TITLE

***pop-1/TCF*, *ref-2/ZIC* and T-box factors regulate the development of anterior cells in the *C. elegans* embryo**

AUTHORS

Jonathan D. Rumley[1], Elicia A. Preston[1], Dylan Cook[1], Felicia L. Peng[1], Amanda L. Zacharias[2,3], Lucy Wu[1], Ilona Jileaeva[1], John Isaac Murray[1,4]

1. **Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104**
2. **Division of Developmental Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229**
3. **Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH 45267**
4. **Corresponding author - jmurr@pennmedicine.upenn.edu**

ABSTRACT

Patterning of the anterior-posterior axis is fundamental to animal development. The Wnt pathway plays a major role in this process by activating the expression of posterior genes in animals from worms to humans. This observation raises the question of whether the Wnt pathway or other regulators control the expression of the many anterior-expressed genes. We found that the expression of five anterior-specific genes in *Caenorhabditis elegans* embryos depends on the Wnt pathway effectors *pop-1/TCF* and *sys-1/*β-catenin. We focused further on one of these anterior genes, *ref-2/ZIC,* a conserved transcription factor expressed in multiple anterior lineages. Live imaging of *ref-2* mutant embryos identified defects in cell division timing and position in anterior lineages. *Cis*-regulatory dissection identified three *ref-2* transcriptional enhancers, one of which is necessary and sufficient for anterior-specific expression. This enhancer is activated by the T-box transcription factors TBX-37 and TBX-38, and surprisingly, concatemerized TBX-37/38 binding sites are sufficient to drive anterior-biased expression alone, despite the broad expression of TBX-37 and TBX-38. Taken together, our results highlight the diverse mechanisms used to regulate anterior expression patterns in the embryo.

INTRODUCTION

Anterior-posterior patterning regulation

Proper anterior-posterior patterning is critical for animal embryonic development and requires the Wnt pathway across all bilaterian animals from worms to humans. Typically, posterior cell identities are induced by posteriorly-produced Wnt ligands signaling through the canonical Wnt pathway [1]. In this pathway, signal transduction in response to Wnt activates posterior-expressed gene transcription through the transcription factor TCF and its co-activator β-catenin [2]. Although this conserved role for the Wnt pathway in regulating the expression of posterior genes is well documented [3], much less is known about how genes expressed in anterior cells are regulated.

Caenorhabditis elegans has been widely used to study anterior-posterior patterning because it is genetically tractable, its embryonic lineage is invariant, and most of its developmental regulators are highly conserved in vertebrates and other animals [4–10]. Also, the majority of embryonic cell divisions are oriented along the anterior-posterior axis, and require the Wnt pathway to differentiate the two daughter fates [11–13]. Following each anterior-posterior cell division, the Wnt pathway is activated in the more posterior daughter cell, leading to high nuclear SYS-1/β-catenin and low nuclear POP-1/TCF, whereas the anterior daughter cell has low nuclear SYS-1 and high nuclear POP-1[11–13]. The combination of high nuclear SYS-1 and low nuclear POP-1 in posterior cells makes it stoichiometrically favorable for POP-1 to be bound to posterior-expressed targets as a complex with SYS-1, thereby driving their expression [14–16]. In contrast, anterior cells have low nuclear SYS-1 and high nuclear POP-1; for posterior-expressed targets, this results in POP-1 being bound in the absence of SYS-1, allowing it to recruit co-repressors, including UNC-37/groucho [17] (Fig 1A). Loss of POP-1/TCF results in posterior genes having either reduced expression in posterior cells or increased expression in anterior cells, indicative of its dual roles as an activator and repressor. Loss of SYS-1/β-catenin results only in down-regulation of posterior genes in posterior cells, consistent with an exclusive role as a posterior activator [13]. These observations leave open the question of how anterior genes are regulated in *C. elegans*.

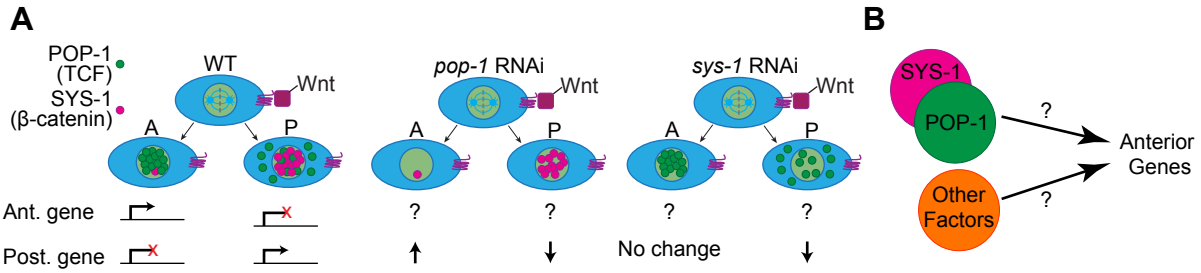


Fig 1. How are anterior genes regulated?
(A) The Wnt/β-catenin asymmetry pathway regulates anterior-posterior patterning by producing asymmetric nuclear concentrations of POP-1/TCF and SYS-1/β-catenin in anterior and posterior sister cells. Low nuclear concentrations of POP-1 and high nuclear concentrations of SYS-1 in the posterior cell activate expression of posterior genes and are associated with the lack of expression of anterior genes. Conversely, high nuclear concentrations of POP-1 and low nuclear concentrations of SYS-1 in the anterior cell repress posterior genes and are associated with expression of anterior genes. Effects of *pop-1* and *sys-1* RNAi on posterior genes are shown; effects on most anterior genes are unknown. (B) In this work, we test whether *pop-1*/TCF and *sys-1*/β-catenin regulate anterior genes in the *C. elegans* embryo, and ask what other factors also regulate anterior expression.

Previous work suggests at least three possible mechanisms by which anterior genes could be regulated. 1) They may be regulated by the Wnt pathway indirectly, with posteriorly-expressed canonical Wnt targets repressing the expression of anterior genes in posterior cells. 2) They may be directly regulated by Wnt pathway components acting in a modified (or "opposite") manner. 3) They may be regulated by a non-Wnt-related mechanism (Fig 1B).

An example of indirect Wnt pathway regulation through a posterior repressor is seen in the *C. elegans* embryonic EMS lineage, in which cells derived from the posterior daughter of EMS, E, express the POP-1-activated, gut-specifying transcription factor *end-1*. In turn, *end-1* represses the transcription factor *ceh-51*, which is normally expressed in and specifies the lineage derived from the anterior daughter of EMS, MS [18,19].

Conversely, the *C. elegans* anterior gene *ttx-3* is directly regulated by Wnt pathway components. *ttx-3* helps specify the AIY neuron class and is regulated by POP-1 and SYS-1 in a manner dependent on the sole *C. elegans* ZIC family transcription factor REF-2 [20]. After the AIY grandmother divides, *ttx-3* is expressed in the anterior daughter, the AIY mother, but not in the posterior daughter. In the AIY mother, the Wnt pathway is inactive, with low nuclear SYS-1 and

high nuclear POP-1, and both *pop-1* and *ref-2* are required to activate *ttx-3* expresion. In its posterior sister, the Wnt pathway is active, with high nuclear SYS-1 levels and low nuclear POP-1 levels, and *sys-1* is required to repress expression of *ttx-3*. Furthermore, POP-1 and REF-2 can directly interact, suggesting that REF-2 and POP-1 bind as a complex to activate *ttx-3* expression in the AIY mother. The role of SYS-1 is less clear but some evidence suggests SYS-1 may bind to the POP-1/REF-2 complex to repress expression, or may sequester away the limited POP-1, such that POP-1 cannot interact with REF-2 [20,21]. Similar regulation, termed "opposite Wnt regulation", has also been seen in other organisms such as *Drosophila* [22].

Because of the role of POP-1 and SYS-1 in regulating the anterior-specific expression of *ttx-3*, we hypothesized that these genes may also regulate other anterior genes. POP-1 and SYS-1 could act as general anterior expression regulators, interacting with lineage-restricted co-regulators to ensure appropriate expression of anterior genes in the correct lineages [20,21]. We tested this by mining single-cell-resolution expression data to identify transcription factors expressed in anterior-specific patterns, and show that of five we tested, all require *pop-1* and/or *sys-1* for either anterior expression or posterior repression. We focus in more detail on a single anterior-specific gene, *ref-2/ZIC*. By automated lineage analysis of mutants we find that *ref-2* is required for normal cell division timing and cell position in *ref-2* expressing lineages. Embryonic expression of *ref-2* is driven by at least three developmental enhancers, two of which drive early embryonic expression and one of which drives expression in later-stage embryos. The most distal enhancer (3.9 kb upstream of the transcription start site) drives highly anterior-biased expression. Functional dissection revealed unexpected redundancy within this enhancer, and a role for the T-box transcription factors *tbx-37* and *tbx-38* in driving anterior-specific expression. Surprisingly, concatemers of a single T-box binding site reiterate much of the anterior-specific expression pattern of *ref-2* in early embryos, suggesting a key role for these genes in driving anterior expression.

RESULTS

Many early embryonic transcription factors have anterior-specific expression

We used large-scale single-cell expression databases derived from time-lapse 4D imaging of embryos expressing reporter genes to identify anterior-expressed genes (Fig 2A). We define "anterior" or "posterior" lineages as sets of related cells derived from either the more-anterior or more-posterior daughter after a specific cell division (Fig 2B). Notably, this is distinct from physical position; for example, a cell that divides near the posterior pole of the embryo gives rise to both an anterior and posterior daughter lineage, both of which are located in the posterior half of the embryo (Fig 2B, left side). Following convention [5,23], all lineage trees in this manuscript are drawn with the more anterior daughters on the left and the posterior daughters on the right of each bifurcation. Earlier studies identified a strong tendency for genes to be expressed in either multiple anterior or multiple posterior lineages during early and mid-embryogenesis; these frequently appear as "lineally repetitive" patterns, with multiple related anterior or posterior cousin lineages all expressing the same gene (e.g. *nob-1* and *ref-2* in Fig 2D,E) [24,25]. Both anterior-specific and posterior-specific gene expression were similarly common. While posterior specific genes (e.g. Fig 2D) appear to be largely regulated by canonical Wnt signaling [13], it is less clear how anterior-specific patterns are regulated. We mined existing literature and large-scale databases of fluorescent reporter expression patterns (Fig 2A) [24–26] to identify 19 conserved genes expressed in anterior lineage-specific patterns (Fig 2C); the anterior-biased expression of these genes is also seen in embryonic single-cell RNA-seq data [27].

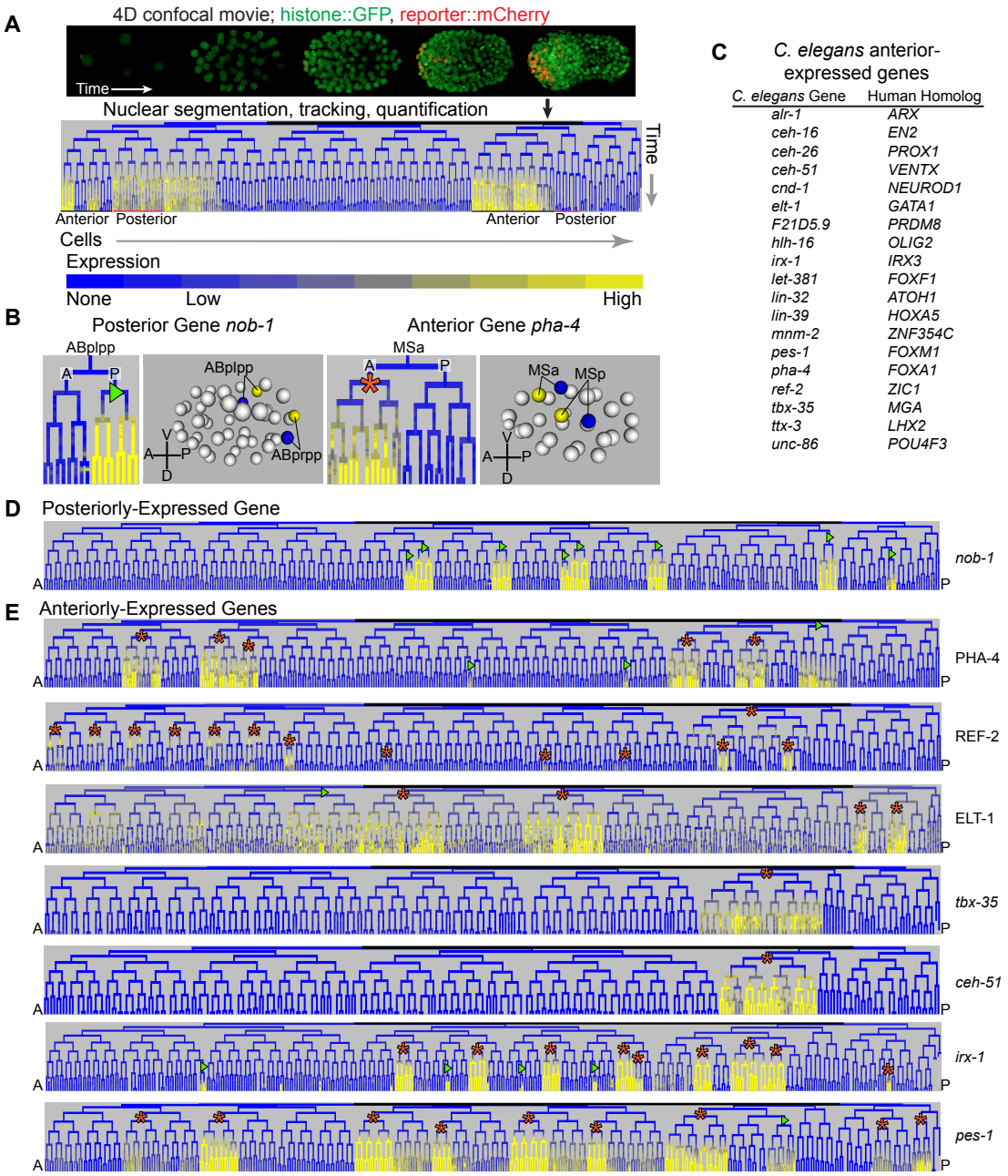


Fig 2. Identification of genes with anterior-biased expression.

(A) Automated lineaging traces lineages and quantifies expression from 4D confocal movies. In the lineage trees, vertical lines represent cells, and horizontal lines represent cell divisions. Most cells divide along the anterior-posterior axis, with anterior cells depicted on the left branch and posterior cells on the right branch. (B) Posterior and anterior genes are denoted as such based on their expression in cells descended from either a posterior or anterior sister cell, respectively, following an anterior-posterior cell division. Anterior and posterior genes can generally be expressed in cells from any part of the embryo. Anterior and posterior founder cells are labeled with orange asterisks/green triangles respectively. (C) List of anterior-biased genes in the EPiC database along with their predicted human homologs. (D) Expression pattern of a posterior gene, *nob-1/Hox9-13*. (E) Expression patterns of several anterior genes; most are expressed in unique combinations of mostly anterior lineages.

We confirmed the anterior-biased expression of seven of these by live confocal imaging followed by automated cell tracking and lineage tracing (Fig 2A, E, Fig S1, Fig S2, Fig S3) [26,28]. For each gene, we collected 3D timelapse movies of transgenic embryos expressing a histone-mCherry (*pha-4*, *ref-2*, *elt-1*, *tbx-35*, *irx-1*, and *pes-1*), or GFP (*ceh-51*) reporter under the control of upstream regulatory sequences of each gene (referred to here as "promoter reporters"). For five genes we also collected images for fosmid transgenes (*irx-1, elt-1, pha-4* and *ref-2)* or CRISPR knock-ins (*ceh-51*) in which the protein is fused to GFP and surrounded by its normal genomic context ("protein reporters"). Each embryo also contained a ubiquitously expressed second-color, histone-GFP or histone-mCherry fusion, used for cell tracking. We identified each nucleus at each time point and traced them across time using StarryNite [28,29] cell tracking software, and used AceTree [30,31] for subsequent manual error correction and validation that the extracted lineages were correct. Finally, we quantified the expression of the reporter in each nucleus at each time point (Fig 2A) [26]. The results confirm the anterior-specific expression of each gene, and, in some cases, identify previously unknown expression patterns (Fig S1, Fig S2, Fig S3). For each promoter reporter, the anterior-specific patterns were consistent, but the protein fusion reporters all had additional expressing lineages and dynamics indicative of distal enhancers and post-transcriptional regulation (Fig S1, Fig S2, Fig S3).

The Wnt pathway effectors *pop-1* and *sys-1* are required for normal anterior-specific expression

Most genes expressed preferentially in posterior lineages require the Wnt effector transcription factor POP-1/TCF either for activation in posterior cells (together with its coactivator SYS-1/β-catenin) or for repression in anterior cells, and at least some are direct targets [12,13,32–35]. Although *pop-1* is required for the normal expression of some anterior genes [20,32,36], it is unclear whether *pop-1* and *sys-1* are generally required for anterior-specific expression, as they are for posterior-specific expression.

To test this, we measured the expression of five anterior gene reporters (*tbx-35*, *ceh-51*, *elt-1*, *irx-1*, and *ref-2*) after *pop-1* and *sys-1* RNAi by live 3D imaging and lineage tracing as described above. We used promoter reporters for *tbx-35*, *ceh-51*, *irx-1*, and *ref-2* because our wild-type

190 expression data indicated these reporters are sufficient for anterior regulation (Fig S1, Fig S2, Fig
191 S3). For *elt-1*, anterior-specific expression was more robust for the protein reporter so we tested
192 that reporter's dependence on *pop-1* and *sys-1* (Fig 3; Fig S4).
193
194 Reporters of two of the genes we tested, *tbx-35* and *ceh-51*, are expressed solely in the anterior
195 lineage MS (Fig S2; Fig S4A, B), and previous qualitative expression analyses indicated that
196 expression persists in the absence of *pop-1* [19,36,37]. Consistent with this, we detected expression
197 of both genes in the MS lineage after *pop-1* RNAi, however this expression was significantly
198 decreased. We also observed increased expression of both genes in the E lineage (posterior sister
199 of MS), although this expression remained lower than both the wild-type and *pop-1* RNAi MS
200 levels (Fig 3A, B; Fig S4A, B; Table S4). We also observe low-level ectopic expression of *tbx-35*
201 in the E lineage under *sys-1* RNAi. These results indicate that both *tbx-35* and *ceh-51* require *pop-*
202 *1* for full expression in the MS lineage and for repression in the E lineage (as shown previously
203 [19,36]), and that *sys-1* is required to repress *tbx-35* in the E lineage.
204
205 *elt-1* reporters are expressed in several early anterior lineages that primarily give rise to ectodermal
206 fates (ABpla, ABpra, Caa and Cpa), and in one posterior lineage (ABarp) (Fig S1; Fig S4C). In
207 the Caa and Cpa lineages, anterior lineages that express both the promoter and protein reporters,
208 *pop-1* RNAi results in decreased *elt-1* reporter expression. However, RNAi for neither *pop-1* nor
209 *sys-1* affects expression in their posterior sister lineages Cap and Cpp (Fig 3C; Fig S4C; Table S4).
210 Similarly, *pop-1* RNAi results in reduced expression of the protein reporter in the ABpla and
211 ABpra lineages, while RNAi knockdown of neither *pop-1* nor *sys-1* has much effect on their
212 posterior sister lineages ABplp and ABprp (Fig S4C; Table S4). Thus, in both the ABp and C
213 lineages, *pop-1* is required for anterior expression of *elt-1*, but neither *pop-1* nor *sys-1* is required
214 for posterior repression.
215
216

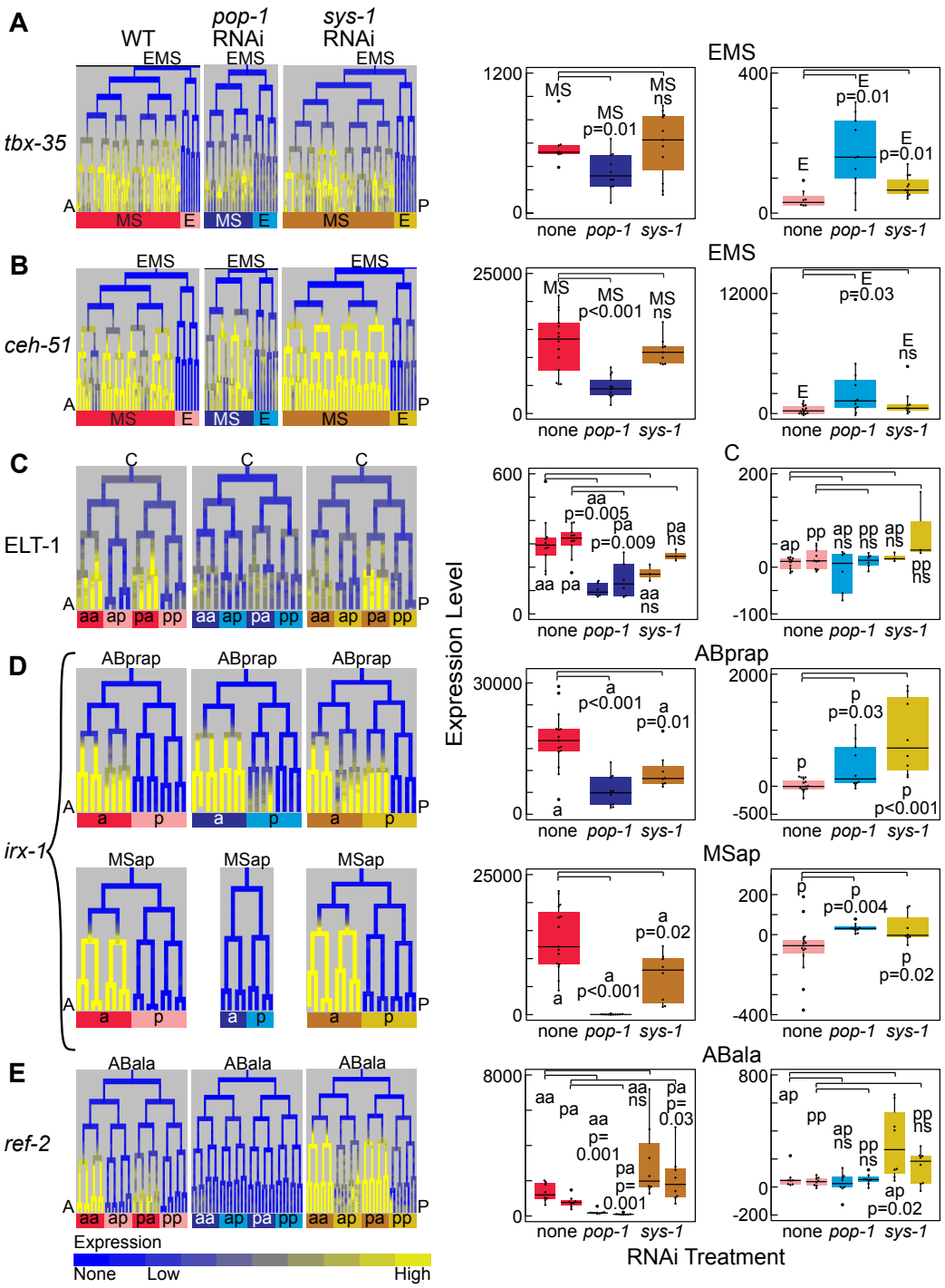


Fig 3. Regulation of anterior genes by Wnt effectors *pop-1* and *sys-1*.
(A-E) Expression pattern (left) and quantification (right) of reporters for the anterior genes *tbx-35* (A), *ceh-51* (B), ELT-1 (C), *irx-1* (D), and *ref-2* (E) in specific anterior lineages and their posterior sisters. Expression is shown under WT conditions and following *pop-1* or *sys-1* RNAi. Expression quantification (right) is the mean expression across all measurements within that lineage from when the reporter is normally first detected until the last time point indicated in the lineage (details in Methods).

irx-1 reporters are expressed early in four related anterior sublineages of ABp, and in three anterior sublineages of MS, while in later embryos they are expressed in both anterior and posterior lineages (Fig S3; Fig S4D). Like for *elt-1, pop-1* RNAi reduces *irx-1* reporter expression in anterior ABp-derived sublineages, and two of the MS-derived expressing sublineages lose expression nearly completely. After both *pop-1* and *sys-1* RNAi, the *irx-1* reporter expression expands into several posterior sublineages of ABp whose anterior sisters normally express *irx-1* (Fig 3D; Fig S4D; Table S4). These results indicate that *irx-1* requires *pop-1* for expression in anterior lineages and both *pop-1* and *sys-1* for repression in posterior lineages.

Reporters for *ref-2* are expressed in the early embryo in all MS descendants and in several related anterior sublineages of ABa that arise at the 50-cell stage (Fig S1; Fig S4E). In later embryos, expression occurs in two anterior sublineages of MS, in several anterior ABp-derived sublineages, and in the Pn ventral epidermal blast cells, consistent with previous studies [38]. After *pop-1* RNAi, expression of the *ref-2* promoter reporter is significantly reduced in anterior sublineages of ABala (Fig 3E). As for other genes, *sys-1* RNAi caused de-repression of *ref-2* in several posterior sisters of the normally expressing sublineages, but we did not detect de-repression in these sublineages after *pop-1* RNAi (Fig 3E; Fig S4E; Table S4). Thus, *ref-2* requires *pop-1* for expression in anterior lineages and *sys-1* for repression in posterior lineages.

To summarize these perturbations, each anterior gene tested requires *pop-1* for full expression in anterior lineages. Several genes also require *pop-1* (*tbx-35*, *ceh-51* and *irx-1*) or *sys-1* (*tbx-35, irx-1* and *ref-2*) for repression in some posterior lineages. Thus, it appears that the expression of most anterior genes requires the Wnt pathway components POP-1 and SYS-1, but this dependency is complex as was seen previously for posterior genes [13].

The anterior-specific transcription factor *ref-2/ZIC* is required for normal patterns of cell division and cell position in expressing lineages.

The striking anterior-specific expression patterns observed for early embryonic TFs raises the question of whether these factors regulate the development of anterior lineages. To address this question, we focused on one anterior gene, *ref-2*. *ref-2* encodes a zinc finger transcription factor

256 homologous to mammalian ZIC genes and *Drosophila* Odd-Paired. Previous work showed that
257 *ref-2* is required to prevent "Pn" epidermal blast cells from fusing with the major hypodermal
258 syncytium in larval stages [38]. More recently, REF-2 was shown to activate the anterior-specific
259 expression of the neural regulator *ttx-3* by binding the *ttx-3 cis*-regulatory region and recruiting
260 POP-1 in a manner independent of POP-1 DNA binding activity in the AIY mother cell [20]. Since
261 *ref-2* is expressed in many other anterior lineages in early embryos, we investigated whether *ref-2*
262 mutants have developmental defects in those lineages.
263
264 We tested for embryonic lineage phenotypes at single-cell resolution by analyzing 3D time-lapse
265 movies of eight embryos homozygous for the deletion allele *ref-2(gk178)* and expressing a
266 ubiquitous histone-mCherry fusion for cell tracking. We used StarryNite to track the lineage of
267 each embryo, and used AceTree to curate the lineages through the >550-cell stage (six embryos)
268 or >350-cell stage (two embryos). We compared the position and cell cycle length of each cell in
269 each embryo to those from a reference of 17 wild-type embryos [39,40] to identify outliers
270 ("defects") (Fig 4A).
271
272

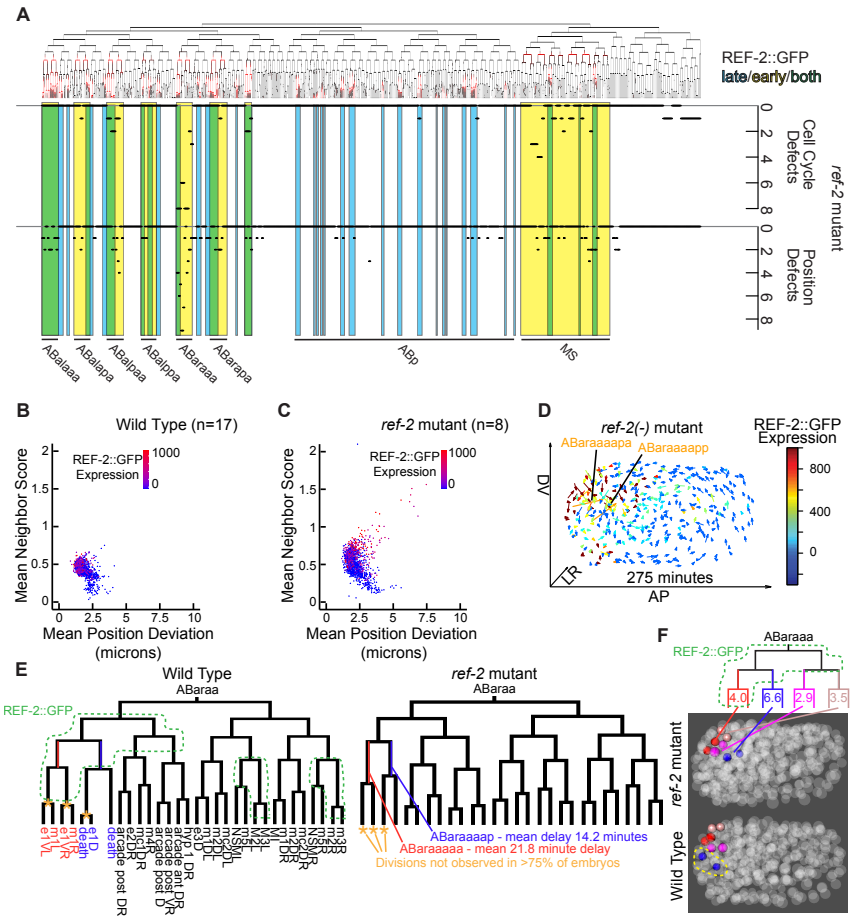


Fig 4. *ref-2* is required for proper division timing and positioning for many embryonic cells that are descended from cells that express *ref-2*.

(A) Lineage tree depicting the REF-2 protein expression pattern (top), and chart indicating the number of embryos out of eight homozygous *ref-2* mutant embryos that exhibit cell cycle defects and position defects in indicated cell lineages. Yellow indicates lineages that express *ref-2* early, blue indicates lineages that express *ref-2* late, and the overlap (green) indicates both. A corresponding analysis of wild-type embryos gives 0-1 defects per cell in all of 17 embryos tested [39]. (B-C) Scatter plot of mean neighbor score vs mean position deviation of cells of 17 wild-type (B) and eight *ref-2* mutant (C) embryos. Neighbor score is the ratio of each cell's distances to its ten closest wild-type neighbors between mutant and wild-type embryos. Points representing cells are colored based on WT REF-2::GFP expression levels [39]. (D) Plot of cell position deviations with arrows starting at the average WT position of cells and pointing to the average position in *ref-2* mutants. Arrows are colored by the WT expression levels of REF-2::GFP for each cell. The labeled cells (ABaraaaapa and ABaraaaapp) have the greatest mean cell position deviations. (E) ABaraa lineages of WT and *ref-2* mutant embryos, with cells expressing REF-2::GFP outlined on the WT tree. Two delayed and three missed cell divisions are highlighted on the *ref-2* mutant tree. (F) ABaraaa lineage with cells expressing REF-2::GFP outlined. The average position deviations in microns of terminal sister cells are indicated on the terminal branches of the lineage tree. Three dimensional projections of a WT and a *ref-2* mutant embryo are shown with the positions of the terminal ABaraaa lineage cells highlighted. ABaraaaapa and ABaraaaapp are outlined in the WT embryo projection as the cells with the greatest mean position deviation in the *ref-2* mutant.

We identified a total of 55 cells with early, late or missing divisions, and 54 of these were derived from REF-2::GFP expressing lineages ($P < 10^{-15}$). Overall 95% of cell cycle defects were delays or missed divisions, indicating a role for *ref-2* in promoting cell division. Similarly, we identified 68 position defects, of which 56 were in *ref-2* expressing lineages (Fig 4B, C, $P = 1.1 * 10^{-14}$, chi-squared test). Ten *ref-2*-expressing cells, and no non-expressing cells, had position defects in two or more embryos. The cells with the strongest position and cell cycle length defects were from the *ref-2*-expressing anterior ABaraaa lineage, which produces primarily anterior pharynx cells (e1, m1 and arcade cells). The most frequent cell cycle length defects were in ABaraaaaa and ABaraaaap (both defective in 8 of 8 embryos), which were delayed by 21.8 minutes and 14.2 minutes, respectively (Fig 4E). The largest position defects were in ABaraaaapa and ABaraaaapp, which were mispositioned on average 6 microns posterior of their wild-type position (Fig 4D,F). However, other defects were broadly distributed across *ref-2*-expressing lineages, in particular MS- and ABa-derived anterior sublineages (Fig 4A). In summary, the anterior lineage-expressed TF *ref-2* is required for the normal development of anterior lineages, although the low penetrance of many defects suggests that *ref-2* may act with other partially redundant regulators, similar to other early zygotic TFs [39,41–46].

Modular enhancers control *ref-2* embryonic expression

The anterior expression and phenotypes of *ref-2* raise the question of what sequences and regulators control this expression. To identify genomic sequences responsible for *ref-2* anterior expression, we first compared expression driven by the 31.9 kb genomic fosmid protein reporter to the shorter 4.1kb upstream ("promoter") reporter (Fig 5A, B, C; Fig S6). In general, the patterns were similar; both are expressed broadly in the MS lineage and in multiple anterior ABa-derived sublineages. In the ABa lineage, the promoter reporter shows background expression in some posterior sister sublineages suggesting the existence of repressive sequences outside of this promoter region. Additionally, the promoter fusion reporter is detected persistently in both the MS and AB lineages, whereas protein reporter expression is detected transiently, likely reflecting the use of a stable histone-mCherry reporter for the promoter reporter. Importantly, both the REF-2::GFP protein and *ref-2* promoter reporters drive expression in progenitors of most cells defective in *ref-2* mutants.

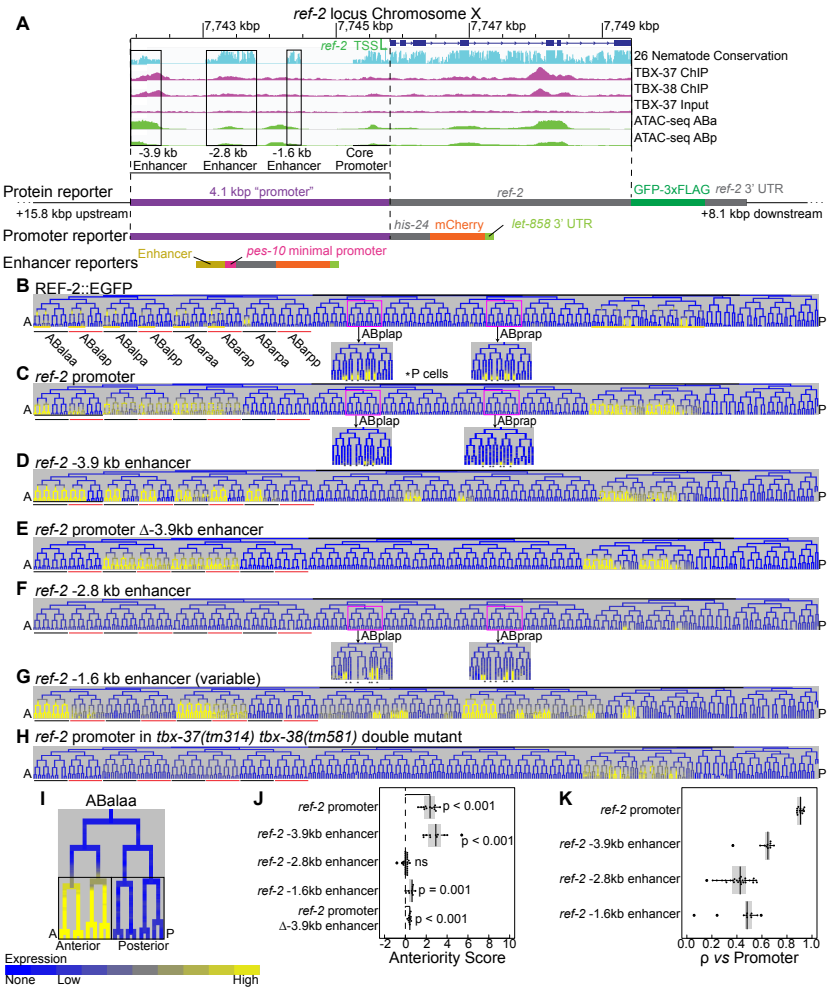


Fig 5. Identification of *ref-2* enhancers.

(A) Genome browser screenshot showing the *ref-2* locus on the X chromosome. Displayed tracks include the gene model, sequence conservation among 26 *Caenorhabditis* species [49], ChIP-seq data for TBX-37 and TBX-38 binding, and ATAC-seq data for cells in the ABa and ABp lineages [50]. Boxes indicate candidate enhancer regions as determined by conservation. Below the genome browser screenshot are models of the fluorescent reporter expression constructs we examined. (B-H) Lineage trees showing the expression patterns of *ref-2* protein (B), promoter (C), -3.9 kb enhancer reporter (D), promoter lacking the -3.9 kb enhancer (E), -2.8 kb enhancer reporter (F), -1.6 kb enhancer reporter (G), and promoter in the *tbx-37 tbx-38* double mutant (H) as determined by time-lapse confocal microscopy. Sublineages used in anteriority analyses are labeled in B and underlined in black (anterior) and red (posterior) under each lineage tree. Sublineages with transient early REF-2::EGFP (protein reporter) expression are underlined in yellow under the lineage tree in B. Pn ventral epidermal cells are indicated with asterisks in insets of panels B, C, and F. (I) Example lineage showing which cells were included in anteriority score analyses. Reporter expression levels were averaged across cells in the anterior or posterior lineages for the indicated cell generations. (J) Anteriority scores of the *ref-2* promoter, enhancers, and promoter lacking the -3.9 kb enhancer. Lineages used for analysis are ABala, ABalaa, ABalap, ABalpa, ABalpp, ABaraa, ABarap, ABarpa, and ABarpp (see Fig 5B). (K) Full somatic lineage correlation (Spearman's ρ) of each reporter's expression pattern with the mean expression pattern across embryos expressing the promoter reporter.

To identify the sequences in this region that regulate expression in these lineages, we sought to identify minimal portions of the region with enhancer activity. We identified three portions of the *ref-2* promoter conserved among *Caenorhabditis* species, suggesting they encode important functions (Fig 5A) [47–49]. The most distal of these regions is in open chromatin in ABa-derived cells as measured in a recently published ATAC-seq data set [50]. Also, this region is bound by the broadly-expressed ABa lineage transcription factors TBX-37 and TBX-38, as determined by a recently published ChIP-seq data set (Fig 5A) [50]. We fused each putative enhancer to a reporter cassette consisting of a minimal *pes-10* promoter and a *his-24*::mCherry reporter. The *pes-10* promoter is widely used, drives no consistent embryonic expression on its own [13,51,52] and is compatible with a wide variety of enhancers [13,51]. We introduced each reporter into worms and used StarryNite to determine in which cells each putative enhancer drives expression.

The most distal enhancer, located 3.9 kb upstream of the likely transcription start site, (-3.9 kb enhancer; 449 base pairs) drives early expression (during gastrulation) in several anterior sublineages of ABa and in anterior sublineages of MS (Fig 5D; Fig S6). This expression occurs in nearly identical lineages to the early portion of the protein reporter expression pattern, but while the protein reporter expression in these lineages is transient, the enhancer reporter persists, again likely because of the use of a stable histone-mCherry reporter. The -3.9 kb enhancer expression is also well correlated with the pattern driven by the full promoter of *ref-2* (Spearman's $\rho > 0.6$, Fig 5K). Deleting this enhancer from the *ref-2* promoter reporter results in loss of the strong anterior-biased expression in the ABalaa, ABalap, and ABarpa lineages, and a loss in the weak anterior-biased expression in the ABalpa, ABalpp, ABaraa, and ABarap lineages. All remaining consistent expression occurs in the lineages derived from the Notch-signaled cells ABalp and ABara and has little anterior bias (Fig 5E, J; Fig S5; Fig S6) [46]. The significance of this expression is unclear; it can be seen in the promoter reporter but not the protein reporter, suggesting that it is normally repressed by additional sequences outside the promoter. We conclude that the -3.9 kb enhancer is necessary and sufficient for anterior-biased expression in the ABa lineage.

A second 745 bp enhancer, located at -2.8 kb, drives expression in later embryos (bean stage) in some anterior sublineages of MS and in the ABp-derived Pn ventral epidermal blast cells that were previously reported to require *ref-2* [38]. These cells also robustly express the *ref-2* protein reporter

and weakly express the *ref-2* promoter reporter (Fig 5B, C, F, J, K). Finally, the most proximal enhancer candidate (at -1.6 kb; 206 base pairs) drives variable, weakly anterior-biased, expression in multiple lineages, most consistently in ABala and ABara (Fig 5G, J, K; Fig S7) and also drives late expression in some ABp sublineages (Fig 5G; Fig S7). Much of this expression is in cells or at stages that do not express the full length reporters, suggesting that the activity of this sequence differs in its normal genomic context.

To measure anterior expression bias we developed an "anteriority score" based on the log2 ratio between anterior and posterior sister lineage expression. Because expression tends to accumulate over several cell cycles in each expressing lineage, we calculated this score based on the mean expression of the progeny 2-3 cell divisions after the birth of each anterior lineage (Fig 5I). Using this metric, we detected robust and significant anterior-biased expression for both the *ref-2* promoter and -3.9 kb enhancer reporters in ABa-derived sublineages, and reduced or no anterior-biased expression for the -2.8 kb and -1.6 kb enhancer reporters and for the promoter Δ -3.9 kb enhancer reporter in the same sublineages (Fig 5J). Comparing the expression driven by each enhancer to the full promoter showed each is positively, but imperfectly, correlated, consistent with each driving a subset of the full pattern (Fig 5K).

ref-2 **promoter expression requires the T-box transcription factors** ***tbx-37*** **and** ***tbx-38***

A dominant feature of the early *ref-2* expression pattern is reiterated expression in six of the eight anterior sister cells of the ectodermal ABa lineage at the 50 cell stage (when there are 16 ABa descendants). This expression is not observed in the fluorescent reporter strains until at least the following cell generation (when there are 32 ABa descendents) because of the time required for the fluorophores of mCherry or GFP to mature. This raises the question of whether *ref-2* expression requires the ABa-specific transcription factors *tbx-37* and *tbx-38*. These genes encode redundant paralogous T-box family transcription factors which are expressed throughout the ABa lineage and are required for multiple cell fate decisions within ABa [44,50,53]. Confirming previous reports [44,50], a TBX-38::GFP knock-in reporter made by CRISPR [50] is expressed throughout the ABa lineage and is detectable at least from the AB16 to AB128 stages (Fig S8). To test whether *tbx-37* and *tbx-38* are necessary for *ref-2* expression, we measured expression of the *ref-2* promoter

reporter in embryos carrying homozygous deletions of both *tbx-37* and *tbx-38*. We found that in the absence of *tbx-37* and *tbx-38,* nearly all *ref-2* promoter expression in the ABa lineage is lost (Fig 5H; Fig S9), while expression in the MS lineage was maintained. Thus, *tbx-37* and *tbx-38* are necessary for the anterior-biased expression of *ref-2* in ABa.

The -3.9 kb enhancer contains three non-overlapping sub-enhancers that independently drive anterior expression in the ABa lineage

Since the *ref-2* -3.9 kb enhancer drives expression in a similar pattern to the early expression of the full promoter and protein reporters, and its deletion from the *ref-2* promoter results in a reduction of anterior-biased expression in portions of the ABa lineage, it is a primary enhancer responsible for driving the early anterior-biased expression of *ref-2*. To map features necessary or sufficient for anterior expression, we tested the activity of eight truncated versions of the *ref-2* -3.9 kb enhancer to identify minimal regions of the enhancer sufficient to drive anterior expression in the ABa lineage (Fig 6, Fig S10). Instead of a single minimal region, we found three non-overlapping regions within the -3.9 kb enhancer that each are sufficient to drive anterior-biased enhancer activity in the ABa lineage: a promoter-distal region (bases 1 to 200), a medial region (bases 227 to 304) and a promoter-proximal region (bases 320 to 449) (Fig 6B, C). Of these regions, the proximal region drives the most consistent expression and this expression is most similar to the full length -3.9 kb enhancer construct (Fig 6D), whereas the medial and distal fragments drive more variable expression and in a subset of these lineages. Thus, although integration of information between these sub-enhancers is likely necessary for robustness of expression pattern, the 130 bp proximal fragment provides a model system for further dissection to identify anterior expression regulators.

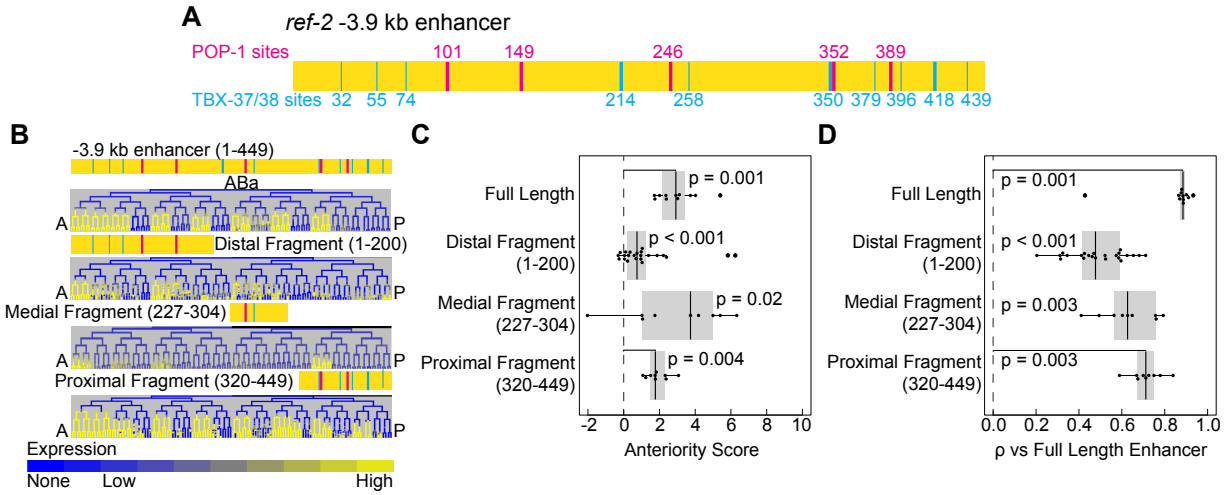


Fig 6. Three non-overlapping fragments of the -3.9 kb enhancer are each sufficient to drive anterior-biased expression in the ABa lineage of the early embryo.
(A) Model of the *ref-2* -3.9 kb enhancer with predicted TBX-37/38 sites (cyan, predicted high affinity sites indicated by thick line and predicted low-affinity by a thin line), and predicted high affinity POP-1 sites (magenta). (B) Expression patterns driven by full-length *ref-2* -3.9 kb enhancer and by minimal fragments. (C-D) Box plots displaying the anteriority scores (C) and Spearman's ρ (D) for the full-length *ref-2* -3.9 kb enhancer and minimal fragments. Lineages used to determine anteriority scores are the same as in Fig 5J. Spearman's ρ analysis uses the full ABa lineage and was calculated relative to the mean expression of the full-length -3.9 kb enhancer. A complete set of enhancer truncations tested is displayed in Fig S10.

TBX-37/38 binding sites are required for anterior expression of the proximal region

The proximal fragment exhibits broad sequence conservation with related nematodes, suggesting it contains multiple important TF binding sites (Fig 5A). To identify sequences necessary for anterior expression, we scrambled each 15 base-pair region along the length of the enhancer fragment and tested the resulting sequences for enhancer activity (Fig S11B, C, D). Scrambling a single region comprising base pairs 410-424 (coordinates relative to the full length enhancer) resulted in a complete loss of the anterior-biased expression in the ABa lineage, indicating that this region is necessary for this expression (Fig 7B; Fig S11B). Scrambling the most proximal regions 425-439 and 435-449 resulted in a loss of expression in subsets of the ABa lineage (in ABar and ABalp) but maintained anterior expression in the ABala lineage (Fig S11B). Expression was largely maintained when other regions were mutated, suggesting sequences in these regions are not individually necessary for anterior-specific expression despite their conservation.

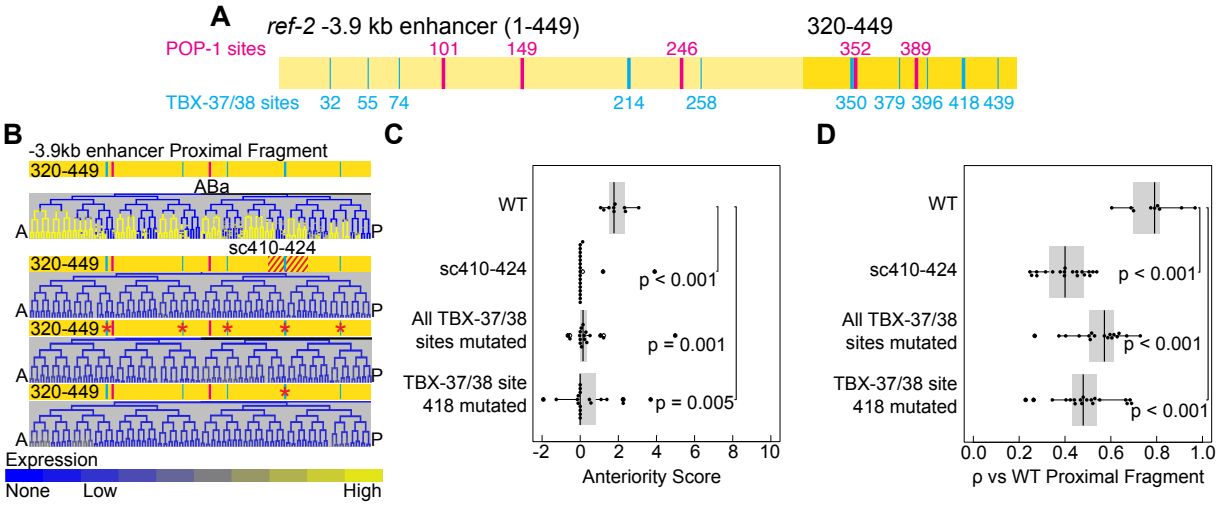


Fig 7. The ABa-expressed transcription factors *tbx-37* and *tbx-38* are required for the anterior-biased expression of *ref-2* in the ABa lineage.
(A) Model of the *ref-2* -3.9 kb enhancer with proximal fragment (base pairs 320-449) highlighted. Predicted TBX-37/38 and POP-1 sites are indicated as in Fig 6. (B) Expression patterns driven by the WT proximal fragment of the *ref-2* -3.9 kb enhancer, with base pairs 410-424 scrambled, with all predicted TBX-37/38 sites mutated, and TBX_{418} mutated. Expression driven by the proximal fragment with additional regions scrambled or with two predicted POP-1 sites mutated are displayed in Fig S11. (C-D) Anteriority scores (C) and Spearman's ρ (D) for the WT proximal fragment, with base pairs 410-424 scrambled, with all predicted TBX-37/38 sites mutated, and with TBX_{418} mutated. Lineages used to determine anteriority scores are the same as in Fig 5J. Spearman's ρ analysis uses the full ABa lineage and was calculated relative to the mean expression of the WT -3.9 kb enhancer proximal fragment.

In this fragment there are five predicted TBX-37/38 sites, two of which are predicted to have high affinity for TBX-37 and TBX-38. Two of the sites overlap with either the 410-424 region that is essential for ABa-specific expression or the 425-449 region that is required for a subset of ABa-specific expression (Fig 7A, B; Fig S11A, B). To determine whether this region is bound by TBX-37 and TBX-38 *in vivo,* we mined a recently published dataset that measured genome-wide TBX-37 and TBX-38 binding by ChIP-seq. Both factors show binding in the proximal region of the -3.9 kb enhancer (Fig 5A) [50].

To determine whether the TBX-37/38 sites are required for anterior expression, we measured the activity of the proximal fragment after mutating the two central nucleotides of all TBX-37/38 sites. The mutated enhancer fragment did not drive anterior-specific expression in the ABa lineage (Fig

7B, C, D; Fig S11B, C, D; Fig S12), confirming the importance of these sites. Additionally, the loss of anterior-biased expression following the mutation of these sites indicates that the two predicted high affinity POP-1 sites in this fragment are not sufficient to drive anterior-biased expression. Also, these sites are not necessary for the anterior-biased expression driven by this enhancer fragment, as the loss of these sites does not result in a loss of the anterior-biased expression (Fig S11B, C, D). In fact, concatemers of POP-1 sites are sufficient to drive posterior-biased expression (Fig S11F) [13,51].

Since the high affinity site beginning at position 418 (TBX_{418}) overlaps the region (410-424) required for expression, we hypothesized that this specific TBX-37/38 site may be required for anterior expression in the ABa lineage. To test this we tested the activity of the proximal fragment with only the central two nucleotides of TBX_{418} mutated (Fig 7B, Fig S11B, E). This mutation resulted in a loss of robust anterior-biased expression driven by the proximal fragment. Therefore, TBX-37 and TBX-38 bind to the *ref-2* -3.9 kb enhancer proximal fragment and TBX_{418} is necessary for it to drive robust anterior-biased expression (Fig 5A; Fig 7B, C, D; Fig S11B, C, D; Fig S12), suggesting that TBX-37 and TBX-38 directly regulate *ref-2* expression.

Concatemerized TBX sites are sufficient to drive anterior-specific expression in the ABa lineage

Given that *tbx-37* and *tbx-38* are expressed in both anterior and posterior cells within the ABa lineage, we hypothesized that site TBX_{418} would be sufficient to drive broad expression in ABa, with other sequences in the enhancer required for anterior specificity. To test this, we tested the activity of synthetic enhancers comprising concatemerized copies of TBX_{418} separated by either short (6-7 bp) or long (11-15 bp) flanking sequences from its endogenous context (Fig 8A). Surprisingly, both of these reporters drove anterior-specific expression in the ABa lineage similar to that driven by the full proximal fragment (Fig 8B, C, D). This indicates that either TBX_{418} or its flanking sequences are sufficient both for ABa lineage expression, and for anterior-specific expression within this lineage. To test the possibility that the flanking sequences are responsible for anterior-specificity, we scrambled those sequences while leaving the central TBX_{418} site intact in the context of the long-flanking-sequence construct. The resulting reporter was again expressed

only in anterior sublineages similar to the full proximal fragment (Fig 8A, B, C, D). Thus, when multimerized, *TBX*$_{418}$ is sufficient both for ABa expression and for anterior specificity without additional non-overlapping sequences.

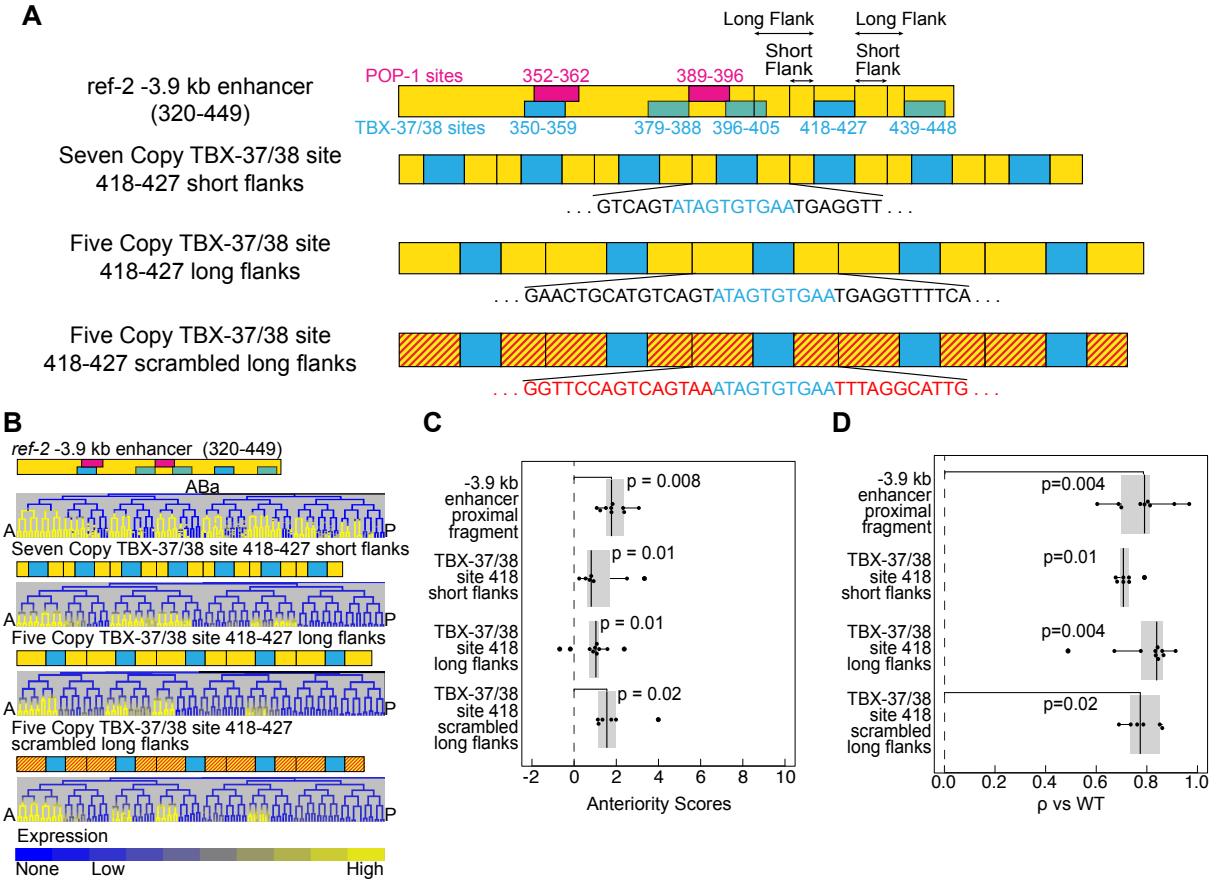


Fig 8. *TBX*$_{418}$ in the *ref-2* -3.9 kb enhancer is sufficient when concatemerized to drive anterior-biased expression in the ABa lineage.
(A) Models of the proximal fragment of the *ref-2* -3.9 kb enhancer, the *TBX*$_{418}$ seven-site concatemer with 6-7 base pair flanking regions, the *TBX*$_{418}$ five-site concatemer with 11-15 base pair flanking regions, and the *TBX*$_{418}$ five-site concatemer with 11-15 base pair scrambled flanking regions. (B) Expression patterns driven by the proximal fragment of the *ref-2* -3.9 kb enhancer and each TBX-37/38 site concatemer. (C-D) Anteriority scores (C) and Spearman's ρ (D) for the proximal fragment of the *ref-2* -3.9 kb enhancer and each TBX-37/38 binding site concatemer. Lineages used to determine anteriority scores are the same as in Fig 5J. Spearman's ρ analysis uses the full ABa lineage expression pattern and was calculated relative to the WT proximal fragment.

DISCUSSION

Previous work has definitively shown that the Wnt pathway components POP-1 and SYS-1 regulate posteriorly expressed genes in the *C. elegans* embryo via the canonical Wnt asymmetry pathway [12,13,35]. What has been less clear is how the similar number of anteriorly expressed genes are regulated. We have demonstrated that a set of anteriorly expressed genes, *tbx-35*, *ceh-51*, *elt-1*, *irx-1*, and *ref-2*, require POP-1 and SYS-1 for their anterior-biased expression. Therefore, POP-1 and SYS-1 regulate the expression of anterior-biased genes. The mechanism by which POP-1 and SYS-1 mediate this regulation remains to be determined. Previous work demonstrated that posterior genes can be regulated by POP-1 binding directly to POP-1 binding sites in their upstream regulatory regions [51]. Thus, a simple indirect model of POP-1 regulation of anterior genes would be POP-1 activating the expression of repressors in posterior cells. This mechanism would predict that anterior genes would be expressed later than posterior genes, or that anterior genes would be expressed early in both anterior and posterior cells, but maintained only in anterior cells. This can happen; for example, in the C lineage *elt-1* is expressed in both anterior and posterior sister cells at the 4C stage, but is only maintained in the anterior granddaughters of C (Caa and Cpa), perhaps due to its repression by the Wnt target *hlh-1* in the descendents of the posterior granddaughters of C (Cap and Cpp) [32]. What we observe in many cases, however, is that anterior gene expression is activated at about the same time as posterior gene expression in sister lineages. This observation suggests that POP-1 could regulate some anterior genes directly. Although some *pop-1* dependence of anterior genes could be explained by upstream anterior-to-posterior cell fate conversions following *pop-1* RNAi, our results as a whole cannot be explained by the known pattern of fate transformations [44].

Indirect *pop-1*-dependent regulation of anterior genes, however, does not explain all of its role in anterior gene expression regulation. For instance, previous work showed that loss of muscle-specifying TFs (MRFs) expressed in Cap and Cpp (*hlh-1*, *unc-120*, and *hnd-1*) is not sufficient to permit ectopic expression of *elt-1* in these cells or to convert them to an epidermal fate [32,54]. Therefore, another factor, perhaps *pop-1* acting directly, must be responsible for restricting *elt-1* to the anterior C granddaughters. Our results identified a requirement for pop-1 in anterior expression of elt-1 but not posterior repression in the C lineage. The proposed role of the MRFs in

repressing *elt-1* expression [32,54,55] suggests that this effect could partly be due to misexpression of MRFs in the anterior C granddaughters, which could be tested in the future.

Direct activation of genes when Wnt is absent by POP-1/TCF, termed "opposite" Wnt target regulation, occurs in other systems. For example, in *Drosophila* TCF bound directly to DNA at non-optimal sites can activate transcription when unbound by β-catenin, and repress transcription when bound to β-catenin. This difference in the regulatory activity of TCF is mediated by an allosteric conformational change in the structure of TCF [22]. In *C. elegans*, POP-1 bound together with REF-2 at the *ttx-3 cis*-regulatory region, may similarly undergo allosteric changes, leading to opposite target regulation [20,21]. Our findings are consistent with a similar mechanism regulating *ref-2* itself.

Our observation that five copies of a single 10 base-pair TBX-37/38 binding site are sufficient to drive anterior-biased expression in the ABa lineage suggests that a POP-1 site is likely unnecessary for anterior expression. However, *tbx-37* and *tbx-38* cannot explain this anterior bias alone, as these factors are expressed throughout the ABa lineage. We showed that *pop-1* is required for both anterior expression and posterior repression of the *ref-2* promoter. If this role of POP-1 is direct, it likely acts by binding to other factors such as TBX-37 and TBX-38, rather than to DNA directly. Consistent with this, TBX-38 and POP-1 were previously found to physically interact with each other [56]. Intriguingly, several other T-box factors are expressed early in the AB lineage [50,57]; future work should determine if these play a role in restricting the activity of these sites to anterior lineages. For instance, TBX_{418} is predicted to be bound by several other *C. elegans* T-box factors. Other T-box factors expressed in the early ABa lineage could compete for binding at this site, with POP-1 altering the relative activity of these factors between anterior and posterior cells. Additionally, it is unclear whether *tbx-37* and *tbx-38* also regulate posterior gene expression in the ABa lineage; intriguingly, the expression of *pha-4*, another anterior gene, in the ABa lineage is also dependent on *tbx-37* and *tbx-38* [44].

In contrast with our observation that TBX_{418} concatemers are sufficient to drive anterior-biased expression, Murgan et al. 2015 found that a REF-2 binding site concatemer is not sufficient to drive expression in the anterior AIY mother cell. Instead, Murgan et al. identified binding sites for

helix-loop-helix family transcription factors that are sufficient to drive anterior expression in the AIY mother and its posterior sister cell, and that combining these with REF-2 binding sites restricts expression to the AIY mother. Thus, REF-2 acting through ZIC binding sites is not sufficient to activate expression without cooperating factors binding other sites whereas TBX-37 and TBX-38 may be sufficient to do so. Other factors could also play a role, for example we used a pes-10 minimal promoter, whereas Murgan et al. did not include a minimal promoter in their constructs, and there could be differences in sensitivity between the microscopy techniques.

Classic work in many species, including flies and vertebrates, has found that developmental genes are often regulated both by partially redundant enhancers and modular enhancers that are responsible for regulating distinct portions of the genes' expression patterns [58,59]. The extent to which *C. elegans* genes are regulated by such distal enhancers vs promoter proximal elements has been unclear. We found that the predicted enhancers in the promoter region of *ref-2* function modularly, such that each enhancer drives different expression patterns, presumably providing multiple inputs to fine-tune the expression pattern of *ref-2*. Combined with other recent studies this adds to evidence that enhancer-mediated regulation is widespread during *C. elegans* embryonic development [52,60–64]. Additionally, we observe modularity within the -3.9 kb enhancer, with three non-overlapping regions of the enhancer each sufficient to drive anterior-biased expression. Since the protein reporter is expressed in more lineages than the promoter reporter, there are likely even more enhancers outside the promoter region that regulate expression of *ref-2*.

Several studies have identified evidence for multiple functions of *ref-2*. *ref-2* was first identified as being required for the production of the Pn.p ventral epidermal cells and for the inhibition of their fusion to the epidermal syncytium hyp7 [38]. *ref-2* is also required for the initiation of the differentiation of the cholinergic neuron AIY [65]. Also, *ref-2* has been found to be required for female fate during sexual development [66]. We have demonstrated a role for *ref-2* in embryonic development. *ref-2* is required for robust WT cell positioning and cell division timing in anterior lineages. Intriguingly, an insect ortholog of *ref-2*, *odd-paired*, acts as a regulator of anterior-posterior expression and is an anterior determinant in several species [67,68], suggesting the function of *ref-2* in anterior fate regulation may be ancestral. Because the defects in cell position

and division timing are only partially penetrant, there are likely other, partially redundant regulators of cell position and cell division timing that act with the sole *C. elegans* ZIC homolog *ref-2* during embryonic development. Further experiments, including suppressor and enhancer screens and co-mutation of other transcription factors will need to be performed to identify other factors that regulate cell position and division timing in conjunction with *ref-2*.

METHODS

C. elegans culture and strain generation

C. elegans strains (Table S1) were maintained at standard growth temperatures on OP50 *E. coli* on NGM plates (Table S2). RNAi knockdown was performed by feeding, as previously described [69,70]. We validated the efficacy of *pop-1* RNAi by measuring the cell cycle delay resulting from the transformation of the anterior MS lineage into an E-like lineage, and we validated the efficacy of *sys-1* RNAi by failure of morphogenesis and embryonic death (Table S2) [13,33]. Enhancer reporter strains were generated by microinjection into RW10029, the GFP histone strain used for lineage tracing. Injection cocktails consisted of reporter DNA construct at 10 ng/µL, with 5 ng/µL *myo-2*p::GFP, and 135 ng/µL pBluescript vector (or highest concentration possible whenever 135 ng/µL was impossible to make) and were injected using a Narishige MN-151 micromanipulator with Tritech microinjector system. Other strains were created through crosses using standard approaches (Table S1).

Generation of transgenes

Candidate enhancers were amplified by PCR from *C. elegans* N2 strain genomic DNA with Phusion HF polymerase (New England Biosciences) and gel purified (Qiagen). Enhancer fragments, mutated enhancers, and binding site concatemers were ordered as either gBlocks or Ultramers from Integrated DNA Technologies (Coralville, Iowa) and were amplified with Phusion HF polymerase (New England Biosciences) with overhangs for stitching and either gel or PCR purified (Qiagen). Enhancer reporters were produced by fusing via PCR stitching these constructs to a *pes-10* minimal promoter::HIS-24::mCherry::let-858 3'UTR fragment amplified from the

POPTOP plasmid [51] (Addgene #34848). The enhancer reporters were purified with a PureLink PCR purification kit (ThermoFisher) and/or gel purified. The desired product was determined by its size based on gel electrophoresis. We identified putative transcription factor binding sites using CIS-BP (http://cisbp.ccbr.utoronto.ca) [71,72].To mutate the TBX-37/38 sites, the central two nucleotides were altered (eg. ATAG**TG**TGAA changed to ATAG**GT**TGAA for *TBX*$_{418}$). For the construct with all TBX-37/38 sites mutated, after mutating the five primary sites, subsequent analysis revealed two weaker sites. All of these predicted TBX-37/38 sites were mutated. The reporter construct for the *ref-2* promoter lacking the -3.9 kb enhancer was amplified from N2 genomic DNA by PCR and PCR stitched to the *pes-10* minimal promoter::HIS-24::mCherry reporter, and, thus, drives expression through the *pes-10* minimal promoter. The *pes-10* promoter drives no consistent embryonic expression on its own prior to elongation [13,51,52] and is compatible with a wide variety of enhancers [13,51] (Table S5).

Quantitative comparisons

All quantitative comparisons were performed using R version 4.0.3 (The R Foundation for Statistical Computing), Microsoft Excel version 16.16.27, or Python version 2.7.16. For the main figures, a minimum N of 6 was used for all experiments, except for ELT-1::GFP expression after *sys-1* RNAi, for which N = 3. For the supplemental figures, a minimum N of 2 was used (largely for perturbations which appeared similar to wild-type at this level). Code and raw data for these analyses are available at https://github.com/jisaacmurray/ref-2_paper.

Expression and phenotypic analysis by 4D imaging

We obtained confocal micrographs using a Leica TCS SP5 or Stellaris scanning confocal microscope (67 z planes at 0.5 µm spacing and 1.5 minute time spacing, with laser power increasing by 4-fold through the embryo depth to account for attenuation of signal with depth). Embryos obtained from self-fertilized hermaphrodites were mounted in egg buffer/methyl cellulose with 20µm beads used as spacers and imaged at 22°C using a stage temperature controller (Brook Industries, Lake Villa, IL) [73]. We used StarryNite software to automatically annotate nuclei and trace lineages [29]. We corrected errors from the automated analysis and quantified

694 reporter expression in each nucleus relative to the local background (using the "blot" background
695 correction technique) with AceTree software as previously described (Table S3) [24,26].
696
697 *Anterior gene WT to RNAi within lineage comparisons*
698
699 To compare the expression levels of the anterior genes *tbx-35*, *ceh-51*, ELT-1, *irx-1*, and *ref-2*
700 between WT and RNAi embryos within the expressing lineages and their posterior sisters, we
701 calculated for each embryo the mean reporter expression across all measurements within each
702 lineage (all descendant cells and time points) starting with the cell stage that expresses the reporter
703 at the time cells are born or shortly afterwards to the last time point in the indicated lineage. The
704 expression analyzed for the MS lineage starts at an earlier stage than that for other lineages, since
705 following *pop-1* RNAi the cell divisions are so dramatically delayed that reporter expression
706 comes on in earlier cell generations in these embryos. The cell generations used in each of these
707 comparisons are noted in Table S4.
708
709 *Anteriority Scores*
710
711 To measure the magnitude of anterior expression bias, we developed a measure that we refer to as
712 the anteriority score. To determine the anteriority scores, first we chose a minimum expression
713 cutoff and set all values below this cutoff to 0 and subtracted the minimum cutoff from all
714 expression values greater than or equal to the cutoff (an expression value of 200 was used as the
715 cutoff for all anteriority score analyses, except a value of zero was used for the TBX-38::GFP
716 reporter, which has very low fluorescence intensity compared to the other reporters analyzed).
717 Next, we averaged the expression levels of the reporter of interest in the cells of each anterior and
718 each posterior sister lineage of interest. For the *ref-2* promoter and enhancer constructs, as well as
719 the TBX-37/38 site concatemers, we analyzed anterior and posterior lineages descended from the
720 cells ABala, ABalaa, ABalap, ABalpa ABalpp ABaraa, ABarap, ABarpa, and ABarpp. We used
721 expression values starting at the generation with ABaxxxxxx cells (AB128 stage) until the time
722 point of the last division of AB128 stage cells in the ABa lineage. Next, we calculated the means
723 of the average expression levels of all of the anterior lineages and all of the posterior lineages
724 analyzed for each embryo. We then added a pseudocount of 1 to each of the anterior and posterior

725 embryo means. For each experimental group, the adjusted average anterior expression is divided
726 by the adjusted average posterior expression, and the base 2 logarithm of the result is reported--
727 Anteriority Score = $\log_2$((anterior mean + 1)/(posterior mean +1)).
728
729 *Correlation analysis*
730
731 Correlation analysis was done by Spearman's ρ, comparing each embryo's expression to the
732 average expression of the control group. Spearman's ρ was calculated in R using the "cor"
733 function. For Spearman's ρ we used all the data from the lineages of interest without excluding
734 any cell generations or having a minimum value cutoff. We included cells born before the onset
735 of expression here to include information about the cell generation of expression onset in the
736 analysis. For the analysis comparing the *ref-2* promoter with its enhancers we used the full somatic
737 lineage starting with the blastomeres AB, EMS, C, and D. For all other analyses we used only the
738 ABa lineage.
739
740 *Mutant cell position and cell division analysis*
741
742 We identified cell position and cell cycle defects as previously described [39]. We corrected for
743 differences in global division rates, and considered divisions as defective if they deviated from the
744 wild-type cell cycle length by at least five minutes and had a z-score greater than three. Cell
745 positions were corrected for differences in embryo size and rotation, and considered defective if
746 they deviated from the expected wild type position by at least five microns, had a z-score greater
747 than five, and a nearest neighbor score greater than 0.8.
748
749 *Statistical analyses*
750
751 Statistical significance of differences in anterior or posterior expression between untreated and
752 RNAi-treated embryos was tested using the two-tailed Wilcoxon Ranked Sum test. Statistical
753 significance of differences in Spearman's ρ or Anteriority Score were tested using the one-tailed
754 Wilcoxon Ranked Sum test to test for differences from either the control group or from 0, as
755 indicated in each figure. The p-values calculated by the Wilcoxon Ranked Sum test were corrected

for multiple comparisons using the false discovery rate method in R. Significance of overlap between positional or cell division defects and expression status were assessed using a Chi-squared test.

FIGURES

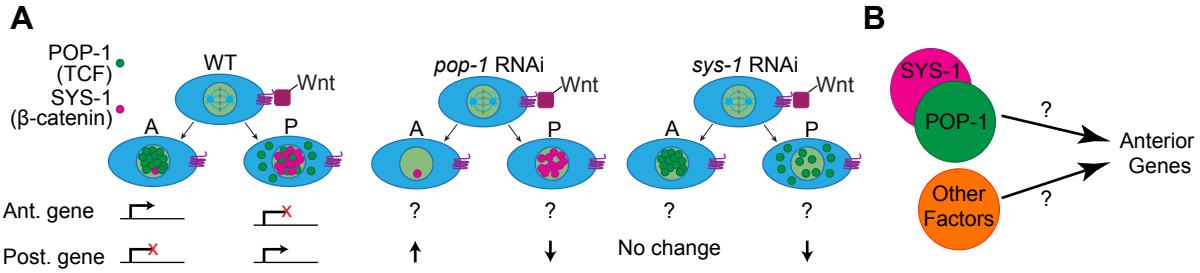


Fig 1. How are anterior genes regulated?

(A) The Wnt/β-catenin asymmetry pathway regulates anterior-posterior patterning by producing asymmetric nuclear concentrations of POP-1/TCF and SYS-1/β-catenin in anterior and posterior sister cells. Low nuclear concentrations of POP-1 and high nuclear concentrations of SYS-1 in the posterior cell activate expression of posterior genes and are associated with the lack of expression of anterior genes. Conversely, high nuclear concentrations of POP-1 and low nuclear concentrations of SYS-1 in the anterior cell repress posterior genes and are associated with expression of anterior genes. Effects of *pop-1* and *sys-1* RNAi on posterior genes are shown; effects on most anterior genes are unknown. (B) In this work, we test whether *pop-1*/TCF and *sys-1*/β-catenin regulate anterior genes in the *C. elegans* embryo, and ask what other factors also regulate anterior expression.

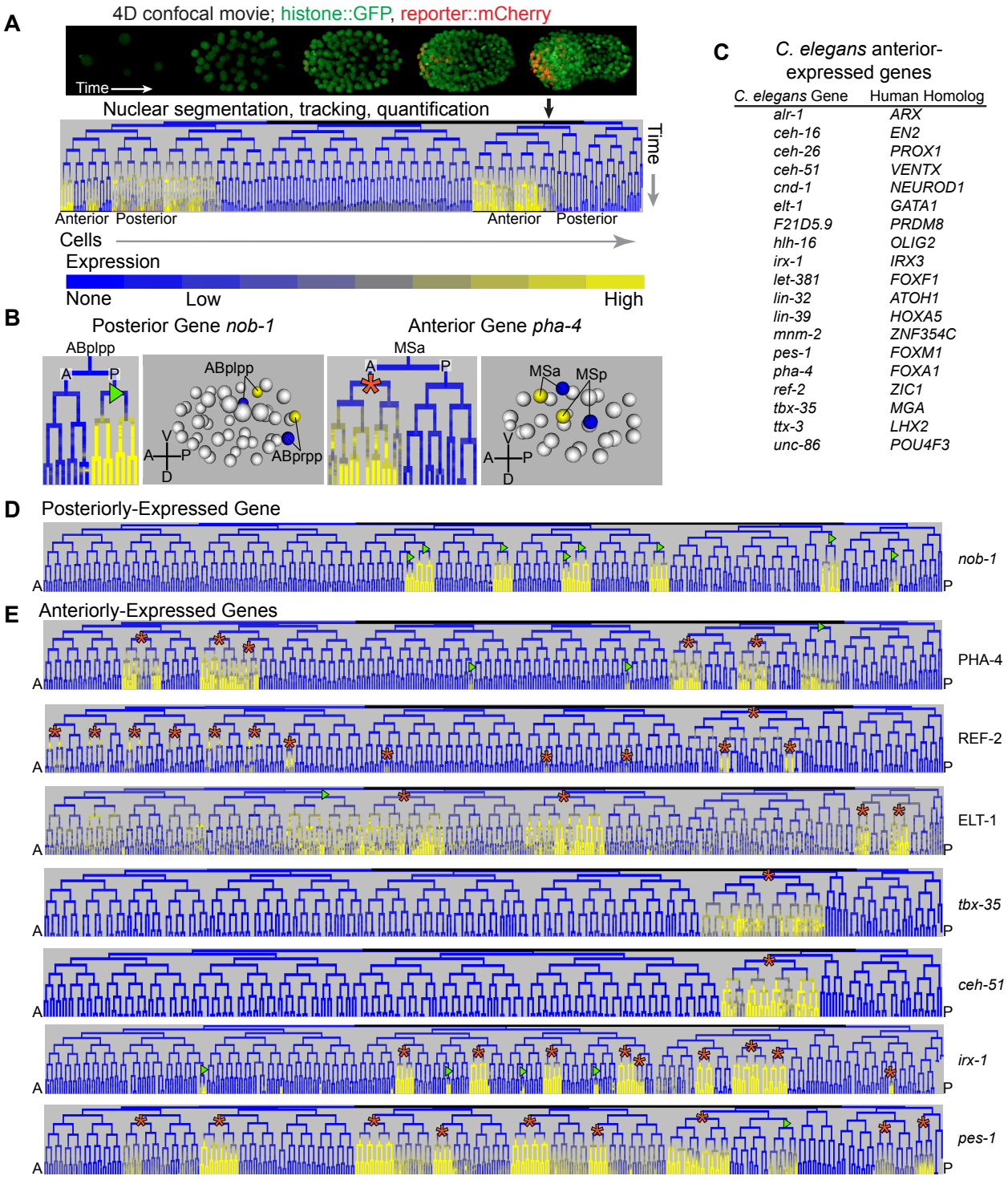


Fig 2. Identification of genes with anterior-biased expression.
(A) Automated lineaging traces lineages and quantifies expression from 4D confocal movies. In the lineage trees, vertical lines represent cells, and horizontal lines represent cell divisions. Most cells divide along the anterior-posterior axis, with anterior cells depicted on the left branch and posterior cells on the right branch. (B) Posterior and anterior genes are denoted as such based on their expression in cells descended from either a posterior or anterior sister cell, respectively,

following an anterior-posterior cell division. Anterior and posterior genes can generally be expressed in cells from any part of the embryo. Anterior and posterior founder cells are labeled with orange asterisks/green triangles respectively. (C) List of anterior-biased genes in the EPiC database along with their predicted human homologs. (D) Expression pattern of a posterior gene, *nob-1/Hox9-13*. (E) Expression patterns of several anterior genes; most are expressed in unique combinations of mostly anterior lineages.

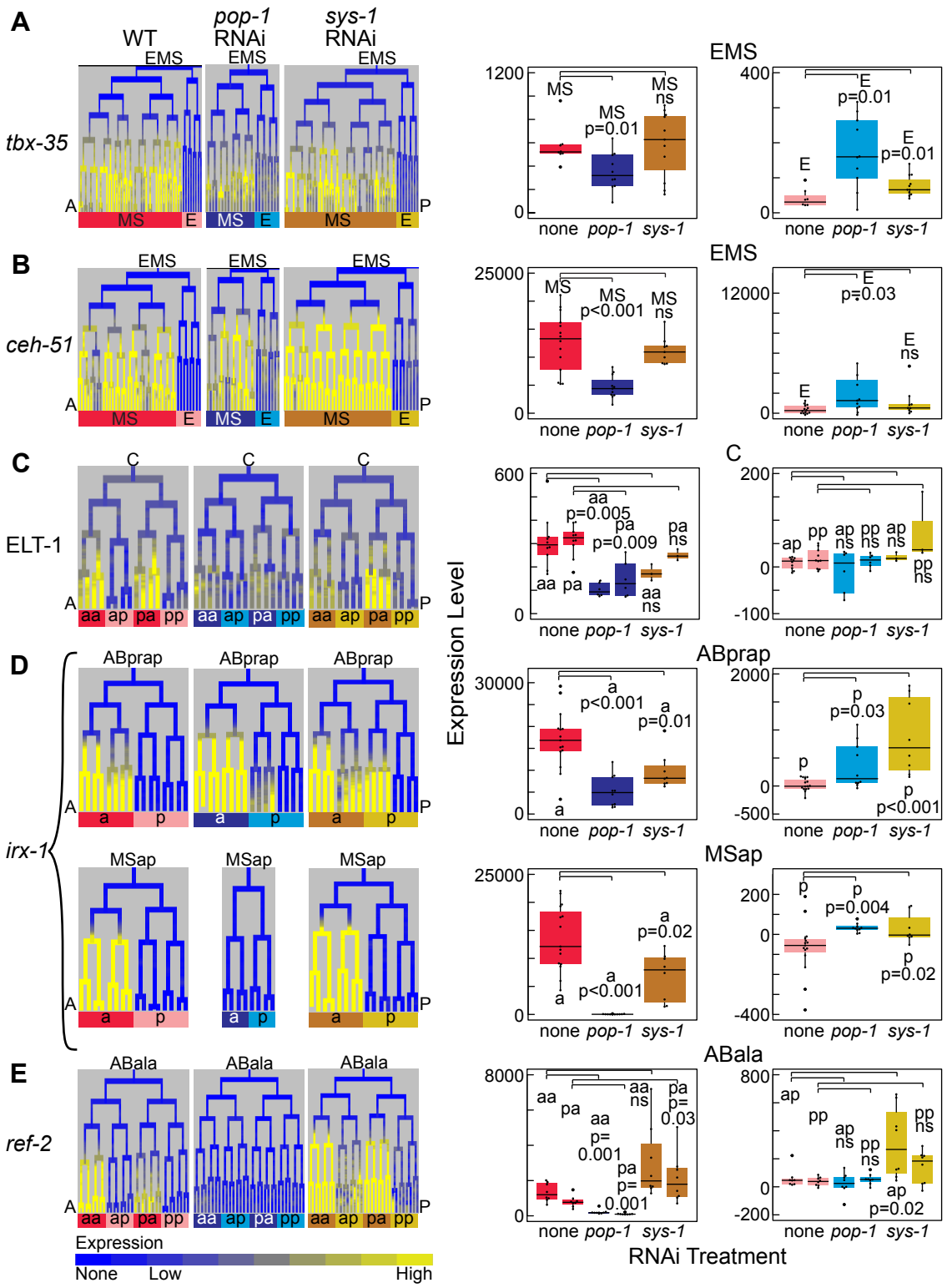


Fig 3. Regulation of anterior genes by Wnt effectors *pop-1* and *sys-1*.

(A-E) Expression pattern (left) and quantification (right) of reporters for the anterior genes *tbx-35* (A), *ceh-51* (B), ELT-1 (C), *irx-1* (D), and *ref-2* (E) in specific anterior lineages and their posterior sisters. Expression is shown under WT conditions and following *pop-1* or *sys-1* RNAi. Expression

quantification (right) is the mean expression across all measurements within that lineage from when the reporter is normally first detected until the last time point indicated in the lineage (details in Methods).

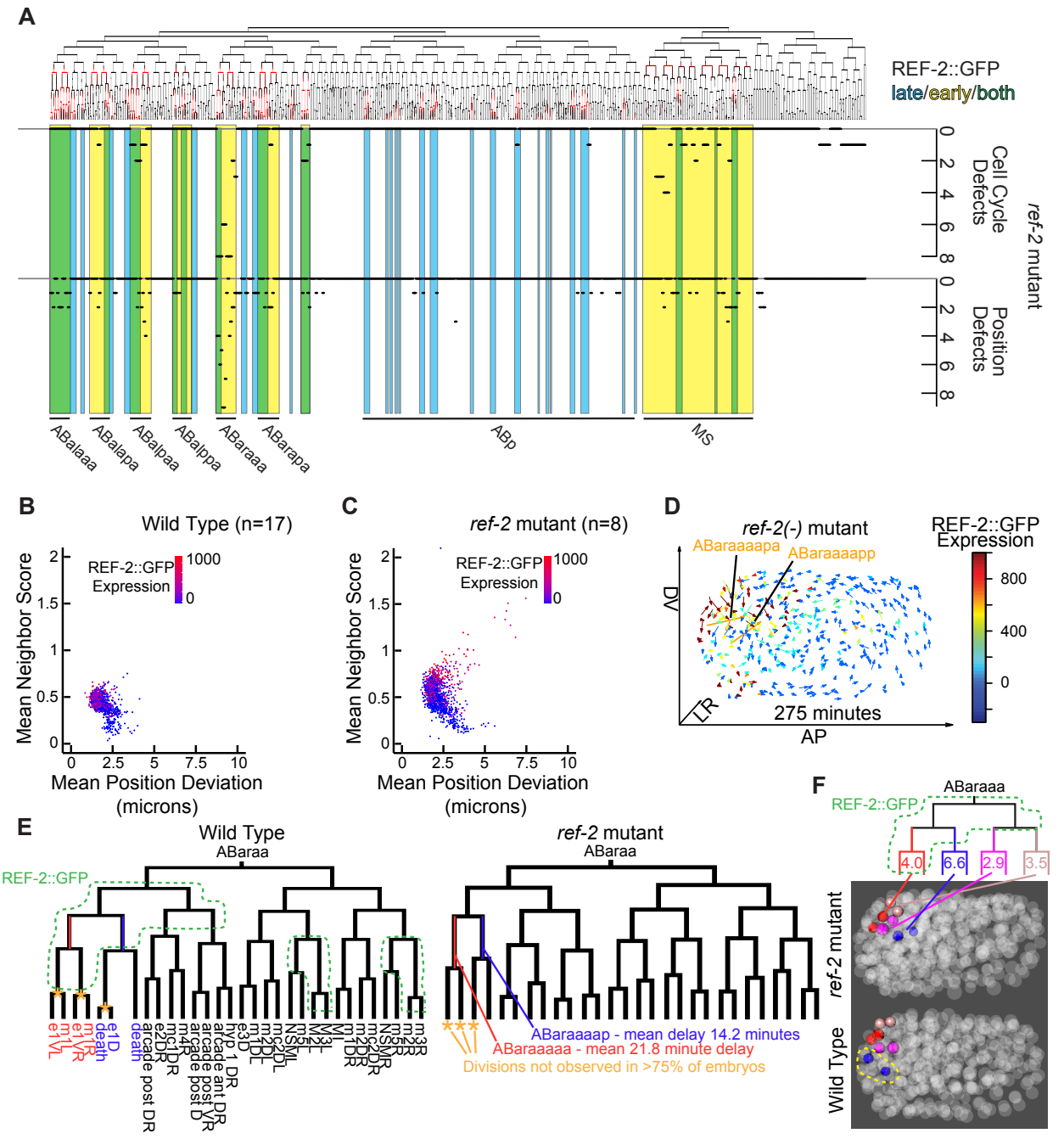


Fig 4. *ref-2* is required for proper division timing and positioning for many embryonic cells that are descended from cells that express *ref-2*.

(A) Lineage tree depicting the REF-2 protein expression pattern (top), and chart indicating the number of embryos out of eight homozygous *ref-2* mutant embryos that exhibit cell cycle defects and position defects in indicated cell lineages. Yellow indicates lineages that express *ref-2* early, blue indicates lineages that express *ref-2* late, and the overlap (green) indicates both. A corresponding analysis of wild-type embryos gives 0-1 defects per cell in all of 17 embryos tested [39]. (B-C) Scatter plot of mean neighbor score vs mean position deviation of cells of 17 wild-type (B) and eight *ref-2* mutant (C) embryos. Neighbor score is the ratio of each cell's distances to its

ten closest wild-type neighbors between mutant and wild-type embryos. Points representing cells are colored based on WT REF-2::GFP expression levels [39]. (D) Plot of cell position deviations with arrows starting at the average WT position of cells and pointing to the average position in *ref-2* mutants. Arrows are colored by the WT expression levels of REF-2::GFP for each cell. The labeled cells (ABaraaaapa and ABaraaaapp) have the greatest mean cell position deviations. (E) ABaraa lineages of WT and *ref-2* mutant embryos, with cells expressing REF-2::GFP outlined on the WT tree. Two delayed and three missed cell divisions are highlighted on the *ref-2* mutant tree. (F) ABaraaa lineage with cells expressing REF-2::GFP outlined. The average position deviations in microns of terminal sister cells are indicated on the terminal branches of the lineage tree. Three dimensional projections of a WT and a *ref-2* mutant embryo are shown with the positions of the terminal ABaraaa lineage cells highlighted. ABaraaaapa and ABaraaaapp are outlined in the WT embryo projection as the cells with the greatest mean position deviation in the *ref-2* mutant.

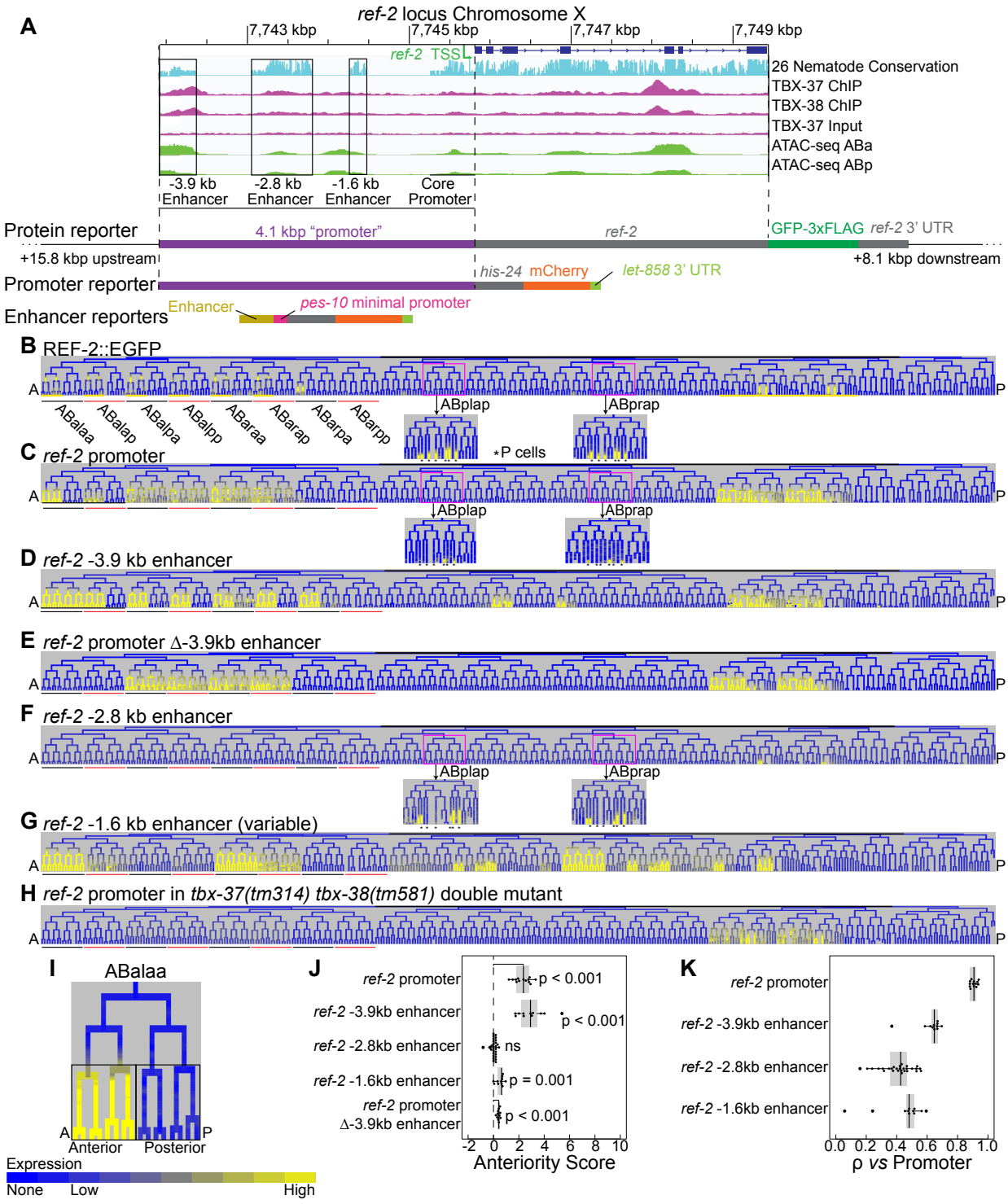


Fig 5. Identification of *ref-2* enhancers.

(A) Genome browser screenshot showing the *ref-2* locus on the X chromosome. Displayed tracks include the gene model, sequence conservation among 26 *Caenorhabditis* species [49], ChIP-seq data for TBX-37 and TBX-38 binding, and ATAC-seq data for cells in the ABa and ABp lineages [50]. Boxes indicate candidate enhancer regions as determined by conservation. Below the genome browser screenshot are models of the fluorescent reporter expression constructs we examined. (B-

H) Lineage trees showing the expression patterns of *ref-2* protein (B), promoter (C), -3.9 kb enhancer reporter (D), promoter lacking the -3.9 kb enhancer (E), -2.8 kb enhancer reporter (F), -1.6 kb enhancer reporter (G), and promoter in the *tbx-37 tbx-38* double mutant (H) as determined by time-lapse confocal microscopy. Sublineages used in anteriority analyses are labeled in B and underlined in black (anterior) and red (posterior) under each lineage tree. Sublineages with transient early REF-2::EGFP (protein reporter) expression are underlined in yellow under the lineage tree in B. Pn ventral epidermal cells are indicated with asterisks in insets of panels B, C, and F. (I) Example lineage showing which cells were included in anteriority score analyses. Reporter expression levels were averaged across cells in the anterior or posterior lineages for the indicated cell generations. (J) Anteriority scores of the *ref-2* promoter, enhancers, and promoter lacking the -3.9 kb enhancer. Lineages used for analysis are ABala, ABalaa, ABalap, ABalpa, ABalpp, ABaraa, ABarap, ABarpa, and ABarpp (see Fig 5B). (K) Full somatic lineage correlation (Spearman's ρ) of each reporter's expression pattern with the mean expression pattern across embryos expressing the promoter reporter.

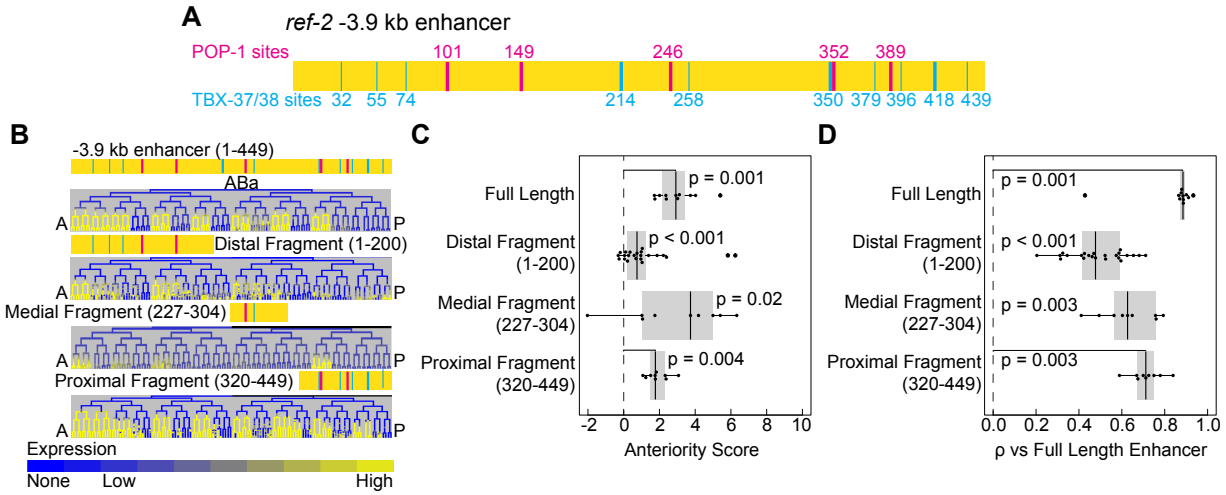


Fig 6. Three non-overlapping fragments of the -3.9 kb enhancer are each sufficient to drive anterior-biased expression in the ABa lineage of the early embryo.

(A) Model of the *ref-2* -3.9 kb enhancer with predicted TBX-37/38 sites (cyan, predicted high affinity sites indicated by thick line and predicted low-affinity by a thin line), and predicted high affinity POP-1 sites (magenta). (B) Expression patterns driven by full-length *ref-2* -3.9 kb enhancer and by minimal fragments. (C-D) Box plots displaying the anteriority scores (C) and Spearman's ρ (D) for the full-length *ref-2* -3.9 kb enhancer and minimal fragments. Lineages used to determine anteriority scores are the same as in Fig 5J. Spearman's ρ analysis uses the full ABa lineage and was calculated relative to the mean expression of the full-length -3.9 kb enhancer. A complete set of enhancer truncations tested is displayed in Fig S10.

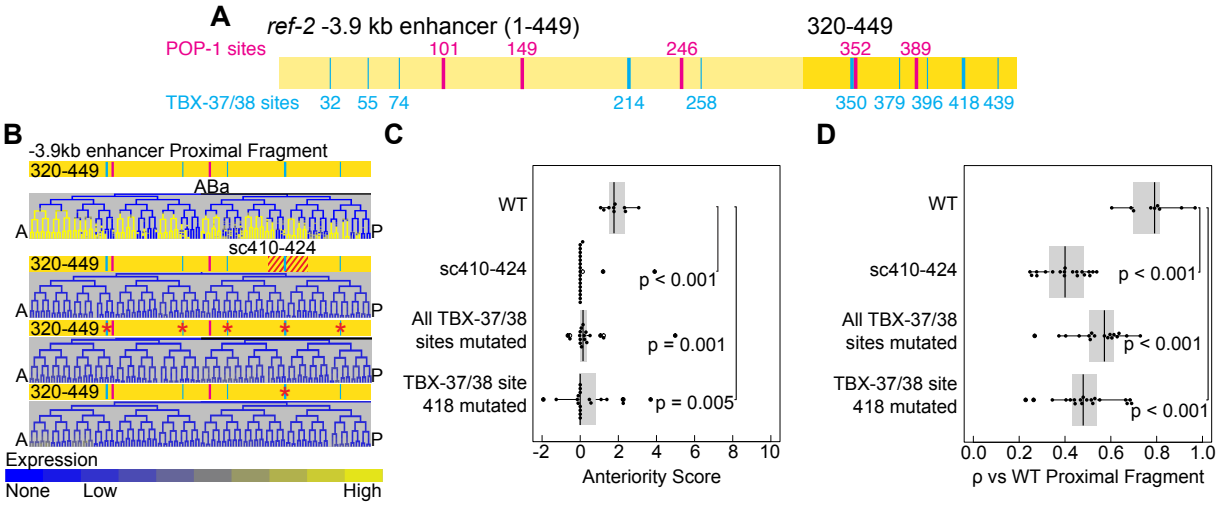


Fig 7. The ABa-expressed transcription factors *tbx-37* and *tbx-38* are required for the anterior-biased expression of *ref-2* in the ABa lineage.

(A) Model of the *ref-2* -3.9 kb enhancer with proximal fragment (base pairs 320-449) highlighted. Predicted TBX-37/38 and POP-1 sites are indicated as in Fig 6. (B) Expression patterns driven by the WT proximal fragment of the *ref-2* -3.9 kb enhancer, with base pairs 410-424 scrambled, with all predicted TBX-37/38 sites mutated, and TBX_{418} mutated. Expression driven by the proximal fragment with additional regions scrambled or with two predicted POP-1 sites mutated are displayed in Fig S11. (C-D) Anteriority scores (C) and Spearman's ρ (D) for the WT proximal fragment, with base pairs 410-424 scrambled, with all predicted TBX-37/38 sites mutated, and with TBX_{418} mutated. Lineages used to determine anteriority scores are the same as in Fig 5J. Spearman's ρ analysis uses the full ABa lineage and was calculated relative to the mean expression of the WT -3.9 kb enhancer proximal fragment.

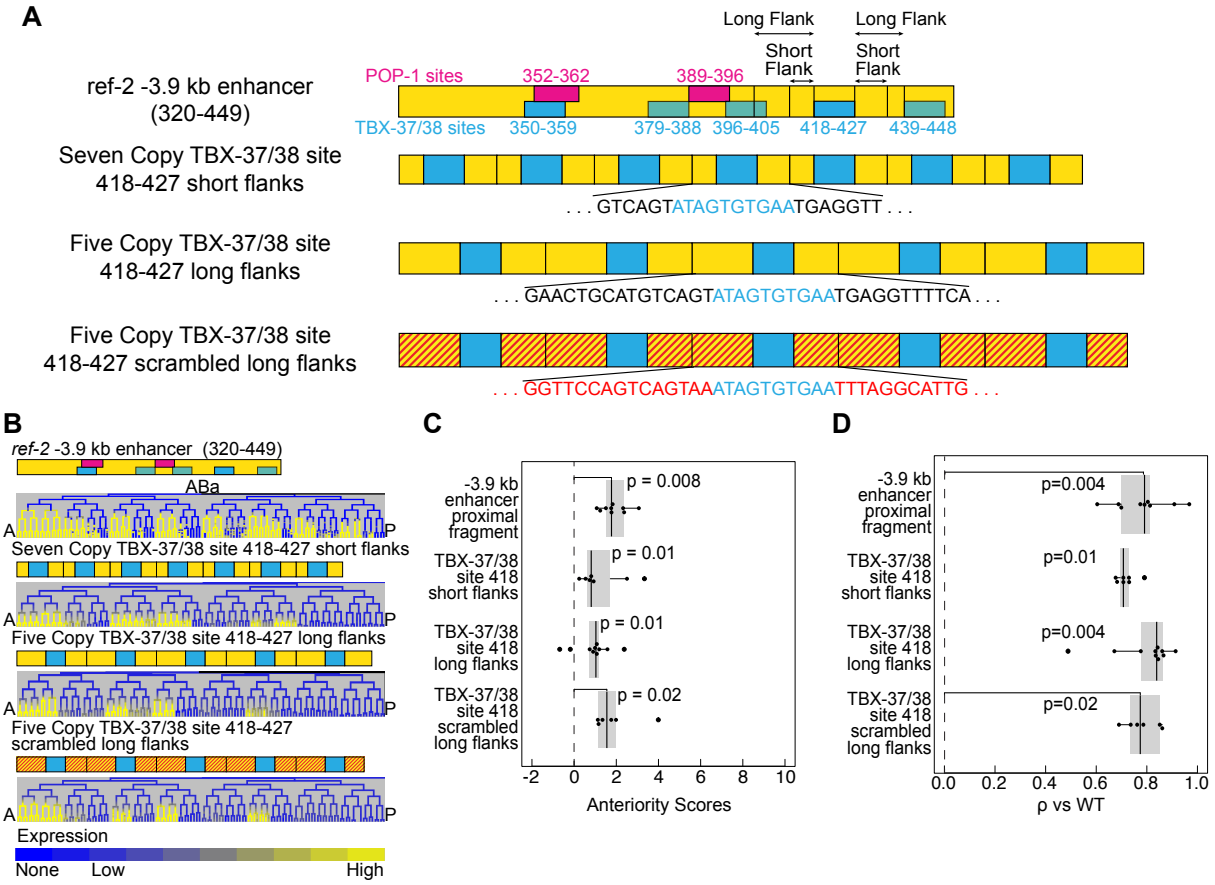
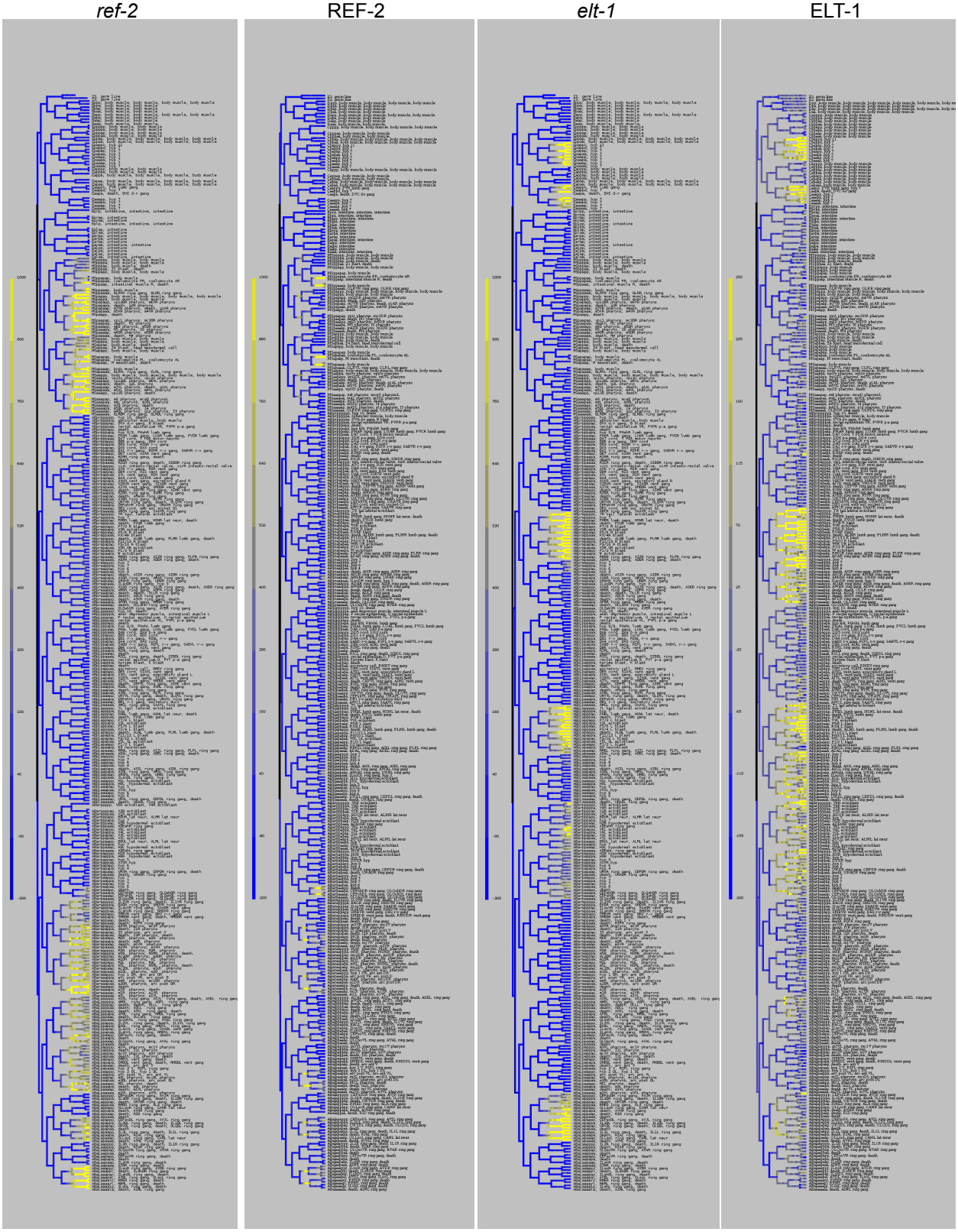


Fig 8. *TBX$_{418}$* in the *ref-2* -3.9 kb enhancer is sufficient when concatemerized to drive anterior-biased expression in the ABa lineage.

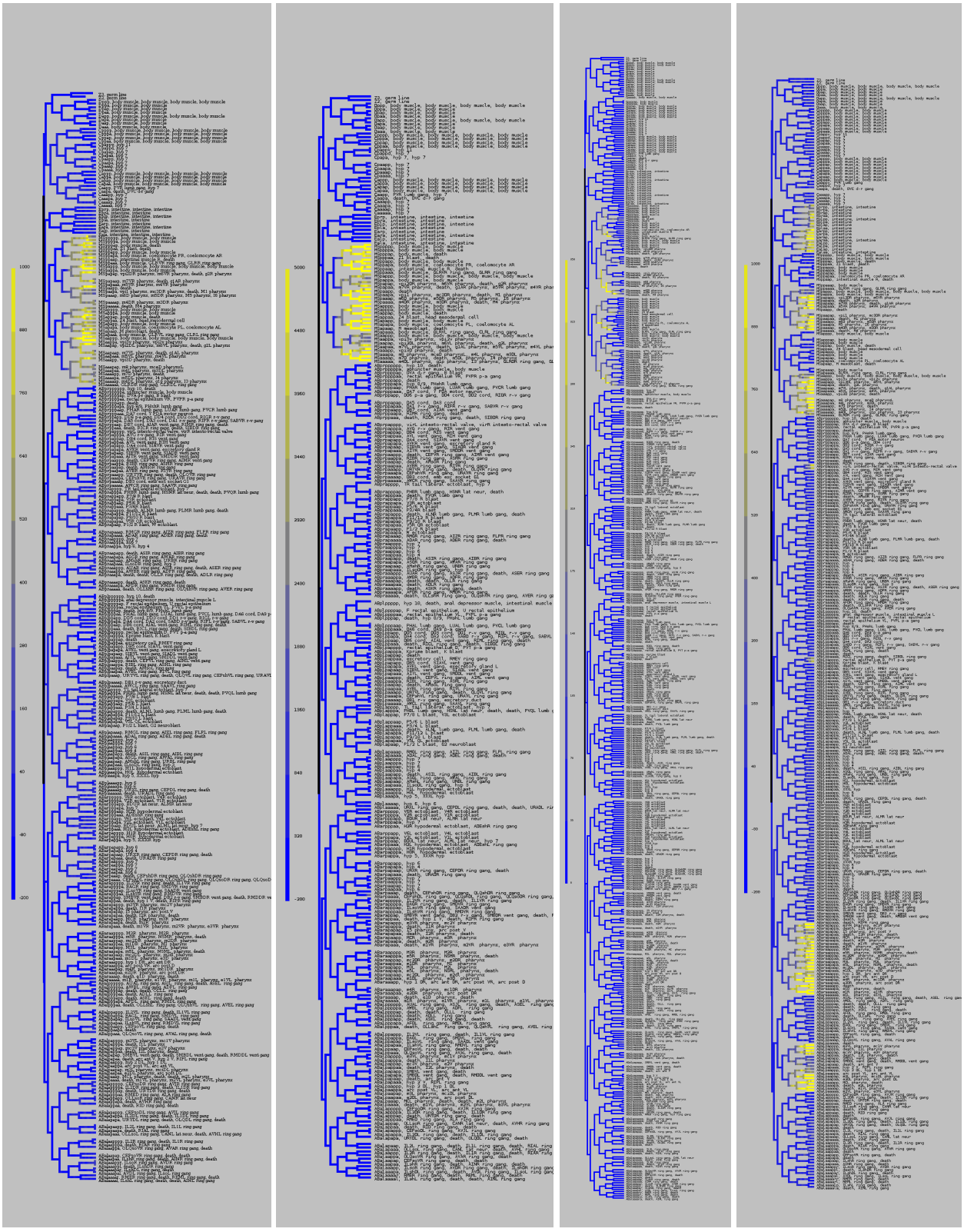
(A) Models of the proximal fragment of the *ref-2* -3.9 kb enhancer, the *TBX$_{418}$* seven-site concatemer with 6-7 base pair flanking regions, the *TBX$_{418}$* five-site concatemer with 11-15 base pair flanking regions, and the *TBX$_{418}$* five-site concatemer with 11-15 base pair scrambled flanking regions. (B) Expression patterns driven by the proximal fragment of the *ref-2* -3.9 kb enhancer and each TBX-37/38 site concatemer. (C-D) Anteriority scores (C) and Spearman's ρ (D) for the proximal fragment of the *ref-2* -3.9 kb enhancer and each TBX-37/38 binding site concatemer. Lineages used to determine anteriority scores are the same as in Fig 5J. Spearman's ρ analysis uses the full ABa lineage expression pattern and was calculated relative to the WT proximal fragment.

ref-2 REF-2 *elt-1* ELT-1



898 **Fig S1. Several early embryonic genes are expressed with anterior bias.**
899 Full lineage expression patterns are shown for integrated transgenic promoter reporters of *ref-2*
900 and *elt-1*; and for integrated transgenic protein reporters of REF-2 and ELT-1.

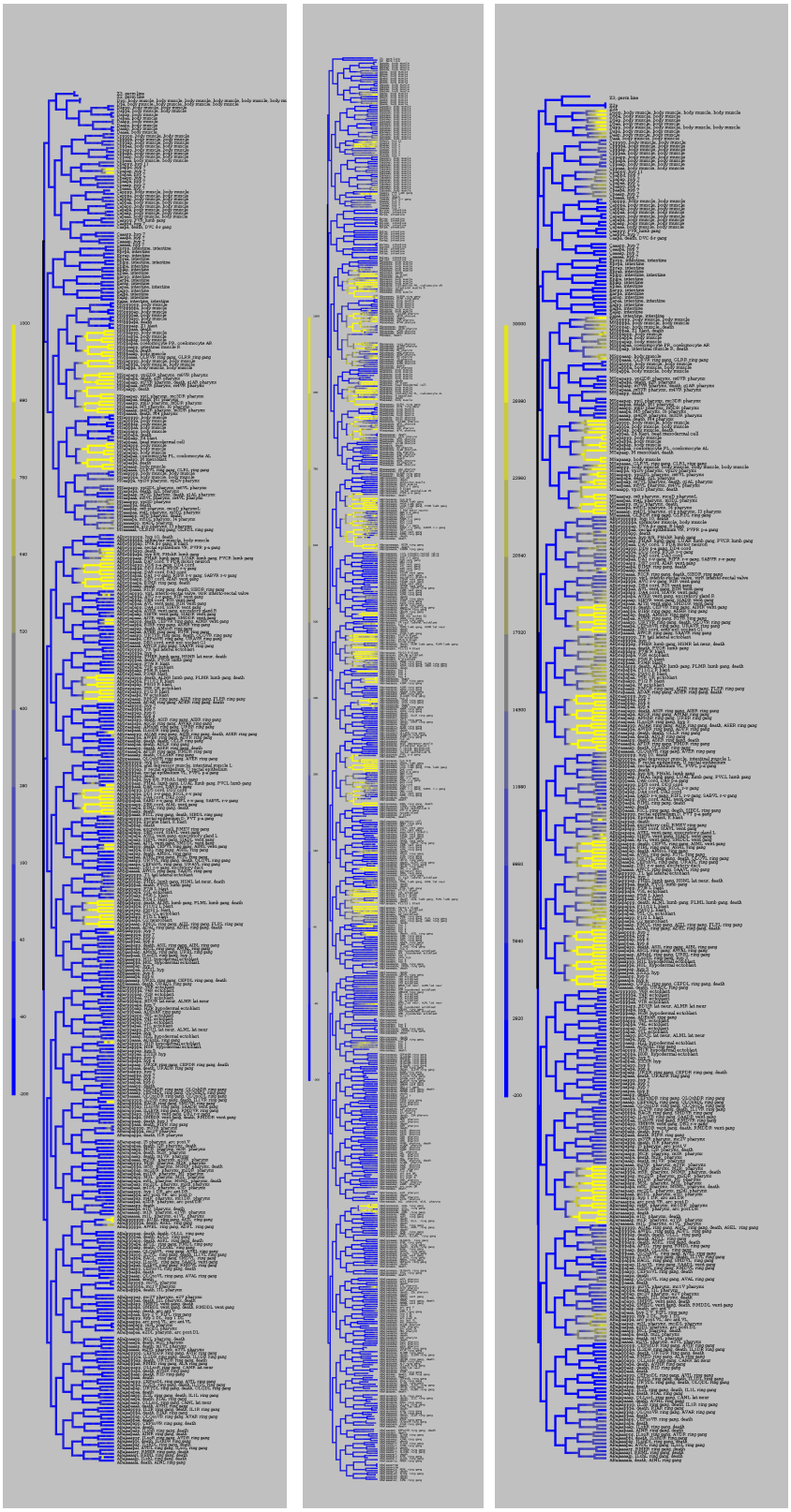
tbx-35 *ceh-51* CEH-51 PHA-4



901
902
903

Fig S2. Several early embryonic genes are expressed with anterior bias.
Full lineage expression patterns are shown for integrated transgenic promoter reporters of *tbx-35* and *ceh-51*; for an integrated transgenic protein reporter of PHA-4; and for a CRISPR knock-in protein reporter of CEH-51.

irx-1 IRX-1 *pes-1*



908
909
910

911 **Fig S3. Several early embryonic genes are expressed with anterior bias.**
912 Full lineage expression patterns are shown for integrated transgenic promoter reporters of *irx-1*
913 and *pes-1*; and for an integrated transgenic protein reporter of IRX-1.
914
915

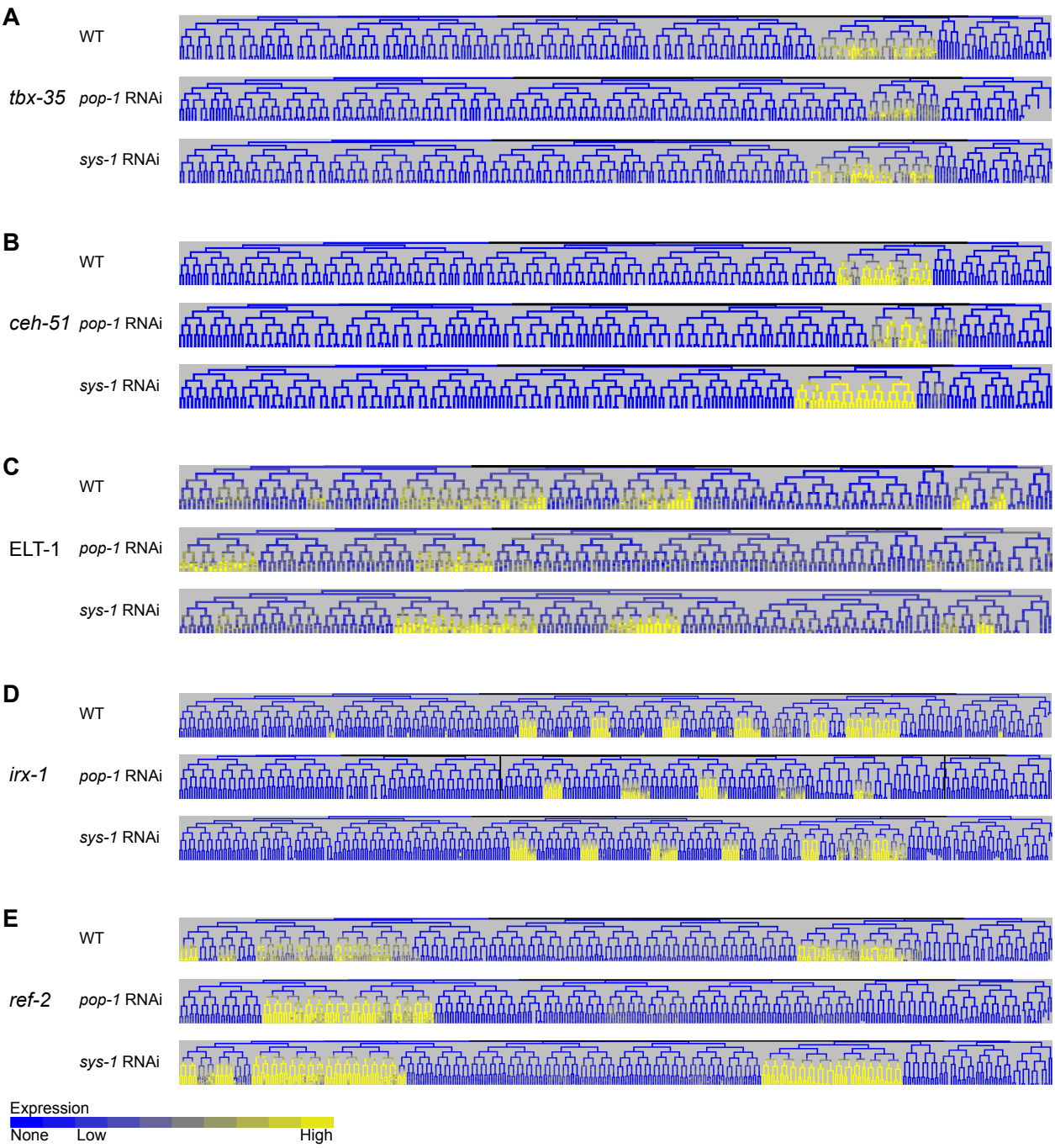


Fig S4. Anterior genes are regulated by Wnt effectors *pop-1* and *sys-1*.
Full example WT, *pop-1* RNAi, and *sys-1* RNAi lineages are shown for integrated transgenic promoter reporters of *tbx-35* (A), *ceh-51* (B), *irx-1* (D), and *ref-2* (E) and for an integrated transgenic protein reporter of ELT-1 (C).

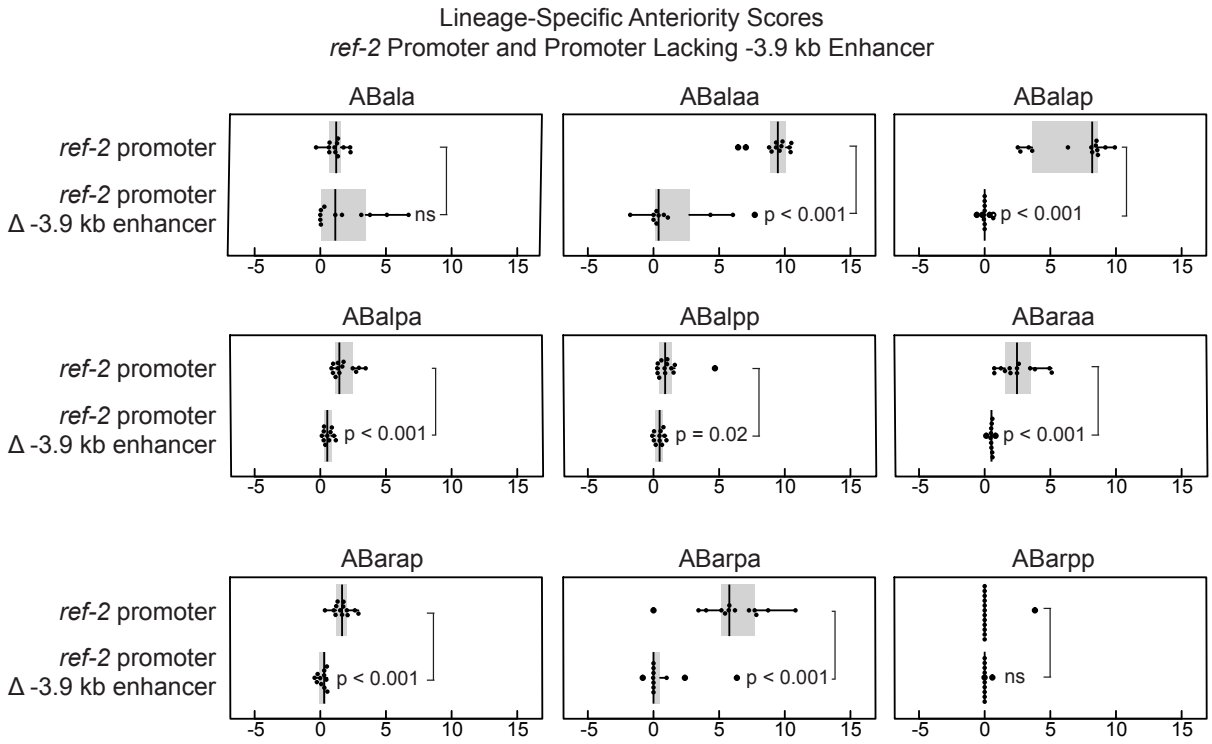


Fig S5. Deletion of the -3.9 kb enhancer from the *ref-2* promoter causes a reduction in the anterior bias in the expression driven by the promoter in most ABa sublineages in which the full-length promoter drives anterior-biased expression.

Box plots display anteriority scores of lineages descended from indicated cells.

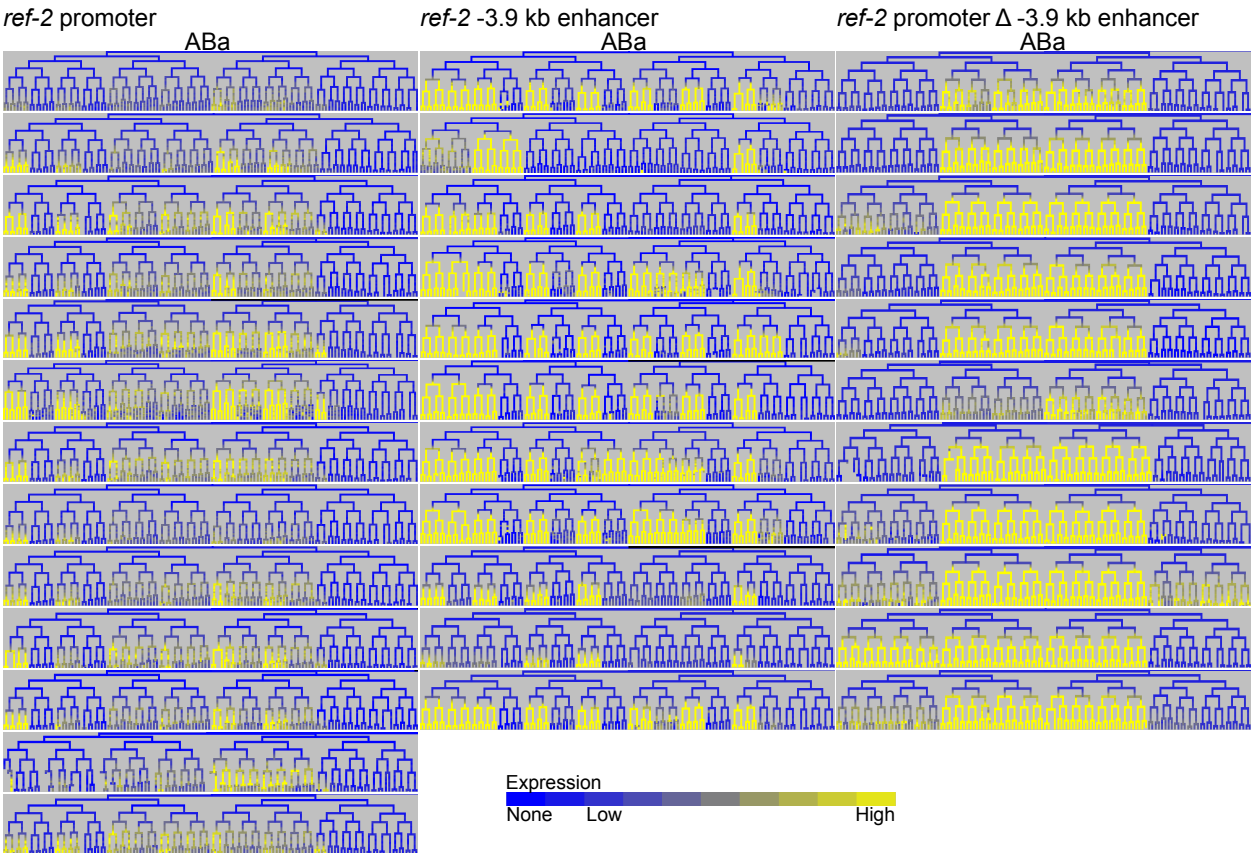
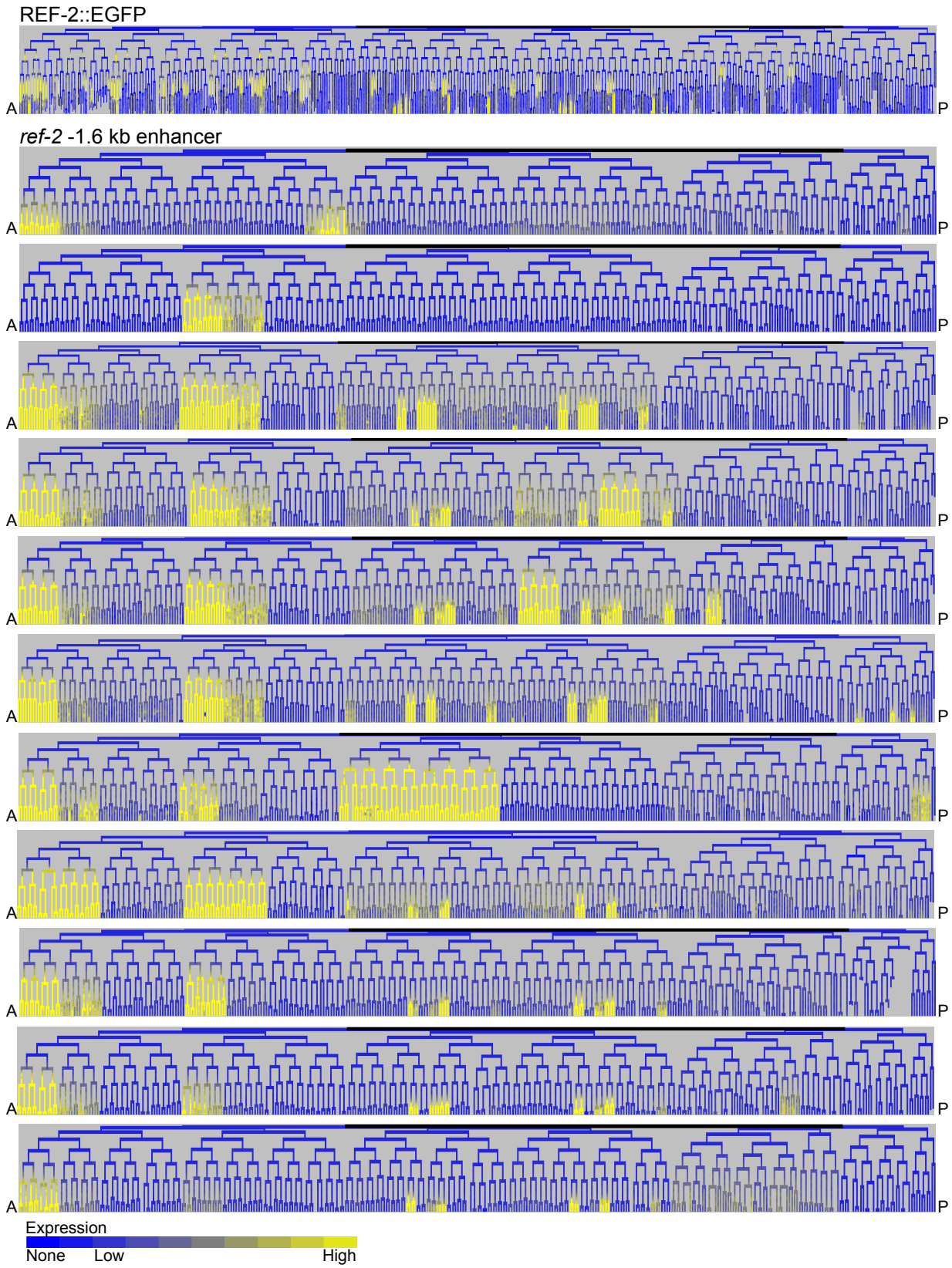
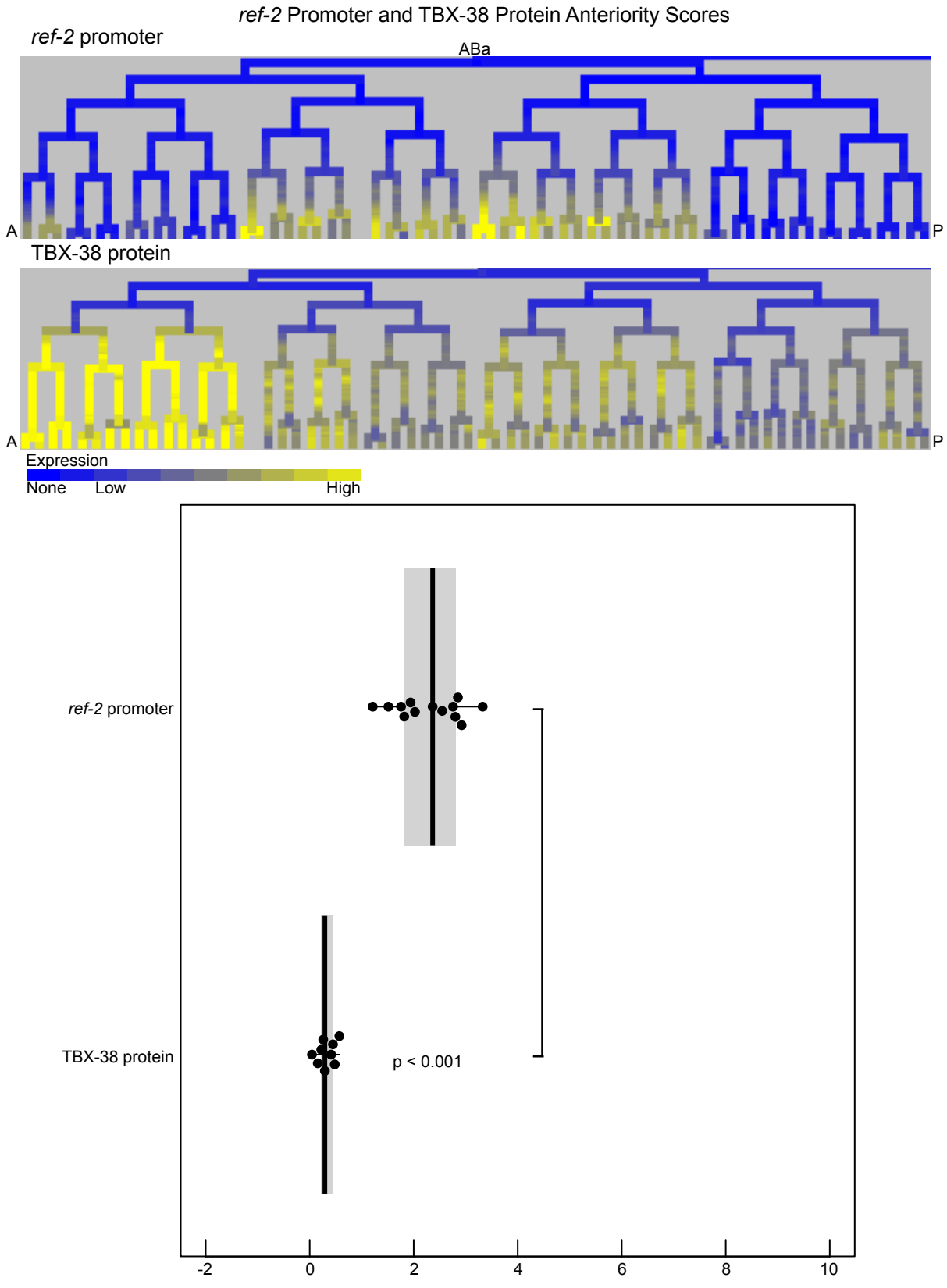


Fig S6. Expression patterns driven by the *ref-2* promoter, the -3.9 kb enhancer, and the promoter lacking the -3.9 kb enhancer are somewhat variable.
Displayed here are partial lineages for all analyzed embryos bearing these reporters to show the variation in their expression.



941 **Fig S7. Expression patterns driven by the *ref-2* -1.6 kb enhancer are variable.**
942 Displayed here are full lineages for all analyzed embryos bearing this reporter to show the variation
943 in its expression.
944



945
946

947 **Fig S8. A TBX-38::GFP CRISPR knock-in protein reporter has minimally anterior-biased**
948 **expression in the lineages in which the *ref-2* promoter drives anterior-biased expression.**
949 Displayed are partial lineages of the P*ref-2*::*his-24*::mCherry promoter reporter expression and the
950 TBX-38::GFP CRISPR knock-in expression. The box plot shows the mean anteriority scores for
951 the *ref-2* promoter reporter and the TBX-38::GFP CRISPR knock-in for the same lineages as in
952 Fig 5J.
953

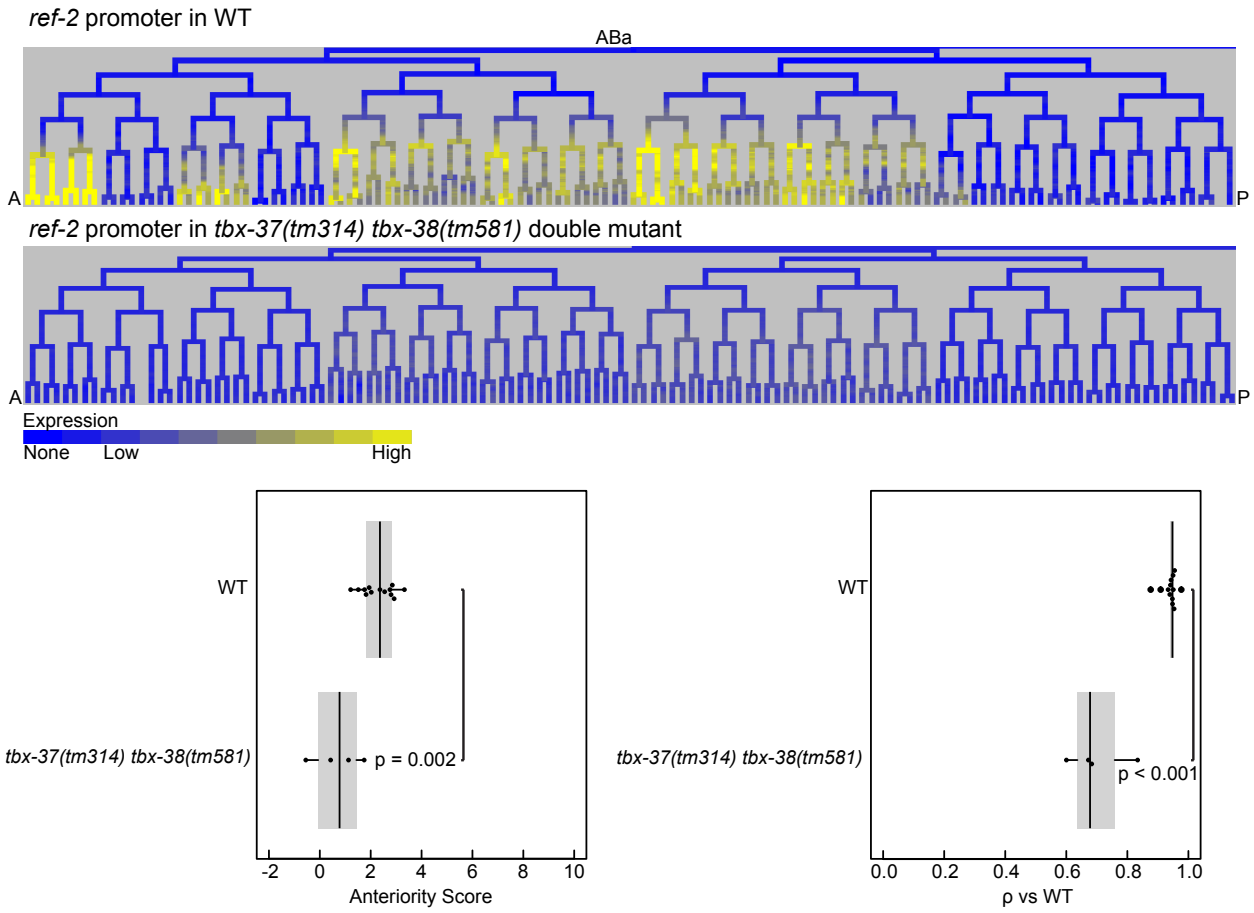


Fig S9. *tbx-37* and *tbx-38* are required for *ref-2* promoter-driven anterior-biased expression in the ABa lineage.

Displayed are partial lineages for the expression driven by the *ref-2* promoter in a WT background and in a *tbx-37(tm314) tbx-38(tm581)* double mutant background. Also shown are box plots of anteriority scores and Spearman's ρ for each of these groups, indicating a loss of anterior bias and correlation to WT for the *ref-2* promoter expression pattern in *tbx-37 tbx-38* double mutant embryos. Lineages used to determine anteriority scores are the same as in Fig 5J Spearman's ρ analysis uses the full ABa lineage expression pattern and was calculated relative to the *ref-2* promoter in the WT background.

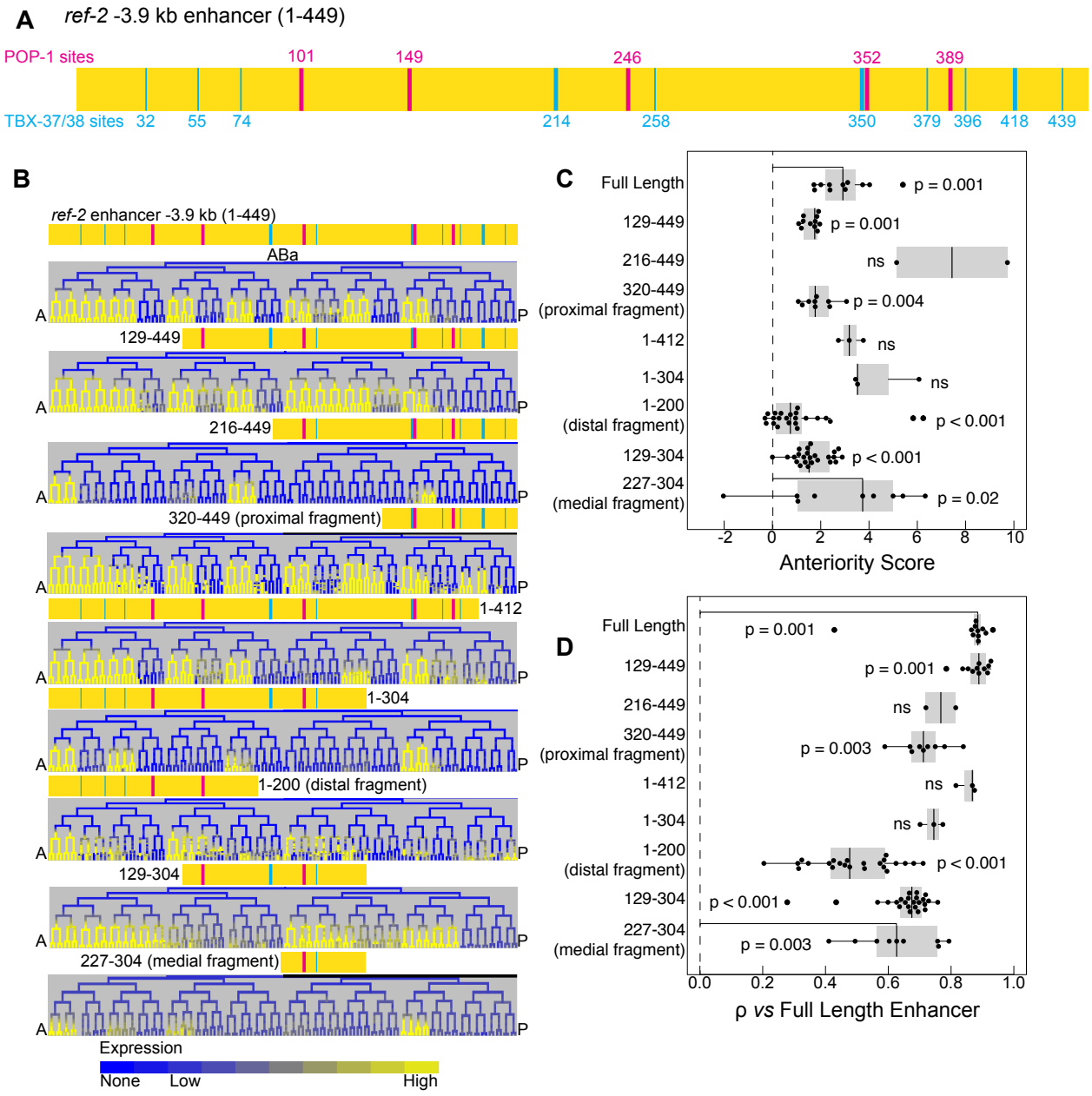


Fig S10. Three non-overlapping fragments of the -3.9 kb enhancer are each sufficient to drive anterior-biased expression in the ABa lineage of the early embryo.
(A) Model of the *ref-2* -3.9 kb enhancer with predicted TBX-37/38 and POP-1 sites indicated as in Fig 6. (B) Expression patterns driven by full-length *ref-2* -3.9 kb enhancer and by all tested fragments. (C-D) Box plots displaying the anteriority scores (C) and Spearman's ρ (D) for the full-length *ref-2* -3.9 kb enhancer and all tested fragments. Lineages used to determine anteriority scores are the same as in Fig 5J. Spearman's ρ analysis uses the full ABa lineage and was calculated relative to the full-length -3.9 kb enhancer.

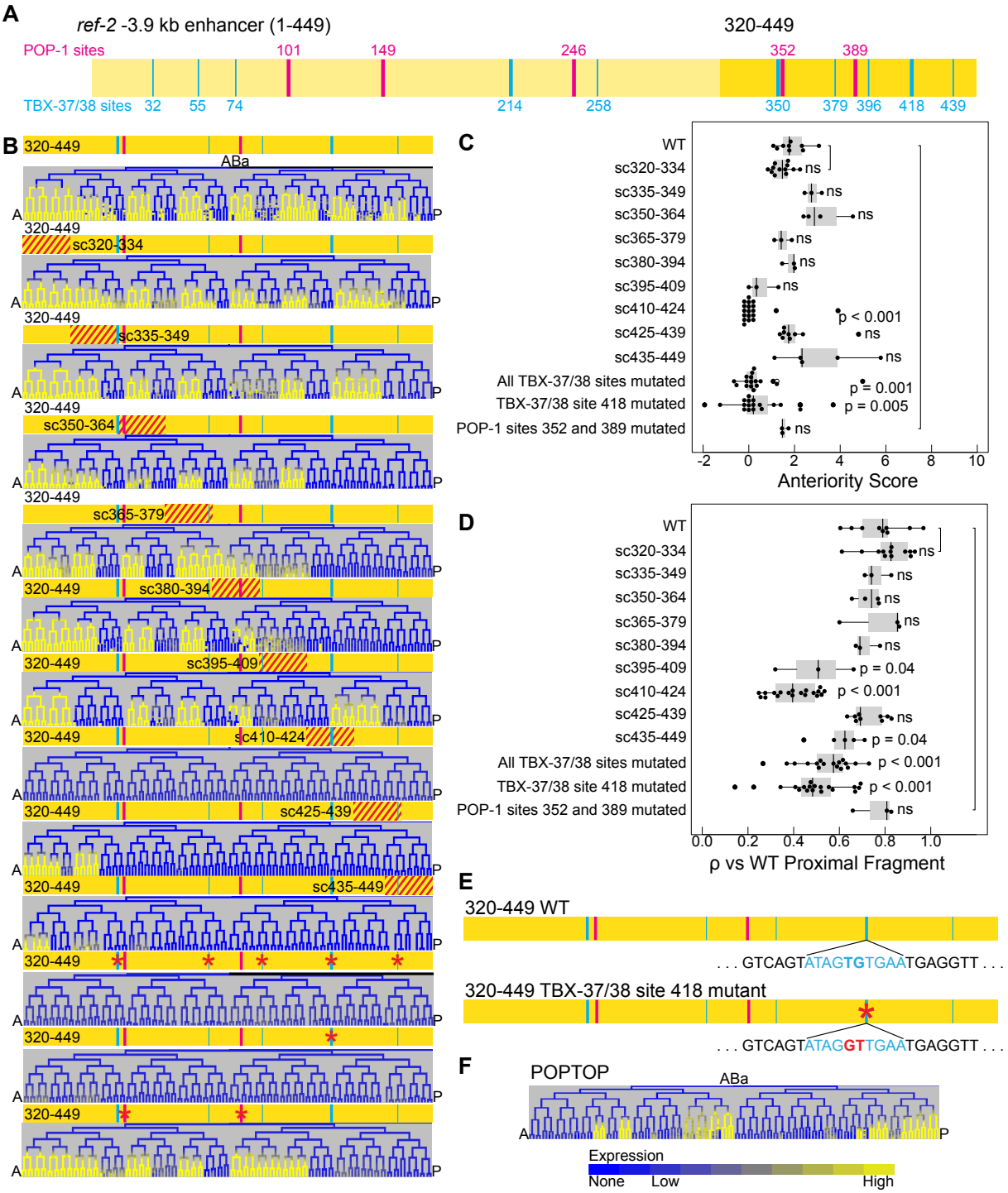


Fig S11. The broadly-expressed transcription factors *tbx-37* and *tbx-38* are required for the anterior-biased expression of *ref-2* in the ABa lineage.

(A) Model of the *ref-2* -3.9 kb enhancer with proximal fragment (base pairs 320-449) highlighted. Predicted TBX-37/38 and POP-1 sites are indicated as in Fig 6. (B) Expression patterns driven by WT proximal fragment of the *ref-2* -3.9 kb enhancer, with each 15 base-pair segment of the fragment scrambled, with all predicted TBX-37/38 sites mutated, with TBX_{418} mutated, and with

985 the two predicted POP-1 sites mutated. (C-D) Box plots displaying the anteriority scores (C) and
986 Spearman's ρ (D) for the WT proximal fragment, with each 15 base-pair segment of the fragment
987 scrambled, with all predicted TBX-37/38 sites mutated, with *TBX*$_{418}$ mutated, and with the two
988 predicted POP-1 sites mutated. Lineages used to determine anteriority scores are the same as in
989 Fig 5J. Spearman's ρ analysis uses the full ABa lineage and was calculated relative to the WT
990 proximal fragment. (E) Models of the proximal fragment of the *ref-2* -3.9 kb enhancer with the
991 WT sequence of *TBX*$_{418}$ and the mutated sequence used in the *TBX*$_{418}$ mutant reporters. (F)
992 Expression pattern driven by the seven-copy POP-1 site concatemer POPTOP [13,51].
993

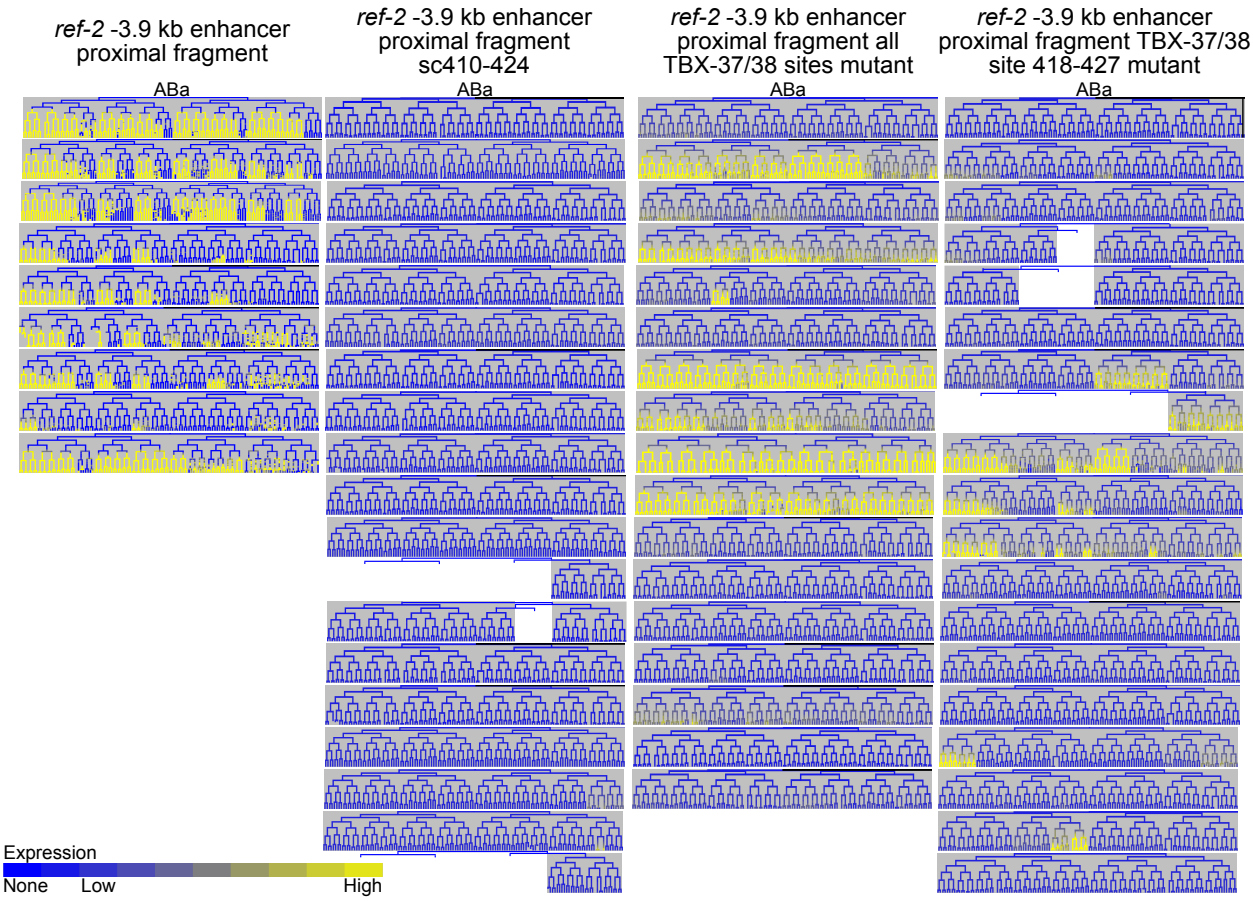


Fig S12. *ref-2* -3.9 kb enhancer WT proximal fragment, proximal fragment with all TBX-37/38 sites mutated, and proximal fragment with TBX_{418} mutated exhibit variation in their expression patterns.

Displayed are partial lineages for all analyzed embryos expressing reporters for the -3.9 kb enhancer WT proximal fragment, the *ref-2* -3.9 kb enhancer proximal fragment with base pairs 410-424 scrambled, the *ref-2* -3.9 kb enhancer proximal fragment with all TBX-37/38 sites mutated, and the *ref-2* -3.9 kb enhancer proximal fragment with TBX_{418} mutated to show the variation in the expression pattern for each of these reporters.

Table S1:
Caenorhabditis elegans strain list

Table S2:
Escherichia coli strain list

Table S3:
Quantitative analyses raw data

Table S4:
Anterior gene summary expression data

Table S5:
Recombinant DNA list

Acknowledgements:
We thank J. Archibald Millar for help with data analysis, Julia Richards for collecting and lineaging the PHA-4 movie, and Shaili Patel for her contributions as a laboratory technician. We thank Meera Sundaram for kindly permitting us to use her injection equipment. We thank Jonathan Rumley's PhD thesis committee members for helpful discussions, including committee chair Christopher Brown and committee members Michael Atchison, Peter Klein, Mary Mullins, and Meera Sundaram. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). We also thank the Waterston lab, the Maduro lab, and the Cochella lab for kindly providing us with *C.elegans* strains. We thank the *C. elegans* Reverse Genetics Core Facility at the University of British Columbia, which is part of the international *C. elegans* Gene Knockout Consortium, which produced the *ref-2* mutant allele (*ref-2(gk178))* used in this study. This work was funded by T32GM008216, F31GM123737, and R35GM130357.

References Cited

1. Hikasa H, Sokol SY. Wnt signaling in vertebrate axis specification. Cold Spring Harb Perspect Biol. 2013;5: a007955. doi:10.1101/cshperspect.a007955

2. Archbold HC, Yang YX, Chen L, Cadigan KM. How do they do Wnt they do?: regulation of transcription by the Wnt/β-catenin pathway. Acta Physiol Oxf Engl. 2012;204: 74–109. doi:10.1111/j.1748-1716.2011.02293.x

3. Martin BL, Kimelman D. Wnt signaling and the evolution of embryonic posterior development. Curr Biol CB. 2009;19: R215–R219. doi:10.1016/j.cub.2009.01.052

4. Zacharias AL, Murray JI. Combinatorial decoding of the invariant C. elegans embryonic lineage in space and time. genesis. 2016;54: 182–197. doi:10.1002/dvg.22928

5. Sulston JE, Schierenberg E, White JG, Thomson JN. The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev Biol. 1983;100: 64–119. doi:10.1016/0012-1606(83)90201-4

6. Bürglin TR, Finney M, Coulson A, Ruvkun G. Caenorhabditis elegans has scores of homoeobox-containing genes. Nature. 1989;341: 239–243. doi:10.1038/341239a0

7. Chen W, Lim HH, Lim L. A new member of the ras superfamily, the rac1 homologue from Caenorhabditis elegans. Cloning and sequence analysis of cDNA, pattern of developmental expression, and biochemical characterization of the protein. J Biol Chem. 1993;268: 320–324. doi:10.1016/S0021-9258(18)54152-1

8. van den Heuvel S. Cell-cycle regulation. WormBook. 2005. pp. 1–16. doi:10.1895/wormbook.1.28.1

9. Shaye DD, Greenwald I. OrthoList: A Compendium of C. elegans Genes with Human Orthologs. PLOS ONE. 2011;6: e20085. doi:10.1371/journal.pone.0020085

10. Kim W, Underwood RS, Greenwald I, Shaye DD. OrthoList 2: A New Comparative Genomic Analysis of Human and Caenorhabditis elegans Genes. Genetics. 2018;210: 445–461. doi:10.1534/genetics.118.301307

11. Herman M. C. elegans POP-1/TCF functions in a canonical Wnt pathway that controls cell migration and in a noncanonical Wnt pathway that controls cell polarity. Development. 2001;128: 581–590. doi:10.1242/dev.128.4.581

12. Mizumoto K, Sawa H. Two βs or not two βs: regulation of asymmetric division by β-catenin. Trends Cell Biol. 2007;17: 465–473. doi:10.1016/j.tcb.2007.08.004

13. Zacharias AL, Walton T, Preston E, Murray JI. Quantitative Differences in Nuclear β-catenin and TCF Pattern Embryonic Cells in C. elegans. Sternberg PW, editor. PLOS Genet. 2015;11: e1005585. doi:10.1371/journal.pgen.1005585

14. Lin R, Hill RJ, Priess JR. POP-1 and anterior-posterior fate decisions in C. elegans embryos. Cell. 1998;92: 229–239.

15. Maduro MF, Lin R, Rothman JH. Dynamics of a developmental switch: recursive intracellular and intranuclear redistribution of Caenorhabditis elegans POP-1 parallels Wnt-inhibited transcriptional repression. Dev Biol. 2002;248: 128–142. doi:10.1006/dbio.2002.0721

16. Huang S, Shetty P, Robertson SM, Lin R. Binary cell fate specification during C. elegans embryogenesis driven by reiterated reciprocal asymmetry of TCF POP-1 and its coactivatorβ-catenin SYS-1. Development. 2007;134: 2685–2695. doi:10.1242/dev.008268

17. Calvo D, Victor M, Gay F, Sui G, Po-Shan Luke M, Dufourcq P, et al. A POP-1 repressor complex restricts inappropriate cell type-specific gene transcription during Caenorhabditis elegans embryogenesis. EMBO J. 2001;20: 7197–7208. doi:10.1093/emboj/20.24.7197

18. Maduro MF, Kasmir JJ, Zhu J, Rothman JH. The Wnt effector POP-1 and the PAL-1/Caudal homeoprotein collaborate with SKN-1 to activate C. elegans endoderm development. Dev Biol. 2005;285: 510–523. doi:10.1016/j.ydbio.2005.06.022

19. Owraghi M, Broitman-Maduro G, Luu T, Roberson H, Maduro MF. Roles of the Wnt effector POP-1/TCF in the C. elegans endomesoderm specification gene network. Dev Biol. 2010;340: 209–221. doi:10.1016/j.ydbio.2009.09.042

20. Murgan S, Kari W, Rothbächer U, Iché-Torres M, Mélénec P, Hobert O, et al. Atypical Transcriptional Activation by TCF via a Zic Transcription Factor in C. elegans Neuronal Precursors. Dev Cell. 2015;33: 737–745. doi:10.1016/j.devcel.2015.04.018

21. Murgan S, Bertrand V. How targets select activation or repression in response to Wnt. Worm. 2015;4: e1086869. doi:10.1080/21624054.2015.1086869

22. Zhang CU, Blauwkamp TA, Burby PE, Cadigan KM. Wnt-Mediated Repression via Bipartite DNA Recognition by TCF in the Drosophila Hematopoietic System. Desplan C, editor. PLoS Genet. 2014;10: e1004509. doi:10.1371/journal.pgen.1004509

23. Sulston JE, Horvitz HR. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. Dev Biol. 1977;56: 110–156. doi:10.1016/0012-1606(77)90158-0

24. Murray JI, Boyle TJ, Preston E, Vafeados D, Mericle B, Weisdepp P, et al. Multidimensional regulation of gene expression in the C. elegans embryo. Genome Res. 2012;22: 1282–1294. doi:10.1101/gr.131920.111

25. Ma X, Zhao Z, Xiao L, Xu W, Kou Y, Zhang Y, et al. A 4D single-cell protein atlas of transcription factors delineates spatiotemporal patterning during embryogenesis. Nat Methods. 2021;18: 893–902. doi:10.1038/s41592-021-01216-1

26. Murray JI, Bao Z, Boyle TJ, Boeck ME, Mericle BL, Nicholas TJ, et al. Automated analysis of embryonic gene expression with cellular resolution in C. elegans. Nat Methods.

2008;5: 703–709. doi:10.1038/nmeth.1228

27. Packer Jonathan S., Zhu Qin, Huynh Chau, Sivaramakrishnan Priya, Preston Elicia, Dueck Hannah, et al. A lineage-resolved molecular atlas of C. elegans embryogenesis at single-cell resolution. Science. 2019;365: eaax1971. doi:10.1126/science.aax1971

28. Bao Z, Murray JI, Boyle T, Ooi SL, Sandel MJ, Waterston RH. Automated cell lineage tracing in Caenorhabditis elegans. Proc Natl Acad Sci U S A. 2006;103: 2707. doi:10.1073/pnas.0511111103

29. Santella A, Du Z, Nowotschin S, Hadjantonakis A-K, Bao Z. A hybrid blob-slice model for accurate and efficient detection of fluorescence labeled nuclei in 3D. BMC Bioinformatics. 2010;11: 580. doi:10.1186/1471-2105-11-580

30. Boyle TJ, Bao Z, Murray JI, Araya CL, Waterston RH. AceTree: a tool for visual analysis of Caenorhabditis elegans embryogenesis. BMC Bioinformatics. 2006;7: 275. doi:10.1186/1471-2105-7-275

31. Katzman B, Tang D, Santella A, Bao Z. AceTree: a major update and case study in the long term maintenance of open-source scientific software. BMC Bioinformatics. 2018;19: 121. doi:10.1186/s12859-018-2127-0

32. Yanai I, Baugh LR, Smith JJ, Roehrig C, Shen-Orr SS, Claggett JM, et al. Pairing of competitive and topologically distinct regulatory modules enhances patterned gene expression. Mol Syst Biol. 2008;4. doi:10.1038/msb.2008.6

33. Lin R, Thompson S, Priess JR. pop-1 encodes an HMG box protein required for the specification of a mesoderm precursor in early C. elegans embryos. Cell. 1995;83: 599–609. doi:10.1016/0092-8674(95)90100-0

34. Kidd AR, Miskowski JA, Siegfried KR, Sawa H, Kimble J. A β-Catenin Identified by Functional Rather Than Sequence Criteria and Its Role in Wnt/MAPK Signaling. Cell. 2005;121: 761–772. doi:10.1016/j.cell.2005.03.029

35. Phillips BT, Kidd AR, King R, Hardin J, Kimble J. Reciprocal asymmetry of SYS-1/beta-catenin and POP-1/TCF controls asymmetric divisions in Caenorhabditis elegans. Proc Natl Acad Sci. 2007;104: 3231–3236. doi:10.1073/pnas.0611507104

36. Broitman-Maduro G, Owraghi M, Hung WWK, Kuntz S, Sternberg PW, Maduro MF. The NK-2 class homeodomain factor CEH-51 and the T-box factor TBX-35 have overlapping function in C. elegans mesoderm development. Dev Camb Engl. 2009;136: 2735–2746. doi:10.1242/dev.038307

37. Broitman-Maduro G, Lin KT-H, Hung WWK, Maduro MF. Specification of the C. elegans MS blastomere by the T-box factor TBX-35. Dev Camb Engl. 2006;133: 3097–3106. doi:10.1242/dev.02475

38. Alper S, Kenyon C. The zinc finger protein REF-2 functions with the Hox genes to inhibit

cell fusion in the ventral epidermis of C. elegans. Development. 2002;129: 3335–3348. doi:10.1242/dev.129.14.3335

39. Walton T, Preston E, Nair G, Zacharias AL, Raj A, Murray JI. The Bicoid Class Homeodomain Factors ceh-36/OTX and unc-30/PITX Cooperate in C. elegans Embryonic Progenitor Cells to Regulate Robust Development. PLOS Genet. 2015;11: e1005003. doi:10.1371/journal.pgen.1005003

40. Richards JL, Zacharias AL, Walton T, Burdick JT, Murray JI. A quantitative model of normal Caenorhabditis elegans embryogenesis and its disruption after stress. Dev Biol. 2013;374: 12–23. doi:10.1016/j.ydbio.2012.11.034

41. Maduro MF, Meneghini MD, Bowerman B, Broitman-Maduro G, Rothman JH. Restriction of Mesendoderm to a Single Blastomere by the Combined Action of SKN-1 and a GSK-3β Homolog Is Mediated by MED-1 and -2 in C. elegans. Mol Cell. 2001;7: 475–485. doi:10.1016/S1097-2765(01)00195-2

42. Maduro MF, Hill RJ, Heid PJ, Newman-Smith ED, Zhu J, Priess JR, et al. Genetic redundancy in endoderm specification within the genus Caenorhabditis. Dev Biol. 2005;284: 509–522. doi:10.1016/j.ydbio.2005.05.016

43. Boeck ME, Boyle T, Bao Z, Murray J, Mericle B, Waterston R. Specific roles for the GATA transcription factors end-1 and end-3 during C. elegans E-lineage development. Spec Sect Hist Dev Biol. 2011;358: 345–355. doi:10.1016/j.ydbio.2011.08.002

44. Good K, Ciosk R, Nance J, Neves A, Hill RJ, Priess JR. The T-box transcription factors TBX-37 and TBX-38 link GLP-1/Notch signaling to mesoderm induction in C. elegans embryos. Development. 2004;131: 1967–1978. doi:10.1242/dev.01088

45. Andachi Y. Caenorhabditis elegans T-box genes tbx-9 and tbx-8 are required for formation of hypodermis and body-wall muscle in embryogenesis. Genes Cells. 2004;9: 331–344. doi:10.1111/j.1356-9597.2004.00725.x

46. Neves A, Priess JR. The REF-1 Family of bHLH Transcription Factors Pattern C. elegans Embryos through Notch-Dependent and Notch-Independent Pathways. Dev Cell. 2005;8: 867–879. doi:10.1016/j.devcel.2005.03.012

47. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 2005;15: 1034–1050. doi:10.1101/gr.3715005

48. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res. 2010;20: 110–121. doi:10.1101/gr.097857.109

49. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The Human Genome Browser at UCSC. Genome Res. 2002;12: 996–1006. doi:10.1101/gr.229102

50. Charest J, Daniele T, Wang J, Bykov A, Mandlbauer A, Asparuhova M, et al. Combinatorial Action of Temporally Segregated Transcription Factors. Dev Cell. 2020;55: 483-499.e7. doi:10.1016/j.devcel.2020.09.002

51. LaBonty M, Szmygiel C, Byrnes LE, Hughes S, Woollard A, Cram EJ. CACN-1/Cactin plays a role in Wnt signaling in C. elegans. PloS One. 2014;9. doi:10.1371/journal.pone.0101945

52. Murray JI, Preston E, Crawford JP, Rumley JD, Amom P, Anderson BD, et al. The anterior Hox gene ceh-13 and elt-1/GATA activate the posterior Hox genes nob-1 and php-3 to specify posterior lineages in the C. elegans embryo. bioRxiv. 2021; 2021.02.09.430385. doi:10.1101/2021.02.09.430385

53. Poole RJ, Hobert O. Early Embryonic Programming of Neuronal Left/Right Asymmetry in C. elegans. Curr Biol. 2006;16: 2279–2292. doi:10.1016/j.cub.2006.09.041

54. Fukushige T, Brodigan TM, Schriefer LA, Waterston RH, Krause M. Defining the transcriptional redundancy of early bodywall muscle development in C. elegans: evidence for a unified theory of animal muscle development. Genes Dev. 2006;20: 3395–3406. doi:10.1101/gad.1481706

55. Fukushige T, Krause M. The myogenic potency of HLH-1 reveals wide-spread developmental plasticity in early C. elegans embryos. Dev Camb Engl. 2005;132: 1795–1805. doi:10.1242/dev.01774

56. Simonis N, Rual J-F, Carvunis A-R, Tasan M, Lemmens I, Hirozane-Kishikawa T, et al. Empirically controlled mapping of the Caenorhabditis elegans protein-protein interactome network. Nat Methods. 2009;6.

57. Tintori SC, Osborne Nishimura E, Golden P, Lieb JD, Goldstein B. A Transcriptional Lineage of the Early C. elegans Embryo. Dev Cell. 2016;38: 430–444. doi:10.1016/j.devcel.2016.07.025

58. Whiting J, Marshall H, Cook M, Krumlauf R, Rigby PW, Stott D, et al. Multiple spatially specific enhancers are required to reconstruct the pattern of Hox-2.6 gene expression. Genes Dev. 1991;5: 2048–2059. doi:10.1101/gad.5.11.2048

59. Epstein DJ, McMahon AP, Joyner AL. Regionalization of Sonic hedgehog transcription along the anteroposterior axis of the mouse central nervous system is regulated by Hnf3-dependent and -independent mechanisms. Development. 1999;126: 281–292. doi:10.1242/dev.126.2.281

60. Teng Y, Girard L, Ferreira HB, Sternberg PW, Emmons SW. Dissection of cis-regulatory elements in the C. elegans Hox gene egl-5 promoter. Dev Biol. 2004;276: 476–492. doi:10.1016/j.ydbio.2004.09.012

61. Streit A, Kohler R, Marty T, Belfiore M, Takacs-Vellai K, Vigano M-A, et al. Conserved Regulation of the Caenorhabditis elegans labial/Hox1 Gene ceh-13. Dev Biol. 2002;242:

96–108. doi:10.1006/dbio.2001.0544

62. Landmann F, Quintin S, Labouesse M. Multiple regulatory elements with spatially and temporally distinct activities control the expression of the epithelial differentiation gene lin-26 in C. elegans. Dev Biol. 2004;265: 478–490. doi:10.1016/j.ydbio.2003.09.009

63. Chen RA-J, Down TA, Stempor P, Chen QB, Egelhofer TA, Hillier LW, et al. The landscape of RNA polymerase II transcription initiation in C. elegans reveals promoter and enhancer architectures. Genome Res. 2013;23: 1339–1347. doi:10.1101/gr.153668.112

64. Serizay J, Dong Y, Jänes J, Chesney M, Cerrato C, Ahringer J. Distinctive regulatory architectures of germline-active and somatic genes in C. elegans. Genome Res. 2020;30: 1752–1765. doi:10.1101/gr.265934.120

65. Bertrand V, Hobert O. Linking asymmetric cell division to the terminal differentiation program of postmitotic neurons in C. elegans. Dev Cell. 2009;16: 563–575. doi:10.1016/j.devcel.2009.02.011

66. Kuersten S, Segal SP, Verheyden J, LaMartina SM, Goodwin EB. NXF-2, REF-1, and REF-2 Affect the Choice of Nuclear Export Pathway for tra-2 mRNA in C. elegans. Mol Cell. 2004;14: 599–610. doi:10.1016/j.molcel.2004.05.004

67. Clark E, Akam M. Odd-paired controls frequency doubling in Drosophila segmentation by altering the pair-rule gene regulatory network. Wittkopp PJ, editor. eLife. 2016;5: e18215. doi:10.7554/eLife.18215

68. Yoon Y, Klomp J, Martin-Martin I, Criscione F, Calvo E, Ribeiro J, et al. Embryo polarity in moth flies and mosquitoes relies on distinct old genes with localized transcript isoforms. Wittkopp PJ, Pick L, Popadic A, Akam M, editors. eLife. 2019;8: e46711. doi:10.7554/eLife.46711

69. Kamath RS, Fraser AG, Dong Y, Poulin G, Durbin R, Gotta M, et al. Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature. 2003;421: 231–237. doi:10.1038/nature01278

70. Vora S, Phillips BT. Centrosome-Associated Degradation Limits β-Catenin Inheritance by Daughter Cells after Asymmetric Division. Curr Biol. 2015;25: 1005–1016. doi:10.1016/j.cub.2015.02.020

71. Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, et al. Determination and inference of eukaryotic transcription factor sequence specificity. Cell. 2014;158: 1431–1443. doi:10.1016/j.cell.2014.08.009

72. Narasimhan K, Lambert SA, Yang AWH, Riddell J, Mnaimneh S, Zheng H, et al. Mapping and analysis of transcription factor sequence specificities. eLife. 2015;4. doi:10.7554/eLife.06967

73. Bao Z, Murray JI. Mounting Caenorhabditis elegans Embryos for Live Imaging of

1250 Embryogenesis. Cold Spring Harb Protoc. 2011;2011: pdb.prot065599.
1251 doi:10.1101/pdb.prot065599