Review

Beyond induced mutants: using worms to study natural variation in genetic pathways

Jan E. Kammenga[1], Patrick C. Phillips[2], Mario De Bono[3] and Agnieszka Doroszuk[1]

[1] Laboratory of Nematology, Wageningen University, 6709 PD, Wageningen, The Netherlands
[2] Center for Ecology and Evolutionary Biology, University of Oregon, Eugene, OR 97403-5289, USA
[3] Cell Biology Division, Medical Research Council, Laboratory of Molecular Biology, Cambridge CB2 2QH, UK

Induced mutants in the nematode *Caenorhabditis elegans* are used to study genetic pathways of processes ranging from aging to behavior. The effects of such mutations are usually analyzed in a single wildtype background: N2. However, studies in other species demonstrate that the phenotype(s) of induced mutations can vary widely depending on the genetic background. Moreover, induced mutations in one genetic background do not reveal the allelic effects that segregate in natural populations and contribute to phenotypic variation. Because other wildtype *Caenorhabditis* spp., including *C. elegans*, are now available, we review how current mapping resources and methodologies within and between species support the use of *Caenorhabditis* spp. for studying genetic variation, with a focus on pathways associated with human disease.

The worm as a model for human biology: induced mutations versus natural allelic variants

Over the past 40 years, the tiny (~1 mm) nematode *Caenorhabditis elegans* [1] has been widely used as a tractable model organism. Populations can be easily maintained and observed on agar plates covered with the bacterium *Escherichia coli* on which *C. elegans* feeds. Because of its short generation time (3.5 days at 20 °C) and abundant reproduction rate (~300 eggs at once), thousands of worms can be obtained over a short period. The hermaphroditic mode of reproduction rapidly drives *C. elegans* to homozygosity, which effectively makes individual worms within a strain genetically identical to one another. This property makes it relatively easy to isolate recessive mutants on all chromosomes. Males occur at <1% of total worm density and can be maintained through mating with hermaphrodites. An important advantage in using worms for genetics research is that *C. elegans* homozygotes do not suffer from inbreeding depression or hybrid vigor (see Glossary) [2] (in fact they display outbreeding depression [3]), suggesting that being highly inbred is a natural state for these worms. These characteristics, together with its transparent body, make *C. elegans* a highly suitable model for developmental biology and genetics. A further advantage of this system in general is that several related species within the *Caenorhabditis* genus share many of these characteristics.

A principle rationale for working with *C. elegans* is that it provides a fast and inexpensive way to learn more about human biology. Worms and humans share many biological characteristics including tissue development, a nervous system, muscles and a feeding gut. Worms also produce sperm and eggs, and their cells undergo senescence. Comparative genomic studies have shown that many genetic pathways including those involved in cellular signaling and apoptosis are conserved between worms and mammals

Glossary

Balancer chromosome: A chromosome that contains a large inversion whose purpose is to suppress meiotic recombination. Ideally, balancer chromosomes are homozygous lethal and contain a dominant marker that can be easily identified phenotypically.
Dauer larvae: Under adverse conditions, e.g. food deprivation, L1 and L2 larvae can form an arrested developmental stage termed the dauer that is specialized for survival and dispersal.
Ectotherm: An organism that regulates its body temperature by exchanging heat with its ambient surroundings.
EMS: Ethyl methanesulphonate, a chemical widely used as a powerful mutagenic agent.
Forward genetics: Studying the phenotype of a gene by means of induced mutation (i.e. from the phenotype to the genotype).
Genetical genomics: Genetic mapping of gene expression. Determine gene expression in genetically different individuals, treat the transcript abundance of each gene in each individual as a quantitative trait, use molecular markers to genotype the individuals and use linkage analysis or association analysis methods to identify regulating loci (expression quantitative trait loci; eQTL).
Gonochoristic: A sexually reproducing species in which there are two distinct sexes.
Hybrid vigor: The tendency of offspring to have qualities superior to those of either parent.
In cis: Local regulatory variation because of a polymorphism (or polymorphisms) in the gene itself.
In trans: Distant regulatory variation because of a polymorphism (or polymorphisms).
NIL: Near isogenic line. A line that carries a single segment from one strain on the genetic background of the other; also called congenic or introgression lines.
Null mutation: A mutation in a gene that does not lead to the production of a functional protein product.
QTL: Genetic loci where allelic variation (polymorphism) is associated with variation in a quantitative trait, e.g. leaf length in plants, body mass, number of offspring but also gene transcription (see *Genetical genomics*).
Reverse genetics: Studying the phenotype of a gene by means of targeted gene knock-outs (i.e. from the genotype to the phenotype).
RIL: Recombinant inbred line. These lines contain unique, nearly equal proportions of genetic loci from two progenitor lines.
RNA interference (RNAi): A mechanism preventing gene expression by causing the degradation of specific RNA molecules or hindering the transcription of specific genes.
Transposon: Stretches of DNA that can move and multiply within the genome of an organism. In *Caenorhabditis elegans, Tc1* is the most abundant transposon.

Corresponding author: Kammenga, J.E. (Jan.Kammenga@wur.nl).

0168-9525/$ – see front matter doi:10.1016/j.tig.2008.01.001 Available online 5 March 2008

and that gene homology with humans can approach 80% [4]. *C. elegans* research has helped to unravel genetic pathways underlying processes ranging from aging to behavior [5,6], as well as the characterization of genes underlying human cancer [7]. Although worms have a relatively short lifespan, the genetic mechanisms underlying life span extension within these nematodes are also highly conserved in mammals. By example, the genes connecting the insulin/insulin-like growth factor-1 (IGF-1) receptor to FOXO (forkhead transcription factor) are key longevity regulators in worms and mammals [8]. Recent findings show that inhibiting IGF-1 signaling prolongs lifespan and increases tumor resistance in *C. elegans* and most likely also in mammals [9]. The worm has also been studied widely for biomedical research, e.g. in the functional characterization of novel human drug targets [4]. One important discovery in this field centered on the transmembrane protein presenilin, which is encoded by *sel-12* (suppressor/enhancer of *lin-12*) in *C. elegans* [10]. *sel-12* mutations cause neuronal deficiencies. Transgenic expression of human *PRESENILIN1*, which is associated with Alzheimer disease (AD), rescues *sel-12* mutants [11,12], indicating a functional conservation between *C. elegans* and humans in this pathway and a possible use for *C. elegans* in identifying potential AD targets.

Genetic pathway analysis in *C. elegans* is rooted in forward genetic screens: the standard laboratory wildtype strain N2 (Box 1) is exposed to mutagenic agents, e.g. ethyl methanesulphonate (EMS) or UV radiation, and specific phenotypes are sought in the F1, F2 or F3 generations. Induced mutants with altered phenotypes (e.g. prolonged lifespan, altered cellular development) are outcrossed for several generations, and homozygous mutant strains are obtained. If a mutation confers lethality, it is maintained in heterozygotes, usually *in trans* to a balancer chromosome. Thousands of alleles have been isolated in this way. Induced mutations include null alleles, weak loss-of-function alleles with varying degrees of severity, gain-of-function mutations and dominant negative mutations. By example, detailed knowledge is available for a variety of allelic mutations in *let-60* (*let*hal, an ortholog to human *RAS*), which affect vulva development [13], and the *daf-2* insulin-like receptor (abnormal *da*uer larvae *f*ormation, an ortholog to the human IGF-I/insulin receptor family), which influences lifespan [14]. These mutations have been studied almost exclusively in the N2 strain.

Induced mutations of this kind are usually selected by researchers on the basis of strong phenotypic effects, which facilitate studying gene functions in signaling pathways. However, geneticists working in many diverse fields, ranging from plant biology to human disease studies, have become increasingly aware that studying natural genetic variation is crucial for understanding gene function; indeed, it often underlies phenotypic differences among individuals. Allelic interactions in different genetic backgrounds play an important role in determining the genetic basis of complex traits. Experimentation that relies solely on induced mutants and gene knockouts or on reverse genetic studies using RNA interference (RNAi) carried out in a single wildtype strain limits our ability to explore how naturally varying alleles alter signaling pathways.

Here we explore the potential use of *Caenorhabditis* spp. to study inter- and intraspecific natural variation. We present an overview of current genetic mapping resources in *Caenorhabditis* spp. and provide detailed examples of suitable mapping strategies. We discuss applications of whole pathway analysis and provide examples of pathways, discovered in worms, that underlie complex human diseases.

Box 1. A brief history of *Caenorhabditis elegans* wildtype strain N2

Nobel laurate Sydney Brenner introduced N2 as a model wildtype strain in developmental biology and genetics. The strain was obtained by picking one hermaphrodite from a stock population originally extracted from mushroom compost near Bristol, England, and maintained within the laboratory in axenic conditions for nearly a decade before being identified as a likely candidate for a new model system by Brenner in 1964. Its progeny was used to establish a line of hermaphrodites that propagate by self-fertilization and a line that contains males [81]. Brenner induced many mutations in N2 that affected behavior and morphology. Since then, N2 has been the standard strain in which many important discoveries have been made and that have served as platforms for future research. For example, the elucidation of the complete N2 cell lineage from egg to adult [82] was a landmark that provided the basis for current *Caenorhabditis elegans* research in developmental biology. The complete N2 genome sequence [69] was the first sequenced metazoan and has laid the foundation for genome-wide analyses of gene function. This was facilitated by the discovery of RNA interference (RNAi) in N2, which involves double-stranded RNA that includes at least 150 bp of the coding sequence of a particular gene. RNAi causes a specific loss-of-function phenotype for many of the genes tested [83].

Genetic pathways in *Caenorhabditis* spp.

The first steps *C. elegans* workers made toward understanding the effects of variation on gene function were comparative in nature. To date, the genus *Caenorhabditis* includes 23 species. Because *Caenorhabditis* spp. can be frozen as viable stocks, genetic drift within natural isolates can be effectively controlled by periodically returning to frozen references. A large and growing number of wildtype strains of *C. elegans*, *C. briggsae* and *C. remanei* have been collected from various locations covering diverse habitats including compost heaps, rotting fruit, commensal insects, garden soils and woodlands [15]. Comparative approaches for studying interspecific variation in pathways have focused primarily on these three species. For instance, Rudel and Kimble [16] investigated genetic diversity in *lin-12* and *glp-1*, which encode Notch-like receptors. The Notch-signaling pathway is present in all animals and regulates the development of a wide variety of tissues and structures [17]. Phylogenetic analysis revealed that LIN-12 orthologs grouped together and the GLP-1 orthologs grouped together. This finding confirms that the *lin-12*/*glp-1* duplication occurred before the divergence of these three *Caenorhanditis* species. RNAi knockdown of *lin-12* in both *C. remanei* and *C. briggsae* showed a phenotype that is typical of *C. elegans lin-12 glp-1* double mutants but that is not observed in *C. elegans* single mutants [16]. This finding suggests that the Notch pathway has diverged within these species, perhaps through subfunctionalization of the *lin-12*/*glp-1* gene duplication.

Perhaps even more revealing from a comparative standpoint are the genetic pathways involved in sex determination and sexual performance, because both *C. elegans* and *C. briggsae* are primarily selfing hermaphrodites, whereas *C. remanei* is an outcrosser with separate sexes. Surprisingly, selfing seems to have evolved independently within *C. elegans* and *C. briggsae*. Although the same set of core sex determination genes are present within these species, they are separately modified within each group [18]. In *C. elegans*, hermaphrodite spermiogenesis requires germline translational repression of the female-promoting gene, *tra-2*, which leads to de-repression of male-promoting *fem* genes. Inhibition of hermaphrodite spermiogenesis requires *fem-3* translational repression. Conversely, *C. briggsae* requires neither *fem-2* nor *fem-3* for hermaphrodite development, and XO (normally male) *C. briggsae fem-2/3* worms are transformed into hermaphrodites, not into females as in *C. elegans*. These observations led Hill *et al.* [18] to conclude that the morphologically similar hermaphrodites in *C. elegans* and *C. briggsae* achieve self-fertility through intervention at different points in the sex determination pathway.

Another well-understood developmental pathway in *C. elegans* is the vuval patterning system. Although *C. elegans*, *C. briggsae* and *C. remanei* are virtually indistinguishable morphologically (including the vulva), this similarity is actually achieved by subtle changes of the underlying cell–cell signaling pathways within each species [19]. This finding demonstrates the strength of probing interspecific variation in a complex developmental pathway by using detailed information from *C. elegans* as a starting point.

As might be expected, these studies reveal that the genetic background in which any particular gene functions changes during evolution; moreover, these changes can be uncovered through comparative studies using *C. elegans* as a focal species. However, although comparisons across species provide valuable insights into the evolution of genetic pathways, variation within species is more informative for allelic interactions and is a more applicable model for human genetic variation.

Natural genetic variation in *C. elegans*

Although an emerging set of studies investigated genetic variation in other species within the *Caenorhabditis* genus [20–23], we will concentrate on work performed only in *C. elegans*. Hodgkin and Doniach [24] were among the first to study natural variation in *C. elegans*, reporting variation among 32 wildtypes for copulatory plug formation which relied on transposon (*Tc1*) polymorphisms. This work was followed by mapping single nucleotide polymorphisms (SNPs) and microsatellite loci among various wildtypes, showing that worldwide, genetic diversity exists across wild isolates [25–29]. The most diverged wildtype strains identified to date [30], N2 and CB4856, exhibit a genetic variation of one SNP per 840 bp in the nuclear genome [26,31]. Another *C. elegans* wildtype, CB4857, has a somewhat lower SNP diversity of one SNP per 1445 bp [25]. Compared to *Drosophila melanogaster* (one SNP per 175 bp) [32], the genetic variation among different *C. elegans* wildtypes is relatively low, likely because of selfing coupled with broadscale migration [33]. By comparison, *C. remanei*, an outcrossing species, harbors substantially more nucleotide variation [20,21].

Caenorhabditis elegans genomic variation research has recapitulated some interesting findings from human studies. Recent work has identified an unexpectedly large number of insertions and deletions in the human genome, ranging from binary insertion/deletion events of short sequences (majority < 10 bp in size) to large-scale structural variation (50 kb to 5 Mb) [34]. The frequency (up to 20%) [35] of these types of polymorphisms suggests that they might be as important as SNPs ($\sim$5% frequency in the human genome) [36] for generating natural phenotypic variation in humans. A comparison of three *C. elegans* isolates, N2, CB4856 and JU258, showed many deletions in CB4856 and JU258 compared with those in N2. Of a total of 20 873 genes in N2, 531 and 670 are deleted in the CB4856 and JU258 strains, respectively [37]. This finding points at large differences in the functional gene content among these isolates. The frequency of deletions was especially high for the MATH-BTB, F-box (both are protein–protein interaction motifs), C-type lectin (proteins involved in sugar recognition) and Srz chemoreceptor (proteins involved in chemosensation) families, suggesting long-term evolutionary stability in these gene families [37].

Although we are beginning to understand the underlying genetic variation within *C. elegans* populations, less is known about phenotypic variation, such as the response of fecundity to changes in temperature in seven natural isolates; at high temperatures (maximum, 25 °C), fecundity decreased in all strains, primarily from a reduction in the number of functional sperm [38]. Similarly, Jovelin *et al.* [21] showed that there is a low level of genetic variation for chemosensory behavior among natural isolates, and variation is also observed in male competitive ability [39].

Thus far, however, few studies have managed to tie natural variation across a genetic pathway directly to phenotypic effects. An amino acid change in the neuropeptide Y–related receptor, NPR-1, modifies foraging behavior between two wildtypes [40]. Natural variation in *npr-1* reorganizes the relative contribution of different sensory inputs to behavioral output [41]. Most ectotherms grow larger at lower temperatures, a phenomenon known as the temperature–size rule. Wildtype N2 complies with the temperature–size rule, whereas another wildtype strain, CB4856, defies the rule [42]. Likewise, an SNP within CB4856 *tra-3* (sexual transformer, ortholog to human calpain-5) leads to a F96L substitution in its protein product; this alteration attenuates the ability of CB4856 to grow larger at low temperature.

One way to try to close the gap between molecular and phenotypic variation is to draw on the tremendous genomic resources available in *C. elegans* to study functional intermediates. For example, intraspecific genetic variation has been studied at the genomic level by aiming to characterize complex regulatory pathways and entire networks that control gene expression. Work using *C. elegans* has been at the forefront of such efforts, building on some of the first large-scale analyses of genetic networks using comprehensive microarray experiments [43] and yeast two-hybrid

analyses [44]. Natural variation in novel worm gene regulation pathways has also been studied on a genome-wide level using quantitative trait loci (QTL) mapping (also termed 'genetical genomics'). Experiments using an N2 × CB4856 recombinant inbred population in combination with high-throughput gene transcription profiling for mapping transcriptional regulation suggests that much of the transcription variation results from to the action of genes located elsewhere in the genome [45]. The comparison of gene transcription variation among natural isolates to a neutral model of transcriptome evolution suggests that the evolution of transcriptional change for many *C. elegans* genes has been dominated by strong stabilizing selection [46].

Genetic mapping in different backgrounds of *Caenorhabditis* spp.

Although genetic diversity among *C. elegans* isolates is relatively low, efficient mapping strategies are possible. The mouse *Mus musculus* serves as a good example for this, because its classical inbred strains have even lower levels of genetic variation (estimated to be one SNP per 1000 bp between pairs of inbred strains [47]), yet they have successfully been used for high-resolution mapping of complex traits [48]. However, compared with the mouse, fewer mapping populations are available for *Caenorhabditis* spp. Johnson and Wood [49,50] were the first to develop such tools for lifespan analysis in *C. elegans*; they created a set of recombinant inbred lines (RILs) from a cross between wildtypes N2 and Bergerac. Ebert *et al.* [51] also crossed the same parental strains followed by random mating for four generations and inbreeding for seven generations. Others [52–56] have used a comparable population to map various traits including halothane sensitivity, reproduction, lifespan and growth. The genetic markers used in these studies were transposable elements (Tc1). However, these Tc1 markers do not evenly cover the full genome, and QTL detection is biased to regions where markers are present. Ayyadevara *et al.* [57] used a Bergerac × RC301 recombinant inbred panel also with Tc1 markers to map loci associated with lifespan. However, many of the RILs obtained from a cross with Bergerac are difficult to maintain, presumably because of the detrimental effects of Tc1 insertions on gene function [58].

Complete marker coverage of wildtype genotypes now is available in an N2 × CB4856 recombinant inbred population [45], which contains 80 RILs and a SNP map of 121 evenly spaced markers. The average distance between two markers is 835 kbp or 2.38 cM, with an average chromosomal coverage of 95.9%. This population of genotyped RILs was used to identify QTLs associated with reproduction, maturation time, body size and gene expression [42,45,59,60]. More than 900 RILs obtained from this cross await genotyping and are available for phenotyping. Similarly, a recombinant inbred population consisting of 40 RILs obtained from a cross between wildtypes N2 and DR1350 was used to map genetic variation in dauer larve formation [61].

In *C. briggsae*, RILs have also been constructed from a cross between the AF16 and HK104 strains for mapping male tail ray pattern variation [22]. Allelic variants affecting the size of the first intron of *Cb-egl-5* are likely candidates associated with the morphology of the rays. Similar lines have been used to generate a final assembly of the *C. briggsae* genome [62].

In addition to the RIL populations, near isogenic lines (NILs) constitute a powerful resource to study gene polymorphisms responsible for natural variation in complex phenotypes (Box 2). The genome of a NIL is composed of the recipient (or background) genome, contributed by one of the parental strains, and a short, homozygous donor (introgressed) segment contributed by another, genetically distinct, parental strain. Genome-wide libraries of NILs are available for many model species, including *M. musculus* and *Arabidopsis thaliana* [63,64], and they have proved effective and efficient in the identification of natural polymorphisms for many traits [65–68]. Only 11 NILs have thus far been developed for *C. elegans* [54,57]. These particular NILs were derived from a parental cross other than N2 × CB4856 (the genetically most diverse strains); however, an N2 × CB4856 RIL population [45] can be used to develop a set of NILs. Ideally, each NIL would carry a different single marker-defined, homozygous genome segment of strain CB4856 introgressed into the N2 background. All NILs together would cover the full CB4856 genome.

The detection of natural polymorphisms is facilitated by the existence of completely sequenced genomes. Thus far, only the genomes of the *C. elegans* N2 [69] and *C. briggsae* AF16 genomes have been fully assembled [70]. Other *C. elegans* strains and other *Caenorhabditis* species are being sequenced (see (http://www.genome.wustl.edu/ for an update on progress). The availability of these genomes will enable the study of the genetic mechanisms underlying phenotypic differences among strains. In contrast to the self-fertilizing congeners *C. elegans* and *C. briggsae*, the out-crosser *C. remanei* exhibits a 20-fold higher diversity in orthologous loci [71]. *C. remanei*, whose genome is currently being sequenced, therefore offers great potential for mapping complex traits in recombinant populations and for gene mapping based on natural polymorphisms. Viable populations can be inbred for up to 10 generations using full-sibling mating [3], after which population viability decreases. This is a rapid decline compared with 40 generations of inbreeding that is possible in *D. melanogaster* [72]. This high level of inbreeding depression makes line construction more problematic within *C. remanei* and highlights the value of easy selfing within *C. elegans* and *C. briggsae*.

Caenorhabditis spp. as models for studying genetic variation in complex human disease pathways

Complex human diseases including cancer and neurodegenerative diseases arise from interactions between many different genes and the environment. Evidence is mounting that suggests that genetic background has a profound impact on a wide variety of complex disease phenotypes. For instance, different genotypes are associated with somatic gene expression data in human breast tumors [73]. However, the exact genetic mechanisms underlying these background modifiers in humans are largely unknown.

Box 2. Mapping natural polymorphisms using near-isogenic lines and mutants

Natural polymorphisms can be mapped in *Caenorhabditis elegans* using near isogenic lines (NILs) and mutants. A hypothetical N2/CB4856 NIL library and clumping behavior is used here as an example. N2 is a solitary feeder, whereas CB4856 feeds in groups (termed clumping behavior) [40].

Complementation with a mutant allele

This strategy allows mapping putative quantitative trait loci (QTLs) to candidate genes [42]. An NIL with the donor segment covering the QTL region for clumping behavior is crossed with a candidate gene mutant (Figure Ia). In the F1 generation, the mutation (*m1*) is combined with the inrogression in *trans*, whereas the genetic background is isogenic (N2). Finding a phenotype different from that observed in N2 implies that the mutation does not complement the introgression. This finding demonstrates that the QTL for clumping behavior maps to the candidate gene.

Mapping natural polymorphisms underlying background effects

The interactions between a mutation and genetic background can form the basis for the identification of natural polymorphisms. As an example, a mutation of interest (*m1*) from the N2 strain is introgressed into the CB4856 strain (Figure Ib). The difference in phenotype between the newly created line and the mutant in N2 indicates a genetic background effect. This effect can result from a natural polymorphism in a single gene or in multiple genes that interact with the mutated gene. To identify these genes, a mapping procedure using NILs can be applied. There are two main ways to proceed: (i) focus on the candidate genes known to interact with or operate in the same pathway as the mutated gene or (ii) perform genome-wide interaction mapping, for which no prior knowledge of candidate genes is required. In both cases, the mutant (in the N2 background) is crossed with NILs (N2/CB4856). In the candidate gene approach, the crosses are made with NILs carrying donor segments at the locations corresponding to the candidate genes. In the genome-wide approach, a genome-wide library of NILs is used for the crosses. The derived strains, homozygous for the mutation and for the donor genomic segment, are subsequently phenotyped. The strains that display a different phenotype than the 'N2 mutant' define, through the position of the donor segment, the region where the interacting naturally polymorphic gene is localized. The fine-mapping by back-crossing NILs with N2 and the involvement of candidate genes can be assessed using standard genetic approaches.

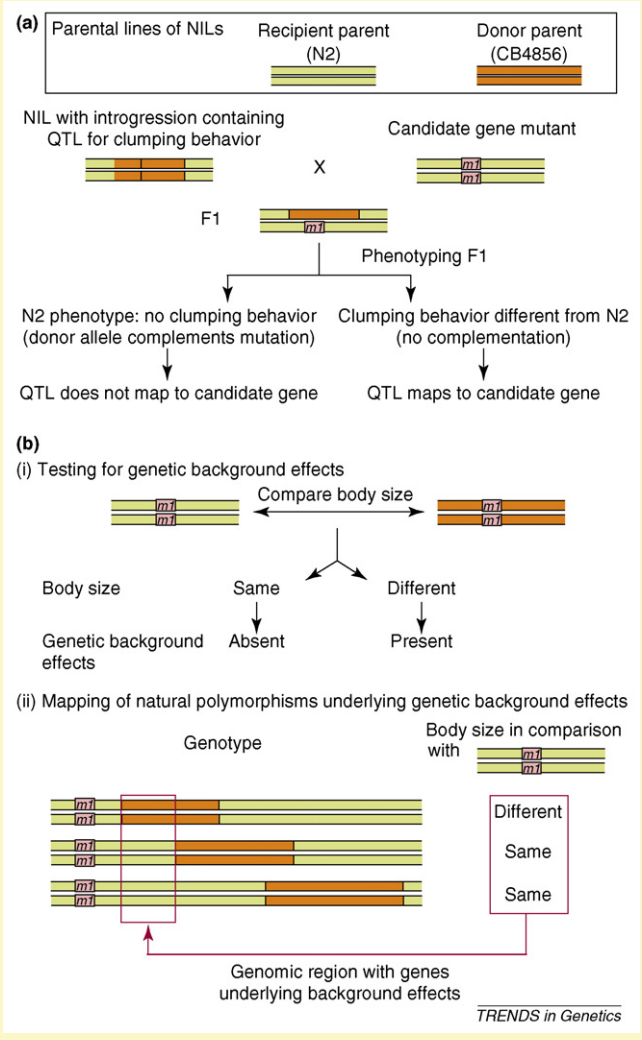

Figure I. Mapping natural polymorphisms using near-isogenic lines and mutants. The genotypes are indicated with color: green indicates N2 genomic background and orange bars indicate CB4856 donor segments. Mutations are indicated as pink boxes labeled '*m1*'. **(a)** Complementation with a mutant allele. **(b)** Mapping natural polymorphisms underlying background effects.

The mouse has been used extensively as a quantitative genetic model to study pathways associated with human diseases and disorders [74]. These studies focus on the detection of modifier genes that are capable of altering the phenotype of a mutant gene. A clear example of a modifier gene was reported for tubby hearing 1 gene (*moth-1*), a gene in which mutations lead to hearing loss [75]. These researchers positionally cloned a QTL, the modifier of *moth-1*, and found that wildtype alleles from three different strains protect *moth-1* (a SNP in a natural allele) mice from hearing loss. These polymorphisms change the binding efficiency of MTAP1A (microtubule-associated protein 1a) to postsynaptic density molecule 95 (PSD95), a core component in the structure of synapses. This finding indicates that the observed polymorphisms are functionally important. Using a similar approach, QTLs associated with hematopoietic stem cell (HSC) numbers in a mouse RIL population were mapped [76]. Using a combination of fine mapping with an NIL approach, gene expression and protein analysis, they identified *Latexin* (*Lxn*), whose expression inversely correlates with HSC number. The differential transcription of *Lxn* was associated with the allelic differences among the RILs. They identified SNP clusters upstream of the *Lxn* transcriptional start site; at least two were associated with potential binding sites for stem cell regulatory transcription factors. Thus, promoter polymorphisms between the two parental alleles might affect *Lxn* gene expression and consequently influence the number of HSCs.

These mouse examples show that background effects can be mechanistically studied. This holds great promise for *C. elegans* because its transparent body and well-described human disease pathways, in combination with a defined cell lineage and completed genome sequence, allow detailed mechanistic studies of genetic background effects. As an example, it would be useful to study genetic background effects in pathways underlying vulva development in *C. elegans*, because these pathways are strongly associated with human cancer formation [77] (Figure 1).

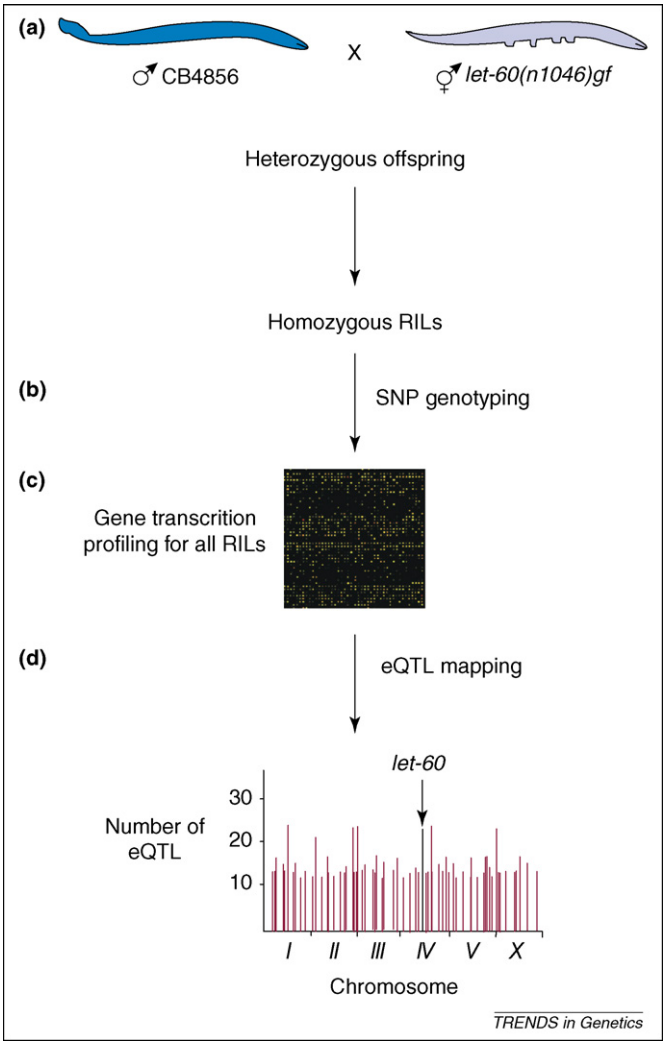

Figure 1. Mapping natural alleles linked with the *let-60* pathway. Shown is an example of an experimental setup illustrating how genes and natural genetic background alleles linked to *let-60* can be detected. In *Caenorhabditis elegans* strain N2, the *let-60(n1046)* gain-of-function (*gf*) induced mutant is similar to Ras mutations detected in many human cancer cells [79]. In *C. elegans,* this mutation leads to a multiple vulva phenotype with variable protrusions. *Let-60(n1046)gf* hermaphrodites usually have variably sized ventral protrusions either anterior or posterior to a functional vulva. *Let-60(n1046)gf* is a homozygous viable mutant of the wildtype N2 and mates successfully with wildtype CB4856 males. To identify candidate genes and natural alleles affecting *let-60,* the following steps can be taken: **(a)** Crossing wildtype CB4856 with *let-60(n1046)gf*. The F1 heterozygotes can be easily selfed to obtain a set of homozygous recombinant inbred lines (RILs) [80]. Each RIL is a recombinant of both parents and one might expect to see genetic variation in gene transcription [45]. **(b)** Genotyping CB4856 × *let-60(n1046)gf* RILs. For mapping, RILs are required, which are all homozygous for *let-60* (this can be assessed by performing rapid SNP genotyping of *let-60*). The RILs can than be genotyped across all chromosomes with evenly spaced SNP markers. **(c)** RIL phenotyping. RILs are phenotyped for gene transcription profiling using whole genome microarrays covering all *C. elegans* transcripts [45]. **(d)** Quantitative trait loci (QTL) mapping. Expression QTLs (eQTLs) can be mapped using a genetical genomics approach [45]. As an example, we highlight several gene transcripts which map to *let-60* (located on chromosome *IV*) in the CB4856 genetic background. The figure shows that ~23 transcripts map to *let-60,* and most likely, many of these transcripts have not been detected using classic mutagenic or knock-down studies.

Concluding remarks and future perspectives

In this review, we highlighted the importance of understanding the role that genetic background plays in structuring the function of various signaling pathways. How might these ideas be implemented in future work conducted within the *Caenorhabditis* system? We recommend an initial focus on developmental processes, including dauer formation and vulva development, that have proven particularly useful for genetic dissection in *Caenorhabditis* spp. because of the ease and detail with which they can be studied and because they are regulated by conserved signaling pathways (Box 3). Two pathways controlling dauer formation involve an insulin-like cascade and transforming growth factor-β (TGF-β) signaling. Vulva development is regulated by four signaling pathways: an epidermal growth factor (EGF)/renin-angiotensin system (RAS)/mitogen-activated protein kinase (MAPK) pathway, the Notch/LIN-12 pathway, a Wnt pathway and the NuRD nucleosomal remodeling complex [78]. Each of these pathways has been implicated in human diseases and are targets for drug development: the insulin pathway in diabetes, obesity and aging; the EGF, TGF-β, Wnt, NuRD and Notch pathways in cancer. For each pathway, a variety of induced mutant alleles in different signaling components have been identified and studied in N2 animals. Introgressing some of these alleles into other natural *C. elegans* isolates or using RNAi to induce loss-of-function phenotypes in a panel of wild strains could provide powerful entry points to study how natural variation can modify these pathways. Both vulva induction and dauer formation can be studied as quantitative traits in which either the number of induced cells of each type or the number of dauers formed are counted.

Box 3. Future directions for the analysis of worm pathways associated with human disease and aging

What are the phenotypic effects of mutations in different wildtype backgrounds?

Perhaps the first and most straightforward question is how mutations derived in N2 'behave' in different backgrounds. For instance, what is the phenotype of mutations in genes affecting lifespan and genes involved in neurodegenerative pathways when they are expressed in different wildtypes? If the ultimate goal is to extrapolate results from worm to human, it is important to study the interaction between wildtype genetic backgrounds and the phenotype of induced mutations.

Which wildtype alleles affect lifespan in natural populations?

Caenorhabditis elegans wildtype N2 has been a workhorse in aging research. More than 70 genes in different pathways have been identified and characterized, with genes in the insulin/insulin-like growth factor signaling pathway being the best studied [84]. The forkhead transcription factor DAF-16 has been identified as the key regulator of this pathway; it can modify lifespan in response to environmental and gonadal stimuli. It is thus far unknown, however, whether natural alleles in *daf-16* are also associated with lifespan in natural worm populations. To answer this question, association mapping is required where the aim is to associate *daf-16* polymorphisms with variation in lifespan among a set of different wildtype populations.

Can we unravel the mechanistic details of wildtype alleles affecting pathways related to human disease?

At the moment, the phenotype of almost all induced mutations and their interactions are studied in the wildtype N2. Although research in yeast has shown genome-wide interactions or epistasis between natural polymorphisms affecting gene expression of many pathways [85], this information is lacking for disease pathways in worms. In aging research, for example, it is unclear how and to what extent natural alleles affect *daf-16* transcription. It also remains unknown how natural alleles affect the Ras and Wnt pathways that direct vulva development.

Caenorhabditis elegans has proven to be one of the most powerful genetic systems yet studied. Part of that power is certainly derived from the fact that nearly every researcher uses the same genetic background for their studies. This practice facilitates sharing of mutants, comparison of results across laboratories, and uniformity in the overall interpretation of genetic effects. Why advocate investigating the effects of other genetic backgrounds? First, it is clear that all laboratories do not in fact study the same strain, because unless steps are taken to ensure that only frozen stocks are used for every experiment, N2 does evolve in the laboratory. This has not been an issue for most mutations of large effects but is detectable in more subtle phenotypes. As we move from the dissection of genetic pathways using mutagenesis and knockdown approaches toward an understanding of how variation within these pathways promotes differences among individuals, the importance of genetic background becomes more central. These issues have received a great deal of attention within the mouse and fly communities and are vitally important for understanding the complex basis for disease and other traits within human populations. The time has come to use the power of the nematode model system to unravel the genetic basis of such complex effects.

References

1 Maupas, E. (1899) La mue e l'enkystement chez les nématodes. *Arch. Zool. Exp. Gén.* 7, 563–628
2 Johnson, T.E. and Hutchinson, E.W. (1993) Absence of strong heterosis for life-span and other life-history traits in *Caenorhabditis elegans*. *Genetics* 134, 465–474
3 Dolgin, E.S. *et al.* (2007) Inbreeding and outbreeding depression in *Caenorhabditis* nematodes. *Evolution Int. J. Org. Evolution* 61, 1339–1352
4 Kaletta, T. and Hengartner, M.O. (2006) Finding function in novel targets: *C. elegans* as a model organism. *Nat. Rev. Drug Discov.* 5, 387–399
5 Braeckman, B.P. and Vanfleteren, J.R. (2007) Genetic control of longevity in *C. elegans*. *Exp. Gerontol.* 42, 90–98
6 Schafer, W.R. (2005) Deciphering the neural and molecular mechanisms of *C. elegans* behavior. *Curr. Biol.* 15, R723–R729
7 Saito, R.M. and Van den Heuvel, S. (2002) Malignant worms: what cancer research can learn from *C. elegans*. *Cancer Invest.* 20, 264–275
8 Kenyon, C. (2005) The plasticity of aging: Insights from long-lived mutants. *Cell* 120, 449–460
9 Pinkston-Gosse, J. and Kenyon, C. (2007) DAF-16/FOXO targets genes that regulate tumor growth in *Caenorhabditis elegans*. *Nat. Genet.* 39, 1403–1409
10 Sundaram, M. and Greenwald, I. (1993) Suppressors of a *lin-12* hypomorph define genes that interact with both *lin-12* and *glp-1* in *Caenorhabditis elegans*. *Genetics* 135, 765–783
11 Wittenburg, N. *et al.* (2000) Presenilin is required for proper morphology and function of neurons in *C. elegans*. *Nature* 406, 306–309
12 Levitan, D. and Greenwald, I. (1995) Facilitation of *lin-12*-mediated signalling by *sel-12*, a *Caenorhabditis elegans* S182 Alzheimer's disease gene. *Nature* 377, 351–354
13 Han, M. and Sternberg, P.W. (1990) *Let-60*, a gene that specifies cell fates during *C. elegans* vulval induction, encodes a RAS protein. *Cell* 63, 921–931
14 Gems, D. *et al.* (1998) Two pleiotropic classes of *daf-2* mutation affect larval arrest, adult behavior, reproduction and longevity in *Caenorhabditis elegans*. *Genetics* 150, 129–155
15 Kiontke, K. and Sudhaus, W. (2006) Ecology of *Caenorhabditis species* (January 09, 2006). In *WormBook* (The *C. elegans* Research Community, ed.), The *C. elegans* Research Community
16 Rudel, D. and Kimble, J. (2002) Evolution of discrete Noth-like receptors from a distant gene duplication in *Caenorhabditis*. *Evol. Dev.* 4, 319–333
17 Podbilewicz, B. (2003) How does a cell anchor and invade an organ? *Dev. Cell* 5, 5–7
18 Hill, R.C. *et al.* (2006) Genetic flexibility in the convergent evolution of hermaphroditism in *Caenorhabditis* nematodes. *Dev. Cell* 10, 531–538
19 Félix, M.A. (2007) Cryptic quantitative evolution of the vulva intercellular signaling network in *Caenorhabditis*. *Curr. Biol.* 17, 103–114
20 Graustein, A. *et al.* (2002) Levels of DNA polymorphism vary with mating system in the nematode genus *Caenorhabditis*. *Genetics* 161, 99–107
21 Jovelin, R. *et al.* (2003) Molecular evolution and quantitative variation for chemosensory behaviour in the nematode genus *Caenorhabditis*. *Mol. Ecol.* 12, 1325–1337
22 Baird, S.E. *et al.* (2005) The genetics of ray pattern variation in Caenorhabditis briggsae. *BMC Evol. Biol.* 5, 3
23 Cutter, A.D. *et al.* (2006) Patterns of nucleotide polymorphism distinguish temperate and tropical wild isolates of *Caenorhabditis briggsae*. *Genetics* 173, 2021–2031
24 Hodgkin, J. and Doniach, T. (1997) Natural variation and copulatory plug formation in *Caenorhabditis elegans*. *Genetics* 146, 149–164
25 Koch, R. *et al.* (2000) Single nucleotide polymorphisms in wild isolates of *Caenorhabditis elegans*. *Genome Res.* 10, 1690–1696
26 Wicks, S.R. *et al.* (2001) Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. *Nat. Genet.* 28, 160–164
27 Sivasundar, A. and Hey, J. (2003) Population Genetics of *Caenorhabditis elegans*: the paradox of low polymorphism in a widespread species. *Genetics* 163, 147–157
28 Cutter, A.D. (2006) Nucleotide polymorphism and linkage disequilibrium in wild populations of the partial selfer *Caenorhabditis elegans*. *Genetics* 172, 171–184
29 Barriere, A. and Feìlix, M.A. (2005) High local genetic diversity and low outcrossing rate in *Caenorhabditis elegans* natural populations. *Curr. Biol.* 15, 1176–1184
30 Denver, D.R. *et al.* (2003) Phylogenetics in *Caenorhabditis elegans*: an analysis of divergence and outcrossing. *Mol. Biol. Evol.* 20, 393–400
31 Swan, K.A. *et al.* (2002) High-throughput gene mapping in *Caenorhabditis elegans*. *Genome Res.* 12, 1100–1105
32 Berger, J. *et al.* (2001) Genetic mapping with SNP markers in *Drosophila*. *Nat. Genet.* 29, 475–481
33 Phillips, P.C. (2006) One perfect worm. *Trends Genet.* 22, 405–407
34 Sharp, A.J. *et al.* (2006) Structural variation of the human genome. Completing the map of human genetic variation. *Annu. Rev. Genomics Hum. Genet.* 7, 407–442
35 The Human Genome Structural Variation Working Group (2007) Completion of the map of human genetic variation. *Nature* 447, 161–165
36 Hinds, D.A. *et al.* (2005) Whole-genome patterns of common DNA variation in three human populations. *Science* 307, 1072–1079
37 Maydan, J.S. *et al.* (2007) Efficient high-resolution deletion discovery *in Caenorhabditis elegans* by array comparative genomic hybridization. *Genome Res.* 17, 337–347
38 Harvey, S.C. and Viney, M.E. (2007) Thermal variation reveals natural variation between isolates of *Caenorhabditis elegans*. *J. Exp. Zool. (Mol. Dev. Evol.)* 308B, 409–416
39 Teotonio, H. *et al.* (2007) Genetic variation for outcrossing among *Caenorhabditis elegans* isolates. *Evolution Int. J. Org. Evolution* 60, 1300–1305
40 De Bono, M. and Bargmann, C. (1998) Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. *Cell* 94, 679–689
41 Rogers, C. *et al.* (2003) Inhibition of *Caenorhabditis elegans* social feeding by FMRFamide-related peptide activation of NPR-1. *Nat. Neurosci.* 6, 1178–1185
42 Kammenga, J.E. *et al.* (2007) A *Caenorhabditis elegans* wild-type defies the temperature-size rule owing to a single nucleotide polymorphism in *tra-3*. *PLoS Genet.* 3, 358–366
43 Kim, S.K. *et al.* (2001) A gene expression map for *Caenorhabditis elegans*. *Science* 293, 2087–2092
44 Li, S. *et al.* (2004) A map of the interactome network of the metazoan *C. elegans*. *Science* 303, 540–543
45 Li, Y. *et al.* (2006) Mapping determinants of gene expression plasticity by genetical genomics in *Caenorhabditis elegans*. *PLoS Genet.* 2, 2155–2166

46 Denver, D.R. *et al.* (2005) The transcriptional consequences of mutation and natural selection in *Caenorhabditis elegans*. *Nat. Genet.* 37, 544–548
47 Lindblad-Toh, K. *et al.* (2000) Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. *Nat. Genet.* 24, 381–386
48 Yalcin, B. *et al.* (2004) Genetic dissection of a behavioral quantitative trait locus shows that Rgs2 modulates anxiety in mice. *Nat. Genet.* 36, 1197–1202
49 Johnson, T.E. and Wood, W.B. (1982) Genetic analysis of life-span in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U. S. A.* 79, 6603–6607
50 Johnson, T.E. (1987) Aging can be genetically dissected into component processes using long-lived lines of *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U. S. A.* 84, 3777–3781
51 Ebert, R.H. *et al.* (1993) Longevity-determining genes in *Caenorhabditis elegans*: chromosomal mapping of multiple noninteractive loci. *Genetics* 135, 1003–1010
52 Shook, D.R. *et al.* (1996) Mapping quantitative trait loci affecting life history traits in the nematode *Caenorhabditis elegans*. *Genetics* 142, 801–817
53 Van Swinderen, B. *et al.* (1997) Quantitative trait loci controlling halothane sensitivity in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U. S. A.* 94, 8232–8237
54 Knight, C.G. *et al.* (2001) Testing life-history pleiotropy in *Caenorhabditis elegans*. *Evolution Int. J. Org. Evolution* 55, 1795–1804
55 Shook, D.R. and Johnson, T.E. (1999) Quantitative trait loci affecting survival and fertility-related traits in *Caenorhabditis elegans* show genotype-environment interactions, pleiotropy and epistasis. *Genetics* 153, 1233–1243
56 Ayyadevara, S. *et al.* (2003) Genetic loci modulating fitness and life span in *Caenorhabditis elegans*: categorical trait interval mapping in CL2a x Bergerac-BO recombinant-inbred worms. *Genetics* 163, 557–570
57 Ayyadevara, S. *et al.* (2001) Genetic mapping of quantitative trait loci governing longevity of *Caenorhabditis elegans* in recombinant-inbred progeny of a Bergerac-BO × RC301 interstrain cross. *Genetics* 157, 655–666
58 Eide, D. and Anderson, P. (1985) Transpositioin of Tc1 in the nematode *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. U. S. A.* 82, 1756–1760
59 Gutteling, E.W. *et al.* (2007) Environmental influence on the genetic correlations between life-history traits in *Caenorhabditis elegans*. *Heredity* 98, 206–213
60 Gutteling, E.W. *et al.* (2007) Mapping phenotypic plasticity and genotype-environment interactions affecting life history traits in *Caenorhabditis elegans*. *Heredity* 98, 28–37
61 Viney, M.E. *et al.* (2003) Variation in *Caenorhabditis elegans* dauer larva formation. *Dev. Growth Differ.* 45, 389–396
62 Hillier, L.W. *et al.* (2007) Comparison of *C. elegans* and *C. briggsae* genome sequences reveals extensive conservation of chromosome organization and synteny. *PLoS Biol.* 5, e167
63 Iakoubova, O.A. *et al.* (2001) Genome-tagged mice (GTM): Two sets of genome-wide congenic strains. *Genomics* 74, 89–104
64 Keurentjes, J.J.B. *et al.* (2007) Development of a near-isogenic line population of *Arabidopsis thaliana* and comparison of mapping power with a recombinant inbred line population. *Genetics* 175, 891–905
65 Fridman, E. *et al.* (2002) Two tightly linked QTLs modify tomato sugar content via different physiological pathways. *Mol. Genet. Genomics* 266, 821–826
66 Semel, Y. *et al.* (2006) Overdominant quantitative trait loci for yield and fitness in tomato. *Proc. Natl. Acad. Sci. U. S. A.* 103, 12981–12986
67 Ashikari, M. *et al.* (2005) Cytokinin oxidase regulates rice grain production. *Science* 309, 741–745
68 Clee, S.M. *et al.* (2006) Positional cloning of Sorcs1, a type 2 diabetes quantitative trait locus. *Nat. Genet.* 38, 688–693
69 The *C. elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282, 2012–2018
70 Stein, L.D. *et al.* (2003) The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PLoS Biol.* 1, 166–192
71 Cutter, A.D. *et al.* (2006) High nucleotide polymorphism and rapid decay of linkage disequilibrium in wild populations of *Caenorhabditis remanei*. *Genetics* 174, 901–913
72 Swindell, W.R. and Bouzat, J.L. (2006) Inbreeding depression and male survivorship in *Drosophila*: Implications for senescence theory. *Genetics* 172, 317–327
73 Kristensen, V.N. *et al.* (2006) Genetic variation in putative regulatory loci controlling gene expression in breast cancer. *Proc. Natl. Acad. Sci. U. S. A.* 103, 7735–7740
74 Flint, J. *et al.* (2005) Strategies for mapping and cloning quantitative trait genes in rodents. *Nat. Rev. Genet.* 6, 271–286
75 Ikeda, A. *et al.* (2002) Microtubule-associated protein 1A is a modifier of tubby hearing (*moth1*). *Nat. Genet.* 30, 401–405
76 Liang, Y. *et al.* (2007) The quantitative trait gene latexin influences the size of the hematopoietic stem cell population in mice. *Nat. Genet.* 39, 178–188
77 Jannot, G. and Simard, M.J. (2006) Tumour-related microRNAs functions in *Caenorhabditis elegans*. *Oncogene* 25, 6197–6201
78 The *C. elegans* Research Community (2008) *WormBook*, The C. *elegans* Research Community
79 Barbacid, M. (1987) Ras genes. *Annu. Rev. Biochem.* 56, 779–827
80 Nanji, M. *et al.* (2005) LET-60 RAS modulates effects of insulin/IGF-1 signaling on development and aging in *Caenorhabditis elegans*. *Aging Cell* 4, 235–245
81 Brenner, S. (1974) The genetics of *Caenorhabditis elegans*. *Genetics* 77, 71–94
82 Sulston, J.E. and Horvitz, H.R. (1977) Post-embryonic cell lineages of the nematode *Caenorhabditis elegans*. *Dev Biol.* 56, 110–156
83 Fire, A. *et al.* (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–811
84 Braeckman, B.P. and Vanfleteren, J.R. (2006) Genetic control of longevity in *C. elegans*. *Exp. Gerontol.* 42, 90–98
85 Brem, R.B. *et al.* (2005) Genetic interactions between polymorphisms that affect gene expression in yeast. *Nature* 436, 701–703