

# Rates of deleterious mutation and the evolution of sex in *Caenorhabditis*

A. D. CUTTER & B. A. PAYSEUR

Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA

## Keywords:

breeding system;  
*Caenorhabditis briggsae*;  
*Caenorhabditis elegans*;  
*Caenorhabditis remanei*;  
deleterious mutation;  
evolution of sex.

## Abstract

A variety of models propose that the accumulation of deleterious mutations plays an important role in the evolution of breeding systems. These models make predictions regarding the relative rates of protein evolution and deleterious mutation in taxa with contrasting modes of reproduction. Here we compare available coding sequences from one obligately outcrossing and two primarily selfing species of *Caenorhabditis* to explore the potential for mutational models to explain the evolution of breeding system in this clade. If deleterious mutations interact synergistically, the mutational deterministic hypothesis predicts that a high genomic deleterious mutation rate ( $U$ ) will offset the reproductive disadvantage of outcrossing relative to asexual or selfing reproduction. Therefore, *C. elegans* and *C. briggsae* (both largely selfing) should both exhibit lower rates of deleterious mutation than the obligately outcrossing relative *C. remanei*. Using a comparative approach, we estimate  $U$  to be equivalent (and  $<1$ ) among all three related species. Stochastic mutational models, Muller's ratchet and Hill–Robertson interference, are expected to cause reductions in the effective population size in species that rarely outcross, thereby allowing deleterious mutations to accumulate at an elevated rate. We find only limited support for more rapid molecular evolution in selfing lineages. Overall, our analyses indicate that the evolution of breeding system in this group is unlikely to be explained solely by available mutational models.

## Introduction

Theory indicates that the accumulation of deleterious mutations in individuals and in populations may play an important role in the evolution of sex. Different mutational models focus on the influence of genetic linkage (Hill & Robertson, 1966), on stochastic population genetic factors (Muller, 1964; Gabriel *et al.*, 1993; Lynch *et al.*, 1993), on the role of selection or competition among lineages (Birky, 1999), and on deterministic effects of deleterious mutation in relation to the two-fold cost of sex (Kondrashov, 1988; Charlesworth, 1990). Most of the work on this topic has focused on the

likelihood of invasion of asexually reproducing individuals in an otherwise outcrossing population. However, these models also apply to comparisons of different types of breeding system. Many features of the theoretical predictions hold for the problem of self-fertilizing individuals invading an outcrossing population, such as expected differences in rates of nucleotide substitution. We use DNA sequence data in a comparative context to evaluate the degree to which these mutational models may contribute to the maintenance or loss of outcrossing in a clade of rhabditid nematodes.

A mutational model that has received much attention in recent years is the mutational deterministic (MD) hypothesis, which posits that selection against deleterious mutations is responsible for the origin and/or maintenance of sex (Kondrashov, 1988). Under synergistic epistasis, where each additional deleterious mutation causes a greater-than-additive reduction in fitness,

Correspondence: Asher D. Cutter, Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA.  
Tel.: +1-520-621-1494; fax: +1-520-621-9190;  
e-mail: acutter@email.arizona.edu

sexual populations can purge combinations of deleterious mutations more efficiently than asexual populations. When the genomic deleterious mutation rate per generation ( $U$ ) is at least 1, the resulting decreased mutational load of outcrossing sexual populations can offset the two-fold disadvantage of meiosis (Charlesworth, 1990). Therefore, the MD hypothesis predicts that  $U$  will be reasonably high (of order 1) in outcrossing species (Kondrashov, 1988; Charlesworth, 1990). Like asexually reproducing individuals, self-fertilizing individuals experience a two-fold reproductive advantage over outcrossers (Lloyd, 1980). However, the fitness of a purely selfing individual at mutation–selection balance ( $\sim e^{-\frac{3}{4}U}$ ) is less than that of an average individual in an obligately outcrossing population under the assumption of synergistic epistasis ( $\sim e^{-\frac{1}{2}U}$ ) (Kimura & Maruyama, 1966; Kondrashov, 1985; Charlesworth, 1990). Consistent with the notion that highly selfing populations will be resistant to invasion by an asexual genotype (Charlesworth, 1980), selfers also experience greater fitness than asexuals ( $\sim e^{-U}$ ) at mutation–selection balance (Kimura & Maruyama, 1966). Note that the well-known  $1\frac{1}{2}$ -fold reproductive advantage of asexuals over random-mating hermaphrodites does not hold for obligate selfers, which completely recover the cost of meiosis (Lloyd, 1980). Thus, akin to the prediction for the typical comparison between outcrossing and asexual reproduction, the MD model predicts that outcrossing will offer a net advantage over selfing only when  $U$  is sufficiently large (specifically, when  $U > \sim 1.4$ ). Consequently, we expect to find higher values of  $U$  in obligately outcrossing members of a clade relative to members that rarely outcross if the MD hypothesis is a general explanation for the maintenance of sex (i.e. obligate outcrossing relative to other breeding systems) and enough time has elapsed for  $U$  to evolve (Kondrashov, 1985; Charlesworth, 1990; Wright *et al.*, 2002).

Kondrashov & Crow (1993) proposed an approach to estimate  $U$  using interspecific DNA sequence comparisons by contrasting the total numbers of constrained and unconstrained sites in a genome, conditioned on divergence time in generations. Keightley & Eyre-Walker (2000) used this approach to estimate  $U$  for a variety of animal species and suggested that  $U$  is often  $< 1$ , challenging the generality of the MD hypothesis. This comparative method of estimating  $U$  based on a nontrivial sample of the genome is becoming more feasible as genomic data accumulate in more species and it allows the derivation of genomic deleterious mutation rate estimates independent from experimental estimates (e.g. Wright *et al.*, 2002) – we use this approach among *Caenorhabditis* species.

The interaction between deleterious mutation and genetic linkage in finite populations also can influence the evolution of breeding system. Interference among selected, linked loci can lead to a reduction in effective population size (Hill & Robertson, 1966) and the

subsequent accumulation of deleterious mutations. Therefore, the rate of amino acid substitution (corrected for mutation rate heterogeneity) should be higher in nonoutcrossing lineages relative to outcrossing lineages when a large fraction of amino acid mutations are slightly deleterious (i.e. selection coefficients on the order of the reciprocal of the population size; Ohta, 1973) and Hill–Robertson interference is a potent force. Muller's ratchet (Muller, 1964), which is stronger in small populations, provides another stochastic mutational explanation for the evolutionary advantage to outcrossing. As the genomes of progeny in outcrossing populations can be reconstituted to contain fewer deleterious mutations than their parents (by virtue of recombinant gametes), nonoutcrossing populations should experience greater fitness reductions and increased probabilities of extinction due to a higher load of deleterious mutations (Muller, 1964; Gabriel *et al.*, 1993; Lynch *et al.*, 1993). Like the predictions that follow from Hill–Robertson interference, nonoutcrossing lineages that persist should exhibit higher amino acid substitution rates than closely related outcrossing species. Additionally, we might expect such lineages to evolve reduced mutation rates to slow the accumulation of deleterious mutations, which will be promoted by selection or competition among nonoutcrossing lineages (Birky, 1999).

Members of the bacteriophagous, soil-dwelling nematode genus *Caenorhabditis* provide an excellent system in which to evaluate explanations for the evolution of outcrossing, and the MD hypothesis in particular. First, the complete genomic sequence of *C. elegans* (The *C. elegans* Sequencing Consortium, 1998) and of large stretches of contiguous sequence in the closely related species *C. briggsae* (Kent & Zahler, 2000) make it possible to compare sequences at many loci, as has been done to estimate levels of selective constraint on intergenic sequences (Shabalina & Kondrashov, 1999). Furthermore, the complete sequence of *C. elegans* allows fairly accurate estimation of parameters relevant to  $U$ , including the total number of coding sites in the genome. Secondly, there is variation in the mode of reproduction among *Caenorhabditis* species. The androdioecious species, *C. elegans* and *C. briggsae*, reproduce largely via self-fertilization of hermaphrodites and outcross with males only rarely (Fitch & Thomas, 1997; Riddle *et al.*, 1997) – *C. elegans* males are produced spontaneously at a frequency of  $\sim 0.2\%$  under laboratory conditions (Hodgkin *et al.*, 1979; Hodgkin & Doniach, 1997) and indirect evidence suggests that outcrossing occurs at a frequency  $< 2\%$  in nature (Cutter & Payseur, 2003). In contrast, the closely related *C. remanei* exhibits an obligately outcrossing breeding system with an even sex ratio (Baird *et al.*, 1994) and much higher levels of genetic variation among isolates (Graustein *et al.*, 2002; Jovelin *et al.*, 2003). The variation in breeding system among *Caenorhabditis* species contrasts with the conservation in morphology (Fitch & Thomas, 1997; Sudhaus & Fitch, 2001).

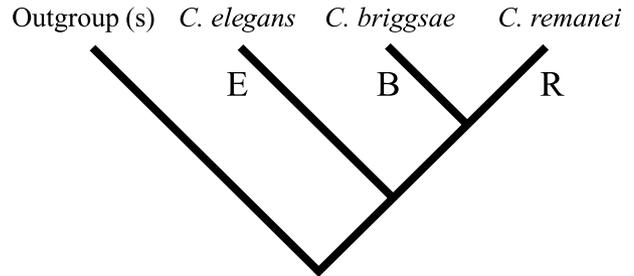
As phylogenetic evidence indicates that obligate outcrossing represents the ancestral state for this group (Fitch & Thomas, 1997) and self-fertilizing hermaphrodites may have evolved independently in *C. elegans* and *C. briggsae* (E. Haag, personal communication), the issue is simplified to the maintenance of obligate outcrossing, to the exclusion of its origin (Lenski, 1999). Finally, selfing has evolved independently several times in the Rhabditidae (Fitch & Thomas, 1997), indicating that the topic of breeding system evolution is a general problem in nematodes and not a peculiarity of the *Caenorhabditis* genus.

Here, we use coding sequences available for caenorhabditids to compare rates of deleterious mutation ( $U$ ) and site substitution ( $K_A$ ) for obligate outcrossing and highly selfing species. We concentrate on the MD hypothesis, which has received the most attention in recent literature. If the mutational models hold generally, then we expect that these species will vary in rates of deleterious mutation and protein evolution with respect to their breeding system – *C. remanei* should exhibit disproportionately high values of  $U$  and low rates of protein evolution. Selfing and outcrossing species of *Arabidopsis* have been shown to exhibit no significant difference in  $U$  (Wright *et al.*, 2002), but such a comparison between related animals that vary in breeding system has not been made previously. We examine the effect of variation in the parameters used to estimate  $U$  and argue that  $U$  is probably low and quite similar for *C. elegans*, *C. briggsae*, and *C. remanei*. We also find little support for the predictions of Hill–Robertson and Muller’s ratchet models in that the rate of molecular evolution is rarely higher in primarily selfing species. These results suggest that the evolution of breeding system in this group is unlikely to be explained by available mutational models alone.

## Methods

### Confirmation of phylogeny

We confirmed the phylogenetic topology for *C. elegans*, *C. briggsae*, and *C. remanei* based on published nuclear DNA sequences for three genes. This analysis served two purposes: it provided us with a framework in which to perform relative-rates tests (Sarich & Wilson, 1973) and sequence-based estimation of  $U$  for each *Caenorhabditis* species-pair. We constructed molecular phylogenies of nuclear genes with the PAUPsearch Wisconsin Package Version 10.2 [Genetics Computer Group (GCG), Madison, WI, USA] software based on coding nucleotide sequences for calmodulin (*cal-1*; outgroup *Strongyloides stercoralis*), globin (outgroup *Nippostrongylus brasiliensis*), and a homeobox gene (*ceh-13*; outgroups *S. ratti* and *S. stercoralis*). Sequences for *C. elegans* were obtained from <http://www.wormbase.org>. Other sequences were extracted from GenBank (<http://www.ncbi.nlm.nih.gov>), from



**Fig. 1** Cladogram for *Caenorhabditis elegans*, *C. briggsae*, and *C. remanei* based on three nuclear loci (globin, *cal-1* and *ceh-13*). Branches B and R are ~78% the length of branch E, as inferred from minimum evolution distance statistics.

the paper by Thomas & Wilson (1991) (*cal-1*), or from A. Streit (personal communication) (*ceh-13*). We used this set of loci for phylogenetic reconstruction because outgroup sequences for additional loci were unavailable. Analyses using minimum evolution distance, maximum parsimony, and maximum likelihood optimality criteria all yielded a single tree with a topology that hypothesizes a sister relationship between *C. briggsae* and *C. remanei* with 62–100% bootstrap support (Fig. 1). This topology is consistent with distance measures for several other loci (Haag & Kimble, 2000; Chen *et al.*, 2001; Rudel & Kimble, 2001; Jovelin *et al.*, 2003) and the pattern of interspecific hybrid formation among these species (Baird *et al.*, 1992). In contrast, slowly evolving ribosomal genes (Fitch *et al.*, 1995; Baldwin *et al.*, 1997) and high similarity among morphological characters have contributed to conflicting and unresolved species relationships in previous phylogenetic analyses (Sudhaus & Kiontke, 1996; Baldwin *et al.*, 1997; Sudhaus & Fitch, 2001); 18S rDNA and male tail characters have been described as insufficient to differentiate species relationships among these three taxa (Fitch *et al.*, 1995; Sudhaus & Kiontke, 1996). Consequently, we rely on the growing consensus hypothesis of species relationships supported by our molecular data (Fig. 1). We inferred that the relative length of branches B and R is 78% that of branch E based on the average minimum evolution branch length distances of these three loci (Fig. 1). Accordingly, we reduced the divergence time between *C. briggsae* and *C. remanei* by 22% relative to *C. elegans* in calculations of  $U$  (Fig. 1).

### Three-species comparisons

We analysed substitution rates in coding sequences for 10 loci (~15 kb) with orthologs in *C. elegans*, *C. briggsae* and *C. remanei* (syn. *vulgaris*) (Table 1). We obtained *C. elegans* sequences from <http://www.wormbase.org> and other sequences from GenBank (<http://www.ncbi.nlm.nih.gov>). *Caenorhabditis remanei* sequence from two loci that were not available in Genbank were provided by

**Table 1** Unadjusted substitution rates, codon usage bias and relative-rates test statistics.

<i>C. elegans</i> locus	<i>C. briggsae</i> , <i>C. remanei</i> accessions	<i>C. elegans</i> $F_{op}$	Wu & Li's relative-rates test		<i>briggsae-remanei</i>		<i>elegans-briggsae</i>		<i>elegans-remanei</i>	
			$z$	$P^*$	$K_A$	$K_S$	$K_A$	$K_S$	$K_A$	$K_S$
<i>cal-1</i>	(see Methods)	0.31	0.27	n.s.	0.00	0.49	0.00	0.95	0.00	0.84
<i>ceh-13</i>	(see Methods)	0.36	0.45	n.s.	0.10	1.48	0.13	1.81	0.13	2.04
<i>fem-2</i>	AF054982, AF507019	0.39	0.69	n.s.	0.30	1.65	0.27	1.89	0.27	1.73
<i>fog-3</i>	AF354169, AF354170	0.45	0.59	n.s.	0.27	1.31	0.34	1.41	0.30	1.75
<i>globin</i>	U48289, U48294	0.58	-1.06	n.s.	0.05	0.84	0.09	0.70	0.08	1.01
<i>glp-1</i>	AF315554, AF315556	0.37	3.39	0.0004	0.26	1.43	0.36	2.71	0.32	1.72
<i>lin-14</i>	AH010072, AF231036	0.35	1.26	n.s.	0.11	0.83	0.17	1.01	0.16	0.91
<i>mec-3</i>	L02878, X63956	0.27	1.78	0.037	0.08	0.94	0.10	1.68	0.06	1.97
<i>rpl-3</i>	AF247847, AF247852	0.80	1.74	0.041	0.00	0.05	0.02	0.21	0.01	0.11
<i>tra-2</i>	U59879, AF187965	0.33	0.46	n.s.	0.42	1.48	0.50	1.72	0.49	1.66

\*Multiple-test correction  $\alpha = 0.005$ .

A. Streit (the homeobox gene *ceh-13*) and the paper by Thomas & Wilson (1991) for the calmodulin gene *cal-1*. We aligned each locus using GCG Pileup on the corresponding amino acid sequences. We then back-translated into nucleotides these aligned amino acid sequences and ran GCG Diverge on each locus pair to calculate rates of synonymous ( $K_S$ ) and nonsynonymous ( $K_A$ ) site substitution. Diverge uses the improved method of Li (Li *et al.*, 1985; Li, 1993; Pamilo & Bianchi, 1993) that utilizes Kimura's two-parameter method (Kimura, 1980) to correct for multiple hits and to take transition and transversion rate differences into account. Given *C. elegans* as outgroup, we tested for differences in rates of molecular evolution between *C. briggsae* and *C. remanei* for each locus and for concatenated sequences with relative-rates tests using the K2WuLi program (Wu & Li, 1985; Muse & Weir, 1992; <http://jcsmr.anu.edu.au/dmm/humgen/lars/k2wulisub.htm>), which also employs Kimura's two-parameter model (Kimura, 1980).

Our analyses assume that  $K_S$  is a robust indicator of the neutral mutation rate. In an effort to account for possible selection on synonymous sites (Akashi, 1995) we adjusted our synonymous substitution rate estimates for the effects of codon bias (see below). We refer to these adjusted estimates as  $\delta_s$ .

We estimated rates of total ( $M$ ) and deleterious ( $U$ ) mutation per generation with  $K_A$  and  $\delta_s$ , following Kondrashov & Crow (1993) as modified by Eyre-Walker & Keightley (1999). Using a method weighted by gene length ( $L$ ),

$$M = Z \frac{\sum L \times \delta_s}{\sum L} \quad \text{and} \quad U = M - Z \frac{\sum L \times K_A}{\sum L},$$

where all summations are across loci and  $Z = 2$  (genomes)  $\times 19\,099$  (genes)  $\times 1644$  (nucleotides/gene)  $\times (1/G$  (generations/year))  $\times (1/Y$  (divergence time)) is a constant that includes the number of genes (The *C. elegans* Sequencing Consortium, 1998), the length of genes

(mean gene length of 1406 *C. elegans* loci, see below), generation time ( $G$ ), and divergence time ( $Y$ ). The gene length ( $L$ ) used in our summations is the number of nucleotides processed by Diverge. This method measures the average deleterious mutation rate of the two species used to generate the  $K_A$  and  $\delta_s$  values. Consequently, the prediction of the MD model is that  $U_{\text{elegans-briggsae}} < U_{\text{briggsae-remanei}} \equiv U_{\text{elegans-remanei}}$ . As synonymous sites are saturated ( $\delta_s \geq 1$ ), on average, we also made estimates of  $U$  by using a constant value of the neutral mutation rate, where  $M = 2 \times 19\,099 \times 1644 \times 9.1 \times 2.25 \times 10^{-10}$  (Drake *et al.*, 1998). We calculated 95% confidence limits of  $M$  and  $U$  via bootstrap analysis with 10 000 replicates (Manly, 1997).

As the average generation time of *Caenorhabditis* species in nature is unknown, and no fossil records provide definitive information regarding the timing of the *C. elegans-briggsae* divergence, we calculated  $M$  and  $U$  for a broad range of these two parameters. Other studies (Denver *et al.*, 2000) assumed 4-day egg-to-egg generation duration ( $\sim 90$  generations/year), which is close to the growth rate of *C. elegans* in the laboratory at 20 °C (Byerly *et al.*, 1976). Under laboratory conditions, *C. elegans* worms develop in  $\sim 2.5$  days ( $\sim 150$  generations/year) at 25 °C and in  $\sim 6$  days ( $\sim 60$  generations/year) at 15 °C (Byerly *et al.*, 1976), although the facultative and developmentally quiescent dauer larval pathway can persist for weeks to months (Riddle, 1988). Fecundity (Byerly *et al.*, 1976) and population growth rate (Venette & Ferris, 1997) are both maximal at  $\sim 20$  °C. Typical estimates (based on one or a few genes) for the time to most recent common ancestor of *C. elegans* and *C. briggsae* vary between 23 and 60 million years (My) (Prasad & Baillie, 1989; Heschl & Baillie, 1990; Lee *et al.*, 1992; Kennedy *et al.*, 1993; Thacker *et al.*, 1999). A more recent analysis based on chromosomal rearrangements estimates the date to be 50–120 My ago (Coghlan & Wolfe, 2002).

### C. *elegans*–*C. briggsae* comparisons

We also calculated rates of substitution for a larger sample of 1406 putative orthologs of *C. elegans* and *C. briggsae*. We extracted the coding regions for gene predictions based on the full *C. elegans* genomic sequence and ~13 Mb of *C. briggsae* genomic sequence from ACEDB (A *C. elegans* Database, <http://www.wormbase.org>). We used GCG BLASTN to identify locus pairs between the two species that provided the best match as potential orthologs. We translated into amino acids the nucleotide sequence for the coding regions of each locus for alignment of best-match pairs using GCG Gap with default parameters. We then back-translated into nucleotides these aligned amino acid sequences and ran GCG Diverge on each locus pair to calculate rates of synonymous ( $K_S$ ) and nonsynonymous ( $K_A$ ) substitutions. In further analyses, we included only the 1406 locus pairs for which the BLAST score  $\leq 10^{-3}$ , both  $K_A$  and  $K_S$  could be calculated, and >40% of codons were processed by Diverge (each locus from each species is included only once). This procedure resulted in 1.94 Mb of sequence processed by Diverge.

We calculated  $F_{op}$ , an index of codon usage bias based on *C. elegans* optimal codons, for each coding sequence for both *C. elegans* and *C. briggsae* (Stenico *et al.*, 1994; Sharp & Bradnam, 1997; Marais & Duret, 2001). Both  $K_A$  and  $K_S$  demonstrated significant negative relationships with species-averaged  $F_{op}$  ( $K_A = 0.363 - 0.454 \times F_{op}$ ,  $F_{1,1404} = 97.7$ ,  $P < 0.0001$ ,  $r^2 = 0.06$ ;  $K_S = 2.48 - 2.67 \times F_{op}$ ,  $F_{1,1404} = 709.5$ ,  $P < 0.0001$ ,  $r^2 = 0.34$ ). Consequently, we corrected  $K_S$  with the  $K_S$ – $F_{op}$  regression equation [at mean  $\bar{F}_{op} = 0.36$  where codon usage is expected to be unbiased (Stenico *et al.*, 1994);  $K_S|_{\bar{F}_{op}=0.36} = 1.52$ ]. We refer to these  $F_{op}$ -adjusted substitution rates as  $\delta_s$ . This procedure did not dramatically influence our  $U$  estimates (for the *C. elegans*–*C. briggsae* comparison, mean  $\delta_s$  is greater than  $K_S$  by 0.16). We used these same regression equations in our adjustment of  $K_S$  for codon bias in analyses involving *C. remanei*.

## Results

Our comparison of loci between the predominantly selfing *C. elegans* and *C. briggsae* and the obligately outcrossing *C. remanei* yields estimates of protein evolutionary rates and deleterious mutation that are surprisingly similar among taxa. The ratio of the rates of nonsynonymous to synonymous nucleotide substitution ( $K_A/\delta_s$ ) do not differ significantly between the three species-pair comparisons ( $F_{2,27} = 0.008$ , n.s.; Table 2). Rates of nonsynonymous site substitution ( $K_A$ ) also did not differ significantly ( $F_{2,27} = 0.16$ , n.s.; Table 2). These results also hold for comparisons of substitution rates unadjusted for codon usage bias. Furthermore, controlling for nonindependence of correlated histories with relative-rates tests yields no significant difference in the

**Table 2** Summary of average rates of evolution at nonsynonymous ( $K_A$ ) and synonymous ( $\delta_s$ ) sites.

Species	Loci	$K_A/\delta_s$	$K_A$	$\delta_s$
<i>elegans</i> – <i>briggsae</i>	1406	0.112 (0.003)	0.175 (0.005)	1.517 (0.011)
<i>elegans</i> – <i>briggsae</i>	10	0.117 (0.029)	0.196 (0.051)	1.571 (0.172)
<i>elegans</i> – <i>remanei</i>	10	0.113 (0.031)	0.182 (0.051)	1.536 (0.142)
<i>briggsae</i> – <i>remanei</i>	10	0.118 (0.030)	0.159 (0.045)	1.214 (0.140)

Standard errors are given in parentheses.

rate of substitution in the selfing *C. briggsae* compared with the outcrossing *C. remanei* in nine of the 10 loci, after correcting for multiple tests (Table 1). The one gene with significant differences among lineages (*glp-1*) has the second highest rate of nonsynonymous substitution in comparisons with *C. elegans*, and demonstrates a higher substitution rate in *C. briggsae*. This gene appears to drive the significance of the relative-rates test performed on concatenated sequence from all 10 genes ( $z = 3.11$ ,  $P = 0.0009$ ) because when *glp-1* is excluded from the concatenated sequence, the relative-rates test fails to achieve statistical significance ( $z = 1.56$ ,  $P = 0.06$ ).

Estimates of the genomic deleterious mutation rate do not differ among the species-comparisons: 95% bootstrap confidence intervals of  $U$  for each species-pair all overlap with each other. For example, assuming 90 generations/year for all species and the time to the most recent common ancestor of *C. elegans* and *C. briggsae/remanei* is 50 My and of *C. briggsae* and *C. remanei* is 44.5 My, then estimates of the deleterious mutation rate for these comparisons are all  $U \sim 0.02$  with overlapping 95% confidence intervals (Table 3). Lower values of  $U$  result if more distant times of divergence or more rapid turnover of generations are assumed (Fig. 2).

The mean (and distributions) of  $K_A$  and  $\delta_s$  for the large number of sequence comparisons between *C. elegans* and *C. briggsae* coincide closely for the set of 10 loci that were compared among all three species of *Caenorhabditis* (Table 2). This suggests that no obvious bias exists in the sample of loci used for the estimation of  $U$  and in relative-rates tests. An analysis of the ANOVA design for testing among estimates of  $K_A$  indicates that a difference of 0.058 in  $K_A/\delta_s$  can be detected with power >80% at  $\alpha = 0.05$ . This power analysis affirms that these data are sufficient to identify modest differences in substitution rates, although the relative-rates tests are phylogenetically more appropriate.

These data also allow the estimation of the proportion of the genome subject to selective constraint. The degree of constraint among coding regions may be estimated as  $1 - K_A/\delta_s$  and as  $U/M$  (Eyre-Walker & Keightley, 1999). There are approximately 20.5 Mb (19 099 genes  $\times$   $0.652 \times 1644$  sites/gene) of nonsynonymous sites and 97 Mb of total DNA in the *C. elegans* genome (The *C. elegans* Sequencing Consortium, 1998). Our estimate

**Table 3** Rates of total genomic ( $M$ ) and deleterious ( $U$ ) mutation per generation, and genomic constraint ( $U/M$ ). Divergence times for *C. elegans-briggsae/remanei* are given as the subscript  $Y$  (for *C. briggsae-remanei*, 0.78Y).  $U$  calculations are based on  $K_A$  and  $\delta_s$  and 90 generations/year.  $U'$  assumes  $M = 0.1286$  (see Methods).

Species pair	Loci	$U/M$	$M_{Y=50 \text{ My}}$	$U_{Y=50 \text{ My}}$	$M_{Y=120 \text{ My}}$	$U_{Y=120 \text{ My}}$	$U'_{Y=120 \text{ My}}$
<i>elegans-briggsae</i>	1406	0.886	0.0211 (0.0208–0.0215)	0.0187 (0.01841–0.0190)	0.0262 (0.1260–0.1263)	0.0088 (0.0086–0.0090)	0.0078 (0.0077–0.0079)
<i>elegans-briggsae</i>	10	0.831	0.0255 (0.0185–0.0318)	0.0212 (0.0154–0.0278)	0.1243 (0.1229–0.1266)	0.0088 (0.0064–0.0116)	0.1268 (0.1262–0.1278)
<i>elegans-remanei</i>	10	0.812	0.0223 (0.0188–0.0247)	0.0181 (0.0153–0.0217)	0.1243 (0.1230–0.1267)	0.0075 (0.0063–0.0091)	0.1268 (0.1262–0.1278)
<i>briggsae-remanei</i>	10	0.803	0.0238 (0.0190–0.0266)	0.0191 (0.0160–0.0222)	0.1239 (0.1224–0.1264)	0.0080 (0.0066–0.0092)	0.1266 (0.1260–0.1277)

Bootstrap 95% confidence intervals are given in parentheses.

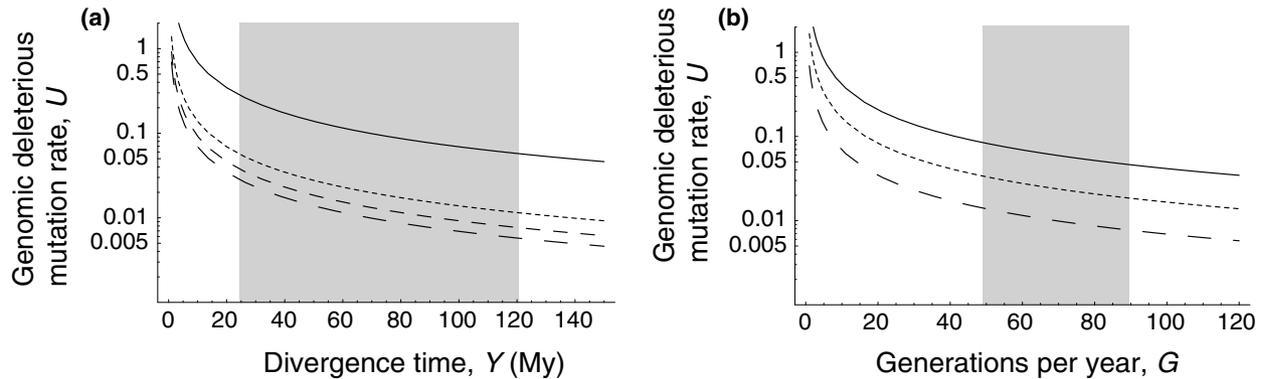
of constraint in coding regions ( $\sim 0.85$ ) therefore implies that  $\sim 18\%$  ( $0.85 \times 20.5 \text{ Mb}/97 \text{ Mb}$ ) of the genome is affected by strong purifying selection. This result can be contrasted with a recent estimate from humans of about 2% (Nachman & Crowell, 2000) using a similar approach, although both results are likely underestimates because they consider selection only on non-synonymous sites. A previous analysis of constraint in  $\sim 150 \text{ kb}$  of syntenic *C. elegans* and *C. briggsae* sequence concluded that at least 32% of sites in their genomes are functionally conserved (Shabalina & Kondrashov, 1999), based on 72.2% invariant exonic sites as well as invariant intronic and intergenic nucleotides. Consistent with the findings of Shabalina & Kondrashov (1999), we calculate 73.9% invariant coding region base pairs for the  $\sim 2 \text{ Mb}$  of sequence from the 1406 loci we compared between *C. elegans* and *C. briggsae*. Such a large extent of the genome subject to purifying selection, in conjunction with the partially selfing mode of reproduction, suggests that the action of background selection (Charlesworth *et al.*, 1993) may be likely to be a potent force influencing patterns of polymorphism in these species (Cutter & Payseur, 2003; Sivasundar & Hey, 2003).

## Discussion

### Selfing and outcrossing lineages do not differ in deleterious mutation rate

Our analyses demonstrate that estimates of the genomic deleterious mutation rate  $U$  from pairwise comparisons of coding sequences do not differ among the species *Caenorhabditis elegans*, *C. briggsae* and *C. remanei*. In all cases,  $U$  is  $< 1$  and likely falls in the range of 0.005–0.05, given a broad range of reasonable estimates of divergence and generation time. However, the absolute magnitude of the  $U$  estimates is immaterial to the argument that genomic deleterious mutation rates do not differ among lineages. Relative-rates tests, used to control for shared histories, confirm the conclusion that *C. remanei* genes generally exhibit rates of evolution no different than *C. briggsae* genes. These observations are particularly compelling in light of the different breeding systems represented in this clade: *C. remanei* is an obligately outcrossing species, whereas *C. elegans* and *C. briggsae* reproduce largely via self-fertilization. Mutational models applied to the relative fitness of selfing and outcrossing predict higher genomic deleterious mutation rates and lower rates of protein evolution in outcrossing lineages (Kondrashov, 1988; Charlesworth, 1990). Consequently, we infer that mutational models are unlikely to fully explain the evolution of outcrossing in this clade.

Alternative explanations for these results seem incompatible with available data for these species, although they are difficult to reject entirely. It is conceivable that



**Fig. 2** Dependence of the genomic deleterious mutation rate per generation ( $U$ ) on divergence time ( $Y$ ) and generation time ( $G$ ). Parameters are based on *C. elegans*–*briggsae*, but the curves for the other species comparisons are indistinguishable by eye. Curves in (a) show  $G = 12, 60, 90, 120$  generations/year descending from top to bottom. Curves in (b) show  $Y = 20, 50, 120$  My descending from top to bottom. Gray regions indicate values of  $Y$  or  $G$  that are consistent with empirical data.

longer generation times in *C. remanei* coupled with a higher mutation rate (and therefore a higher genomic deleterious mutation rate) could produce the same pattern that we report. However, laboratory data suggest that development times of these species do not differ dramatically (Byerly *et al.*, 1976) and relative-rates tests fail to detect elevated substitution rates along the *C. remanei* lineage. Differences in mutation rate among the three species could be masked by saturation at synonymous sites, which could compromise our conclusions if the mutation rate in *C. remanei* is much higher than in the other species. However, the facts that (1)  $K_A$  and  $\delta_s$  positively covary (Spearman's  $\rho = 0.45$ ;  $n = 1406$ ;  $P < 0.0001$ ) and that (2)  $\delta_s$  interpreted as a distance statistic is lower in the *C. briggsae*–*remanei* comparison (consistent with the tree topology) argue that  $\delta_s$  does capture biologically meaningful variation in mutation rates. Alternatively, one could contend that insufficient time has elapsed for mutation rates to evolve to different levels along each lineage (or, equivalently, that breeding system evolution occurred very recently). However, available estimates of divergence time between *C. elegans* and *C. briggsae* seem sufficiently large (23–120 My; Prasad & Baillie, 1989; Heschl & Baillie, 1990; Lee *et al.*, 1992; Kennedy *et al.*, 1993; Thacker *et al.*, 1999; Coghlan & Wolfe, 2002) for mutation rates to have evolved. Finally, no differences in  $U$  or  $K_A$  would be expected *a priori* if *C. elegans* and *C. briggsae* were cryptic obligate outcrossers, but this is unlikely given their low levels of genetic variation (Thomas & Wilson, 1991; Egilmez *et al.*, 1995; Koch *et al.*, 2000; Wicks *et al.*, 2001; Graustein *et al.*, 2002; Jovelin *et al.*, 2003; Sivasundar & Hey, 2003), low predicted rate of outcrossing (Hodgkin *et al.*, 1979; Hodgkin & Doniach, 1997; Chasnov & Chow, 2002; Cutter & Payseur, 2003; Cutter *et al.*, 2003) and negligible inbreeding depression or heterosis (Johnson & Hutchinson, 1993; Chasnov & Chow, 2002).

### Implications of equivalent deleterious mutation rates

Our conclusion that the MD hypothesis on its own does not explain the maintenance of outcrossing in caenorhabditids has three main implications. First, deleterious mutations may not have synergistic effects on fitness, on average, in this species group. Consistent with this implication, a recent test for synergistic epistasis between mutations affecting *C. elegans* life-history traits found only a nonsignificant trend for this mode of gene interaction (Peters & Keightley, 2000). The lack of a dominant role of synergistic interactions also has been shown in *Escherichia coli* (Elena & Lenski, 1997), viruses (Elena & Moya, 1999; de la Pena *et al.*, 2000), and *Saccharomyces cerevisiae* (Wloch *et al.*, 2001), although experiments with *Chlamydomonas* (de Visser *et al.*, 1996) and *Drosophila melanogaster* (Mukai, 1964; Kitagawa, 1967) have supported the notion of synergistic epistatic interactions. Theoretical work has also demonstrated that population structure (Agrawal & Chasnov, 2001; Otto & Barton, 2001), environmental heterogeneity (Lenormand & Otto, 2000), and variation in the sign of epistasis among loci (Otto & Feldman, 1997) can dramatically affect the expected benefit of recombination in purging deleterious mutations. However, when most deleterious mutations are recessive, as believed for *D. melanogaster* (Muller, 1950), the assumption of synergistic epistasis need not be met for the advantages of outcrossing to outweigh the costs (Chasnov, 2000). Furthermore, ecological conditions in nature could induce epistatic gene interactions that are unobservable in the laboratory (Peters & Keightley, 2000), and the degree of epistasis in the *C. briggsae* and *C. remanei* genomes remains unknown.

Secondly, factors other than selection against deleterious mutations are likely to be important in the evolution of outcrossing in caenorhabditids. Although one locus shows evidence for an elevated rate of protein evolution in

the selfing *C. briggsae* relative to the outcrossing *C. remanei*, stochastic mutational models (Muller, 1964; Hill & Robertson, 1966) seem unlikely to strongly influence the evolution of outcrossing in these species. The effect of Muller's ratchet will be weak when effective population size ( $N_e$ ) exceeds  $\sim 1000$  (Maynard Smith, 1978), and we suspect that these nematodes will generally experience very large population sizes and densities in nature (Mikola & Sulkava, 2001). The observation of significant levels of codon usage bias in *C. elegans* (Stenico *et al.*, 1994; Marais & Duret, 2001), for which selection coefficients ( $s$ ) are likely to be small (Akashi, 1995), also supports the notion of a large effective size of *C. elegans* populations (i.e.  $N_e \sim |1/s| \sim 10^6$ ). Rough calculations of  $N_e$  based on single nucleotide polymorphism data (Cutter & Payseur, 2003) and microsatellites under a stepwise mutation model (Sivasundar & Hey, 2003) give a range of  $N_e \sim 10^4$ – $10^6$ . Consequently, it is unlikely that the operation of Muller's ratchet or Hill–Robertson interference in selfing lineages could sufficiently offset the two-fold cost of sex.

The major alternative class of adaptive hypotheses designed to explain the origin and maintenance of sexual reproduction, but which also apply to the relative fitness of selfing and outcrossing, have been termed environmental–ecological models (Kondrashov, 1988; West *et al.*, 1999), of which Red Queen hypotheses are most favoured (Bell, 1982; West *et al.*, 1999). Red Queen models propose that biotic interactions, typically parasitic or pathogenic (Jaenike, 1978; Hamilton *et al.*, 1990), drive an antagonistic coevolution that leads to selection for different allelic combinations in the interacting species (Bell, 1982). Like the MD hypothesis, these models depend on epistatic interactions among alleles, although the specific form of epistasis depends on the particular model (Barton, 1995; West *et al.*, 1999). *Caenorhabditis elegans*, *C. briggsae* and *C. remanei* are known to participate in phoretic, necromenic relationships (the nematodes cling to a host and wait for it to die before consuming the bacteria that grow on the carcass) with slugs, snails and/or terrestrial isopods (Baird *et al.*, 1994; Baird, 1999), and *C. elegans* succumbs to some pathogenic bacterial species in the laboratory (Aballay *et al.*, 2000; Hodgkin *et al.*, 2000). However, it remains an open question as to the degree to which antagonistic coevolutionary scenarios involving caenorhabditids may be evolutionarily important in nature. This leaves open the possibility that both mutational and environmental–ecological models for the evolution of sex may be important simultaneously in obligately outcrossing caenorhabditids like *C. remanei*, by accounting for the 'cost of sex' in combination (West *et al.*, 1999).

Third, the consistent observation that  $U$  is  $< 1$  in this species group (Keightley & Caballero, 1997; Keightley & Bataillon, 2000; Vassilieva *et al.*, 2000) suggests that the magnitude of  $U$  may be more strongly influenced by factors unrelated to the evolution of breeding system. The notion that the number of germline cell divisions

contributes to the genomic deleterious mutation rate provides one alternative explanation for variation in  $U$  across taxa (Lynch *et al.*, 1999). Our data are consistent with, but not a direct test of, the germline cell division idea in that all three *Caenorhabditis* species exhibit a similar number of germline cell divisions and similar estimates of  $U$ . The germline cell division model also receives support from other studies: the number of mutations per genome per cell division appears to be relatively constant across many eukaryote species (Drake *et al.*, 1998), the per generation mutation rate increases with the age of reproduction of human males (Crow, 1993), and  $U$  scales with generation time (which correlates with germline cell division number) across species (Keightley & Eyre-Walker, 2000).

### **U estimates in *Caenorhabditis***

Of the species included in our analyses, genomic deleterious mutation rates have been estimated for phenotypes related to fitness only in *C. elegans*. Estimates of  $U$  for life-history characters from large mutation-accumulation experiments range between 0.002 (Keightley & Caballero, 1997) and 0.03 (Vassilieva & Lynch, 1999; Vassilieva *et al.*, 2000). A recent re-analysis of these published data indicates that the values may be closer to  $U \sim 0.005$  (Keightley & Bataillon, 2000), but because these mutation accumulation studies may fail to detect  $\sim 96\%$  of deleterious mutations (Davies *et al.*, 1999), these estimates may need to be scaled upward by a factor of  $\sim 25$  (i.e.  $U \sim 0.125$ ). Consequently, the magnitude of  $U$  for fitness in *C. elegans* is still considered an open question (Kondrashov, 2001). Our comparative sequence-based approach to estimating  $U$  in *Caenorhabditis*, given literature-based assumptions regarding generation and divergence time, yields values (0.005–0.05) that are comparable with those derived from the mutation accumulation studies. Although it is interesting to make use of this independent means of estimating  $U$ , it remains difficult to place great confidence in any particular estimate. Both experimental and comparative methods for calculating  $U$  yield minimum estimates, so these studies provide a lower bound for the genomic deleterious mutation rate. Potential sources of error for the comparative approach (in addition to uncertainty in divergence and generation times) include the inability to account for insertion/deletion mutations or mutations in noncoding regions, and the possibility that our proxy for neutral mutation rate is inaccurate. Combining extreme values for these sources of error into a single estimate provides a potential upper bound to  $U$  of  $\sim 0.82$  (23 My divergence, 50 generations/year, per site mutation rate based on  $9.1 \times 2.25 \times 10^{-10}$  mutations per generation (Drake *et al.*, 1998), dividing by 0.6 to account for constrained noncoding regions (Shabalina & Kondrashov, 1999), based on the 1406 *C. elegans*–*C. briggsae* loci).

Irrespective of the absolute magnitude of  $U$ , our evaluation of the MD hypothesis is probably not compromised by its underestimation or uncertainty in our estimates of other parameters because it is the relative values of  $U$  that are important for our analyses. The phylogenetic arrangement of the three caenorhabditids used here mitigates much of the uncertainty in the comparison of  $U$  values. Given that *C. briggsae* and *C. remanei* are sister species relative to *C. elegans*, the *C. elegans*–*C. briggsae* and *C. elegans*–*C. remanei* divergence times are equivalent. This motivated our test of the relative rates of evolution between *C. briggsae* and *C. remanei*, conditioned on the outgroup status of *C. elegans*. *Caenorhabditis remanei* does not demonstrate elevated rates of protein evolution, according with the observation that confidence intervals for the  $U$  estimates overlap.

## Conclusions

This study provides one of the first comparative tests of mutational models for the evolution of outcrossing in a clade that exhibits variation in breeding system (Wright *et al.*, 2002). Very few estimates of  $U$  are available in animal taxa that do not outcross obligately, or from clades that contain species that exhibit a variety of breeding systems. However, such a comparative approach should prove fruitful among the lineages with varying modes of reproduction in brassicaceous plants related to *Arabidopsis* and rotifers. In *A. thaliana*,  $U$  has been estimated by several means: a mutation accumulation study estimated  $U \sim 0.1$  (Schultz *et al.*, 1999), levels of inbreeding depression provide a value of  $\sim 0.5$  (Charlesworth *et al.*, 1990), and sequence data indicate  $U$  in *Arabidopsis* to be 0.2–0.6 (Wright *et al.*, 2002). Consistent with our findings in *Caenorhabditis*, Wright *et al.* (2002) found no significant difference in  $U$  estimates between the selfing *A. thaliana* and the outcrossing *A. lyrata*. Additionally, the ability for mutational models to explain patterns of variation in breeding system should depend on the frequency of outcrossing and the fitness function describing synergistic epistasis (Kimura & Maruyama, 1966). Given that a small amount of outcrossing may confer benefits under a variety of different models (Hurst & Peck, 1996), and that highly selfing, facultative outcrossers provide natural examples of the evolution of breeding system in progress, a thorough treatment of theoretical models that predict when mutational processes may be appropriate in partially selfing taxa will be useful (Kondrashov, 1985).

## Acknowledgments

We thank T. Harris and L. Stein for gracious database assistance and A. Streit (F. Muller and F. Gautron) for kindly providing unpublished *C. remanei* sequence. We also thank E. Haag for making preliminary data available to us. The constructive comments of L. Avilés, C.W.

Birky, H. Ochman, and several anonymous reviewers improved previous drafts of the manuscript. A Department of Defense National Defense Science and Engineering Graduate fellowship to A.D.C., an NSF Integrative Training Grant (IGERT) fellowship in Biology, Physics, and Mathematics to B.A.P., and NSF IGERT fellowships in Genomics to A.D.C. and B.A.P. facilitated this research.

## References

- Aballay, A., Yorgey, P. & Ausubel, F.M. 2000. *Salmonella typhimurium* proliferates and establishes a persistent infection in the intestine of *Caenorhabditis elegans*. *Curr. Biol.* **10**: 1539–1542.
- Agrawal, A.F. & Chasnov, J.R. 2001. Recessive mutations and the maintenance of sex in structured populations. *Genetics* **158**: 913–917.
- Akashi, H. 1995. Inferring weak selection from patterns of polymorphism and divergence at silent sites in *Drosophila* DNA. *Genetics* **139**: 1067–1076.
- Baird, S.E. 1999. Natural and experimental associations of *Caenorhabditis remanei* with *Trachelipus rathkii* and other terrestrial isopods. *Nematology* **1**: 471–475.
- Baird, S.E., Sutherland, M.E. & Emmons, S.W. 1992. Reproductive isolation in Rhabditidae (Nematoda, Secernentea) – mechanisms that isolate 6 species of 3 genera. *Evolution* **46**: 585–594.
- Baird, S.E., Fitch, D.H.A. & Emmons, S.W. 1994. *Caenorhabditis vulgaris* sp. n. (Nematoda, Rhabditidae) – a necromenic associate of pill bugs and snails. *Nematologica* **40**: 1–11.
- Baldwin, J.G., GiblinDavis, R.M., Eddleman, C.D., Williams, D.S., Vida, J.T. & Thomas, W.K. 1997. The buccal capsule of *Aduncospiculum halicti* (Nemata: Diplogasterina): an ultrastructural and molecular phylogenetic study. *Canadian Journal of Zoology-Revue Canadienne de Zoologie* **75**: 407–423.
- Barton, N.H. 1995. A general model for the evolution of recombination. *Genet. Res.* **65**: 123–144.
- Bell, G. 1982. *Masterpiece of Nature: The Evolution and Genetics of Sexuality*. Croom Helm, London.
- Birky, C.W. 1999. An even broader perspective on sex and recombination – commentary. *J. Evol. Biol.* **12**: 1013–1016.
- Byerly, L., Cassada, R.C. & Russell, R.L. 1976. Life-cycle of nematode *Caenorhabditis elegans*. 1. Wild-type growth and reproduction. *Dev. Biol.* **51**: 23–33.
- Charlesworth, B. 1980. The cost of sex in relation to mating system. *J. Theor. Biol.* **84**: 655–671.
- Charlesworth, B. 1990. Mutation–selection balance and the evolutionary advantage of sex and recombination. *Genet. Res.* **55**: 199–221.
- Charlesworth, B., Charlesworth, D. & Morgan, M.T. 1990. Genetic loads and estimates of mutation-rates in highly inbred plant-populations. *Nature* **347**: 380–382.
- Charlesworth, B., Morgan, M.T. & Charlesworth, D. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**: 1289–1303.
- Chasnov, J.R. 2000. Mutation selection balance, dominance and the maintenance of sex. *Genetics* **156**: 1419–1425.
- Chasnov, J.R. & Chow, K.L. 2002. Why are there males in the hermaphroditic species *Caenorhabditis elegans*? *Genetics* **160**: 983–994.

- Chen, P.J., Cho, S., Jin, S.W. & Ellis, R.E. 2001. Specification of germ cell fates by FOG-3 has been conserved during nematode evolution. *Genetics* **158**: 1513–1525.
- Coghlan, A. & Wolfe, K.H. 2002. Fourfold faster rate of genome rearrangement in nematodes than in *Drosophila*. *Genome Res.* **12**: 857–867.
- Crow, J.F. 1993. How much do we know about spontaneous human mutation-rates. *Environ. Mol. Mutagenesis* **21**: 122–129.
- Cutter, A.D. & Payseur, B.A. 2003. Selection at linked sites in the partial selfer *Caenorhabditis elegans*. *Mol. Biol. Evol.* **20**: 665–673.
- Cutter, A.D., Avilés, L. & Ward, S. 2003. The proximate determinants of sex ratio in *C. elegans* populations. *Genet. Res.* **81**: 91–102.
- Davies, E.K., Peters, A.D. & Keightley, P.D. 1999. High frequency of cryptic deleterious mutations in *Caenorhabditis elegans*. *Science* **285**: 1748–1751.
- Denver, D.R., Morris, K., Lynch, M., Vassilieva, L.L. & Thomas, W.K. 2000. High direct estimate of the mutation rate in the mitochondrial genome of *Caenorhabditis elegans*. *Science* **289**: 2342–2344.
- Drake, J.W., Charlesworth, B., Charlesworth, D. & Crow, J.F. 1998. Rates of spontaneous mutation. *Genetics* **148**: 1667–1686.
- Egilméz, N.K., Ebert, R.H. & Reis, R.J.S. 1995. Strain evolution in *Caenorhabditis elegans* – transposable elements as markers of interstrain evolutionary history. *J. Mol. Evol.* **40**: 372–381.
- Elena, S.F. & Lenski, R.E. 1997. Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390**: 395–398.
- Elena, S.F. & Moya, A. 1999. Rate of deleterious mutation and the distribution of its effects on fitness in vesicular stomatitis virus. *J. Evol. Biol.* **12**: 1078–1088.
- Eyre-Walker, A. & Keightley, P.D. 1999. High genomic deleterious mutation rates in hominids. *Nature* **397**: 344–347.
- Fitch, D.H.A. & Thomas, W.K. 1997. Evolution. In: *C. elegans II* (D. L. Riddle, T. Blumenthal, B. J. Meyer & J. R. Priess, eds), pp. 815–850. Cold Spring Harbor Laboratory Press, New York.
- Fitch, D.H.A., Bugajgaweda, B. & Emmons, S.W. 1995. 18S ribosomal-RNA gene phylogeny for some Rhabditiidae related to *Caenorhabditis*. *Mol. Biol. Evol.* **12**: 346–358.
- Gabriel, W., Lynch, M. & Burger, R. 1993. Muller's ratchet and mutational meltdown. *Evolution* **47**: 1744–1757.
- Graustein, A., Gaspar, J.M., Walters, J.R. & Palopoli, M.F. 2002. Levels of DNA polymorphism vary with mating system in the nematode genus *Caenorhabditis*. *Genetics* **161**: 99–107.
- Haag, E.S. & Kimble, J. 2000. Regulatory elements required for development of *Caenorhabditis elegans* hermaphrodites are conserved in the tra-2 homologue of *C. remanei*, a male/female sister species. *Genetics* **155**: 105–116.
- Hamilton, W.D., Axelrod, R. & Tanese, R. 1990. Sexual reproduction as an adaptation to resist parasites (a review). *Proc. Natl. Acad. Sci. USA* **87**: 3566–3573.
- Heschl, M.F.P. & Baillie, D.L. 1990. Functional elements and domains inferred from sequence comparisons of a heat-shock gene in 2 nematodes. *J. Mol. Evol.* **31**: 3–9.
- Hill, W.G. & Robertson, A. 1966. Effect of linkage on limits to artificial selection. *Genet. Res.* **8**: 269–294.
- Hodgkin, J. & Doniach, T. 1997. Natural variation and copulatory plug formation in *Caenorhabditis elegans*. *Genetics* **146**: 149–164.
- Hodgkin, J., Horvitz, H.R. & Brenner, S. 1979. Nondisjunction mutants of the nematode *Caenorhabditis elegans*. *Genetics* **91**: 67–94.
- Hodgkin, J., Kuwabara, P.E. & Corneliusen, B. 2000. A novel bacterial pathogen, *Microbacterium nematophilum*, induces morphological change in the nematode *C. elegans*. *Curr. Biol.* **10**: 1615–1618.
- Hurst, L.D. & Peck, J.R. 1996. Recent advances in understanding of the evolution and maintenance of sex. *Trends Ecol. Evol.* **11**: 310–310.
- Jaenike, J. 1978. An hypothesis to account for the maintenance of sex within populations. *Evol. Theor.* **3**: 191–194.
- Johnson, T.E. & Hutchinson, E.W. 1993. Absence of strong heterosis for life-span and other life-history traits in *Caenorhabditis elegans*. *Genetics* **134**: 465–474.
- Jovelin, R., Ajie, B.C. & Phillips, P.C. 2003. Molecular evolution and quantitative variation for chemosensory behaviour in the nematode genus *Caenorhabditis*. *Mol. Ecol.* **12**: 1325–1337.
- Keightley, P.D. & Bataillon, T.M. 2000. Multigeneration maximum-likelihood analysis applied to mutation-accumulation experiments in *Caenorhabditis elegans*. *Genetics* **154**: 1193–1201.
- Keightley, P.D. & Caballero, A. 1997. Genomic mutation rates for lifetime reproductive output and lifespan in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **94**: 3823–3827.
- Keightley, P.D. & Eyre-Walker, A. 2000. Deleterious mutations and the evolution of sex. *Science* **290**: 331–333.
- Kennedy, B.P., Aamodt, E.J., Allen, F.L., Chung, M.A., Heschl, M.F.P. & McGhee, J.D. 1993. The gut esterase gene (*ges-1*) from the nematodes *Caenorhabditis elegans* and *Caenorhabditis briggsae*. *J. Mol. Biol.* **229**: 890–908.
- Kent, W.J. & Zahler, A.M. 2000. Conservation, regulation, synteny, and introns in a large-scale *C. briggsae*-*C. elegans* genomic alignment. *Genome Res.* **10**: 1115–1125.
- Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide-sequences. *J. Mol. Evol.* **16**: 111–120.
- Kimura, M. & Maruyama, T. 1966. Mutational load with epistatic gene interactions in fitness. *Genetics* **54**: 1337–1351.
- Kitagawa, O. 1967. Interaction in fitness between lethal genes in heterozygous condition in *Drosophila melanogaster*. *Genetics* **57**: 809–820.
- Koch, R., van Luenen, H.G.A.M., van der Horst, M., Thijssen, K.L. & Plasterk, R.H.A. 2000. Single nucleotide polymorphisms in wild isolates of *Caenorhabditis elegans*. *Genome Res.* **10**: 1690–1696.
- Kondrashov, A.S. 1985. Deleterious mutations as an evolutionary factor. 2. Facultative apomixis and selfing. *Genetics* **111**: 635–653.
- Kondrashov, A.S. 1988. Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**: 435–440.
- Kondrashov, A.S. 2001. Sex and U. *Trends Genet.* **17**: 75–77.
- Kondrashov, A.S. & Crow, J.F. 1993. A molecular approach to estimating the human deleterious mutation-rate. *Hum. Mutat.* **2**: 229–234.
- Lee, Y.H., Huang, X.Y., Hirsh, D., Fox, G.E. & Hecht, R.M. 1992. Conservation of gene organization and transsplicing in the glyceraldehyde-3-phosphate dehydrogenase-encoding genes of *Caenorhabditis briggsae*. *Gene* **121**: 227–235.
- Lenormand, T. & Otto, S.P. 2000. The evolution of recombination in a heterogeneous environment. *Genetics* **156**: 423–438.
- Lenski, R.E. 1999. A distinction between the origin and maintenance of sex – commentary. *J. Evol. Biol.* **12**: 1034–1035.
- Li, W.-H. 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J. Mol. Evol.* **36**: 96–99.

- Li, W.-H., Wu, C.I. & Luo, C.C. 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* **2**: 150–174.
- Lloyd, D.G. 1980. Benefits and handicaps of sexual reproduction. *Evol. Biol.* **13**: 69–111.
- Lynch, M., Burger, R., Butcher, D. & Gabriel, W. 1993. The mutational meltdown in asexual populations. *J. Hered.* **84**: 339–344.
- Lynch, M., Blanchard, J., Houle, D., Kibota, T., Schultz, S., Vassilieva, L. & Willis, J. 1999. Perspective: spontaneous deleterious mutation. *Evolution* **53**: 645–663.
- Manly, B.F.J. 1997. *Randomization, Bootstrap and Monte Carlo Methods in Biology*, 2nd edn. Chapman & Hall, New York.
- Marais, G. & Duret, L. 2001. Synonymous codon usage, accuracy of translation, and gene length in *Caenorhabditis elegans*. *J. Mol. Evol.* **52**: 275–280.
- Maynard Smith, J. 1978. *The Evolution of Sex*. Cambridge University Press, New York.
- Mikola, J. & Sulkava, P. 2001. Responses of microbial-feeding nematodes to organic matter distribution and predation in experimental soil habitat. *Soil Biol. Biochem.* **33**: 811–817.
- Mukai, T. 1964. Genetic structure of natural populations of *Drosophila melanogaster*. I. Spontaneous mutation rate of polygenes controlling viability. *Genetics* **50**: 1–9.
- Muller, H.J. 1950. Our load of mutations. *Am. J. Hum. Genet.* **2**: 111–176.
- Muller, H.J. 1964. The relation of recombination to mutational advance. *Mutat. Res.* **1**: 2–9.
- Muse, S.V. & Weir, B.S. 1992. Testing for equality of evolutionary rates. *Genetics* **132**: 269–276.
- Nachman, M.W. & Crowell, S.L. 2000. Estimate of the mutation rate per nucleotide in humans. *Genetics* **156**: 297–304.
- Ohta, T. 1973. Slightly deleterious mutant substitutions in evolution. *Nature* **246**: 96–98.
- Otto, S.P. & Barton, N.H. 2001. Selection for recombination in small populations. *Evolution* **55**: 1921–1931.
- Otto, S.P. & Feldman, M.W. 1997. Deleterious mutations, variable epistatic interactions, and the evolution of recombination. *Theor. Popul. Biol.* **51**: 134–147.
- Pamilo, P. & Bianchi, N.O. 1993. Evolution of the *zfx* and *zfy* genes – rates and interdependence between the genes. *Mol. Biol. Evol.* **10**: 271–281.
- de la Pena, M., Elena, S.F. & Moya, A. 2000. Effect of deleterious mutation-accumulation on the fitness of RNA bacteriophage MS2. *Evolution* **54**: 686–691.
- Peters, A.D. & Keightley, P.D. 2000. A test for epistasis among induced mutations in *Caenorhabditis elegans*. *Genetics* **156**: 1635–1647.
- Prasad, S.S. & Baillie, D.L. 1989. Evolutionarily conserved coding sequences in the *dpy-20-unc-22* region of *Caenorhabditis elegans*. *Genomics* **5**: 185–198.
- Riddle, D.L. 1988. The dauer larva. In: *The Nematode Caenorhabditis elegans* (W. B. Wood, ed.), pp. 393–412. Cold Spring Harbor Laboratory Press, New York.
- Riddle, D.L., Blumenthal, T., Meyer, B.J. & Priess, J.R. 1997. Introduction to *C. elegans*. In: *C. elegans II* (D. L. Riddle, T. Blumenthal, B. J. Meyer & J. R. Priess, eds), pp. 1–22. Cold Spring Harbor Laboratory Press, New York.
- Rudel, D. & Kimble, J. 2001. Conservation of *glp-1* regulation and function in nematodes. *Genetics* **157**: 639–654.
- Sarich, V.M. & Wilson, A.C. 1973. Generation time and genomic evolution in primates. *Science* **179**: 1144–1147.
- Schultz, S.T., Lynch, M. & Willis, J.H. 1999. Spontaneous deleterious mutation in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **96**: 11393–11398.
- Shabalina, S.A. & Kondrashov, A.S. 1999. Pattern of selective constraint in *C. elegans* and *C. briggsae* genomes. *Genet. Res.* **74**: 23–30.
- Sharp, P.M. & Bradnam, K.R. 1997. Codon usage in *C. elegans*. In: *C. elegans II* (D. L. Riddle, T. Blumenthal, B. J. Meyer & J. R. Priess, eds), pp. 1053–1057. Cold Spring Harbor Laboratory Press, New York.
- Sivasundar, A. & Hey, J. 2003. Population genetics of *Caenorhabditis elegans*: the paradox of low polymorphism in a widespread species. *Genetics* **163**: 147–157.
- Stenico, M., Lloyd, A.T. & Sharp, P.M. 1994. Codon usage in *Caenorhabditis elegans* – delineation of translational selection and mutational biases. *Nucleic Acids Res.* **22**: 2437–2446.
- Sudhaus, W. & Fitch, D.H.A. 2001. Comparative studies on the phylogeny and systematics of the Rhabditidae (Nematoda). *J. Nematol.* **33**: 1–70.
- Sudhaus, W. & Kiontke, K. 1996. Phylogeny of *Rhabditis* subgenus *Caenorhabditis* (Rhabditidae, Nematoda). *J. Zool. Syst. Evol. Res.* **34**: 217–233.
- Thacker, C., Marra, M.A., Jones, A. & Baillie, D.L. 1999. Functional genomics in *Caenorhabditis elegans*: an approach involving comparisons of sequences from related nematodes. *Genome Res.* **9**: 348–359.
- The *C. elegans* Sequencing Consortium 1998. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**: 2012–2018.
- Thomas, W.K. & Wilson, A.C. 1991. Mode and tempo of molecular evolution in the nematode *Caenorhabditis* – cytochrome oxidase-II and calmodulin sequences. *Genetics* **128**: 269–279.
- Vassilieva, L.L. & Lynch, M. 1999. The rate of spontaneous mutation for life-history traits in *Caenorhabditis elegans*. *Genetics* **151**: 119–129.
- Vassilieva, L.L., Hook, A.M. & Lynch, M. 2000. The fitness effects of spontaneous mutations in *Caenorhabditis elegans*. *Evolution* **54**: 1234–1246.
- Venette, R.C. & Ferris, H. 1997. Thermal constraints to population growth of bacterial-feeding nematodes. *Soil Biol. Biochem.* **29**: 63–74.
- de Visser, J.A.G.M., Hoekstra, R.F. & van den Ende, H. 1996. The effect of sex and deleterious mutations on fitness in *Chlamydomonas*. *Proc. R. Soc. Lond. B* **263**: 193–200.
- West, S.A., Lively, C.M. & Read, A.F. 1999. A pluralist approach to sex and recombination. *J. Evol. Biol.* **12**: 1003–1012.
- Wicks, S.R., Yeh, R.T., Gish, W.R., Waterston, R.H. & Plasterk, R.H.A. 2001. Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. *Nat. Genet.* **28**: 160–164.
- Wloch, D.M., Borts, R.H. & Korona, R. 2001. Epistatic interactions of spontaneous mutations in haploid strains of the yeast *Saccharomyces cerevisiae*. *J. Evol. Biol.* **14**: 310–316.
- Wright, S.I., Lauga, B. & Charlesworth, D. 2002. Rates and patterns of molecular evolution in inbred and outbred *Arabidopsis*. *Mol. Biol. Evol.* **19**: 1407–1420.
- Wu, C.I. & Li, W.H. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. *Proc. Natl. Acad. Sci. USA* **82**: 1741–1745.

Received 16 January 2003; revised 16 June 2003; accepted 19 June 2003