



Glial expression of a steroidogenic enzyme underlies natural variation in hitchhiking behavior

Heeseung Yang^a, Daehan Lee^{b,c}, Heekyeong Kim^d, Daniel E. Cook^b, Young-Ki Paik^d, Erik C. Andersen^{b,1}, and Junho Lee^{a,e,2}

Affiliations are included on p. 9.

Edited by Anne Villeneuve, Stanford University, Stanford, CA; received November 27, 2023; accepted June 10, 2024

Phoresy is an interspecies interaction that facilitates spatial dispersal by attaching to a more mobile species. Hitchhiking species have evolved specific traits for physical contact and successful phoresy, but the regulatory mechanisms involved in such traits and their evolution are largely unexplored. The nematode *Caenorhabditis elegans* displays a hitchhiking behavior known as nictation during its stress-induced developmental stage. Dauer-specific nictation behavior has an important role in natural *C. elegans* populations, which experience boom-and-bust population dynamics. In this study, we investigated the nictation behavior of 137 wild *C. elegans* strains sampled throughout the world. We identified species-wide natural variation in nictation and performed a genome-wide association mapping. We show that the variants in the promoter of *nta-1*, encoding a putative steroidogenic enzyme, underlie differences in nictation. This difference is due to the changes in *nta-1* expression in glial cells, which implies that glial steroid metabolism regulates phoretic behavior. Population genetic analysis and geographic distribution patterns suggest that balancing selection maintained two *nta-1* haplotypes that existed in ancestral *C. elegans* populations. Our findings contribute to further understanding of the molecular mechanism of species interaction and the maintenance of genetic diversity within natural populations.

Caenorhabditis elegans | nictation | natural variation | steroidogenic enzyme | balancing selection

In natural ecosystems, interactions between species can mediate the dispersal of organisms. Phoresy is a species interaction in which less mobile organisms travel efficiently by riding on another, usually larger animal (1–6). This commensal interaction requires direct physical contact between organisms, which can be promoted by adaptive traits of hitchhiking species (7, 8). The free-living nematode *Caenorhabditis elegans* exhibits a hitchhiking behavior known as nictation (9, 10). Nictation involves the nematode standing and waving its body to enhance its chances of riding another organism (11). This behavior is only performed by dauers, which is an alternative developmental stage induced under unfavorable conditions. Dauers undergo significant anatomical remodeling and acquire distinct physiological and behavioral properties including nictation (12–15). In the fluctuating natural habitats, *C. elegans* experience boom-and-bust population dynamics (16). During the bust phase, which involves a population collapse, the survival and dispersal of dauer have important roles in initiating a new boom cycle (17). By hitchhiking onto carrier animals such as isopods and slugs, dauer larvae can disperse to a new favorable habitat where they can resume their reproductive growth.

Although phoretic behavior is an ecologically important species dispersal strategy and is observed in many animal species, very little is known regarding the molecular and neural mechanisms of this behavior and its evolution. In this context, *C. elegans*, which has well-established genetic tools and connectome, has been explored as a model system for the investigation of phoretic behavior at all stages, from the genes involved to the resulting neural circuitry to evolution. For example, the core components of the neuronal circuits and various modulatory pathways involved in nictation have been uncovered (9, 10, 18–22). In addition, wild *C. elegans* strains and their whole-genome sequences, which are available through the *Caenorhabditis* Natural Diversity Resource (CaeNDR) (23), provide an opportunity to study the evolution of this behavior. Moreover, after testing a small number of wild *C. elegans* strains, our group previously described natural variation in nictation behavior (10). Given the ecological importance of nictation behavior, this natural variation in this trait among species is likely subject to natural selection. In this study, we examined 137 wild *C. elegans* strains to explore natural variation in nictation behavior on the species level. Our approach, which integrates behavioral, quantitative, and population genetic analyses, identified the molecular basis of variation in nictation behavior and its origin.

Significance

Ecosystems contain diverse species, and interspecific interactions are important for population dynamics and evolutionary change. Many cases of phoresy, a commensal interaction that promotes species dispersal, have been described, but little is known regarding regulatory mechanisms and evolution of specialized traits. Here, we used *Caenorhabditis elegans* nictation behavior as a model for elucidating the mechanism and evolution of phoretic behavior. We characterized the genetic architecture of the natural differences in nictation across 137 *C. elegans* wild strains and uncovered the molecular and cellular basis involved. Moreover, by combining population genetic and phenotypic analyses, we provided insights into the origin and maintenance of genetic diversity.

Author contributions: H.Y. and J.L. designed research; H.Y. performed research; H.K., D.E.C., Y.-K.P., and E.C.A. contributed new reagents/analytic tools; H.Y. and D.L. analyzed data; and H.Y., D.L., and J.L. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2024 the Author(s). Published by PNAS. This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Although PNAS asks authors to adhere to United Nations naming conventions for maps (<https://www.un.org/geospatial/mapsgeo>), our policy is to publish maps as provided by the authors.

¹Present address: Department of Biology, Johns Hopkins University, Baltimore, MD 21218.

²To whom correspondence may be addressed. Email: eleegans@snu.ac.kr.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2320796121/-DCSupplemental>.

Published July 3, 2024.

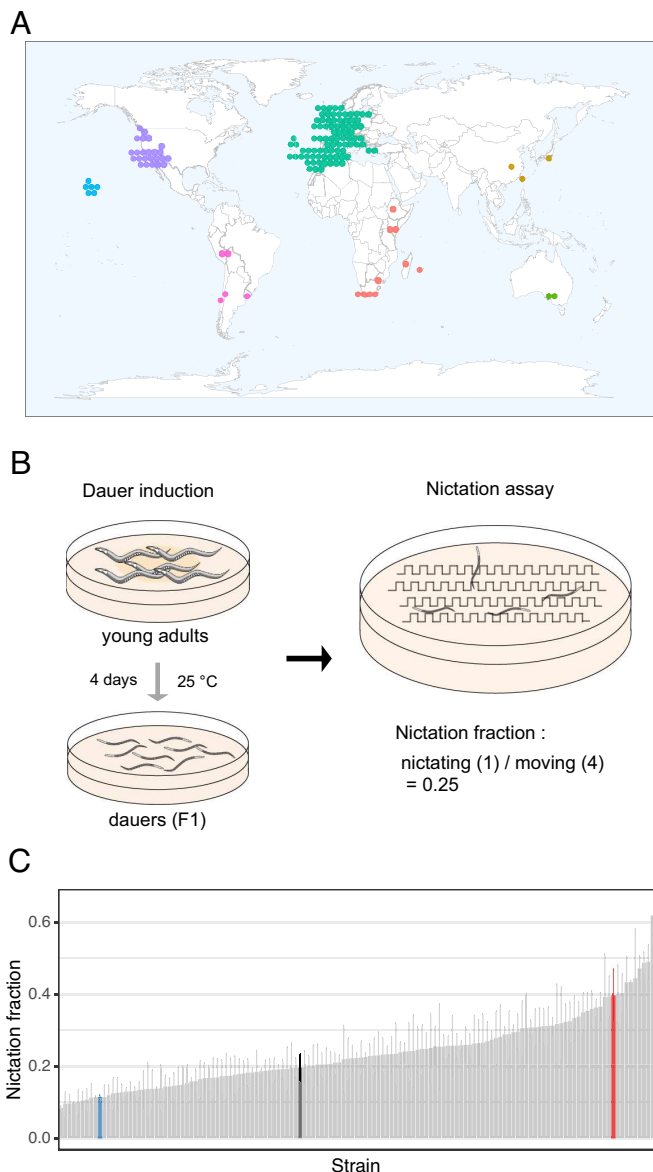


Fig. 1. Natural variation in the nictation behavior of *C. elegans*. (A) Global distribution of 135 of 137 wild isolates used for the nictation assay. Dots are colored by continents (i.e., purple: North America, pink: South America, green: Europe, orange: Africa, yellow: Asia, light green: Australia) except for Hawaii (blue). (B) Schematic of the nictation assay. (C) Nictation fraction of 137 wild isolates. All strains were measured using three replicates. Error bars show SD. The CB4856, N2, and MY23 strains are highlighted in blue, black, and red, respectively.

Results

Genome-Wide Association (GWA) Mapping Reveals the *nict-2* Quantitative Trait Locus (QTL) for Natural Variation in Nictation Behavior. To explore the natural genetic variation in the nictation behavior of *C. elegans*, we examined dauers of 137 wild strains sampled from different sites throughout the world (Fig. 1A). We then quantified nictation behavior on the population level; to do so, we measured the proportion of nictating dauers at a given time three times and averaged these values as the nictation fraction (Fig. 1B, *Methods*). Using this assay, we found that no wild strains were defective in nictation and showed nictation fractions ranging from 0.08 to 0.62 (Fig. 1C and *Dataset S1*). For example, the JU2522 strain exhibited a nictation fraction of only 8%, whereas 62% of JU1395 dauers were found to be nictating. The

narrow-sense heritability (h^2) of variation in nictation fraction was estimated to be 41%, suggesting that substantial genetic variation contributes to the observed behavioral variation. To uncover the genetic architecture of variation in nictation behavior, we performed a GWA mapping (*Dataset S2*) (23, 24). We identified a single QTL on the left arm of chromosome II (Fig. 2A and *SI Appendix, Fig. S1*). This QTL, henceforth *nict-2*, spans from position 3,732,152 to 4,766,883 (1.03 Mb) on chromosome II (*Dataset S3*). Moreover, a single-nucleotide variant (SNV) at the peak position (II: 4,240,841) explains 24.7% of the phenotypic variance. Wild strains sharing the reference (N2 strain)-like (REF) allele display lower nictation fraction than strains with the alternative (ALT) allele (Fig. 2B).

Next, to examine the effect of this locus on nictation behavior, we produced near-isogenic lines (NIL) using CB4856 and MY23 as parental strains. CB4856 harbors the REF allele at this location and exhibits a low nictation fraction (rank: 128/137), whereas MY23 has the ALT allele and displays a high nictation fraction (rank: 10/137). The *nict-2* locus was introgressed from CB4856 into the MY23 genetic background (LJ1301: *nict-2* QTL, CB4856 > MY23) and vice versa (LJ1317: *nict-2* QTL, MY23 > CB4856) (Fig. 2C). The behavior phenotypes of the NILs confirmed the effect of the *nict-2* QTL. Introducing CB4856 *nict-2* QTL led to a decrease in the nictation fraction of the MY23 background, whereas the MY23 *nict-2* QTL in the CB4856 background increased the nictation fraction (Fig. 2C).

To specify the genomic locus underlying nictation variation more precisely, we created three additional NILs (LJ1318–LJ1320) from LJ1317 by generating recombinants across the *nict-2* QTL (Fig. 3A). These NILs harbor genomic regions of MY23 in the CB4856 background. The QTL in LJ1317 was split into two loci of 660 kb (LJ1317 to LJ1318) and 91 kb (LJ1318 to LJ1319) in length, suggesting polygenic contribution. We focused on the 91 kb loci (II:3,708,886 to 3,799,634). While LJ1319 exhibited a nictation fraction similar to the CB4856 parental control, LJ1318 displayed a significant increase in nictation fraction. LJ1318 carries a 91-kb longer MY23 genome than LJ1319. As this 91-kb region showed a major QTL effect even at its shorter range, and also was proximal to the GWA peak, we proceeded to determine the causal genetic variation within this QTL.

A Hydroxysteroid Dehydrogenase Underlies the *nict-2* QTL Effect.

The 91-kb *nict-2* locus contained 31 protein-coding genes, five pseudogenes, two ncRNAs, and one 21U-RNA gene (Fig. 3B). Of these, the *F12E12.11* gene was previously reported to reside within a dauer formation QTL, and its expression was altered under dauer-inducing conditions (26, 27). Noting that nictation is a dauer-specific behavior, we tried to test the effect of this gene in nictation. We performed RNAi knockdown along with several other genes and found that knockdown of *F12E12.11* reduces nictation (*SI Appendix, Fig. S2*). We also deleted *F12E12.11* using CRISPR/Cas9 genome editing (Fig. 3C) and found that deletion of this gene increased the nictation fraction in both CB4856 and N2, both of which share the REF allele at the *nict-2* QTL peak position (Fig. 3D). Because we were unable to exclude the possibility that RNAi has disadvantages such as off-target effects and differences in effects between tissues, we decided to pursue the phenotype of CRISPR deletion mutants and conducted further experiments, which confirmed that the CRISPR mutation produced a consistent phenotype. We named this gene *nta-1* (*nta*; NicTation Altered). The *nta-1* encodes an ortholog of a human *HSD17B14* (hydroxysteroid 17-beta dehydrogenase 14), an enzyme involved in steroid hormone metabolism (28). The increased nictation fraction in *nta-1* mutants suggests that this

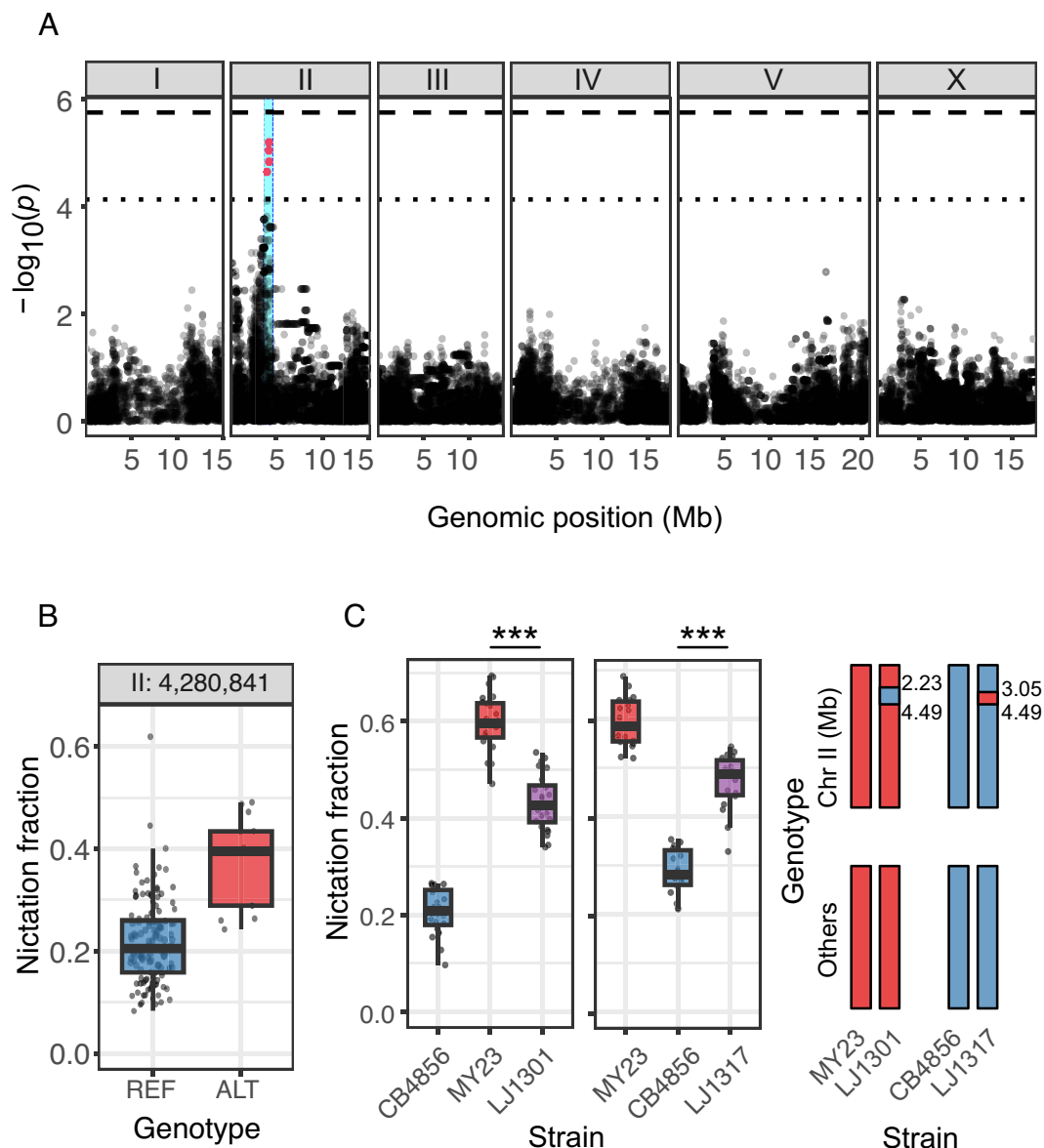


Fig. 2. GWA mapping reveals a QTL for nictation behavior, *nict-2*. (A) Manhattan plot for the GWA mapping of nictation fraction. The upper horizontal dashed line corresponds to the Bonferroni-corrected threshold using all markers, while the lower horizontal dashed line corresponds to the EIGEN threshold, which is the Bonferroni-corrected threshold for the number of independent markers (i.e., the genome-wide eigen-decomposition threshold). SNVs were colored red if they passed the second threshold. A cyan rectangle represents the region of interest for the QTL. (B) Phenotype by genotype split for the QTL region. Strains were divided into two groups, with the peak SNV(chrII: 4,280,841) used as a marker. REF indicates the reference group, which has the same SNV as the reference strain N2. ALT indicates the alternative group. (C) Nictation fraction (Left) and genotype (Right) of two NILs, LJ1301 (*nict-2*, CB4856>MY23) and LJ1317 (*nict-2*, MY23>CB4856). Statistics: *** $P < 0.001$. Multiple comparison (Tukey's HSD).

gene suppresses the nictation behavior. Notably, when the MY23 version of *nta-1* was deleted in LJ1318, no phenotypic effect was observed. We hypothesized that MY23 harbors a loss-of-function allele of *nta-1*, and may therefore enhance the nictation behavior of this strain. Thus, no additional effect was observed in the *nta-1* deletion experiments in LJ1318.

To test the causality of the *nta-1* gene on the *nict-2* QTL effect as well as the functionality of *nta-1* natural alleles, we performed a reciprocal hemizygosity F1 test (29) (Fig. 3E). We found that two homozygous strains, harboring the CB4856 allele or MY23 alleles for *nta-1*, displayed significantly different nictation fractions (Fig. 3E). This result demonstrates that *nta-1* polymorphism indeed underlies the natural variation in nictation behavior. In addition, the hemizygous strain with the MY23 allele showed a similar phenotype to the *nta-1* null mutant, further supporting the hypothesis that MY23 contains a loss-of-function

allele for *nta-1*. The phenotype of hetero-F1 was similar to CB4856, suggesting CB4856 allele is dominant to MY23 allele.

Lack of *nta-1* Expression in Glial Cells Enhances Nictation Behavior. The *nta-1* locus is rich in SNVs and structural variants (Fig. 4A). We identified two missense SNVs within the coding sequence of the MY23 *nta-1* allele. In addition, we also identified over 100 noncoding SNVs in the region encompassing the gene and its promoter (0.85 kb), as well as alterations in intronic structure (Fig. 4A). To investigate whether genetic changes in the coding or promoter region are responsible for differences in behavior, we performed promoter swap experiments. While transgene expression driven by *nta-1p*^{CB4856} significantly reduced the nictation of the *nta-1* mutant to levels similar to the wild-type (CB4856) irrespective of the protein-coding sequence, transgene expression driven by

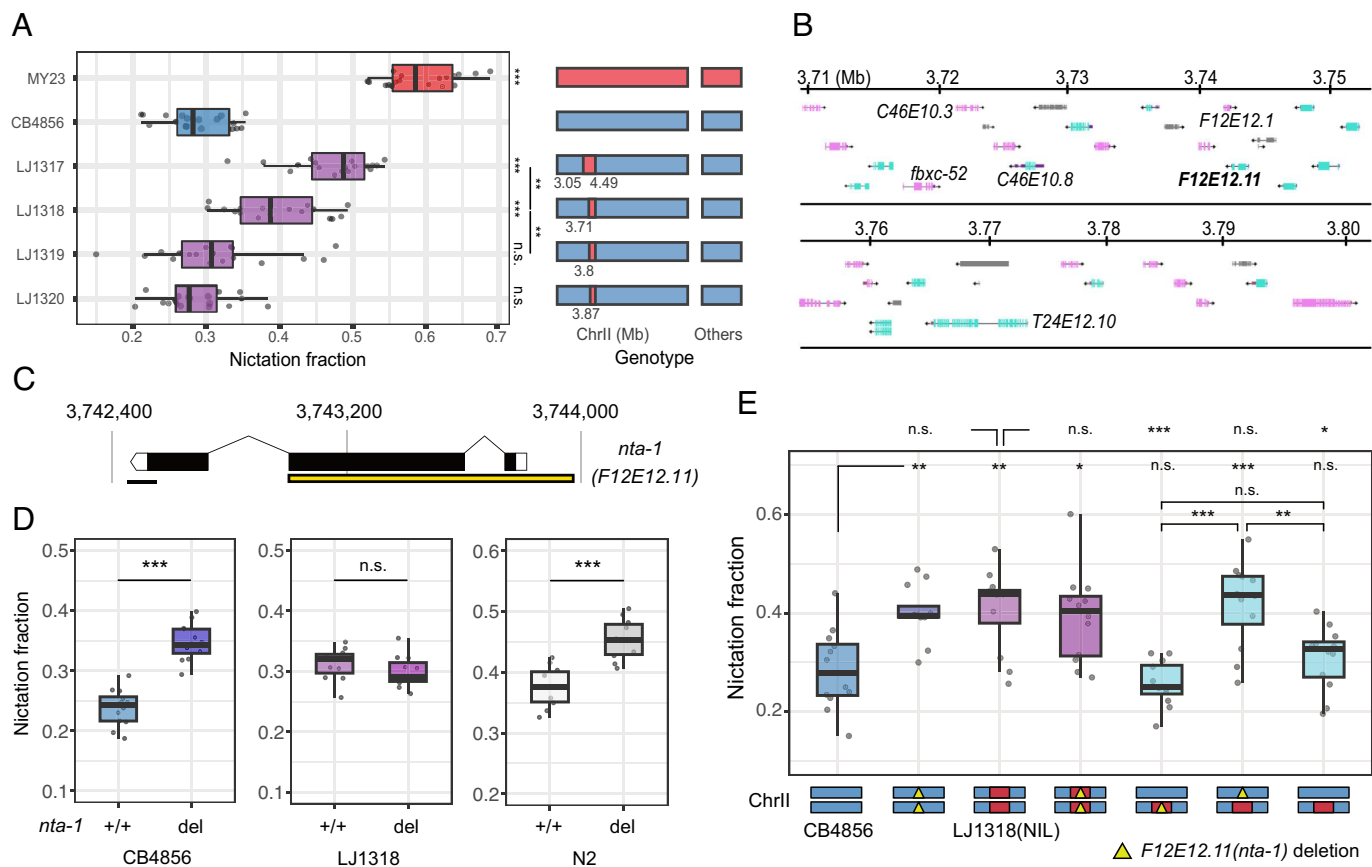


Fig. 3. Variants of a putative steroid metabolism enzyme, *nta-1*, underlie natural variation in nictation of *C. elegans*. (A) Nictation fraction of NILs in the CB4856 background used to map the 91-kb *nict-2* locus. Statistics: *** $P < 0.001$, ** $P < 0.01$. Multiple comparison (Tukey's HSD). (B) The *nict-2* locus. Protein coding genes are colored pink or cyan based on their direction. Other regions are colored gray. Images obtained from JBrowse 2 of Wormbase were modified to generate this figure (25). (C) A schematic plot for the genetic structure of *nta-1*. The yellow box represents the CRISPR-Cas9-mediated deletion site. (D) Nictation fraction of *nta-1* mutants of the CB4856 and LJ1318 NIL and N2. Statistics: *** $P < 0.001$. Pairwise t test. (E) Nictation fractions of *nta-1* mutants and reciprocal F1 progenies. CB4856, LJ1318, and *nta-1* mutants of both strains were used. The bottom of the figure shows the chromosome II. Blue and red boxes represent the genomes of CB4856 and MY23, respectively. The yellow triangle represents the *nta-1* deletion allele. Statistics: *** $P < 0.001$. Multiple comparison (Tukey's HSD).

nta-1^{MY23} did not affect nictation as much (Fig. 4B). Moreover, *nta-1* expression driven by *nta-1*^{CB4856}, not by *nta-1*^{MY23}, also reduced the nictation fraction of MY23 strain (SI Appendix, Fig. S3). These results demonstrate that genetic differences in the *nta-1* promoter, rather than in the coding region, explains behavior differences between CB4856 and MY23.

Next, we investigated how sequence divergence at the *nta-1* promoter led to differences in gene expression. Using a published RNA-seq dataset (30), we found that *nta-1* expression was significantly lower in wild strains with *nta-1*^{MY23} than in strains with *nta-1*^{CB4856} (Fig. 4C). Consistently, when we examined the spatial expression patterns of these two promoters, we found weaker signals from *nta-1*^{MY23}. Both *nta-1*^{CB4856} and *nta-1*^{MY23} drive expression in a variety of tissues, including pharyngeal gland cells, pharyngeal muscles, coelomocytes, body wall muscles, the head mesodermal cell (HMC), and in intestinal cells (Fig. 4D and SI Appendix, Table S1). However, compared to *nta-1*^{CB4856}, *nta-1*^{MY23} expression was absent in head muscles and GLR cells in dauer (Fig. 4D and SI Appendix, Figs. S4–S6) (31–33). The phenotype of the mutants and the expression data suggest that additional expression driven by the CB4856 promoter lowers nictation (Fig. 4E).

GLR cells are glial cells whose cell bodies are located just behind the nerve ring, the major neuropil in *C. elegans*. There they extend sheet-like processes anteriorly into the inner region

of the nerve ring (31, 32, 34) and form intensive gap junctions with head muscles and motor neurons (RMEs), implying that they play a role in the fine-tuning of head movements (34). Notably, the chemical synapses from IL2 neurons to GLR cells are strengthened at the dauer stage (SI Appendix, Fig. S7) (35). Given the essential role IL2 neurons play during nictation (9, 18), we hypothesized that *nta-1* expression in GLR cells may regulate nictation. To test this possibility, we expressed *nta-1*(CB4856 *nta-1* coding region) in GLR cells under two independent promoters (i.e., *nep-2p* and *egl-6p*, from N2 strain) (33). We found that expressing *nta-1* in GLR cells was sufficient to rescue *nta-1* mutant phenotype (Fig. 4F). Moreover, *nta-1* expression by GLR promoter also reduced the nictation fraction of MY23 strain (SI Appendix, Fig. S8). Since *nta-1* expression was also different in head muscles, we also tried expressing *nta-1* under the head muscle promoter (i.e., *ser-2p*, from N2 strain), but it had no effect (SI Appendix, Fig. S8). Taken together, these results demonstrate that glial *nta-1* expression underlies natural variation in nictation behavior.

Ancient Balancing Selection Has Maintained Two Versions of the *nta-1* Promoter. To better understand the evolution of the *nta-1* promoter on a species-wide scale, we extended our analysis to the 548 wild *C. elegans* genomes available from CaenorNDR (23). We found that most wild genomes shared their *nta-1* haplotype either

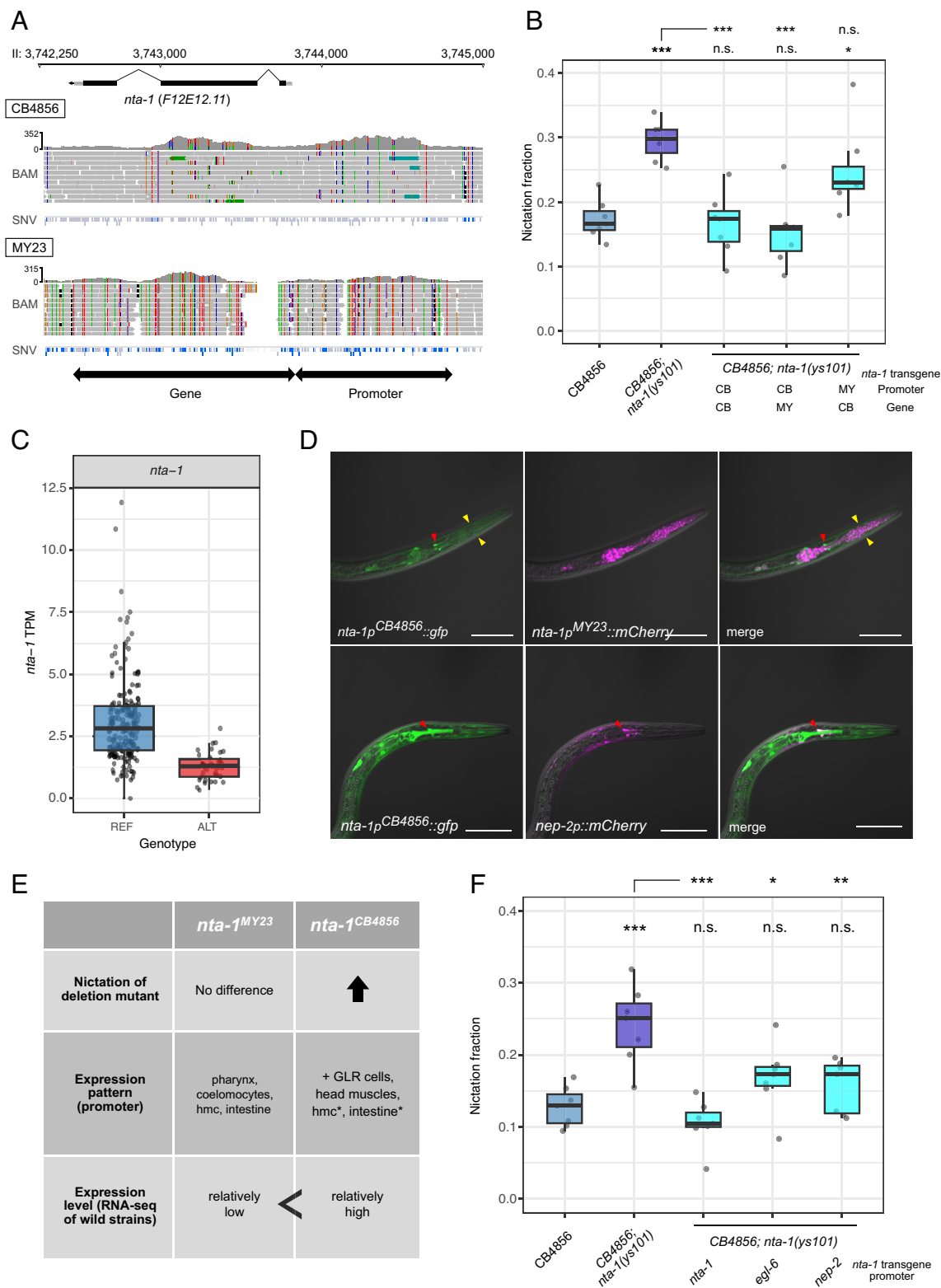


Fig. 4. *nta-1* expression in GLR glial cells regulates nictation. (A) Variants in *nta-1*. BAM: Short-red alignments of CB4856 and MY23 to the N2 reference genome (WS245) around *nta-1*. SNV: SNVs of each strain compared to the N2 reference genome are colored blue. The image obtained from CaenDR was modified. (B) Nictation fraction of worms expressing the *nta-1* transgene of different promoter-gene combinations in the *nta-1* mutant in CB4856 background. CB and MY indicate sequences from CB4856 and MY23, respectively. Statistics: *** $P < 0.001$, * $P < 0.05$. Multiple comparison (Tukey's HSD). (C) Phenotype by genotype split for *nta-1* TPM from RNAseq data of wild isolates. Each dot represents a strain. RNAseq data were obtained from ref. 30. *nta-1* promoter haplotypes are defined in Fig. 5A. (D) Expression patterns of dauers driven by each promoter (Red arrows: GLR cells, Yellow arrows: head muscles). Dauer induction conditions caused slight mCherry aggregation. (Scale bars, 50 μ m.) (E) A summary table of phenotypic differences between two divergent *nta-1* promoters. (F) Nictation fractions of worms expressing *nta-1* transgenes driven by its own and GLR promoters in the *nta-1* mutant in CB4856 background. Statistics: *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$. Multiple comparison (Tukey's HSD).

with CB4856 or MY23. Specifically, 409 wild genomes with *nta-1p*^{CB4856}, 133 with *nta-1p*^{MY23}, and eight remaining wild genomes did not carry either of the two major haplotypes (Fig. 5A and

SI Appendix, Table S2). As for the peak SNV of *nict-2* QTL, the *nta-1* promoter haplotype was clearly associated with the nictation variation (Fig. 5B).

Next, we investigated the geographic distribution of the two major *nta-1p* haplotypes (Dataset S4). Both haplotypes are found worldwide and often coexisted within the same geographic areas or local populations (Fig. 5C). For example, 60% of the Hawaiian strains carry *nta-1p*^{MY23} and 35% possess *nta-1p*^{CB4856} and these two haplotypes have been found to be colocalize throughout the Hawaiian islands (Fig. 5D). Similar colocalization and balanced frequency patterns for these two haplotypes were also observed throughout the western continental United States (SI Appendix, Fig. S9). This geographic pattern implies that different *nta-1p* haplotypes are maintained in local populations via balancing selection. Indeed, we detected a pronounced signature of long-term balancing

selection at the *nta-1* locus, as evidenced by Tajima's *D* estimates being substantially higher than in adjacent regions (Fig. 5E). Previous studies revealed greater genetic diversity of *C. elegans* in Hawaii than in other regions of the world, suggesting that this population is among the oldest in the wild (36–38). Notably, we identified both *nta-1p*^{CB4856} and *nta-1p*^{MY23} in wild strains of the Hawaiian Divergent population (Fig. 5F and SI Appendix, Table S3), which has been shown to have diverged early and therefore isolated from later gene flow events (36). These results imply that the two *nta-1p* haplotypes existed in ancestral *C. elegans* populations prior to divergence and have been preserved over many generations via ancient balancing selection.

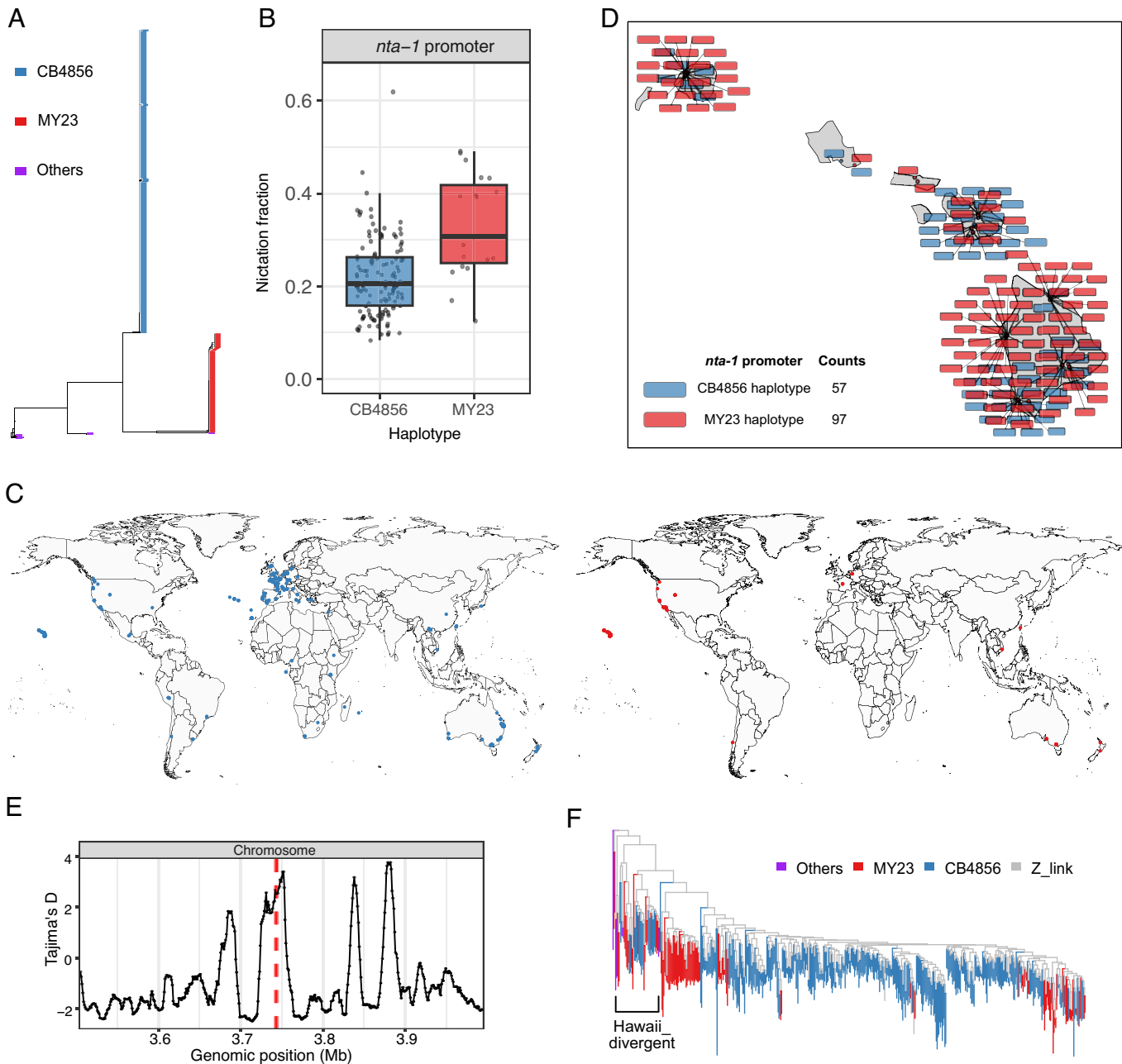


Fig. 5. Ancient balancing selection has maintained two versions of the *nta-1* promoter. (A) Haplotype tree of the *nta-1* promoter from 550 *C. elegans* isolates. The *nta-1* promoter haplotypes of CB4856 (*nta-1p*^{CB4856}) and MY23 (*nta-1p*^{MY23}) are shown in blue and red, respectively. (B) Phenotype by genotype split for the nictation fractions of wild isolates. Strains are split based on the haplotype of the *nta-1* promoter. (C) Global geographic distribution of the *nta-1* promoter haplotypes. Each dot represents a strain. Shown are strains with the CB4856 *nta-1* haplotype (Left, blue dots) and strains with MY23 *nta-1* haplotype (Right, red dots). (D) Geographic distribution of the *nta-1* promoter haplotype in Hawaii. (E) Tajima's *D* statistics across the *nta-1* locus. The dashed red line indicates the location of *nta-1*. (F) Genome-wide tree of 550 *C. elegans* isolates. Color indicates the *nta-1* promoter haplotype. The Hawaii divergent group was reconstructed based on the range of the distribution of isolates in previously divided groups as reported by Crombie et al. (36).

Since nictation behavior facilitates dispersal to new habitats through phoresy (10), the MY23 *nta-1* allele that enhances nictation likely offers a fitness advantage. However, the balance between *nta-1p^{MY23}* and *nta-1p^{CB4856}* suggests more complex fitness implications for these alleles. When we assessed its effect on reproduction, while two *nta-1* alleles did not show differences in total fecundity (SI Appendix, Fig. S10), we found differences in the postdauer reproductive success. Specifically, when comparing the brood sizes up to 48 h after dauers were transferred to a new plate with food, only deletion of the CB4856 *nta-1* allele delayed postdauer reproduction, thereby resulting in a significantly lower brood size at 48 h (SI Appendix, Fig. S11). This delayed reproduction was specific to postdauers; the brood sizes in nematodes up to 72 h from egg that bypassed the dauer stage did not differ between strains. The pleiotropic effect of *nta-1* variation might have influenced fitness in multiple ways, which in case affect balancing selection of *nta-1*.

Discussion

Our GWA mapping uncovered a single large-effect QTL underlying natural differences in nictation behavior. Using NILs, we found that the QTL effect was split into two loci. We traced the effect of the major short (91-kb) QTL to a specific quantitative trait gene, *nta-1*. The detailed molecular function of *NTA-1* remains elusive, but sequence homology provides some clues. *nta-1* is homologous to human *HSD17B14*, an enzyme that is known to have a function involved in the metabolism of steroids at C17 (for example, oxidizing estradiol to estrone) (28). Intriguingly, SNPs in human *HSD17B14* are associated with athletic performance (39). In *C. elegans*, sterol metabolism plays a key role in dauer development (40, 41). For example, dafachronic acid, a specific steroid hormone, is produced by the *DAF-9* cytochrome P450 enzyme and acts to prevent dauer formation by interacting with the nuclear hormone receptor, *DAF-12* (42). We hypothesize that *nta-1* is involved in the steroid hormone signaling that regulates dauer-specific traits, including nictation behavior.

Next, the promoter swap experiments revealed that coding variants do not impact behavior. Rather, we observed that expression of this gene in a specific glial cell type, GLR, underlies variations in nictation behavior (Fig. 6). Despite differences in GLR expression, we also found that two promoters drive common expression in various other tissues. Expression of the *nta-1* transgene by *nta-1p^{MY23}* had no effect on nictation, suggesting that *NTA-1* in

these cells may have a function distinct from regulating nictation. Our findings exemplify how cell-specific function of a gene and changes in their expression can contribute to differences in behavior. Since GLR cells establish strengthened synapse-like connections with IL2 neurons that initiate nictation, cell-autonomous alterations in GLR cells may influence the function of IL2 neurons and their connected neural circuits. Alternatively, changes in *nta-1* expression in GLR may exert cell nonautonomous impacts; for example, GLR cells may regulate the local sterol hormone environment in the nerve ring during the neural circuit remodeling phase of dauer development. Recently, it has been demonstrated that glial steroid signaling in *Drosophila melanogaster* regulates sleep behavior (43), and glial hormone esterase governs caste-specific behavior in ants (44), indicating that the glial response to steroid hormones, which regulate plastic behavior, might be widely conserved.

By leveraging whole-genome sequence data from 550 wild *C. elegans* genomes, we explored the origin of variation in *nta-1* expression. Our analysis demonstrated that two major haplotypes of *nta-1* promoter (i.e., *nta-1p^{CB4856}*: 74%, *nta-1p^{MY23}*: 24%) arose in an ancient *C. elegans* population and have been maintained via long-term balancing selection. Although wild strains containing *nta-1p^{MY23}* display a higher nictation ratio compared to those with *nta-1p^{CB4856}*, *nta-1^{CB4856}* allele have advantage in postdauer reproduction rate. This implies that modification in steroid hormone signaling due to changes in *nta-1* expression may exert pleiotropic effects, thereby leading to a context-dependent balancing selection. Although both haplotypes are distributed worldwide, we found that the balance between the two haplotypes was not as prevalent outside of the Pacific region as it was within Pacific populations (SI Appendix, Fig. S9). Although it might seem counterintuitive that the *nta-1p^{CB4856}*, which reduces nictation behavior, has spread worldwide, our findings align with a previously proposed hypothesis, which suggested that *C. elegans* likely had its origins in the Pacific region and thereafter expanded beyond the Pacific via human-mediated migration (36–38). For example, the transport of agricultural products might reduce the reliance on phoresy related to other animals. Furthermore, wild *C. elegans* strains are often found in human-associated environments such as orchards, gardens, and compost (16, 45). Such nutrient-rich settings might favor quick dauer exit and postdauer reproduction instead of a prolonged dispersal stage. While human activities are thought to influence dauer entry (46), our study implies that they have also reshaped the evolution of dauer dispersal and postdauer reproduction.

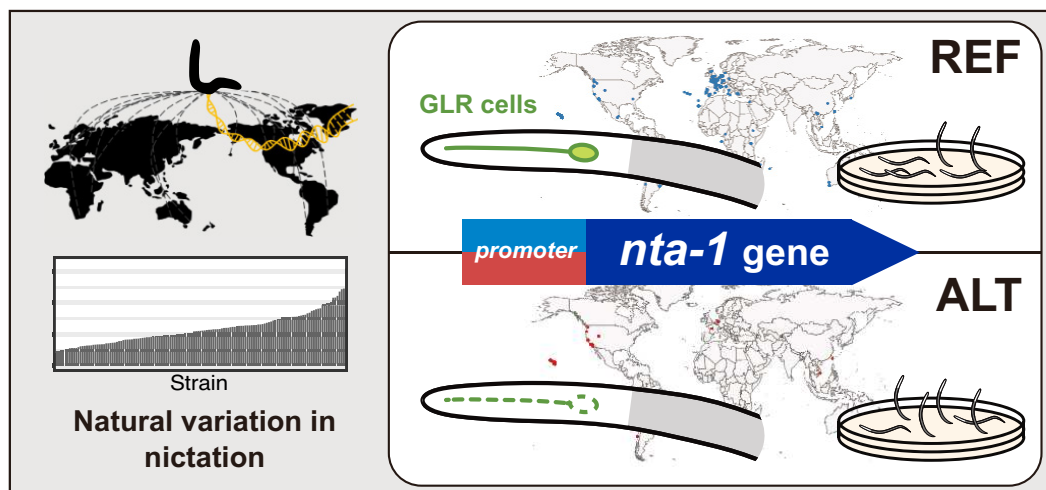


Fig. 6. Graphical summary.

Methods

C. elegans Strains. Worms were grown at 15 or 20 °C under standard conditions (47), except for dauer induction. A total of 137 *C. elegans* wild strains used in GWA mapping were obtained from Caenor (Dataset S1) (23). Information about other strains made in this study (NILs, mutants, transgenic) is summarized in SI Appendix, Table S4.

Dauer Induction. Approximately 15 young adults were transferred to the pheromone plates seeded with *Escherichia coli* OP50. Plates were then incubated at 25 °C. Pheromone plates contained agar (10 g/L), agarose (7 g/L), NaCl (2 g/L), KH₂PO₄ (3 g/L), K₂HPO₄ (0.5 g/L), cholesterol (8 mg/L), and synthetic pheromone-ascarosides C7 (i.e., daumone 1, ascaroside 3), C6 (i.e., daumone 2, ascaroside 1), and C9 (i.e., daumone 3, ascaroside 2) (2 mg/L each) (48, 49). After 4 d, dauers among progeny were identified by their constricted and thin bodies.

Nictation Assays. To assess nictation, a 3.5% agar solution was first poured onto a poly-dimethylsiloxane (PDMS) mold to prepare microdirt chips. A microwave was then used to dissolve the agar powder. After the agar had solidified, it was removed from the PDMS mold and dried at 37 °C for 90 min. Dauers on the pheromone plates were then collected by glass capillary (Kimble Chase Life Science and Research Products LLC.) using M9 buffer and transferred onto the dried microdirt chip. After 30 to 40 min, when most of the dauers were moving, the proportion of the nictating dauers among the moving dauers was measured. For each biological replicate, the nictation fraction was measured three times in a row. The mean of these three technical replicates represented the nictation fraction of a biological replicate. All nictation assays were performed in a closed room with a thermo-hygrostat set to maintain a temperature of 25 °C and a humidity of 30%. One hundred thirty-six wild isolates were tested alongside a CB4856 control, and their nictation fractions were normalized to that of CB4856. For all the other nictation assays, nictation fractions were regressed by fitting a linear model (i.e., nictation fraction ~ assay trial), and the (total) mean raw nictation fraction was added to each residual.

Calculation of Narrow-Sense Heritability. The nictation fractions of 137 wild isolates were then used to calculate an estimate of narrow-sense heritability (h^2). The mmer function as implemented by the sommer R package was used to calculate the additive variance. Then, the estimate of h^2 was calculated using the pin function of the same package (i.e., $h^2 \sim V1/V1 + V2$).

GWA Mapping. Before mapping, the nictation fractions of 137 wild isolates were rescaled to have a mean of zero and a SD of one. GWA mapping and fine mapping of QTL were performed in Caenor using NemaScan (24), using genotype data from the latest VCF release (i.e., Release 20220216) on Caenor (23, 24). Hard-filter VCF and imputed VCF files were used for GWA mapping and fine mapping, respectively. For the GWA mapping, regions of interest were extended as intervals of ± 150 SNVs from the significant marker above the eigen-decomposition significance threshold and were taken to define a QTL. Regions of interest within 1,000 SNVs were grouped together as a single QTL.

Generation of NILs. All NILs were generated from CB4856 and MY23 parental strains. To do so, CB4856 was backcrossed to the MY23 parents ten times while selecting the CB4856 genome near the QTL interval on chromosome II by checking for single-nucleotide polymorphisms (SNPs) at II:3,768,531 and II:4,286,410 (SI Appendix, Table S5). To genotype 3,768,531, we used EcoRV restriction fragment length polymorphism (RFLP) of single-worm PCR products amplified using the following primers: 5'-CGGTTCGGCTGGAATTGG-3' and 5'-CAGTGTCTCCGTTAAAGCG-3'. To genotype 4,286,410, we used HaeIII RFLP of single-worm PCR products amplified using the following primers: 5'-TCGTAAATCCACACCTGCGG-3' and 5'-CACGTGGAAACCGCATCTG-3'. MY23 was then backcrossed to the CB4856 parents. Next, the genomes of the generated NILs were analyzed by short-read sequencing. The remaining unwanted genomes in chromosome II were removed by further backcrossing to generate LJ1301 and LJ1317. Finally, LJ1318, LJ1319, and LJ1320 were generated by further backcrossing of LJ1317.

RNAi Nictation Assays. For RNAi nictation assays, we used a *C. elegans* RNAi feeding library from Julie Ahringer's group. Each target cell in this library was streaked and cultured in LB broth containing ampicillin. Next, RNAi plates

and RNAi pheromone plates containing 1 mM IPTG were prepared for the *rrf-3*(pk1426) mutant. To avoid OP50 being transferred alongside nematodes, we transferred worms three times. First, L4 worms on the OP50 plates were transferred to empty fresh plates without bacteria and were then allowed to move for 30 min. These same nematodes were then transferred to RNAi plates. After 30 min, they were again transferred to new RNAi plates, which were then incubated at 25 °C. After 3 d, approximately 15 young adults (F1) were transferred to the RNAi pheromone plates. After another 4 d, dauers were collected, and their nictation fractions were measured using the procedure described above.

CRISPR/Cas9-Mediated Mutagenesis. To induce deletion in *nta-1*, we designed two sgRNA sequences 5'-ACCTTCTCAAGACCTGGTGG-3' and 5'-GTGATCATGTGTGCCT GCGG-3' that together can induce deletion of most of the *nta-1* sequence. pJL3021 and pJL3022, each containing one of these two sgRNAs, were derived from the pJW1219 CRISPR/Cas9 plasmid containing sgRNA(F+E) and were edited using the Q5 site-directed mutagenesis kit (E0554; New England Biolabs). CRISPR-edited worms were identified by co-CRISPR using a PJA50 plasmid and an AF-JA-76 repair template that induce *unc-58*(e665) (50). The injection mixture included a final concentration of 30 ng/μL of pJA50, 20 ng/μL of the *unc-58* repair template, and 50 ng/μL of both pJL3021 and pJL3022. This was then injected into young adults of the target background strains. F1 progenies with the *unc* phenotype were then singled out to new plates and incubated until enough eggs were laid. After laying eggs, F1 progenies were genotyped after single-worm lysis. Genotyping was accomplished by determining the length of the PCR product, including the *nta-1* deletion site, that was amplified by the following primers: 5'-CTGTCGTGGGATGTTGA-3' and 5'-GGCTCAACCCGTAGAAATG-3'. F2 progenies were genotyped in the same way, and the F2 progenies with homozygous *nta-1* mutations were then backcrossed to each background strain for at least three generations to avoid off-target editing.

F1 Hemizygous Nictation Assays. A total of 25 to 30 of day one adult CB4856 or LJ1321 hermaphrodites were cultured to lay eggs. During these days, worms were transferred to a new plate each day to check for whether eggs had been laid. After 3 to 4 d, when their egg-laying had nearly ceased, they were then mated with LJ1318 or LJ1322 males. The next day, 18 of the mated hermaphrodites were transferred to a pheromone plate. The mated hermaphrodites were distinguished by the formation of copulatory plugs around the vulva. Four days later, the nictation fraction of the F1 dauers was measured. After the completion of these assays, the dauers on the nictation chips were collected again and transferred to new fresh plates containing OP50 food. The next day, the male-to-hermaphrodite ratio of the recovered dauers was counted to validate that the measured dauers represented the pool of F1 progenies.

Generation of the Transgene Construct. To generate all *promoter::gfp* or *promoter::mCherry* plasmids, promoter PCR templates were inserted into the GFP vector pPD95.77 or the mCherry vector pPD117.01. All plasmids for rescue experiments (i.e., *promoter::nta-1::sl2::gfp*) were generated by modification of the pEM1 vector (51). All constructs were made using a Gibson assembly cloning kit (E5510; New England Biolabs).

Generation of Transgenic Animals. Plasmid constructs were injected into the gonads of the young adult animals as previously described (52). The injection mixture contained plasmids for fluorescent protein expression at 100 ng/μL, and/or plasmids for *nta-1* rescue at 50 ng/μL, and/or plasmids expressing a *rol-6*(*su1006*) transgenic marker at 100 ng/μL.

Microscopy. A ZEISS LSM700 confocal microscope and ZEN software (Carl Zeiss) were used to obtain images of fluorescence protein expression. When using a confocal microscope, animals were fixed on a 3% agar pad on a slide glass and were paralyzed using 3 mM levamisole.

Population Genetics. A VCF file containing 58 biallelic SNVs from 550 wild *C. elegans* genomes was converted to the PHYLIP file format to generate a *nta-1* promoter tree. A distance matrix and pseudorecorded (ECA2199) neighbor-joining tree were produced from this PHYLIP file using dist.ml and the NJ function as implemented by the phangorn (version 2.11.1) R package. This tree was then visualized using the ggtree (version 3.6.2) R package. Using

this tree, we classified two major haplotypes at the *nta-1* promoter locus. To visualize the geographic distribution of these *nta-1* haplotypes, the latest strain data from the CaenDR (Release 20220216) were used to get collection information for each isolate (Dataset S4). A genome-wide neighbor-joining tree for 550 wild isotypes was then obtained from the CaenDR (23). This tree was pseudorooted with ECA2199 and was visualized using the ggtree. Tajima's D statistics for the *nta-1* locus were obtained from ref. 38, where Tajima's D was calculated from the allele frequency spectrum of SNVs within a 10 kb window across 328 wild genomes.

Brood Size Assays. To obtain synchronized embryos, approximately 30 day one adults were transferred onto fresh NGM plates seeded with OP50; these were incubated at 20 °C. After an hour, adult worms were fully removed from the plate. After 2 d, the synchronized L3–L4 worms were individually transferred to different plates. After reaching adulthood and starting to lay eggs, the adult worms were transferred to a new plate every day for 2 d. The number of progenies in each plate was then counted after 2 d after transfer; these values were then summed to obtain the total brood size of an individual worm. For postdauer assays, 30 dauers were transferred onto fresh NGM plates seeded with OP50. Dauers tended to move out of the OP50 lawns, so we checked on them for about 30 min to transfer them back onto the OP50 lawn. After 40 h, the worms were transferred to new plates and were kept there for another 8 h, after which they were again removed. The number of worms remaining was counted at 40 and 48 h. After 2 d, the number of progeny in each plate was counted. The number of offspring on the first plate was divided by the number of animals remaining after 40 h, and the number of offspring on the second

plate was divided by the number of animals remaining after 48 h. Finally, these two were added together to calculate the mean brood size up to 48 h after dauer transfer. Next, for the brood size assay started from eggs, synchronized embryos were obtained as described above. After 2 d, 30 synchronized L3–L4 F1 worms were transferred onto a new plate. The next day (72 h from the eggs laid), the number of remaining F1 worms was counted, and they were removed from the plate. The number of F2 progeny on each plate was counted after 2 d when they were grown up. Brood sizes were normalized by fitting a linear model, (i.e., brood size ~ test date).

Data, Materials, and Software Availability. All study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. We thank Hyunsoo Yim for helping with the connectome data and Jun Kim for helping with CB4856 long-read sequencing data. We appreciate all members of Lee Lab for helpful discussion. Some *C. elegans* strains were kindly provided by the Caenorhabditis Genetics Center. This research was funded by the Samsung Science and Technology Foundation (SSTF-BA1501-52). H.Y. was supported by the BK21 program.

Author affiliations: ^aDepartment of Biological Sciences, Seoul National University, Seoul 08826, Republic of Korea; ^bDepartment of Molecular Biosciences, Northwestern University, Evanston, IL 60208; ^cDepartment of Biological Sciences, Sungkyunkwan University, Suwon 16419, Republic of Korea; ^dYonsei Proteome Research Center, Yonsei University, Seoul 03722, Republic of Korea; and ^eResearch Institute of Basic Sciences, Seoul National University, Seoul 08826, Republic of Korea

1. C. Darwin, *On the Origins of Species by Means of Natural Selection* (Murray, London, UK, 1859).
2. C. Darwin, On the dispersal of freshwater bivalves. *Nature* **25**, 529–530 (1882).
3. L. C. S. Lopez, P. J. F. P. Rodrigues, R. I. Rios, Frogs and snakes as phoretic dispersal agents of bromeliad ostracods (Limnocytheridae: Elpidium) and annelids (Naididae: Dero)? *Biotropica* **31**, 705–708 (1999).
4. H. Proctor, I. Owens, Mites and birds: Diversity, parasitism and coevolution. *Trends Ecol. Evol.* **15**, 358–364 (2000).
5. L. S. Saul-Gershenz, J. G. Millar, Phoretic nest parasites use sexual deception to obtain transport to their host's nest. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14039–14044 (2006).
6. J. Magis-Castillo *et al.*, Phoretic dispersal of armored scale crawlers (Hemiptera: Diaspididae). *J. Econ. Entomol.* **103**, 1172–1179 (2010).
7. P. S. White, L. Morran, J. de Roode, Phoresy. *Curr. Biol.* **27**, R578–R580 (2017).
8. A. W. Bartlow, S. J. Agosta, Phoresy in animals: Review and synthesis of a common but understudied mode of dispersal. *Biol. Rev. Camb. Philos. Soc.* **96**, 223–246 (2021).
9. H. Lee *et al.*, Nictation, a dispersal behavior of the nematode *Caenorhabditis elegans*, is regulated by IL2 neurons. *Nat. Neurosci.* **15**, 107–112 (2011).
10. D. Lee *et al.*, The genetic basis of natural variation in a phoretic behavior. *Nat. Commun.* **8**, 273 (2017).
11. K. Kiontke, W. Sudhaus, Ecology of *Caenorhabditis* species. *WormBook* **9**, 1–14 (2006).
12. R. C. Cassada, R. L. Russell, The dauerlarva, a post-embryonic developmental variant of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **46**, 326–342 (1975).
13. M. Klass, D. Hirsh, Non-ageing developmental variant of *Caenorhabditis elegans*. *Nature* **260**, 523–525 (1976).
14. P. S. Albert, D. L. Riddle, Developmental alterations in sensory neuroanatomy of the *Caenorhabditis elegans* dauer larva. *J. Comp. Neurol.* **219**, 461–481 (1983).
15. J. W. Golden, D. L. Riddle, The *Caenorhabditis elegans* dauer larva: Developmental effects of pheromone, food, and temperature. *Dev. Biol.* **102**, 368–378 (1984).
16. L. Frézal, M.-A. Félix, *C. elegans* outside the Petri dish. *eLife* **4**, e05849 (2015).
17. M.-A. Félix, C. Braendle, The natural history of *Caenorhabditis elegans*. *Curr. Biol.* **20**, R965–R969 (2010).
18. N. E. Schroeder *et al.*, Dauer-specific dendrite arborization in *C. elegans* is regulated by KPC-1/Furin. *Curr. Biol.* **23**, 1527–1535 (2013).
19. J. S. Lee *et al.*, FMRFamide-like peptides expand the behavioral repertoire of a densely connected nervous system. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E10726–E10735 (2017).
20. D. Lee, H. Lee, N. Kim, D. S. Lim, J. Lee, Regulation of a hitchhiking behavior by neuronal insulin and TGF- β signaling in the nematode *Caenorhabditis elegans*. *Biochem. Biophys. Res. Commun.* **484**, 323–330 (2017).
21. S. Ahn *et al.*, The *C. elegans* regulatory factor X (RFX) DAF-19M module: A shift from general ciliogenesis to cell-specific ciliary and behavioral specialization. *Cell Rep.* **39**, 110661 (2022).
22. B. Cockx *et al.*, Mass spectrometry-driven discovery of neuropeptides mediating nictation behavior of nematodes. *Mol. Cell. Proteomics* **22**, 100479 (2023).
23. T. A. Crombie *et al.*, CaenDR, the *Caenorhabditis* Natural Diversity Resource. *Nucleic Acids Res.* **52**, D850–D858 (2024).
24. S. J. Widmayer, K. S. Evans, S. Zdravljic, E. C. Andersen, Evaluating the power and limitations of genome-wide association studies in *Caenorhabditis elegans*. *G3* **12**, jkac114 (2022).
25. C. Diesh *et al.*, JBrowse 2: A modular genome browser with views of synteny and structural variation. *Genome Biol.* **24**, 74 (2023).
26. S. C. Harvey, A. Shorto, M. E. Viney, Quantitative genetic analysis of life-history traits of *Caenorhabditis elegans* in stressful environments. *BMC Evol. Biol.* **8**, 1–16 (2008).
27. S. C. Harvey, G. L. A. Barker, A. Shorto, M. E. Viney, Natural variation in gene expression in the early development of dauer larvae of *Caenorhabditis elegans*. *BMC Genomics* **10**, 1–11 (2009).
28. P. Lukacik *et al.*, Structural and biochemical characterization of human orphan DHRS10 reveals a novel cytosolic enzyme with steroid dehydrogenase activity. *Biochem. J.* **402**, 419–427 (2007).
29. D. L. Stern, Identification of loci that cause phenotypic variation in diverse species with the reciprocal hemizygosity test. *Trends Genet.* **30**, 547–554 (2014).
30. G. Zhang, N. M. Roberto, D. Lee, S. R. Hahnel, E. C. Andersen, The impact of species-wide gene expression variation on *Caenorhabditis elegans* complex traits. *Nat. Commun.* **13**, 3462 (2022).
31. J. G. White, E. Southgate, J. N. Thomson, S. Brenner, The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos. Trans. R. Soc. Lond. B, Biol. Sci.* **314**, 1–340 (1986).
32. G. Oikonomou, S. Shaham, The glia of *Caenorhabditis elegans*. *Glia* **59**, 1253–1263 (2011).
33. L. Meng, A. Zhang, Y. Jin, D. Yan, Regulation of neuronal axon specification by glia-neuron gap junctions in *C. elegans*. *eLife* **5**, e19510 (2016).
34. Z. F. Altun, D. H. Hall, Muscle system. GLR cells. *WormAtlas* (2009), 10.3098/wormatlas.1.9.
35. H. Yim *et al.*, Comparative connectomics of dauer reveals developmental plasticity. *Nat. Commun.* **15**, 1546 (2024).
36. T. A. Crombie *et al.*, Deep sampling of Hawaiian *Caenorhabditis elegans* reveals high genetic diversity and admixture with global populations. *eLife* **8**, e50465 (2019).
37. E. C. Andersen *et al.*, Chromosome-scale selective sweeps shape *Caenorhabditis elegans* genomic diversity. *Nat. Genet.* **44**, 285–290 (2012).
38. D. Lee *et al.*, Balancing selection maintains hyper-divergent haplotypes in *Caenorhabditis elegans*. *Nat. Ecol. Evol.* **5**, 794–807 (2021).
39. C. Pickering *et al.*, A genome-wide association study of sprint performance in elite youth football players. *J. Strength Cond. Res.* **33**, 2344–2351 (2019).
40. N. Fienelbach, A. Antebi, *C. elegans* dauer formation and the molecular basis of plasticity. *Genes Dev.* **22**, 2149–2165 (2008).
41. M. G. Zhang, P. W. Sternberg, Both entry to and exit from diapause arrest in *Caenorhabditis elegans* are regulated by a steroid hormone pathway. *Development* **149**, dev200173 (2022).
42. D. L. Motola *et al.*, Identification of ligands for DAF-12 that govern dauer formation and reproduction in *C. elegans*. *Cell* **124**, 1209–1223 (2006).
43. Y. Li, P. Haynes, S. L. Zhang, Z. Yue, A. Sehgal, Ecdysone acts through cortex glia to regulate sleep in *Drosophila*. *eLife* **12**, e81723 (2023).
44. L. Ju *et al.*, Hormonal gatekeeping via the blood-brain barrier governs caste-specific behavior in ants. *Cell* **186**, 4289–4309.e23 (2023).
45. J. Hodgkin, T. Doniach, Natural variation and copulatory plug formation in *Caenorhabditis elegans*. *Genetics* **146**, 149–164 (1997).
46. D. Lee *et al.*, Selection and gene flow shape niche-associated variation in pheromone response. *Nat. Ecol. Evol.* **3**, 1455–1463 (2019).
47. S. Brenner, The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
48. P. Y. Jeong *et al.*, Chemical structure and biological activity of the *Caenorhabditis elegans* dauer-inducing pheromone. *Nature* **433**, 541–545 (2005).
49. R. A. Butcher, M. Fujita, F. C. Schroeder, J. Clardy, Small-molecule pheromones that control dauer development in *Caenorhabditis elegans*. *Nat. Chem. Biol.* **3**, 420–422 (2007).
50. J. A. Arribere *et al.*, Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in *Caenorhabditis elegans*. *Genetics* **198**, 837–846 (2014).
51. E. Z. Macosko *et al.*, A hub-and-spoke circuit drives pheromone attraction and social behaviour in *C. elegans*. *Nature* **458**, 1171–1175 (2009).
52. C. C. Mello, J. M. Kramer, D. Stinchcomb, V. Ambros, Efficient gene transfer in *C. elegans* extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10**, 3959–3970 (1991).