The anterior Hox gene *ceh-13* and *elt-1/GATA* activate the posterior Hox genes *nob-1* and *php-3* to specify posterior lineages in the *C. elegans* embryo

John Isaac Murray[2], Elicia Preston[2], Jeremy P. Crawford[1], Jonathan D. Rumley[2], Prativa Amom[1], Breana D. Anderson[1], Priya Sivaramakrishnan[2], Shaili D. Patel[2], Barrington Alexander Bennett[2], Teddy D. Lavon[2], Felicia Peng[2], Amanda L. Zacharias[1,3,4]

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

Abstract

Hox transcription factors play a conserved role in specifying positional identity during animal development, with posterior Hox genes typically repressing the expression of more anterior Hox genes. Here, we dissect the regulation of the posterior Hox genes *nob-1* and *php-3* in the nematode *C. elegans*. We show that *nob-1* and *php-3* are co-expressed in gastrulation-stage embryos in cells that express the anterior Hox gene *ceh-13*. This expression is controlled by several partially redundant transcriptional enhancers. Surprisingly, these enhancers require *ceh-13* for expression, providing an example of an anterior Hox gene positively regulating a posterior Hox gene. Several other regulators also act positively through *nob-1/php-3* enhancers, including *elt-1/GATA*, *ceh-20/ceh-40/Pbx*, *unc-62/Meis*, *pop-1/TCF*, *ceh-36/Otx* and *unc-30/Pitx*. We identified defects in both cell position and cell division patterns in *ceh-13* and *nob-1;php-3* mutants, suggesting that these factors regulate lineage identity in addition to positional identity. Together, our results highlight the complexity and remarkable flexibility of Hox gene regulation and function.

Introduction

Hox genes encode conserved transcription factors famously expressed along the anterior-posterior axis during animal development to specify axial position. Mutations in Hox genes cause a wide variety of developmental defects in both model organisms and humans. Hox gene regulation is complex and includes both transcriptional and post-transcriptional control. In most animals, Hox genes are found in genomic clusters, and their expression along the A-P axis is collinear with their genomic position. The Hox genes often show "posterior dominance," where posterior Hox genes repress the expression of more anterior Hox genes to establish the expression boundaries of individual Hox genes (Reviewed in (Duboule and Morata, 1994; Noordermeer and Duboule, 2013; Gaunt, 2018)).

The genome of the nematode *Caenorhabditis elegans* encodes a single set of six Hox genes, loosely organized into three degenerate "clusters" that each contain two adjacent genes

on Chromosome III (Aboobaker and Blaxter, 2003; Gaunt, 2018). In larval stages, these genes are expressed in specific positions along the A-P axis and regulate both positional differences in cell fate and function, similar to their homologs in other animals (Harris et al., 1996; Ferreira et al., 1999; Maloof et al., 1999; Shemer and Podbilewicz, 2002; Takács-Vellai et al., 2007; Tihanyi et al., 2010; Yu et al., 2010), and directly regulate terminal fates (Zheng et al., 2015).

Three of these, the anterior Hox gene *ceh-13/HOX1* and the posterior Hox genes *nob-1/HOX9-13* and *php-3/HOX9-13,* are also expressed in early embryogenesis, during gastrulation (Wittmann et al., 1997; Brunschwig et al., 1999; Van Auken et al., 2000, 2000; Zhao et al., 2010). *ceh-13* mutants have severe defects in head morphology (Brunschwig et al., 1999), while mutants lacking the functionally redundant genes *nob-1* and *php-3* have severe tail defects (Van Auken et al., 2000). Partial cell lineage tracing of *ceh-13* and *nob-1*;*php-3* mutants identified defects in cell position but not in division patterns, leading to the hypothesis that these genes regulate positional identity, rather than lineage identity (Brunschwig et al., 1999; Van Auken et al., 2000; Aboobaker and Blaxter, 2003).

Intriguingly, despite their apparently opposite roles in anterior vs posterior embryonic morphogenesis, both *ceh-13* and *nob-1* are expressed in overlapping posterior lineages during gastrulation and both require the Wnt pathway for their expression (Streit et al., 2002; Zacharias et al., 2015). *ceh-13* is transiently expressed in the progeny of 7 of the 8 posterior sister cells derived from the (largely ectoderm-producing) AB blastomere at the 24-cell stage (Wittmann et al., 1997). *nob-1* is expressed 1-2 cell cycles later, in the posterior daughters or grand-daughters of four of these ceh-13-expressing cells (Zhao et al., 2010). In addition, both *ceh-13* and *nob-1* are expressed in the posterior daughter of the intestinal blastomere E ("Ep"). This raises the question of whether *ceh-13* activates *nob-1* in these lineages. Expression of both factors at later embryonic stages is regulated by feedback mechanisms; early *ceh-13* activity is required for later ceh-13 expression (Streit et al., 2002), and early *nob-1* negatively regulates later *nob-1* expression through the microRNA *mir-57* (Zhao et al., 2010).

Here we analyzed the *cis*-regulatory control of *ceh-13* and *nob-1* expression in the early embryo. We find that nob-1 expression is regulated by several partially redundant distal enhancers, including at least three that drive overlapping patterns during gastrulation. Two of these enhancers are positively regulated by *ceh-13*, while a third appears to be regulated by a *ceh-13*-independent mechanism. We further identified the GATA family transcription factor *elt-1*, previously known as a specifier of epidermal fate, as one of several additional positive regulators of *nob-1* expression. Detailed analysis of cell positions and cell division timing identifies both position and cell cycle defects in *ceh-13* and *nob-1*/*php-3* mutants. This positive regulation of a posterior Hox gene by an anterior Hox gene suggests a novel role of Hox genes in early lineage specification.

Results

The anterior Hox gene *ceh-13* and the posterior Hox gene *nob-1* are expressed sequentially in gastrulating embryos

To better understand the expression dynamics of the *ceh-13, nob-1,* and *php-3* Hox genes during embryogenesis (Figure 1), we collected 3D time-lapse movies (~1.5 minute temporal resolution) of embryos expressing either *ceh-13, nob-1* or *php-3* tagged with GFP at their C-termini at the endogenous locus by CRISPR/Cas9 genome editing. The same embryos also expressed a ubiquitously expressed mCherry-tagged histone transgene to allow for cell

lineage tracing. We traced cells from soon after fertilization (4-8 cell stage) through the last round of cell divisions for most cells (bean stage) by using StarryNite automated cell tracking software, and quantified reporter expression in each cell across time (Bao et al., 2006; Murray et al., 2006; Richards et al., 2013; Santella et al., 2014). Note, each Hox::GFP fusion line is homozygous viable, fertile, and displays no obvious phenotypes, demonstrating that the fusion proteins largely function like the wild type proteins.

Consistent with previous studies using shorter reporters (Wittmann et al., 1997), CEH-13 expression is first seen in eight Wnt-signaled posterior sister cells and their progeny starting at the 26-cell stage, during early gastrulation (Figure 1B,C, Supplemental Figure 2). These include the AB lineage-derived cells ABalap, ABalpp, ABarpp, ABplap, ABplpp, ABprap and ABprpp, and the posterior endoderm progenitor Ep. The ABalap and ABalpp expression was barely detectable and quickly faded, while the other lineages had more robust and persistent expression. By the 200-350 cell stage, many of the granddaughters of the initially expressing cells have lost CEH-13 expression, while a few cells fated to become blast cells or neurons sustained or increased their expression. Also at this stage, some cells in the D and MS lineages as well as a few additional AB-derived cells activate CEH-13 expression (Figure 1D). We compared the reporter expression pattern to that of the endogenous mRNA as measured in a lineage-resolved single cell RNA-seq dataset and found that the mRNA and reporter expression patterns were consistent (Packer et al., 2019). In comparison, a rescuing 35 kb CEH-13::GFP fosmid transgene had brighter expression in the cells expressing the CRISPR-tagged protein, plus additional weak expression detectable in some epidermal precursors from the C lineage (possibly detectable due to higher copy number). A shorter reporter containing only 8.2 kb of upstream sequence lacked expression in the MS lineage, emphasizing the role of distal regulatory elements in regulating *C. elegans* Hox gene expression (Kuntz et al., 2008; Ho et al., 2017) (Supplemental Figure 1D).

Next, we defined the embryonic expression patterns of the posterior Hox genes *php-3* and *nob-1* tagged with GFP at the endogenous locus. As expected, *nob-1* and *php-3*, which are adjacent in the genome with *php-3* directly downstream of *nob-1*, are expressed in identical patterns in both our GFP imaging data (Supplemental Figure 1A) and in the single-cell RNA-seq data (Packer et al., 2019). NOB-1/PHP-3 expression begins during mid-gastrulation in cells derived from the cells ABplapp, ABplppp, ABprapp, ABprppp, Ep, and the four posterior great-granddaughters of the C blastomere. Expression becomes stronger one cell cycle later in the AB-derived lineages with the exception of ABp(l/r)appa. Additional late embryonic expression occurs in the ABp(l/r)papppp lineages starting at the onset of morphogenesis ("bean" stage). NOB-1-expressing lineages give rise to fates including neurons, hypodermis, seam cells, epithelial cells, intestine, and death, and have posterior positions clustered in and near the developing tail at comma stage (Sulston, 1983). Like CEH-13, early NOB-1 expression persists through the terminal cell divisions in only a subset of cells (Figure 1D). A previously described NOB-1::GFP rescuing transgene that includes 9kb of sequence upstream of the *nob-1* transcript expresses GFP in a similar pattern, except it is not robustly expressed in the C great-granddaughter lineages, suggesting this expression requires additional regulatory sequences outside of this region (Zhao et al., 2010) (Figure 1D,E, Supplemental Figure 1).

Early CEH-13::GFP expression occurs in cells that are distributed broadly along the anterior-posterior axis (Figure 1B), while NOB-1 and PHP-3 expression is more limited to the posterior of the embryo. In the AB lineage, NOB-1/PHP-3 are expressed exclusively in cells that were the posterior sister after cell division and whose mother expressed CEH-13 (Figure 1D-H). In addition, these three genes are co-expressed at similar onset times in the posterior intestine

lineage derived from Ep. In contrast, in late embryos, high levels of CEH-13::GFP and NOB-1::GFP/PHP-3::GFP are largely mutually exclusive (Figure 1D, Supplemental Figure 1F, G), consistent with classic models of Hox expression and posterior dominance.

Several overlapping lineage specific enhancers regulate *nob-1* embryonic expression

To identify regulators of *nob-1* embryonic expression, we tested nearby genomic sequences for embryonic *cis*-regulatory activity (Figure 2A,B). Previous work showed that other *C. elegans* Hox genes are regulated by distal enhancers located as much as 20kb from a given gene's promoter (Streit et al., 2002; Teng et al., 2004; Kuntz et al., 2008) but enhancers for *nob-1* and *php-3* have not been identified. We took advantage of existing reporters of different lengths to narrow the sequence search space for *nob-1* enhancers (Figure 2C). A transcriptional reporter containing just 5.3 kb of upstream sequence driving histone-mCherry reporter expression, and the rescuing 9kb NOB-1::GFP reporter are expressed in most of the same lineages as endogenous NOB-1::GFP, indicating they contain regulatory sequences sufficient for expression in full set of expressing lineages. However, while the protein fusion reporters are expressed at fairly similar levels in each lineage, the shorter 5.3kb transcriptional reporter has much lower expression in the ABp(l/r)ppp lineages compared to the other expressing lineages (Figure 2C, D). This suggests that additional sequences between -5.3kb and -9kb are required for the full endogenous expression levels in ABp(l/r)ppp.

Although other *Caenorhabditis* species have large (12-20kb) intergenic regions upstream of *nob-1*, there is no detectable sequence conservation in the 5.3kb region upstream of *nob-1*, and only two short conserved stretches (85 and 275 nucleotides long) between 5.3 and 9kb upstream. This indicates that any regulatory elements in this region have diverged substantially at the sequence level during *Caenorhabditis* evolution, limiting utility evolutionary conservation to identify enhancers. Chromatin marks typically used to identify transcriptional enhancers such as H3K27ac, H3K4me, or chromatin accessibility from whole embryos did not show strong peaks near *nob-1* nor 20 additional genes expressed in lineage-specific patterns at the same stage as *nob-1* (Daugherty et al., 2017; Ho et al., 2017; Jänes et al., 2018).

We hypothesized that these signals are diminished in early embryonically expressed genes because the experiments used embryos of mixed stages where later stage nuclei dramatically outnumber nuclei from earlier stages. Therefore, we searched for other factors with published genomic binding patterns that could mark embryonic enhancers. One pattern stood out as preferentially bound near genes expressed with lineage-specific patterns in the early embryo: binding of an NHR-2::GFP fusion protein (Niu et al., 2011; Araya et al., 2014). NHR-2::GFP binds in the intergenic sequences upstream of 19 of these 20 genes, and 16 genes have multiple clustered NHR-2 binding sites, compared with 22% and 8% of randomly chosen genes, respectively ($p < 0.001$). NHR-2 is a nuclear hormone receptor distantly related to mammalian thyroid and PPAR receptors, and NHR-2::GFP is expressed in most or all somatic cells from the ~50-cell to ~200-cell stage. The functional importance of NHR-2 binding is unclear; *nhr-2* RNAi causes embryonic and larval arrest, but partial deletion alleles are viable suggesting the RNAi phenotype could be non-specific. However, regardless of its function, we hypothesized that these clustered NHR-2 binding sites might be useful proxies for accessible chromatin and could predict enhancer activity in the early embryo.

We tested four NHR-2::GFP-bound regions upstream of *nob-1* for enhancer activity by generating transgenic worms expressing histone-mCherry under the control of each candidate enhancer placed upstream of a *pes-10* minimal promoter, which drives no consistent embryonic

expression on its own (Figure 2A,B). We also tested four additional sequences for which we observed no NHR-2::GFP binding but which contained putative binding motifs for POP-1/TCF, which is required for expression of the *nob-1* transcriptional reporter (Zacharias et al., 2015). We identified embryonic cells expressing each enhancer reporter by confocal time lapse imaging and StarryNite. Only three of the regions (-3.4kb, -4.3kb and -8.3kb) tested showed consistent enhancer activity that overlapped with the endogenous NOB-1::GFP expression pattern, suggesting they represent functional enhancers (Figure 2C, Supplemental Figure 2). All three functional enhancers were identified on the basis of NHR-2 binding.

The -3.4kb enhancer recapitulates most of the *nob-1* early embryonic expression pattern with some differences in the C lineage and drives much weaker expression in ABp(l/r)ppp than in ABp(l/r)app, similar to the *nob-1* -5.3 kb upstream transcriptional reporter (Figure 2C-E). The -4.3kb enhancer drives expression in ABp(l/r)app and ABp(l/r)ppp at a very high level (Figure 2D) and also variable misexpression in cells that don't normally express NOB-1 (Figure 2C, Supplemental Figure 2A). When this enhancer was cloned downstream of the reporter, it showed less ectopic expression and stronger expression in ABp(l/r)app relative to ABp(l/r)ppp, similar to the -3.4 kb enhancer. This confirms the -4.3kb region as a *bona fide* enhancer capable of acting at a distance and suggests that the placement of this enhancer relative to the promoter influences its activity differently in different lineages and may be important for specificity. Embryos carrying multiple copies of the -4.3kb reporter transgene occasionally displayed the "no backend" phenotype observed in *nob-1/php-3* mutant embryos, indicating this enhancer binds key regulators of *nob-1* that may be drawn away from the endogenous locus resulting in a loss of function phenotype (Supplemental Figure 2B). A third *nob-1* enhancer (-8.3kb) drives early embryonic expression only in ABp(l/r)ppp, as well as later expression in ABp(l/r)papppp. This most-distal enhancer is included in -9kb NOB-1::GFP transgene but not the shorter -5.3kb transcriptional reporter and includes the only substantial stretch of conserved sequence in the *nob-1* promoter region. This region thus can explain why ABp(l/r)ppp expression is relatively stronger in the translational reporter compared to the shorter transcriptional reporter.

As the -5.3kb *nob-1* transcriptional reporter includes both the -3.4kb and -4.3kb enhancers, we deleted each enhancer from this construct to test their necessity for *nob-1* expression (Figure 2C). Deleting both enhancers led to a complete loss of early embryonic expression in both AB lineages, suggesting there are no additional enhancers present sufficient for early embryonic expression in these lineages. The deletion of the -3.4 kb enhancer did not disrupt expression in ABp(l/r)app, ABp(l/r)ppp and in the absence of the -4.3 kb enhancer, this enhancer was not sufficient to rescue expression in these lineages. This indicates that while the -3.4kb enhancer may be sufficient to drive expression when placed directly next to the promoter, it is unable to drive expression in the endogenous context. Conversely, the -4.3 kb enhancer is a key driver of expression in these lineages. The expression remaining when both enhancers are lost indicates that additional sequences that drive the expression in the Ep and Cpapp lineages must exist within the -5.3 kb upstream region.

ceh-13 **and Hox cofactors activate** ***nob-1*** **expression**

Because *ceh-13* expression precedes that of *nob-1* in many lineages, we asked whether *ceh-13* is required for *nob-1* expression. To evaluate this, we examined *nob-1* reporter expression in embryos homozygous for the likely null mutation *ceh-13(sw1)* (Brunschwig et al., 1999). Expression of the *nob-1* -5.3kb transcriptional reporter is eliminated in ABp(l/r)ppp ($p < 0.001$) and decreased by 25% in ABp(l/r)ap, and 40% in Ep ($p < 0.02$) (Figure 3A-C). Depletion of *ceh-13* by RNAi gave similar results with the transcriptional reporter ($p < 0.002$). In contrast,

ceh-13 RNAi does not significantly alter expression of the -9kb NOB-1::GFP transgene, indicating that the additional cis-regulatory elements in this reporter which are *ceh-13* independent (Supplemental Figure 3A).

To determine which enhancers require *ceh-13*, we measured the activity of individual *nob-1* enhancer reporters in worms lacking *ceh-13* (Figure 3A, Supplemental Figure 3A). In *ceh-13* mutants, expression driven by the -3.4 kb enhancer decreased significantly in all expressing cells, excluding the C lineage ($p < 0.003$) with an average decrease of 90% and complete loss in ABp(l/r)ppp. *ceh-13* loss also reduces activity of the -4.3 kb enhancer in ABp(l/r)ppp (47% decrease) but expression in ABp(l/r)ap is unchanged. This indicates that other factors besides *ceh-13* activate this enhancer in ABp(l/r)ap and suggests that this enhancer may be responsible for the residual expression of the 5.3kb transcriptional regulation in this lineage in *ceh-13* mutants. The distal -8.3kb enhancer also lost expression in ABp(l/r)ppp in the *ceh-13* mutant($p < 0.047$), indicating its activity also requires *ceh-13*. Motif analysis identified six putative *ceh-13*-binding motifs in the -3.4kb enhancer; mutating these sites abolished expression in the AB and C lineage ($p < 0.02$, $p < 0.03$), but did not alter expression in the Ep lineage indicating *ceh-13* regulation of *nob-1* may be indirect in the E lineage (Figure 3A, Supplemental Figure 3B). These results show that multiple *ceh-13*-dependent enhancers work together to regulate *nob-1* expression in gastrulating embryos.

Expression of classically defined Hox targets often require Hox cofactors such as *homothorax* or *extradenticle*. These cofactors form larger TF complexes with Hox factors to increase binding specificity and may also have Hox-independent functions (Van Auken et al., 2002; Slattery et al., 2011; Abe et al., 2015). To determine when and where Hox cofactors are expressed in early embryos, we used StarryNite to trace the expression of translational reporters for *unc-62*, the *C. elegans homothorax/Meis* homolog, and of *ceh-20* and *ceh-40*, the homologs of *extradenticle/Pbx*. A third *extradenticle* homolog, *ceh-60*, is only expressed in later development (>200 minutes) and does not overlap with *ceh-13*, so was not investigated further (Packer et al., 2019). We found that each of these genes has specific and dynamic expression patterns that overlap with each other and with *ceh-13* and *nob-1* expression (Figure 4A-D, Supplemental Figure 5). In particular, all three cofactors are coexpressed with CEH-13 in the AB-derived cells that will later express NOB-1. Other CEH-13 and NOB-1 expressing cells also express all three cofactors, except only CEH-20 is expressed in the early E lineage. We conclude that Hox cofactors are expressed in cells where *ceh-13* activates *nob-1* expression.

We tested whether Hox cofactors regulate *nob-1* expression by examining *nob-1* reporter expression after loss of each gene. We found that depleting *unc-62* by RNAi led to significant loss of expression of the *nob-1* -5.3kb reporter in all AB lineages (80% decrease, $p=0.001$), with near complete loss in the ABp(l/r)apap and ABp(l/r)ppp lineages (Figure 4E). The -9kb NOB-1:GFP protein reporter showed similar expression decreases, although this was only significant in ABp(l/r)appa and incompletely penetrant (Supplemental Figure 4B-C). In contrast, CEH-13::GFP reporter expression was unchanged after *unc-62* RNAi (Supplemental Figure 4D). These results show that *nob-1* is regulated by *unc-62*. The fact that *unc-62* RNAi has a stronger phenotype in ABp(l/r)appa than loss of *ceh-13* suggests that *unc-62* may regulate *nob-1* independently of *ceh-13* in these cells.

To investigate the role of the *extradenticle* homologs, we examined the *nob-1* transcriptional reporter in a *ceh-40(gk159)* null mutant background in embryos of worms treated with control or ceh-20 RNAi. We found that the combination of *ceh-20* RNAi with the *ceh-40* mutation caused a significant loss of expression in virtually all expressing cells including Ep and

Cpap as compared to control RNAi (97% decrease, $p < 0.004$)(Figure 4F). This indicates that the *exd* homologs *ceh-20* and *ceh-40* are required to activate *nob-1* expression and that they are at least partially independent of *ceh-13* in the ABp(l/r)ap, Ep, and Cpap lineages.

Multiple lineage-specific embryonic TFs including *elt-1/GATA*, *ceh-36/OTX* and *unc-30/PITX* are required for early *nob-1* expression

Since some NOB-1::GFP reporter expression remains in the absence of *ceh-13*, especially in the ABp(l/r)ap lineage, other factors must also activate *nob-1* in these cells. Indeed, we previously identified the Wnt effectors *pop-1* and *sys-1* as activators of *nob-1* expression (Zacharias et al., 2015). We took advantage of databases of TF expression to search for additional candidate regulators (Mace et al., 2013; Murray et al., 2012; Packer et al., 2019).

Only two of the three *nob-1* enhancers (-3.4kb and -4.3kb) drive expression in the ABp(l/r)ap lineage, so to identify additional potential regulators of *nob-1* in the ABp(l/r)ap lineage, we used two criteria. First, we searched for TFs expressed specifically in ABp(l/r)ap. Second, we used a TF binding site scanning approach to identify which of these TFs have predicted binding motifs in the -3.4kb and -4.3kb enhancers (Weirauch et al., 2014; Narasimhan et al., 2015). From this analysis, the GATA Zinc finger TF encoded by *elt-1* emerged as the strongest candidate. An ELT-1::GFP fosmid reporter is expressed prior to gastrulation in the ABp(l/r)ap progenitors ABpla and ABpra, in the other major epidermis-producing lineages ABarp, Caa and Cpa, and in the primarily neuronal lineages ABalap and ABalppp (Figure 5A, Supplemental Figure 5). ELT-1::GFP expression is later downregulated in cells that do not adopt epidermal fates (Figure 5B). ELT-1::GFP expression in the ABpla/ABpra lineages could be detected at least 20 minutes prior to *nob-1* reporter expression in ABp(l/r)ap (Figure 5B,C). GATA motifs predicted to bind ELT-1 are enriched in the ABp(l/r)ap enhancers -3.4kb (8 sites) and -4.3kb (7 sites) enhancers compared to the -8.3 enhancer (1 site), which does not drive expression in ABp(l/r)ap.

To test whether *elt-1* regulates *nob-1* expression, we measured the activity of *nob-1* reporters in the *elt-1(ok1002)* null mutant background (Figure 5E-G, Supplemental Figure 6A,B). We found that *elt-1* mutants showed significantly decreased expression of the transcriptional reporter in the ABp(l/r)ap lineage (87% decrease, $p < 0.014$), and the Cpap lineage (92% decrease, not significant), in which *elt-1* precedes *nob-1* expression, but not in ABp(l/r)pp, where ELT-1::GFP is not expressed. Similarly, we observed decreased expression in the same lineages for the -3.4kb enhancer reporter (ABp(l/r)ap: 72% decrease, $p < 0.008$; C: 71% decrease, $p < 0.03$) and the -4.3kb enhancer reporter (ABp(l/r)ap: 31% decrease, not significant). To determine if *elt-1* directly regulates the -3.4kb enhancer we mutated the predicted ELT-1 binding sites, which have a characteristic GATA motif and observed loss of reporter expression in all lineages ($p < 0.001$). We conclude that *elt-1* activates *nob-1* expression in the ABp(l/r)ap lineage through at least two enhancers and is at least partially redundant with other factors that may bind GATA motifs.

We previously showed that two homeodomain family TFs, *ceh-36/OTX* and *unc-30/PITX* are redundantly required for proper development of the ABp(l/r)p lineage (Walton et al., 2015). Since *ceh-13* and *elt-1* mutants still have some *nob-1* reporter expression in the ABp(l/r)p descendants, we hypothesized that *ceh-36* and *unc-30* might be important for *nob-1* expression in these lineages. To test this, we examined NOB-1::GFP transgene expression in eight *ceh-36;unc-30* double mutant embryos and found that two of the eight embryos (25%) lost

expression in this lineage (Supplemental Figure 7A). Furthermore, the ratio of ABp(l/r)ap to ABp(l/r)ppp expression significantly increased ($p < 0.002$) in the remaining double mutant embryos, consistent with a decrease in ABp(l/r)ppp expression (Supplemental Figure 7B). In conclusion, at least seven partially redundant lineage specific transcription factors positively regulate *nob-1* expression: *ceh-13, ceh-20, ceh-40, unc-62, elt-1*, *ceh-36* and *unc-30*.

ceh-13 and nob-1 are required for correct cell position and division patterns in the early embryo

Previous work suggested that *nob-1/php-3* and *ceh-13* are required for cell position but not other aspects of cell fate specification (Van Auken et al., 2000). However, mutating lineage identity regulators often leads to cell cycle defects (Boeck et al., 2011; Walton et al., 2015) that could have been missed in previous studies. To test for such defects, we measured cell positions and cell division timing in six *nob-1/php-3 (*referred to as *nob-1* for simplicity*)* and *ceh-13* mutant embryos by time-lapse microscopy and StarryNite cell tracking and compared each to a database of 17 wild type embryos (Richards et al., 2013; Walton et al., 2015) (Figure 6).

Consistent with previous work (Van Auken et al., 2002), we observed striking global patterns of cell position defects in *nob-1* mutant embryos. The most severe and highest penetrance defects were in cells from NOB-1-expressing lineages (Figure 6A,C,D,G, Supplemental Figure 8)). In total, 48% (62 of 127) of cells descended from NOB-1::GFP expressing lineages had severe position defects (>5 micron deviation, z score > 3.5) in at least two embryos compared to 1.2% (14 of 1094) of cells from non-expressing lineages ($p << 10^{-20}$). However, a large number of cells from non-expressing lineages were moderately mis-positioned, suggesting that many defects are not cell-autonomous. Globally, dorsal cells that normally express NOB-1 fail to migrate to the posterior at 230 minutes post-fertilization; by 300 minutes non-expressing ventral cells compensate by inappropriately moving to the posterior, resulting in a global counterclockwise rotation of cell positions when the embryo is viewed from the left aspect (Figure 6G). The position defects we observed are consistent with the severe posterior defects observed in rare surviving larvae.

We also identified broad cell position defects in *ceh-13* mutant embryos (Figure 6B,E,F,H,I). Intriguingly, the global patterns of cell position defects differed from *nob-1* mutants. The number of severe position defects was lower for *ceh-13* than for *nob-1*, with 4% (15/357) of cells descended from CEH-13::GFP expressing lineages severely displaced in at least two embryos, as opposed to 0.5% (5/874) of cells from non-expressing lineages ($p << 10^{-20}$). At 210 minutes, (200 cell stage), the most strongly mispositioned cells were *ceh-13* expressing cells that do not express *nob-1* in the posterior ventral region of the embryo. At this stage, cells on the dorsal surface were posteriorly mispositioned, such that cells in the posterior half of the embryo were rotationally mispositioned clockwise, the opposite direction to that seen in *nob-1* mutants (Figure 6H). Cells at the anterior ventral surface (both *ceh-13*-expressing and non-expressing) were mispositioned towards the posterior, resulting in a counterclockwise rotational defect in the anterior half of the embryo. At later stages, these global defects were largely resolved, and after the onset of morphogenesis the position defects were largely limited to a small number of strongly mispositioned cells (Figure 6I). The most extreme were the DA and SAB motorneurons after 300 minutes (bean stage), which were mispositioned several cell diameters anterior of their wild-type position (Supplemental Figure 8B). Given the recovery of most cell positions by the beginning of morphogenesis, it is unclear if cell position defects contribute to the severe anterior morphogenesis phenotypes in *ceh-13* mutants that hatch

(Brunschwig et al., 1999), as the earlier cell position defects could disrupt normal cell-cell signaling interactions required for fate specification.

In contrast to previous studies (Brunschwig et al., 1999; Van Auken et al., 2000), we also identified numerous cell cycle defects in both *nob-1* and *ceh-13* mutants, and these were also heavily restricted to expressing lineages (Figure 6A,B). 13% of *nob-1* expressing cells and 5% of *ceh-13* expressing cells had severe cell division timing defects in at least two of the corresponding mutant embryos (>5 minute deviation from wild type, and z-score >5), vs none in non-expressing cells ($p << 10^{-20}$). The most common cell cycle defect in *nob-1* mutants was the PLM/ALN neuroblast mother (ABp(l/r)apappp), which failed to divide in 8 of 12 observations (Figure 6J). The neuroblast that produces the PHA, PVC and LUA neurons instead underwent programmed cell death in two of 12 mutant lineages. Other cells that had multiple cell cycle delays in *nob-1* mutants included the PVQ neuroblast, PHB/HSN neuroblast, and the mother of the epidermal cells PHsh, hyp8 and hyp9. All of these cells also were mispositioned (mean 5 micron deviation from expected position, Supplemental Figure 8C). The most striking defects in *ceh-13* mutants were in the cells ABarppaaa(a/p), which normally differentiate into the ADE sheath (glial) cell and H2 epidermal cell, but instead each divided inappropriately in 4 of 6 mutant embryos (Figure 6K). In addition, ABplppaaap, which normally undergoes programmed cell death, divided in two *ceh-13* mutant embryos. Several neuroblasts including those producing the DA, DD and SAB motorneurons also had cell division delays in *ceh-13* mutants. We conclude that both *ceh-13* and *nob-1/php-3* are required not only for proper cell positioning, but also for normal division patterns, suggesting a broader role in fate specification. Furthermore, the incomplete penetrance of both cell cycle and cell position defects in *ceh-13* and *nob-1* mutants indicate redundancy in developmental programming that contributes to the incredible robustness observed in *C. elegans* embryonic development (Sulston, 1983).

Discussion

Classic work defined the concept of Posterior Dominance, in which posterior Hox genes repress the expression and activity of more anterior Hox genes (Duboule and Morata, 1994). Indeed, previous work in later developmental stages showed that *nob-1* represses *ceh-13* in the posterior ventral nerve cord, consistent with posterior dominance (Van Auken et al., 2002). In contrast, we show that *ceh-13* activates *nob-1/php-3* expression through at least two enhancers during *C. elegans* gastrulation. To our knowledge, this is the first example of an anterior Hox gene positively regulating the expression of a posterior Hox gene during development. *ceh-13* and *nob-1/php-3* are also expressed at an earlier phase in development (pre-gastrulation for *ceh-13* and mid-gastrulation for *nob-1/php-3*) than in many other organisms. Since *C. elegans* is an unsegmented organism, this may have enabled these Hox genes to be co-opted to play earlier “lineage identity” roles in development than is typical for genes in the cluster. This demonstrates an unexpected flexibility in the Hox cluster.

Several studies have highlighted that conserved transcription factors often “moonlight,” with the same factor playing apparently distinct roles in early progenitors and in terminal cell types. For example, the OTX homeodomain TF *ceh-36*, and the PITX homeodomain TF *unc-30*, which both specify specific terminal neuron types, each act redundantly 6-7 cell divisions earlier in the ABp(l/r)p lineages to specify broad features of lineage identity (Jin et al., 1994; Lanjuin et al., 2003; Chang et al., 2003; Walton et al., 2015), and *skn-1*, which is required maternally to specify the endomesodermal blastomere EMS at the 4-cell stage, is required postembryonically for oxidative stress resistance and longevity (Bowerman et al., 1992; An et al., 2005). Our work shows that *ceh-13* and *nob-1/php-3* are another example of this phenomenon. While these were

previously shown to regulate positional differences between cells, our identification of cell division defects in each mutant demonstrates they also play a role in lineage specification. Previous work showed that at least one neuron (the PLM touch receptor) is defective in less severe nob-1 mutants (Zheng et al., 2015); our work shows that the PLM neuroblasts do not divide normally in *nob-1* mutants, indicating an earlier function in PLM progenitor lineages. Similarly, *ceh-13* mutants have cell division defects, and *ceh-13* is required for the lineage-specific activity of specific *nob-1* enhancers. Finally, while *elt-1* has a well-defined known role in epidermal fate specification (Gilleard and McGhee, 2001), our work shows that it is also required for *nob-1* enhancer activity in non-epidermal cells. This is consistent with other recent work identifying defects in *elt-1* mutants in cell division patterns and the expression of the neurogenic TF *lin-32* (Ma et al., 2020). The ability of developmental transcription factors to "moonlight' is likely facilitated by complex regulation by multiple enhancers as seen for *nob-1/php-3*, and *ceh-13* since enhancer evolution can enable genes to be expressed at different locations and times in development (Carroll, 2005; Levine, 2010).

While early reporter studies found that the expression of many genes expressed in terminal cells is well approximated using just the promoter-proximal region (Dupuy et al., 2004; Hunt-Newbury et al., 2007; Reece-Hoyes et al., 2007), open chromatin mapping studies have identified many distal regions of open chromatin, many of which can act as enhancers (Daugherty et al., 2017; Ho et al., 2017; Jänes et al., 2018). Our work reinforces these studies and emphasizes the importance of enhancers in *C. elegans* gene regulation. Intriguingly, other genes expressed in lineage-specific patterns around the same time as *nob-1* have similar tendency to have multiple nearby regions of NHR-2::GFP binding and lineage-specific accessible chromatin as measured by ATAC-seq (Charest et al., 2020), suggesting they may also be regulated by multiple distal enhancers. This could reflect these genes' need for more complex regulation. The *nob-1/php-3* locus contains at least two or three enhancers responsible for expression in each expression domain during gastrulation. This is reminiscent of so-called "shadow enhancers" identified in other organisms, for which redundancy appears to confer robustness in the face of environmental variability (Hong et al., 2008; Frankel et al., 2010; Perry et al., 2010). This high level of redundancy also appears to extend to the number of transcription factors regulating each enhancer. Current data implicate at least eight TFs as activators of *nob-1* enhancers including *ceh-13*, *elt-1*, *ceh-36*, *unc-30*, the Wnt effectors *pop-1*, and the Hox co-factors *ceh-20*, *ceh-40* and *unc-62* (Supplemental Figure 9). In some cases, the same binding motif may be regulated by different factors in different lineages. For example, mutating six GATA motifs in the -3.4 kb enhancer caused the loss of expression in ABp(l/r)ap, potentially due to loss of ELT-1 binding. However, this mutant also lost expression in Ep, which does not express *elt-1*, but which does express other GATA factors with similar binding specificity including *end-1*, *end-3*,and *elt-7* (Maduro, 2017). Similarly, the more severe loss of *nob-1* enhancer expression when CEH-13 motifs are mutated than in *ceh-13* mutant animals indicates that other TFs may also bind these sites. In sum, multiple modes of regulatory redundancy could provide a mechanism for the remarkable developmental robustness observed in *C. elegans.*

Methods

Strain generation and propagation

Worm strains (Supplemental Table 1) were maintained at 21-23°C on OP50 *E. coli* on NGM plates. Enhancer reporter strains were generated by injection into RW10029, a GFP histone strain used for lineage tracing. Injection cocktails consisted of reporter DNA construct at 10ng/µL, with 5 ng/µL myo-2p::GFP and 135 ng/µL pBluescript vector and were injected using a

Narishige MN-151 micromanipulator with Tritech microinjector system. The *nob-1* promoter deletion strains were generated by injection and compared to a wild-type *nob-1* promoter strain created by injection of pJIM20::nob-1, JIM518. The *ceh-13* rescuing fosmid, *ujIs153*, was created by bombardment of WRM0622C_C06 (Sarov et al., 2012) as previously described (Walton et al., 2015). Other strains were created through crosses using standard approaches. Lethal alleles were maintained as balanced heterozygotes (*ceh-13(sw1))* or rescued by free duplications (*nob-1(ct223)*), with homozygous mutants recognized by characteristic high-penetrance morphology defects. Note, the *ct223* allele disrupts both *nob-1* and *php-3*. Lethal *ceh-36;unc-30* double mutants were maintained as homozygotes rescued by an extrachromosomal array carrying *ceh-36*::GFP; lack of rescue was scored by absence of GFP in the ABpxpa lineage. CEH-13::GFP and NOB-1::GFP CRISPR knock-in alleles were created under contract by SUNY Biotech (Fuzhou, China).

Molecular biology

Candidate enhancers were amplified with Phusion HF polymerase (New England Biosciences) from either pJIM20::nob-1 or N2 genomic DNA, with overhangs for stitching, which were then either gel or PCR purified (Qiagen). Putative enhancers were attached to a *pes-10* minimal promoter::HIS-24::mCherry::let-858 3'UTR fragment amplified from POPTOP plasmid (Green et al., 2008)(Addgene #34848) by using PCR stitching to create an enhancer reporter which was sequence verified and purified with a PureLink PCR purification kit (ThermoFisher) for injection. The *pes-10* minimal promoter drives no consistent early embryonic expression alone but has previously been shown to facilitate expression driven by different enhancers (Fire et al., 1990; Green et al., 2008; Wei et al., 2012). For the downstream enhancer experiment, the -4.3kb enhancer was cloned immediately downstream of the stop codon. Tested enhancer sequences can be found in Supplemental Table 2. Putative transcription factor binding sites were identified using CIS-BP (cisbp.ccbr.utoronto.ca)(Weirauch et al., 2014; Narasimhan et al., 2015) (Supplemental Table 3-5). DNA fragments with desired mutations were synthesized (Integrated DNA Technologies) and PCR stitched as before.

Imaging/Lineaging

We acquired confocal images with a Leica TCS SP5, Stellaris or Nikon A1RSi resonance scanning confocal microscope (67 z planes at 0.5 µm spacing and 1.5 minute time spacing, with laser power increasing by 4-fold through the embryo depth to account for attenuation of signal with depth). Embryos were mounted in egg buffer/methyl cellulose with 20µm beads as spacers (Bao and Murray, 2011) and imaged at 22°C using a stage temperature controller (Brook Industries, Lake Villa, IL). We used StarryNite software to automatically annotate nuclei and trace lineages, and AceTree software to identify and fix any errors from the automated analysis, and quantified reporter expression in each nucleus relative to local background (using the "blot" background correction technique) as previously described (Bao et al., 2006; Boyle et al., 2006; Murray et al., 2006, 2008; Santella et al., 2014).

Quantitative comparisons

Cell averages of nuclear fluorescence were computed for each cell based on all (typically >20) measurements across its lifetime. For each control condition, the average value of each expressing cell was computed and used to determine the fold change for each control and mutant cell, which were displayed as boxplots using R 3.5.1 (The R Foundation for Statistical Computing). Background-corrected expression values were rounded up to zero if values fell below zero. To determine if changes were significant across the lineage, the summed expression for each cell in each lineage was calculated for all control and mutant embryos and

the two sets of values were compared using a Wilcoxon Ranked Sum test using R. To evaluate the enrichment of NHR-2 binding sites, a chi-squared test was used.

Mutant cell position analysis

Cell position defects were identified as previously (Walton et al., 2015). Briefly, we corrected for differences in global division rates (which did not differ dramatically from wild-type) and considered divisions as defective/outliers if they deviated from the wild-type cell cycle length by at least five minutes and had a z-score greater than three. Cell positions were corrected for differences in embryo size and rotation and considered defective if they deviated from the expected wild-type position by at least five microns, had a z-score greater than five, and a nearest neighbor score greater than 0.8 (defined empirically based on the distribution of wild-type scores).

Acknowledgements

We thank Brian Gebelein and members of the Murray, Zacharias, and Gebelein laboratories and the Philly Worm Group for helpful discussion and comments on the manuscript. We thank Bob Waterston and Adrian Streit for strains. Some strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). We thank Swati Mundre, Tom Sesterhenn, Yannis Belloucif, Ilona Jileaeva for cloning assistance and technical assistance with lineage tracing and worm genetics. We thank Meera Sundaram for the use of injection equipment and reagents, Stephen Nehrbass for software assistance, Matt Kofron, Evan Meyer, and the Confocal Imaging Core at Cincinnati Children's, a Nikon Center of Excellence, for their help with imaging, and Matt Batie for designing and building a cooling stage. This work was supported by NIH grants R35GM127093 (to J.I.M.), R00GM111825 (to A.L.Z.), and F31GM123737 (to J.D.R.).

Competing interests

The authors affirm that they have no competing interests related to this research.

References

Abe, N., Dror, I., Yang, L., Slattery, M., Zhou, T., Bussemaker, H.J., Rohs, R., Mann, R.S., 2015. Deconvolving the recognition of DNA shape from sequence. Cell 161, 307–318. https://doi.org/10.1016/j.cell.2015.02.008

Aboobaker, A., Blaxter, M., 2003. Hox gene evolution in nematodes: novelty conserved. Curr. Opin. Genet. Dev. 13, 593–598. https://doi.org/10.1016/j.gde.2003.10.009

An, J.H., Vranas, K., Lucke, M., Inoue, H., Hisamoto, N., Matsumoto, K., Blackwell, T.K., 2005. Regulation of the Caenorhabditis elegans oxidative stress defense protein SKN-1 by glycogen synthase kinase-3. Proc. Natl. Acad. Sci. U. S. A. 102, 16275–16280. https://doi.org/10.1073/pnas.0508105102

Araya, C.L., Kawli, T., Kundaje, A., Jiang, L., Wu, B., Vafeados, D., Terrell, R., Weissdepp, P., Gevirtzman, L., Mace, D., Niu, W., Boyle, A.P., Xie, D., Ma, L., Murray, J.I., Reinke, V.,

Waterston, R.H., Snyder, M., 2014. Regulatory analysis of the C. elegans genome with spatiotemporal resolution. Nature 512, 400–405. https://doi.org/10.1038/nature13497
Bao, Z., Murray, J.I., 2011. Mounting Caenorhabditis elegans Embryos for Live Imaging of Embryogenesis. Cold Spring Harb. Protoc. 2011, pdb.prot065599. https://doi.org/10.1101/pdb.prot065599
Bao, Z., Murray, J.I., Boyle, T.J., Ooi, S.-L., Sandel, M.J., Waterston, R.H., 2006. Automated cell lineage tracing in Caenorhabditis elegans. Proc. Natl. Acad. Sci. 103, 2707–2712. https://doi.org/10.1073/pnas.0511111103
Boeck, M.E., Boyle, T., Bao, Z., Murray, J., Mericle, B., Waterston, R., 2011. Specific roles for the GATA transcription factors end-1 and end-3 during C. elegans E-lineage development. Dev. Biol. 358, 345–355. https://doi.org/10.1016/j.ydbio.2011.08.002
Bowerman, B., Eaton, B.A., Priess, J.R., 1992. skn-1, a maternally expressed gene required to specify the fate of ventral blastomeres in the early C. elegans embryo. Cell 68, 1061–1075. https://doi.org/10.1016/0092-8674(92)90078-q
Boyle, T.J., Bao, Z., Murray, J.I., Araya, C.L., Waterston, R.H., 2006. AceTree: a tool for visual analysis of Caenorhabditis elegans embryogenesis. BMC Bioinformatics 7, 275. https://doi.org/10.1186/1471-2105-7-275
Brunschwig, K., Wittmann, C., Schnabel, R., Bürglin, T.R., Tobler, H., Müller, F., 1999. Anterior organization of the Caenorhabditis elegans embryo by the labial-like Hox gene ceh-13. Dev. Camb. Engl. 126, 1537–1546.
Carroll, S.B., 2005. Evolution at two levels: on genes and form. PLoS Biol. 3, e245. https://doi.org/10.1371/journal.pbio.0030245
Chang, S., Johnston, R.J., Hobert, O., 2003. A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of C. elegans. Genes Dev. 17, 2123–2137. https://doi.org/10.1101/gad.1117903
Charest, J., Daniele, T., Wang, J., Bykov, A., Mandlbauer, A., Asparuhova, M., Röhsner, J., Gutiérrez-Pérez, P., Cochella, L., 2020. Combinatorial Action of Temporally Segregated Transcription Factors. Dev. Cell 55, 483-499.e7. https://doi.org/10.1016/j.devcel.2020.09.002
Daugherty, A.C., Yeo, R.W., Buenrostro, J.D., Greenleaf, W.J., Kundaje, A., Brunet, A., 2017. Chromatin accessibility dynamics reveal novel functional enhancers in C. elegans. Genome Res. 27, 2096–2107. https://doi.org/10.1101/gr.226233.117
Duboule, D., Morata, G., 1994. Colinearity and functional hierarchy among genes of the homeotic complexes. Trends Genet. TIG 10, 358–364. https://doi.org/10.1016/0168-9525(94)90132-5
Dupuy, D., Li, Q.-R., Deplancke, B., Boxem, M., Hao, T., Lamesch, P., Sequerra, R., Bosak, S., Doucette-Stamm, L., Hope, I.A., Hill, D.E., Walhout, A.J.M., Vidal, M., 2004. A First Version of the Caenorhabditis elegans Promoterome. Genome Res. 14, 2169–2175. https://doi.org/10.1101/gr.2497604
Ferreira, H.B., Zhang, Y., Zhao, C., Emmons, S.W., 1999. Patterning of Caenorhabditis elegans posterior structures by the Abdominal-B homolog, egl-5. Dev. Biol. 207, 215–228. https://doi.org/10.1006/dbio.1998.9124
Fire, A., Harrison, S.W., Dixon, D., 1990. A modular set of lacZ fusion vectors for studying gene expression in Caenorhabditis elegans. Gene 93, 189–198. https://doi.org/10.1016/0378-1119(90)90224-f
Frankel, N., Davis, G.K., Vargas, D., Wang, S., Payre, F., Stern, D.L., 2010. Phenotypic robustness conferred by apparently redundant transcriptional enhancers. Nature 466, 490–493. https://doi.org/10.1038/nature09158
Gaunt, S.J., 2018. Hox cluster genes and collinearities throughout the tree of animal life. Int. J. Dev. Biol. 62, 673–683. https://doi.org/10.1387/ijdb.180162sg

Gilleard, J.S., McGhee, J.D., 2001. Activation of hypodermal differentiation in the Caenorhabditis elegans embryo by GATA transcription factors ELT-1 and ELT-3. Mol. Cell. Biol. 21, 2533–2544. https://doi.org/10.1128/MCB.21.7.2533-2544.2001

Green, J.L., Inoue, T., Sternberg, P.W., 2008. Opposing Wnt pathways orient cell polarity during organogenesis. Cell 134, 646–656. https://doi.org/10.1016/j.cell.2008.06.026

Harris, J., Honigberg, L., Robinson, N., Kenyon, C., 1996. Neuronal cell migration in C. elegans: regulation of Hox gene expression and cell position. Dev. Camb. Engl. 122, 3117–3131.

Ho, M.C.W., Quintero-Cadena, P., Sternberg, P.W., 2017. Genome-wide discovery of active regulatory elements and transcription factor footprints in Caenorhabditis elegans using DNase-seq. Genome Res. 27, 2108–2119. https://doi.org/10.1101/gr.223735.117

Hong, J.-W., Hendrix, D.A., Levine, M.S., 2008. Shadow enhancers as a source of evolutionary novelty. Science 321, 1314. https://doi.org/10.1126/science.1160631

Hunt-Newbury, R., Viveiros, R., Johnsen, R., Mah, A., Anastas, D., Fang, L., Halfnight, E., Lee, D., Lin, J., Lorch, A., McKay, S., Okada, H.M., Pan, J., Schulz, A.K., Tu, D., Wong, K., Zhao, Z., Alexeyenko, A., Burglin, T., Sonnhammer, E., Schnabel, R., Jones, S.J., Marra, M.A., Baillie, D.L., Moerman, D.G., 2007. High-throughput in vivo analysis of gene expression in Caenorhabditis elegans. PLoS Biol. 5, e237. https://doi.org/10.1371/journal.pbio.0050237

Jänes, J., Dong, Y., Schoof, M., Serizay, J., Appert, A., Cerrato, C., Woodbury, C., Chen, R., Gemma, C., Huang, N., Kissiov, D., Stempor, P., Steward, A., Zeiser, E., Sauer, S., Ahringer, J., 2018. Chromatin accessibility dynamics across C. elegans development and ageing. eLife 7. https://doi.org/10.7554/eLife.37344

Jin, Y., Hoskins, R., Horvitz, H.R., 1994. Control of type-D GABAergic neuron differentiation by C. elegans UNC-30 homeodomain protein. Nature 372, 780–783. https://doi.org/10.1038/372780a0

Kuntz, S.G., Schwarz, E.M., DeModena, J.A., De Buysscher, T., Trout, D., Shizuya, H., Sternberg, P.W., Wold, B.J., 2008. Multigenome DNA sequence conservation identifies Hox cis-regulatory elements. Genome Res. 18, 1955–1968. https://doi.org/10.1101/gr.085472.108

Lanjuin, A., VanHoven, M.K., Bargmann, C.I., Thompson, J.K., Sengupta, P., 2003. Otx/otd homeobox genes specify distinct sensory neuron identities in C. elegans. Dev. Cell 5, 621–633. https://doi.org/10.1016/s1534-5807(03)00293-4

Levine, M., 2010. Transcriptional enhancers in animal development and evolution. Curr. Biol. CB 20, R754-763. https://doi.org/10.1016/j.cub.2010.06.070

Ma, X., Zhao, Z., Xiao, L., Xu, W., Wang, Y., Zhang, Y., Wu, G., Du, Z., 2020. Single-Cell Protein Atlas of Transcription Factors Reveals the Combinatorial Code for Spatiotemporal Patterning the C. elegans Embryo. bioRxiv 2020.06.30.178640. https://doi.org/10.1101/2020.06.30.178640

Mace, D.L., Weisdepp, P., Gevirtzman, L., Boyle, T., Waterston, R.H., 2013. A high-fidelity cell lineage tracing method for obtaining systematic spatiotemporal gene expression patterns in Caenorhabditis elegans. G3 Bethesda Md 3, 851–863. https://doi.org/10.1534/g3.113.005918

Maduro, M.F., 2017. Gut development in C. elegans. Semin. Cell Dev. Biol. 66, 3–11. https://doi.org/10.1016/j.semcdb.2017.01.001

Maloof, J.N., Whangbo, J., Harris, J.M., Jongeward, G.D., Kenyon, C., 1999. A Wnt signaling pathway controls hox gene expression and neuroblast migration in C. elegans. Dev. Camb. Engl. 126, 37–49.

Murray, J.I., Bao, Z., Boyle, T.J., Boeck, M.E., Mericle, B.L., Nicholas, T.J., Zhao, Z., Sandel, M.J., Waterston, R.H., 2008. Automated analysis of embryonic gene expression with cellular resolution in C. elegans. Nat. Methods 5, 703–709. https://doi.org/10.1038/nmeth.1228

Murray, J.I., Bao, Z., Boyle, T.J., Waterston, R.H., 2006. The lineaging of fluorescently-labeled Caenorhabditis elegans embryos with StarryNite and AceTree. Nat. Protoc. 1, 1468–1476. https://doi.org/10.1038/nprot.2006.222

Murray, J.I., Boyle, T.J., Preston, E., Vafeados, D., Mericle, B., Weisdepp, P., Zhao, Z., Bao, Z., Boeck, M., Waterston, R.H., 2012. Multidimensional regulation of gene expression in the C. elegans embryo. Genome Res. 22, 1282–1294. https://doi.org/10.1101/gr.131920.111

Narasimhan, K., Lambert, S.A., Yang, A.W.H., Riddell, J., Mnaimneh, S., Zheng, H., Albu, M., Najafabadi, H.S., Reece-Hoyes, J.S., Fuxman Bass, J.I., Walhout, A.J.M., Weirauch, M.T., Hughes, T.R., 2015. Mapping and analysis of Caenorhabditis elegans transcription factor sequence specificities. eLife 4. https://doi.org/10.7554/eLife.06967

Niu, W., Lu, Z.J., Zhong, M., Sarov, M., Murray, J.I., Brdlik, C.M., Janette, J., Chen, C., Alves, P., Preston, E., Slightham, C., Jiang, L., Hyman, A.A., Kim, S.K., Waterston, R.H., Gerstein, M., Snyder, M., Reinke, V., 2011. Diverse transcription factor binding features revealed by genome-wide ChIP-seq in C. elegans. Genome Res. 21, 245–254. https://doi.org/10.1101/gr.114587.110

Noordermeer, D., Duboule, D., 2013. Chromatin architectures and Hox gene collinearity. Curr. Top. Dev. Biol. 104, 113–148. https://doi.org/10.1016/B978-0-12-416027-9.00004-8

Packer, J.S., Zhu, Q., Huynh, C., Sivaramakrishnan, P., Preston, E., Dueck, H., Stefanik, D., Tan, K., Trapnell, C., Kim, J., Waterston, R.H., Murray, J.I., 2019. A lineage-resolved molecular atlas of C. elegans embryogenesis at single-cell resolution. Science 365. https://doi.org/10.1126/science.aax1971

Perry, M.W., Boettiger, A.N., Bothma, J.P., Levine, M., 2010. Shadow enhancers foster robustness of Drosophila gastrulation. Curr. Biol. CB 20, 1562–1567. https://doi.org/10.1016/j.cub.2010.07.043

Reece-Hoyes, J.S., Shingles, J., Dupuy, D., Grove, C.A., Walhout, A.J.M., Vidal, M., Hope, I.A., 2007. Insight into transcription factor gene duplication from Caenorhabditis elegans Promoterome-driven expression patterns. BMC Genomics 8, 27. https://doi.org/10.1186/1471-2164-8-27

Richards, J.L., Zacharias, A.L., Walton, T., Burdick, J.T., Murray, J.I., 2013. A quantitative model of normal Caenorhabditis elegans embryogenesis and its disruption after stress. Dev. Biol. 374, 12–23. https://doi.org/10.1016/j.ydbio.2012.11.034

Santella, A., Du, Z., Bao, Z., 2014. A semi-local neighborhood-based framework for probabilistic cell lineage tracing. BMC Bioinformatics 15, 217. https://doi.org/10.1186/1471-2105-15-217

Sarov, M., Murray, J.I., Schanze, K., Pozniakovski, A., Niu, W., Angermann, K., Hasse, S., Rupprecht, M., Vinis, E., Tinney, M., Preston, E., Zinke, A., Enst, S., Teichgraber, T., Janette, J., Reis, K., Janosch, S., Schloissnig, S., Ejsmont, R.K., Slightam, C., Xu, X., Kim, S.K., Reinke, V., Stewart, A.F., Snyder, M., Waterston, R.H., Hyman, A.A., 2012. A genome-scale resource for in vivo tag-based protein function exploration in C. elegans. Cell 150, 855–866. https://doi.org/10.1016/j.cell.2012.08.001

Shemer, G., Podbilewicz, B., 2002. LIN-39/Hox triggers cell division and represses EFF-1/fusogen-dependent vulval cell fusion. Genes Dev. 16, 3136–3141. https://doi.org/10.1101/gad.251202

Slattery, M., Riley, T., Liu, P., Abe, N., Gomez-Alcala, P., Dror, I., Zhou, T., Rohs, R., Honig, B., Bussemaker, H.J., Mann, R.S., 2011. Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins. Cell 147, 1270–1282. https://doi.org/10.1016/j.cell.2011.10.053

Streit, A., Kohler, R., Marty, T., Belfiore, M., Takacs-Vellai, K., Vigano, M.-A., Schnabel, R., Affolter, M., Müller, F., 2002. Conserved regulation of the caenorhabditis elegans labial/Hox1 gene ceh-13. Dev. Biol. 242, 96–108.

Sulston, J., 1983. The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev. Biol. 100, 64–119. https://doi.org/10.1016/0012-1606(83)90201-4
Takács-Vellai, K., Vellai, T., Chen, E.B., Zhang, Y., Guerry, F., Stern, M.J., Müller, F., 2007. Transcriptional control of Notch signaling by a HOX and a PBX/EXD protein during vulval development in C. elegans. Dev. Biol. 302, 661–669. https://doi.org/10.1016/j.ydbio.2006.09.049
Teng, Y., Girard, L., Ferreira, H.B., Sternberg, P.W., Emmons, S.W., 2004. Dissection of cis-regulatory elements in the C. elegans Hox gene egl-5 promoter. Dev. Biol. 276, 476–492. https://doi.org/10.1016/j.ydbio.2004.09.012
Tihanyi, B., Vellai, T., Regos, A., Ari, E., Müller, F., Takács-Vellai, K., 2010. The C. elegans Hox gene ceh-13 regulates cell migration and fusion in a non-colinear way. Implications for the early evolution of Hox clusters. BMC Dev. Biol. 10, 78. https://doi.org/10.1186/1471-213X-10-78
Van Auken, K., Weaver, D., Robertson, B., Sundaram, M., Saldi, T., Edgar, L., Elling, U., Lee, M., Boese, Q., Wood, W.B., 2002. Roles of the Homothorax/Meis/Prep homolog UNC-62 and the Exd/Pbx homologs CEH-20 and CEH-40 in C. elegans embryogenesis. Dev. Camb. Engl. 129, 5255–5268.
Van Auken, K., Weaver, D.C., Edgar, L.G., Wood, W.B., 2000. Caenorhabditis elegans embryonic axial patterning requires two recently discovered posterior-group Hox genes. Proc. Natl. Acad. Sci. U. S. A. 97, 4499–4503. https://doi.org/10.1073/pnas.97.9.4499
Walton, T., Preston, E., Nair, G., Zacharias, A.L., Raj, A., Murray, J.I., 2015. The Bicoid class homeodomain factors ceh-36/OTX and unc-30/PITX cooperate in C. elegans embryonic progenitor cells to regulate robust development. PLoS Genet. 11, e1005003. https://doi.org/10.1371/journal.pgen.1005003
Wei, X., Potter, C.J., Luo, L., Shen, K., 2012. Controlling gene expression with the Q repressible binary expression system in Caenorhabditis elegans. Nat. Methods 9, 391–395. https://doi.org/10.1038/nmeth.1929
Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., Zheng, H., Goity, A., van Bakel, H., Lozano, J.-C., Galli, M., Lewsey, M.G., Huang, E., Mukherjee, T., Chen, X., Reece-Hoyes, J.S., Govindarajan, S., Shaulsky, G., Walhout, A.J.M., Bouget, F.-Y., Ratsch, G., Larrondo, L.F., Ecker, J.R., Hughes, T.R., 2014. Determination and inference of eukaryotic transcription factor sequence specificity. Cell 158, 1431–1443. https://doi.org/10.1016/j.cell.2014.08.009
Wittmann, C., Bossinger, O., Goldstein, B., Fleischmann, M., Kohler, R., Brunschwig, K., Tobler, H., Müller, F., 1997. The expression of the C. elegans labial-like Hox gene ceh-13 during early embryogenesis relies on cell fate and on anteroposterior cell polarity. Dev. Camb. Engl. 124, 4193–4200.
Yu, H., Seah, A., Sternberg, P.W., 2010. Re-programming of C. elegans male epidermal precursor fates by Wnt, Hox, and LIN-12/Notch activities. Dev. Biol. 345, 1–11. https://doi.org/10.1016/j.ydbio.2010.05.008
Zacharias, A.L., Walton, T., Preston, E., Murray, J.I., 2015. Quantitative Differences in Nuclear β-catenin and TCF Pattern Embryonic Cells in C. elegans. PLoS Genet. 11, e1005585. https://doi.org/10.1371/journal.pgen.1005585
Zhao, Z., Boyle, T.J., Liu, Z., Murray, J.I., Wood, W.B., Waterston, R.H., 2010. A negative regulatory loop between microRNA and Hox gene controls posterior identities in Caenorhabditis elegans. PLoS Genet. 6, e1001089. https://doi.org/10.1371/journal.pgen.1001089
Zheng, C., Jin, F.Q., Chalfie, M., 2015. Hox Proteins Act as Transcriptional Guarantors to Ensure Terminal Differentiation. Cell Rep. 13, 1343–1352. https://doi.org/10.1016/j.celrep.2015.10.044

Figure legends

Figure 1: *ceh-13* and *nob-1/php-3* are expressed broadly in an overlapping pattern in the early *C. elegans* embryo. A) Time lapse images of transgenic *C. elegans* embryos carrying two transgenes, a ubiquitous fluorescent histone to mark all nuclei (shown in green) and a reporter of interest (shown in red), which can be a cis-regulatory element driving a fluorescent histone (transcriptional) or a fluorescently tagged transcription factor protein (translational). Image analysis software identifies nuclei, and quantifies reporter intensity within the nuclei, which can be displayed as a lineage tree colored by expression as in (C). B) 3D projections of nuclei expressing a CEH-13::GFP CRISPR reporter (red), NOB-1::GFP or PHP-3::GFP CRISPR reporter (blue), with overlap shown in magenta, at 120 minutes (50 cell stage), 150 minutes (100 cell stage), and 280 minutes (400 cell stage) post fertilization. NOB-1::GFP and PHP-3::GFP CRISPR reporters were expressed in indistinguishable patterns (Supplemental Figure 1), but PHP-3::GFP was slightly brighter so is shown in the following panels. C) Lineage tree though the 100-cell stage, showing early expression of CEH-13 and PHP-3 CRISPR reporters, colored as in (B). D) Expressing lineages showing CEH-13 and/or PHP-3::GFP CRISPR reporter expression to 350 minutes of development, colors as in (B). Note that CEH-13::GFP precedes PHP-3::GFP and NOB-1::GFP in all lineages except ABp(l/r)papppp (asterisk - expression is consistent in ABprpapppp and variable in ABplpapppp). (E-H) Quantitative detail for highlighted lineages, showing nuclear fluorescence intensity of CEH-13::GFP (fosmid) and NOB-1::GFP transgene reporters across developmental time for the cells leading to ABp(l/r)appaaa (E), ABp(l/r)appppp (F), ABp(l/r)pppppppp (G), and Ep(l/r)p (H). Nuclear fluorescence intensity is in arbitrary units. Grey bars mark cell divisions.

Figure 2: Regions upstream of *nob-1/php-3* can recapitulate its expression pattern. A) Genome browser view of the *nob-1*/*php-3* locus showing the genes (black), candidate enhancers tested (brown), sequence conservation with other nematodes (grey), and NHR-2 ChIP-seq trace (dark blue) from modENCODE (Araya et al., 2014) and ATAC-seq traces (red: ABa lineage, green: ABp lineage) from Charest et al, 2020 (Charest et al., 2020). B) Schematics of the *nob-1* reporter constructs examined, shown in alignment with (A). Enhancers were primarily tested in an orientation 5' to the pes-10 minimal promoter, but a downstream orientation was also used for the -4.3 kb enhancer. C) Lineage trees colored to show expression patterns for the various reporters and tested enhancers with relevant reproducible activity, using a rainbow color scale to increase visible dynamic range. Major expressing lineages are underlined: ABp(l/r)ap: pink, ABp(l/r)ppp: cyan, Ep: black, Cpapp: orange, ectopic: purple. Note the changes in expression driven by the -4.3 kb enhancer depending on its position relative to the *pes-10* minimal promoter. D) Average nuclear fluorescence values for cells in the ABp(l/r)ap (pink) and ABp(l/r)ppp (blue) lineages from at least four embryos, shown on a log scale. E) Ratio of expression in average ABp(l/r)ap cell to the average ABp(l/r)ppp cell at the 350 cell stage for at least 4 embryos. F) Lineage trees showing expression driven by the nob-1 -5.3 kb upstream region reporter with the -3.4 and/or -4.3 enhancers are deleted. Lineages where expression is lost are underlined with colors as in (C).

Figure 3: Expression of cis-regulatory elements upstream of *nob-1* and *php-3* depends on *ceh-13*. A) Fold change in expression level relative to mean wild-type control for *nob-1* promoter and enhancers in *ceh-13* mutant and RNAi conditions. *P* values determined by Wilcoxon Ranked Sum Test.. B) Lineage view of wild-type *nob-1* -5.3kb upstream reporter ("promoter") expression in specified lineages. C) -5.3kb *nob-1* promoter reporter expression in *ceh-13(sw1)* null mutant.

Figure 4: Hox co-factors precede *nob-1* and regulate its expression. A) Trees of NOB-1::GFP expressing lineages, ABp, E, and C, showing the expression of *ceh-13* (fosmid) and *nob-1* (GFP transgene) reporters, and fosmid GFP transgene reporters for the Hox co-factors, *ceh-20*, *ceh-40* and *unc-62*. Color thresholds were adjusted for each gene to show all expressing cells. Highlighted branches are shown in graphs B, C and D. B-D) Average TF reporter nuclear fluorescence intensity across embryos (n >= 2) and for left/right symmetric cells across developmental time for the cells leading to B) ABp(l/r)apapa , C) ABp(l/r)pppaa and D) Cpapp. Fluorescence intensity is in arbitrary units and grey bars mark cell divisions. E, F) Fold change values for -5.3 kb *nob-1* promoter expression in untreated and *unc-62* RNAi treated (E) or *ceh-40(gk159)* mutant embryos treated with *ges-1* (control) or *ceh-20* RNAi (F) in specified lineages. Significant changes marked by asterisk; *P* values determined by Wilcoxon Ranked Sum Test.

Figure 5: *elt-1* regulates *nob-1* expression in ABp(l/r)ap lineage. A) Partial lineage trees showing ELT-1::GFP (Green), NOB-1::GFP (Blue) or overlapping (Cyan) expression in the ABp and C lineages. Arrowheads indicate branches shown in subsequent panels. B-D) Average ELT-1::GFP and NOB-1::GFP nuclear fluorescence intensity (arbitrary units) in cells leading to B) ABp(l/r)apppap (neural fate) C), ABp(l/r)appppa (hypodermal fate), and D) Cpappd (hypodermal fate). Grey bars mark cell divisions. E) Fold change values for the *nob-1* promoter and -3.4 kb and -4.3 kb wild-type or mutated enhancers in the ABp(l/r)ap lineage in the control and *elt-1(ok1002)* mutant conditions. F-G) Fold-change of *nob-1* reporter intensity in *elt-1(ok1002)* relative to mean wild-type control level in the E) ABp(l/r)pp, F) ABp(l/r)ppp, and G) Cpap lineages. *P* values determined by Wilcoxon Ranked Sum Test.

Figure 6: Cell defects are observed in the posterior of both *ceh-13* and *nob-1* mutant embryos. A) Lineage showing *ceh-13* expression with plots of frequency of cell cycle defects and cell position defects for each cell plotted below. Cells expressing *ceh-13* during their development are highlighted in pink (ABalap and ABalpp, which express very low levels of CEH-13::GFP, are denoted with lighter shading). Cell cycle defects are defined as missed, ectopic or significant change in division timing of at least 5 minutes and cell position defects are defined as significant deviation of at least 5 μm from the positions expected from wild type (see methods). Defects are listed cumulatively for terminal cells, so defects in both a parent and daughter cell would be scored as two defects. B) Lineage showing *nob-1* and *php-3* expression overlap with plots of frequency of cell cycle defects and cell position defects for each cell plotted below. Cells expressing both *nob-1* and *php-3* are highlighted in cyan on the plots. C, D) Plot of mean deviation in microns vs. mean neighbor score for each cell in the embryo, colored by level of NOB-1::GFP expression, for *nob-1(ct223)* mutant which disrupts both *nob-1* and *php-3* (C) and wild-type (D) embryos. Neighbor score is a measure of whether a cell is close to its normal neighbors; values above 0.8 indicate aberrant neighbor relationships. E, F) Plot of mean deviation in microns vs. mean neighbor score for each cell in the embryo, colored by level of CEH-13::GFP expression, for ceh-13(sw1) mutant (E) and wild-type (F) embryos. (G-I) Three-dimensional plot of cell position deviations - arrows point from average wild-type location to average mutant location for each cell, colors indicate the level of NOB-1::GFP (G) or CEH-13::GFP(H-I) expression: G) *nob-1(ct223)* at 335 minutes of development. H) *ceh-13(sw1)* at 230 minutes of development) *ceh-13(sw1)* at 322 minutes of development. J) wild-type and *nob-1(ct223)* mutant lineages for the ABp(l/r)ap and ABp(l/r)pp lineages, showing examples of cell division defects. Green stars indicate the normal divisions of neuroblasts, orange Xs indicate the failed divisions of these cells with the frequency observed noted below. Blue underline indicates the expression of NOB-1::GFP. K) Wild-type and *ceh-13(sw1)* mutant lineages for the ABarp lineage. Green stars indicate the normal development of ADEshL and H2L cells, orange Xs

indicate the ectopic division of these cells with the frequency observed noted below. Red underline indicates the expression of CEH-13::GFP.

Supplemental Figure 1: Hox gene are expressed in overlapping patterns. A) Overlap in the expression of NOB-1::GFP and PHP-3::GFP CRISPR knock-in alleles. Expression is nearly identical except *php-3* is expressed more consistently in ABplpapppp. B, C) Early (B) and late (C) overlap of a CEH-13::GFP fosmid translational reporter transgene with a -9kb NOB-1::GFP rescuing translational reporter transgene. The expression patterns are nearly identical to the CRISPR knock-in alleles except the 9kb NOB-1::GFP reporter has reduced expression in the C lineage. D) Expression pattern of a CEH-13::GFP transgene that contains 8.2 kb of upstream sequence plus the first intron. It lacks expression in the MS lineage compared to the other CEH-13 reporters. E) Correlation between average PHP-3::GFP CRISPR and NOB-1::GFP CRISPR reporters for each cell during embryonic development. F) Correlation between average CEH-13::GFP CRISPR and NOB-1::GFP CRISPR reporters for each cell during embryonic development. Note that no cells express high levels of both proteins. G) Correlation between average CEH-13::GFP and PHP-3::GFP intensity levels for each cell at each (1.5 minute) time point. Note that no cells express high levels of both proteins.

Supplemental Figure 2: Some *nob-1* enhancer constructs drive variable ectopic expression. A) Representative examples showing the expression variability for each of the enhancers tested. Colored lines indicate the lineages of interest. Some lineages with no expression are shown as partial trees. B) Brightfield image of a hatched L1 larva carrying the -4.3kb enhancer reporter and showing the “no backend” phenotype characteristic of *nob-1/php-3* mutants. Nearby bead is 20µm in diameter.

Supplemental Figure 3: *ceh-13* regulates enhancers of *nob-1*. A) Graphs showing the effect of *ceh-13* RNAi on NOB-1::GFP rescuing transgene as well as the control values for the *nob-1* enhancer reporters (*sw1* mutant values same as reported in Fig. 3). Red * indicates $p<0.05$ in Wilcoxon Ranked sum test. B) Lineage trees showing expression changes relative to wild-type embryos for the -3.4kb enhancer with mutated HOX sites.

Supplemental Figure 4: Hox co-factors affect *nob-1* expression in specific lineages. A) Tree showing effects of *unc-62i* on *nob-1* promoter transcriptional reporter expression in the ABpl lineage. Lineages with loss of expression are underlined. The ABp(l/r)apap lineage was most affected in 13/14 lineages tested. B) Trees showing the effect of *unc-62i* on NOB-1::GFP translational reporter expression, with two examples shown. Lineages affected are underlined. C) Graphs for all embryos tested (eight) for the same lineages shown in (B). Red * indicates $p<0.05$ in Wilcoxon Ranked sum test. D) Effect of *unc-62* RNAi on CEH-13::GFP fosmid reporter in lineages where the two genes are both expressed. No significant changes were detected.

Supplemental Figure 5: Full Lineages for *nob-1* regulators. Full lineages for CEH-20::GFP, CEH-40::GFP, UNC-62::GFP, and ELT-1::GFP, all fosmid translational reporters. All lineages are shown to at least the 350 cell stage - selected lineages are shown later to identify additional expression or dynamics.

Supplemental Figure 6: *nob-1* enhancer expression requires *elt-1*. A) Fold change values of *nob-1* enhancers in control and *elt-1* mutant conditions in the ABp(l/r)ap and Cpap lineages (if expressed in control). Mutant values are the same as reported in Fig. 5. Red * indicates $p<0.05$ in Wilcoxon Ranked sum test. B) Fold change values of *nob-1* enhancers in control and *elt-1*

mutant conditions in the ABp(l/r)ppp and Ep lineages (if expressed in control) where ELT-1::GFP is not expressed. Red * indicates p<0.05 in Wilcoxon Ranked sum test. C) Expression of the -3.4 kb enhancer and a version from which all GATA sites have been mutagenized in the ABp, E and C lineages. Note: the mutagenesis also disrupted one *ceh-20/40* predicted site, two *nob-1* predicted sites, and one *pop-1* predicted site, as these were fully overlapping with GATA sites.

Supplemental Figure 7: *ceh-36* and *unc-30* contribute to *nob-1* expression. A) wild-type and *ceh-36(ok795);unc-30(ok613)* mutant embryos expressing the NOB-1::GFP reporter transgene. Expression is shown in the ABp lineage at the ~200 cell stage. B) Boxplot showing the ratio of expression in the ABp(l/r)ap lineage to the ABp(l/r)ppp lineage for wild-type and *ceh-36(ok795);unc-30(ok613)* mutant at the 350 cell stage for at least 4 embryos. Red * indicates p<0.05 in Wilcoxon Ranked sum test.

Supplemental Figure 8: Defective positions of specific cells in *nob-1* mutants. A) The position of cells that normally express *nob-1* are highlighted in the context of the whole embryo at ~320 minutes (~570 cells, early morphogenesis). As compared to the wild-type average (left), the *nob-1(ct223)* mutant (right) cells are disorganized and displaced anteriorly, particularly ABp(l/r)ppp and some ABp(l/r)app cells (ABp(l/r)appp). B) The positions of the SAB neuron and neuroblasts (red) and DA motorneurons (blue) are shown relative to the intestine (green) in wild-type (left) and *ceh-13(sw1)* mutant (right), showing the anterior displacement of these cells. C) The positions of the neuroblasts that fail to divide the *nob-1(ct223)* mutant as compared to control (left), showing that these are dramatically mispositioned.

Supplemental Figure 9: Regulation of *nob-1/php-3* at the lineage and enhancer levels. A) Diagram summarizing the transcriptional regulation of *nob-1/php-3* expression in the specific lineages noted. Asterisk indicates indirect regulation. Note: although *nob-1/php-3* are expressed in other lineages in C, only Cpap was possible to analyze with the transgenic reporters examined. Regulation by *pop-1/sys-1* from Zacharias et al., 2015 (Zacharias et al., 2015) B) Diagram summarizing enhancer-level transcriptional regulation of *nob-1/php-3* by the indicated factors. Specific enhancers regulated by *ceh-20/ceh-40*, *ceh-36/unc-30*, *pop-1/sys-1* and *unc-62* were not defined, but their activity can be localized to the regions marked by brackets. Based on our results, additional *cis*-regulatory elements likely exist within the blue, green and yellow regions (yellow encompasses the rest of the genome).

Supplemental Tables found in SupplementalTables1-9.xlsx
Supplemental Table 1: Strain list
Supplemental Table 2: Tested enhancer sequences
Supplemental Tables 3-5: Motifs found in nob-1 enhancers -3.4kb, -4.3kb, -8.3kb
Supplemental Table 6: Cell averaged reporter values for all embryos analyzed
Supplemental Tables 7-9: Cell division times for wild-type, *ceh-13(sw1),* and *nob-1(ct223)* embryos

Supplemental Tables found in SupplementalTables10-12.xlsx
Can be downloaded from https://www.dropbox.com/s/lvxh80n8s0ntp2l/SupplementalTables10-12.xlsx?dl=0
Supplemental Table 10-12: Cell positions for wild-type, *ceh-13(sw1),* and *nob-1(ct223)* embryos

Figure 1

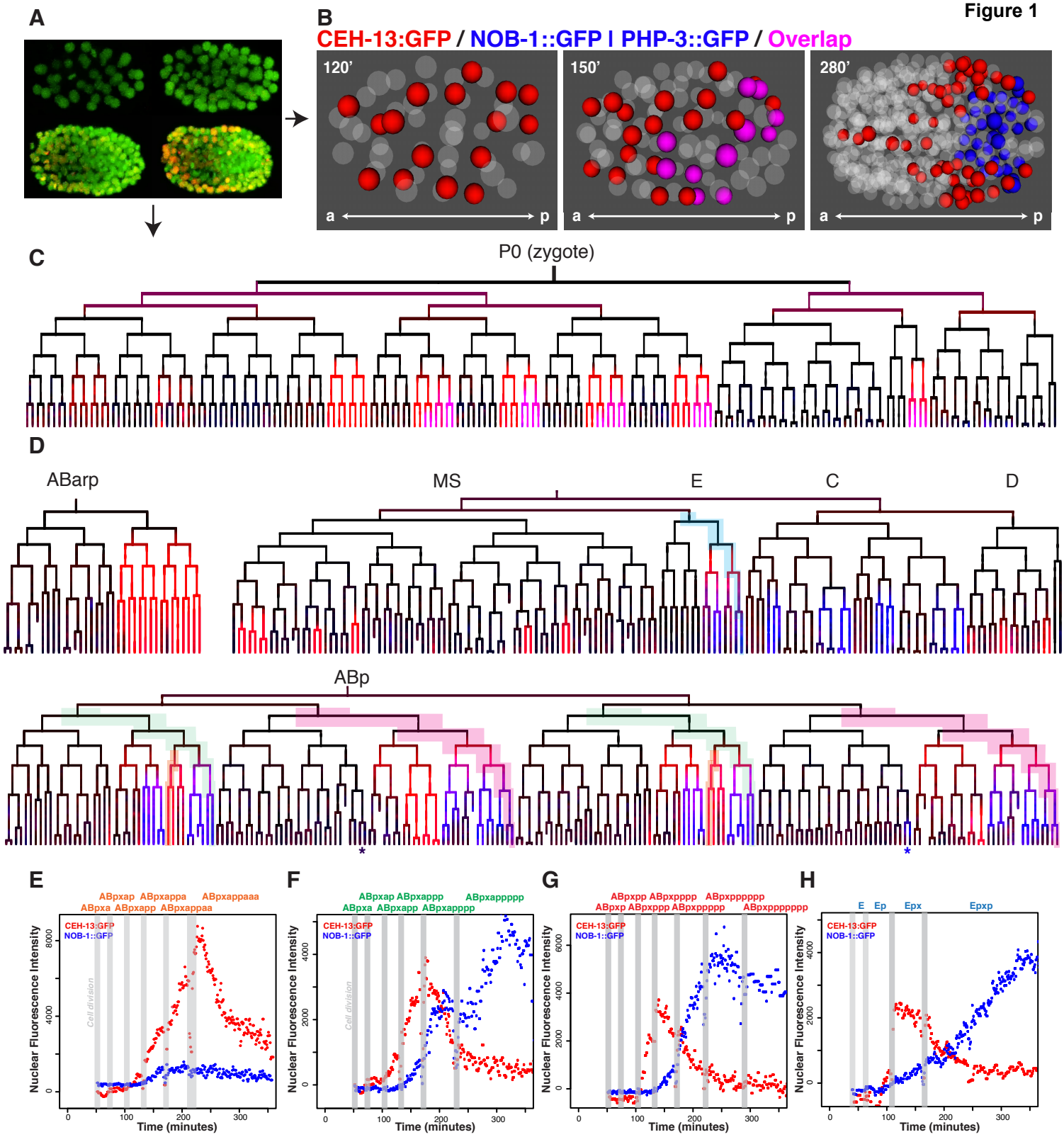

A
B
CEH-13:GFP / NOB-1::GFP I PHP-3::GFP / Overlap
120'
150'
280'
a
p
C
P0 (zygote)
D
ABarp
MS
E
C
D
ABp
E
F
G
H
ABpxap ABpxappa ABpxappaaa
ABpxa ABpxapp ABpxappaa
ABpxap ABpxappp ABpxappppp
ABpxa ABpxapp ABpxappppp
ABpxpp ABpxpppp ABpxppppp
ABpxp ABpxppp ABpxpppp ABpxpppppp
E Ep Epx Epxp
CEH-13:GFP
NOB-1::GFP
Cell division
Nuclear Fluorescence Intensity
Time (minutes)

Figure 2

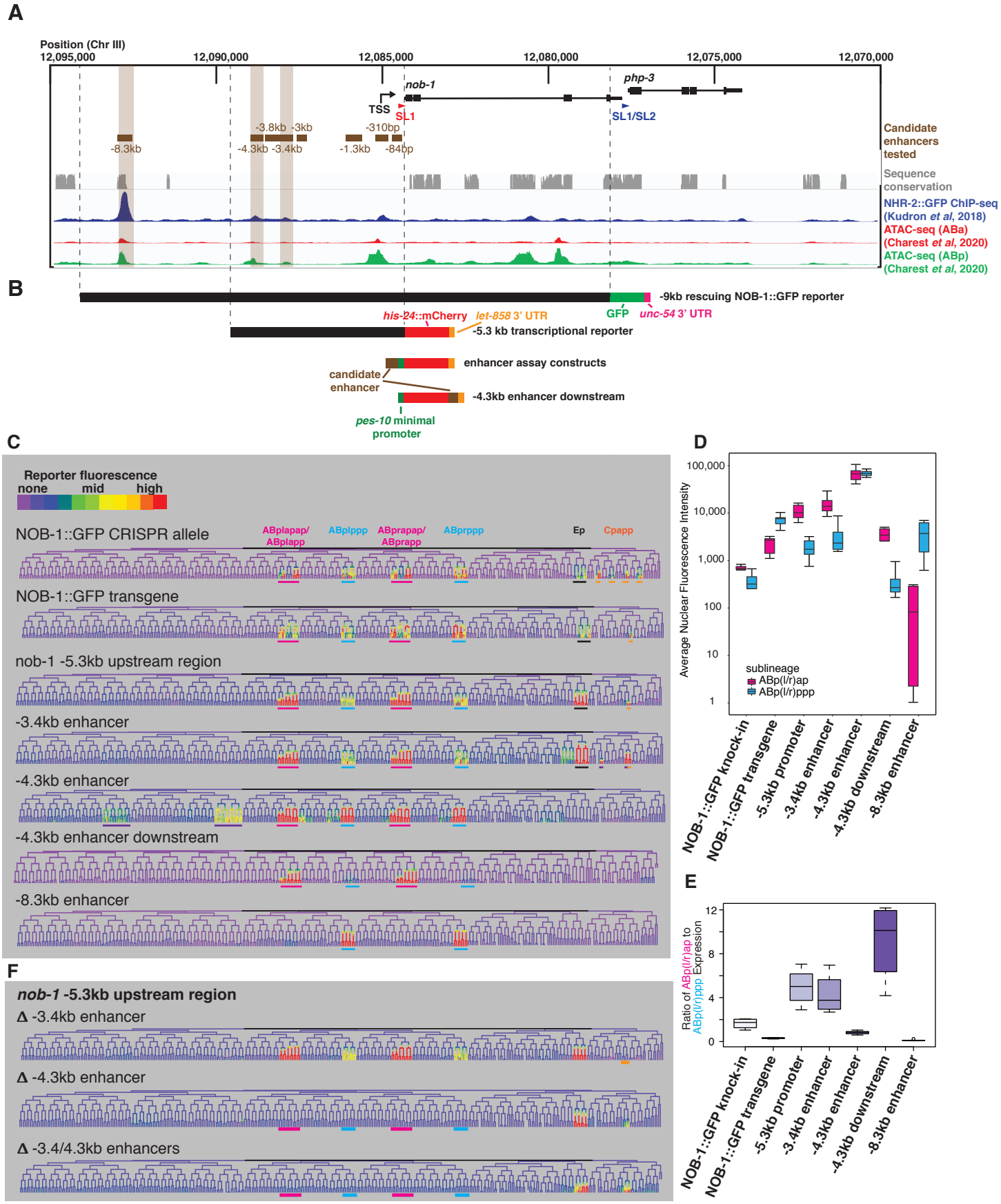

A
Position (Chr III)
12,095,000
12,090,000
12,085,000
12,080,000
12,075,000
12,070,000
nob-1
php-3
TSS
SL1
SL1/SL2
-8.3kb
-3.8kb
-3kb
-4.3kb
-3.4kb
-310bp
-1.3kb
-84bp
Candidate enhancers tested
Sequence conservation
NHR-2::GFP ChIP-seq (Kudron et al, 2018)
ATAC-seq (ABa) (Charest et al, 2020)
ATAC-seq (ABp) (Charest et al, 2020)
B
-9kb rescuing NOB-1::GFP reporter
his-24::mCherry
let-858 3' UTR
GFP
unc-54 3' UTR
-5.3 kb transcriptional reporter
enhancer assay constructs
candidate enhancer
-4.3kb enhancer downstream
pes-10 minimal promoter
C
Reporter fluorescence
none
mid
high
NOB-1::GFP CRISPR allele
ABplapap/ ABplapp
ABplppp
ABprapap/ ABprapp
ABprppp
Ep
Cpapp
NOB-1::GFP transgene
nob-1 -5.3kb upstream region
-3.4kb enhancer
-4.3kb enhancer
-4.3kb enhancer downstream
-8.3kb enhancer
D
Average Nuclear Fluorescence Intensity
100,000
10,000
1,000
100
10
1
sublineage
ABp(l/r)ap
ABp(l/r)ppp
NOB-1::GFP knock-in
NOB-1::GFP transgene
-5.3kb promoter
-3.4kb enhancer
-4.3kb enhancer
-4.3kb downstream
-8.3kb enhancer
E
Ratio of ABp(l/r)ap to ABp(l/r)ppp Expression
12
8
4
0
NOB-1::GFP knock-in
NOB-1::GFP transgene
-5.3kb promoter
-3.4kb enhancer
-4.3kb enhancer
-4.3kb downstream
-8.3kb enhancer
F
nob-1 -5.3kb upstream region
Δ -3.4kb enhancer
Δ -4.3kb enhancer
Δ -3.4/4.3kb enhancers

Figure 3

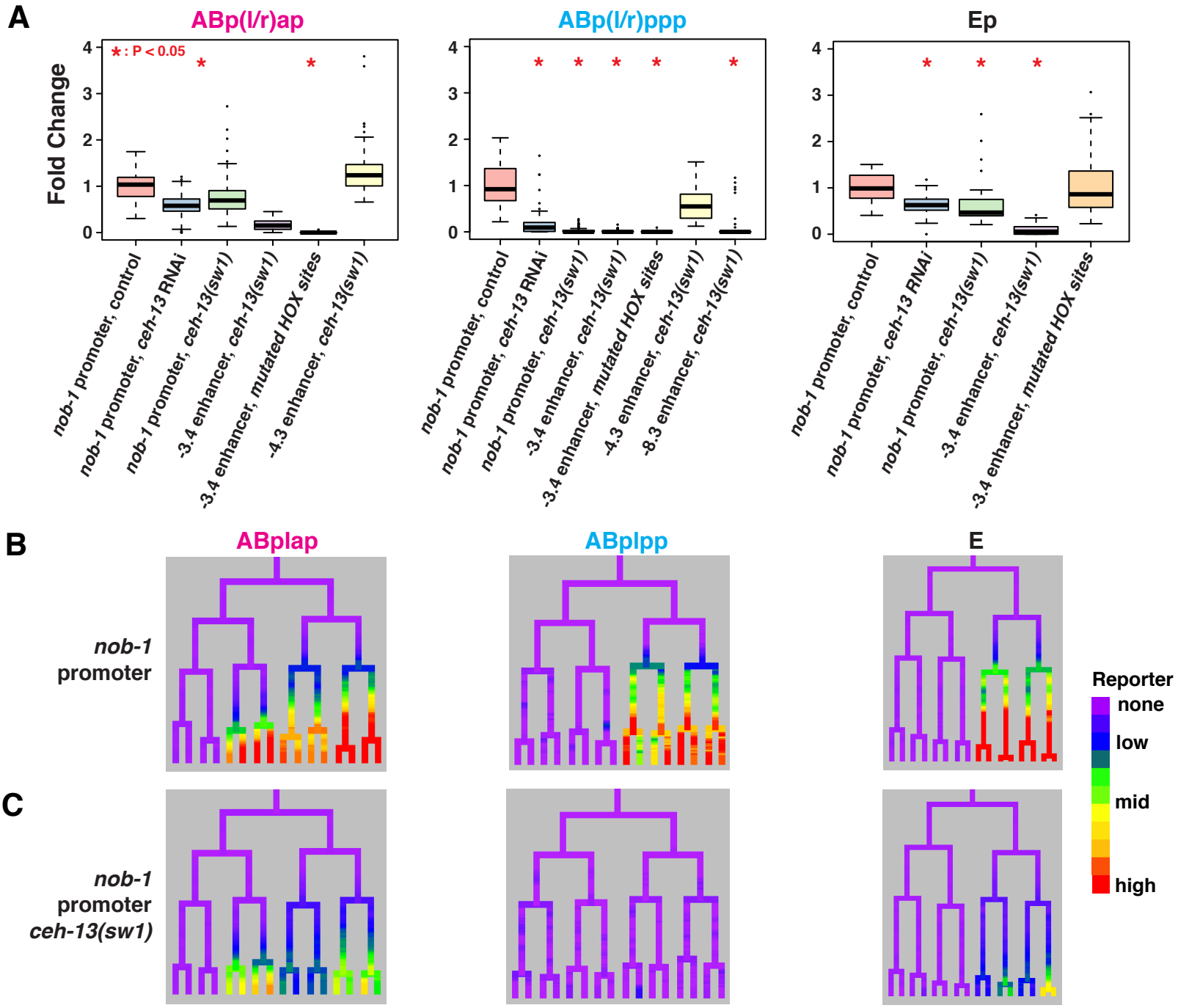

A
ABp(l/r)ap
ABp(l/r)ppp
Ep
Fold Change
*: P < 0.05
nob-1 promoter, control
nob-1 promoter, ceh-13 RNAi
nob-1 promoter, ceh-13(sw1)
-3.4 enhancer, ceh-13(sw1)
-3.4 enhancer, mutated HOX sites
-4.3 enhancer, ceh-13(sw1)
-8.3 enhancer, ceh-13(sw1)
B
ABplap
ABplpp
E
nob-1 promoter
C
nob-1 promoter ceh-13(sw1)
Reporter
none
low
mid
high

Figure 4

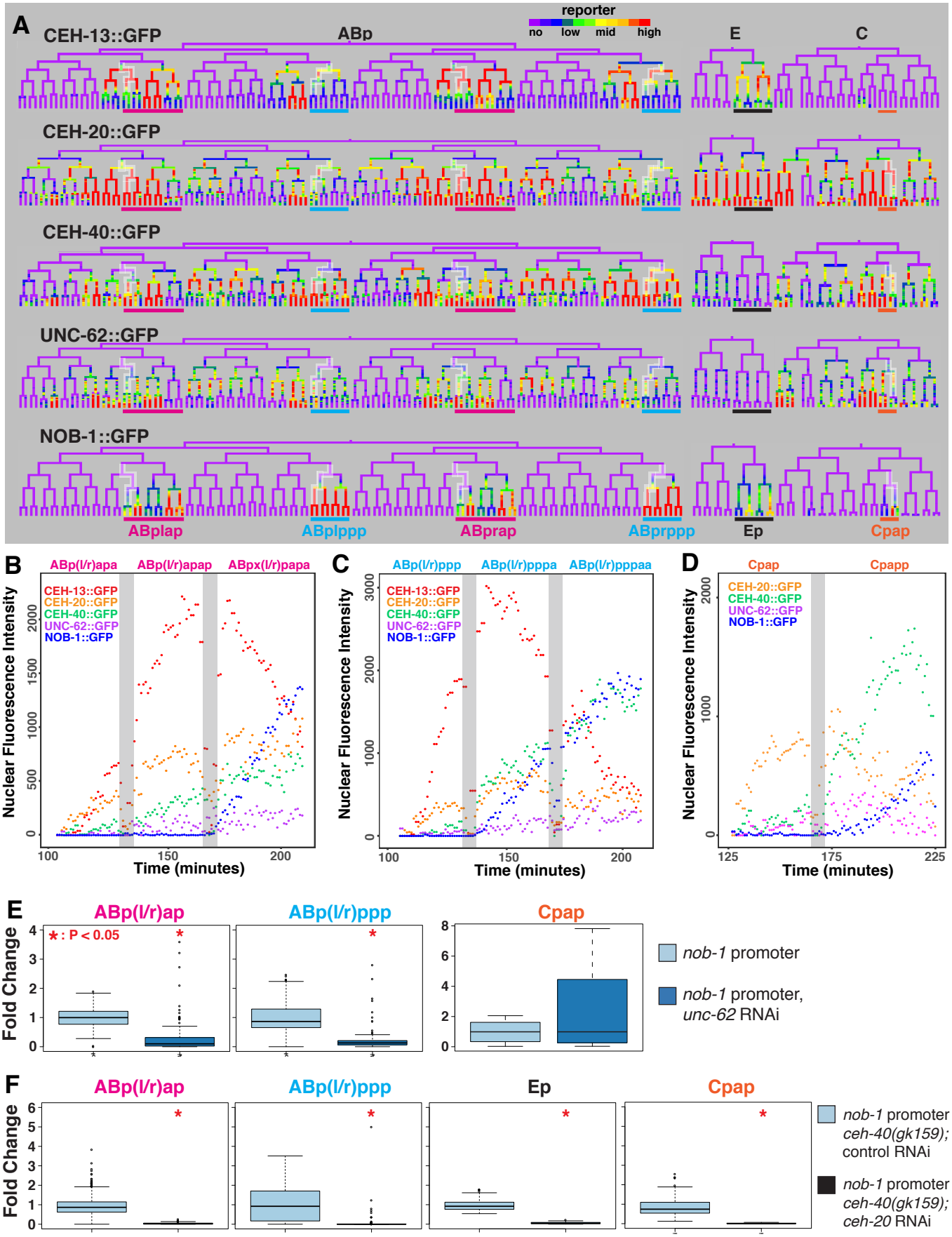

A
CEH-13::GFP
ABp
reporter
no
low
mid
high
E
C
CEH-20::GFP
CEH-40::GFP
UNC-62::GFP
NOB-1::GFP
ABplap
ABplppp
ABprap
ABprppp
Ep
Cpap
B
ABp(l/r)apa
ABp(l/r)apap
ABpx(l/r)papa
CEH-13::GFP
CEH-20::GFP
CEH-40::GFP
UNC-62::GFP
NOB-1::GFP
Nuclear Fluorescence Intensity
Time (minutes)
C
ABp(l/r)ppp
ABp(l/r)pppa
ABp(l/r)pppaa
D
Cpap
Cpapp
E
ABp(l/r)ap
ABp(l/r)ppp
Cpap
*: P < 0.05
Fold Change
nob-1 promoter
nob-1 promoter, unc-62 RNAi
F
ABp(l/r)ap
ABp(l/r)ppp
Ep
Cpap
nob-1 promoter ceh-40(gk159); control RNAi
nob-1 promoter ceh-40(gk159); ceh-20 RNAi

Figure 5

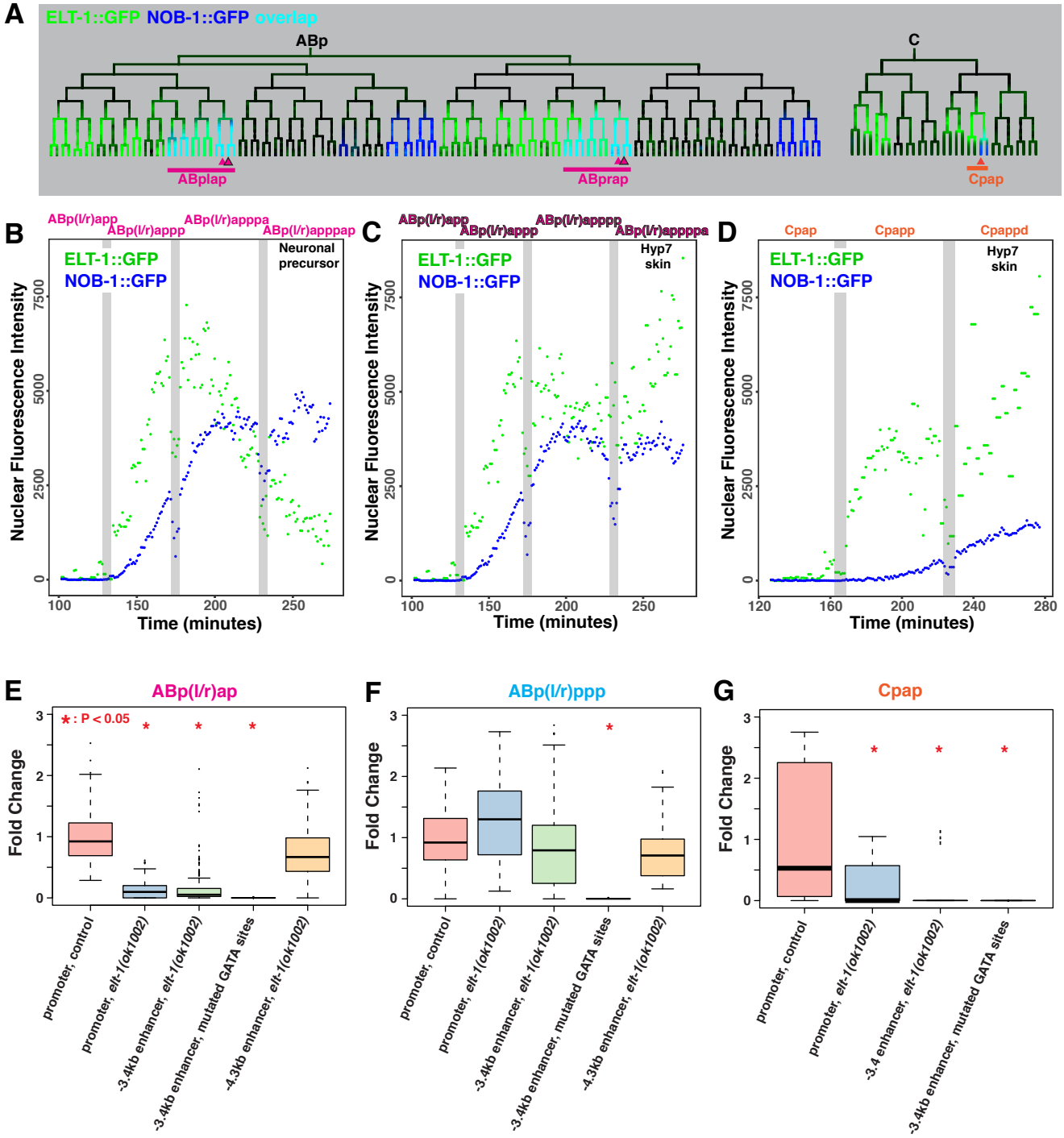
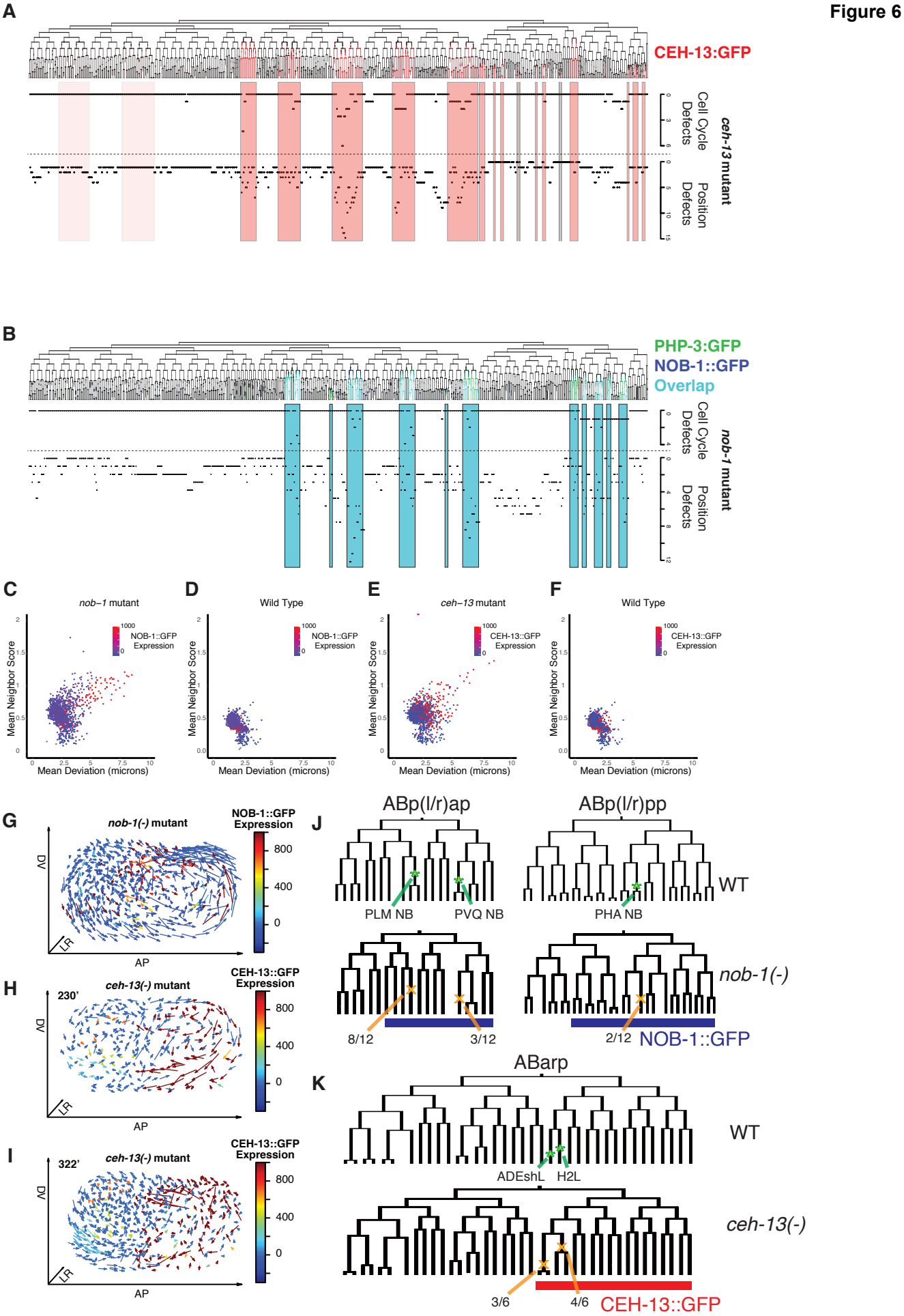
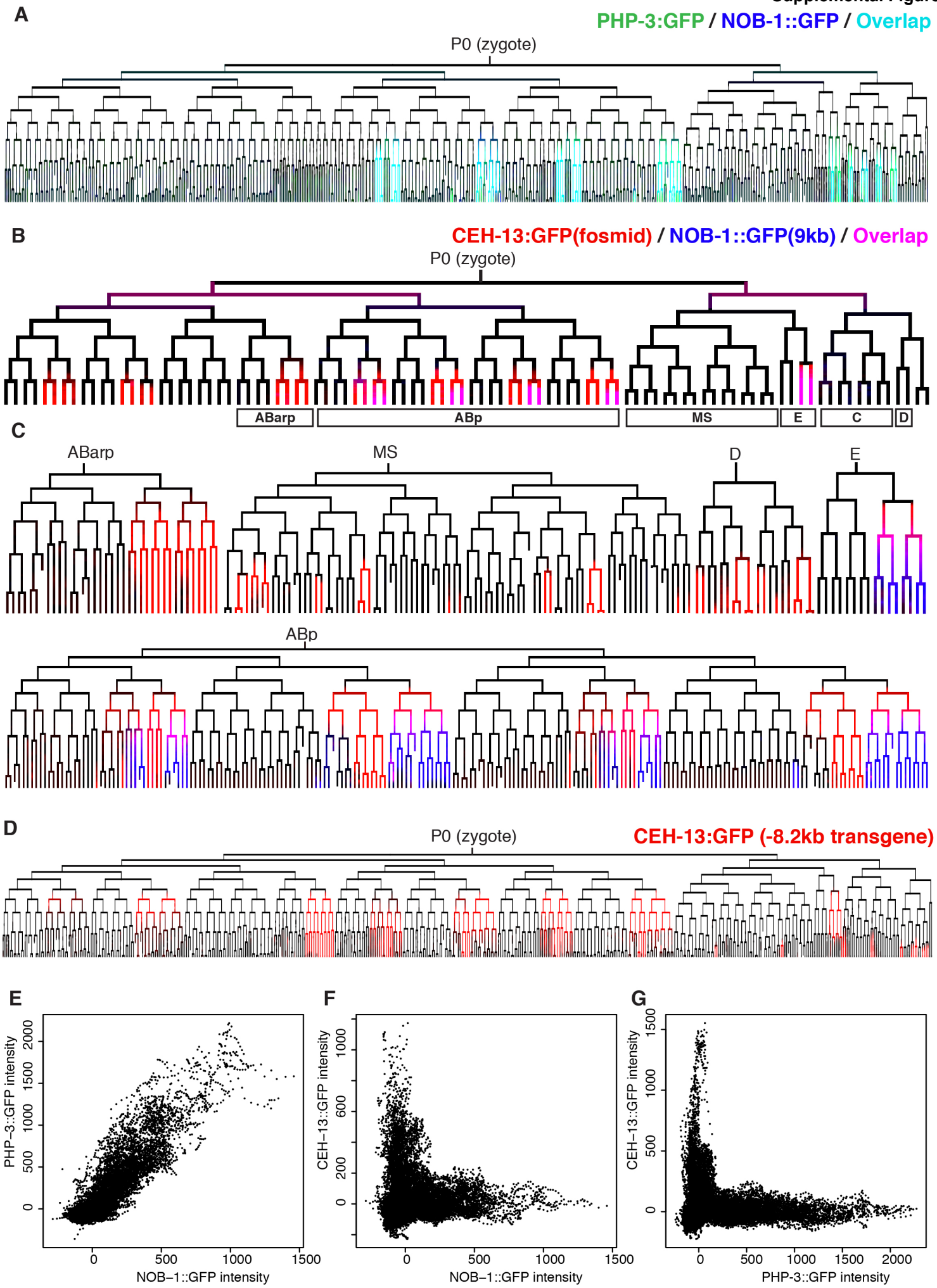
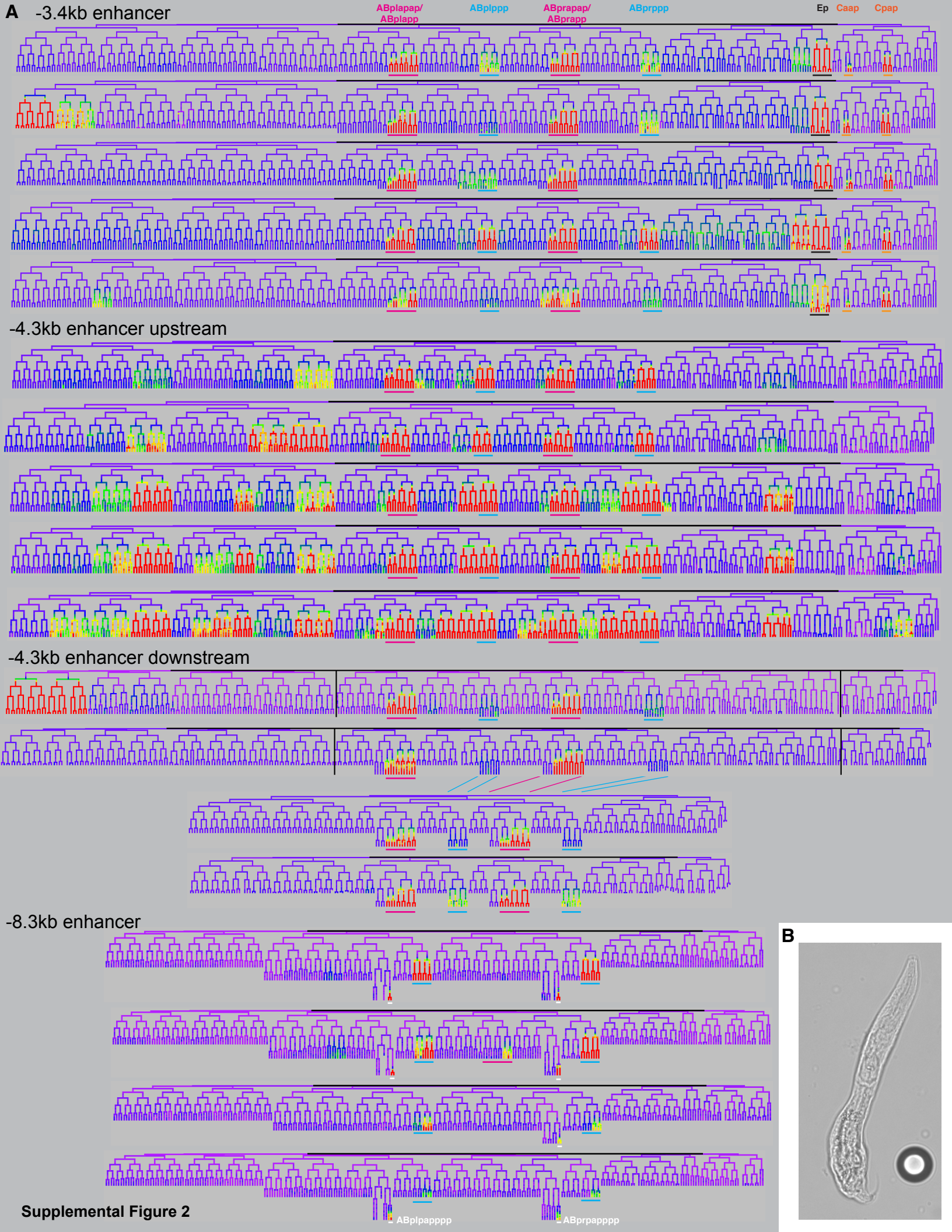

A
ELT-1::GFP NOB-1::GFP overlap
ABp
C
ABplap
ABprap
Cpap
B
ABp(l/r)app
ABp(l/r)appp
ABp(l/r)apppa
ABp(l/r)apppap
ELT-1::GFP
NOB-1::GFP
Neuronal precursor
Nuclear Fluorescence Intensity
Time (minutes)
C
ABp(l/r)app
ABp(l/r)appp
ABp(l/r)apppp
ABp(l/r)appppa
Hyp7 skin
D
Cpap
Cpapp
Cpappd
Hyp7 skin
E
ABp(l/r)ap
* : P < 0.05
Fold Change
promoter, control
promoter, elt-1(ok1002)
-3.4kb enhancer, elt-1(ok1002)
-3.4kb enhancer, mutated GATA sites
-4.3kb enhancer, elt-1(ok1002)
F
ABp(l/r)ppp
G
Cpap
-3.4 enhancer, elt-1(ok1002)

Figure 6


A
CEH-13:GFP
Cell Cycle Defects
Position Defects
ceh-13 mutant
B
PHP-3:GFP
NOB-1::GFP
Overlap
Cell Cycle Defects
Position Defects
nob-1 mutant
C
nob−1 mutant
NOB-1::GFP Expression
Mean Neighbor Score
Mean Deviation (microns)
D
Wild Type
NOB-1::GFP Expression
E
ceh−13 mutant
CEH-13::GFP Expression
F
Wild Type
CEH-13::GFP Expression
G
nob-1(-) mutant
NOB-1::GFP Expression
DV
LR
AP
H
ceh-13(-) mutant
230'
CEH-13::GFP Expression
I
ceh-13(-) mutant
322'
CEH-13::GFP Expression
J
ABp(l/r)ap
ABp(l/r)pp
PLM NB
PVQ NB
PHA NB
WT
nob-1(-)
8/12
3/12
2/12
NOB-1::GFP
K
ABarp
ADEshL
H2L
WT
ceh-13(-)
3/6
4/6
CEH-13::GFP

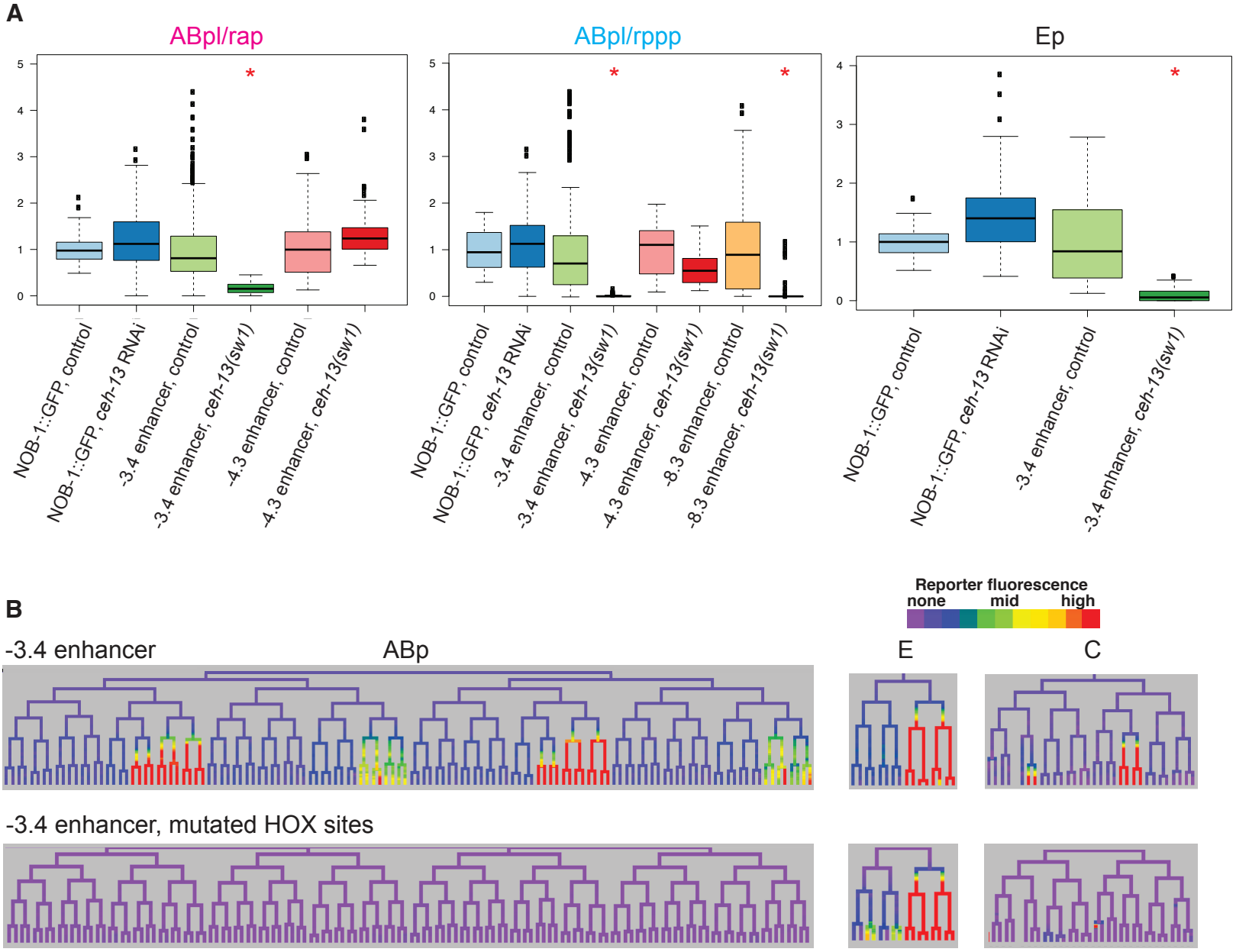
Supplemental Figure 1





Supplemental Figure 2

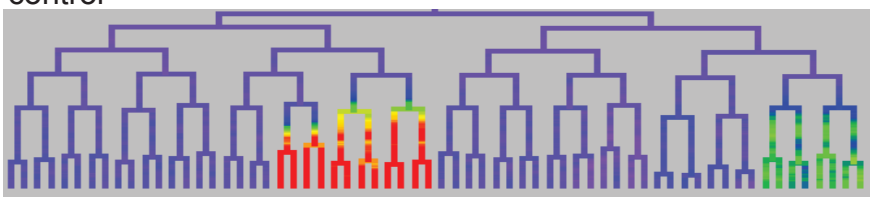
Supplemental Figure 3



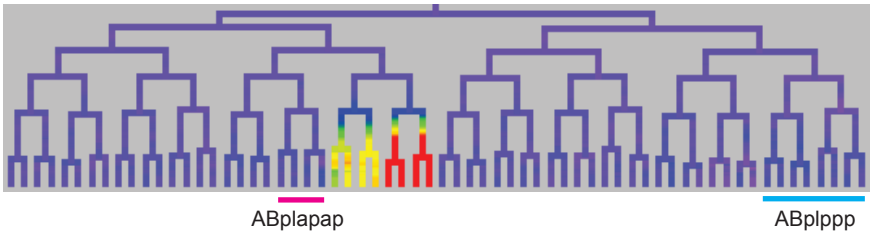
Supplemental Figure 4

A *nob-1* promoter, ABpl sublineage

control

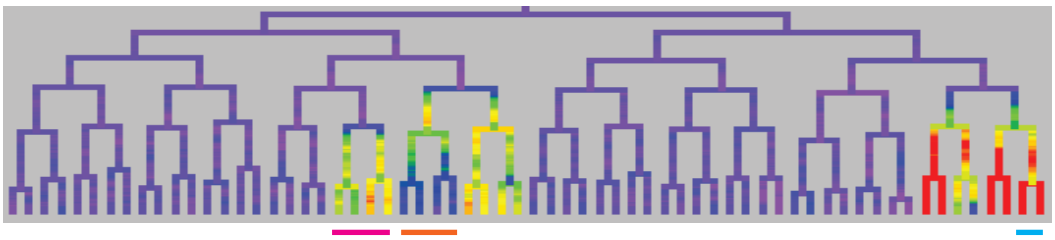


unc-62 RNAi

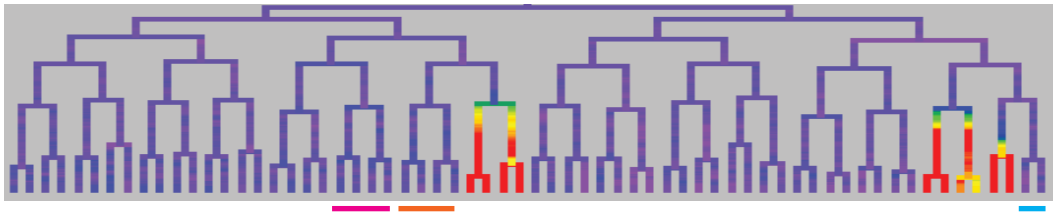


B NOB-1::GFP, ABpr sublineage

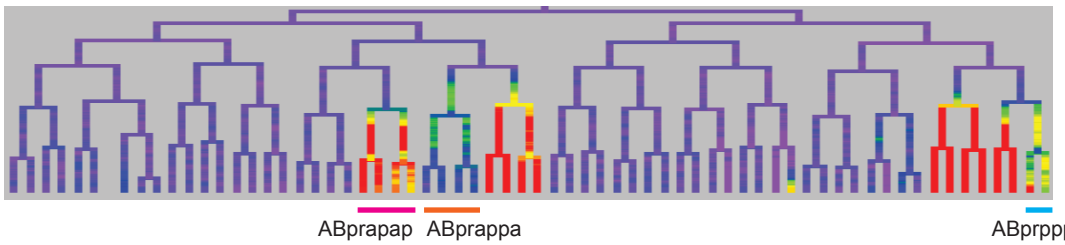
control



unc-62 RNAi

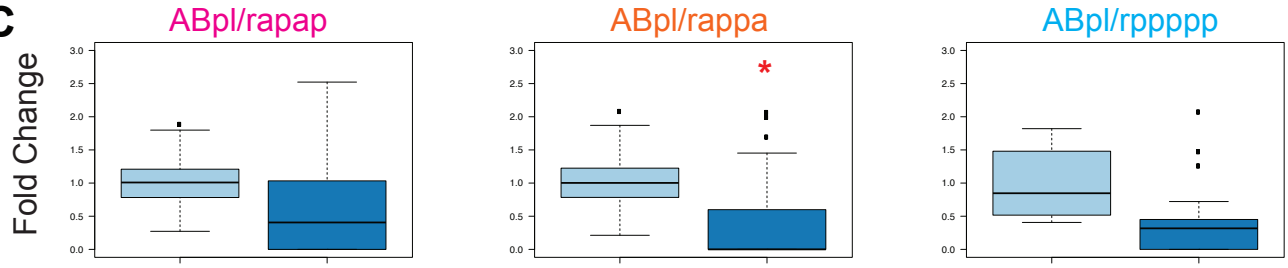


example #1

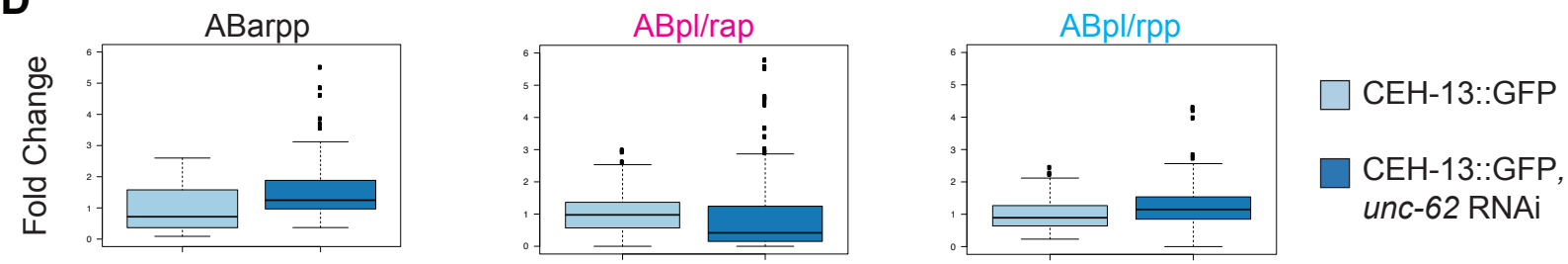


example #2

C

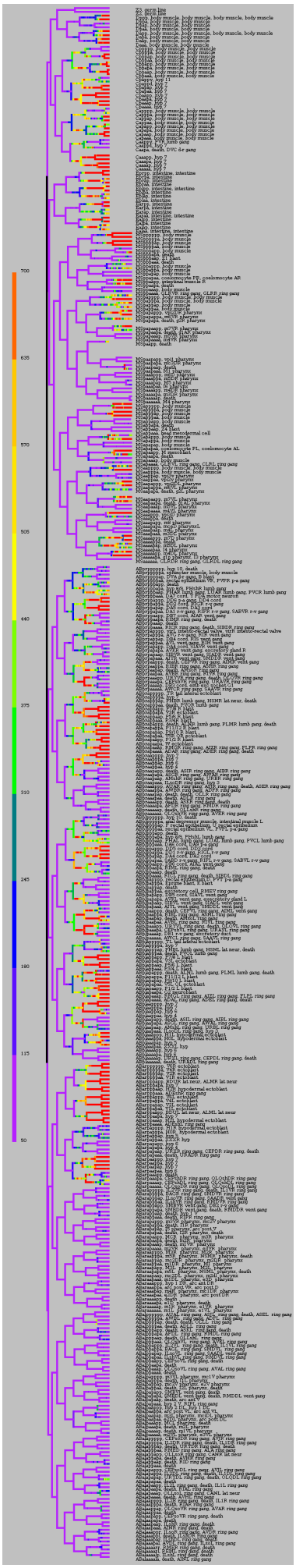


D

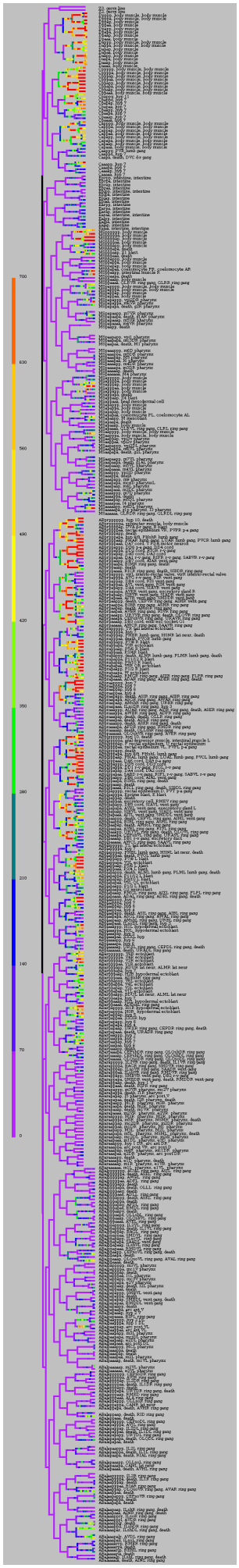


Supplemental Figure 5

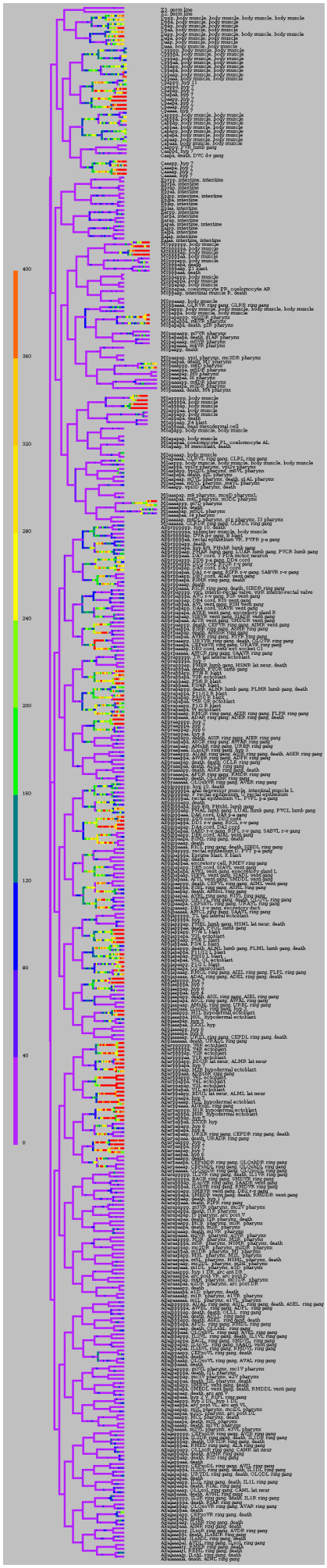
CEH-20::GFP



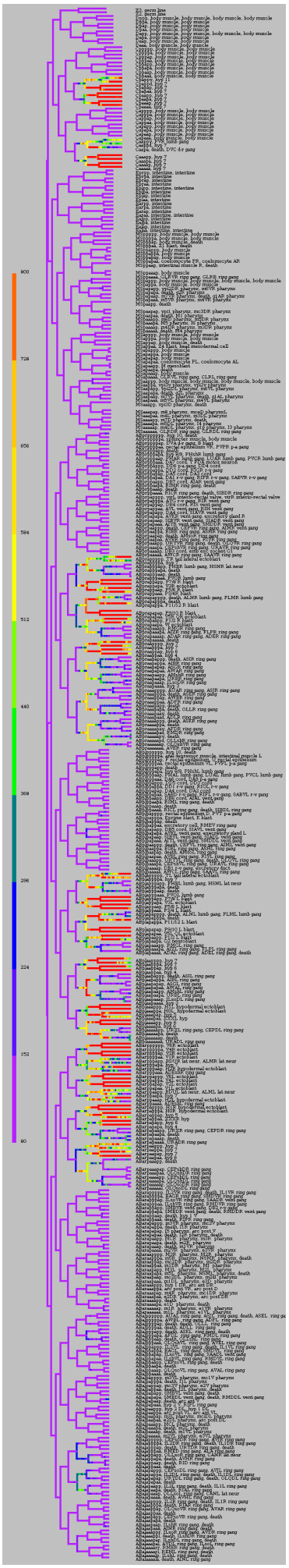
CEH-40::GFP



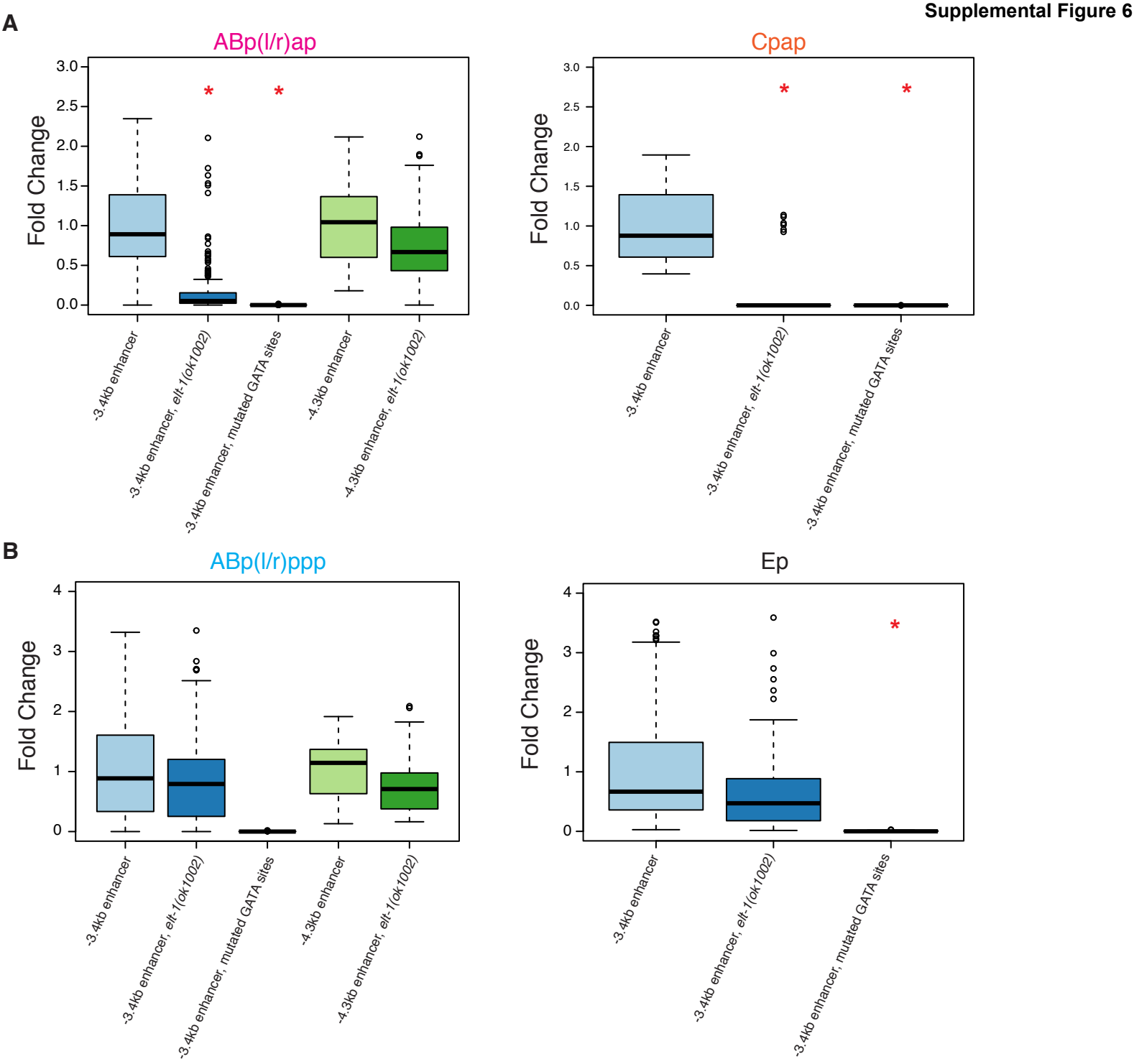
UNC-62::GFP

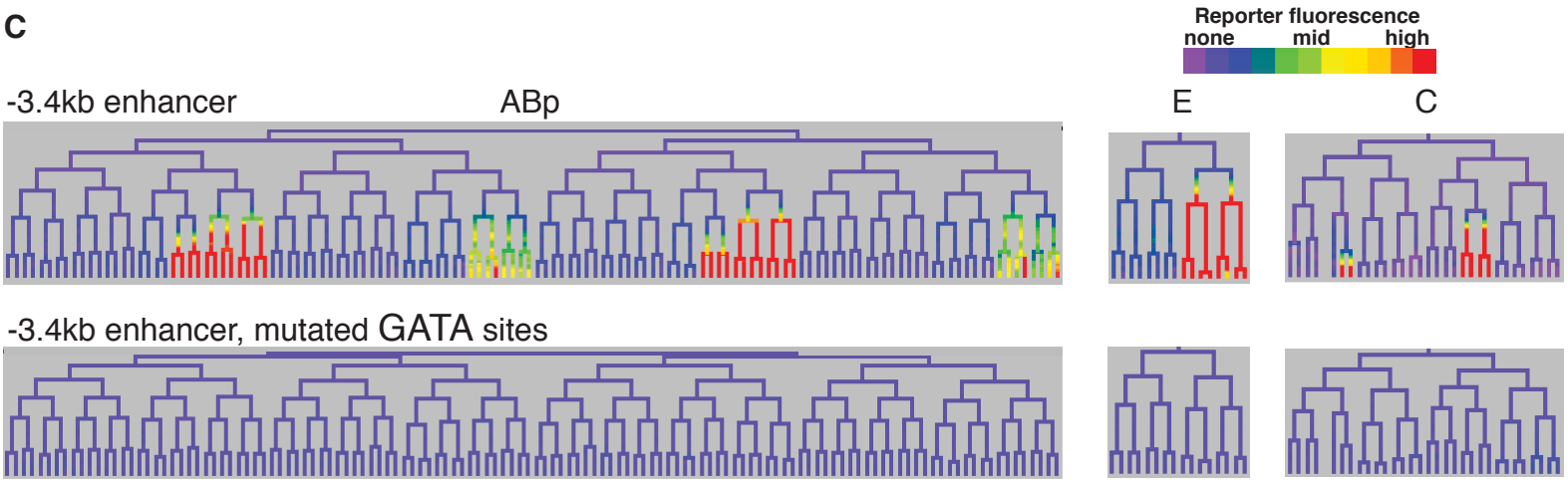


ELT-1::GFP



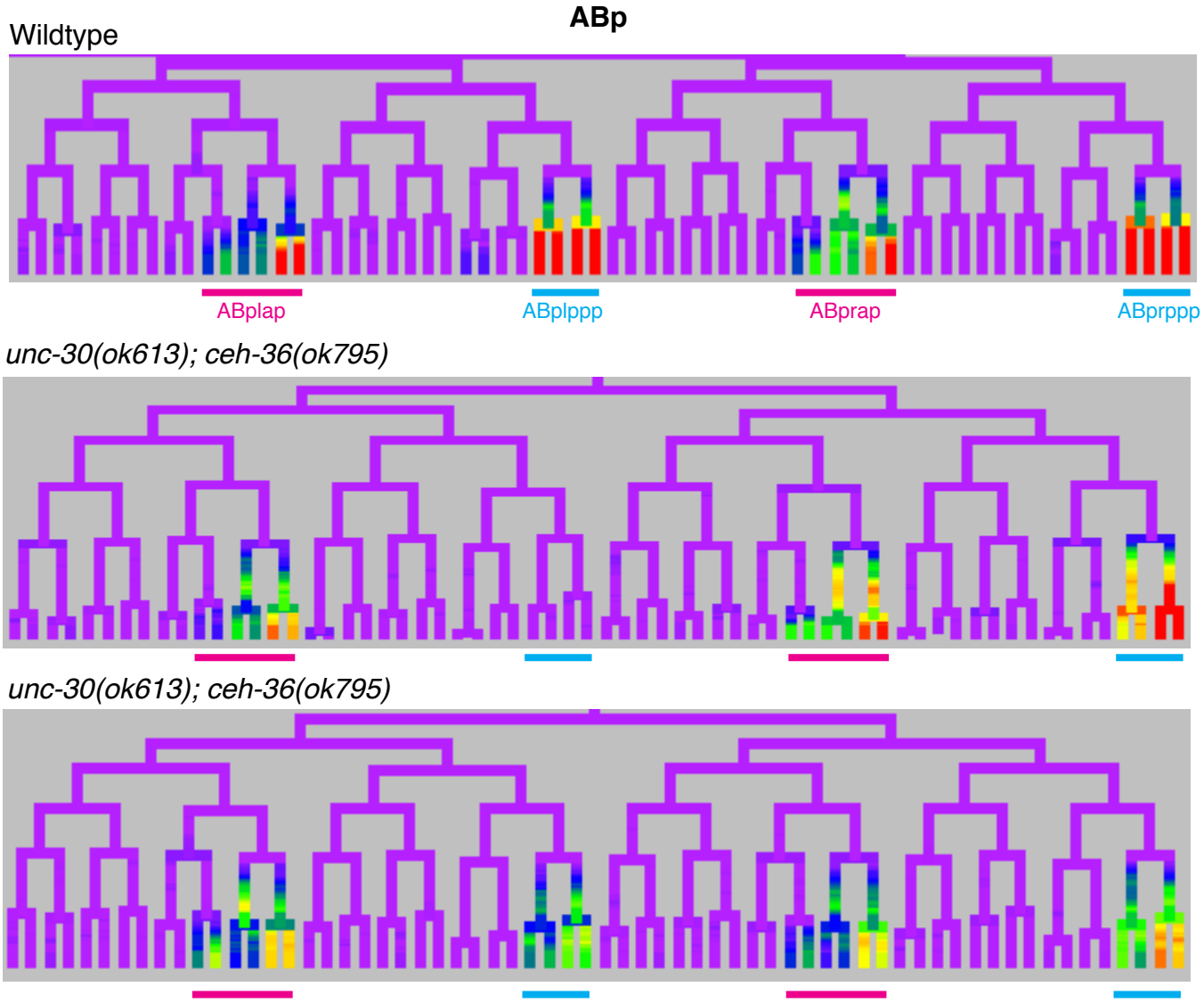
Supplemental Figure 6



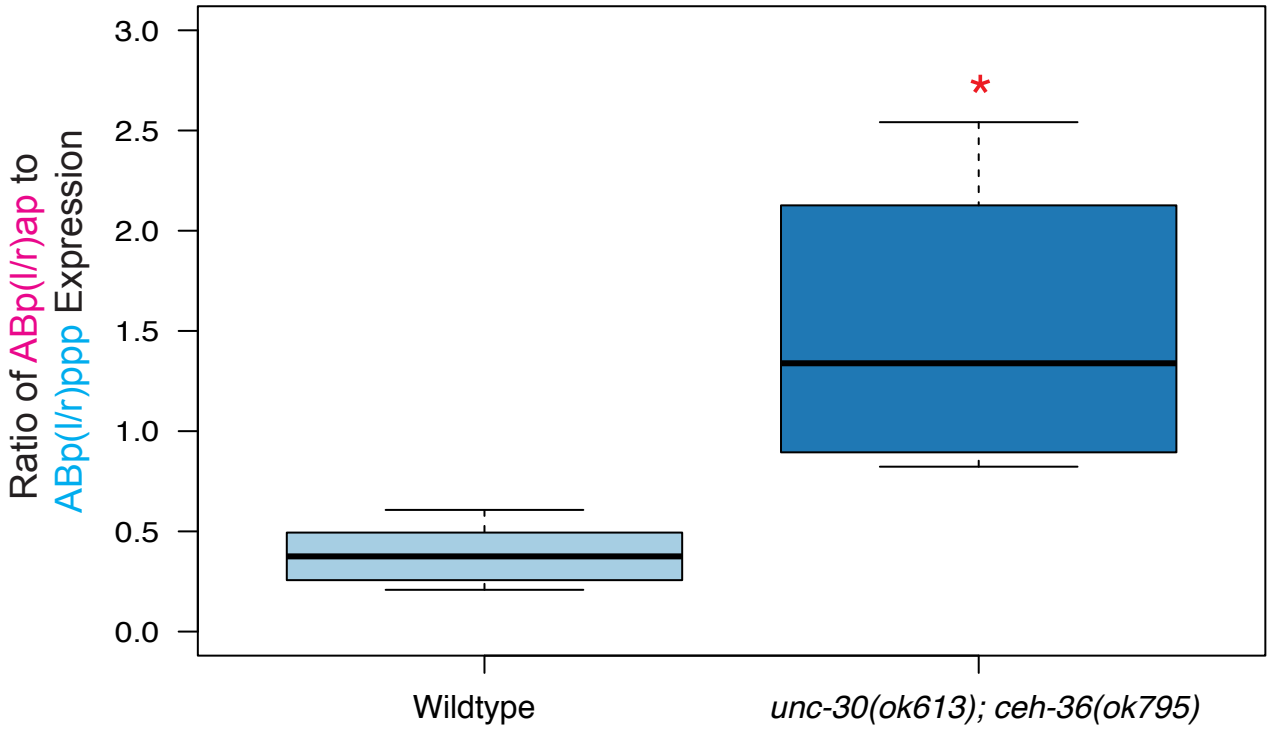


Supplemental Figure 7

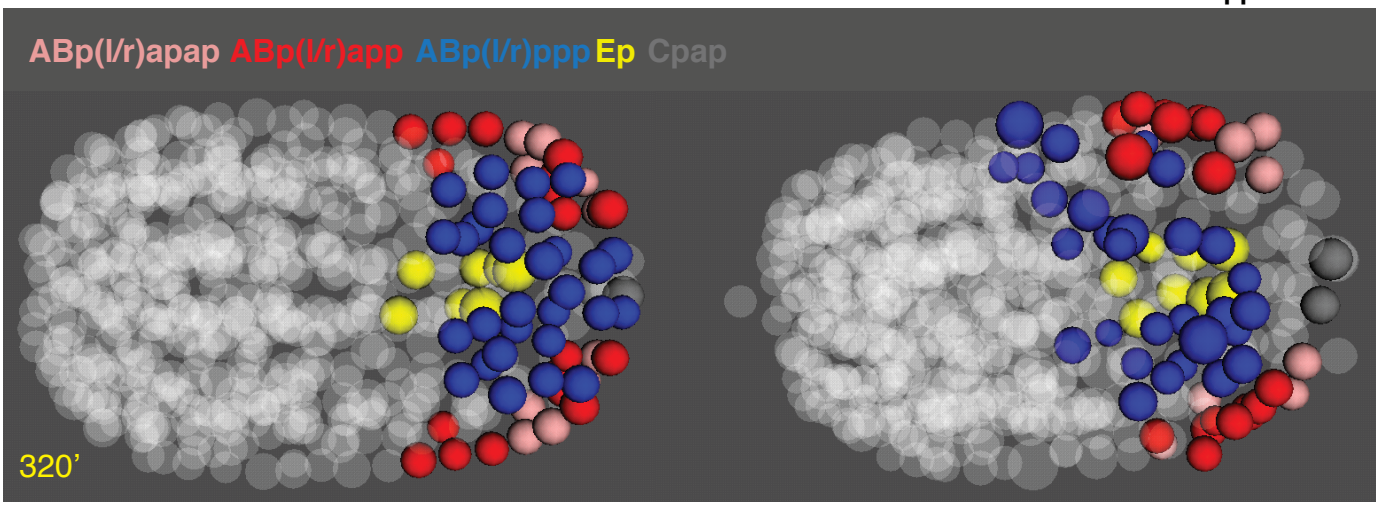
A



B

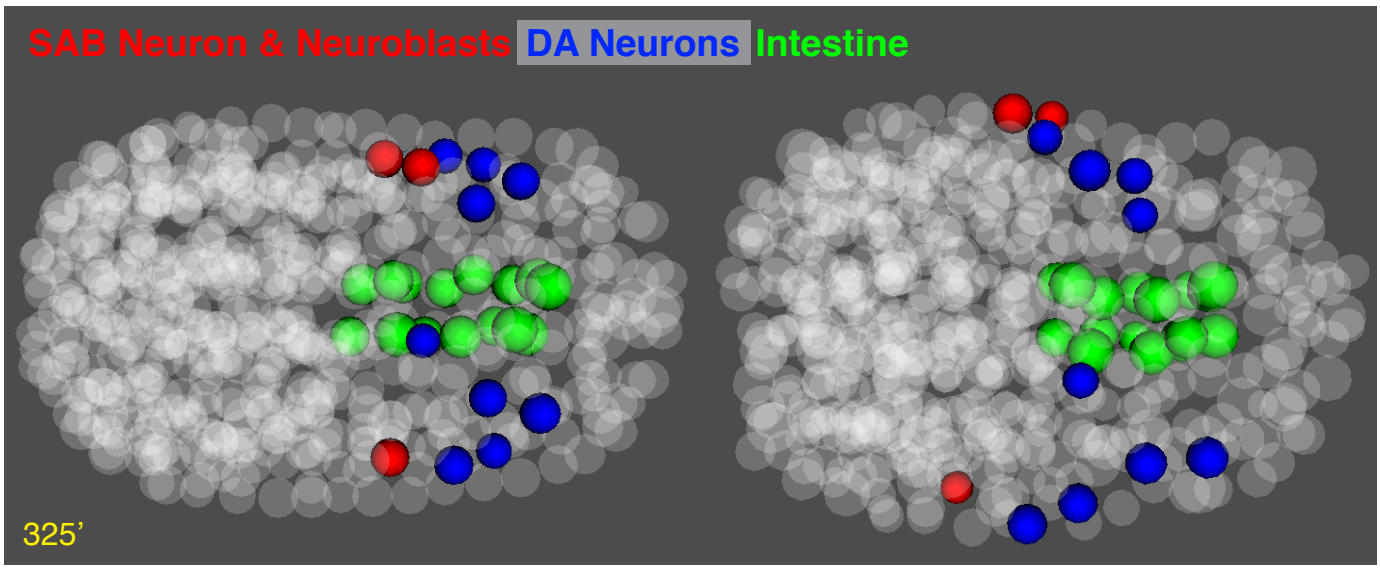


A



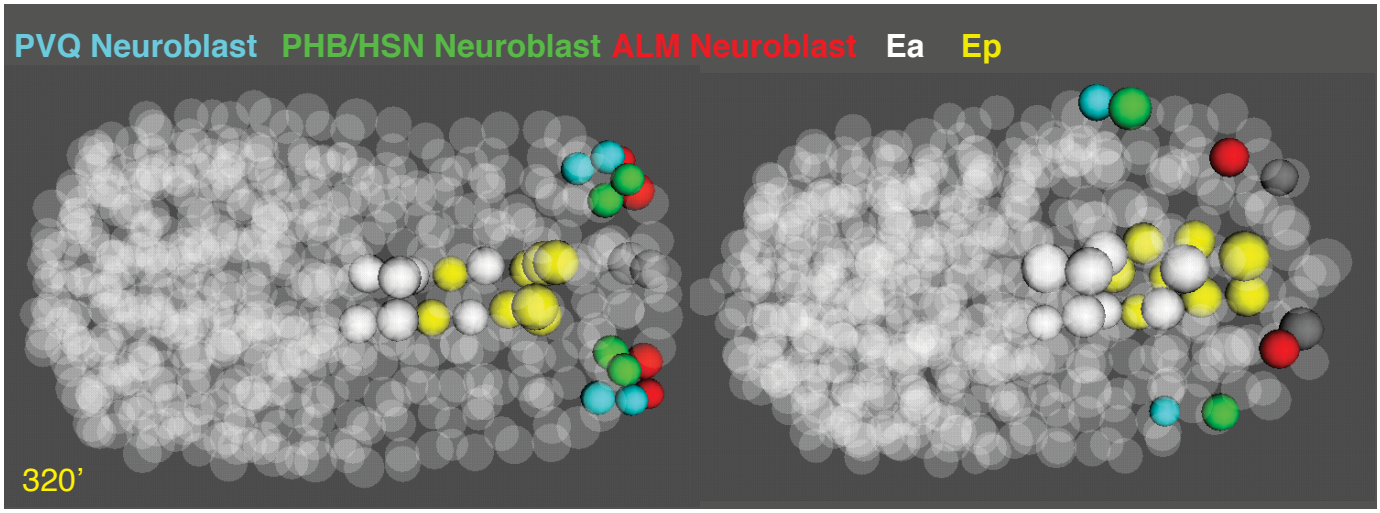
Wild Type | *nob-1(ct223)*

B



Wild Type | *ceh-13(sw1)*

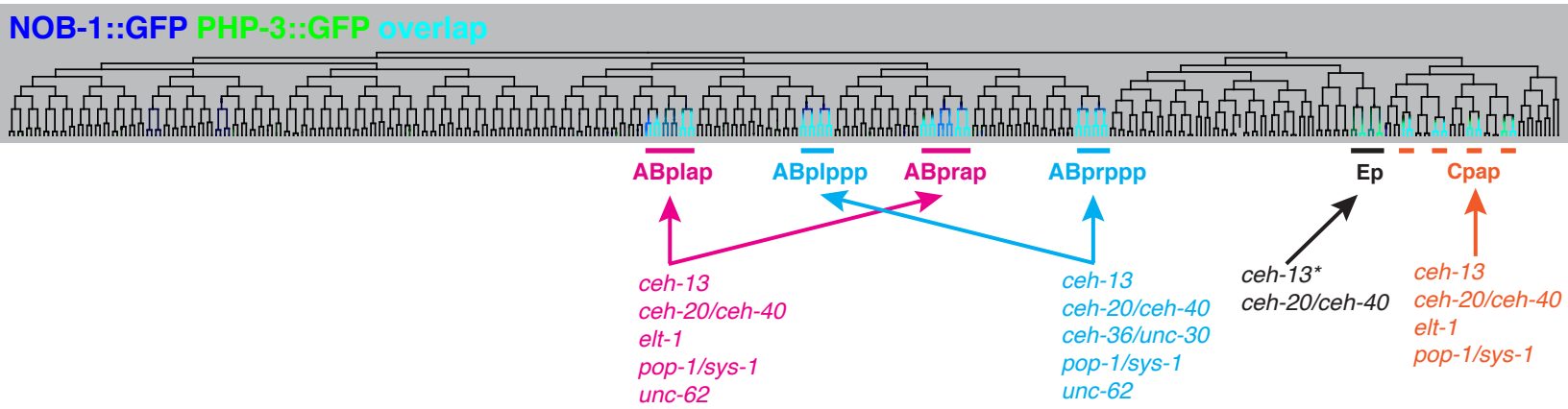
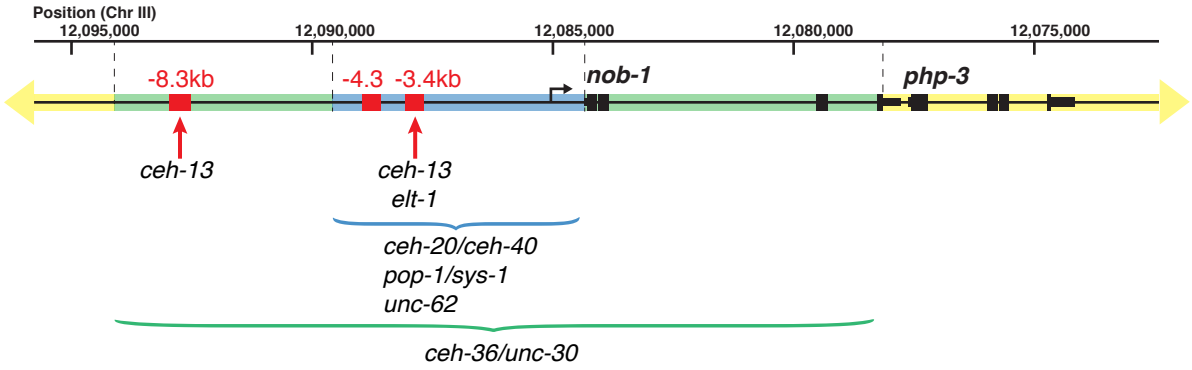
C



Wild Type | *nob-1(ct223)*

A Lineage level regulation

Supplemental Figure 9



B Enhancer level regulation