Published in final edited form as:

Trends Genet. 2015 May ; 31(5): 224–231. doi:10.1016/j.tig.2015.02.009.

The laboratory domestication of *Caenorhabditis elegans*

Mark G. Sterken^{1,2}, L. Basten Snoek¹, Jan E. Kammenga^{1,*}, and Erik C. Andersen^{3,*}

¹Laboratory of Nematology, Droevendaalsesteeg 1, 6708 PB, Wageningen University, The Netherlands ²Laboratory of Virology, Droevendaalsesteeg 1, 6708 PB, Wageningen University, The Netherlands ³Department of Molecular Biosciences, Northwestern University, Evanston, Illinois, United States of America

Abstract

Model organisms are of great importance to understanding basic biology and to making advances in biomedical research. However, the influence of laboratory cultivation on these organisms is underappreciated, especially how that environment can affect research outcomes. Recent experiments led to insights into how the widely used laboratory reference strain of the nematode *Caenorhabditis elegans* compares to natural strains. Here, we describe potential selective pressures that led to fixation of laboratory-derived alleles for the genes *npr-1*, *glb-5*, and *nath-10*. These alleles influence a large number of traits, resulting in behaviors that affect experimental interpretations. Furthermore, strong phenotypic effects caused by these laboratory-derived alleles hinder the discovery of natural alleles. We highlight strategies to reduce the influences of laboratory-derived alleles and to harness the full power of *C. elegans*.

Keywords

C. elegans; laboratory-derived; selection; *npr-1*; aerotaxis

Model organisms pave the road to biological discovery

Sustained progress in the biological sciences is facilitated by discoveries using organisms amenable to laboratory investigation. They have large numbers of offspring, are small, and easy to maintain. Many different species have these attributes. For example, the single-cell eukaryote *Saccharomyces cerevisiae* (baker's yeast) is immensely powerful as a genetic model organism for conserved cellular processes [1] and for quantitative genetics using large populations [2]. The fruit fly *Drosophila melanogaster* contributed extensively to our understanding of signal-transduction pathways and developmental patterning [3]. The free-living nematode *Caenorhabditis elegans* is a widely used model organism in studies of

© 2015 Published by Elsevier Ltd.

*corresponding authors (jan.kammenga@wur.nl; erik.andersen@northwestern.edu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

development [4, 5], mechanistic neurobiology [6], aging [7], and small RNAs [8, 9]. When the results from experimental studies of model organisms are tabulated, it is obvious that they facilitated much of what we know about conserved biological processes.

Quantitative geneticists often use tractable model organisms to identify loci and (sometimes) genetic variants that influence phenotypic differences among populations. To elucidate the underlying genetic basis of complex traits, recombinant offspring are generated and their traits measured. Organisms that give rise to large (preferably clonal) populations and are easy to manipulate experimentally enable these approaches. These attributes make *S. cerevisiae* the most powerful eukaryotic organism for quantitative genetics [2, 10]. However, as a metazoan genetic system, *C. elegans* is unmatched [11]. It has an extremely rapid life cycle (3.5 days at 20°C), produces 200–300 offspring per hermaphrodite individual, possesses a small and well-annotated genome, can be cryopreserved, and transgenic strains are easily obtained. Wild strains isolated from nature can be phenotyped and genotyped to perform genome-wide association studies (GWAS, see Glossary) [12–15]. This combination of studies on natural allelic variation paired with analyses of mutations using the laboratory strain offer a powerful approach to broaden our understanding of how genetic background contributes to phenotype. However, characterization of the behaviors and genomes of wild *C. elegans* strains led to suspicions about laboratory adaptation in the widely used N2 laboratory strain [16]. Indeed, wild-type strains used in other model organisms have laboratory-derived variants that result in large pleiotropic effects, including cell clumping in *S. cerevisiae* [17, 18] and plant growth in *Arabidopsis thaliana* [19].

Here, we review documented examples of *C. elegans* laboratory-derived alleles in the commonly used Bristol (or N2) strain and the effects on its phenotype. We first describe how the laboratory strain N2 is different from wild strains of *C. elegans* and discuss the laboratory history of this nematode to gain insight into genetic bottlenecks and possible laboratory selection. Then, we discuss three known laboratory-derived alleles and their effects on *C. elegans* biology, including implications for the interpretation of observations that can confound experimental outcomes.

N2 is distinct from all wild strains of *C. elegans*

Since its introduction to the research community by Sydney Brenner in 1974 [20], the Bristol (or N2) strain has been used in many laboratories and became the canonical wild-type strain (for more history, see Box 1). From the Laboratory of Molecular Biology in Cambridge, N2 spread across the world via trainees from the Brenner laboratory, resulting in massive clonal amplification of N2 around the world. However, before its dissemination and cryogenic preservation, the N2 strain was propagated for many generations leading to the accumulation and selection of random mutations. We do not know how often the strain was transferred to new cultures during this early propagation. Conservatively, the strain could have been passaged every two months. At the other extreme, the strain could have been passaged every four days. Therefore, the strain underwent approximately 300 to 2000 generations from 1951 to 1969 (Box 1). Given the germline mutation rate of 2.7×10^{-9} mutations per site per generation [21], up to a thousand neutral mutations could have accumulated before cryogenic preservation. Furthermore, additional genetic differences

among N2 strains from different laboratories arose after dispersal of this strain around the world, causing differences in the phenotypes of these standard wild-type strains [22, 23].

The cultivation history of the Bristol N2 strain provides only a few opportunities to identify mutations that accumulated during the early culturing period (Box 1; Figure 1). Clues come from a strain that diverged from N2 sometime before 1963 while in the Dougherty laboratory [24] up to 12 years after initial isolation from nature. This strain was mislabeled as *C. briggsae* – a mistake that was corrected later [25]. In 1995 and 2009, hermaphrodites were removed from axenic culture and frozen as the LSJ1 and LSJ2 strains [25, 26], respectively. Comparison by sequencing revealed an estimate of approximately 100 accumulated variants in N2 [25]. However, no strains are currently known that diverged from N2 before the LSJ1 and LSJ2 strains diverged. Therefore, it is impossible to identify the mutations that accumulated in the initial decade after isolation.

From analysis of the genotypes and phenotypes of wild strains, we understand a great deal about variation in nature. Notably, many *C. elegans* strains reported to be isolated from nature were contaminated by the N2 strain (Table 1) [16]. Initial characterizations of natural phenotypic variation were confounded by these contaminated strains [27, 28]. Fortunately, recent sampling and genotyping of true wild strains have made it possible to study natural variation in *C. elegans* [12, 29, 30]. The genomes of most wild *C. elegans* strains isolated from nature are highly related, sharing nearly two thirds of the genome. This high degree of sharing is likely the effect of advantageous alleles that swept through the population reducing linked variation [12] and background selection that eliminates variation linked to deleterious alleles [31]. However, we still cannot identify all of the alleles that accumulated during laboratory propagation of the N2 strain even with these genotype data. The N2 genome contains private alleles found in the strain when isolated from nature along with mutations that accumulated during laboratory propagation. Together, this mix of alleles makes it impossible to identify laboratory-derived alleles from sequence information alone.

Selective pressures: nature versus laboratory

To understand how laboratory conditions could influence *C. elegans*, we need to know more about its natural habitat and ecology. Although progress has been made in recent years, the ecology of *C. elegans* is still largely unknown [15, 32, 33]. Despite frequent use of the statement in the literature, it is unlikely that *C. elegans* is a soil nematode. Soil samples harbor *C. elegans* only when in close proximity to rotting vegetation or fruit [15, 34] and recent successful sampling suggests that its natural habitat is rotting material. Wild strains were isolated successfully from rotting hogweed [15, 29], rotting fruits [15, 29, 32, 34], and compost [29, 32]. Additionally, strains have been isolated from “carrier” species, such as snails or terrestrial isopods [15]. Current observations indicate that *C. elegans* occupies short-lived microbiota-rich habitats. In this niche, it establishes a population quickly and is thought to compete for bacterial food with other species [15, 32]. When food is limiting and population density is high, *C. elegans* enters a long-lived alternative larval stage called dauer. These dauers likely endure periods without food while dispersing to new habitats [35]. By contrast, laboratory cultivation provides a much more constant environment (Box 2).

When animals are removed from their natural environments and transported to the laboratory, species undergo strong selective pressures that ultimately can change the organism. The impact of a laboratory environment on an organism is significant: environmental conditions are kept nearly constant; breeding regimes are strictly enforced; and food is readily available (Box 2). Additionally, researchers impose novel pressures by the culturing system, *e.g.* transferring individual animals to start a new culture (bottlenecks). The substrate on which animals are grown should be considered. Agar plates offer a two-dimensional substrate, whereas rotting fruit is a three-dimensional environment [15]. This laboratory propagation results in evolution through artificial selection, which inevitably affects genotypic and phenotypic characteristics of model organisms, including *C. elegans*.

From two studies, it is clear that the N2 genotype exhibits higher fitness in laboratory conditions than wild strains [26, 36]. The phenotype of N2 is distinct from wild strains in several ways, including aggregation behavior, maturation time, fecundity, body size, and many other traits [26, 27, 29, 36–43]. The atmospheric oxygen concentration on agar plates is substantially higher than levels preferred by wild strains [44–46], and laboratory oxygen concentration is a strong selective pressure on the organism. This oxygen concentration affects the growth and physiology of the animal profoundly because many behaviors are altered, including how the animals consume bacterial food. These oxygen-dependent effects are so profound that two out of three confirmed laboratory-derived alleles are associated with altered behaviors at higher oxygen concentrations [16, 41]. The effects of these alleles and possibly other laboratory-derived alleles are pleiotropic, so they could have been selected by additional unexplained pressures.

Laboratory-derived alleles in the *C. elegans* N2 strain and their functional consequences

During the first 18 years that the N2 strain was grown in the laboratory, many mutations arose that might not have conferred any selective advantage [21]. However, we know that laboratory propagation of this strain led to the fixation of several alleles that confer a strong selective advantage under these conditions [26, 36]. Laboratory-derived alleles are random mutations that increase the fitness of the organism in laboratory conditions. At least three genes in the N2 strain have laboratory-derived variation: *npr-1*, *glb-5*, and *nath-10* [16, 36, 41]. For each of these genes, the N2 genome contains a different variant than found among all *bona fide* wild strains. Furthermore, the two N2-diverged strains, LSJ1 and LSJ2 (Figure 1), carry the same alleles as wild strains. These results provide further evidence for the laboratory origin of the alleles, because LSJ1 and LSJ2 were separated from the N2 strain at least six years before cryopreservation [16, 25, 26].

The neuropeptide receptor encoding gene *npr-1*: laboratory adaptation abnormally represses the *C. elegans* nervous system

A seven transmembrane neuropeptide receptor encoded by *npr-1* was first identified as a master regulator of a behavioral dimorphism where animals either aggregate or remain solitary in the presence of bacterial food [27]. This aggregation behavior mapped to an amino-acid substitution within the third intracellular loop of the NPR-1 receptor. Wild

strains of *C. elegans* contain the 215F allele, with which the NPR-1 receptor responds to the neuropeptide FLP-21. By contrast, the laboratory strain N2 contains the 215V allele, which leads to a neomorphic gain-of-function sensitivity of NPR-1 to FLP-18 in addition to sensitivity to FLP-21 [47]. This gain-of-function sensitivity creates an abnormally repressed neural circuit through inactivation of the RMG interneuron [48], affecting a large number of behaviors (Table 2) [16, 27, 31, 39, 41, 44–56].

A modified aerotaxis response is one of the central drivers of the behavioral differences caused by variation in NPR-1 (Figure 2). Wild-type *C. elegans* strongly prefer oxygen concentrations lower than ambient levels [44–46]. On agar plate cultures, this behavior manifests as taxis to oxygen concentrations of approximately 10%, which is often found at the border of the bacterial lawn [44–46]. Aggregation of animals decreases the local oxygen concentration even further [46]. This reduction in oxygen concentration caused by aggregation reinforces the further formation of aggregates, which in turn decreases available food as animals compete in close proximity. The reduction in growth rate and offspring production observed in wild *C. elegans* strains is likely caused by a mild starvation state in aggregates [41]. Additionally, these animals could experience higher levels of pheromones, potentially signaling a stress state that reduces growth rate and offspring production [41]. The attraction of wild strains to the border of the lawn increases the exposure to bacteria [39]. When these bacteria are pathogenic, strains with the 215F allele will be exposed more extensively to the pathogen and succumb faster to infection than the N2 strain with the 215V allele [41, 49, 50].

Most traits related to *npr-1* variation are linked to aerotaxis behavior (Table 2). Traits not linked to aerotaxis include heat avoidance [53], ethanol tolerance [51], carbon dioxide avoidance [16, 52], and pheromone response [48, 54]. However, these traits still might be linked to aerotaxis via the RMG neuron, but these connections have not been characterized extensively. For example, the ethanol response might be regulated through FLP-18 [51] or might be sensed in the nociception neuron ASH, which is connected to RMG [48]. RMG is also connected to the pheromone sensing ADL neuron. As the aggregation observed in wild strains likely causes higher pheromone exposure and lower oxygen concentrations, it is difficult to distinguish the contributions of both factors [41]. Many traits where *npr-1* variation is implicated in the behavior have not been directly connected to aerotaxis behaviors by empirical evidence. Most of these traits, however, are likely caused by variation in the aerotaxis responses and differences in food consumption mediated by RMG through its role as a “hub and spoke” neuron [48].

Other laboratory-derived alleles found in the N2 strain affect *nath-10* and *glb-5*

Together with NPR-1, the neuronal globulin domain protein GLB-5 affects a behavioral response to changes in carbon dioxide and oxygen concentrations. The causal variant in *glb-5* is a duplication/insertion of 765 base pairs, leading to a 179 amino-acid truncation and a 40 amino-acid substitution in the N2 strain [16, 56]. The combination of laboratory-adapted alleles at the *glb-5* and *npr-1* loci leads to opposing responses to changes in carbon dioxide concentration, as compared to the wild alleles. Wild strains move more quickly and make more turns when they sense a simultaneous decrease in carbon dioxide concentration

and increase in oxygen concentration. By contrast, N2 animals move more quickly and make more turns when they sense an increase in carbon dioxide concentration. Furthermore, the N2 allele of *glb-5* desensitizes the URX neuron (which is also connected to RMG) to small fluctuations in oxygen levels, leading to reduced responses to in oxygen concentrations [16, 56]. The *npr-1* and *glb-5* alleles exhibit a genetic interaction. A strain with the natural alleles at both loci displays a different phenotype than the strains with only one natural allele. If only the N2 allele of *npr-1* is present, animals will react to oxygen in a concentration dependent manner, whereas the N2 *glb-5* allele on its own renders them insensitive to fluctuations in oxygen concentrations. If animals carry both alleles, however, they react strongly to minor shifts in oxygen and carbon dioxide concentrations around the atmospheric oxygen concentration [16, 56]. These discoveries related to oxygen and carbon dioxide preferences led to original observations of the derived nature of the N2 strain [16].

Variation in the human N-acetyltransferase homolog gene (*nath-10*) causes variation in vulval cell-fate specification and shows pleiotropic effects on fecundity and growth rate [36]. The laboratory-derived allele encodes a putative substitution of methionine 746 with isoleucine in a highly conserved region of the N-acetyltransferase domain. This laboratory-derived allele was identified because of specific effects on variation in vulval cell-fate specification. Variation in *nath-10* causes visible effects on vulval development only when additional mutations sensitize the *let-60* Ras pathway activity. The laboratory-derived allele of *nath-10* partially suppresses a lower level of vulval cell-fate induction caused by a reduction-of-function mutation in the gene encoding an EGF receptor (*let-23*) and enhances the level of vulval cell-fate induction caused by a gain-of-function mutation in the gene encoding Ras (*let-60*), indicating that the laboratory-derived allele of *nath-10* stimulates Ras pathway activity. This allele also affects the age at maturity, brood size, and egg-laying speed through an increase in the production of sperm. Given this large effect on fitness, the N2 allele of *nath-10* causes a selective advantage when animals are grown in laboratory competition assays [36].

The effects of natural allelic variation is obscured by propagation of strains in the laboratory

To investigate the effects of laboratory alleles, we analyzed the *C. elegans* linkage mapping results from the last decade for linkage to *npr-1*, *glb-5*, and *nath-10* genomic regions (Table 3). A large number of linkage mapping experiments detected a quantitative trait locus (QTL) with a confidence interval that includes the *npr-1* locus, including dauer formation [57, 58], body size [38, 41], lifespan [59], and vulval index [36]. Laboratory-derived alleles have large effects when strains are grown in laboratory conditions. To estimate this effect, we compared the broad-sense heritability (H^2) and the variance explained by the *npr-1* QTL. This comparison indicates how much of the genetic differences among strains are influenced by the *npr-1* QTL. Variation at the *npr-1* locus explains 30–82% of the variance contributed by genetic factors for a variety of traits [38, 41] – a large phenotypic effect. However, not all traits consistently detect a QTL at the *npr-1* locus. For example, one expression QTL study detected a *trans*-band at *npr-1* [31], but three other studies did not [60–62]. Similarly, one study detected the *npr-1* QTL for fecundity [41], but two other studies did not [37, 38]. We

suggest that studies that failed to detect a QTL nearby *npr-1* likely had different laboratory culture conditions, including the number of animals in the culture and the assay temperature. The differences in population density and variation at *npr-1* interact with large effects on the growth and physiology of the organism [41]. With increased culture density comes more crowding of animals at the edge of the bacterial lawn. This crowding causes a chronic low-level starvation state, which has large phenotypic effects. Additionally, differences in rearing temperature result in altered growth rates and population densities with similar effects.

The *nath-10* locus has been associated with several traits, including age at maturity [36, 37], an expression QTL *trans*-band [61], and dauer formation [58, 63]. It is difficult to assess the effect of the locus in general, as only one study reports the contribution to heritable variation (52%) [37]. The *trans*-band associated with the *nath-10* locus was measured in L4 and reproductive animals at 24°C. Therefore, it is likely that the developmental differences caused by *nath-10* result in gene-expression differences. Because *nath-10* is a pleiotropic locus implicated in fecundity and growth rate, these correlated QTL are not surprising.

In summary, the contribution of laboratory-derived alleles to heritable variation is large (30–82%) and seems to be environment-dependent. It is important to consider the context in which traits are measured. Given that both *npr-1* and *glb-5* affect behavior at atmospheric oxygen concentrations [16, 41, 44], many behavioral studies using the N2 strain in standard laboratory conditions might be difficult to interpret with respect to a normal behavioral circuit and natural behaviors.

Where do we go from here?

C. elegans is an essential model organism used to understand human biology. However, we need to be aware of the large and pleiotropic phenotypic effects caused by laboratory-derived alleles, especially those alleles present in the reference strain N2. These alleles can influence our conclusions and could alter the interpretations of results for understanding human biology, as it alters the natural physiology of *C. elegans*. However, investigators should not abandon the N2 strain. The large experimental toolkit and the decades of results obtained by study of this one strain are invaluable. These advantages need to be tempered with the knowledge that the N2 strain has been bred in a single environment for a long time prior to cryopreservation. Laboratories whose research focuses exclusively on the N2 strain and mutant derivatives should consider expanding to more natural *C. elegans* strains, especially when the focus includes traits that are influenced by population density (e.g. metabolism).

Newly isolated *C. elegans* that are cryopreserved as soon as possible after arrival in the laboratory are an untapped resource of genetic variants to expand the experimental power of *C. elegans* and its applicability to humans. For example, these strains can be added to the panel currently used for GWAS [12–14]. Additionally, new recombinant inbred line collections can be constructed using natural strains, which will greatly benefit quantitative genetic studies. One of the major strengths of *C. elegans* is the combination of wild strains and the accumulated knowledge from the study of the laboratory strain N2. This

combination allows for rapid screening of causal genes to understand evolutionary and ecological genetics along with making a larger impact on biomedical science. The *C. elegans* research community is ready for the next round of rapid and important progress once natural strains are integrated into the existing genetic toolkit.

Acknowledgements

We thank Rachel Ankeny, Bob Horvitz, Patrick McGrath, John Sulston, and Diana Wall for discussions, further documentation, and additional information about the history of the *C. elegans* N2 strain. We are grateful to Roel Bevers, Daniel Cook, Patrick McGrath, Anne Morbach, Lisa van Sluijs, Robyn Tanny, and Stefan Zdraljevic for editorial comments on the manuscript. MGS was funded by the Graduate School Production Ecology & Resource Conservation (PE&RC). LBS was funded by ALW grant 823.01.001. JEK received support from the EU-funded project PANACEA (222936). ECA received support from the Pew Charitable Trusts and NIH grant R01GM107227.

Glossary

Axenic culture	Conditions where organisms are grown in completely synthetic media. In the case of <i>C. elegans</i> , media is based on a liver extract [64, 65].
Dauer	Second larval stage (L2) animals enter this alternate larval stage at high culture density, low food abundance, and high temperature. [66] These L3 dauer larvae can survive stressful conditions and are thought to disperse to new locations in nature.
Ecological niche	The specific environment in which an organism lives and competes for resources. <i>C. elegans</i> are most often found in decaying material and not in soil [15, 32].
FLP-18	One of two FMRFamide neuropeptides encoded by the <i>C. elegans</i> genome that can bind to the NPR-1 neuropeptide receptor. FLP-18 can activate the NPR-1(215V) allele (found in N2 animals) but not the NPR-1(215F allele) (found in all wild strains) [16, 47].
FLP-21	The second of two FMRFamide neuropeptides encoded by the <i>C. elegans</i> genome that can bind to the NPR-1 neuropeptide receptor. FLP-21 is the natural ligand of NPR-1 [16, 41]
GLB-5	A globin domain protein that modifies behavioral responses to oxygen and oxygen/carbon dioxide stimuli [16, 56]
GWAS	Genome-wide association study, a technique used on natural populations to identify genomic regions correlated with differences in phenotypic traits [12–14].
Heritability	The amount of trait variation in a population that can be explained by genetic factors.
NATH-10	A vertebrate N-acetyltransferase homolog that has been shown to affect vulval induction in <i>C. elegans</i> . The 746I allele (N2) also results in faster maturation and fitness under laboratory conditions [36].

Neomorphic	Describes a type of phenotype caused by an alteration of gene function that is novel and different from the normal function of the gene.
NPR-1	a G protein-coupled neuropeptide receptor normally activated by FLP-21. The 215V allele (N2) gained the ability to respond to FLP-18. The 215V allele affects many different traits [16, 27, 31, 39, 41, 44–56].
Private alleles	Specific alleles occurring only in a single strain.
QTL	Quantitative trait locus, a locus correlated with quantitative trait variation. For example, the <i>npr-1</i> QTL is correlated with variation in aggregation behavior.
QTN	Quantitative trait nucleotide, the variant site that causes variation in the quantitative trait.
RMG interneuron	The central neuron involved in most behaviors mediated by NPR-1 [48].
Trans-band	In expression QTL studies, when variation of expression in many genes is correlated with the same genomic locus [31, 60–62].
Vulval cell induction	When any of six hypodermal cells located on the ventral surface of the hermaphrodite are specified and divide to become vulval cells.

References

1. Dunham MJ, Fowler DM. Contemporary, yeast-based approaches to understanding human genetic variation. *Current opinion in genetics & development*. 2013; 23:658–664. [PubMed: 24252429]
2. Bloom JS, et al. Finding the sources of missing heritability in a yeast cross. *Nature*. 2013; 494:234–237. [PubMed: 23376951]
3. Ejsmont RK, Hassan BA. The Little Fly that Could: Wizardry and Artistry of *Drosophila* Genomics. *Genes*. 2014; 5:385–414. [PubMed: 24827974]
4. Jackson BM, Eisenmann DM. beta-catenin-dependent Wnt signaling in *C. elegans*: teaching an old dog a new trick. *Cold Spring Harbor perspectives in biology*. 2012; 4:a007948. [PubMed: 22745286]
5. Felix MA, Barkoulas M. Robustness and flexibility in nematode vulva development. *Trends in genetics : TIG*. 2012; 28:185–195. [PubMed: 22325232]
6. Husson SJ, et al. Optogenetic manipulation of neural activity in *C. elegans*: from synapse to circuits and behaviour. *Biology of the cell / under the auspices of the European Cell Biology Organization*. 2013; 105:235–250. [PubMed: 23458457]
7. Kenyon C. The first long-lived mutants: discovery of the insulin/IGF-1 pathway for ageing. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. 2011; 366:9–16.
8. Fire A, et al. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*. 1998; 391:806–811. [PubMed: 9486653]
9. Kaufman EJ, Miska EA. The microRNAs of *Caenorhabditis elegans*. *Seminars in cell & developmental biology*. 2010; 21:728–737. [PubMed: 20637886]
10. Ehrenreich IM, et al. Genetic dissection of complex traits in yeast: insights from studies of gene expression and other phenotypes in the BYxRM cross. *Cold Spring Harbor symposia on quantitative biology*. 2009; 74:145–153.

11. Gaertner BE, Phillips PC. *Caenorhabditis elegans* as a platform for molecular quantitative genetics and the systems biology of natural variation. *Genetics research*. 2010; 92:331–348. [PubMed: 21429266]
12. Andersen EC, et al. Chromosome-scale selective sweeps shape *Caenorhabditis elegans* genomic diversity. *Nature genetics*. 2012; 44:285–290. [PubMed: 22286215]
13. Rockman MV, Kruglyak L. Recombinational landscape and population genomics of *Caenorhabditis elegans*. *PLoS genetics*. 2009; 5:e1000419. [PubMed: 19283065]
14. Ashe A, et al. A deletion polymorphism in the *Caenorhabditis elegans* RIG-I homolog disables viral RNA dicing and antiviral immunity. *eLife*. 2013; 2:e00994. [PubMed: 24137537]
15. Felix MA, Braendle C. The natural history of *Caenorhabditis elegans*. *Current biology : CB*. 2010; 20:R965–R969. [PubMed: 21093785]
16. McGrath PT, et al. Quantitative mapping of a digenic behavioral trait implicates globin variation in *C. elegans* sensory behaviors. *Neuron*. 2009; 61:692–699. [PubMed: 19285466]
17. Yvert G, et al. Trans-acting regulatory variation in *Saccharomyces cerevisiae* and the role of transcription factors. *Nature genetics*. 2003; 35:57–64. [PubMed: 12897782]
18. Brem RB, et al. Genetic dissection of transcriptional regulation in budding yeast. *Science*. 2002; 296:752–755. [PubMed: 11923494]
19. van Zanten M, et al. The many functions of ERECTA. *Trends in plant science*. 2009; 14:214–218. [PubMed: 19303350]
20. Brenner S. The genetics of *Caenorhabditis elegans*. *Genetics*. 1974; 77:71–94. [PubMed: 4366476]
21. Denver DR, et al. A genome-wide view of *Caenorhabditis elegans* base-substitution mutation processes. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106:16310–16314. [PubMed: 19805298]
22. Vergara IA, et al. Polymorphic segmental duplication in the nematode *Caenorhabditis elegans*. *BMC genomics*. 2009; 10:329. [PubMed: 19622155]
23. Gems D, Riddle DL. Defining wild-type life span in *Caenorhabditis elegans*. *The journals of gerontology. Series A, Biological sciences and medical sciences*. 2000; 55:B215–B219.
24. Dougherty EC. Letter from Ellsworth C. Dougherty to Sydney Brenner. *CSHL Archives Repository*. 1963
25. McGrath PT, et al. Parallel evolution of domesticated *Caenorhabditis* species targets pheromone receptor genes. *Nature*. 2011; 477:321–325. [PubMed: 21849976]
26. Weber KP, et al. Whole genome sequencing highlights genetic changes associated with laboratory domestication of *C. elegans*. *PloS one*. 2010; 5:e13922. [PubMed: 21085631]
27. de Bono M, Bargmann CI. Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. *Cell*. 1998; 94:679–689. [PubMed: 9741632]
28. Reiner DJ, et al. *C. elegans* anaplastic lymphoma kinase ortholog SCD-2 controls dauer formation by modulating TGF-beta signaling. *Current biology : CB*. 2008; 18:1101–1109. [PubMed: 18674914]
29. Volkers RJ, et al. Gene-environment and protein-degradation signatures characterize genomic and phenotypic diversity in wild *Caenorhabditis elegans* populations. *BMC biology*. 2013; 11:93. [PubMed: 23957880]
30. Thompson O, et al. The million mutation project: a new approach to genetics in *Caenorhabditis elegans*. *Genome research*. 2013; 23:1749–1762. [PubMed: 23800452]
31. Rockman MV, et al. Selection at linked sites shapes heritable phenotypic variation in *C. elegans*. *Science*. 2010; 330:372–376. [PubMed: 20947766]
32. Petersen C, et al. The prevalence of *Caenorhabditis elegans* across 1.5 years in selected North German locations: the importance of substrate type, abiotic parameters, and *Caenorhabditis* competitors. *BMC ecology*. 2014; 14:4. [PubMed: 24502455]
33. Petersen C, et al. Why we need more ecology for genetic models such as *C. elegans*. *Trends in genetics : TIG*. 2015
34. Kiontke KC, et al. A phylogeny and molecular barcodes for *Caenorhabditis*, with numerous new species from rotting fruits. *BMC evolutionary biology*. 2011; 11:339. [PubMed: 22103856]

35. Lee H, et al. Nictation, a dispersal behavior of the nematode *Caenorhabditis elegans*, is regulated by IL2 neurons. *Nature neuroscience*. 2012; 15:107–112.
36. Duveau F, Felix MA. Role of pleiotropy in the evolution of a cryptic developmental variation in *Caenorhabditis elegans*. *PLoS biology*. 2012; 10:e1001230. [PubMed: 22235190]
37. Gutteling EW, et al. Mapping phenotypic plasticity and genotype-environment interactions affecting life-history traits in *Caenorhabditis elegans*. *Heredity*. 2007; 98:28–37. [PubMed: 16955112]
38. Gutteling EW, et al. Environmental influence on the genetic correlations between life-history traits in *Caenorhabditis elegans*. *Heredity*. 2007; 98:206–213. [PubMed: 17203010]
39. Bendesky A, et al. Catecholamine receptor polymorphisms affect decision-making in *C. elegans*. *Nature*. 2011; 472:313–318. [PubMed: 21412235]
40. Bendesky A, et al. Long-range regulatory polymorphisms affecting a GABA receptor constitute a quantitative trait locus (QTL) for social behavior in *Caenorhabditis elegans*. *PLoS genetics*. 2012; 8:e1003157. [PubMed: 23284308]
41. Andersen EC, et al. A variant in the neuropeptide receptor *npr-1* is a major determinant of *Caenorhabditis elegans* growth and physiology. *PLoS genetics*. 2014; 10:e1004156. [PubMed: 24586193]
42. Kammenga JE, et al. A *Caenorhabditis elegans* wild type defies the temperature-size rule owing to a single nucleotide polymorphism in *tra-3*. *PLoS genetics*. 2007; 3:e34. [PubMed: 17335351]
43. Snoek LB, et al. A rapid and massive gene expression shift marking adolescent transition in *C. elegans*. *Scientific reports*. 2014; 4:3912. [PubMed: 24468752]
44. Gray JM, et al. Oxygen sensation and social feeding mediated by a *C. elegans* guanylate cyclase homologue. *Nature*. 2004; 430:317–322. [PubMed: 15220933]
45. Chang AJ, et al. A distributed chemosensory circuit for oxygen preference in *C. elegans*. *PLoS biology*. 2006; 4:e274. [PubMed: 16903785]
46. Rogers C, et al. Behavioral motifs and neural pathways coordinating O₂ responses and aggregation in *C. elegans*. *Current biology : CB*. 2006; 16:649–659. [PubMed: 16581509]
47. Rogers C, et al. Inhibition of *Caenorhabditis elegans* social feeding by FMRFamide-related peptide activation of NPR-1. *Nature neuroscience*. 2003; 6:1178–1185.
48. Macosko EZ, et al. A hub-and-spoke circuit drives pheromone attraction and social behaviour in *C. elegans*. *Nature*. 2009; 458:1171–1175. [PubMed: 19349961]
49. Styer KL, et al. Innate immunity in *Caenorhabditis elegans* is regulated by neurons expressing NPR-1/GPCR. *Science*. 2008; 322:460–464. [PubMed: 18801967]
50. Reddy KC, et al. A polymorphism in *npr-1* is a behavioral determinant of pathogen susceptibility in *C. elegans*. *Science*. 2009; 323:382–384. [PubMed: 19150845]
51. Davies AG, et al. Natural variation in the *npr-1* gene modifies ethanol responses of wild strains of *C. elegans*. *Neuron*. 2004; 42:731–743. [PubMed: 15182714]
52. Hallem EA, Sternberg PW. Acute carbon dioxide avoidance in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105:8038–8043. [PubMed: 18524955]
53. Glauser DA, et al. Heat avoidance is regulated by transient receptor potential (TRP) channels and a neuropeptide signaling pathway in *Caenorhabditis elegans*. *Genetics*. 2011; 188:91–103. [PubMed: 21368276]
54. Jang H, et al. Neuromodulatory state and sex specify alternative behaviors through antagonistic synaptic pathways in *C. elegans*. *Neuron*. 2012; 75:585–592. [PubMed: 22920251]
55. Choi S, et al. Analysis of NPR-1 reveals a circuit mechanism for behavioral quiescence in *C. elegans*. *Neuron*. 2013; 78:869–880. [PubMed: 23764289]
56. Persson A, et al. Natural variation in a neural globin tunes oxygen sensing in wild *Caenorhabditis elegans*. *Nature*. 2009; 458:1030–1033. [PubMed: 19262507]
57. Harvey SC, et al. Quantitative genetic analysis of life-history traits of *Caenorhabditis elegans* in stressful environments. *BMC evolutionary biology*. 2008; 8:15. [PubMed: 18211672]
58. Green JW, et al. Genetic mapping of variation in dauer larvae development in growing populations of *Caenorhabditis elegans*. *Heredity*. 2013; 111:306–313. [PubMed: 23715016]

59. Doroszuk A, et al. A genome-wide library of CB4856/N2 introgression lines of *Caenorhabditis elegans*. *Nucleic acids research*. 2009; 37:e110. [PubMed: 19542186]
60. Li Y, et al. Mapping determinants of gene expression plasticity by genetical genomics in *C. elegans*. *PLoS genetics*. 2006; 2:e222. [PubMed: 17196041]
61. Vinuela A, et al. Genome-wide gene expression regulation as a function of genotype and age in *C. elegans*. *Genome research*. 2010; 20:929–937. [PubMed: 20488933]
62. Li Y, et al. Global genetic robustness of the alternative splicing machinery in *Caenorhabditis elegans*. *Genetics*. 2010; 186:405–410. [PubMed: 20610403]
63. Green JW, et al. Highly Polygenic Variation in Environmental Perception Determines Dauer Larvae Formation in Growing Populations of *Caenorhabditis elegans*. *PloS one*. 2014; 9:e112830. [PubMed: 25393108]
64. Hansen ELY, E A, Nicholas WL, Sayre FW. Differential nutritional requirements for reproduction of two strains of *caenorhabditis elegans* in axenic culture. *Nematologica*. 1960; 5:27–31.
65. Nicholas WL, Mc EM. A technique for obtaining axenic cultures of rhabditid nematodes. *Journal of helminthology*. 1957; 31:135–144. [PubMed: 13475757]
66. Altun ZF, Herndon LA, Crocker C, Lints R, Hall DH. *Wormatlas*. 2002–2014
67. Ankeny RA. The natural history of *Caenorhabditis elegans* research. *Nature reviews. Genetics*. 2001; 2:474–479.
68. Brown, A. In the beginning was the worm : finding the secrets of life in a tiny hermaphrodite. Columbia University Press; 2003.
69. Riddle, DL. *C. elegans II*. Cold Spring Harbor Laboratory Press; 1997.
70. Adenle AA, et al. Review of the results from the International *C. elegans* first experiment (ICE-FIRST). *Advances in space research : the official journal of the Committee on Space Research*. 2009; 44:210–216.
71. Haber M, et al. Evolutionary history of *Caenorhabditis elegans* inferred from microsatellites: evidence for spatial and temporal genetic differentiation and the occurrence of outbreeding. *Molecular biology and evolution*. 2005; 22:160–173. [PubMed: 15371529]
72. Shmookler Reis RJ, et al. Quantitative trait loci define genes and pathways underlying genetic variation in longevity. *Experimental gerontology*. 2006; 41:1046–1054. [PubMed: 16919411]
73. Chandler CH. Cryptic intraspecific variation in sex determination in *Caenorhabditis elegans* revealed by mutations. *Heredity*. 2010; 105:473–482. [PubMed: 20502478]
74. Elvin M, et al. A fitness assay for comparing RNAi effects across multiple *C. elegans* genotypes. *BMC genomics*. 2011; 12:510. [PubMed: 22004469]
75. Gaertner BE, et al. More than the sum of its parts: a complex epistatic network underlies natural variation in thermal preference behavior in *Caenorhabditis elegans*. *Genetics*. 2012; 192:1533–1542. [PubMed: 23086219]
76. Francesconi M, Lehner B. The effects of genetic variation on gene expression dynamics during development. *Nature*. 2014; 505:208–211. [PubMed: 24270809]
77. Snoek LB, et al. Widespread genomic incompatibilities in *Caenorhabditis elegans*. *G3*. 2014; 4:1813–1823. [PubMed: 25128438]

Box 1: *C. elegans*: the journey from nature to the bench

Most *C. elegans* research laboratories use the strain named N2, which was collected in 1951 from mushroom compost in Bristol, England. Like most model organisms, the journey from nature to the laboratory was circuitous (Figure 1). The compost was collected by L.N. Staniland, who brought the sample to a short course on plant nematology organized by the British Ministry of Agriculture and Fisheries [16]. From this sample, Bristol *C. elegans* was isolated by Warwick Nicholas who cultured the animal first on petri dishes containing nutrient agar with bacterial contaminants as food [16]. Later, Warwick Nicholas developed axenic liquid cultures from the nutrient agar cultures, as these conditions required less frequent sub-culturing [65]. In 1957, the nematodes were transported to the laboratory of Ellsworth Dougherty at the Kaiser Foundation Research Institute in Richmond, California in a liquid axenic culture [16]. In the Dougherty laboratory, Eder Hansen cultured the nematodes. Two types of cultures were established: nutrient agar slants seeded with *E. coli* in test tubes and liquid axenic culture based on liver extract [64].

Concurrently, Sydney Brenner sought an organism suitable for neurobiology research [67]. He corresponded with Ellsworth Dougherty and even isolated nematodes from his own garden [68]. This nematode culture was called the N1 strain. Sydney Brenner requested the Bristol strain from Ellsworth Dougherty, and it was sent in 1963 [20, 24, 67]. In the Brenner laboratory, the liquid axenic culture was transferred to agar plates containing *E. coli*. After several passages of a population containing both males and hermaphrodites, a single hermaphrodite was selected. This strain, which was used for all subsequent work, was called N2 [68]. The populations were kept in culture on *E. coli* monoxenic agar plates, and the hermaphrodite strain was eventually frozen in 1969 by John Sulston [69].

Box 2: Living conditions of *C. elegans* in the laboratory

The life cycle of *C. elegans* consists of an embryonic stage, followed by four larval stages (L1-L4) and an adult stage. The N2 strain completes one generation every 3.5 days at 20°C. Alternatively, *C. elegans* can enter a long-term survival stage (dauer) as an alternative to the standard L3 larval stage [66].

Axenic culture

Axenic cultures do not contain other organisms as a food source and can be chemically defined or contain extracts of organic material (*e.g.* liver). Such cultures can be either in a solid state (*e.g.* nutrient agar) or in liquid.

Nowadays, axenic cultures are not often used for keeping *C. elegans* with the exception of transport into space [70]. In the early days of *Caenorhabditis* sp. research, much time was invested to establish a defined axenic medium to grow nematodes [64, 65] for two major reasons. First, it required a lower frequency of sub-culturing. Before the cryopreservation method was developed, infrequent sub-culture requirements were a great advantage. Second, axenic culture offered the ability to chemically define the medium, which allows the researcher to alter components and investigate nutritional requirements.

Monoxenic culture

Monoxenic cultures contain one organism as a food source. In the case of *C. elegans*, the nematode is almost exclusively cultured on media containing *E. coli*.

There are two main methods for monoxenic culture of *C. elegans*: either in liquid or on solid medium. In liquid culture, animals are grown with agitation in solution. On solid media, the animals are kept on nematode growth medium (NGM) agar plates seeded with an *E. coli* strain [20].

Highlights

1. Laboratory-derived alleles in the *C. elegans* reference strain Bristol N2 cause large phenotypic effects.
2. These alleles were selected after many generations of growth in the laboratory.
3. Natural strains are more representative of *C. elegans* behaviors and physiology than the laboratory strain N2.

Figure 1. The history of the Bristol (N2) lineage

Monoxenic (gray) or axenic (black) cultures are denoted by colored boxes. The gradients indicate uncertainty in when the culture type was switched. The dates show the year of isolation or when the strain was moved to another laboratory. Underscored dates mark the dates of cryogenic preservation. In 1951, Bristol was isolated by L.N. Staniland and Warwick Nicholas. The strain was kept in Liverpool first as a monoxenic culture and later as an axenic culture [16]. In 1957, it was shipped to the Kaiser Foundation Research Institute in Richmond, California. During early laboratory propagation, both axenic and monoxenic

cultures were maintained. Later, the monoxenic culture was discontinued and the axenic culture continued [24]. The LSJ1/LSJ2 strains originate from this axenic culture [16, 26], as does N2. It is unclear when exactly the LSJ1/LSJ2 lineage split from N2 [25]. In 1963, Brenner received an axenic culture containing *C. elegans* [24], which was cryogenically preserved by John Sulston in 1969 [69]. The LSJ1 strain was cryogenically preserved in 1995 [16, 26] and the LSJ2 strain in 2009 [16].

Figure 2. The aerotaxis effects of *npr-1*

Natural *C. elegans* strains aggregate at the edges of the bacterial lawn (orange) when propagated in laboratory conditions on monoxenic agar plates (left diagram). The edges of the bacterial lawn have lower than ambient oxygen concentrations (approximately 13%, center diagram). Wild *C. elegans* strains respond to this oxygen gradient and prefer lower oxygen concentrations in the presence of bacterial food [44]. The abnormal N2 strain is less sensitive to oxygen concentrations and does not aggregate at the edges of the bacterial lawn (right diagram). This difference in aerotaxis or oxygen preference leads to different aggregation and lawn leaving behaviors [27, 39, 44, 47, 48]. Because of these behavioral changes, strains also differ in exposure to pathogens [41, 49, 50]. Additionally, the aggregation behavior might cause a chronic mild starvation state, which results in a reduced growth rate [41], reduced fecundity [41], altered gene expression [31, 41], increased crawling speed [56], and reduced quiescence during lethargus [55].

Table 1

A large number of “wild” *C. elegans* strains are actually mislabeled N2 strains or recombinant strains derived from N2.

Strain	Genotype*	Ref.
CB3191	N2	16, 71
CB3192	N2	16, 71
CB3193	N2	16, 71
CB3194	N2	16, 71
CB3195	N2	16, 71
CB4507	N2	16
CB4555	N2×CB4851 recombinant	16
DH424	N2×CB4851 recombinant	16
DR1349	N2×CB4851 recombinant	16
PX176	N2	16
TR388	N2	16, 71
TR389	N2	16, 71
TR403	N2×CB4851 recombinant	16

* - All strains with the N2 genotype have N2 markers at 1,453 of 1,454 markers spread throughout the genome. The N2×CB4851 recombinant strains are largely or completely N2 for chromosomes I, II, III, and X and CB4851 for chromosomes IV and V. Strain TR389 has N2 for 1,453 markers but harbors the CB4856-like *glb-5* deletion allele.

Table 2

The laboratory-derived allele of *npr-1* causes a large number of phenotypic effects.

Trait	Phenotypic effect	Related to aerotaxis?	Ref.
Aggregation	Lower	Yes	27, 44, 47, 48
Taxis to low oxygen	Lower	Yes	41, 44–46
Pathogen avoidance	Higher	Yes	41, 49, 50
Lifetime fecundity	Higher	Yes	41
Body size	Larger	Yes	41
Gene-expression regulation	NA	Yes	31, 41
Ethanol tolerance	Lower	Not tested	51
Carbon dioxide avoidance	Higher	Not tested	16, 52
Heat avoidance	Higher	Not tested	53
Hermaphrodite leaving	Lower	Yes	39
Pheromone responses	Repulsed	Not tested	48, 54
Lethargus quiescence	Higher	Yes	55
Crawling speed	Lower	Not tested	56

Table 3

Many linkage mapping genetic studies identified the *npr-1*, *glb-5*, or *nath-10* locus.

Strains	Trait	Interval detected*			Identified Causal gene	Ref.
		<i>nath-10</i>	<i>glb-5</i>	<i>npr-1</i>		
N2×BO	Lifespan		yes			72
	Oxidative stress response		yes			
N2×CB4856	Age at maturity, 24°C	yes				37
N2×CB4856	Body mass, 12°C			yes		38
	Body mass, 24°C			yes		
N2×DR1350	Dauer formation, high food, 19°C			yes		57
	Dauer formation, food, plasticity			yes		
N2×CB4856	Pathogen susceptibility			yes	<i>npr-1</i>	50
N2×CB4856	Lifespan			yes		59
N2×CB4856	Carbon dioxide upshift, oxygen downshift		yes	yes	<i>glb-5</i> , <i>npr-1</i>	16
	Carbon dioxide downshift, oxygen upshift		yes	yes	<i>glb-5</i> , <i>npr-1</i>	
N2×CB4856	Oxygen sensing and response		yes	yes	<i>glb-5</i> , <i>npr-1</i>	56
N2 [#] ×CB4856	Male tail phenotype, 13°C	yes				73
N2×CB4856	Gene expression, L4, 24°C	yes				61
	Gene expression, L4 and reproductive, 24°C	yes				
N2×CB4856	Gene expression, young adult, 20°C			yes		31
N2×CB4856	Population growth on RNAi (8/11 genes)	yes			<i>ppw-1</i>	74
N2×CB4856	Lawn leaving			yes	<i>tyra-3</i>	39
N2×CB4856	Heat avoidance			yes	<i>npr-1</i>	53
JU605×JU606	Vulval induction, 20°C			yes		36
	Vulval induction, 25.5°C	yes			<i>nath-10</i>	
	Vulval induction, plasticity	yes			<i>nath-10</i>	
N2×CB4856	Bordering		yes		<i>exp-1</i>	40
N2×CB4856	Thermal preference			yes		75
	Isothermal dispersion			yes		

Strains	Trait	Interval detected*			Identified Causal gene	Ref.
		<i>nath-10</i>	<i>glb-5</i>	<i>npr-1</i>		
N2×CB4856	Dauer formation	yes		yes		58
N2×CB4856	Gene expression			yes		76
N2×CB4856	Lifetime fecundity			yes	<i>npr-1</i>	41
	Adult body size		yes	yes	<i>npr-1</i>	
	Susceptibility to <i>S. aureus</i>			yes	<i>npr-1</i>	
	Gene expression			yes	<i>npr-1</i>	
N2×CB4856	Dauer formation (pheromone exposure)	yes	yes			58
	Dauer formation (food exhaustion)	yes			<i>nath-10</i>	
N2×CB4856	Embryonic development	yes				77

* – When these intervals are not detected, this negative result could be caused by several factors. For example, this interval is not involved in the trait. Alternatively, the lack of detection could be caused by technical reasons, including presence of markers, statistical power, etc.

- The strain used for constructing the recombinant inbred population was CB5362, a strain containing the *tra-2(ar221)* and the *xol-1(y9)* mutations in an N2 background