Inferring infection processes of a parasitic nematode using population genetics

S. PATERSON*†, M. C. FISHER‡ *and* M. E. VINEY†

Division of Biological Sciences, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK

(*Received 27 May 1999; revised 29 August 1999; accepted 30 August 1999*)

SUMMARY

The distribution of genetic differentiation in a population of the parasitic nematode *Strongyloides ratti* divided between rat hosts was determined. We applied population genetic theory to these data to determine the source of new infections. We estimate the rate at which a rat acquires a new infection from (a) the existing subpopulation of parasites within that rat ('self-reinfection') versus (b) the wider environment ('immigration'). We find that the observed levels of genetic diversity and differentiation in the study population are consistent with low to moderate rates of self-reinfection and inconsistent with high rates of self-reinfection.

Key words: host–parasite interactions, helminths, epidemiology.

INTRODUCTION

We do not understand the process or the pattern by which hosts acquire new helminth infections. Measuring rates of infection in nature is a difficult process (Smith *et al*. 1995; Anderson, 1998). This is especially true for helminth parasites which do not spread through a host population purely by infection of naive individuals, but are endemic in the host population and can therefore superinfect existing hosts (Anderson & May, 1992). Modern molecular genetic markers have provided new tools with which to examine natural infection processes and offer the potential to dissect infections at the genetic, rather than microscopic, level (Anderson, Romero-Abal & Jaenike, 1995; Blouin *et al*. 1995). Population genetic studies of helminth populations have now become widespread (for a review of nematode population genetics see Anderson, Blouin & Beech, 1998). Indeed, epidemiology (the study of the dynamics of infection within a population) and population genetics (the study of the dynamics of inheritance within a population) are analogous processes and the 2 disciplines have much to offer each other (Anderson, May & Gupta, 1989).

Endemic populations of helminth parasites persist within host populations and are typically distributed between hosts in an aggregated manner, with the majority of helminths present in a minority of hosts (Shaw & Dobson, 1996). Heterogeneities in helminth transmission and host susceptibility are known to be important in generating such aggregated distributions (Anderson & May, 1992; Lively & Apanius, 1995; Mollison & Levin, 1995), although the relative importance of these 2 processes is unclear. A host typically acquires a helminth infection by environmental contamination with infective stages. For any individual host these infective stages can be derived either from parasites already in that same individual host or from parasites in other hosts in the population. We term the former process 'self-reinfection' and define it to include (i) when parasite development may proceed in the external environment but where the infective stages acquired by a host are the progeny of adult parasites within that host, and (ii) when all stages leading to the infective stage are completed within or upon host tissues without entering the external environment. Self-reinfection as defined by (i) will occur in spatially-structured host populations with home territories, separate nest sites etc. in which the opportunity exists for infection of a host by progeny of its own parasites (Boulinier, Ives & Danchin, 1996). Self-reinfection as defined by (ii) has only been reported for *Strongyloides stercoralis* (Grove, 1999) and *Capillaria philippinensis* (Cross, 1990). Self-reinfection will result in heterogeneity between hosts in transmission, since more heavily infected hosts will be exposed to a greater number of infective stages compared to lightly infected hosts (Hudson & Dobson, 1995).

A helminth population is compartmentalized between individual hosts, and thus the parasites of an

* Corresponding author: School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK. Tel: +44 117 928 7489. Fax: +44 117 925 7374. E-mail: S.Paterson@bris.ac.uk

† Present address: School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK

‡ Present address: Department of Plant and Microbial Biology, 111 Koshland Hall, University of California, Berkeley, California 94720–3102, USA

https://doi.org/10.1017/S0031182099005417 Published online by Cambridge University Press

individual host can be considered a subpopulation of the total parasite population (Anderson, Romero-Abal & Jaenike, 1993; Fisher & Viney, 1998). Self-reinfection has the potential to lead to genetic differentiation of the parasite population since allele frequencies in each parasite subpopulation would be expected to drift independently of each other. Two consequences will result. Firstly, as allele frequencies diverge the genetic distance between parasite subpopulations will increase. Secondly, the genetic diversity of each subpopulation will be reduced as the rate of self-reinfection increases (Hartl & Clark, 1989).

It is important to understand the process of helminth infection for a number of reasons. Firstly, such information is fundamental to basic helminth epidemiology. Secondly, understanding the nature of the infectious process will add to our understanding of helminth population genetics and this will therefore inform and allow us to predict the spread of genetic traits such as drug-resistance (Anderson, May & Gupta, 1989; Saul, 1995). For example, under high rates of self-reinfection, the frequency of drug-resistant alleles would, intuitively, be expected to rise rapidly after just a single drug treatment. However, high rates of self-reinfection will subdivide the parasite population and may therefore retard the spread of drug-resistant alleles through the parasite, and thus host, populations. Self-reinfection also has the potential to influence the efficacy of novel anti-helminthic vaccines, since parasite genes allowing parasites to evade the immune effects of a vaccine will behave similarly under high rates of self-reinfection.

Here we analyse a wild population of rats naturally infected with the nematode parasite *Strongyloides ratti*. We consider the distribution of parasites found in rats caught within the same locale and determine the genetic differentiation between parasite subpopulations in order to estimate rates of self-reinfection.

MATERIALS AND METHODS

Study system

S. ratti is an intestinal nematode parasite of rats. Its life-cycle has 2 adult generations. The obligate parasitic generation is female only and reproduces by mitotic parthenogenesis (Viney, 1994). The facultative free-living adult generation is dioecious and reproduces by conventional meiosis and syngamy (Viney, Matthews & Walliker, 1993). Eggs that are produced by the parasitic female pass out of the host in the faeces. These eggs can develop by one of 2 routes, called homogonic and heterogonic. In the former, the larvae moult through 2 stages into infective 3rd-stage larvae (iL3s). In heterogonic development, the eggs passed by the host moult through 4 larval stages into free-living adult males and females. These mate and their progeny moult through 2 larval stages into iL3s, as in homogonic development. Heterogonic development is restricted to a single generation. All iL3s are committed to infecting a host, which they do by penetration of the skin. From isolates sampled throughout the world, variation in the proportion of larvae developing by either route is observed, with this developmental choice under both genetic and environmental control (Viney, 1999). Isolates from the UK, however, develop almost exclusively via the homogonic route (Viney, Matthews & Walliker, 1992; Fisher & Viney, 1998).

Sampling

Brown rats (*Rattus norvegicus*) were trapped from a single rural farm in Berkshire as previously described (Fisher & Viney, 1998). Post-capture, animals were maintained individually and 6 g of faecal material was cultured and maintained at 19 °C for 3 days, after which the number of iL3s present was determined and the intensity of infection calculated as the number of iL3s/g faeces. In addition, up to 20 iL3s were harvested, washed twice in distilled water and stored individually in 5 μl of distilled water at −20 °C. DNA preparations of single iL3s were made as previously described (Viney, 1994) and genotyped at the polymorphic Actin and BSP-8 loci (Fisher & Viney, 1998). Infective L3s from 22 rats were successfully genotyped for Actin and iL3s from 21 rats were successfully genotyped for BSP-8 (Fisher & Viney, 1998).

Data analysis

F statistics. The population structure of *S. ratti* within the Berkshire sample was defined as containing 2 levels; (i) 'parasite subpopulation', the population of worms found within an individual rat host and (ii) 'total parasite population', all worms found within all the rats sampled.

Genetic differentiation of parasite populations was assayed using Wright's (1951) F statistics defined as F_{IT}, the inbreeding coefficient of the total parasite population, F_{ST}, the inbreeding coefficient due to population subdivision ('coancestry') of worms within individual rat hosts, i.e. the genetic distance between parasite subpopulations, and F_{IS}, the inbreeding coefficient within parasite subpopulations. These statistics are related by:

$$(1-F_{IT})=(1-F_{IS})(1-F_{ST}). \quad (1)$$

Of the 3 F statistics, F_{ST} is of particular interest because it defines the level of genetic differentiation due to population subdivision. F statistics were

https://doi.org/10.1017/S0031182099005417 Published online by Cambridge University Press

estimated following the method of Weir & Cockerham (1984) and implemented in Matlab V (MathWorks, Inc., Natick, MA).

Because *S. ratti* parasitic females reproduce by mitotic parthenogenesis, all progeny are genetically identical to their parent. Thus, the parasite population within hosts can be assayed by sampling and genotyping larvae passed from infected hosts. The intensity of infection (number of iL3s/g faeces) is directly related to the intestinal parasite burden (Fisher, 1997). Thus, the intensity of infection will be used as the effective size of each subpopulation, N, throughout.

A host can acquire new infections that are derived from (i) parasites in other hosts, which we term 'immigration' and (ii) the host's extant parasite subpopulation which we term 'self-reinfection'. Immigration occurs at rate m, and self-reinfection at rate $1-m$. Population subdivision, as measured by F_{ST}, will be expected to increase with greater self-reinfection, $1-m$. At equilibrium, the level of population subdivision, $\hat{F}$, can be given as:

$$\hat{F}=\frac{1}{1+2N\left(\frac{1}{(1-m)^2}-1\right)}, \tag{2}$$

where N is the effective subpopulation size and m is the proportion of immigrants per generation in the effective population. For small values of m, $\hat{F}$ is commonly given by the approximation:

$$\hat{F}\approx\frac{1}{4Nm+1} \tag{3}$$

(Hartl & Clark, 1997), however, for estimates of self-reinfection rate (given by $1-m$) at intermediate levels of self-reinfection equation (2) is more appropriate.

Estimation of the rate of self-reinfection ($1-m$) using equation (2) can be made given F_{ST} values and N for each subpopulation. However, F_{ST}, cannot be made linear for m and so a joint distribution, $P(\mathbf{y}, \mathbf{N}, \hat{\mathbf{F}}, m, \tau)$, of observed F_{ST} values ($\mathbf{y}$), intensity of infection $\mathbf{N}$, expected F_{ST} values ($\hat{\mathbf{F}}$), immigration rate m and variance of observed around expected F_{ST} values (τ^{-1}) was constructed. m and τ were given normal and gamma prior distributions, respectively. Gibb's sampling (Hastings, 1970; Gilks, Richardson & Spielgelhalter, 1996) was then used to condition on the data to obtain a posterior distribution of the unknown parameters m and τ. The posterior distribution was then marginalized over the parameters using numerical integration to obtain estimates of m and τ. Gibb's sampling was implemented using WinBUGS (http://www.mrc-bsu.cam.ac.uk/bugs).

Simulation. The effect of self-reinfection upon the genetic differentiation of parasite subpopulations, as measured by F_{ST}, was examined by computer simulation. Due to the very low frequency ($<1\%$) of free-living males and females (Viney, Matthews & Walliker, 1992; Fisher & Viney, 1998) that reproduce sexually in *S. ratti*, iL3s are produced almost exclusively by mitotic parthenogenesis and will thus be genetically identical to their parents. The genotypes of both immigrant iL3s and self-reinfecting iL3s were therefore calculated from the observed genotype frequencies rather than inferred from Hardy–Weinberg proportions of allele frequencies. Genotype frequencies in the total parasite population were assumed to be constant whilst genotype frequencies within each parasite subpopulation were allowed to vary stochastically following either immigration from the total parasite population at rate m or self-reinfection at rate $1-m$. The distribution of worms within each parasite subpopulation was given by the observed intensity of infection, N_k, found for each rat, k, of the Berkshire sample.

We used 2 simulation schemes to model the infection process. Simulation 1 assumed equilibrium conditions, i.e. that the rate of new infections was balanced by the rate at which infections were lost. A discrete time Markov chain (Grimmitt & Stizaker, 1992) was used to model this infection process assuming N_k adult worms in each parasite subpopulation, each subpopulation k at time t occupying state $S_{ik}=\{n_{ik1}, n_{ik2}, \ldots, n_{ikl}, \ldots, n_{ikL}\}$ where $n_{..1}$ is the number of worms of genotype g_l with $N_{.k}=\Sigma_l n_{.kl}$. At time $t+1$ it is assumed that a single adult worm is lost and replaced by another single worm to give state $S_{jk}=\{n_{jk1}, n_{jk2}, \ldots, n_{jkl}, \ldots, n_{jkL}\}$ such that $N_{ik}=N_{jk}$ and $\Sigma_l \mid n_{ikl}-n_{jkl} \mid \in\{0, 2\}$. The probability of loss of an adult worm of genotype $g_{l'}$ is given by:

$$P_{ikl}(\text{loss } g_l)=\frac{n_{ikl}}{N_k}, \tag{4}$$

and successful infection by a new worm of genotype $g_{l'}$ is given by:

$$P_{ikl'}(\text{gain } g_{l'})=\begin{cases} mf^*_{l'}+(1-m)\dfrac{n_{ikl'}}{N_k-1}, & l\neq l' \\ mf^*_{l}+(1-m)\dfrac{n_{ikl}-1}{N_k}, & l=l' \end{cases} \tag{5}$$

where $f^*_{\cdot}$ gives genotype frequencies of the total parasite population.

At each time-point, t, of the Markov chain, a rat was selected at random with a probability $p_k=N_k/\Sigma N$, and moved from state S_{ik} to state S_{jk} with gain and loss of a single worm as described by equations (4) and (5). This Markov chain was used to follow progressive decreases in immigration rates with 800 samples taken to estimate F_{ST} after a burn-in of 200 samples, or until convergence was reached.

https://doi.org/10.1017/S0031182099005417 Published online by Cambridge University Press

F_{ST} was estimated from random multinomial samples from the parasite subpopulations. This process mimics the sampling of iL3s from faeces used here (Fisher & Viney, 1998). These samples were taken every 1000 iterations, except for $m < 0{\cdot}05$ when the Markov chain moved more slowly and samples were taken every 2000 iterations.

Simulation 2 modelled the infection process before equilibrium was reached, i.e. the infection of naive hosts. Under this scheme, only gains of new infections were considered and no infections were lost. A single worm from the genotype frequencies of the total parasite population, $f_{\cdot}^{*}$, was randomly assigned to each host. Subsequent infections followed from either self-reinfection or immigration until the observed, N_k, number of worms in each host was reached, i.e. the probability of successful infection by a new worm of genotype $g_{l'}$ at time t is given by:

$$P_{ikl'}(\text{gain } g_{l'}) = \begin{cases} f_{l'}^{*}, & t = 0 \\ mf_{l'}^{*} + (1-m)\dfrac{n_{ikl'}}{\sum_{l} n_{ikl}}, & 0 < t < N_k \\ 0, & t = N_k \end{cases} \quad (6)$$

using the same nomenclature as for Simulation 1. This process was repeated 1000 times for each level of self-reinfection and F_{ST} values calculated from each replicate as for Simulation 1.

Number of genotypes. The numbers of unique genotypes present within each parasite subpopulation, U_k, were calculated and a regression of $U_k - 1$ against $\log_{10} N$ was conducted using linear least squares regression without an intercept. $U_k - 1$ was used since it allowed regression through the origin. The observed value of this regression coefficient was compared against regression coefficients generated by simulation (below).

The simulation was used to determine the number of unique genotypes, U_k, for each parasite subpopulation under different rates of self-reinfection. As with the simulation of F_{ST} values, the observed intensity of infection was used as the distribution of worms within hosts. Samples of U_k generated by simulation were used to estimate the regression of U_k against intensity of infection as described above.

RESULTS

Distribution of parasites

In total 27 rats were sampled, 24 of which were found to be infected. Rats had a mean of 37·67 iL3s/g faeces with variance $\sigma^2 = 1932$. The distribution of iL3s found in the faeces is shown in Fig. 1 and was found to be significantly overdispersed relative to a Poisson distribution ($\chi^2_{\nu=26} = 685$, $P < 0{\cdot}001$) but did approximate to a negative binomial distribution ($\chi^2_{\nu=25} = 27{\cdot}32$, $P = 0{\cdot}39$) with overdispersion parameter $k = 0{\cdot}695$ ($\pm 0{\cdot}188$ S.E.).

Genetic differentiation and self-reinfection

F_{ST} compares the genetic differentiation of a subpopulation with the total population. Observed F_{ST} values, together with associated F_{IS} and F_{IT} values, for Actin, BSP-8 and both loci combined are shown in Table 1. This shows that the observed genetic differentiation between parasite subpopulations is low ($F_{ST} < 0{\cdot}05$). An excess of heterozygotes is indicated by negative values of the inbreeding coefficients F_{IS} and F_{IT}. This result has been reported previously and is consistent with low levels of sexual development and recombination in *S. ratti* (Viney, Matthews & Walliker, 1992; Fisher & Viney, 1998). F_{ST} will be unaffected by lack of recombination since it measures divergence of allele frequencies rather than inbreeding.

To examine the effect of self-reinfection upon genetic differentiation, simulations were performed to generate the distribution of F_{ST} values under varying levels of self-reinfection. We used 2 simulation schemes. Firstly (Simulation 1), we assumed equilibrium conditions, i.e. that the rate of gain of parasites is balanced by the rate of loss of parasites and that F_{ST} values have reached their equilibrium levels. In the wild, however, rats may not live long enough for the infection process, and for F_{ST} values, to reach equilibrium. To address this (Simulation 2), we also modelled the infection of naive hosts (non-equilibrium conditions), considering only the gain of infections and calculated F_{ST} values when all hosts had received the observed numbers of parasites. The results of these simulations are shown in Fig. 2. Under equilibrium conditions for the Actin and BSP-8 loci or both loci combined (Fig. 2A–C), we found that with low to moderate levels of self-reinfection values of F_{ST} remained low ($< 0{\cdot}05$). With high levels of self-reinfection values of F_{ST} rose rapidly to $> 0{\cdot}2$. This pattern is also seen under non-equilibrium conditions (Fig. 2D–F). This indicates that by the time the observed levels of infection have been built up in a host population, genetic differentiation between parasite subpopulations could also be established, under high levels of self-reinfection. We note, however, that the plot of F_{ST} versus self-reinfection rate is more concave under equilibrium conditions than under non-equilibrium conditions. Because of this, equilibrium conditions are to be preferred in estimating rates of self-reinfection from observed F_{ST} values since these give conservative estimates relative to non-equilibrium conditions.

The values of F_{ST} that we observed were $< 0{\cdot}04$ and this is consistent with low to moderate levels of self-reinfection. However, the 95% confidence limits

https://doi.org/10.1017/S0031182099005417 Published online by Cambridge University Press

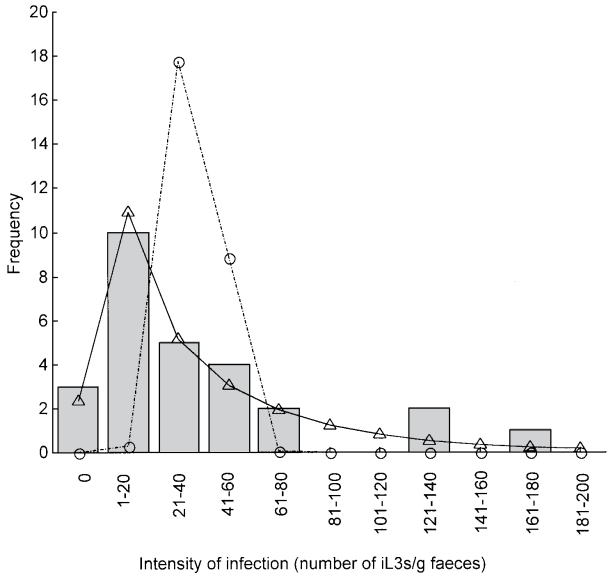

Fig. 1. Intensity of infection (number of iL3s/g faeces) found in 27 rats sampled in Berkshire (bars) and the predicted values from Poisson (○) and negative binomial (△) distributions.

Table 1. Wright's F statistics for genetic diversity of iL3s partitioned between rats within the Berkshire population

(n, Number of parasite subpopulations sampled.)

	n	F_{IT}	F_{ST}	F_{IS}
Actin	22	−0·5730	0·0346	−0·6294
BSP-8	21	−0·0424	0·0145	−0·0578
Combined	21	−0·4952	0·0329	−0·5460

associated with these simulated estimates of F_{ST} under equilibrium conditions do not allow us to exclude the possibility of high rates of self-reinfection. We also note that the 95% confidence limits for BSP-8 are much larger than those for Actin. This is probably due to the allele frequencies at BSP-8 being close to 1.

Genetic differentiation and intensity of infection

The parasite population sampled had an over-dispersed distribution in common with most helminth infections (Fig. 1). The hosts with the high intensities of infection may have acquired these heavy infections by high levels of self-reinfection. If so, then there would be a relationship between F_{ST} (Fig. 2) and intensity of infection. To investigate whether there was evidence of this phenomenon occurring, the observed values of F_{ST} and the intensity of infection (N) for each parasite subpopulation were used to derive estimates for the self-reinfection rate ($1-m$) using equation (2) in a model (Model A) that assumed that there was no relationship between self-reinfection ($1-m$) and intensity of infection (N). Equation (2) assumes equilibrium conditions. We also used the observed values of F_{ST} and the intensity of infection (N) for each parasite subpopulation in a model (Model B) that assumed that there was a relationship between self-reinfection rate ($1-m$) and intensity of infection (N) as:

$$m=\frac{b}{1+\nu \ln N} \tag{7}$$

to estimate b, the underlying immigration rate, and to estimate ν, which is the decrease of the immigration rate with increasing intensity of infection. Parameter estimation was made by Gibb's sampling. This analysis was done for (I) Actin, (II) BSP-8 and (III) both loci combined and the results are shown in Table 2.

Model A assumed that self-reinfection rate, $1-m$, was independent of intensity of infection, i.e. $\nu=0$ and thus $1-m=1-b$. The mean estimate of $1-m$ for both loci combined was 0·35, which indicates that low to moderate levels of self-reinfection occur. This implies that the predominant source of new infections is immigration. This, in turn, would lead to low levels of observed genetic differentiation between parasite subpopulations. $1-m$ was found to have wide 95% confidence limits for BSP-8 and both loci combined with the lower 95% limit lying very close to a self-reinfection rate of zero. (Note that m is bounded between 0 and 1.) The lower 95% limit for Actin ($1-m\geqslant 0{\cdot}16$) initially suggests that self-reinfection rate may be significantly different from zero. However, closer inspection showed that this result was lost after the exclusion of 1 data point and made results for Actin concordant with those BSP-8 and both loci combined. This data point had the lowest estimated intensity of infection ($N=2$). Parameter estimation from equation (2) is particularly sensitive to data points with small N.

Model B allowed self-reinfection rate, $1-m$, to increase with intensity of infection, N, according to equation (7) and the results are shown in Table 2. Estimates of the underlying self-reinfection rate $1-b$ were close to those found for $1-m$ in Model A ($1-b=0{\cdot}23$ for both loci combined). The rate, ν, by which self-reinfection, $1-m$, increases with intensity of infection, N, was 0·279 for both loci combined. The 95% confidence intervals for ν are large and included zero. This suggests that the self-reinfection rate is not affected by intensity of infection, i.e. more heavily infected hosts do not show higher rates of self-reinfection.

The standard deviation $\tau^{-1/2}$ of the distribution of the observed F_{ST} values around the expected F_{ST} values is also shown in Table 2. This shows that for both models $\tau^{-1/2}$ is of the same order of magnitude

https://doi.org/10.1017/S0031182099005417 Published online by Cambridge University Press

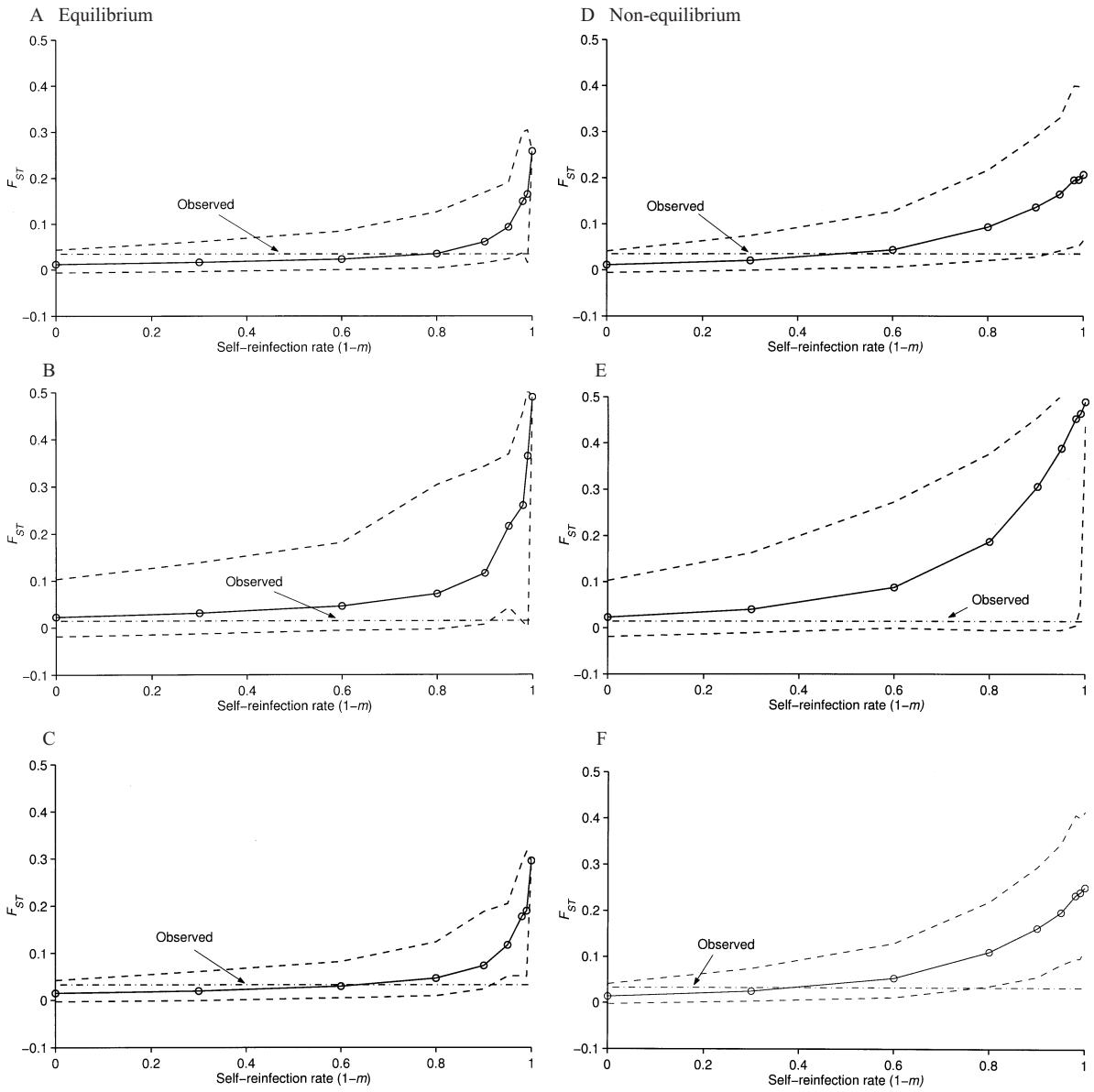

Fig. 2. Median simulated F_{ST} values (—) with 95% confidence intervals (·····) under increasing levels of self-reinfection for (A and D) Actin, (B and E) BSP-8 and (C and F) both loci combined. (A–C) The simulated response of F_{ST} under equilibrium conditions (Simulation 1) i.e where gain of new infections is balanced by the loss of old infections. (D–F) The simulated response of F_{ST} under non-equilibrium conditions (Simulation 2) i.e. the infection of naive animals, where only gain of new infections is considered. Observed F_{ST} values shown by (·—·—·).

as the observed F_{ST} values (Table 1). This indicates that much, if not all, the variation in F_{ST} values between parasite subpopulations is due to sampling variance.

Number of genotypes

We investigated the effect of self-reinfection upon genetic diversity within parasite subpopulations. The measure of genetic diversity used was the number of parasite genotypes present within each subpopulation, which will increase with intensity of infection. Regression of the number of genotypes against log-transformed intensity of infection gave an observed regression coefficient of 0·957 ($\pm$0·266 S.E.) as shown in Fig. 3. The number of genotypes present in each parasite subpopulation will, intuitively, decrease with increasing levels of self-reinfection. Using the simulations described, we determined the number of unique genotypes that would be present in each parasite subpopulation under varying levels of self-reinfection ($1-m$). From this, we derived the regression coefficients of the number of unique genotypes against log intensity of infection (as above) under varying levels of self-reinfection, the results of which are shown in Fig. 4. This shows that at low to moderate levels of self-reinfection, the regression coefficient is high, but that this falls rapidly with high levels of self-

https://doi.org/10.1017/S0031182099005417 Published online by Cambridge University Press

Table 2. Estimation of self-reinfection rate

($1-m$, the self-reinfection rate, was estimated as a simple term (Model A) or with $m=b/(1+\nu \ln N)$ (Model B) (N, Intensity of infection; $\tau^{-1/2}$, variance of observed around expected F_{ST} values. Values in parentheses for Actin in Model A refer to the effect of dropping the data point with the lowest intensity of infection (see text). All estimates were derived by Gibb's sampling.)

	n	Variable	Mean	2·5 %	Median	97·5 %
Model A						
Actin	22	$1-m$	0·429	0·157	0·447	0·569
	(21)		(0·317)	(0·018)	(0·322)	(0·632)
		$\tau^{-1/2}$	0·028	0·021	0·028	0·040
			(0·044)	(0·032)	(0·043)	(0·061)
BSP-8	21	$1-m$	0·239	0·011	0·222	0·564
		$\tau^{-1/2}$	0·084	0·062	0·082	0·118
Combined	21	$1-m$	0·354	0·035	0·380	0·578
		$\tau^{-1/2}$	0·041	0·030	0·040	0·057
Model B						
Actin	22	$1-b$	0·259	0·014	0·254	0·546
		ν	0·293	−0·006	0·254	0·802
		$\tau^{-1/2}$	0·044	0·033	0·043	0·061
BSP-8	21	$1-b$	0·188	0·007	0·162	0·500
		ν	0·307	−0·008	0·254	0·881
		$\tau^{-1/2}$	0·086	0·062	0·084	0·120
Combined	21	$1-b$	0·233	0·011	0·224	0·515
		ν	0·279	−0·007	0·240	0·770
		$\tau^{-1/2}$	0·041	0·030	0·040	0·058

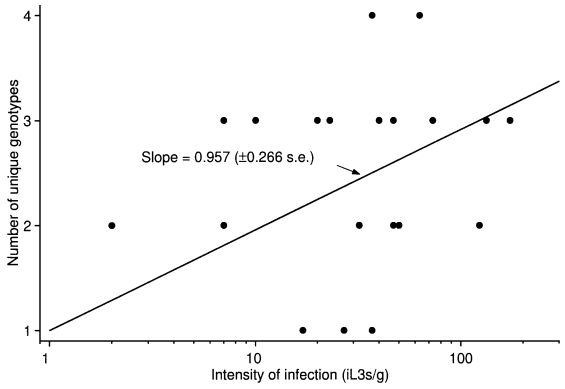

Fig. 3. Regression of the number of unique parasite genotypes within each parasite subpopulation against log intensity of infection (number of iL3s/g faeces).

reinfection. This pattern is observed under both equilibrium and non-equilibrium conditions, although the curve is more convex under equilibrium conditions. Because of this, equilibrium conditions are to be preferred in estimating rates of self-reinfection since these give conservative estimates relative to non-equilibrium conditions.

The observed regression of number of genotypes against intensity of infection (Fig. 3) is consistent with low to moderate levels of self-reinfection ($1-m<0{\cdot}7$). Hence the observed levels of genetic diversity, as measured by the number of unique genotypes, is likely to be a result of rats acquiring new infections predominantly by immigration.

DISCUSSION

Despite the extensive effort and advances that have been made in understanding the dynamics of helminth infections, the process by which hosts become infected in nature is not known. Here we have used a population genetic approach to investigate infection dynamics in a natural population of *S. ratti*. We considered each host as harbouring separate subpopulations of parasites and simulated the dynamics of *S. ratti* infection to investigate the impact of self-reinfection on (i) genetic differentiation between parasite subpopulations and (ii) the level of diversity within parasite subpopulations. We find that under high levels of self-reinfection, a rapid increase in the level of genetic differentiation between parasite subpopulations is expected. Similarly, under high levels of self-reinfection, the genetic diversity in a parasite subpopulation will be low. These simulated results are consistent with established population genetic theory which states that only a few immigrants are sufficient to reduce drastically genetic differentiation and to maintain genetic diversity (Slatkin, 1987).

In the Berkshire sample of *S. ratti*, we found low levels of genetic differentiation ($F_{ST}<0{\cdot}05$) which, by comparison with simulated values, are consistent with low to moderate levels of self-reinfection. In an extension to this analysis, we obtained parameter estimates of the self-reinfection rate as $1-m=0{\cdot}35$ (95 % C.I. 0·03–0·58). This suggests that the pre-

https://doi.org/10.1017/S0031182099005417 Published online by Cambridge University Press

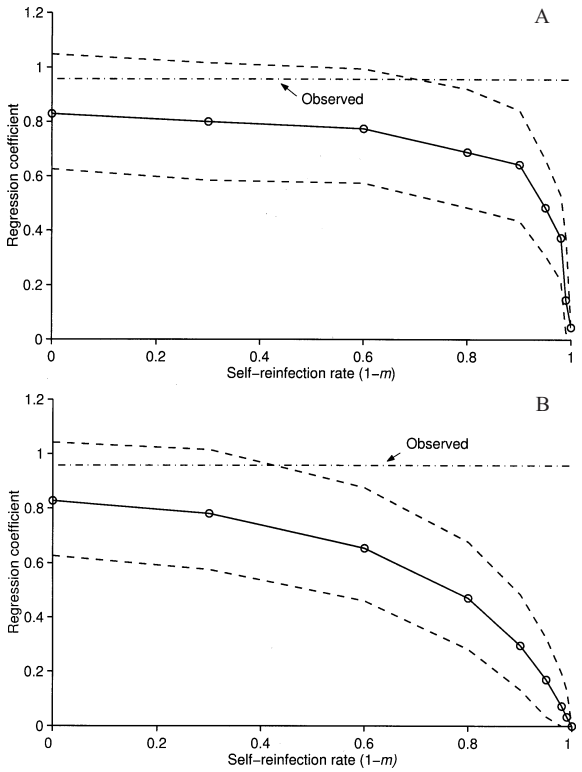

Fig. 4. Median simulated regression coefficients (—) with 95% confidence intervals (·····) of the number of unique parasite genotypes within each parasite subpopulation against log intensity of infection under increasing levels of self-reinfection. (A) The simulated response under equilibrium conditions (Simulation 1) i.e. where gain of new infections is balanced by the loss of old infections. (B) The simulated response under non-equilibrium conditions (Simulation 2) i.e. the infection of naive animals, where only gain of new infections is considered. The observed regression coefficient value is shown by (·—·—·).

dominant source from which a host acquires new infections is immigration, i.e. by environmental contamination from other hosts rather than self-reinfection. We also found no evidence that self-reinfection rate increases with intensity of infection. This suggests that heavily infected hosts have not acquired their heavy infections by self-reinfection. Thus, our results suggest that other factors, such as host predisposition, are responsible for generating the observed pattern of helminth overdispersion.

Genetic diversity, as measured by the number of unique genotypes in each parasite subpopulation, will be expected to rise with increasingly large subpopulations. This will result in a positive regression between the number of genotypes in a parasite subpopulation and the size of that subpopulation. Self-reinfection will reduce both the slope of this regression and the genetic diversity of each parasite subpopulation. We have observed a positive relationship between the number of unique genotypes and log intensity of infection, with a slope of 0·957. This shows that not only do a minority of hosts harbour the majority of parasites, but that they also harbour the majority of the genetic diversity of the total parasite population. Comparison of this observation with simulated regression coefficients under differing levels of self-reinfection indicates that the observed regression coefficient is consistent with low to moderate levels of self-reinfection.

We conclude from these analyses that the source of most new infections of a host is by immigration into that parasite subpopulation rather than by self-reinfection. This will occur by continuous infection from the general environment. Levels of genetic differentiation between parasite subpopulations are low, indicating significant gene-flow within the Berkshire sample. This gene-flow acts to maintain diversity both within parasite subpopulations and in the total parasite population.

The simulation approach used here to investigate the response of genetic differentiation and diversity to self-reinfection is doubtless much simpler than the reality of infection in the field. We used 2 schemes that simulated firstly, an equilibrium between gain of new infections and loss of old infections and, secondly, the infection of naive individuals under non-equilibrium conditions. The infection process in the field probably lies somewhere between these two extremes. Nevertheless, since both simulation schemes gave broadly similar results, we can be confident that our findings also hold for the actual infection process in the field. It's worth noting, however, that there is a slightly less concave/convex curvature under non-equilibrium conditions relative to equilibrium conditions in Figs 2 and 4 (the response of genetic differentiation and diversity to self-reinfection, respectively). In Simulation 2 (non-equilibrium conditions), a single worm is initially assigned at random to each host. This immediately sets up relatively high levels of genetic diferentiation and low levels of genetic diversity (since each parasite subpopulation has genotype frequencies of either zero or 1 and has only a single, unique genotype). Subsequent infections move levels of genetic differentiation and diversity towards their equilibrium values. However, this equilibrium level has not yet been reached by the end of the simulation, resulting in slightly flatter curves under non-equilibrium conditions compared to equilibrium conditions. Because of this, equilibrium conditions are to be preferred in estimating rates of self-reinfection from observed F_{ST} values since these give conservative estimates relative to non-equilibrium conditions.

A feature of all the analyses presented here are the large confidence intervals associated with estimates of self-reinfection rate. The reason for this is easily seen in Figs 2 and 4, which show the response of genetic differentiation and diversity to self-reinfection, respectively. Only high rates of self-reinfection ($1-m>0{\cdot}8$) give rise to appreciable

https://doi.org/10.1017/S0031182099005417 Published online by Cambridge University Press

changes in genetic differentiation and diversity between and within parasite subpopulations. Low to moderate levels of self-reinfection have little detectable effect on the genetic constitution of the parasite subpopulations. While the population genetic analyses presented here rule out high levels of self-reinfection, accurate estimation of low rates of self-reinfection may not be possible by these methods – even with larger sample sizes or a larger number of loci.

The use of greater numbers of loci could, however, allow the possibility of genetic fingerprinting and identification of individual worms. For a species with an asexual parasitic stage, such as *S. ratti*, the genotypes of new infections may be compared directly to the existing infection to test for self-reinfection. For species with a sexual parasitic stage, a measure of the relatedness between new infections and the existing infection could be used as a test for self-reinfection. Such an approach could, in principle, allow increased precision in estimating low rates of self-reinfection. However, it does require sufficiently large numbers of polymorphic loci such that genetic identity and/or relatedness can be assigned with a high level of confidence. At present, these loci are unavailable for *S. ratti* and for most other helminths, but the development of such genetic fingerprinting techniques would represent a valuable research tool in understanding epidemiological processes.

The application of population genetics to parasite epidemiology represents a valuable tool in understanding the processes and patterns of infection. Anderson *et al*. (1989), however, noted a gulf between population genetic models, which deal primarily with changes in genotype frequency, and ecological models, which deal with changes in parasite or host abundance. Bridging the gap between these two disciplines represents a major challenge for both geneticists and epidemiologists, but one that offers great potential when applied to field data.

We would like to thank Andrew Read for helpful discussions and Yvonne Bisset for technical assistance. This work was supported by the Leverhume Trust, the Medical Research Council and the Natural Environment Research Council.

REFERENCES

ANDERSON, R. M. (1998). Evolutionary pressures in the spread and persistence of infectious agents in vertebrate populations. *Parasitology* **111** (Suppl.), S15–S31.

ANDERSON, R. M. & MAY, R. M. (1992). *Infectious Diseases of Humans : Dynamics and Control*. Oxford University Press, Oxford.

ANDERSON, R. M., MAY, R. M. & GUPTA, S. (1989). Non-linear phenomena in host–parasite interactions. *Parasitology* **99**, 59–79.

ANDERSON, T. J. C., BLOUIN, M. S. & BEECH, R. N. (1998). Population biology of parasitic nematodes: applications of genetic markers. *Advances in Parasitology* **41**, 220–283.

ANDERSON, T. J. C., ROMERO-ABAL, M. E. & JAENIKE, J. (1993). Genetic structure and epidemiology of *Ascaris* populations: patterns of host affiliation in Guatemala. *Parasitology* **107**, 319–334.

ANDERSON, T. J. C., ROMERO-ABAL, M. E. & JAENIKE, J. (1995). Mitochondrial DNA and *Ascaris* microepidemiology: the composition of parasite populations from individual hosts, families and villages. *Parasitology* **110**, 221–229.

BLOUIN, M. S., YOWELL, C. A., COURTNEY, C. H. & DAME, B. J. (1995). Host movement and the genetic structure of populations of parasitic nematodes. *Genetics* **141**, 1007–1014.

BOULINIER, T., IVES, A. R. & DANCHIN, E. (1996). Measuring aggregation of parasites at different host population levels. *Parasitology* **112**, 581–587.

CROSS, J. H. (1990). Intestinal Capillariasis. *Parasitology Today* **6**, 26–28.

FISHER, M. C. (1997). Population genetics of the parasitic nematode *Strongyloides ratti*. Ph.D. thesis, University of Edinburgh, UK.

FISHER, M. C. & VINEY, M. E. (1998). The population genetic structure of the facultatively sexual parasitic nematode *Strongyloides ratti* in wild rats. *Proceedings of the Royal Society, B* **265**, 703–709.

GILKS, W. R., RICHARDSON, S. & SPIELGELHALTER, D. J. (1996). *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London.

GRIMMITT, G. R. & STIRZAKER, D. R. (1992). *Probability and Random Processes*. Clarendon Press, Oxford.

GROVE, D. I. (1999). Human Strongyloidiasis. *Advances in Parasitology* **38**, 251–309.

HARTL, D. L. & CLARK, A. G. (1989). *Principles of Population Genetics*. Sinauer Associates, Sunderland, MA.

HASTINGS, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**, 97–109.

HUDSON, P. J. & DOBSON, A. P. (1995). Macroparasites: observed patterns. In *Ecology of Infectious Diseases in Natural Populations* (ed. Grenfell, B. T. & Dobson, A. P.), pp. 144–176. Cambridge University Press, Cambridge.

LIVELY, C. M. & APANIUS, V. (1995). Genetic diversity in host–parasite interactions. In *Ecology of Infectious Diseases in Natural Populations* (ed. Grenfell, B. T. & Dobson, A. P.), pp. 421–449. Cambridge University Press, Cambridge.

MOLLISON, D. & LEVIN, S. A. (1995). Spatial dynamics of parasitism. In *Ecology of Infectious Diseases in Natural Populations* (ed. Grenfell, B. T. & Dobson, A. P.), pp. 384–398. Cambridge University Press, Cambridge.

SAUL, A. (1995). Computer-model of the maintenance and selection of genetic-heterogeneity in polygamous helminths. *Parasitology* **111**, 531–536.

SHAW, D. J. & DOBSON, A. P. (1996). Patterns of

macroparasitic abundance and aggregation in wildlife populations: a quantitative review. *Parasitology* **111** (Suppl.), S111–S133.

SLATKIN, M. (1987). Gene flow and the geographic structure of natural populations. *Science* **236**, 787–792.

SMITH, G., BASANEZ, M. G., DIETZ, K., GEMMELL, M. A., GRENFELL, B. T., KENNEDY, C. R., LLOYD, S., MEDLEY, G., NASELL, I., RANDOLPH, S. E., ROBERTS, M. G., SHAW, D. J. & WOOLHOUSE, M. E. (1995). Macroparasite group report. In *Ecology of Infectious Diseases in Natural Populations* (ed. Grenfell, B. T. & Dobson, A. P.), pp. 209–229. Cambridge University Press, Cambridge.

VINEY, M. E. (1994). A genetic analysis of reproduction in *Strongyloides ratti*. *Parasitology* **109**, 511–515.

VINEY, M. E. (1999). Exploiting the life cycle of *Strongyloides ratti*. *Parasitology Today* **15**, 231–235.

VINEY, M. E., MATTHEWS, B. E. & WALLIKER, D. (1992). On the biological and biochemical nature of cloned populations of *Strongyloides ratti*. *Journal of Helminthology* **66**, 45–52.

VINEY, M. E., MATTHEWS, B. E. & WALLIKER, D. (1993). Mating in the nematode parasite *Strongyloides ratti*: proof of genetic exchange. *Proceedings of the Royal Society, B* **254**, 213–219.

WEIR, B. S. & COCKERHAM, C. C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.

WRIGHT, S. (1951). The genetical structure of populations. *Annals of Eugenics* **15**, 323–354.