Gene duplication and evolutionary plasticity of *lin-12/Notch* gene function in *Caenorhabditis*

Haimeng Lyu,[1] Nicolas D. Moya,[2] Erik C. Andersen,[2] Helen M. Chamberlin ⓘ [1],*

[1]Department of Molecular Genetics, Ohio State University, 484 W 12th Ave, Columbus, OH 43210, USA
[2]Department of Biology, Johns Hopkins University, Bascom UTL 383, 3400 North Charles St., Baltimore, MD 21218, USA

*Corresponding author: Department of Molecular Genetics, Ohio State University, 484 W 12th Ave, Columbus, OH 43210, USA. Email: chamberlin.27@osu.edu

Gene duplication is an important substrate for the evolution of new gene functions, but the impacts of gene duplicates on their own activities and on the developmental networks in which they act are poorly understood. Here, we use a natural experiment of *lin-12/Notch* gene duplication within the nematode genus *Caenorhabditis*, combined with characterization of loss- and gain-of-function mutations, to uncover functional distinctions between the duplicate genes in 1 species (*Caenorhabditis briggsae*) and their single-copy ortholog in *Caenorhabditis elegans*. First, using improved genomic sequence and gene model characterization, we confirm that the *C. briggsae* genome includes 2 complete *lin-12* genes, whereas most other genes encoding proteins that participate in the LIN-12 signaling pathway retain a one-to-one orthology with *C. elegans*. We use CRISPR-mediated genome editing to introduce alleles predicted to cause gain-of-function (*gf*) or loss-of-function (*lf*) into each *C. briggsae* gene and find that the *gf* mutations uncover functional distinctions not apparent from the *lf* alleles. Specifically, *Cbr-lin-12.1(gf)*, but not *Cbr-lin-12.2(gf)*, causes developmental defects similar to those observed in *Cel-lin-12(gf)*. In contrast to *Cel-lin-12(gf)*, however, the *Cbr-lin-12.1(gf)* alleles do not cause dominant phenotypes as compared to the wild type, and the mutant phenotype is observed only when 2 *gf* alleles are present. Our results demonstrate that gene duplicates can exhibit differential capacities to compensate for each other and to interfere with normal development, and uncover coincident gene duplication and evolution of developmental sensitivity to LIN-12/Notch activity.

Keywords: Notch signaling; gene duplication; nematode vulval development; comparative genetics

Introduction

Gene duplication is an important source of material for the evolution of new gene functions and biological novelty (Force *et al.* 1999; Innan and Kondrashov 2010; Birchler and Yang 2022). Following a single duplication event, increased gene dose or the stoichiometric mis-match between functional partners can cause fitness costs, especially for genes involved in essential processes (Birchler *et al.* 2005; Blomme *et al.* 2006; Konrad *et al.* 2018; Zhang *et al.* 2022). The adaptive changes or developmental circumstances that correlate with the stable maintenance of duplication events in such cases are poorly understood. *Notch* genes in the nematode genus *Caenorhabditis* provide a specific example where a developmentally important gene has undergone multiple stable duplication events (Stevens *et al.* 2019). The 2 *Notch* genes in *Caenorhabditis elegans* (*Cel-glp-1* and *Cel-lin-12*) have been genetically characterized, and each has distinct and overlapping developmental functions in vivo (Yochem and Greenwald 1989; Lambie and Kimble 1991). *Cel-lin-12*, for example, has been shown to mediate several developmental decisions yielding 2 distinct cell types, where *Cel-lin-12(lf)* and *Cel-lin-12(gf)* mutations yield opposite phenotypes (Greenwald *et al.* 1983). Previous studies identified a duplication of the *lin-12* gene in *Caenorhabditis briggsae* (Rudel and Kimble 2002; Stevens *et al.* 2019), providing a natural experiment to evaluate the impact of a more recent *Notch* gene duplication event on gene function and on *Notch* developmental networks.

Here, we use available sequence and gene model characterization from the AF16-related strain QX1410 (Stevens *et al.* 2022; Moya *et al.* 2023) to show that the *C. briggsae* genome includes 2 complete *lin-12* genes, whereas most other genes encoding LIN-12 signaling pathway proteins retain a one-to-one orthology with *C. elegans*. Using CRISPR-mediated genome editing, we introduce alleles predicted to cause gain-of-function (*gf*) or loss-of-function (*lf*) for each gene. We find that the wild-type alleles of each gene are able to compensate for loss of the other. *Cbr-lin-12.1(gf)*, but not *Cbr-lin-12.2(gf)*, causes defects similar to those observed in *Cel-lin-12(gf)*. Notably, however, the *Cbr-lin-12.1(gf)* alleles do not cause dominant phenotypes as compared to the wild type, and the mutant phenotype is observed only when 2 *gf* alleles are present. Together, these results demonstrate overlapping but distinct capabilities for the gene duplicates. The results suggest genetic and developmental network conditions that could accommodate the stabilization of duplication events and highlight how functions might distribute among gene duplicates.

Results and discussion

The *C. briggsae* genome includes 2 *lin-12* genes

Since the commonly used *C. briggsae* AF16 reference genome annotates 3 incomplete *Cbr-lin-12* genes (Stein *et al.* 2003; Hillier *et al.* 2007), we used recent chromosome-level genome sequence, RNA

sequencing data, and gene model curation for the closely related strain QX1410 (Stevens *et al.* 2022; Moya *et al.* 2023) to confirm the number and structure of *Cbr-lin-12* genes. Because these strains are highly similar to each other, we anticipated that the vast majority of differences reflect differences between genome assembly or annotation, rather than strain-specific variation. We found that the QX1410 genome annotation includes 2 full-length *Cbr-lin-12* genes, compared to the 3 partial genes in AF16 (Fig. 1; Stein *et al.* 2003; Hillier *et al.* 2007). These 2 genes correspond to the 2 identified previously (Rudel and Kimble 2002) and contained on the sequenced AF16 cosmid CB020L07 (Genbank AC140918.1). We used the nomenclature of Rudel and Kimble, assigning the gene with 9 exons *Cbr-lin-12.1*, and the one with 8, *Cbr-lin-12.2*. Previous analysis of *Caenorhabditis* genomes found only a single *lin-12* ortholog in 17 other species of the *Elegans* supergroup (Stevens *et al.* 2019). Data from the closely related *Caenorhabditis nigoni* (not analyzed in Stevens *et al.* 2019) indicate that the

duplication event predates divergence between *C. nigoni* and *C. briggsae* [Supplementary File 2 (Supplementary Figs. 1–3)]. However, in contrast to *Cbr-lin-12.2*, *Cni-lin-12.2* includes an in-frame stop, suggesting distinct outcomes for the duplicates within this lineage.

We completed a comparison of QX1410 and AF16 to identify any Notch pathway gene changes in *C. briggsae* coincident with the receptor gene duplication (Greenwald and Kovall 2013) (Table 1). Of the 16 genes evaluated, most exhibited one-to-one orthology with each other and with the *C. elegans* N2 reference genome. Notable exceptions relate to the ligand-encoding genes *arg-1* and *dsl-1*. *arg-1* lacks a protein-encoding ortholog in both AF16 and QX1410 (this work and Zhao *et al.* 2008), whereas *dsl-1* has duplicated in *C. briggsae* [Supplementary File 2 (Supplementary Figs. 4 and 5)]. In addition to gene copy number, protein-length accuracy for each gene was estimated (Fig. 2), with improvements observed in QX1410 compared to AF16. Overall, we conclude that, aside from *lin-12*, *arg-1* and *dsl-1*, the *C. elegans* and *C. briggsae* genomes have a one-to-one correspondence for critical Notch pathway members.

CRISPR-mediated genome editing recovers alleles in *Cbr-lin-12* genes predicted to confer gain and loss of gene activity

To investigate the functions associated with each *Cbr-lin-12* gene, we used CRISPR-mediated genome editing to introduce gain-of-function (*gf*) and loss-of-function (*lf*) mutations individually and in combination [Supplementary File 1 (Supplementary Table 2); Supplementary File 2 (Supplementary Figs. 6 and 7)]. We identified at least 1 strain representing each of the possible allele combinations, except for the double *lf* class. All recovered genotypes, except for the *Cbr-lin-12.1(lf)* *Cbr-lin-12.2(gf)* double, were viable and fertile as homozygotes. The *Cbr-lin-12.1(lf)* *Cbr-lin-12.2(gf)* mutants exhibit an arrest phenotype and hindgut defects similar to that observed by Rudel and Kimble following *Cbr-lin-12(RNAi)* treatment [Supplementary File 2 (Supplementary Fig. 8); Rudel and Kimble 2002]. Animals homozygous for a *lf* allele for either gene or *Cbr-lin-12.2(gf)* did not exhibit any overt mutant phenotypes. All strains that included a *Cbr-lin-12.1(gf)* allele exhibited a high frequency of vulva and egg-laying defects. We interpret that either gene is sufficient for grossly normal development and fertility when the other is altered with a *lf* mutation. However, a *gf* mutation in *Cbr-lin-12.1*, but not *Cbr-lin-12.2*, is
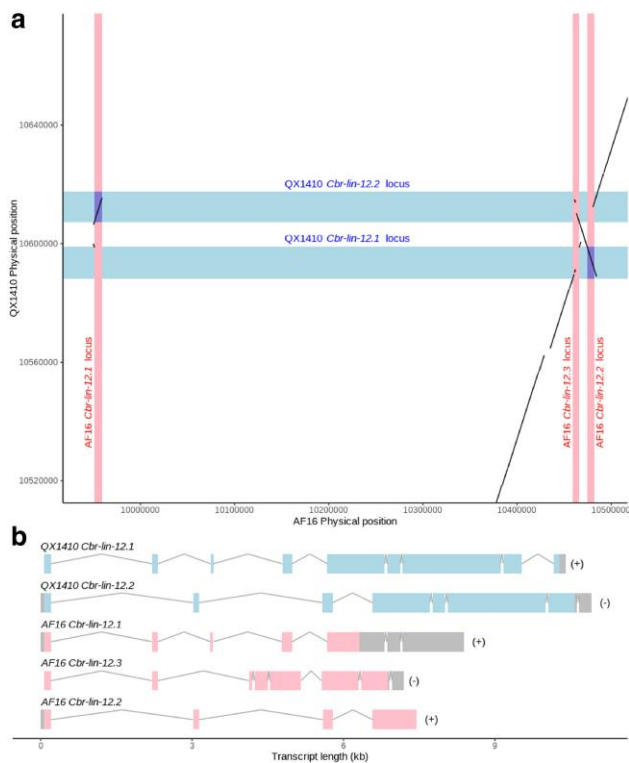

Fig. 1. The *C. briggsae* genome includes 2 *lin-12* genes. a) Plot showing the physical position of *C. briggsae lin-12* orthologs in both the AF16 and QX1410 genomes. All 3 *Cbr-lin-12.1*, *Cbr-lin-12.2*, and *Cbr-lin-12.3* genes (on the X-axis) are in individual regions of the AF16 genome assembly that align to either 1 of 2 individual loci in the QX1410 genome assembly (*Cbr-lin-12.1* and *Cbr-lin-12.2*, on the Y-axis). In the QX1410 genome assembly, the 2 genes are transcribed from opposite strands of the DNA, separated by about 8 kb of sequence. b) Gene models for *Cbr-lin-12* orthologs from AF16 and QX1410 assemblies. The coding sequences of AF16 *Cbr-lin-12* genes appear to be truncated. A local inversion positioned at 10.48 Mb has disrupted the coding structure of the 3′ terminus of AF16 *Cbr-lin-12.2*. The gene model of QX1410 *Cbr-lin-12.2* seems to have corrected the truncation in the gene model of AF16 *Cbr-lin-12.2*. Similarly, the coding sequence of AF16 *Cbr-lin-12.3* is affected by the inversion but also disrupted by a highly fragmented local assembly, leading to a chimeric gene model. The first 2.3 kb of *Cbr-lin-12.3* and *Cbr-lin-12.1* in the AF16 assembly is identical in sequence. The gene model of AF16 *Cbr-lin-12.1* seems to have been partially duplicated from the *Cbr-lin-12.3* locus and mispositioned at 9.95 Mb. The QX1410 *Cbr-lin-12.1* gene collapses the truncated coding sequences of AF16 *Cbr-lin-12.1* and *Cbr-lin-12.3* into a single, nonredundant gene model with full-length coding sequence based on long-read RNA-seq and confirmed with Sanger sequencing of cDNA.

Table 1. Count of *C. briggsae* orthologs of LIN-12/Notch signaling genes.

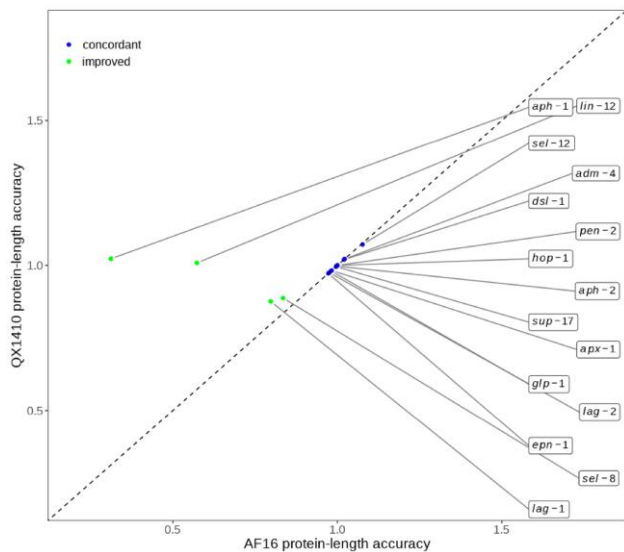| N2	QX1410	AF16
dsl-1	4	1
aph-1	1	2
lin-12	2	3
arg-1	0	0
apx-1	1	1
hop-1	1	1
aph-2	1	1
sup-17	1	1
sel-12	1	1
adm-4	1	1
epn-1	1	1
pen-2	1	1
sel-8	1	1
glp-1	1	1
lag-1	1	1
lag-2	1	1

Fig. 2. Evaluation of LIN-12/Notch signaling genes in *C. briggsae*. Comparisons between the protein-length accuracies of *C. briggsae* genes (QX1410 and AF16 strain genome assemblies) orthologous to *C. elegans* (N2 strain) Notch signaling genes. Protein-length accuracy was calculated by dividing each *C. briggsae* protein sequence length by the length of its respective *C. elegans* N2 ortholog. A *C. briggsae* gene with a protein-length accuracy of 1 is identical in protein sequence length to its *C. elegans* ortholog. A gene with coded protein length that is more similar to that of its *C. elegans* ortholog in the QX1410 genome than AF16 genome is considered improved (above the dashed line). Genes that are identical in protein length between QX1410 and AF16 are considered concordant (aligned with the dashed line).

sufficient to interfere with normal development. *Cbr-lin-12.1* genotype, in contrast, can alter the effect of *Cbr-lin-12.2*(*gf*), as *Cbr-lin-12.1*(+) provides sufficient wild-type function to compensate for any genotype at *Cbr-lin-12.2*, whereas in a *Cbr-lin-12.1*(*lf*) background *Cbr-lin-12.2*(*gf*) is lethal. We hypothesize that this difference between loss- and gain-of-function genetics results because the only distinction is the capacity of each gene (either based on abundance or specific sequence differences) to interfere with the activity of the other when subject to ligand-independent activation. Our analysis demonstrates that probing duplicate gene function using both loss- and gain-of-function alleles uncovers redundancies and functional differences not discovered through loss-of-function experiments alone.

Cbr-lin-12.1(*gf*) animals exhibit defects in egg-laying system development

To understand the cellular basis for the *Cbr-lin-12.1* egg-laying defects, we focused on known roles for *Cel-lin-12* in development of the AC/VU decision in the somatic gonad and patterning of cell fates among the vulval precursor cells (VPCs) (Greenwald *et al.* 1983). Wild-type animals produce a single AC following a *Cel-lin-12*-mediated interaction between 2 initially equivalent precursor cells (Z1.ppp/Z4.aaa; Wilkinson *et al.* 1994). This cell then produces a graded signal (LIN-3/EGF) that "induces" vulva development causing the VPC most proximal to the AC (P6.p) to initiate a developmental program that includes expression of Cel-LIN-12 ligands (Chen and Greenwald 2004). These ligands activate Cel-LIN-12 in 2 neighboring cells (P5.p and P7.p) that divide to produce the 2° cell fate (Sternberg and Horvitz 1989). To assess the function of *Cbr-lin-12* in the AC/VU decision, we evaluated expression of a *lin-3::GFP* transgene and found that, like *Cel-lin-12*(*gf*) mutants, animals homozygous for *Cbr-lin-12.1*(*gf*) did not produce

an AC (Fig. 3). The gonad of animals homozygous for other genotypes includes a single AC. To evaluate the impact of *Cbr-lin-12* mutations on vulval cell development, we used 2 assays: VPC induction (the number of VPCs that divide to produce vulval tissue) and expression in differentiating vulval cells of an *egl-17::gfp* (2°-specific in L4 animals) reporter transgene (Félix 2007). Normal vulval development in *C. briggsae* is similar to that of *C. elegans*, and the production of the vulva is dependent on the AC (Félix 2007). We find that animals homozygous for *Cbr-lin-12.1*(*lf*), *Cbr-lin-12.2*(*lf*), or *Cbr-lin-12.2*(*gf*) exhibit normal vulval development with respect to the numbers of induced VPCs, inferred patterns of cell division, and expression of *egl-17::gfp* (Fig. 4). By contrast, as in *Cel-lin-12*(*gf*) mutants, the VPCs divide to produce vulval tissue in *Cbr-lin-12.1*(*gf*) mutants despite the absence of an AC (Fig. 4a; also apparent in Fig. 3a–n). These mutants can exhibit a multivulva or Muv phenotype where more than 3 VPCs produce vulval tissue, with the number of induced cells ranging from 3 to 6. The number of GFP-positive cells varies considerably across animals (from 2–18), and the pattern of expression and the inferred pattern of cell division frequently does not conform with the wild-type 2° cell lineage. Nevertheless, reliable expression of this marker indicates that, as in *Cel-lin-12*(*gf*) mutants, the VPCs in *Cbr-lin-12.1*(*gf*) mutants can adopt the 2° fate independent of inductive signal from the AC. Altogether, the results show that *Cbr-lin-12.1*(*gf*), but not *Cbr-lin-12.2*(*gf*), interferes with the normal development of the AC and the vulva in a manner similar to that seen in *Cel-lin-12*(*gf*).

Cbr-lin-12.1(*gf*) alleles cause recessive egg-laying defects as compared to the wild type

C. elegans hermaphrodites are highly sensitive to activation of LIN-12, and *Cel-lin-12*(*gf*) alleles like *n137* confer a strong gain-of-function phenotype in heterozygotes (Fig. 5; Greenwald *et al.* 1983). By contrast, we found that, although homozygotes exhibit a strong mutant phenotype, the *Cbr-lin-12.1*(*gf*) phenotype is generally recessive to the *Cbr-lin-12.1*(+) wild-type phenotype, an effect that is not altered by *Cbr-lin-12.2* genotype (Fig. 5). In addition, we find that animals of *Cbr-lin-12.1*(*gf*/*lf*) genotype are grossly similar to *Cbr-lin-12.1*(*gf*/+) animals. This result indicates that the gain-of-function phenotype is caused by the presence of 2 *gf* alleles, rather than competition between the products of the *gf* and wild-type alleles. We conclude that, although *Cbr-lin-12.1*(*gf*) mutant alleles interfere with normal development in a manner similar to *Cel-lin-12*(*gf*) alleles, the sensitivity of the animals to the *gf* variant and the relationship between *gf* and wild-type alleles are remarkably distinct. This tolerance could be because of sequence differences in LIN-12, such that the S872F-related substitutions are not as destabilizing in the *C. briggsae* proteins. Alternatively, it could be that the signaling networks that involve LIN-12 in *C. briggsae* cells are less dependent or more able to compensate for perturbations to LIN-12, or otherwise tuned to respond to a different (higher) dose of active LIN-12. Altogether, we have uncovered functional distinctions for recent *lin-12*/Notch gene duplicates, and identified how these changes coincide with developmental sensitivity to LIN-12/Notch activity. We anticipate that these changes reflect the adaptive constraints for single gene duplicates within essential developmental networks.

Materials and methods
Genome alignment

The *C. briggsae* AF16 (WormBase, WS280; Davis *et al.* 2022) and *C. nigoni* JU1422 (WormBase Parasite, WBPS19; Howe *et al.* 2017)
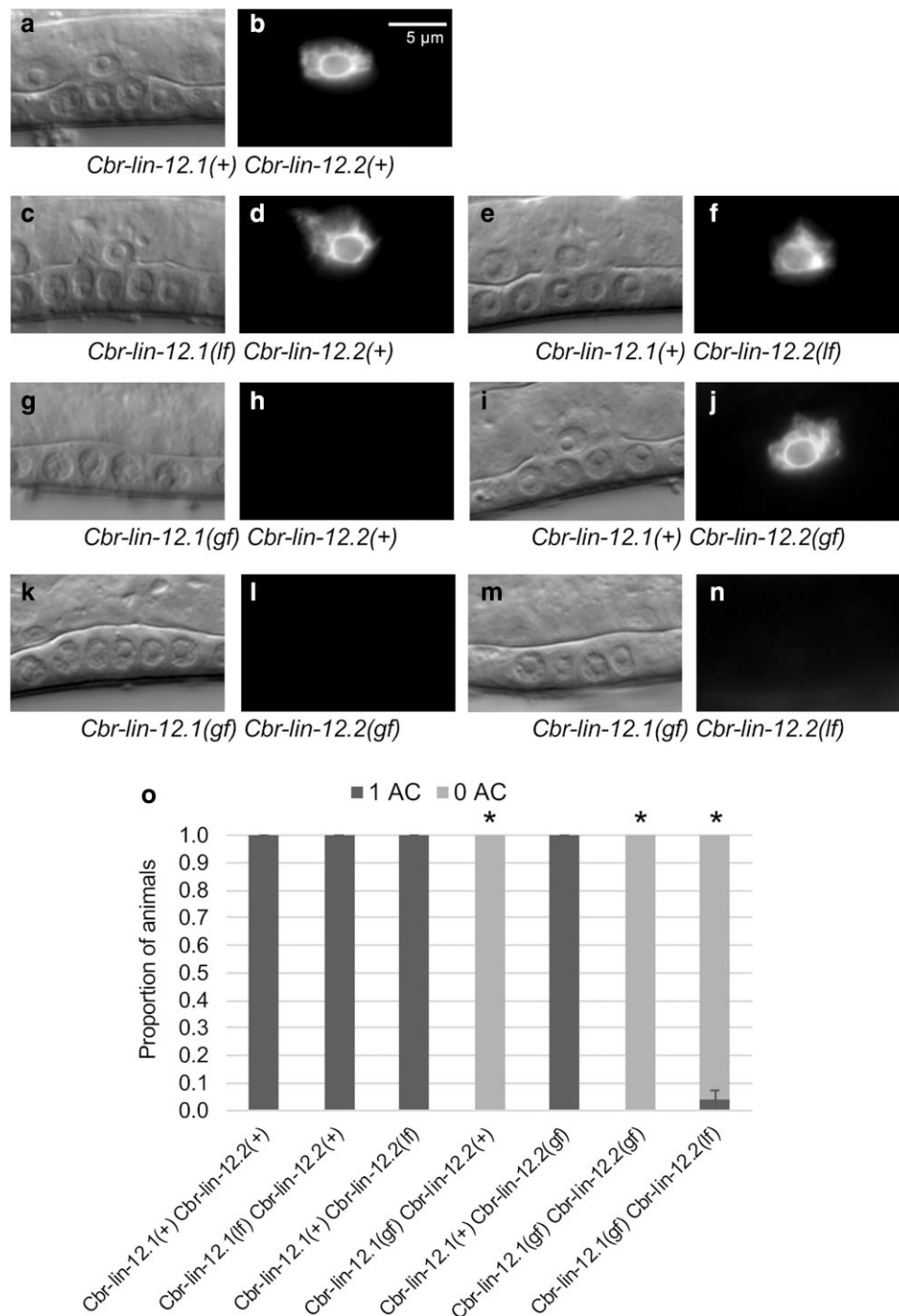
Fig. 3. Animals bearing gain-of-function (*gf*) mutations in *Cbr-lin-12.1* lack an anchor cell (AC). (a–n) DIC and fluorescent micrographs of L3 animals bearing a *Cel-lin-3::gfp* transgene that expresses in the AC. All animals are at the stage where dividing VPCs have completed 2 of 3 rounds of cell division. Wild type and animals homozygous for loss-of-function (*lf*) alleles of either *Cbr-lin-12* gene have a single AC. No animals with 2 ACs were observed. Animals homozygous for *Cbr-lin-12.1(gf)* and with any genotype at *Cbr-lin-12.2* lack an AC, yet the VPCs divide to produce vulval tissue. (o) Quantification of the phenotype. Sample size ≥ 25 for each genotype. *Indicates statistically different from wild type control, $P < 0.05$, Fisher's exact test subject to Bonferroni correction. Full *Cbr-lin-12* genotypes (in order in the figures and on the bar graph): AF16 (wild type); *Cbr-lin-12.1(gu272lf) Cbr-lin-12.2(+)*; *Cbr-lin-12.1(+) Cbr-lin-12.2(gu254lf)*; *Cbr-lin-12.1(gu269gf) Cbr-lin-12.2(+)*; *Cbr-lin-12.1(+) Cbr-lin-12.2(gu260gf)*; *Cbr-lin-12.1(gu259gf) Cbr-lin-12.2(gu298gf)*; *Cbr-lin-12.1(gu258gf) Cbr-lin-12.2(gu297lf)*. Full strain genotypes listed in Supplementary File 1 (Supplementary Table 1).

genomes were aligned against *C. briggsae* QX1410 (NCBI, PRJNA784955) using NUCleotide MUMmer (NUCmer) v3.1, allowing a maximum gap of 500 bp (Marçais *et al.* 2018). Duplicated sequences of the AF16 genomes were identified using R by selecting alignments that had a single set of coordinates in the QX1410 genome but distinct sets of coordinates in the AF16 genome.

Gene model visualization

Gene models were plotted using R with the ggplot2 package (Wickham *et al.* 2016). The physical coordinates of gene features were extracted from Gene Feature Format (GFF) files retrieved from WormBase (WS280 release), WormBase ParaSite (release WBPS19), and NCBI (PRJNA784955). The gene models were
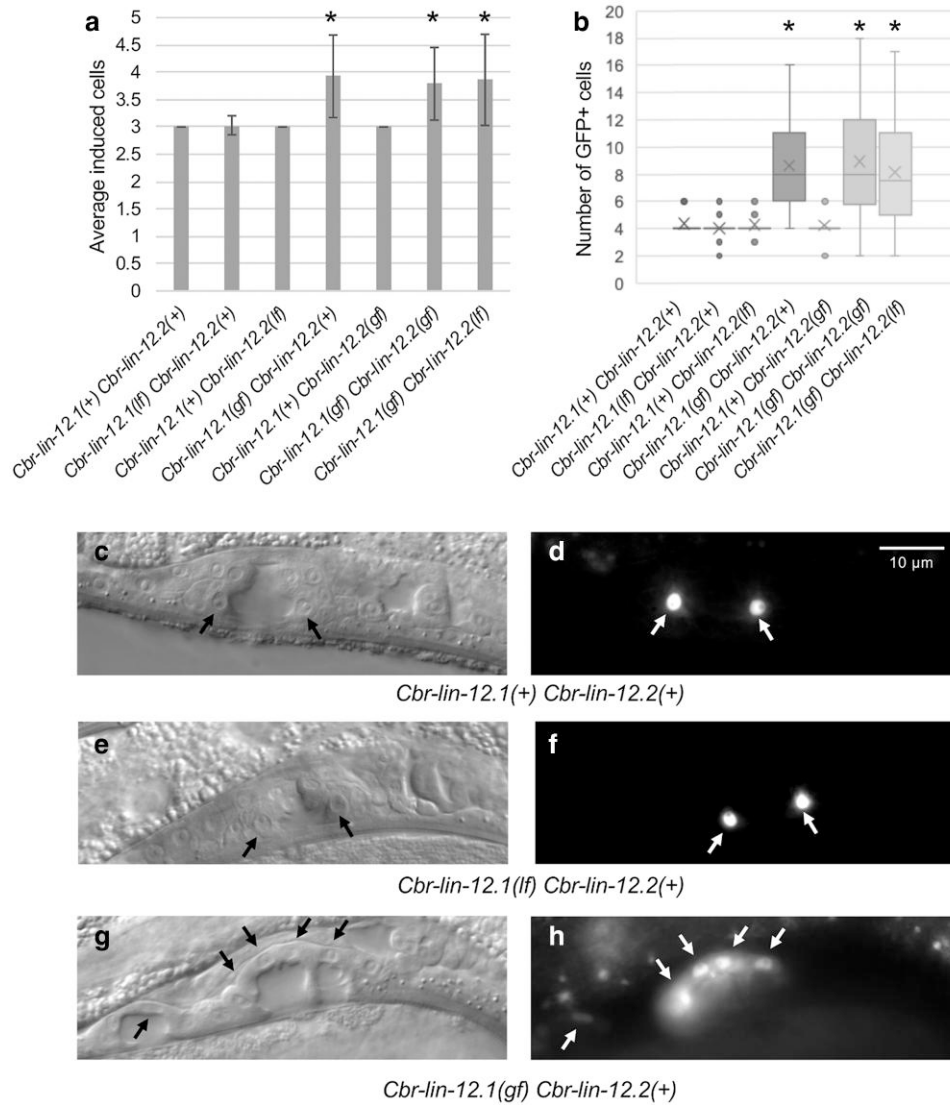
Fig. 4. An excess of vulval precursor cells (VPCs) in animals bearing gain-of-function alleles of *Cbr-lin-12.1* divide to produce vulval cell types that express a 2° lineage-specific marker. a) In wild-type L3 animals, 3 of 6 VPCs divide to produce vulval tissue. In animals homozygous for *Cbr-lin-12.1(gf)*, up to 6 (range 3–6) cells divide to produce vulval tissue, despite the absence of an AC. The VPCs in animals homozygous for other viable *Cbr-lin-12* genotypes divide with a wild-type pattern. Error bars correspond to standard deviation. *Indicates statistically different from wild type control, $P < 0.05$, 1-way ANOVA and Tukey HSD. b) In mid-L4 animals, the reporter *egl-17::gfp* is expressed exclusively in vulval cells that derive from the 2° lineage, normally produced by the P5.p and P7.p cells. In wild type, fluorescence is detected in 4 or sometimes 6 cells, representing the 4 vulC cells (2 from each precursor), or these cells plus the 2 vulD cells (1 from each precursor). In animals homozygous for *Cbr-lin-12.1(gf)*, significantly more vulval cells are GFP-positive, and the pattern can be highly variable. Both the cells that are typically induced (P5.p–P7.p) and the ectopic ones can produce cells that express *egl-17::gfp*. Data represented as box and whisker plot (representing median and quartiles), with mean indicated with an X, and outliers as a dot. The vast majority of animals with genotypes that do not include *Cbr-lin-12.1(gf)* include only 4 GFP-positive vulval cells. *Indicates statistically different from wild type control, $P < 0.05$, Kruskal–Wallis test subject to Bonferroni correction. (c–h) Representative images of vulval cells in mid-L4 animals. In the wild type and *Cbr-lin-12.1(lf)*, the plane showing 1 side of vulC cells is shown (2 cells, 1 from each 2° precursor). In the *Cbr-lin-12.1(gf)*, a comparable plane is shown. Sample size ≥ 30 for all conditions. Full *Cbr-lin-12* genotypes as in Fig. 3.

confirmed by long-read RNA sequencing from QX1410 (Moya *et al.* 2023) and *Cbr-lin-12* models by targeted Sanger sequencing of cDNA generated from AF16 RNA.

Orthology and protein length analysis

Protein sequences for LIN-12/Notch signaling genes were extracted from their respective GFF files using Gffread v0.12.1 (Pertea and Pertea 2020). Only the longest isoform for each gene was kept using the *agat_sp_keep_longest_isoform.pl* script from AGAT v0.8.1 (Dainat *et al.* 2022). Orthologous relationships between protein sequences across *C. elegans*, *C. briggsae*, and *C. nigoni* were drawn using OrthoFinder v2.1.4 (Emms and Kelly 2019) with default parameters.

Protein sequence alignments between *C. briggsae* and *C. nigoni lin-12* genes were visualized using the Clustal Omega website (https://www.ebi.ac.uk/jdispatcher/msa/clustalo).

Maximum-likelihood lin-12 tree

Protein sequences from *C. elegans*, *C. briggsae*, and *C. nigoni lin-12* orthologs were aligned using MAFFT v7.471 (Katoh and Standley 2013). Maximum-likelihood tree was modeled from protein alignment using IQ-TREE v2.2.5 (Nguyen *et al.* 2015) with automatic model selection and 1,000 iterations of ultra-fast bootstrapping to estimate branch support. The model selected by IQ-TREE was WAG + F + G4.
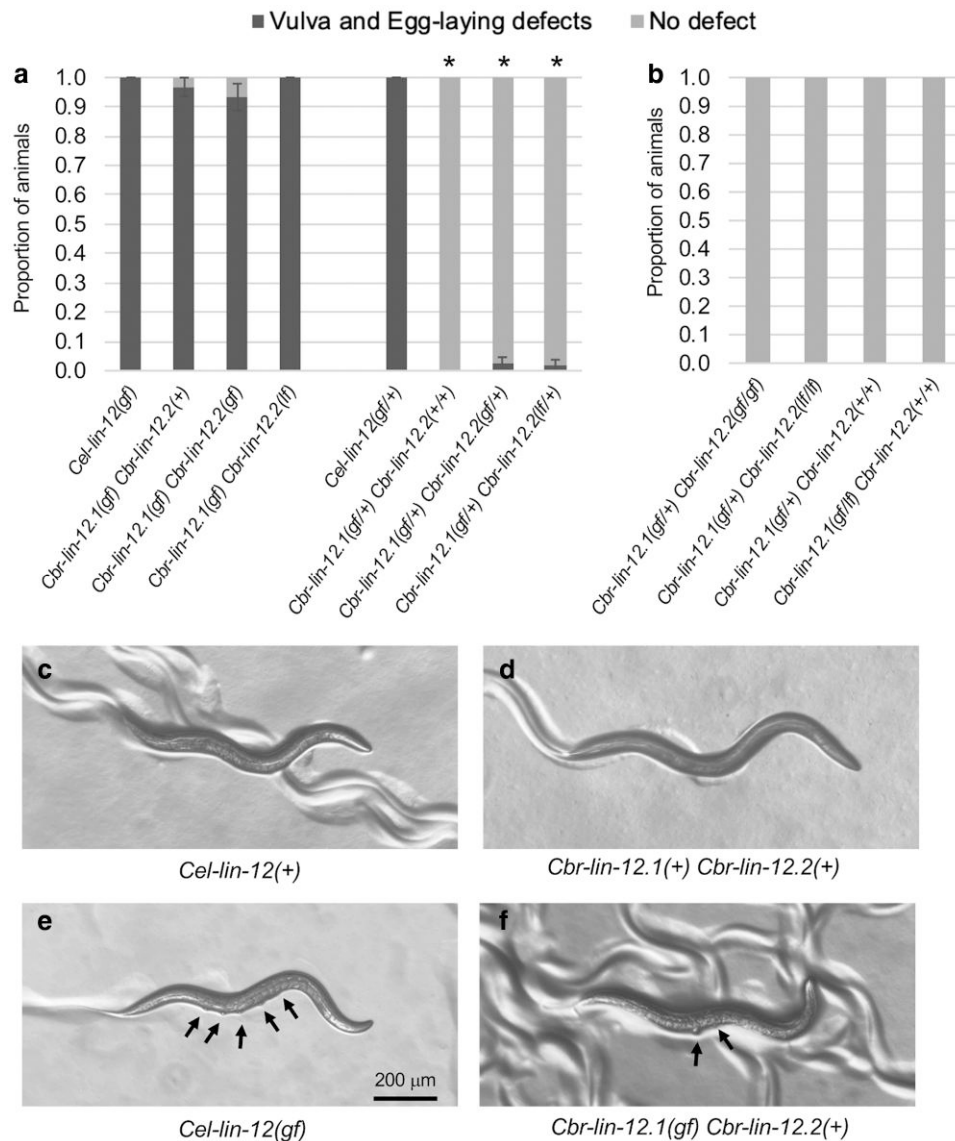
Fig. 5. The *Cbr-lin-12.1*(*gf*) mutations are recessive, and require 2 mutant alleles to confer a phenotype. a) Comparison of phenotype in *Cel-lin-12*(*gf*) and *Cbr-lin-12.1*(*gf*) homozygotes and heterozygotes. Homozygous animals (derived from a heterozygous parent) of all genotypes exhibit a high frequency of vulva or egg-laying defects. *Cel-lin-12*(*gf/+*) heterozygotes are similarly abnormal, whereas most *Cbr-lin-12.1*(*gf/+*) animals exhibit no egg-laying defects. *Indicates statistically different from homozygous control, $P < 0.05$, Fisher's exact test subject to Bonferroni correction. b) *Cbr-lin-12.1*(*gf/+*) animals are grossly normal, irrespective of the genotype of *Cbr-lin-12.2*. *Cbr-lin-12.1*(*gf/lf*) animals are also grossly normal, suggesting the defect results from presence of 2 *Cbr-lin-12.1*(*gf*) gene copies, rather than interference of wild type with mutant gene product. (c–f) Representative images for wild type and *lin-12*(*gf*) homozygotes. *Cel-lin-12*(*gf*) mutants (c and d) typically exhibit 5–6 small ventral protrusions (multivulva or Muv phenotype), whereas (e and f) *Cbr-lin-12.1*(*gf*) homozygotes typically have a single, enlarged protrusion, or not more than 2 protrusions. Error bars in (a) correspond to standard error of the proportion. Sample size ≥ 25 for each condition. Full *lin-12* genotypes (homozygous or heterozygous) for (a) are *Cel-lin-12*(*n137*), *Cbr-lin-12.1*(*gu269gf*) *Cbr-lin-12.2*(*+*); *Cbr-lin-12.1*(*gu259gf*) *Cbr-lin-12.2*(*gu298gf*); *Cbr-lin-12.1*(*gu258gf*) *Cbr-lin-12.2*(*gu297lf*). Full *Cbr-lin-12* genotypes for B. are *Cbr-lin-12.1*(*gu259gf/+*) *Cbr-lin-12.2*(*gu298gf/gu260gf*); *Cbr-lin-12.1*(*gu258gf/+*) *Cbr-lin-12.2*(*gu297lf/gu254lf*); *Cbr-lin-12.1*(*gu269gf/+*) *Cbr-lin-12.2*(*+/+*); *Cbr-lin-12.1*(*gu269gf/gu272lf*) *Cbr-lin-12.2*(*+/+*).

C. nigoni RNA-seq alignment

RNA sequences of control embryos retrieved from PRJNA849332 (Xie *et al.* 2022) were aligned against the *C. nigoni* JU1422 reference genome using STAR v1.5.2 (Dobin *et al.* 2013) in 2-pass mode (–*twopassMode Basic*) using a maximum intron size of 10,000 (–*alignIntronMax* 10000). Alignments were visualized with IGV v2.8.7 (Thorvaldsdóttir *et al.* 2013). RNA alignments that overlapped with the *010.g27411 C. nigoni* gene model were used to manually assemble an extended gene model (*ext4_010.g27411*) that includes 4 additional putative exons and an alternative open-reading frame.

Animal maintenance and genetics

All nematode strains were grown on NGM plates seeded with *Escherichia coli* strain OP50 as a food source, following *C. elegans* protocols (Stiernagle 2006). Strains were grown and experiments performed at 20°C, unless otherwise noted. Strains and genotypes are in Supplementary File 1 (Supplementary Table 1).

CRISPR-mediated genome editing

Mutant alleles of *Cbr-lin-12.1* and *Cbr-lin-12.2* were generated as done previously (Paix *et al.* 2017). A total of 20 ng/μl of a fluorescent marker (*Cel-myo-2::mCherry*, pCFJ90) was included in the injection

mix to identify F1 animals from gametes that received the injected mixture. These mCherry-positive F1 animals were selected and allowed to self-cross. PCR and Sanger sequencing were used to identify plates founded by candidate mutants and to screen for homozygotes. All sgRNA sequences and DNA primers are listed in Supplementary File 1 (Supplementary Table 3). Mutations were backcrossed 2× prior to phenotypic analysis. Sequence details for the mutations are included in Supplementary File 1 (Supplementary Table 2).

Production of transgenes

A *Cel-lin-3::gfp* plasmid (Hwang and Sternberg 2004) was injected into adult *C. briggsae* animals following standard *C. elegans* protocols (Mello *et al.* 1991), in an injection mix containing 50 ng/μl *Cel-lin-3::gfp* (pMC782), 15 ng/μl pCFJ151 [contains *Cbr-unc-119*(+)], 25 ng/μl pCFJ90 (contains *myo-2::mCherry*), and 50 ng/μl 1 kb ladder (NEB; carrier DNA). F1 animals were selected based on mCherry expression, and strains were maintained by picking mCherry-positive animals each generation.

Microscopy

Animals were examined using differential interference contrast (DIC) and fluorescence microscopy. For Fig. 3, L3 larval animals were selected for presence of the transgene (mCherry-positive) and then assessed for presence and number of GFP-positive anchor cells. As in *Cel-lin-12*(*gf*) mutants (Greenwald *et al.* 1983), the P6.p cell divided in all *Cbr-lin-12.1*(*gf*) animals observed, despite the absence of an anchor cell. For Fig. 4a, early L4 larval animals were selected and evaluated for the number of vulval cells and inferred number of VPCs that had divided to produce them (Sternberg and Horvitz 1986). For Fig. 4b, mid-L4 stage animals homozygous for *mfIs5* (*egl-17::gfp*) were selected, and all GFP-positive cells in the vulva epithelium were counted. In the wild-type strain, this result can vary from 4 to 6 cells, depending on whether expression is only in the vulC cells or the vulC and the vulD cells. The number of expressing cells in *Cbr-lin-12.1*(*gf*) mutants is highly variable and does not obviously correlate with the inferred division pattern that produced the vulval cells.

Dominance tests and genetic crosses

For Fig. 5a, a balanced heterozygous strain was used for each *lin-12*(*gf*) strain [relevant genotypes: *Cel-lin-12*(*gf*)/*Cel-unc-32*(*e189*) and *Cbr-lin-12.1*(*gf*)/*Cbr-unc-119*(*st20000*); full strain names and genotypes in Supplementary File 1 (Supplementary Table 1)]. Non-Unc L4 hermaphrodites were selected individually from each strain, aged overnight, and then scored for any egg-laying defects (Muv, Pvl, Egl). Three to 4 days later, the offspring were scored to infer the genotype of the parent. For Fig. 5b, male animals homozygous for the relevant *Cbr-lin-12.1* or *Cbr-lin-12.2* genotype and bearing an extrachromosomal transgene array marked with *myo-2::mCherry* were crossed with hermaphrodites of the appropriate balanced *Cbr-lin-12.1*(*gf*) genotype. mCherry-positive F1 L4 hermaphrodite cross progeny were selected individually, aged overnight, scored for any egg-laying defects, and allowed to produce self-cross offspring to infer the maternal genotype as above. The figure includes data only from the animals that received the *Cbr-lin-12.1*(*gf*) allele from the heterozygous parent.

Data availability

Strains and plasmids are available upon request. Raw input data (genomes and gene annotations) can be retrieved from WormBase (under WS280 release) and NCBI (under accession PRJNA784955). Intermediate files and code used to produce analyses and figures are available in GitHub at https://github.com/AndersenLab/LIN12_notch_MS.

Supplemental material available at GENETICS online.

Conflicts of interest

The author(s) declare no conflicts of interest.

Literature cited

Birchler JA, Riddle NC, Auger DL, Veitia RA. 2005. Dosage balance in gene regulation: biological implications. Trends Genet. 21(4):219–226. doi:10.1016/j.tig.2005.02.010.

Birchler JA, Yang H. 2022. The multiple fates of gene duplications: deletion, hypofunctionalization, subfunctionalization, neofunctionalization, dosage balance constraints, and neutral variation. Plant Cell. 34(7):2466–2474. doi:10.1093/plcell/koac076.

Blomme T, Vandepoele K, De Bodt S, Simillion C, Maere S, Van de Peer Y. 2006. The gain and loss of genes during 600 million years of vertebrate evolution. Genome Biol. 7(5):R43. doi:10.1186/gb-2006-7-5-r43.

Chen N, Greenwald I. 2004. The lateral signal for LIN-12/Notch in *C. elegans* vulval development comprises redundant secreted and transmembrane DSL proteins. Dev Cell. 6(2):183–192. doi:10.1016/S1534-5807(04)00021-8.

Dainat J, Hereñú D. LucileSol, and pascal-git. 2022. NBISweden/AGAT: AGAT-v0.8.1. https://zenodo.org/record/5834795#.Y2jXNNLMKV4.

Davis P, Zarowiecki M, Arnaboldi V, Becerra A, Cain S, Chan J, Chen WJ, Cho J, da Veiga Beltrame E, Diamantakis S, et al. 2022. WormBase in 2022-data, processes, and tools for analyzing Caenorhabditis elegans. Genetics 220:iyac003. doi:10.1093/genetics/iyac003.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 29(1):15–21. doi:10.1093/bioinformatics/bts635.

Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 20(1):238. doi:10.1186/s13059-019-1832-y.

Félix M-A. 2007. Cryptic quantitative evolution of the vulva intercellular signaling network in Caenorhabditis. Curr Biol. 17(2):103–114. doi:10.1016/j.cub.2006.12.024.

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics. 151(4):1531–1545. doi:10.1093/genetics/151.4.1531.

Greenwald I, Kovall R. 2013. Notch signaling: genetics and structure. In: WormBook, editor. The C. elegans Research Community. p. 1–28. doi:10.1895/wormbook.1.10.2.

Greenwald IS, Sternberg PW, Horvitz HR. 1983. The lin-12 locus specifies cell fates in *Caenorhabditis elegans*. Cell. 34(2):435–444. doi:10.1016/0092-8674(83)90377-X.

Hillier LW, Miller RD, Baird SE, Chinwalla A, Fulton LA, Koboldt DC, Waterston RH. 2007. Comparison of *C. elegans* and *C. briggsae* genome sequences reveals extensive conservation of chromosome organization and synteny. PLoS Biol. 5(7):e167. doi:10.1371/journal.pbio.0050167.

Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M. 2017. WormBase ParaSite – a comprehensive resource for helminth genomics. Mol Biochem Parasitol. 215:2–10. doi:10.1016/j.molbiopara.2016.11.005.

Hwang BJ, Sternberg PW. 2004. A cell-specific enhancer that specifies lin-3 expression in the *C. elegans* anchor cell for vulval development. Development. 131(1):143–151. doi:10.1242/dev.00924.

Innan H, Kondrashov F. 2010. The evolution of gene duplications: classifying and distinguishing between models. Nat Rev Genet. 11(2):97–108. doi:10.1038/nrg2689.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780. doi:10.1093/molbev/mst010.

Konrad A, Flibotte S, Taylor J, Waterston RH, Moerman DG, Bergthorsson U, Katju V. 2018. Mutational and transcriptional landscape of spontaneous gene duplications and deletions in *Caenorhabditis elegans*. Proc Natl Acad Sci U S A. 115(28):7386–7391. doi:10.1073/pnas.1801930115.

Lambie EJ, Kimble J. 1991. Two homologous regulatory genes, lin-12 and glp-1, have overlapping functions. Development. 112(1):231–240. doi:10.1242/dev.112.1.231.

Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. 2018. MUMmer4: a fast and versatile genome alignment system. PLoS Comput Biol. 14(1):e1005944. doi:10.1371/journal.pcbi.1005944.

Mello CC, Kramer JM, Stinchcomb D, Ambros V. 1991. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. EMBO J. 10(12):3959–3970. doi:10.1002/j.1460-2075.1991.tb04966.x.

Moya ND, Stevens L, Miller IR, Sokol CE, Galindo JL, Bardas AD, KohESH, Rozenich J, Yeo C, Xu M, *et al.* 2023. Novel and improved *Caenorhabditis briggsae* gene models generated by community curation. BMC Genomics. 24(1):486. doi:10.1186/s12864-023-09582-0.

Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):268–274. doi:10.1093/molbev/msu300.

Paix A, Folkmann A, Seydoux G. 2017. Precision genome editing using CRISPR-Cas9 and linear repair templates in *C. elegans*. Methods. 121-122:86–93. doi:10.1016/j.ymeth.2017.03.023.

Pertea G, Pertea M. 2020. GFF utilities: GffRead and GffCompare. F1000Res. 9:304. doi:10.12688/f1000research.23297.2.

Rudel D, Kimble J. 2002. Evolution of discrete Notch-like receptors from a distant gene duplication in *Caenorhabditis*. Evol Dev. 4(5):319–333. doi:10.1046/j.1525-142X.2002.02027.x.

Stein LD, Bao Z, Blasiar D, Blumenthal T, Brent MR, Chen N, Chinwalla A, Clarke L, Clee C, Coghlan A, *et al.* 2003. The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. PLoS Biol. 1(2):E45. doi:10.1371/journal.pbio.0000045.

Sternberg PW, Horvitz HR. 1986. Pattern formation during vulval development in *C. elegans*. Cell. 44(5):761–772. doi:10.1016/0092-8674(86)90842-1.

Sternberg PW, Horvitz HR. 1989. The combined action of two intercellular signaling pathways specifies three cell fates during vulval induction in *C. elegans*. Cell. 58(4):679–693. doi:10.1016/0092-8674(89)90103-7.

Stevens L, Félix M-A, Beltran T, Braendle C, Caurcel C, Fausett S, Fitch D, Frézal L, Gosse C, Kaur T, *et al.* 2019. Comparative genomics of 10 new *Caenorhabditis* species. Evol Lett. 3(2):217–236. doi:10.1002/evl3.110.

Stevens L, Moya ND, Tanny RE, Gibson SB, Tracey A, Na H, Chitrakar R, Dekker J, Walhout AJM, Baugh LR, *et al.* 2022. Chromosome-level reference genomes for two strains of *Caenorhabditis briggsae*: an improved platform for comparative genomics. Genome Biol. Evol. 14(4):evac042. doi:10.1093/gbe/evac042.

Stiernagle T. 2006. Maintenance of *C. elegans*. In: WormBook, editor. The C. elegans Research Community. p. 1–11. doi:10.1895/wormbook.1.101.1.

Thorvaldsdóttir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 14(2):178–192. doi:10.1093/bib/bbs017.

Wickham H, Navarro D, Pedersen TL. 2016. ggplot2: Elegant Graphics for Data Analysis (3e). https://ggplot2-book.org/.

Wilkinson HA, Fitzgerald K, Greenwald I. 1994. Reciprocal changes in expression of the receptor lin-12 and its ligand lag-2 prior to commitment in a *C. elegans* cell fate decision. Cell. 79(7):1187–1198. doi:10.1016/0092-8674(94)90010-8.

Xie D, Ye P, Ma Y, Li Y, Liu X, Sarkies P, Zhao Z. 2022. Genetic exchange with an outcrossing sister species causes severe genome-wide dysregulation in a selfing *Caenorhabditis* nematode. Genome Res. 32(11-12):2015–2027. doi:10.1101/gr.277205.122.

Yochem J, Greenwald I. 1989. glp-1 and lin-12, genes implicated in distinct cell–cell interactions in *C. elegans*, encode similar transmembrane proteins. Cell. 58(3):553–563. doi:10.1016/0092-8674(89)90436-4.

Zhang D, Leng L, Chen C, Huang J, Zhang Y, Yuan H, Ma C, Chen H, Zhang YE. 2022. Dosage sensitivity and exon shuffling shape the landscape of polymorphic duplicates in *Drosophila* and humans. Nat Ecol Evol. 6(3):273–287. doi:10.1038/s41559-021-01614-w.

Zhao Z, Boyle TJ, Bao Z, Murray JI, Mericle B, Waterston RH. 2008. Comparative analysis of embryonic cell lineage between *Caenorhabditis briggsae* and *Caenorhabditis elegans*. Dev Biol. 314(1):93–99. doi:10.1016/j.ydbio.2007.11.015.

Editor: B. Conradt