RESEARCH ARTICLE

Inbreeding in stochastic subdivided mating systems: the genetic consequences of host spatial structure, aggregated transmission dynamics and life history characteristics in parasite populations

GUHA DHARMARAJAN*

Department of Biological Sciences, Indian Institute of Science Education and Research - Kolkata, Mohanpur 741 246, India

Abstract

Inbreeding in parasite populations can have important epidemiological and evolutionary implications. However, theoretical models have predominantly focussed on the evolution of parasite populations under strong selection or in epidemic situations, and our understanding of neutral gene dynamics in parasite populations at equilibrium has been limited to verbal arguments or conceptual models. This study focusses on how host–parasite population dynamics affects observed levels of inbreeding in a random sample of parasites from an infinite population of hosts by bridging traditional genetic and parasitological processes utilizing a backward–forward branching Markov process embedded within a flexible statistical framework, the logarithmic-poisson mixture model. My results indicate that levels of inbreeding in parasites are impacted by demographic and/or transmission dynamics (subdivided mating, aggregated transmission dynamics and host spatial structure), and that this inbreeding is poorly estimated by 'equilibrium' levels of inbreeding calculated assuming regular systems of mating. Specifically, the model reveals that at low levels of inbreeding ($F \leq 0.1$), equilibrium levels of inbreeding are lower than those observed, while at high levels of inbreeding the opposite pattern occurs. The model also indicates that inbreeding could have important epidemiological implications (e.g., the spread of recessive drug resistance genes) by directly impacting the observed frequency of rare homozygotes in parasite populations. My results indicate that frequencies of rare homozygotes are affected by aggregated transmission dynamics and host spatial structure, and also that an increase in the frequency of rare homozygotes can be caused by a decrease in effective population size solely due to the presence of a subdivided breeding system.

[Dharmarajan G. 2015 Inbreeding in stochastic subdivided mating systems: the genetic consequences of host spatial structure, aggregated transmission dynamics and life history characteristics in parasite populations. *J. Genet*. **94**, 43–53]

Introduction

Understanding the factors that affect levels of inbreeding in parasites is essential from both an epidemiological and evolutionary perspective (Nee *et al.* 2002). For example, an important effect of inbreeding is that it can facilitate the spread of drug resistance in parasite populations (e.g. Dye 1994; Paul *et al.* 1995; Schmidt 1995; Schwab *et al.* 2007; Churcher *et al.* 2008; but also see Mackinnon and Hastings 1998; Mharakurwa 2004). Our understanding of these factors has been greatly improved by theoretical models which have focussed on the evolution of simple Mendelian traits under strong selection (e.g. drug resistance genes; see Anderson *et al.* 1989; Saul 1995), or changes in population genotypic frequencies (relative to Hardy–Weinberg equilibrium, HWE, expectations) in epidemic situations (e.g. Cornell *et al*. 2000, 2003). However, from an evolutionary and genetic perspective, it also is important to elucidate how factors associated with host–parasite population dynamics affect levels of inbreeding in populations at 'equilibrium' (i.e. conditions of consistent variability; *sensu* May 2001; pp. 109). Unfortunately, our current understanding of neutral gene dynamics in parasite populations at equilibrium remains limited to conceptual models (e.g. Criscione and Blouin 2005) or verbal arguments (e.g. Sire *et al.* 2001; Criscione *et al.* 2005; Criscione 2008; Lymbery and Thompson 2012) that extend the theoretical expectations developed for free-living populations to parasites (e.g. by equating parasite transmission dynamics with migration). However, several characters of parasite populations make them demographically distinct from free-living ones. For example, the infrapopulation (IP; i.e.

*E-mail: guha@iiserkol.ac.in; guha.dharmarajan@gmail.com.

Keywords. aggregation; identity by descent; negative binomial distribution; nematode; stochastic model; tick.

parasites within a particular host; Bush *et al.* 1997) cannot be considered similar to 'true' populations (*sensu* Hartl and Clark 1989, pp. 71) in that they formed exclusively by infection (migration) rather than intrahost (intrapopulation) natality (Sirea *et al.* 2001). Consequently, a group of parasites at the level of the component population (i.e. all parasites of a species within a host population) also cannot be considered, at least a priori, to be panmictic. Rather the more appropriate a priori view would be to consider the component population being subdivided into breeding groups (IPs), with random mating within but not between the breeding groups (see Ryman 1994; Criscione and Blouin 2005). Each breeding group subsequently contributes its progeny to a mixed pool of individuals (free-living or infective stages in the environment) which can potentially constitute the next generation.

The current study focusses on how host–parasite population dynamics affects observed levels of inbreeding in a random sample of parasites from an infinite population of hosts. For the purpose of this paper the term parasite is restricted to macro-parasites (i.e. macroscopic parasites which reproduce sexually in the definitive host and in which pathology is intensity-dependent, e.g. ticks and gastrointestinal nematodes; see Lafferty and Kuris 2002). From a genetic perspective, an individual can be considered to be inbred if its parents shared a common ancestor, and the extent of inbreeding is measured in terms of probability of identity by descent (*F*; i.e., the probability that two homologous genes in uniting gametes have descended from a common ancestral gene; Malécot 1969). Barring a few exceptions (e.g. nidicolous ticks; Sonenshine 1993), mating in parasites is constrained to the scale of the IP, and thus inbreeding is driven by the levels of coancestry among parasites at this scale. The current study considers three important factors related to host–parasite population dynamics that can affect the levels of coancestry amongst individuals at the IP scale, and these are: host spatial structure, aggregated transmission dynamics and parasite life history.

First, the host spatial structure increases the probability of coancestry amongst parasites at the IP scale (Nadler 1995; Criscione *et al.* 2005). Because my model assumes an infinite number of hosts, two parasites in an IP cannot share consanguinity if their pedigrees cannot be traced back to a single IP in the previous generation (figure 1a). However, when hosts are spatially structured there is a nonzero probability of an IP being composed of parasites originating from parents infecting a single IP in the previous generation, and these parasites may be consanguineous (with the level of consanguinity being determined by the pedigree of the parents themselves; figure 1b).

Secondly, the aggregation of larvae in the environment (e.g. faecal deposits of infected hosts in the case of helminths) is likely to cause parasites to be 'sampled' by individual hosts as 'packets' of related individuals and such aggregated transmission dynamics can in turn lead to increased potential for consanguineous mating, and thus inbreeding (Cornell *et al*. 2003; also compare figure 1a and 1c). The level of consanguinity among individuals at the IP scale produced by such aggregated transmission dynamics will depend upon both the number and size of infecting kin groups (compare figure 1c and 1d).

Finally, levels of consanguinity among individuals within the infecting kin groups likely are directly influenced by the life history characteristics of the parasites themselves. For

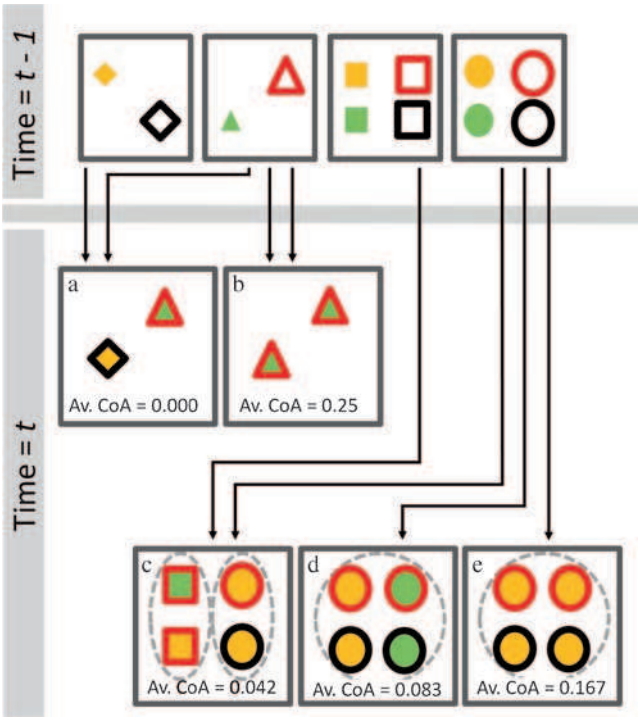

Figure 1. Schematic diagram representing three factors (host spatial structure, aggregated transmission dynamics and parasite life history characteristics) on average levels of coancestry between parasites at the scale of an infrapopulation (IP; black rectangles) at time *t*. Each parasite in the IPs represented at time *t*are assumed to have parents originating from a single IP at time $t-1$. Parasites in unique IPs in the parental generation are represented by unique symbols (diamonds, triangles, squares and circles). The pedigree of each individual at time *t* can be traced using the symbol shape (indicating parental IP), outline colour (indicates father) and fill colour (indicate mother). Coancestry of individuals with different symbols is 0 (see text), while coancestry of individuals that share a symbol increases by 0.125 for each shared fill and/or outline colour (i.e. per parent shared by the individuals). Each host is assumed to undergo discrete infectious cycles (ICs; black arrows), and in each infectious cycle, the host can encounter a single parasite or a kin group of parasites (dashed ellipses). Five scenarios are represented: (a) host undergoes two ICs and 'samples' one offspring each from two distinct IPs at time $t-1$; (b) same as scenario (a), except that due to presence of host spatial structure the two parasites are offspring of a single IP at time $t-1$; (c) same as scenario (a), except that each of the two ICs results in a host contacting a kin group of two parasites with each kin group being produced by panmictic mating of males and females in the parental IP; (d) same as scenario (c), except that the host undergoes only a single IC, and contacts a kin group of four parasites; (e) same as scenario (d), except that all individuals in the single kin group are offspring of a single mother mated panmictically by all males in the parental kin group.

example, eggs in helminths are passed out in the faeces of an infected host, and these faecal deposits act as foci for infection of the next host (Cornell *et al.* 2003). Thus, the genotypic composition of the larval aggregates in the environment can be considered to be the result of panmictic mating among all females and males in the parental IP. Alternatively, in the case of most ticks, mated females detach from the host after engorgement and deposit their eggs (usually several thousands) in the environment (Sonenshine 1991, 1993). Thus, the genotypic composition of the larval aggregates in the environment can be considered to be the result of panmictic mating between a single female and all males in an IP. The two alternative mechanisms described above lead to different expectations with regard to levels of coancestry among individuals in an infecting kin group (compare figure 1d and 1e).

To understand the effects of stochastic host–parasite population dynamics (i.e. host spatial structure and aggregated transmission) on levels of inbreeding in a random sample of parasites the current study utilizes a flexible modelling framework. The model assumes that each individual in the random sample of parasites originates from mating of individuals infecting a single host (designated the focal IP), and focusses on calculating the levels of coancestry among individuals comprising the focal IP using a basic backward–forward simulation procedure. The backward step is a branching, discrete time Markov process which follows the pedigree of all individuals in the focal IP, while the forward step calculates the probabilities of coancestry (C) and F of all individuals in the pedigree up to the individuals in the focal IP. To incorporate the effects of host–parasite population dynamics into the construction of each pedigree I consider the infection process to be a nonhomogeneous Poisson process where the composition of individuals in each IP in the pedigree is considered to be a sample from a Logarithmic-Poisson mixture distribution. Within this statistical framework, hosts are assumed to undergo discrete 'infection events', the number of infection events in an individual host being drawn from a stationary Poisson distribution with a given mean (λ). Additionally, to incorporate the effects of aggregated transmission dynamics, the statistical framework assumes that the number of parasites comprising each infection event is drawn from a logarithmic distribution (with coefficient α). The Logarithmic-Poisson mixture distribution proves to be a convenient statistical framework because this distribution readily reduces to the form of the negative binomial distribution (NBD; Pielou 1976) which is the common distribution used to describe the highly aggregated distribution of parasites among hosts (Shaw and Dobson 1995).

The model reveals that stochastic factors associated with both host (i.e. spatial structure) and parasite (i.e. aggregated distribution of infective stages and life-history) critically affect genetic patterns at the scale of the parasite IP and individual (i.e. C and F, respectively). My analyses also highlight the fact that genetic patterns associated with the stochastic subdivided mating systems observed in parasites do not always match theoretical expectations generated for regular systems of mating (e.g. Hedrick and Cockerham 1986). While, my current model makes numerous simplifying assumptions (e.g. temporal stability of the distribution parameters, direct parasite life cycle, lack of heterogeneity among hosts; detailed below), its relevance lies in the fact that it utilizes a flexible statistical framework (i.e. the negative binomial distribution) long familiar to parasitologists to bridge 'traditional' genetic and parasitological processes.

Materials and methods

Model overview

The objective of the current study was to elucidate factors influencing levels of inbreeding in a random sample of parasites ($n = 10{,}000$). Owing to the fact that the level of inbreeding (F) of each target parasite in the sample was equivalent to the coancestry of its parents, the model focussed on calculating the coancestry among individuals in the parental IP (designated the focal IP) using a backward–forward simulation procedure. The backward step was modelled as a branching discrete time Markov chain process which follows the pedigree of individuals in the focal IP, while the forward step calculated the probabilities of coancestry (C) and F of the individuals in a particular host at each time step up to the individuals in the focal IP. I followed each pedigree for 25 generations, thus the backward step began at time $t = 25$ and terminated at time $t = 0$; at time $t = 0$ I assumed that all individuals were outbred (i.e. $C = F = 0$). The backward–forward simulation to create the pedigree and calculate C/F values was implemented using code written in C++ (Borland, Austin, USA), and the model is described briefly below.

Generating the focal host

In the first step of the simulation, I generated the focal IP from which the target individual was to be sampled. The pattern of parasites among hosts was assumed to follow NBD since this distribution has been shown to fit the aggregated distribution of parasites well (Shaw and Dobson 1995). The NBD is a distribution characterized by two parameters the mean (μ) and the index of dispersion (k), and my simulations evaluated four levels of μ (2.5, 5, 10 and 20 parasites/host) and nine levels of k (50, 5, 1, 0.75, 0.5, 0.4, 0.3, 0.2 and 0.1). To model aggregated transmission dynamics, I modelled the NBD distribution of parasites amongst hosts as a Logarithmic-Poisson mixture distribution (Pielou 1976). Thus, I assumed that the focal host 'sampled' a certain number of kin groups (k_t) at time t where $k_t \sim$Poisson (λ), with the size of the ith kin group ($i = 1$ to k_t) being $s_i \sim$logarithmic (α); where λ and α are the parameters of the Poisson and Logarithmic distribution, respectively. Thus, the total number of individuals 'sampled' by the host ($n_t = \Sigma s_i$) is

distributed as a NBD with a particular μ and k when $\alpha = \mu/(\mu + k)$ and $\lambda = k \times \log_e[(\mu+k)/k]$ (Pielou 1976).

Prior to describing the backward–forward simulation procedure, I detail the numerous simplifying assumptions that the model is based upon. First, I assumed that the genes were neutral and inherited in a Mendelian fashion. The parasites were assumed to have a 1:1 sex ratio, were diploid, reproduced sexually and had a direct life cycle with nonoverlapping generations. I assumed that there are an infinite number of hosts, that there was no heterogeneity among hosts (i.e., the number of infections / host is a Poisson process) and that there was no temporal variability in the pattern of parasite patterns among hosts (i.e. parameters of the NBD are stable within each simulation). Finally, I assumed that there was no density dependence in offspring production at the scale of the IP and that there was no parasite induced host mortality.

The backward simulation

Parasite mating is assumed to take place only at the scale of IP, and thus each k_t kin groups 'sampled' by the focal host at time t was necessarily the product of the mating of individuals within an IP in the previous time step (i.e. $t-1$). In the absence of host spatial structure, the probability (φ) of a particular host at time t 'sampling' multiple kin groups from a single 'parental' IP at time $t-1$ was considered to be negligible (because I assume the presence of infinite number of hosts) and as the host spatial structure increases φ will also increase. I calculated the number of unique parental IPs contributing to the focal IP using the formula $h_t = k_t - n_t$ (where $n_t \sim$Binomial (k_t-1, φ)). Thus, $h_t = k_t$ when $\varphi = 0$ and $h_t = 1$ when $\varphi = 1$, and for the purposes of my simulation I evaluated four levels of φ (0, 0.2, 0.4 and 0.6).

Once the number of unique 'parental' IPs (h_t) contributing offspring to the focal IP at time t was calculated, the backward simulation proceeded by generating these IPs at time $t-1$ (using the procedure used to generate the focal IP; see above), and the k_t kin groups were drawn at random from the h_t parental IPs so generated. The final step that determined the pedigree of the individuals in the focal IP at time t was the parental contributions to each of the k_t kin groups produced by the mating of individuals in the h_t parental IPs at time $t-1$. This parental contribution depended upon the life history characteristics of the parasites (i.e. 'tick-like' versus 'nematode-like' life history). In case of ticks, a single female detaches from a host and lays several thousand eggs on the ground, thus to simulate this life history the individuals comprising each kin group were considered to be the offspring of a single random female mated panmictically with all males within the relevant parental IP (all parents being sampled with replacement). Alternatively, in the case of nematodes, the eggs of all females are passed out together in the faeces of the infected host (the faecal deposit acting as a source of infection for the next host), thus the individuals comprising each kin group were considered to be the offspring of the panmictic mating of all females and males within the parental IP (sampled with replacement). Multiple paternities has been empirically demonstrated in ticks (Hasle *et al.* 2008; Cutulle *et al.* 2010; Ruiz-Lopez *et al.* 2012) and nematodes (Redman *et al.* 2008; Zhou *et al.* 2011). I assumed a direct life cycle and thus the time was only incremented when the definitive host 'samples' the kin group (in the case of an indirect life cycle the attrition of kin group size caused by repeated 'sampling' by intermediate hosts could be considered to be reflected in the coefficient of the logarithmic distribution, α).

Once the parental IPs were generated using the procedure described above, the grandparental IPs (i.e. the parental IPs of the parents) were generated using the same procedure recursively until time 0. At time $t = 0$, all individuals were considered to be outbred, the pedigree was terminated and the forward simulation started.

The forward simulation

Once all the IPs contributing to the pedigree of individuals in the focal IP were generated using the backward simulation procedure described above, the forward simulation started. The forward simulation started at the first time-point (time $t = 0$) and moved in step-wise increments to the last time-point (time $t = 25$). At the first time-point (time $t = 0$), all individuals were considered to be outbred (i.e. level of inbreeding of each individual, and level of coancestry between any pair of individuals, was considered to be zero). At each time step after this (i.e. $t > 0$) in the forward simulation the levels of inbreeding of each individual and the coancestry of each pair of individuals in the pedigree were calculated based on the C and F of individual's parents following standard methods (see Rostron 1978). Briefly, level of inbreeding of any individual Z with dam (D_Z) and sire (S_Z) was calculated as the coancestry of its parents: $F(Z) = C(D_Z S_Z)$. Similarly, the coancestry between any two individuals, X and Y, was calculated as the average consanguinity between their parents: $C(XY) = \frac{1}{4}[C(D_X S_X) + C(D_Y S_Y) + C(D_X D_Y) + C(S_X S_Y)]$. Finally, the coancestry of any individual Z with itself was calculated as $C(ZZ) = \frac{1}{2}[1+F(Z)]$. This recursive forward process was repeated until inbreeding and coancestry values of the focal IP were calculated (i.e. $t = 25$).

Calculating expected levels of inbreeding and homozygosity

Once the inbreeding and coancestry of individuals in the focal IP were calculated using the backward–forward simulation described above, I sampled a random male and female and allowed them to produce a single offspring I and saved the inbreeding coefficient of the offspring (and the coancestry of its parents). I repeated the above backward–forward simulation 10,000 times and calculated the average and variance in observed levels of inbreeding (F_{OBS}) of the individuals thus generated. I also calculated the equilibrium (expected) inbreeding levels (F_{EQ}), which assumed a 'regular' system of

mating based upon the parental coancestry values (see above) using the formula of Hedrick and Cockerham (1986):

$$F_{EQ} = \frac{\sum_j \frac{S_j}{2^j}}{1 - \sum_j S_j \left[1 - \left(\frac{1}{2}\right)^j\right]},$$

where S_j is the proportion of the jth degree of mating ($j = 2$ for full-sib mating and $j = 3$ for half-sib mating).

Cornell *et al.* (2003) have shown that inbreeding increases the frequency of rare homozygotes which may be of epidemiological importance (e.g. drug resistance). Thus, I also calculated the expected frequency of homozygotes (P) expected given an allele frequency of p ($1 \times 10^{-4}, 5 \times 10^{-4}, 1 \times 10^{-3}, 5 \times 10^{-3}$ and 1×10^{-2}) and the observed levels of homozygosity using the formula $P = p^2 + Fp(1 - p)$.

Results and discussion

Inbreeding can have serious consequences on the fitness of individuals by decreasing survival and fecundity, and these negative effects are considered to be predominantly due to the presence of deleterious recessive mutations in the population (Keller and Waller 2002; Charlesworth and Willis 2009). However, an evolutionary perspective inbreeding also can be advantageous because inbreeding prevents breakdown of locally adapted gene complexes (Templeton 1986; Waser and Price 1989). The contradictory effects of inbreeding on individual fitness have long been recognized, leading to the concept of 'optimal inbreeding' (Shields 1982). In parasites, some level of inbreeding may be essential, despite the costs, to enable rapid evolution to changes in host immune response, environmental changes and/or more recently evolution of resistance to drugs. While, understanding the factors affecting gene dynamics under equilibrium conditions is essential, our current understanding remains limited to verbal and/or conceptual models (e.g. Criscione and Blouin 2005; Criscione 2008). The model I present here provides a convenient quantitative approach capable of incorporating host–parasite population dynamics into a genetic framework, and thus helps elucidate factors affecting gene dynamics, specifically levels of inbreeding, in parasite populations under equilibrium conditions.

Deviations from HWE expectations due to a deficit of heterozygotes (HD) does not seem to be uncommon in parasite populations (Luo *et al.* 2003; Picard *et al.* 2004; Chevillon *et al.* 2007; Plantard *et al.* 2008; Wielgoss *et al.* 2008; Dharmarajan *et al.* 2010, 2011; Vilas *et al.* 2012; Ascunce *et al.* 2013; but also see McCoy *et al.* 2005; Caillaud *et al*. 2006; Keeney *et al.* 2007; Thiele *et al.* 2008; see also meta-analysis in Dharmarajan 2008). There is no doubt that such deviations from HWE may be caused by the presence of technical issues associated with PCR amplification of DNA (e.g. de Meeûs *et al.* 2004; Wielgoss *et al.* 2008 or the presence of PCR inhibitory substances in genomic DNA extracted from parasites (e.g. Desloire *et al.* 2006; Dharmarajan and Rhodes 2011). However, it is also possible that factors intrinsically associated with parasite biology also are important contributors to deviations from HWE expectations.

The current study evaluated the effects of three factors associated with host parasite dynamics on levels of inbreeding in parasite populations: host spatial structure, aggregated transmission dynamics and parasite life-history characteristics. In a previous meta-analysis, Dharmarajan (2008) has shown that the magnitude of observed heterozygote deficits (compared to HWE expectations) in ticks is close to double of that observed in the case of nematodes (ticks: $F_{IS} \pm$ SE $= 0.24 \pm 0.03$; nematodes: $F_{IS} \pm$ SE $= 0.11 \pm 0.04$). The results of this study reveal that the life-history characteristics of ticks lead to a higher level of inbreeding compared to that of nematodes (figure 2). Thus, the results support the patterns observed in the meta-analysis of Dharmarajan (2008), and reveal that the observed increase in HD in ticks *cf.* nematodes could be generated by differences in the genetic composition of kin groups due to differences in their life-history characteristics (figure 1). My results also reveal a dramatic increase in levels of inbreeding as host spatial structure increases (figure 2) because host spatial structure increases the probability of parasites from a single IP being 'sampled' by a host in the next generation (figure 1). The fact that highly structured host populations leads to inbreeding in the parasites infecting these hosts is intuitive (Price 1980; Nadler 1995; Criscione *et al.* 2005) because gene flow in parasites is predominantly determined by host movement patterns (Blouin *et al.* 1995; McCoy *et al.* 2003; Dharmarajan *et al.* 2010). From an evolutionary perspective, parasites generally are expected to adapt to local host populations (Dybdahl and Storfer 2003) and my results indicate that inbreeding caused by high levels of host spatial structure should promote the rapid evolution of such locally adapted gene complexes. Akin to host spatial structure, my results revealed that aggregated transmission dynamics also affect levels of inbreeding in intuitive ways. Thus, levels of inbreeding increase as the size of kin groups increase (and number of kin groups decrease) within IPs (figure 2) because of an increase in the probability that two parasites within an IP are consanguineous (see figure 1). The hypothesis that parasites may exhibit such aggregated transmission dynamics has received empirical support from fine-scale genetic analyses at the scale of the IP (Nadler *et al.* 1990; Anderson *et al.* 1995; Guzinski *et al.* 2009; Dharmarajan *et al.* 2010; Vilas *et al.* 2012), and a previous study has revealed that deviations from HWE in parasite populations may be caused by the combined effects of a subdivided breeding system along with aggregated transmission dynamics (Dharmarajan *et al.* 2011). The effects of aggregated transmission dynamics on genotypic frequencies in parasite populations also has been evaluated by previous theoretical models (Smith *et al.* 1999; Cornell *et al.* 2000, 2003), but unlike the work of Cornell and coworkers, the model presented here elucidates how aggregated transmission dynamics affects inbreeding

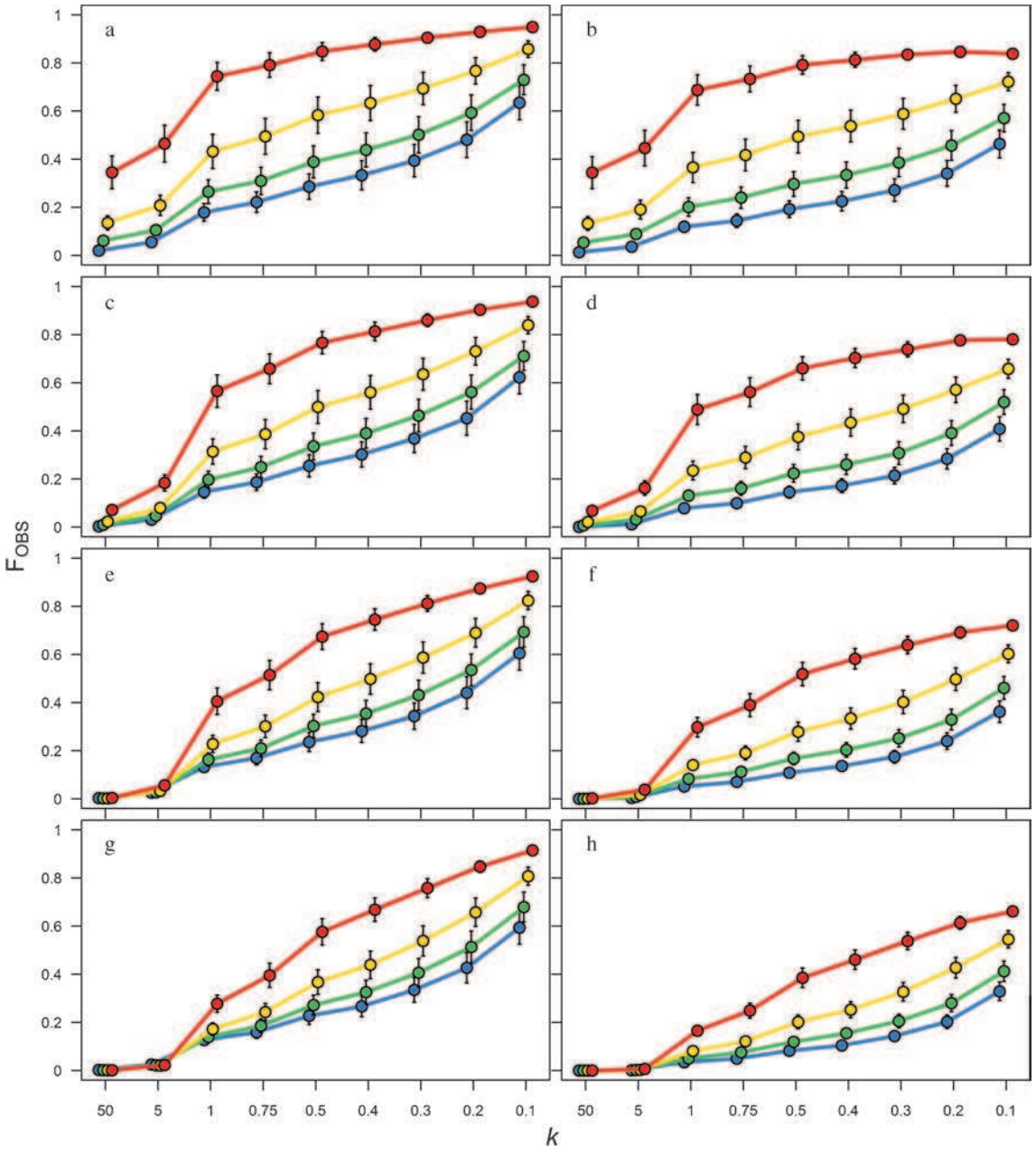

Figure 2. The effects of aggregated transmission dynamics (k) on observed levels of inbreeding (i.e. identity by descent; F) in a random sample of 10,000 parasites (levels of aggregation increase with decreasing k; distribution of parasites among hosts is close to Poisson at $k = 50$). The effects of host spatial structure, represented by four different probabilities (φ) of a host being infected by parasites which are offspring of parents from a single infrapopulation (IP) in the previous generation: 0 (blue), 0.2 (green), 0.4 (yellow) and 0.6 (red). The graphs represent data for four different mean parasite intensity (2.5, a and b; 5, c and d; 10, e and f; 20, g and h), as well as for two different parasite life history characteristics: ticks (i.e. each infecting kin group produced by a single female; a, c, e and g) and nematodes (i.e. each infecting kin group produced by panmictic mating of all males and females in the parental IP; b, d, f and h). Error bars represent variance.

in parasite populations under equilibrium (rather than epidemic) conditions.

Gene dynamics of parasite systems are particularly affected by stochasticity associated with demographic and/or transmission dynamics (Cornell 2005), and thus 'equilibrium' levels of inbreeding calculated assuming regular systems of mating may be poor estimates of observed levels of inbreeding in parasite populations. Indeed, our results reveal a pattern where at low levels of inbreeding ($F_{OBS} \leq 0.1$), equilibrium levels of inbreeding are lower than those observed, while at high levels of inbreeding the opposite pattern occurs (figure 3, a&b). This pattern leads to the so called 'heterozygosity paradox' (Brown 1979) where the observed heterozygosity will be lower than equilibrium expectations in outcrossers and higher than equilibrium expectations in inbreeders will be higher than

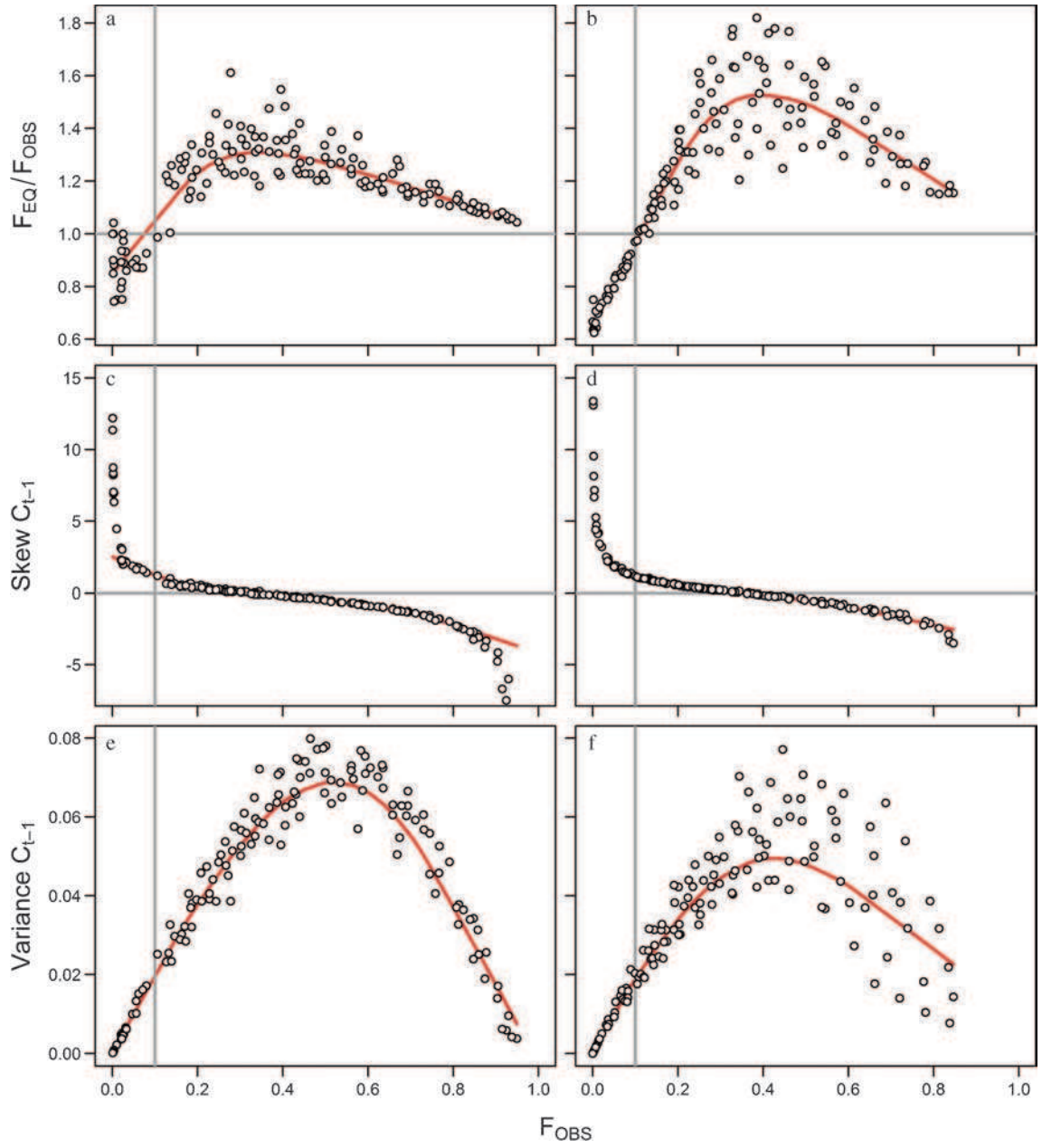

Figure 3. Equilibrium inbreeding (F) and the heterozyosity paradox in the case of two different parasite life history characteristics: ticks (a, c and e) and nematodes (b, d and f). Graphs pool data from all evaluated parameter values relating to aggregation, host spatial structure and mean parasite intensity (see text). The graphs represent: (i) the ratio between the equilibrium (expected) and observed levels of inbreeding (F_{EQ}/F_{OBS}) (a and b). (ii) Skew in coancestry values in the parental generation (skew C_{t-1}) (c and d); (iii) variance in coancestry values in the parental generation (variance C_{t-1}) (e and f). Also shown are the fits generated by locally weighted regression (LOWESS) analysis (red lines). The observed F at which the equilibrium inbreeding exceeds observed F_{IS} indicated in all graphs (gray vertical line).

equilibrium expectations. In the case of outbreeders it is generally thought that this pattern is driven by high levels of relatedness among breeders due to small neighbourhood size (Brown 1979; Hedrick and Cockerham 1986). Alternatively, in inbreeders it has been proposed that both finite population size and variance in outcrossing rates could lead to a reduction in observed inbreeding *cf.* equilibrium expectations (Nei 1975; Brown 1979; Saleem and Gliddon 1986). In the case of parasite populations, the observed patterns seem to be driven by two forces: (i) in outbreeders (i.e. at low levels of inbreeding) the high positive skew in parental coancestry values (figure 3, c&d) leads to a few mating between highly related individuals which leads to an increase in the observed average levels of inbreeding *cf.* equilibrium expectations; (ii) in inbreeders, however the high variance in parental coancestry values (figure 3, e&f) likely leads to the disruption of long pedigrees and thus reduction in the average levels of inbreeding *cf.* equilibrium expectations.

The model results demonstrate that stochastic factors related to transmission dynamics directly affect gene dynamics in parasite populations because they control the numbers and genotypic composition of individuals at the scale of the IP, which is also the scale at which sexual reproduction takes place. Thus, traditional genetic assumptions regarding equilibrium populations are violated in stochastic subdivided mating systems. Parasite systems are inherently stochastic, and thus the equilibrium of parasite populations (as with many biological systems) does not reflect constancy; rather it implies that the system varies consistently (May 2001; pp. 109). It is important to keep in mind that though we have assumed temporal stability this does not imply that the average probability that two parasites in an IP share a parent in the previous generation is constant but rather that it depends on the distributional characteristics of parasites among hosts (albeit with a consistent variance). However, the calculation of 'equilibrium' inbreeding values does not take into consideration the stochasticity of host parasite systems. Thus, even assuming panmixia at the scale of the IP, stochasticity in numbers of parasites at the scale of the IP which contribute offspring to the next generation could decrease (or increase) the number of effective breeders that contribute to the pedigree of any particular individual. As a consequence, levels of coancestry between individuals (and thus inbreeding of individuals) may accrue much more rapidly (or more slowly) in subdivided mating systems compared to 'equilibrium' expectations.

Finally, I turn to the effects of host–parasite population dynamics (i.e. stochastic subdivided breeding, host spatial structure and aggregated transmission dynamics) on the frequency of rare homozyygotes in a random sample of parasites. In populations where genotypic frequencies are HWE, we expect that an allele occurring with frequency q will occur as a homozygote at the frequency of q^2, which should act inhibit the spread of recessive traits (Cornell *et al.* 2003). However, aggregated transmission dynamics could enhance the spread of recessive traits by directly influencing genotypic frequencies in parasite populations such that rare traits occur at a prevalence proportional closer to q rather than q^2 (Cornell *et al.* 2003). My results are generally in agreement with those of Cornell *et al.* (2003), especially when levels of aggregated transmission are high and/or levels of hosts spatial structure are high (figure 4). Of even greater interest, however, are the patterns at low levels of aggregation (i.e. $k = 50$; i.e. distribution of parasites amongst hosts is approximately Poisson) and host spatial structure ($\varphi = 0$), where genetic patterns are predominantly driven by the fact that parasites exhibit a subdivided breeding system. At these low levels of aggregation and host spatial structure, my results

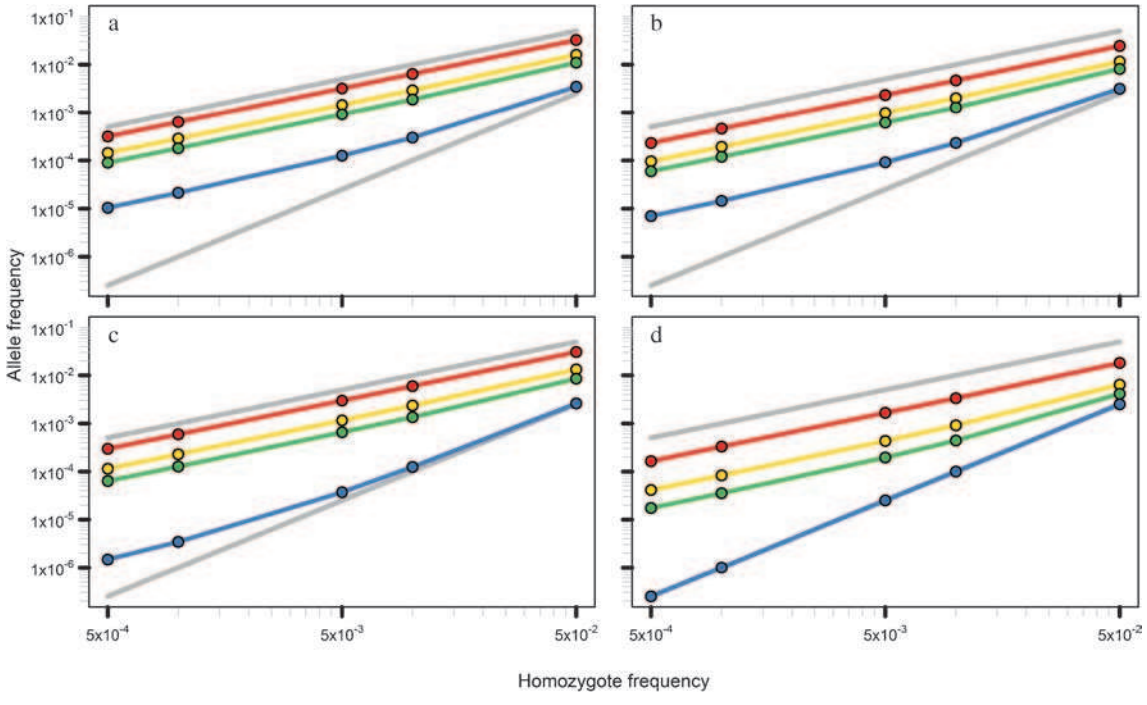

Figure 4. Rare homozygote frequencies in stochastic subdivided mating systems at five different initial allele frequencies (p : 1×10^{-4}, 5×10^{-4}, 1×10^{-3}, 5×10^{-3} and 1×10^{-2}). Lines indicating frequency of the allele (p) and homozygote frequencies under HWE expectations (p^2) are also given for reference (thick grey upper and lower lines, respectively). In each graph, the effects of four levels of aggregated transmission (k) are represented: 50 (blue), 1 (green), 0.5 (yellow) and 0.1 (red). The graphs represent data from simulations with no host spatial structure (i.e. $\varphi = 0$) at two different mean parasite intensities (2.5 parasites/host, a and b; 25 parasites/host, c and d) and two different parasite life history characteristics (ticks, a and c; nematodes, b and d).

indicate a strong negative effect of mean parasite burden on the frequency of rare homozygotes (figure 4). This result is a direct reflection of the effects of reduction in effective (versus census) population sizes in subdivided mating systems on the frequency of rare homozygotes (with the reduction being greater when the average number of individuals/IP is low; figure 4). Thus, my results indicate that while rare homozygote frequencies are affected by both aggregated transmission dynamics and host spatial structure, there may be an even more important intrinsic effect of a subdivided breeding system on the frequency of rare homozygotes.

In building theoretical models, it has long been realized that there is a trade-off between biological realism, generality and tractability (Levins 1966; see also Weisberg 2006). Thus, to build a general but tractable model I have sacrificed certain aspects of biological realism through some simplifying assumptions. In my opinion, of the assumptions made, two are the most critical: homogeneity among hosts and a direct lifecycle. Homogeneity of the infection process among hosts, with aggregation being driven solely by aggregation of parasites in the environment, is clearly unrealistic because hosts are likely to show heterogeneous levels of infection due to numerous factors, such as demography (host age and sex), behaviour and/or immune response (reviewed by Shaw *et al*. 1998). Similarly, the assumption of a direct lifecycle also is unrealistic because the life cycle of most parasites includes one to several intermediate hosts, and the presence of such intermediate hosts facilitates the mixing of parasite genes prior to reproduction in the definitive host (Rauch *et al.* 2005; Keeney *et al.* 2007). However, the flexible modelling framework developed can be easily modified to relax many of these assumptions, and I hope that future work in this direction will further our understanding of gene dynamics in host–parasite populations.

In conclusion, I utilize a flexible statistical distribution (the Logarithmic-Poisson mixture model) to integrate demographic factors associated with host–parasite systems into a genetic framework. My results indicate that deviations from HWE due to inbreeding may be directly due to biological factors intrinsically associated with host parasite dynamics (subdivided mating systems, aggregated transmission dynamics and host spatial structure), and that this inbreeding is poorly estimated by the equilibrium levels of inbreeding calculated based on constant levels of kin mating. My results also indicate that aggregated transmission dynamics may have direct epidemiological implications (e.g. the spread of drug resistance) because they both directly impact the observed frequency of rare homozygotes in parasite populations, a result that is in concordance with those of Cornell *et al.* (2003). However, even more importantly my results demonstrate that the increased frequency of rare homozygotes in parasite populations can be caused directly by the decrease in effective population size solely due to the presence of a subdivided breeding system.

Acknowledgements

I thank Gene Rhodes and Saptarshi Guha for early discussions on the model. I also thank the Department of Science and Technology, Government of India for financial support through a Ramanujan Fellowship Grant.

References

Anderson R. M., May R. M. and Gupta S. 1989 Non-linear phenomena in host-parasite interactions. *Parasitology* **99**, S59–S79.

Anderson T. J. C., Romeroabal M. E. and Jaenike J. 1995 Mitochondrial DNA and *Ascaris* microepidemiology - the composition of parasite populations from individual hosts, families and villages. *Parasitology* **110**, 221–229.

Ascunce M. S., Toups M. A., Kassu G., Fane J., Scholl K. and Reed D. L. 2013 Nuclear genetic diversity in human lice (*Pediculus humanus*) reveals continental differences and high inbreeding among worldwide populations. *PLoS One* **8**, e57619.

Blouin M. S., Yowell C. A., Courtney C. H. and Dame J. B. 1995 Host movement and the genetic structure of populations of parasitic nematodes. *Genetics* **141**, 1007–1014.

Brown A. H. D. 1979 Enzyme polymorphism in plant-populations. *Theor. Pop. Biol.* **15**, 1–42.

Bush A. O., Lafferty K. D., Lotz J. M. and Shostak A. W. 1997 Parasitology meets ecology on its own terms: Margolis *et al.* revisited. *J. Parasitol.* **83**, 575–583.

Caillaud D., Prugnolle F., Durand P., Theron A. and de Meeus T. 2006 Host sex and parasite genetic diversity. *Microb. Infect.* **8**, 2477–2483.

Charlesworth D. and Willis J. H. 2009 Fundamental concepts in genetics: the genetics of inbreeding depression. *Nat. Rev. Genet.* **10**, 783–796.

Chevillon C., Koffi B. B., Barre N., Durand P., Arnathau C. and de Meeus T. 2007 Direct and indirect inferences on parasite mating and gene transmission patterns - pangamy in the cattle tick *Rhipicephalus* (*Boophilus*) *microplus*. *Infect. Genet. Evol.* **7**, 298–304.

Churcher T. S., Schwab A. E., Prichard R. K. and Basanez 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.

Cornell S. 2005 Modelling nematode populations: 20 years of progress. *Trends Parasitol* **21**, 542–545.

Cornell S. J., Isham V. S. and Grenfell B. T. 2000 Drug-resistant parasites and aggregated infection - early-season dynamics. *J. Math. Biol.* **41**, 341–360.

Cornell S. J., Isham V. S., Smith G. and Grenfell B. T. 2003 Spatial parasite transmission, drug resistance, and the spread of rare genes. *Proc. Nat. Acad. Sci. USA* **100**, 7401–7405.

Criscione C. D. 2008 Parasite co-structure: Broad and local scale approaches. *Parasite* **15**, 439–443.

Criscione C. D. and Blouin M. S. 2005 Effective sizes of macroparasite populations: a conceptual model. *Trends Parasitol.* **21**, 212–217.

Criscione C. D., Poulin R. and Blouin M. S. 2005 Molecular ecology of parasites: elucidating ecological and microevolutionary processes. *Mol. Ecol.* **14**, 2247–2257.

Cutulle C., Jonsson N. N. and Seddon J. M. 2010 Multiple paternity in *Rhipicephalus* (*Boophilus*) *microplus* confirmed by microsatellite analysis. *Exp. Appl. Acarol.* **50**, 51–58.

de Meeus T., Humair P. F., Grunau C., Delaye C. and Renaud F. 2004 Non-Mendelian transmission of alleles at microsatellite loci: an example in *Ixodes ricinus*, the vector of Lyme disease. *Int. J. Parasitol.* **34**, 943–950.

Desloire S., Moro C. V., Chauve C. and Zenner L. 2006 Comparison of four methods of extracting DNA from *D. gallinae* (*Acari: Dermanyssidae*). *Vet. Res.* **37**, 725–732.

Dharmarajan G. 2008 Of genes and ticks: the population genetics of *Dermacentor variabilis* and *Ixodes texanus* infecting a wildlife host. Doctoral dissertation submitted to Purdue University, West Lafayette, USA.

Dharmarajan G. and Rhodes Jr O. E. 2011 Evaluating levels of PCR efficiency and genotyping error in DNA extracted from engorged and non-engorged female *Dermacentor variabilis* ticks. *Med. Vet. Entomol.* **25**, 109–112.

Dharmarajan G., Beasely J. C. and Rhodes Jr O. E. 2010 Spatial and temporal factors affecting parasite genotypes encountered by hosts: empirical data from American dog ticks (*Dermacentor variabilis*) parasitizing raccoons (*Procyon lotor*). *Int. J. Parasitol.* **7**, 787–795.

Dharmarajan G., Beasley J. C. and Rhodes Jr O. E. 2011 Heterozygote deficiencies in parasite populations: An evaluation of interrelated hypotheses in the raccoon tick, *Ixodes texanus*. *Heredity* **106**, 253–260.

Dybdahl M. F. and Storfer A. 2003 Parasite local adaptation: red queen versus suicide king. *Trends Ecol. Evol.* **18**, 523–530.

Dye 1994 Models for investigating genetic exchange in protozoan populations. In *Modelling vector-borne and other parasitic diseases* (ed. B. D. Perry and J. W. Hansen), pp. 165–176. ILRAD, Nairobi, Kenya.

Guzinski J., Bull C. M., Donnellan S. C. and Gardner M. G. 2009 Molecular genetic data provide support for a model of transmission dynamics in an Australian reptile tick, *Bothriocroton hydrosauri*. *Mol. Ecol.* **18**, 227–234.

Hartl D. L. and Clark A. G. 1989 *Principles of population genetics.* 2nd edition. Sinauer Associates, Sunderland.

Hasle G., Roed K. H. and Leinaas H. P. 2008 Multiple paternity in *Ixodes ricinus* (Acari: Ixodidae), assessed by microsatellite markers. *J. Parasitol.* **94**, 345–347.

Hedrick P. W. and Cockerham C. C. 1986 Partial inbreeding-equilibrium heterozygosity and the heterozygosity paradox. *Evolution* **40**, 856–861.

Keeney D. B., Waters J. M. and Poulin R. 2007 Clonal diversity of the marine trematode *Maritrema novaezealandensis* within intermediate hosts: the molecular ecology of parasite life cycles. *Mol. Ecol.* **16**, 431–439.

Keller L. F. and Waller D. M. 2002 Inbreeding effects in wild populations. *Trends Ecol. Evol.* **17**, 230–241.

Lafferty K. D. and Kuris A. M. 2002 Trophic strategies, animal diversity and body size. *Trends Ecol. Evol.* **17**, 507–513.

Levins R. 1966 Strategy of model building in population biology. *Am. Sci.* **54**, 421–431.

Luo H. Y., Nie P., Zhang Y. A., Yao W. J. and Wang G. T. 2003 Genetic differentiation in populations of the cestode *Bothriocephalus acheilognathi* (Cestoda, Pseudophyllidea) as revealed by eight microsatellite markers. *Parasitology* **126**, 493–501.

Lymbery A. J. and Thompson R. C. A. 2012 The molecular epidemiology of parasite infections; Tools and applications. *Mol. Biochem. Parasitol.* **181**, 102–116.

Mackinnon M. J. and Hastings I. M. 1998 The evolution of multiple drug resistance in malaria parasites. *Trans. Roy. Soc. Trop. Med. Hyg.* **92**, 188–195.

Malécot G. 1969. *The mathematics of heredity* (translated by D. M. Yermanos). Freeman, San Fransico, California.

May R. M. 2001 *Stability and complexity in model ecosystems*. Princeton University Press, Princeton, New Jersey, USA.

McCoy K. D., Boulinier T. and Tirard C. 2005 Comparative host-parasite population structures: disentangling prospecting and dispersal in the black-legged kittiwake *Rissa tridactyla*. *Mol. Ecol.* **14**, 2825–2838.

McCoy K. D., Boulinier T., Tirard C. and Michalakis Y. 2003 Host-dependent genetic structure of parasite populations: Differential dispersal of seabird tick host races. *Evolution* **57**, 288–296.

Mharakurwa S. 2004 *Plasmodium falciparum* transmission rate and selection for drug resistance: a vexed association or a key to successful control? *Int. J. Parasitol.* **34**, 1483–1487.

Nadler S. A. 1995 Microevolution and the genetic structure of parasite populations. *J. Parasitol.* **81**, 395–403.

Nadler S. A., Hafner M. S., Hafner J. C. and Hafner D. J. 1990 Genetic differentiation among chewing louse populations (*Mallophaga, Trichodectidae*) in a pocket gopher contact zone (*Rodentia, Geomyidae*). *Evolution* **44**, 942–951.

Nee S., West S. A. and Read A. F. 2002 Inbreeding and parasite sex ratios. *Proc. R. Soc. London., B Biol.* **269**, 755–760.

Nei M. 1975 *Molecular population genetics and evolution*. North-Holland, Amsterdam.

Paul R. E. L., Packer M. J., Walmsley M., Lagog M., Ranfordcartwright L. C., Paru R. and Day K. P. 1995 Mating patterns in malaria parasite populations of Papua New Guinea. *Science* **269**, 1709–1711.

Picard D., Scurrah M. and Mugniery D. 2004 Inbreeding and population structure of the potato cyst nematode (*Globodera pallida*) in its native area (Peru). *Mol. Ecol.* **13**, 2899–2908.

Pielou E. C. 1976 *Mathematical ecology*. Wiley, New York, USA.

Plantard O., Picard D., Valette S., Scurrah M., Grenier E. and Mugniery D. 2008 Origin and genetic diversity of Western European populations of the potato cyst nematode (*Globodera pallida*) inferred from mitochondrial sequences and microsatellite loci. *Mol. Ecol.* **17**, 2208–2218.

Price P. W. 1980 *Evolutionary biology of parasites*. Princeton University Press, Princeton, USA.

Rauch G., Kalbe M. and Reusch T. B. H. 2005 How a complex life cycle can improve a parasite's sex life. *J. Evol. Biol.* **18**, 1069–1075.

Redman E., Grillo V., Saunders G., Packard E., Jackson F., Berriman M. and Gilleard J. S. 2008 Genetics of mating and sex determination in the parasitic nematode *Haemonchus contortus*. *Genetics* **180**, 1877–1887.

Rostron J. 1978 Computation of inbreeding coefficients. *Ann. Hum. Genet.* **41**, 469–475.

Ruiz-Lopez M. J., Chaskelson S., Gompper M. E. and Eggert L. S. 2012 Multiple paternity in the american dog tick *Dermacentor variabi*lis (acari: Ixodidae). *J. Parasitol.* **98**, 498–501.

Ryman N. 1994 Supportive breeding and effective population size - differences between inbreeding and variance effective numbers. *Conserv. Biol.* **8**, 888–890.

Saleem M. and Gliddon C. J. 1986 Effect of breeding system on the genotypic structure in trifolium species. *Pakistan J. Bot.* **18**, 183–188.

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

Schmidt K. F. 1995 Malaria research - inbred parasites may spur resistance. *Science* **269**, 1670.

Schwab A. E., Churcher T. S., Schwab A. J., Basanez M. G. and Prichard R. K. 2007 An analysis of the population genetics of potential multi-drug resistance in *Wuchereria bancrofti* due to combination chemotherapy. *Parasitology* **134**, 1025–1040.

Shaw D. J. and Dobson A. P. 1995 Patterns of macroparasite abundance and aggregation in wildlife populations: A quantitative review. *Parasitology* **111**, S111–S133.

Shaw D. J., Grenfell B. T. and Dobson A. P. 1998 Patterns of macroparasite aggregation in wildlife host populations. *Parasitology* **117**, 597–610.

Shields M. W. 1982 *Philopatry, inbreeding, and the evolution of sex*. State University of New York Press, Albany, New York, USA.

Sirea C., Durand P., Pointier J. and Theron A. 2001 Genetic diversity of *Schistosoma mansoni* within and among individual hosts (*Rattus rattus*): infrapopulation differentiation at microspatial scale. *Int. J. Parasitol.* **31**, 1609–1616.

Smith G., Grenfell B. T., Isham V. and Cornell S. 1999 Anthelmintic resistance revisited: under-dosing, chemoprophylactic strategies, and mating probabilities. *Int. J. Parasitol* **29**, 77–91.

Sonenshine D. E. 1991 *Biology of ticks*, vol. 1. Oxford University Press, New York, USA.

Sonenshine D. E. 1993 *Biology of ticks*, vol. 2. Oxford University Press, New York, USA.

Templeton A. R. 1986 Coadaptation and outbreeding depression. In *Conservation biology: the science of scarcity and diversity* (ed. M. E. Soule), pp. 105–116. Sinauer Associates, Sunderland, USA.

Thiele E. A., Sorensen R. E., Gazzinelli A. and Minchella D. J. 2008 Genetic diversity and population structuring of *Schistosoma mansoni* in a Brazilian village. *Int. J. Parasitol.* **38**, 389–399.

Vilas R., Vazquez-Prieto S. and Paniagua E. 2012 Contrasting patterns of population genetic structure of *Fasciola hepatica* from cattle and sheep: Implications for the evolution of anthelmintic resistance. *Infect. Genet. Evol.* **12**, 45–52.

Waser N. M. and Price M. V. 1989 Optimal outcrossing in *Ipomopsis aggregata* - seed set and offspring fitness. *Evolution* **43**, 1097–1109.

Weisberg M. 2006 Forty years of 'The strategy': Levins on model building and idealization. *Biol. Philos.* **21**, 623–645.

Wielgoss S., Taraschewski H., Meyer A. and Wirth T. 2008 Population structure of the parasitic nematode *Anguillicola crassus*, an invader of declining North Atlantic eel stocks. *Mol. Ecol.* **17**, 3478–3495.

Zhou C. H., Yuan K., Tang X. L., Hu N. Y. and Peng W. D. 2011 Molecular genetic evidence for polyandry in *Ascaris suum*. *Parasitol. Res.* **108**, 703–708.

Received 17 June 2014, in revised form 25 September 2014; accepted 1 October 2014
Unedited version published online: 27 October 2014
Final version published online: 3 March 2015