JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.12489

Individual genetic diversity and probability of infection by avian malaria parasites in blue tits (*Cyanistes caeruleus*)

E. S. FERRER*†, V. GARCÍA-NAVAS*†, J. J. SANZ‡ & J. ORTEGO*§

*Grupo de Investigación de la Biodiversidad Genética y Cultural, Instituto de Investigación en Recursos Cinegéticos - IREC (CSIC, UCLM, JCCM), Ciudad Real, Spain
†Departamento de Ciencias Ambientales, Facultad de Ciencias Ambientales y Bioquímica, Universidad de Castilla-La Mancha, Toledo, Spain
‡Departamento de Ecología Evolutiva, Museo Nacional de Ciencias Naturales (CSIC), Madrid, Spain
§Conservation and Evolutionary Genetics Group, Department of Integrative Ecology, Estación Biológica de Doñana, EBD-CSIC, Seville, Spain

Keywords:

avian malaria;
disease;
genetic diversity;
heterozygosity–fitness correlations;
Plasmodium.

Abstract

Understanding the importance of host genetic diversity for coping with parasites and infectious diseases is a long-standing goal in evolutionary biology. Here, we study the association between probability of infection by avian malaria (*Plasmodium relictum*) and individual genetic diversity in three blue tit (*Cyanistes caeruleus*) populations that strongly differ in prevalence of this parasite. For this purpose, we screened avian malaria infections and genotyped 789 blue tits across 26 microsatellite markers. We used two different arrays of markers: 14 loci classified as neutral and 12 loci classified as putatively functional. We found a significant relationship between probability of infection and host genetic diversity estimated at the subset of neutral markers that was not explained by strong local effects and did not differ among the studied populations. This relationship was not linear, and probability of infection increased up to values of homozygosity by locus (*HL*) around 0.15, reached a plateau at values of *HL* from 0.15 to 0.40 and finally declined among a small proportion of highly homozygous individuals ($HL > 0.4$). We did not find evidence for significant identity disequilibrium, which may have resulted from a low variance of inbreeding in the study populations and/or the small power of our set of markers to detect it. A combination of subtle positive and negative local effects and/or a saturation threshold in the association between probability of infection and host genetic diversity in combination with increased resistance to parasites in highly homozygous individuals may explain the observed negative quadratic relationship. Overall, our study highlights that parasites play an important role in shaping host genetic variation and suggests that the use of large sets of neutral markers may be more appropriate for the study of heterozygosity–fitness correlations.

Introduction

Genetic diversity is essential for individuals and populations to cope with novel and changing environmental conditions (Coltman *et al.*, 1999; Willi *et al.*, 2006). A plethora of studies have reported that reduced genetic diversity decreases fitness (David, 1998; Coltman & Slate, 2003; Chapman *et al.*, 2009; Szulkin *et al.*, 2010) and increases extinction risk of natural populations (Saccheri *et al.*, 1998; Spielman *et al.*, 2004). For this reason, understanding the consequences of genetic diversity on different components of fitness is a central question in evolutionary and conservation biology (Keller & Waller, 2002; Chapman *et al.*, 2009; Szulkin *et al.*, 2010). An important negative consequence of reduced levels of individual genetic diversity is a lower resistance to parasites (e.g. Coltman *et al.*, 1999; Hawley *et al.*, 2005; Acevedo-Whitehouse *et al.*, 2006; Luikart

Correspondence: Joaquín Ortego, Conservation and Evolutionary Genetics Group, Department of Integrative Ecology, Estación Biológica de Doñana, EBD-CSIC, Avda. Américo Vespucio s/n, E-41092 Seville, Spain. Tel.: +34 954 232 340; fax: +34 954 621 125; e-mail: joaquin.ortego@csic.es

et al., 2008; Rijks *et al.*, 2008), a phenomenon that can also lead to higher disease susceptibility at the whole population level if severe demographic bottlenecks have depleted genetic variability (Obrien & Evermann, 1988; e.g. Meagher, 1999; Pearman & Garner, 2005; Whiteman *et al.*, 2006; Whitehorn *et al.*, 2011). Different genetic mechanisms can explain the association between individual genetic diversity and increased susceptibility to parasites and diseases (Keller & Waller, 2002). Genetically more diverse individuals are expected to have a higher probability of carrying adaptive alleles implicated in immune response and recognition of pathogens that can confer resistance to a wider array of parasite species or genotypes (Keller & Waller, 2002; Puurtinen *et al.*, 2004; Reid *et al.*, 2007; Fossoy *et al.*, 2009). In turn, high levels of individual genetic diversity can also reduce the susceptibility to diseases through overdominance at genes not directly involved in immune response, but involved in the capacity to clear and survive an infection (Coltman *et al.*, 1999). For instance, high enzyme polymorphism at genes affecting physiological processes such as oxygen consumption, cellular homeostasis and metabolic efficiency is likely to ultimately influence vulnerability to pathogens (Mitton & Grant, 1984; Ferguson & Drahushchak, 1990; Kristensen *et al.*, 2002; Pedersen *et al.*, 2005; Luong *et al.*, 2007). Finally, individuals with low genetic diversity have a higher chance of expressing recessive deleterious alleles, which would decrease their capacity to resist infectious diseases and survive (Coltman *et al.*, 1999; Keller & Waller, 2002). Thus, low genetic diversity can increase susceptibility to pathogens through its influence on different immunological, physiological and biochemical mechanisms (Coltman *et al.*, 1999; Acevedo-Whitehouse *et al.*, 2003).

The relationship between parasitism and host genetic diversity can also potentially vary across different genomic regions not directly involved in immune defence (Szulkin & David, 2011). Microsatellite loci, the most common markers used in HFC studies, have been generally assumed to be evolutionary neutral, but there is growing evidence on their implication in different biological processes (Li *et al.* 2004). Differences in the functionality of the markers employed to estimate heterozygosity can help to elucidate the relative importance of genomewide/inbreeding effects ('general effect hypothesis') or effects at a single or few loci ('local and direct effect hypothesis') on observed heterozygosity–fitness correlations (HFC) (David, 1998; Hansson & Westerberg, 2002). A few recent empirical studies have shown that microsatellite markers located in functional and nonfunctional regions of the genome can behave differentially in terms of signal intensity and patterns of HFC (see Da Silva *et al.*, 2009; Kupper *et al.*, 2010; Olano-Marin *et al.*, 2011a,b; Szulkin & David, 2011; Laine *et al.*, 2012). However, both classes of loci have not been yet employed to study the association between host genetic diversity and resistance to infectious diseases, an approach that could contribute to get a better understanding on the underlying mechanisms behind parasite–host genotype interactions (Szulkin & David, 2011).

The strength of the relationship between fitness and individual genetic diversity or inbreeding has been often found to increase under stressful and harsh environmental conditions (reviewed in Armbruster & Reed, 2005; Fox & Reed, 2011). Previous studies have shown that the stress imposed by parasites can increase the strength of HFC (e.g. Coltman *et al.*, 1999; Voegeli *et al.*, 2012) and exacerbate the negative effects of inbreeding (e.g. Carr *et al.*, 2003). However, no study has yet analysed differences in heterozygosity-dependent parasite resistance across populations experiencing contrasting patterns of parasite pressure. If the strength of the association between parasitism and host genetic diversity increases during disease outbreaks or in populations experiencing high parasitism rates, this could buffer the expected loss of genetic diversity due to parasite-mediated demographic declines and ultimately increase the chance of population viability (e.g. Coltman *et al.*, 1999; see also Forcada & Hoffman, 2014). Thus, studies comparing populations with differences in parasite pressure can provide insight into the population's responses to infectious diseases and help to further understand the demographic and genetic consequences of parasites and its role as mediators of HFC in wild populations (Coltman *et al.*, 1999; Voegeli *et al.*, 2012).

Here, we study the association between individual genetic diversity and probability of avian malaria infections in three Mediterranean blue tit (*Cyanistes caeruleus*) populations that strongly differ in prevalence of these parasites (Ferrer *et al.*, 2012). The negative consequences of avian malaria on different components of fitness have been reported in different species (e.g. Van Riper *et al.*, 1986; Merino *et al.*, 2000; Sol *et al.*, 2003; Ortego *et al.*, 2008; Atkinson & Samuel, 2010; Lachish *et al.*, 2011). We have previously found six different lineages of avian malaria parasitizing our studied blue tit populations, but the lineage SGS1 (*Plasmodium relictum*) is involved in most infections (Ferrer *et al.*, 2012). Previous studies have reported that both chronic and acute infections by *P. relictum* have negative fitness consequences in blue tits (Knowles *et al.*, 2010; Lachish *et al.*, 2011), including effects on mortality and survival probability in both this and other bird species (Atkinson *et al.*, 2000; Atkinson & Samuel, 2010; Lachish *et al.*, 2011). Thus, this parasite probably exerts important evolutionary pressures that may shape host genetic composition and diversity both at specific loci involved in immune defence and at other parts of the genome related to physiological condition or health status (Ortego *et al.*, 2007a; Reid *et al.*, 2007; Fossoy *et al.*, 2009). Mediterranean blue tits often have to cope with the

extreme heterogeneity of fragmented forest habitats (Blondel & Aronson, 1999), show a sedentary behaviour and tend to disperse across very short distances (Blondel *et al.*, 2006; Ortego *et al.*, 2011a), factors that may increase inbreeding rates and favour the detection of heterozygosity–fitness correlations (HFC) (Balloux *et al.*, 2004; Szulkin *et al.*, 2010). Accordingly, previous studies have found HFC in blue tits for individual quality, female fecundity, reproductive performance and survival (Foerster *et al.*, 2003; Garcia-Navas *et al.*, 2009; Olano-Marin *et al.*, 2011a,b). Thus, Mediterranean blue tits constitute an ideal study system to explore the association between individual genetic diversity and probability of infection by avian malaria.

In this study, we have genotyped 789 blue tits across 26 polymorphic microsatellite markers to estimate individual genetic diversity and investigate its association with the probability of infection by avian malaria parasites screened using a highly efficient polymerase chain reaction (Waldenström *et al.*, 2004). We used two different arrays of markers, classified as potentially neutral (14 loci) or functional (12 loci) by considering whether the genomic region where the markers are located is transcribed to RNA (*sensu* Olano-Marin *et al.*, 2011a,b; see also Da Silva *et al.*, 2009; Kupper *et al.*, 2010; Laine *et al.*, 2012). With this information, we tested the following specific hypotheses: we predicted that (i) more heterozygous individuals have a lower probability of infection by avian malaria parasites and (ii) the strength of such relationship increases in populations with a higher prevalence of the studied parasite; (iii) we analysed whether the association between heterozygosity and parasitism reflects a genomewide effect (general effect hypothesis; Weir & Cockerham, 1973; David, 1998) or is explained by single-locus effects due to their functional nature (direct effect hypothesis; David, 1998; Hansson & Westerberg, 2002) or strong linkage disequilibrium with other functional loci (local effect hypothesis; David, 1998; Hansson *et al.*, 2001; Hansson & Westerberg, 2002); and (iv) finally, to further understand the underlying mechanism behind the correlation between heterozygosity and probability of infection, we analysed whether this association is better explained by functional loci, neutral loci or a combination of both. Neutral markers can cause this association either by general effects or by local effects if they happen to be linked to functional loci. Direct effects and strong local effects, however, are only likely to be caused by functional markers (Olano-Marin *et al.*, 2011a, b; Laine *et al.*, 2012).

Materials and methods

Study species and sampling

The study was conducted in three blue tit populations located in Montes de Toledo (central Spain): Quintos de Mora (Toledo province, 39°25′N 4°04′W; 2008–2011 breeding seasons), San Pablo de los Montes (Toledo province, 39°31′N 4°21′W; 2011 breeding season) and Cabañeros (Ciudad Real and Toledo provinces, 39°24′N, 3°35′W; 2008 breeding season) (see Table 1 for more details). Adult birds (Quintos de Mora: 228 females and 175 males; San Pablo de los Montes: 160 females and 119 males; Cabañeros: 54 females and 53 males) were captured when feeding 8-day-old chicks by means of a spring-trap shutting the entrance hole as the bird entered the nest. Individuals were weighed using a pocket balance (accuracy ± 0.1 g), and their wing length measured to the nearest 0.1 mm using a stopped ruler. Birds were sexed by the presence/absence of a brood patch and aged according to Svensson (1992) as juveniles (yearlings) or experienced breeders (second-year and older birds). All birds were individually marked with aluminium rings for further identification. Blood samples (≤ 25 μL) for genetic analyses were obtained by brachial venipuncture and stored in ethanol 96%.

Microsatellite markers

We genotyped blue tits across 27 polymorphic microsatellite markers (Table 2). These markers were classified as presumably functional or neutral as described by Olano-Marin *et al.* (2011a,b) (see also Olano-Marin *et al.*, 2010). We used NucleoSpin Blood kits (Macherey-Nagel, Duren, Germany) to extract and purify genomic DNA from the blood samples. Approximately 5 ng of template DNA was amplified in 10 μL reaction volumes containing 1X reaction buffer (67 mM Tris–HCl, pH 8.3, 16 mM $(NH_4)_2SO_4$, 0.01% Tween-20, Eco-Start Reaction Buffer; Ecogen, Barcelona, Spain), 2 mM $MgCl_2$, 0.2 mM of each dNTP, 0.15 μM of each dye-labelled primer (FAM, PET, VIC or NED) and 0.1 U of *Taq* DNA EcoStart Polymerase (Ecogen). The PCR program used was 9-min denaturation at 95 °C followed by 40 cycles of 30 s at 94 °C, 45 s at the annealing temperature (Table 2) and 45 s at 72 °C, ending with a 10-min final elongation stage at 72 °C. Amplification products were electrophoresed using an ABI 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA), and genotypes were scored using GENEMAPPER 3.7 (Applied Biosystems).

Linkage disequilibrium (LD) between loci and deviations from Hardy–Weinberg equilibrium (HWE) were tested with GENEPOP web version 4.0.10 (Raymond & Rousset, 1995). Probabilities of significance were computed applying a Markov chain method (Guo & Thompson, 1992) using 200 batches and 1000 iterations per batch as implemented in GENEPOP. The LD correlation coefficient (*r*LD) between alleles at different loci was used to estimate the magnitude of LD. *r*LD was computed with the program LinkDos (http://genepop.curtin.edu.au/linkdos.html; Garnier-Gere & Dillmann, 1992). Sequential Bonferroni corrections were applied to account for multiple comparisons (Rice, 1989).

Table 1 Geographical location of the studied populations and subpopulations of blue tits (*Cyanistes caeruleus*) in Montes de Toledo, sample size (*n*), genetic diversity (mean ± SE) estimated as homozygosity by locus (*HL*) at all loci (HL_{Total}) and neutral ($HL_{Neutral}$) and functional loci ($HL_{Functional}$) separately, and prevalence of avian malaria (lineage SGS1).

Population/Subpopulation	Latitude	Longitude	*n*	HL_{Total}	$HL_{Neutral}$	$HL_{Functional}$	Prevalence, %
Quintos de Mora			403	0.218 ± 0.004	0.188 ± 0.005	0.258 ± 0.007	17.12
Gil García	39.37976	−4.12593	215	0.223 ± 0.006	0.186 ± 0.007	0.274 ± 0.009	13.95
Valdeyernos	39.44008	−4.09091	156	0.214 ± 0.007	0.194 ± 0.009	0.240 ± 0.010	18.59
Fuente del Común	39.36316	−4.07830	32	0.201 ± 0.012	0.175 ± 0.016	0.235 ± 0.025	31.25
San Pablo de los Montes			279	0.210 ± 0.005	0.187 ± 0.006	0.242 ± 0.007	41.94
Marchés	39.54413	−4.37361	101	0.220 ± 0.009	0.188 ± 0.010	0.268 ± 0.014	42.57
Morra	39.52895	−4.32851	80	0.207 ± 0.008	0.194 ± 0.011	0.221 ± 0.014	38.75
Fuente Cerecera	39.51296	−4.27795	40	0.203 ± 0.009	0.189 ± 0.014	0.221 ± 0.014	50.00
Fuenlabrada	39.52414	−4.25709	46	0.198 ± 0.010	0.168 ± 0.013	0.240 ± 0.018	34.78
Robledillo	39.50862	−4.35865	12	0.219 ± 0.026	0.202 ± 0.028	0.243 ± 0.030	58.33
Cabañeros			107	0.219 ± 0.007	0.203 ± 0.010	0.241 ± 0.011	13.08
Anchurones	39.34241	−4.36783	33	0.207 ± 0.014	0.182 ± 0.018	0.240 ± 0.022	21.21
Brezoso	39.36307	−4.35798	74	0.225 ± 0.008	0.213 ± 0.012	0.241 ± 0.013	9.46

Table 2 Microsatellite loci used to genotype blue tits: chromosome location in the zebra finch (*Taeniopygia guttata*) genome, category (neutral or functional), number of alleles (*K*), expected heterozygosity (H_E), observed heterozygosity (H_O) (based on 789 individuals) and annealing temperature (*T*) for each locus.

Locus	Chromosome	Category	*K*	H_E	H_O	*T* (°C)	References
Pca7	1	Neutral	17	0.90	0.88	60	Dawson *et al.* (2000)
Pocc6	2	Neutral	25	0.86	0.82	62	Bensch *et al.* (1997)
Pat-MP2-43	2	Neutral	12	0.55	0.54	59	Otter *et al.* (1998)
Pca8	2	Neutral	70	0.96	0.93	53	Dawson *et al.* (2000)
Ase18	3	Neutral	20	0.86	0.83	60	Richardson *et al.* (2000)
Pca3	4	Neutral	31	0.89	0.87	55	Dawson *et al.* (2000)
Pdoμ5	4	Neutral	42	0.91	0.81	46	Griffith *et al.* (1999)
PK12	5	Neutral	25	0.86	0.86	62	GenBank Acc. no.: AF041466
Mcyμ4	5	Neutral	18	0.84	0.82	50	Double *et al.* (1997)
Pca9	7	Neutral	16	0.81	0.80	62	Dawson *et al.* (2000)
Pocc1	7	Neutral	16	0.85	0.84	55	Bensch *et al.* (1997)
Pca4	8	Neutral	19	0.75	0.73	60	Dawson *et al.* (2000)
PK11	Unassigned	Neutral	13	0.81	0.81	52	GenBank Acc. no.: AF041465
Pca2	Unassigned	Neutral	15	0.76	0.67	60	Dawson *et al.* (2000)
CcaTgu8	2	Functional	8	0.61	0.58	63	Olano-Marin *et al.* (2010)
CcaTgu7	2	Functional	8	0.75	0.76	55	Olano-Marin *et al.* (2010)
CcaTgu11	3	Functional	7	0.73	0.71	60	Olano-Marin *et al.* (2010)
TG05-053	5	Functional	9	0.77	0.77	55	Dawson *et al.* (2010)
TG05-046	5	Functional	4	0.49	0.52	55	Dawson *et al.* (2010)
CcaTgu15	5	Functional	9	0.65	0.64	60	Olano-Marin *et al.* (2010)
CcaTgu14	5	Functional	22	0.83	0.78	55	Olano-Marin *et al.* (2010)
Tgu07	6	Functional	7	0.67	0.69	55	Slate *et al.* (2007)
PiJ14	7	Functional	18	0.88	0.87	60	Olano-Marin *et al.* (2010)
CcaTgu19	10	Functional	51	0.93	0.91	60	Olano-Marin *et al.* (2010)
Tg13-017	13	Functional	21	0.86	0.84	60	Dawson *et al.* (2010)
CcaTgu25*	18	Functional	26	0.85	0.63	63	Olano-Marin *et al.* (2010)
CcaTgu28	23_random	Functional	13	0.73	0.72	60	Olano-Marin *et al.* (2010)

*This locus deviated from Hardy–Weinberg equilibrium across all populations/years and was discarded for further analyses.

Genetic structure

We analysed patterns of spatial genetic structure using the Bayesian Markov chain Monte Carlo clustering analysis implemented in the program STRUCTURE 2.3.3 (Pritchard *et al.*, 2000; Falush *et al.*, 2003; Hubisz *et al.*, 2009). STRUCTURE assigns individuals to populations based on their multilocus genotypes (Pritchard *et al.*, 2000; Falush *et al.*, 2003). We ran STRUCTURE assuming correlated allele frequencies and admixture and conducted ten independent runs for each value of $K = 1$–10 to estimate the 'true' number of clusters with

200 000 MCMC cycles, following a burn-in step of 100 000 iterations. The number of populations best fitting the data set was defined using both log probabilities [Pr(X|K)] (Pritchard *et al.*, 2000) and the ΔK method (Evanno *et al.*, 2005), as implemented in STRUCTURE HARVESTER (Earl & vonHoldt, 2012).

Heterozygosity, identity disequilibrium and inbreeding

We used two metrics to estimate individual genetic diversity: (i) uncorrected heterozygosity (H_O), calculated as the proportion of loci at which an individual is heterozygous; (ii) homozygosity by locus (*HL*), a microsatellite-derived measure that improves heterozygosity estimates in open populations by weighting the contribution of each locus to the homozygosity value depending on their allelic variability (Aparicio *et al.*, 2006). H_O and *HL* were calculated using CERNICALIN, an EXCEL spreadsheet available on request.

Identity disequilibrium (ID), defined as the correlation in heterozygosity and/or homozygosity across loci within individuals (Weir & Cockerham, 1973), arises when there is enough variation in individual inbreeding coefficients within the population (Balloux *et al.*, 2004; Szulkin *et al.*, 2010). We used two methods to analyse the presence of identity disequilibrium and test whether heterozygosity measured at our set of microsatellite loci was representative of genomewide inbreeding: (i) we calculated heterozygosity–heterozygosity correlations (HHC) following Balloux *et al.* (2004). If our microsatellites markers carry information about genomewide levels of heterozygosity, then comparing two random subsets of such markers should yield a positive, significant correlation (Balloux *et al.*, 2004). The correlation is obtained by repeatedly and randomly dividing the genotyped loci in half and calculating an estimate of individual multilocus heterozygosity for both sets of loci. The mean correlation between these sets will then yield the HHC (Balloux *et al.*, 2004). We used the R package 'RHH' to run 1000 randomizations of the markers and estimate the average HHC coefficient (r) and the 95% confidence intervals (Alho *et al.*, 2010); (ii) we calculated the parameter g^2 described by David *et al.* (2007), a central measure of identity disequilibrium that quantifies the excess of multiple heterozygotes at the studied loci relative to the expectation under random association (i.e. covariance in heterozygosity) (see Szulkin *et al.*, 2010). We used the software RMES to calculate g^2 and test whether this parameter differs significantly from zero (David *et al.*, 2007).

Pedigree data are only available for Quintos de Mora population because the two other studied populations have been only monitored during one (San Pablo de los Montes; 2001 breeding season) or a few breeding seasons (Cabañeros; 2006–2008 breeding seasons). Individual inbreeding coefficients (f) were estimated from the reconstructed pedigree using PEDIGREE VIEWER (http://www-personal.une.edu.au/~bkinghor/pedigree.htm). Only one recruit to the study populations was highly inbred, and we compared its level of multilocus heterozygosity with those values obtained for other individuals also captured within Quintos de Mora during the same breeding season (e.g. Ortego *et al.*, 2011b).

Screening for avian malaria infections

Each individual was screened for avian malaria infection using a highly efficient nested PCR protocol that amplifies a 524-bp fragment (including primers) of the mitochondrial cytochrome *b* gene of both *Plasmodium* and *Haemoproteous* parasites (Waldenström *et al.*, 2004). This method consists in two rounds, an initial 20 cycles of PCR using the primers HAEMNF and HAEMNR2 and a final 35 cycles of PCR using the internally nested primers HAEMF and HAEMR2 (Waldenström *et al.*, 2004). For the second PCR, 1 μL of the PCR product from the initial PCR was used as template. All PCRs were performed in 25 μL volumes, and we routinely used positive (i.e. DNA from individuals with known malarial infections) and negative controls (i.e. samples with ddH_2O instead of genomic DNA as template) to ascertain that the outcome of each PCR run was not affected by contamination (Waldenström *et al.*, 2004). The PCR program, thermal profile and reagent proportions were as described by Waldenström *et al.* (2004), with the exception of using a 9-min denaturation at 95 °C rather than 3 min at 94 °C because we used a hot-start polymerase (EcoStart; Ecogen). Reagents were the same used for microsatellite amplifications. Amplification of the 26 microsatellite loci above described was used as a control for DNA quality. These amplifications were successful in all cases. Further, negative infections were confirmed by repeated PCR. Positive or negative second-round PCR products (i.e. birds having or not gametocytes or merozoites in their blood stream) were scored by electrophoresis on 2% agarose gels stained with ethidium bromide and determining the presence/absence of a band of the expected size under UV light. PCR products from positive samples were purified using NucleoSpin Extract II (Macherey-Nagel) kits and bidirectionally sequenced on an ABI 310 Genetic Analyser (Applied Biosystems). Sequences were edited and aligned using the program SEQUENCHER 5.0 (GeneCodes Corporation, Ann Arbor, MI, USA).

Statistical analyses: general effects

We analysed the relationship between multilocus heterozygosity and probability of infection by avian malaria parasites using generalized linear mixed models (GLMMs) with a binomial error structure and

a logit link function. We constructed two separate GLMMs, fitting *HL* or H_O as explanatory variables together with nongenetic terms (fixed factors: sex and age; covariates: body mass and wing length) that could potentially influence avian malaria infections. Additionally, we ran GLMMs including as predictors heterozygosity (*HL* or H_O) calculated with the subset of neutral ($HL_{Neutral}$ or $H_{ONeutral}$) and functional markers ($HL_{Functional}$ or $H_{OFunctional}$). Heterozygosity measured with neutral and functional markers are not correlated (see results section) and, for this reason, both variables were included together into the same models. We included breeding pair, year and population and subpopulation identity as random effects in all the models. Initially, each GLMM was constructed with all explanatory terms fitted, including first-order interactions and quadratic effects to account for potential nonlinear relationships. Final models were selected following a backward procedure, by progressively eliminating nonsignificant variables. The significance of the remaining variables was tested again until no additional variable reached significance. The result is the minimal most adequate model for explaining the variability in the response variable, where only the significant explanatory variables are retained. All random effects were retained in the final models. GLMMs were run using the package lme4 in R 3.0.0 (R Development Core Team, 2012).

Statistical analyses: local effects

For those cases where models included an estimate of multilocus heterozygosity, we followed the approach described in Szulkin *et al.* (2010), to test whether single-locus heterozygosity (SLH) effects explain more variance than multilocus heterozygosity (MLH) effects. In brief, we used a *F*-test ratio test to compare the model including MLH (i.e. the model described in the previous section) and a similar model in which we replaced MLH with 'normalized' SLH at all markers incorporated as binary variables (Szulkin *et al.*, 2010; e.g. Olano-Marin *et al.*, 2011a,b; Ruiz-Lopez *et al.*, 2012). Only individuals with complete genotypes for all loci were included in these analyses (e.g. Ruiz-Lopez *et al.*, 2012).

We also explored the effects of SLH by fitting one GLMM per locus, that is 26 models in total. We calculated the effect size for each locus as the partial correlation coefficient obtained from their respective models (Nakagawa & Cuthill, 2007). We used a binomial sign test to analyse whether positive and negative effects occurred equally and χ^2 tests to investigate whether the number of positive/negative or significant/nonsignificant effects differed between the subsets of neutral and functional loci. Finally, we used a general linear model (GLM) analysis to test whether the absolute effect size of SLH was associated with marker diversity (estimated as H_E, H_O and allelic richness; Table 2) or differed between the groups of neutral and functional loci (e.g. Olano-Marin *et al.*, 2011a; Ruiz-Lopez *et al.*, 2012).

Results

Microsatellite data

Observed heterozygosity at each locus ranged from 0.51 to 0.93, with 4–70 alleles per locus (Table 2). After applying sequential Bonferroni corrections to compensate for multiple statistical tests, loci Pdoμ5, CcaTgu8 and CcaTgu15 showed significant deviations from HWE in the Cabañeros population. Locus CcaTgu15 also showed deviations in Quintos de Mora population during a single study year. Locus CcaTgu25 showed deviations across all study years and populations, and it was discarded for further analyses. After sequential Bonferroni corrections, significant LD was detected for loci TG05-053/CcaTgu15 in Cabañeros population and loci Ase18/CcaTgu14 in San Pablo population. For the two loci not assigned to any chromosome (Pca2 and PK11; Table 2), we only found significant LD between PK11 and Pdoμ5 in Quintos de Mora population for one of the study years. As the LD correlation coefficient (*r*LD) between these pairs of loci was very small (range: 0.062–0.074) and we found no consistent LD across study years/populations, none of these markers was discarded (e.g. Olano-Marin *et al.*, 2011a,b).

Genetic structure

STRUCTURE analyses indicated an optimal value of $K = 2$ according to the method of Evanno *et al.* (2005) (Appendix S1), but most individuals and all populations showed a considerable degree of genetic admixture (Appendix S2). The genetic assignment of individuals did not show any clear geographical pattern of genetic structure (Appendix S2), suggesting that the inferred clusters may be due to the presence of individuals with different genetic backgrounds (e.g. originated from genetically differentiated populations; see also Olano-Marin *et al.*, 2011a for a similar pattern).

Heterozygosity, identity disequilibrium and inbreeding

HL and H_O were highly correlated considering all loci or the subsets of neutral and functional markers ($r > 0.95$ and $P < 0.001$ in all cases). Subsequent statistical analyses and figures are only presented for *HL* because this measure is a better estimator of genome-wide heterozygosity and inbreeding in open populations subjected to immigration and admixture (Aparicio *et al.*, 2006). H_O data provided analogous results and are available from the authors. *HL* values for each population and subpopulation are indicated in Table 1. *HL* estimated at the subset of functional markers was not

correlated with *HL* at the subset of neutral markers ($r = 0.02$, $P = 0.567$). Average (± SE) *HL* did not differ among populations (all markers: $F_{2,786} = 0.98$, $P = 0.375$; neutral markers: $F_{2,786} = 0.98$, $P = 0.375$; functional markers: $F_{2,786} = 0.98$, $P = 0.375$). *HL* did not differ among the different subpopulations for all markers or the subset of neutral markers (all *P*s > 0.1). *HL* estimated at the subset of functional markers did not differ among subpopulations within Cabañeros or San Pablo de los Montes. However, *HL* estimated at the subset of functional markers differed among subpopulations within Quintos de Mora ($F_{2,402} = 3.29$, $P = 0.038$), and post hoc Tukey tests showed that these differences were due to a lower *HL* in Valdeyernos than in Gil Garcia ($P = 0.048$) (Table 1).

HHC were nonsignificant (i.e. 95% quantiles crossed zero) when all the studied individuals were considered (all markers: $r = 0.016$, 95% CI = −0.036 to 0.064; neutral markers: $r = -0.017$, 95% CI = −0.064 to 0.038; functional markers: $r = 0.001$, 95% CI = −0.033 to 0.043). We also analysed HHC within each of the three studied populations. HHC were not significant in Quintos de Mora (all markers: $r = 0.017$, 95% CI = −0.054 to 0.088; neutral markers: $r = 0.005$, 95% CI = −0.072 to 0.086; functional markers: $r = 0.010$, 95% CI = −0.056 to 0.069), San Pablo de los Montes (all markers: $r = 0.003$, 95% CI = −0.072 to 0.090; neutral markers: $r = -0.061$, 95% CI = −0.132 to 0.016; functional markers: $r = 0.004$, 95% CI = −0.064 to 0.076) or Cabañeros (all markers: $r = 0.033$, 95% CI = −0.092 to 0.171; neutral markers: $r = 0.014$, 95% CI = −0.104 to 0.152; functional markers: $r = -0.025$, 95% CI = −0.159 to 0.103). The parameter g^2 did not differ significantly from zero when all the individuals were considered (all markers: $g^2 = -0.003$, $P = 0.682$; neutral markers: $g^2 = -0.0004$, $P = 0.645$; functional markers: $g^2 = 0.006$, $P = 0.329$) or when individuals from Quintos de Mora (all markers: $g^2 = 0.002$, $P = 0.372$; neutral markers: $g^2 = 0.0020$, $P = 0.100$; functional markers: $g^2 = 0.010$, $P = 0.266$), Cabañeros (all markers: $g^2 = -0.011$, $P = 0.724$; neutral markers: $g^2 = -1.1617$, $P = 0.465$; functional markers: $g^2 = 0.009$, $P = 0.575$) or San Pablo de los Montes (all markers: $g^2 = -0.008$, $P = 0.748$; neutral markers: $g^2 = -0.0036$, $P = 0.988$; functional markers: $g^2 = 0.013$, $P = 0.438$) were analysed separately.

Only 144 of 619 breeding individuals captured at Quintos de Mora population over the six study years were recruits with both parents known (23.3%). The other individuals were founders, immigrants or recruits whose parents could not be captured or identified (i.e. individuals with *f* assumed to be zero). Of the 144 breeding individuals with known ancestry (at least both parents known), only one (0.7%) resulted from a consanguineous mating involving full siblings ($f = 0.25$). This individual had a significantly lower heterozygosity (all markers: $HL = 0.291$; neutral markers: $HL = 0.218$; functional markers: $HL = 0.387$) than the other individuals ($n = 114$, mean ± SD, all markers: $HL = 0.220 \pm 0.008$, $t = -8.85$, $P < 0.001$; neutral markers: $HL = 0.190 \pm 0.010$, $t = -2.84$, $P = 0.005$; functional markers: $HL = 0.259 \pm 0.011$, $t = -11.20$, $P < 0.001$) also captured during the same breeding season within Quintos de Mora population. We have also found two other crosses between related individuals (a mother–son pair, $f = 0.25$; a grandfather–granddaughter pair, $f = 0.125$), but none of their offspring recruited in the population.

Avian malaria infections

Prevalence of avian malaria (lineage SGS1) at each population and subpopulation is indicated in Table 1. Probability of infection differed among populations (Wald = 59.64, $P < 0.001$), and post hoc analyses showed that these differences were due to a higher probability of infection in San Pablo de los Montes than in Quintos de Mora (Wald = 4.64, $P < 0.001$) and Cabañeros (Wald = 48.64, $P < 0.001$) (see also Ferrer *et al.*, 2012). Probability of infection did not differ among subpopulations within San Pablo de los Montes (Wald = 3.64, $P = 0.456$) or Cabañeros (Wald = 2.64, $P = 0.104$). However, we found that probability of infection slightly differed among subpopulations within Quintos de Mora (Wald = 5.98, $P = 0.050$). A post hoc analysis showed that these differences were due to a higher probability of infection in Fuente del Común than in Gil García (Wald = 5.77, $P = 0.016$) (Table 1).

Association between heterozygosity and parasitism: general effects

We found no association between probability of infection by the lineage SGS1 of avian malaria and heterozygosity estimated at all typed loci (Table 3a) or the subset of functional markers (Table 3b). However, we found a highly significant quadratic relationship between probability of infection and individual heterozygosity estimated at neutral loci (Table 3b). Probability of infection increases with homozygosity up to values of *HL* around 0.15, reaches a plateau at values of *HL* from 0.15 to 0.40 and finally declines in highly homozygous individuals ($HL > 0.4$) (Fig. 1). The other studied variables (sex, age, body mass and wing length) were not significantly associated with probability of infection in any model (Table 3). Other quadratic terms and interactions between independent variables were not significant ($P > 0.1$ in all cases).

Genetic diversity did not significantly differ between the subsets of neutral and functional loci, but tended to be higher in the former (H_E: $F_{1,24} = 3.86$, $P = 0.061$; H_O: $F_{1,24} = 2.66$, $P = 0.116$; allelic richness: $F_{1,24} = 2.83$, $P = 0.105$) (Table 2). We performed an additional analysis to ascertain whether the differences in the HFC

Table 3 GLMMs (binomial error and logit link function; $n = 789$ individuals) for probability of infection by avian malaria (lineage SGS1) in relation to nongenetic terms (covariates: body mass, wing length; fixed factors: sex, age) and homozygosity by locus (HL) estimated for (a) all loci (HL_{Total}) or the (b) subsets of neutral ($HL_{Neutral}$) and functional ($HL_{Functional}$) loci.

	Estimate ± SE	Test	P
(a) All loci			
Rejected terms			
HL_{Total}		$Z = 0.356$	0.722
HL_{Total}^2		$Z = 0.126$	0.899
Sex		$Z = 1.403$	0.161
Age		$Z = -0.238$	0.812
Body mass		$Z = -1.422$	0.155
Wing length		$Z = 0.590$	0.555
(b) Subsets of neutral and functional loci			
Explanatory terms			
Intercept	-2.467 ± 0.578		
$HL_{Neutral}$	10.856 ± 3.242	$Z = 3.349$	<0.001
$HL_{Neutral}^2$	-23.095 ± 7.507	$Z = -3.076$	0.002
Rejected terms			
$HL_{Functional}$		$Z = -1.073$	0.283
$HL_{Functional}^2$		$Z = -1.084$	0.278
Sex		$Z = 1.297$	0.195
Age		$Z = -0.163$	0.871
Body mass		$Z = -1.243$	0.214
Wing length		$Z = 0.622$	0.534

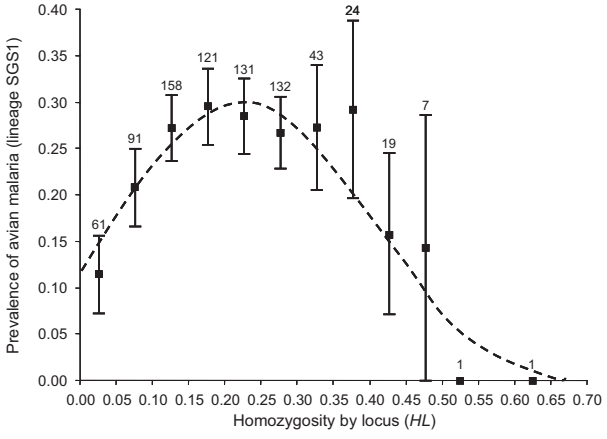

Fig. 1 Mean ± 1 SE prevalence of avian malaria in relation to homozygosity by locus (HL) for the subset of 'neutral' loci. Dashed line shows predicted values from the quadratic model. Numbers above bars indicate sample size for each category.

obtained for MLH estimated at the subsets of neutral and functional markers were due to the differences in genetic diversity and/or the different number of typed markers in each group (14 neutral markers vs. 12 functional markers; Table 2). For this purpose, we removed the two most diverse loci (Pca8 and Pdoμ5; Table 2) from the calculation of MLH at neutral markers and ran another GLMM using this new variable. After the removal of these two loci, genetic diversity did not significantly differ between the subsets of neutral and functional loci (H_E: $F_{1,22} = 2.30$, $P = 0.144$; H_O: $F_{1,22} = 1.67$, $P = 0.209$; allelic richness: $F_{1,22} = 1.06$, $P = 0.313$) and we still found a significant quadratic relationship between probability of infection and HL estimated at the subset of neutral loci (linear term of HL: $Z = 2.454$, $P = 0.014$; quadratic term of HL: $Z = -2.041$, $P = 0.041$).

We also analysed the data for each population separately to see whether the relationship between probability of infection and host genetic diversity differs among the studied populations (Table 1). Although the degree of spatial genetic structure within the whole study area is low, these analyses also help to avoid the possibility that the association between probability of avian malaria infections and individual genetic diversity has merely resulted from population stratification (Slate *et al.*, 2004; Slate & Pemberton, 2006). In Quintos de Mora and San Pablo de los Montes, probability of infection increased (linear term of HL: Quintos de Mora: $Z = 2.794$, $P = 0.005$; San Pablo de los Montes: $Z = 2.046$, $P = 0.040$) and then declined (quadratic term of HL: Quintos de Mora: $Z = -2.340$, $P = 0.019$; San Pablo de los Montes: $Z = -1.944$, $P = 0.051$) with HL estimated at the subset of neutral loci. A similar pattern was found in Cabañeros, but the association was not significant (linear term of HL: $Z = 0.750$, $P = 0.453$; quadratic term of HL: $Z = -1.152$, $P = 0.249$). Probability of infection was not associated with the linear or quadratic term of heterozygosity estimated at all loci or the subset of functional markers in any population (all Ps > 0.06). We also re-analysed the data from Quintos de Mora excluding the subpopulation with higher prevalence to avian malaria (Fuente del Común; Table 1), and we still found that probability of infection increased (linear term of HL: $Z = 3.410$, $P < 0.001$) and then declined (quadratic term of HL: $Z = -2.963$, $P = 0.003$) with HL estimated at the subset of neutral loci. Finally, we re-analysed the complete data set but including population or subpopulation as fixed factors (rather than as random effects) and their respective interactions with HL and its quadratic term. No interaction was significant (all Ps > 0.3), indicating that the relationship between probability of infection and host genetic diversity did not differ among the studied populations/subpopulations.

Association between heterozygosity and parasitism: local effects

The model including SLH at all neutral loci did not improve the variance explained by the model including the significant effect of MLH at this subset of loci ($F_{12,657} = 1.23$, $P = 0.258$). No single-locus effect was significant after sequential Bonferroni correction

(Table 4; Fig. 2). The probability of obtaining positive or negative effects, irrespective of the statistical significance, was not different from 0.5, and the direction and significance of SLH effects did not differ between the subsets of neutral and functional markers (all *P*s > 0.5). Absolute effect size of SLH was not correlated with marker genetic diversity (H_E: $F_{1,24} = 0.07$, $P = 0.79$; H_O: $F_{1,24} = 0.24$, $P = 0.631$; allelic richness: $F_{1,24} = 0.10$, $P = 0.76$) and did not differ between the subsets of neutral and functional loci ($F_{1,24} = 2.18$, $P = 0.153$).

Discussion

We found a significant relationship between probability of infection by lineage SGS1 of avian malaria and individual genetic diversity across the three studied blue tit populations. However, contrary to our expectations, the relationship was not linear and individuals with high and low genetic diversity showed a lower probability of being infected than individuals with an intermediate level of heterozygosity. Further, this association was significant only for MLH estimated at the panel of neutral markers, and we did not find any relationship between parasitism and host genetic diversity estimated at the subset of functional markers. Finally, the association between probability of infection and host genetic diversity did not change across populations with remarkably different levels of malaria prevalence, indicating that contrasting patterns of parasite pressure do not change the shape of the observed relationship in our study system.

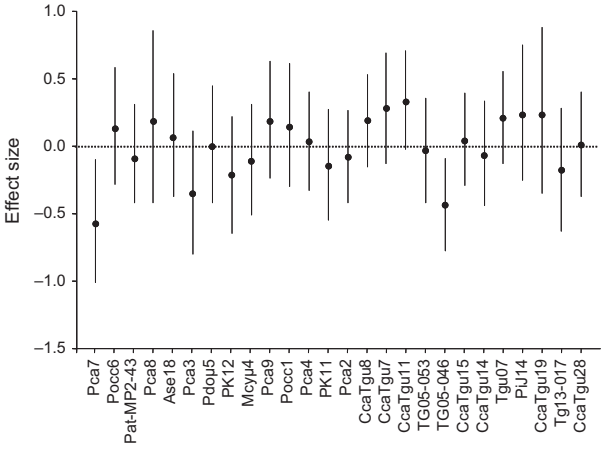

Fig. 2 Effect sizes and 95% confidence intervals of single-locus heterozygosity (SLH) for probability of infection by avian malaria.

Table 4 Test for the effects of single-locus heterozygosity (SLH) on probability of infection by avian malaria (lineage SGS1).

Locus	Category	Effect size	95% CI Lower/Upper	Test	*P*
Pca7	Neutral	−0.576	−1.026/−0.112	$Z = -2.477$	0.013
Pocc6	Neutral	0.127	−0.287/0.561	$Z = 0.588$	0.556
Pat-MP2-43	Neutral	−0.095	−0.419/0.299	$Z = -0.577$	0.564
Pca8	Neutral	0.182	−0.432/0.860	$Z = 0.555$	0.579
Ase18	Neutral	0.061	−0.366/0.510	$Z = 0.275$	0.783
Pca3	Neutral	−0.358	−0.804/0.106	$Z = -1.544$	0.122
Pdoμ5	Neutral	−0.009	−0.433/0.434	$Z = -0.041$	0.967
PK12	Neutral	−0.217	−0.651/0.235	$Z = -0.963$	0.335
Mcyμ4	Neutral	−0.114	−0.516/0.304	$Z = -0.544$	0.587
Pca9	Neutral	0.180	−0.223/0.603	$Z = 0.858$	0.391
Pocc1	Neutral	0.136	−0.304/0.601	$Z = 0.589$	0.556
Pca4	Neutral	0.028	−0.330/0.398	$Z = 0.154$	0.878
PK11	Neutral	−0.149	−0.545/0.262	$Z = -0.724$	0.469
Pca2	Neutral	−0.085	−0.423/0.259	$Z = -0.489$	0.625
CcaTgu8	Functional	0.184	−0.168/0.542	$Z = 1.019$	0.308
CcaTgu7	Functional	0.279	−0.120/0.695	$Z = 1.342$	0.179
CcaTgu11	Functional	0.325	−0.040/0.703	$Z = 1.718$	0.086
TG05-053	Functional	−0.038	−0.417/0.353	$Z = -0.195$	0.845
TG05-046	Functional	−0.438	−0.775/−0.104	$Z = -2.563$	0.010
CcaTgu15	Functional	0.038	−0.299/0.381	$Z = 0.222$	0.825
CcaTgu14	Functional	−0.070	−0.453/0.325	$Z = -0.356$	0.722
Tgu07	Functional	0.204	−0.148/0.565	$Z = 1.119$	0.263
PiJ14	Functional	0.229	−0.258/0.752	$Z = 0.892$	0.373
CcaTgu19	Functional	0.228	−0.362/0.880	$Z = 0.726$	0.468
Tg13-017	Functional	−0.180	−0.615/0.274	$Z = -0.794$	0.427
CcaTgu28	Functional	0.004	−0.360/0.377	$Z = 0.020$	0.984

Table shows effect sizes, 95% confidence intervals and *P*-values. No test was significant after sequential Bonferroni correction.

Nonlinear association between heterozygosity and parasitism

Nonlinear relationships between host genetic diversity and parasitism have been also found in previous studies estimating genetic variability at microsatellite markers (Ortego *et al.*, 2007a; Blanchet *et al.*, 2009; Ruiz-Lopez *et al.*, 2012) and genes involved in immune response (e.g. Wegner *et al.*, 2003a,b; Westerdahl *et al.*, 2005). We found that the positive effects of individual genetic diversity in terms of reduced probability of infection follow a saturation curve, quickly reaching a plateau after which lower levels of heterozygosity do not result in increased risk of infection. This may indicate that the benefits of increased genetic diversity are particularly patent among highly heterozygous individuals and readily stabilize due to a similarly low capacity to clean or avoid infections below a certain threshold of genetic diversity ($HL \sim 0.15$). Only a few (<0.4%) highly homozygous individuals ($HL > 0.4$) are responsible of the negative quadratic pattern. Different factors may explain the unexpected lower probability of infection among these highly homozygous individuals.

A possible explanation for the observed negative quadratic pattern between parasitism and host genetic diversity could be related to the existence of a bias in mortality, in such a way that highly homozygous individuals infected by avian malaria do not survive acute infections and only those that have never been exposed to the parasite (i.e. uninfected) are among the sampled individuals (Westerdahl *et al.*, 2005; Ortego *et al.*, 2007b; Westerdahl, 2007). If this is the case, we would expect (i) that the frequency of infected individuals among highly homozygous individuals is lower than in highly heterozygous individuals and (ii) that a higher mortality rate among highly homozygous individuals reduces its frequency in the population in comparison with that expected in the absence of positive heterozygosity-based selection. However, we found that the frequency of infected individuals did not significantly differ between highly heterozygous ($HL < 0.15$) and highly homozygous ($HL > 0.40$) individuals (2×2 contingency table; $\chi^2 = 0.451$, $P = 0.502$). To see whether the frequency of highly homozygous individuals in the population is lower than expected in the absence of heterozygosity-based selection, we simulated multilocus genotypes for the subset of neutral loci considering the allele frequencies observed in each subpopulation. We found that the observed frequency of highly homozygous individuals tended to be higher than expected in the absence of positive heterozygosity-based selection (2×2 contingency table; $\chi^2 = 2.829$, $P = 0.093$) (Fig. 3). These analyses indicate that a bias in mortality proposed in previous studies to explain a negative quadratic association between probability of infection and individual genetic diversity is not likely to be the reason behind such pattern in our study system

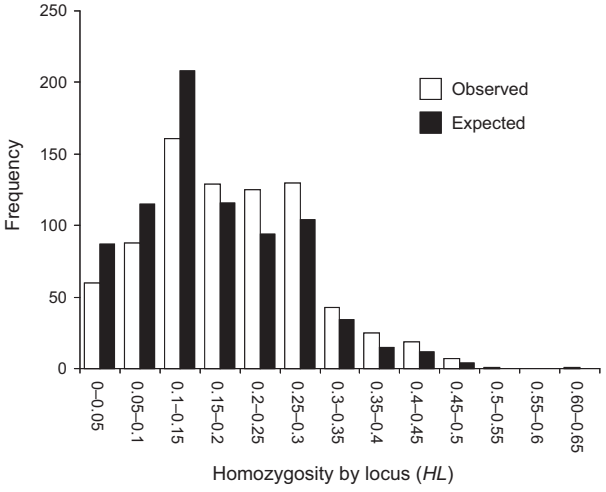

Fig. 3 Observed (open bars) and expected (black bars) distribution of homozygosity by locus (*HL*) for the subset of neutral loci. The expected distribution of *HL* was obtained via simulations of 789 multilocus genotypes (i.e. a figure equal to our total number of sampled individuals) based on observed allele frequencies within each subpopulation and assuming random mating.

(Westerdahl *et al.*, 2005; Westerdahl, 2007). Thus, an alternative explanation for the lower probability of infection among highly homozygous individuals is that it reflects increased resistance to parasites among this small proportion of presumably highly inbreed individuals due to the purging of deleterious recessive alleles (Barrets & Charleswort, 1991; Dudash *et al.*, 1997), which may have uncoupled the expected relationship between individual genetic diversity and resistance to parasites (Wiehn *et al.*, 2002; Haag *et al.*, 2003).

Blanchet *et al.* (2009) and Ruiz-Lopez *et al.* (2012) also reported that hosts with intermediate levels of heterozygosity had higher ectoparasite load than highly homozygous or heterozygous individuals. These studies did not find evidence for strong local effects, but suggested that a combination of both positive and negative small local HFC (irrespective of their statistical significance) may be behind the observed quadratic relationship at the multilocus level and interpreted that the global pattern could be caused by parasite-mediated disruptive selection on genetic diversity (see also Neff, 2004; Kupper *et al.*, 2010; Olano-Marin *et al.*, 2011b; Mueller *et al.*, 2011). Similarly, we did not find any significant single-locus effect, but the proportion of positive and negative local effects was similar (Fig. 2), and this may also explain the negative quadratic relationship between host genetic diversity and parasitism observed in our studied blue tit populations.

General vs. local effects

There is no consensus about whether heterozygosity measured across a set of markers can reflect

genomewide heterozygosity and therefore inbreeding levels (Balloux *et al.*, 2004; Forstmeier *et al.*, 2012). We found no evidence that local effects are the underlying mechanism behind the observed association between probability of infection by avian malaria parasites and individual genetic diversity. The model including SLH at all neutral loci did not improve the variance explained by the model including MLH at this subset of loci, and no single-locus effect was significant (Table 4; Fig. 2). However, we found that HHC (Balloux *et al.*, 2004) and parameter g^2 (David *et al.*, 2007; Szulkin *et al.*, 2010) did not differ significantly from zero, suggesting that heterozygosity measured at our set of microsatellite loci may not be representative of inbreeding. It should be considered that a nonsignificant correlation in heterozygosity among markers (HHC or g^2) cannot be regarded as disproving general effects, as the studied traits can capture the effect of potentially many more loci than the number of typed markers (Szulkin *et al.*, 2010; e.g. Forcada & Hoffman, 2014). Accordingly, a recent simulation-based study has demonstrated that failure to detect ID cannot be taken as strong evidence that inbreeding depression is not behind observed HFC and showed that many studies are likely to detect HFC caused by inbreeding depression, but fail to detect ID (Kardos *et al.*, 2014). This study also suggested that ID is only likely to be detected when variance of inbreeding is high or many loci are used (>100 microsatellites), a situation that is not likely to the case in most studies on natural populations. Our study is well above the typical number of 10-14 microsatellite loci employed in most studies on HFC, but way too far from the number of loci necessary to detect ID in the presence of a moderate to low variance of inbreeding. The very limited dispersal capacity of blue tits has probably resulted in the patterns of fine spatial-scale genetic structure and matings between relatives reported in different populations of this species (Garcia-Navas *et al.*, 2009; Olano-Marin *et al.*, 2011a,b; Ortego *et al.*, 2011). This may increase the variance in genomewide genetic diversity and increase the chance of detecting HFC (Balloux *et al.*, 2004; Slate *et al.*, 2004). Pedigree data were only available for a small proportion of our sampled individuals (<20%), and for most of them (>80%), we exclusively had information on the identity of their parents, which is likely to have resulted in we have failed to identify most inbreed individuals. Despite these limitations of our pedigree data, we have identified two cases of mating between close relatives (Garcia-Navas *et al.*, 2009) and the only inbreed recruit had significantly lower heterozygosity than other supposedly outbreed individuals, which suggests that our set of markers may be reflecting genomewide inbreeding. Thus, inbreeding rates may be enough for HFC arise in our and other studied blue tit populations (Foerster *et al.*, 2003; Garcia-Navas *et al.*, 2009; Olano-Marin *et al.*, 2011a,b), but they are also likely to be below the threshold necessary to detect ID using our set of typed markers (Kardos *et al.*, 2014; see also Forcada & Hoffman, 2014, for a recent empirical study).

Neutral vs. functional markers

We have found an association between probability of avian malaria infections and genetic diversity at neutral loci, but such association was not significant when only considering the subset of functional loci or a combination of both groups of loci. In our case, these differences are not attributable to differences in genetic diversity between our panels of neutral and functional markers (see also Olano-Marin *et al.*, 2011a,b; Szulkin & David, 2011). One possibility to explain this discrepancy is that heterozygosity (or homozygosity) at functional markers may be selected by many processes related with their different specific functions (or those from their closely linked genes), which may reduce their ability to reflect genomewide heterozygosity (see also Mueller *et al.*, 2011; Szulkin & David, 2011). The capacity to survive and clear an infection is likely to be associated with many physiological traits whose heterozygosity may be better reflected by neutral markers. Accordingly, a previous study on blue tits also found that a similar set of neutral loci provides more power to detect HFC and identity disequilibrium than functional loci, indicating that the former may capture better the effects of inbreeding (Olano-Marin *et al.*, 2011a). HFC reported at functional loci (or loci located in functional genomic regions) have been generally interpreted as evidence of local or direct effects (David, 1998; Hansson *et al.*, 2001; Hansson & Westerberg, 2002) and, thus, the ability of these markers to detect HFC may depend on whether their specific functions are related to the particular studied traits (e.g. Da Silva *et al.*, 2009; Kupper *et al.*, 2010; Olano-Marin *et al.*, 2011a,b; Laine *et al.*, 2012).

Conclusions

This study shows that host genetic diversity is associated with probability of infection by avian malaria parasites. The fact that we did not find any evidence for local effects and that the association between probability of infection and individual genetic diversity was exclusively explained by heterozygosity at putatively neutral loci suggests that the observed effects may be mediated by inbreeding and/or of genomewide genetic diversity. Given that heterozygosity is heritable in our study populations, avian malaria could be an important factor shaping host genetic diversity in this system (Garcia-Navas *et al.*, 2009). Different studies have demonstrated that genetic variation in genes involved in parasite defence is not necessarily linked to the levels of variation in the rest of the genome and they can strongly differ in their responses to demographic

changes and parasite-mediated selection pressures (Westerdahl *et al.*, 2005; Oliver *et al.*, 2009; Oliver & Piertney, 2012). Thus, our study and others previously finding an association between parasitism and heterozygosity estimated at putatively neutral markers suggest that genetic diversity at loci not directly involved in immune defence is also relevant for the resistance to infectious diseases (Acevedo-Whitehouse *et al.*, 2006; Ortego *et al.*, 2007a). The fact that the strength and the shape of the relationship between probability of infection and heterozygosity did not differ among populations strongly differing in malaria prevalence suggests that host genetic diversity is likely to similarly respond to infectious diseases under different scenarios of parasite pressure. Overall, our study highlights that parasites play a major role in shaping genetic variation in wild populations and suggests that the use of larger sets of neutral markers not associated with functional genomic regions may be more appropriate to the study of HFC, particularly if functional loci related to the specific studied traits have not been identified. The application of new available high-throughput sequencing technology will help to extraordinarily increase the number of employed loci and improve the estimates of genome-wide heterozygosity in future HFC studies (e.g. Hoffman *et al.*, 2014).

Acknowledgments

We wish to thank Pedro J. Cordero for allowing us to carry out the genetic analyses in his laboratory and Javier Bueno and Rafael Barrientos for their help during sample collection at San Pablo de los Montes. Two anonymous referees provided useful discussion and valuable comments on an earlier draft of this manuscript. We thank the board of the Cabañeros National Park, Centro Quintos de Mora and the Council of San Pablo de Los Montes for the facilities offered during the fieldwork. This study was funded by grants 69/2003 (Organismo Autónomo de Parques Nacionales-Ministerio de Medio Ambiente), GCL2007-61395 (Ministerio de Educación y Ciencia), CGL2010-21933-C02-01 (Ministerio de Ciencia e Innovación) and POIC10-0269-7632 (Junta de Comunidades de Castilla-La Mancha-European Social Fund). ESF was supported by a predoctoral scholarship from Junta de Comunidades de Castilla-La Mancha-European Social Fund, and VGN enjoyed a FPI predoctoral scholarship from Ministerio de Ciencia e Innovación-European Social Fund. J.O was supported by 'Juan de la Cierva' and 'Severo Ochoa-EBD' post-doctoral fellowships (Ministerio de Ciencia e Innovación).

References

Acevedo-Whitehouse, K., Gulland, F., Greig, D. & Amos, W. 2003. Disease susceptibility in California sea lions. *Nature* **422**: 35.

Acevedo-Whitehouse, K., Spraker, T.R., Lyons, E., Melin, S.R., Gulland, F., Delong, R.L. *et al.* 2006. Contrasting effects of heterozygosity on survival and hookworm resistance in California sea lion pups. *Mol. Ecol.* **15**: 1973–1982.

Alho, J.S., Valimaki, K. & Merila, J. 2010. Rhh: an R extension for estimating multilocus heterozygosity and heterozygosity-heterozygosity correlation. *Mol. Ecol. Resour.* **10**: 720–722.

Aparicio, J.M., Ortego, J. & Cordero, P.J. 2006. What should we weigh to estimate heterozygosity, alleles or loci? *Mol. Ecol.* **15**: 4659–4665.

Armbruster, P. & Reed, D.H. 2005. Inbreeding depression in benign and stressful environments. *Heredity* **95**: 235–242.

Atkinson, C.T. & Samuel, M.D. 2010. Avian malaria *Plasmodium relictum* in native Hawaiian forest birds: epizootiology and demographic impacts on 'apapane *Himatione sanguinea*. *J. Avian Biol.* **41**: 357–366.

Atkinson, C.T., Dusek, R.J., Woods, K.L. & Iko, W.M. 2000. Pathogenicity of avian malaria in experimentally-infected Hawaii amakihi. *J. Wildl. Dis.* **36**: 197–204.

Balloux, F., Amos, W. & Coulson, T. 2004. Does heterozygosity estimate inbreeding in real populations? *Mol. Ecol.* **13**: 3021–3031.

Barrets, C.H. & Charleswort, D. 1991. Effects of a change in the level of inbreeding on the genetic load. *Nature* **352**: 522–524.

Bensch, S., Price, T. & Kohn, J. 1997. Isolation and characterization of microsatellite loci in a *Phylloscopus* warbler. *Mol. Ecol.* **6**: 91–92.

Blanchet, S., Rey, O., Berthier, P., Lek, S. & Loot, G. 2009. Evidence of parasite-mediated disruptive selection on genetic diversity in a wild fish population. *Mol. Ecol.* **18**: 1112–1123.

Blondel, J. & Aronson, J. 1999. *Biology and Wildlife of the Mediterranean Region*. Oxford University Press, Oxford, UK.

Blondel, J., Thomas, D.W., Charmantier, A., Perret, P., Bourgault, P. & Lambrechts, M.M. 2006. A thirty-year study of phenotypic and genetic variation of blue tits in Mediterranean habitat mosaics. *Bioscience* **56**: 661–673.

Carr, D.E., Murphy, J.F. & Eubanks, M.D. 2003. The susceptibility and response of inbred and outbred *Mimulus guttatus* to infection by *Cucumber mosaic virus*. *Evol. Ecol.* **17**: 85–103.

Chapman, J.R., Nakagawa, S., Coltman, D.W., Slate, J. & Sheldon, B.C. 2009. A quantitative review of heterozygosity-fitness correlations in animal populations. *Mol. Ecol.* **18**: 2746–2765.

Coltman, D.W. & Slate, J. 2003. Microsatellite measures of inbreeding: a meta-analysis. *Evolution* **57**: 971–983.

Coltman, D.W., Pilkington, J.G., Smith, J.A. & Pemberton, J.M. 1999. Parasite-mediated selection against inbred Soay sheep in a free-living, island population. *Evolution* **53**: 1259–1267.

Da Silva, A., Gaillard, J.M., Yoccoz, N.G., Hewison, A.J.M., Galan, M., Coulson, T. *et al.* 2009. Heterozygosity-fitness correlations revealed by neutral and candidate gene markers in roe deer from a long-term study. *Evolution*, **63**: 403–417.

David, P. 1998. Heterozygosity-fitness correlations: new perspectives on old problems. *Heredity* **80**: 531–537.

David, P., Pujol, B., Viard, F., Castella, V. & Goudet, J. 2007. Reliable selfing rate estimates from imperfect population genetic data. *Mol. Ecol.* **16**: 2474–2487.

Dawson, D.A., Hanotte, O., Greig, C., Stewart, I.R.K. & Burke, T. 2000. Polymorphic microsatellites in the blue tit *Parus caeruleus* and their cross-species utility in 20 songbird families. *Mol. Ecol.* **9**: 1941–1944.

Dawson, D.A., Horsburgh, G.J., Kupper, C., Stewart, I.R.K., Ball, A.D., Durrant, K.L. *et al.* 2010. New methods to identify conserved microsatellite loci and develop primer sets of high cross-species utility - as demonstrated for birds. *Mol. Ecol. Resour.* **10**: 475–494.

Double, M.C., Dawson, D., Burke, T. & Cockburn, A. 1997. Finding the fathers in the least faithful bird: a microsatellite-based genotyping system for the superb fairy-wren *Malurus cyaneus*. *Mol. Ecol.* **6**: 691–693.

Dudash, M.R., Carr, D.E. & Fenster, C.B. 1997. Five generations of enforced selfing and outcrossing in *Mimulus guttatus*: inbreeding depression variation at the population and family level. *Evolution* **51**: 54–65.

Earl, D.A. & vonHoldt, B.M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**: 359–361.

Evanno, G., Regnaut, S. & Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**: 2611–2620.

Falush, D., Stephens, M. & Pritchard, J.K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**: 1567–1587.

Ferguson, M.M. & Drahushchak, L.R. 1990. Disease resistance and enzyme heterozygosity in rainbow-trout. *Heredity* **64**: 413–417.

Ferrer, E.S., Garcia-Navas, V., Sanz, J.J. & Ortego, J. 2012. Molecular characterization of avian malaria parasites in three Mediterranean blue tit (*Cyanistes caeruleus*) populations. *Parasitol. Res.* **111**: 2137–2142.

Foerster, K., Delhey, K., Johnsen, A., Lifjeld, J.T. & Kempenaers, B. 2003. Females increase offspring heterozygosity and fitness through extra-pair matings. *Nature* **425**: 714–717.

Forcada, J. & Hoffman, J.I. 2014. Climate change selects for heterozygosity in a declining fur seal population. *Nature* **511**: 462–465.

Forstmeier, W., Schielzeth, H., Mueller, J.C., Ellegren, H. & Kempenaers, B. 2012. Heterozygosity-fitness correlations in zebra finches: microsatellite markers can be better than their reputation. *Mol. Ecol.* **21**: 3237–3249.

Fossoy, F., Johnsen, A. & Lifjeld, J.T. 2009. Cell-mediated immunity and multi-locus heterozygosity in bluethroat nestlings. *J. Evol. Biol.* **22**: 1954–1960.

Fox, C.W. & Reed, D.H. 2011. Inbreeding depression increases with environmental stress: An experimental study and meta-analysis. *Evolution* **65**: 246–258.

Garcia-Navas, V., Ortego, J. & Sanz, J.J. 2009. Heterozygosity-based assortative mating in blue tits (*Cyanistes caeruleus*): implications for the evolution of mate choice. *Proc. R. Soc. Lond. B* **276**: 2931–2940.

Garnier-Gere, P. & Dillmann, C. 1992. A computer-program for testing pairwise linkage disequilibria in subdivided populations. *J. Hered.* **83**: 239–239.

Griffith, S.C., Stewart, I.R.K., Dawson, D.A., Owens, I.P.F. & Burke, T. 1999. Contrasting levels of extra-pair paternity in mainland and island populations of the house sparrow (*Passer domesticus*): is there an 'island effect'? *Biol. J. Linn. Soc.* **68**: 303–316.

Guo, S.W. & Thompson, E.A. 1992. A monte-carlo method for combined segregation and linkage analysis. *Am. J. Hum. Genet.* **51**: 1111–1126.

Haag, C.R., Sakwinska, O. & Ebert, D. 2003. Test of synergistic interaction between infection and inbreeding in *Daphnia magna*. *Evolution* **57**: 777–783.

Hansson, B. & Westerberg, L. 2002. On the correlation between heterozygosity and fitness in natural populations. *Mol. Ecol.* **11**: 2467–2474.

Hansson, B., Bensch, S., Hasselquist, D. & Akesson, M. 2001. Microsatellite diversity predicts recruitment of sibling great reed warblers. *Proc. R. Soc. Lond. B* **268**: 1287–1291.

Hawley, D.M., Sydenstricker, K.V., Kollias, G.V. & Dhondt, A.A. 2005. Genetic diversity predicts pathogen resistance and cell-mediated immunocompetence in house finches. *Biol. Lett.* **1**: 326–329.

Hoffman, J.I., Simpson, F., David, P., Rijks, J.M., Kuiken, T., Thorne, M.A.S. *et al.* 2014. High-throughput sequencing reveals inbreeding depression in a natural population. *Proc. Natl. Acad. Sci. USA* **111**: 3775–3780.

Hubisz, M.J., Falush, D., Stephens, M. & Pritchard, J.K. 2009. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour.* **9**: 1322–1332.

Kardos, M., Allendorf, F.W. & Luikart, G. 2014. Evaluating the role of inbreeding depression in heterozygosity-fitness correlations: how useful are tests for identity disequilibrium? *Mol. Ecol. Resour.* **14**: 519–530.

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

Knowles, S.C.L., Palinauskas, V. & Sheldon, B.C. 2010. Chronic malaria infections increase family inequalities and reduce parental fitness: experimental evidence from a wild bird population. *J. Evol. Biol.* **23**: 557–569.

Kristensen, T.N., Dahlgaard, J. & Loeschcke, V. 2002. Inbreeding affects Hsp70 expression in two species of *Drosophila* even at benign temperatures. *Evol. Ecol. Res.* **4**: 1209–1216.

Kupper, C., Kosztolanyi, A., Augustin, J., Dawson, D.A., Burke, T. & Szekely, T. 2010. Heterozygosity-fitness correlations of conserved microsatellite markers in Kentish plovers *Charadrius alexandrinus*. *Mol. Ecol.* **19**: 5172–5185.

Lachish, S., Knowles, S.C.L., Alves, R., Wood, M.J. & Sheldon, B.C. 2011. Fitness effects of endemic malaria infections in a wild bird population: the importance of ecological structure. *J. Anim. Ecol.* **80**: 1196–1206.

Laine, V.N., Herczeg, G., Shikano, T. & Primmer, C.R. 2012. Heterozygosity-behaviour correlations in nine-spined stickleback (*Pungitius pungitius*) populations: contrasting effects at random and functional loci. *Mol. Ecol.* **21**: 4872–4884.

Li, Y.C., Korol, A.B., Fahima, T. & Nevo, E. 2004. Microsatellites within genes: Structure, function, and evolution. *Mol. Biol. Evol.* **21**: 991–1007.

Luikart, G., Pilgrim, K., Visty, J., Ezenwa, V.O. & Schwartz, M.K. 2008. Candidate gene microsatellite variation is associated with parasitism in wild bighorn sheep. *Biol. Lett.* **4**: 228–231.

Luong, L.T., Heath, B.D. & Polak, M. 2007. Host inbreeding increases susceptibility to ectoparasitism. *J. Evol. Biol.* **20**: 79–86.

Meagher, S. 1999. Genetic diversity and *Capillaria hepatica* (Nematoda) prevalence in Michigan deer mouse populations. *Evolution* **53**: 1318–1324.

Merino, S., Moreno, J., Sanz, J.J. & Arriero, E. 2000. Are avian blood parasites pathogenic in the wild? A medication

experiment in blue tits (*Parus caeruleus*). *Proc. R. Soc. Lond. B* **267**: 2507–2510.

Mitton, J.B. & Grant, M.C. 1984. Associations among protein heterozygosity, growth-rate, and developmental homeostasis. *Annu. Rev. Ecol. Syst.* **15**: 479–499.

Mueller, J.C., Hermisson, J., Olano-Marin, J., Hansson, B. & Kempenaers, B. 2011. Linking genetic mechanisms of heterozygosity-fitness correlations to footprints of selection at single loci. *Evol. Ecol.* **25**: 1–11.

Nakagawa, S. & Cuthill, I.C. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol. Rev.* **82**: 591–605.

Neff, B.D. 2004. Stabilizing selection on genomic divergence in a wild fish population. *Proc. Natl. Acad. Sci. USA* **101**: 2381–2385.

Obrien, S.J. & Evermann, J.F. 1988. Interactive influence of infectious disease and genetic diversity in natural populations. *Trends Ecol. Evol.* **3**: 254–259.

Olano-Marin, J., Dawson, D.A., Girg, A., Hansson, B., Ljungqvist, M., Kempenaers, B. *et al.* 2010. A genome-wide set of 106 microsatellite markers for the blue tit (*Cyanistes caeruleus*). *Mol. Ecol. Resour.* **10**: 516–532.

Olano-Marin, J., Mueller, J.C. & Kempenaers, B. 2011a. Correlations between heterozygosity and reproductive success in the blue tit (*Cyanistes caeruleus*): an analysis of inbreeding and single locus effects. *Evolution* **65**: 3175–3194.

Olano-Marin, J., Mueller, J.C. & Kempenaers, B. 2011b. Heterozygosity and survival in blue tits (*Cyanistes caeruleus*): contrasting effects of presumably functional and neutral loci. *Mol. Ecol.* **20**: 4028–4041.

Oliver, M.K. & Piertney, S.B. 2012. Selection maintains MHC diversity through a natural population bottleneck. *Mol. Biol. Evol.* **29**: 1713–1720.

Oliver, M.K., Lambin, X., Cornulier, T. & Piertney, S.B. 2009. Spatiotemporal variation in the strength and mode of selection acting on major histocompatibility complex diversity in water vole (*Arvicola terrestris*) metapopulations. *Mol. Ecol.* **18**: 80–92.

Ortego, J., Aparicio, J.M., Calabuig, G. & Cordero, P.J. 2007a. Risk of ectoparasitism and genetic diversity in a wild lesser kestrel population. *Mol. Ecol.* **16**: 3712–3720.

Ortego, J., Cordero, P.J., Aparicio, J.M. & Calabuig, G. 2007b. No relationship between individual genetic diversity and prevalence of avian malaria in a migratory kestrel. *Mol. Ecol.* **16**: 4858–4866.

Ortego, J., Cordero, P.J., Aparicio, J.M. & Calabuig, G. 2008. Consequences of chronic infections with three different avian malaria lineages on reproductive performance of lesser kestrels (*Falco naumanni*). *J. Ornithol.* **149**: 337–343.

Ortego, J., Garcia-Navas, V., Ferrer, E.S. & Sanz, J.J. 2011a. Genetic structure reflects natal dispersal movements at different spatial scales in the blue tit, *Cyanistes caeruleus*. *Anim. Behav.* **82**: 131–137.

Ortego, J., Yannic, G., Shafer, A.B.A., Mainguy, J., Festa-Bianchet, M., Coltman, D.W. *et al.* 2011b. Temporal dynamics of genetic variability in a mountain goat (*Oreamnos americanus*) population. *Mol. Ecol.* **20**: 1601–1611.

Otter, K., Ratcliffe, L., Michaud, D. & Boag, P.T. 1998. Do female black-capped chickadees prefer high-ranking males as extra-pair partners? *Behav. Ecol. Sociobiol.* **43**: 25–36.

Pearman, P.B. & Garner, T.W.J. 2005. Susceptibility of Italian agile frog populations to an emerging strain of *Ranavirus parallels* population genetic diversity. *Ecol. Lett.* **8**: 401–408.

Pedersen, K.S., Kristensen, T.N. & Loeschcke, V. 2005. Effects of inbreeding and rate of inbreeding in *Drosophila melanogaster* – Hsp70 expression and fitness. *J. Evol. Biol.* **18**: 756–762.

Pritchard, J.K., Stephens, M. & Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, **155**: 945–959.

Puurtinen, M., Hytonen, M., Knott, K.E., Taskinen, J., Nissinen, K. & Kaitala, V. 2004. The effects of mating system and genetic variability on susceptibility to trematode parasites in a freshwater snail, *Lymnaea stagnalis*. *Evolution* **58**: 2747–2753.

R Development Core Team 2012. *R: A Language and Environment for Statistical Computing. Version 3.0.2*. R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org.

Raymond, M. & Rousset, F. 1995. GENEPOP (Version-1.2) – population-genetics software for exact tests and ecumenicism. *J. Hered.* **86**: 248–249.

Reid, J.M., Arcese, P., Keller, L.F., Elliott, K.H., Sampson, L. & Hasselquist, D. 2007. Inbreeding effects on immune response in free-living song sparrows (*Melospiza melodia*). *Proc. R. Soc. Lond. B* **274**: 697–706.

Rice, W.R. 1989. Analyzing tables of statistical tests. *Evolution* **43**: 223–225.

Richardson, D.S., Jury, F.L., Dawson, D.A., Salgueiro, P., Komdeur, J. & Burke, T. 2000. Fifty Seychelles warbler (*Acrocephalus sechellensis*) microsatellite loci polymorphic in *Sylviidae* species and their cross-species amplification in other passerine birds. *Mol. Ecol.* **9**: 2226–2231.

Rijks, J.M., Hoffman, J.I., Kuiken, T., Osterhaus, A.D.M.E. & Amos, W. 2008. Heterozygosity and lungworm burden in harbour seals (*Phoca vitulina*). *Heredity* **100**: 587–593.

Ruiz-Lopez, M.J., Monello, R.J., Gompper, M.E. & Eggert, L.S. 2012. The effect and relative importance of neutral genetic diversity for predicting parasitism varies across parasite taxa. *PLoS ONE* **7**: e45404.

Saccheri, I., Kuussaari, M., Kankare, M., Vikman, P., Fortelius, W. & Hanski, I. 1998. Inbreeding and extinction in a butterfly metapopulation. *Nature* **392**: 491–494.

Slate, J. & Pemberton, J. 2006. Does reduced heterozygosity depress sperm quality in wild rabbits (*Oryctolagus cuniculus*)? *Curr. Biol.* **16**: R790–R791.

Slate, J., David, P., Dodds, K.G., Veenvliet, B.A., Glass, B.C., Broad, T.E. *et al.* 2004. Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. *Heredity* **93**: 255–265.

Slate, J., Hale, M.C. & Birkhead, T.R. 2007. Simple sequence repeats in zebra finch (*Taeniopygia guttata*) expressed sequence tags: a new resource for evolutionary genetic studies of passerines. *BMC Genomics* **8**: 52.

Sol, D., Jovani, R. & Torres, J. 2003. Parasite mediated mortality and host immune response explain age-related differences in blood parasitism in birds. *Oecologia* **135**: 542–547.

Spielman, D., Brook, B.W. & Frankham, R. 2004. Most species are not driven to extinction before genetic factors impact them. *Proc. Natl. Acad. Sci. USA* **101**: 15261–15264.

Svensson, L. 1992. *Identification Guide to European Passerines*. Lars Svensson, Stockholm.

Szulkin, M. & David, P. 2011. Negative heterozygosity-fitness correlations observed with microsatellites located in functional areas of the genome. *Mol. Ecol.* **20**: 3949–3952.

Szulkin, M., Bierne, N. & David, P. 2010. Heterozygosity-fitness correlations: a time for reappraisal. *Evolution* **64**: 1202–1217.

Van Riper, C., Vanriper, S.G., Goff, M.L. & Laird, M. 1986. The epizootiology and ecological significance of malaria in hawaiian land birds. *Ecol. Monogr.* **56**: 327–344.

Voegeli, B., Saladin, V., Wegmann, M. & Richner, H. 2012. Parasites as mediators of heterozygosity-fitness correlations in the great tit (*Parus major*). *J. Evol. Biol.* **25**: 584–590.

Waldenström, J., Bensch, S., Hasselquist, D. & Ostman, O. 2004. A new nested polymerase chain reaction method very efficient in detecting *Plasmodium* and *Haemoproteus* infections from avian blood. *J. Parasitol.* **90**: 191–194.

Wegner, K.M., Kalbe, M., Kurtz, J., Reusch, T.B.H. & Milinski, M. 2003a. Parasite selection for immunogenetic optimality. *Science* **301**: 1343–1343.

Wegner, K.M., Reusch, T.B.H. & Kalbe, M. 2003b. Multiple parasites are driving major histocompatibility complex polymorphism in the wild. *J. Evol. Biol.* **16**: 224–232.

Weir, B.S. & Cockerham, C.C. 1973. Mixed self and random mating at 2 loci. *Gen. Res.* **21**: 247–262.

Westerdahl, H. 2007. Passerine MHC: genetic variation and disease resistance in the wild. *J. Ornithol.* **148**: S469–S477.

Westerdahl, H., Waldenstrom, J., Hansson, B., Hasselquist, D., von Schantz, T. & Bensch, S. 2005. Associations between malaria and MHC genes in a migratory songbird. *Proc. R. Soc. Lond. B* **272**: 1511–1518.

Whitehorn, P.R., Tinsley, M.C., Brown, M.J.F., Darvill, B. & Goulson, D. 2011. Genetic diversity, parasite prevalence and immunity in wild bumblebees. *Proc. R. Soc. Lond. B* **278**: 1195–1202.

Whiteman, N.K., Matson, K.D., Bollmer, J.L. & Parker, P.G. 2006. Disease ecology in the Galapagos hawk (*Buteo galapagoensis*): host genetic diversity, parasite load and natural antibodies. *Proc. R. Soc. Lond. B* **273**: 797–804.

Wiehn, J., Kopp, K., Rezzonico, S., Karttunen, S. & Jokela, J. 2002. Family-level covariation between parasite resistance and mating system in a hermaphroditic freshwater snail. *Evolution* **56**: 1454–1461.

Willi, Y., Van Buskirk, J. & Hoffmann, A.A. 2006. Limits to the adaptive potential of small populations. *Annu. Rev. Ecol. Syst.* **37**: 433–458.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Results of Bayesian clustering analyses in STRUCTURE.

Appendix S2 Probability of population membership of the studied individuals to each genetic cluster inferred by Bayesian clustering analyses in STRUCTURE.

Data deposited at Dryad: doi:10.5061/dryad.0r7g4

Received 25 May 2014; accepted 1 September 2014