Fifteen Years of Evolutionary Genomics in *Caenorhabditis elegans*

Richard Jovelin, *Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada*

Alivia Dey, *Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada*

Asher D Cutter, *Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada*

Advanced article

Article Contents

- Introduction
- Genome Organisation
- Gene Number and Functional Diversification
- Regulation of Genome Function
- Mutation Rates
- *C. elegans* in Nature
- Conclusion

Online posting date: 15th May 2013

The nematode worm *Caenorhabditis elegans*, introduced by Sydney Brenner for the genetic analysis of nervous system formation, is now a powerful model organism for studying nearly all aspects of biology, from development to diseases to evolution. Sequencing and analysis of the worm genome revealed intriguing nonrandom patterns of genome organisation and unusual features such as abundant operons. Surprising upon first discovery, worms and humans have a similar number of genes that are comprised of a similar proportion of transcription factors to regulate their genomes. However, differences in small ribonucleic acid content may contribute to differences in organismal complexity. In nature, the bacterivorous *C. elegans* is found primarily on rotting vegetation in temperate regions across the world. Natural selection, combined with the low effective recombination rate associated with selfing, strongly reduces nucleotide variation across the genome, yielding similarly low polymorphism to other selfing hermaphrodite species of *Caenorhabditis*. The genus *Caenorhabditis* provides a superb model system for ecological and evolutionary genetics, benefiting from *C. elegans* tools and information when applied to investigations of species with a higher polymorphism and better known ecological context.

Introduction

Following major advances in molecular biology in the 1950s and 1960s – the discovery of the DNA structure, the nature of the genetic code and the basis of gene regulation – experimental embryology and genetics merged into the field of developmental genetics to tackle the principles of multicellular organisation. Long-standing issues such as the morphogenesis of an adult phenotype from an undifferentiated egg cell and the genetic programmes regulating organ formation could now be approached with the successful mechanistic and reductionist methodology of molecular biology. As developmental genetics matured and later reconnected with modern evolutionary theory in the 1980s, evolutionary biologists and developmental biologists then focused to identify the genes causing morphological variation and diversity in the natural world (evo-devo).

The fruit fly was the model of choice for investigating the genetics of development because of the long history in *Drosophila* genetics initiated in the early twentieth century by Thomas Hunt Morgan, and later proved to generate fundamental insights into the principles of early animal embryogenesis. However, Sydney Brenner wanted a simple organism to decipher the genetic control of the nervous system. He settled on *Caenorhabditis elegans* because this small and transparent nematode is easy to rear and to manipulate in the laboratory, has a short generation time, a large brood size, a tractable cell lineage and only 302 neurons. A peculiar feature that makes *C. elegans* so amenable to genetic analysis is its mating system: *C. elegans* is androdioecious with most individuals being selfing hermaphrodites that can also mate with the few males present in populations. This mostly selfing mode of reproduction facilitates the identification and maintenance of new mutations identified in genetic screens, and crosses with males enable the exchange of mutations on different genetic backgrounds.

The neurobiology, neuron connectivity and the genetic mechanisms of neuron differentiation are now relatively well understood (Bargmann, 1998; Hobert, 2008). The worm community has grown to several hundred laboratories since Brenner first introduced *C. elegans* as a model organism, placing the worm at the forefront of biological research with Nobel prizes awarded to six researchers in the

eLS subject area: Evolution & Diversity of Life

How to cite:
Jovelin, Richard; Dey, Alivia; and Cutter, Asher D (May 2013) Fifteen Years of Evolutionary Genomics in *Caenorhabditis elegans*. In: eLS. John Wiley & Sons, Ltd: Chichester.
DOI: 10.1002/9780470015902.a0022897

last decade for pioneering work on cell apoptosis and organ development (2002), the discovery of ribonucleic acid (RNA) interference (2006) and the development of the green fluorescent protein for *in vivo* visualisation of gene activity (2008).

C. elegans eventually became the first metazoan to have a complete sequenced genome. The publication, soon thereafter, of the genome for the related species *Caenorhabditis briggsae* opened the door for comparative genomics. Most of the features that make *C. elegans* a great model organism for molecular genetics also attracted biologists interested in using the worm for evolutionary genetic studies. However, the peculiar mating system, although interesting in its own right from the standpoint of the evolution of sex, leaves pronounced marks in the genome sequence. This has turned out to be a double-edged sword – on the one hand comparisons between *C. elegans* and related gonochoristic (outcrossing male and female) species enable testing of predictions regarding the evolution of genomic features from population genetic models; on the other hand, the low nucleotide diversity and low effective recombination rate associated with selfing impose some limits for *C. elegans* as a model if the goal is to understand the connection between natural genetic and phenotypic variation. Moreover, little is known about the ecology of the worm, making it difficult to link life history traits and genomic patterns with environmental conditions. Here, we review recent findings in genetics and comparative genomics as they help us understand patterns of genome evolution, and highlight how the ease of genetics and wealth of functional information in *C. elegans* combined with the analysis of related species with higher levels of nucleotide diversity make the genus *Caenorhabditis* as a whole a powerful model system in evolutionary biology. Readers can refer to the Further Reading section for the recent detailed reviews on specific topics: ecology, epigenetics, molecular evolution, nucleotide diversity and quantitative genetics. **See also**: *Caenorhabditis elegans* as an Experimental Organism; *Caenorhabditis elegans* Genome Project; *Caenorhabditis elegans* Neural Development; Cell Death in *C. elegans*; Green Fluorescent Protein (GFP); RNA Interference (RNAi) and MicroRNAs

Genome Organisation

Sex determination in *Caenorhabditis* results from the ratio between five pairs of autosomes to the number of copies of the unique sex chromosome such that females, and hermaphrodites in selfing species, have two copies of the X chromosome ($2 \times 5A|\ XX$) and males have a single copy ($2 \times 5A|\ X\varnothing$). In *C. elegans*, this chromosomal difference is compensated by halving expression from both X chromosomes in hermaphrodites to equalise expression levels between sexes (Meyer, 2010) and by upregulating X-linked genes in both the sexes to restore expression levels between autosomes and the sex chromosome (Deng *et al.*, 2011).

Caenorhabditis elegans chromosomes are holocentric, a condition that may be common to all nematodes and that arose multiple times in animal and plant evolution (Melters *et al.*, 2012). Instead of having a single centromere to which kinetochore proteins attach spindle microtubules, *C. elegans* chromosomes distribute centromeric activity along their length in regions that are rendered permissive by the lack of germline and early embryonic transcription (Gassmann *et al.*, 2012). However, holocentric meiosis poses some challenges during the first meiotic division for releasing cohesion between homologous chromosomes while maintaining sister chromatid cohesion, orienting sister chromatids towards the same pole and restricting attachment of microtubules for proper chromosome segregation. As a solution to these issues, the site of crossing over acts as an organising centre to direct the proper loss of cohesion, and the adequate orientation and segregation of chromosomes (Schvarzstein *et al.*, 2010). The distribution and number of crossovers are tightly regulated in *C. elegans* such that a single crossover located near a chromosome end (on the 'arms') occurs during meiosis (Tsai *et al.*, 2008; Wagner *et al.*, 2010; Meneely *et al.*, 2012), with some aspects of this regulation involving different genes among the *Caenorhabditis* species (Meneely *et al.*, 2012). Thus, the mechanisms that ensure proper segregation of holocentric chromosomes also create a bias in the distribution of recombination events, generating much higher recombination rates in the chromosome arms of both *C. elegans* (Rockman and Kruglyak, 2009) and *C. briggsae* (Hillier *et al.*, 2007; Ross *et al.*, 2011).

This pattern of recombination has profound effects on the overall genome evolution. Fewer nucleotide and structural polymorphisms and stronger linkage disequilibrium are found in the chromosome centres, because their low recombination rate allows selection to remove linked variants over a broader genomic span (Cutter and Payseur, 2003; Rockman and Kruglyak, 2009; Cutter and Choi, 2010; Andersen *et al.*, 2012; **Figure 1**). Higher recombination rates on the arms also might drive the more abundant intrachromosomal rearrangements found in the arms and the greater conserved synteny in the chromosome centres (Hillier *et al.*, 2007; Rockman and Kruglyak, 2009; Ross *et al.*, 2011). Several other features such as the density of transposable elements, euchromatin marks, protein coding genes, repetitive elements, pseudogenes and operons also are distributed nonrandomly along the chromosomes (**Figure 1**), but it is less clear as to how the interactions between selection, linkage, mutation and genetic drift might generate these patterns.

A peculiar feature of the *C. elegans* genome, relative to most eukaryote genomes, is the large proportion of genes ($\sim$15%) that are transcribed as polycistronic pre-messenger RNAs (pre-mRNAs) from operons (Blumenthal *et al.*, 2002). Operon genes in *Caenorhabditis* are processed into individual mature mRNAs before translation through the *trans*-splicing machinery common in eukaryotes, but with the specialised SL2 splice leader (Guiliano and Blaxter, 2006). Operons are enriched in

eLS © 2013, John Wiley & Sons, Ltd. www.els.net

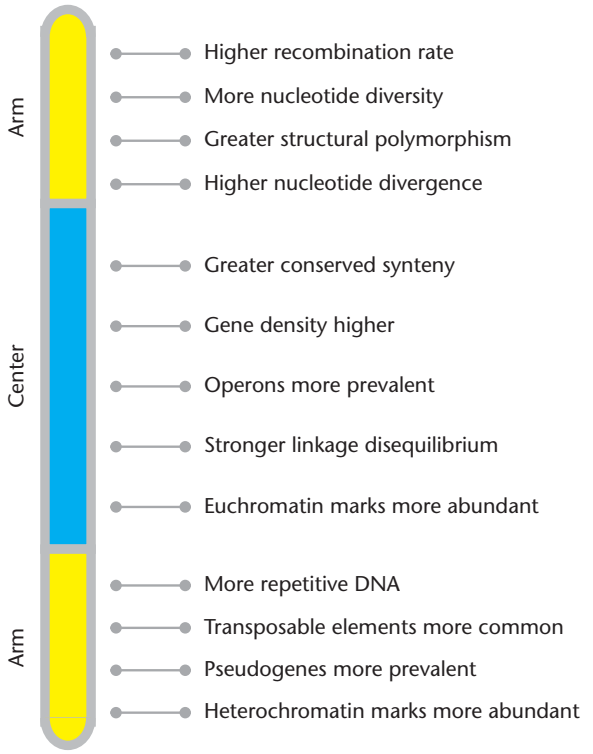

Figure 1 Nonrandom patterns in the organisation of the *C. elegans* genome. Chromosomes are typically partitioned into arms (yellow) and centres (blue) that differ by the relative distribution of several genomic features. Each domain makes ~30% of the chromosome length, although arms are asymmetric relative to centre and the position of centres varies among chromosomes. Here, features are mapped to a chromosome domain (grey line), arm or centre, when their abundance in this domain is greater than on the other domain. For instance, protein-coding genes are preferentially distributed in chromosome centres and recombination occurs more frequently in chromosome arms. Note that the distribution of genomic features in the arms does not differentiate between left and right ends (see Cutter *et al.*, 2009 and references therein; Gerstein *et al.*, 2010).

genes participating in basic cellular processes including transcription, translation, splicing, protein folding and RNA degradation as well as proteins for the mitochondrion. However, genes located in the same operon do not necessarily function in the same pathways (Blumenthal and Gleason, 2003). Nearly all operon genes are expressed in the germline, albeit nonexclusively, with the exception of spermatogenesis-related genes, which are almost entirely excluded from operons (Reinke and Cutter, 2009).

Operons are remarkably well conserved among the *Caenorhabditis* species, with 96% of *C. elegans* operons shared with *C. briggsae* (Stein *et al.*, 2003). However, operons are not evolutionarily frozen and operon losses via inversion, relocation and gene insertion have been reported in the genus (Qian and Zhang, 2008). Indeed, the present-day distribution of *C. elegans* operons fits well into a birth–death model describing how gene movement can account for operon formation (Cutter and Agrawal, 2010). Why is operon organisation so important in *C. elegans* and related nematodes but rare in other eukaryotes? One possibility is that the effective population size of most eukaryotes is too low for selection to effectively counter the disruption of operons and/or the loss of the operon splicing machinery by genetic drift (Lynch, 2007).

Alternatively, or in addition, operons may be restricted to those species for which they confer a strong selective benefit. The enrichment of growth-related and expression-related genes in operons suggests that signal-dependent regulation corresponding to specific circumstances could favour the persistence of operons. Zaslaver *et al.* (2011) hypothesise that operon organisation is advantageous in species that undergo growth arrest during their lifecycle because it allows fast recovery by decreasing competition between promoters for limited transcriptional resources. During harsh conditions, *C. elegans* can arrest growth at the first larval stage (L1) and can either resume normal development if conditions improve, or enter the dauer stage, an alternative nonfeeding and long-lived larval stage enabling the worm to disperse and seek better conditions. Transcriptional machinery is at low levels during arrest, presumably because the energetic resources are scarce (Zaslaver *et al.*, 2011). Importantly, on recovery from L1 and dauer arrest, the expression of operons increases while expression of non-operon genes remains mostly constant, with the exception of upregulated monocistronic genes involved in energy metabolism, transcription and translation. The hypothesis that operons may increase fitness in the face of low transcriptional resources is attractive because it is fully consistent with several features of operon organisation. The 5′ region of operons is shorter than that of monocistronic genes lacking germline expression (Reinke and Cutter, 2009), suggesting that a simplified transcriptional control of operons answers the demand for the expression of operon genes when transcriptional resources are limited, while the internal promoters (Huang *et al.*, 2007) may permit specific gene expression during the nonrecovery phase. Until the four-cell stage, embryonic expression is mostly maternal, and thus many transcripts need to be packaged into oocytes. That nearly all operons are expressed in the germline (Reinke and Cutter, 2009) may be a way to maximise expression while reducing energetic costs. Spermatogenesis-related genes are almost entirely absent from operons (Reinke and Cutter, 2009), perhaps because no similar selection pressure to optimise transcriptional resources exists during sperm formation. Similarly, a lack of selection to maintain operon organisation or even selection against operons because of expression silencing of the X chromosome in the germline (Kelly *et al.*, 2002) may explain why operons are absent from the X chromosome (Blumenthal *et al.*, 2002). **See also**: Gene Clustering in Eukaryotes

Gene Number and Functional Diversification

The *C. elegans* genome encodes 21 733 protein-coding genes (Gerstein *et al.*, 2010), only ~6% less than the

human genome despite the great disparity in the number of cell types and organismal complexity. For instance, the nervous system of *C. elegans* is relatively simple with only 302 neurons compared with the billions of neurons of the human brain. Yet, *C. elegans* can detect a similar variety of chemical compounds as mammals using only ~30 chemosensory neurons because, unlike in mammals, *C. elegans* neurons express multiple chemoreceptors (Troemel *et al.*, 1997). Chemosensation is the primary way through which the worm interacts with its environment, so it is perhaps not surprising that a large proportion of the gene complement is dedicated to this function and includes large gene families of G-proteins, guanylyl cyclases and seven-transmembrane receptors that are diversified by gene duplication (Bergamasco and Bazzicalupo, 2006).

Gene duplication is a major source of evolutionary innovation, increasing proteome diversity and gene network complexity. Individual members of multigene families that originate by duplication often evolve specialised physiological and developmental functions. Rapid expression and protein sequence divergence between gene duplicates can evolve via relaxed selection and/or positive selection facilitated by the initial functional redundancy following gene duplication (Li *et al.*, 2005). Gene duplication is particularly relevant to the evolution of the *C. elegans* genome because it is estimated that ~10.5% of the genes have paralogs (Lynch and Conery, 2000). However, gradual divergence of gene duplicates by subfunction partitioning and/or neofunctionalisation (Force *et al.*, 1999) is only one pattern of duplicate gene evolution. Half of the *C. elegans* young gene duplicates have a heterogeneous protein sequence because of partial and chimeric duplications, indicating that functional divergence may be frequent at birth. Moreover, older gene duplicates show similar proportions of structural heterogeneity suggesting that duplicates with structural differences are not disproportionally lost and may contribute to the long-term evolution of duplicate genes (Katju and Lynch, 2003, 2006). Nevertheless, whether transient or evolutionarily stable (Tischler *et al.*, 2006; Vavouri *et al.*, 2008; Li *et al.*, 2010), functional redundancy between gene duplicates contributes to genetic robustness, the ability of organisms to withstand deleterious mutations (Gu *et al.*, 2003; Hsiao and Vitkup, 2008) and provides protection against stochastic developmental errors (Burga *et al.*, 2011). For instance, *C. elegans* genes with paralogs are less likely to have a detrimental phenotype when knocked down by RNA interference (RNAi), presumably because of functional compensation (Kamath *et al.*, 2003). Similarly, the proportion of genes with fitness defects when knocked down by RNAi is ~100% for singleton genes but is only ~70% when genes with paralogs are included, highlighting the protective role of gene duplicates for an organism's fitness (Ramani *et al.*, 2012).

Another way for an organism to increase proteome diversity is through what may be called the 'more-for-less strategy', whereby multiple transcripts are produced from a single gene. This is particularly relevant in human because ~95% of human genes are alternatively spliced and often in a tissue-specific manner (Wang *et al.*, 2008). Alternative splicing is nevertheless efficient in promoting functional diversification in *C. elegans*, as ~25% of its genes produce several transcripts by differential regulation of splice sites. Moreover, 30% of the alternatively spliced genes are differentially regulated during development, further suggesting that distinct isoforms have unique functions (Ramani *et al.*, 2010). Interestingly, alternative splicing patterns and their regulation are well conserved between *C. elegans* and *C. briggsae*, suggesting that most isoforms were established in the most common ancestor of the two species and have remained stable since their divergence (Rukov *et al.*, 2007).

Regulation of Genome Function

Dynamic regulation of the genome is essential for organisms to respond to physiological and developmental demands. Transcription factors (TFs) are regulatory proteins that repress or activate gene expression in a time and tissue-specific manner through their interaction with *cis*-regulatory DNA motifs and protein cofactors. The *C. elegans* genome encodes 988 TFs, falling into more than 50 families based on the type of DNA-binding domain. Most of these families are conserved in humans, albeit with different relative abundances owing to differential gene duplication and retention. Moreover, a similar proportion of protein-coding genes encode TFs regulating the worm (5%) and the human genome (6–8%) (Reece-Hoyes *et al.*, 2005; Haerty *et al.*, 2008; Vaquerizas *et al.*, 2009). *Caenorhabditis* TF protein sequences diverge rapidly. TFs involved in the differentiation of chemosensory neurons evolve particularly fast and harbour substantial amounts of nucleotide variation in natural populations, highlighting the potential importance of TF sequence changes in developmental evolution (Haerty *et al.*, 2008; Jovelin, 2009; Jovelin *et al.*, 2009). **See also**: Evolution of Transcription Factors in *Caenorhabditis*

Small noncoding RNAs (ncRNAs) have emerged as an important component of gene regulatory networks. MicroRNA genes (miRNAs), the most studied and well understood class of ncRNAs, encode small 20–22 nucleotide-long RNAs processed from a hairpin precursor transcript. Mature miRNAs inhibit translation of mRNAs by binding to a 7 bp-motif called the seed in the 3′ untranslated region (UTR) of their targets. In contrast to TFs, miRNA content is markedly different in worms and humans (respectively, ~2% and 9% relative to the protein-coding gene complement, mirbase.org); it has been suggested that the origin of new miRNAs may facilitate the evolution of morphological complexity (Heimberg *et al.*, 2008). Sequencing of small RNAs expressed in *Caenorhabditis* species has revealed a similar number of miRNAs in *C. elegans*, *C. briggsae* and *Caenorhabditis remanei*. However, seed shifting, hairpin shifting (the formation of a new hairpin with sequence upstream or downstream of the

eLS © 2013, John Wiley & Sons, Ltd. www.els.net

miRNA), and arm-switching of the miRNA sequence within the precursor hairpin all contribute to the functional diversification of miRNAs among these species (de Wit *et al*., 2009). Moreover, the level of nucleotide diversity in nature for *C. briggsae* miRNAs is relatively high, and mutations that are predicted to alter function based on principles of miRNA target recognition and RNA folding suggest that miRNA sequence variation could play a role in the phenotypic differences among individuals (Jovelin and Cutter, 2011).

Epigenetic remodelling of the chromatin state through histone modifications and the incorporation of histone variants contributes to gene expression control by defining regions that are inaccessible to transcription (heterochromatin) and regions associated with gene activity (euchromatin). Chromatin organisation is an essential component of cell differentiation during development (Nakano *et al*., 2011), and can both regulate global expression profiles and fine-tune the expression of a small subset of developmental regulatory genes such as the Hox TFs (Studencka *et al*., 2011). Interestingly, transgenerational epigenetic inheritance has been documented in *C. elegans*, despite the absence of DNA methylation machinery (Bird, 2002). For example, longevity (Greer *et al*., 2011) and gene silencing by both small interfering RNAs (siRNAs) (Burton *et al*., 2011) and piwi-interacting RNAs (Ashe *et al*., 2012) can be transmitted nongenetically across several generations. As yet, we are only beginning to appreciate the role of epigenetic inheritance in evolution. A key point is that variation in epigenetic states among similar or identical genotypes increases phenotypic variance and can in turn affect the population pool of genotypes because of the action of genetic drift or selection (Richards, 2008; Feinberg and Irizarry, 2010). For instance, *C. elegans* individuals passing through the facultative dauer developmental stage have lower genome-wide profiles of the euchromatin histone modifications H4ac (pan-acetylation of histone H4) and H3K4me3 (trimethylation of histone H3 at lysine 4) than individuals with the same genotype that do not experience dauer growth arrest. Importantly, these differences in chromatin states are directly associated with variation in gene expression levels and also with increased longevity and larger brood size in postdauer individuals (Hall *et al*., 2010). What remains to be elucidated is how genome-wide patterns of chromatin marks are conserved among the *Caenorhabditis* species, how variable is epigenetic information within populations, and to what extent variation in chromatin organisation contributes to phenotypic variation among *Caenorhabditis* lineages.

A major challenge now is to build regulatory networks that integrate these different modes of regulation and to determine how these networks evolve between related species. Some efforts have been made in this direction with integrated TF and miRNA networks (Martinez *et al*., 2008; Gerstein *et al*., 2010), and one of the results from the modENCODE project is the prediction of protein-coding gene and miRNA expression levels from epigenetic chromatin states (Gerstein *et al*., 2010). An indirect way to examine the evolution of regulatory networks is through the comparison of transcriptomes, which has not only uncovered the variation in gene regulation during the development of different *C. elegans* strains (Capra *et al*., 2008) and *Caenorhabditis* species (Levin *et al*., 2012), but also has revealed that nematode development is characterised by a stage of low expression divergence analogous to the phylotypic stages of chordates and arthropods (Levin *et al*., 2012). Comparative transcriptomics between *C. elegans* and *C. briggsae* has also revealed that coexpression between gene neighbours in one species is generally not conserved in the other species, and that expression divergence is further increased when the genomic neighbourhoods are not conserved (Yanai and Hunter, 2009). Essential orthologous genes (i.e. with a lethal RNAi phenotype) have more similar expression profiles regardless of the conservation of their genomic neighbourhoods (Yanai and Hunter, 2009). High-throughput sequencing of the transcriptomes of several *Caenorhabditis* species has also shown that the evolution of hermaphroditism from an ancestral outcrossing mode of reproduction in *C. elegans* and *C. briggsae* was accompanied by a reduction in transcriptome size, due to the preferential loss of genes with female-biased expression (Thomas *et al*., 2012).

Mutation Rates

The rate at which new mutations get introduced into the genome is fundamental for understanding phenotypic and genomic evolution because mutations create all genetic variation. Biases in the origin of new variants also can strongly affect adaptive trajectories and drive populations to local fitness optima (Yampolsky and Stoltzfus, 2001; Stoltzfus and Yampolsky, 2009). Therefore, it is critical to obtain accurate estimates of mutation rates and the distribution of mutational effects. Rapid progress permitted by high-throughput sequencing of mutation accumulation (MA) lines has provided direct estimates of mutation rates. In MA experiments, a single *C. elegans* hermaphrodite individual (two individuals for gonochoristic species) is propagated at each new generation. This extreme form of bottleneck maximises the power of genetic drift over natural selection to permit the fixation of even severely deleterious but nonlethal mutations. Thus, a classic second and indirect way of estimating the mutation rate is to measure MA lines for fitness decline across generations. Although this approach underestimates the mutation rate because some mutations contribute little to fitness detectable in laboratory conditions, it is nevertheless essential for quantifying the distribution of mutational effects on fitness. These two complementary approaches have yielded some clear patterns, although causality has sometimes been difficult to disentangle from the various processes that could generate them.

Direct estimates of the gene duplication rate are high in *C. elegans* (1.2×10^{-7} and 3.4×10^{-7} /gene/generation,

respectively, for complete and partial duplications) (Lipinski *et al*., 2011) but are consistent with the abundance of duplicates in the genome (10.5%) (Lynch and Conery, 2000). Surprisingly, the spontaneous rate of gene duplication is two orders of magnitude greater than the direct estimates of the single base mutation rate in *C. elegans* MA lines (Denver *et al*., 2009, 2012), further highlighting the relevance of gene duplication in *C. elegans* genome evolution. However, the gene duplication rate inferred from MA lines is also two orders of magnitude higher than a previous estimate based on the distribution of young duplicates in the *C. elegans* genome (Lynch and Conery, 2000). In addition, the median duplication size in the MA lines (7 kb) is greater than both the median duplication size of young gene duplicates (1.4 kb) (Katju and Lynch, 2003) and of indels segregating in *C. elegans* natural isolates (2.7 kb) (Maydan *et al*., 2010), suggesting that the purifying selection in nature removes a substantial fraction of newly arising gene duplicates, and preferentially the larger ones.

The rate of base substitution also is much higher in the mitochondrial genome than in the nuclear genome of both *C. elegans* and *C. briggsae* (Denver *et al*., 2000, 2009, 2012; Howe *et al*., 2010; **Figure 2**), a pattern similar to *Drosophila* (Haag-Liautard *et al*., 2007, 2008; Keightley *et al*., 2009). The mitochondrial DNA (mtDNA) nucleotide mutation rate does not differ between *C. elegans* and *C. briggsae*, although the rate of deletions is 10-fold higher in the mitochondrial genome of *C. briggsae*. Two *C. briggsae* natural isolates do not differ from each other in base substitution and in deletion rates either (**Figure 2a**; Howe *et al*., 2010). However, the mutational spectra of the mitochondrial genome can vary significantly within and among lineages. In *Drosophila*, the mtDNA mutation rate is biased towards G/C→A/T transitions, consistent with a mitochondrial genome rich in A + T (82%) (Haag-Liautard *et al*., 2008). Despite a similar A + T richness (76%), the mitochondrial genome of *C. elegans* is biased towards A/T→G/C transitions (Denver *et al*., 2000); two different *C. briggsae* natural isolates also have distinct mutational biases (either towards G/C→T/A transversions or G/C→A/T transitions) (Howe *et al*., 2010).

As the mutations accumulate in a neutral fashion in MA experiments, comparison of patterns of mutation in MA lines and wild isolates can inform as to how natural selection shapes allelic variation in nature. The base mutation rate does not seem to differ among chromosomes or among chromosomal regions within *C. elegans* and *C. briggsae* (Denver *et al*., 2009, 2012). This homogeneity in mutational input along the chromosomes contrasts starkly with the biased chromosomal distribution of SNPs towards the arms that segregate in *C. elegans* natural isolates (**Figure 1**). This indicates that the mutational input is not the causal factor for the low abundance of allelic variation in the chromosome centres, and points to a role of selection at linked sites in these low-recombination chromosomal regions (Cutter and Payseur, 2003). Consistent with this hypothesis, selective sweeps coupled with low recombination rates due to the selfing mode of reproduction appear to wipe out nucleotide variation along many megabases of *C. elegans* chromosomes (Andersen *et al*., 2012). These effects of selective sweeps accrue on top of the persistent force of background selection that removes neutral polymorphisms linked to deleterious mutations, with genome-wide effects on nucleotide diversity demonstrated in *C. briggsae* and *C. elegans* (Rockman and Kruglyak, 2009; Cutter and Choi, 2010; Rockman *et al*., 2010). Comparison of the mutational spectra in MA lines and wild isolates also reveals the action of purifying selection in shaping nucleotide content in the genome. The frequencies of mutation type vary significantly in *C. elegans* and *C. briggsae* MA lines. Two types of mutations, G/C→A/T transitions and G/C→T/A transversions, are the most

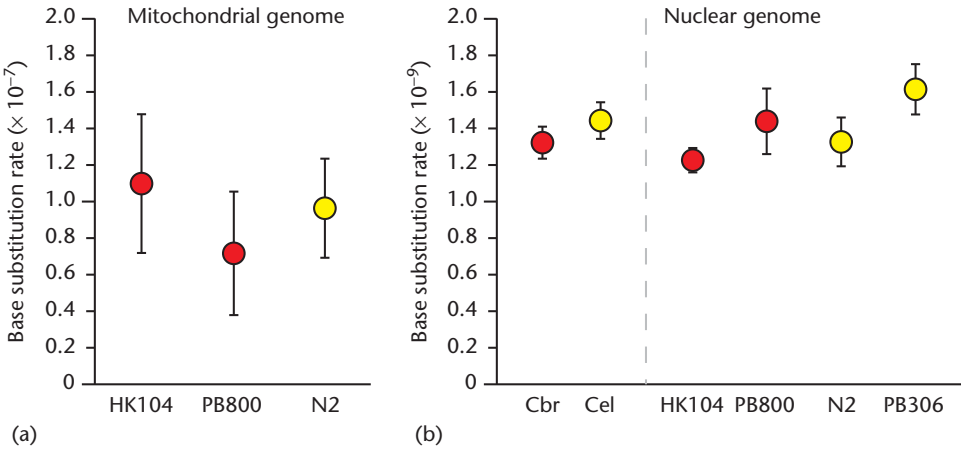

Figure 2 The base substitution rate (per nucleotide and per generation) is similar within and between *C. elegans* and *C. briggsae* for the mitochondrial genome (a) and the nuclear genome (b). In both species the mitochondrial mutation rate is 100-fold larger than the nuclear mutation rate (note the scale difference of the *y*-axis in (a) and (b)). Means are represented with ± 1 standard error of the mean, in red for *C. briggsae* (Cbr) and in yellow for *C. elegans* (Cel). Data for the mutation rate in the mitochondrial genome of *C. elegans* is from Denver *et al.* (2000) and from Howe *et al.* (2010) for *C. briggsae*. In (b), the mutation rate for each *C. elegans* and *C. briggsae* MA lines (Denver *et al*., 2012) is also pooled to obtain an average estimate for each species.

eLS © 2013, John Wiley & Sons, Ltd. www.els.net

common (Denver *et al*., 2009, 2012). Moreover, transversion mutations dominate in MA lines whereas transitions comprise most SNP mutations in wild isolates, suggesting that purifying selection eliminates deleterious transversions in natural populations. Interestingly, strain CB4856 from Hawaii does not exhibit any SNP transition bias (Solorzano *et al*., 2011), which could reflect either differences in the mutation process or different selective constraints in this genetically isolated strain (Rockman and Kruglyak, 2009; Andersen *et al*., 2012).

Applications of MA experiments presume that mutational processes are similar in laboratory and natural environments. Comparison of fitness decline in MA lines between the *Caenorhabditis* species with different mating types suggests that the genomic deleterious mutation rate could be higher in outcrossing species (*C. remanei*, *Caenorhabditis brenneri* and *Caenorhabditis* sp. 5) than in selfing species (*C. elegans*), and that the selection for a lower mutation rate could be more efficient in selfing species (Baer *et al*., 2010). Unfortunately, this pattern could also result as an experimental artefact of a stronger inbreeding depression in the gonochoristic species. This result is also at odds with the inverse relationship between the mutation rate and the effective population size (N_e) observed across diverse organisms (Lynch, 2010), which predicts that outcrossing species with their larger N_e would evolve lower mutation rates. Direct estimates of the mutation rate by sequencing MA lines from outcrossing and selfing progenitor species may help address this issue. For instance, analyses of fitness decline between *C. elegans* and *C. briggsae* MA lines suggest differences in the mutation rate between these two species (Baer *et al*., 2005), whereas no differences in the mutation rate could be detected when single nucleotide mutations were quantified in the MA line genomes (**Figure 2b**; Denver *et al*., 2012). One explanation is that species differ in the types of mutations, such as indel mutations (Phillips *et al*., 2009). A second possibility is that the deleterious effect of mutations on fitness and/or the mutation rate may vary among the species depending on the environmental conditions. Indeed, the mutation rate at microsatellite loci is increased and mutations have smaller deleterious effects on fitness (but are more numerous and contribute to a larger fitness decline) at a higher temperature in *C. elegans* and in *C. briggsae*. Moreover, these effects are stronger in *C. elegans* because of the low-fitness stressful conditions for *C. elegans* at one of the two temperatures assayed (Matsuba *et al*., 2012).

MA experiments are powerful because they provide a nearly unbiased view of mutational input, but they are nevertheless limited because of the few mutations they accumulate per line. Patterns of nucleotide variation in natural populations can help test for context-dependent mutational effects, such as in indel-associated mutation (Tian *et al*., 2008). Nucleotide polymorphisms within populations of *C. remanei* and *Caenorhabditis* sp. 23, and substitutions between these two-closely related species, cluster together with indels (Jovelin and Cutter, personal communication). Moreover, the single nucleotide mutation rate is elevated in the proximity of indels and is biased towards transversions, a signature of error-prone DNA polymerases. Together with additional analysis, this suggests a mechanism in which simple repeat sequences increase the incidence of indels, with subsequent recruitment of low-fidelity polymerases for the DNA repair, resulting in a locally elevated single-nucleotide mutation rate and a local bias towards transversions (Jovelin and Cutter, personal communication). **See also**: Mutation Rates: Evolution

C. elegans in Nature

Understanding the drivers of *Caenorhabditis* species abundance and distribution is essential for placing patterns of genome evolution into an ecological context. Unfortunately, little is known about the ecology of the worm, although sampling efforts in recent years have yielded new insights into the worm's natural habitat. Taxonomic representation in the genus also has improved markedly as a result of field collection, with more than half of 27 *Caenorhabditis* species currently maintained in culture having been isolated in the last decade (Kiontke *et al*., 2011).

Most species are confined either to temperate or tropical regions, with the exception of the cosmopolitan species *C. briggsae* and *C. elegans*. Even *C. elegans*, however, is found disproportionately in temperate regions (Kiontke *et al*., 2011). Climate is likely to be an important factor in the distribution of *Caenorhabditis*: species differ in thermotolerance in the laboratory in a manner that correlates with geographical origin (Kiontke *et al*., 2011), and tropical and temperate *C. briggsae* isolates differ in temperature-dependent fitness (Prasad *et al*., 2011). Species richness can be high even at small geographical scales, in particular in tropical areas unperturbed by human activity (Félix *et al*., 2013), and several species are sometimes found on the same substrate (Kiontke *et al*., 2011; Félix and Duveau, 2012; Félix *et al*., 2013). However, even when species ranges overlap, seasonal temperature variation may define ecological niches, relieving species from high levels of competition (Félix and Duveau, 2012).

What is clear from these collection efforts is that *Caenorhabditis* does not often live in the soil but instead feeds on bacteria within the rotting vegetal matter (Kiontke *et al*., 2011; Félix and Duveau, 2012; Félix *et al*., 2013). Populations may bloom ephemerally on a given resource and may be maintained in a local area in the long term through metapopulation dynamics (Barriere and Félix, 2007). Several *Caenorhabditis* species, including *C. elegans*, *C. briggsae* and *C. remanei* have generalist phoretic associations with various invertebrates, using them for dispersal. *Caenorhabditis japonica*, however, forms an intimate and specialised phoretic association with the burrower bug *Parastrachia japonensis* (Kiontke and Sudhaus, 2006; Okumura *et al*., 2012), and nematodes of the genus *Pristionchus* also specialise on distinct insect species

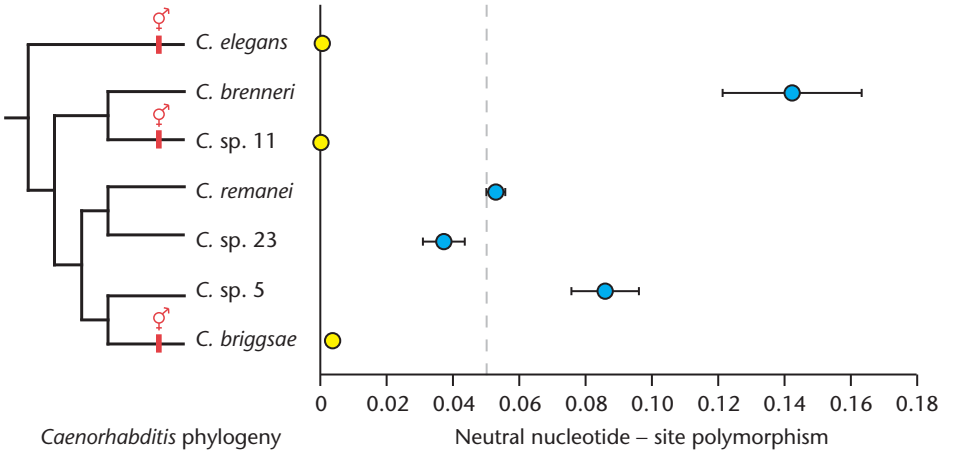

Figure 3 A selfing mode of reproduction has evolved independently three times (in red) from a gonochoristic ancestor in the genus *Caenorhabditis* (left). The level of nucleotide variation is strongly affected by the mating type and is greater in gonochoristic outcrossing *Caenorhabditis* species (blue) than in hermaphroditic selfing species (yellow) (right). The average level of neutral site diversity (at silent or synonymous sites) is plotted for each species. Error bars indicate the standard error of the mean with values smaller than the symbol size not shown. Updated from Cutter *et al.* (2013) with data for *C. briggsae* (Félix *et al.*, 2013), *Caenorhabditis* sp. 23 (Dey *et al.*, 2012), and *Caenorhabditis* sp. 11 (Gimond *et al.*, personal communication). The 5% threshold identifies hyperdiverse species (Cutter *et al.*, 2013).

(Herrmann *et al.*, 2007). The association of *Caenorhabditis* with invertebrate hosts may be complex in the wild, ranging from phoretic to opportunistic necromenic and even facultative entomopathogenic. Although, *C. briggsae* populations are commonly found on decaying vegetation, one entomopathogenic strain of *C. briggsae* has been isolated. This strain can enter and reproduce in the insect host, killing it using the bacteria *Serratia* sp. lining its gut (Abebe *et al.*, 2010, 2011). Interestingly, other *Caenorhabditis* species cultivated with *Serratia* sp., including *C. elegans*, were able to kill the insect host (Abebe *et al.*, 2010), and some types of *Serratia* are pathogenic to *C. elegans* (Morran *et al.*, 2009b), raising the possibility that an association with different bacteria may result in variation in life history strategies and facultative entomopathogenesis, as live bacteria are often found in the guts of wild strains of *C. elegans* and *C. briggsae* (Félix and Duveau, 2012).

Males are rare in both dauer and proliferating natural populations of *C. elegans* (Félix and Duveau, 2012), consistent with low outcrossing rates estimated from nucleotide polymorphism and linkage disequilibrium (reviewed in Anderson *et al.*, 2010). Nevertheless, *C. elegans* has heritable variation in the ability of isogenic populations to maintain males (Anderson *et al.*, 2010), suggesting that outcrossing rates can also vary depending on the genetic background, perhaps in response to selection in different environments. Dauer-inducing starvation increases the survival of males during dauer and increases outcrossing rates in populations recovering from dauer, indicating that environmental stress may favour biparental reproduction (Morran *et al.*, 2009a). Moreover, controlled experiments demonstrated that outcrossing in *C. elegans* populations slows the fixation of deleterious mutations compared with selfing populations, and accelerates adaptation (Morran *et al.*, 2009b). Several pathogens, including viruses, fungi, bacteria and microsporidia are commonly found in *C. elegans* and *C. briggsae* populations (Félix and Duveau, 2012), and the coevolution between *C. elegans* and pathogenic bacteria maintains higher rates of outcrossing in the long term (Morran *et al.*, 2011; Masri *et al.*, 2013). Thus, changing environmental conditions, in particular biotic pressures, may select for reproductive strategies favouring outcrossing, resulting in varying outcrossing rates among distinct populations.

The selfing-dominated reproduction of *C. elegans* nevertheless leaves a potent imprint in the genome sequence. The three hermaphroditic species, *C. elegans*, *C. briggsae* and *Caenorhabditis* sp. 11, all have orders of magnitude lower nucleotide polymorphism compared with their outcrossing relatives (**Figure 3**). As this difference is much greater than the twofold reduction expected due to selfing alone (Nordborg, 2000), factors other than the mating system *per se* must also influence nucleotide diversity in these species. Purifying selection (Rockman and Kruglyak, 2009), recurrent selective sweeps (Andersen *et al.*, 2012), and demographic fluctuation through metapopulation dynamics (Pannell and Charlesworth, 2000) are likely additional explanations for the low nucleotide diversity in *C. elegans*.

Conclusion

As a classical model organism having roots in the reductionist methodology of molecular biology, *C. elegans* has geared up for a systems biology level of investigation of gene function, the fitness effect of mutations and network analysis. With this wealth of genetic information and the rapid development of genomic resources in related species, the entire *Caenorhabditis* genus is emerging as a powerful

eLS © 2013, John Wiley & Sons, Ltd. www.els.net

model system for the application of the comparative method and for the investigation of broad questions related to ecological and evolutionary genetics.

References

Abebe E, Abebe-Akele F, Morrison J, Cooper V and Thomas WK (2011) An insect pathogenic symbiosis between a *Caenorhabditis* and *Serratia*. *Virulence* **2**: 158–161.

Abebe E, Jumba M, Bonner K *et al*. (2010) An entomopathogenic *Caenorhabditis briggsae*. *Journal of Experimental Biology* **213**: 3223–3229.

Andersen EC, Gerke JP, Shapiro JA *et al*. (2012) Chromosome-scale selective sweeps shape *Caenorhabditis elegans* genomic diversity. *Nature Genetics* **44**: 285–290.

Anderson JL, Morran LT and Phillips PC (2010) Outcrossing and the maintenance of males within *C. elegans* populations. *Journal of Heredity* **101**: S62–S74.

Ashe A, Sapetschnig A, Weick EM *et al*. (2012) piRNAs can trigger a multigenerational epigenetic memory in the germline of *C. elegans*. *Cell* **150**: 88–99.

Baer CF, Joyner-Matos J, Ostrow D *et al*. (2010) Rapid decline in fitness of mutation accumulation lines of gonochoristic (outcrossing) *Caenorhabditis* nematodes. *Evolution* **64**: 3242–3253.

Baer CF, Shaw F, Steding C *et al*. (2005) Comparative evolutionary genetics of spontaneous mutations affecting fitness in rhabditid nematodes. *Proceedings of the National Academy of Sciences of the USA* **102**: 5785–5790.

Bargmann CI (1998) Neurobiology of the *Caenorhabditis elegans* genome. *Science* **282**: 2028–2033.

Barriere A and Félix MA (2007) Temporal dynamics and linkage disequilibrium in natural *Caenorhabditis elegans* populations. *Genetics* **176**: 999–1011.

Bergamasco C and Bazzicalupo P (2006) Chemical sensitivity in *Caenorhabditis elegans*. *Cellular and Molecular Life Sciences* **63**: 1510–1522.

Bird A (2002) DNA methylation patterns and epigenetic memory. *Genes & Development* **16**: 6–21.

Blumenthal T, Evans D, Link CD *et al*. (2002) A global analysis of *Caenorhabditis elegans* operons. *Nature* **417**: 851–854.

Blumenthal T and Gleason KS (2003) *Caenorhabditis elegans* operons: form and function. *Nature Genetics Reviews* **4**: 112–120.

Burga A, Casanueva MO and Lehner B (2011) Predicting mutation outcome from early stochastic variation in genetic interaction partners. *Nature* **480**: 250–253.

Burton NO, Burkhart KB and Kennedy S (2011) Nuclear RNAi maintains heritable gene silencing in *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the USA* **108**: 19683–19688.

Capra EJ, Skrovanek SM and Kruglyak L (2008) Comparative developmental expression profiling of two *C. elegans* isolates. *PLoS One* **3**: e4055.

Cutter AD and Agrawal AF (2010) The evolutionary dynamics of operon distributions in eukaryote genomes. *Genetics* **185**: 685–693.

Cutter AD and Choi JY (2010) Natural selection shapes nucleotide polymorphism across the genome of the nematode *Caenorhabditis briggsae*. *Genome Research* **20**: 1103–1111.

Cutter AD, Dey A and Murray RL (2009) Evolution of the *Caenorhabditis elegans* genome. *Molecular Biology and Evolution* **26**: 1199–1234.

Cutter AD, Jovelin R and Dey A (2013) Molecular hyperdiversity and evolution in very large populations. doi:10.1111/mec.12281.

Cutter AD and Payseur BA (2003) Selection at linked sites in the partial selfer *Caenorhabditis elegans*. *Molecular Biology and Evolution* **20**: 665–673.

Deng X, Hiatt JB, Nguyen DK *et al*. (2011) Evidence for compensatory upregulation of expressed X-linked genes in mammals, *Caenorhabditis elegans* and *Drosophila melanogaster*. *Nature Genetics* **43**: 1179–1185.

Denver DR, Dolan PC, Wilhelm LJ *et al*. (2009) A genome-wide view of *Caenorhabditis elegans* base-substitution mutation processes. *Proceedings of the National Academy of Sciences of the USA* **106**: 16310–16314.

Denver DR, Morris K, Lynch M, Vassilieva LL and Thomas WK (2000) High direct estimate of the mutation rate in the mitochondrial genome of *Caenorhabditis elegans*. *Science* **289**: 2342–2344.

Denver DR, Wilhelm LJ, Howe DK *et al*. (2012) Variation in base-substitution mutation in experimental and natural lineages of *Caenorhabditis* nematodes. *Genome Biology and Evolution* **4**: 513–522.

Dey A, Jeon Y, Wang GX and Cutter AD (2012) Global population genetic structure of *Caenorhabditis remanei* reveals incipient speciation. *Genetics* **191**: 1257–1269.

Feinberg AP and Irizarry RA (2010) Evolution in health and medicine Sackler colloquium: stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proceedings of the National Academy of Sciences of the USA* **107**(suppl. 1): 1757–1764.

Félix MA and Duveau F (2012) Population dynamics and habitat sharing of natural populations of *Caenorhabditis elegans* and *C. briggsae*. *BMC Biology* **10**: 59.

Félix MA, Jovelin R, Ferrari C *et al*. (2013) Species richness, distribution and genetic diversity of *Caenorhabditis* nematodes in a remote tropical forest. *BMC Evolutionary Biology* **13**: 10.

Force A, Lynch M, Pickett FB *et al*. (1999) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**: 1531–1545.

Gassmann R, Rechtsteiner A, Yuen KW *et al*. (2012) An inverse relationship to germline transcription defines centromeric chromatin in *C. elegans*. *Nature* **484**: 534–537.

Gerstein MB, Lu ZJ, Van Nostrand EL *et al*. (2010) Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. *Science* **330**: 1775–1787.

Greer EL, Maures TJ, Ucar D *et al*. (2011) Transgenerational epigenetic inheritance of longevity in *Caenorhabditis elegans*. *Nature* **479**: 365–371.

Gu Z, Steinmetz LM, Gu X *et al*. (2003) Role of duplicate genes in genetic robustness against null mutations. *Nature* **421**: 63–66.

Guiliano DB and Blaxter ML (2006) Operon conservation and the evolution of trans-splicing in the phylum Nematoda. *PloS Genetics* **2**: e198.

Haag-Liautard C, Coffey N, Houle D *et al*. (2008) Direct estimation of the mitochondrial DNA mutation rate in *Drosophila melanogaster*. *PloS Biology* **6**: e204.

Haag-Liautard C, Dorris M, Maside X *et al*. (2007) Direct estimation of per nucleotide and genomic deleterious mutation rates in *Drosophila*. *Nature* **445**: 82–85.

Haerty W, Artieri C, Khezri N, Singh RS and Gupta BP (2008) Comparative analysis of function and interaction of transcription factors in nematodes: extensive conservation of orthology coupled to rapid sequence evolution. *BMC Genomics* **9**: 399.

Hall SE, Beverly M, Russ C, Nusbaum C and Sengupta P (2010) A cellular memory of developmental history generates phenotypic diversity in *C. elegans*. *Current Biology* **20**: 149–155.

Heimberg AM, Sempere LF, Moy VN, Donoghue PC and Peterson KJ (2008) MicroRNAs and the advent of vertebrate morphological complexity. *Proceedings of the National Academy of Sciences of the USA* **105**: 2946–2950.

Herrmann M, Mayer WE, Hong RL *et al*. (2007) The nematode *Pristionchus pacificus* (Nematoda: Diplogastridae) is associated with the oriental beetle *Exomala orientalis* (Coleoptera: Scarabaeidae) in Japan. *Zoological Science* **24**: 883–889.

Hillier LW, Miller RD, Baird SE *et al*. (2007) Comparison of *C. elegans* and *C. briggsae* genome sequences reveals extensive conservation of chromosome organization and synteny. *PloS Biology* **5**: e167.

Hobert O (2008) Regulatory logic of neuronal diversity: terminal selector genes and selector motifs. *Proceedings of the National Academy of Sciences of the USA* **105**: 20067–20071.

Howe DK, Baer CF and Denver DR (2010) High rate of large deletions in *Caenorhabditis briggsae* mitochondrial genome mutation processes. *Genome Biology and Evolution* **2**: 29–38.

Hsiao TL and Vitkup D (2008) Role of duplicate genes in robustness against deleterious human mutations. *PloS Genetics* **4**: e1000014.

Huang P, Pleasance ED, Maydan JS *et al*. (2007) Identification and analysis of internal promoters in *Caenorhabditis elegans* operons. *Genome Research* **17**: 1478–1485.

Jovelin R (2009) Rapid sequence evolution of transcription factors controlling neuron differentiation in *Caenorhabditis*. *Molecular Biology and Evolution* **26**: 2373–2386.

Jovelin R and Cutter AD (2011) MicroRNA sequence variation potentially contributes to within-species functional divergence in the nematode *Caenorhabditis briggsae*. *Genetics* **189**: 967–976.

Jovelin R, Dunham JP, Sung FS and Phillips PC (2009) High nucleotide divergence in developmental regulatory genes contrasts with the structural elements of olfactory pathways in *Caenorhabditis*. *Genetics* **181**: 1387–1397.

Kamath RS, Fraser AG, Dong Y *et al*. (2003) Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421**: 231–237.

Katju V and Lynch M (2003) The structure and early evolution of recently arisen gene duplicates in the *Caenorhabditis elegans* genome. *Genetics* **165**: 1793–1803.

Katju V and Lynch M (2006) On the formation of novel genes by duplication in the *Caenorhabditis elegans* genome. *Molecular Biology and Evolution* **23**: 1056–1067.

Keightley PD, Trivedi U, Thomson M *et al*. (2009) Analysis of the genome sequences of three *Drosophila melanogaster* spontaneous mutation accumulation lines. *Genome Research* **19**: 1195–1201.

Kelly WG, Schaner CE, Dernburg AF *et al*. (2002) X-chromosome silencing in the germline of *C. elegans*. *Development* **129**: 479–492.

Kiontke K and Sudhaus W (2006) Ecology of *Caenorhabditis* species. *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.37.1, http://www.wormbook.org.

Kiontke KC, Felix MA, Ailion M *et al*. (2011) A phylogeny and molecular barcodes for *Caenorhabditis*, with numerous new species from rotting fruits. *BMC Evolutionary Biology* **11**: 339.

Levin M, Hashimshony T, Wagner F and Yanai I (2012) Developmental milestones punctuate gene expression in the *Caenorhabditis* embryo. *Developmental Cell* **22**: 1101–1108.

Li J, Yuan Z and Zhang Z (2010) The cellular robustness by genetic redundancy in budding yeast. *PloS Genetics* **6**: e1001187.

Li WH, Yang J and Gu X (2005) Expression divergence between duplicate genes. *Trends in Genetics* **21**: 602–607.

Lipinski KJ, Farslow JC, Fitzpatrick KA *et al*. (2011) High spontaneous rate of gene duplication in *Caenorhabditis elegans*. *Current Biology* **21**: 306–310.

Lynch M (2007) *The Origins of Genome Architecture*. Sunderland, MA: Sinauer.

Lynch M (2010) Evolution of the mutation rate. *Trends in Genetics* **26**: 345–352.

Lynch M and Conery JS (2000) The evolutionary fate and consequences of duplicate genes. *Science* **290**: 1151–1155.

Martinez NJ, Ow MC, Barrasa MI *et al*. (2008) A *C. elegans* genome-scale microRNA network contains composite feedback motifs with high flux capacity. *Genes & Development* **22**: 2535–2549.

Masri L, Schulte RD, Timmermeyer N *et al*. (2013) Sex differences in host defence interfere with parasite-mediated selection for outcrossing during host-parasite coevolution. *Ecology Letters*. doi:10.1111/ele.12068.

Matsuba C, Ostrow DG, Salomon MP, Tolani A and Baer CF (2012) Temperature, stress and spontaneous mutation in *Caenorhabditis briggsae* and *Caenorhabditis elegans*. *Biology Letters* **9**: 20120334.

Maydan JS, Lorch A, Edgley ML, Flibotte S and Moerman DG (2010) Copy number variation in the genomes of twelve natural isolates of *Caenorhabditis elegans*. *BMC Genomics* **11**: 62.

Melters DP, Paliulis LV, Korf IF and Chan SW (2012) Holocentric chromosomes: convergent evolution, meiotic adaptations, and genomic analysis. *Chromosome Research* **20**: 579–593.

Meneely PM, McGovern OL, Heinis FI and Yanowitz JL (2012) Crossover distribution and frequency are regulated by *him-5* in *Caenorhabditis elegans*. *Genetics* **190**: 1251–1266.

Meyer BJ (2010) Targeting X chromosomes for repression. *Current Opinion in Genetics and Development* **20**: 179–189.

Morran LT, Cappy BJ, Anderson JL and Phillips PC (2009a) Sexual partners for the stressed: facultative outcrossing in the self-fertilizing nematode *Caenorhabditis elegans*. *Evolution* **63**: 1473–1482.

Morran LT, Parmenter MD and Phillips PC (2009b) Mutation load and rapid adaptation favour outcrossing over self-fertilization. *Nature* **462**: 350–352.

Morran LT, Schmidt OG, Gelarden IA, Parrish RC 2nd and Lively CM (2011) Running with the Red Queen: host–parasite coevolution selects for biparental sex. *Science* **333**: 216–218.

Nakano S, Stillman B and Horvitz HR (2011) Replication-coupled chromatin assembly generates a neuronal bilateral asymmetry in *C. elegans*. *Cell* **147**: 1525–1536.

Nordborg M (2000) Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. *Genetics* **154**: 923–929.

Okumura E, Tanaka R and Yoshiga T (2012) Species-specific recognition of the carrier insect by dauer larvae of the nematode *Caenorhabditis japonica*. *Journal of Experimental Biology* **216**: 568–572.

Pannell JR and Charlesworth B (2000) Effects of metapopulation processes on measures of genetic diversity. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. **355**: 1851–1864.

Phillips N, Salomon M, Custer A, Ostrow D and Baer CF (2009) Spontaneous mutational and standing genetic (co)variation at dinucleotide microsatellites in *Caenorhabditis briggsae* and *Caenorhabditis elegans*. *Molecular Biology and Evolution* **26**: 659–669.

Prasad A, Croydon-Sugarman MJ, Murray RL and Cutter AD (2011) Temperature-dependent fecundity associates with latitude in *Caenorhabditis briggsae*. *Evolution* **65**: 52–63.

Qian W and Zhang J (2008) Evolutionary dynamics of nematode operons: easy come, slow go. *Genome Research* **18**: 412–421.

Ramani AK, Calarco JA, Pan Q *et al*. (2010) Genome-wide analysis of alternative splicing in *Caenorhabditis elegans*. *Genome Research* **21**: 342–348.

Ramani AK, Chuluunbaatar T, Verster AJ *et al*. (2012) The majority of animal genes are required for wild-type fitness. *Cell* **148**: 792–802.

Reece-Hoyes JS, Deplancke B, Shingles J *et al*. (2005) A compendium of *Caenorhabditis elegans* regulatory transcription factors: a resource for mapping transcription regulatory networks. *Genome Biology* **6**: R110.

Reinke V and Cutter AD (2009) Germline expression influences operon organization in the *Caenorhabditis elegans* genome. *Genetics* **181**: 1219–1228.

Richards EJ (2008) Population epigenetics. *Current Opinion in Genetics and Development* **18**: 221–226.

Rockman MV and Kruglyak L (2009) Recombinational landscape and population genomics of *Caenorhabditis elegans*. *PloS Genetics* **5**: e1000419.

Rockman MV, Skrovanek SS and Kruglyak L (2010) Selection at linked sites shapes heritable phenotypic variation in *C. elegans*. *Science* **330**: 372–376.

Ross JA, Koboldt DC, Staisch JE *et al*. (2011) *Caenorhabditis briggsae* recombinant inbred line genotypes reveal inter-strain incompatibility and the evolution of recombination. *PloS Genetics* **7**: e1002174.

Rukov JL, Irimia M, Mork S *et al*. (2007) High qualitative and quantitative conservation of alternative splicing in *Caenorhabditis elegans* and *Caenorhabditis briggsae*. *Molecular Biology and Evolution* **24**: 909–917.

Schvarzstein M, Wignall SM and Villeneuve AM (2010) Coordinating cohesion, co-orientation, and congression during meiosis: lessons from holocentric chromosomes. *Genes & Development* **24**: 219–228.

Solorzano E, Okamoto K, Datla P *et al*. (2011) Shifting patterns of natural variation in the nuclear genome of *Caenorhabditis elegans*. *BMC Evolutionary Biology* **11**: 168.

Stein LD, Bao Z, Blasiar D *et al*. (2003) The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PloS Biology* **1**: 166–192.

Stoltzfus A and Yampolsky LY (2009) Climbing mount probable: mutation as a cause of nonrandomness in evolution. *Journal of Heredity* **100**: 637–647.

Studencka M, Wesolowski R, Opitz L *et al*. (2011) Transcriptional repression of Hox genes by *C. elegans* HP1/HPL and H1/HIS-24. *PloS Genetics* **8**: e1002940.

Thomas CG, Li R, Smith HE *et al*. (2012) Simplification and desexualization of gene expression in self-fertile nematodes. *Current Biology* **22**: 2167–2172.

Tian D, Wang Q, Zhang P *et al*. (2008) Single-nucleotide mutation rate increases close to insertions/deletions in eukaryotes. *Nature* **455**: 105–108.

Tischler J, Lehner B, Chen N and Fraser AG (2006) Combinatorial RNA interference in *Caenorhabditis elegans* reveals that redundancy between gene duplicates can be maintained for more than 80 million years of evolution. *Genome Biology* **7**: R69.

Troemel ER, Kimmel BE and Bargmann CI (1997) Reprogramming chemotaxis responses: sensory neurons define olfactory preferences in *C. elegans*. *Cell* **91**: 161–169.

Tsai CJ, Mets DG, Albrecht MR *et al*. (2008) Meiotic crossover number and distribution are regulated by a dosage compensation protein that resembles a condensin subunit. *Genes & Development* **22**: 194–211.

Vaquerizas JM, Kummerfeld SK, Teichmann SA and Luscombe NM (2009) A census of human transcription factors: function, expression and evolution. *Nature Genetics Reviews* **10**: 252–263.

Vavouri T, Semple JI and Lehner B (2008) Widespread conservation of genetic redundancy during a billion years of eukaryotic evolution. *Trends in Genetics* **24**: 485–488.

Wagner CR, Kuervers L, Baillie DL and Yanowitz JL (2010) *xnd-1* regulates the global recombination landscape in *Caenorhabditis elegans*. *Nature* **467**: 839–843.

Wang ET, Sandberg R, Luo S *et al*. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**: 470–476.

de Wit E, Linsen SE, Cuppen E and Berezikov E (2009) Repertoire and evolution of miRNA genes in four divergent nematode species. *Genome Research* **19**: 2064–2074.

Yampolsky LY and Stoltzfus A (2001) Bias in the introduction of variation as an orienting factor in evolution. *Evolution & Development* **3**: 73–83.

Yanai I and Hunter CP (2009) Comparison of diverse developmental transcriptomes reveals that coexpression of gene neighbors is not evolutionarily conserved. *Genome Research* **19**: 2214–2220.

Zaslaver A, Baugh LR and Sternberg PW (2011) Metazoan operons accelerate recovery from growth-arrested states. *Cell* **145**: 981–992.

Further Reading

Cutter AD, Jovelin R and Dey A (2013) Molecular hyperdiversity and evolution in very large populations. *Molecular Ecology*. doi:10.1111/mec.12281.

Felix MA and Braendle C (2010) The natural history of *Caenorhabditis elegans*. *Current Biology* **20**: R965–R969.

Gaertner BE and Phillips PC (2010) *Caenorhabditis elegans* as a platform for molecular quantitative genetics and systems biology of natural variation. *Genetical Research* **92**: 331–348.

Wenzel D, Palladino F and Jedrusik-Bode M (2011) Epigenetics in *C. elegans*: facts and challenges. *Genesis* **49**: 647–661.