed a similar variability to that observed with MEC-18 immunostaining (Table S1). [We tried to test whether addition of mec-3 (13, 41) whereas its maintained expression requires both MEC-3 and UNC-86 (13). For example, mec-3 mutants expressed P_mec-3 gfp less often (~60%) and less intensely than wild type did at hitting (Table S1). This expression disappeared as animals matured. This initial mec-3-independent expression was also alr-1–independent (Fig. 3B and Table S1). Moreover, expression from unc-86 promoter-gfp fusions was different with and without alr-1 (Fig. S7).

The extended requirement for alr-1 in touch-sensory implicates it in the maintenance of mec-3 expression. To test this hypothesis, we measured P_mec-3 praja::gfp expression in adult or alr-1 mutant animals, expressing ME8–regulated alr-1 (P_alr-1 intron 9::alr-1) either early or late in development (before or after the L2 stage). Expression of alr-1 only before the L2 stage was not sufficient for maximum levels of P_mec-3 praja::gfp expression in adults, but expression of alr-1 after the L2 stage was (Fig. 3C).

ALR-1 Enhances MEC-3/UNC-86–Induced Transcription from the mec-3 Promoter. A likely hypothesis is that ALR-1 directly activates mec-3 expression. To investigate this possibility, we examined its stimulation of β-galactosidase production from several promoters in yeast cells. We inserted P_mec-3 or P_unc-119 upstream of the yeast minimal promoter P_cyc-1 to drive the expression of β-galactosidase in the presence of different combinations of MEC-3, UNC-86, and ALR-1. MEC-3 or UNC-86 alone did not stimulate β-galactosidase expression from the basal promoter (P_cyc-1), the mec-3 promoter (P_mec-3 P_cyc-1), or the pan-neuronal unc-119 promoter (P_unc-119 P_cyc-1). Together, MEC-3 and UNC-86 increased β-galactosidase activity 30-fold from P_mec-3 P_cyc-1, but not from the other promoters (Fig. 4A). ALR-1 also increased expression from P_mec-3 P_cyc-1 and not the other promoters, but only slightly (1.4-fold) (Fig. 4A) and it did not have any effect on transcription in combination with either MEC-3 or UNC-86. ALR-1, however, had a greater effect in the presence of both MEC-3 and UNC-86, increasing expression to 60-fold (Fig. 4A). These results show that ALR-1 can enhance the bulk mec-3 expression and echo the in vivo requirement for alr-1 for TRN-specific gene expression.

Our results demonstrate that mec-3 autoregulation without alr-1 is not sufficient for maintenance of mec-3 expression; alr-1 provides a second positive feedback loop that enhances mec-3 expression and maintains TRN differentiation (Fig. 4B).

TRN Expression of mec-3 Is More Variable in alr-1 Mutants. Our yeast expression studies demonstrate that ALR-1 can enhance mec-3 expression, but these assays provide a bulk assessment of ALR-1 activity and do not indicate how ALR-1 increases mec-3 expression. Using single-molecule fluorescence in situ hybridiza-

**Fig. 2.** ALR-1 is needed for TRN expression of mec-17::gfp but not for the expression of P_unc-119 gfp. (A) mec-17::gfp expression in wild-type (Upper) and oy42 mutant (Lower) animals. GFP levels remain unaffected at the PVM neuron where alr-1 is not expressed. (Scale bars: 20 μm.) (B) P_unc-119 gfp is expressed equally in TRNs (ALM, Left; PLM, Right) of wild-type and alr-1 (oy42) adults. (Scale bars: 5 μm.)
tion (SM-FISH) (42) that renders each mRNA visible as a single fluorescent spot, we found that ALR-1 normally reduces the variability of the expression of mec-3 (Fig. 5A and B). mec-3 expression is much more variable in ALM and PLM neurons in alr-1 first-stage larvae than in wild-type larvae (coefficient of variation: 13.7 and 27.8 for wild-type and alr-1 PLM neurons and 13.3 and 26.3 for wild-type and alr-1 ALM neurons). The maximum number of mec-3 mRNAs is the same for both strains, but the range of values is greater in alr-1 mutant TRNs. Moreover, consistent with cell-autonomous activity of alr-1, we found that the number of mec-3 mRNAs in individual ALM and PLM neurons varied independently (Table S3 and S4). Thus, ALR-1 restricts the stochastic expression of mec-3.

The reduced mec-3 expression and the greater variability of mec-3 expression in alr-1 mutants reflect the phenotypic variability seen in these mutants. Nonetheless, the maximum level of expression was the same for both strains. In other words, the presence of alr-1(+) restricts mec-3 expression to the high end of its range, but does not increase the range. We call this process refinement to distinguish it from enhancement (Fig. 5C). In contrast, expression of mec-3 in FLP neurons, which do not detectably express alr-1, was the same in wild-type and mutant animals (Fig. 5B). These results indicate that alr-1 increases mec-3 expression by restricting mec-3 expression variability and suggest that the phenotypic variability observed in alr-1 mutants correlates with variability in mec-3 gene expression in the TRNs.

Discussion

Incomplete penetrance and variable expressivity are often characteristic features of mutant phenotypes (1–3). In this study, we provide evidence that links stochastic fluctuations in gene expression to phenotypic variability. In addition, we provide an explicit demonstration of how transcription factors can reduce variability in gene expression and ensure development.

Earlier studies have shown the cooperative role of the MEC-3 and UNC-86 transcription factors in defining the terminal fate of the TRNs and in maintaining the expression of MEC-3 (13–18). Here we demonstrate that, due to stochastic fluctuations, MEC-3 autoregulation is insufficient to maintain TRN differentiation.

The second positive feedback loop provided by ALR-1 enhances mec-3 expression by reducing the variability in mec-3 expression. We propose that regulation of the variability in gene expression is important for controlling differentiation and that genes like
**alr-1** function to overcome stochastic fluctuations in gene expression and ensure differentiation.

**ALR-1 Maintains Expression of MEC-3.** Our results suggest that ALR-1 is a transcriptional activator that controls TRN activity by maintaining optimal *mec-3* expression throughout larval development. First, ALR-1 enhanced *mec-3* expression in mec-3(proximal promoter) expression by providing the *lacZ* mutant phenotype. The hypothesis that ALR-1 enhances transcription from the *mec-3* promoter is further supported by ChIP-Seq experiments, demonstrating a strong association of ALR-1 with the *mec-3* proximal promoter (25th of 7,009 potential targets) (43).

The interactions of ALR-1 and MEC-3 provide a positive feedback loop: MEC-3 controls the expression of *alr-1*, and ALR-1 regulates the maintenance of *mec-3* expression. Even though it is expressed downstream of selector genes, ALR-1 appears to participate in the maintenance of *mec-3* expression by providing a second positive feedback loop parallel to the autoregulatory *mec-3* feedback loop (Fig. 4B). Regulatory interactions involving LIM-type homeodomain transcription factors and *aristaless* homodomain transcription factors have also been described in pretarsus development in *Drosophila* (AL and LIM1) (27, 29). Chemosensory neuron development in *C. elegans* (ALR-1 and the LIM1 ortholog *lin-11*) (20) and in the neocortex and ganglionic eminence (ARX and Lhx6) and thalamic eminence (ARX and LHX9) in mice (23).

**alr-1** does not function as a general transcription factor in the TRNs, but appears to activate the *mec-3*-dependent pathway. In addition to amplifying *mec-3* expression, which would indirectly increase the production from *mec-3* targets, *alr-1* may also directly affect the expression of those targets. The ChIP-Seq analysis of ALR-1 (43) shows a strong association of the protein with *mec-2*, *mec-7*, *mec-4*, *mec-9*, *mec-10*, *mec-14*, and *mec-17* and a low association with *mec-1*, *mec-6*, and *mec-12*, suggesting that ALR-1 may directly affect their transcription. These observations suggest that ALR-1 participates in *mec-3* expression and in the expression of at least some of the *mec-3*-target genes (Fig. 4B).

**ALR-1 Prevents Phenotypic Variability by Controlling Variable Gene Expression.** Unlike MEC-3, ALR-1 does not act as a terminal selector gene (6); it ensures but does not determine the fate of the TRNs. TRNs in *alr-1* mutants still express *mec-3* target genes, albeit at lower levels. In addition, the PVM neurons, which do not express detectable *alr-1* in wild-type animals, nonetheless share differentiated features with the other TRNs. The PVM cells express *mec-3* target genes at lower amounts (e.g., Fig. 2A), and they contain fewer of the TRN-specific microtubules than the other TRNs (44).

Our results link stochastic fluctuations in gene expression to incomplete penetrance and variable expressivity and provide a mechanism to explain these genetic phenomena. In our model, removal of a stabilizing component (e.g., ALR-1) results in stochastic expression that is so broad that some, but not all, cells express the selector gene (*mec-3*) at levels below the threshold level needed for cell function (touch sensitivity). Wild-type ALR-1 stabilizes gene expression and ensures that it is above the threshold needed for function.

These results also provide an explicit demonstration for how variability may be prevented during development when cell fate accuracy is essential. Because incomplete penetrance, phenotypic variability and variable expression of transcription factors have been reported for many mutants (1, 3, 45), we suggest that regulation of stochastic gene expression may be important for cell differentiation and function.

Although touch insensitivity of *alr-1* mutants differs from animal to animal, individual animals maintain the same level of touch insensitivity through the larval period. Thus, the stochastic processes that define *mec-3* expression levels in *alr-1* mutants appear to be set early and remain stable during the development of the TRNs. The temperature-shift experiments, however, demonstrate that ALR-1 can reset the expression level of *mec-3* at later times. Our experiments indicate that MEC-3 (with UNC-86) sets the maximum levels of *mec-3* expression and ALR-1 ensures (by overcoming stochastic events) that *mec-3* expression will always be sufficient. In general, the effectiveness of any transcription factor will depend on its nuclear concentration, its capacity to bind DNA, and the efficiency with which it activates transcription. If the concentration of the transcription factor is not saturating, increasing the concentration should increase transcription and shift the range of expression to a higher position. Alternatively, if the concentration is already maximal, ensuring the efficiency of transcription (either by securing the association with the promoter or by adding to the intrinsic efficiency of activation) should shift expression to the top of the range (refinement in Fig. 5C). We feel that the alteration in *mec-3* expression caused by ALR-1 may be of this latter type.

Because variability similar to that seen in touch sensitivity was described for the Dyf, Osm, Daf-d, chemosensation phenotypes, and *lin-11* expression in AWA and ASG neurons of *alr-1* mutants (20), the control of the variability of gene expression may be a general function for ALR-1. *alr-1* does not function as a classical selector gene, but it plays the critical role of fine-tuning the autoregulation of selector genes to ensure differentiation. This hypothesis is further supported by our preliminary results showing that ALR-1 reduces the variability of expression of another selector gene, *unc-30*, that defines the fate of the DD and VD GABAergic neurons (46) (Fig. S8). In addition, ChIP-Seq results (43) show that ALR-1 associates with *unc-30* and with some of the UNC-30 downstream targets (*unc-25, unc-46, unc-47*). These observations suggest that ALR-1 plays a more general role in ensuring neuronal terminal differentiation by securing the efficiency of the expression of selector genes and probably of their downstream targets.

Because *alr-1*, *Drosophila aristaeless*, and mouse Arx regulate the LIM homeobox genes *lin-11* (20), *Lim1* (29, 30), and *Lhx6/Lhx9* (23), respectively, we hypothesize that Aristaeless proteins may have a conserved role across species in regulating and restricting the variability of the expression of LIM-homeodomain transcription factors and, perhaps, other selector genes. Unfortunately, the null phenotypes of these genes are so severe that variability cannot be easily assessed (23, 47).

**Materials and Methods**

**General Procedures.** Unless otherwise indicated, strains were maintained and studied at 20° according to Brenner (48) on OP50 *Escherichia coli*. Detailed information on the strains used, reporter plasmids, plasmid constructions, microinjections, quantitative real-time PCR, and microscopy and immunofluorescence is given in SI Materials and Methods.

**Touch Assays.** We assayed gentle touch sensitivity in blind tests as described (9). We quantified the response by counting the number of responses to 10 touches delivered alternately near the head and tail in 20 animals (20 stable transformants for the rescuing experiments).

**SM-FISH.** We designed oligonucleotide probes and performed SM-FISH (42) on L1 stage animals as described at http://www.singlemoleculefish.com/. Forty-three 20-nucleotide probes for *mec-3* mRNA were synthesized by BioSearch Technologies and coupled to Cy5 (GE Amersham). We imaged the animals using a Nikon TE 2000 microscope equipped with a Princeton Instruments Micromax 1024B camera and appropriate filters for Cy5. We collected stacks of 30–40 images spaced 0.3 μm apart for each individual neuron and counted the number of fluorescent spots per neuron using ImageJ (http://rsb.info.nih.gov/ij/).

**Yeast Transcription Assays.** We constructed yeast strains expressing *P_mec-3* > lacZ, *P_mec-3* > lacZ, and *P_mec-12* > lacZ by digesting plasmids with these DNAs with Apal and integrating them in the URA locus of the yeast strain YPH499.
A gene for all seasons.

Development 132:313–324.