Levels of DNA Polymorphism Vary With Mating System in the Nematode Genus Caenorhabditis

Andrew Graustein, John M. Gaspar, James R. Walters and Michael F. Palopoli[1]

Department of Biology, Bowdoin College, Brunswick, Maine 04011

Manuscript received September 26, 2001
Accepted for publication February 1, 2002

ABSTRACT

Self-fertilizing species often harbor less genetic variation than cross-fertilizing species, and at least four different models have been proposed to explain this trend. To investigate further the relationship between mating system and genetic variation, levels of DNA sequence polymorphism were compared among three closely related species in the genus Caenorhabditis: two self-fertilizing species, *Caenorhabditis elegans* and *C. briggsae*, and one cross-fertilizing species, *C. remanei*. As expected, estimates of silent site nucleotide diversity were lower in the two self-fertilizing species. For the mitochondrial genome, diversity in the selfing species averaged 42% of diversity in *C. remanei*. Interestingly, the reduction in genetic variation was much greater for the nuclear than for the mitochondrial genome. For two nuclear genes, diversity in the selfing species averaged 6 and 13% of diversity in *C. remanei*. We argue that either population bottlenecks or the repeated action of natural selection, coupled with high levels of selfing, are likely to explain the observed reductions in species-wide genetic diversity.

A fundamental goal of population genetics is to understand the forces maintaining genetic variation in natural populations. Since different evolutionary processes are expected to have different effects on the genetic variation found within a species, it is possible to use trends in patterns of DNA sequence variation to identify the forces that drive evolution at the molecular level (see Kimura 1983; Li 1997).

For example, studies of *Drosophila melanogaster* have revealed that genes situated in regions of the genome with greatly reduced rates of recombination (crossing over) are much less variable than genes in regions with normal rates of recombination (Aguadé *et al.* 1989; Berry *et al.* 1991; Begun and Aquadro 1991, 1992; Langley *et al.* 1993). Subsequent work has shown that this positive correlation between recombination and variation is a characteristic shared by a wide range of taxa, including humans (Nachman 1997, 2001; Dvořák *et al.* 1998; Kraft *et al.* 1998; Nachman *et al.* 1998; Stephan and Langley 1998). One model proposed to explain this pattern is genetic hitchhiking, in which positive selection occasionally "sweeps away" polymorphisms in regions of low recombination, since it is here that the largest segments will be carried along with every advantageous allele that goes to fixation (Maynard Smith and Haigh 1974; Kaplan *et al.* 1989; Braverman *et al.* 1995). An alternative model is background selection, in which a steady rain of deleterious mutations drives variation out of regions of low recombination, since it is here that the largest segments will be dragged to eventual loss along with every deleterious allele that is eliminated (Charlesworth *et al.* 1993; Hudson and Kaplan 1995; Nordborg *et al.* 1996). Much current research in molecular population genetics is focused on testing patterns of DNA sequence variation against detailed predictions that arise from these models (for recent reviews see Aquadro 1997; Charlesworth and Charlesworth 1998).

Genetic diversity also varies in a consistent manner between species with divergent mating systems: Self-fertilizing species often harbor less genetic variation than cross-fertilizing species (Hamrick and Godt 1990, 1996; Schoen and Brown 1991; Jarne and Städler 1995; Charlesworth and Yang 1998; Liu *et al.* 1998, 1999; Baudry *et al.* 2001). This trend across mating systems provides another opportunity to study the forces that drive evolution at the molecular level (see Charlesworth and Wright 2001 for a recent review).

Reproduction by self-fertilization can be considered an extreme form of inbreeding, resulting in high levels of homozygosity. For neutral alleles, this is expected to decrease the effective population size (N_e) for autosomal genes by a factor of $(2 - s)/2$, where s is the selfing rate (Pollak 1987). In a population of complete selfers ($s = 1$), this will decrease N_e by 50% relative to cross-fertilizers. According to the neutral theory of molecular evolution, equilibrium levels of neutral variability should be proportional to N_e (Kimura 1971; Hudson 1990), so the neutral model predicts that a population

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. AF491458–AF491469, AF491508–491543, AF491452–AF491457, AF491470–AF491507, AF492686–AF492701.

[1]*Corresponding author:* Department of Biology, Bowdoin College, 6500 College Station, Brunswick, ME 04011.
E-mail: mpalopol@bowdoin.edu

of complete selfers will harbor ~50% of the diversity present in a similar population of cross-fertilizers (*e.g.*, Nordborg 2000).

At least three additional models have been proposed to account for the reduced levels of variation observed in self-fertilizing species. First, extreme bottlenecks in population size may be more common in self-fertilizers, since a single individual can found a new population, and self-fertilizing species are often associated with the founding of isolated populations (Baker 1955, 1967; Cox 1989; Schoen and Brown 1991). Population bottlenecks result in the loss of genetic variation (Nei *et al.* 1975), so repeated bottlenecks could explain the reduced diversity observed in selfing species. Second, if there are balanced polymorphisms that are favored by overdominant selection (heterozygote advantage), these may be lost more easily in highly selfing populations, since there are few heterozygous individuals around to reap the benefits of heterozygosity (Kimura and Ohta 1971; Charlesworth and Charlesworth 1998). The loss of these balanced polymorphisms, along with any linked variation that might have accumulated between alleles, could also explain the reduced diversity in selfers. Third, the increase in homozygosity that results from selfing is expected to reduce the effectiveness of recombination, which depends on the frequency of double heterozygotes (Narrain 1966). This reduction in effective recombination rate, in turn, should increase the impact that genetic hitchhiking and background selection have on levels of variation (Hedrick 1980; Charlesworth *et al.* 1997). If selection indeed has a greater effect on neutral variation in selfing species, this could also explain the observed reduction in diversity.

Although there have been several recent studies of molecular variation in highly selfing species (Bergelson *et al.* 1998; Dvorák *et al.* 1998; Kraft *et al.* 1998; Liu *et al.* 1998, 1999; Savolainen *et al.* 2000; Baudry *et al.* 2001), it has proven difficult to distinguish between these different models. For example, in the plant genus Leavenworthia, populations of self-fertilizers harbor much less diversity at nuclear genes than do closely related populations of cross-fertilizers (Charlesworth and Yang 1998; Liu *et al.* 1998, 1999). The observed reduction in diversity is too great to explain by simply incorporating high levels of selfing into the neutral model, but it is not clear which additional evolutionary mechanism(s) should be invoked.

One way to begin distinguishing between models is to compare patterns of variation across mating systems for mitochondrial *vs.* nuclear genes. Mitochondrial DNA (mtDNA) exhibits strict maternal inheritance in many species (Birky 1978). As a result, the effective population size experienced by mtDNA (N_{mt}) is proportional to the number of breeding females (Birky *et al.* 1983). In gonochoristic species, this is often ~50% of the breeding adults; in self-fertilizing species, however, it may approach 100% of the breeding adults. All other things being equal, therefore, neutral models do not predict a consistent reduction in the N_{mt} of self-fertilizers relative to cross-fertilizers; instead, the increased number of individuals that transmit mitochondria to progeny should tend to result in an *increase* in N_{mt}. Furthermore, since mtDNA is cytoplasmic and effectively haploid, N_{mt} should not be affected by the tendency for selfers to lose balanced polymorphisms due to the loss of fit heterozygotes. Thus mating system differences are expected to cause reduced diversity primarily in nuclear genes. These discrepancies between the predictions for mitochondrial *vs.* nuclear genes may provide a means to begin distinguishing between alternative models.

The purpose of this study was to examine the relationship between mating system and DNA sequence variation in the genus Caenorhabditis and to compare these results for nuclear *vs.* mitochondrial genes. Members of the genus Caenorhabditis are free-living, soil nematodes with remarkably similar morphologies but divergent mating systems (Sudhaus and Kiontke 1996; Fitch and Thomas 1997). Phylogenetic analyses suggest that gonochorism (males and females) is the ancestral mating system in this genus and that hermaphroditism (males and self-fertilizing hermaphrodites) either evolved twice independently in the *elegans* species group or evolved once in the common ancestor of this group and was subsequently lost in one lineage (Sudhaus and Kiontke 1996; Baldwin *et al.* 1997; Fitch and Thomas 1997; Rudel and Kimble 2001). Because *Caenorhabditis elegans* is a model organism with a fully sequenced genome (*C. elegans* Sequencing Consortium 1998), it is possible to study genes that have well-characterized genomic environments in at least one species.

Genetic diversity was sampled in three closely related species: hermaphroditic *C. elegans*, hermaphroditic *C. briggsae*, and gonochoristic *C. remanei*. Polymorphism was quantified for *tra-2*, a gene that plays a central role in the sex-determination pathway (Kuwabara and Kimble 1995; Haag and Kimble 2000); *glp-1*, a *Notch*-related receptor that is required for certain induction events during embryogenesis as well as germline proliferation in adults (Austin and Kimble 1989; Rudel and Kimble 2001); *spe-9*, a sperm transmembrane protein that is required for fertilization and appears to function in sperm-egg interactions (Singson *et al.* 1998, 1999); and *COII* (cytochrome oxidase subunit II), a mitochondrial gene involved in cellular respiration (Thomas and Wilson 1991). The nuclear genes studied were all located on autosomes, and they were chosen to span a wide range of local recombination rates in *C. elegans* (Barnes *et al.* 1995). Levels of DNA sequence polymorphism were compared between species to test current models for the effect of mating system on genetic diversity.

MATERIALS AND METHODS

Nematode samples: *C. elegans*, *C. briggsae*, and *C. remanei* are members of the *elegans* species group (Sudhaus and Kiontke 1996; Fitch and Thomas 1997). All three are widespread

TABLE 1

Nematode strains sampled to estimate diversity

Species	Strain	Origin
C. elegans	AB1	Adelaide, Australia
	AB2	Adelaide, Australia
	CB4507	Palm Canyon, CA
	CB4851	Bergerac, France
	CB4852	Rothamsted, England
	CB4853	Altadena, CA
	CB4854	Altadena, CA
	CB4855	Palo Alto, CA
	CB4856	Hawaii
	CB4857	Claremont, CA
	CB4858	Pasadena, CA
	CB4932	Taunton, England
	DH424	El Prieto Canyon, CA
	DR1344	Bergerac, France
	N2	Bristol, England
	PB303	Unknown[a]
	PB305	Unknown[a]
	PB306	Unknown[b]
	BP307	Unknown[b]
	TR388	Madison, WI
C. briggsae	AF16	Ahmedabad, India
	HK104	Okayama, Japan
	HK105	Sendai, Japan
	PB800	Dayton, OH
	PB826	Dayton, OH
	VT847	Hawaii
C. remanei	CR1014	Gloucester, MA
	CR1415	Gloucester, MA
	CR2124	Gloucester, MA
	EM464	Brooklyn, NY
	PB205	Dayton, OH
	PB206	Dayton, OH
	PB212	Dayton, OH
	PB219	Dayton, OH
	PB228	Dayton, OH
	PB229	Dayton, OH
	SB146	Freiburg, Germany
	VT733	New Haven, CT

[a] Associated with an isopod (*Porcellio scaber*) purchased from Ward's Biological and Lab Supplies.

[b] Associated with an isopod (*Porcellio scaber*) purchased from Connecticut Valley Biological Supply.

geographically, and strains from around the world were chosen to sample much of the variation present within each species (Table 1). Regardless of the strain examined, *C. elegans* and *C. briggsae* cultures consisted almost entirely of self-fertilizing hermaphrodites, with rare males arising spontaneously at a frequency of $\ll 1\%$ (Honda 1925; data not shown); all *C. remanei* cultures, however, consisted of males and females at approximately equal frequencies (data not shown). Stocks were obtained from the Caenorhabditis Genetics Center, or collected from the wild, and maintained under standard conditions (Wood 1988). Mating tests were conducted to confirm the species identity of all strains used (data not shown). All strains of *C. remanei* were bottlenecked through single-pair, brother-sister matings for at least five generations prior to sampling to reduce within-strain heterozygosity.

DNA isolation, PCR amplification, and DNA sequencing: For each strain, total DNA was isolated from a pool of individuals using the DNeasy tissue kit (QIAGEN, Valencia, CA). Portions of the *tra-2*, *glp-1*, *spe-9*, and *COII* genes were amplified using the polymerase chain reaction (PCR) under standard reaction conditions (Saiki *et al.* 1988). Species-specific, oligonucleotide primers (Life Technologies) were designed, on the basis of published sequences, to amplify portions of the *tra-2*, *glp-1*, and *spe-9* genes (primer sequences are available upon request). Universal primers were used to amplify a portion of the mitochondrial *COII* gene (Thomas and Wilson 1991). All nuclear gene fragments were chosen to include both exons and introns (the mitochondrial *COII* gene does not contain introns). Published data were used to estimate nucleotide polymorphism for the mitochondrial *COII* gene in *C. elegans* (Thomas and Wilson 1991). Polymorphism levels for the *spe-9* gene were measured only in *C. elegans*, with the goal of obtaining an estimate of nucleotide diversity in this species for a gene situated in a region of the genome with relatively high recombination rates (Barnes *et al.* 1995). All amplified fragments were sequenced with the ABI PRISM dRhodamine Terminator cycle sequencing ready reaction kit (Perkin-Elmer, Norwalk, CT) and separated on an ABI Prism 377 automated DNA sequencer (Perkin-Elmer) at the Genotyping and Sequencing Center of the University of Chicago. Overlapping forward- and reverse-strand sequences were obtained in all cases.

Data analysis: Nucleotide sequences were compiled using EDITSEQ, SEQMAN II, and MEGALIGN software (DNASTAR, Madison, WI). Sequences were aligned using the MEGALIGN ClustalW algorithm (Thompson *et al.* 1994). Some *C. remanei* strains appeared to be heterozygous for polymorphic nucleotides at the two nuclear genes sampled (data not shown). None of these intrastrain differences identified polymorphisms that were not apparent on the basis of comparisons among the rest of the (homozygous) strains. Overall, a heterozygous nucleotide was observed in four strains for the *tra-2* gene (CR1415, PB206, PB228, and PB229) and in three strains for the *glp-1* gene (PB205, PB206, and SB146). In each case, to be conservative when testing the prediction that diversity would be greatest in *C. remanei*, the heterozygous nucleotide was recorded as the most common of the two possibilities. The DnaSP version 3.53 software (Rozas and Rozas 1999) was used to estimate population genetic parameters (Nei 1987) and to perform neutrality tests based on the frequency distributions of segregating sites (Tajima 1989; Fu and Li 1993; Fu 1997). Neutrality tests based on haplotype number (K) and haplotype diversity (H) were conducted by comparing observed values of K and H to their $\sim 95\%$ confidence intervals under strict neutrality (Depaulis and Veuille 1998). For each gene fragment and each species, nucleotide diversity was estimated for the entire fragment sequenced (π_t) and for silent sites separately (π_{si}). Silent sites were defined to include both synonymous coding sites and intron sites. Approximate 95% confidence intervals were obtained for π_{si} using Monte Carlo simulations based on the coalescent process, as implemented in DnaSP version 3.53. These simulations assumed a neutral, infinite-sites model, with a large and constant population size and no recombination. All simulations were conducted by fixing the number of segregating sites to that observed in the sample. The empirical distribution of the statistic was generated by simulating the evolution of 10,000 independent replicate populations, and this distribution was used to determine approximate confidence intervals. Local rates of recombination for nuclear genes were estimated by fitting polynomial regressions to Marey plots (results not shown) on the basis of genetic map data available in Wormbase (release WS48: August 3, 2001; Stein *et al.* 2001) and using the Mathematica version 4.0 software (Wolfram Research, Champaign, IL). A small number of polymorphic insertions/deletions were detected in each species, but were excluded from the analyses.

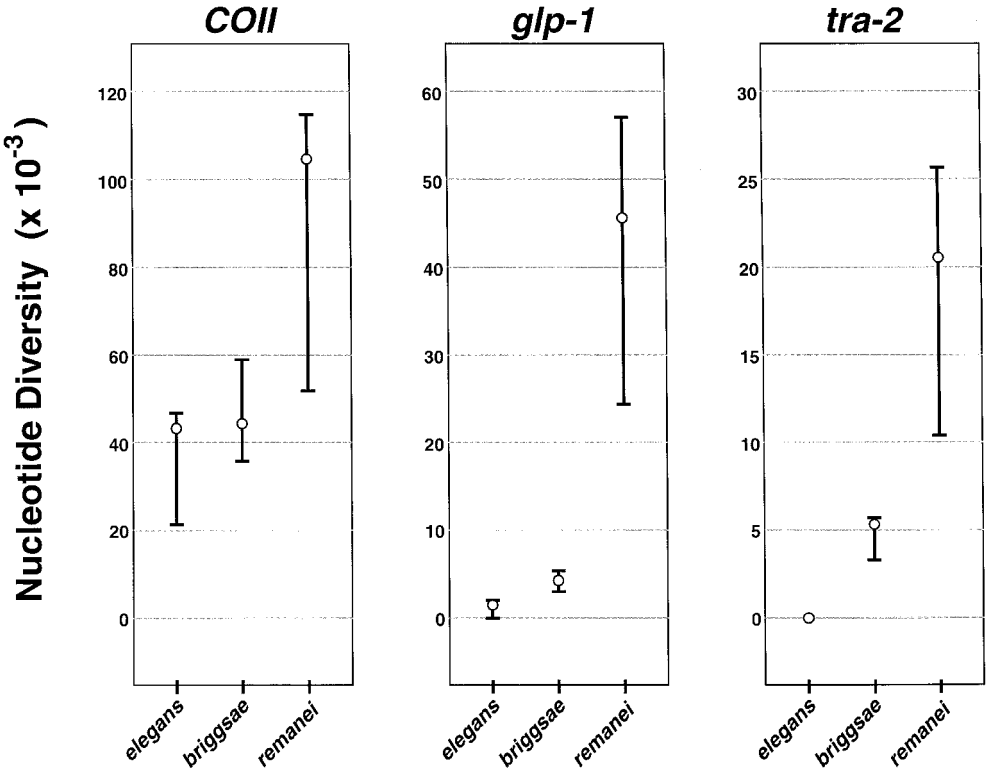

FIGURE 1.—Silent site nucleotide diversity for three genes (*COII*, *glp-1*, and *tra-2*) in three species of nematode (*C. elegans*, *C. briggsae*, and *C. remanei*). Open circles represent estimates of silent site nucleotide diversity that are based on pairwise sequence comparisons. Error bars represent the ~95% confidence intervals for this estimate, determined by Monte Carlo simulation of the coalescent process as described in MATERIALS AND METHODS. The self-fertilizing species (*C. elegans* and *C. briggsae*) always harbor less variation than the cross-fertilizing species (*C. remanei*). This difference between mating systems is more pronounced for the two nuclear genes (*glp-1* and *tra-2*) than for the mitochondrial gene (*COII*).

To conduct a nonparametric test of the hypothesis that the cross-fertilizing species (*C. remanei*) harbors more genetic variation overall than either of the self-fertilizing species (*C. elegans* and *C. briggsae*), the Mann-Whitney *U*-test was performed using Statview version 5 software (SAS Institute) on the basis of silent site nucleotide diversity estimates for each species and gene fragment. Neighbor-joining trees (SAITOU and NEI 1987) were drawn using the MEGA version 2.1 software (KUMAR *et al.* 2000), and the resulting diagrams were then combined with MacDraw Pro version 1.5 software (Claris).

RESULTS

Nucleotide diversity: Overall, the data revealed a significant relationship between mating system and nucleotide diversity ($P = 0.018$, Mann-Whitney *U*-test): Estimates of nucleotide diversity were always greater in the cross-fertilizing species than in either of the self-fertilizing species (Figure 1 and Table 2). The approximate 95% confidence intervals for π_{si} in *C. remanei* did not overlap with those of the other two species in any instance except for the *COII* gene in *C. briggsae* (Figure 1).

The reduction in π_{si} associated with selfing was much greater for the two nuclear genes than for the mitochondrial gene (Figure 1). For the mitochondrial *COII* gene, π_{si} in the selfing species averaged 42% of that in *C. remanei*. In contrast, for the nuclear genes *glp-1*and *tra-2*, π_{si} in the selfing species averaged 6 and 13%, respectively, of that in *C. remanei*.

Neutrality tests: None of the neutrality tests indicated a significant deviation from the neutral model based on the frequency distributions of segregating sites (TAJIMA 1989; FU and LI 1993; FU 1997) or haplotypes (DEPAULIS and VEUILLE 1998). Estimates of Tajima's *D* statistic for each sample are provided in Table 2.

Patterns of haplotype variation: For each nuclear gene sampled in *C. remanei*, multiple recombination events could be inferred to have happened within the interval sequenced, on the basis of the observed DNA sequences (a minimum of three such events for *tra-2* and at least four for *glp-1*). In contrast, for these same genes, it was not possible to detect recombination events for the sequences sampled from *C. elegans* or *C. briggsae*, since the number of observed haplotypes was too low in each case (the minimum number of distinct haplotypes needed to infer recombination in the history of a sample is four).

Nuclear gene sequences from the two selfing species always clustered into a small number of distinct haplotypes, with little or no sequence variation apparent within each cluster (Figure 2). In *C. elegans*, diversity was extremely low for all three nuclear genes sampled: No polymorphism of any sort was detected for *tra-2*; only one polymorphism was detected for *glp-1*; and of six nucleotide polymorphisms detected for *spe-9* (the unrooted tree is not shown in Figure 2 because *spe-9* sequences were obtained only for *C. elegans*), five of these were singletons (in four cases, the minority nucleotide was present only in CB4856, and in one case it was present only in AB2). In *C. briggsae*, although the level of variation observed for nuclear genes was slightly higher, the strains still clustered into a small number

TABLE 2

Comparisons of nucleotide diversity between nematode species with different mating systems

Gene	Species	Mating system	No. strains	Length (bp)	Silent sites[a]	No. haplotypes	*P/S/R*[b]	Tajima's *D*[c]	$\pi_t \times 10^{-3}$	$\pi_{si} \times 10^{-3}$
tra-2	*elegans*	Self	20	670	297	1	0/0/0	—	0	0
	briggsae	Self	6	760	500	2	9/0/4	1.3	6.3	5.3
	remanei	Cross	12	418	184	12	12/2/2	0.4	11.2	20.6
glp-1	*elegans*	Self	20	547	353	2	1/1/0	1.3	0.9	1.4
	briggsae	Self	6	670	222	3	3/1/1	0.9	2.3	4.2
	remanei	Cross	12	426	137	10	20/12/3	0.4	18.8	45.7
COII	*elegans*	Self	11	696	151	4	15/14/1	1.7	10.2	43.2
	briggsae	Self	6	686	149	5	16/16/0	−0.4	9.6	44.3
	remanei	Cross	12	686	147	8	38/37/1	1.0	22.8	104.8
spe-9	*elegans*	Self	16	3385	1869	4	6/2/2	−1.3	0.3	0.5

[a] Total number of synonymous coding sites plus intron sites compared between sequences, on average.
[b] Total number of polymorphic (*P*) sites, as well as the number of polymorphisms that were synonymous (*S*) and replacement (*R*) changes in coding sequences. The number of polymorphic intron sites = $P - (S + R)$.
[c] No significant departures from the frequency distribution expected under neutrality were observed.

of distinct haplotypes (Figure 2). For the *tra-2* gene in *C. briggsae*, two haplotypes were observed (Figure 2); although there was no sequence variation segregating within each type, the divergence between haplotypes was substantial and included four amino acid changes.

In contrast, the sequences that were sampled from the cross-fertilizing species *C. remanei* did not cluster into low-diversity groups of haplotypes for either of the nuclear genes sampled (Figure 2). Within-population sequence diversity appeared to be high in *C. remanei*, as divergent sequences were obtained from three strains of *C. remanei* that were originally isolated at the same time from a single site in Gloucester, Massachusetts (these strains are marked with an asterisk in Figure 2).

The mitochondrial gene *COII* did not show marked differences between species in patterns of haplotype variation (Figure 2). Considerable haplotype diversity was observed for the mitochondrial genome in each species: In *C. elegans*, four haplotypes were observed among 11 samples; in *C. briggsae*, five haplotypes were observed among 6 samples; and in *C. remanei*, eight haplotypes were observed among 12 samples.

In *C. elegans*, local recombination rates for the three nuclear genes *tra-2*, *glp-1*, and *spe-9* were estimated to be 0.7, 2.1, and 4.4 cM/Mb, respectively. Estimates of π_{si} in *C. elegans* for these same genes were 0, 0.0014, and 0.0005, respectively (Table 2).

DISCUSSION

Many previous comparisons have found that self-fertilizing species tend to harbor less genetic variation than cross-fertilizing species (Hamrick and Godt 1990, 1996; Schoen and Brown 1991; Jarne and Staedler 1995; Awadalla and Ritland 1997; Charlesworth and Yang 1998; Liu *et al.* 1998, 1999; Baudry *et al.* 2001; but see Awadalla and Ritland 1997; Bergelson *et al.* 1998; Savolainen *et al.* 2000). We have observed a similar pattern in comparisons among closely related species in the genus Caenorhabditis (Figure 1 and Table 2).

The observed reductions in silent site nucleotide diversity were not predicted by the neutral model for completely selfing species (Figure 1). For nuclear genes, the effective population size is expected to be reduced by 50%, due to increased homozygosity (Pollak 1987); the observed reductions in nucleotide diversity at *tra-2* and *glp-1* were much greater than the neutral prediction. For mitochondrial genes, the effective population size is expected to double, due to the increased number of breeding individuals that transmit mitochondria to the next generation (Birky *et al.* 1983); the reductions in nucleotide diversity at *COII* were in the opposite direction to the neutral prediction.

In the present study, N_{mt} in selfers was estimated to be only ≈40% of that in cross-fertilizers. Since the neutral model prediction is that N_{mt} in selfers will be approximately double that in cross-fertilizers, it appears that N_{mt} has been reduced by a factor of $\sim 2/0.4 = 5$ in these highly selfing species. If the factor(s) responsible for the observed reduction in N_{mt} tended to affect the nuclear genes similarly, then we would expect the diversity in nuclear genes to be reduced ~10-fold altogether (2-fold due to the mating system alone and 5-fold more due to whatever additional factors reduced N_{mt} in selfers).

Interestingly, the average reduction in π_{si} observed for the two nuclear genes in the selfers agreed well with this prediction. Together, the estimates of nucleotide diversity in the two nuclear genes in the two selfing species averaged 9.5% of those in *C. remanei*, which is extremely close to the 10-fold reduction in nuclear diversity predicted by the mtDNA results. This close

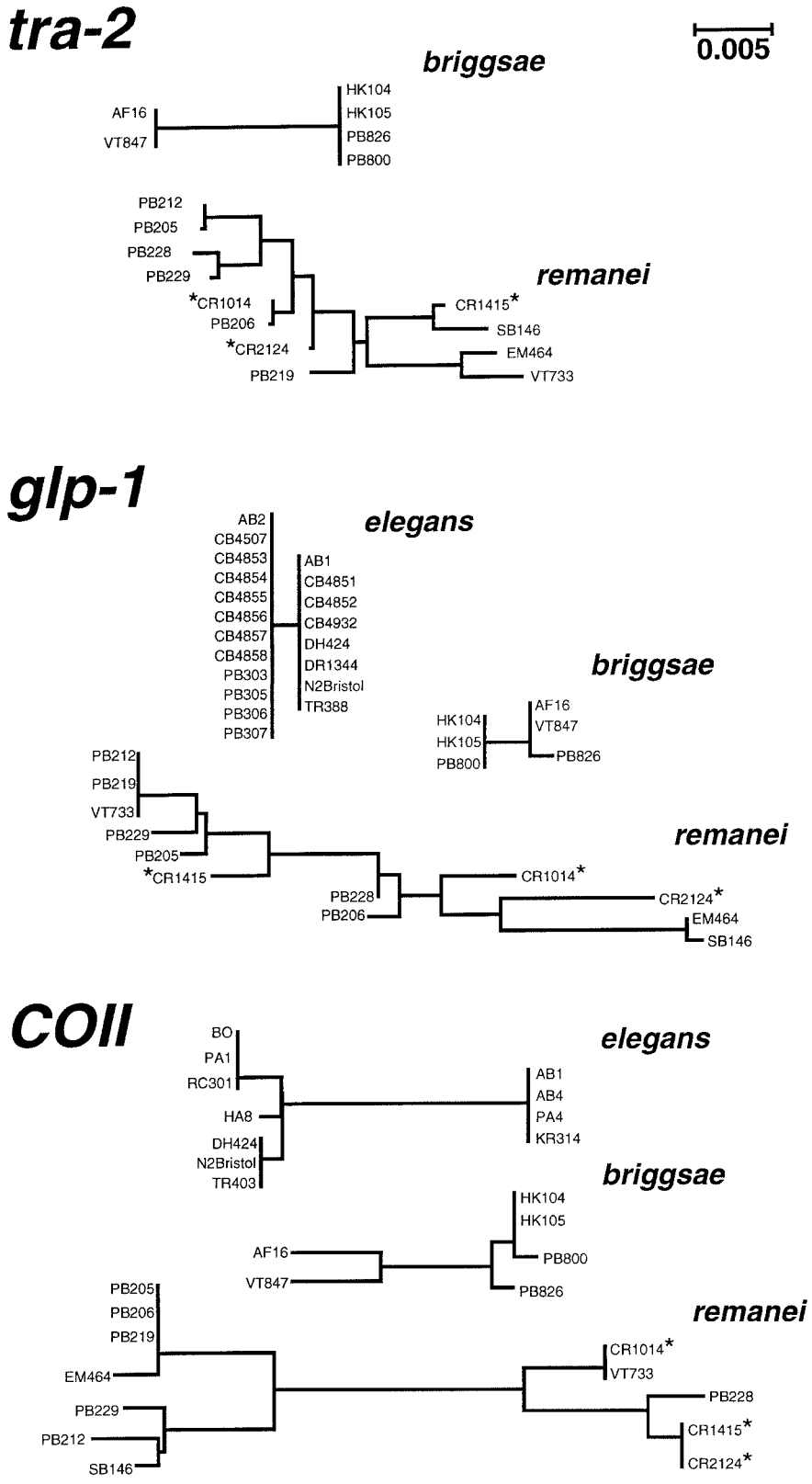

FIGURE 2.—Silent site divergence estimated between different strains within each species (*C. elegans*, *C. briggsae*, and *C. remanei*), for each of three genes (*tra-2*, *glp-1*, and *COII*), depicted in the form of unrooted trees. At the top right is the scale used for horizontal distances in all trees (the bar represents divergence between two sequences of 0.005, which is approximately one substitution per 200 silent sites). No tree is depicted for *tra-2* in *C. elegans* because no variation was observed in the sample of *tra-2* sequences for this species. The two nuclear gene sequences (*tra-2* and *glp-1*) from the two selfing species (*C. elegans* and *C. briggsae*) always clustered into a small number of distinct haplotypes, with little or no sequence variation apparent within each haplotype group. The same was not true for the nuclear gene sequences obtained from the cross-fertilizing species (*C. remanei*). The mitochondrial gene sequences (*COII*) also did not tend to cluster into a small number of distinct haplotypes in any species. Asterisks denote three *C. remanei* strains that were recovered at one site (within a 10-m radius) on the same day. The *C. elegans COII* sequences are from THOMAS and WILSON (1991).

agreement suggests that the same evolutionary mechanism could account for the observed patterns of variation at both nuclear and mitochondrial genes in these selfing species. As discussed above, the additional mechanisms that have been proposed include (1) repeated population bottlenecks, (2) the loss of balanced polymorphisms that would have been favored by overdominant selection (heterozygote advantage), and (3) enhanced effects of selection at linked sites (either genetic hitchhiking or background selection) due to reduced effective recombination rates.

Since mtDNA is cytoplasmic and effectively haploid,

N_{mt} should not be reduced by the loss of balanced polymorphisms in self-fertilizing species. The observed reductions in nucleotide diversity for the mtDNA suggest that the loss of balanced polymorphisms is not the primary factor that has reduced genetic variation in these two selfing species.

With the possible exception of *tra-2* in *C. elegans*, none of the patterns of haplotype diversity were consistent with a recent, species-wide, selective sweep at or near the genes studied (Table 2 and Figure 2). Indeed, although there were trends in this direction, none of the neutrality tests indicated a significant excess of rare alleles (Table 2). For this reason, if selective sweeps due to hitchhiking are responsible for the reduced variation observed in the selfing species, then either the selected alleles must be loosely linked to the loci studied here or alternative alleles must have been favored in different populations.

Most of the processes invoked in models to explain reduced diversity in selfing species are expected to reduce primarily the within-population diversity (Charlesworth *et al.* 1997). These effects may extend to the entire species, depending on the degree of population subdivision; high levels of population isolation may sometimes allow high genetic diversity to persist within the species, despite reductions in within-population diversity (reviewed in Charlesworth *et al.* 1997; Whitlock and Barton 1997). Unfortunately, we have little information on the structures of natural populations in these nematodes. Given this difficulty in defining a population, the species-wide diversity was sampled instead. This is conservative with regard to a comparison of levels of variation across mating systems; it is likely that the trends observed would be even more striking if well-defined, local populations could have been compared between these species. The observed species-wide differences suggest that, at least in these nematodes, the extent of population subdivision is not great enough to retain high levels of species-wide variation despite high levels of selfing.

Little is known about the breeding structure of natural populations of these species. For example, it is possible that the frequency of selfing in the hermaphroditic species is much lower in natural populations than it is in the laboratory. If males are indeed frequent in natural populations, then these hermaphroditic species are only partially selfing and are expected to experience smaller changes in population size (and hence levels of standing genetic variation) relative to the cross-fertilizing species. Hence, since our predictions were based on the assumption that these hermaphroditic species are close to 100% selfing, our test was conservative; to the extent that these species actually tend to cross-fertilize in nature, any effects of selfing on genetic variation should be reduced.

The estimates of nucleotide diversity reported here for *C. elegans* autosomal genes (Table 2) agreed with previous studies of this species (Thomas and Wilson 1991; Koch *et al.* 2000). For example, Thomas and Wilson (1991) sampled a total of 155 silent sites among 11 alleles of the *calmodulin* (*cal-1*) gene in *C. elegans* and observed no polymorphisms. Like the *tra-2* gene, the *cal-1* gene is located in an autosomal region with relatively low rates of recombination (local recombination rate for *cal-1* is estimated to be 1.0 cM/Mb). More recently, in a large-scale search for single nucleotide polymorphisms to be used for genetic mapping experiments, random genomic sequences from four wild isolates were compared to the published sequence from the standard laboratory strain, N2 (Koch *et al.* 2000). On the basis of a comparison of ~730 kb of sequence in this manner, a total of 313 single nucleotide mutations were identified. These results correspond to an average nucleotide diversity of $\sim\pi_t = 313/730{,}000 = 0.0004$. This estimate agrees with the average value, $\pi_t = 0.0006$, observed across three nuclear genes in our study (Table 2).

Koch *et al.* (2000) also found that polymorphism levels are much higher in the lateral regions (arms) of each autosome than in the central regions and interpret this as evidence that genes on the autosomal arms experience more rapid evolution. It is important to note, however, that estimates of recombination rates are also much higher in the lateral arms (Barnes *et al.* 1995). A positive correlation between regional rates of recombination and levels of variation has been observed in several other taxa and seems more likely to result from either genetic hitchhiking or background selection than to differences in evolutionary rates (for recent reviews see Aquadro 1997; Charlesworth and Charlesworth 1998). In our study, autosomal genes were surveyed from both low- and high-recombination regions. The *tra-2* gene is situated in a region with relatively low recombination rates in *C. elegans* (estimated to be 0.7 cM/Mb), and no DNA sequence polymorphism was observed for this gene. In contrast, low levels of polymorphism were observed for both *glp-1* and *spe-9*, and these genes are situated in regions with higher recombination rates (estimated to be 2.1 and 4.4 cM/Mb, respectively). These differences among loci were consistent with the pattern reported by Koch *et al.* (2000).

The relationship between mating system and levels of genetic diversity has been tested most thoroughly in plants, where it is often found that populations of self-fertilizers tend to display significantly less DNA variation than do populations of cross-fertilizers. The estimates of nucleotide diversity reported in this study suggest that this trend is also found in at least one group of nematodes. The reduction in variation was too great to be explained entirely by the incorporation of selfing into the neutral model of molecular evolution. The observation that variation was reduced in the mitochondrial genome suggests that the loss of balanced polymorphisms is not the primary factor reducing variation in these selfing nematodes. As expected, the reduction in genetic variation due to selfing was much greater for the nuclear than for the mitochondrial genome. Some

combination of population bottlenecks and the action of natural selection on nuclear genes probably explains the reduced genetic variation in these species. Application of this approach to other comparisons across mating systems should allow us to model the effect of mating system evolution on genetic variation with greater confidence.

We are grateful to Deborah Charlesworth, two anonymous reviewers, Patsy Dickinson, and the students of Biology 317 (Fall 2001) for making suggestions to improve the manuscript. Deborah Charlesworth, in particular, saved us from committing an error in our interpretation of these results. We thank the members of the Biology Department for many stimulating discussions on these and related topics. Also, we thank Scott Baird for generously sharing nematode stocks prior to publication and the Caenorhabditis Genetics Center for providing many nematode stocks. This work was supported by Bowdoin College and NSF grant 0110994 to M.F.P.

LITERATURE CITED

Aguadé, M., N. Miyashita and C. H. Langley, 1989 Reduced variation in the *yellow-achaete-scute* region in natural populations of *Drosophila melanogaster*. Genetics **122:** 607–615.

Aquadro, C. F., 1997 Insights into the evolutionary process from patterns of DNA sequence variability. Curr. Opin. Genet. Dev. **7:** 835–840.

Austin, J. A., and J. E. Kimble, 1989 Transcript analysis of *glp-1* and *lin-12*, homologous genes required for cell interactions during development of *C. elegans*. Cell **58:** 565–571.

Awadalla, P., and K. Ritland, 1997 Microsatellite variation and evolution in the *Mimulus guttatus* species complex with contrasting mating systems. Mol. Biol. Evol. **14:** 1023–1034.

Baker, H. G., 1955 Self-compatibility and establishment after 'long-distance' dispersal. Evolution **9:** 347–348.

Baker, H. G., 1967 Support for Baker's law—as a rule. Evolution **21:** 853–856.

Baldwin, J. G., R. M. Giblin-Davis, C. D. Eddleman, T. S. Williams, J. T. Vida *et al.*, 1997 The buccal capsule of *Aduncospiculum halicti* (Nemata:Diplogasterina): an ultrastructural and molecular phylogenetic study. Can. J. Zool. **75:** 407–423.

Barnes, T. M., Y. Kohara, A. Coulson and S. Hekimi, 1995 Meiotic recombination, noncoding DNA and genomic organization in *Caenorhabditis elegans*. Genetics **141:** 159–179.

Baudry, E., C. Kerdelhué, H. Innan and W. Stephan, 2001 Species and recombination effects on DNA variability in the tomato genus. Genetics **158:** 1725–1735.

Begun, D. J., and C. F. Aquadro, 1991 Molecular population genetics of the distal portion of the X chromosome in Drosophila: evidence for genetic hitchhiking of the *yellow-achaete* region. Genetics **129:** 1147–1158.

Begun, D. J., and C. F. Aquadro, 1992 Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. Nature **356:** 519–520.

Bergelson, J., E. Stahl, S. Dudek and M. Kreitman, 1998 Genetic variation within and among populations of *Arabidopsis thaliana*. Genetics **148:** 1311–1323.

Berry, A. J., J. W. Ajioka and M. Kreitman, 1991 Lack of polymorphism on the Drosophila fourth chromosome resulting from selection. Genetics **129:** 1111–1117.

Birky, Jr., C. W., 1978 Transmission genetics of mitochondria and chloroplasts. Annu. Rev. Genet. **12:** 471–512.

Birky, Jr., C. W., T. Maruyama and P. Fuerst, 1983 An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. Genetics **103:** 513–527.

Braverman, J. M., R. N. Hudson, N. L. Kaplan, C. H. Langley and W. Stephan, 1995 The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. Genetics **140:** 783–796.

C. elegans Sequencing Consortium, 1998 Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science **282:** 2012–2018.

Charlesworth, B., M. T. Morgan and D. Charlesworth, 1993 The effect of deleterious mutations on neutral molecular variation. Genetics **134:** 1289–1303.

Charlesworth, B., M. Nordborg and D. Charlesworth, 1997 The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. Genet. Res. **70:** 155–174.

Charlesworth, D., and B. Charlesworth, 1998 Sequence variation: looking for the effects of genetic linkage. Curr. Biol. **8:** R658–R661.

Charlesworth, D., and Z. Yang, 1998 Allozyme diversity in *Leavenworthia* populations with different inbreeding levels. Heredity **81:** 453–461.

Cox, P. A., 1989 Baker's law: plant breeding systems and island colonization, pp. 209–224 in *The Evolutionary Ecology of Plants*, edited by J. Bock and Y. Linhart. Westview Press, Boulder, CO.

Cummings, M. P., and M. T. Clegg, 1998 Nucleotide sequence diversity at the alcohol dehydrogenase 1 locus in wild barley (*Hordeum vulgare* ssp. *spontaneum*): an evaluation of the background selection hypothesis. Proc. Natl. Acad. Sci. USA **95:** 5637–5642.

Depaulis, F., and M. Veuille, 1998 Neutrality tests based on the distribution of haplotypes under an infinite-site model. Mol. Biol. Evol. **15:** 1788–1790.

Dvorák, J., M.-C. Luo and Z.-L. Yang, 1998 Restriction fragment length polymorphism and divergence in the genomic regions of high and low recombination in self-fertilizing and cross-fertilizing Aegilops species. Genetics **148:** 423–434.

Filatov, D. A., and D. Charlesworth, 1999 DNA polymorphism, haplotype structure and balancing selection in the Leavenworthia PgiC locus. Genetics **153:** 1423–1434.

Filatov, D. A., F. Monéger, I. Negrutiu and D. Charlesworth, 2000 Low variability in a Y-linked plant gene and its implications for Y-chromosome evolution. Nature **404:** 388–390.

Fitch, D. H. A., and W. K. Thomas, 1997 Evolution, pp. 815–850 in *C. elegans II*, edited by D. L. Riddle, T. Blumenthal, B. J. Meyer and J. R. Priess. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Fitch, D. H. A., B. Bugaj-Gaweda and S. W. Emmons, 1995 18S ribosomal RNA gene phylogeny for some Rhabditidae related to *Caenorhabditis*. Mol. Biol. Evol. **12:** 346–358.

Fu, Y. X., 1997 Statistical tests of neutrality against population growth, hitchhiking and background selection. Genetics **147:** 915–925.

Fu, Y. X., and W. H. Li, 1993 Statistical tests of neutrality of mutations. Genetics **133:** 693–709.

Haag, E. S., and J. Kimble, 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.

Hamrick, J. L., and M. J. Godt, 1990 Allozyme diversity in plant species, pp. 43–63 in *Plant Population Genetics, Breeding, and Genetic Resources*, edited by A. H. D. Brown, M. T. Clegg, A. L. Kahler and B. S. Weir. Sinauer Associates, Sunderland, MA.

Hamrick, J. L., and M. J. Godt, 1996 Effects of life history traits on genetic diversity in plant species. Philos. Trans. R. Soc. Lond. B Biol. Sci. **351:** 1291–1298.

Hedrick, P. W., 1980 Hitch-hiking: a comparison of linkage and partial selfing. Genetics **94:** 791–808.

Honda, H., 1925 Experimental and cytological studies on bisexual and hermaphrodite free living nematodes, with special reference to problems of sex. J. Morph. Physiol. **40:** 191–233.

Hudson, R. R., 1990 Gene genealogies and the coalescent process. Oxf. Surv. Evol. Biol. **7:** 1–45.

Hudson, R. R., and N. L. Kaplan, 1995 Deleterious background selection with recombination. Genetics **141:** 1605–1617.

Jarne, P., and T. Städler, 1995 Population genetic structure and mating system evolution in freshwater pulmonates. Experientia **51:** 482–497.

Kaplan, N. L., R. R. Hudson and C. H. Langley, 1989 The 'hitchhiking effect' revisited. Genetics **123:** 887–899.

Kimura, M., 1971 Theoretical foundations of population genetics at the molecular level. Theor. Popul. Biol. **2:** 174–208.

Kimura, M., 1983 *The Neutral Theory of Molecular Evolution.* Cambridge University Press, New York.

Kimura, M., and T. Ohta, 1971 *Theoretical Topics in Population Genetics.* Princeton University Press, Princeton, NJ.

Koch, R., H. G. A. M. van Luenen, M. van der Horst, K. L. Thijssen and R. H. A. Plasterk, 2000 Single nucleotide polymorphisms in wild isolates of *Caenorhabditis elegans*. Genome Res. **10:** 1690–1696.

Kraft, T., T. Säll, I. Magnusson-Rading, N.-O. Nilsson and C. Halldén, 1998 Positive correlation between recombination rates and levels of genetic variation in natural populations of sea beet (*Beta vulgaris* subsp. *maritima*). Genetics **150:** 1239–1244.

Kumar, S., K. Tamura, I. Jakobsen and M. Nei, 2000 *MEGA, Molecular Evolutionary Genetics Analysis*, version 2.1. Institute of Molecular Evolutionary Genetics, The Pennsylvania State University, University Park, PA.

Kuwabara, P. E., and J. E. Kimble, 1995 A predicted membrane protein, TRA-2A, directs hermaphrodite development in *Caenorhabditis elegans*. Development **121:** 2995–3004.

Langley, C. H., J. Macdonald, N. Miyashita and M. Aguadé, 1993 Lack of correlation between interspecific divergence and intraspecific polymorphism at the suppressor of forked region in *Drosophila melanogaster* and *Drosophila simulans*. Proc. Natl. Acad. Sci. USA **90:** 1800–1803.

Li, W.-H., 1997 *Molecular Evolution.* Sinauer Associates, Sunderland, MA.

Liu, F., L. Zhang and D. Charlesworth, 1998 Genetic diversity in *Leavenworthia* populations with different inbreeding levels. Proc. R. Soc. Lond. Ser. B Biol. Sci. **265:** 293–301.

Liu, F., D. Charlesworth and M. Kreitman, 1999 The effect of mating system differences on nucleotide diversity at the phosphoglucose isomerase locus in the plant genus Leavenworthia. Genetics **151:** 343–357.

Maynard Smith, J., and J. Haigh, 1974 The hitch-hiking effect of a favorable gene. Genet. Res. **23:** 23–35.

Nachman, M. W., 1997 Patterns of DNA variability at X-linked loci in *Mus domesticus*. Genetics **147:** 1303–1316.

Nachman, M. W., 2001 Single nucleotide polymorphisms and recombination rate in humans. Trends Genet. **17:** 481–485.

Nachman, M. W., V. L. Bauer, S. L. Crowell and C. F. Aquadro, 1998 DNA variability and recombination rates at X-linked loci in humans. Genetics **150:** 1133–1141.

Narrain, P., 1966 Effect of linkage on homozygosity of a population under mixed selfing and random mating. Genetics **54:** 303–314.

Nei, M., 1987 *Molecular Evolutionary Genetics.* Columbia University Press, New York.

Nei, M., T. Maruyama and R. Chakraborty, 1975 The bottleneck effect and genetic variability in populations. Evolution **29:** 1–10.

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

Nordborg, M., B. Charlesworth and D. Charlesworth, 1996 The effect of recombination on background selection. Genet. Res. **67:** 159–174.

Palopoli, M. F., and C.-I Wu, 1996 Rapid evolution of a coadapted gene complex: evidence from the *Segregation Distorter* (*SD*) system of meiotic drive in *Drosophila melanogaster*. Genetics **143:** 1675–1688.

Pollak, E., 1987 On the theory of partially inbreeding finite populations. I. Partial selfing. Genetics **117:** 353–360.

Rozas, J., and R. Rozas, 1999 DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics **15:** 174–175.

Rudel, D., and J. Kimble, 2001 Conservation of *glp-1* regulation and function in nematodes. Genetics **157:** 639–654.

Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi *et al.*, 1988 Primer-directed enzymatic amplification of DNA with a thermostable polymerase. Science **239:** 487–491.

Saitou, N., and M. Nei, 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:** 406–425.

Savolainen, O., C. H. Langley, B. P. Lazzaro and H. Freville, 2000 Contrasting patterns of nucleotide polymorphism at the alcohol dehydrogenase locus in the outcrossing *Arabidopsis lyrata* and the selfing *Arabidopsis thaliana*. Mol. Biol. Evol. **17:** 645–655.

Schoen, D. J., and A. H. D. Brown, 1991 Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. Proc. Natl. Acad. Sci. USA **88:** 4494–4497.

Singson, A. K., K. B. Mercer and S. W. L'Hernault, 1998 The *C. elegans spe-9* gene encodes a sperm transmembrane protein that contains EGF-like repeats and is required for fertilization. Cell **93:** 71–79.

Singson, A. K., K. L. Hill and S. W. L'Hernault, 1999 Sperm competition in the absence of fertilization in *Caenorhabditis elegans*. Genetics **152:** 201–208.

Stein, L., P. Sternberg, R. Durbin, J. Thierry-Mieg and J. Spieth, 2001 WormBase: network access to the genome and biology of *Caenorhabditis elegans*. Nucleic Acids Res. **29:** 82–86.

Stephan, W., and C. H. Langley, 1989 Molecular genetic variation in the centromeric region of the X chromosome in three *Drosophila ananassae* populations. I. Contrasts between the *vermilion* and *forked* loci. Genetics **121:** 89–99.

Stephan, W., and C. H. Langley, 1998 DNA polymorphism in *Lycopersicon* and crossing-over per physical length. Genetics **150:** 1585–1593.

Stephan, W., and S. J. Mitchell, 1992 Reduced levels of DNA polymorphism and fixed between-population differences in the centromeric region of *Drosophila ananassae*. Genetics **132:** 1039–1045.

Sudhaus, W., and K. Kiontke, 1996 Phylogeny of *Rhabditis* subgenus *Caenorhabditis* (Rhabditidae, Nematoda). J. Zool. Syst. Evol. Res. **34:** 217–233.

Tajima, F., 1989 Statistical method for testing the neutral mutation hypothesis. Genetics **123:** 585–595.

Thomas, W. K., and A. C. Wilson, 1991 Mode and tempo of molecular evolution in the nematode Caenorhabditis: cytochrome oxidase II and calmodulin sequences. Genetics **128:** 269–279.

Thompson, J. D., D. G. Higgins and T. J. Gibson, 1994 ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:** 4673–4680.

Van Treuren, R., H. Kuittinen, K. Kärkkäinen, E. Baena-Gonzalez and O. Savolainen, 1997 Evolution of microsatellites in *Arabis petraea* and *Arabis lyrata*, outcrossing relatives of *Arabidopsis thaliana*. Mol. Biol. Evol. **14:** 220–229.

Wood, W. B., 1988 Introduction to *C. elegans* biology, pp. 1–16 in *The Nematode Caenorhabditis elegans*, edited by W. B. Wood and the Community of *C. elegans* Researchers. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Communicating editor: D. Charlesworth