

# Testing the hypothesis of recent population expansions in nematode parasites of human-associated hosts

DA Morrison and J Höglund

Department of Parasitology (SWEPAR), National Veterinary Institute and Swedish University of Agricultural Sciences, 751 89 Uppsala, Sweden

It has been predicted that parasites of human-associated organisms (eg humans, domestic pets, farm animals, agricultural and silvicultural plants) are more likely to show rapid recent population expansions than are parasites of other hosts. Here, we directly test the generality of this demographic prediction for species of parasitic nematodes that currently have mitochondrial sequence data available in the literature or the public-access genetic databases. Of the 23 host/parasite combinations analysed, there are seven human-associated parasite species with expanding populations and three without, and there are three non-human-associated parasite species with expanding populations and 10 without. This statistically significant pattern confirms the

prediction. However, it is likely that the situation is more complicated than the simple hypothesis test suggests, and those species that do not fit the predicted general pattern provide interesting insights into other evolutionary processes that influence the historical population genetics of host–parasite relationships. These processes include the effects of postglacial migrations, evolutionary relationships and possibly life-history characteristics. Furthermore, the analysis highlights the limitations of this form of bioinformatic data-mining, in comparison to controlled experimental hypothesis tests.

*Heredity* (2005) **94**, 426–434. doi:10.1038/sj.hdy.6800623

Published online 26 January 2005

**Keywords:** demographic expansion; population growth; mitochondrial genes; nematoda; host–parasite relationships

## Introduction

The population genetics of parasites has traditionally been a comparatively neglected field (Grant, 1994), and it is only in the last decade that it has received any serious attention (Anderson *et al*, 1998; Gasser and Chilton, 2001; Hu *et al*, 2004). This means that some quite basic predictions about the structure of parasite populations remain largely untested. One of these predictions concerns the evolutionary stability of population sizes in parasites of human-associated host species compared to those of naturally occurring host species. It has been suggested that parasites of human-associated organisms are more likely to show rapid recent demographic expansions than are other types of parasites (Blouin *et al*, 1995; Donnelly *et al*, 2001; Mes, 2003). Such demographic expansions may include increases in the numbers of individuals in the species and/or increases in the geographic range of the species, for example.

By human-associated species we mean humans themselves, domestic pets, farm animals, agricultural and silvicultural plants, and possibly other species that are now more common in human-disturbed (eg urban and agricultural) environments than elsewhere. The prediction of population expansions is based on the idea that

certain aspects of the parasite biology often closely track those of their hosts (Cox, 1993; Anderson *et al*, 1998; Poulin, 1998), as part of the coevolutionary process, and the hosts in this case have usually undergone population expansions relatively recently as a result of their own association with the rapid demographic expansion of the human population over the past 10 000 years (Cavalli-Sforza *et al*, 1994). For example, a number of pathogens have apparently undergone recent expansions in concert with that of humans, including the causal agents of plague, malaria, AIDS, tuberculosis and Japanese encephalitis (Conway and Baum, 2002). Furthermore, many of the human-associated host species are also subject to large-scale spatial movements, particularly for farm animals (compared to the usually more restricted dispersal of wild hosts), and this will also contribute to a population expansion in the associated parasite species when they come into contact with new host individuals (Blouin *et al*, 1995; Anderson *et al*, 1998; McCoy *et al*, 2003; Criscione and Blouin, 2004). Note that the hypothesis makes no specific prediction about how frequent population expansions will be among parasites of non-human-associated hosts, but merely predicts that they will be *more* frequent among parasites of human-associated hosts. There are likely to be many other factors that can potentially create demographic expansions in parasite populations, and the process of testing the hypothesis is likely to provide deeper insights into the population genetics of naturally occurring species as well.

It is important to test this prediction, and to assess its generality, because it has enormous practical as well as

Correspondence: DA Morrison, Department of Parasitology (SWEPAR), National Veterinary Institute and Swedish University of Agricultural Sciences, 751 89 Uppsala, Sweden. E-mail: David.Morrison@vmm.slu.se  
Received 25 February 2004; accepted 18 October 2004; published online 26 January 2005

theoretical implications. From the theoretical point of view, the inter-relationship between parasites and their hosts is one of the most important aspects of parasite biology, and it thus dominates most scientific studies of parasites; indeed, this distinguishes such studies from those of the hosts themselves, as the latter usually ignore the parasites entirely. From the practical viewpoint, the epidemiology of medically and veterinarily important diseases is clearly affected by the demography of the associated parasites. Furthermore, large-scale efforts are under way to control parasites of most human-associated species, notably livestock, companion animals and agricultural plants, as well as humans, particularly via past development of drugs and the more recent interest in vaccines. The success or failure of these projects will clearly be determined to a large extent by how successfully they deal with any changes occurring in parasite population sizes (eg the current spread of resistance to anthelmintic drugs). However, the possibility of population expansions has rarely been quantitatively tested for parasite species (eg Hughes and Verra, 2001).

Here, we attempt to test directly the generality of this prediction using as many species of parasitic nematodes (phylum Nematoda) as possible, by analysing the data currently available in the literature and the public-access genetic databases.

## Materials and methods

The best of the current methods for directly testing population expansions require nucleotide sequence data (although good methods have recently been developed also for microsatellite data; Beaumont, 1999). Unfortunately, this requirement excludes from consideration much of the population-genetics data for nematodes, which have historically been based on allozyme, RAPD or RFLP analyses (with occasional use of microsatellites and AFLP). Also, the test methods assume that there has been no recombination in the recent evolutionary history of the population, and so we chose to restrict the data set to mitochondrial gene sequences, since there is little evidence for recombination in nematode mitochondria (Mes, 2003). This has the added advantages of still including the most variable genes (Blouin, 2000b), making all of the population estimates directly comparable (ie there are no confounding effects of inheritance patterns or ploidy), and making the sequence alignment simple because there is little length variation of these genes. Finally, the test methods require data for at least four individuals from the population and with at least one polymorphic site. This requirement excludes data based on consensus sequences from pooled isolates, cDNA libraries and laboratory-maintained strains. Within these constraints, we performed literature searches as well as searches of nucleotide databases, in order to locate suitable data sets. We have assumed that all of the genetic markers used are neutral or nearly neutral, rather than subject to some form of selection – this is hard to test, but our results using synonymous sites (for the protein-coding genes) are not different from those using all nucleotide sites, for example.

For each data set, the degree of demographic expansion was first assessed using maximum likelihood estimates of the population growth rate based on coalescent simulations. Calculations were performed

using the Fluctuate v1.4 program (Kuhner *et al*, 1998), which estimates the exponential growth rate ( $g$ ) of the population within an explicitly genealogical framework. Since the implementation of these calculations is sensitive to the parameters specified for the coalescent model (Beaumont, 1999; Mes, 2003; Criscione and Blouin, 2004), the nucleotide frequencies, transition:transversion ratio and proportion of invariable sites were all estimated quantitatively, as these have very nonstandard values for nematode mitochondria (Blouin, 2000a) – there is a large AT bias, a high transition:transversion ratio and most of the sites are invariable. These three sets of estimates were produced simultaneously via maximum likelihood using the Tree-Puzzle v5.0 program (Strimmer and von Haesler, 1996) – the exact likelihood function was used with parameters estimated by quartet sampling plus the neighbor-joining tree, based on the HKY nucleotide-substitution model. In Fluctuate, the Tree-Puzzle estimated values were then prespecified, and the Watterson estimate was used as the starting value of theta and zero as the starting value of  $g$ , along with a random starting tree. The Metropolis-Hastings sampler used 10 short chains with 1000 steps of increment 20, followed by three long chains with 20 000 steps of increment 20. If this search strategy was insufficient to produce a stable result, then we tried (in order): an alternative random-number seed, doubling the number of short chains and doubling the number of long chains.

As an independent assessment of population expansion for each data set, we used the statistic from the  $F_s$  test of selective neutrality (Fu, 1997). This has been shown to be one of the most powerful of the tests based on summary statistics (Ramos-Onsins and Rozas, 2002), being based simply on the haplotype distribution within the population. Calculations were performed using the DnaSP v4.0 program (Rozas *et al*, 2003), and statistical significance was assessed using 10 000 coalescent simulations based on segregating sites, with all of the parameters set to the observed values.

For both sets of analyses, the calculations were performed separately for each gene sequence for each combination of parasite species and host species. The calculations were also performed separately for suitable subsets (eg at least four individuals with at least one polymorphic site) of the data, such as geographically separated subpopulations. For Fluctuate, the analysis substitution parameters were estimated independently for each host/parasite combination; however, for each subpopulation within a host, the values estimated for the whole population were used.

The relationship between whether or not there has been a recent population expansion in the parasites and whether or not their hosts are human-associated was tested with a log-likelihood  $2 \times 2$  contingency test with Williams correction (Sokal and Rohlf, 1994). Decisions as to whether the particular host populations sampled were human-associated or not were based on the information provided in the original publications – for example, the muskoxen sampled for *Teladorsagia boreoarcticus* were from natural and historically isolated populations, whereas muskoxen have been deliberately introduced into other locations, and the white-tailed deer sampled for *Mazamastrongylus odocoilei* were human-associated – and all decisions were made before the data analyses were carried out.

## Results

Suitable data were located for 18 nematode species, covering 23 host/parasite combinations (Table 1). No usable data were found relating to plant-parasitic nematodes, and the data set is restricted almost entirely to parasites of vertebrates. Some other potentially suitable publications were found but could not be used, either because the sequence data are not available in public databases (Leignel and Humbert, 2001), the available sequence data cannot be attributed directly to the individual nematodes (Hawdon *et al.*, 2001; Leignel *et al.*, 2002; van der Veer *et al.*, 2003) or the sequence data are invariant among the individuals (Perlman *et al.*, 2003).

The analyses performed using Fluctuate appear to be more powerful than those using Fs, as all of the population expansions strongly detected using Fs were also detected using Fluctuate, but Fluctuate also detected several weaker expansions (Table 1). This result could also be a by-product of the tendency towards upward bias that can occur in the values of  $g$  (Kuhner *et al.*, 1998). Furthermore, Fluctuate also detected some population contractions (indicated as negative  $g$ -values in Table 1), which Fs does not directly test. Our classification of the nematode species as having a recent demographic expansion or not is based on the agreement between the two analyses.

For those data sets where several loci were sampled, we combined the results using a weighted average of the values from each locus, weighted by the respective sample sizes. For the  $g$ -values, we calculated a standardized statistic for testing purposes, calculated as  $g$  divided by its standard error. We thus produced a single value of Fs and standardized  $g$  for each host/parasite combination (Figure 1).

Of the 23 host/parasite combinations, there are seven human-associated parasite species with expanding populations and three without, and there are three non-human-associated parasite species with expanding populations and 10 without (Figure 1). The contingency statistical test for these data yields  $G = 0.90$ ,  $P = 0.027$ . So, we can reject the null hypothesis that there is no relationship between parasite population expansions and human-associated hosts. However, one of the four expected frequency values is  $< 5$  (ie 4.4), and so a cross-check with the more conservative Fisher exact test is appropriate. This yields  $P = 0.040$ , thus confirming the result. Nevertheless, the retrospective power of the  $G$ -test is relatively low, at 0.64, and thus relatively small changes in the composition of the data would have a notable effect on the outcome – 18 host/parasite combinations would be needed in both the human-associated and nonassociated groups in order to raise the power to 0.80 (calculated using the DStplan v4.2 program; Brown *et al.*, 2000).

## Discussion

A clear test of the hypothesis is evident in the data currently available: nematode parasites of human-associated hosts are more likely to show evidence of rapid population expansions in their recent evolutionary history than are nematode parasites of naturally occurring hosts. This conclusion seems to apply to a wide taxonomic diversity of nematodes, and it is based on a

consistent methodology applied across all of the taxa, without having to assume a common evolutionary history for the species. However, the available data restrict this conclusion largely to nematodes with vertebrate hosts, and the hypothesis that the same situation applies to nematode parasites of plants, for example, remains to be tested. It also remains to be tested for other candidate parasites, such as those of the phyla Platyhelminthes, Acari and Insecta. Nevertheless, the consistent nature of the pattern suggests that it may be true for these other parasites as well (but see below).

There are several specific patterns within the nematode species that also confirm this general result in their details. First, samples were taken for *Oesophagostomum bifurcum* from both human and Mona monkey hosts in northern Ghana, and the human sample shows clear evidence of a recent population expansion while the non-human sample does not. While the two sample sizes are not large, this is precisely the sort of detailed result that the theory predicts. Second, for those species with sufficient data, the geographical subpopulations usually show the same pattern as that of the population as a whole. This provides evidence that the observed pattern is a general one for those species, at least at the locations sampled, and not an artefact of the particular sample taken. This confirmation is important because the maximum likelihood coalescent method used by Fluctuate is designed to deal with panmictic (ie undivided) populations, and strong geographic subdivision can produce demographic patterns that are opposite to those of population expansion.

## Cautionary notes

In contrast, it should be noted that, for those parasite species where multiple mitochondrial genes have been sampled, the results based on different genes are sometimes not congruent (Table 1). This is unfortunate, because consistent patterns across loci would provide stronger evidence for or against a demographic expansion. In particular, it proved to be impossible to combine the data for multiple genes for any species except one – Fluctuate would not produce estimates of the standard error in any of these cases. This indicates that the loci are producing very different estimates of the rate of change in the population from which they came, which is also obvious in most cases simply by comparing the  $g$ -values shown for each gene in Table 1. So, accurate estimates of the magnitude of the demographic expansions cannot be obtained even when there is consistent evidence that an expansion (or contraction) has occurred. Other recent studies of population expansions have not been confronted with this problem, because they have been based on only one locus (eg Lessa *et al.*, 2003; Mes, 2003). If an overall estimate is needed, it might thus be more useful to produce a weighted average of the individual estimates (which is the strategy we adopted here) rather than trying to combine the data first and then produce a single estimate.

There are several other considerations that indicate that the results here should be treated with at least some caution. As with all studies of unobservable historical events for which manipulative experiments cannot be performed, there are a series of untestable assumptions any one of which might lead to a failure to detect the

**Table 1** Summary details of the data included in the analyses, indicating the nematode species sampled and its taxonomic classification, the host from which it was collected, the gene sequenced, the number of individuals sampled (*n*), the estimated rate of population change prior to the time of sampling based on exponential growth (*g*) with its standard error (SE) and the result of the test of selective neutrality (*F<sub>s</sub>*) with its associated probability (*P*)

Classification	Host	Gene	Population	n	<i>g</i> ( $\pm$ SE)	<i>F<sub>s</sub></i> ( <i>P</i> )	Source of data
Class Enoplea							
Order Trichinellida							
Family Trichuridae							
<i>Capillaria gastrica</i>	Native rodents	<i>cox1</i>	Overall	9	$-14 \pm 7$	7.65 (0.997)	Zhu <i>et al</i> (2000)
<i>Capillaria</i> sp 2	Marsupials	<i>cox1</i>	Overall	13	$-8 \pm 2$	19.43 ( $> 0.999$ )	Zhu <i>et al</i> (2000)
			Quoll*	4	$-11 \pm 55$	2.60 (0.287)	
			Bandicoot*	9	$-18 \pm 149$	1.59 (0.820)	
Class Chromadorea							
Order Rhabditida							
Superfamily Ascaridoidea							
Family Anisakidae							
<i>Contraecum osculatum</i>	Seals	<i>cox1</i>	Overall	39	$-4 \pm 2$	11.76 (0.997)	Hu <i>et al</i> (2001)
			Bearded*	14	$-11 \pm 16$	3.44 (0.961)	
			Harp*	15	$-23 \pm 8$	9.35 ( $> 0.999$ )	
			Grey*	10	$-2 \pm 9$	2.93 (0.923)	
		<i>rrnL</i>	Overall	40	$-129 \pm 118$	2.88 (0.913)	Hu <i>et al</i> (2001)
			Bearded*	14	$-470 \pm 1446$	1.14 (0.784)	
			Harp*	15	$-479 \pm 991$	1.25 (0.847)	
			Grey*	11	$-346 \pm 258$	2.26 ( $> 0.999$ )	
		<i>rrnS</i>	Overall	41	$-20 \pm 11$	5.53 (0.971)	Hu <i>et al</i> (2001)
			Harp*	15	$33 \pm 47$	0.32 (0.636)	
			Grey*	12	$24 \pm 75$	0.87 (0.729)	
<i>Contraecum osculatum baicalensis</i>	Baikal seal	<i>cox1</i>	Overall	15	$-45 \pm 35$	3.99 (0.991)	Hu <i>et al</i> (2001)
		<i>rrnS</i>	Overall	15	$-79 \pm 388$	1.12 (0.766)	Hu <i>et al</i> (2001)
<i>Contraecum</i> sp D/E	Weddell seal	<i>cox1</i>	Overall	8	$12 \pm 4$	$-0.16$ (0.409)	Hu <i>et al</i> (2001)
			D	5	$0 \pm 2$	1.82 (0.857)	
		<i>rrnL</i>	Overall	10	$204 \pm 57$	0.11 (0.572)	Hu <i>et al</i> (2001)
			E	5	$-24 \pm 69$	1.69 ( $> 0.999$ )	
		<i>rrnS</i>	Overall	10	$38 \pm 51$	0.77 (0.701)	Hu <i>et al</i> (2001)
			D	5	$18 \pm 55$	1.22 (0.849)	
			E	5	$-24 \pm 61$	2.20 (0.838)	
Order Strongylida							
Superfamily Heterorhabditoidea							
Family Heterorhabditidae							
<i>Heterorhabditis marelatus</i>	Soil insects	<i>nad4</i>	Overall	58	$-41 \pm 12$	13.29 ( $> 0.999$ )	Blouin <i>et al</i> (1999)
			Newport	10	$-25 \pm 9$	4.58 (0.979)	
			Florence	10	$-25 \pm 9$	4.58 (0.979)	
			Bullard	10	$42 \pm 234$	0.42 (0.629)	
			Blanco	10	$-33 \pm 13$	3.67 (0.955)	
		<i>nad4</i>	Overall	5	$-5 \pm 2$	5.88 ( $> 0.999$ )	Liu <i>et al</i> (1999)
Superfamily Ancylostomatoidea							
Family Ancylostomatidae							
<i>Ancylostoma caninum</i>	Domestic dog	<i>cox1</i>	Overall	38	$-6 \pm 2$	14.43 ( $> 0.999$ )	Hu <i>et al</i> (2002a)
<i>Ancylostoma duodenale</i>	Human	<i>cox1</i>	Overall	27	$3 \pm 14$	0.85 (0.686)	Hu <i>et al</i> (2002a)
		<i>rrnL</i>	Overall	4	$413 \pm 138$	2.15 (0.768)	Hu <i>et al</i> (2003)
<i>Necator americanus</i>	Human	<i>cox1</i>	Overall	28	$-3 \pm 6$	1.59 (0.773)	Hu <i>et al</i> (2002a)
			Togo	9	$12 \pm 6$	$-0.04$ (0.394)	
			China	19	$-16 \pm 7$	2.63 (0.885)	
		<i>rrnL</i>	Overall	25	$154 \pm 23$	$-5.66$ (0.030)	Hu <i>et al</i> (2003)
			Togo	14	$7116 \pm 251$	$-7.51$ (0.001)	
			China	11	$76 \pm 33$	$-1.86$ (0.146)	
Superfamily Strongyloidea							
Family Chabertiidae							
<i>Oesophagostomum bifurcum</i>	Primates	<i>cox1</i>	Overall	10	$-1 \pm 1$	$-1.75$ (0.081)	de Gruijter <i>et al</i> (2002)
			Human*	6	$61 \pm 16$	$-3.32$ (0.009)	
			Mona monkey*	4	$-1 \pm 1$	0.97 (0.326)	
Superfamily Trichostrongyloidea							
Family Heligmosomatidae							
<i>Longistriata caudabullata</i>	Shrews	<i>nad4</i>	Overall	28	$144 \pm 5$	$-20.76$ ( $< 0.001$ )	Brant and Ortí (2003b)
			Brevicauda*	13	$54 \pm 6$	$-5.20$ (0.007)	
			Hylophaga*	110	$190 \pm 8$	$-5.39$ (0.002)	

Table 1 Continued

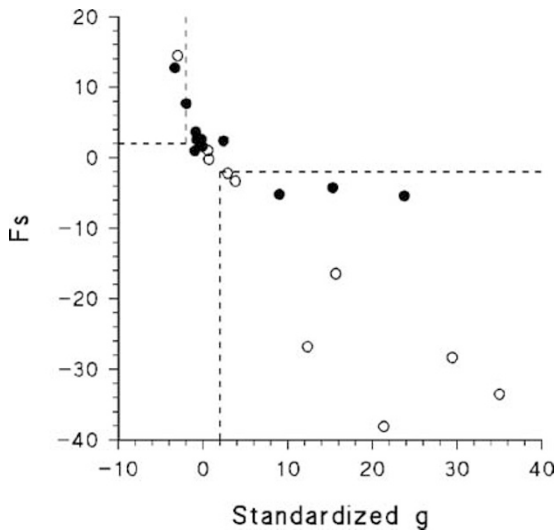
Classification	Host	Gene	Population	n	g ( $\pm$ SE)	Fs (P)	Source of data
Family Dictyocaulidae							
<i>Dictyocaulus viviparus</i>	Domestic cattle	<i>cox1</i>	Overall	252	13 $\pm$ 19	-0.22 (0.544)	Hu <i>et al</i> (2002b)
			Farm1	29	-99 $\pm$ 99	1.13 (>0.999)	
			Farm2	10	-44 $\pm$ 29	2.20 (0.887)	
			Farm3	8	-35 $\pm$ 19	4.78 (0.982)	
			Farm4	9	-37 $\pm$ 78	2.08 (0.885)	
			Farm9	39	-54 $\pm$ 26	5.91 (0.990)	
			Farm10	10	-39 $\pm$ 19	4.71 (0.986)	
			Farm14	5	-16 $\pm$ 22	3.02 (0.905)	
			Farm15	34	-53 $\pm$ 36	3.50 (0.959)	
Family Cooperiidae							
<i>Cooperia oncophora</i>	Domestic cattle	<i>atp6</i>	Overall	50	161 $\pm$ 10	-26.87 (<0.001)	van der Veer (2003)
		<i>cox3</i>	Overall	50	130 $\pm$ 7	-33.35 (<0.001)	
		<i>nad4</i>	Overall	50	115 $\pm$ 3	-20.46 (<0.001)	
		<i>nad5</i>	Overall	50	363 $\pm$ 7	-27.84 (<0.001)	
		<i>rrnS</i>	Overall	50	289 $\pm$ 13	-33.21 (<0.001)	
Family Haemonchidae							
<i>Haemonchus contortus</i>	Domestic sheep	<i>nad4</i>	Overall	50	37 $\pm$ 3	-26.81 (<0.001)	Blouin <i>et al</i> (1998)
			OH	10	58 $\pm$ 6	-2.68 (0.053)	
			TN	10	32 $\pm$ 6	-2.97 (0.027)	
			VA	10	120 $\pm$ 8	-3.08 (0.025)	
			WY	10	49 $\pm$ 11	-1.72 (0.137)	
			TX	10	-3 $\pm$ 18	1.22 (0.724)	
<i>Haemonchus placei</i>	Domestic cattle	<i>nad4</i>	Overall	40	47 $\pm$ 3	-16.49 (<0.001)	Blouin <i>et al</i> (1998)
			FL	10	47 $\pm$ 8	-2.60 (0.053)	
			GA	10	20 $\pm$ 11	-0.55 (0.354)	
			LA	10	42 $\pm$ 13	-2.01 (0.122)	
			NC	10	38 $\pm$ 10	-1.69 (0.156)	
<i>Teladorsagia circumcincta</i>	Domestic sheep	<i>nad4</i>	Overall	40	140 $\pm$ 4	-33.53 (<0.001)	Blouin <i>et al</i> (1998)
			TN	10	119 $\pm$ 15	-3.33 (0.022)	
			VA	10	113 $\pm$ 19	-3.16 (0.031)	
			WY	10	118 $\pm$ 9	-3.13 (0.032)	
<i>Teladorsagia boreoarcticus</i>	Muskox	<i>nad4</i>	Overall	11	153 $\pm$ 10	-4.24 (0.015)	Hoberg <i>et al</i> (1999)
			Island	5	119 $\pm$ 7	-2.18 (0.071)	
			Mainland	6	380 $\pm$ 55	-1.45 (0.155)	
<i>Mazamastrongylus odocoilei</i>	Deer	<i>nad4</i>	Overall	51	64 $\pm$ 3	-38.10 (<0.001)	Blouin <i>et al</i> (1998)
			SC1	11	65 $\pm$ 12	-4.02 (0.024)	
			SC2	10	25 $\pm$ 12	-3.76 (0.015)	
			FL1	9	39 $\pm$ 10	-1.49 (0.135)	
			FL2	8	52 $\pm$ 7	-1.26 (0.134)	
			AR+	13	91 $\pm$ 12	-3.14 (0.044)	

The classification is based on the phylogenetic classification of De Ley and Blaxter (2002), except that they place the Strongylida and its contents one level lower in the taxonomic hierarchy. The overall population refers to the pooled sample for the entire species, whereas named subpopulations are local geographical areas, except those marked with an asterisk, which indicate different hosts. The genes partially sequenced for each study are: *atp6* = adenosine triphosphate synthase subunit 6; *cox1* = cytochrome *c* oxidase subunit 1; *cox3* = cytochrome *c* oxidase subunit 3; *nad4* = nicotinamide dehydrogenase subunit 4; *nad5* = nicotinamide dehydrogenase subunit 5; *rrnL* = ribosomal RNA large subunit; *rrnS* = ribosomal RNA small subunit. The *g*-values range from negative (population contraction) through zero (stable population) to positive (population growth), while the range of the *Fs* values is the inverse of this pattern.

predicted pattern or to false detection of a nonexistent pattern. For example, the demographic expansion might follow linear growth, logistic growth or instantaneous growth rather than the exponential growth that was tested here, and therefore go undetected (Beaumont, 1999). Alternatively, the genetic markers used might not be neutral or nearly neutral but instead be subject to some form of selection, which can produce patterns indistinguishable from those produced by demographic changes. Furthermore, it is likely that much of the data analysed here do not fit the infinite-sites mutational model on which the *Fs* analyses (but not the Fluctuate analyses) are based, although Mes (2003) demonstrated considerable robustness to this assumption for the nematode data that he analysed. It is for these reasons

that the general congruence of the *Fs* and Fluctuate analyses is important, as they provide relatively independent evidence for population growth (ie they are based on different model assumptions). Moreover, the strong concordance of the observed patterns across the diverse range of taxa militates against alternative *ad hoc* explanations.

Furthermore, it is important to note that because the experiment performed here is basically a descriptive one, there is little control over potentially confounding factors concerning those data appearing in the analysis. The most notable confounding factor in this particular instance is the choice of mitochondrial gene(s) used in the various source studies. Most of those studies that produced evidence for demographic expansions in the



**Figure 1** Relationship between the weighted average of the  $F_s$  values for the different loci and the weighted average of the standardized  $g$ -values (ie  $g/SE$ ). Each symbol represents one host-parasite combination, with open symbols for human-associated species and filled symbols for the other species. The dashed lines represent the boundary of the values considered here to represent agreement between the  $F_s$  and  $g$  analyses regarding demographic expansion (to the lower right) and contraction (to the upper left).

parasites of human-associated hosts were based on the *nad4* gene, while most of the studies that produced no evidence for demographic expansions in the parasites of non-human-associated hosts were based on the *cox1* gene (Table 1). It is therefore entirely conceivable that the overall pattern observed here is nothing more than a by-product of the differential usefulness of these two genes for detecting population growth. However, while these two genes are usually considered to have different degrees of evolutionary conservation (Blouin, 2000b), there is currently no evidence that this difference affects the detection of demographic growth. So, population growth remains the simplest (ie most parsimonious) explanation for the departures from neutrality observed here, because it provides a single explanation (predicted *a priori*) for most of the observed patterns.

Finally, we note that failure to use a suitable nucleotide-substitution model when calculating the maximum likelihood estimates of the population growth rate based on coalescent simulations (ie the Fluctuate analysis) results in useless estimates for these mitochondrial data. In particular, use of the default program settings is inappropriate, because these settings do not yield a realistic model, and therefore the estimated values are very different from those found here – they are much larger, apparently exaggerating the propensity of the method to overestimate the population growth rate. As previously reported (eg for nematodes: Blouin *et al*, 1998; Blouin, 2000a; Mes, 2003), mitochondrial genes can have a large AT bias, a high transition:transversion ratio and most of the sites can be invariable. Therefore, it is essential to estimate simultaneously appropriate values for these parameters in the model independently of Fluctuate (as we did here), because Fluctuate does not do this itself – Fluctuate will then produce more conservative and more realistic demographic values. This issue

can also be problematic when trying to combine the data from different genes, as the parameter estimates are different for different genes and thus any attempt to pool the data results in a poor compromise for the combined estimates. The results of studies that do not report attempts to provide realistic estimates of the model parameters should therefore be interpreted with care, especially if the analyses produce very large values of  $g$ .

#### Other factors affecting population expansions

There is little new to be said about the individual results relating to those parasites with expanding populations and human-associated hosts and those with relatively stable populations and non-human-associated hosts, since this is the pattern predicted by the theory. Instead, it is more relevant to further examine some of those nematode species that do not fit this predicted general pattern, in order to assess the relative effect of other potential influences on parasite demography.

*Dictyocaulus viviparus* is the lungworm species in domesticated cattle worldwide, and in the analysis performed here all of the farms sampled from Sweden show no evidence at all of recent demographic expansion (Table 1). This result is in accord with other population data for this nematode species in Sweden, as strong population structure has been confirmed using AFLP data (Höglund *et al*, 2004), with most of the nematode genetic variation being between farms rather than within farms. It is unclear whether this result is a general one for this species in Europe, or indeed elsewhere. However, it is in marked contrast to the situation encountered for other nematode species of domestic ruminant hosts, where most of the variation is observed to be within farms rather than between farms both in North America (Blouin *et al*, 1995) and Europe (Leignel and Humbert, 2001). Clearly, a more detailed comparison of lungworms with related nematodes is warranted, especially a comparison of their demographic structure on different continents. Furthermore, parasitologists will need to be careful when developing vaccines or anthelmintics, because the strong population structure means that results from one location should not be directly extrapolated to another (cf Le Jambre, 1993).

Neither of the two *Ancylostoma* (hookworm) species showed consistent evidence of recent changes in population size (Table 1), in spite of the fact that they were sampled from different locations and that the other hookworm species (*Necator americanus*) did do so. This result is particularly unexpected for the nematode of the dog host, which actually shows a small but significant population contraction (Table 1 and Figure 1), since this host is a recent introduction into Australia (where those particular samples were taken) and has therefore itself recently had a considerable population expansion on that continent. Several possible explanations for this situation present themselves. First, this result could simply be related to the possible confounding effect of using the *cox1* gene, as noted above. This possibility can only be tested by examining other mitochondrial genes. Alternatively, there could be something unusual about the host–parasite relationship for dogs in Australia.

As a heuristic test of this second idea, we analysed the mitochondrial *rrnS* data of Skerrat *et al* (2002) for the ectoparasitic mite *Sarcoptes scabiei* (Acari) in Australia,

using Fluctuate. While the sample sizes are not large, there is moderate evidence of population expansion within the parasite species as a whole ( $g \pm SE = 12 \pm 3$ ), but there is no convincing evidence for any of the three host/parasite combinations individually: humans ( $18 \pm 10$ ), dogs ( $-9 \pm 10$ ) and wombats ( $-4 \pm 17$ ). It is likely in this case that any population expansion that exists is related as much to switches between the three host species as to expansion within any one host. We thus cannot reject the hypothesis that population expansions are unusual among dog parasites in Australia; further tests of this hypothesis may therefore be worth pursuing. One possible explanation is that the 200 years since European-domesticated dogs were introduced into Australia, and apparently much less time for the parasites (Walton *et al.*, 2004), are not long enough to leave a genetic footprint of population expansion in the parasites – not enough is yet known about the demography of these parasites to address this possibility quantitatively.

In contrast, *T. boreoarcticus* is an abomasal parasite of wild muskoxen, while some of its close relatives are parasites of ruminant livestock. It might thus be predicted (from the hypothesis being tested here) that these relatives would show recent demographic expansions while *T. boreoarcticus* would not. In fact, all of these particular nematode species show population expansions in North America (where these samples were taken) (Table 1). This may simply indicate that the muskoxen sampled in northern Canada have been human-associated to a greater extent than previously thought. Alternatively, Hoberg *et al.* (1999) attribute the current population structure of *T. boreoarcticus* in North America to muskox migration patterns since the last glaciation. If this is so, then the results of Lessa *et al.* (2003) predict that *T. boreoarcticus* should, in fact, also show a demographic expansion reflecting population growth as a result of its arctic and subarctic distribution in North America. Our results are therefore likely to have been caused by two different mechanisms operating at somewhat different temporal scales: nematode population growth due to the range expansion of wild host species during the late Quaternary and nematode population growth due to the expansion of domesticated host species during the more recent European occupation of North America.

*Longistriata caudabullata* is an intestinal parasite of several species of wild short-tailed shrew in North America, and it is therefore not explicitly predicted to show a demographic expansion although it does so in each of the two host species sampled here (Table 1). Brant and Ortí (2003b) note that there appears to be considerable gene flow within this nematode species across both host species, indicating that it forms one panmictic population. Thus, it might be true in this case, as well, that any population expansion that exists is related as much to switches between the two host species as to expansion within any one host.

However, in this particular case, it is possible to test directly the congruence between the parasite biology and host biology that is predicted by our hypothesis. This is because Brant and Ortí (2003b) also sequenced the mitochondrial control region (or D loop) of the host individuals from which the parasites were collected. We thus analysed the data for the two host species using

Fluctuate. There is considerable evidence of population expansion within each of the host species: *Blarina brevicaudata*,  $g \pm SE = 40 \pm 11$ ; *Blarina hylophaga*,  $g \pm SE = 45 \pm 8$ . Thus, for this characteristic, the parasite biology does appear to be tracking that of the hosts, providing an additional test of our hypothesis. Brant and Ortí (2003a) attribute the population expansion in *B. brevicaudata*, at least, to range expansion following the last glaciation.

Alternatively, it is possible that most trichostrongylid nematodes have undergone recent demographic expansions irrespective of their host status – for the data analysed here, of the taxa in the superfamily Trichostrongyloidea only *D. viviparus* fails to show evidence of population growth (see above). Indeed, a  $2 \times 2$  contingency statistical test comparing the frequency of population expansions in this superfamily to that of the other superfamilies yields  $G = 12.82$ ,  $P < 0.001$ . However, this comparison is confounded by the fact that most of the human-associated taxa are in this superfamily, with most of the remainder in the superfamily Ancylostomatoidea, as well as being an *a posteriori* hypothesis test rather than an *a priori* one (ie the hypothesis arises after examining the data and so the data cannot be a valid test of that hypothesis). Furthermore, no one has yet suggested a model that would explain why demographic expansions would be associated with particular taxonomic groups, although the taxonomic distribution of life-cycle characteristics (eg presence or absence of intermediate hosts, mode of reproduction) may play a part (Viney, 1998; Blouin *et al.*, 1999; Criscione and Blouin, 2004). Nevertheless, should this alternative idea turn out to be correct, then the data may be providing evidence that *D. viviparus* is currently placed in an inappropriate taxonomic group, a suggestion that has already been made based on phylogenetic analyses (Höglund *et al.*, 2003).

Overall, it seems likely that the situation is more complicated than our simple hypothesis test suggests. In particular, other causes of recent population expansions in the hosts clearly can confound the effects of human association. Also, the existence of direct (single host) and indirect (alternate hosts, with different life-history stages in different types of host) life cycles may have large effects, particularly if one of the hosts of a multihost parasite is human-associated while the other is not. This situation will become even more complex for taxa like the platyhelminths, for example, where different parasitic species can have one, two or three alternating hosts – in the latter cases, rarely is more than one of the hosts human-associated. Testing the hypothesis for platyhelminths, incidentally, has another difficulty as there is currently a dearth of data for non-human-associated host species – almost all of the data sets currently available are for parasites of humans, livestock or salmonid fishes.

The results observed here also show some consistent indications of population contractions rather than expansions among nematodes from non-human-associated hosts, particularly among the *Capillaria*, *Contracaecum* and *Heterorhabditis* species (Figure 1). However, most of the rate estimates have large standard errors and so these results are not necessarily reliable (Table 1). For example, the baikal seal has shown a population decline (Reijnders *et al.*, 1993), and so its parasite species can be expected to show a decline also, but this is not clear from the data analysed here because of the large

standard error of the estimate. Nevertheless, the possibility of population decline among nematode species clearly bears looking into.

### Future prospects

Our analysis aimed to investigate the prospects for using sequence data to test general hypotheses about the evolution of parasite populations. From this perspective, the results are encouraging, and bioinformatic data-mining of the currently available information can clearly be productive. However, our study also highlights the limitations of this approach, which is basically a descriptive one in which there is little control over potentially confounding factors such as the taxonomy and life-history characteristics of those parasites appearing in the analysis (cf Blouin, 1998; Criscione and Blouin, 2004). In particular, it may not be worth pursuing this line much further for this particular hypothesis, although a similar strategy could be used for parasites in the Platyhelminthes, Acari and Insecta for example, as the scale of resolution is too poor and therefore further *ad hoc* comparisons may not yield any more-useful data. The indications provided by such analyses are promising, but they do not provide a definitive test of the hypothesis.

Further advances will therefore require more controlled comparisons of parasite species, directed at testing predictions that are more specific. For example, more critical tests of the hypothesis examined here could involve assessing: (i) the same parasite species in both human-associated and non-human-associated hosts (eg as carried out here for *O. bifurcum*); (ii) closely related pairs of parasite species in both human-associated and non-human-associated hosts (eg as carried out here for the *Teladorsagia* species); (iii) the same parasite species in the same host species both in areas where the host has been introduced and in areas where it has not (eg host species such as roe deer in different parts of Europe); (iv) the same parasite species in the same host species in areas where the host has been introduced or domesticated at different times (eg hosts in Europe *vs* those in North America or Australia); (v) different parasite species in the same host individuals; (vi) closely related parasites with different life-history characteristics (to assess alternative influences on the demographic patterns); or (vii) simultaneous sampling of both the hosts and the parasites to confirm that the population behaviour is the same in both species (eg as carried out here for *L. caudabullata* and its shrew hosts).

### Acknowledgements

We thank Jens Mattsson for comments on an earlier draft of the manuscript.

### References

Anderson TJC, Blouin MS, Beech RN (1998). Population biology of parasitic nematodes: applications of genetic markers. *Adv Parasitol* **41**: 219–283.  
Beaumont MA (1999). Detecting population expansion and decline using microsatellites. *Genetics* **153**: 2013–2029.  
Blouin MS (1998). Mitochondrial DNA diversity in nematodes. *J Helminthol* **72**: 285–289.  
Blouin MS (2000a). Neutrality tests on mtDNA: unusual results from nematodes. *J Hered* **91**: 156–158.

Blouin MS (2000b). Molecular prospecting for cryptic species of nematodes: mitochondrial DNA *versus* internal transcribed spacer. *Int J Parasitol* **32**: 527–531.  
Blouin MS, Liu J, Berry RE (1999). Life cycle variation and the genetic structure of nematode populations. *Heredity* **83**: 253–259.  
Blouin MS, Yowell CA, Courtney CH, Dame JB (1995). Host movement and the genetic structure of populations of parasitic nematodes. *Genetics* **141**: 1007–1014.  
Blouin MS, Yowell CA, Courtney CH, Dame JB (1998). Substitution bias, rapid saturation, and the use of mtDNA for nematode systematics. *Mol Biol Evol* **15**: 1719–1727.  
Brant SV, Ortí G (2003a). Phylogeography of the northern short-tailed shrew, *Blarina brevicauda* (Insectivora: Soricidae): past fragmentation and postglacial recolonization. *Mol Ecol* **12**: 1435–1449.  
Brant SV, Ortí G (2003b). Evidence for gene flow in parasitic nematodes between two host species of shrews. *Mol Ecol* **12**: 2853–2859.  
Brown BW, Brauner C, Chan A, Gutierrez D, Herson J, Lovato J *et al* (2000). *DStplan. Calculations for Sample Sizes and Related Problems*, Version 4.2. MD Anderson Cancer Center, Department of Biomathematics, University of Texas, Houston, TX.  
Cavalli-Sforza LL, Menozzi P, Piazza A (1994). *The History and Geography of Human Genes*. Princeton University Press: Princeton, NJ.  
Conway DJ, Baum J (2002). In the blood – the remarkable ancestry of *Plasmodium falciparum*. *Trends Parasitol* **18**: 351–355.  
Cox FEG (ed) (1993). *Modern Parasitology* 2nd edn Blackwell Scientific Publications: Oxford.  
Criscione CD, Blouin MS (2004). Life cycles shape parasite evolution: comparative population genetics of salmon trematodes. *Evolution* **58**: 198–202.  
de Gruijter JM, Polderman AM, Zhu XQ, Gasser RB (2002). Screening for haplotypic variability within *Oesophagostomum bifurcatum* (Nematoda) employing a single-strand conformation polymorphism approach. *Mol Cell Probes* **16**: 185–190.  
De Ley P, Blaxter M (2002). Systematic position and phylogeny. In: Lee DL (ed) *The Biology of Nematodes*. Taylor & Francis: London, pp 1–30.  
Donnelly MJ, Licht MC, Lehmann T (2001). Evidence for recent population expansion in the evolutionary history of the malaria vectors *Anopheles arabiensis* and *Anopheles gambiae*. *Mol Biol Evol* **18**: 1353–1364.  
Fu Y-X (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**: 915–925.  
Gasser RB, Chilton NB (2001). Applications of single-strand conformation polymorphism (SSCP) to taxonomy, diagnosis, population genetics and molecular evolution of parasitic nematodes. *Vet Parasitol* **101**: 201–213.  
Grant WN (1994). Genetic variation in parasitic nematodes and its implications. *Int J Parasitol* **24**: 821–830.  
Hawdon JM, Li T, Zhan B, Blouin MS (2001). Genetic structure of populations of the human hookworm, *Necator americanus*, in China. *Mol Ecol* **10**: 1433–1437.  
Hoberg EP, Monsen KJ, Kutz S, Blouin MS (1999). Structure, biodiversity, and historical biogeography of nematode faunas in Holarctic ruminants: morphological and molecular diagnoses for *Teladorsagia boreoarcticus* n.sp. (Nematoda: Ostertagiinae), a dimorphic cryptic species in muskoxen (*Ovibos moschatus*). *J Parasitol* **85**: 910–934.  
Höglund J, Engström A, Morrison DA, Mattsson JG (2004). Genetic diversity assessed by AFLP analysis of the parasitic nematode *Dictyocaulus viviparus*, the lungworm of cattle. *Int J Parasitol* **34**: 475–484.  
Höglund J, Morrison DA, Divina BP, Wilhemsson E, Mattsson JG (2003). Phylogeny of *Dictyocaulus* (lungworms) from eight species of ruminants based on analyses of ribosomal RNA data. *Parasitology* **127**: 179–187.



- Hughes AL, Verra F (2001). Very large long-term effective population size in the virulent human malaria parasite *Plasmodium falciparum*. *Proc R Soc London B* **268**: 1855–1860.
- Hu M, Abs El-Osta YG, Gasser RB (2003). Comparative analysis of mitochondrial genome data for *Necator americanus* from two endemic regions reveals substantial genetic variation. *Int J Parasitol* **33**: 955–963.
- Hu M, Chilton NB, Gasser RB (2004). The mitochondrial genomics of parasite nematodes of socio-economic importance: recent progress, and implications for population genetics and systematics. *Adv Parasitol* **56**: 134–212.
- Hu M, Chilton NB, Zhu X, Gasser RB (2002a). Single-stranded conformation polymorphism-based analysis of mitochondrial cytochrome *c* oxidase subunit 1 reveals significant substructuring in hookworm populations. *Electrophoresis* **23**: 27–34.
- Hu M, D'Amelio S, Zhu Z, Paggi L, Gasser R (2001). Mutation scanning for sequence variation in three mitochondrial DNA regions for members of the *Contraecum osculatum* (Nematoda: Ascaridoidea) complex. *Electrophoresis* **22**: 1069–1075.
- Hu M, Höglund J, Chilton NB, Zhu X, Gasser RB (2002b). Mutation scanning analysis of mitochondrial cytochrome *c* oxidase subunit 1 reveals limited gene flow among bovine lungworm subpopulations in Sweden. *Electrophoresis* **23**: 3357–3363.
- Kuhner MK, Yamato J, Felsenstein J (1998). Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics* **149**: 429–434.
- Leignel V, Cabaret J, Humbert JF (2002). New molecular evidence that *Teladorsagia circumcincta* (Nematoda: Trichostrongylidae) is a species complex. *J Parasitol* **88**: 135–140.
- Leignel V, Humbert J-F (2001). Mitochondrial DNA variation in benzimidazole-resistant and -susceptible populations of the small ruminant parasite *Teladorsagia circumcincta*. *J Hered* **92**: 503–506.
- Le Jambre LF (1993). Molecular variation in trichostrongylid nematodes from sheep and cattle. *Acta Trop* **53**: 331–343.
- Lessa EP, Cook JA, Patton JL (2003). Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. *Proc Natl Acad Sci USA* **100**: 10331–10334.
- Liu J, Berry RE, Blouin MS (1999). Molecular differentiation and phylogeny of entomopathogenic nematodes (Rhabditida: Heterorhabditidae) based on ND4 gene sequences of mitochondrial DNA. *J Parasitol* **85**: 709–715.
- McCoy KD, Boulinier T, Tirard C, Michalakakis Y (2003). Host-dependent genetic structure of parasite populations: differential dispersal of seabird tick host races. *Evolution* **57**: 288–296.
- Mes THM (2003). Demographic expansion of parasitic nematodes of livestock based on mitochondrial DNA regions that conflict with the infinite-sites model. *Mol Ecol* **12**: 1555–1566.
- Pearlman SJ, Spicer GS, Shoemaker D, Jaenike J (2003). Associations between mycophagous *Drosophila* and their *Howardula* nematode parasites: a worldwide phylogenetic shuffle. *Mol Ecol* **12**: 237–249.
- Poulin R (1998). *Evolutionary Ecology of Parasites*. Chapman & Hall: London.
- Ramos-Onsins SE, Rozas J (2002). Statistical properties of new neutrality tests against population growth. *Mol Biol Evol* **19**: 2092–2100.
- Reijnders PJH, Brasseur SMJ, van der Toorn J, van der Wolf P, Boyd I, Harwood J *et al* (1993). *Seals, Fur Seals, Sea Lions, and Walrus: Survey Status and Conservation Action Plan*. IUCN Seal Specialist Group: Gland, Switzerland.
- Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.
- Skerratt LF, Campbell NJH, Murrell A, Walton S, Kemp D, Barker SC (2002). The mitochondrial 12S gene is a suitable marker of populations of *Sarcoptes scabiei* from wombats, dogs and humans in Australia. *Parasitol Res* **88**: 376–379.
- Sokal RR, Rohlf FJ (1994). *Biometry: The Principles and Practice of Statistics in Biological Research*, 3rd edn. Freeman: San Francisco.
- Strimmer K, von Haesler A (1996). Quartet puzzling: a quartet maximum likelihood method for reconstructing tree topologies. *Mol Biol Evol* **13**: 964–969.
- van der Veer M (2003). Studies on the genetic population structure of *Cooperia oncophora*. PhD thesis, Department of Parasitology and Tropical Veterinary Medicine, Faculty of Veterinary Medicine, Utrecht University, The Netherlands.
- van der Veer M, Kanobana K, Ploeger HW, de Vries E (2003). Cytochrome oxidase *c* subunit 1 polymorphisms show significant differences in distribution between a laboratory maintained population and a field isolate of *Cooperia oncophora*. *Vet Parasitol* **116**: 231–238.
- Viney ME (1998). Nematode population genetics. *J Helminthol* **72**: 281–283.
- Walton SF, Dougall A, Pizzutto S, Holt D, Taplin D, Arlian LG *et al* (2004). Genetic epidemiology of *Sarcoptes scabiei* (Arcari: Sarcoptidae) in northern Australia. *Int J Parasitol* **34**: 839–849.
- Zhu X, Spratt DM, Beveridge I, Haycock P, Gasser RB (2000). Mitochondrial DNA polymorphism within and among species of *Capillaria sensu lato* from Australian marsupials and rodents. *Int J Parasitol* **30**: 933–938.