

# Levels of DNA Polymorphism in Various Mating Systems in the Nematode Genus *Caenorhabditis*

Andrew G. Tate, John M. Galloway, James R. Watson and Michael F. Li<sup>1</sup>

*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

*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







TABLE 2

C m a i n f n c t i d e d i e t b e e n n e m a t e d e e c i e i n d i f f e r e n t m a t i n g t e m

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

<sup>a</sup> 6sbp) us coding seq-333.ng number of intron site26s

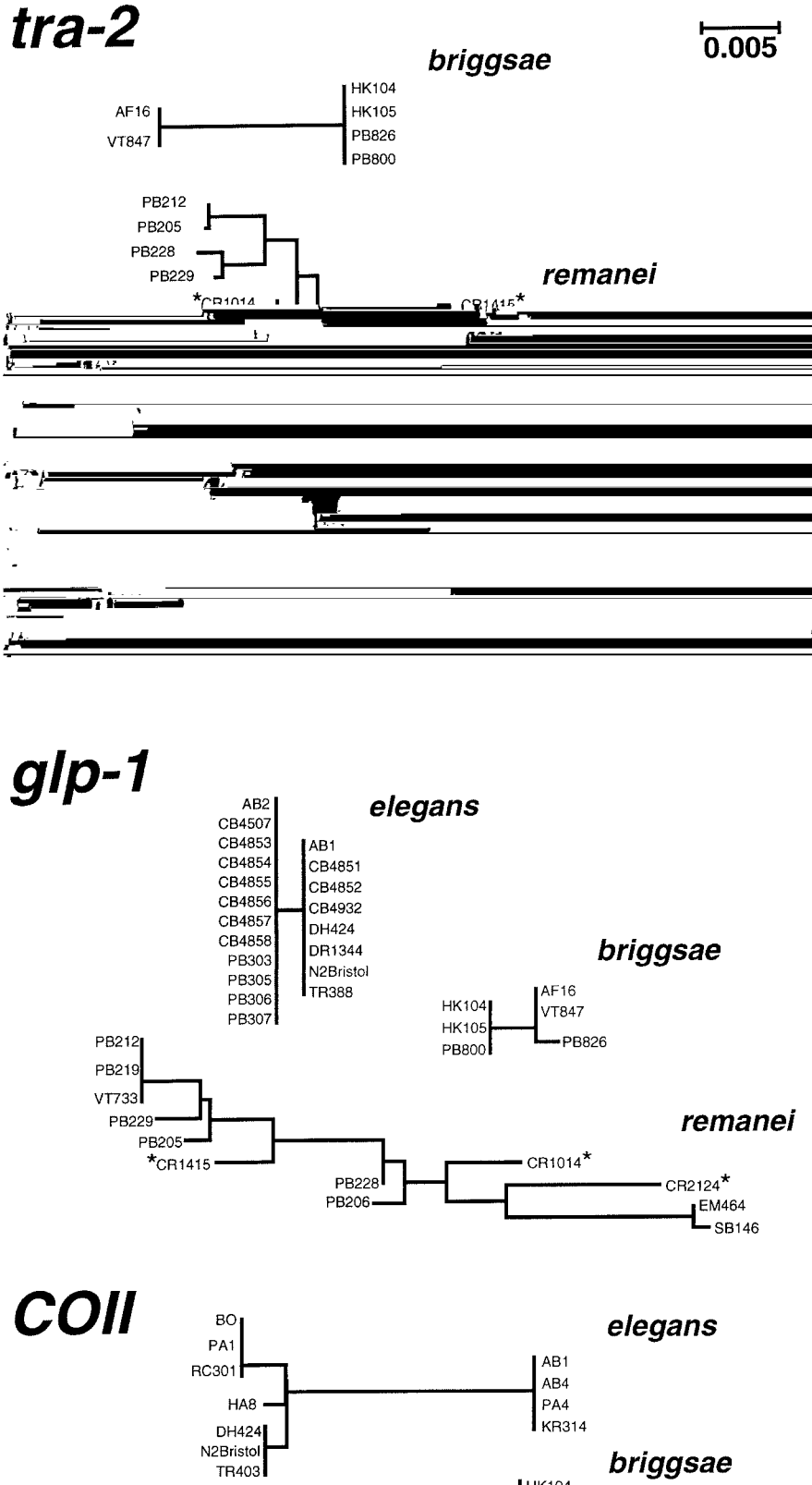


Figure 2.—Silent site divergence estimated between different strains within each species (*C. elegans*, *C. briggsae*, and *C. remanei*), for each of three genes (*tra-2*, *glp-1*, and *COII*), depicted in the form of unrooted trees. At the top right is the scale used for horizontal distances in all 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 poly-

morphisms 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 deg4300(deg4300(dem4j/F5 9.948 0 0a1e)-300(((manner,giuv).)Tjg1 -1.e5iyi1-wner,giu.e.o(

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 expe-

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* 155: 105–116.
- 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. Negruțiu 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.



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

Kimura, M., and T. Ohta

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

