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 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] (Figure 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*.

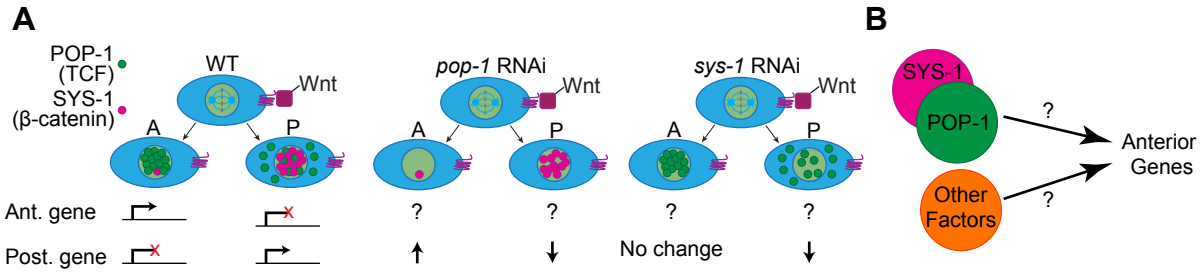


Figure 1: How are anterior genes regulated? A) The Wnt/β-catenin asymmetry pathway regulates anterior-posterior patterning by producing asymmetric nuclear concentrations of POP-1/TCF and SYS-1/β-catenin in anterior and posterior sister cells. Low nuclear concentrations of POP-1 and high nuclear concentrations 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 (Figure 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* appears to be 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

92 and high nuclear POP-1, and both *pop-1* and *ref-2* are required to activate *ttx-3* expression. In the
93 posterior sister, the Wnt pathway is active, with high nuclear SYS-1 levels and low nuclear POP-
94 1 levels, and *sys-1* is required to repress expression of *ttx-3*. POP-1 and REF-2 can directly interact,
95 suggesting that REF-2 and POP-1 bind as a complex to activate *ttx-3* expression in the AIY mother.
96 The role of SYS-1 is less clear but some evidence suggests SYS-1 may bind to the POP-1/REF-2
97 complex to repress expression, or may sequester away the limited POP-1, such that POP-1 cannot
98 interact with REF-2 [20,21]. Similar regulation has also been seen in other species. For example,
99 *Drosophila* TCF/Pangolin can act in a "reverse transcriptional switch" such that when TCF is
100 bound to a non-optimal site without binding to β-catenin/Armadillo it activates expression, and
101 when it binds along with β-catenin it represses expression [22].
102
103 Because of the role of POP-1 and SYS-1 in regulating the anterior-specific expression of *ttx-3*, we
104 hypothesized that these Wnt pathway components may also regulate other anterior genes. POP-1
105 and SYS-1 could act as general anterior expression regulators, interacting with lineage-restricted
106 co-regulators to ensure appropriate expression of anterior genes in the correct lineages [20,21]. We
107 tested this by mining single-cell-resolution expression data to identify transcription factors
108 expressed in anterior-specific patterns [23–25], and show that of five we tested, all require *pop-1*
109 and/or *sys-1* for either anterior expression or posterior repression. We focus in more detail on a
110 single anterior-specific gene, *ref-2/ZIC*. By automated lineage analysis of mutants, we find that
111 *ref-2* is required for normal cell division timing and cell position in *ref-2* expressing lineages.
112 Embryonic expression of *ref-2* is driven by at least three developmental enhancers, two of which
113 drive early embryonic expression and one of which drives expression in later-stage embryos. The
114 most distal enhancer (3.9 kb upstream of the transcription start site) drives highly anterior-biased
115 expression. Functional dissection revealed striking redundancy within this enhancer, reminiscent
116 of enhancer redundancy that facilitates robust gene expression in other species [26]. We also
117 identified a role for the T-box transcription factors *tbx-37* and *tbx-38* in driving anterior-specific
118 expression. Surprisingly, concatemers of a single T-box binding site reiterate much of the anterior-
119 specific expression pattern of *ref-2* in early embryos, suggesting a key role for these genes in
120 driving anterior expression.
121
122

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 (Figure 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 (Figure 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 (Figure 2B). Following convention [5,27], 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 Figure 2D, E) [23,24]. Both anterior-specific and posterior-specific gene expression were similarly common. While posterior specific genes (e.g. Figure 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 (Figure 2A) [23–25] to identify 19 conserved genes expressed in anterior lineage-specific patterns (Figure 2C).

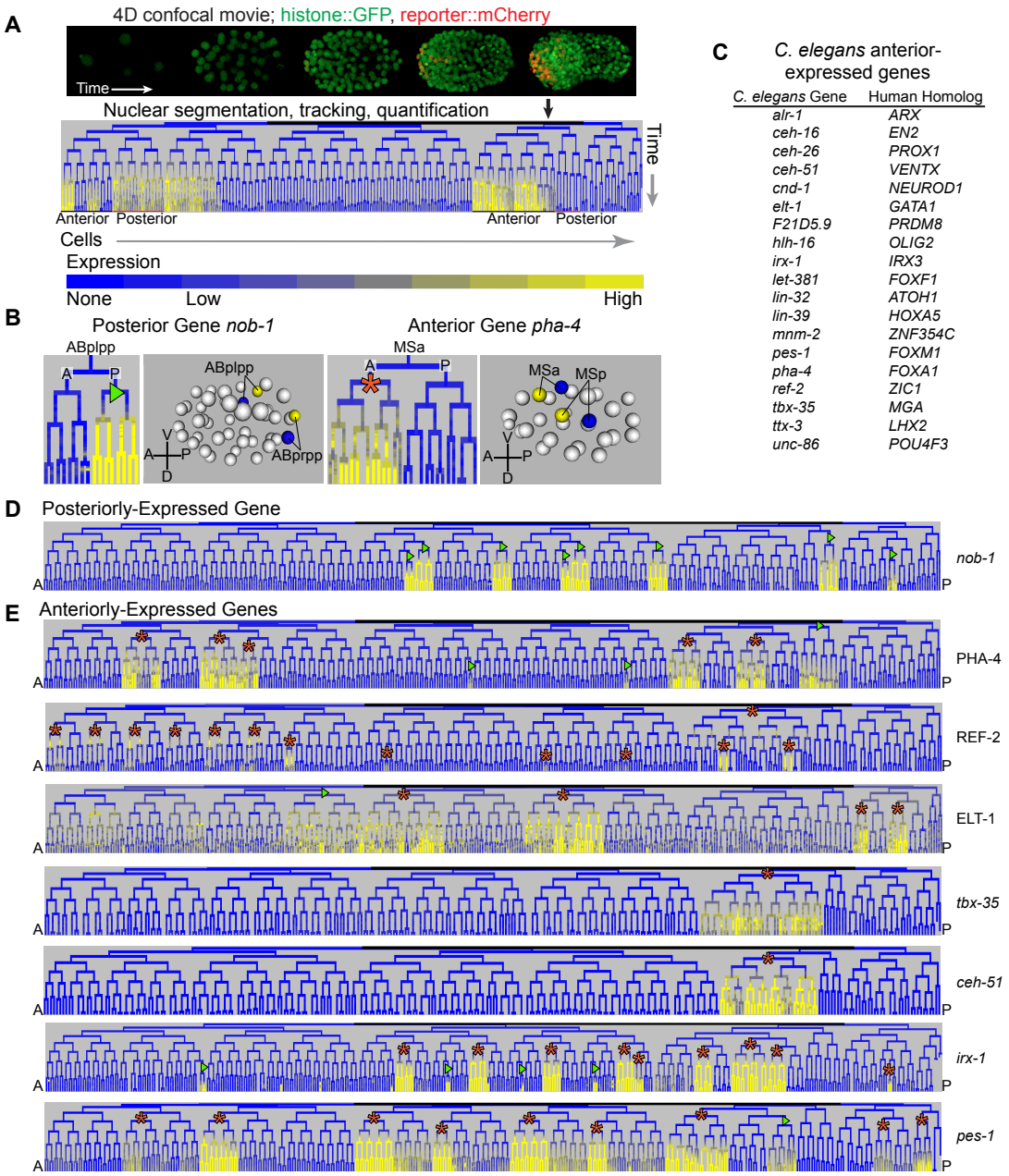


Figure 2: Identification of genes with anterior-biased expression. A) Automated lineaging traces lineages and quantifies expression from 4D confocal images. In the lineage trees, vertical lines represent cells, and horizontal lines represent cell divisions. Most cells divide along the anterior-posterior 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 seven anterior genes; most are expressed in unique combinations of mostly anterior lineages.

158 We measured cellular resolution expression profiles of seven of these by live confocal imaging
159 followed by automated cell tracking and lineage tracing (Figure 2A, E, Figure S1, Figure S2,
160 Figure S3) [25,28]. For each gene, we collected 4D time-lapse images of transgenic embryos
161 expressing a histone-mCherry (*pha-4*, *ref-2*, *elt-1*, *tbx-35*, *irx-1*, and *pes-1*), or GFP (*ceh-51*)
162 reporter under the control of upstream regulatory sequences of each gene (referred to here as
163 "promoter reporters"). For five genes we also collected images for fosmid transgenes (*irx-1, elt-1,*
164 *pha-4* and *ref-2)* or CRISPR knock-ins (*ceh-51*) in which the protein is fused to GFP and
165 surrounded by its normal genomic context ("protein reporters"). Each embryo also contained a
166 ubiquitously expressed second-color, histone-GFP or histone-mCherry fusion, used for cell
167 tracking. We identified each nucleus at each time point and traced them across time using
168 StarryNite [28,29] cell tracking software, and used AceTree [30,31] for subsequent manual error
169 correction and validation that the extracted lineages were correct. Finally, we quantified the
170 expression of the reporter in each nucleus at each time point (Figure 2A) [25]. The results confirm
171 the anterior-specific expression of each gene, and, in some cases, identify previously unknown
172 expression patterns (Figure S1, Figure S2, Figure S3). For each promoter reporter, the anterior-
173 specific patterns were consistent, but the protein fusion reporters all had additional expressing
174 lineages and dynamics indicative of distal enhancers and post-transcriptional regulation (Figure
175 S1, Figure S2, Figure S3). The anterior-biased expression of these genes is also seen in embryonic
176 single-cell RNA-seq data (Figures S4 and S5) [32].
177
178 **The Wnt pathway effectors *pop-1* and *sys-1* are required for normal anterior-specific**
179 **expression**
180
181 Most genes expressed preferentially in posterior lineages require the Wnt effector transcription
182 factor POP-1/TCF either for activation in posterior cells (together with its coactivator SYS-1/β-
183 catenin) or for repression in anterior cells, and at least some are direct targets [12,13,33–36].
184 Although *pop-1* is required for the normal expression of some anterior genes [20,33,37], it is
185 unclear whether *pop-1* and *sys-1* are generally required for anterior-specific expression, as they
186 are for posterior-specific expression.
187

188 To test this, we measured the expression of five anterior gene reporters (*tbx-35*, *ceh-51*, *elt-1*, *irx-*
189 *1*, and *ref-2*) after *pop-1* and *sys-1* RNAi by live 4D imaging and lineage tracing as described
190 above. We used promoter reporters for *tbx-35*, *ceh-51*, *irx-1*, and *ref-2* because our wild-type
191 expression data indicated these reporters are sufficient for anterior regulation (Figure S1, Figure
192 S2, Figure S3). For *elt-1*, anterior-specific expression was more robust for the protein reporter so
193 we tested that reporter's dependence on *pop-1* and *sys-1* (Figure 3; Figure S6).
194
195 Reporters of two of the genes we tested, *tbx-35* and *ceh-51*, are expressed solely in the anterior
196 lineage MS (Figure S2; Figure S6A, B), and previous qualitative expression analyses indicated
197 that expression persists in the absence of *pop-1* [19,37,38]. Consistent with this, we detected
198 expression of both genes in the MS lineage after *pop-1* RNAi, however this expression was
199 significantly decreased. We also observed increased expression of both genes in the E lineage
200 (posterior sister of MS), although this expression remained lower than both the wild-type and *pop-*
201 *1* RNAi MS levels (Figure 3A, B; Figure S6A, B; Table S4). We also observe low-level ectopic
202 expression of *tbx-35* in the E lineage under *sys-1* RNAi. These results indicate that both *tbx-35* and
203 *ceh-51* require *pop-1* for full expression in the MS lineage and for repression in the E lineage (as
204 shown previously [19,37]), and that *sys-1* is required to repress *tbx-35* in the E lineage.
205
206 *elt-1* reporters are expressed in several early anterior lineages that primarily give rise to ectodermal
207 fates (ABpla, ABpra, Caa and Cpa), and in one posterior lineage (ABarp) (Figure S1; Figure S6C).
208 In the Caa and Cpa lineages, anterior lineages that express both the promoter and protein reporters,
209 *pop-1* RNAi results in decreased *elt-1* reporter expression. However, RNAi for neither *pop-1* nor
210 *sys-1* affects expression in their posterior sister lineages Cap and Cpp (Figure 3C; Figure S6C;
211 Table S4). Similarly, *pop-1* RNAi results in reduced expression of the protein reporter in the ABpla
212 and ABpra lineages, while RNAi knockdown of neither *pop-1* nor *sys-1* has much effect on their
213 posterior sister lineages ABplp and ABprp (Figure S6C; Table S4). Thus, in both the ABp and C
214 lineages, *pop-1* is required for anterior expression of *elt-1*, but neither *pop-1* nor *sys-1* is required
215 for posterior repression.

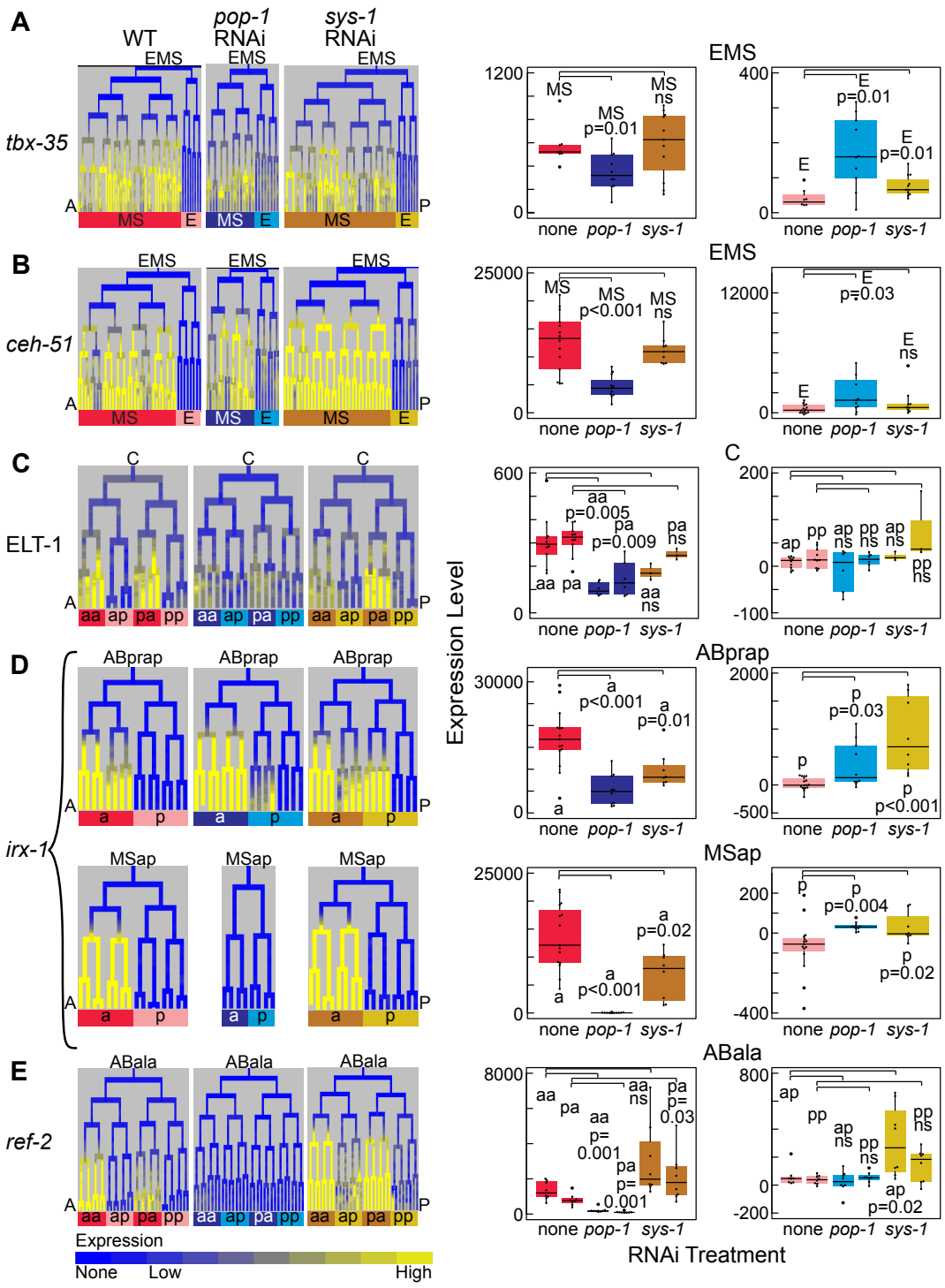


Figure 3: Regulation of anterior genes by Wnt effectors *pop-1* and *sys-1*. A-E) Expression pattern (left) and quantification (right) of reporters for the anterior genes *tbx-35* (A), *ceh-51* (B), ELT-1 (C), *irx-1* (D), and *ref-2* (E) in specific anterior lineages and their posterior sisters. Expression 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).

224 *irx-1* reporters are expressed early in four related anterior sublineages of ABp, and in three anterior
225 sublineages of MS, while in later embryos they are expressed in both anterior and posterior
226 lineages (Figure S3; Figure S6D). Like for *elt-1, pop-1* RNAi reduces *irx-1* reporter expression in
227 anterior ABp-derived sublineages, and two of the MS-derived expressing sublineages lose
228 expression nearly completely. After both *pop-1* and *sys-1* RNAi, the *irx-1* reporter expression
229 expands into several posterior sublineages of ABp whose anterior sisters normally express *irx-1*
230 (Figure 3D; Figure S6D; Table S4). These results indicate that *irx-1* requires *pop-1* for expression
231 in anterior lineages and both *pop-1* and *sys-1* for repression in posterior lineages.
232
233 Reporters for *ref-2* are expressed in the early embryo in all MS descendants and in several related
234 anterior sublineages of ABa that arise at the 50-cell stage (Figure S1; Figure S6E). In later
235 embryos, expression occurs in two anterior sublineages of MS, in several anterior ABp-derived
236 sublineages, and in the Pn ventral epidermal blast cells, consistent with previous studies [39]. After
237 *pop-1* RNAi, expression of the *ref-2* promoter reporter is significantly reduced in anterior
238 sublineages of ABala (Figure 3E). As for other genes, *sys-1* RNAi caused de-repression of *ref-2*
239 in several posterior sisters of the normally expressing sublineages, but we did not detect de-
240 repression in these sublineages after *pop-1* RNAi (Figure 3E; Figure S6E; Table S4). Thus, *ref-2*
241 requires *pop-1* for expression in anterior lineages and *sys-1* for repression in posterior lineages.
242
243 To summarize these perturbations, each anterior gene tested requires *pop-1* for full expression in
244 anterior lineages. Several genes also require *pop-1* (*tbx-35*, *ceh-51* and *irx-1*) or *sys-1* (*tbx-35, irx-*
245 *1* and *ref-2*) for repression in some posterior lineages. Thus, it appears that the expression of most
246 anterior genes requires the Wnt pathway components POP-1 and SYS-1, but this dependency is
247 complex as was seen previously for posterior genes [13]. Another layer of complexity is that in
248 some cases, we cannot rule out that the loss of expression after *pop-1* RNAi is, at least in part, due
249 to indirect effects of *pop-1* loss, such as an earlier anterior to posterior fate conversion. For
250 example, for *irx-1* expression in descendants of the MS lineage, an MS to E fate conversion could
251 explain the loss of expression after *pop-1* RNAi. However, in this case, we observe consistent
252 reduction in *irx-1* reporter expression after *pop-1* RNAi even in embryos without a strong MS to
253 E conversion (as assayed by cell cycle timing). Expression changes of *irx-1* in the AB lineage

254 cannot be explained by known fate conversions, indicating that the expression changes are at least
255 in part due to a direct effect of *pop-1*.
256
257 **The anterior-specific transcription factor *ref-2/ZIC* is required for normal patterns of cell**
258 **division and cell position in expressing lineages.**
259
260 The striking anterior-specific expression patterns observed for early embryonic TFs raises the
261 question of whether these factors regulate the development of anterior lineages. To address this
262 question, we focused on one anterior gene, *ref-2*. *ref-2* encodes a zinc finger transcription factor
263 homologous to mammalian ZIC genes and *Drosophila* Odd-Paired. Previous work showed that
264 *ref-2* is required to prevent progeny of "Pn" epidermal blast cells from fusing with the major
265 hypodermal syncytium in larval stages [39] and a predicted null allele of *ref-2*, *ref-2(gk178)*,
266 causes larval lethality [40]. Recently, REF-2 was shown to activate the anterior-specific expression
267 of the neural regulator *ttx-3* in the AIY mother cell by binding the *ttx-3 cis*-regulatory region and
268 recruiting POP-1 independent of POP-1 DNA binding activity [20]. Since *ref-2* is expressed in
269 many other anterior lineages in early embryos, we investigated whether *ref-2* mutants have
270 developmental defects in those lineages.
271
272 We tested for embryonic lineage phenotypes at single-cell resolution by analyzing 4D time-lapse
273 images of eight embryos homozygous for the deletion allele *ref-2(gk178)* and expressing a
274 ubiquitous histone-mCherry fusion for cell tracking. We used StarryNite to track the lineage of
275 each embryo, and used AceTree to curate the lineages through the >550-cell stage (six embryos)
276 or >350-cell stage (two embryos). We compared the position and cell cycle length of each cell in
277 each embryo to those from a reference of 17 wild-type embryos [41,42] to identify outliers
278 ("defects") (Figure 4A).
279

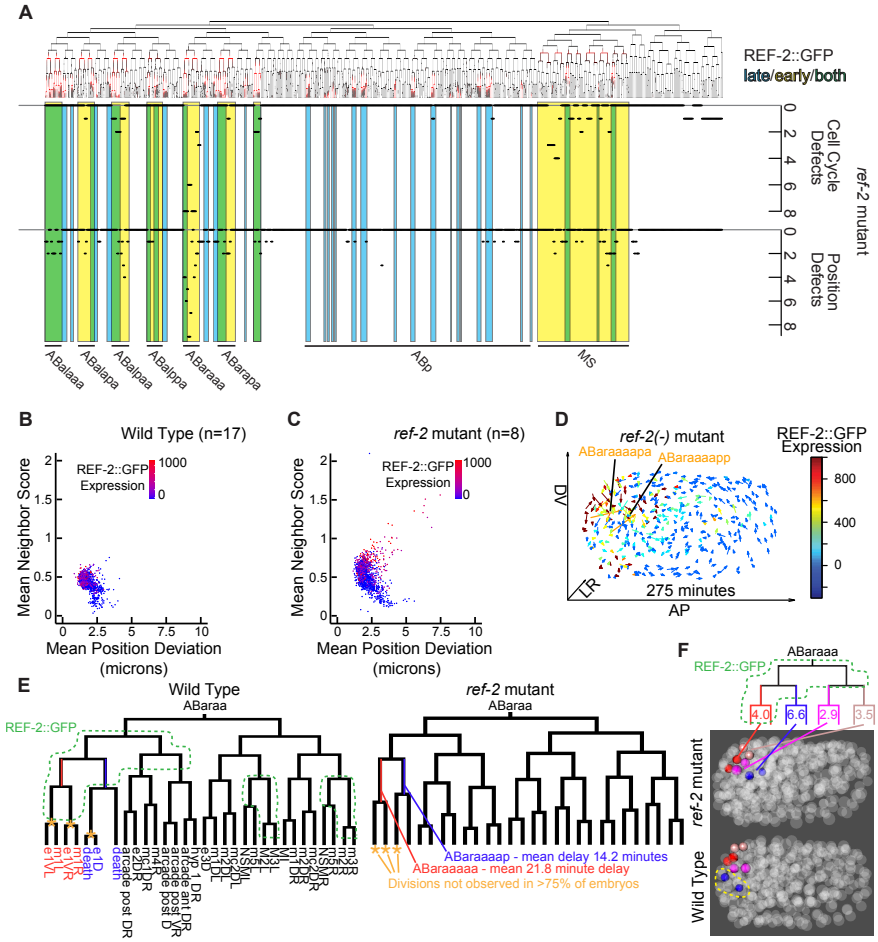


Figure 4: *ref-2* is required for proper division timing and positioning for many embryonic cells that are descended from cells that express *ref-2*. A) Lineage tree depicting the REF-2 protein expression pattern (top), and chart indicating the number of embryos out of eight homozygous *ref-2* mutant 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 out of all 17 embryos tested [41]. B, C) Scatter plots 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 [41]. 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.

302 We identified a total of 55 cells with early, late or missing divisions, and 54 of these were derived
303 from REF-2::GFP expressing lineages ($P < 10^{-15}$). Overall 95% of cell cycle defects were delays
304 or missed divisions, indicating a role for *ref-2* in promoting cell division. Similarly, we identified
305 68 position defects, of which 56 were in *ref-2* expressing lineages (Figure 4B, C, $P = 1.1 * 10^{-14}$,
306 chi-squared test). Ten *ref-2*-expressing cells, and no non-expressing cells, had position defects in
307 two or more embryos. The cells with the strongest position and cell cycle length defects were from
308 the *ref-2*-expressing anterior ABaraaa lineage, which produces primarily anterior pharyngeal cells
309 (e1, m1 and arcade cells). The most frequent cell cycle length defects were in ABaraaaaa and
310 ABaraaaap (both defective in 8 of 8 embryos), which were delayed by 21.8 minutes and 14.2
311 minutes, respectively (Figure 4E). The largest position defects were in ABaraaaapa and
312 ABaraaaapp, which were mispositioned on average 6 microns posterior of their wild-type position
313 (Figure 4D, F). However, other defects were broadly distributed across *ref-2*-expressing lineages,
314 in particular MS- and ABa-derived anterior sublineages (Figure 4A). In summary, the anterior
315 lineage-expressed TF *ref-2* is required for the normal development of anterior lineages, although
316 the low penetrance of many defects suggests that *ref-2* may act with other partially redundant
317 regulators, similar to other early zygotic TFs [41,43–48].
318
319 **Modular enhancers control *ref-2* embryonic expression**
320
321 The anterior expression and phenotypes of *ref-2* raise the question of what sequences and
322 regulators control this expression. To identify genomic sequences responsible for *ref-2* anterior
323 expression, we first compared expression driven by the 31.9 kb genomic fosmid protein reporter
324 to the shorter 4.1kb upstream ("promoter") reporter (Figure 5A, B, C; Figure S8). In general, the
325 patterns were similar; both are expressed broadly in the MS lineage and in multiple anterior ABa-
326 derived sublineages. In the ABa lineage, the promoter reporter shows background expression in
327 some posterior sister sublineages suggesting the existence of repressive sequences outside of this
328 promoter region. Additionally, the promoter fusion reporter is detected persistently in both the MS
329 and ABa lineages, whereas protein reporter expression is detected transiently, likely reflecting the
330 use of a stable histone-mCherry reporter for the promoter reporter. Importantly, both the REF-
331 2::GFP protein and *ref-2* promoter reporters drive expression in progenitors of most cells defective
332 in *ref-2* mutants.

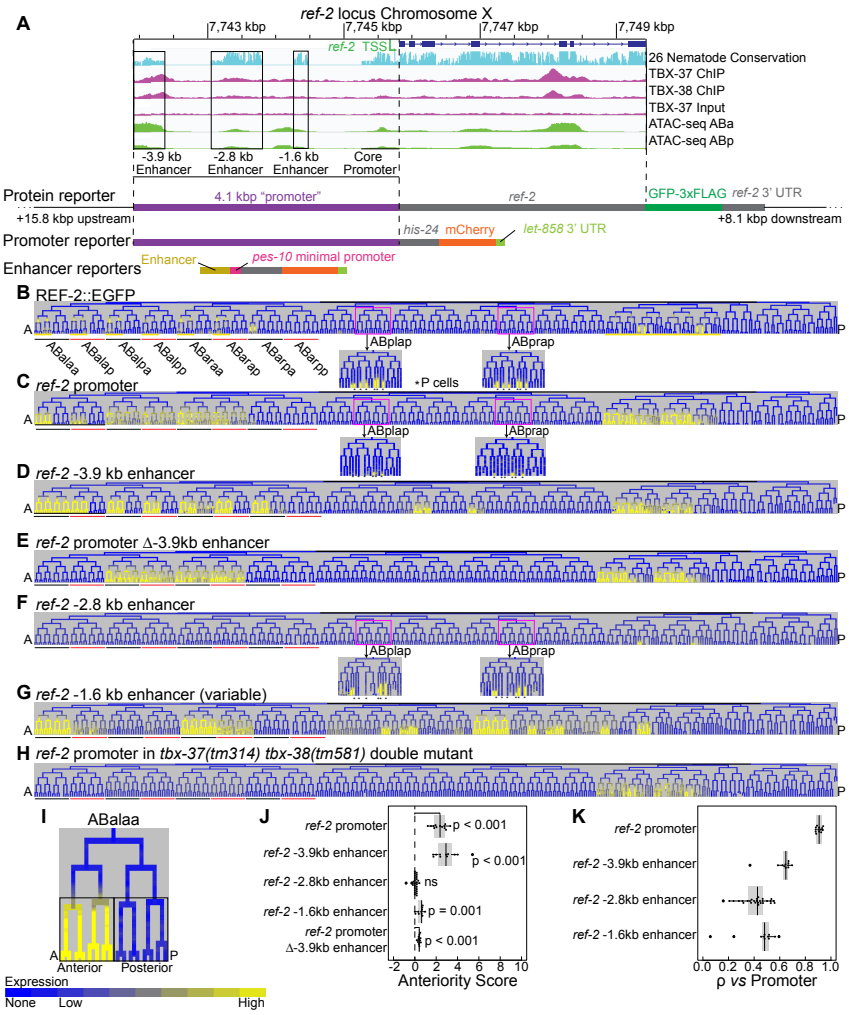


Figure 5: Identification of *ref-2* enhancers. A) Genome browser screenshot showing the *ref-2* locus on the X chromosome. Displayed tracks include the gene model, sequence conservation among 26 *Caenorhabditis* species [49], ChIP-seq data for TBX-37 and TBX-38 binding, and ATAC-seq data for cells in 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 Figure 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.

353 To identify the sequences in this region that regulate expression in these lineages, we sought to
354 identify minimal portions of the region with enhancer activity. We identified three portions of the
355 *ref-2* promoter conserved among *Caenorhabditis* species, suggesting they encode important
356 functions (Figure 5A) [49,51,52]. The most distal of these regions is in open chromatin in ABa-
357 derived cells as measured in a recently published ATAC-seq data set [50]. Also, this region is
358 bound by the broadly-expressed ABa lineage transcription factors TBX-37 and TBX-38, as
359 determined by a recently published ChIP-seq data set (Figure 5A) [50]. We fused each putative
360 enhancer to a reporter cassette consisting of a minimal *pes-10* promoter and a *his-24*::mCherry
361 reporter. The *pes-10* promoter is widely used, drives no consistent embryonic expression on its
362 own [13,53,54] and is compatible with a wide variety of enhancers [13,53]. We introduced each
363 reporter into worms and used StarryNite to determine in which cells each putative enhancer drives
364 expression.
365
366 The most distal enhancer, located 3.9 kb upstream of the likely transcription start site, (-3.9 kb
367 enhancer; 449 base pairs) drives early expression (during gastrulation) in several anterior
368 sublineages of ABa and in anterior sublineages of MS (Figure 5D; Figure S8). This expression
369 occurs in nearly identical lineages to the early portion of the protein reporter expression pattern,
370 but while the protein reporter expression in these lineages is transient, the enhancer reporter
371 persists, again likely because of the use of a stable histone-mCherry reporter. The -3.9 kb enhancer
372 expression is also well correlated with the pattern driven by the full promoter of *ref-2* (Spearman's
373 $\rho > 0.6$, Figure 5K). Deleting this enhancer from the *ref-2* promoter reporter results in loss of the
374 strong anterior-biased expression in the ABalaa, ABalap, and ABarpa lineages, and a loss in the
375 weak anterior-biased expression in the ABalpa, ABalpp, ABaraa, and ABarap lineages. All
376 remaining consistent expression occurs in the lineages derived from the Notch-signaled cells
377 ABalp and ABara and has little anterior bias (Figure 5E, J; Figure; S7; Figure S8) [48]. The
378 significance of this expression is unclear; it can be seen in the promoter reporter but not the protein
379 reporter, suggesting that it is normally repressed by additional sequences outside the promoter. We
380 conclude that the -3.9 kb enhancer is necessary and sufficient for anterior-biased expression in the
381 ABa lineage.
382

383 A second 745 bp enhancer, located at -2.8 kb, drives expression in later embryos (bean stage) in
384 some anterior sublineages of MS and in the ABp-derived Pn ventral epidermal blast cells that were
385 previously reported to require *ref-2* [39]. These cells also robustly express the *ref-2* protein reporter
386 and weakly express the *ref-2* promoter reporter (Figure 5B, C, F, J, K). Finally, the most proximal
387 enhancer candidate (at -1.6 kb; 206 base pairs) drives variable, weakly anterior-biased, expression
388 in multiple lineages, most consistently in ABala and ABara (Figure 5G, J, K; Figure S9) and also
389 drives late expression in some ABp sublineages (Figure 5G; Figure S9). Much of this expression
390 is in cells or at stages that do not express the full-length reporters, suggesting that the activity of
391 this sequence differs in its normal genomic context.
392
393 To measure anterior expression bias, we developed an "anteriority score" based on the log2 ratio
394 between anterior and posterior sister lineage expression. Because expression tends to accumulate
395 over several cell cycles in each expressing lineage, we calculated this score based on the mean
396 expression of the progeny 2-3 cell divisions after the birth of each anterior lineage (Figure 5I).
397 Using this metric, we detected robust and significant anterior-biased expression for both the *ref-2*
398 promoter and -3.9 kb enhancer reporters in ABa-derived sublineages, and reduced or no anterior-
399 biased expression for the -2.8 kb and -1.6 kb enhancer reporters and for the promoter Δ -3.9 kb
400 enhancer reporter in the same sublineages (Figure 5J). Comparing the expression driven by each
401 enhancer to the full promoter showed each is positively, but imperfectly, correlated, consistent
402 with each driving a subset of the full pattern (Figure 5K).
403
404 ***ref-2* promoter expression requires the T-box transcription factors *tbx-37* and *tbx-38***
405
406 A dominant feature of the early *ref-2* expression pattern is reiterated expression in six of the eight
407 anterior sister cells of the ectodermal ABa lineage at the 50-cell stage (AB32 stage--when there
408 are 16 ABa descendants). This expression is first observed in the fluorescent reporter strains in
409 pairs of sister cells in the following cell generation (AB64 stage--when there are 32 ABa
410 descendants). scRNAseq data are consistent with mRNA being expressed in the parent cell, with
411 reporter detection presumably delayed by the time required for the fluorophores of mCherry or
412 GFP to mature (Figure S5). This raises the question of whether *ref-2* expression requires the ABa-
413 specific transcription factors *tbx-37* and *tbx-38*. These genes encode redundant paralogous T-box

414 family transcription factors which are expressed throughout the ABa lineage and are required for
415 multiple cell fate decisions within ABa [46,50,55]. Confirming previous reports [46,50], a TBX-
416 38::GFP knock-in reporter made by CRISPR [50] is expressed throughout the ABa lineage and is
417 detectable at least from the AB16 to AB128 stages (Figure S10) (*tbx-37* and *tbx-38* RNA are
418 detectable by the AB4 stage [56]). To test whether *tbx-37* and *tbx-38* are necessary for *ref-2*
419 expression, we measured expression of the *ref-2* promoter reporter in embryos carrying
420 homozygous deletions of both *tbx-37* and *tbx-38*. We found that in the absence of *tbx-37* and *tbx-*
421 *38,* nearly all *ref-2* promoter expression in the ABa lineage is lost (Figure 5H; Figure S11), while
422 expression in the MS lineage was maintained. Thus, *tbx-37* and *tbx-38* are necessary for the
423 anterior-biased expression of *ref-2* in ABa.
424
425 **The -3.9 kb enhancer contains three non-overlapping sub-enhancers that independently**
426 **drive anterior expression in the ABa lineage**
427
428 Since the *ref-2* -3.9 kb enhancer drives expression in a similar pattern to the early expression of
429 the full promoter and protein reporters, and its deletion from the *ref-2* promoter results in a
430 reduction of anterior-biased expression in portions of the ABa lineage, it is a primary enhancer
431 responsible for driving the early anterior-biased expression of *ref-2*. To map features necessary or
432 sufficient for anterior expression, we tested the activity of eight truncated versions of the *ref-2* -
433 3.9 kb enhancer to identify minimal regions of the enhancer sufficient to drive anterior expression
434 in the ABa lineage (Figure 6, Figure S12). Instead of a single minimal region, we found three non-
435 overlapping regions within the -3.9 kb enhancer that each are sufficient to drive anterior-biased
436 enhancer activity in the ABa lineage: a promoter-distal region (bases 1 to 200), a medial region
437 (bases 227 to 304) and a promoter-proximal region (bases 320 to 449) (Figure 6B, C). Of these
438 regions, the proximal region drives the most consistent expression and this expression is most
439 similar to the full length -3.9 kb enhancer construct (Figure 6D), whereas the medial and distal
440 fragments drive more variable expression and in a subset of these lineages. Thus, although
441 integration of information between these sub-enhancers is likely necessary for robustness of
442 expression pattern, the 130 bp proximal fragment provides a model system for further dissection
443 to identify anterior expression regulators.
444

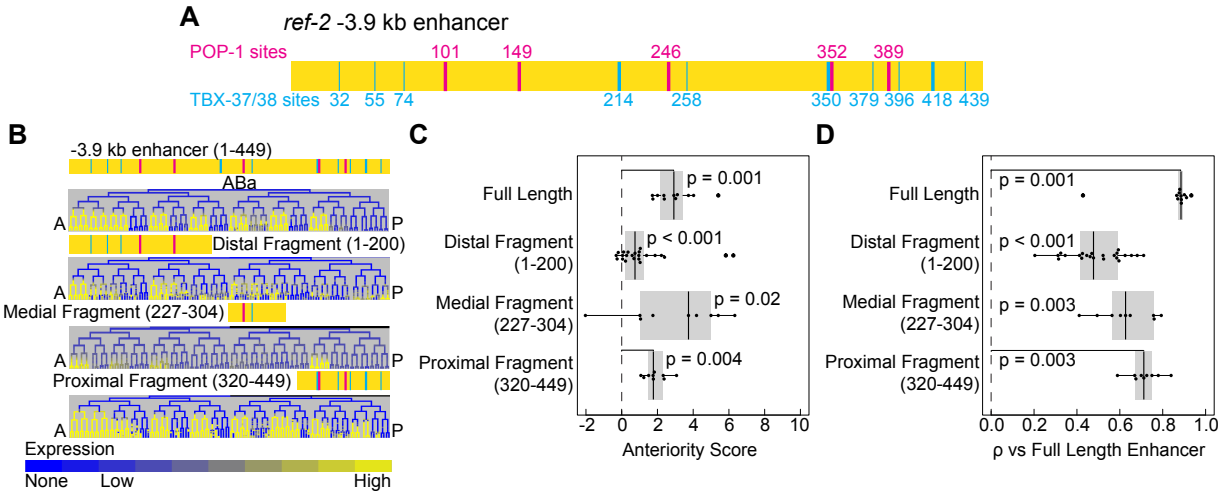


Figure 6: Three non-overlapping fragments of the -3.9 kb enhancer are each sufficient to drive anterior-biased expression in the ABa lineage of the early embryo. A) Model of the *ref-2* -3.9 kb enhancer with predicted TBX-37/38 sites (cyan, predicted high affinity sites indicated by thick line and predicted 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 Figure 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 Figure S12.

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 (Figure 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 (Figure S13B, 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 (Figure 7B; Figure S13B). 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 (Figure S13B). 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.

472

473 The proximal region includes five predicted TBX-37/38 sites, two of which are predicted to have
474 high affinity for TBX-37 and TBX-38. Two of the sites overlap with either the 410-424 region that
475 is essential for ABa-specific expression or the 425-449 region that is required for a subset of ABa-
476 specific expression (Figure 7A, B; Figure S13A, B). To determine whether this region is bound by
477 TBX-37 and TBX-38 *in vivo,* we mined a recently published dataset that measured genome-wide
478 TBX-37 and TBX-38 binding by ChIP-seq. Both factors show binding in the proximal region of
479 the -3.9 kb enhancer (Figure 5A) [50].
480

481 To determine whether the TBX-37/38 sites are required for anterior expression, we measured the
482 activity of the proximal fragment after mutating the two central nucleotides of all TBX-37/38 sites.
483 The mutated enhancer fragment did not drive anterior-specific expression in the ABa lineage
484 (Figure 7B, C, D; Figure S13B, C, D; Figure S14), confirming the importance of these sites.
485 Additionally, the loss of anterior-biased expression following the mutation of these sites indicates
486 that the two predicted high affinity POP-1 sites in this fragment are not sufficient to drive anterior-
487 biased expression. Also, these sites are not necessary for the anterior-biased expression driven by
488 this enhancer fragment, as the loss of these sites does not result in a loss of the anterior-biased
489 expression (Figure S13B, C, D). In fact, concatemers of POP-1 sites are sufficient to drive
490 posterior-biased expression (Figure S13F) [13,53].
491

492 Since the high affinity site beginning at position 418 (TBX_{418}) overlaps the region (410-424)
493 required for expression, we hypothesized that this specific TBX-37/38 site may be required for
494 anterior expression in the ABa lineage. To test this, we tested the activity of the proximal fragment
495 with only the central two nucleotides of TBX_{418} mutated (Figure 7B, Figure S13B, E). This
496 mutation resulted in a loss of robust anterior-biased expression driven by the proximal fragment.
497 Therefore, TBX-37 and TBX-38 bind to the *ref-2* -3.9 kb enhancer proximal fragment and TBX_{418}
498 is necessary for it to drive robust anterior-biased expression (Figure 5A; Figure 7B, C, D; Figure
499 S13B, C, D; Figure S14), suggesting that TBX-37 and TBX-38 directly regulate *ref-2* expression.
500

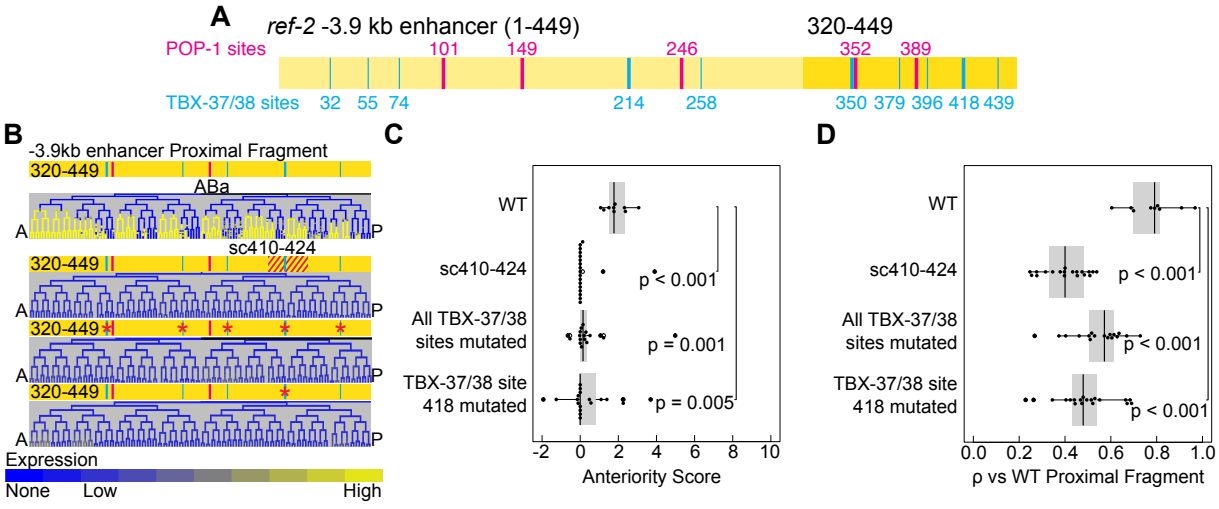


Figure 7: The ABa-expressed transcription factors *tbx-37* and *tbx-38* are required for the anterior-biased expression of *ref-2* in the ABa lineage. A) Model of the *ref-2* -3.9 kb enhancer with proximal fragment (base pairs 320-449) highlighted. Predicted TBX-37/38 and POP-1 sites are indicated as in Figure 6. B) Expression patterns driven by the proximal fragment of the *ref-2* -3.9 kb enhancer with its WT sequence, with base pairs 410-424 scrambled, with all predicted TBX-37/38 sites mutated, and with TBX_{418} mutated. Expression driven by the proximal fragment with additional regions scrambled or with two predicted POP-1 sites mutated are displayed in Figure S13. C-D) Anteriority scores (C) and Spearman's ρ (D) for the proximal fragment with its WT sequence, 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 Figure 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.

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 (Figure 8A). Surprisingly, both of these reporters drove anterior-specific expression in the ABa lineage similar to that driven by the full proximal fragment (Figure 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

528 for anterior-specificity, we scrambled those sequences while leaving the central TBX_{418} site intact
529 in the context of the long-flanking-sequence construct. The resulting reporter was again expressed
530 only in anterior sublineages similar to the full proximal fragment (Figure 8A, B, C, D). Thus, when
531 multimerized, TBX_{418} is sufficient both for ABa expression and for anterior specificity without
532 additional non-overlapping sequences.
533
534 **Anteriorly-biased expression of concatemerized TBX sites is regulated by *pop-1***
535
536 *ref-2* anterior-specific expression is dependent on *pop-1* and *sys-1*, and a concatemerized TBX-
537 37/38 binding site recapitulates part of this expression. This raises the question of whether the
538 TBX-37/38 binding site expression is *pop-1*-dependent. To test this, we examined the expression
539 of the TBX_{418} concatemer with scrambled flanking regions after *pop-1* RNAi. Knockdown of *pop-*
540 *1* results in changes in the expression pattern, including a significant decrease in anterior
541 expression bias in the ABalap lineage and a significant decrease in expression level in the ABalaa
542 lineage (Figure 8E, F, G). We also tested the expression effects of *lit-1* RNAi, which causes an
543 increase in nuclear POP-1 levels in posterior cells [16,57,58]. Knockdown of *lit-1* results in a
544 significant decrease in the anterior bias of the TBX_{418} concatemer expression across several
545 sublineages of ABa, including a significant increase in expression in the ABalap lineage (Figure
546 8E, F, G). Some of these changes mirror the fate transformations known to occur after *pop-1* or
547 *lit-1* RNAi, but others, including the increase in ABalap expression after *lit-1* RNAi, do not
548 [14,59]. Thus, our results indicate that either POP-1 or its targets influence both the lineage
549 specificity and anterior specificity of expression driven by TBX_{418}.
550

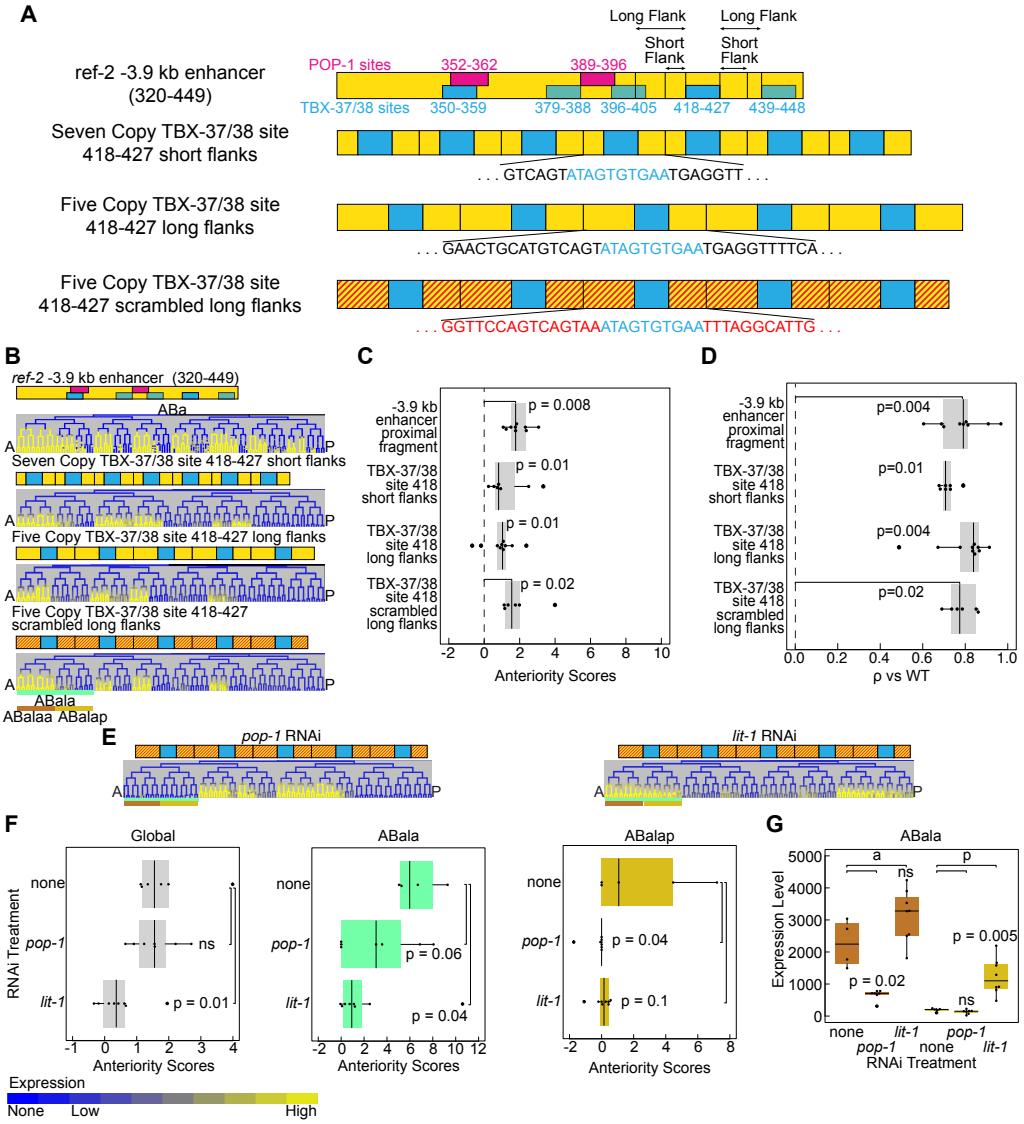


Figure 8: *TBX$_{418}$* in the *ref-2* -3.9 kb enhancer is sufficient when concatemerized to drive anterior-biased expression in the ABa lineage. A) Models of the proximal fragment of the *ref-2* -3.9 kb enhancer, the *TBX$_{418}$* seven-site concatemer with 6-7 base pair flanking regions, the *TBX$_{418}$* 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. E) Expression patterns driven by the *TBX$_{418}$* five-site concatemer with 11-15 base pair scrambled flanking regions after *pop-1* RNAi or *lit-1* RNAi. F) Anteriority scores for the *TBX$_{418}$* five-site concatemer with 11-15 base pair scrambled flanking regions in indicated lineages following indicated RNAi treatments. G) Quantification of reporter expression for the *TBX$_{418}$* five-site concatemer with 11-15 base pair scrambled flanking regions in the ABala lineage. Expression is shown under WT conditions and following *pop-1* or *lit-1* RNAi. Lineages used to determine global anteriority scores are the same as in Figure 5J. Spearman's ρ analysis uses the full ABa lineage expression pattern and was calculated relative to the WT proximal fragment.

568 **DISCUSSION**

569

570 Previous work has definitively shown that the Wnt pathway components POP-1 and SYS-1
571 regulate posteriorly expressed genes in the *C. elegans* embryo via the canonical Wnt asymmetry
572 pathway [12,13,36]. What has been less clear is how the similar number of anteriorly expressed
573 genes are regulated. We have demonstrated that a set of anteriorly expressed genes, *tbx-35*, *ceh-*
574 *51*, *elt-1*, *irx-1*, and *ref-2*, require POP-1 and SYS-1 for their anterior-biased expression. Therefore,
575 POP-1 and SYS-1 regulate the expression of anterior-biased genes. The mechanism by which
576 POP-1 and SYS-1 mediate this regulation remains to be determined. Previous work demonstrated
577 that posterior genes can be regulated by POP-1 binding directly to POP-1 binding sites in their
578 upstream regulatory regions [53]. Thus, a simple model of POP-1 regulation of anterior genes
579 would be indirect, with POP-1 activating the expression of repressors in posterior cells. For
580 example, *pop-1* RNAi is known to cause the anterior MS cell to adopt the fate of its posterior sister
581 E due in part to derepression in MS of the E-specific classic Wnt target *end-1*. *end-1* (or its targets)
582 could then act to repress expression of MS-specific genes. This mechanism would predict that
583 anterior genes would be expressed later than posterior genes, or that anterior genes would be
584 expressed early in both anterior and posterior cells, but maintained only in anterior cells. This can
585 happen. For example, in the C lineage *elt-1* is expressed in both anterior and posterior sister cells
586 at the 4C stage, but is only maintained in the anterior granddaughters of C (Caa and Cpa), perhaps
587 due to its repression by the Wnt target *hlh-1* in the descendants of the posterior granddaughters of
588 C (Cap and Cpp) [33]. What we observe in most cases, however, is that anterior gene expression
589 is activated at about the same time as posterior gene expression in sister lineages (Figures S1, S2,
590 S3, S4, S5, and S6, Zacharias et al. 2015 [13]). This observation suggests that POP-1 could regulate
591 some anterior genes directly. Some *pop-1* dependence of anterior genes certainly could be
592 explained by upstream anterior-to-posterior cell fate conversions following *pop-1* RNAi; however,
593 our results as a whole cannot be explained by the known pattern of fate transformations [46].

594

595 Indirect *pop-1*-dependent regulation of anterior genes does not explain all of its role in anterior
596 gene expression regulation. For instance, previous work showed that loss of muscle-specifying
597 TFs (MRFs) expressed in Cap and Cpp (*hlh-1*, *unc-120*, and *hnd-1*) is not sufficient to permit
598 ectopic expression of *elt-1* in these cells or to convert them to an epidermal fate [33,60]. Therefore,

599 another factor, perhaps *pop-1* acting directly, must be responsible for restricting *elt-1* to the
600 anterior C granddaughters. Our results identified a requirement for *pop-1* in anterior expression of
601 *elt-1* but not posterior repression in the C lineage. The proposed role of the MRFs in repressing
602 *elt-1* expression [33,60,61] suggests that this effect could partly be due to misexpression of MRFs
603 in the anterior C granddaughters, which could be tested in the future.
604
605 Direct activation of genes by POP-1/TCF when Wnt is absent occurs in other systems. For
606 example, in *Drosophila* TCF/Pangolin can act in a reverse transcriptional switch such that when it
607 is bound directly to DNA at non-optimal sites it can activate transcription when unbound by β-
608 catenin, and repress transcription when bound to β-catenin. This difference in the regulatory
609 activity of TCF is mediated by an allosteric conformational change in the structure of TCF [22].
610 In *C. elegans*, POP-1 bound together with REF-2 at the *ttx-3 cis*-regulatory region, may similarly
611 undergo allosteric changes, leading to the opposite regulation of target genes, such that POP-1
612 activates the expression of *ttx-3* when bound with REF-2 in the absence of SYS-1 and represses
613 the expression of *ttx-3* when bound with REF-2 and SYS-1 [20,21]. Our findings are consistent
614 with a similar mechanism regulating *ref-2* itself.
615
616 Our observation that five copies of a single 10 base-pair TBX-37/38 binding site are sufficient to
617 drive anterior-biased expression in the ABa lineage suggests that a POP-1 site is likely unnecessary
618 for anterior expression. However, *tbx-37* and *tbx-38* cannot explain this anterior bias alone, as
619 these factors are expressed throughout the ABa lineage. We showed that *pop-1* is required for both
620 anterior expression and posterior repression of the *ref-2* promoter and for anterior expression of a
621 TBX-37/38 binding site concatemer. If this role of POP-1 is direct, it likely acts by binding to other
622 factors, possibly TBX-37 and TBX-38, rather than to DNA directly since the concatemer contains
623 no sequences with high affinity for POP-1. Consistent with this, TBX-38 and POP-1 were
624 previously found to physically interact with each other in a yeast 2-hybrid assay [62]. Intriguingly,
625 several other T-box factors are expressed early in the AB lineage [50,56]; future work should
626 determine if these play a role in restricting the activity of these sites to anterior lineages. For
627 instance, TBX_{418} is predicted to be bound by several other *C. elegans* T-box factors. Other T-box
628 factors expressed in the early ABa lineage could compete for binding at this site, with POP-1
629 altering the relative activity of these factors between anterior and posterior cells. Additionally, it

630 is unclear whether *tbx-37* and *tbx-38* also regulate posterior gene expression in the ABa lineage;
631 intriguingly, the expression of *pha-4*, another anterior gene, in the ABa lineage is also dependent
632 on *tbx-37* and *tbx-38* [46].
633
634 In contrast with our observation that TBX_{418} concatemers are sufficient to drive anterior-biased
635 expression, Murgan et al. 2015 found that a REF-2 binding site concatemer is not sufficient to
636 drive expression in the anterior AIY mother cell. Instead, Murgan et al. identified binding sites for
637 helix-loop-helix family transcription factors that are sufficient to drive expression in the AIY
638 mother and its posterior sister cell, and that combining these with REF-2 binding sites restricts
639 expression to the AIY mother. Thus, REF-2 acting through ZIC binding sites is not sufficient to
640 activate expression without cooperating factors binding other sites whereas TBX-37 and TBX-38
641 may be sufficient to do so. Other factors could also play a role, for example we used a *pes-10*
642 minimal promoter, whereas Murgan et al. did not include a minimal promoter in their constructs,
643 and there could be differences in sensitivity between the microscopy techniques [20].
644
645 Classic work in many species, including flies and vertebrates, has found that developmental genes
646 are often regulated both by partially redundant enhancers and modular enhancers that are
647 responsible for regulating distinct portions of the genes' expression patterns [63,64]. The extent to
648 which *C. elegans* genes are regulated by such distal enhancers vs promoter proximal elements has
649 been unclear. We found that the predicted enhancers in the promoter region of *ref-2* function
650 modularly, such that each enhancer drives different expression patterns, presumably providing
651 multiple inputs to fine-tune the expression pattern of *ref-2*. Combined with other recent studies
652 this adds to evidence that enhancer-mediated regulation is widespread during *C. elegans*
653 embryonic development [54,65–69]. Additionally, we observe modularity within the -3.9 kb
654 enhancer, with three non-overlapping regions of the enhancer each sufficient to drive anterior-
655 biased expression. Since the protein reporter is expressed in more lineages than the promoter
656 reporter, there are likely even more enhancers outside the promoter region that regulate expression
657 of *ref-2*.
658
659 Several studies have identified evidence for multiple functions of *ref-2*. *ref-2* was first identified
660 as being required for the production of the Pn.p ventral epidermal cells and for the inhibition of

661 their fusion to the epidermal syncytium hyp7 [39]. *ref-2* is also required for the initiation of the
662 differentiation of the cholinergic neuron AIY [40]. Also, *ref-2* has been found to be required for
663 female fate during sexual development [70]. We have demonstrated a role for *ref-2* in embryonic
664 development. *ref-2* is required for robust WT cell positioning and cell division timing in anterior
665 lineages. Intriguingly, an insect ortholog of *ref-2*, *odd-paired*, acts as a regulator of anterior-
666 posterior expression and is an anterior determinant in several species [71,72], suggesting the
667 function of *ref-2* in anterior fate regulation may be ancestral. Because the defects in cell position
668 and division timing are only partially penetrant, there are likely other, partially redundant
669 regulators of cell position and cell division timing that act with the sole *C. elegans* ZIC homolog
670 *ref-2* during embryonic development. Further experiments, including suppressor and enhancer
671 screens and co-mutation of other transcription factors will need to be performed to identify other
672 factors that regulate cell position and division timing in conjunction with *ref-2*.
673
674 **METHODS**
675
676 ***C. elegans* culture and strain generation**
677
678 *C. elegans* strains (Table S1) were maintained at standard growth temperatures on OP50 *E. coli*
679 on NGM plates (Table S2). RNAi knockdown was performed by feeding, as previously described
680 [73,74]. We validated the efficacy of *pop-1* RNAi by measuring the cell cycle delay resulting from
681 the transformation of the anterior MS lineage into an E-like lineage, the efficacy of *sys-1* RNAi by
682 failure of morphogenesis and embryonic death, and the efficacy of *lit-1* RNAi by measuring the
683 cell cycle shortening resulting from the transformation of the posterior E lineage into an MS-like
684 lineage (Table S2) [13,34,59]. Enhancer reporter strains were generated by microinjection into
685 RW10029, the GFP histone strain used for lineage tracing. Injection cocktails used to create
686 extrachromosomal array transgenes consisted of a linear reporter DNA construct at 10 ng/μL, with
687 5 ng/μL *myo-2*p::GFP, and 135 ng/μL pBluescript vector (or highest concentration possible
688 whenever 135 ng/μL was impossible to make) and were injected using a Narishige MN-151
689 micromanipulator with Tritech microinjector system. Enhancer reporters were maintained in
690 extrachromosomal arrays by screening for worms expressing the *myo-2*p::GFP reporter, which
691 have a green fluorescent pharynx. Most enhancer reporters, including all for which loss of

692 expression was observed, were analyzed in at least two strains to control for differences in copy
693 number and extrachromosomal array artifacts. Other strains were created through crosses using
694 standard approaches (Table S1).
695
696 **Generation of transgenes**
697
698 Candidate enhancers were amplified by PCR from *C. elegans* N2 strain genomic DNA with
699 Phusion HF polymerase (New England Biosciences) and gel purified (Qiagen). Enhancer
700 fragments, mutated enhancers, and binding site concatemers were ordered as either gBlocks or
701 Ultramers from Integrated DNA Technologies (Coralville, Iowa) and were amplified with Phusion
702 HF polymerase (New England Biosciences) with overhangs for stitching and either gel or PCR
703 purified (Qiagen). Enhancer reporters were produced by fusing via PCR stitching these constructs
704 to a *pes-10* minimal promoter::HIS-24::mCherry::let-858 3'UTR fragment amplified from the
705 POPTOP plasmid [53] (Addgene #34848). The enhancer reporters were purified with a PureLink
706 PCR purification kit (ThermoFisher) and/or gel purified. The desired product was determined by
707 its size based on gel electrophoresis. We identified putative transcription factor binding sites using
708 CIS-BP (http://cisbp.ccbr.utoronto.ca) [75,76].To mutate the TBX-37/38 sites, the central two
709 nucleotides were altered (e.g. ATAG**TG**TGAA changed to ATAG**GT**TGAA for TBX_{418}). For the
710 construct with all TBX-37/38 sites mutated, after mutating the five primary sites, subsequent
711 analysis revealed two weaker sites. All of these predicted TBX-37/38 sites were mutated. The
712 reporter construct for the *ref-2* promoter lacking the -3.9 kb enhancer was amplified from N2
713 genomic DNA by PCR and PCR stitched to the *pes-10* minimal promoter::HIS-24::mCherry
714 reporter, and, thus, drives expression through the *pes-10* minimal promoter. The *pes-10* promoter
715 drives no consistent embryonic expression on its own prior to elongation [13,53,54] and is
716 compatible with a wide variety of enhancers [13,53] (Table S5).
717
718 **Quantitative comparisons**
719
720 All quantitative comparisons were performed using R version 4.0.3 (The R Foundation for
721 Statistical Computing), Microsoft Excel version 16.16.27, or Python version 2.7.16. For the main
722 figures, a minimum N of 6 was used for all experiments, except for ELT-1::GFP expression after

723 *sys-1* RNAi, for which N = 3. For the supplemental figures, a minimum N of 2 was used (largely
724 for perturbations which appeared similar to wild-type at this level). Code and raw data for these
725 analyses are available at https://github.com/jisaacmurray/ref-2_paper.
726
727 *Expression and phenotypic analysis by 4D imaging*
728
729 We obtained confocal micrographs using a Leica TCS SP5 or Stellaris scanning confocal
730 microscope (67 z planes at 0.5 µm spacing and 1.5 minute time spacing, with laser power
731 increasing by 4-fold through the embryo depth to account for attenuation of signal with depth).
732 Embryos obtained from self-fertilized hermaphrodites were mounted in egg buffer/methyl
733 cellulose with 20µm beads used as spacers and imaged at 22°C using a stage temperature controller
734 (Brook Industries, Lake Villa, IL) [77]. We used StarryNite software to automatically annotate
735 nuclei and trace lineages [29]. We corrected errors from the automated analysis and quantified
736 reporter expression in each nucleus relative to the local background (using the "blot" background
737 correction technique) with AceTree software as previously described (Table S3) [23,25].
738
739 *Anterior gene WT to RNAi within lineage comparisons*
740
741 To compare the expression levels of the anterior genes *tbx-35*, *ceh-51*, ELT-1, *irx-1*, and *ref-2* and
742 the TBX_{418} concatemer with 11-15 base pair scrambled flanking regions between WT and RNAi
743 embryos within the expressing lineages and their posterior sisters, we calculated for each embryo
744 the mean reporter expression across all measurements within each lineage (all descendant cells
745 and time points) starting with the cell stage that expresses the reporter at the time cells are born or
746 shortly afterwards to the last time point in the indicated lineage. For expression level comparisons
747 of extrachromosomal array transgenes, we normalized expression values to the 75th percentile
748 expression value of the strain with the higher 75th percentile expression value across all measured
749 cells in the indicated lineage within each RNAi treatment. The expression analyzed for the MS
750 lineage starts at an earlier stage than that for other lineages, since following *pop-1* RNAi the cell
751 divisions are so dramatically delayed that reporter expression comes on in earlier cell generations
752 in these embryos. The cell generations used in each anterior gene comparison are noted in Table
753 S4.

754

755 *Anteriority Scores*

756

757 To measure the magnitude of anterior expression bias, we developed a measure that we refer to as
758 the anteriority score. To determine the anteriority scores, first we chose a minimum expression
759 cutoff and set all values below this cutoff to 0 and subtracted the minimum cutoff from all
760 expression values greater than or equal to the cutoff (an expression value of 200 was used as the
761 cutoff for all anteriority score analyses, except a value of zero was used for the TBX-38::GFP
762 reporter, which has very low fluorescence intensity compared to the other reporters analyzed).
763 Next, we averaged the expression levels of the reporter of interest in the cells of each anterior and
764 each posterior sister lineage of interest. For the *ref-2* promoter and enhancer constructs, as well as
765 the TBX-37/38 site concatemers, we analyzed anterior and posterior lineages descended from the
766 cells ABala, ABalaa, ABalap, ABalpa ABalpp ABaraa, ABarap, ABarpa, and ABarpp. We used
767 expression values starting at the generation with ABaxxxxxx cells (AB128 stage) until the time
768 point of the last division of AB128 stage cells in the ABa lineage. Next, we calculated the means
769 of the average expression levels of all of the anterior lineages and all of the posterior lineages
770 analyzed for each embryo. We then added a pseudocount of 1 to each of the anterior and posterior
771 embryo means. For each experimental group, the adjusted average anterior expression is divided
772 by the adjusted average posterior expression, and the base 2 logarithm of the result is reported--
773 Anteriority Score = $\log_2$((anterior mean + 1)/(posterior mean +1)).

774

775 *Correlation analysis*

776

777 Correlation analysis was done by Spearman's ρ, comparing each embryo's expression to the
778 average expression of the control group. Spearman's ρ was calculated in R using the "cor"
779 function. For Spearman's ρ we used all the data from the lineages of interest without excluding
780 any cell generations or having a minimum value cutoff. We included cells born before the onset
781 of expression here to include information about the cell generation of expression onset in the
782 analysis. For the analysis comparing the *ref-2* promoter with its enhancers we used the full somatic
783 lineage starting with the blastomeres AB, EMS, C, and D. For all other analyses we used only the
784 ABa lineage.

785

786 *Mutant cell position and cell division analysis*

787

788 We identified cell position and cell cycle defects as previously described [41]. We corrected for
789 differences in global division rates, and considered divisions as defective if they deviated from the
790 wild-type cell cycle length by at least five minutes and had a z-score greater than three. Cell
791 positions were corrected for differences in embryo size and rotation, and considered defective if
792 they deviated from the expected wild type position by at least five microns, had a z-score greater
793 than five, and a nearest neighbor score greater than 0.8.

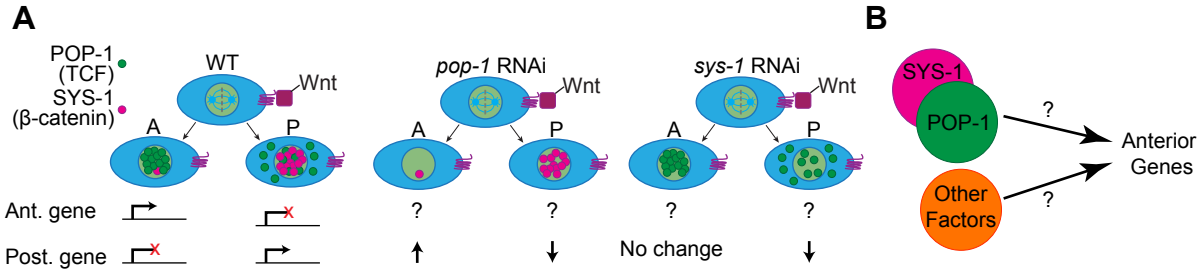
794

795 *Statistical analyses*

796

797 Statistical significance of differences in anterior or posterior expression between untreated and
798 RNAi-treated embryos was tested using the two-tailed Wilcoxon Ranked Sum test. Statistical
799 significance of differences in Spearman's ρ or Anteriority Score were tested using the one-tailed
800 Wilcoxon Ranked Sum test to test for differences from either the control group or from 0, as
801 indicated in each figure. The p-values calculated by the Wilcoxon Ranked Sum test were corrected
802 for multiple comparisons using the false discovery rate method in R. Significance of overlap
803 between positional or cell division defects and expression status were assessed using a Chi-squared
804 test.

805 **FIGURE LEGENDS**
806



807
808
809 **Figure 1: How are anterior genes regulated?** A) The Wnt/β-catenin asymmetry pathway
810 regulates anterior-posterior patterning by producing asymmetric nuclear concentrations of POP-
811 1/TCF and SYS-1/β-catenin in anterior and posterior sister cells. Low nuclear concentrations of
812 POP-1 and high nuclear concentrations of SYS-1 in the posterior cell activate expression of
813 posterior genes and are associated with the lack of expression of anterior genes. Conversely, high
814 nuclear concentrations of POP-1 and low nuclear concentrations of SYS-1 in the anterior cell
815 repress posterior genes and are associated with expression of anterior genes. Effects of *pop-1* and
816 *sys-1* RNAi on posterior genes are shown; effects on most anterior genes are unknown. B) In this
817 work, we test whether *pop-1*/TCF and *sys-1*/β-catenin regulate anterior genes in the *C. elegans*
818 embryo, and ask what other factors also regulate anterior expression.
819

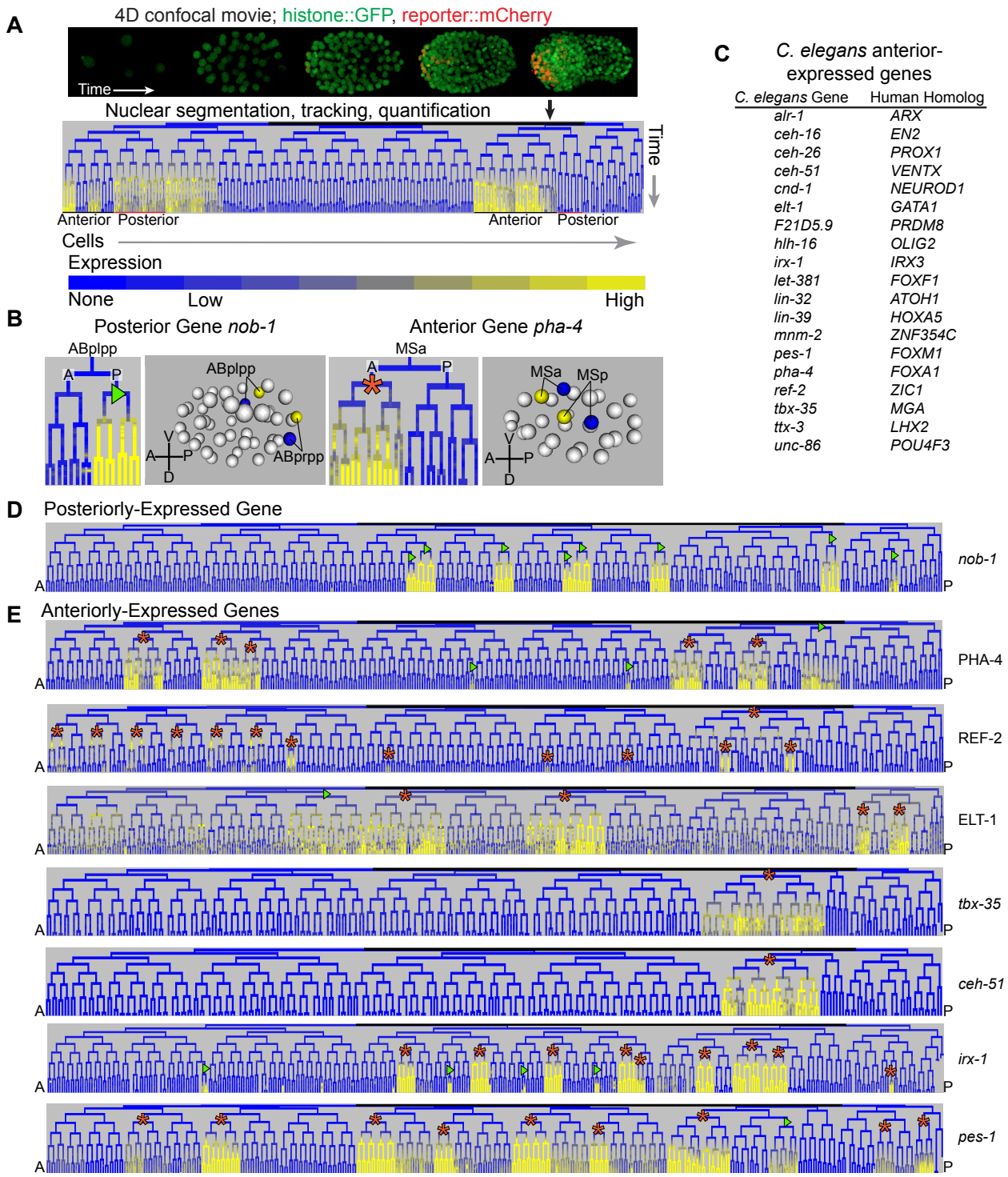


Figure 2: Identification of genes with anterior-biased expression. A) Automated lineaging traces lineages and quantifies expression from 4D confocal images. In the lineage trees, vertical lines represent cells, and horizontal lines represent cell divisions. Most cells divide along the anterior-posterior 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-

828 posterior cell division. Anterior and posterior genes can generally be expressed in cells from any
829 part of the embryo. Anterior and posterior founder cells are labeled with orange asterisks/green
830 triangles respectively. C) List of anterior-biased genes in the EPiC database along with their
831 predicted human homologs. D) Expression pattern of a posterior gene, *nob-1/Hox9-13*. E)
832 Expression patterns of seven anterior genes; most are expressed in unique combinations of mostly
833 anterior lineages.
834

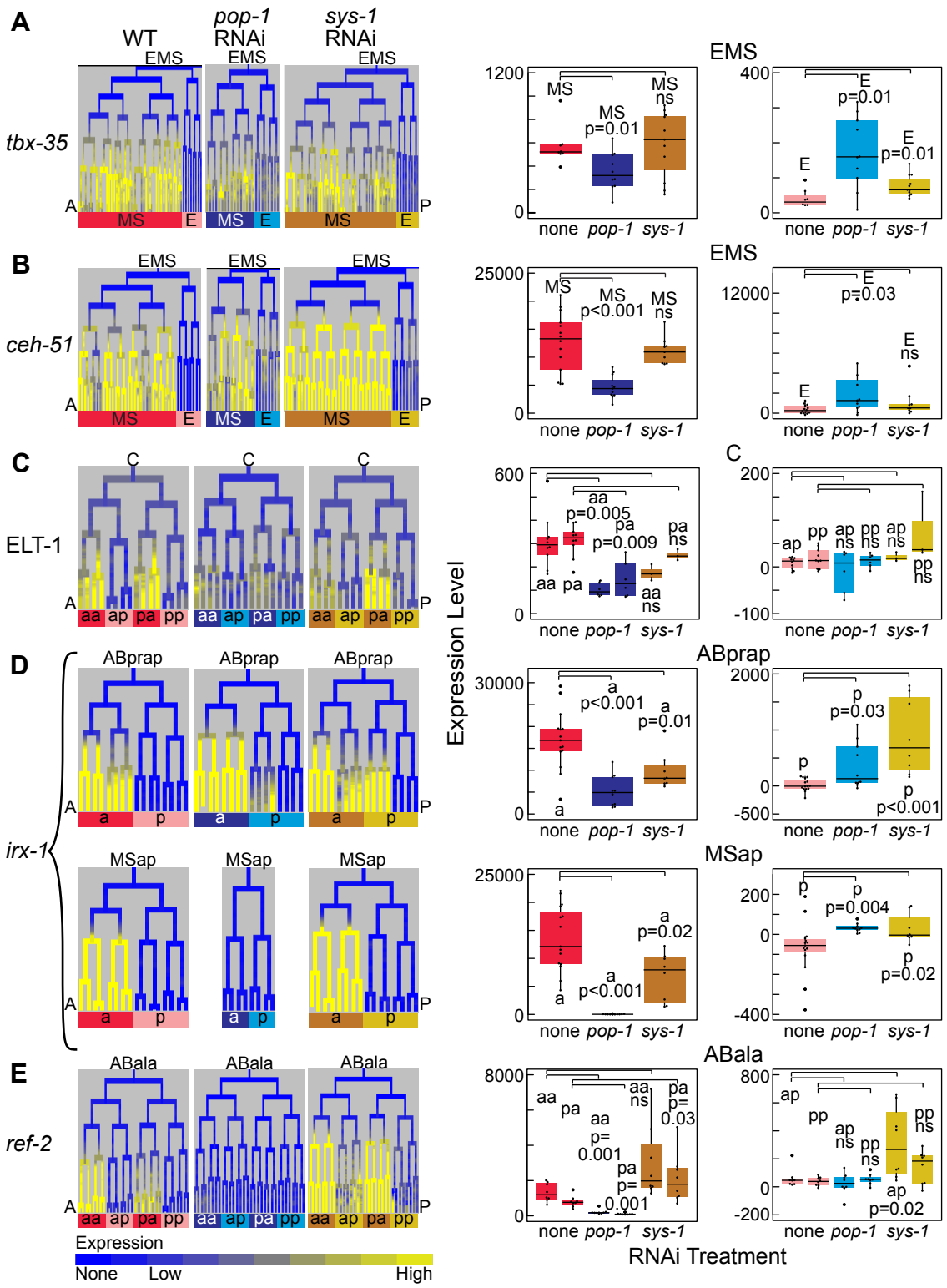


Figure 3: Regulation of anterior genes by Wnt effectors *pop-1* and *sys-1*. A-E) Expression pattern (left) and quantification (right) of reporters for the anterior genes *tbx-35* (A), *ceh-51* (B), ELT-1 (C), *irx-1* (D), and *ref-2* (E) in specific anterior lineages and their posterior sisters. Expression is shown under WT conditions and following *pop-1* or *sys-1* RNAi. Expression

841 quantification (right) is the mean expression across all measurements within that lineage from
842 when the reporter is normally first detected until the last time point indicated in the lineage (details
843 in Methods).
844

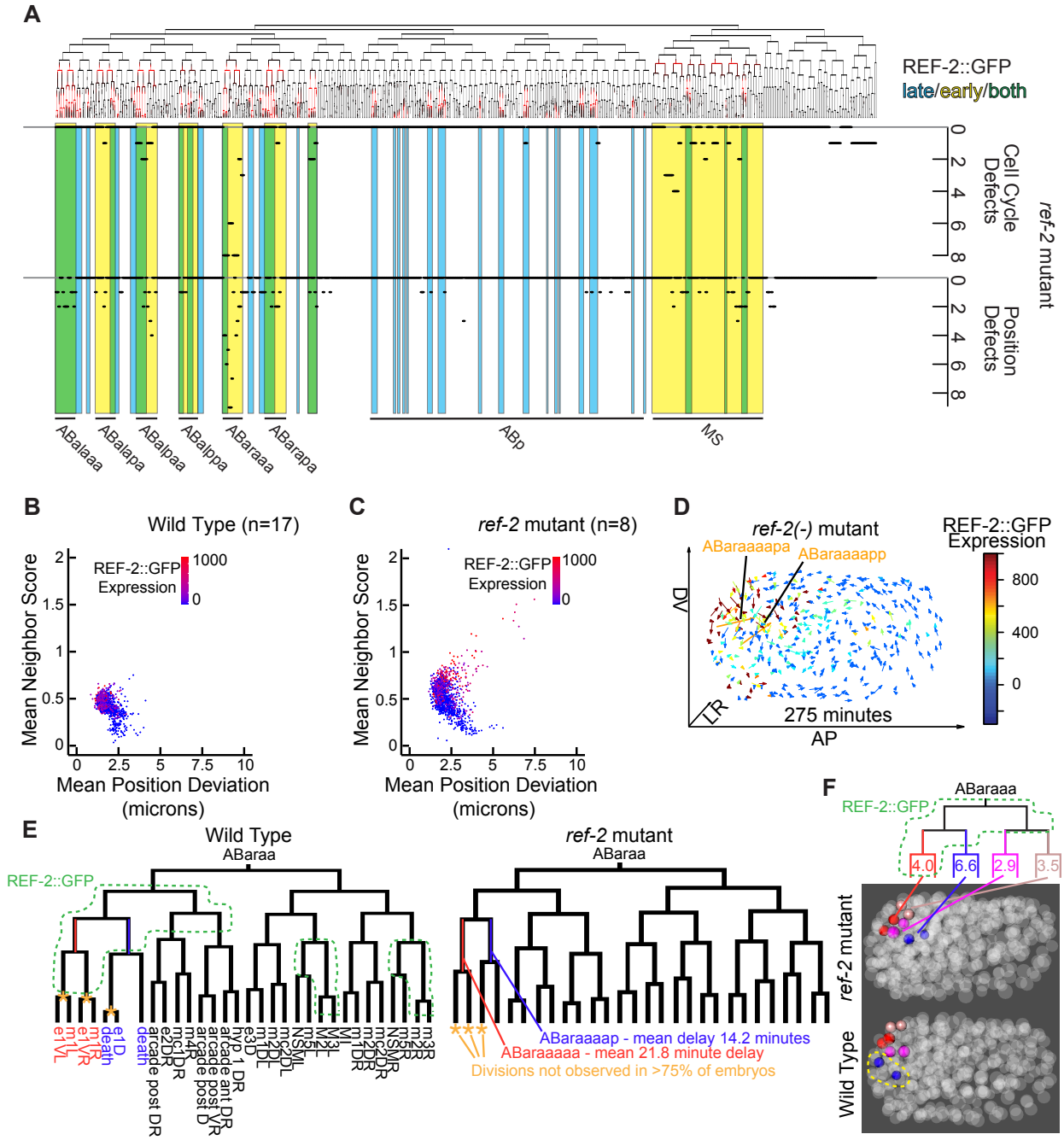


Figure 4: *ref-2* is required for proper division timing and positioning for many embryonic cells that are descended from cells that express *ref-2*. A) Lineage tree depicting the REF-2 protein expression pattern (top), and chart indicating the number of embryos out of eight homozygous *ref-2* mutant 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 out of all 17 embryos tested [41]. 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 [41]. 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.

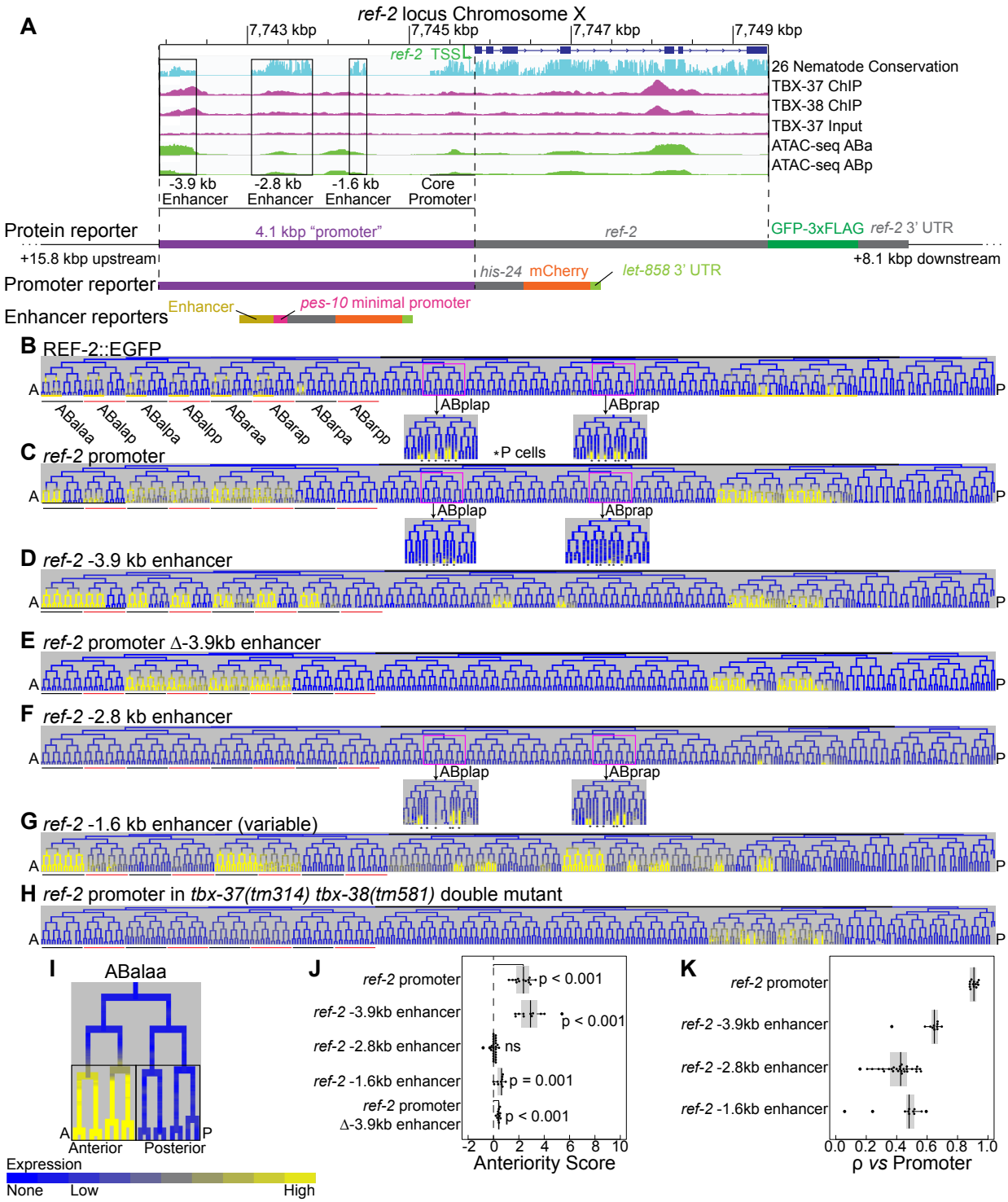


Figure 5: Identification of *ref-2* enhancers. A) Genome browser screenshot showing the *ref-2* locus on the X chromosome. Displayed tracks include the gene model, sequence conservation among 26 *Caenorhabditis* species [49], ChIP-seq data for TBX-37 and TBX-38 binding, and ATAC-seq data for cells in the ABa and ABp lineages [50]. Boxes indicate candidate enhancer regions as determined by conservation. Below the genome browser screenshot are models of the

876 fluorescent reporter expression constructs we examined. B-H) Lineage trees showing the
877 expression patterns of *ref-2* protein (B), promoter (C), -3.9 kb enhancer reporter (D), promoter
878 lacking the -3.9 kb enhancer (E), -2.8 kb enhancer reporter (F), -1.6 kb enhancer reporter (G), and
879 promoter in the *tbx-37 tbx-38* double mutant (H) as determined by time-lapse confocal microscopy.
880 Sublineages used in anteriority analyses are labeled in B and underlined in black (anterior) and red
881 (posterior) under each lineage tree. Sublineages with transient early REF-2::EGFP (protein
882 reporter) expression are underlined in yellow under the lineage tree in B. Pn ventral epidermal
883 cells are indicated with asterisks in insets of panels B, C, and F. I) Example lineage showing which
884 cells were included in anteriority score analyses. Reporter expression levels were averaged across
885 cells in the anterior or posterior lineages for the indicated cell generations. J) Anteriority scores of
886 the *ref-2* promoter, enhancers, and promoter lacking the -3.9 kb enhancer. Lineages used for
887 analysis are ABala, ABalaa, ABalap, ABalpa, ABalpp, ABaraa, ABarap, ABarpa, and ABarpp
888 (see Figure 5B). K) Full somatic lineage correlation (Spearman's ρ) of each reporter's expression
889 pattern with the mean expression pattern across embryos expressing the promoter reporter.
890

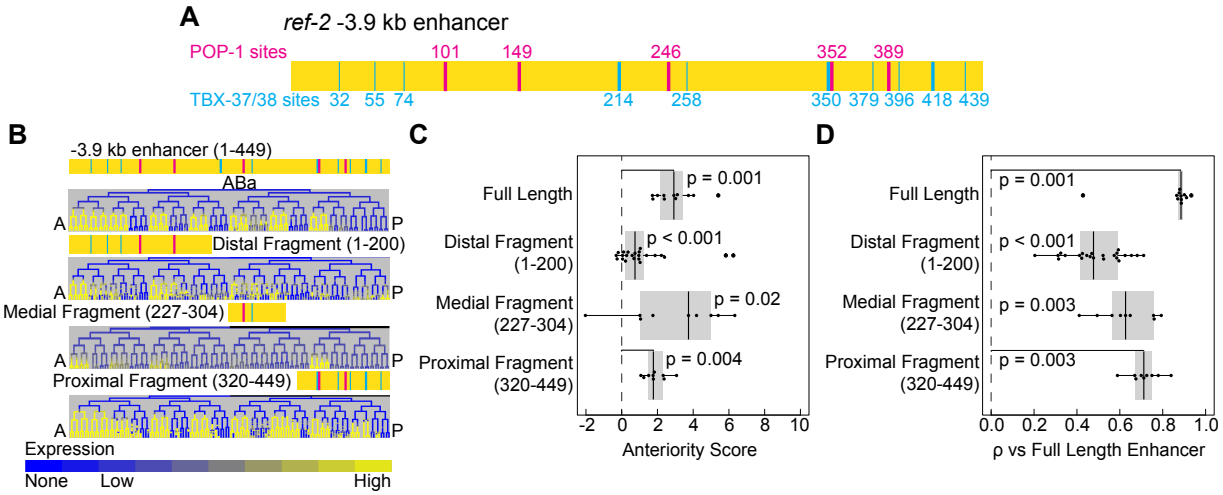


Figure 6: Three non-overlapping fragments of the -3.9 kb enhancer are each sufficient to drive anterior-biased expression in the ABa lineage of the early embryo. A) Model of the *ref-2* -3.9 kb enhancer with predicted TBX-37/38 sites (cyan, predicted high affinity sites indicated by thick line and predicted 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 Figure 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 Figure S12.

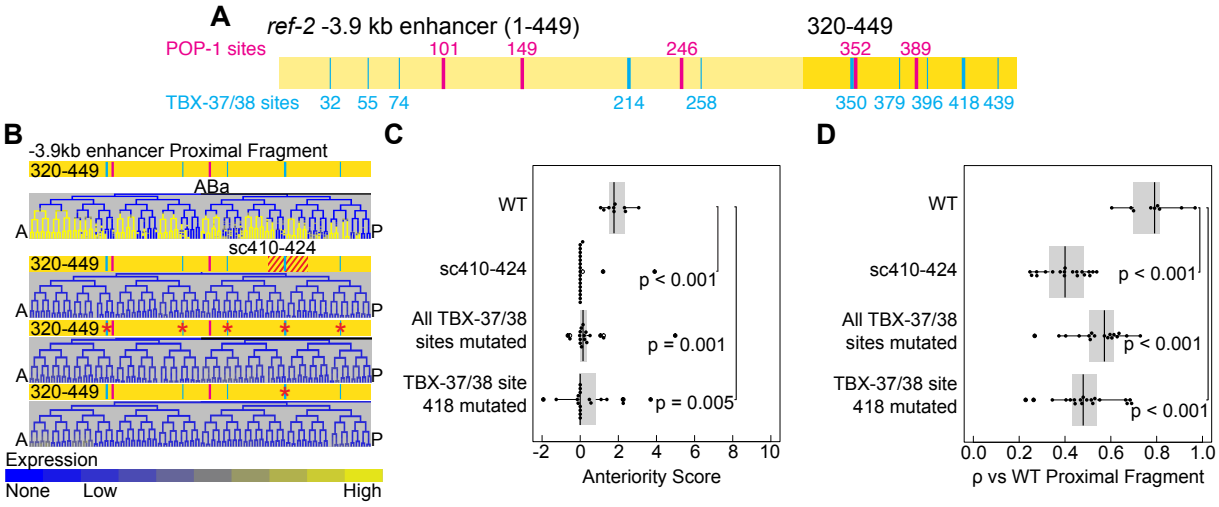


Figure 7: The ABa-expressed transcription factors *tbx-37* and *tbx-38* are required for the anterior-biased expression of *ref-2* in the ABa lineage. A) Model of the *ref-2* -3.9 kb enhancer with proximal fragment (base pairs 320-449) highlighted. Predicted TBX-37/38 and POP-1 sites are indicated as in Figure 6. B) Expression patterns driven by the proximal fragment of the *ref-2* -3.9 kb enhancer with its WT sequence, with base pairs 410-424 scrambled, with all predicted TBX-37/38 sites mutated, and with TBX_{418} mutated. Expression driven by the proximal fragment with additional regions scrambled or with two predicted POP-1 sites mutated are displayed in Figure S13. C-D) Anteriority scores (C) and Spearman's ρ (D) for the proximal fragment with its WT sequence, 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 Figure 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.

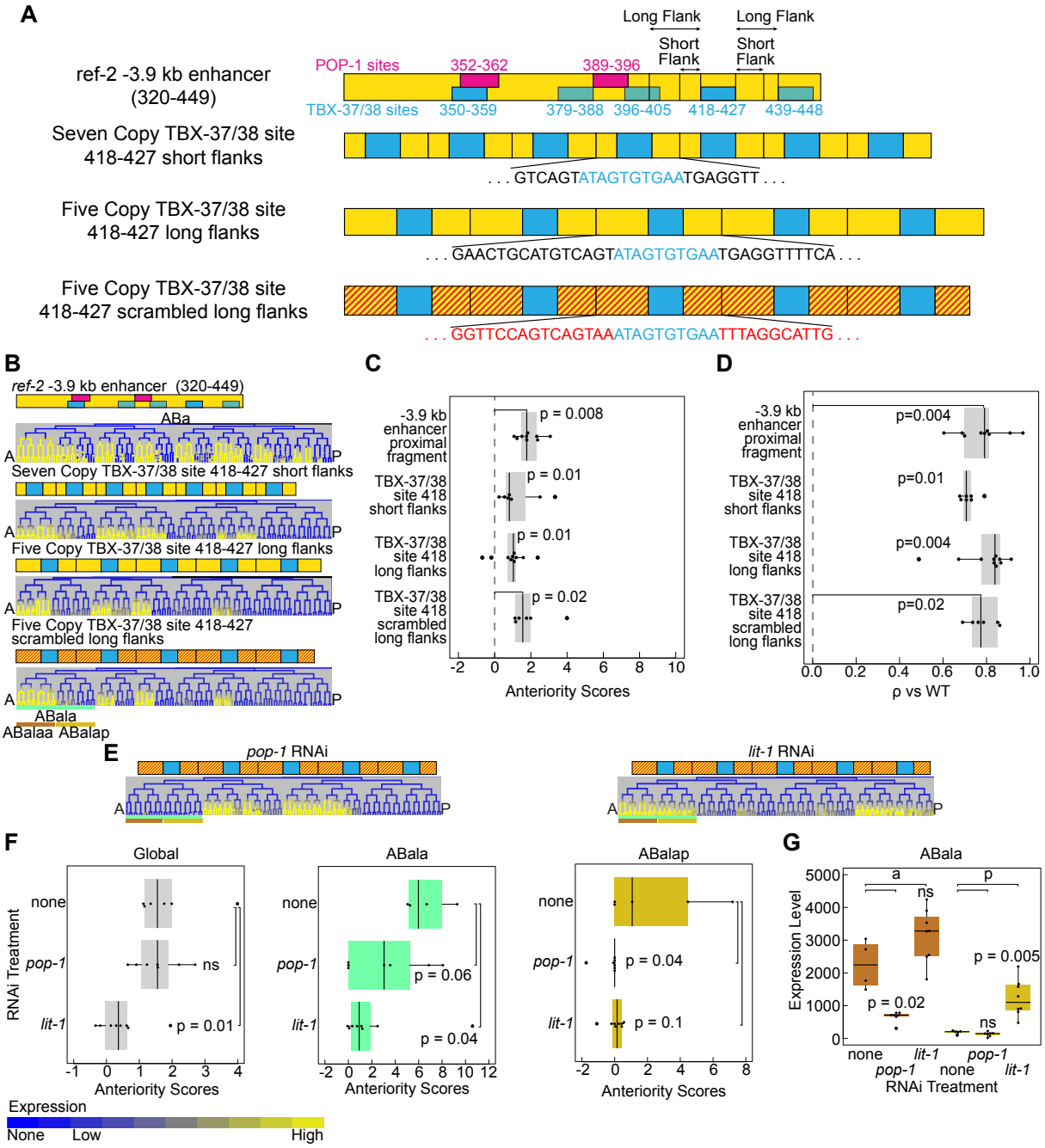
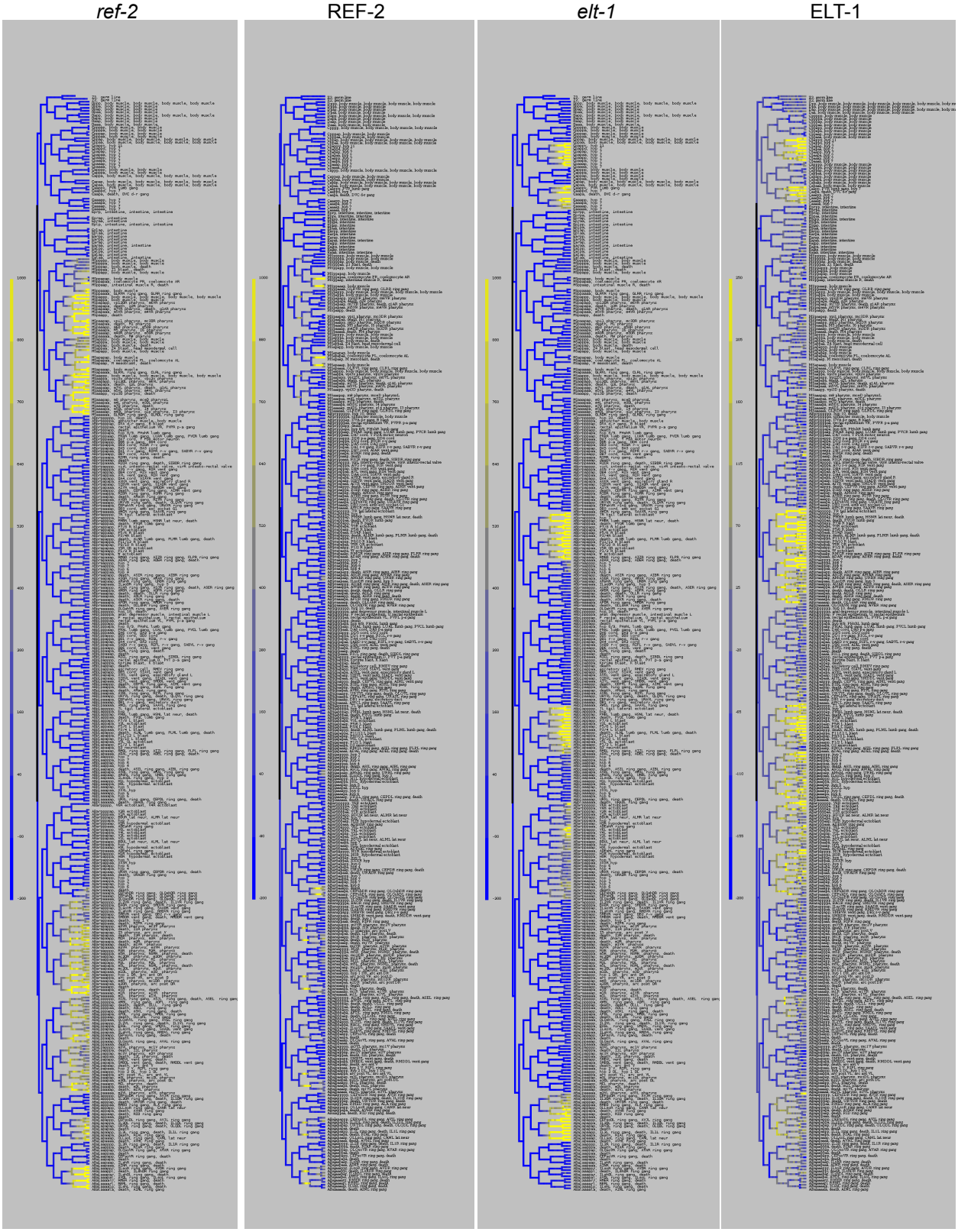


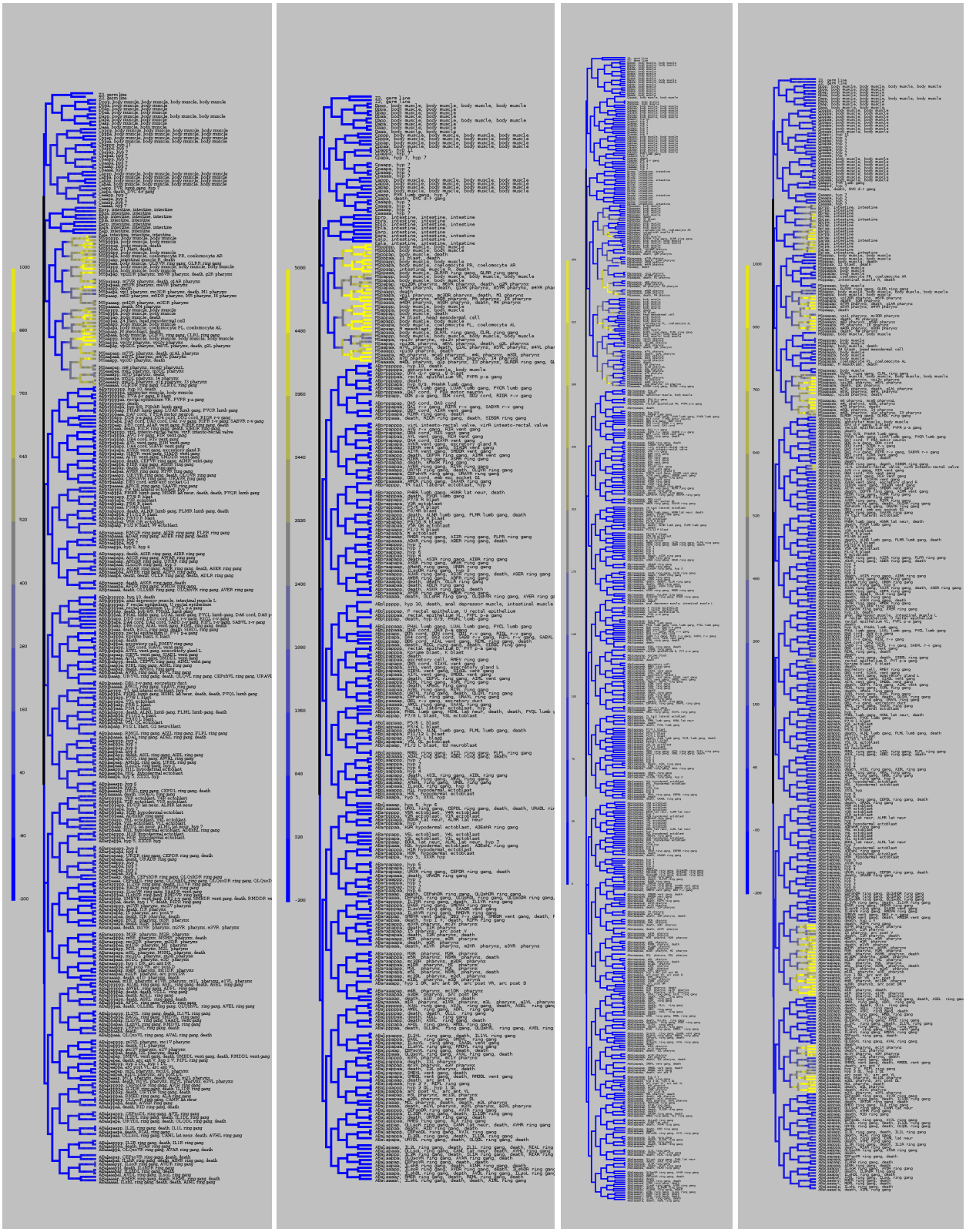
Figure 8: *TBX$_{418}$* in the *ref-2* -3.9 kb enhancer is sufficient when concatemerized to drive anterior-biased expression in the ABa lineage. A) Models of the proximal fragment of the *ref-2* -3.9 kb enhancer, the *TBX$_{418}$* seven-site concatemer with 6-7 base pair flanking regions, the *TBX$_{418}$* 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. E) Expression patterns driven by the *TBX$_{418}$* five-site concatemer with 11-15 base pair scrambled flanking regions after *pop-1* RNAi or

930 *lit-1* RNAi. F) Anteriority scores for the TBX_{418} five-site concatemer with 11-15 base pair
931 scrambled flanking regions in indicated lineages following indicated RNAi treatments. G)
932 Quantification of reporter expression for the TBX_{418} five-site concatemer with 11-15 base pair
933 scrambled flanking regions in the ABala lineage. Expression is shown under WT conditions and
934 following *pop-1* or *lit-1* RNAi. Lineages used to determine global anteriority scores are the same
935 as in Figure 5J. Spearman's ρ analysis uses the full ABa lineage expression pattern and was
936 calculated relative to the WT proximal fragment.
937


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

940 **Figure S1: Several early embryonic genes are expressed with anterior bias.** Full lineage
941 expression patterns are shown for integrated transgenic promoter reporters of *ref-2* and *elt-1*; and
942 for integrated transgenic protein reporters of REF-2 and ELT-1.

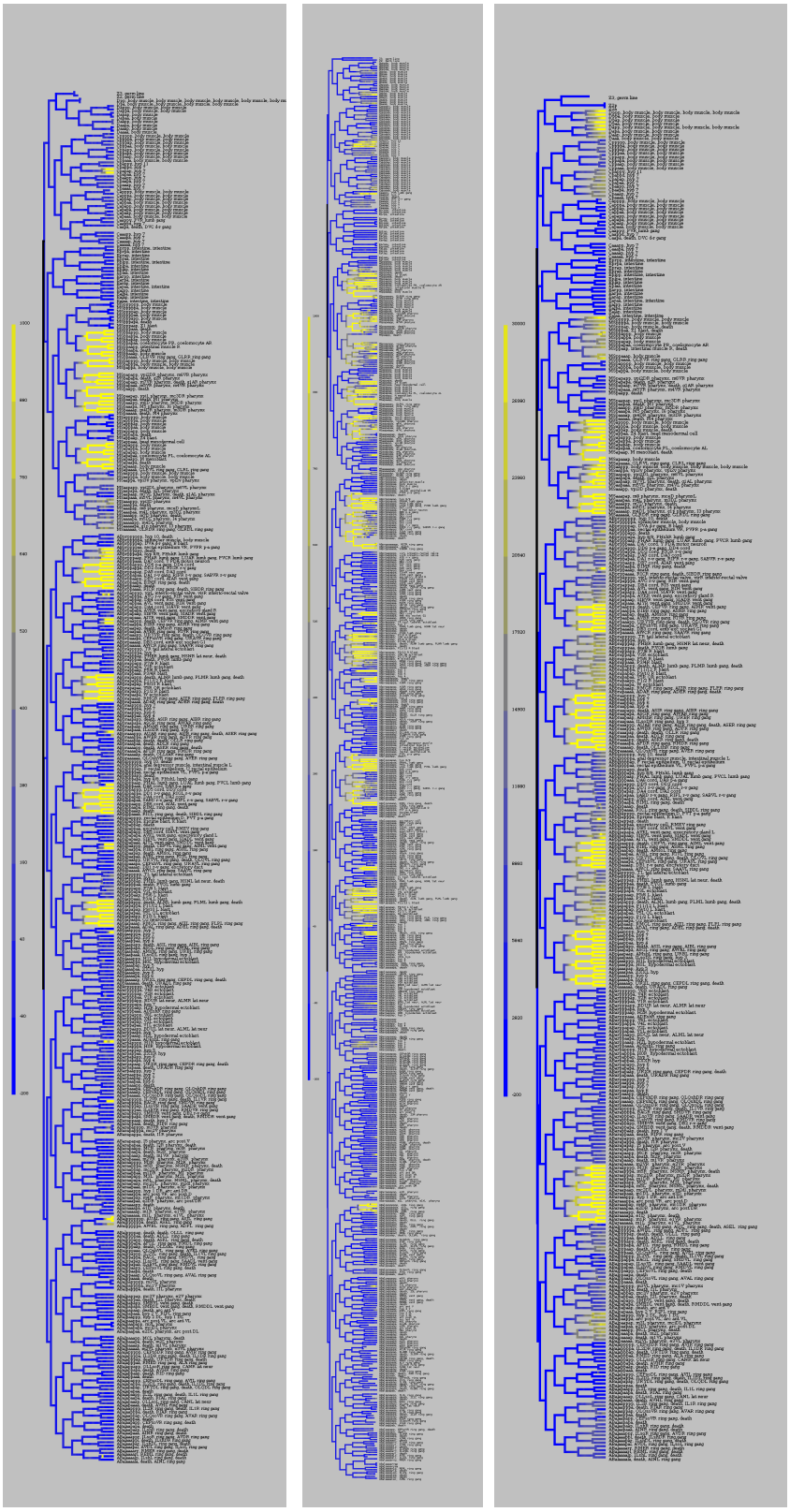
tbx-35 *ceh-51* CEH-51 PHA-4



943
944

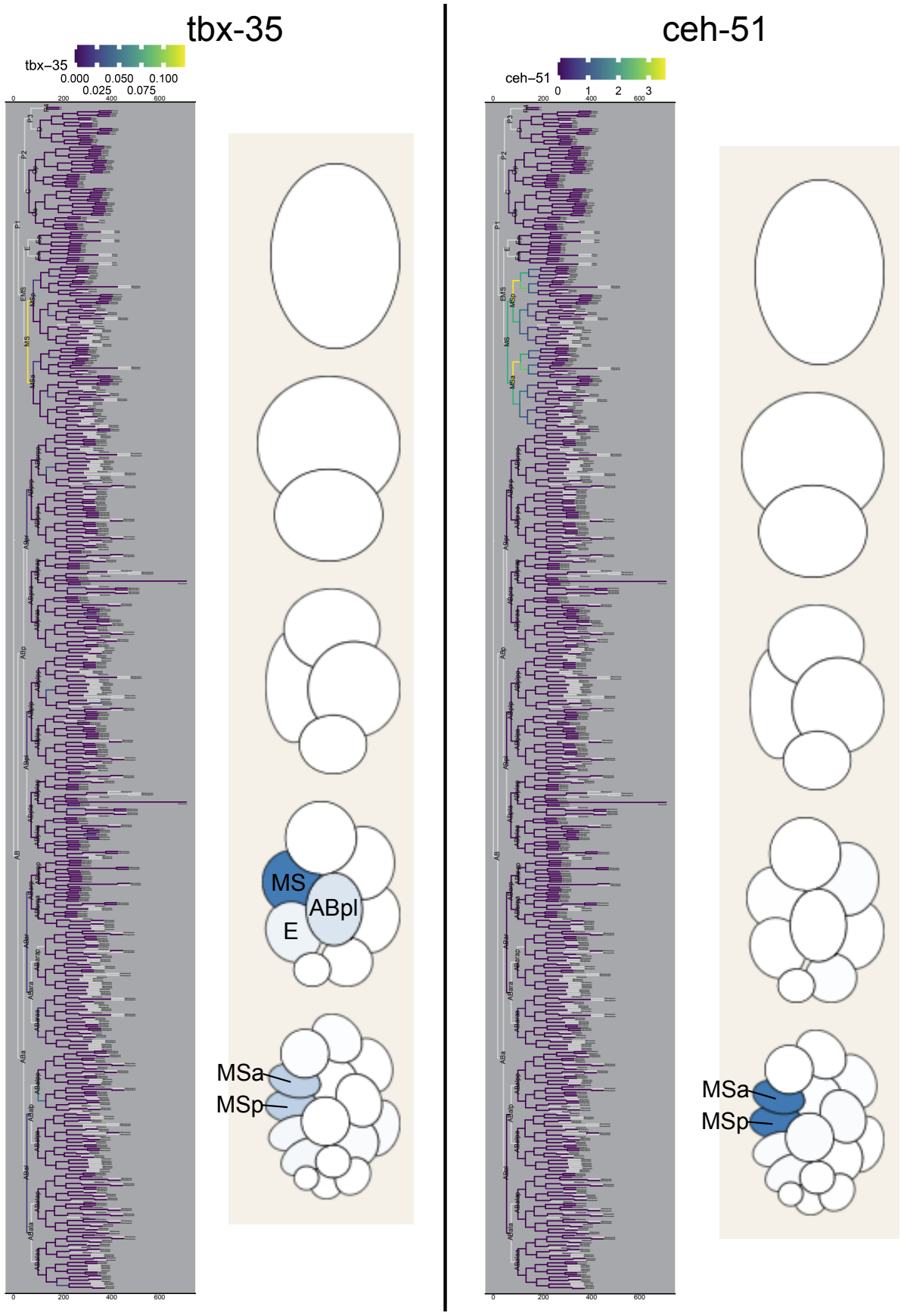
945 **Figure S2: Several early embryonic genes are expressed with anterior bias.**
946 Full lineage expression patterns are shown for integrated transgenic promoter reporters of *tbx-35*
947 and *ceh-51*; for an integrated transgenic protein reporter of PHA-4; and for a CRISPR knock-in
948 protein reporter of CEH-51.
949

irx-1 IRX-1 *pes-1*

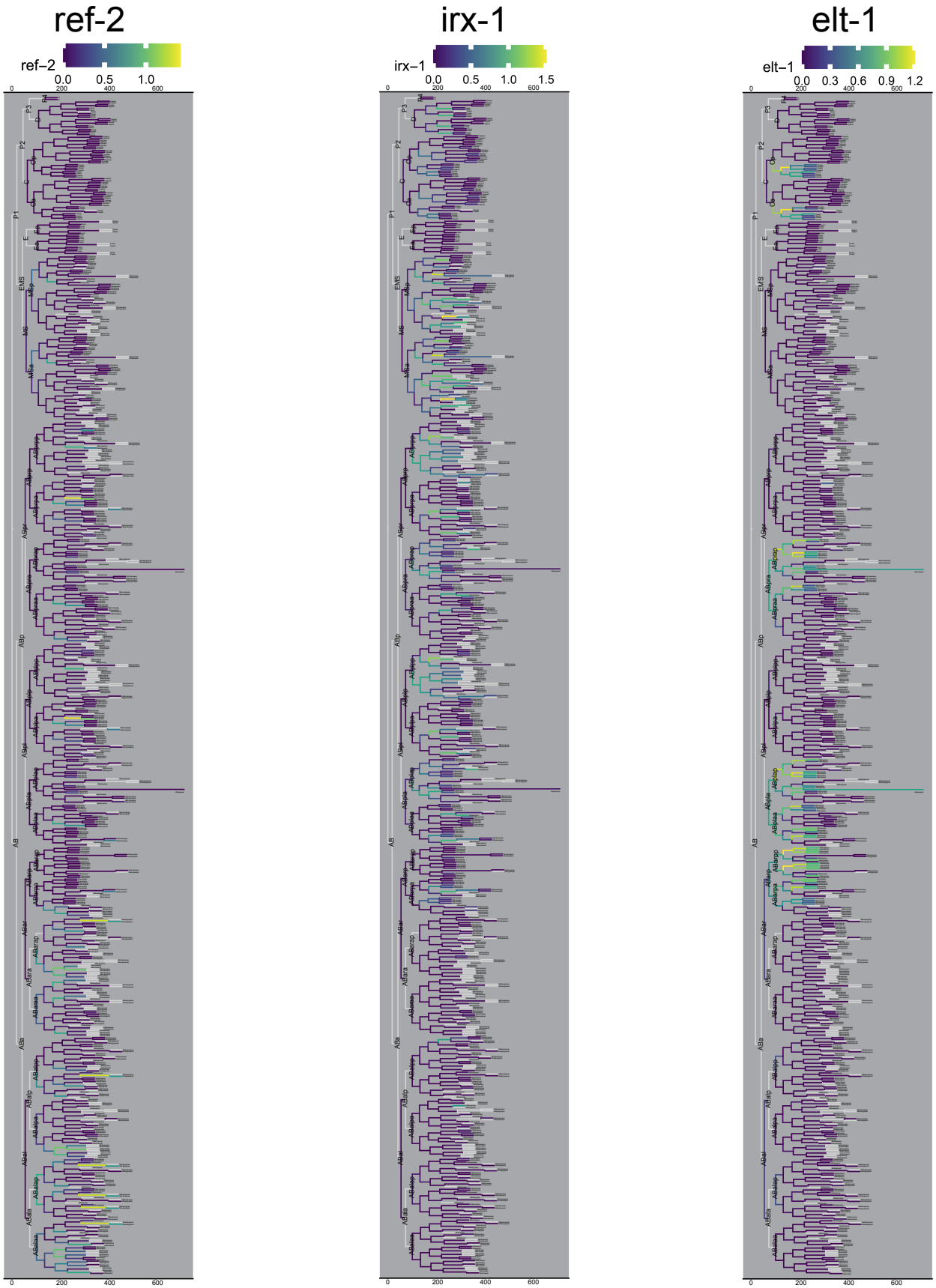


950
951

952 **Figure S3: Several early embryonic genes are expressed with anterior bias.**
953 Full lineage expression patterns are shown for integrated transgenic promoter reporters of *irx-1*
954 and *pes-1*; and for an integrated transgenic protein reporter of IRX-1.
955


tbx-35
tbx−35
0.000
0.025
0.050
0.075
0.100
MS
ABpl
E
MSa
MSp
ceh-51
ceh−51
0
1
2
3
MSa
MSp

957 **Figure S4: scRNA-seq data confirm imaging-based expression data.** scRNA-seq expression
958 data for *tbx-35* and *ceh-51*. *tbx-35* expression lineage derived from Packer et al. 2019 indicates
959 expression in the daughters of MS (MSa and MSp) [32]. Early *C. elegans* embryo diagram derived
960 from Tintori et al. 2016 indicates *tbx-35* expression in MS and its daughters, as well as low-level
961 expression in E and ABpl [56]. *ceh-51* expression lineage derived from Packer et al. 2019 indicates
962 expression in the descendants of MS from the MS2 through the MS16 generation (MS daughters
963 through MS great-great granddaughters) [32]. Early *C. elegans* embryo diagram derived from
964 Tintori et al. 2016 indicates *ceh-51* expression in the daughters of MS [56]. In both cases scRNA-
965 seq expression data agree with imaging-based expression data, albeit with expression being
966 detected earlier and persisting for fewer cell generations than for the fluorescent reporters (Figure
967 3 and Figure S6). Expressing cells in early embryo diagrams are labeled with their names. Gray
968 branches on lineage trees indicate cells for which distinct transcriptomes could not be determined.
969



970

Figure S5: scRNA-seq data confirm imaging-based expression data. scRNA-seq expression data for *ref-2*, *irx-1*, and *elt-1* derived from Packer et al. 2019 [32]. scRNA-seq expression data largely agree with imaging-based expression data, albeit with expression being detected earlier and persisting for fewer cell generations than for the fluorescent reporters (Figures 3 and 5 and Figure S6). Note that while most anterior expression events are resolved in these data, a few of the ABa lineage *ref-2* expression domains are not (because the anterior vs posterior sister cells were not separately annotated at the stage when *ref-2* is expressed). However, the data that are resolved for *ref-2* are consistent with the promoter and fosmid reporter expression patterns. Gray branches on lineage trees indicate cells for which distinct transcriptomes could not be determined.

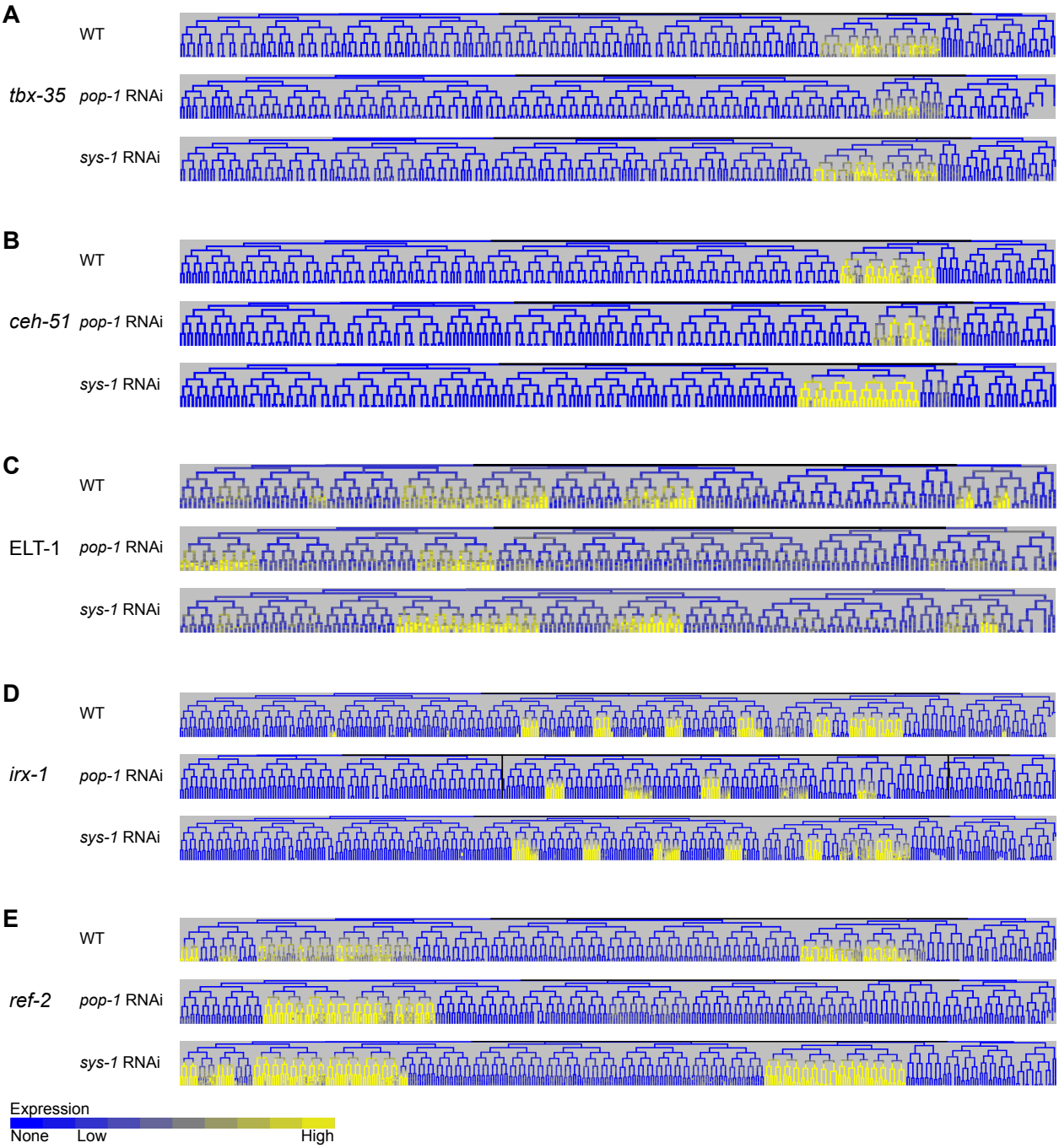


Figure S6: Anterior genes are regulated by Wnt effectors *pop-1* and *sys-1*. Full example WT, *pop-1* RNAi, and *sys-1* RNAi lineages are shown for integrated transgenic promoter reporters of *tbx-35* (A), *ceh-51* (B), *irx-1* (D), and *ref-2* (E) and for an integrated transgenic protein reporter of ELT-1 (C).

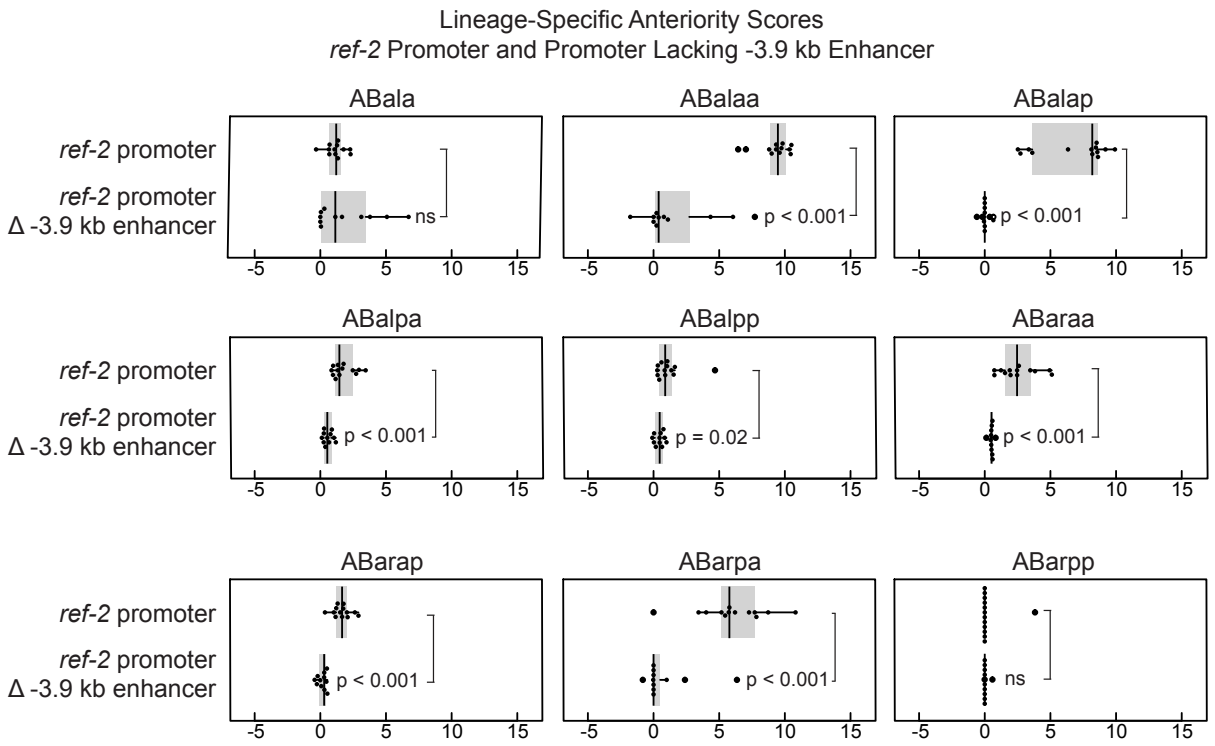


Figure S7: Deletion of the -3.9 kb enhancer from the *ref-2* promoter causes a reduction in the anterior bias in the expression driven by the promoter in most ABa sublineages in which the full-length promoter drives anterior-biased expression. Box plots display anteriority scores of lineages descended from indicated cells.

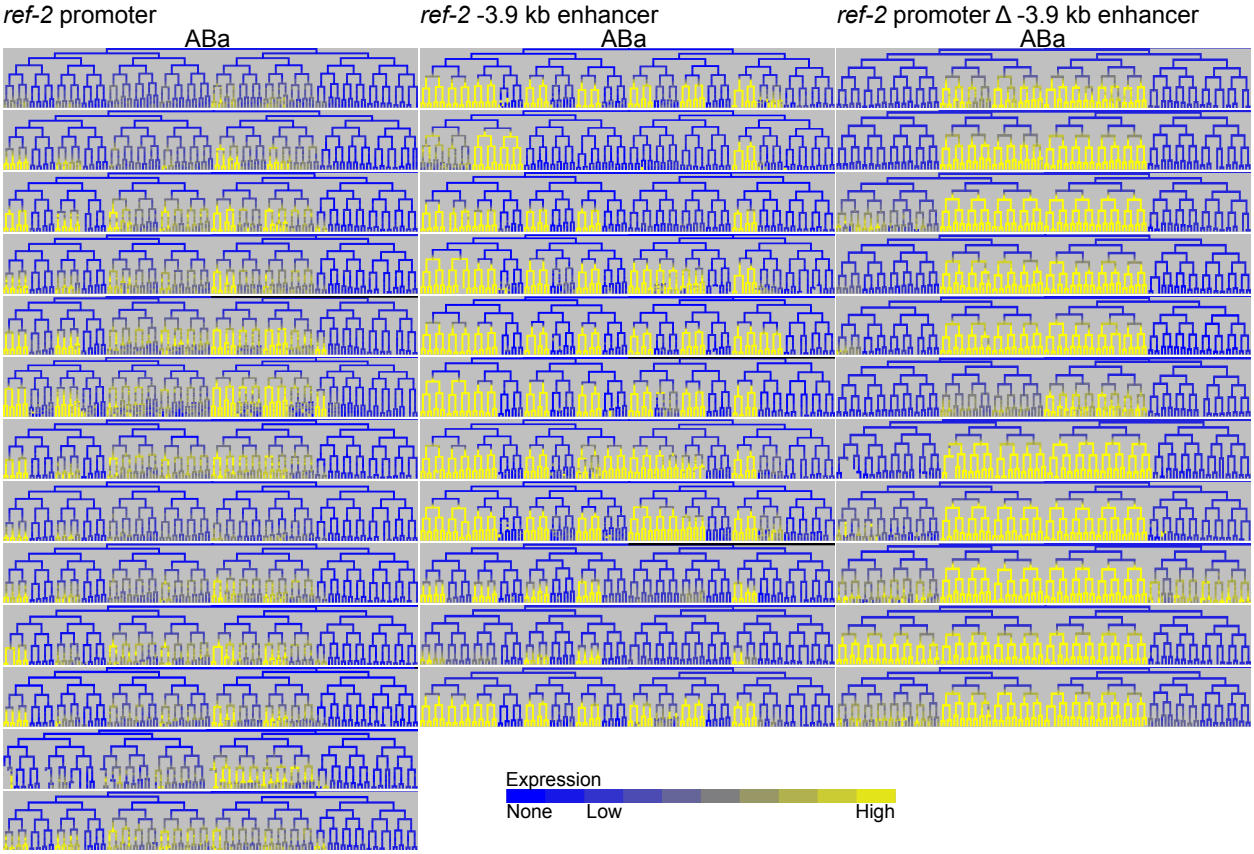
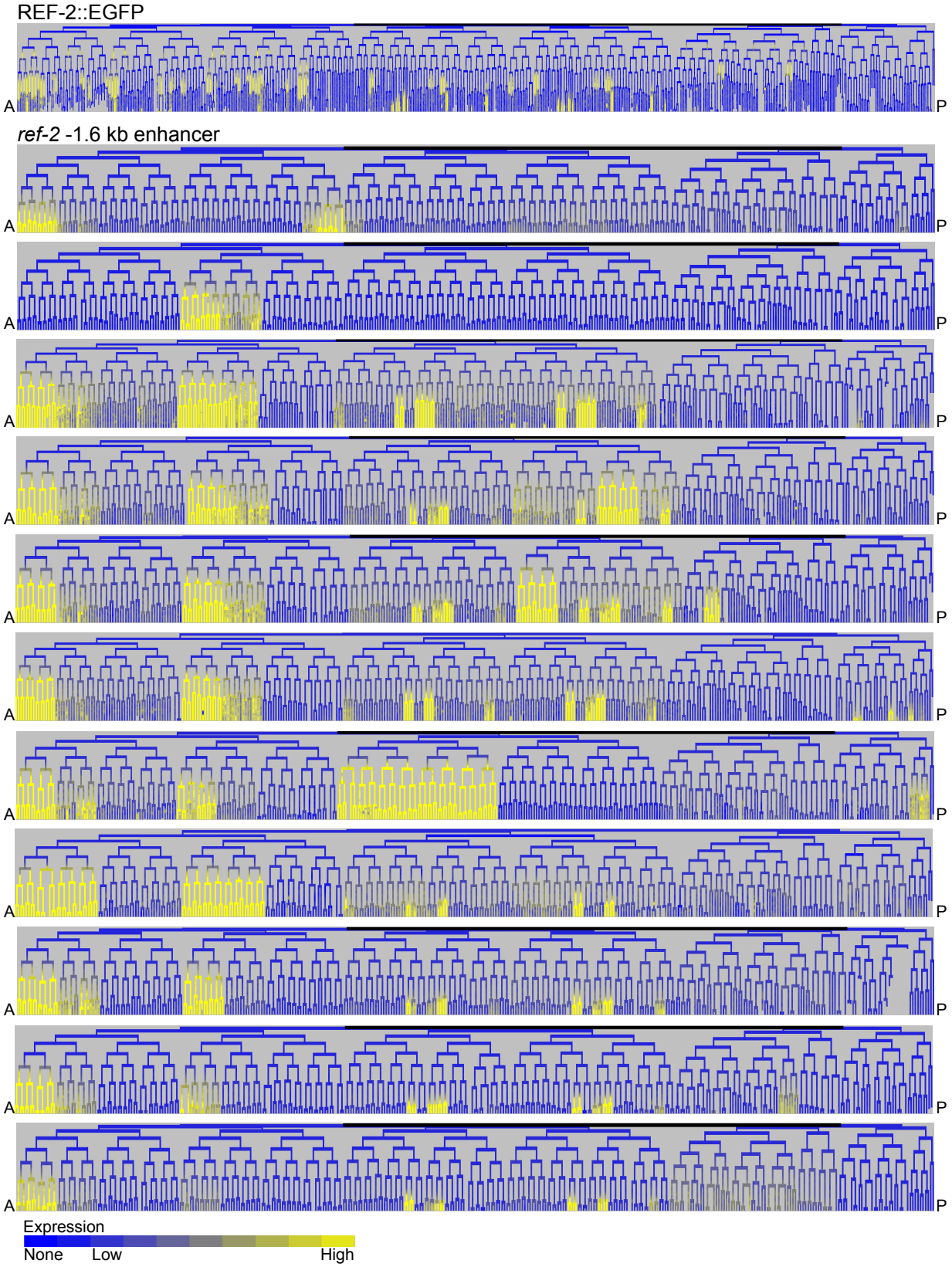
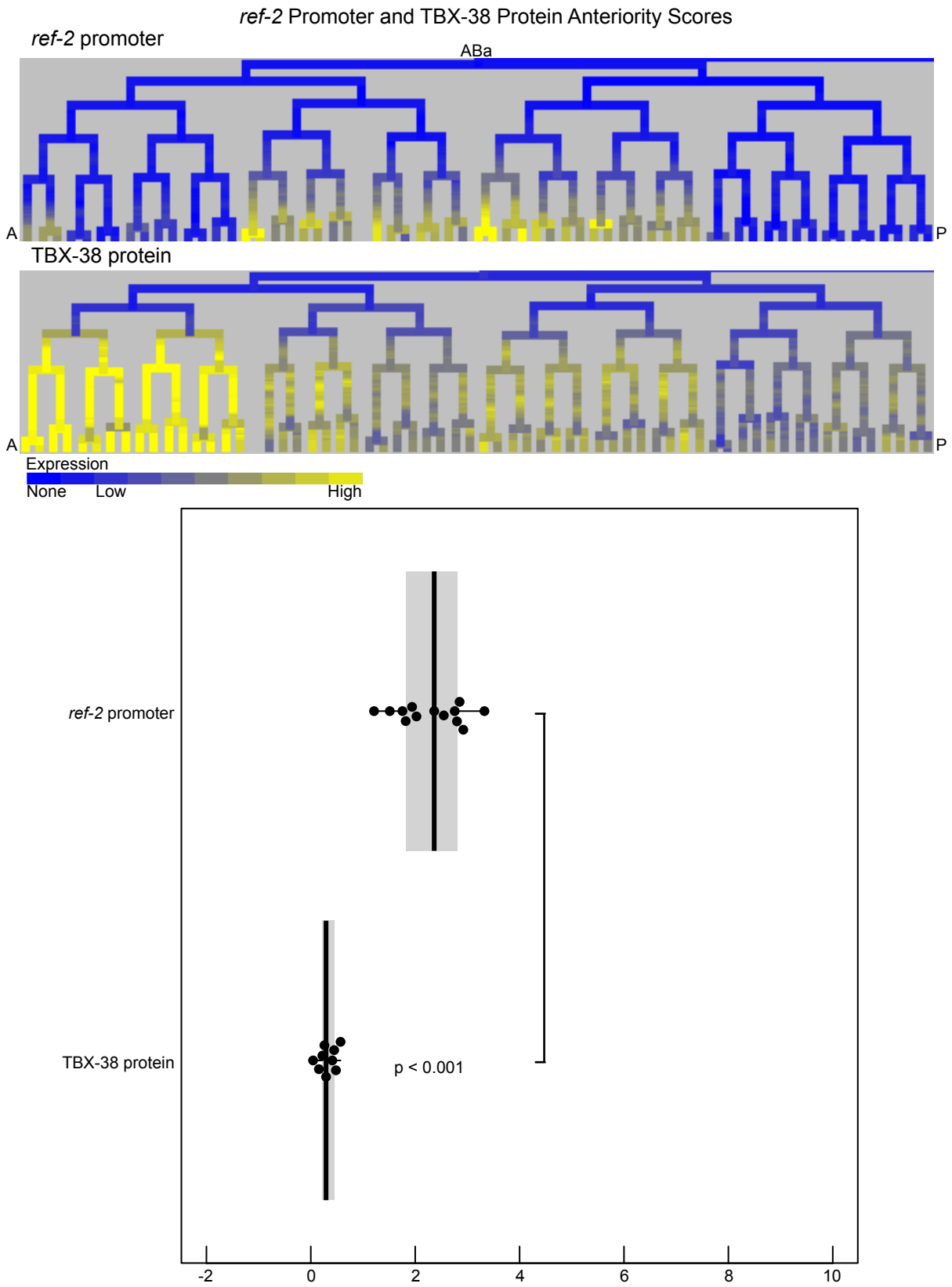


Figure S8: Expression patterns driven by the *ref-2* promoter, the -3.9 kb enhancer, and the promoter lacking the -3.9 kb enhancer are somewhat variable. Displayed here are partial lineages for all analyzed embryos bearing these reporters to show the variation in their expression.



999
1000

1001 **Figure S9: Expression patterns driven by the *ref-2* -1.6 kb enhancer are variable.** Displayed
1002 here are full lineages for all analyzed embryos bearing this reporter to show the variation in its
1003 expression. The full lineage expression pattern of the REF-2::GFP protein reporter is shown for
1004 comparison.
1005


ref-2 Promoter and TBX-38 Protein Anteriority Scores
ref-2 promoter
ABa
A
P
TBX-38 protein
A
P
Expression
None
Low
High
ref-2 promoter
TBX-38 protein
p < 0.001
-2
0
2
4
6
8
10

1006
1007

1008 **Figure S10: A TBX-38::GFP CRISPR knock-in protein reporter has minimally anterior-**
1009 **biased expression in the lineages in which the *ref-2* promoter drives anterior-biased**
1010 **expression.** Displayed are partial lineages of the P*ref-2*::*his-24*::mCherry promoter reporter
1011 expression and the TBX-38::GFP CRISPR knock-in expression. The box plot shows the mean
1012 anteriority scores for the *ref-2* promoter reporter and the TBX-38::GFP CRISPR knock-in for the
1013 same lineages as in Figure 5J.
1014

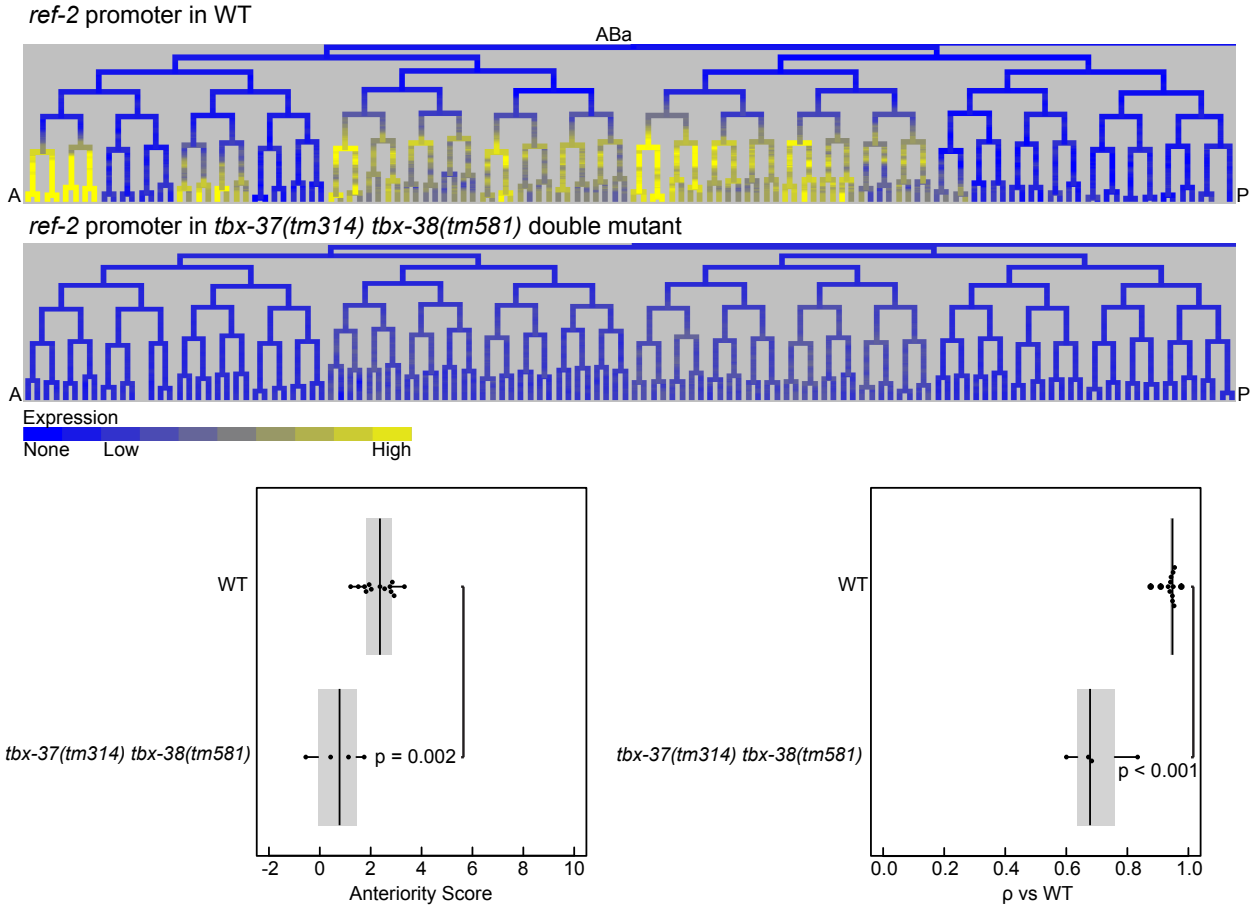


Figure S11: *tbx-37* and *tbx-38* are required for *ref-2* promoter-driven anterior-biased expression in the ABa lineage. Displayed are partial lineages for the expression driven by the *ref-2* promoter in a WT background and in a *tbx-37(tm314) tbx-38(tm581)* double mutant background. Also shown are 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 Figure 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.

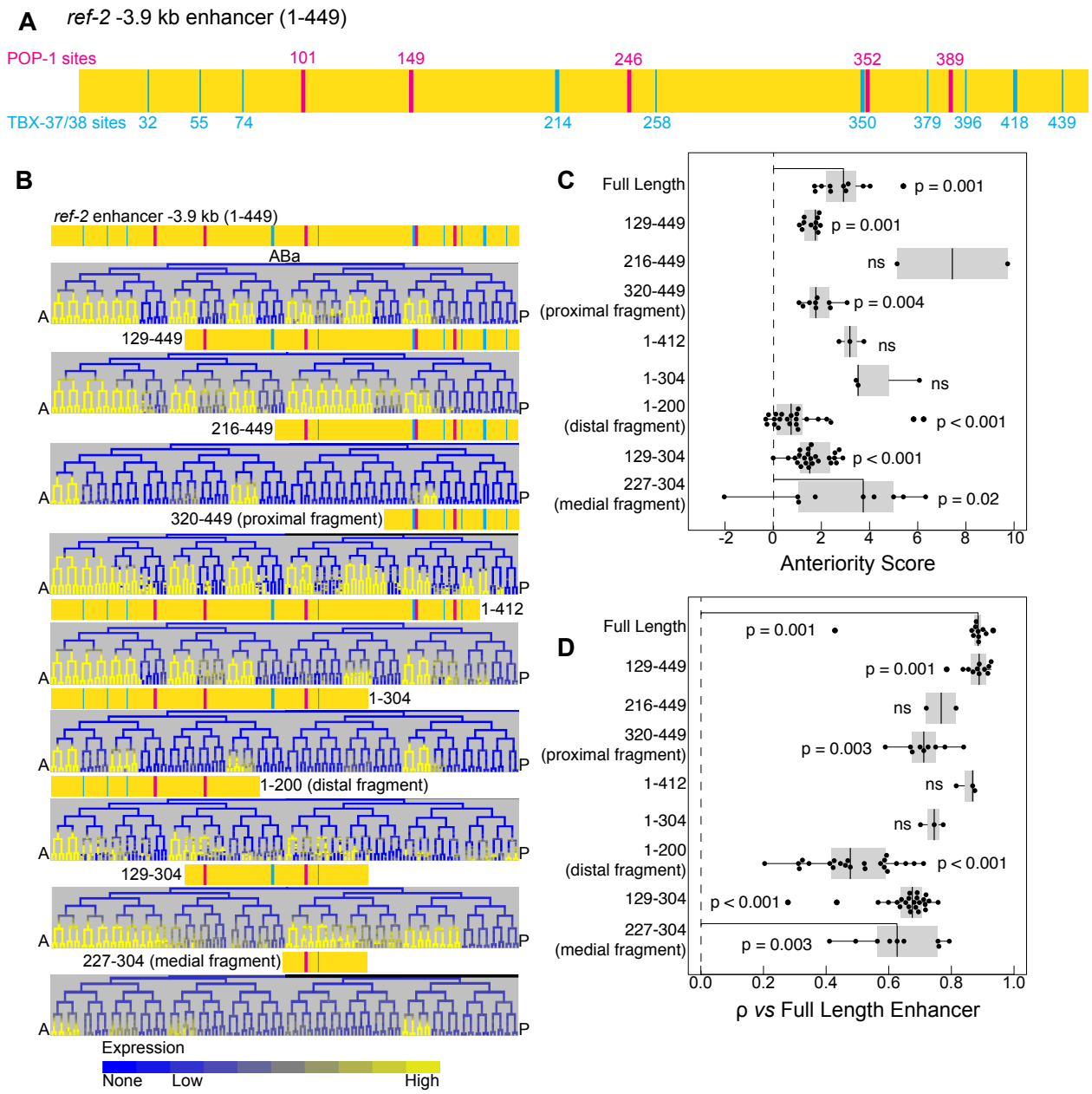


Figure S12: Three non-overlapping fragments of the -3.9 kb enhancer are each sufficient to drive anterior-biased expression in the ABa lineage of the early embryo. A) Model of the *ref-2* -3.9 kb enhancer with predicted TBX-37/38 and POP-1 sites indicated as in Figure 6. B) Expression patterns driven 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 Figure 5J. Spearman's ρ analysis uses the full ABa lineage and was calculated relative to the full-length -3.9 kb enhancer.

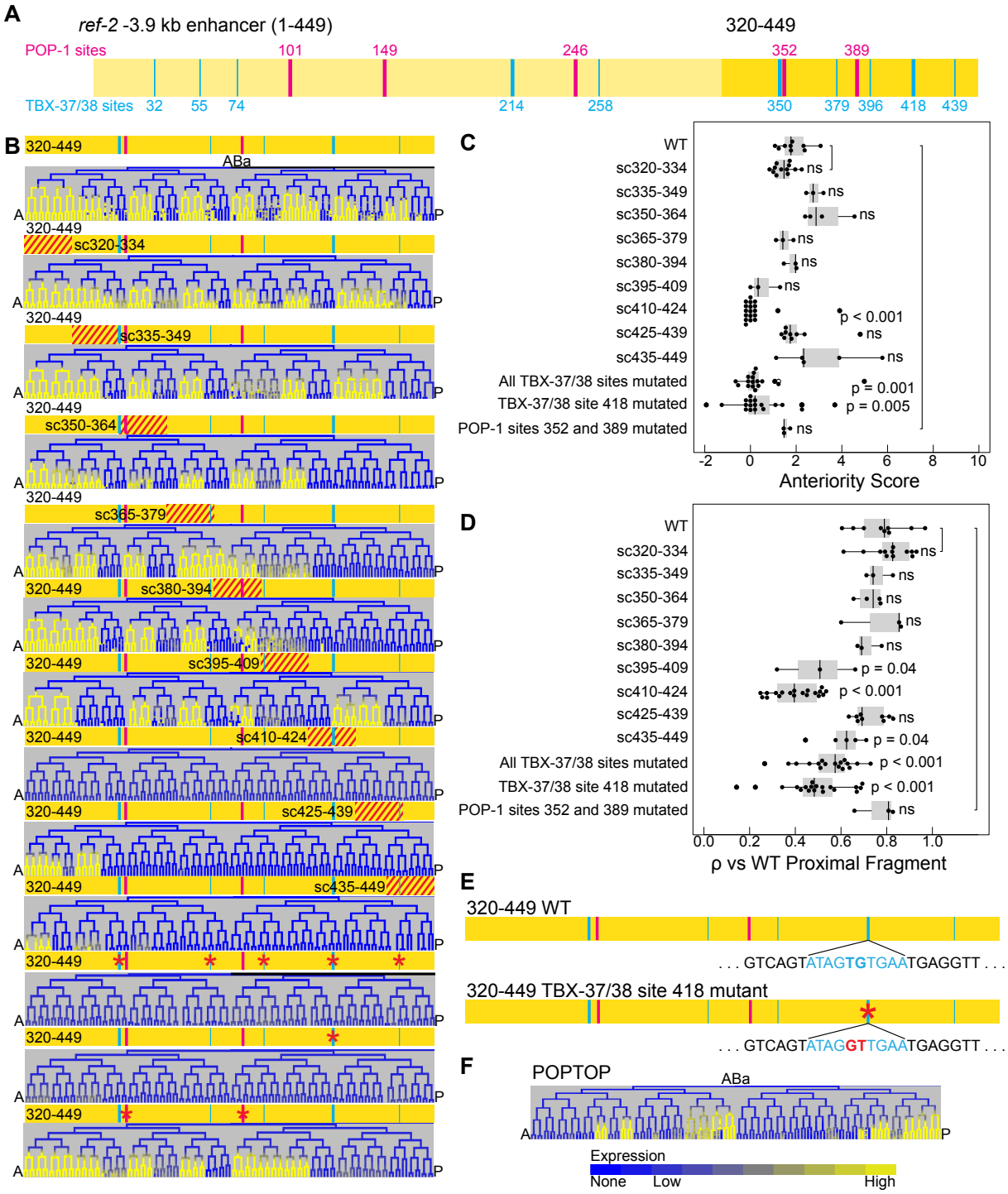


Figure S13: The broadly-expressed transcription factors *tbx-37* and *tbx-38* are required for the anterior-biased expression of *ref-2* in the ABa lineage. A) Model of the *ref-2* -3.9 kb enhancer with proximal fragment (base pairs 320-449) highlighted. Predicted TBX-37/38 and POP-1 sites are indicated as in Figure 6. B) Expression patterns driven by the proximal fragment of the *ref-2* -3.9 kb enhancer with its WT sequence, with each 15 base-pair segment of the fragment scrambled, with all predicted TBX-37/38 sites mutated, with TBX_{418} mutated, and with the two

1045 predicted POP-1 sites mutated. C-D) Box plots displaying the anteriority scores (C) and
1046 Spearman's ρ (D) for the proximal fragment with its WT sequence, with each 15 base-pair segment
1047 of the fragment scrambled, with all predicted TBX-37/38 sites mutated, with TBX_{418} mutated, and
1048 with the two predicted POP-1 sites mutated. Lineages used to determine anteriority scores are the
1049 same as in Figure 5J. Spearman's ρ analysis uses the full ABa lineage and was calculated relative
1050 to the WT proximal fragment. E) Models of the proximal fragment of the *ref-2* -3.9 kb enhancer
1051 with the WT sequence of TBX_{418} and the mutated sequence used in the TBX_{418} mutant reporters.
1052 F) Expression pattern driven by the seven-copy POP-1 site concatemer POPTOP [13,53].
1053

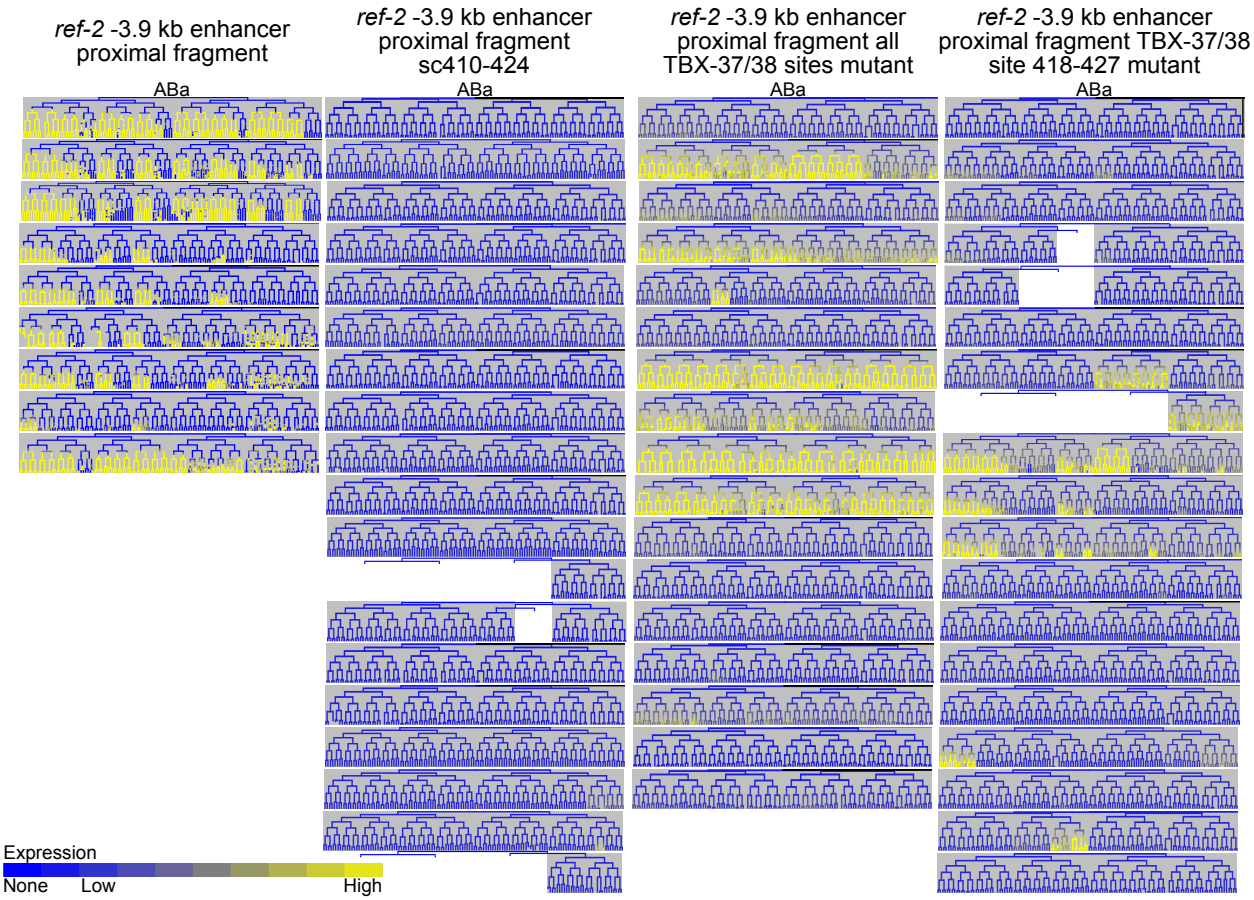


Figure S14: *ref-2* -3.9 kb enhancer WT proximal fragment, proximal fragment with all TBX-37/38 sites mutated, and proximal fragment with *TBX*$_{418}$ mutated exhibit variation in their expression patterns. Displayed are partial lineages for all analyzed embryos expressing reporters for the *ref-2* -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.

TABLES

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

ACKNOWLEDGMENTS

We thank J. Archibald Millar, Julia Richards and Shaili Patel for technical contributions to data collection and analysis. We thank Meera Sundaram for use of her injection equipment. We thank Christopher Brown, Meera Sundaram, Michael Atchison, Peter Klein, and Mary Mullins for helpful discussions. 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 *C. elegans* strains. We thank the *C. elegans* Reverse Genetics Core Facility at the University of British Columbia, part of the International *C. elegans* Gene Knockout Consortium, which produced the *ref-2* mutant allele (*ref-2(gk178))* used in this study. We thank the National Bioresource Project, Tokyo, Japan, which is part of the International *C. elegans* Gene Knockout Consortium, which produced the *tbx-37* and *tbx-38* mutant alleles (*tbx-37(tm314)* and *tbx-38(tm581)*) used in this study. This work was funded by T32GM008216, F31GM123737, and R35GM130357.

1099 **References Cited**

1100 1. Hikasa H, Sokol SY. Wnt signaling in vertebrate axis specification. Cold Spring Harb
1101 Perspect Biol. 2013;5: a007955. doi:10.1101/cshperspect.a007955

1102 2. Archbold HC, Yang YX, Chen L, Cadigan KM. How do they do Wnt they do?: regulation of
1103 transcription by the Wnt/β-catenin pathway. Acta Physiol Oxf Engl. 2012;204: 74–109.
1104 doi:10.1111/j.1748-1716.2011.02293.x

1105 3. Martin BL, Kimelman D. Wnt signaling and the evolution of embryonic posterior
1106 development. Curr Biol CB. 2009;19: R215–R219. doi:10.1016/j.cub.2009.01.052

1107 4. Zacharias AL, Murray JI. Combinatorial decoding of the invariant C. elegans embryonic
1108 lineage in space and time. genesis. 2016;54: 182–197. doi:10.1002/dvg.22928

1109 5. Sulston JE, Schierenberg E, White JG, Thomson JN. The embryonic cell lineage of the
1110 nematode Caenorhabditis elegans. Dev Biol. 1983;100: 64–119. doi:10.1016/0012-
1111 1606(83)90201-4

1112 6. Bürglin TR, Finney M, Coulson A, Ruvkun G. Caenorhabditis elegans has scores of
1113 homoeobox-containing genes. Nature. 1989;341: 239–243. doi:10.1038/341239a0

1114 7. Chen W, Lim HH, Lim L. A new member of the ras superfamily, the rac1 homologue from
1115 Caenorhabditis elegans. Cloning and sequence analysis of cDNA, pattern of developmental
1116 expression, and biochemical characterization of the protein. J Biol Chem. 1993;268: 320–
1117 324. doi:10.1016/S0021-9258(18)54152-1

1118 8. van den Heuvel S. Cell-cycle regulation. WormBook. 2005. pp. 1–16.
1119 doi:10.1895/wormbook.1.28.1

1120 9. Shaye DD, Greenwald I. OrthoList: A Compendium of C. elegans Genes with Human
1121 Orthologs. PLOS ONE. 2011;6: e20085. doi:10.1371/journal.pone.0020085

1122 10. Kim W, Underwood RS, Greenwald I, Shaye DD. OrthoList 2: A New Comparative
1123 Genomic Analysis of Human and Caenorhabditis elegans Genes. Genetics. 2018;210: 445–
1124 461. doi:10.1534/genetics.118.301307

1125 11. Herman M. C. elegans POP-1/TCF functions in a canonical Wnt pathway that controls cell
1126 migration and in a noncanonical Wnt pathway that controls cell polarity. Development.
1127 2001;128: 581–590. doi:10.1242/dev.128.4.581

1128 12. Mizumoto K, Sawa H. Two βs or not two βs: regulation of asymmetric division by β-catenin.
1129 Trends Cell Biol. 2007;17: 465–473. doi:10.1016/j.tcb.2007.08.004

1130 13. Zacharias AL, Walton T, Preston E, Murray JI. Quantitative Differences in Nuclear β-catenin
1131 and TCF Pattern Embryonic Cells in C. elegans. Sternberg PW, editor. PLOS Genet.
1132 2015;11: e1005585. doi:10.1371/journal.pgen.1005585

1133 14. Lin R, Hill RJ, Priess JR. POP-1 and anterior-posterior fate decisions in C. elegans embryos.
1134 Cell. 1998;92: 229–239.

1135 15. Maduro MF, Lin R, Rothman JH. Dynamics of a developmental switch: recursive
1136 intracellular and intranuclear redistribution of Caenorhabditis elegans POP-1 parallels Wnt-
1137 inhibited transcriptional repression. Dev Biol. 2002;248: 128–142.
1138 doi:10.1006/dbio.2002.0721

1139 16. Huang S, Shetty P, Robertson SM, Lin R. Binary cell fate specification during C. elegans
1140 embryogenesis driven by reiterated reciprocal asymmetry of TCF POP-1 and its
1141 coactivatorβ-catenin SYS-1. Development. 2007;134: 2685–2695. doi:10.1242/dev.008268

1142 17. Calvo D, Victor M, Gay F, Sui G, Po-Shan Luke M, Dufourcq P, et al. A POP-1 repressor
1143 complex restricts inappropriate cell type-specific gene transcription during Caenorhabditis
1144 elegans embryogenesis. EMBO J. 2001;20: 7197–7208. doi:10.1093/emboj/20.24.7197

1145 18. Maduro MF, Kasmir JJ, Zhu J, Rothman JH. The Wnt effector POP-1 and the PAL-1/Caudal
1146 homeoprotein collaborate with SKN-1 to activate C. elegans endoderm development. Dev
1147 Biol. 2005;285: 510–523. doi:10.1016/j.ydbio.2005.06.022

1148 19. Owraghi M, Broitman-Maduro G, Luu T, Roberson H, Maduro MF. Roles of the Wnt
1149 effector POP-1/TCF in the C. elegans endomesoderm specification gene network. Dev Biol.
1150 2010;340: 209–221. doi:10.1016/j.ydbio.2009.09.042

1151 20. Murgan S, Kari W, Rothbächer U, Iché-Torres M, Mélénec P, Hobert O, et al. Atypical
1152 Transcriptional Activation by TCF via a Zic Transcription Factor in C. elegans Neuronal
1153 Precursors. Dev Cell. 2015;33: 737–745. doi:10.1016/j.devcel.2015.04.018

1154 21. Murgan S, Bertrand V. How targets select activation or repression in response to Wnt.
1155 Worm. 2015;4: e1086869. doi:10.1080/21624054.2015.1086869

1156 22. Zhang CU, Blauwkamp TA, Burby PE, Cadigan KM. Wnt-Mediated Repression via
1157 Bipartite DNA Recognition by TCF in the Drosophila Hematopoietic System. Desplan C,
1158 editor. PLoS Genet. 2014;10: e1004509. doi:10.1371/journal.pgen.1004509

1159 23. Murray JI, Boyle TJ, Preston E, Vafeados D, Mericle B, Weisdepp P, et al. Multidimensional
1160 regulation of gene expression in the C. elegans embryo. Genome Res. 2012;22: 1282–1294.
1161 doi:10.1101/gr.131920.111

1162 24. Ma X, Zhao Z, Xiao L, Xu W, Kou Y, Zhang Y, et al. A 4D single-cell protein atlas of
1163 transcription factors delineates spatiotemporal patterning during embryogenesis. Nat
1164 Methods. 2021;18: 893–902. doi:10.1038/s41592-021-01216-1

1165 25. Murray JI, Bao Z, Boyle TJ, Boeck ME, Mericle BL, Nicholas TJ, et al. Automated analysis
1166 of embryonic gene expression with cellular resolution in C. elegans. Nat Methods. 2008;5:
1167 703–709. doi:10.1038/nmeth.1228

1168 26. Kvon EZ, Waymack R, Gad M, Wunderlich Z. Enhancer redundancy in development and
1169 disease. Nat Rev Genet. 2021;22: 324–336. doi:10.1038/s41576-020-00311-x

1170 27. Sulston JE, Horvitz HR. Post-embryonic cell lineages of the nematode, Caenorhabditis
1171 elegans. Dev Biol. 1977;56: 110–156. doi:10.1016/0012-1606(77)90158-0

1172 28. Bao Z, Murray JI, Boyle T, Ooi SL, Sandel MJ, Waterston RH. Automated cell lineage
1173 tracing in Caenorhabditis elegans. Proc Natl Acad Sci U S A. 2006;103: 2707.
1174 doi:10.1073/pnas.0511111103

1175 29. Santella A, Du Z, Nowotschin S, Hadjantonakis A-K, Bao Z. A hybrid blob-slice model for
1176 accurate and efficient detection of fluorescence labeled nuclei in 3D. BMC Bioinformatics.
1177 2010;11: 580. doi:10.1186/1471-2105-11-580

1178 30. Boyle TJ, Bao Z, Murray JI, Araya CL, Waterston RH. AceTree: a tool for visual analysis of
1179 Caenorhabditis elegans embryogenesis. BMC Bioinformatics. 2006;7: 275.
1180 doi:10.1186/1471-2105-7-275

1181 31. Katzman B, Tang D, Santella A, Bao Z. AceTree: a major update and case study in the long
1182 term maintenance of open-source scientific software. BMC Bioinformatics. 2018;19: 121.
1183 doi:10.1186/s12859-018-2127-0

1184 32. Packer Jonathan S., Zhu Qin, Huynh Chau, Sivaramakrishnan Priya, Preston Elicia, Dueck
1185 Hannah, et al. A lineage-resolved molecular atlas of C. elegans embryogenesis at single-cell
1186 resolution. Science. 2019;365: eaax1971. doi:10.1126/science.aax1971

1187 33. Yanai I, Baugh LR, Smith JJ, Roehrig C, Shen-Orr SS, Claggett JM, et al. Pairing of
1188 competitive and topologically distinct regulatory modules enhances patterned gene
1189 expression. Mol Syst Biol. 2008;4. doi:10.1038/msb.2008.6

1190 34. Lin R, Thompson S, Priess JR. pop-1 encodes an HMG box protein required for the
1191 specification of a mesoderm precursor in early C. elegans embryos. Cell. 1995;83: 599–609.
1192 doi:10.1016/0092-8674(95)90100-0

1193 35. Kidd AR, Miskowski JA, Siegfried KR, Sawa H, Kimble J. A β-Catenin Identified by
1194 Functional Rather Than Sequence Criteria and Its Role in Wnt/MAPK Signaling. Cell.
1195 2005;121: 761–772. doi:10.1016/j.cell.2005.03.029

1196 36. Phillips BT, Kidd AR, King R, Hardin J, Kimble J. Reciprocal asymmetry of SYS-1/beta-
1197 catenin and POP-1/TCF controls asymmetric divisions in Caenorhabditis elegans. Proc Natl
1198 Acad Sci. 2007;104: 3231–3236. doi:10.1073/pnas.0611507104

1199 37. Broitman-Maduro G, Owraghi M, Hung WWK, Kuntz S, Sternberg PW, Maduro MF. The
1200 NK-2 class homeodomain factor CEH-51 and the T-box factor TBX-35 have overlapping
1201 function in C. elegans mesoderm development. Dev Camb Engl. 2009;136: 2735–2746.
1202 doi:10.1242/dev.038307

1203 38. Broitman-Maduro G, Lin KT-H, Hung WWK, Maduro MF. Specification of the C. elegans
1204 MS blastomere by the T-box factor TBX-35. Dev Camb Engl. 2006;133: 3097–3106.
1205 doi:10.1242/dev.02475

1206 39. Alper S, Kenyon C. The zinc finger protein REF-2 functions with the Hox genes to inhibit
1207 cell fusion in the ventral epidermis of C. elegans. Development. 2002;129: 3335–3348.
1208 doi:10.1242/dev.129.14.3335

1209 40. Bertrand V, Hobert O. Linking asymmetric cell division to the terminal differentiation
1210 program of postmitotic neurons in C. elegans. Dev Cell. 2009;16: 563–575.
1211 doi:10.1016/j.devcel.2009.02.011

1212 41. Walton T, Preston E, Nair G, Zacharias AL, Raj A, Murray JI. The Bicoid Class
1213 Homeodomain Factors ceh-36/OTX and unc-30/PITX Cooperate in C. elegans Embryonic
1214 Progenitor Cells to Regulate Robust Development. PLOS Genet. 2015;11: e1005003.
1215 doi:10.1371/journal.pgen.1005003

1216 42. Richards JL, Zacharias AL, Walton T, Burdick JT, Murray JI. A quantitative model of
1217 normal Caenorhabditis elegans embryogenesis and its disruption after stress. Dev Biol.
1218 2013;374: 12–23. doi:10.1016/j.ydbio.2012.11.034

1219 43. Maduro MF, Meneghini MD, Bowerman B, Broitman-Maduro G, Rothman JH. Restriction
1220 of Mesendoderm to a Single Blastomere by the Combined Action of SKN-1 and a GSK-3β
1221 Homolog Is Mediated by MED-1 and -2 in C. elegans. Mol Cell. 2001;7: 475–485.
1222 doi:10.1016/S1097-2765(01)00195-2

1223 44. Maduro MF, Hill RJ, Heid PJ, Newman-Smith ED, Zhu J, Priess JR, et al. Genetic
1224 redundancy in endoderm specification within the genus Caenorhabditis. Dev Biol. 2005;284:
1225 509–522. doi:10.1016/j.ydbio.2005.05.016

1226 45. Boeck ME, Boyle T, Bao Z, Murray J, Mericle B, Waterston R. Specific roles for the GATA
1227 transcription factors end-1 and end-3 during C. elegans E-lineage development. Spec Sect
1228 Hist Dev Biol. 2011;358: 345–355. doi:10.1016/j.ydbio.2011.08.002

1229 46. Good K, Ciosk R, Nance J, Neves A, Hill RJ, Priess JR. The T-box transcription factors
1230 TBX-37 and TBX-38 link GLP-1/Notch signaling to mesoderm induction in C. elegans
1231 embryos. Development. 2004;131: 1967–1978. doi:10.1242/dev.01088

1232 47. Andachi Y. Caenorhabditis elegans T-box genes tbx-9 and tbx-8 are required for formation
1233 of hypodermis and body-wall muscle in embryogenesis. Genes Cells. 2004;9: 331–344.
1234 doi:10.1111/j.1356-9597.2004.00725.x

1235 48. Neves A, Priess JR. The REF-1 Family of bHLH Transcription Factors Pattern C. elegans
1236 Embryos through Notch-Dependent and Notch-Independent Pathways. Dev Cell. 2005;8:
1237 867–879. doi:10.1016/j.devcel.2005.03.012

1238 49. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The Human
1239 Genome Browser at UCSC. Genome Res. 2002;12: 996–1006. doi:10.1101/gr.229102

1240 50. Charest J, Daniele T, Wang J, Bykov A, Mandlbauer A, Asparuhova M, et al. Combinatorial
1241 Action of Temporally Segregated Transcription Factors. Dev Cell. 2020;55: 483-499.e7.
1242 doi:10.1016/j.devcel.2020.09.002

1243 51. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, et al.
1244 Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome
1245 Res. 2005;15: 1034–1050. doi:10.1101/gr.3715005

1246 52. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution
1247 rates on mammalian phylogenies. Genome Res. 2010;20: 110–121.
1248 doi:10.1101/gr.097857.109

1249 53. LaBonty M, Szmygiel C, Byrnes LE, Hughes S, Woollard A, Cram EJ. CACN-1/Cactin
1250 plays a role in Wnt signaling in C. elegans. PloS One. 2014;9.
1251 doi:10.1371/journal.pone.0101945

1252 54. Murray JI, Preston E, Crawford JP, Rumley JD, Amom P, Anderson BD, et al. The anterior
1253 Hox gene ceh-13 and elt-1/GATA activate the posterior Hox genes nob-1 and php-3 to
1254 specify posterior lineages in the C. elegans embryo. PLOS Genet. 2022;18: e1010187.
1255 doi:10.1371/journal.pgen.1010187

1256 55. Poole RJ, Hobert O. Early Embryonic Programming of Neuronal Left/Right Asymmetry in
1257 C. elegans. Curr Biol. 2006;16: 2279–2292. doi:10.1016/j.cub.2006.09.041

1258 56. Tintori SC, Osborne Nishimura E, Golden P, Lieb JD, Goldstein B. A Transcriptional
1259 Lineage of the Early C. elegans Embryo. Dev Cell. 2016;38: 430–444.
1260 doi:10.1016/j.devcel.2016.07.025

1261 57. Rocheleau CE, Yasuda J, Shin TH, Lin R, Sawa H, Okano H, et al. WRM-1 Activates the
1262 LIT-1 Protein Kinase to Transduce Anterior/Posterior Polarity Signals in C. elegans. Cell.
1263 1999;97: 717–726. doi:10.1016/S0092-8674(00)80784-9

1264 58. Lo M-C, Gay F, Odom R, Shi Y, Lin R. Phosphorylation by the β-Catenin/MAPK Complex
1265 Promotes 14-3-3-Mediated Nuclear Export of TCF/POP-1 in Signal-Responsive Cells in C.
1266 elegans. Cell. 2004;117: 95–106. doi:10.1016/S0092-8674(04)00203-X

1267 59. Kaletta T, Schnabel H, Schnabel R. Binary specification of the embryonic lineage in
1268 Caenorhabditis elegans. Nature. 1997;390: 294–298. doi:10.1038/36869

1269 60. Fukushige T, Brodigan TM, Schriefer LA, Waterston RH, Krause M. Defining the
1270 transcriptional redundancy of early bodywall muscle development in C. elegans: evidence
1271 for a unified theory of animal muscle development. Genes Dev. 2006;20: 3395–3406.
1272 doi:10.1101/gad.1481706

1273 61. Fukushige T, Krause M. The myogenic potency of HLH-1 reveals wide-spread
1274 developmental plasticity in early C. elegans embryos. Dev Camb Engl. 2005;132: 1795–
1275 1805. doi:10.1242/dev.01774

1276 62. Simonis N, Rual J-F, Carvunis A-R, Tasan M, Lemmens I, Hirozane-Kishikawa T, et al.
1277 Empirically controlled mapping of the Caenorhabditis elegans protein-protein interactome
1278 network. Nat Methods. 2009;6.

1279 63. Whiting J, Marshall H, Cook M, Krumlauf R, Rigby PW, Stott D, et al. Multiple spatially
1280 specific enhancers are required to reconstruct the pattern of Hox-2.6 gene expression. Genes
1281 Dev. 1991;5: 2048–2059. doi:10.1101/gad.5.11.2048

1282 64. Epstein DJ, McMahon AP, Joyner AL. Regionalization of Sonic hedgehog transcription
1283 along the anteroposterior axis of the mouse central nervous system is regulated by Hnf3-
1284 dependent and -independent mechanisms. Development. 1999;126: 281–292.
1285 doi:10.1242/dev.126.2.281

1286 65. Teng Y, Girard L, Ferreira HB, Sternberg PW, Emmons SW. Dissection of cis-regulatory
1287 elements in the C. elegans Hox gene egl-5 promoter. Dev Biol. 2004;276: 476–492.
1288 doi:10.1016/j.ydbio.2004.09.012

1289 66. Streit A, Kohler R, Marty T, Belfiore M, Takacs-Vellai K, Vigano M-A, et al. Conserved
1290 Regulation of the Caenorhabditis elegans labial/Hox1 Gene ceh-13. Dev Biol. 2002;242: 96–
1291 108. doi:10.1006/dbio.2001.0544

1292 67. Landmann F, Quintin S, Labouesse M. Multiple regulatory elements with spatially and
1293 temporally distinct activities control the expression of the epithelial differentiation gene lin-
1294 26 in C. elegans. Dev Biol. 2004;265: 478–490. doi:10.1016/j.ydbio.2003.09.009

1295 68. Chen RA-J, Down TA, Stempor P, Chen QB, Egelhofer TA, Hillier LW, et al. The landscape
1296 of RNA polymerase II transcription initiation in C. elegans reveals promoter and enhancer
1297 architectures. Genome Res. 2013;23: 1339–1347. doi:10.1101/gr.153668.112

1298 69. Serizay J, Dong Y, Jänes J, Chesney M, Cerrato C, Ahringer J. Distinctive regulatory
1299 architectures of germline-active and somatic genes in C. elegans. Genome Res. 2020;30:
1300 1752–1765. doi:10.1101/gr.265934.120

1301 70. Kuersten S, Segal SP, Verheyden J, LaMartina SM, Goodwin EB. NXF-2, REF-1, and REF-
1302 2 Affect the Choice of Nuclear Export Pathway for tra-2 mRNA in C. elegans. Mol Cell.
1303 2004;14: 599–610. doi:10.1016/j.molcel.2004.05.004

1304 71. Clark E, Akam M. Odd-paired controls frequency doubling in Drosophila segmentation by
1305 altering the pair-rule gene regulatory network. Wittkopp PJ, editor. eLife. 2016;5: e18215.
1306 doi:10.7554/eLife.18215

1307 72. Yoon Y, Klomp J, Martin-Martin I, Criscione F, Calvo E, Ribeiro J, et al. Embryo polarity in
1308 moth flies and mosquitoes relies on distinct old genes with localized transcript isoforms.
1309 Wittkopp PJ, Pick L, Popadic A, Akam M, editors. eLife. 2019;8: e46711.
1310 doi:10.7554/eLife.46711

1311 73. Kamath RS, Fraser AG, Dong Y, Poulin G, Durbin R, Gotta M, et al. Systematic functional
1312 analysis of the Caenorhabditis elegans genome using RNAi. Nature. 2003;421: 231–237.
1313 doi:10.1038/nature01278

1314 74. Vora S, Phillips BT. Centrosome-Associated Degradation Limits β-Catenin Inheritance by
1315 Daughter Cells after Asymmetric Division. Curr Biol. 2015;25: 1005–1016.
1316 doi:10.1016/j.cub.2015.02.020

1317 75. Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, et al.
1318 Determination and inference of eukaryotic transcription factor sequence specificity. Cell.
1319 2014;158: 1431–1443. doi:10.1016/j.cell.2014.08.009

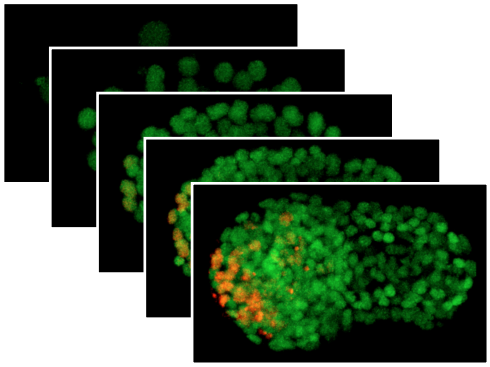
1320 76. Narasimhan K, Lambert SA, Yang AWH, Riddell J, Mnaimneh S, Zheng H, et al. Mapping
1321 and analysis of transcription factor sequence specificities. eLife. 2015;4.
1322 doi:10.7554/eLife.06967

1323 77. Bao Z, Murray JI. Mounting Caenorhabditis elegans Embryos for Live Imaging of
1324 Embryogenesis. Cold Spring Harb Protoc. 2011;2011: pdb.prot065599.
1325 doi:10.1101/pdb.prot065599

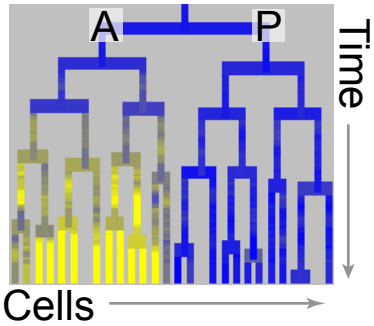
1326

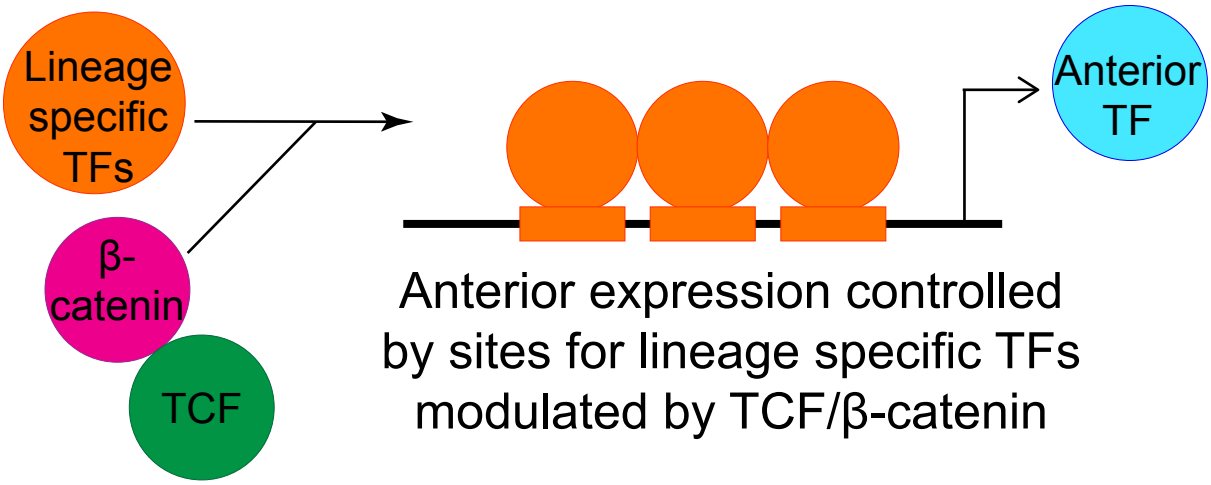
Graphical Abstract

Quantitative expression profiling by 4D imaging



Identification of anterior genes and regulatory dissection





Anterior expression controlled by sites for lineage specific TFs modulated by TCF/β-catenin