

## ORIGINAL ARTICLE

Bayesian paternity analysis and mating patterns in a parasitic nematode, *Trichostrongylus tenuis*PCD Johnson<sup>1,2</sup>, JD Hadfield<sup>3</sup>, LMI Webster<sup>1,4</sup>, A Adam<sup>1</sup>, BK Mable<sup>1</sup> and LF Keller<sup>1,5</sup><sup>1</sup>Division of Environmental and Evolutionary Biology, Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow, UK;<sup>2</sup>Robertson Centre for Biostatistics, Faculty of Medicine, University of Glasgow, Glasgow, UK; <sup>3</sup>Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, UK; <sup>4</sup>Institute of Biological and Environmental Sciences, School of Biological Sciences, University of Aberdeen, Aberdeen, UK and <sup>5</sup>Zoological Museum, University of Zürich, Zürich, Switzerland

Mating behaviour is a fundamental aspect of the evolutionary ecology of sexually reproducing species, but one that has been under-researched in parasitic nematodes. We analysed mating behaviour in the parasitic nematode *Trichostrongylus tenuis* by performing a paternity analysis in a population from a single red grouse host. Paternity of the 150 larval offspring of 25 mothers (sampled from one of the two host caeca) was assigned among 294 candidate fathers (sampled from both caeca). Each candidate father's probability of paternity of each offspring was estimated from 10-locus microsatellite genotypes. Seventy-six (51%) offspring were assigned a father with a probability of >0.8, and the estimated number of unsampled males was 136 (95% credible interval (CI)

77–219). The probability of a male from one caecum fathering an offspring in the other caecum was estimated as 0.024 (95% CI 0.003–0.077), indicating that the junction of the caeca is a strong barrier to dispersal. Levels of promiscuity (defined as the probability of two of an adult's offspring sharing only one parent) were high for both sexes. Variance in male reproductive success was moderately high, possibly because of a combination of random mating and high variance in post-copulatory reproductive success. These results provide the first data on individual mating behaviour among parasitic nematodes.

*Heredity* (2010) **104**, 573–582; doi:10.1038/hdy.2009.139; published online 14 October 2009

**Keywords:** parasitic nematodes; *Trichostrongylus tenuis*; paternity analysis; pedigree reconstruction; microsatellites; mating system

## Introduction

Animal parasitic nematodes are among the most widespread and diverse animal groups (Blaxter, 2001). Their hidden lifestyle, however, presents difficulties in observing their behaviour in the wild, and many areas of their ecology remain unexplored (Criscione *et al.*, 2005). One area in which progress has been made is in the application of molecular ecology methods to the inference of large-scale host-mediated ecological processes, such as colonization and dispersal (Criscione *et al.*, 2005; Nejsun *et al.*, 2005; Nieberding *et al.*, 2005; Troell *et al.*, 2006; Grillo *et al.*, 2007; Webster *et al.*, 2007; Wielgoss *et al.*, 2008). However, no studies have yet used molecular tools to analyse the patterns of parasite mating behaviour at the individual level. Mating behaviour has important consequences for understanding how this group of animals has evolved as well as for predicting how parasite populations will respond to changes in their environment. For example, deviation from random mating in the form of reproductive skew reduces the

effective population size and increases the opportunity for sexual selection (Crow, 1958).

In the absence of specific data, parasitic nematodes have generally been assumed to mate randomly and promiscuously (for example, Dobson *et al.*, 1987; Barnes *et al.*, 1995; Saul, 1995; Smith *et al.*, 1999; Churcher and Basáñez, 2008, but for exceptions see Braselton *et al.*, 2004; Churcher *et al.*, 2008). In nonparasitic taxa, predictions about mating systems and the distribution of reproductive success have been successfully tested through parentage analysis using molecular markers. Indeed, molecular markers have transformed the study of animal mating systems, often overturning conclusions based on observational evidence (for example, Jones and Avise, 2001; Griffith *et al.*, 2002), and would seem to be ideal tools for studying the concealed mating behaviour of animals such as parasitic nematodes. Paternity analysis (the most common type of parentage analysis) is conventionally performed by analysing the genotypes of a sample of mothers, their known offspring and, ideally, all plausible candidate fathers. When applied to a parasitic nematode, paternity analysis presents particular technical and statistical challenges, including raising offspring *in vitro*, obtaining DNA of sufficient quantity and quality from microscopic larvae and accurately assigning paternities among potentially hundreds or thousands of candidate fathers, some of whom might already have been shed from the host. These difficulties

Correspondence: Dr PCD Johnson, Robertson Centre for Biostatistics, Boyd Orr Building, University of Glasgow, Glasgow G12 8QQ, UK.

E-mail: paulj@stats.gla.ac.uk

Received 27 February 2009; revised 23 June 2009; accepted 7 August 2009; published online 14 October 2009

may explain the current lack of parentage analyses in parasitic nematodes.

We conducted a paternity analysis in a single host-population (or infrapopulation) of *Trichostrongylus tenuis*, a parasitic nematode of galliform and anseriform birds. Our aim was to provide the first data on patterns of parentage in a sexually reproducing animal parasitic nematode, and to shed light on the assumption that mating in these parasites is random.

*T. tenuis* has a direct life cycle and obligate sexual reproduction. Adult *T. tenuis* reside and mate in the two caeca of the host; these are blind guts of about 75 cm in length that extend from the junction of the small and large intestines (Hudson, 1992). Candidate fathers for a *T. tenuis* offspring fall into three categories. The first and most obvious category is of males present at the time of sampling in the caecum in which the mothers were sampled. Second, males from the neighbouring caecum must be considered, although we know nothing about the movement of adults between caeca. The third category is of unsampled males, including candidate fathers shed before sampling and those present in the caeca that had escaped detection. Information about the number of unsampled males is essential for paternity assignment, which in *T. tenuis* will depend on both the longevity of males and the duration of sperm storage. *T. tenuis* infecting naive captive hosts can live for more than 2 years (Shaw and Moss, 1989), and most sexually reproducing nematodes are thought to store sperm (Bird and Bird, 1991), but we know nothing about these parameters in wild populations. Given perfect genotypic data, the number of unsampled males could be estimated relatively easily by partially inferring the genotypes of missing fathers from the genotypes of mothers and offspring. In practice, the presence of data errors makes this challenging (Emery *et al.*, 2001).

In this study we jointly estimated paternity, the number of unsampled males and the probability of mating across caeca using a Bayesian method tolerant of genotyping error (Hadfield *et al.*, 2006). Our aim was to analyse four questions relevant to mating and reproduction in *T. tenuis*: (1) How are paternities distributed among sampled and unsampled males? (2) Among sampled males, what proportion of fathers was sampled in the same caecum as the mother? (3) How promiscuous are males and females? (4) Is reproductive success randomly distributed among males?

## Materials and methods

### Sampling of study population

A single male red grouse was harvested at a grouse-shooting estate in Lauderdale, Scotland at 1200 h on 29

October 2004 (as the timing of events, such as the death of the host and the isolation of the females, influences the interpretation of the results, we give the local time (British Summer Time) at each stage). The two caeca were removed (1230 h) and stored in M9 buffer (Hope, 1999) at approximately 30 °C during transport to the laboratory. On arrival (1500 h), the diluted contents and the mucosal and submucosal surface of one caecum (designated the *local* caecum, as distinct from the *neighbouring* caecum) were examined for *T. tenuis* adults under a dissecting microscope. The neighbouring caecum was stored at –20 °C. All 108 female and 122 male adult *T. tenuis* found in the local caecum were removed. Male *T. tenuis* are easily distinguished from females by their shorter length and the presence of bursate claspers. In all, 51 live females were isolated in 250 µl M9 buffer in separate wells of a covered transparent polystyrene 96-well flat-bottom microtitre plate (Greiner Bio-One, Stonehouse, UK) at 25 °C (1730 h). After 2.5 h, all 51 females were transferred to a fresh microtitre plate containing M9 (2000 h), allowing eggs and larvae to be categorized as laid early (1730–2000 h) or late (after 2000 h). Both plates were incubated at 25 °C for 48 h (by which time egg laying had ceased) and stored overnight at 4 °C. The number of eggs laid was 987, of which 154 (16%) hatched and moulted to stage 3 infective larvae (Table 1). Approximately equal numbers of eggs were laid in each time period, but early eggs were four times more likely to develop to stage 3 infective larvae than late eggs. A number of factors could explain the low hatch rate, including shedding of unfertilized eggs and mortality of fertilized eggs *in vitro*. Similarly, the decline in hatch rate over time could be related to either the worsening condition (of the mother or of stored sperm) or a dwindling supply of stored sperm. There was a substantial variation in fecundity among females, whether measured in terms of the numbers of eggs (mean 19.3, s.d. 13.1, range 0–45) or larvae (mean 3.0, s.d. 4.1, range 0–16). The 29 mothers that had produced at least one larva and all of their 154 larval offspring were then removed from the microtitre plate for immediate DNA extraction.

The neighbouring caecum was searched for adult *T. tenuis*, as described for the local caecum, leading to the recovery of 183 male and 200 female *T. tenuis*. In total, 305 males, 122 and 183 from the local and neighbouring caeca, respectively, were preserved for genotyping in 95% ethanol at 4 °C.

### Microsatellite genotyping

All females, males and offspring were genotyped at 10 microsatellite loci, including replicate genotyping of DNA-poor samples likely to incur genotyping errors, as

**Table 1** The numbers of eggs laid and larvae hatched by 51 female *Trichostrongylus tenuis* within 2.5 h of isolation (early) and in the subsequent 48 h (late)

Laying period	n (%), females laying eggs	n, eggs laid	n (%), females producing larvae	n, larvae hatched	Mean (s.d.) hatch rate <sup>a</sup>
Early	38 (74.5%)	517	22 (43.1%)	122	0.20 (0.24)
Late	43 (84.3%)	470	12 (23.5%)	32	0.06 (0.13)
Combined	43 (84.3%)	987	29 (56.9%)	154	0.15 (0.19)

<sup>a</sup>Mean (s.d.) hatch rate was calculated using only females that laid eggs in both laying periods ( $n = 38$ ), and differed significantly between periods ( $P = 0.001$ , paired  $t$ -test).

**Table 2** Genetic diversity at the 10 microsatellite marker loci assayed

Locus	n	N <sub>a</sub>	H <sub>e</sub>	PIC
Tte002	453	14	0.72	0.68
Tte003	440	21	0.77	0.74
Tte057	428	18	0.78	0.76
Tte102	416	9	0.78	0.75
Tte134	437	15	0.85	0.83
Tte211	451	12	0.69	0.65
Tte254	420	19	0.80	0.77
Tte303	439	11	0.63	0.60
Tte335	458	9	0.75	0.72
Tte365	439	10	0.60	0.53
Mean	438	14	0.74	0.70

Abbreviations: H<sub>e</sub>, expected heterozygosity (Nei, 1978); n, number of individuals typed; N<sub>a</sub>, number of alleles observed; PIC, polymorphic information content (Botstein *et al.*, 1980).

follows. Template DNA for PCR was prepared from the mothers, offspring and candidate fathers using the worm lysis method (Williams *et al.*, 1992; Grillo *et al.*, 2006). Males and larvae were lysed whole, whereas only the heads (the anterior 10–20%) of females were used, to minimize the risk of contamination from sperm or fertilized eggs (Anderson *et al.*, 2003). Live larvae were ex-sheathed under a dissecting microscope in 0.2% sodium hypochlorite solution. Females, males and larvae were rinsed in double-distilled water and then placed into either 10 µl (larvae and females) or 20 µl (males) of lysis buffer (50 mM potassium chloride, 10 mM Tris pH 8.0, 2.5 mM magnesium chloride, 0.45% Nonidet P-40, 0.45% Tween-20, 0.01% gelatine and 120 µg/ml proteinase K) in 96-well PCR plates. The plates were frozen at –80 °C for 10 min to lyse the tissues, incubated overnight at 65 °C and then heated to 95 °C for 15 min. The resulting 488 lysates were diluted with double-distilled water by a factor of five (females and larvae) or 10 (males) and stored at –20 °C.

The *T. tenuis* females, larvae and males were genotyped using 10 autosomal microsatellite markers (Table 2), as described by Johnson *et al.* (2006), with the exception that primers were redesigned for two loci. For Tte303 the reverse primer was 5'-ACGTTCCCTGGCCTA AATAC-3' and for Tte365 the primers were 5'-GGTGTC TTTTGGCGTGTAGTG-3' (forward) and 5'-GATCGTCA GCAGCCTCG-3' (reverse). Microsatellite genotyping was attempted for all 488 lysates. The rates of PCR failure and genotyping error were high in the genotypes from females and larvae but low in males, possibly because of the differences in relative tissue quantities available (males are around 5 mm long, compared with 1 mm for the heads of females and 0.5 mm for larvae). An alternative method of nematode lysis (Floyd *et al.*, 2002) was optimized but did not reduce error rates. Therefore, to reduce the number of missing genotypes and allow accurate estimation of error rates, all females and larvae were genotyped thrice, and 68 (23%) males were genotyped twice. Allele lengths were measured using an ABI3730 DNA analyzer (Applied Biosystems, Foster City, CA, USA) by the Sequencing Service (University of Dundee, Dundee, UK) and analysed using GeneMapper 4.0 (Applied Biosystems).

## Paternity analysis

The approach to paternity assignment followed in this study is fractional allocation of paternity (Jones and Ardren, 2003). For each offspring a probability of paternity is estimated for every candidate male, which, together with the probability that the father was not sampled, must sum to one. This approach is contrasted with categorical allocation, most widely implemented in the software CERVUS (Marshall *et al.*, 1998), in which the paternity is assigned to the single most likely candidate that passes a likelihood threshold. Categorical allocation is useful when a single 'accepted' pedigree is required, but because it ignores the uncertainty in estimating the pedigree, it results in optimistic assessment of precision and biased parameter estimates (Jones and Ardren, 2003; Hadfield *et al.*, 2006). In contrast, fractional methods aim to avoid this systematic bias and optimism by effectively averaging over all possible pedigrees in proportion to their probability. The method used in this study further reduces bias by simultaneously estimating individual paternity probabilities with the overall effect of the caecum of origin on the probability of paternity (Hadfield *et al.*, 2006).

For each offspring, we estimated the probability of paternity by each of 294 candidate fathers (the 305 sampled males minus 11 from the neighbouring caecum that provided no genotypic data) as well as the probability that the true father was not sampled. For the purpose of illustration, the informativeness of these 295 probabilities can be intuitively summarized by the maximum posterior probability. A high maximum probability of paternity (for example, >0.9) implies that a single male (or the unsampled males) has been 'assigned' paternity of the offspring with high confidence. A low maximum probability of paternity (for example, <0.5) implies that no single male is a strong candidate.

Paternity probabilities, the number of unsampled males (Koch *et al.*, 2008) and the probability of paternity from the neighbouring caecum were estimated from the microsatellite genotypes (including all replicate genotypes) by a Bayesian Markov chain Monte Carlo approach implemented in MASTERBAYES (available at <http://www.R-project.org>) according to Hadfield *et al.* (2006).

MASTERBAYES estimates the total number of unsampled males, not unsampled fathers. The data only contain information about unsampled fathers, but we can extrapolate to the number of unsampled males by assuming that the proportion of unsampled males with no offspring is the same as the proportion observed among the sampled males. The probability of paternity from the neighbouring caecum is the probability that a male from the neighbouring caecum will gain paternity ahead of an otherwise identical male from the local caecum, and effectively quantifies the degree to which free interbreeding is restricted by the subdivision of the population into two caeca. A probability of 0.5 would suggest that there is no barrier to mating between the caeca, whereas a probability close to zero would indicate that mating between *T. tenuis* from opposite caeca is very unlikely. For a precise definition of these two parameters and how they are estimated, see Hadfield *et al.* (2006).

A particularly difficult problem for paternity assignment can occur when the true father has not been

sampled and has close relatives among the sampled males. MASTERBAYES could potentially be led astray by this problem, although it is robust to the presence of related males among the sampled males. To assess the frequency of close relationships (for example, parent–offspring pairs, full-sibs and half-sibs) among the sampled males, we estimated their pairwise relatedness ( $r$ ), according to Lynch and Ritland (1999), using the software GENALEX (Peakall and Smouse, 2006).

Hadfield *et al.* (2006) account for stochastic genotyping errors by modelling two separate classes of error: allelic dropouts (in which one allele fails to amplify, causing a heterozygote to be recorded as a homozygote) and false alleles (any other stochastic genotyping error; Wang, 2004). Allelic dropout and false allele genotyping error rates were estimated according to Hadfield *et al.* (2006) with the exception that, rather than assuming that genotyping errors are independently distributed across all loci and individuals, we estimated two separate pairs of error rates, one for males and the other for females and larvae. Two categories of error rate were required because of the wide disparity in error rates observed between these two groups. The prior distributions of the four error rates were uniform in the range of 0–1.

For the paternity analysis, the Markov chain was run for 250 million iterations with a thinning rate of 225 000 and burn-in of 25 million. The prior distribution of the probability of paternity across sampled and unsampled males was left unspecified. Estimation of the probability of paternity from the neighbouring caecum was based on a model with a normally distributed prior on the logit scale with a mean of zero and a variance of 3.06. This is the closest logit parameterisation to a uniform prior on the probability scale, although it differs from the uniform distribution in having very low density for extreme probabilities. In consequence, probabilities outside the range of 0.01–0.99 are heavily penalized and probabilities of zero or one are impossible.

The distributions of further parameters of interest (see below) were estimated directly from 1000 Markov Chain Monte Carlo samples from the posterior distribution of the pedigree, and summarized by the median and the range between the quantiles of 2.5 and 97.5% (referred to as a 95% credible interval or 95% CI).

### Reproductive promiscuity

Promiscuous mating systems are those in which pair bonds are not formed. Promiscuous mating can be easily diagnosed from a pedigree. However, it would be more useful to be able to quantify the degree of promiscuity. (Because we are unable to observe mating directly, but only the pedigree, in this study we are referring to ‘reproductive’ promiscuity or the tendency to produce offspring by different mates.) An intuitive measure of reproductive promiscuity would be the number of mates that contributed to a mother’s brood (for simplicity, we describe only female promiscuity, although the same arguments apply to males). This measure is unsatisfactory because it is related to brood size: a mother might seem more promiscuous simply because we sampled a larger number of her offspring. We propose a measure of reproductive promiscuity that can be estimated from pedigree data and that is not biased by brood size; the probability,  $p$ , that two siblings chosen at random from

a mother’s offspring have different parents, or are half-sibs. We can treat the observed set of genotyped offspring as a random sample from a larger population consisting of all offspring that could have been produced by the mother during the sampling period. A sample of size  $n$  from this population of offspring consists of  $k$  full sibships (by  $k$  fathers), each sibship having frequency (proportion)  $x_i$ ,  $i = 1, 2, \dots, k$ . This is analogous to a classic scenario in population genetics, in which a sample of  $n$  chromosomes consisting of  $k$  different alleles at frequencies  $x_i$  is drawn from a larger gene pool. The probability,  $p$ , that any two sibs have different fathers is analogous to the expected heterozygosity of a gene locus, which is the probability that two randomly sampled alleles are different. We, estimate  $p$  as

$$\hat{p} = \frac{n}{n-1} \left( 1 - \sum x_i^2 \right)$$

(Nei and Roychoudhury, 1974). This estimate is unbiased at any  $n$ , which is important as broods could be as small as  $n = 2$ . Unless broods are large, individual estimates will be very variable, but averaging over a large pedigree will give a reliable pedigree-wide estimate of reproductive promiscuity.

### Distribution of male reproductive success

The observed distribution of paternities among males was compared with the distribution expected under a null hypothesis of random allocation. Treating paternities as random events, the number of paternities allocated to a male under the null hypothesis is a Poisson random variable with parameter  $\lambda$ , in which  $\lambda$  is the average number of paternities per male. Deviation from random mating was estimated by the standardized variance in reproductive success,  $I$ , which is the variance in reproductive success divided by the mean (Boness *et al.*, 1993), and is equivalent to the index of dispersion (Krebs, 1989). A wide variety of statistics is available to quantify variance in reproductive success (Kokko *et al.*, 1999). For the purpose of quantifying deviation from randomly allocated reproductive success,  $I$  is a natural choice because of its relationship with the Poisson distribution. Estimates of  $I$  in males can be related to a continuum from strict monogamy ( $I = 0$ ) to random allocation ( $I = 1$ , that is, the mean equals the variance, as expected under the null hypothesis of Poisson-distributed reproductive success) to high variance in reproductive success ( $I \gg 1$ ). Because  $I$  is sensitive to mean reproductive success, to allow comparison with other studies, we also estimated the standardized variance in reproductive success as the variance divided by the square of the mean,  $I_s$ , which is also known as an index of the opportunity for sexual selection (Crow, 1958; Wade and Arnold, 1980).

As variation in male fecundity (the number of offspring resulting from a single copulation) will produce a nonrandom distribution of paternities even when copulation is random, we also tested the allocation of copulations for deviation from randomness as described above for paternities. To infer the number of copulations, it was necessary to make the assumption that all multiple paternities by a given male with a given female arose from a single copulation. This assumption seems reason-

able for two reasons. First, all nematodes are thought to have the capacity to store sperm (Bird and Bird, 1991), even if the longevity of stored sperm is unknown in *T. tenuis*, which increases the potential for multiple fertilizations from a single copulation. Second, at 613 adults, the population within the sampled host was large enough that females were likely to encounter many potential mates.

We did not assess variation in female reproductive success as this is likely to reflect variation in egg and larva viability *in vitro*.

Except where otherwise stated, data analyses were performed in R (R Development Core Team, 2008). The analysis data set is available as supplementary information on the Heredity website (<http://www.nature.com/hdy>).

## Results

### Genotyping

Of the 488 adults and larvae genotyped, 469 (96%) yielded genotypic data from at least one locus (Table 3). No genotypes were recovered from four larvae and 11 of the males sampled from the neighbouring caecum. Each of the four larvae was the sole offspring of its mother, and hence, in effect, four mothers were also lost from the paternity analysis, although their genotypes nevertheless contributed to the estimation of the population allele frequencies and genotyping error rates. Thus, the main task of the analysis was to allocate the paternity of 150 offspring of 25 mothers among 294 candidate fathers and an unknown number of unsampled males.

Data quality varied widely between males on one hand and mothers and larvae on the other (Table 3). Among males, 97% of loci on average yielded genotypic data, rising to 98% when including repeat genotypes. Both allelic dropout and false allele error rates were low (<2%) in males. Averaged across loci, 84% of mothers and larvae yielded genotypic data from at least one of the three repeat genotypes. Mothers and larvae incurred high allelic dropout (18%) but relatively low false allele (1.9%) error rates.

### Pairwise relatedness

Median pairwise relatedness ( $r$ ) among the 294 candidate fathers was 0.00. The 2.5 and 97.5% quantiles were -0.19 and 0.19 and the maximum and minimum values were

-0.45 and 0.43, respectively. The distribution of relatedness was smooth, symmetrical and typical of normally distributed noise, with a single mode at zero and no modes characteristic of first- ( $r=0.5$ ) or second-degree relatives ( $r=0.25$ ). We conclude that the problems caused by the presence of closely related candidate fathers are unlikely to have affected this analysis.

### Patterns of paternity

The posterior probability of paternity from the neighbouring caecum was 0.024 (95% CI 0.003–0.077). Even this low probability may be an overestimate given the strong prior odds against probabilities very close to zero or one. The 95% CI for the number of unsampled males ranged from 77 to 219, with a median of 136 (Figure 1). Combining this estimate with the 294 sampled males reveals that 21–43% (median 32%) of adult males went missing between mating and sampling. Even if we assume that 10% of males were lost during sampling, the bounds of the credible interval fall only slightly to 13–37% (median 25%).

Two patterns emerge from the inferred distribution of paternity confidence (Figure 2). First, it shows a strong bias towards high probabilities, indicating that paternity has been determined with high probability for a large proportion of offspring. Of the 150 larvae, 112 (75%) were assigned a 'probable' father, that is, the maximum probability of paternity was >0.5, whereas 76 (51%) were assigned a father with a probability of >0.8. Second, no male from the neighbouring caecum was a probable father of any of the larvae, reflecting the lack of support in the data for cross-caeca paternity.

Although fractional assignment of paternity yields no single 'accepted' pedigree, for illustration a pedigree constructed using the paternity assignments with the highest probabilities is presented in Figure 3. This pedigree shows two main features: (1) a high degree of promiscuity among both sexes and (2) a moderately high variance in reproductive success, indicated by the observation that although many males have contributed offspring to the next generation, only a few (for example, M13) have fathered considerably more than the average.

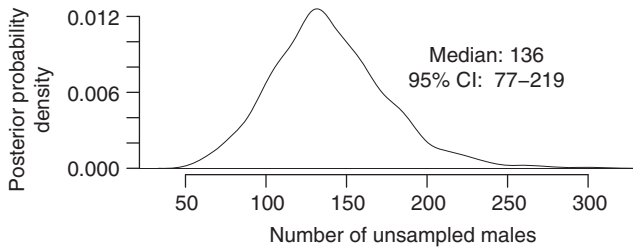
### Promiscuity

Mean reproductive promiscuity was 0.75 (95% CI 0.69–0.80) for females and 0.64 (95% CI 0.52–0.75) for males,

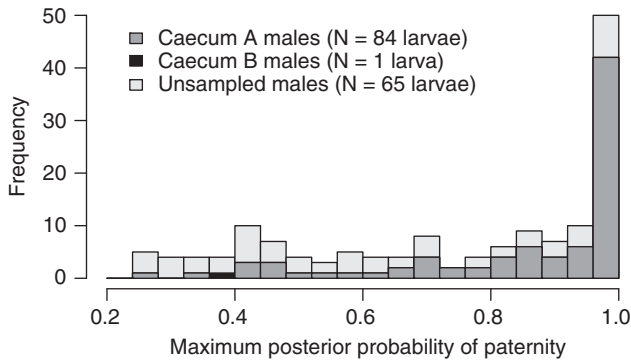
**Table 3** The quantity and quality of microsatellite genotype data from *Trichostrongylus tenuis*: number of samples ( $n$ ), number of genotypes attempted, proportion of loci and individuals successfully genotyped and genotyping error rates (with 95% credible intervals)

Generation	Caecum	n	n genotypes assayed (mean n per individual)	Proportion of loci typed, individuals (%)	Proportion of loci typed, n = 887 genotypes (%)	n (%) individuals typed at ≥5 loci	Allelic dropout rate per allele (95% CI)	False allele rate per allele (95% CI)
Mother	Local	25 <sup>a</sup>	75 (3.0)	84	53	22 (88%)	0.177 (0.163–0.191)	0.019 (0.015–0.023)
Larval offspring	Local	150	450 (3.0)	84	55	133 (89%)		
Candidate father	Local	122	148 (1.2)	96	96	122 (100%)		
Candidate father	Neighbouring	172	214 (1.2)	98	97	170 (99%)		
All	Both	469	887 (1.9)	92	72	449 (96%)	0.082 (0.075–0.090)	0.013 (0.010–0.016)

<sup>a</sup>An additional four mothers whose offspring yielded no genotype data have been omitted from the table. They were nevertheless retained in the analysis because they provided additional information for estimating allele frequencies and error rates. Eleven males from the neighbouring caecum and four larvae that failed at all loci have been excluded from the table.



**Figure 1** Posterior probability density of the number of unsampled males.



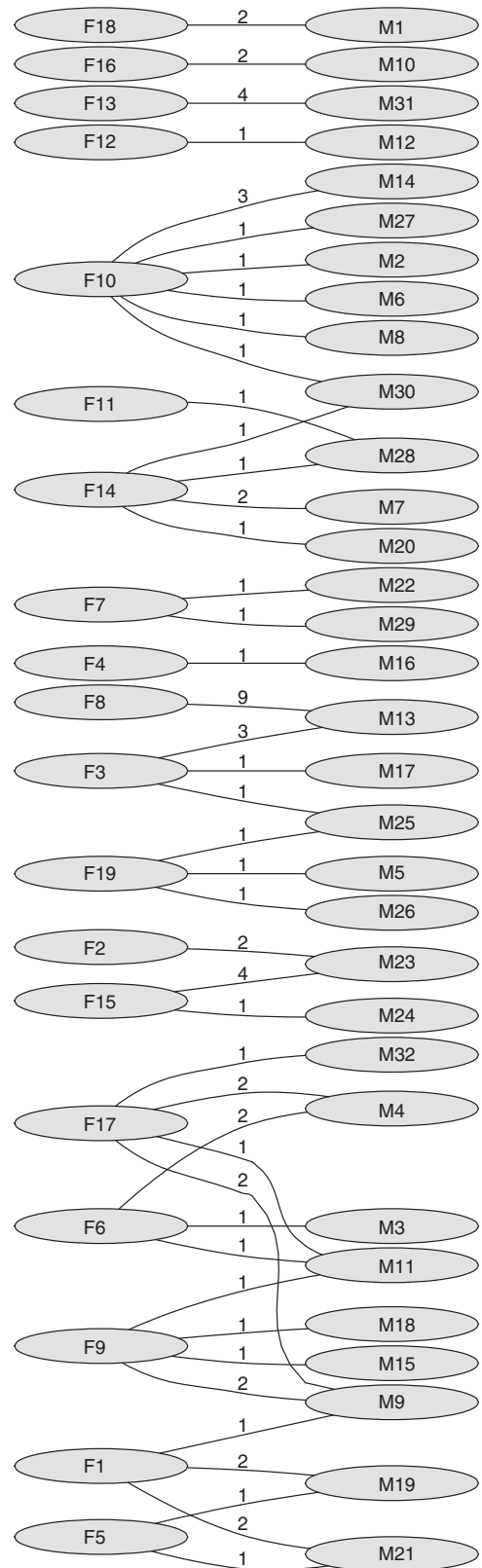
**Figure 2** The distribution of the maximum posterior probability of paternity among 150 *T. tenuis* larvae.

and was 17% (95% CI 1–42%) higher in females when compared with males.

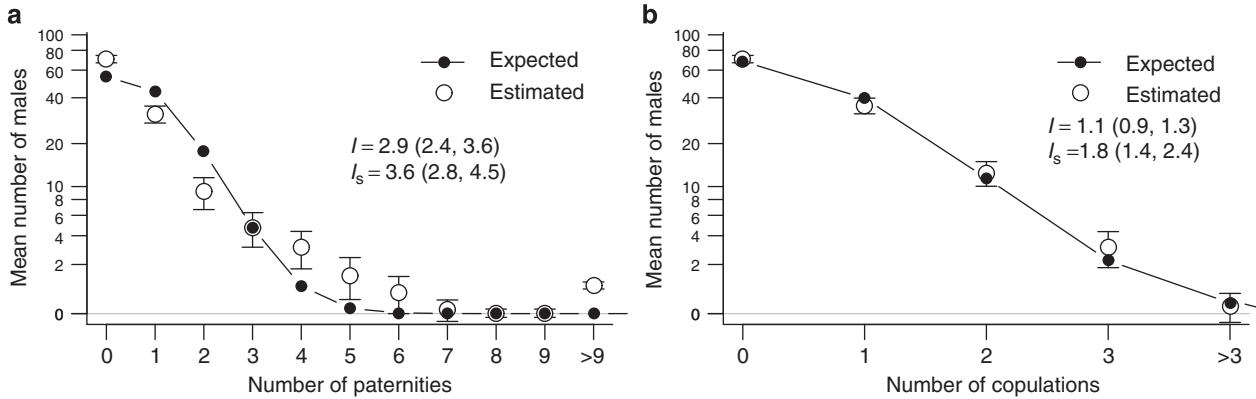
#### Distribution of male reproductive success

We analysed the distribution of reproductive success among the 122 sampled males from the local caecum, measured in terms of both paternities and copulations. Only males from the local caecum were considered because the opportunity to father offspring seems to be largely restricted to males sharing a caecum with the mothers (see above) and, therefore, to include males from the neighbouring caecum would conflate the effects of inter- and intra-caecal barriers to random mating. The distribution of the paternities among the 122 sampled males is shown in Figure 4a. The distribution is overdispersed and clearly deviates from the null distribution under random allocation, there being more males than expected with no paternities, fewer with 1–2 paternities and more males with at least four paternities than expected. This pattern of high variance in reproductive success is confirmed by an *I* estimate of >1 (2.9; 95% CI 2.4–3.6).

If the deviation from random allocation of paternities was caused by fecundity differences, the estimate of *I* will be artificially inflated by variation in female fecundity and by differential mortality because of the *in vitro* conditions. To analyse the impact of these potentially confounding factors, we recalculated *I* after setting variance in female reproductive success to zero. This was achieved by including in the pedigree only females with the mean number of offspring. Females with fewer offspring were excluded and a random sample of offspring was used from females with more



**Figure 3** A pedigree showing the parentage of 72 larvae by 19 females (left) and 32 sampled males (right). Paternity assignments with a maximum probability of paternity exceeding 0.64 are shown, resulting in a mean paternity assignment confidence of 90%. Parentage is indicated by solid lines linking males and females, with the numbers of offspring given above the lines.



**Figure 4** The observed distribution of reproductive success among sampled males in the local caecum, and the expected distribution under the null hypothesis of a random (Poisson) allocation. The mean ( $\pm$  s.d.) number of males achieving a given number of units of reproductive success (measured in paternities (a) and copulations (b)) was estimated from 1000 Markov Chain Monte Carlo (MCMC) samples. The statistics  $I$  and  $I_s$  are standardized variances in reproductive success (95% credible interval (CI)) as defined by Boness *et al.* (1993) and Wade and Arnold (1980), respectively.

offspring (mean  $n$  offspring per female = 4.4; mean  $n$  females included = 11). Error due to sampling offspring was removed by averaging over 100 random samples in each of the 1000 pedigrees in the Markov Chain Monte Carlo output. Because the mean number of offspring was generally not an integer, the  $I$  estimate was a weighted average of estimates obtained by rounding up and down (for example, if the average number of offspring was 4.4, the  $I$  estimate would be  $0.6I^{(4)} + 0.4I^{(5)}$  in which the superscripts denote the number of offspring per female used to estimate  $I$ ). The resulting estimate of 1.6 (95% CI 1.4–1.8) is considerably lower than the unadjusted estimate, and indicates that variation in female reproductive success inflated the  $I$  estimate by about 80%. Nevertheless, this caveat applies only to the size of the estimate and not to the conclusion that allocation is nonrandom, because if paternities are allocated randomly there will be no female-level correlation between paternity assignments, and hence the expected value of  $I$  will be one regardless of variation in female reproductive success.

In contrast, when reproductive success was defined as the minimum possible number of copulations that could have given rise to the observed distribution of paternities (that is, assuming that all full-sib larvae arose from a single copulation), it did not deviate significantly from random allocation (Figure 4b). The  $I$  estimate for the number of copulations of 1.1 (95% CI 0.9–1.3) is close to one, showing no substantial deviation from random mating. Thus, the distribution of paternities is consistent with random mating combined with post-copulatory differences in reproductive success.

## Discussion

We have conducted the first paternity analysis in a parasitic nematode population, the four main outcomes of which are: (1) a large fraction of candidate fathers were not sampled, probably having been ejected by the host; (2) mating was predominantly (and possibly exclusively) within, not between, caeca; (3) both sexes mated promiscuously; and (4) variance in male reproductive success was higher than expected under random allocation, possibly as a result of some males receiving

more mating opportunities, but more likely because of differences in male fecundity after random mating. An additional outcome of this study is the demonstration that paternity analysis in parasitic nematodes is feasible, in spite of the technical and statistical challenges imposed by their endoparasitic lifestyle, microscopic larvae, large population of candidate fathers and unknown number of unsampled males.

Perhaps the most surprising finding was the large proportion (21–43%) of males that were not sampled. Could we have overestimated the proportion of unsampled males? A potential source of overestimation is violation of the assumption that all samples are equally prone to genotyping error. Although error parameters were specific to mothers and larvae on one hand and fathers on the other, sample quality and therefore error rates would have varied within these groups. However, analysis of simulated data in which the samples were mixtures of two unobserved and highly disparate error rates suggested that even extreme error rate variation results in only a moderate (14%) upward bias in the estimated number of unsampled males (JD Hadfield, unpublished data). Another possible source of overestimation is the assumption that sampled and unsampled males contribute equally to paternity. If unsampled males have a higher probability of gaining paternity, then those offspring assigned to unsampled males will be distributed across fewer males than expected. This will result in overestimation of the unsampled population size even though the number of offspring assigned to unsampled versus sampled males should be accurate. A tendency for reproductive success to increase with age (for example, due to a period of sexual immaturity in newly moulted adult males) would therefore bias the estimate upwards, assuming that older males are more likely to have been lost from the caeca. Although there are no data on how male fecundity changes over time in *T. tenuis*, senescence and declining mating frequency with age is suspected in males of *Onchocerca volvulus* (Karam *et al.*, 1987). Older populations of *T. tenuis* produce fewer eggs (Shaw and Moss, 1989), but this is likely to be caused by declining female fecundity. Bias aside, the unsampled male population will also tend to be high if male longevity is low and the

duration of sperm storage is high. Experimental infections of *T. tenuis* in captive red grouse can survive for more than 2 years without substantial mortality (Shaw and Moss, 1989), although their longevity in wild populations is unknown. The duration of sperm storage is also unknown in any parasitic nematode. Thus, at present there is insufficient information to explain the surprisingly large number of unsampled males in this study.

Among the sampled males, it was less surprising to discover that a male sharing a caecum with a female is much likelier to gain paternity over her offspring than a male from the other caecum. This result suggests that the passage of adults or sperm between caeca occurs very rarely, if at all. The capacity for oriented movement is thought to be widespread among nematodes (Burr and Robinson, 2004), and hence it is plausible that, given sufficient motivation, *T. tenuis* adults would be capable of moving between caeca. Such motivation may therefore have been absent in the population of 613 adult *T. tenuis* from which our sample was taken, although populations at extremes of population density—burdens of *T. tenuis* can range from fewer than ten to tens of thousands of adults—may experience greater pressures to emigrate in search of mates or food.

In addition to the population-level processes discussed above, we made inferences from the estimated pedigree about the mating behaviour of individual nematodes. Obligately sexual parasitic nematodes have long been thought to mate randomly and promiscuously, and this assumption of randomness is generally extended to cover allocation of reproductive success (Dobson *et al.*, 1987; Barnes *et al.*, 1995; Saul, 1995; Smith *et al.*, 1999; Churcher and Basáñez, 2008). Deviation from random mating will usually reduce the effective population size, which will in turn reduce the amount of genetic variation that a population can maintain, accelerate genetic drift and reduce the effectiveness of selection on traits such as drug resistance. Our results from *T. tenuis* are consistent with random mating and promiscuity, but not with random allocation of reproductive success. However, the deviation from random allocation observed in this study should be interpreted with caution because it is effectively instantaneous. Ideally, measures of reproductive success are taken over an organism's lifetime, whereas this study was restricted to observing a few hours of reproductive output from the time of host sampling until the females ceased laying viable eggs *in vitro*. It is possible that the degree of variance in male reproductive success observed in this study might average out over the lifespan of a *T. tenuis* to produce a distribution of male reproductive success more indicative of random allocation of paternities, although such data would be practically impossible to acquire in an endoparasite. In species in which lifelong observations are feasible, short-term estimates of standardized variance in male reproductive success ( $I_s$ ) have been found to be unreliable indicators of long-term  $I_s$ , either by underestimating (Partridge, 1988) or overestimating it (Fincke, 1988).

It should also be noted that only a single host was sampled, and that the parameters estimated in this study might vary between hosts, possibly in association with host-level factors, in particular, parasite population size. It should be further borne in mind that the posterior

distribution of parentage is likely to be biased towards random (Poisson) mating because the Bayesian analysis assumes that parentage follows a Poisson process, conditioning on the fixed effects (in this case, the caecum). One option for dealing with this bias would be to take a random effects approach to model overdispersion in male fecundity, or even to model mate fidelity through overdispersion at the level of parental combination, although this would add further complexity to an already computationally complex analysis.

Fully random and promiscuous mating, in which mate choice and pair bonding are absent, is expected to be rare because both sexes should benefit from the ability to choose mates that bring direct (for example, high fecundity) and indirect benefits ('good genes'). However, promiscuity can be favoured under certain conditions: when there is no variation in fitness, or in unpredictable environments in which selective forces favour different genes in different generations (Jennions and Petrie, 2000) or when monopolization of mates or resources is not feasible. Any of these factors might explain the degree of promiscuity that we have observed in *T. tenuis*. A lack of variation in fitness seems the least likely factor, for a variety of reasons. First, a *T. tenuis* population must have the ability to evade the potentially varied immune defences of their hosts, and variation in immune evasion by parasitic nematodes is likely to have a significant genetic component, as suggested by variation in immunogenicity and other life-history traits between strains (Grant, 1994; Paterson and Barber, 2007). In addition, variation in fitness is likely to be influenced by non-genetic factors such as age, which is associated with declining fecundity in female *T. tenuis* (Shaw and Moss, 1989). Other explanations are therefore more plausible. *T. tenuis*, in common with most parasites of vertebrates, live in an unpredictable environment in the sense that each new generation is likely to encounter a new host with a different set of immune defences. Such unpredictability should also make promiscuous mating a beneficial strategy by maximizing the genetic variation among offspring and improving the chances of creating optimal genotypes for an unknown future environment.

Absence of mate choice could also result from mechanistic limitations. It is not known to what extent parasitic nematodes are able to distinguish between mates, or if either sex has the ability to control access to matings. Pheromone-mediated sex attraction of females by males and of males by females is widespread in animal parasitic nematodes (Bone and Shorey, 1978), and hence pheromones provide a potential mechanism through which mate choice could operate. The existence of sperm storage organs also provides the opportunity for cryptic post-copulatory mate choice in females and sperm competition between males. All nematodes have large amoeboid sperm that are thought to compete for access to the spermatheca. Indeed, the evolution of amoeboid sperm in nematodes is thought to have been driven by sperm competition (LaMunyon and Ward, 1999; Snook, 2005).

In summary, we have used paternity analysis of a *T. tenuis* population to reveal a large number of missing males and patterns of mating consistent with random and promiscuous mating within caeca followed by post-copulatory variance in male reproductive success. The



extent to which these patterns reflect general processes in parasitic nematodes—and in particular whether the assumption of random mating is justified—is difficult to discover by studying populations in wild hosts, but could perhaps more easily be tackled by genotyping the shed eggs of experimentally infected captive hosts.

## Note added in proof

The authors have become aware of a related paper which has recently been published: Redman *et al.*, *Genetics* **180**: 1877–1887 (December 2008).

## Acknowledgements

We are grateful to the gamekeeper and the estate owner for access to the sampling site and assistance with sampling. We also thank Malcolm Kennedy, Steve Paterson, Victoria Grillo, Erica Packard, John Gilleard, Mike Stear and Terry Burke for advice and discussions. This paper benefited considerably from the comments of Christine Müller-Graf and two anonymous reviewers. This work was funded by a Leverhulme Trust research project grant (to PCDJ, AA and LFK) and an NERC grant awarded to T Burke in Sheffield (to JDH).

## References

- Anderson JD, Williams-Blangero S, Anderson TJC (2003). Spurious genotypes in female nematodes resulting from contamination with male DNA. *J Parasitol* **89**: 1232–1234.
- Barnes EH, Dobson RJ, Barger IA (1995). Worm control and anthelmintic resistance: adventures with a model. *Parasitol Today* **11**: 56–63.
- Bird AF, Bird J (1991). *The Structure of Nematodes*. Academic Press: San Diego.
- Blaxter ML (2001). Molecular analysis of nematode evolution. In: Kennedy MW, Harnett W (eds). *Parasitic Nematodes: Molecular Biology, Biochemistry and Immunology*. CABI Publishing: Wallingford, UK, pp 139–166.
- Bone LW, Shorey HH (1978). Nematode sex pheromones. *J Chem Ecol* **4**: 595–612.
- Boness DJ, Bowen WD, Francis JM (1993). Implications of DNA fingerprinting for mating systems and reproductive strategies of pinnipeds. *Symp Zool Soc Lond* **66**: 61–93.
- Botstein D, White RL, Skolnick M, Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment polymorphisms. *Am J Hum Genet* **32**: 314–331.
- Braselton J, Huband J, Braselton L (2004). Incorporating mating preferences into a host–parasite model. *Math Biosci* **192**: 1–18.
- Burr A, Robinson AF (2004). Locomotion behaviour. In: Gaugler R, Bilgrami AL (eds). *Nematode Behaviour*. CABI Publishing: Wallingford, UK, pp 25–62.
- Churcher TS, Basáñez M-G (2008). Density dependence and the spread of anthelmintic resistance. *Evolution* **62**: 528–537.
- Churcher TS, Schwab AE, Prichard RK, Basáñez M-G (2008). An analysis of genetic diversity and inbreeding in *Wuchereria bancrofti*: implications for the spread and detection of drug resistance. *PLoS Negl Trop Dis* **2**: e211.
- Criscione CD, Poulin R, Blouin MS (2005). Molecular ecology of parasites: elucidating ecological and microevolutionary processes. *Mol Ecol* **14**: 2247–2257.
- Crow JF (1958). Some possibilities of measuring selection intensities in man. *Hum Biol* **30**: 1–13.
- Dobson RJ, Griffiths DA, Donald AD, Waller PJ (1987). A genetic model describing the evolution of levamisole resistance in *Trichostrongylus colubriformis*, a nematode parasite of sheep. *IMA J Math Appl Med Biol* **4**: 279–293.
- Emery AM, Wilson IJ, Craig S, Boyle PR, Noble LR (2001). Assignment of paternity groups without access to parental genotypes: multiple mating and developmental plasticity in squid. *Mol Ecol* **10**: 1265–1278.
- Fincke OM (1988). Sources of variation in lifetime reproductive success in a nonterritorial damselfly (Odonata: Coenagrionidae). In: Clutton-Brock TH (ed). *Reproductive Success: Studies of Individual Variation in Contrasting Breeding Systems*. University of Chicago Press: Chicago, pp 24–43.
- Floyd R, Abebe E, Papert A, Blaxter M (2002). Molecular barcodes for soil nematode identification. *Mol Ecol* **11**: 839–850.
- Grant WN (1994). Genetic variation in parasitic nematodes and its implications. *Int J Parasitol* **24**: 821–830.
- Griffith SC, Owens IPF, Thuman KA (2002). Extra pair paternity in birds: a review of interspecific variation and adaptive function. *Mol Ecol* **11**: 2195–2212.
- Grillo V, Jackson F, Cabaret J, Gilleard JS (2007). Population genetic analysis of the ovine parasitic nematode *Teladorsagia circumcincta* and evidence for a cryptic species. *Int J Parasitol* **37**: 435–447.
- Grillo V, Jackson F, Gilleard JS (2006). Characterisation of *Teladorsagia circumcincta* microsatellites and their development as population genetic markers. *Mol Biochem Parasitol* **148**: 181–189.
- Hadfield JD, Richardson DS, Burke T (2006). Towards unbiased parentage assignment: combining genetic, behavioural and spatial data in a Bayesian framework. *Mol Ecol* **15**: 3715–3730.
- Hope IA. (ed) (1999). *C. Elegans: A Practical Approach*. Oxford University Press: Oxford.
- Hudson PJ (1992). *Grouse in Space and Time*. The Game Conservancy Trust: Fordingbridge, UK.
- Jennions MD, Petrie M (2000). Why do females mate multiply? A review of the genetic benefits. *Biol Rev Camb Philos Soc* **75**: 21–64.
- Johnson PCD, Webster LMI, Adam A, Buckland R, Dawson DA, Keller LF (2006). Abundant variation in microsatellites of the parasitic nematode *Trichostrongylus tenuis* and linkage to a tandem repeat. *Mol Biochem Parasitol* **148**: 210–218.
- Jones AG, Ardren WR (2003). Methods of parentage analysis in natural populations. *Mol Ecol* **12**: 2511–2523.
- Jones AG, Avise JC (2001). Mating systems and sexual selection in male-pregnant pipefishes and seahorses: insights from microsatellite-based studies of maternity. *J Hered* **92**: 150–158.
- Karam M, Schulz-Key H, Remme J (1987). Population dynamics of *Onchocerca volvulus* after 7 to 8 years of vector control in West Africa. *Acta Trop* **44**: 445–457.
- Koch M, Hadfield JD, Sefc KM, Sturmbauer C (2008). Pedigree reconstruction in wild cichlid fish populations. *Mol Ecol* **17**: 4500–4511.
- Kokko H, Mackenzie A, Reynolds JD, Lindstrom J, Sutherland WJ (1999). Measures of inequality are not equal. *Am Nat* **154**: 358–382.
- Krebs CJ (1989). *Ecological Methodology*. Harper & Row: New York.
- LaMunyon CW, Ward S (1999). Evolution of sperm size in nematodes: sperm competition favours larger sperm. *Proc R Soc London B Biol Sci* **266**: 263–267.
- Lynch M, Ritland K (1999). Estimation of pairwise relatedness with molecular markers. *Genetics* **152**: 1753–1766.
- Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998). Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol* **7**: 639–655.
- Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583–590.
- Nei M, Roychoudhury AK (1974). Sampling variances of heterozygosity and genetic distance. *Genetics* **76**: 379–390.

- Nejsum P, Frydenberg J, Roepstorff A, Davis Parker JE (2005). Population structure in *Ascaris suum* (Nematoda) among domestic swine in Denmark as measured by whole genome DNA fingerprinting. *Hereditas* **142**: 7–14.
- Nieberding C, Libois R, Douady CJ, Morand S, Michaux JR (2005). Phylogeography of a nematode (*Heligmosomoides polygyrus*) in the western Palearctic region: persistence of northern cryptic populations during ice ages? *Mol Ecol* **14**: 765–779.
- Partridge L (1988). Lifetime reproductive success in *Drosophila*. In: Clutton-Brock TH (ed). *Reproductive Success: Studies of Individual Variation in Contrasting Breeding Systems*. University of Chicago Press: Chicago. pp 11–23.
- Paterson S, Barber R (2007). Experimental evolution of parasite life-history traits in *Strongyloides ratti* (Nematoda). *Proc R Soc Biol Sci Ser B* **274**: 1467–1474.
- Peakall R, Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* **6**: 288–295.
- R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna.
- Saul A (1995). Computer model of the maintenance and selection of genetic heterogeneity in polygamous helminths. *Parasitology* **111**: 531–536.
- Shaw JL, Moss R (1989). The role of parasite fecundity and longevity in the success of *Trichostrongylus tenuis* in low-density red grouse populations. *Parasitology* **99**: 253–258.
- Smith G, Grenfell BT, Isham V, Cornell S (1999). Anthelmintic resistance revisited: under-dosing, chemoprophylactic strategies, and mating probabilities. *Int J Parasitol* **29**: 77–91.
- Snook RR (2005). Sperm in competition: not playing by the numbers. *Trends Ecol Evol* **20**: 46–53.
- Troell K, Engstrom A, Morrison DA, Mattsson JG, Hoglund J (2006). Global patterns reveal strong population structure in *Haemonchus contortus*, a nematode parasite of domesticated ruminants. *Int J Parasitol* **36**: 1305–1316.
- Wade MJ, Arnold SJ (1980). The intensity of sexual selection in relation to male sexual-behavior, female choice, and sperm precedence. *Anim Behav* **28** (MAY): 446–461.
- Wang J (2004). Sibship reconstruction from genetic data with typing errors. *Genetics* **166**: 1963–1979.
- Webster LMI, Johnson PCD, Adam A, Mable BK, Keller LF (2007). Macrogeographic population structure in a parasitic nematode with avian hosts. *Vet Parasitol* **144**: 93–103.
- Wielgoss S, Taraschewski H, Meyer A, Wirth T (2008). Population structure of the parasitic nematode *Anguillicola crassus*, an invader of declining North-Atlantic eel stocks. *Mol Ecol* **17**: 3478–3495.
- Williams BD, Schrank B, Huynh C, Shownkeen R, Waterston RH (1992). A genetic mapping system in *Caenorhabditis elegans* based on polymorphic sequence-tagged sites. *Genetics* **131**: 609–624.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)