ARTICLE

doi:10.1038/nature19848

Balancing selection shapes density-dependent foraging behaviour

Joshua S. Greene[1], Maximillian Brown[1], May Dobosiewicz[1], Itzel G. Ishida[1], Evan Z. Macosko[1], Xinxing Zhang[2], Rebecca A. Butcher[2], Devin J. Cline[3], Patrick T. McGrath[3]§ & Cornelia I. Bargmann[1]§

The optimal foraging strategy in a given environment depends on the number of competing individuals and their behavioural strategies. Little is known about the genes and neural circuits that integrate social information into foraging decisions. Here we show that ascaroside pheromones, small glycolipids that signal population density, suppress exploratory foraging in *Caenorhabditis elegans*, and that heritable variation in this behaviour generates alternative foraging strategies. We find that natural *C. elegans* isolates differ in their sensitivity to the potent ascaroside icas#9 (IC-asc-C5). A quantitative trait locus (QTL) regulating icas#9 sensitivity includes *srx-43*, a G-protein-coupled icas#9 receptor that acts in the ASI class of sensory neurons to suppress exploration. Two ancient haplotypes associated with this QTL confer competitive growth advantages that depend on ascaroside secretion, its detection by *srx-43* and the distribution of food. These results suggest that balancing selection at the *srx-43* locus generates alternative density-dependent behaviours, fulfilling a prediction of foraging game theory.

The success of a particular foraging strategy varies according to the behaviour of competitors. Balancing selection can therefore favour the co-existence of multiple strategies within a species[1,2]. The pioneering example of strategic competition is the natural genetic variation seen at the *foraging* (*for*) gene in *Drosophila melanogaster* larvae[3,4]. Two *for* alleles for active (rover) or sedentary (sitter) behaviour are maintained in a population because of frequency-dependent balancing selection against larvae with the more common foraging strategy[5]. This example, and others like it, suggests that animals could benefit from detecting and responding to competitors in real time and modifying foraging behaviour accordingly. However, little is known about the genes and neural circuits that incorporate information about conspecifics into foraging strategies.

Owing to the well-characterized nature of its foraging circuits and intraspecific pheromone system, the nematode *C. elegans* provides an opportunity to address this question. The *C. elegans* strategy for foraging bacterial food spontaneously alternates between an exploratory behaviour (roaming) and a less active behaviour (dwelling), with each persisting for several minutes per episode[6,7]. Transitions between roaming and dwelling are regulated by distributed neuromodulatory systems that link internal cues, such as nutritional status, to locomotion circuits[7,8]. *C. elegans* senses population density using a family of secreted pheromones called ascarosides, molecules that control the developmental decision to enter the starvation-resistant dauer larva stage[9] and regulate behaviours such as aggregation and male attraction to hermaphrodites[10,11]. We show here that physiological levels of certain ascarosides also regulate foraging by suppressing roaming behaviours. By characterizing differences in pheromone sensitivity in natural *C. elegans* isolates, we identify a pheromone receptor that shapes alternative foraging strategies and affects fitness depending on the structure of the food environment.

Pheromones regulate foraging behaviour

The effects of ascaroside pheromones on *C. elegans* foraging were examined by quantifying long-term exploration of a bacterial lawn by individual wild-type N2 strain animals (Fig. 1a). To mimic the effects of high population density on these isolated animals, we conducted the assay in the presence of natural pheromone extracts. The pheromones strongly suppressed exploration (Fig. 1b), as did several pure synthetic ascarosides at concentrations at or below those that induced dauer larva development (Fig. 1c, d). However, ascr#5, a potent regulator of dauer development, only weakly suppressed exploration (Fig. 1d). Thus a subset of ascarosides regulates foraging behaviour at biologically relevant concentrations.

The exploration assay is an indirect measure of the relative time *C. elegans* spends in roaming and dwelling states[7,8]. Quantitative behavioural analysis of video recordings showed that the potent ascarosides icas#9 and ascr#8 decreased the fraction of time spent

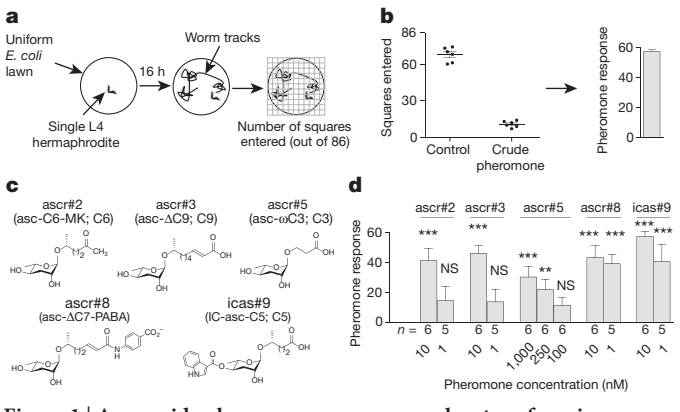


Figure 1 | Ascaroside pheromones suppress exploratory foraging behaviour. a, Schematic of the exploration assay. **b**, Wild-type N2 response to crude pheromone extract, showing exploration scores and a pheromone-response index, presented as mean ± s.e.m. **c**, Structures and names of selected ascarosides. **d**, N2 response to individual ascarosides, presented as mean ± s.e.m.. $^{**}P < 0.01$, $^{***}P < 0.001$ by ANOVA with Dunnett's correction; NS, not significant.

[1]Howard Hughes Medical Institute, Lulu and Anthony Wang Laboratory of Neural Circuits and Behavior, The Rockefeller University, New York, New York 10065, USA. [2]Department of Chemistry, University of Florida, Gainesville, Florida 32611, USA. [3]School of Biological Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332, USA.
§These authors jointly supervised this work.

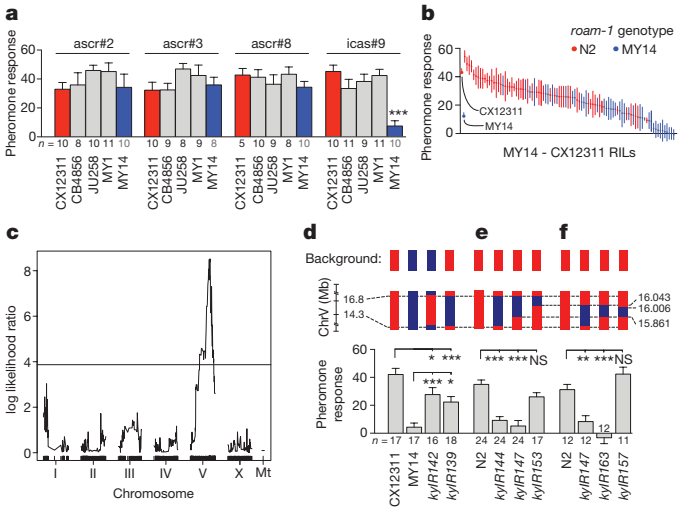


Figure 2 | Natural genetic variation in pheromone sensitivity. **a**, Response of wild-type *C. elegans* strains (*x* axis) to synthetic ascarosides. ***$P < 0.001$ by ANOVA with Dunnett's correction. **b**, icas#9 response of 94 CX12311-MY14 RILs and parental strains. **c**, QTL analysis of RILs shown in **b**. Horizontal line denotes $P < 0.05$ genome-wide significance threshold. Mt, mitochondrial. **d**–**f**, Mean icas#9 response of NILs used to map the *roam-1* locus, in MY14 and CX12311 (**d**) or N2 (**e**, **f**) backgrounds. **$P < 0.01$, ***$P < 0.001$ by ANOVA with Dunnett's correction. All data are presented as mean ± s.e.m.

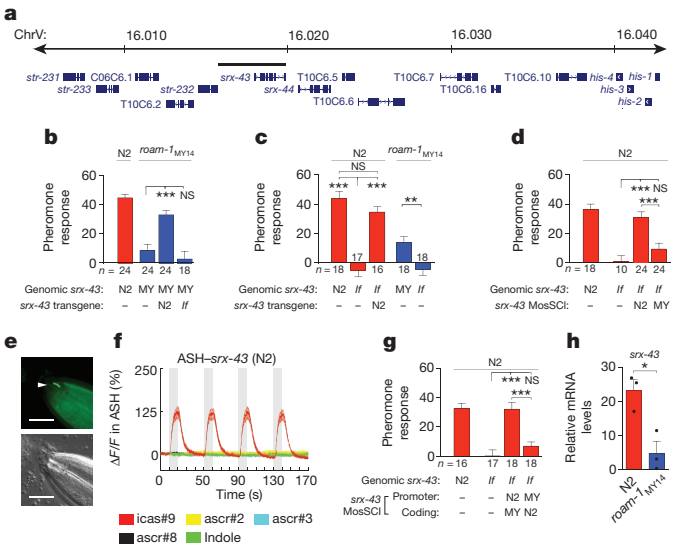


Figure 3 | The *roam-1* QTL includes the icas#9 receptor SRX-43. a, The *roam-1* locus. **b**–**d**, Sensitivity to icas#9, as affected by high-copy *srx-43* transgenes (**b**), *srx-43(lf)* mutations (**c**) or single-copy integrated *srx-43* transgenes (**d**). **e**, Image showing *srx-43::GFP* translational reporter (top) and Nomarski image (bottom). Arrowhead indicates an ASI sensory cilium. Scale bar, 10 μm. **f**, ASH calcium responses in *ASH::srx-43* transgenic strain. **g**, icas#9 sensitivity after exchanging promoters in single-copy *srx-43* transgenes. **h**, Endogenous *srx-43* mRNA levels ($n = 3$ assays each, indicated). All data are presented as mean ± s.e.m. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$ by ANOVA with Dunnett's correction or *t*-test.

roaming, the duration of roaming states and the speed of locomotion during roaming, but did not affect dwelling states (Extended Data Fig. 1).

A QTL for pheromone sensitivity

A variety of genetically diverse wild-type *C. elegans* strains, including the control N2-like strain CX12311, responded to the presence of ascarosides with a suppression of exploration[12] (Fig. 2a). The wild-type German strain MY14 failed to respond to 10 nM icas#9 in the exploration assay, although it responded normally to the presence of ascr#2, ascr#3, and ascr#8 (Fig. 2a and Extended Data Fig. 2). Coupled alterations in the levels of pheromone signalling and detection ability can contribute to reproductive isolation during incipient speciation[13]. However, MY14 and CX12311 were found to produce similar levels of icas#9 and another 16 ascarosides (Extended Data Fig. 3), indicating that the change in the icas#9 response in MY14 was independent of icas#9 production.

To determine the genetic basis of icas#9 insensitivity, 94 recombinant inbred lines (RILs) were generated from intercrosses between MY14 and CX12311 strains. A continuous distribution of icas#9 sensitivity was observed in the exploration behaviour of the RILs (Fig. 2b), suggesting that two or more loci contribute to icas#9 sensitivity. The 94 RILs were genotyped at ~185-kb resolution across the genome by low-coverage whole-genome sequencing[14] (Supplementary Table 1). QTL analysis identified a single significant QTL of genome-wide significance that accounted for 34.9% of the total variance between the RILs, which we term *roam-1* (Fig. 2c). Covariate analysis failed to find additional QTLs that were either additive or interactive with *roam-1* (Extended Data Fig. 4).

The impact of *roam-1* on foraging was confirmed by creating near-isogenic lines (NILs) that had small genetic regions substituted between the strains. NILs in which the 2.5 Mb surrounding *roam-1* were reciprocally exchanged between CX12311 and MY14 were intermediate in their icas#9 sensitivity compared to the parental strains (Fig. 2d). Depending on the direction of the introgression, the *roam-1* locus accounted for 39–46.4% of the genetic variance between the two parental strains.

To simplify further mapping, the *roam-1* region from MY14 (*kyIR139*) was crossed into the N2 laboratory strain. The resulting NIL (*kyIR144*, Fig. 2e) facilitated further mapping that localized *roam-1* to 182 kb (Fig. 2e). High-density mapping of 2,600 F2 progeny of a cross between N2 and the NIL *kyIR147* yielded 12 informative recombinants in this 182-kb region that mapped the *roam-1* QTL to a 37-kb region on chromosome V (Fig. 2f and Methods).

roam-1 affects the icas#9 receptor SRX-43

The 37-kb *roam-1* region contains 16 protein-coding genes, including 5 genes that encode predicted G-protein-coupled chemoreceptors in the *srx* or *str* gene families (Fig. 3a). We thought it possible that one or more of the chemoreceptors could be icas#9-receptors, with reduced activity in MY14. We therefore introduced N2-derived sequences that overlapped the chemoreceptor genes into the *roam-1*$_{MY14}$ NIL strain *kyIR163*. N2-derived transgenes that covered *srx-43* (Fig. 3a, bar) conferred sensitivity to icas#9 in *roam-1*$_{MY14}$ organisms (Fig. 3b), whereas transgenes with nonsense mutations disrupting the coding region of *srx-43* did not (Fig. 3b).

The function of *srx-43* was examined further by characterizing loss-of-function mutations in the endogenous *srx-43* gene (*srx-43(lf)*; Methods). We found that *srx-43(lf)* mutants in both N2 and *roam-1*$_{MY14}$ genetic backgrounds were insensitive to icas#9 (Fig. 3c). Sensitivity to icas#9 in the N2 *srx-43(lf)* mutant was restored by the introduction of an N2 *srx-43* transgene (Fig. 3c). These results indicate that *srx-43* is necessary for the icas#9 response in both N2 and MY14 strains and is essential for the behavioural difference between them.

The activity of the N2 and MY14 *srx-43* genes was compared by targeting a single copy of *srx-43* from each strain to a defined locus. This was done using the Mos1 transposase in an *srx-43(lf)* mutant, so that the single-copy transgene was the sole source of *srx-43*. The N2 *srx-43* genomic region fully rescued the icas#9 response, whereas the MY14 region did not (Fig. 3d). The different effects induced by the introduction of different single-copy transgenes indicate that MY14 *srx-43* possesses reduced activity compared to N2 *srx-43*.

Reporter genes with N2 or MY14 *srx-43* sequences driving expression of green fluorescent protein (GFP) were expressed selectively in the ASI sensory neurons (Extended Data Fig. 5a), which promote roaming

behaviour[8,15]. A genomic clone with GFP fused to the C terminus of the SRX-43 protein was enriched in ASI sensory cilia, the site of sensory transduction (Fig. 3e), suggesting that SRX-43 is a chemoreceptor. We investigated the effects of icas#9 on ASI activity using *in vivo* calcium imaging, but did not observe a response. This negative result is consistent with studies of dauer formation, in which ascarosides regulate gene expression in ASI neurons and not acute calcium levels[16–19].

To investigate whether SRX-43 was an icas#9 receptor, *srx-43* cDNA was expressed in the ASH class of sensory neurons, which are normally insensitive to ascarosides[17] (Extended Data Fig. 5b). Ascaroside-induced calcium flux was then monitored using genetically encoded calcium indicators[17,20]. ASH neurons expressing SRX-43 responded to 10 nM icas#9 with calcium transients, but this was not observed in response to other ascarosides or indole (Fig. 3f). Although the MY14 strain was largely insensitive to icas#9 in foraging assays, MY14 SRX-43 also detected icas#9 when expressed in ASH (Extended Data Fig. 5c).

We observed expression of a MY14-derived *srx-43::GFP* reporter gene in ASI neurons, but this appeared to be weaker than expression of the N2 *srx-43::GFP* reporter (Extended Data Fig. 5a). To investigate which sequences led to the difference in N2 and MY14 *srx-43* activity, we exchanged the *srx-43* promoter and coding regions of both strains and tested these constructs as Mos1-mediated single-copy insertion (MosSCI) *srx-43* transgenes. A transgene with the promoter region of the N2 strain and the coding region of MY14 rescued icas#9 sensitivity in *srx-43(lf)* mutants but the opposite did not, localizing the difference to the *srx-43* promoter (Fig. 3g). Quantitative measurements of endogenous *srx-43* mRNA levels demonstrated that *srx-43* was expressed at a level fivefold lower in *roam-1*$_{MY14}$ than in N2 (Fig. 3h). Therefore, the natural variation in *srx-43* promoter activity between N2 and MY14 affects *srx-43* gene expression and behavioural sensitivity to icas#9.

Ascarosides promote dauer larva development in part by suppressing the transcription of *daf-7*, a gene expressed in ASI neurons that encodes a secreted TGFβ-related peptide[19]. As with animals treated with ascarosides, *daf-7(lf)* mutants exhibit reduced levels of roaming[7]. Expression of a *daf-7*::*GFP* reporter was significantly reduced by treatment with icas#9, indicating that *daf-7* may be a target for icas#9; animals bearing a *daf-7* mutation were also less responsive to icas#9 than controls (Extended Data Fig. 5d, e). Behaviour was only influenced by icas#9 after several hours of exposure (Extended Data Fig. 5f), a delay that is in agreement with the slow transcriptional regulation of the *daf-7* signalling pathway. Together, these results suggest that icas#9 acts as a primer pheromone[21,22] that regulates foraging via transcription and endocrine signalling.

Balancing selection at a foraging QTL

To understand the population genetics of *roam-1*, we examined the genomic sequence of a 20-kb region centred around *srx-43* in 39 additional wild-type *C. elegans* isolates sequenced by the Million Mutation Project[23]. Two discrete, highly divergent haplotypes for the *roam-1* region were found: one resembling N2 that was present in 34 strains and another resembling MY14 that was present in 7 strains with different geographical origins and genetic backgrounds (Fig. 4a, b and Extended Data Fig. 6a). The N2 and MY14 haplotypes differed at 2.64% of all positions over the 20-kb *srx-43* region, a figure 12 times that of the genome-wide average[24]. These data were, however, derived using Illumina sequencing, a method that can fail to align highly divergent sequences. Targeted Sanger sequencing revealed that MY14 and N2 actually differed at 19.7% of all positions in the *srx-43* promoter and coding regions (Extended Data Fig. 6b). A phylogeny constructed for *srx-43* and the most closely related genes in *C. elegans*, *C. briggsae*, and *C. remanei* confirmed that the divergent *srx-43* alleles represent the same gene (Extended Data Fig. 6c). All seven tested MY14-like strains were relatively resistant to icas#9 compared to N2-like strains (Fig. 4c). These results suggest that naturally occurring resistance to

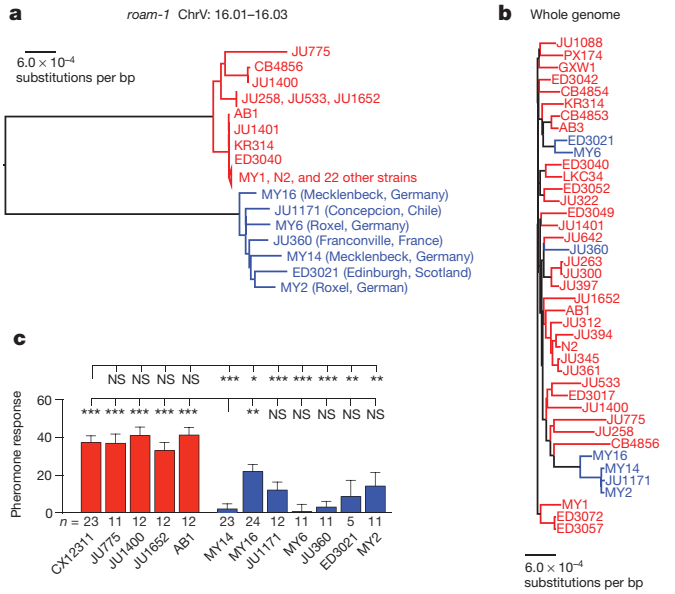


Figure 4 | Population genetics of the *roam-1* locus and icas#9 sensitivity. **a**, A dendrogram across 41 natural isolates representing a 20-kb region surrounding *srx-43*. The two major *roam-1* haplotypes from N2 and MY14 strains are indicated in red and blue, respectively. **b**, Whole-genome dendrogram for the same strains as **a**, showing relationships among strains. Red and blue colours follow *roam-1* haplotypes in **a**. **c**, Responses to icas#9 in natural isolates with the MY14 haplotype (blue) are consistently lower than those with the N2 haplotype (red). Pheromone response shown as mean ± s.e.m. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ by ANOVA with Dunnett's correction; NS, not significant.

icas#9 is associated with a highly divergent *roam-1* haplotype that includes *srx-43*.

The marked allelic divergence of the *roam-1* region, together with the observation that most genes in the interval including *srx-43* have a low d_N/d_S ratio (Extended Data Table 1), suggests that these alternative haplotypes might be subject to balancing selection. To examine this possibility, we analysed a large database of wild-type strains assembled and sequenced by the Andersen laboratory at Northwestern University (CeNDR; http://www.elegansvariation.org). The MY14 haplotype was present in 21 of the 152 unique strains, with the remainder being of the N2 haplotype. Both haplotypes were found globally: the rarer MY14 haplotype was found in Europe, the United States, New Zealand and Chile. In almost all cases in which individuals with the MY14 haplotype were isolated, individuals of the N2 haplotype were isolated from proximal environments at the same time—a distribution that is consistent with balancing selection.

Sequence features of the *roam-1* region also fit the criteria for a region under balancing selection. The region encompassing *roam-1* had a relatively high Tajima's *D* of 1.01, a figure unusual both at the genomic level (at which <3.4% of bins had a higher value) and in the centre of chromosome V where *srx-43* lies (at which <3.6% of bins had a higher value). A phylogenetic analysis of the 152 strains revealed that the *roam-1* haplotype extends by approximately 30 kb before being disrupted by recombination events (Extended Data Fig. 7). Given the low outcrossing rate in wild-type *C. elegans* populations[25,26], this short *roam-1* haplotype suggests co-occurrence of both haplotypes within interbreeding populations over many generations.

To assess directly the possibility that selection could act on *roam-1*, we designed competition experiments to compare the relative fitness of N2 and *roam-1*$_{MY14}$ strains. Experiments were conducted under high-density conditions to permit the accumulation and detection of endogenously produced icas#9, with competition applied by growing cultures past the point of starvation (that is, with limiting food). The first competition experiments were conducted on a simple lawn of *Escherichia coli* OP50 bacteria with a population of 20 N2 and 20

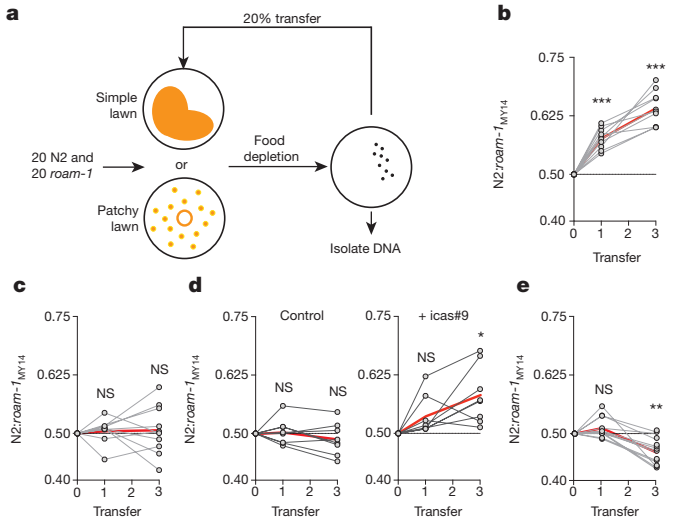


Figure 5 | Bidirectional competitive selection at the *roam-1* locus. **a**, Diagram of 'boom–bust' competition experiments, with food depletion followed by 48 h of starvation before transfer. **b**–**d**, Competition on a 'simple lawn' showing allele ratio of DNA harvested at transfers 1 and 3. N2 versus *roam-1*$_{MY14}$ NIL (**b**), N2 *srx-43(lf)* versus *roam-1*$_{MY14}$ *srx-43(lf)* (**c**) and N2 *daf-22(lf)* versus *roam-1*$_{MY14}$ *daf-22(lf)*, without or with 10 nM exogenous icas#9 (**d**). **e**, 'Patchy lawn' competition between N2 and *roam-1*$_{MY14}$ NIL. Grey points represent individual competition experiments; red line indicates the mean. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ compared to an expected value of 0.5 by *t*-test with Bonferroni correction.

roam-1$_{MY14}$ age-matched adults (Fig. 5a). These conditions resulted in a growth advantage for the N2 genotype over *roam-1*$_{MY14}$ in the first cycle of competition that continued in subsequent cycles (Fig. 5b).

The two tested strains differ in the 182-kb *roam-1* region, which encompasses 81 genes. To investigate whether this competitive advantage required *srx-43*, the experiment was repeated using N2 *srx-43(lf)* and *roam-1*$_{MY14}$ *srx-43(lf)* strains. The competitive N2 advantage disappeared in this setting, indicating that the icas#9 receptor SRX-43 is essential for the competitive effect (Fig. 5c).

We assessed the role of endogenous pheromones by repeating the competition experiments with N2 and *roam-1*$_{MY14}$ strains with loss-of-function mutations in the gene *daf-22*, which is required for the secretion of ascarosides including icas#9. Mutations in *daf-22* eliminated the competitive advantage of the N2 strain over *roam-1*$_{MY14}$ (Fig. 5d). A partial advantage for N2 *daf-22(lf)* was recovered upon addition of exogenous 10 nM icas#9 (Fig. 5d). These results thus indicate that selection on the *roam-1* locus depends on pheromones.

The increased roaming of *roam-1*$_{MY14}$ animals at high population densities could be expected to cause greater exploration of a patchy food environment. In a second competition design, we seeded a patchy environment consisting of 16 small bacterial lawns with 20 N2 and 20 *roam-1*$_{MY14}$ adults (Fig. 5a). In these conditions, the N2 advantage was lost and instead a moderate but significant selection favoured *roam-1*$_{MY14}$ over N2 animals (Fig. 5e). Together, these results demonstrate that *roam-1* has a bidirectional effect on fitness that is dependent on *srx-43* expression, pheromone production and food distribution.

Discussion

Conspecific individuals are informative elements of an animal's natural environment, in part because they represent competition for resources. Our results demonstrate that conspecific pheromones alter long-term foraging strategies and that natural variation in this behaviour stems from altered expression of the icas#9 receptor SRX-43. The complement of ascarosides produced by *C. elegans* varies with sex, age and feeding status[27,28]. The specificity of receptors such as SRX-43 provides a mechanism through which this information can be detected by the nervous system, allowing it to regulate different behaviours and physiological responses. *srx-43* is expressed in ASI sensory neurons, which are targets of internal neuromodulators that regulate roaming and dwelling[8], and represents a site of integration for internal and external influences on foraging behaviour. Although we do not know the suite of pheromones that are produced by *C. elegans* in the wild, the presence of secreted icas#9 in dense culture supernatants at concentrations 100-fold above those that suppress roaming suggest that it is a relevant regulator of foraging and that altered sensitivity to this molecule could affect animals' overall sensitivity to secreted ascarosides.

The *roam-1* QTL that encompasses *srx-43* has sequence features consistent with an area under balancing selection. While *srx-43* is an essential component of *roam-1*, it may not be the only gene in this QTL, or the only gene that is under balancing selection, as the haplotype extends for ~30 kb to include several other genes. Moreover, the behaviour identified here need not be the most important one in natural settings; it may represent one of several behavioural and physiological responses that facilitate adaptation to different environments. Balancing selection may be fairly common throughout the *C. elegans* genome: a recent report identified 61 highly divergent regions consistent with balancing selection that segregate among wild strains of *C. elegans*, including a second region 200 kb from *roam-1* (refs 24, 29; Extended Data Fig. 8). The composition of these regions is biased towards particular gene classes, including chemoreceptors, which may act as evolutionary hotspots.

At a conceptual level, behavioural genetics in animals, including humans, is dominated by evidence for gene–environment interactions[30,31]. Our results ground this abstraction in observation, showing that natural trait variation acts explicitly at the intersection of innate circuits and environment cues, with genetic changes allowing the differential incorporation of environmental information into innate foraging behaviours.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 27 March; accepted 14 September 2016.
Published online 31 October 2016.

1. Smith, J. M. The theory of games and the evolution of animal conflicts. *J. Theor. Biol.* **47,** 209–221 (1974).
2. Dugatkin, L. A. & Reeve, H. K. *Game Theory and Animal Behavior.* (Oxford Univ. Press, 1998).
3. Sokolowski, M. B., Pereira, H. S. & Hughes, K. Evolution of foraging behavior in *Drosophila* by density-dependent selection. *Proc. Natl Acad. Sci. USA* **94,** 7373–7377 (1997).
4. Osborne, K. A. *et al.* Natural behavior polymorphism due to a cGMP-dependent protein kinase of *Drosophila*. *Science* **277,** 834–836 (1997).
5. Fitzpatrick, M. J., Feder, E., Rowe, L. & Sokolowski, M. B. Maintaining a behaviour polymorphism by frequency-dependent selection on a single gene. *Nature* **447,** 210–212 (2007).
6. Fujiwara, M., Sengupta, P. & McIntire, S. L. Regulation of body size and behavioral state of *C. elegans* by sensory perception and the EGL-4 cGMP-dependent protein kinase. *Neuron* **36,** 1091–1102 (2002).
7. Ben Arous, J., Laffont, S. & Chatenay, D. Molecular and sensory basis of a food related two-state behavior in C. elegans. *PLoS One* **4,** e7584 (2009).
8. Flavell, S. W. *et al.* Serotonin and the neuropeptide PDF initiate and extend opposing behavioral states in *C. elegans*. *Cell* **154,** 1023–1035 (2013).
9. Jeong, P.-Y. *et al.* Chemical structure and biological activity of the *Caenorhabditis elegans* dauer-inducing pheromone. *Nature* **433,** 541–545 (2005).
10. Srinivasan, J. *et al.* A modular library of small molecule signals regulates social behaviors in *Caenorhabditis elegans*. *PLoS Biol.* **10,** e1001237 (2012).
11. Macosko, E. Z. *et al.* A hub-and-spoke circuit drives pheromone attraction and social behaviour in *C. elegans*. *Nature* **458,** 1171–1175 (2009).
12. McGrath, P. T. *et al.* Quantitative mapping of a digenic behavioral trait implicates globin variation in *C. elegans* sensory behaviors. *Neuron* **61,** 692–699 (2009).
13. Greenberg, A. J., Moran, J. R., Coyne, J. A. & Wu, C.-I. Ecological adaptation during incipient speciation revealed by precise gene replacement. *Science* **302,** 1754–1757 (2003).
14. Andolfatto, P. *et al.* Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome Res.* **21,** 610–617 (2011).

15. You, Y.-J., Kim, J., Raizen, D. M. & Avery, L. Insulin, cGMP, and TGF-β signals regulate food intake and quiescence in *C. elegans*: a model for satiety. *Cell Metab.* **7,** 249–257 (2008).
16. Kim, K. *et al.* Two chemoreceptors mediate developmental effects of dauer pheromone in *C. elegans*. *Science* **326,** 994–998 (2009).
17. McGrath, P. T. *et al.* Parallel evolution of domesticated *Caenorhabditis species* targets pheromone receptor genes. *Nature* **477,** 321–325 (2011).
18. Schackwitz, W. S., Inoue, T. & Thomas, J. H. Chemosensory neurons function in parallel to mediate a pheromone response in *C. elegans*. *Neuron* **17,** 719–728 (1996).
19. Ren, P. *et al.* Control of *C. elegans* larval development by neuronal expression of a TGF-β homolog. *Science* **274,** 1389–1391 (1996).
20. Akerboom, J. *et al.* Optimization of a GCaMP calcium indicator for neural activity imaging. *J. Neurosci.* **32,** 13819–13840 (2012).
21. Sorensen, P. W., Stacey, N. E. & Chamberlain, K. J. Differing behavioral and endocrinological effects of two female sex pheromones on male goldfish. *Horm. Behav.* **23,** 317–332 (1989).
22. Le Conte, Y. & Hefetz, A. Primer pheromones in social hymenoptera. *Annu. Rev. Entomol.* **53,** 523–542 (2008).
23. Thompson, O. *et al.* The million mutation project: a new approach to genetics in *Caenorhabditis elegans*. *Genome Res.* **23,** 1749–1762 (2013).
24. Thompson, O. A. *et al.* Remarkably divergent regions punctuate the genome assembly of the *Caenorhabditis elegans* Hawaiian strain CB4856. *Genetics* **200,** 975–989 (2015).
25. Barrière, A. & Félix, M.-A. High local genetic diversity and low outcrossing rate in *Caenorhabditis elegans* natural populations. *Curr. Biol.* **15,** 1176–1184 (2005).
26. Barrière, A. & Félix, M. A. Natural variation and population genetics of *Caenorhabditis elegans*. *WormBook* http://dx.doi.org/10.1895/wormbook.1.43.1 (2005).
27. Zhang, X. *et al.* Acyl-CoA oxidase complexes control the chemical message produced by *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **112,** 3955–3960 (2015).
28. Artyukhin, A. B. *et al.* Succinylated octopamine ascarosides and a new pathway of biogenic amine metabolism in *Caenorhabditis elegans*. *J. Biol. Chem.* **288,** 18778–18783 (2013).
29. Ghosh, R., Andersen, E. C., Shapiro, J. A., Gerke, J. P. & Kruglyak, L. Natural variation in a chloride channel subunit confers avermectin resistance in *C. elegans*. *Science* **335,** 574–578 (2012).
30. Bendesky, A. & Bargmann, C. I. Genetic contributions to behavioural diversity at the gene–environment interface. *Nat. Rev. Genet.* **12,** 809–820 (2011).
31. Kendler, K. S. *et al.* Stressful life events, genetic liability, and onset of an episode of major depression in women. *Am. J. Psychiatry* **152,** 833–842 (1995).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank E. Andersen for sharing the unpublished CeNDR database; the *Caenorhabditis* Genetics Center (NIH P40 OD010440) and the Million Mutation Project for strains; S. Flavell and A. Lopez for advice and insight; and P. Sengupta, M. O'Donnell, X. Jin and A. Sordillo for comments on the manuscript. R.A.B. and X.Z. were supported by the Research Corporation for Science Advancement (Cottrell Scholar Award, 22844). P.T.M. was supported by NIH grant R01GM114170 and the Ellison Medical Foundation. J.S.G. was supported by the NIH grant F30 MH101931-03. C.I.B. is an investigator of the HHMI. This work was supported by the Ellison Medical Foundation.

Author Contributions J.S.G. designed and performed the genetic, molecular biology, and behavioural experiments, together with M.B. for RIL analyses and I.G.I. for competition experiments. M.D. performed calcium imaging experiments. E.Z.M. discovered the effect of pheromones on foraging. X.Z. and R.A.B. analysed pheromone production and synthesized pure pheromones. D.J.C. and P.T.M. performed population genetic analysis. J.S.G, P.T.M. and C.I.B. analysed and interpreted data. J.S.G. and C.I.B. wrote the manuscript, with input from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.I.B. (cori@rockefeller.edu).

Reviewer Information *Nature* thanks E. Haag, S. Lockery and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Nematode culture. All strains were grown at 21–22 °C on nematode growth-medium plates seeded with *E. coli* OP50 bacteria[32]. For OP50 cultures, a single colony was inoculated into 100 ml of LB and grown for 48 h at 21–22 °C. Transgenic lines were generated by standard injection methods and included the desired transgene, a fluorescent co-injection marker and an empty vector, bringing the total DNA concentration up to 100 ng μl^{-1}. For each transgene, three independent extrachromosomal lines that propagated the transgene at high rates were tested in parallel to account for variability typical of such strains. All mutagenized strains were back-crossed 5–7 times before characterization.

Natural isolates and origin of wild-type strains. N2 (Bristol, UK), CX12311 *kyIR1* (chrV,CB4856 > N2) V; *qqIR1* (chrX, CB4856 > N2), CB4856 (HW) (Hawaii, USA), JU258 (Riberio Frio, Madeira), MY1 (Lingen, Germany), MY14 (Mecklenbeck, Germany), JU775 (Lisbon, Portugal), JU1400 (Sevilla, Spain), JU1652 (Montevideo, Uruguay), AB1 (Adelaide, Australia), MY16 (Mecklenbeck, Germany), JU1171 (Concepcion, Chile), MY6 (Roxel, Germany), JU360 (Franconville, France), ED3021 (Edinburgh, Scotland), MY2 (Roxel, Germany).

Strain CX12311 bears ancestral alleles of the *npr-1* and *glb-5* genes, which affect oxygen sensitivity and are mutated in the N2 laboratory strain[17]; it is therefore used as a comparison strain for wild-type strains bearing the ancestral alleles.

MY14–CX12311 RILs. CX14697-CX14712, CX14731-CX14748, CX14750-CX14757, CX14783, CX14784, CX14786-CX14820, CX14822-CX14839. Genotypes inferred from low-coverage genomic sequence and behavioural data are included as Supplementary Table 1.

Near-isogenic lines. CX15881 *kyIR142* (chrV:∼14.3–16.8 Mb, CX12311 > MY14), CX15878 *kyIR139* (chrV:∼14.3–16.8 Mb, MY14 > CX12311), CX15883 *kyIR144* (chrV:∼14.3–16.8 Mb, MY14 > N2), CX16075 *kyIR147* (chrV:∼15.861–16.8 Mb, MY14 > N2), CX16140 *kyIR153* (chrV:∼16.043–16.8 Mb, MY14 > N2), CX16300 *kyIR163* (chrV:∼15.861–16.043 Mb, MY14 > N2; *roam-1*$_{MY14}$), CX16294 *kyIR157* (chrV:∼15.861–16.006 Mb, MY14 > N2).

Transgenic lines. DNAs are N2-derived unless otherwise noted. CX16884 *kyIR163 V; kyEx5851* (*Psrx-43*^{N2}*::srx-43*^{N2}*::sl2::GFP*, 2.5 ng μl^{-1}; *Pmyo3::mcherry*, 5 ng μl^{-1}), CX17202 *kyIR163 V; kyEx6012* (*Psrx-43*^{N2}*::srx-43*^{N2}*(nonsense)::sl2::GFP*, 2.5 ng μl^{-1}, *Pmyo3::mcherry*, 5 ng μl^{-1}), CX16881 *srx-43(gk922634)* V; *kyEx5848* (*srx-43*^{N2}, 2.5 ng μl^{-1}, *Pmyo3::mcherry*, 5 ng μl^{-1}); *gk922634* changes R160 to an opal stop codon, CX17204 *kyEx6013* (*Psrx-43*^{N2}*::srx-43*^{N2}*::GFP*, 50 ng μl^{-1}, *Pelt-2::GFP*, 5 ng μl^{-1}), CX16943 *kyIR163* V; *kyEx5894* (*Psrx-43*^{MY14}*::srx-43*^{MY14}*::sl2::GFP*, 2.5 ng μl^{-1}, *Pmyo3*::mcherry, 5 ng μl^{-1}), CX16425 *kyIs602* (*Psra-6::GCaMP3.0*, 75 ng μl^{-1}; *Pcoel::GFP*, 10 ng μl^{-1}); *kyEx5594* (*Psra-6::srx-43*^{N2}, 50 ng μl^{-1}, *Pmyo3::mcherry*, 5 ng μl^{-1}), CX16931 *kyIs602*; *kyEx5885* (*Psra-6::srx-43*^{MY14}, 50 ng μl^{-1}, *Pmyo3::mcherry*, 5 ng μl^{-1}), CX17196 *kySi66* (*MosSCI Psrx-43*^{N2}*::srx-43*^{N2}*) II; srx-43*(*gk922634*) *V*, outcrossed 4×, CX17198 *kySi68* (*MosSCI Psrx-43*^{MY14}*::srx-43*^{MY14}) *II; srx-43*(*gk922634*) *V*, outcrossed 4×, CX17201 *kySi71* (*MosSCI Psrx-43*^{N2}*::srx-43*^{MY14}) *II; srx-43*(*gk922634*) *V*, outcrossed 4×, CX17203 *kySi72* (*MosSCI Psrx-43*^{MY14}*::srx-43*^{N2}) *II; srx-43*(*gk922634*) *V*, outcrossed 4×, FK181 *ksIs2* (*Pdaf-7::GFP + rol-6*(*su1006*)), CX16958 *kyIR163 V; ksIs2*

Mutants. CX16849 *srx-43*(*gk922634*) *V*, outcrossed 5× to N2. *gk922634* is R160opal. This mutation was provided by the Million Mutation Project[23].

CX16935 *kyIR163 srx-43(ky1019) V*. *ky1019* is a CRISPR/Cas9-induced indel mutation that causes a frame-shift mutation after the first transmembrane domain (insertion (TCACTGAGTTCGAAT), deletion (CCCCG), final sequence TCGCAGCTCTCAAGTTCACTGAGTTCGAATTTCGGAATTCTC; insertion is underlined). We used the coCRISPR protocol described previously[33]. Young adults were injected with a mix of plasmids containing Cas9, guide RNA targeting *rol-6*, and guide RNA targeting the location of the desired mutation, as well as a ssDNA template for inducing a dominant *rol-6(su1006)* mutation. F1 animals with a roller phenotype were isolated and allowed to lay eggs before secondary screening for the target mutation by Sanger sequencing.

Previously described mutants included JT5464 *daf-7(e1372) III; daf-3(e1376) X*, CX17307 *daf-7(e1372) III; kyIR163 V; daf-3(e1376) X*, CX13846 *daf-22(ok693) II*, CX17082 *daf-22(ok693) II*; *kyIR163 V*.

Behavioural analysis. Exploration assays[8] were conducted on 35-mm Petri dishes evenly seeded with 100 µl of OP50 *E. coli* bacteria 24 h before the start of the assay. Individual two-day-old L4 hermaphrodites were placed in the centre of the plate. After 16 h, plates were placed on a grid containing 86 squares and the number of full or partial squares that contained tracks was quantified by an investigator blinded to the genotype. Pheromones or control solvent were mixed into the agar. A pheromone response for each animal on an ascaroside plate was determined with respect to the behaviour of control animals that were tested on ascaroside-free plates on the same day.

Individual pheromone response was defined as the mean number of squares entered by controls tested on the same day, subtracted by the number of squares entered by the test animal. Group pheromone response was defined as the mean pheromone response of all individuals tested across days. For statistical analysis, *n* was defined as the total number of animals in the ascaroside group.

N2-derived strains were tested in 21% oxygen (Figs 1, 2e, f, 3, 5 and Extended Data Figs 1, 5). All naturally isolated strains and CX12311-derived strains bearing ancestral alleles of *npr-1* and *glb-5* were tested in 8% oxygen to suppress the oxygen-dependent roaming behaviour of ancestral *npr-1* alleles[12,34] (Figs 2a, b, d, 4 and Extended Data Fig. 2),

Direct examination of roaming and dwelling was modified from previously reported techniques[8]. At 14.5 h before testing, 25 L4 larvae were picked and placed onto 150-mm test plates thinly seeded with 1.5 ml of OP50 bacteria with or without synthetic pheromone. Video recording was conducted under red light to minimize behavioural response to imaging conditions. 1.5-h-long videos were recorded at 3 frames s^{-1} using Streampix software (Norpix Inc.) and a 6.6 MP PL-B781F CMOS camera (PixeLINK.). Custom Matlab scripts[8] were employed to determine worm trajectories and conduct a two-state hidden Markov model determining the most probable state path for each animal and thereby measure roaming- and dwelling-state durations.

The low basal exploration rate in *daf-7(lf)* mutants[7] prevented a direct assessment of the effect of icas#9 on foraging behaviour. Instead, we examined *daf-7(lf) daf-3(lf)* double mutants, as *daf-7* canonically acts by antagonizing *daf-3*, which encodes a co-SMAD transcriptional regulator. We found that *daf-3* mutations suppressed the low basal exploration rate of *daf-7* mutants. N2 *daf-7(lf) daf-3(lf)* animals explored control plates to a greater degree than the wild type, so larger (10 cm) exploration assay plates were used to score these strains.

Statistics. No statistical methods were used to predetermine sample size.

Most experiments were repeated on three separate days. For exploration assays, the standard group size on a single day was six; this ensured sufficient power to detect moderate effects, while also limiting the influence of daily variation. All plates with a healthy adult animal at the end of the assay were scored and included in the analysis. Randomization was ensured using the following or similar approach: at the start of each exploration assay, six animals were placed on a pick at a time. They were then transferred individually to three control plates and then to three icas#9 plates in the order the animals came off the pick. Assays were scored by an experimenter blinded to the condition or genotype.

Most statistical comparisons were performed using ANOVA with Dunnett's correction for multiple comparisons or using a (two-sided) *t*-test, as noted in the figure legends. The normality of the data was tested with D'Agostino–Pearson omnibus test. Bartlett's test was used to check for differences in variance between groups being statistically compared. N2 groups in Fig. 3c and Extended Data Figure 5e did not pass a normality test. As the *n* number was large (>15), ANOVA was still an appropriate test. Moreover, the findings were still significant when a nonparametric test was used (Kruskal–Wallis with Dunn multiple comparison test).

Ascaroside quantification. We grew 150 ml unsynchronized worm cultures for 9 days and fed on *E. coli* (HB101 or OP50), as described[27]. Extracts were generated from the culture medium and analysed by liquid chromatography–tandem mass spectroscopy (LC–MS/MS), as described previously[27], and analysed on a Thermo Scientific TSQ Quantum Access MAX, with the collision gas pressure set to 1 mTorr. Ascaroside concentrations present in the culture were quantified using the corresponding synthetic standards, with the following exceptions: synthetic ascr#18 was used to quantify ascr#22 and ascr#26 and synthetic icas#3 was used to quantify icas#1 and icas#10.

Recombinant inbred lines. The MY14–CX12311 recombinant inbred lines were generated by crossing MY14 males with CX12311 hermaphrodites and CX12311 males with MY14 hermaphrodites, to ensure the mitochondrial DNA from both strains were equally represented in the RILs. In total, 94 F2 animals were individually picked, placed onto plates and inbred through self-fertilization for 10 generations. RIL genotyping was conducted by low-coverage shotgun sequencing[14]. Genomic DNA was fragmented and attached to sequencing adapters with a Nextera DNA Library Prep Kit (Illumina). Samples were pooled and sequenced on an Illumina HiSeq 2000. Sequencing reads from each strain were mapped to the WS235 release of the *C. elegans* genome using the Burrows–Wheeler Aligner to create bam files for further analysis[35]. The set of MY14/N2 single nucleotide variants (SNVs) identified in the Million Mutation Project were used for genotyping purposes[23]. Each genetic variant was genotyped in each strain. Owing to the low coverage, the majority of SNVs were not genotyped. To improve the data coverage, we grouped 200 neighbouring SNV genotypes together to create a consensus genotype for 540 bins (either N2, MY14 or heterozygous). These genotypes were used for QTL mapping.

QTL mapping. The pheromone response index was used as the phenotype in combination with the 540 genotype bins from above. R/qtl was used to perform a one-dimensional scan using marker regression on all 540 markers. The significance

threshold was determined using 1,000 permutation tests. The effect-size of the *roam-1* locus was estimated using the fitqtl function with a single QTL. The peak of the *roam-1* locus (chromosome V: 16,451,686–16,579,457) was used as an additive and interactive covariate for additional one-dimensional scans, assuming a normal model. The significance threshold for these two tests was also determined using 1,000 permutation tests.

NIL mapping. Before the detailed QTL mapping by sequencing described above, the *roam-1* QTL was localized to 2.5 Mb (chrV: 14.3–16.8 Mb) by examining 14 high-confidence phenotypically extreme RILs (Supplementary Table 1). This result, which was confirmed by the full analysis, guided the initial generation of NILs. The NIL *kyIR142* was produced by backcrossing the RIL CX14816 nine times to MY14, maintaining N2 alleles at chrV: 14.3 and chrV: 16.8 Mb at each generation. The NIL *kyIR139* was produced by backcrossing the RIL CX14708 nine times to CX12311, maintaining MY14 alleles at chrV: 14.3 and chrV: 16.8 Mb. The NIL: *kyIR144* was produced by crossing *kyIR139* with N2 and isolating recombinants with the N2 allele of *glb-5* (chrV: 5.56 Mb), the MY14 alleles at chrV:14.3 and 16.8 Mb, and the N2 allele of *npr-1* on chrX. The NILs *kyIR147* and *kyIR153* were created by crossing *kyIR144* with N2 and identifying progeny with the N2 allele at chrV: 14.3 Mb and the MY14 allele at chrV: 16.8 Mb.

High-density recombination mapping. We crossed *kyIR147* with males from CX16290, a N2 strain with an integrated fluorescent marker at chrV: 15.83 Mb. F1 progeny were identified by fluorescence, picked to growth plates, and allowed to lay eggs for 12 h. Following 3 days of growth, ~2,600 non-fluorescent F2 animals were sorted individually into wells of 96-well plates by a worm sorter (COPAS Biosort Systems, Union Biometrica). These F2 were grown in 200 μl of S Basal buffer (5.85 g NaCl, 1 g K_2 HPO_4, 6 g KH_2PO_4, 5 mg cholesterol per litre) with cholesterol, supplemented with OP50 bacteria. A fraction of the F3 progeny from each isolate were lysed and genotyped at chrV: 16.069 Mb. Those with an N2 allele at chrV: 16.069 Mb were genotyped at chrV: 15.861 Mb. Twelve recombinants with an N2 allele at chrV: 16.069 Mb and a MY14 allele at chrV: 15.861 Mb were isolated and characterized behaviourally, among which were *kyIR163* and *kyIR157* (Fig. 2f). The N2 NIL with *kyIR163* (182 kb of MY14 sequence) is referred to as *roam-1*$_{MY14}$.

Imaging. Calcium imaging experiments were performed and analysed as described previously[36]. In brief, young adult animals were placed into custom-made 3 mm^2 microfluidic polydimethylsiloxane devices that permit rapid changes in stimulus solution. Each device contains two arenas, allowing for simultaneous imaging of two genotypes with approximately ten animals each. Animals were transferred to the arenas in S-Basal buffer and paralyzed for 80–100 min in 1 mM (−)-tetramisole hydrochloride. Experiments consisted of four pulses of 10 s of stimulus separated by 30 s of buffer, with an additional 60 s between stimulus types. Tiff stacks were acquired at 10 frames s^{-1} at 5× magnification (Hamamatsu Orca Flash 4 sCMOS), with 10 ms pulsed illumination every 100 ms (Sola, Lumencor; 470/40 nm excitation).

Fluorescence levels were analysed using a custom ImageJ script that integrates and background-subtracts fluorescence levels of the ASH neuron cell body (4 × 4 pixel region of interest). Using MATLAB, the calcium responses were normalized for each stimulus type by dividing fluorescence levels by the baseline fluorescence, defined as the average fluorescence of the 10 s preceding the first pulse of the stimulus. Each experiment was performed a total of four times over two separate days. Animals were pooled together by strain to calculate population mean and standard error (N2 *srx-43* allele, 23 animals; MY14 *srx-43* allele, 30 animals; array negative control, 19 animals). Experiments were conducted on two days.

For GFP expression studies, live adult animals were mounted on 2% agarose pads containing 5 mM sodium azide. Images were collected with a 100× objective on a Zeiss Axio Imager.Z1 Apotome microscope with a Zeiss AxioCam MRm CCD camera. For *daf-7* reporter studies, expression was quantified 16–24 h after L4 animals were placed on exploration assay plates. Images were processed in Metamorph and ImageJ to generate a maximum-intensity *Z*-projection. Reporter values were assessed as the mean grey value for a 16-pixel-radius circle centred over the cell body minus the mean background intensity. Both ASI neurons were analysed in each animal; experiments were performed over three days.

Digital PCR. Digital PCR was conducted on a QuantStudio 3D digital PCR platform (Thermo Fisher), and analysed on the QuantStudio 3D AnalysisSuite Cloud.

The *srx-43* mRNA expression studies were conducted on synchronized L4 worms 48 h after laying. RNA was collected on RNeasy Mini columns (Qiagen) and treated with DNase (Qiagen). SuperScript III First-Strand Synthesis System (Thermo Fisher) was used to create cDNA libraries. Custom TaqMan Expression Assays (Thermo Fisher) were used for *srx-43* quantification, and the tubulin gene, *tbb-1*, was used for normalization of digital PCR.

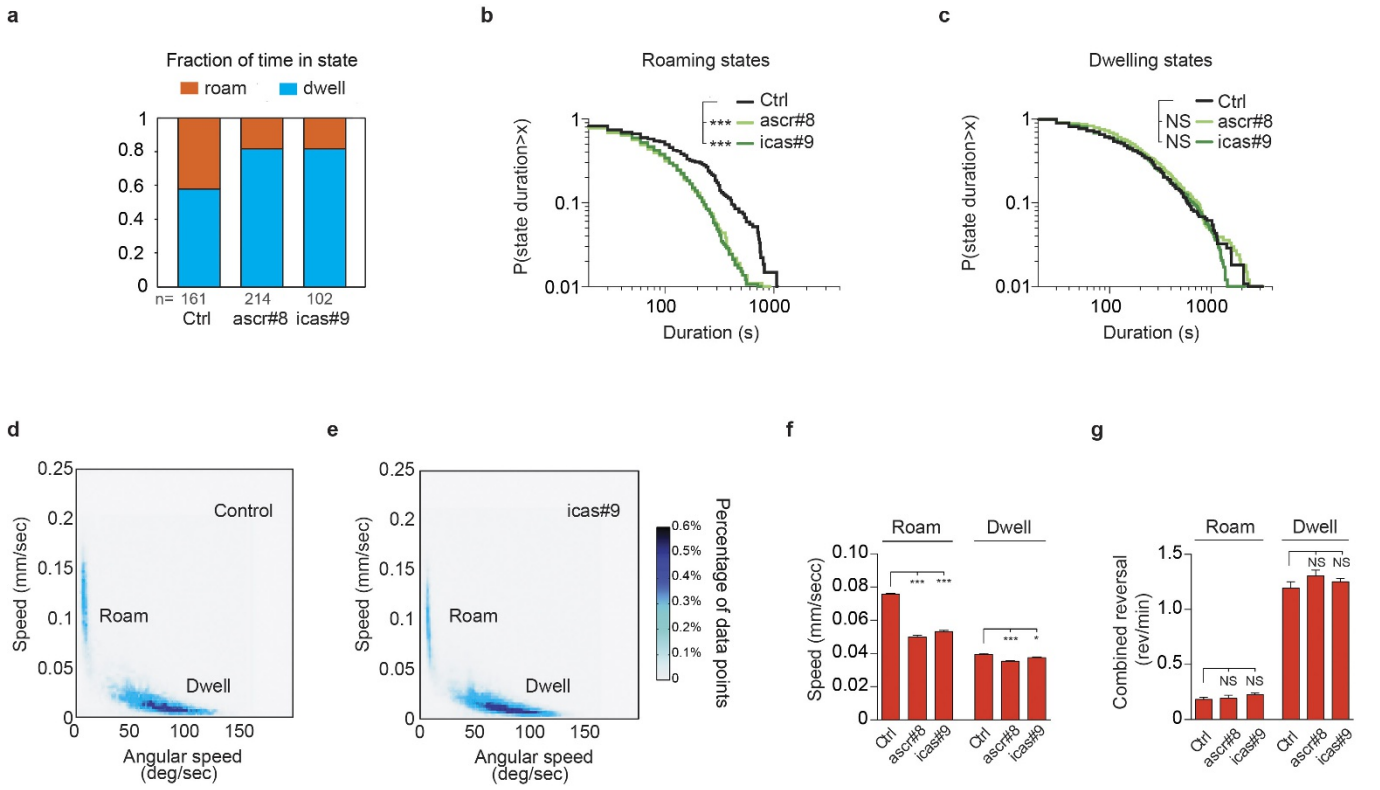
For quantitative analysis of the competition experiments, DNA was extracted with a standard phenol–chloroform protocol. Custom TaqMan SNP Genotyping Assays (Thermo Fisher) were used to determine the relative ratio of N2 versus *roam-1*$_{MY14}$ DNA by digital PCR. The assay was validated with known ratios of N2 to *roam-1*$_{MY14}$ DNA (Extended Data Fig. 9).

Population genetics. To create the gene and organism phylogenies, we used SNV data downloaded from the Million Mutation Project (http://genome.sfu.ca/mmp/) or the CeNDR resource (http://www.elegansvariation.org). For the CeNDR dataset, MY14 was assumed to be clonal or near-clonal with MY23, as was suggested by RAD sequencing. Software was written in Python using the Biopython module to create a neighbour joining tree. For the *roam-1* locus, SNVs on chrV between 16,010,000 and 16,030,000 were used. For the *glc-1* locus, SNVs on chrV between 16,181,000 and 16,222,000 were used. All SNVs were used to construct the whole-genome strain tree. Number of genetic variants and Tajima's *D* were calculated on 5-kb bins using vcftools[37]. d_N/d_S was calculated by counting using custom Python scripts analysing variants between MY23 and the N2 reference. Phylogenies of *srx-43* and closely related genes were performed using protein sequences obtained previously[38].

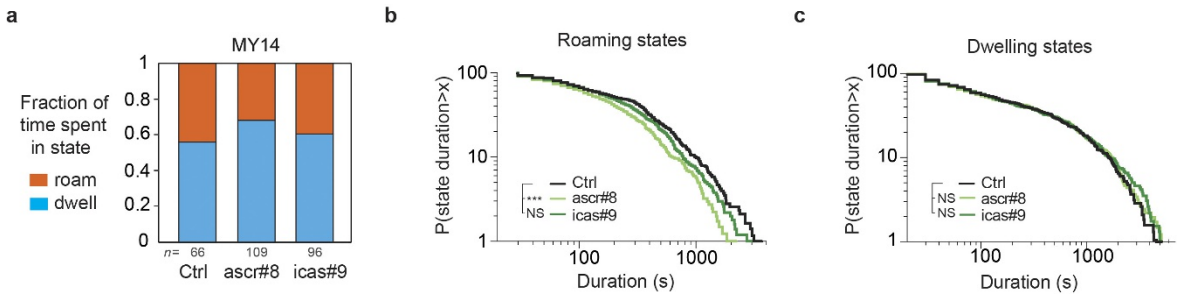
Fitness assays. Competition experiments consisted of three boom–bust cycles. During the boom phase, population growth led to the rapid depletion of food, initiating the bust phase, which lasted for two days. Simple lawn competition experiments were conducted on 100-mm NGM agar plates with a single lawn formed from 800 μl of saturated OP50 culture. Patchy lawn competition experiments were conducted on 150 mm NGM agar plates with a 200 μl ring-shaped OP50 lawn in the centre of the plate surrounded by 15 small 40-μl lawns (Fig. 5a); at the assay start and at transfers animals were placed in the centre of the plate.

Populations were initiated from 20 N2-type and 20 *roam-1*$_{MY14}$-type age-synchronized young adult animals. The initial population depleted food within 4 days, and on day 6 animals were washed into M9 media. 20% of the suspension was transferred to a new plate and the remainder was lysed for quantitative DNA analysis. For the second and third boom–bust cycle, food resources were depleted in 2 days and the plates were kept starved for an additional 2 days. Following the second bust phase, 20% of the animals were transferred to a new plate; following the third bust phase, the entire population was harvested for DNA extraction.

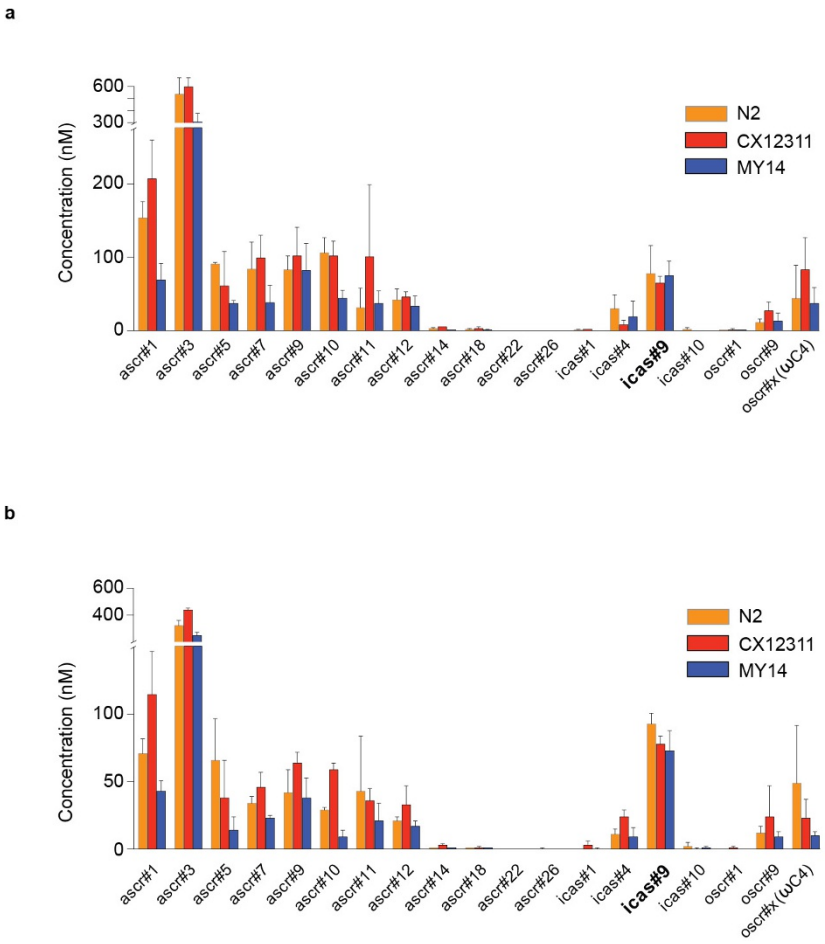
32. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77,** 71–94 (1974).
33. Arribere, J. A. *et al.* Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in *Caenorhabditis elegans*. *Genetics* **198,** 837–846 (2014).
34. Cheung, B. H. H., Cohen, M., Rogers, C., Albayram, O. & de Bono, M. Experience-dependent modulation of *C. elegans* behavior by ambient oxygen. *Curr. Biol.* **15,** 905–917 (2005).
35. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26,** 589–595 (2010).
36. Larsch, J., Ventimiglia, D., Bargmann, C. I. & Albrecht, D. R. High-throughput imaging of neuronal activity in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **110,** E4266–E4273 (2013).
37. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27,** 2156–2158 (2011).
38. Thomas, J. H. & Robertson, H. M. The *Caenorhabditis* chemoreceptor gene families. *BMC Biol.* **6,** 42 (2008).



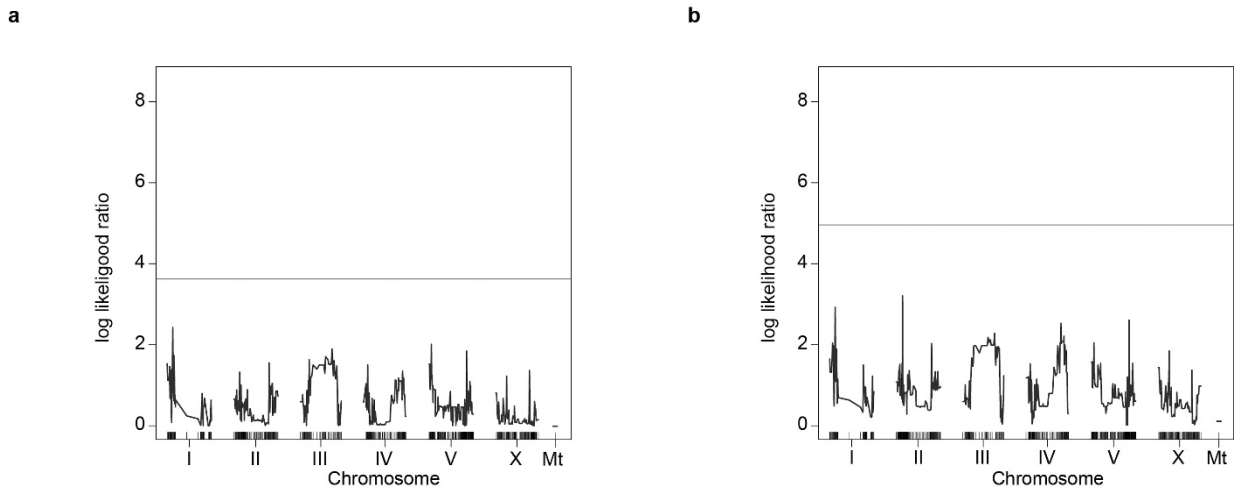
Extended Data Figure 1 | Roaming and dwelling states in the presence of ascarosides. a, Roaming and dwelling behaviours scored from video analysis. $n = 102–214$ tracks per data point. **b**, **c**, Cumulative distribution of roaming (**b**) and dwelling (**c**) state durations for animals in **a**. ***$P < 0.001$ by log-rank test; ns, not significant. **d**, **e**, Scatter plot of average speed and angular speed (a measure of turning rate) in 10 s intervals taken from 1.5-h-long video recordings of wild-type animals in control (**d**) and icas#9 (**e**) conditions. Roaming animals move quickly and turn infrequently compared with dwelling animals. Note the bimodal distribution defining distinct behavioural states. Control, 161 tracks; icas#9, 102 tracks. **f**, **g**, Speed following a reversal (**f**) and the reversal rate (**g**) for roaming or dwelling animals. Roaming speed is slightly slower in ascarosides (**e**, **f**). Data presented as mean ± s.e.m. *$P < 0.05$, ***$P < 0.001$ by ANOVA with Dunnett correction; ns, not significant.



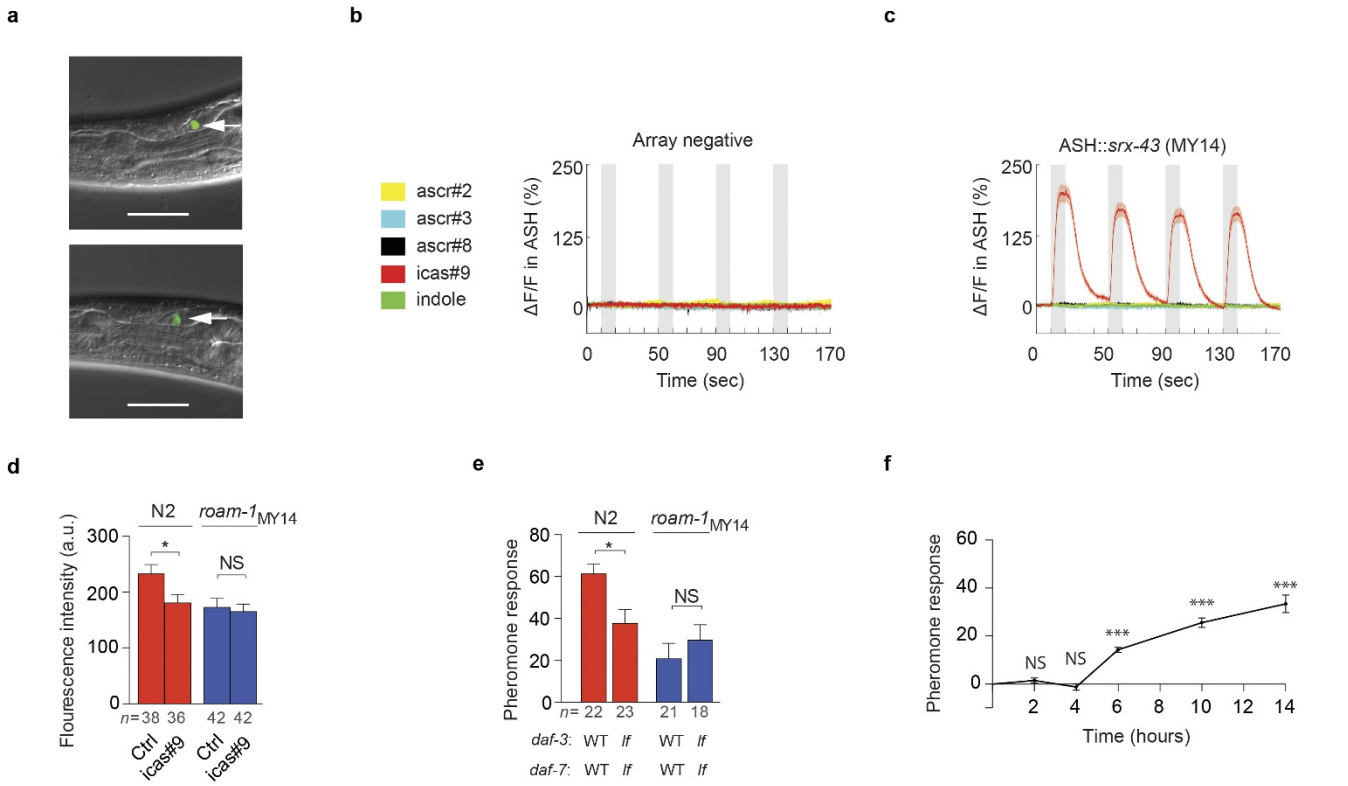
Extended Data Figure 2 | Roaming and dwelling behaviour of MY14. **a**, Fraction of time that MY14 animals spend roaming or dwelling in control, ascr#8, and icas#9 conditions; $n = 66–109$ tracks per data point. Assays were conducted in 8% O_2. **b**, **c**, Cumulative distribution of roaming (**b**) and dwelling (**c**) state durations for MY14 animals scored in **a**. Roaming states are significantly shorter in the presence of ascr#8 ($t½ = \sim 150$ s, versus ~ 220 s in controls), but are not significantly affected by icas#9 ($t½ = \sim 190$ s). Roaming states may also be longer at baseline in MY14 than in N2 (see Extended Data Fig. 1). $^{***}P < 0.001$ by log-rank test.



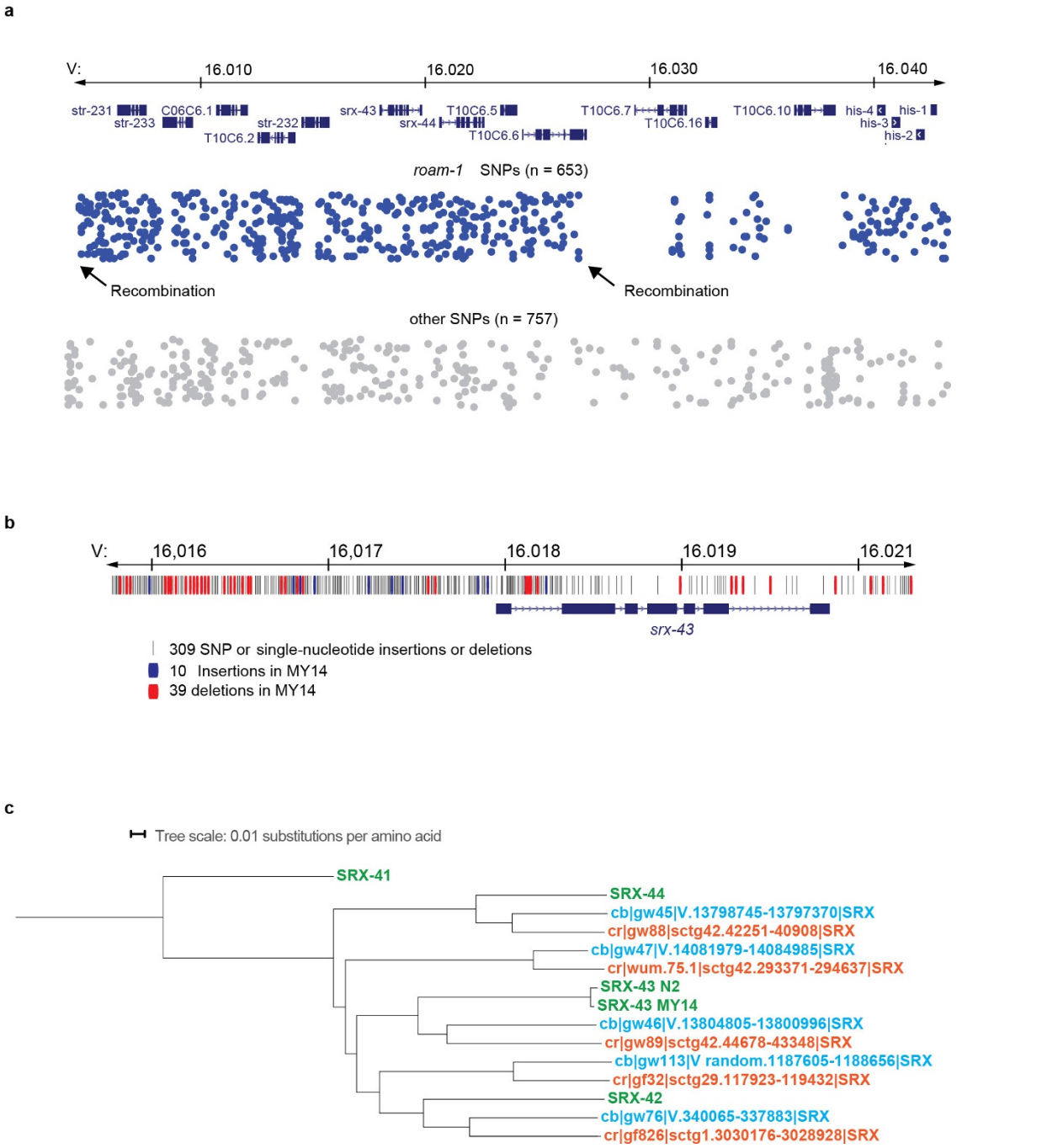
Extended Data Figure 3 | Ascarosides produced by wild-type strains. a, **b**, LC–MS/MS analysis of ascarosides secreted by N2, CX12311 and MY14 strains grown on OP50 (**a**) or HB101 (**b**) bacteria. icas#9 is produced at similar levels by icas#9-sensitive and icas#9-resistant strains. $n = 2$ (**a**) or 3 (**b**) culture extracts per genotype.



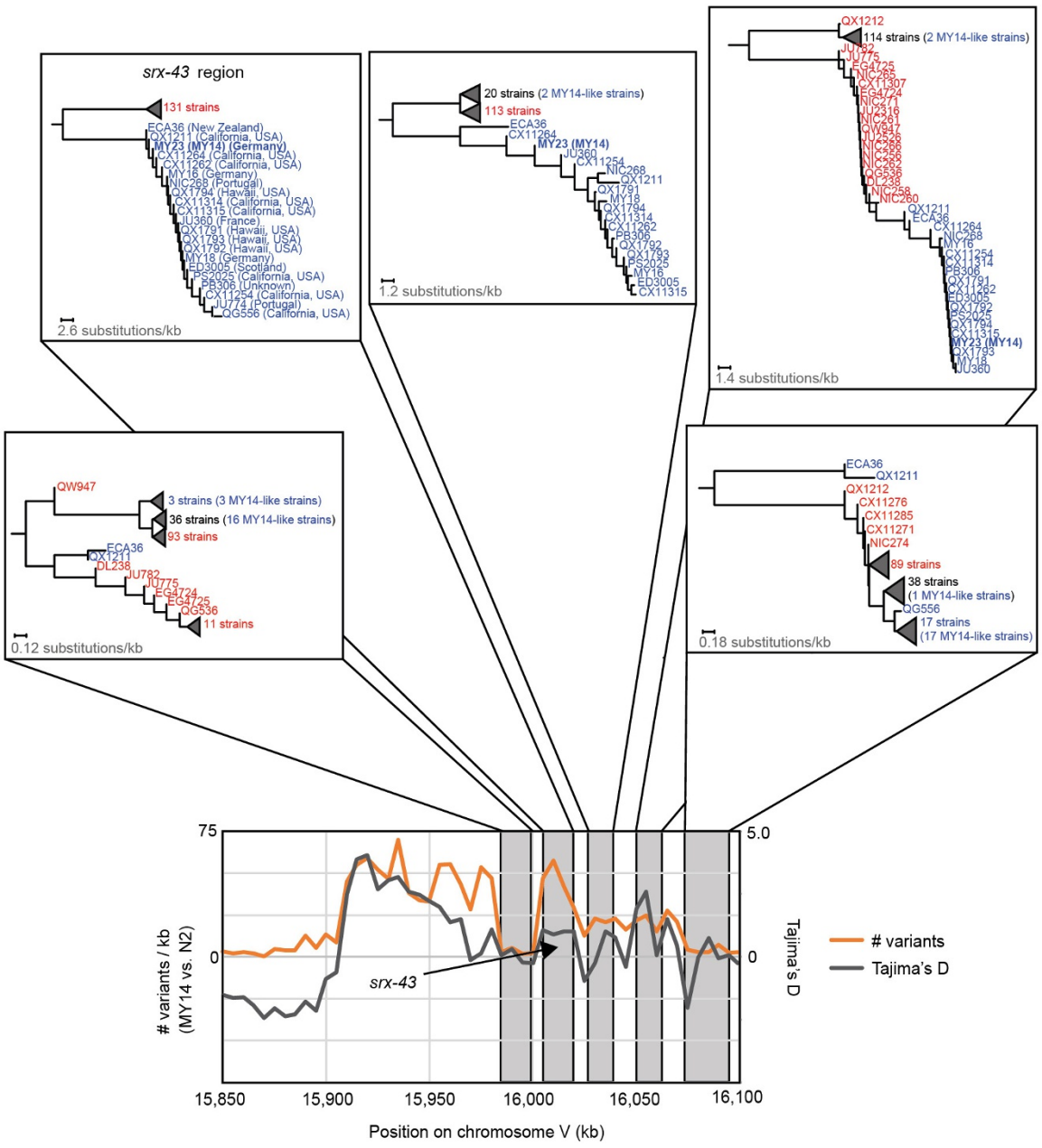
Extended Data Figure 4 | Covariate analysis of 94 RILs. a, **b**, Covariate analysis controlling for the *roam-1* genotype, testing for additive (**a**) or interactive (**b**) QTLs at other loci. The horizontal line denotes the $P < 0.05$ genome-wide significance threshold. LOD, log likelihood ratio.



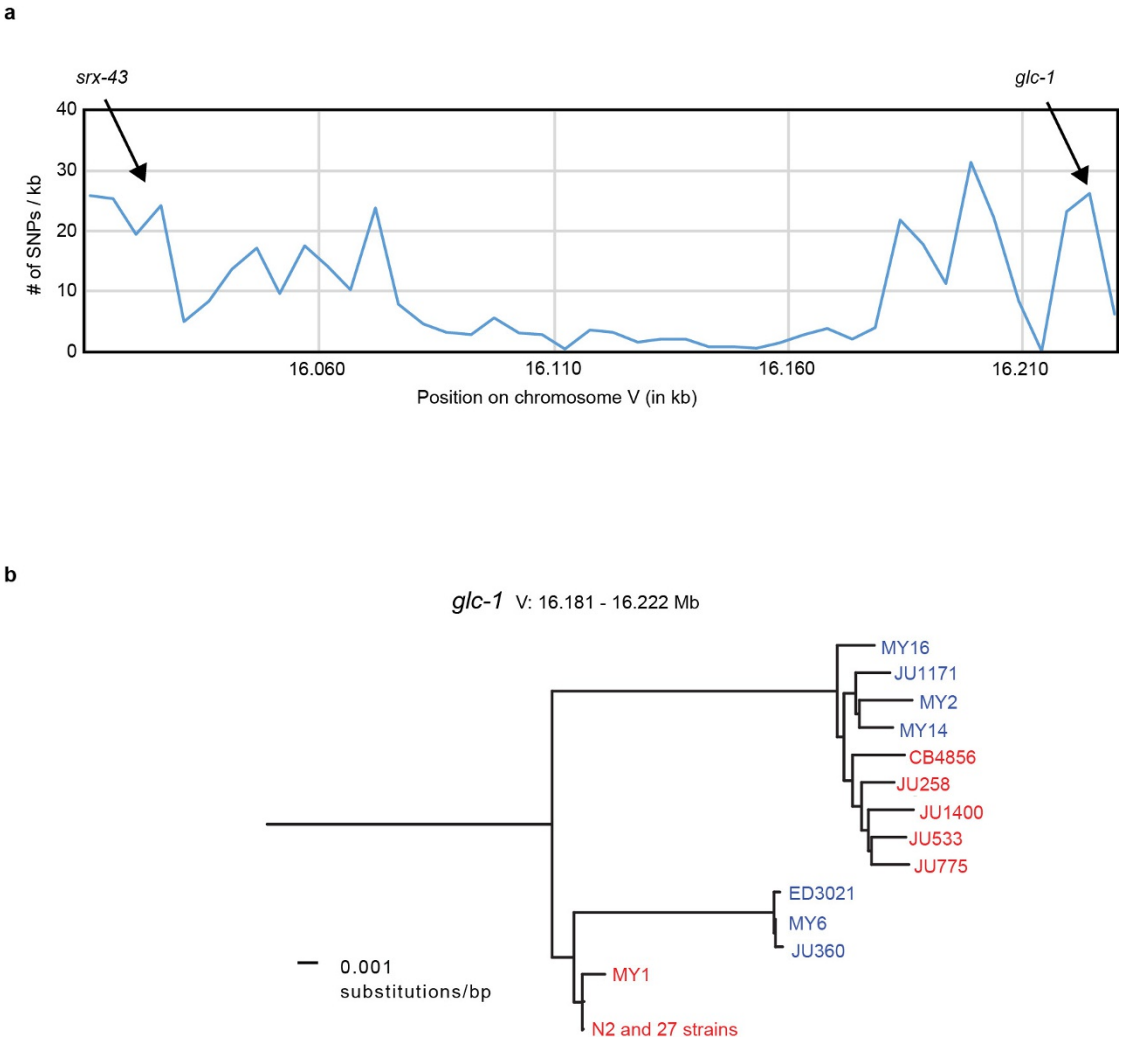
Extended Data Figure 5 | Signal transduction by SRX-43. a, Expression of *Psrx-43::srx-43::SL2::GFP* bicistronic reporter transgenes bearing N2 (top) or MY14 (bottom) *srx-43* sequences. Arrows indicate cell bodies of ASI sensory neurons. Scale bars, 50 μm. **b**, ASH sensory neurons are insensitive to multiple ascarosides. ASH calcium imaging with GCaMP3 in control animals that do not express the *srx-43* transgene, isolated as non-transgenic siblings of transgenic animals tested in Fig. 3f ($n = 19$). Ascarosides tested at 10 nM. **c**, SRX-43 from MY14 confers icas#9 sensitivity on ASH neurons. Compare SRX-43 from N2 in Fig. 3f. **d**, icas#9 decreases *daf-7::GFP* expression in ASI neurons of N2 but not *roam-1*$_{MY14}$ adults. Bars indicate mean fluorescence intensity ± s.e.m. $^*P < 0.05$ by ANOVA with Tukey's multiple comparisons test. **e**, Responses to icas#9 of *daf-7(lf)* mutants are attenuated in N2 but not in *roam-1*$_{MY14}$ genetic backgrounds. Modified exploration assays were conducted on strains including *daf-3(lf)* alleles (see methods). $^*P < 0.05$ by *t*-test. Data presented as mean ± s.e.m. **f**, Time course for icas#9 response in exploration assay. Pheromone response expressed as mean ± s.e.m. for 2, 4, 6, 10, and 14 h following initiation of exploration assay. $^{***}P < 0.001$ by *t*-test with Bonferonni correction comparing squares entered in control plates versus 10 nm icas#9 plates; $n = 12$ for all time points.



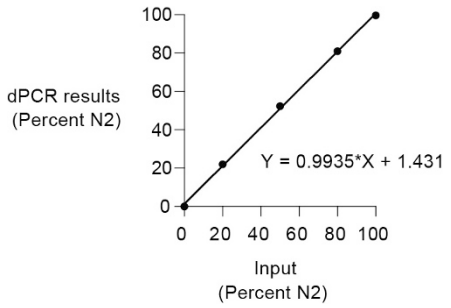
Extended Data Figure 6 | Alternative *roam-1* alleles have high sequence variability. a, The *roam-1* QTL region (top). *roam-1* SNPs are SNPs, when compared to the N2 reference genome, that are shared by JU360, MY2, MY14, ED3021, JU1171, MY16 and MY6 and not by any other strains, according to the Million Mutation Project. This defines the *roam-1*$_{MY14}$ haplotype. Other SNPs denote all other SNPs with respect to the N2 reference genome found in any of the 40 wild isolates in the Million Mutation Project. **b**, Polymorphisms of the *srx-43* promoter and coding region revealed by Sanger sequencing. Despite the high rate of polymorphism, there are only four non-synonymous mutations in the MY14 coding sequence detected by Sanger sequencing; three of these four were detected by the Million Mutation Project (Extended Data Table 1). We confirmed that the MY14 and N2 sequences are alleles of the same gene by examining sequence reads of the MY14-like strain MY23 in the CeNDR data set (http://www.elegansvariation.org) and aligning each read to N2 and MY14 sequence for the *srx-43* region as determined by Sanger sequencing. We observed that 7,272 of the MY23 (MY14) reads better matched the MY14 Sanger sequence and 4 of the reads better matched the N2 reference sequence, as would be expected if MY14 and N2 each bear one alternative allele of the gene. **c**, Phylogeny constructed for *srx-43* and related genes in *C. elegans*, *C. briggsae* and *C. remanei* demonstrates that the *srx-43* alleles in N2 and MY14 are closely related alleles of a single gene. Genes are colour-coded by species (green, *C. elegans*; blue, *C. briggsae*; orange, *C. remanei*). Protein sequences and gene names are as previously described[38].



Extended Data Figure 7 | Substantial recombination between *roam-1* and surrounding regions. Top, phylogenies constructed with 152 diverse wild-type isolates revealing differences for the region surrounding *srx-43* and the regions immediately to the left and right of the 30-kb haplotype. Bottom, graph showing the number of variants and Tajima's *D* score calculated for 5-kb bins across a 250-kb region. The bin containing *srx-43* has 250 polymorphisms and a Tajima's *D* of 1.01, which is high both at the genomic level (<3.4% of bins had a higher value) and for the chromosomal location of *srx-43* (<3.6% of bins had a higher value).



Extended Data Figure 8 | Recombination between *srx-43* and *glc-1* in natural isolates. a, The *glc-1* gene has previously been shown to be subject to balancing selection[33] and is chromosomally near *srx-43*. The blue line shows the number of SNPs per kb for N2 and MY14 averaged over 5-kb intervals for the region spanning *srx-43* and *glc-1*. The large region of low heterozygosity between *srx-43* and *glc-1* indicates that balancing selection on *glc-1* is unlikely to account for the high heterozygosity near *srx-43*. **b**, Dendrogram for the *glc-1* region for the strains shown in Fig. 4a. The clades for $roam\text{-}1_{MY14}$ and *glc-1* are not identical.



Extended Data Figure 9 | Standard curve for digital PCR experiments. Best-fit line of digital PCR results for known ratios of N2 to *roam-1*$_{MY14}$ DNA created by mixing different ratios of genomic DNA extracted from independent N2 or *roam-1*$_{MY14}$ populations.

Extended Data Table 1 | d_N/d_S for *srx-43* and other genes in the *roam-1* region

Gene	Chr.	Start	Stop	substitutions	Non-synonymous	Synonymous	dN/dS
str-231	V	16006304	16007605	36	9	27	0.097
str-233	V	16008332	16009664	85	24	61	0.118
C06C6.1	V	16010710	16012112	72	22	50	0.128
T10C6.2	V	16012537	16014226	60	24	36	0.195
str-232	V	16014509	16015752	60	20	40	0.145
srx-43	V	16017971	16019862	30	3	27	0.033
srx-44	V	16020342	16022652	25	5	20	0.071
T10C6.5	V	16023351	16024185	10	1	9	0.03
T10C6.6	V	16024326	16027454	27	2	25	0.025
T10C6.7	V	16027449	16031713	12	10	2	1.26
T10C6.16	V	16032471	16034961	28	22	6	0.943
T10C6.10	V	16036413	16038293	39	22	17	0.357
his-4	V	16040029	16040548	6	0	6	0
his-3	V	16040771	16041288	13	0	13	0
his-2	V	16041828	16042283	3	0	3	0
his-1	V	16042486	16043013	7	0	7	0
T10C6.15	V	16044296	16045874	19	9	10	0.228
ZK285.2	V	16046873	16047446	8	3	5	0.175
str-198	V	16047808	16049147	5	3	2	0.401
best-12	V	16049492	16051554	13	6	7	0.227
F14H3.3	V	16051923	16053510	32	22	10	0.522
F14H3.4	V	16053900	16055072	5	3	2	0.388
F14H3.5	V	16055806	16057054	11	10	1	2.711
F14H3.6	V	16057322	16058892	11	9	2	1.24
F14H3.15	V	16059781	16060235	11	9	2	1.37
fbxa-100	V	16060526	16062093	11	6	5	0.303
F14H3.9	V	16067949	16068763	9	7	2	0.83
cyp-35D1	V	16069238	16071301	94	27	67	0.113
nhr-176	V	16071325	16072738	15	5	10	0.127