Identification of a spontaneously arising variant affecting thermotaxis behavior in a recombinant inbred *Caenorhabditis elegans* line

Jihye Yeon,[1,*] Charmi Porwal,[1] Patrick T. McGrath,[2] Piali Sengupta (ID) [1,*]

[1]Department of Biology, Brandeis University, Waltham, MA 02454, USA
[2]School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA

*Corresponding author: Department of Biology, Brandeis University, Waltham, MA 02454, USA. Email: sengupta@brandeis.edu; *Corresponding author: Department of Biology, Brandeis University, Waltham, MA 02454, USA. Email: jihyeyeon@brandeis.edu

Analyses of the contributions of genetic variants in wild strains to phenotypic differences have led to a more complete description of the pathways underlying cellular functions. Causal loci are typically identified via interbreeding of strains with distinct phenotypes in order to establish recombinant inbred lines (RILs). Since the generation of RILs requires growth for multiple generations, their genomes may contain not only different combinations of parental alleles but also genetic changes that arose *de novo* during the establishment of these lines. Here, we report that in the course of generating RILs between *Caenorhabditis elegans* strains that exhibit distinct thermotaxis behavioral phenotypes, we identified spontaneously arising variants in the *ttx-1* locus. *ttx-1* encodes the terminal selector factor for the AFD thermosensory neurons, and loss-of-function mutations in *ttx-1* abolish thermotaxis behaviors. The identified genetic changes in *ttx-1* in the RIL are predicted to decrease *ttx-1* function in part via specifically affecting a subset of AFD-expressed *ttx-1* isoforms. Introduction of the relevant missense mutation in the laboratory *C. elegans* strain via gene editing recapitulates the thermotaxis behavioral defects of the RIL. Our results suggest that spontaneously occurring genomic changes in RILs may complicate identification of loci contributing to phenotypic variation, but that these mutations may nevertheless lead to the identification of important causal molecules and mechanisms.

Keywords: *Caenorhabditis elegans*; recombinant inbred line; thermotaxis; *ttx-1*

Introduction

Animals must adapt their behaviors in response to environmental changes for optimal survival. Forward and reverse genetic approaches in lab-derived or domesticated strains have been extremely powerful in identifying genetic loci associated with behaviors, but the analysis of only laboratory reference strains is insufficient to identify genetic variants driving adaptive evolution in diverse natural environments (Sterken et al. 2015; Niepoth and Bendesky 2020; Jourjine and Hoekstra 2021). Characterization of phenotypic traits in wild strains and causal association of these traits with underlying genetic variation have begun to provide a deeper and richer understanding of the genetic architecture underlying nervous system function and behavioral adaptation (de Bono and Bargmann 1998; Félix and Duveau 2012; Baldwin et al. 2014; Bendesky et al. 2017; Seeholzer et al. 2018; Andersen and Rockman 2022). This approach has been widely used in multiple experimental systems to identify a range of molecules and mechanisms that represent substrates for evolutionary change to modulate neuron and circuit properties.

A simple analysis to identify loci that quantitatively contribute to trait variation involves the interbreeding of 2 strains exhibiting distinct phenotypes to generate a panel of recombinant inbred lines (RILs) (Rockman and Kruglyak 2008; Pollard 2012). The generation of RILs homozygous for different combinations of parental alleles requires growth for multiple generations. Even under controlled laboratory conditions, prolonged cultivation increases the probability of spontaneously arising genetic changes in these RILs (Cook et al. 2006; Zhao et al. 2020). Consequently, the genomic architecture of RILs may not only represent a mosaic of parental alleles but may also include new variants that arose during the process of RIL generation. Although these novel variants complicate analyses of the contributions of parental genotypes to the trait under study, these changes have nevertheless allowed for the identification of relevant genetic changes that also influence phenotypic differences (Cook et al. 2006; Zhao et al. 2020).

Caenorhabditis elegans provides a particularly robust experimental system in which to study how genetic changes contribute to natural phenotypic variation (Félix and Braendle 2010; Frezal and Felix 2015). Wild inbred *C. elegans* strains have been isolated from different regions around the world, and their genomes have been sequenced (Andersen et al. 2012; Cook et al. 2016; Cook et al. 2017). Analyses of quantitative trait loci (QTLs) contributing to behavioral variation between the N2 laboratory reference and wild strains have led to a detailed understanding of the genetic basis of natural variation and have allowed for the

identification and characterization of new pathways underlying phenotypic and developmental adaptations (Evans et al. 2021).

Caenorhabditis elegans is highly sensitive to temperature changes in its environment and exhibits complex and experience-dependent navigation behaviors on spatial thermal gradients (Hedgecock and Russell 1975; Ryu and Samuel 2002). Thermotaxis behavior is driven primarily, although not exclusively, by the single AFD thermosensory neuron pair which responds with high sensitivity to temperature changes in an experience-dependent manner (Mori and Ohshima 1995; Kimura et al. 2004; Clark et al. 2006). A subset of wild *C. elegans* strains has been reported to exhibit differences in their temperature preferences on thermal gradients, but these preferences remain experience-dependent (Anderson et al. 2007; Jurado et al. 2010; Anderson et al. 2011; Gaertner et al. 2012). The analysis of the genetic variants contributing to behavioral differences in thermal preference between 2 *C. elegans* strains suggested complex epistatic interactions among multiple loci whose molecular identities have not been identified (Gaertner et al. 2012).

Here, we examine thermotaxis in a panel of *C. elegans* strains and identify a strain that shows enhanced thermotaxis behavior. In the course of generating RILs to identify the genetic changes leading to this behavioral difference, we identified an RIL that is atactic and also exhibits defects in locomotory behavior. The thermotaxis behavioral defect in this RIL arises due to variants in at least 2 independently segregating loci. We identify 1 of these loci as the *ttx-1* Otx homeodomain gene that acts as a terminal selector factor for the AFD thermosensory neurons (Satterlee et al. 2001). We show that these variants reduce *ttx-1* function and result in both locomotory and thermotaxis defects and that these changes are not present in either parental strain. Our results suggest that spontaneously arising genetic changes in the course of RIL generation can identify loci critical for driving variation in complex phenotypic traits.

Materials and methods
Strains
All nematode strains were grown on *E. coli* OP50 and cultivated at 20°C. Complete lists of all analyzed strains and strains generated in this work are provided in Supplementary Tables 1 and 3, respectively. Transgenic strains containing *gcy-8p::gfp* in extrachromosomal arrays were generated using PSAA685 (*gcy-8p::gfp*) at 5 ng/μl with the *unc-122::dsRed* coinjection marker at 50 ng/μl. All behavioral assays were performed using adult animals that were grown with ample bacterial food for at least 3 generations.

Generation of RILs
RILs were generated from a cross between CC1 and N2 males carrying the *kyIs37*(*str-2p::gfp*) marker. Thirty F2 hermaphrodites from a single heterozygous F1 from this cross were individually placed on plates and allowed to self-fertilize for 8 generations (see Supplementary Fig. 1a). A single hermaphrodite was transferred in each generation to obtain 30 RILs. A similar approach was used to generate RILs from PY12237. Ten F2 progeny from a cross between PY12237 and N2 were placed on individual plates and allowed to self-fertilize for 8 generations.

Whole-genome sequencing and analysis
Animals were grown on 10-cm High Growth Medium plate (HGMA) until the bacterial food was nearly depleted. Genomic DNA was isolated using the Qiagen DNeasy Blood & Tissue kit (catalog number: 69504). Preparation of DNA libraries and whole-genome sequencing was performed by BGI (www.bgi.com). Each genome was sequenced at ~30× coverage.

Sequence analysis
The reads were aligned to the WS260 reference genome using Burrows–Wheeler Aligner (BWA v0.7.17 (Li and Durbin 2009). BAM files were deduplicated and processed using SAMtools v1.9 (Li et al. 2009) and Picard (http://broadinstitute.github.io/picard/). Single-nucleotide variations (SNVs were called by Freebayes and annotated by SnpEff (Cingolani et al. 2012; Garrison and Marth 2012). Variant analyses sequence files can be found at Figshare (https://doi.org/10.6084/m9.figshare.23565756). Custom Python scripts (https://github.com/ptmcgrat/Sequencer/blob/master/NGS_Nematode.py) using the pysam library (https://github.com/pysam-developers/pysam) were used to identify regions of the genome with a large number of clipped and chimeric reads. Read depths were visualized using Integrative Genomics Viewer (IGV (Robinson et al. 2011).

CRISPR/Cas9-mediated genome editing
To generate *ttx-1*(*oy184*), a donor oligonucleotide (IDT: Integrated DNA Technologies) containing the ATG to ACG change along with 25 bp homology arms and a mutated PAM site (TCC to TGC) was injected along with sgRNA (100 ng/μl; IDT) and Cas9 (25 ng/μl; IDT) into N2. *dpy-10* crRNA (20 ng/μl; IDT) was used as the co-CRISPR injection marker (Arribere et al. 2014). F1 animals with Rol and Dpy phenotypes were isolated, and genome editing was confirmed via sequencing. F2 progeny were further screened for homozygosity.

To generate Y102A11.1(*oy185*), a donor oligonucleotide (IDT) containing the CAA (Gln) to TAA (STOP) change along with 35 bp homology arms was injected along with crRNA (100 ng/μl; IDT), tracrRNA (100 ng/μl; IDT) and Cas9 (25 ng/μl; IDT), and *unc-122p::dsRed* (50 ng/μl) as the coinjection markers into N2. F1 animals with *unc-122p::dsRed* fluorescence were isolated and confirmed via sequencing. F2 progeny were further screened for homozygosity.

Sequences of sgRNA and donor oligonucleotides:

sgRNA [*ttx-1*(*oy184*)]: 5′- CCT//TGACGTCTTCCTCAGCTCCA -3′

Donor oligonucleotide [*ttx-1*(*oy184*)]: 5′- TCATTTGTC ATATATTCGTGCTTCACGTCCTTGACGTCTTCCTCAGCTCCAT -3′

crRNA [Y102A11.1(*oy185*)]: 5′- CCT//CCTCAGTTGTTCAAT CAACA -3′

Donor oligonucleotide: [Y102A11.1(*oy185*)]: 5′- CCCTCCG GCTGCTCAGCCAGCTCCACAATAAATTATGATACCTCCTCAGTT-GTTCAATCAACAAATGCCTCAATTACAAAACCCACAGCAACAC-AATC -3′.

Thermotaxis behavior
Thermotaxis behavioral assays were performed as described previously (Takeishi et al. 2020). In brief, a thermal gradient (ranging from 23 to 28°C at 0.5°C/cm) was established on 10-cm NGM agar plates using Peltier thermoelectric temperature controllers (Oven Industries). Fifteen animals grown at 20°C were placed at the center of the assay plate at the start of the assay. The temperature of the edges of NGM agar was measured at the beginning of the assay with a 2-probe digital thermometer (Fluke Electronics). Animal behavior was recorded at a rate of 1 Hz for 35 min using PixeLink CCD cameras controlled by custom written scripts in MATLAB. Videos were analyzed using WormLab (MBF Bioscience) and custom written scripts in MATLAB as described previously (Beverly et al. 2011; Yu et al. 2014).

Exploration assay

The grid exploration assay was performed essentially as described previously (Flavell et al. 2013). The day prior to the assay, a well-fed young adult animal was picked onto a 35-mm agar plate uniformly seeded with 100 μl of *Escherichia coli* OP50. After 16 h, the animal was removed, and the plate was superimposed on a grid containing 3.5-mm squares. The number of squares entered by the worm as assessed by the presence of tracks in the bacterial food was manually counted. Twenty animals were tested per data point for each genotype.

Chemotaxis behavior

Chemotaxis behavioral assays were performed as described previously (Bargmann et al. 1993). Well-fed young adult animals were washed twice with S-Basal twice and once with water and placed at the center of a 10-cm NGM agar plate. One μl of the diluted odorant in ethanol or ethanol alone were placed on either end of the plate along with 1 μl of 1 M sodium azide. Chemicals used were isoamyl alcohol (Fisher A393-500) and diacetyl (Sigma B85307). Animals at the odorant, diluent, and throughout the plate were counted after 1 h.

Fluorescence reporter quantification

Well-fed young adults grown at 20°C were paralyzed with 10 mM levamisole and mounted on 3% agarose pads on a glass coverslip. Animals were examined using a ×63 oil objective (NA 1.4) on a Zeiss Axio Imager M2 epifluorescence microscope. z-stack images of individual animals were acquired using Zeiss Pro software (resolution, 1280 × 1024 pixels; exposure time, 120 ms; step size, 0.25 μm). Images were processed in ImageJ (NIH). An ROI was drawn around a single AFD soma per animal, and the fluorescence intensities in 5 optical slices per animal were quantified and averaged. To account for expression variation from extrachromosomal arrays carrying *gcy-8p::gfp*, 2 independent transgenic lines were examined for each strain; expression levels were similar between lines.

Statistical analyses

Statistical analyses were performed using Prism v9.5.1. Comparisons to N2 were performed with *t*-test or 1-way ANOVA. Corrections for multiple comparisons were performed using Tukey's or Bonferroni tests. The statistical test used is indicated in the legend to each figure.

Results

The CC1 *C. elegans* strain exhibits enhanced negative thermotaxis behavior

To assess whether wild *C. elegans* strains collected from multiple locations across the world or different laboratory N2-derived strains exhibit variations in thermotaxis behavior, we grew a panel of 30 strains (Supplementary Table 1) (Cook et al. 2017; Lee et al. 2021) at 20°C and examined their behavior on a spatial thermal gradient ranging from 23 to 28°C (Fig. 1a. Since the laboratory N2 strain contains a gain-of-function allele of the *npr-1* neuropeptide receptor as compared with the allele present in wild *C. elegans* strains (de Bono and Bargmann 1998), we also examined the behavior of *npr-1(ad609)* loss-of-function mutants. Under these conditions, as expected, the N2 laboratory reference strain moved toward the colder side of the gradient (negative thermotaxis; Fig. 1b). The negative thermotaxis behavior of the majority of examined strains was not significantly different from that of N2,

with the exception of a subset of strains that exhibited decreased negative thermotaxis (Fig. 1b). In contrast, the CC1 strain exhibited significantly enhanced negative thermotaxis (Fig. 1b–c). The velocity of CC1 animals was similar to that of N2 on an isothermal plate (N2, 136.9 ± 3.2 μm/s; CC1, 129.5 ± 7.2 μm/s; 15 animals each at 20°C). CC1 is an N2-derived strain that was frozen following 4 years of growth in liquid axenic media (Fig. 1d) (Szewczyk et al. 2003; McGrath et al. 2011). An additional strain (LSJ1) that was grown in axenic media for >30 years before being frozen down (Fig. 1d) (McGrath et al. 2009; McGrath et al. 2011; Sterken et al. 2015) did not exhibit a similar behavioral phenotype (Fig. 1b), suggesting that the genetic changes in CC1 leading to enhanced negative thermotaxis are unlikely to have arisen due to selective pressure upon growth in liquid acting directly on this trait.

Enhanced negative thermotaxis behavior in CC1 is unlikely to have arisen due to a loss-of-function mutation at a single locus

To identify causal changes underlying the improved negative thermotaxis behavior of CC1, we compared the genome sequences of the reference N2 and CC1 strains (McGrath et al. 2011) (wormbase.org), focusing primarily on variants in protein-coding sequences. This analysis led to the identification of missense variants in 8 protein-coding genes in CC1 compared with N2 (Supplementary Table 2). A subset of these changes is predicted to alter residues that may disrupt protein function (Fig. 2a; Supplementary Table 2) in a range of molecules including a predicted receptor guanylyl cyclase and a metabotropic glutamate receptor (Fig. 2a; Supplementary Table 2). Since the genomes of the strains used in our experiments likely differ from those of the reference strains, we verified the presence of these changes in the N2 and CC1 strains used in our behavioral assays. We confirmed the presence of 6 of these changes in our CC1 but not in our N2 strain, whereas 1 mutation was present in both strains (Supplementary Table 2).

The neuronal circuit driving negative thermotaxis includes the AFD, AWC, and ASI neurons, as well as the AIY and AIZ first layer interneurons (Mori and Ohshima 1995; Biron et al. 2008; Kuhara et al. 2008; Beverly et al. 2011; Hawk et al. 2018; Yeon et al. 2021). Cross-referencing with the CeNGEN neuron gene expression database (Taylor et al. 2021) indicated that *gcy-12*, *ufd-3*, *mgl-3*, and *scd-2* may be expressed in a subset of these neurons. However, since the gene expression profile of AFD is regulated by the temperature experience of the animal (Yu et al. 2014; Kobayashi et al. 2016; Harris et al. 2023), additional AFD-expressed genes may not have been identified in the CeNGEN analysis. We found that the negative thermotaxis phenotype of animals carrying predicted null mutations in each of the 6 identified genes in the N2 background did not phenocopy that of CC1 (Fig. 2b), suggesting that enhanced negative thermotaxis in CC1 could arise due to the specific missense mutation at an individual locus, variants in multiple genes, and/or other sequence changes in the CC1 genetic background.

Identification of an RIL between N2 and CC1 with decreased negative thermotaxis behavior and locomotion defects

To determine whether the variant(s) in CC1 driving improved negative thermotaxis can be genetically segregated, we generated a panel of 30 RILs from a cross between N2 and CC1 (Supplementary Fig. 1a) without performing behavioral selection. We subsequently examined their thermotaxis behavior. Although no RIL exhibited increased negative thermotaxis similar to CC1,
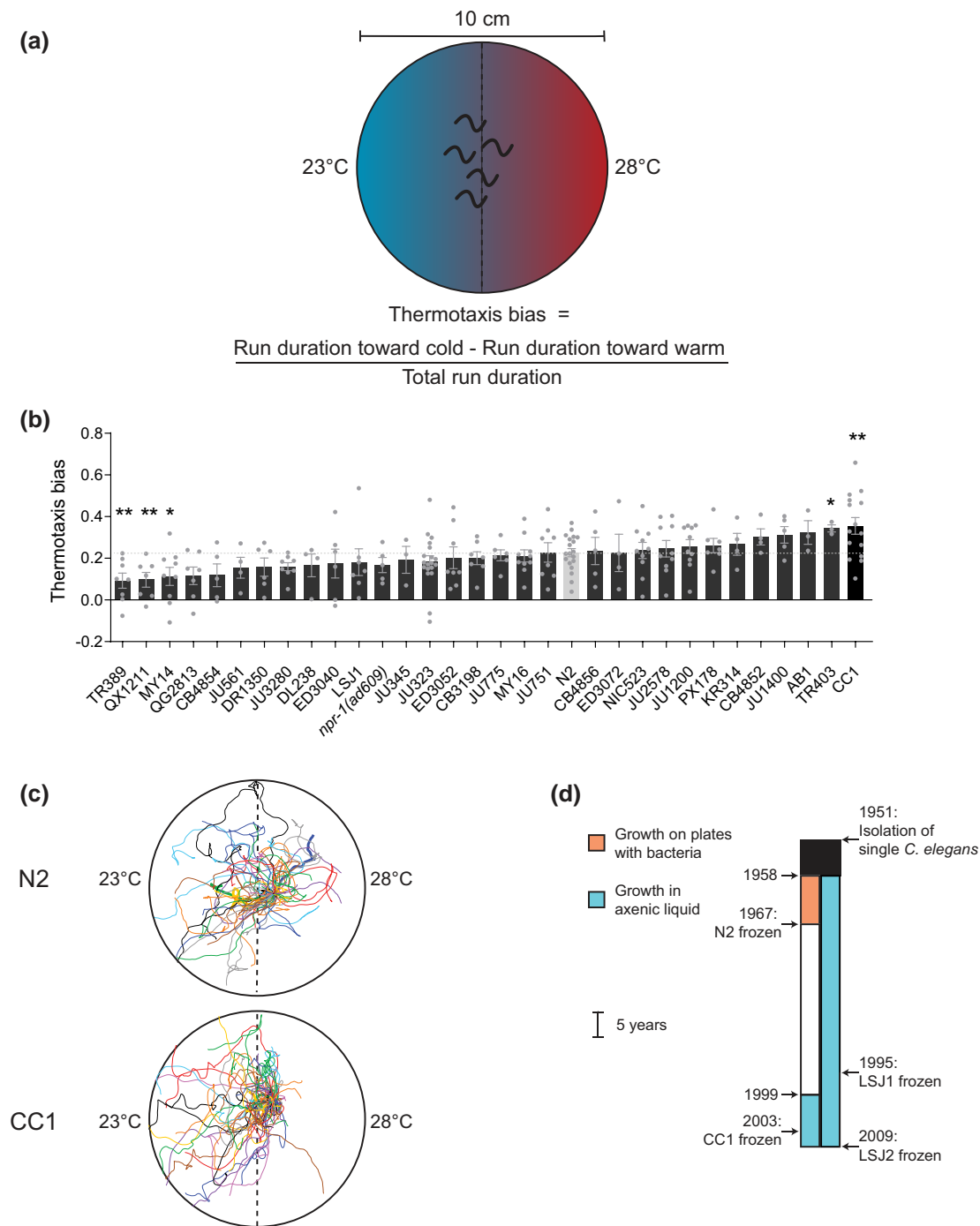
(a)

10 cm

23°C   28°C

Thermotaxis bias =

$$\frac{\text{Run duration toward cold - Run duration toward warm}}{\text{Total run duration}}$$

(b)

(c)

N2   23°C   28°C

CC1   23°C   28°C

(d)

Growth on plates with bacteria

Growth in axenic liquid

5 years

1951: Isolation of single *C. elegans*

1958 →
1967: N2 frozen
1999 →
2003: CC1 frozen

1995: LSJ1 frozen
2009: LSJ2 frozen

Fig. 1. CC1 exhibits improved negative thermotaxis behavior. a) Cartoon of the thermotaxis behavioral assay. Animals grown at 20°C are placed at the center (dotted line) at the start of the assay (see Materials and Methods for additional details). b) Mean thermotaxis bias of shown strains. Each dot is the thermotaxis bias of a single assay with 15 animals. Data shown are from at least 3 independent experiments. Animals were grown at 20°C and assayed on a 0.05°C/cm gradient from 23 to 28°C (a). Errors are SEM. * and ** indicate different from N2 at $P < 0.05$ and 0.001, respectively (t-test). c) Representative trajectories of N2 and CC1 animals in a single thermotaxis assay. Trajectories of individual animals are color coded. d) Isolation and growth history of N2, CC1, LSJ1, and LSJ2. Modified from McGrath *et al.* (2011).

we identified a strain (PY12237; Supplementary Table 3) that exhibited strong defects in negative thermotaxis (Fig. 3a).

Closer examination of the locomotory behavior of PY12237 showed that these animals circled extensively on the thermal gradient and did not exhibit directed movement toward the colder side of the gradient (Fig. 3b). This strain exhibited a similar locomotory phenotype on an isothermal plate (Fig. 3b) and also failed to move toward attractive odorants (Supplementary Fig. 1b). However,

PY12237 animals explored robustly on bacteria-containing plates over a prolonged time period (Fig. 3c) (Flavell et al. 2013), indicating that the circling phenotype does not preclude exploratory behavior. Since neither N2 nor CC1 exhibited similar phenotypes (Fig. 1c; Supplementary Fig. 1b), we conclude that PY12237 contains spontaneously arising variant(s) driving locomotory and/or thermotaxis behavioral defects. Alternatively, these phenotypes are masked in the CC1 genetic background.
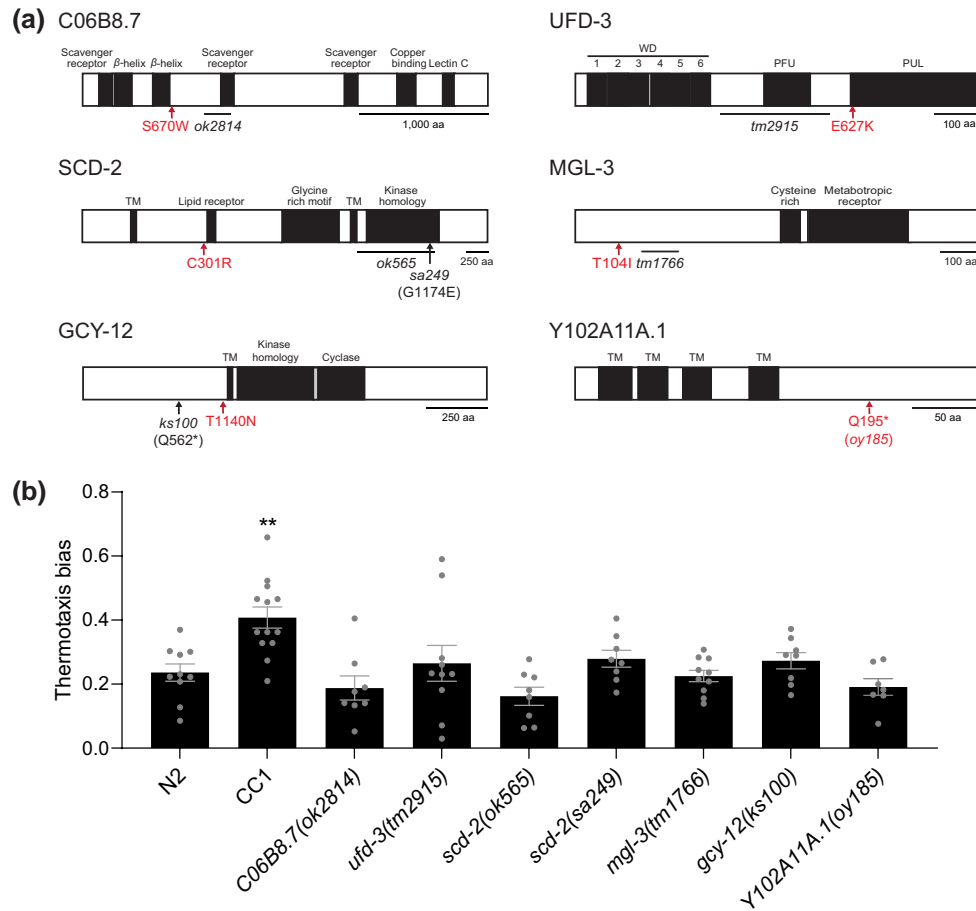
Fig. 2. Enhanced negative thermotaxis in CC1 may not be due to loss of function of a single gene. a) Molecular identities and locations of missense variants (red in protein-coding genes present in CC1 but not in N2. Predicted structural domains in each of these proteins are indicated. The extent of the deletions and the location of nonsense mutations in putative null alleles examined for effects on thermotaxis behavior in (b) are shown. The Q195* change in Y102A11.1 present in CC1 was engineered in the N2 background to generate *oy185*. TM, transmembrane domain. b) Mean thermotaxis bias of N2 animals carrying the indicated mutations. Each dot is the thermotaxis bias of a single assay with 15 animals. Data shown are from at least 3 independent experiments. Animals were grown at 20°C and assayed on a 0.05°C/cm gradient from 23 to 28°C. Errors are SEM. ** indicates different at $P < 0.01$ compared with N2 (t-test).

Independently segregating variants contribute to the altered negative thermotaxis and locomotory behaviors of PY12237

We considered multiple hypotheses to account for the negative thermotaxis phenotype of PY12237. The defects in thermotaxis behavior could simply be a secondary consequence of the locomotory defects of this strain. Alternatively, the locomotory and thermotaxis phenotypes of PY12237 could arise due to either a single or multiple independent genomic changes. To distinguish among these possibilities, we crossed PY12237 to N2 and segregated 10 independent lines following self-fertilization for 8 generations. Two of these lines (PY12270 and PY12272) exhibited defects in negative thermotaxis similar to PY12237 (Fig. 4a). The examination of their locomotion showed that while PY12270 circled similarly to PY12237, PY12272 did not (Fig. 4b), indicating that the locomotory and thermotaxis defects are at least partly genetically separable. Consistently, PY12270 but not PY12272 also showed chemotaxis behavioral defects (Supplementary Fig. 2). PY12265, which performed robust negative thermotaxis (Fig. 4a), also did not exhibit circling on the assay plate (Fig. 4b). We infer that PY12237 contains at least 1 variant that results in locomotory defects with possibly secondary or parallel consequences on directed navigation behavior and an

independently segregating variant that also affects negative thermotaxis.

Spontaneously arising genetic variants in *ttx-1* contribute to altered locomotion and thermotaxis behavior in the CC1-derived PY12237 strain

To identify the causal variants regulating the locomotory and/or thermotaxis behavioral phenotypes of PY12237, we subjected PY12237, PY12270, and PY12272 to whole-genome sequencing. We also sequenced the genome of the PY12265 line that exhibits wild-type locomotion and negative thermotaxis (Fig. 4a–b). Since it is possible that these variants are also present in CC1 but are masked by epistasis, we additionally sequenced the genomes of our CC1 and N2 strains. We reasoned that variants causal to the circling behavior (with or without associated effects on thermotaxis) would be present in PY12237 and PY12270 but not in PY12265 or PY12272, whereas variants responsible for decreased negative thermotaxis alone would be present in PY12237 and PY12272 but not in PY12265 (Fig. 5a). The latter variants may or may not also be present in PY12270 since the locomotory defect of this strain precludes direct analysis of its thermotaxis behavior.
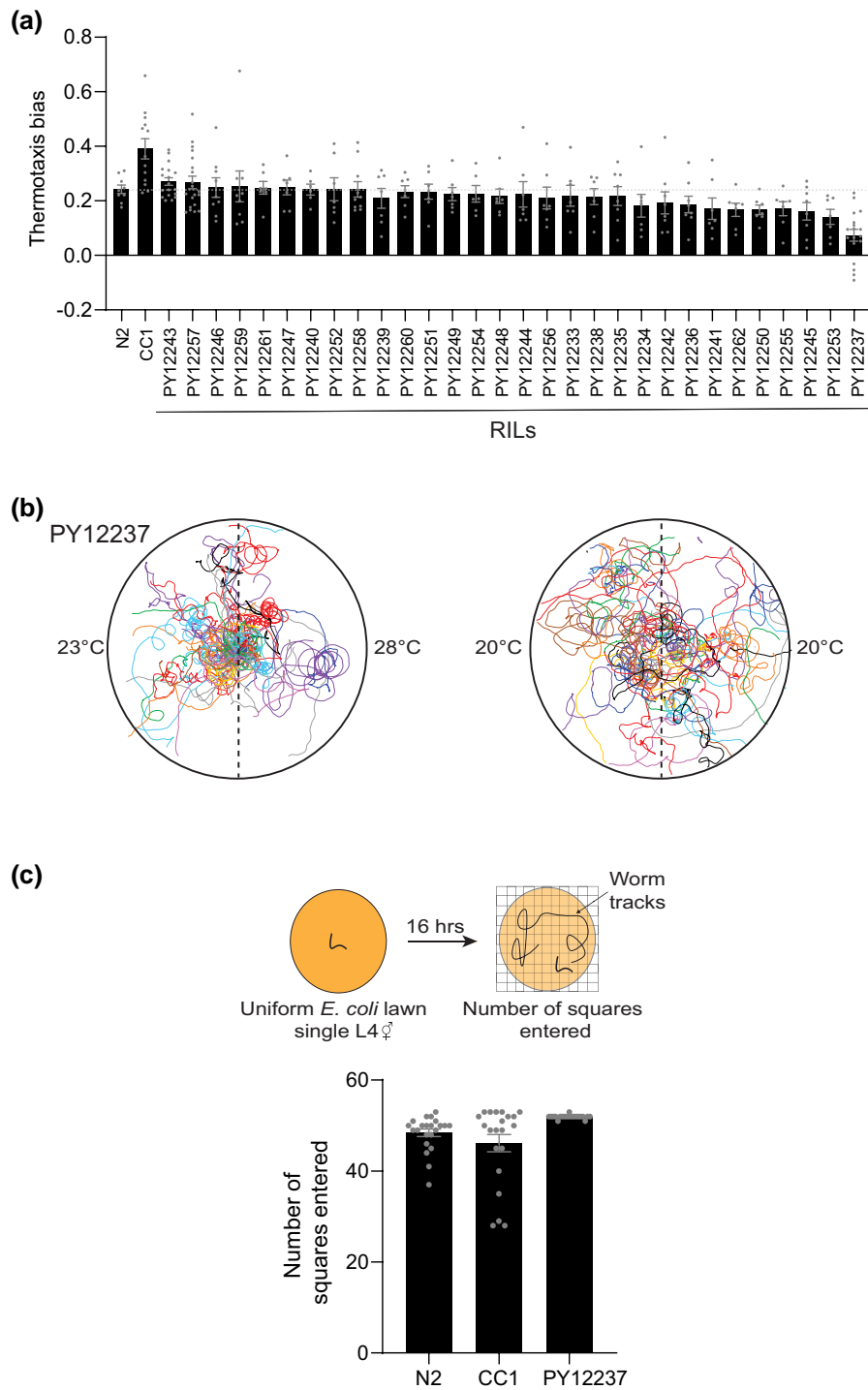
Fig. 3. An RIL generated from interbreeding N2 and CC1 exhibits decreased negative thermotaxis and locomotory defects. a) Mean thermotaxis bias of animals from N2×CC1 RILs. Each dot is the thermotaxis bias of a single assay with 15 animals. Data shown are from at least 3 independent experiments. Animals were grown at 20°C and assayed on a 0.05°C/cm gradient from 23 to 28°C. Errors are SEM. b) Representative trajectories of PY12237 animals on a thermotaxis assay plate (left) and isothermal plate at 20°C (right). Trajectories of individual animals are color coded. c) Number of squares entered (top) by N2, CC1, or PY12237 animals during a grid exploration assay performed in the presence of bacterial food (see Materials and Methods) (Flavell et al. 2013). Each dot represents the movement of a single animal; assays were performed over multiple days. Errors are SEM.

Focusing first on possible variants in protein-coding genes specifically affecting thermotaxis but not locomotory behavior, we identified a missense mutation only in *efn-4* present in both PY12237 and PY12272 but not in PY12270 or PY12265 (Fig. 5b; Table 1). This mutation was not present in additional RIL strains derived from PY12237 that exhibited normal thermotaxis or in CC1 (Table 1). *efn-4* encodes an ephrin homolog that has previously been shown to regulate neurite outgrowth in the AIY interneurons, a major component of the circuit driving thermotaxis (Chin-Sang et al. 1999; Schwieterman et al. 2016). The identified variant is a D47G missense mutation in a highly conserved residue in the β-strand C of EFN-4 predicted to be involved in the
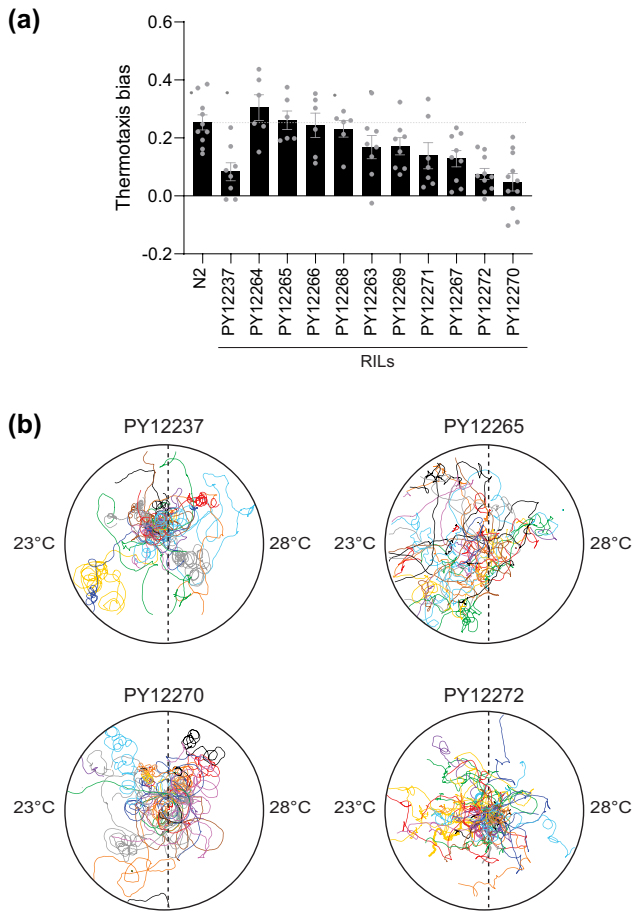
Fig. 4. The thermotaxis behavioral defect of PY12237 is genetically separable from its locomotory defects. a) Mean thermotaxis bias of animals from strains generated via backcrossing PY12237 with N2. Each dot is the thermotaxis bias of a single assay with 15 animals. Data shown are from at least 3 independent experiments. Animals were grown at 20°C and assayed on a 0.05°C/cm gradient from 23 to 28°C. Errors are SEM. b) Representative trajectories of animals from the indicated strains on a thermotaxis assay plate in a single assay. Trajectories of individual animals are color coded.

ligand–receptor tetramerization interface (Supplementary Fig. 3) (Chin-Sang et al. 1999). However, *efn-4(bx80)* null mutants exhibited robust negative thermotaxis (Supplementary Fig. 3), indicating that loss of function of *efn-4* is unlikely to underlie the thermotaxis defects of PY12237 and PY12272. No variants in protein-coding sequences were shared among PY12237, PY12270, and PY12272 (Fig. 5b). Since sequence changes present in PY12237 may also be present in CC1 but masked phenotypically, we also attempted to identify variants shared between CC1 and PY12237, PY12270, and PY12272 but not PY12265. However, no variants in protein-coding sequences were shared among these strains. We hypothesize that the thermotaxis defects of PY12272 may arise from noncoding variants or from changes at multiple loci.

We next identified changes underlying the locomotory and/or thermotaxis defect by identifying protein-coding variants present in PY12237 and PY12270 but not in PY12272 or PY12265 (Fig. 5b). Of the variants that fulfilled these criteria, we noted that 1 was associated with the *ttx-1* locus (Fig. 5b; Table 1). *ttx-1* encodes an OTX homeodomain transcription factor that serves as a terminal selector factor for specification of AFD thermosensory neuron fate (Satterlee et al. 2001). Interestingly, animals carrying loss-of-function

mutations in *ttx-1* have been reported to not only exhibit strong thermotaxis behavioral defects but also to circle during directed navigation on bacteria-free plates (Ryu and Samuel 2002). Consistently, *ttx-1(p767)* exhibited atactic and circling behavior on thermal gradients similar to PY12237 and PY12270 animals (Fig. 6d).

We noted chimeric and clipped reads in a ~1 kb region upstream of the *ttx-1* locus in both PY12237 and PY12270, suggesting a possible complex genomic structural change (Fig. 6a; Supplementary Fig. 4). The *ttx-1* locus is predicted to encode 4 distinct protein isoforms; each of these isoforms contains the conserved homeodomain but differ at their N-termini (Fig. 6a). In PY12237 and PY12270, we also identified a missense mutation converting the start codon of the *ttx-1.a* and *ttx-1.b* isoforms from Met to Thr; the *ttx-1.c* and *ttx-1.d* isoforms that include sequences encoding the homeodomain are predicted to be unaffected by this mutation (Fig. 6a). The *ttx-1*-associated variants were not present in either N2 or in CC1 (Table 1), suggesting that these changes arose spontaneously during the generation of PY12237. Recent bulk RNA-Seq of sorted individual neuron types showed that *ttx-1.b* is the predominantly expressed isoform in AFD (Taylor et al. 2021), and we previously showed that expression of this isoform alone can rescue AFD fate specification phenotypes (Satterlee et al. 2001). However, AFD also expresses *ttx-1.a* and *ttx-1.d* at lower levels (Barrett et al. 2022) (Fig. 6b), suggesting that the identified missense mutation is likely to decrease although not fully abolish *ttx-1* function.

TTX-1 regulates the expression of AFD-expressed terminal differentiation genes including the *gcy-8*-encoded receptor guanylyl cyclase (Satterlee et al. 2001). A *gcy-8p::gfp* reporter was expressed at significantly lower levels in PY12237 and PY12270 as compared with levels in N2 or CC1, although at levels higher than in a *ttx-1(p767)* loss-of-function mutant (Fig. 6c). Similar effects were observed in multiple transgenic strains carrying independent *gcy-8p::gfp*-containing extrachromosomal arrays. In addition, the introduction of only the missense mutation into the N2 background via gene editing [*ttx-1(oy184)*] resulted in thermotaxis and locomotory behavioral defects similar to those exhibited by PY12237 and PY12270 (Fig. 6d). We conclude that the missense mutation in *ttx-1*, possibly together with decreased expression due to the presumptive upstream genomic structural change, is the predominant driver of both the altered locomotion and negative thermotaxis behavior of PY12237 and PY12270 but that 1 or more independently segregating mutations likely contribute to decreased negative thermotaxis in PY12237 and PY12272.

Discussion

Here, we identify spontaneously arising sequence changes that contribute to the thermotaxis and locomotory behavioral phenotypes of an RIL generated from intercrossing the N2 and CC1 *C. elegans* strains. A missense mutation is predicted to affect 2 of 4 predicted isoforms of the *ttx-1* Otx homeobox gene, whereas a structural change lies in the upstream regulatory region of the *ttx-1* locus and may further affect *ttx-1* expression. Additional spontaneous sequence change(s) in PY12237 that segregate independently of the *ttx-1* variants may also contribute to its thermotaxis behavioral phenotype. These observations together with previous reports (Cook et al. 2006; Zhao et al. 2020) indicate that mutations arising *de novo* during the generation of RILs may complicate the analysis of QTLs contributing to phenotypic differences between strains.

Two lines of evidence support the notion that the identified spontaneous sequence changes at the *ttx-1* locus reduce protein
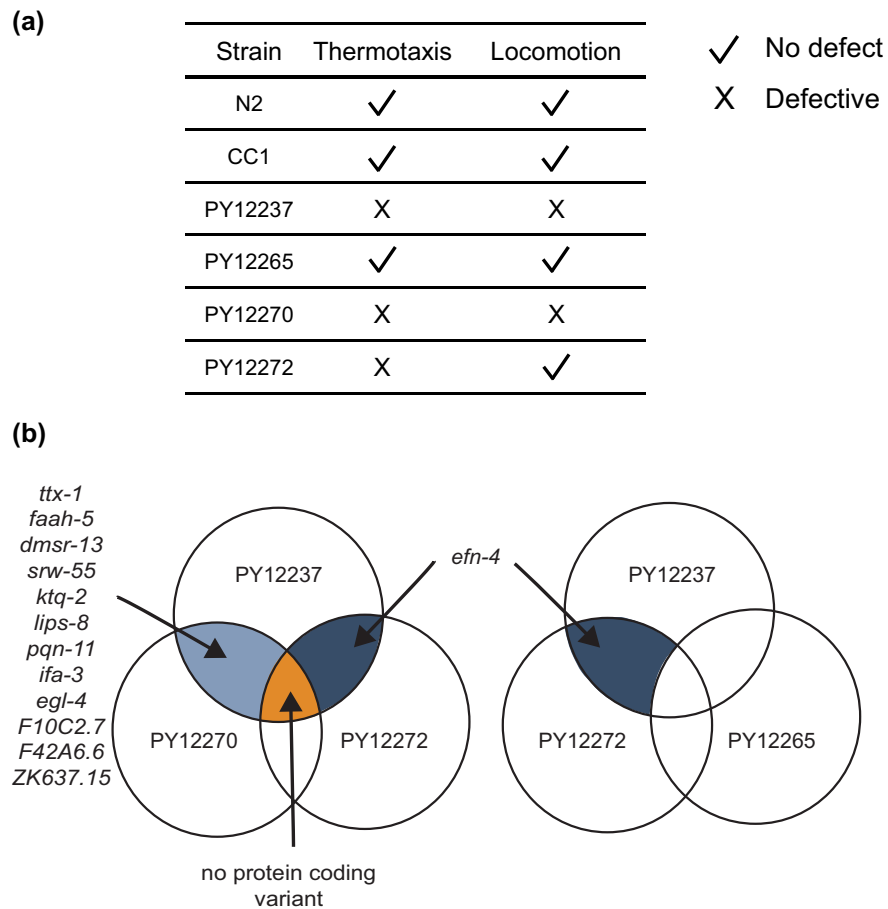
(a)

Strain	Thermotaxis	Locomotion
N2	✓	✓
CC1	✓	✓
PY12237	X	X
PY12265	✓	✓
PY12270	X	X
PY12272	X	✓

✓ No defect
X Defective

(b)

Fig. 5 . Identification of shared protein-coding variants in RILs. a) Summary of thermotaxis and/or locomotory phenotypes of the indicated strains. b) Shared sequence variants present in the genomes of the indicated strains as identified via whole-genome sequencing. Only protein-coding variants are indicated (see Table 1).

Table 1. Summary of presence or absence of protein-coding variants in examined strains.

Gene	Encoded protein	Chr (position)[a]	Nucleotide change[b]	Protein change	Presence in N2[c]	CC1[c]	PY12237	PY12265	PY12270	PY12272
efn-4	Ephrin	IV (565523)	A to G	Missense	No	No	Yes	No	No	Yes
ZK637.15	Uncharacterized protein	III (8921065)	A to C	Missense	No	No	Yes	No	Yes	No
faah-5	Amidase domain–containing protein	III (11887252)	C to T	Missense	No	No	Yes	No	Yes	No
egl-4	cGMP-dependent protein kinase	IV (1875890)	AGACTATGG AGA TAAGGA ACGTTTGGCT CAAGT to TCCCT	Indel	No	No	Yes	No	Yes	No
F42A6.6	DUF775 domain–containing protein	IV (3336613)	C to T	Missense	No	No	Yes	No	Yes	No
dmsr-13	GPCR	V (6819161)	A to C	Missense	No	No	Yes	No	Yes	No
F10C2.7	MFS domain–containing protein	V (12055146)	TC to GT	Missense	No	No	Yes	No	Yes	No
srw-55	GPCR	V (15939563)	A to G	Missense	No	No	Yes	No	Yes	No
ttx-1	OTX homeodomain protein	V (20064174)	T to C	Missense	No	No	Yes	No	Yes	No
kqt-2	Voltage-gated potassium channel	X (8245825)	C to T	Missense	No	No	Yes	No	Yes	No
lips-8	Lipase	X (14899594)	G to A	Missense	No	No	Yes	No	Yes	No
pqn-11	Prion-like domain–containing protein	X (15966176)	A to C	Missense	No	No	Yes	No	Yes	No
ifa-3	Intermediate filament protein	X (16269688)	C to T	Missense	No	No	Yes	No	Yes	No

[a] wormbase.org.
[b] Present in reference N2 and CC1 strains.
[c] Reference N2 and CC1 strains; the absence or presence of the *ttx-1* variant was confirmed in the N2 and CC1 strains, respectively, used in this work.
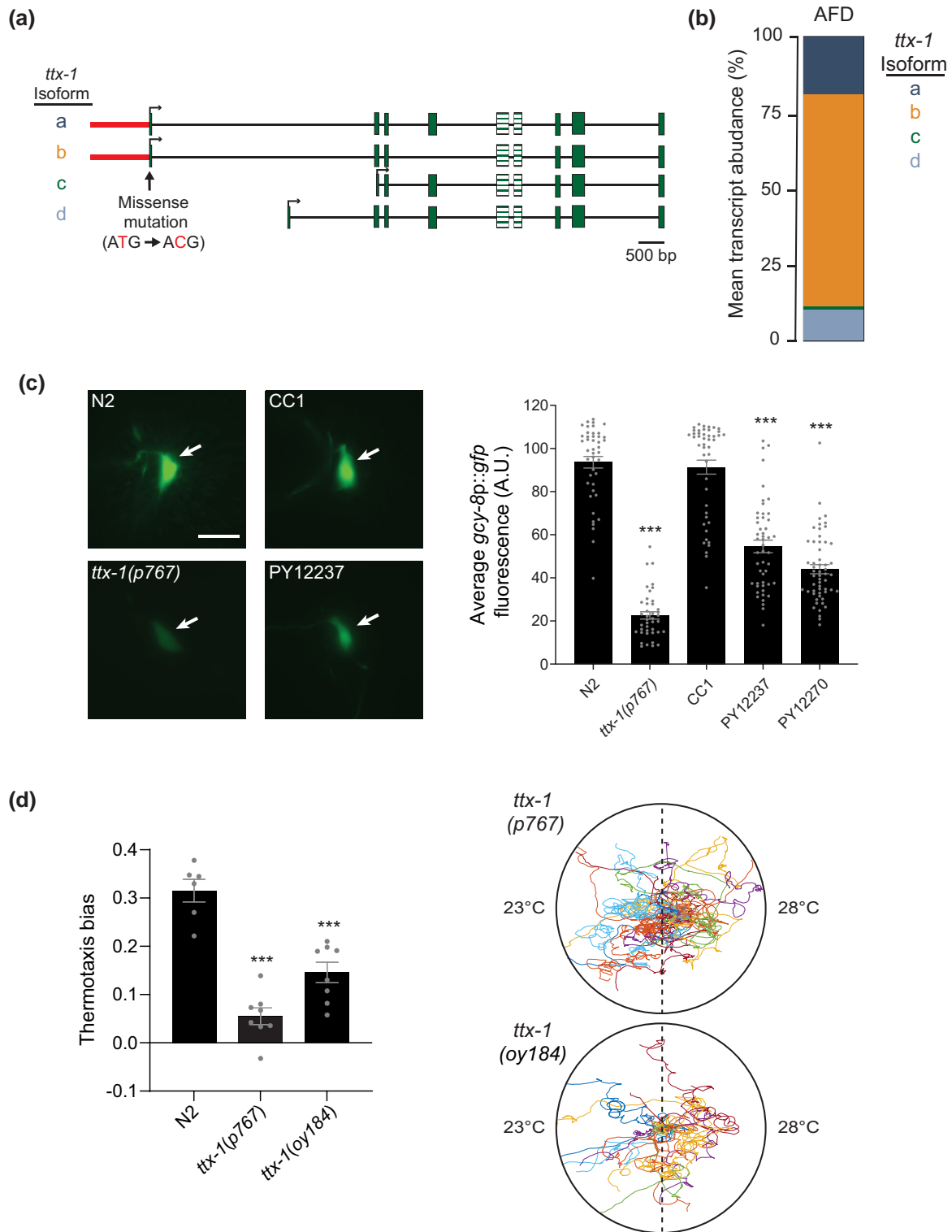
Fig. 6. Sequence variants at the *ttx-1* locus in part underlie the locomotory and thermotaxis phenotypes of PY12237 and derived strains. a) Shown are the genomic structures of 4 alternatively spliced *ttx-1* isoforms. The missense mutation present in PY12237 and PY12270 but not in N2 or CC1 is indicated. Hatched boxes indicate sequences encoding the homeodomain. The presence of upstream chimeric and clipped reads is indicated in red. b) Diagrammatic representation of *ttx-1* transcript abundance in AFD from neuronal RNA-Seq data (Barrett et al. 2022). c) Representative images (left) and quantification of *gcy-8p::gfp* reporter fluorescence (right) in AFD in the indicated strains. Anterior is at the left in all images. AFD soma are indicated by arrows. Scale bar: 10 μm. Each dot is the fluorescence level in a single neuron. Animals from 2 independently generated transgenic lines were examined for each strain. Errors are SEM. *** indicates different at $P < 0.001$ from N2 (1-way ANOVA and Tukey's test). d) Mean thermotaxis bias (left) and representative trajectories (right) of N2 animals carrying indicated *ttx-1* alleles. The missense mutation shown in (a) was introduced into the *ttx-1* locus in the N2 background to generate *ttx-1(oy184)*. Each dot is the thermotaxis bias of a single assay with 15 animals. Data shown are from at least 3 independent experiments. Animals were grown at 20°C and assayed on a 0.05°C/cm gradient from 23 to 28°C. Trajectories of individual animals are color coded. Errors are SEM. *** indicates $P < 0.001$, compared with N2 (1-way ANOVA and Tukey's test).

function. First, the expression level of the AFD-specific and TTX-1-regulated *gcy-8* gene is decreased in RILs carrying the identified sequence changes, although the phenotype is weaker than in a strong *ttx-1* loss-of-function mutant. Second, while we were unable to fully resolve the complex sequence changes upstream of the *ttx-1* coding sequence, the introduction of the missense mutation present in the derived RILs into the *ttx-1* locus in N2 results in thermotaxis and locomotory phenotypes similar to those observed in PY12237 and PY12270. The missense mutation in *ttx-1* is predicted to specifically affect the *ttx-1.a* and *ttx-1.b* isoforms while sparing the remaining 2 isoforms. Further phenotypic analyses of animals expressing specific *ttx-1* transcripts in AFD may identify isoform-specific functions of this conserved protein.

Although large genomic structural rearrangements are often not maintained in populations due to negative fitness consequences (Lupski 2007; Katju and Bergthorsson 2013; Konrad et al. 2018), changes leading to gene duplication followed by functional divergence can be beneficial and maintained under positive selection (Taylor and Raes 2004; Kuzmin et al. 2022). *Caenorhabditis elegans* has been reported to have a high rate of spontaneous structural variants both within the laboratory and in wild strains (Maydan et al. 2010; Farslow et al. 2015; Konrad et al. 2018; Kim et al. 2019). The structural variant in the *ttx-1* upstream sequences may reduce *ttx-1* expression levels by affecting critical transcriptional regulatory sites. The divergence of transcriptional regulatory mechanisms can drive functional specialization of paralogous genes (Ohno 1970; Castillo-Davis et al. 2004; Huminiecki and Wolfe 2004; Li et al. 2005; Loker and Mann 2022). In rodents, Otx2 can functionally substitute for Otx1 in many tissues but the spatio-temporal expression pattern of these genes is distinct (Acampora et al. 1999; Suda et al. 1999) Similarly, while the 3 Otx type factors encoded by the *C. elegans* genome can functionally substitute for each other in examined contexts, these genes are expressed in different sensory neuron types in the adult (Satterlee et al. 2001; Lanjuin et al. 2003). Analyses of spontaneous mutations affecting regulatory sequences may allow for the identification of variants that drive the evolution of functional complexity in the wild.

This work began with the identification of enhanced negative thermotaxis behavior by the CC1 strain. However, we were unable to identify the causal genetic loci for this improved behavior. Thermotaxis behavior in *C. elegans* is influenced by multiple external and internal variables. For instance, navigation behavior toward the preferred temperature is regulated not only by the animal's recent temperature experience but also by the assay conditions including the steepness of the gradient as well as the animal's satiety state (Hedgecock and Russell 1975; Yamada and Ohshima 2003; Kodama et al. 2006; Ramot et al. 2008; Takeishi et al. 2020). While mutations in many individual loci including those required for AFD development or function can disrupt negative thermotaxis, improved behavior may arise from complex interactions among multiple loci (Gaertner et al. 2012). Loss-of-function mutations in individual genes in which we identified missense changes in CC1 did not exhibit altered thermotaxis behavior, raising the possibility that the specific missense mutations in 1 or multiple genes may underlie the observed enhanced negative thermotaxis. It is also possible that noncoding variants may contribute to the phenotypic difference. However, despite the relatively low sequence divergence between N2 and CC1, we did not identify an RIL among 30 examined strains that exhibited improved negative thermotaxis similar to CC1. It remains to be determined whether epistatic interactions among multiple loci affecting both coding and noncoding sequences underlie the thermotaxis behavioral phenotype of CC1.

While the analysis of spontaneously arising mutations influencing a particular trait allows for the identification of important contributing genetic loci, the presence of such changes even under highly regulated environments without obvious selection pressures on the trait under study suggests caution about assumptions regarding the genomic structures of RILs. Although many mutations arising *de novo* are unlikely to significantly affect the trait being studied, others may adversely influence the analysis of QTLs contributing to the phenotypic differences. The analysis of multiple RILs possibly generated from independent founder crosses as well as careful comparison of genomic sequences in RILs may address this concern.

Data availability

Whole-genome sequences for all parent strains and RILs described in this work have been deposited at SRA (project accession number PRJNA987746). Underlying data for all behavioral experiments and variant analyses sequence files have been deposited to Figshare (https://doi.org/10.6084/m9.figshare.23565756). Custom Python scripts used for sequence analysis can be found at https://github.com/ptmcgrat/Sequencer/blob/master/NGS_Nematode.py. All other data necessary for confirming the conclusions of this article are represented fully within the article and its tables and figures.

Supplemental material available at G3 online.

Conflicts of interest statement

The authors declare no conflict of interest.

Literature cited

Acampora D, Avantaggiato V, Tuorto F, Barone P, Perera M, Choo D, Wu D, Corte G, Simeone A. Differential transcriptional control as the major molecular event in generating *otx1-/-* and *otx2-/-* divergent phenotypes. Development 1999;126(7):1417–1426. doi:10.1242/dev.126.7.1417.

Andersen EC, Gerke JP, Shapiro JA, Crissman JR, Ghosh R, Bloom JS, Félix MA, Kruglyak L. Chromosome-scale selective sweeps shape *Caenorhabditis elegans* genomic diversity. Nat Genet. 2012;44(3):285–290. doi:10.1038/ng.1050.

Andersen EC, Rockman MV. Natural genetic variation as a tool for discovery in *Caenorhabditis* nematodes. Genetics 2022;220(1):iyab156. doi:10.1093/genetics/iyab156.

Anderson JL, Albergotti L, Ellebracht B, Huey RB, Phillips PC. Does thermoregulatory behavior maximize reproductive fitness of natural isolates of *Caenorhabditis elegans*? BMC Evol Biol. 2011;11(1):157. doi:10.1186/1471-2148-11-157.

Anderson JL, Albergotti L, Proulx S, Peden C, Huey RB, Phillips PC. Thermal preference of *Caenorhabditis elegans*: a null model and

empirical tests. J Exp Biol. 2007;210(Pt 17):3107–3116. doi:10.1242/jeb.007351.

Arribere JA, Bell RT, Fu BX, Artiles KL, Hartman PS, Fire AZ. Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in *Caenorhabditis elegans*. Genetics 2014;198(3):837–846. doi:10.1534/genetics.114.169730.

Baldwin MW, Toda Y, Nakagita T, O'Connell MJ, Klasing KC, Misaka T, Edwards SV, Liberles SD. Sensory biology. Evolution of sweet taste perception in hummingbirds by transformation of the ancestral umami receptor. Science 2014;345(6199):929–933. doi:10.1126/science.1255097.

Bargmann CI, Hartwieg E, Horvitz HR. Odorant-selective genes and neurons mediate olfaction in *C. elegans*. Cell 1993;74(3):515–527. doi:10.1016/0092-8674(93)80053-H.

Barrett A, Varol E, Weinreb A, Taylor SR, McWhirter R, Cros C, Basaravaju M, Poff A, Tipps JA, Majeed M, *et al.* Integrating bulk and single cell RNA-Seq refines transcriptomic profiles of specific *C. elegans* neurons. bioRxiv. 2022.04.05.487209. doi:10.1101/2022.04.05.487209, preprint: not peer reviewed. 2022.

Bendesky A, Kwon YM, Lassance JM, Lewarch CL, Yao S, Peterson BK, He MX, Dulac C, Hoekstra HE. The genetic basis of parental care evolution in monogamous mice. Nature 2017;544(7651):434–439. doi:10.1038/nature22074.

Beverly M, Anbil S, Sengupta P. Degeneracy and signaling within a sensory circuit contributes to robustness in thermosensory behaviors in *C. elegans*. J Neurosci. 2011;31(32):11718–11727. doi:10.1523/JNEUROSCI.1098-11.2011.

Biron D, Wasserman SM, Thomas JH, Samuel AD, Sengupta P. An olfactory neuron responds stochastically to temperature and modulates *C. elegans* thermotactic behavior. Proc Natl Acad Sci USA. 2008;105(31):11002–11007. doi:10.1073/pnas.0805004105.

Castillo-Davis CI, Hartl DL, Achaz G. Cis-regulatory and protein evolution in orthologous and duplicate genes. Genome Res. 2004;14(8):1530–1536. doi:10.1101/gr.2662504.

Chin-Sang ID, George SE, Ding M, Moseley SL, Lynch AS, Chisholm AD. The ephrin *vab-2/efn-1* functions in neuronal signaling to regulate epidermal morphogenesis in *C. elegans*. Cell 1999;99(7):781–790. doi:10.1016/S0092-8674(00)81675-X.

Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly (Austin). 2012;6(2):80–92. doi:10.4161/fly.19695.

Clark DA, Biron D, Sengupta P, Samuel ADT. The AFD sensory neurons encode multiple functions underlying thermotactic behavior in *C. elegans*. J Neurosci. 2006;26(28):7444–7451. doi:10.1523/JNEUROSCI.1137-06.2006.

Cook DN, Whitehead GS, Burch LH, Berman KG, Kapadia Z, Wohlford-Lenane C, Schwartz DA. Spontaneous mutations in recombinant inbred mice: mutant toll-like receptor 4 (Tlr4) in BXD29 mice. Genetics 2006;172(3):1751–1755. doi:10.1534/genetics.105.042820.

Cook DE, Zdraljevic S, Roberts JP, Andersen EC. CeNDR, the *Caenorhabditis elegans* natural diversity resource. Nucleic Acids Res. 2017;45(D1):D650–D657. doi:10.1093/nar/gkw893.

Cook DE, Zdraljevic S, Tanny RE, Seo B, Riccardi DD, Noble LM, Rockman MV, Alkema MJ, Braendle C, Kammenga JE, *et al.* The genetic basis of natural variation in *Caenorhabditis elegans* telomere length. Genetics 2016;204(1):371–383. doi:10.1534/genetics.116.191148.

de Bono M, Bargmann CI. Natural variation in a neuropeptide y receptor homolog modifies social behavior and food response in *C. elegans*. Cell 1998;94(5):679–689. doi:10.1016/S0092-8674(00)81609-8.

Evans KS, van Wijk MH, McGrath PT, Andersen EC, Sterken MG. From QTL to gene: *C. elegans* facilitates discoveries of the genetic

mechanisms underlying natural variation. Trends Genet. 2021;37(10):933–947. doi:10.1016/j.tig.2021.06.005.

Farslow JC, Lipinski KJ, Packard LB, Edgley ML, Taylor J, Flibotte S, Moerman DG, Katju V, Bergthorsson U. Rapid increase in frequency of gene copy-number variants during experimental evolution in *Caenorhabditis elegans*. BMC Genomics. 2015;16(1):1044. doi:10.1186/s12864-015-2253-2.

Félix MA, Braendle C. The natural history of *Caenorhabditis elegans*. Curr Biol. 2010;20(22):R965–R969. doi:10.1016/j.cub.2010.09.050.

Félix MA, Duveau F. Population dynamics and habitat sharing of natural populations of *Caenorhabditis elegans* and *C. briggsae*. BMC Biol. 2012;10(1):59. doi:10.1186/1741-7007-10-59.

Flavell SW, Pokala N, Macosko EZ, Albrecht DR, Larsch J, Bargmann CI. Serotonin and the neuropeptide pdf initiate and extend opposing behavioral states in *C. elegans*. Cell 2013;154(5):1023–1035. doi:10.1016/j.cell.2013.08.001.

Frezal L, Felix MA. *C. elegans* outside the petri dish. eLife 2015;4:e05849. doi:10.7554/eLife.05849.

Gaertner BE, Parmenter MD, Rockman MV, Kruglyak L, Phillips PC. More than the sum of its parts: a complex epistatic network underlies natural variation in thermal preference behavior in *Caenorhabditis elegans*. Genetics 2012;192(4):1533–1542. doi:10.1534/genetics.112.142877.

Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. arXiv:12073907. https://doi.org/10.48550/arXiv.1207.3907. 2012.

Harris N, Bates SG, Zhuang Z, Bernstein M, Stonemetz JM, Hill TJ, Yu YV, Calarco JA, Sengupta P. Molecular encoding of stimulus features in a single sensory neuron type enables neuronal and behavioral plasticity. Curr Biol. 2023;33(8):1487–1501.e7. doi:10.1016/j.cub.2023.02.073.

Hawk JD, Calvo AC, Liu P, Almoril-Porras A, Aljobeh A, Torruella-Suarez ML, Ren I, Cook N, Greenwood J, Luo L, *et al.* Integration of plasticity mechanisms within a single sensory neuron of *C. elegans* actuates a memory. Neuron. 2018;97(2):356–367.e4. doi:10.1016/j.neuron.2017.12.027.

Hedgecock EM, Russell RL. Normal and mutant thermotaxis in the nematode *Caenorhabditis elegans*. Proc Natl Acad Sci USA. 1975;72(10):4061–4065. doi:10.1073/pnas.72.10.4061.

Huminiecki L, Wolfe KH. Divergence of spatial gene expression profiles following species-specific gene duplications in human and mouse. Genome Res. 2004;14(10A):1870–1879. doi:10.1101/gr.2705204.

Jourjine N, Hoekstra HE. Expanding evolutionary neuroscience: insights from comparing variation in behavior. Neuron. 2021;109(7):1084–1099. doi:10.1016/j.neuron.2021.02.002.

Jurado P, Kodama E, Tanizawa Y, Mori I. Distinct thermal migration behaviors in response to different thermal gradients in *Caenorhabditis elegans*. Genes Brain Behav. 2010;9(1):120–127. doi:10.1111/j.1601-183X.2009.00549.x.

Katju V, Bergthorsson U. Copy-number changes in evolution: rates, fitness effects and adaptive significance. Front Genet. 2013;4:273. doi:10.3389/fgene.2013.00273.

Kim C, Kim J, Kim S, Cook DE, Evans KS, Andersen EC, Lee J. Long-read sequencing reveals intra-species tolerance of substantial structural variations and new subtelomere formation in *C. elegans*. Genome Res. 2019;29(6):1023–1035. doi:10.1101/gr.246082.118.

Kimura KD, Miyawaki A, Matsumoto K, Mori I. The *C. elegans* thermosensory neuron afd responds to warming. Curr Biol. 2004;14(14):1291–1295. doi:10.1016/j.cub.2004.06.060.

Kobayashi K, Nakano S, Amano M, Tsuboi D, Nishioka T, Ikeda S, Yokoyama G, Kaibuchi K, Mori I. Single-cell memory regulates a neural circuit for sensory behavior. Cell Rep. 2016;14(1):11–21. doi:10.1016/j.celrep.2015.11.064.

Kodama E, Kuhara A, Mohri-Shiomi A, Kimura KD, Okumura M, Tomioka M, Iino Y, Mori I. Insulin-like signaling and the neural circuit for integrative behavior in *C. elegans*. Genes Dev. 2006; 20(21):2955–2960. doi:10.1101/gad.1479906.

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

Kuhara A, Okumura M, Kimata T, Tanizawa Y, Takano R, Kimura KD, Inada H, Matsumoto K, Mori I. Temperature sensing by an olfactory neuron in a circuit controlling behavior of *C. elegans*. Science 2008;320(5877):803–807. doi:10.1126/science.1148922.

Kuzmin E, Taylor JS, Boone C. Retention of duplicated genes in evolution. Trends Genet. 2022;38(1):59–72. doi:10.1016/j.tig.2021.06.016.

Lanjuin A, VanHoven MK, Bargmann CI, Thompson JK, Sengupta P. *Otx/otd* homeobox genes specify distinct sensory neuron identities in *C. elegans*. Dev Cell. 2003;5(4):621–633. doi:10.1016/S1534-5807(03)00293-4.

Lee D, Zdraljevic S, Stevens L, Wang Y, Tanny RE, Crombie TA, Cook DE, Webster AK, Chirakar R, Baugh LR, *et al.* Balancing selection maintains hyper-divergent haplotypes in *Caenorhabditis elegans*. Nature Ecol Evol. 2021;5(6):794–807. doi:10.1038/s41559-021-01435-x.

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25(14): 1754–1760. doi:10.1093/bioinformatics/btp324.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; Genome Project Data Processing S. The sequence alignment/map format and samtools. Bioinformatics 2009;25(16):2078–2079. doi:10.1093/bioinformatics/btp352.

Li WH, Yang J, Gu X. Expression divergence between duplicate genes. Trends Genet. 2005;21(11):602–607. doi:10.1016/j.tig.2005.08.006.

Loker R, Mann RS. Divergent expression of paralogous genes by modification of shared enhancer activity through a promoter-proximal silencer. Curr Biol. 2022;32(16):3545–3555.e4. doi:10.1016/j.cub.2022.06.069.

Lupski JR. Genomic rearrangements and sporadic disease. Nat Genet. 2007;39(S7):S43–S47. doi:10.1038/ng2084.

Maydan JS, Lorch A, Edgley ML, Flibotte S, Moerman DG. Copy number variation in the genomes of twelve natural isolates of *Caenorhabditis elegans*. BMC Genomics. 2010;11(1):62. doi:10.1186/1471-2164-11-62.

McGrath PT, Rockman MV, Zimmer M, Jang H, Macosko EZ, Kruglyak L, Bargmann CI. Quantitative mapping of a digenic behavioral trait implicates globin variation in *C. elegans* sensory behaviors. Neuron. 2009;61(5):692–699. doi:10.1016/j.neuron.2009.02.012.

McGrath PT, Xu Y, Ailion M, Garrison JL, Butcher RA, Bargmann CI. Parallel evolution of domesticated *Caenorhabditis* species targets pheromone receptor genes. Nature 2011;477(7364):321–325. doi:10.1038/nature10378.

Mori I, Ohshima Y. Neural regulation of thermotaxis in *Caenorhabditis elegans*. Nature 1995;376(6538):344–348. doi:10.1038/376344a0.

Niepoth N, Bendesky A. How natural genetic variation shapes behavior. Annu Rev Genomics Hum Genet. 2020;21(1):437–463. doi:10.1146/annurev-genom-111219-080427.

Ohno S. Evolution by Gene Duplication. Heidelberg: Springer-Verlag; 1970.

Pollard DA. Design and construction of recombinant inbred lines. Methods Mol Biol. 2012;871:31–39. doi:10.1007/978-1-61779-785-9_3.

Ramot D, MacInnis BL, Lee HC, Goodman MB. Thermotaxis is a robust mechanism for thermoregulation in *Caenorhabditis elegans* nematodes. J Neurosci. 2008;28(47):12546–12557. doi:10.1523/JNEUROSCI.2857-08.2008.

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24–26. doi:10.1038/nbt.1754.

Rockman MV, Kruglyak L. Breeding designs for recombinant inbred advanced intercross lines. Genetics 2008;179(2):1069–1078. doi:10.1534/genetics.107.083873.

Ryu WS, Samuel AD. Thermotaxis in *Caenorhabditis elegans* analyzed by measuring responses to defined thermal stimuli. J Neurosci. 2002;22(13):5727–5733. doi:10.1523/JNEUROSCI.22-13-05727.2002.

Satterlee JS, Sasakura H, Kuhara A, Berkeley M, Mori I, Sengupta P. Specification of thermosensory neuron fate in *C. elegans* requires *ttx-1*, a homolog of *otd/otx*. Neuron. 2001;31(6):943–956. doi:10.1016/S0896-6273(01)00431-7.

Schwieterman AA, Steves AN, Yee V, Donelson CJ, Bentley MR, Santorella EM, Mehlenbacher TV, Pital A, Howard AM, Wilson MR, *et al.* The *Caenorhabditis elegans* ephrin EFN-4 functions non-cell autonomously with heparan sulfate proteoglycans to promote axon outgrowth and branching. Genetics 2016;202(2): 639–660. doi:10.1534/genetics.115.185298.

Seeholzer LF, Seppo M, Stern DL, Ruta V. Evolution of a central neural circuit underlies *Drosophila* mate preferences. Nature 2018; 559(7715):564–569. doi:10.1038/s41586-018-0322-9.

Sterken MG, Snoek LB, Kammenga JE, Andersen EC. The laboratory domestication of *Caenorhabditis elegans*. Trends Genet. 2015; 31(5):224–231. doi:10.1016/j.tig.2015.02.009.

Suda Y, Nakabayashi J, Matsuo I, Aizawa S. Functional equivalency between *otx2* and *otx1* in development of the rostral head. Development 1999;126(4):743–757. doi:10.1242/dev.126.4.743.

Szewczyk NJ, Kozak E, Conley CA. Chemically defined medium and *Caenorhabditis elegans*. BMC Biotechnol. 2003;3(1):19. doi:10.1186/1472-6750-3-19.

Takeishi A, Yeon J, Harris N, Yang W, Sengupta P. Feeding state functionally reconfigures a sensory circuit to drive thermosensory behavioral plasticity. eLife 2020;9:e61167. doi:10.7554/eLife.61167.

Taylor JS, Raes J. Duplication and divergence: the evolution of new genes and old ideas. Annu Rev Genet. 2004;38(1):615–643. doi:10.1146/annurev.genet.38.072902.092831.

Taylor SR, Santpere G, Weinreb A, Barrett A, Reilly MB, Xu C, Varol E, Oikonomou P, Glenwinkel L, McWhirter R, *et al.* Molecular topography of an entire nervous system. Cell 2021;184(16):4329–4347.e23. doi:10.1016/j.cell.2021.06.023.

Yamada Y, Ohshima Y. Distribution and movement of *Caenorhabditis elegans* on a thermal gradient. J Exp Biol. 2003;206(Pt 15): 2581–2593. doi:10.1242/jeb.00477.

Yeon J, Takeishi A, Sengupta P. Chronic vs acute manipulations reveal degeneracy in a thermosensory neuron network. MicroPubl Biol. 2021;2021:10.17912/micropub.biology.000355. doi:10.17912/micropub.biology.000355.

Yu YV, Bell HW, Glauser DA, Goodman MB, Van Hooser SD, Sengupta P. Camki-dependent regulation of sensory gene expression mediates experience-dependent plasticity in the operating range of a thermosensory neuron. Neuron. 2014;84(5):919–926. doi:10.1016/j.neuron.2014.10.046.

Zhao Y, Long L, Wan J, Biliya S, Brady SC, Lee D, Ojemakinde A, Andersen EC, Vannberg FO, Lu H, *et al.* A spontaneous complex structural variant in *rcan-1* increases exploratory behavior and laboratory fitness of *Caenorhabditis elegans*. PLoS Genet. 2020;16(2): e1008606. doi:10.1371/journal.pgen.1008606.

Editor: J. Ward