Molecular Ecology (2007) **16**, 1267–1280 doi: 10.1111/j.1365-294X.2006.03222.x

Distinct patterns of genetic variation in *Pristionchus pacificus* and *Caenorhabditis elegans*, two partially selfing nematodes with cosmopolitan distribution

HANS ZAUNER, WERNER E. MAYER, MATTHIAS HERRMANN, ANDREAS WELLER, MIRJA ERWIG and RALF J. SOMMER
Max Planck Institute for Developmental Biology, Department of Evolutionary Biology, 72076 Tübingen, Germany

Abstract

Hermaphroditism has evolved several times independently in nematodes. The model organism *Caenorhabditis elegans* and *Pristionchus pacificus* are self-fertile hermaphrodites with rare facultative males. Both species are members of different families: *C. elegans* belongs to the Rhabditidae and *P. pacificus* to the Diplogastridae. Also, both species differ in their ecology: *C. elegans* is a soil-dwelling nematode that is often found in compost heaps. In contrast, field studies in Europe and North America indicate that *Pristionchus* nematodes are closely associated with scarab beetles. In *C. elegans*, several recent studies have found low genetic diversity and rare out-crossing events. Little is known about diversity levels and population structure in free-living hermaphroditic nematodes outside the genus *Caenorhabditis*. Taking a comparative approach, we analyse patterns of molecular diversity and linkage disequilibrium in 18 strains of *P. pacificus* from eight countries and four continents. Mitochondrial sequence data of *P. pacificus* isolates reveal a substantially higher genetic diversity on a global scale when compared to *C. elegans*. A mitochondrial-derived hermaphrodite phylogeny shows little geographic structuring, indicating several worldwide dispersal events. Amplified fragment length polymorphism and single strand conformation polymorphism analyses demonstrate a high degree of genome-wide linkage disequilibrium, which also extends to the mitochondrial genome. Together, these findings indicate distinct patterns of genetic variation of the two species. The low level of genetic diversity observed in *C. elegans* might reflect a recent human-associated dispersal, whereas the *P. pacificus* diversity might reflect a long-lasting and ongoing insect association. Thus, despite similar lifestyle characteristics in the laboratory, the reproductive mode of hermaphroditism with rare facultative males can result in distinct genetic variability patterns in different ecological settings.

Keywords: *Pristionchus pacificus*, population genetics, hermaphroditism, mtDNA

Received 24 September 2006; revision received 22 October 2006; accepted 30 October 2006

Introduction

The nematode *Pristionchus pacificus* is cultivated as a laboratory organism to study the evolution of development, behaviour and ecology in comparison to the model organism *Caenorhabditis elegans* (Hong & Sommer 2006). This comparative approach is leading towards insights into alterations of developmental mechanisms within the nematode bauplan. For example, *P. pacificus* differs from *C. elegans* in the development of the vulva and the gonad (Rudel *et al*. 2005; Schlager *et al*. 2006) and the buccal cavity (the mouth) of diplogastrid nematodes, to which *P. pacificus* belongs, contains movable teeth and exposed fangs, structures that are not present in eurhabditid nematodes (Lieven Fürst von & Sudhaus 2000; Hong & Sommer 2006). Such morphological differences demand explanations equally from the perspectives of development, ecology

Correspondence: Ralf J. Sommer, Max Planck Institut für Entwicklungsbiologie, Abteilung für Evolutionsbiologie, Spemannstr. 37–39, 72076 Tübingen, Germany. Fax: +49-7071-601498; E-mail: ralf.sommer@tuebingen.mpg.de

and evolution. A variety of genetic and genomic tools are now available to integrate such research in *P. pacificus*: a genetic linkage map with approximately 550 single-nucleotide polymorphism (SNP) markers and bacterial artificial chromosome (BAC) libraries are available for positional cloning and for quantitative trait loci (QTL) studies (Srinivasan *et al*. 2002 and www.pristionchus.org). Whole genome sequencing is approaching completion and will provide a basis for a detailed comparison with the fully sequenced genome of *C. elegans* (www.nhgri.nih.gov/12511858).

Morphological differences notwithstanding, the lifestyles of *P. pacificus* and *C. elegans* resemble each other, at least under laboratory conditions: both nematodes can be kept on *Escherichia coli* bacteria and have a generation time of about 3.5 days at 20 °C. Before reaching adulthood, they pass four juvenile stages. Under starvation, larvae can enter 'dauer' stage, an alternative, nonfeeding larval stage that can be maintained for several weeks. Importantly, *P. pacificus* and *C. elegans* also share the same mode of reproduction: both nematodes are self-fertilizing hermaphrodites with a rare occurrence of spontaneous males by X-chromosomal nondisjunction. In both species, sex determination is regulated by the number of X-chromosomes; XX — animals develop as hermaphrodites and X0 — animals are males (Hodgkin & Brenner 1977; Pires-DaSilva & Sommer 2004). Spontaneous males are fertile, with cross-fertilized offspring being 50% male and 50% hermaphroditic. However, males never reach high frequencies in standard laboratory cultures because self-fertilizing hermaphrodites outcompete cross-fertilizing males (Chasnov & Chow 2002; Stewart & Phillips 2002; Cutter *et al*. 2003).

Despite these commonalities in the laboratory, however, it cannot be inferred how lifestyle parameters like outcrossing rates or generation times are fine-tuned in nature. Indeed, it is essential to study natural populations of the chosen species in their ecological context to understand the evolutionary processes that did shape the biology of these organisms. Field studies in Western Europe and the Eastern United States of America revealed more than 12 *Pristionchus* species that are associated with scarab beetles and the Colorado potato beetle (Herrmann *et al*. 2006a, 2006b) However, *P. pacificus* was not frequently observed during these studies. Therefore, a potential host of *P. pacificus* is so far unknown and most isolates stem from soil samples. Currently, field studies are ongoing in additional continents to identify the natural host of *P. pacificus* (Herrmann, Weller and Sommer, ongoing studies).

Similarly, the natural habitat of the model organism *C. elegans* is still debated. Most *C. elegans* isolates stem from anthropogenic sources, namely compost heaps, where the species is sometimes found in association with various invertebrates such as millipedes, isopods, insects, snails, and slugs (Barrière & Félix 2005a; Kiontke & Sudhaus 2005). The quest for the nematodes' habitats immediately raises issues related to their lifestyle. How frequent are males in nature? What are the vectors they use for dispersal over longer distances? Although in nematodes it is notoriously difficult to infer these parameters by direct observation, different lifestyles and demographies leave different molecular signatures.

In *C. elegans*, several studies addressed the level of local and global molecular variation and linkage disequilibrium (LD) to deduce out-crossing and migration rates as well as effective population sizes (reviewed in Barrière & Félix 2005b). These studies revealed that global molecular diversity is low and out-crossing is a rare event in the wild (Koch *et al*. 2000; Graustein *et al*. 2002; Denver *et al*. 2003; Sivasundar & Hey 2003, 2005; Barrière & Félix 2005a; Haber *et al*. 2005; Cutter 2006). Interestingly, a high fraction of the worldwide variation is present within individual localities (Barrière & Félix 2005a). A possible interpretation of these data sees *C. elegans* as a colonizer with high migration rates and local demographic bottlenecks (Barrière & Félix 2005a).

Pristionchus pacificus has the same lifestyle characteristics as *C. elegans* under laboratory conditions and also has a worldwide distribution. In this study, we show that molecular diversity is higher in *P. pacificus* than in *C. elegans*. *Pristionchus pacificus* displays high levels of genome-wide LD, indicating low out-crossing, but also substantially higher molecular diversity as revealed by mitochondrial DNA (mtDNA) sequence data. Several, not mutually exclusive scenarios might explain this difference. We discuss the potential influence of differences in mutation rates and population substructure in the light of the ecology of the two species.

Materials and methods

Strains

Strains used in this study can be classified according to collection history: the majority of our strains (SB5880, PS1843, JU482, JU723, RS106, PS312, JU138, RS5171, JU150, RS5200, RS5160) are isolated from soil samples taken randomly over the last 10 years by various collectors (see Table 1). In contrast, strains from Northern America (RS5131,RS5134, RS5138) and South Africa (RS5202, RS5203, RS5204, RS5205) are the result of field trips that were specifically planned to search the natural habitat of *Pristionchus* and comprised sampling of soil as well as various beetles (Herrmann *et al*. 2006b). Isolation procedures and taxonomic determination of nematode species from these field trips are described in detail in Herrmann *et al*. (2006a). In brief, soil samples are spread on 10 cm NGM (nematode growth medium) agar plates with *Escherichia coli* OP50 and inspected regularly during the

Table 1 *Pristionchus pacificus* strains

Strain	Location	Origin	Collector
RS5134*	Wooster (Ohio), USA	*Phyllophaga* sp. (Coleoptera: Scarabaeidae)	Herrmann
SB5880*	New York, USA	Soil	Sudhaus
RS5138	Indian Cave state park (Nebraska), USA	*Cyclocephala* sp. (Coleoptera: Scarabaeidae)	Herrmann
PS1843*	PortAngeles (Washington), USA	Soil	Carta
RS5160/RS5153*	Enoshima, Japan	Soil	Sudhaus
JU482*	Hakone, Japan	Soil	Félix
RS5131*	Carver (Massachusetts), USA	*Lichnanthe vulpina* (Coleoptera: Scarabaeidae)	Herrmann
RS5200	Kalimpong, India	Compost heap	Sahm
JU138*	Captain Cook (Hawaii), USA	Soil	Félix
PS312*	Pasadena (California), USA	Soil	Tang/Lamb
RS106	Augustow, Poland	Soil	Schierenberg
JU723*	Longsheng (Guangxi), PR China	Soil	Félix
RS5171*	Tivat, Montenegro	grape marc	Grbic
JU150*	Antananarivo, Madagascar	Soil	Félix
RS5202	Pretoria, South Africa	*Pseudoderopeltis* sp. (Blattodea: Blattidae)	Weller
RS5203	Pretoria, South Africa	*Pseudoderopeltis* sp. (Blattodea: Blattidae)	Weller
RS5204	Pretoria, South Africa	Burrower bug (Heteroptera: Cydnidae)	Weller
RS5205	Pretoria, South Africa	*Onitis* sp. (Coleoptera: Scarabaeidae)	Weller

Legend: Sampling sites, origin and collectors of 18 worldwide strains of *Pristionchus pacificus*. Note that most isolates of *P. pacificus* originate from soil samples, in particular the strains from North America. Only three *P. pacificus* isolates were obtained from beetle samples in Ohio, Nebraska and Massachusetts, respectively. Four isolates have been obtained from South Africa and three of these four are derived from nonbeetle insects. A beetle host for *P. pacificus* has yet to be identified. However, given the close association of other *Pristionchus* nematodes with scarab beetles in Western Europe and North America, such an association is likely. Asterisks indicate strains that were chosen for AFLP and SSCP analyses. RS5160 and RS5153 are substrains of the same original isolate.

following one to three weeks with a dissecting scope. Insects are cut in half before putting them on NGM plates. A dissecting scope and a Zeiss Axioplan 2 microscope are used to determine nematodes to the family and to the genus level, respectively. Morphological methods are not sufficient to unambiguously distinguish *Pristionchus* species. Thus, species identity was assessed on the basis of identical SSU sequences as described in Herrmann *et al*. (2006a, b) and confirmed by mating tests with the laboratory strain PS312. Due to long inbreeding in the laboratory, strains are expected to be homozygous at nearly all loci. Strains originating from the field trips to North America and South Africa are iso-female lines, i.e. they each go back to one individual hermaphrodite. Thus, each strain contains only one mitochondrial haplotype. We note that the available strains provide a basis to investigate molecular divergence patterns on a global scale, but do not allow inferring the genetic structure of local populations.

Mitochondrial sequences

Blasting the *Caenorhabditis elegans* mitochondrial sequence (accession no. X54252) to an early version of the *Pristionchus pacificus* genome (preliminary assembly 'Ppa-latestChris2' on www.pristionchus.org) we obtained a contig (ctg482.5) containing, among others, the *P. pacificus* mitochondrial genes ND6, ND4L, ND1, ND2 and cytochrome *b*.

Based on this sequence information we designed the following polymerase chain reaction (PCR) primers:

ND6 and ND4L: HZ11896 TCCAGGCAAAATTTATTATT-TACAA and HZ11897 CCATTTTAAACCATCTTAAACCA,
ND1: HZ11898 CATGAAGCGAGTACAACTCGTAG and HZ11899 AACAGATCTCAAAGGTAGTCTAGCA,
ND2: HZ11900 CGCAAAAGATATACGCCAAT and HZ11998 AACATGATATATTTCACCAGAAAAGTT,
cytochrome B: HZ11902 TCAGCATTAACATTTAGATATTG-GTT and HZ11999 TCTGGAGCAAAACAAACTTTATCA

For the strain JU150 the cytochrome *b* primer set did not work. New primers were designed to amplify the gene from this strain:

HZ12197 CTCAAATAAGGTTTTGAGCTTCAG and HZ12198 ATGTTGTTAAACCCAAAGAACCA.

DNA was prepared with one of the following methods: Either worms from one to three overgrown 6-cm plates were washed three times in M9 medium. DNA was isolated with the AquaPure Genomic DNA isolation kit (Bio-Rad) and the DNA concentration adjusted to 20 ng/μL. Alternatively, around 10 worms were put in 10 μL worm lysis buffer (50 mM KCl, 10 mM Tris-HCl, pH 8.3, 2.5 mM $MgCl_2$, 0.45% NP40, 0.45% Tween and 0.05 mg/mL Proteinase K), frozen for at least 15 min at –80 °C and then lysed for 1 h at 65 °C, followed by de-activating the

enzyme at 95 °C for 10 min. The lysate was diluted with 20–40 μL ddH_2O; 4 μL of diluted lysate was used in a 20-μL PCR. PCR was performed in 20 μL 1× PCR buffer (Amersham Biosciences, Freiburg) containing 1 U *Taq* DNA Polymerase (Amersham), 0.5 μM of each primer, 0.2 mM of each deoxynucleotide triphosphate and 4 μL DNA preparation. PCRs were performed as follows: initial denaturation at 94 °C for 2 min, followed by 35 cyles of denaturation at 94 °C for 30 s, primer annealing at 52 °C (or 54 °C for some primer pairs) for 30 s and extension at 72 °C for 70–90 s, depending on the product length.

A final incubation step at 72 °C was performed for 7 min. PCR products were diluted 1:20 with ddH20 and sequenced without further purification from both ends (Big Dye terminator protocol, Applied Biosystems)

Sequences have been submitted to the National Center for Biotechnology Information (NCBI) under accession nos DQ885129–DQ885191.

Single strand conformation polymorphism analysis

Single strand conformation polymorphism (SSCP) markers were taken from the *P. pacificus* genetic linkage map (Srinivasan *et al*. 2002 and www.pristionchus.org). Initially, five to six markers per chromosome were selected in such a way that, for each chromosome, marker pairs with distances varying from 0 to 5 cM up to around 30–40 cM are included. PCR and fragment detection were performed as described previously (Srinivasan *et al*. 2002). Band classification was carried out by eye, only considering clearly distinct genotypes. At each locus, alleles are labelled with letters in alphabetical order according to their frequency. Other than that, naming of alleles at different loci is independent, that is, alleles with the same name at different loci only share the property that A is the most common, etc. Out of initially 35 markers, three (S151, S170 and S158) were excluded from the analysis because they did not amplify a reliable product in several of the tested strains. In the remaining data set, four out of 352 data points (1.1%) are missing, either due to failure of the PCR or due to polymorphisms in the affected strains that inhibit amplification of a product.

Amplified fragment length polymorphism analysis

Amplified fragment length polymorphism (AFLP) (Vos *et al*. 1995) was performed according to manufacturer's instructions (AFLP Small Genome Primer Kit Cat. # 10719-011, Life Technologies) as described previously (Srinivasan *et al*. 2001). We used two primer combinations that proved to be suitable in previous experiments: E-AT/M-CTC and E-TG/M-CTA. Only strong bands were counted and scored as present (1) or absent (0).

Data analysis

Sequence data were assembled with SEQMAN and MEGALIGN software (DNAStar Inc.). Primer sequences and unreliable sequences at the ends were trimmed by eye. For phylogenetic analysis we used the PAUP*4.0b10 program (Swofford 2002). For presentation, trees were rooted at midpoint. MODELTEST 3.7 software (Posada & Crandall 1998) was used to select a substitution model for maximum-likelihood (ML) analysis. For maximum parsimony (MP) and ML trees, we performed 500 bootstrap iterations each.

Amino acid sequences of parts of the mitochondrial genes ND6, ND4L, ND2 and cytochrome *b* from *P. pacificus* PS312 and *C. elegans* N2 were aligned with MEGALIGN to identify homologous sequence positions between the two species. There are no insertion/deletion mutations in these sequences except for ND2, which contains a 1-amino-acid indel. Next, the corresponding DNA sequence of PS312 was aligned to the same 14 *P. pacificus* strains that were also included in the phylogenetic analysis. All of these 15 *P. pacificus* isolates stem from different sampling sites. The homologous *C. elegans* N2 sequence was aligned to 14 *C. elegans* sequences from Denver *et al*. (2003) that were obtained via NCBI (accession nos AY171133.1–AY171147.1 and AY171163.1–AY171192.1). Some of the sequences within this set represent several strains that share identical haplotypes. Thus, the 15 NCBI sequences represent 27 *C. elegans* strains from 18 localities (see Denver *et al*. 2003 and Table S1, Supplementary material for details). To obtain a data set that is better comparable to *P. pacificus* we randomly chose one *C. elegans* isolate per sampling site in the four cases where several isolates stem from the same locality (Table S1). Given the close proximity of the sampling sites 'Pasadena' and 'Altadena', especially compared to the distances between the *P. pacificus* isolates, we included only one strain from these two locations. Sequences from the following 17 *C. elegans* strains were used to obtain the data in Tables 2 and 3: AB3, DH424, CB4507, CB4851, CB4852, CB4855, CB4856, CB4857,CB4858, CB4932, KR314, LSJ1, N2,PB303, PB306, TR389, and RC301. Divergence parameters, GC content and values for neutrality tests were obtained for the homologous *C. elegans* and *P. pacificus* data sets with the program DNASP version 4.18 (Rozas *et al*. 2003). Nucleotide diversity π is defined as the average number of nucleotide differences per site between two sequences (Hartl & Clark 1997). The population parameter Θ is calculated from the number of segregating sites, S, and the number of samples, n ($\theta = S/[\sum_{i=1}^{n-1}(1/i)]$ under the infinite-sites model; Watterson 1975). We report values of Θ and π on a per-site basis for the entire data set and also separately for synonymous and nonsynonymous sites.

To determine if the choice of strains influenced the *C. elegans* analysis in those cases where several isolates per sampling site are available, we compared 10 sequence

Table 2 Sequence characteristics of the five mitochondrial genes ND6, ND4L, ND1, ND2 and cyt *b*

	Length	Haplotypes	GC(c)	GC(3)	S	η
Pristionchus pacificus (*n* = 15)						
ND6	381	12	0.19	0.14	43	46
ND4L	231	12	0.19	0.13	17	19
ND1	572	12	0.26	0.17	56	57
ND2	705	12	0.18	0.14	81	83
cyt *b*	480	12	0.27	0.13	46	47
concatenated	2369	15	0.22	0.15	243	252
Caenorhabditis elegans (*n* = 17)						
ND6	381	3	0.2	0.12	7	7
ND4L	231	4	0.19	0.12	4	4
ND1	572	5	0.27	0.15	12	12
ND2	708	5	0.21	0.15	11	11
cyt *b*	480	4	0.26	0.17	12	12
concatenated	2372	7	0.23	0.15	46	46

Legend: *n*, number of strains in the alignment; length, Sequence length in base pairs; GC(c), GC content at all coding positions; GC(3), GC content at 3rd codon positions; S, number of segregating sites; η, total number of mutations. Higher values for η than S in *Pristionchus pacificus* indicate the presence of sites with more than two variants.

sets, each of them with one randomly chosen strain per sampling site (Table S1). These randomizations do not alter our conclusions in terms of the comparison of molecular diversity between *P. pacificus* and *C. elegans*. In particular, variations of Θ and π due to strain choice are only a minute fraction of the difference in these parameters between *C. elegans* and *P. pacificus* (Table S2, Supplementary material). Similarly, Tajima's *D* is positive in each of the 10 analyses. However, we note that the numerical value (and, in case of ND6, the statistical significance) of Tajima's *D* varies depending on strain composition.

LD between and among mtDNA, SSCP and AFLP markers was analysed with ARLEQUIN version 2.0. Since we are dealing with highly inbred, selfing strains, markers are treated as haplotypic data with known gametic phase. Statistical significance is assessed under ARLEQUIN, which uses an extension of the Fisher exact probability test on contingency tables (Slatkin 1994; ARLEQUIN manual section 7.1.4.1 of Schneider *et al*. 2000).

Results

Pristionchus pacificus *is a cosmopolitan species*

The first strains of *Pristionchus pacificus* were isolated at the West coast of the USA from Pasadena, California (PS312) and Port Angeles, Washington (PS1843) (Sommer *et al*. 1996). In the last 10 years, additional strains were obtained from eight countries in Africa, Asia, Europe and North America (Table 1; see also Materials and methods). Thus, *P. pacificus* has a cosmopolitan distribution and is so far the only species of the genus *Pristionchus* for which isolates have been obtained from several continents. However, most of the strains are unique and represent erratic findings within a sampled locality.

Most isolates originate from soil samples, but occasionally *P. pacificus* has been found in association with insects, in particular beetles (Table 1). Specifically, three *P. pacificus* isolates were obtained from beetle samples in Ohio,

Table 3 Divergence estimates of mitochondrial genes in *Pristionchus pacificus* and *Caenorhabditis elegans*

P. pacificus (*n* = 15)								
	π	π(s)	π(a)	Θ	Θ(s)	Θ(a)	Tajima's *D*	Fu & Li's *D**
ND6	0.03	0.11	0.0065	0.035	0.13	0.01	−0.83	−0.64
ND4L	0.023	0.094	0.0029	0.023	0.11	0.0017	−0.35	−0.18
ND1	0.024	0.097	0.0032	0.03	0.13	0.0035	−0.94	−1.1
ND2	0.028	0.11	0.0078	0.035	0.14	0.0093	−1	−0.97
cyt *b*	0.027	0.096	0.0049	0.029	0.1	0.0075	−0.51	−0.48
concatenated	0.027	0.1	0.0054	0.032	0.12	0.007	−0.84	−0.81
C. elegans (*n* = 17)								
ND6	0.0089	0.027	0.0035	0.0054	0.017	0.002	2.07*	1.31
ND4L	0.0059	0.023	0.0018	0.0051	0.013	0.0032	0.48	0.23
ND1	0.0092	0.041	0	0.0062	0.028	0	1.8	0.69
ND2	0.0071	0.026	0.0021	0.0046	0.016	0.0016	1.99*	1.02
cyt *b*	0.0088	0.031	0	0.0074	0.029	0	0.68	−0.18
concatenated	0.0081	0.031	0.0014	0.0057	0.021	0.0011	1.68	0.63

Legend: π, diversity from pairwise differences, considering all sites; π(s), π(a), π from synonymous and nonsynonymous sites, respectively; Θ, diversity from the number of segregating sites, S; Θ(s), Θ(a), Θ from synonymous and nonsynonymous sites, respectively. Tajima's *D* and Fu & Li's *D** are test statistics for the assumption of neutral sequence evolution. *: significant at 5% level.

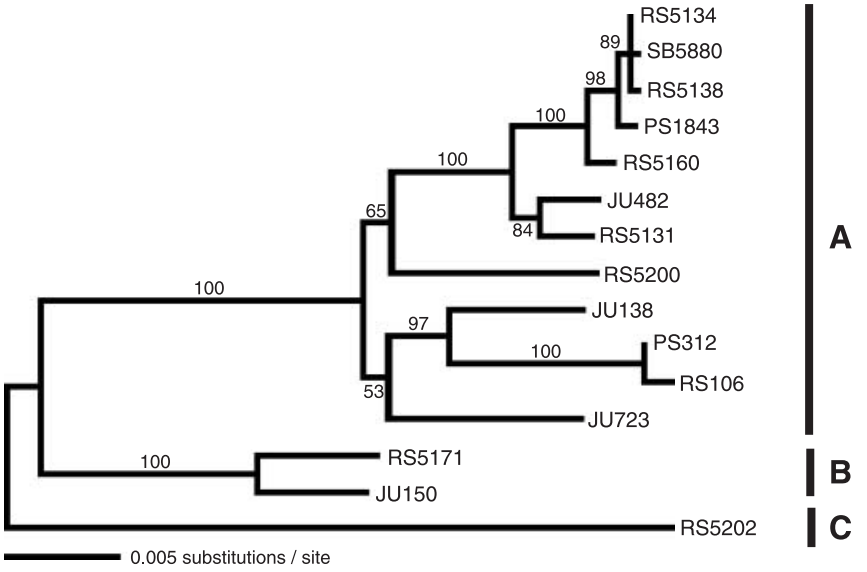


Fig. 1 Maximum likelihood tree of *Pristionchus pacificus* strains. Concatenated sequences of 15 *P. pacificus* strains (see grey boxes, Fig. 2). Numbers indicate bootstrap support (500 iterations). The substitution model was selected by the MODELTEST 3.7 software (Posada & Crandall 1998). The likelihood settings, corresponding to the TIM + I + G model, were base frequencies of A = 0.4028, C = 0.1118, G = 0.1046, T = 0.3808; substitution rates of A-C = 1.0000, A-G = 17.4529, A-T = 0.5257, C-G = 0.5257, C-T = 13.7244, G-T = 1.0000; proportion of invariable sites (I) = 0.6136, gamma distribution shape parameter = 0.9561. A heuristic search was performed. The tree is rooted at midpoint for presentation.

Nebraska and Massachusetts, respectively (Herrmann *et al*. 2006b). However, these are coincidental findings — an insect species that can be called a host for *P. pacificus* has still to be identified. Finally, we noted that the sampling of *Pristionchus* in South Africa differs strongly from the findings from Western Europe and North America. Only four *Pristionchus* isolates have been obtained from insects in South Africa after intense sampling efforts (data not shown). Surprisingly, all of them represent strains of *P. pacificus*. Furthermore, the isolates from South Africa are unique in that three out of four isolates are from the same location, a sport field in Pretoria.

Mitochondrial sequences reveal high level of genetic diversity between P. pacificus *isolates*

To study molecular diversity and the extent of LD within the *P. pacificus* isolates, we obtained and analysed three molecular data sets. Mitochondrial DNA sequences were used to construct a phylogeny of hermaphrodite lineages and to compare diversity parameters with homologous sequences of *Caenorhabditis elegans*. We genotyped SSCP markers from the *P. pacificus* genetic linkage map (Srinivasan *et al*. 2002) to estimate LD in relation to genetic map distance and to search for traces of historic recombination. Finally, AFLP data provide a genome-wide view of diversity and LD.

Mitochondria are predominantly maternally inherited, thus a phylogeny derived from mitochondrial data shows relationships of hermaphroditic lineages, irrespective of potential recombination events in the nuclear genome. We sequenced 3344 bp of mtDNA, containing sequences of the genes ND1, ND2, ND4L, ND6, cytochrome *b* and flanking regions (Fig. S1, Supplementary material). Sequence composition is characterized by a low GC content, especially at synonymous positions (Table 2), and a bias toward transitions, similar to what is seen in mitochondrial sequences of most animals, including *C. elegans* (Thomas & Wilson 1991). In total, 162 phylogenetically informative sites are used to construct maximum likelihood and maximum parsimony trees. Both methods result in the same tree topology.

On the basis of these mitochondrial trees we can distinguish three clades (Fig. 1). Clade A contains 11 strains from Asia, Europe and North America, Clade B is represented by two strains from Montenegro and Madagascar and Clade C is represented by the strains from Pretoria, South Africa. We obtained the full sequence data set for one of the four available African strains. More detailed sequence analysis of the ND6 and the ND4l genes, as well as SSCP gel analysis, of all four isolates from South Africa showed that these isolates are very similar at the SSCP level and form a distinct clade when mitochondrial haplotypes are compared (Fig. S2 and data not shown; see below for more detailed analysis of the South African isolates).

Within clade A, four of the North American strains from Ohio (RS5134), New York (SB5880), Nebraska (RS5138) and Washington (PS1843) have similar haplotypes and build a subclade with exclusively American members (Fig. 1). The two Japanese isolates (JU482 and RS5160) are associated with five of the six isolates from the North American continent. Surprisingly, PS312 from California, the laboratory strain that is used for genetic, developmental and molecular analyses, is not part of this subclade. Previous studies already indicated that PS312 is identical in its AFLP pattern to the RS106 strain from Poland (Srinivasan *et al*. 2002). The mitochondrial sequence data confirm this result and reveal that PS312 and RS106 have almost identical haplotypes. Thus, the mitochondrial haplotype of PS312

differs from all other continental strains of North America and is most closely related to RS106 from Poland and JU138 from Hawaii. Taken together, the mitochondrial phylogeny indicates the existence of substantial genetic diversity between the worldwide *P. pacificus* isolates and the existence of at least three distinct clades. Furthermore, these data suggest several dispersal events between North America, Europe and Asia.

Comparison of P. pacificus *and* C. elegans *mitochondrial sequence data*

A comparison of our tree with the *C. elegans* mitochondrial tree from Denver *et al*. (2003) suggests that the worldwide isolates of *P. pacificus* are more divergent than those of *C. elegans*. To investigate this difference quantitatively, we obtained divergence parameters in homologous alignments of mitochondrial coding sequences of our data set and 17 strains from the data set of Denver and coworkers (Figs 2 and 3, Tables 2 and 3). With the exception of the four African strains, we possess only one *P. pacificus* isolate per sampling site. Thus, we also included only one of the strains from Pretoria in the divergence analysis. Similarly, in order to have a *C. elegans* data set that is comparable to the scattered sample structure in *P. pacificus*, we include only one randomly chosen *C. elegans* sequence in those cases where the data set of Denver and co-workers contains several strains from the same locality (see also Material and Methods and Table S1). This subsampling scheme may approximate the assumptions of the neutral coalescent process if an island model with many demes, which are connected by migration, adequately describes the population in the 'collecting phase' of a genealogy (Wakeley & Lessard 2003; Lessard & Wakeley 2004; Cutter 2006).

We report diversity estimates for all sites, as well as separately for synonymous and nonsynonymous sites (Table 3), for five mitochondrial genes and the concatenated data set. Estimates of the average pairwise distance per site (π) are around three times higher in *P. pacificus* (e.g. $\pi_s = 0.1$) compared to *C. elegans* ($\pi_s = 0.031$). The population parameter Θ ($\Theta = 2N_e\mu$ for mitochondrial sequences at equilibrium, N_e = effective population size of hermaphrodites, μ = mitochondrial mutation rate) can be derived from sequence data by using the number of segregating sites (Watterson 1975). We obtained values for Θ that are five to six times higher in *P. pacificus* ($\Theta_s = 0.12$ for the concatenated data set) compared to *C. elegans* ($\Theta_s = 0.021$; Table 3). Possible explanations for the higher diversity estimates in *P. pacificus* will be discussed below and comprise, among others, different population substructures, different actual census population sizes and a higher mitochondrial mutation rate in *P. pacificus*.

The data presented here do not indicate a strong correlation of genetic and geographic distance. For example, we find no fixed differences between groups when we compare the strains from Eurasia and continental North America. However, our knowledge of local population structures and genetic diversity is limited and multiple *P. pacificus* isolates from the same locality are only available from Pretoria, South Africa. The four strains from Pretoria have similar haplotypes and form a distinct clade (Fig. S2 and data not shown). Furthermore, within 760 bp of ND6 and ND4L sequence we counted 14 fixed nucleotide differences between the strains from North America and Africa and 5 between Eurasia and Africa (Tables 4 and 5). This example suggests that it might be more likely to find a correlation of geographic and genetic structure of *P. pacificus* populations on a local rather than a *trans*-continental scale. However, further sampling is needed to see whether this holds true on a larger scale and in other localities. Taken together, we conclude that the global meta-population of *P. pacificus* shows a substantially higher degree of mtDNA diversity than *C. elegans*. However, like in *C. elegans*, there is little evidence for geographic structuring on a global scale, indicating repeated *trans*-continental dispersals of the species.

One way to quantify deviations from a neutral scenario of sequence evolution, either caused by selection or demographic processes, is to compare the two diversity parameters Θ and π (Tajima 1989). At equilbrium, Θ, calculated

Table 4 Nucleotide diversity in geographic subgroups

	n	S	π
North America	6	22	0.011
Eurasia	6	45	0.024
Pretoria	4	20	0.014
Total	16	71	0.027

Legend: n, number of strains; S, segregating sites; π, Nucleotide diversity; Continental North America: SB5880, RS5134, RS5138, PS1843, RS5131, PS312; Eurasia: JU482, JU723, RS106, RS5171, RS5200, RS5160; Pretoria (South Africa): RS5202, RS5203, RS5204, RS5205. The basis for the analyses in Tables 4 and 5 is an alignment of 760 bp of the genes ND6 and ND4L.

Table 5 Genetic differentiation of scattered worldwide strains compared to the local isolates from Pretoria, South Africa

	Fixed differences	F_{ST}
North America — Pretoria	14	0.71
Eurasia — Pretoria	5	0.5
Eurasia — North America	0	0.076

Legend: Fixed differences, number of sites at which all of the sequences in one population are different from all of the sequences in the second population.

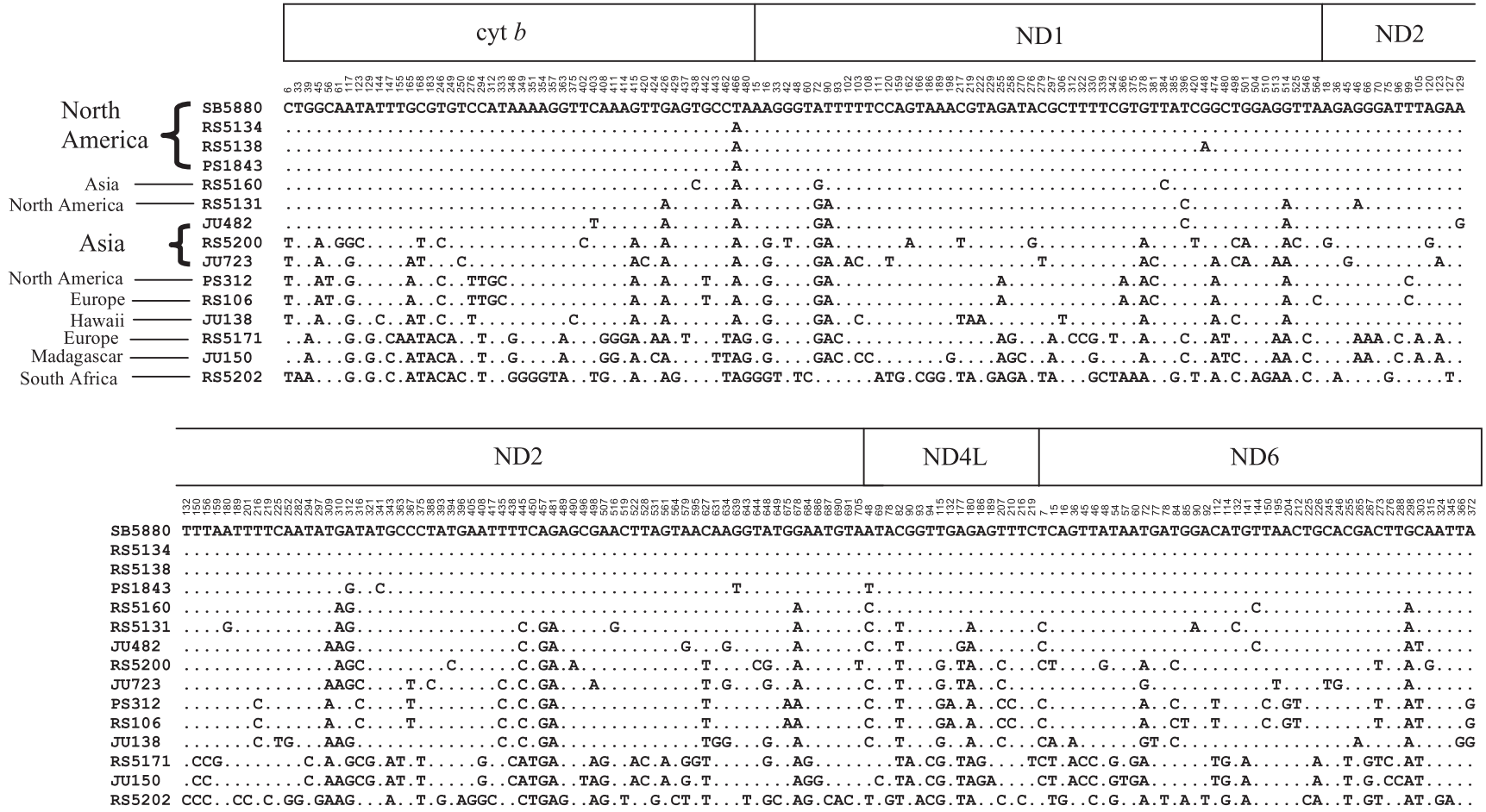
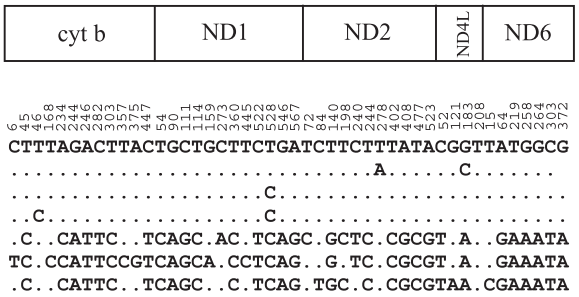


Fig. 2 Multi-gene haplotypes of *Pristionchus pacificus* strains. 2369 bp concatenated coding sequence of the mitochondrial genes cyt *b*, ND1, ND2, ND4L and ND6 reveal 243 segregating sites in 15 *P. pacificus* strains. Numbers above the nucleotides indicate the position of polymorphic sites within the alignment of the respective genes.




```
                6 45 46 68 234 244 246 262 303 357 374 447 54 90 111 114 159 273 360 415 425 528 546 567 72 84 140 198 200 244 278 402 407 477 523 521 1183 208 15 64 219 258 304 303 372
H1  CTTTAGACTTACTGCTGCTTCTGATCTTCTTTATACGGTTATGGCG
H2  ..............................A......C.......
H3  ......................C.........................
H4  ..C..................C.........................
H5  .C..CATTC..TCAGC.AC.TCAGC.GCTC.CGCGT.A..GAAATA
H6  TC.CCATTCCGTCAGCA.CCTCAG..G.TC.CGCGT.A..GAAATA
H7  .C..CATTC..TCAGC..C.TCAG.TGC.C.CGCGTAA.CGAAATA
```

Fig. 3 Multi-gene haplotypes of *Caenorhabditis elegans*, 2372 bp of mtDNA sequence homologous to the one analysed in *P. pacificus* (Fig. 2) reveal 46 segregating sites. Haplotypes represent the following strains: H1: AB3, PB306, KR314; H2: CB4855; H3: CB4858, CB4857; H4: CB4852; H5: CB4507, DH424, LSJ1, TR389, CB4851, N2, CB4932; H6: CB4856; H7: PB303, RC301.

from the number of segregating sites, and π, based on pairwise differences, should yield similar values. Tajima's *D* is a test statistic for this expectation. In *P. pacificus*, we observe negative values of *D* for all five mitochondrial genes. Although not significant, these could be indicative of population expansion or purifying selection. In contrast, *D* values in *C. elegans* are positive, in case of ND6 and ND2 significantly at the 5% level. Values of *D* are positive in *C. elegans* regardless of the choice of strains used in this scattered sample approach. However, we note that the numerical value (and statistical significance) of D in the *C. elegans* sample varies depending on the sample composition (Materials and Methods, Tables S1 and S2). The different signs of *D* in the mitochondrial samples of *P. pacificus* (*D* negative) and *C. elegans* (*D* positive) might be a first indication for different evolutionary processes acting on these genes.

SSCP and AFLP genotyping indicate rare recombination

Pristionchus pacificus propagates as a self-fertilizing hermaphrodite, but males appear spontaneously at low frequency under laboratory conditions, similar to *C. elegans*. Thus, out-crossing events between different strains are possible. Frequent out-crossing should lead to a breakdown of LD over genetic distance (Nordborg 2000; Charlesworth & Wright 2001; McVean 2002; Nordborg *et al*. 2002). In particular, strong LD should be rare across chromosomes as compared to within chromosomes. To assess the rate of LD in relation to genetic map distance we genotyped 11 of the *P. pacificus* isolates for 32 SSCP markers taken from the genetic linkage map (Fig. 4) (Srinivasan *et al*. 2002 and www.pristionchus.org). Out of 496 pairwise comparisons, 293 (59%) show significant LD ($P < 0.05$) (Fig. 5). After Bonferroni correction, 159 pairwise comparisons (32%) are significant. When we compared within-chromosome and across-chromosome

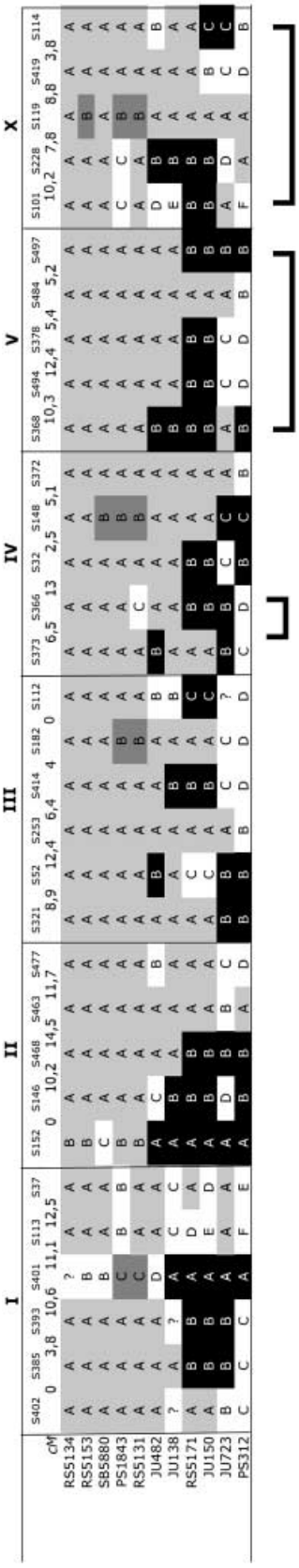


Fig. 4 SSCP genotypes of 11 homozygous *Pristionchus pacificus* strains. SSCP markers are located on the five autosomes (I–V) and on the sex chromosome (X). Genetic distances between markers (cM) have been obtained from an F2 meiotic mapping panel with 42 animals (Srinivasan *et al*. 2002). Alleles are labelled with capital letters according to frequency. Brackets show pairs of loci on the same chromosome with all four possible gamete types.

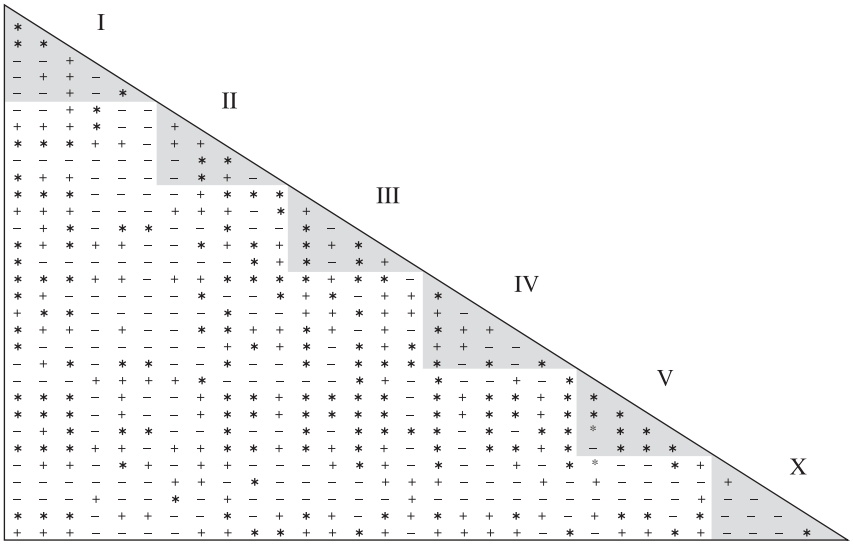


Fig. 5 Pairwise linkage disequilibrium between SSPC markers derived from data in Fig. 6. +, significant with $P = 0.05$; *significant ($P = 0.05$) after Bonferroni correction. Gray shading indicates within-chromosome comparisons.

LD, we found no differences: 43/70 (61%) pairwise comparisons of within-chromosome markers are significant ($P < 0.05$), virtually the same fraction as in across-chromosome comparisons. We conclude that the effective recombination rate between these strains is low. One caveat of the SSCP data set is that the markers are taken from the genetic linkage map and are thus preselected to be polymorphic between the strains PS312 and PS1843 (Srinivasan *et al*. 2002). To cover an unselected genome-wide group of random markers, we also compared levels of genetic diversity and LD in a set of 77 randomly chosen polymorphic AFLP (Vos *et al*. 1995) markers (Table 6; Fig. S3, Supplementary material) in the same 11 *P. pacificus* strains. Sixty-nine out of these 77 markers (90%) are polymorphic in the selected 11 strains. For comparison, in a previous AFLP study including only four strains (Srinivasan *et al*. 2001), 68% of the characters were found to be polymorphic. The average proportion of loci that differ between pairs of strains studied here is 33% (Nei's AFLP diversity; Nei 1987). These data indicate a higher AFLP diversity in *P. pacificus* than was found in *C. elegans*. In addition, we determined LD among and between marker types (sequence data, SSCP and AFLP) and genomic origin (nuclear and mitochondrial) (Table 6). We find high levels of LD, irrespective of the type of comparison. In particular, out of 2208 pairwise comparisons between AFLP and SSCP loci, 738 pairs of loci (33%) show significant LD ($P < 0.01$). Similarly, by analysing 1258 pairs of polymorphic sites of the mitochondrial genes ND6 and cyt *b*, we find that 458 pairs of loci (36%) show significant LD ($P < 0.01$). If outcrossing occurred regularly between polymorphic strains, LD between maternally inherited mitochondrial and recombining nuclear loci should be lower than between loci from the same genomic origin. However, we find that, in total, 2304 out of 7171 pairwise comparisons between mitochondrial and nuclear sites (32%) indicate significant LD ($P < 0.01$), which is only marginally lower than the values obtained from comparisons within the mitochondrial and nuclear data sets (Table 6). Thus, SSCP and AFLP data sets indicate high levels of genome-wide LD that

Table 6 LD among and between nuclear and mitochondrial loci

	Marker	Loci	p.c.	**	%
a) within type	cyt *b* (mtDNA)	37	666	296	44%
	ND6 (mtDNA)	34	561	143	25%
	SSCP (nuclear)	32	496	206	42%
	AFLP (nuclear)	69	2346	721	31%
b) mtDNA-mtDNA	cyt *b* vs. ND6	37 + 34	1258	458	36%
c) nuclear-nuclear	SSCP vs. AFLP	32 + 69	2208	738	33%
d) mtDNA-nuclear	cyt *b* vs. SSCP	37 + 32	1184	417	35%
	cyt *b* vs. AFLP	37 + 69	2553	883	35%
	ND6 vs. SSCP	34 + 32	1088	339	31%
	ND6 vs. AFLP	34 + 69	2346	665	28%
	Total	71 + 101	7171	2304	32%

Legend: a) comparisons within marker type b), c) comparisons between polymorphic loci of the same genomic origin; d) comparisons between loci of different genomic origin; marker: cyt *b* and ND6 are mitochondrial genes, SSCP and AFLP patterns cover the entire genome; loci: number of polymorphic loci within a marker type. Polymorphisms are segregating sites in mtDNA sequences, PCR fragment migration differences in SSCP gels and presence/absence of bands in AFLP patterns. p.c., number of pairwise comparisons; **number of pairs with significant LD at $P < 0.01$. The following 11 strains are used in the analysis: RS5134, RS5153, SB5880, PS1843, RS5131, JU482, JU138, RS5171, JU150, JU723 and PS312.

also extends to the maternally inherited mitochondrial genome. This result suggests rare recombination between isolates from distant sampling sites.

Similar to *C. elegans*, *P. pacificus* generates males with a low frequency in the laboratory; however, the frequency of males in laboratory cultures is higher in *P. pacificus* than *C. elegans* (unpublished observation). Therefore, we asked whether there are traces of recombination events in our sample, which would suggest a role for males in nature. One indication of historic recombination is the presence of all four possible gamete types, given two alleles at two loci and neglecting recurrent mutations (Hudson & Kaplan 1985; Stumpf & McVean 2003). In our SSCP data, we found 16 (partially redundant) pairs of loci with all four gamete types (Table S3, Supplementary material). Three out of these 16 suggest cross over events on chromosomes IV (S373-S366), V (S368-S497) and X (S101-S114), respectively (brackets in Fig. 4). These results provide some evidence for rare recombination and suggest a possible role for males in nature. We also tried to find evidence for recombination at a local scale by analysing 30 of the SSCP markers described above in the four local strains from South Africa. Only three markers were polymorphic and no evidence for out-crossing has been obtained from this limited set of local isolates (data not shown).

Discussion

In this study, we have reported substantially higher levels of molecular diversity in *Pristionchus pacificus* compared to *Caenorhabditis elegans*, as inferred from mtDNA. We used mitochondrial sequences to compare population parameters between *C. elegans* and *P. pacificus* because of the opportunity to assign homologous sequences at the amino-acid level and, at the same time, have markers that are highly variable within species at the nucleotide level. It should be emphasized that, due to the lack of recombination, data from different mitochondrial genes are not independent and might be affected simultaneously by non-neutral processes like hitch-hiking (Ballard & Rand 2005). However, we find high levels of LD between nuclear and mitochondrial loci, which could be seen as indication that mitochondrial and nuclear data sets yield congruent answers. The mitochondrial mutation rate for base substitutions has been estimated in *C. elegans* mutation accumulation lines to be 9.7×10^{-8} per site per generation (Denver *et al*. 2000). Mutation rates for *P. pacificus*, nuclear or mitochondrial, have not been estimated yet, but studies in other nematodes suggest that there is some variability of mutation rates within nematodes (Baer *et al*. 2005). Thus, one explanation for the higher molecular diversity of *P. pacificus* strains might be a higher mitochondrial mutation rate.

However, other explanations could also account for the observed differences between *P. pacificus* and *C. elegans*. Under the assumption that the mitochondrial mutation rate is not significantly different between both species, estimations of the effective population size of hermaphrodites ($N_{e(\text{hermaphrodites})} = \Theta/2\,\mu_{(\text{mtDNA})}$ at mutation-drift equilibrium) are roughly five times higher in the *P. pacificus* than in the *C. elegans* populations that are represented by the compared samples. This difference could be interpreted at face value: The actual census size of the worldwide population of *P. pacificus* might be higher than that of *C. elegans*.

It should be noted that the estimation of N_e from Θ is only meaningful if the scattered sample approach taken here approximates the assumptions of a neutral genealogy. In particular, we have to assume that populations of *P. pacificus* and *C. elegans* both can be adequately described by an island model of migration with a large number of demes that are connected by migration (Lessard & Wakeley 2004; Cutter 2006). Indeed, one reason for our observation of higher mtDNA sequence diversity in *P. pacificus* could be that there is little migration between *P. pacificus* subpopulations when compared to *C. elegans*. In general, the relation between effective population size and actual census size is influenced by various factors: Recent population bottlenecks, for example, lead to drastic reductions of N_e (Hartl & Clark 1997). Dealing with a partially selfing species, the influence of the breeding system is of particular importance: complete inbreeding reduces $N_{e(\text{nuclear})}$ by $^1/_2$ (Nordborg 2000; Charlesworth & Wright 2001). Direct effects of inbreeding should only affect diversity levels of nuclear genes and not mitochondria; however, several processes have been proposed that can lead to further reductions of N_e, affecting nuclear and organelle genomes. Genetic hitch-hiking effects, for example, extend over larger genetic distances in selfers due to reduced recombination (Nordborg 2000; Charlesworth *et al*. 2003).

Could a higher out-crossing rate in *P. pacificus* explain the higher molecular diversity compared to *C. elegans*? In *C. elegans*, estimates of out-crossing rates have been obtained by several approaches. First, minor incongruence between mitochondrial and nuclear phylogenies indicates infrequent out-crossing in the wild (Denver *et al*. 2003; Barrière & Félix 2005a). Second, from levels of LD, effective out-crossing rates have been calculated in the range of one in several thousand generations (Barrière & Félix 2005a; Cutter 2006). Finally, individual heterozygous animals have been isolated at low frequencies in compost heaps, leading to estimates of 1.3% out-crossing in local populations (Barrière & Félix 2005a). In *P. pacificus* we find indications for occasional recombination, but strong genome-wide LD on a global scale, as revealed by AFLP, SSCP and mtDNA analysis, suggests a predominantly self-fertilizing mode of reproduction in nature. Nonetheless, strong LD can also arise by strict isolation of populations over long time periods. Most of the available *P. pacificus* strains go back to individuals from distant localities, making it difficult to

rule out this possibility. In conclusion, substantially higher effective recombination rates are unlikely to be a sufficient explanation for the higher molecular diversity in *P. pacificus* for two reasons: First, diversity of mitochondrial sequences should not be affected directly by the breeding system. However, like the nuclear genetic diversity, the mitochondrial genetic diversity may be drastically reduced due to selective sweeps in a selfing species. Second, high levels of interchromosomal LD indicate rare recombination between strains from distant localities, either because local populations have little secondary contact with each other or because males are rare.

To understand the different diversity patterns in the two nematode species it will be essential to study the ecologies of *P. pacificus* and *C. elegans*. What are their host species, how abundant are they and how are they dispersed over long distances? Differences in these parameters might influence, for example, migration rates between subpopulations and meta-population events like extinctions followed by recolonization. Both processes, migration and extinction/recolonization, can lead to reduced structuring and low molecular diversity (Ingvarsson 2002). On the other hand, high levels of diversity can be maintained, even in a selfing species, in the presence of population structure. The selfing plant *Arabidopsis thaliana*, for example, shows a polymorphism pattern that would, generally speaking, not be unusual for a sexually reproducing species (Nordborg *et al*. 2005). To our knowledge, no other selfing nematode species outside the genus *Caenorhabditis* have been studied. *Caenorhabditis elegans* isolates used in the study of Denver *et al*. (2003), which is the basis for our comparison, but also most of the other available strains, stem from habitats closely associated with human activity, namely compost heaps. Thus one possible explanation for the low diversity in this species is that high, human-caused migration rates together with frequent bottlenecks led to a homogenization of the global *C. elegans* meta-population (Barrière & Félix 2005a). However, it has also been discussed previously (e.g. Cutter 2006) that no *C. elegans* strains from Asia, Africa or South America have been obtained yet, which raises the possibility that the global diversity of this species might be seriously underestimated.

In contrast, *P. pacificus* is found in soil samples with a much lower frequency than *C. elegans*, and not preferentially in human-associated habitats. To find potential hosts of *Pristionchus* nematodes, large sampling efforts have been carried out, including the examination of more than 8.000 beetles from Europe and North America (Herrmann *et al*. 2006a, b). These studies revealed that members of the genus *Pristionchus* are associated with scarab beetles. However, a host of *P. pacificus* has not been identified in these field studies. Specifically, only three out of 285 *Pristionchus* strains isolated from Scarab beetles in the USA and none of 371 strains from Western Europe represent *P. pacificus* (Herrmann *et al*. 2006a, b). Currently, further sampling in other locations, i.e. in Asia, is in progress. For two hermaphroditic members of the genus *Pristionchus* a clear beetle association could be shown (Herrmann *et al*. 2006a). *Pristionchus maupasi* is associated with the chockcafer *Melolontha melolontha* and *Pristionchus maupasi entomophagus* with the dung beetle *Geotrupes stercorarius*. *Pristionchus* nematodes have a necromenic association with their beetle hosts — that is, the nematodes, while sitting on the beetle as a dauer larva, do not harm it as long as the beetle is alive, but feed on the growing microbes on the carcass after its death (Kiontke & Sudhaus 2005). We speculate that *P. pacificus* lives in a similar association with a yet unidentified beetle species. Additional sampling will be necessary to investigate in how far differences in ecology might be responsible for the higher genetic diversity of *P. pacificus* that is revealed in this study. In fact, our data would be consistent with a rather ancient population subdivision (e.g. between the African and American strains) as well as more recent colonization events (e.g. regarding the molecularly identical strains from California and Poland and the Japanese strains within the 'American' clade; see Fig. 1). In particular, the fact that PS312 from California and RS106 from Poland share almost identical mitochondrial haplotypes and AFLP patterns indicates a recent *trans*-continental dispersal of *P. pacificus*. We speculate that this might have happened in association with an insect host. However, detailed ecological studies will be necessary to support this claim. We conclude from our study on *P. pacificus* and the comparison to *C. elegans* that natural populations of these two species might differ substantially in respect to their mutation rates and subpopulation structures, despite similar lifestyles and reproduction modes in the laboratory.

Acknowledgements

We thank Isabella D'Anna for help with SSCP gel analysis. We thank Dr A. Cutter for critically reading the manuscript and members of the Sommer laboratory for discussions. HZ was supported by a predoctoral fellowship of the Boehringer Ingelheim Fonds.

Supplementary material

The supplementary material is available from
http://www.blackwellpublishing.com/products/journals/suppmat/MEC/MEC3222/MEC3222sm.htm

Fig. S1 Mitochondrial gene order. Arrows indicate coding regions of ND6, ND4L, ND2, ND1, cyt *b* and COXIII. Gray boxes show regions used for tree construction. Lines indicate regions that have been aligned to *Caenorhabditis elegans* and were used to obtain divergence parameters in both species. Numbers are interspecific K_a values (the number of nonsynonymous substitutions per nonsynonymous site) of these regions for the comparison of *C. elegans* N2 and *Pristionchus pacificus* PS312.

Fig. S2 ND6-ND4l haplotypes of 18 *Pristionchus pacificus* strains, including four from South Africa (boxed).

Fig. S3 AFLP genotypes. 69 bands were scored as present (1) or absent (0) in the same 11 inbred strains that were used for SSCP analysis.

Table S1 Strains of *Caenorhabditis elegans* used to compare divergence parameters in mitochondrial genes

Table S2 Diversity estimates and Tajima's *D* for the 10 randomized data sets shown in Table S1

Table S3 Pairs of loci showing all four possible gamete types, as revealed by SSCP

References

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

Ballard JWO, Rand DM (2005) The population biology of mitochondrial DNA and its phylogenetic implications. *Annual Review of Ecology Evolution and Systematics*, **36**, 621–642.

Barrière A, Félix MA (2005a) High local genetic diversity and low outcrossing rate in *Caenorhabditis elegans* natural populations. *Current Biology*, **15**, 1176–1184.

Barrière A, Félix MA (2005b) Natural variation and population genetics of *C. elegans*. In: *Wormbook*. The *C. elegans* Research Community (www.wormbook.org).

Charlesworth D, Wright SI (2001) Breeding systems and genome evolution. *Current Opinions in Genetics and Development*, **11**, 685–690.

Charlesworth B, Charlesworth D, Barton NH (2003) The effects of genetic and geographic structure on neutral variation. *Annual Review of Ecology Evolution and Systematics*, **34**, 99–125.

Chasnov JR, Chow KL (2002) Why are there males in the hermaphroditic species *Caenorhabditis elegans*? *Genetics*, **160**, 983–994.

Cutter AD (2006) Nucleotide polymorphism and linkage disequilibrium in wild populations of the partial selfer *Caenorhabditis elegans*. *Genetics*, **172**, 171–184.

Cutter AD, Aviles L, Ward S (2003) The proximate determinants of sex ratio in *C. elegans* populations. *Genetics Research*, **81**, 91–102.

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

Denver DR, Morris K, Thomas WK (2003) Phylogenetics in *Caenorhabditis elegans*: an analysis of divergence and outcrossing. *Molecular Biology and Evolution*, **20**, 393–400.

Graustein A, Gaspar JM, Walters JR, Palopoli MF (2002) Levels of DNA polymorphism vary with mating system in the nematode genus caenorhabditis. *Genetics*, **161**, 99–107.

Haber M, Schungel M, Putz A *et al.* (2005) Evolutionary history of *Caenorhabditis elegans* inferred from microsatellites: evidence for spatial and temporal genetic differentiation and the occurrence of outbreeding. *Molecular Biology and Evolution*, **22**, 160–173.

Hartl D, Clark A (1997) *Principles of Population Genetics*, 3rd edn. Sinauer Associates, Sunderland, Masachusetts.

Herrmann M, Mayer WE, Sommer RJ (2006a) Nematodes of the genus *Pristionchus* are closely associated with scarab beetles and the Colorado potato beetle in Western Europe. *Zoology (Jena)*, **109**, 96–108.

Herrmann M, Mayer WE, Sommer RJ (2006b) Sex, bugs and Haldane's rule: the nematode genus *Pristionchus* in the United States. *Frontiers in Zoology*, **3**, 14; doi:10.1186/1742-9994-3-14.

Hodgkin J, Brenner S (1977) Mutations causing transformation of sexual phenotype in the nematode *Caenorhabditis elegans*. *Genetics*, **86**, 275–287.

Hong RL, Sommer RJ (2006) *Pristionchus pacificus*: a well-rounded nematode. *BioEssays*, **28**, 651–659.

Hudson RR, Kaplan N (1985) Inferring the number of recombination events in the history of a sample. *Biometrics*, **41**, 572–572.

Ingvarsson PK (2002) A metapopulation perspective on genetic diversity and differentiation in partially self-fertilizing plants. *Evolution; International Journal of Organic Evolution*, **56**, 2368–2373.

Kiontke K, Sudhaus W (2005) Ecology of *Caenorhabditis* species. In: *Wormbook*. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.37.1, http://www.wormbook.org.

Koch R, van Luenen HGAM, van der Horst M, Thijssen KL, Plasterk RHA (2000) Single nucleotide polymorphisms in wild isolates of *Caenorhabditis elegans*. *Genome Research*, **10**, 1690–1696.

Lessard S, Wakeley J (2004) The two-locus ancestral graph in a subdivided population: convergence as the number of demes grows in the island model. *Journal of Mathematical Biology*, **48**, 275–292.

Lieven Fürst von A, Sudhaus W (2000) Comparative and functional morphology of the buccal cavity of Diplogastrina (Nematoda) and a first outline of the phylogeny of this taxon. *Journal of Zoological Systematics and Evolutionary Research*, **38**, 37–63.

McVean GT (2002) A genealogical interpretation of linkage disequilibrium. *Genetics*, **162**, 987–991.

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

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

Nordborg M, Borevitz JO, Bergelson J *et al.* (2002) The extent of linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics*, **30**, 190–193.

Nordborg M, Hu TT, Ishino Y *et al.* (2005) The pattern of polymorphism in *Arabidopsis thaliana*. *PloS Biology*, **3**, 1289–1299.

Pires-DaSilva A, Sommer RJ (2004) Conservation of the global sex determination gene tra-1 in distantly related nematodes. *Genes and Development*, **18**, 1198–1208.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Rozas J, Sanchez-DeI, Barrio JC, Messeguer X, Rozas R (2003) DNASP: DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Rudel D, Riebesell M, Sommer RJ (2005) Gonadogenesis in *Pristionchus pacificus* and organ evolution: development, adult morphology and cell-cell interactions in the hermaphrodite gonad. *Developmental Biology*, **277**, 200–221.

Schlager B, Roseler W, Zheng M, Gutierrez A, Sommer RJ (2006) HAIRY-like transcription factors and the evolution of the nematode vulva equivalence group. *Current Biology*, **16**, 1386–1394.

Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN: A Software for Population Genetic Data Analysis*, version 2.000. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland.

Sivasundar A, Hey J (2003) Population genetics of *Caenorhabditis elegans*: the paradox of low polymorphism in a widespread species. *Genetics*, **163**, 147–157.

Sivasundar A, Hey J (2005) Sampling from natural populations with RNAI reveals high outcrossing and population structure in *Caenorhabditis elegans*. *Current Biology*, **15**, 1598–1602.

Slatkin M (1994) Linkage disequilibrium in growing and stable populations. *Genetics*, **137**, 331–336.

Sommer RJ, Carta LK, Kim SY, Sternberg PW (1996) Morphological, genetic and molecular description of *Pristionchus pacificus* sp.n. (Nematoda: Neodiplogastridae). *Fundamental and Applied Nematology*, **19**, 511–521.

Srinivasan J, Pires-daSilva A, Gutierrez A *et al.* (2001) Microevolutionary analysis of the nematode genus *Pristionchus* suggests a recent evolution of redundant developmental mechanisms during vulva formation. *Evolution and Development*, **3**, 229–240.

Srinivasan J, Sinz W, Lanz C *et al.* (2002) A bacterial artificial chromosome-based genetic linkage map of the nematode *Pristionchus pacificus*. *Genetics*, **162**, 129–134.

Stewart AD, Phillips PC (2002) Selection and maintenance of androdioecy in *Caenorhabditis elegans*. *Genetics*, **160**, 975–982.

Stumpf MP, McVean GA (2003) Estimating recombination rates from population-genetic data. *Nature Reviews Genetics*, **4**, 959–968.

Swofford DL (2002) *PAUP: Phylogenetic Analysis Using Parsimony, and Other Methods*, Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Thomas WK, Wilson AC (1991) Mode and tempo of molecular evolution in the nematode *Caenorhabditis* — Cytochrome Oxidase-II and Calmodulin sequences. *Genetics*, **128**, 269–279.

Vos P, Hogers R, Bleeker M *et al.* (1995) Aflp — a new technique for DNA-fingerprinting. *Nucleic Acids Research*, **23**, 4407–4414.

Wakeley J, Lessard S (2003) Theory of the effects of population structure and sampling on patterns of linkage disequilibrium applied to genomic data from humans. *Genetics*, **164**, 1043–1053.

Watterson G (1975) Number of segregating sites in genetic models without recombination. *Theoretical Population Biology*, **7**, 256–276.

The authors are at the Max Planck Institute for Developmental Biology in Tuebingen, Germany. The laboratory of Ralf J. Sommer has established *Pristionchus pacificus* as a satellite organism for evolutionary and developmental biology. Initially focussing on macro-evolutionary comparisons of development between *C. elegans* and *P. pacificus*, more recently the research agenda of the laboratory expanded to include, among others, studies of micro-evolutionary differences of development within the genus *Pristionchus* (Hans Zauner), the ecology and biogeography of the nematode-beetle interaction (Matthias Herrmann, Andreas Weller) and the phylogeny of the genus *Pristionchus* (Werner E. Mayer).
