MOLECULAR ECOLOGY

Molecular Ecology (2016) doi: 10.1111/mec.13759

Landscape-scale variation in an anthropogenic factor shapes immune gene variation within a wild population

CATALINA GONZALEZ-QUEVEDO,*† RICHARD G. DAVIES,* KARL P. PHILLIPS,*‡ LEWIS G. SPURGIN* and DAVID S. RICHARDSON*

*School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK, †Grupo Ecología y Evolución de Vertebrados, Instituto de Biología, Facultad de Ciencias Exactas y Naturales, Universidad de Antioquia, Calle 70 No. 52-21, Medellín, Colombia, ‡Evolutionary Biology Group, Institute of Environmental Biology, Faculty of Biology, Adam Mickiewicz University, Poznań, Poland

Abstract

Understanding the spatial scale at which selection acts upon adaptive genetic variation in natural populations is fundamental to our understanding of evolutionary ecology, and has important ramifications for conservation. The environmental factors to which individuals of a population are exposed can vary at fine spatial scales, potentially generating localized patterns of adaptation. Here, we compared patterns of neutral and major histocompatibility complex (MHC) variation within an island population of Berthelot's pipit (*Anthus berthelotii*) to assess whether landscape-level differences in pathogen-mediated selection generate fine-scale spatial structuring in these immune genes. Specifically, we tested for spatial associations between the distribution of avian malaria, and the factors previously shown to influence that distribution, and MHC variation within resident individuals. Although we found no overall genetic structure across the population for either neutral or MHC loci, we did find localized associations between environmental factors and MHC variation. One MHC class I allele (ANBE48) was directly associated with malaria infection risk, while the presence of the ANBE48 and ANBE38 alleles within individuals correlated (positively and negatively, respectively) with distance to the nearest poultry farm, an anthropogenic factor previously shown to be an important determinant of disease distribution in the study population. Our findings highlight the importance of considering small spatial scales when studying the patterns and processes involved in evolution at adaptive loci.

Keywords: Berthelot's pipit, major histocompatibility complex, malaria, principal components of neighbour matrices

Received 10 April 2015; revision received 23 June 2016; accepted 5 July 2016

Introduction

Understanding the spatial scale at which selection acts upon adaptive genetic variation in natural populations provides information on the degree of local adaptation of populations, and thus, potentially, on initial steps in the speciation process (Chave 2013). Furthermore, assessment of the spatial scale of evolutionary processes provides information on the epidemiology of wildlife diseases, mechanisms involved in the maintenance of genetic variation, and patterns of dispersal (DeAngelis & Mooij 2005), and should provide background information for delineating conservation strategies. When different groups of individuals evolve in different environments, each may become adapted to the local conditions. There is a large amount of empirical evidence for such local adaptation (e.g. Kawecki & Ebert 2004; Vincent *et al.* 2013), but studies have generally been carried out at coarse scales, where differences in environment are conspicuous and limited gene flow does not counteract the effects of selection (Lenormand 2002).

Correspondence: David S. Richardson, Fax: +44 1603 592250; E-mail: david.richardson@uea.ac.uk

© 2016 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

However, the environment can vary at fine spatial scales (Wood *et al.* 2007; Soto-Centeno *et al.* 2013). Selection and adaptation at scales at which gene flow is not likely to be limited have been increasingly studied in recent years (reviewed in Richardson *et al.* 2014), and there is growing evidence that fine-scale evolutionary divergence is more common than previously thought (Svensson & Sinervo 2004; Ray & King 2006; Milá *et al.* 2010).

Loci involved in disease resistance/resilience are of particular interest when studying local adaptation. Pathogens can be strong selective agents in wild host populations (Haldane 1949; Fumagalli *et al.* 2011), and their distribution is highly dependent on environmental factors (Ostfeld *et al.* 2005). For example, both climatic and anthropogenic factors can affect the distribution of pathogenic disease (e.g. Bradley & Altizer 2007; LaPointe *et al.* 2010). Furthermore, adaptation of host immune genes to local parasite assemblages appears to be widespread (Murray *et al.* 1995; Dionne *et al.* 2007; Ekblom *et al.* 2007; Evans *et al.* 2010; Eizaguirre *et al.* 2012; Lenz *et al.* 2013). Assessment of the scale of selection pressures exerted by pathogens on immune genes is important for understanding the patterns of disease epidemiology and transmission in the landscape.

The genes of the major histocompatibility complex (MHC), with their extraordinary levels of variation and key role in the vertebrate acquired immune response, have become a classic model for studying natural selection at the genetic level (reviewed in Bernatchez & Landry 2003; Spurgin & Richardson 2010). These loci encode cell surface receptors that bind peptides derived from pathogens (antigens) via the peptide binding region (PBR; Wakelin 1996). High variation at the MHC is thought to be driven largely by pathogen-mediated selection (PMS; Spurgin & Richardson 2010), although sexual selection (Ejsmond *et al.* 2014) and other mechanisms may also play a role (Van Oosterhout 2009). In an effort to understand the role of selection mechanisms at the MHC, many studies have investigated among-population structure of MHC genes (reviewed in Bernatchez & Landry 2003; Babik *et al.* 2008). Nevertheless, despite it being clear that the distribution of pathogens within an environment can vary greatly at small spatial scales (Eisen & Wright 2001; Wood *et al.* 2007), studies which assess the effects of selection on the MHC at these scales are lacking. Such studies may provide understanding of how selection acts within a population to maintain overall levels of variation, insight which may be obscured if we only focus on coarser patterns of variation.

Fine-scale genetic structure can result from two processes: differential selection pressures (extrinsic factors) that result from fine-scale environmental variation, and demographic processes (intrinsic factors) particular to the studied species (Legendre & Legendre 2012). These demographic processes include dispersal patterns, kin structure, mating system and genetic drift (Legendre & Legendre 2012; Wagner & Fortin 2013; Richardson *et al.* 2014). Differentiating between these extrinsic and intrinsic processes is vital if we are to draw conclusions about the causes and consequences of fine-scale genetic structure. To distinguish between these alternatives, we can contrast patterns of neutral genetic variation with functional genetic variation, for instance, using correlations between environment and genetic variation. Several approaches have been proposed to do this (Wagner & Fortin 2005; Dray *et al.* 2006; Jombart *et al.* 2008, 2009). One such approach uses principal components of neighbour matrices (PCNMs) to model patterns of genetic structure that are not accounted for by measured environmental gradients and would otherwise contribute to spatial autocorrelation in environmental model residuals (Borcard & Legendre 2002). This approach has been used in several recent studies (Manel *et al.* 2010; Garroway *et al.* 2013; Pavlova *et al.* 2013). PCNM analysis involves the generation and testing of a range of eigenvector maps as candidate spatial predictor variables that allow components of variation in a response variable, which may be dependent on geographic position alone, to be accounted for (see Dray *et al.* 2006; Legendre & Legendre 2012 for details). Another approach, spatial principal components analysis (sPCA), detects the components of genetic structuring that simultaneously show high variation and spatial autocorrelation. Hence, sPCA finds spatially dependent gradients in genetic variation without distinguishing between gradients that are the result of environmental or intrinsic factors. Both sPCA and PCNM methods are useful means of deriving spatially mapped visualizations of the fine-scale genetic variation they represent and predict, respectively.

Here, we investigated patterns of fine-scale genetic structure at neutral markers and MHC class I loci within a population of Berthelot's pipit (*Anthus berthelotii*) on Tenerife, in the Canary Islands. This population is isolated from other conspecific populations but widespread and abundant across the Tenerife landscape. Importantly, it exhibits a high and spatially varying prevalence of avian malaria (Spurgin *et al.* 2012; Gonzalez-Quevedo *et al.* 2014) which has already been shown to be associated with fine-scale variation in the environment (Gonzalez-Quevedo *et al.* 2014). Here, we (i) assessed neutral genetic structure in the population, (ii) estimated fine-scale genetic structure at the MHC and (iii) tested for associations between the spatial distribution of MHC alleles and malaria infection risk and other spatially variable environmental factors known to influence disease distribution.

© 2016 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

Methods

Study species and sampling

Berthelot's pipit (*Anthus berthelotii*) is a small, insectivorous passerine endemic to the Macaronesian archipelagos, where it inhabits all 12 islands. The species is normally distributed in open areas with dispersed shrubs from sea level up to 3700 m above sea level (asl) (Martín & Lorenzo 2001). The Berthelot's pipit population size is estimated to be up to 100 000 breeding pairs across its entire range (BirdLife International 2016), with up to 50 000 pairs estimated on Tenerife (J. C. Illera, personal communication). Gene flow among islands is thought to be very limited based on analysis of patterns of microsatellite variation (Illera *et al.* 2007; Spurgin *et al.* 2014). Field observations indicate that adult individuals are very territorial during the breeding season, and tend to stay in close proximity (ca 500 m) to the area where they were first observed year-round, although young individuals may disperse more widely (J. C. Illera, personal communication). The species has an estimated generation time of 3.7 years and a socially monogamous breeding system with a maximum of two clutches of 2–4 eggs per clutch laid during the annual breeding season, which in Tenerife spans February–April depending on altitude (Garcia-del-Rey & Cresswell 2007; BirdLife International 2016). Despite the species being territorial and sedentary, the population on Tenerife (Fig. 1) is likely to be interconnected because much of the habitat they use is continuous. The exception to this are pipits found on the top of the mountain of El Teide (above 2000 m asl) that may be isolated from the rest of the island population by the band of pine forest extending from 1600 to 2000 m asl that is not inhabited by the pipit. The surface area of Tenerife is 2034 km^2. To obtain a representative sample of the pipit over its entire range and across all environmental gradients on Tenerife, a 1-km^2 grid was laid over a map of the island obtained from Google Earth in ARCGIS version 10 (Esri 2011, Redlands, CA, www.esri.com). The majority of accessible square kilometres that contained habitat suitable for pipits were visited and, where pipits were present, an attempt was made to catch at least one per km^2 using clap nets baited with *Tenebrio molitor* larvae. The GPS coordinates of all visited sites were recorded. Each captured bird was ringed and a blood sample was taken by brachial venipuncture and stored in absolute ethanol in screwcap microcentrifuge tubes at room temperature. DNA was extracted using a salt extraction method following Richardson *et al.* (2001).

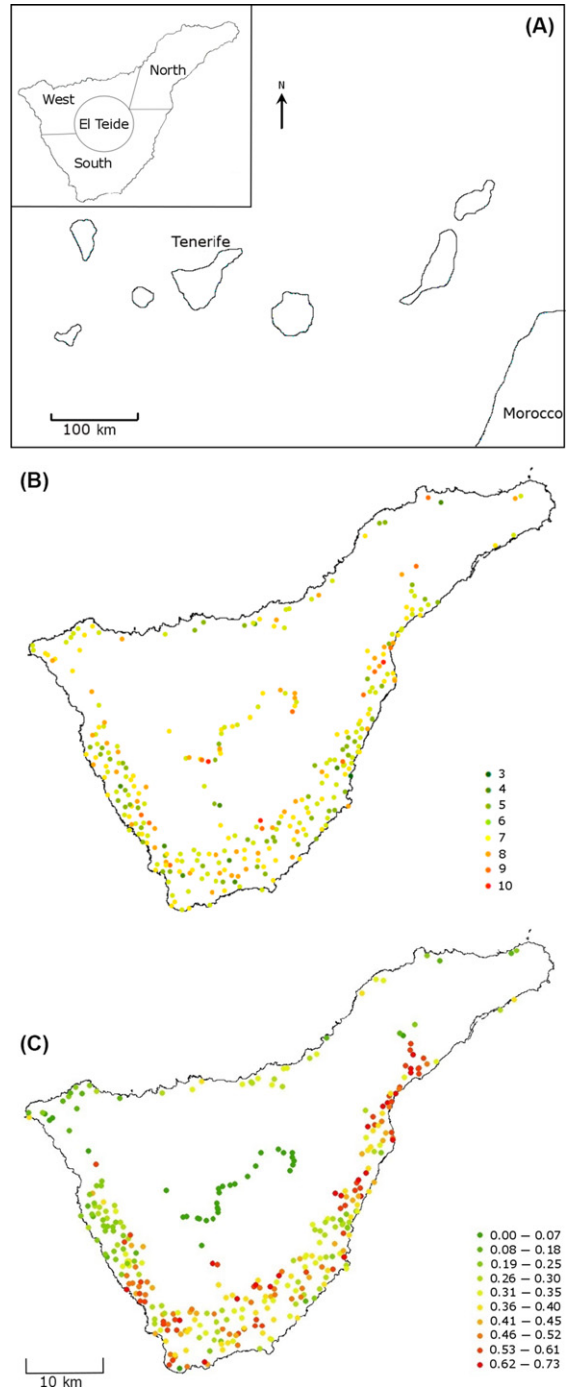

Fig. 1 The distribution of Berthelot's pipits (*Anthus berthelotii*) sampled across Tenerife. (A) Location of Tenerife within the Canary Islands in the North Atlantic. Inset shows our division of Tenerife into four subpopulations according to climatic and topographic differences: the south (dry), the north (wet), the west (narrow coastlines and high cliffs) and El Teide (high altitude). (B) The spatial distribution of the number of MHC class I exon 3 alleles identified per individual. (C) The malaria risk level (see key) for the exact location where each of the 388 individuals was sampled. Each dot represents an individual.

© 2016 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

Genotyping

A total of 388 pipits were genotyped at 21 microsatellite markers (Table S1, Supporting Information), following Spurgin *et al.* (2014). Of these birds, 310 were genotyped at the MHC class I by sequencing 240 bp (of 270 bp) of the exon 3, which codes for most of the peptide binding region (PBR), using 454 sequencing. Briefly, individuals were genotyped in duplicate to assess variant calling accuracy, and amplification efficiency (based on the number of reads of the variant per individual) was calculated for each variant to avoid allele dropout. The repeatability of the genotyping method was 96.1% (SE = 5.45). For detailed methods of the amplification procedure and the bioinformatics analyses performed for the MHC genotyping, see Gonzalez-Quevedo *et al.* (2015). One sample had poor coverage in the MHC screening and was discarded; therefore, analyses are based on 309 pipits. As in most passerines (Miller & Lambert 2004; Balakrishnan *et al.* 2010; Wutzler *et al.* 2012), because there are linked duplicated loci within the class I MHC, alleles cannot be assigned to specific loci for the pipit; however for simplicity, all variants identified are termed 'alleles' hereafter. Therefore, we characterize all the class I alleles an individual carries irrespective of locus, and use the total number as a measure that reflects individual heterozygosity across the MHC loci (hereafter termed 'MHC diversity'). The presence of supertypes (groups of MHC alleles with similar binding properties) could potentially confound the analyses; however, in Gonzalez-Quevedo *et al.* (2015) we found that each allele used in this study has different chemical properties (and so could represent its own supertype), and therefore, we did not consider supertypes in further analyses.

Genetic variation and overall population structure

MICRO-CHECKER 2.2.3 (Van Oosterhout *et al.* 2004) was used to check for microsatellite null alleles and scoring errors. Allele frequencies, observed heterozygosity and expected heterozygosity for each microsatellite locus were calculated using CERVUS 3.0.3 (Marshall *et al.* 1998), and an exact test of Hardy–Weinberg equilibrium was performed using 1000 dememorization steps, 100 batches and 1000 iterations per batch in GENEPOP 4.0.10 (Rousset 2008). We estimated individual diversity at microsatellites by calculating homozygosity by loci (HL), a measure that weighs the contribution of each locus to the homozygosity value depending on its allelic variability, using Cernicalin (Aparicio *et al.* 2006). To test how well HL might represent genomewide heterozygosity, and thus, potentially, inbreeding, we also calculated the g_2 measure of identity disequilibrium (see Szulkin *et al.* 2010) in the software RMES (David *et al.* 2007).

We first divided Tenerife into four subpopulations, chosen according to climatic and topographic differences: the south (dry), the north (wet), the west (narrow coastlines and high cliffs) and El Teide (high altitude; Fig. 1). We assessed coarse patterns of genetic structure among subpopulations using analysis of molecular variance (AMOVA) and *F*-statistics, performed in ARLEQUIN 3.1 (Excoffier & Lischer 2010). Pairwise F_{ST} values were calculated by entering the allele sequences and number of individuals with each allele in each population as haplotype data. Significance of F_{ST} was evaluated using 50 000 permutations. Second, STRUCTURE 2.3.3 (Falush *et al.* 2003) was used to infer the number of genetic groups (*K*) with individual genotype-based clustering methods using microsatellite data. We used the admixture model and correlated allele frequencies with 100 000 Markov chain steps and a burn-in of 10 000 steps (the runs achieved convergence within these steps, and no further steps were needed) and performed four independent runs for each value of *K* from 1 to 4. STRUCTURE HARVESTER (Earl & von Holdt 2012) was used to visualize the results and to select the *K* value with highest support. Third, we tested whether pairwise genetic distance, based on microsatellites or MHC, correlates with pairwise geographic distance. Microsatellite genetic distance was obtained by calculating pairwise relatedness (Queller & Goodnight 1989) in COANCESTRY (Wang 2011). MHC genetic distance between each pair of individuals was calculated as the mean amino acid p-distance between all alleles present in the pair, as amino acid differences between molecules are likely to reflect functional differences between MHC alleles. We correlated microsatellite and MHC distance matrices using a Mantel test, performed using the R package 'ECODIST' (Goslee & Urban 2007) with 1000 permutations and 500 bootstrap iterations.

Landscape genetics analyses

All landscape analyses were performed in R 3.1.0 (R Development Core Team 2011) unless stated otherwise. To assess how MHC and microsatellite allele distribution varied across the landscape at a fine scale, we performed a spatial principal components analysis (sPCA; Jombart *et al.* 2008) implemented in the R package 'ADEGENET' (Jombart 2008). The sPCA assesses spatial patterns of genetic variability by finding synthetic components (eigenvectors) that maximize the product of the variance in the data and Moran's I, the latter being a measure of the spatial dependency (or autocorrelation) associated with that gradient of variation. Each eigenvector captures either positive or negative

© 2016 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

autocorrelation, and is hence referred to as either a 'global' or 'local' structure, respectively (Jombart *et al.* 2008). Local structures (negative autocorrelation) result from greater genetic differences among neighbours than expected, reflecting processes such as dispersal for inbreeding avoidance. Global structures (positive autocorrelation) arise where there are discrete clusters of genetic similarity or spatial gradients reflecting either isolation by distance or isolation by adaptation. For the sPCA, we defined neighbouring sites by building a connection network using the minimum distance that would keep all points (individual birds) in the network connected, and not leaving any unconnected points. We mapped the scores at each sampling point for the most important sPCA eigenvectors as a means of visualizing the spatial genetic structures.

Environmental variables

In a previous study of the spatial distribution of avian malaria in Berthelot's pipits in Tenerife, one predominant strain of malaria, *Plasmodium* LK6, was found to infect 36% of the pipit population (Gonzalez-Quevedo *et al.* 2014). Infection was best predicted by minimum temperature of the coldest month, distance to artificial water sources and distance to poultry farms, in decreasing order of importance (Gonzalez-Quevedo *et al.* 2014). We took the predicted values of this model, hence the predicted probability of an individual being infected with malaria based on the location inhabited (hereafter referred to as 'malaria risk') for each bird, and used this as a continuous response variable. Model predictions are a more appropriate index of malaria risk than our original infection data, because they better account for estimated environmental variation in average risk of infection. We also directly assessed the separate effects of the following continuous environmental variables on the MHC: minimum temperature of the coldest month, slope of terrain, density of pipits, distance to artificial water sources, distance to poultry farms, distance to other livestock farms and distance to urbanized sites. These variables were chosen on the basis of known effects on disease distribution (Harvell *et al.* 2002; Bradley & Altizer 2007; LaPointe *et al.* 2010; Gonzalez-Quevedo *et al.* 2014). All calculations of environmental variables for each sampling site were carried out in ARCGIS 10 and R (R Development Core Team 2011; details in Appendix S1, Supporting Information).

Models of malaria risk and infection

All models were run in R 3.1.0 (R Development Core Team 2011). We assessed whether either individual malaria risk and actual malaria infection status was associated with microsatellites or MHC characteristics using a general linear model (LM) and a generalized linear model with binomial error distribution (GLM), respectively. Malaria risk was normalized by logit-transformation prior to fitting the LMs. To test for genomewide effects of genetic diversity on malaria infection risk, we ran LMs of individual microsatellite diversity vs. malaria risk. Our expectation was that if inbred individuals with reduced diversity are more susceptible to mortality from malaria, then we would find higher microsatellite diversity in the individuals surviving in higher malaria risk areas. We then tested the association between malaria risk and MHC class I diversity by running an LM of malaria risk as the response variable and MHC diversity as an explanatory variable. We also assessed which MHC class I alleles best explained malaria risk using a multipredictor model with malaria risk as response variable and including all MHC alleles as binary explanatory variables (hence a full model). Any allele highlighted as significant by this multipredictor test was retested in a single-predictor model. We checked for spatial autocorrelation in model residuals by constructing correlograms with a 1000 m distance increment and resampling 1000 times at each distance class, implemented in the R package 'NCF' (Bjornstad 2012). Where residual spatial autocorrelation was present, we used simultaneous autoregressive (SAR) models (Kissling & Carl 2008) implemented in the R package 'SPDEP' (Bivand 2012), specifying an appropriate neighbourhood size within which autocorrelation is accounted for.

Models of MHC and microsatellite variation in the landscape

We used environmental and spatial predictors, respectively, to assess the extent to which extrinsic and intrinsic factors explain MHC and microsatellite variation (in terms of both diversity and occurrence of specific alleles). We selected our spatial predictors from a set of principal components of neighbour matrices (PCNMs; Dray *et al.* 2006), computed using a principal coordinates analysis (PCoA) performed on the spatial coordinates of sample locations using the R package 'PCNM'. To avoid unnecessarily inflating the number of spatial variables tested (Nakagawa 2004), we performed two redundancy analyses (RDAs) as a means of selecting PCNMs that significantly influence overall MHC and microsatellite allelic variation, respectively. The two resulting selections of PCNMs were then included as spatial predictors in subsequent modelling of the diversity and presence/absence of MHC and microsatellite alleles, respectively (Borcard *et al.* 2011; Legendre *et al.* 2013; further details below and in Appendix S2, Supporting Information).

© 2016 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

To test whether separate environmental variables were associated with specific MHC class I alleles, we performed multipredictor generalized linear models for each MHC allele in turn as a response variable, fitting all possible combinations of environmental variables and previously selected PCNMs as predictors. For this purpose, we used a model selection approach (Burnham & Anderson 2001) and compared the relative fit of models using the Akaike information criterion (AIC). MHC alleles were not in linkage disequilibrium with each other; therefore, we were also able to test each allele separately. To estimate the relative importance of predictors, we performed model averaging on the models with $\Delta AIC \leq 2$ relative to the best model. We also compared the fit of the models relative to the null model, a model including only the intercept. All model selection calculations were performed using the R package 'MUMIN' (Barton 2013). Model selection is a valuable alternative to traditional null hypothesis testing that is being increasingly used in studies of disease ecology (Moore & Borer 2012; Manzoli *et al.* 2013). On all occasions when multipredictor models were built, we checked our final models by comparing them with a series of single-predictor models to ensure consistency of results. To test whether the explanatory power of our MHC effects could be expected by chance, we ran single-predictor GLMs for each of the 107 microsatellite alleles (binomial response) at polymorphic loci and each of the seven environmental variables (predictors), resulting in a total of 749 models. We extracted the adjusted R^2 for every model and compared the distribution of R^2 values for the MHC and microsatellite models. Our inference from the model selection approaches were only based on comparisons of AICs and model fit (R^2). We therefore do not report *P*-values and do not apply significance thresholds. For these analyses, as for the characterization of the MHC alleles, we excluded very rare (<2%) alleles (Gonzalez-Quevedo *et al.* 2015), and alleles at >98% frequency were deemed to be fixed and not included.

Results

Neutral and MHC genetic diversity

Two of the 21 microsatellite markers showed evidence of homozygote excess and null alleles and were excluded from further analyses. The other microsatellite loci were all in Hardy–Weinberg equilibrium. The number of alleles per microsatellite locus ranged from 2 to 17, and observed heterozygosity from 0.023 to 0.860 (Table S2, Supporting Information). Identity disequilibrium was significant ($g_2 = 0.012$, SD = 0.005, $P = 0.003$), suggesting that the heterozygosity of our microsatellite panel correlates with genomewide heterozygosity. At the MHC, a total of 22 class I alleles were identified (for details, see Gonzalez-Quevedo *et al.* 2015). Two of these alleles, with frequencies (per cent individuals with the allele) of 0.84 and 0.82, had very low amplification efficiencies and are therefore likely to suffer from allelic dropout. This could mean that they are present in most, if not all, individuals but missed by the screening process in some cases. To avoid this uncertainty confounding the spatial analysis, these two alleles were excluded from the analysis. After removing these two alleles, pipits in Tenerife each had between three and ten MHC class I alleles (Fig. 1). Allele ANBE7 was also excluded from the spatial analysis because it was fixed in the population.

Overall population genetic structure

At microsatellites, we found very low but significant levels of differentiation among pipits inhabiting the four predefined zones of Tenerife ($F_{ST} = 0.008$, $P = 0.003$). Levels of MHC differentiation were even lower, and nonsignificant ($F_{ST} = 0.001$, $P = 0.770$). The STRUCTURE analysis indicated that $K = 1$ was the most likely number of genetic clusters. Pairwise genetic distance between individuals was not significantly associated with geographic distance (microsatellites: $r = -0.015$, $P = 0.339$; MHC: $r = 0.038$, $P = 0.104$).

Landscape genetics analyses

The sPCA of microsatellite genotypes showed no evidence of global structure ($P = 0.437$). While the test for local structure was not formally significant ($P = 0.055$), the plot of the eigenvector with the largest negative eigenvalue suggested short-scale spatial structure indicative of some dissimilarity among neighbours (Fig. S1, Supporting Information). The sPCA of MHC genotypes did not reveal significant patterns of global ($P = 0.587$) or local ($P = 0.732$) structure. Furthermore, when visually scrutinized, none of the global (positive eigenvalue) or local (negative eigenvalue) sPCA axes revealed any obvious spatial structuring in the distribution of MHC allelic composition. None of the alleles made an important contribution to sPCA axes. Overall, this indicates that the major axes of variation in composition of MHC genotypes across the pipit population are not spatially structured (but see PCNM results below).

Malaria risk/infection status in relation to genetic characteristics

Malaria risk was not associated with either microsatellite diversity ($P = 0.327$, $R^2 = 0.001$) or MHC class I diversity

© 2016 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

($P = 0.277$, $R^2 = 0.001$; Table S3a, Supporting Information). However, when testing the association between all MHC class I alleles and malaria risk in a multipredictor model, the presence of the ANBE48 allele was significantly negatively correlated with malaria risk (Table S3b, Supporting Information), although the explanatory power was low (single-predictor model with ANBE48 as predictor: coefficient = −0.113 ($P < 0.05$), $R^2 = 0.010$).

Models of spatial and environmental MHC and microsatellite variation in the landscape

The PCoA performed on spatial coordinates of samples with MHC genotypes (309) identified 176 PCNMs of which 89 had significant Moran's I values, the first six being positive, indicating positive spatial autocorrelation, while the remaining 83 were negative, indicating negative spatial autocorrelation. The PCoA performed on samples with microsatellite genotypes (388) identified 214 PCNMs of which 103 had significant Moran's I values, the first nine of which were positive while the remaining 94 were negative. The redundancy analyses (RDAs) performed on MHC alleles and on microsatellite genotypes and the corresponding PCNMs indicated that there was a weak association between these PCNMs and the MHC or microsatellite allele distribution (MHC alleles: adjusted $R^2 = 0.022$, $P = 0.030$; microsatellites: $R^2 = 0.012$, $P = 0.035$). This result adds further insight to the sPCA findings, by indicating that a small minority of MHC and microsatellite alleles are spatially structured. The forward selection procedure retained seven PCNMs for MHC: numbers 3, 4, 9, 13, 23, 24 and 87 (adjusted $R^2 = 0.028$, $P = 0.005$), which accounted for spatial variation at different scales (Fig. S2, Supporting Information). PCNMs 3 and 4 are positive and reflect relatively large-scale spatial structures, while the other five PCNMs are negative and reflect intermediate- to fine-scale spatial structures. For microsatellites, the forward selection retained eight PCNMs: numbers 1, 2, 3, 4, 10, 11, 16 and 66, all of which were positive and reflect large-scale spatial structures (adjusted $R^2 = 0.013$, $P = 0.005$).

In the multipredictor model of MHC diversity as the response variable, 'slope of terrain' and 'distance to poultry farm' were the only significant environmental predictors ($P = 0.002$ and 0.014, respectively; adjusted R^2 of full model (all predictor variables) = 0.034; Table 1). In combination, the PCNMs accounted for an additional 3% of variance in the response, and residual spatial autocorrelation was successfully removed. Nevertheless, overall explanatory power remained low and none of the PCNMs showed important associations with MHC diversity. In the multipredictor model of HL (microsatellite diversity) as the response variable, pipit density was the only significant predictor ($P = 0.025$, R^2 of the full model 0.013, P-value of full model = 0.174; Table 1).

Model selection and model averaging results of the multipredictor GLMs investigating the presence/absence of MHC alleles (response variables) in relation to environmental variables and significant PCNMs as

Table 1 Summary of general linear models predicting MHC diversity (No. MHC alleles) and microsatellite diversity (HL) per individual in Berthelot's pipits (*Anthus berthelotii*) in Tenerife, fitting all environmental variables and corresponding PCNMs in a multipredictor model

No. MHC alleles			HL		
Predictor	Coefficient	Adj-R^2	Predictor	Coefficient	Adj-R^2
Poultry	**−0.513***	0.034	Poultry	1.44×10^{-2}	0.013
Water	0.924		Water	1.14×10^{-3}	
Temperature	0.191		Temperature	2.19×10^{-3}	
Slope	**0.054****		Slope	1.82×10^{-3}	
Farm	0.251		Farm	9.75×10^{-2}	
Pipit density	−0.114		Pipit density	$\mathbf{7.69 \times 10^{-2}}$*	
Urbanization	0.030		Urbanization	1.68×10^{-3}	
PCNM3	-1.91×10^{-5}		PCNM1	-6.33×10^{-9}	
PCNM4	2.61×10^{-5}		PCNM2	-6.07×10^{-7}	
PCNM9	-3.67×10^{-5}		PCNM3	-2.51×10^{-7}	
PCNM13	-4.98×10^{-5}		PCNM4	-9.26×10^{-7}	
PCNM23	2.41×10^{-5}		PCNM10	-7.81×10^{-8}	
PCNM24	2.66×10^{-5}		PCNM11	1.12×10^{-6}	
PCNM87	-1.74×10^{-5}		PCNM16	-5.04×10^{-7}	
			PCNM66	4.00×10^{-6}	

Significant coefficients are in bold, significance of predictors is designated by asterisks (* <0.05, ** <0.01) and model fit is expressed as the adjusted R^2 (Adj-R^2).

© 2016 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

predictors showed different patterns for different alleles. The highest reduction in AIC compared to the null model was obtained for the model predicting ANBE48 (ΔAIC = 18.5; Table 2). Model averaging shows that distance to poultry farms has a relative importance of 1.00 in explaining ANBE48 distribution, while the other environmental variables each had a relative importance of 0.10. Of all the models for individual alleles, the best-fit model for ANBE48 showed the highest explanatory power (adjusted $R^2 = 0.310$) and included a positive association with distance to poultry farms and PCNMs 3 and 23 (Table 2; Fig. 2). The next highest explanatory power (adjusted $R^2 = 0.210$) was obtained for the best model of ANBE38 which included a negative association with distance to poultry farms, as well as a positive correlation with slope, and PCNMs 3 and 4 (Table 2; Fig. 2). The explanatory power of the multipredictor models for the other alleles was relatively low (adjusted $R^2 \leq 0.170$; Table S4, Supporting Information), indicating that associations of either environmental or spatial predictors with other alleles are lower than those identified for ANBE38 and ANBE48. Single-predictor GLMs for all 107 microsatellite alleles yielded a mean ± SE R^2 of 0.007 ± 0.0004. Importantly, none of the 749 microsatellite GLMs had as much explanatory power as the single-predictor GLM of distance to poultry farms on the presence of allele ANBE48, which was a clear outlier (Fig. S3, Supporting Information). Seven of the microsatellite (~1%) GLMs had a higher R^2 than the model of distance to poultry farms on the presence of ANBE38.

Single-predictor models broadly supported the results of multipredictor models in this study, confirming, for example, the relative importance of poultry farms in explaining ANBE48 and ANBE38 distribution ($R^2 = 0.154$ and 0.051, respectively), and the relatively low explanatory power of many of the individual predictor–response relationships investigated (Table S5, Supporting Information).

Table 2 Summary of model averaging performed on multipredictor generalized linear models (with ΔAIC ≤ 2 when compared to the best-fit model) of Berthelot's pipits (*Anthus berthelotii*) MHC class I alleles ANBE38 and ANBE48 in Tenerife as response variables. Relative importance of each variable in explaining variation in the presence of each allele after model averaging is shown in brackets

	ANBE38	ANBE48
Poultry	−2.03 (1.00)	5.27 (1.00)
Water	0.98 (0.75)	−0.24 (0.08)
Slope	0.12 (1.00)	0.02 (0.07)
Temperature	−0.01 (0.04)	−0.06 (0.09)
Farm	1.30 (0.38)	0.52 (0.09)
Urbanization	−0.72 (0.31)	0.50 (0.09)
Pipit density	−0.04 (0.04)	1.02 (0.08)
PCNM3	-8.63×10^{-5} (0.20)	4.20×10^{-4} (1.00)
PCNM4	2.61×10^{-4} (1.00)	3.36×10^{-5} (0.04)
PCNM9	2.83×10^{-5} (0.04)	1.79×10^{-4} (0.12)
PCNM13	3.34×10^{-4} (1.00)	-9.87×10^{-6} (0.04)
PCNM23	-1.82×10^{-4} (0.20)	-1.78×10^{-4} (1.00)
PCNM24	-7.86×10^{-5} (0.05)	2.05×10^{-4} (0.10)
PCNM87	-7.33×10^{-6} (0.04)	5.53×10^{-4} (0.44)
AIC Null*	139.30	90.30
AIC best†	125.36	71.78
Δ AIC	13.94	18.52
R^2 best‡	0.21	0.31

*AIC of the model with only the intercept.
†AIC of the model with the lowest AIC compared to all other possible models.
‡Explanatory power of the model with the lowest AIC.

Discussion

In this study, we assessed fine-scale structure at neutral markers and at MHC class I loci in relation to environmental factors within a population of Berthelot's pipit. We found no evidence of consistent population-wide genetic structure for either set of markers. Nevertheless, when assessing genetic characteristics in relation to specific environmental factors we did find a weak negative association between malaria infection risk and one specific MHC allele (ANBE48). Furthermore, when taking into account spatial processes independent of environmental gradients, we found stronger, and opposing, associations between two MHC alleles

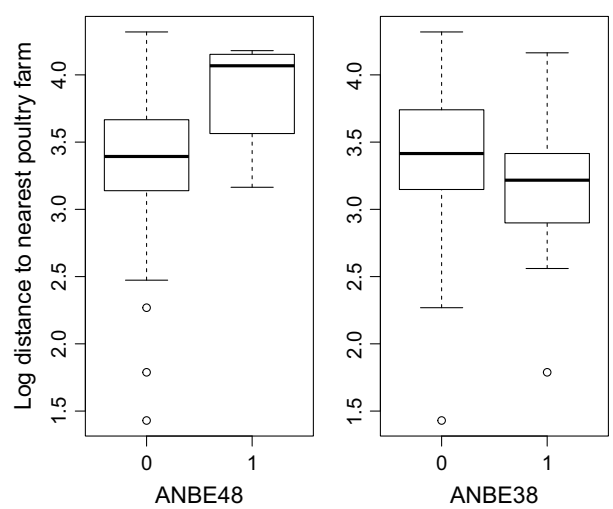

Fig. 2 Distance to nearest poultry farms (log km) across individual Berthelot's pipits carrying (1) or not (0) the ANBE38 and ANBE48 MHC class I alleles. Centre lines show the medians. Box limits indicate the 25th and 75th percentiles, and outliers are represented by open dots.

© 2016 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

(ANBE48 and ANBE38) and an anthropogenic environmental variable (distance to poultry farms) already known to be important in disease distribution in the pipit.

The STRUCTURE and F_{ST} analyses based on four predefined populations revealed no evidence of neutral or MHC structure within the population of Berthelot's pipits on Tenerife. There was also no signature of isolation by distance for either type of marker. Berthelot's pipits in Tenerife thus seem to be a panmictic population. These results were confirmed by the spatial PCAs, which found no predominant overall pattern of spatial genetic structure for either the microsatellite or MHC markers. This study therefore indicates that neither climatic differences nor apparent barriers to dispersal across Tenerife impede gene flow between different areas of the island. The opposite has been found in other systems where intrapopulation spatial clines and global (positively spatially autocorrelated) structures in allele frequencies of adaptive loci have been described at relatively small spatial scales (Garroway *et al.* 2013; reviewed in Richardson *et al.* 2014).

The lack of overall neutral and MHC-wide local or global structure was broadly supported by the very low explanatory power of PCNMs in the RDA of microsatellite genotypes or MHC genotypes. Nevertheless, the PCNMs did reveal some spatial structure, albeit for a small proportion of allelic variation. This was not detected by sPCA which optimizes the product of major components of genetic variation and spatial autocorrelation (Jombart *et al.* 2008). Our results echo those of other studies in which PCNMs are considered to be more sensitive in detecting weak spatial genetic patterns compared with sPCA (Galpern *et al.* 2014). The seven forward-selected PCNMs associated with MHC genotypes accounted for spatial variation at different scales: two of them were positive and reflected relatively large-scale spatial structures, while the other five PCNMs were negative and reflected intermediate- to fine-scale spatial structures. This suggests that a small subset of variation at the MHC decomposes into a number of spatial structures, indicating different scales of influence on different alleles or groups of alleles. This may be expected if a few specific alleles are under selective pressure, while others are evolving under neutrality, which might occur if only some MHC alleles confer resistance or susceptibility to the particular pathogens to which Berthelot's pipits are currently exposed in Tenerife. This pattern of structure at specific MHC alleles is swamped when performing analyses with genotypes, and hence, single allele effects cannot be revealed. Consequently, analyses performed on individual alleles are needed in any fine-scale genetic structure analysis.

Individual MHC alleles produce molecules which are able to bind subsets of specific pathogen-derived peptides and thus trigger the appropriate immune response to those pathogens (Wakelin 1996). Thus, MHC characteristics can be linked with pathogens in two ways. First, specific MHC alleles can confer resistance or susceptibility to a specific pathogen, and under this scenario, we would expect a correlation between allele presence/absence and disease among individuals and populations (e.g. Meyer-Lucht & Sommer 2005; Bonneaud *et al.* 2006; Zhang & He 2013). Second, individuals with greater allelic diversity may be better at responding both to individual pathogens and to the diversity of pathogens in the environment; if this is the case we expect MHC diversity to be negatively associated with disease (Westerdahl *et al.* 2005; e.g. Kloch *et al.* 2010; but see Radwan *et al.* 2012). In the present study, we found no association between malaria infection risk and individual MHC diversity. However, we did find that one allele (ANBE48) was negatively associated with increased malaria risk (Table S3b, Supporting Information), although by itself this relationship was too weak ($R^2 = 0.010$ for the single-predictor model) to be able to draw definitive conclusions (but see below). However, while our explanatory power was relatively low, these results do concur with other studies that have found MHC alleles that confer resistance/susceptibility to malaria (Hill *et al.* 1991; Bonneaud *et al.* 2006; Westerdahl *et al.* 2013).

Spatially variable selection on specific MHC alleles has been reported at large scales (Landry & Bernatchez 2001; Ekblom *et al.* 2007). Interestingly, in Berthelot's pipits we found that distribution of allele ANBE48 was positively associated with distance to poultry farms, a variable previously found to have a negative association with malaria infection in pipits in Tenerife (Gonzalez-Quevedo *et al.* 2014). In short, the closer an individual pipit was to a poultry farm (where malaria transmission has been shown to be higher, Gonzalez-Quevedo *et al.* 2014) the less likely it was to be carrying the ANBE48 allele (Fig. 2). Distance to a poultry farm explained 15% of the variation in the distribution of ANBE48, and was the most important variable in the best model for this allele. Another allele, ANBE38, was negatively associated with distance to poultry farms (Fig. 2), although the amount of variation in ANBE38 explained by this variable was not as large (5%) as for ANBE48. Comparison with the R^2 estimates for all possible bivariate microsatellite allele–environmental predictor relationships indicated that the chance that any neutral marker showed similar strength of environmental association as MHC allele ANBE48 was nil. Moreover, that we found no association between microsatellite or MHC diversity and these factors suggests that this MHC result is not

© 2016 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

explained by genomewide diversity but is directly associated with the alleles identified. This lack of a general genomewide effect is perhaps more noteworthy given that significant identity disequilibrium (g_2) indicates that variability in our microsatellite panel may correlate reasonably well with genomewide variability.

Given that (i) ANBE48 only explained 1% variation in an index of malaria risk, (ii) ANBE38 was not significantly associated with malaria risk and (iii) both ANBE48 and ANBE38 were much more significantly associated with distance to poultry farms, we hypothesize that these alleles are linked with the incidence of diseases other than just malaria. The relationship that poultry farms have with the transmission of other avian diseases that may interact with the MHC, and possibly affect the survival of pipits, has yet to be explored. However, we do know that other pathogens exist within this population, such as avian pox and haemosporidians of the genus *Leucocytozoon* (Spurgin *et al.* 2012). Therefore, assessment of the interaction of these pathogens with the MHC and of their association with poultry farms might give some insight into the mechanisms behind the association we found of ANBE48 and ANBE38 with distance to poultry farms. The relatively low explanatory power of the detected associations between environmental variables/disease risk and MHC variation is perhaps not surprising. The vertebrate immune system is an extremely complex, multifaceted, interacting system underpinned by numerous genes (Ekblom *et al.* 2010). Further exploration of variation at other key immune loci, such as Toll-like receptors and beta-defensins which play important roles in the innate immune system, would help resolve this.

Previous work has provided considerable evidence that selection (as well as drift) has shaped MHC class I variation in Berthelot's pipit (Spurgin *et al.* 2011; Gonzalez-Quevedo *et al.* 2015). In particular, selection appears to be focused on 15 sites within the exon 3 (Brown *et al.* 1993) that are involved in encoding the key peptide binding region (PBR; Gonzalez-Quevedo *et al.* 2015). Among the alleles found in the Berthelot's pipits in Tenerife, allele ANBE48 has a unique PBR, different from that of allele ANBE38 at nine amino acids, suggesting these two alleles have very different binding properties (Gonzalez-Quevedo *et al.* 2015). This is in line with our finding of opposing effects associated with these alleles. The most logical explanation we can put forward for the results that we find is that carrying the ANBE48 allele renders an individual susceptible to a pathogen that exists (or is at higher levels) around poultry farms, while ANBE38 is a nonsusceptible alternative allele. Whatever the specific pathogen, we hypothesize that birds that live close to poultry farms have a higher risk of contracting a pathogenic disease if they carry ANBE48.

Understanding the mechanisms that drive fine-scale genetic structure at adaptive loci is vital in evolutionary research (Richardson *et al.* 2014). To our knowledge, this is the first study to show an effect of a physical environmental variable on MHC variation at the intrapopulation level. The fact that this variable (the presence of poultry farms) is anthropogenic has considerable implications for understanding evolution in the context of global environmental change and human impact on disease transmission in wild populations (Daszak *et al.* 2001). Moreover, this study highlights the importance of considering fine spatial scales, in addition to coarse scales, when assessing patterns of selection at adaptive loci. Key patterns and associations may be overlooked when we lump together within-population variation to assess differences only at greater scales, potentially undermining our understanding of the factors and mechanisms that drive the evolution of the loci and species in question. Furthermore, understanding the scales, speed and causes of local adaptation within a species can have important implications for conservation, particularly when populations are challenged by new factors induced by environmental changes, be they due to habitat disturbance, agricultural changes or even conservation actions.

Acknowledgements

The research was funded by a PhD grant to CGQ from COLCIENCIAS (Departamento Administrativo de Ciencia, Tecnología e Innovación, Colombia), by a Natural Environment Research Council (UK) grant (NE/H006818/1) to DSR and by the School of Biological Sciences at the University of East Anglia. Fieldwork was partly funded by a grant from the John and Pamela Salter Trust. We would like to thank the Spanish government for providing the permits to work in Tenerife and H Bellman, JC Illera and D Padilla for field assistance. We also thank JC Illera for providing important information on the natural history of Berthelot's pipit. None of the co-authors have declared a conflict of interest.

References

Aparicio JM, Ortego J, Cordero PJ (2006) What should we weigh to estimate heterozygosity, alleles or loci? *Molecular Ecology*, **15**, 4659–4665.

Babik W, Pabijan M, Radwan J (2008) Contrasting patterns of variation in MHC loci in the Alpine newt. *Molecular Ecology*, **17**, 2339–2355.

Balakrishnan CN, Ekblom R, Voelker M *et al.* (2010) Gene duplication and fragmentation in the zebra finch major histocompatibility complex. *BMC Biology*, **8**, 29.

Barton K (2013) MuMIn: Multi-model inference. R package version 1.9.13.

Bernatchez L, Landry C (2003) MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *Journal of Evolutionary Biology*, **16**, 363–377.

© 2016 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

BirdLife International (2016) Species factsheet: *Anthus berthelotii*. Downloaded from http://www.birdlife.org

Bivand R (2012) spdep: Spatial dependence: weighting schemes, statistics and models. R package version 0.5-46.

Bjornstad ON (2012) ncf: spatial nonparametric covariance functions. R package version 1.1-4.

Bonneaud C, Perez-Tris J, Federici P, Chastel O, Sorci G (2006) Major histocompatibility alleles associated with local resistance to malaria in a passerine. *Evolution*, **60**, 383–389.

Borcard D, Legendre P (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling*, **153**, 51–68.

Borcard D, Gillet F, Legendre P (2011) *Numerical Ecology with R*. Springer, London.

Bradley CA, Altizer S (2007) Urbanization and the ecology of wildlife diseases. *Trends in Ecology & Evolution*, **22**, 95–102.

Brown JH, Jardetzky TS, Gorga JC *et al.* (1993) 3-Dimensional structure of the human class II-histocompatibility antigen HLA-DR1. *Nature*, **364**, 33–39.

Burnham KP, Anderson DR (2001) *Model Selection and Multi-Model Inference: A Practical Information–Theoretic Approach*. Springer, New York, New York.

Chave J (2013) The problem of pattern and scale in ecology: what have we learned in 20 years? *Ecology Letters*, **16**, 4–16.

Daszak P, Cunningham A, Hyatt A (2001) Anthropogenic environmental change and the emergence of infectious diseases in wildlife. *Acta Tropica*, **78**, 103–116.

David P, Pujol B, Viard F, Castellas D, Goudet J (2007) Reliable selfing rate estimates from imperfect population genetic data. *Molecular Ecology*, **16**, 2474–2487.

DeAngelis DL, Mooij WM (2005) Individual-based modeling of ecological and evolutionary processes. *Annual Review of Ecology Evolution and Systematics*, **36**, 147–168.

Dionne M, Miller KM, Dodson JJ, Caron F, Bernatchez L (2007) Clinal variation in MHC diversity with temperature: evidence for the role of host-pathogen interaction on local adaptation in Atlantic salmon. *Evolution*, **61**, 2154–2164.

Dray S, Legendre P, Peres-Neto PR (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, **196**, 483–493.

Earl DA, von Holdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Eisen RJ, Wright NM (2001) Landscape features associated with infection by a malaria parasite (*Plasmodium mexicanum*) and the importance of multiple scale studies. *Parasitology*, **122**, 507–513.

Eizaguirre C, Lenz TL, Kalbe M, Milinski M (2012) Divergent selection on locally adapted major histocompatibility complex immune genes experimentally proven in the field. *Ecology Letters*, **15**, 723–731.

Ejsmond MJ, Radwan J, Wilson AB (2014) Sexual selection and the evolutionary dynamics of the major histocompatibility complex. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20141662.

Ekblom R, Saether SA, Jacobsson P *et al.* (2007) Spatial pattern of MHC class II variation in the great snipe (*Gallinago media*). *Molecular Ecology*, **16**, 1439–1451.

Ekblom R, French L, Slate J, Burke T (2010) Evolutionary analysis and expression profiling of zebra finch immune genes. *Genome Biology and Evolution*, **2**, 781–790.

Evans ML, Neff BD, Heath DD (2010) MHC-mediated local adaptation in reciprocally translocated Chinook salmon. *Conservation Genetics*, **11**, 2333–2342.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

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

Fumagalli M, Sironi M, Pozzoli U *et al.* (2011) Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genetics*, **7**, e1002355.

Galpern P, Peres-Neto PR, Polfus J, Manseau M (2014) MEMGENE: spatial pattern detection in genetic distance. *Methods in Ecology and Evolution*, **5**, 1116–1120.

Garcia-del-Rey E, Cresswell W (2007) The breeding biology of the endemic Berthelot's Pipit, *Anthus berthelotii*, in a harsh oceanic island environment (Tenerife, Canary Islands). *Ostrich*, **78**, 583–589.

Garroway CJ, Radersma R, Sepil I *et al.* (2013) Fine-scale genetic structure in a wild bird population: the role of limited dispersal and environmentally based selection as causal factors. *Evolution*, **67**, 3488–3500.

Gonzalez-Quevedo C, Davies RG, Richardson DS (2014) Predictors of malaria infection in a wild bird population: landscape-level analyses reveal climatic and anthropogenic factors. *Journal of Animal Ecology*, **83**, 1091–1102.

Gonzalez-Quevedo C, Phillips K, Spurgin L, Richardson D (2015) 454 screening of individual MHC variation in an endemic island passerine. *Immunogenetics*, **67**, 149–162.

Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, **22**, 1–19.

Haldane JBS (1949) Disease and evolution. *La Ricerca Scientifica Suplemento*, **19**, 1–11.

Harvell CD, Mitchell CE, Ward JR *et al.* (2002) Climate warming and disease risks for terrestrial and marine biota. *Science*, **296**, 2158–2162.

Hill AVS, Allsopp CEM, Kwiatkowski D *et al.* (1991) Common West African HLA antigens are associated with protection from severe malaria. *Nature*, **352**, 595–600.

Illera JC, Emerson BC, Richardson DS (2007) Population history of Berthelot's pipit: colonization, gene flow and morphological divergence in Macaronesia. *Molecular Ecology*, **16**, 4599–4612.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

Jombart T, Devillard S, Dufour AB, Pontier D (2008) Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity*, **101**, 92–103.

Jombart T, Dray S, Dufour A-B (2009) Finding essential scales of spatial variation in ecological data: a multivariate approach. *Ecography*, **32**, 161–168.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225–1241.

Kissling WD, Carl G (2008) Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography*, **17**, 59–71.

© 2016 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

Kloch A, Babik W, Bajer A, Sinski E, Radwan J (2010) Effects of an MHC-DRB genotype and allele number on the load of gut parasites in the bank vole *Myodes glareolus*. *Molecular Ecology*, **19**, 255–265.

Landry C, Bernatchez L (2001) Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic salmon (*Salmo salar*). *Molecular Ecology*, **10**, 2525–2539.

LaPointe DA, Goff ML, Atkinson CT (2010) Thermal constraints to the sporogonic development and altitudinal distribution of avian malaria *Plasmodium relictum* in Hawaii. *Journal of Parasitology*, **96**, 318–324.

Legendre P, Legendre L (2012) *Numerical Ecology*. Elsevier, Amsterdam, the Netherlands.

Legendre P, Borcard D, Blanchet G, Dray S (2013) PCNM: MEM spatial eigenfunction and principal coordinate analyses. R package version 2.1-2/r109. Available from http://R-Forge.R-project.org/projects/sedar/.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, **17**, 183–189.

Lenz TL, Eizaguirre C, Rotter B, Kalbe M, Milinski M (2013) Exploring local immunological adaptation of two stickleback ecotypes by experimental infection and transcriptome-wide digital gene expression analysis. *Molecular Ecology*, **22**, 774–786.

Manel S, Poncet BN, Legendre P, Gugerli F, Holderegger R (2010) Common factors drive adaptive genetic variation at different spatial scales in *Arabis alpina*. *Molecular Ecology*, **19**, 3824–3835.

Manzoli DE, Antoniazzi LR, Saravia MJ *et al.* (2013) Multi-level determinants of parasitic fly infection in forest passerines. *PLoS One*, **8**, e67104.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*, **7**, 639–655.

Martín A, Lorenzo JA (2001) *Aves del Archipiélago Canario*. F. Lemus, La Laguna, Canary Islands.

Meyer-Lucht Y, Sommer S (2005) MHC diversity and the association to nematode parasitism in the yellow-necked mouse (*Apodemus flavicollis*). *Molecular Ecology*, **14**, 2233–2243.

Milá B, Warren BH, Heeb P, Thebaud C (2010) The geographic scale of diversification on islands: genetic and morphological divergence at a very small spatial scale in the Mascarene grey white-eye (Aves: *Zosterops borbonicus*). *BMC Evolutionary Biology*, **10**, 158.

Miller HC, Lambert DM (2004) Gene duplication and gene conversion in class II MHC genes of New Zealand robins (Petroicidae). *Immunogenetics*, **56**, 178–191.

Moore SM, Borer ET (2012) The influence of host diversity and composition on epidemiological patterns at multiple spatial scales. *Ecology*, **93**, 1095–1105.

Murray B, Malik S, White B (1995) Sequence variation at the major histocompatibility complex locus DQ beta in beluga whales (*Delphinapterus leucas*). *Molecular Biology and Evolution*, **12**, 582–593.

Nakagawa S (2004) A farewell to Bonferroni: the problem of low statistical power and publication bias. *Behavioral Ecology*, **15**, 1044–1045.

Ostfeld RS, Glass GE, Keesing F (2005) Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends in Ecology & Evolution*, **20**, 328–336.

Pavlova A, Amos JN, Joseph L *et al.* (2013) Perched at the mito-nuclear crossroads: divergent mitochondrial lineages correlate with environment in the face of ongoing nuclear gene flow in an Australian bird. *Evolution*, **67**, 3412–3428.

Queller DC, Goodnight KF (1989) Estimating relatedness using genetic markers. *Evolution*, **43**, 258–275.

R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Radwan J, Zagalska-Neubauer M, Cichon M *et al.* (2012) MHC diversity, malaria and lifetime reproductive success in collared flycatchers. *Molecular Ecology*, **21**, 2469–2479.

Ray JM, King RB (2006) The temporal and spatial scale of microevolution: fine-scale colour pattern variation in the Lake Erie watersnake, *Nerodia sipedon insularum*. *Evolutionary Ecology Research*, **8**, 915–928.

Richardson DS, Jury FL, Blaakmeer K, Komdeur J, Burke T (2001) Parentage assignment and extra-group paternity in a cooperative breeder: the Seychelles warbler (*Acrocephalus sechellensis*). *Molecular Ecology*, **10**, 2263–2273.

Richardson JL, Urban MC, Bolnick DI, Skelly DK (2014) Microgeographic adaptation and the spatial scale of evolution. *Trends in Ecology & Evolution*, **29**, 165–176.

Rousset F (2008) GENEPOP '007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Soto-Centeno JA, Barrow LN, Allen JM, Reed DL (2013) Reevaluation of a classic phylogeographic barrier: new techniques reveal the influence of microgeographic climate variation on population divergence. *Ecology and Evolution*, **3**, 1603–1613.

Spurgin LG, Richardson DS (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B-Biological Sciences*, **277**, 979–988.

Spurgin LG, van Oosterhout C, Illera JC *et al.* (2011) Gene conversion rapidly generates major histocompatibility complex diversity in recently founded bird populations. *Molecular Ecology*, **20**, 5213–5225.

Spurgin LG, Illera JC, Padilla DP, Richardson DS (2012) Biogeographical patterns and co-occurrence of pathogenic infection across island populations of Berthelot's pipit (*Anthus berthelotii*). *Oecologia*, **168**, 691–701.

Spurgin LG, Illera JC, Jorgensen TH, Dawson DA, Richardson DS (2014) Genetic and phenotypic divergence in an island bird: isolation by distance, by colonisation or by adaptation? *Molecular Ecology*, **23**, 1028–1039.

Svensson EI, Sinervo B (2004) Spatial scale and temporal component of selection in side-blotched lizards. *American Naturalist*, **163**, 726–734.

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

Van Oosterhout C (2009) Trans-species polymorphism, HLA-disease associations and the evolution of the MHC. *Communicative and Integrative Biology*, **2**, 408–410.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Vincent B, Dionne M, Kent MP, Lien S, Bernatchez L (2013) Landscape genomics in Atlantic salmon (*Salmo salar*):

© 2016 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.

searching for gene-environment interactions driving local adaptation. *Evolution*, **67**, 3469–3487.

Wagner HH, Fortin MJ (2005) Spatial analysis of landscapes: concepts and statistics. *Ecology*, **86**, 1975–1987.

Wagner HH, Fortin M-J (2013) A conceptual framework for the spatial analysis of landscape genetic data. *Conservation Genetics*, **14**, 253–261.

Wakelin D (1996) *Immunity to Parasites: How Parasitic Infections are Controlled*, 2nd edn. Cambridge University Press, Cambridge.

Wang J (2011) COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Molecular Ecology Resources*, **11**, 141–145.

Westerdahl H, Waldenstrom J, Hansson B *et al.* (2005) Associations between malaria and MHC genes in a migratory songbird. *Proceedings of the Royal Society B-Biological Sciences*, **272**, 1511–1518.

Westerdahl H, Stjernman M, Raberg L, Lannefors M, Nilsson J-A (2013) MHC-I affects infection intensity but not infection status with a frequent avian malaria parasite in blue tits. *PLoS One*, **8**, e72647.

Wood MJ, Cosgrove CL, Wilkin TA *et al.* (2007) Within-population variation in prevalence and lineage distribution of avian malaria in blue tits, *Cyanistes caeruleus*. *Molecular Ecology*, **16**, 3263–3273.

Wutzler R, Foerster K, Kempenaers B (2012) MHC class I variation in a natural blue tit population (*Cyanistes caeruleus*). *Genetica*, **140**, 349–364.

Zhang M, He H (2013) Parasite-mediated selection of major histocompatibility complex variability in wild Brandt's voles (*Lasiopodomys brandtii*) from Inner Mongolia, China. *BMC Evolutionary Biology*, **13**, 149.

D.S.R. initiated and managed the overall project. D.S.R., C.G.Q., R.G.D. and L.G.S. all contributed to designing the study. Fieldwork was undertaken by C.G.Q. with help from D.S.R., L.G.S. and K.P.P. C.G.Q. did the molecular work and associated analysis with the assistance of L.G.S. and K.P.P. Spatial analyses were devised by C.G.Q. and R.G.D. and undertaken by C.G.Q. with input from the other authors. C.G.Q. wrote the manuscript to which all the authors contributed.

Data accessibility

Sequences of MHC class I alleles have been deposited in GenBank, accession nos KM593305–KM59331. MHC and microsatellite genotype files have been deposited in Dryad DOI http://dx.doi.org/10.5061/dryad.40vv8.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Summary of the 21 markers used for assessing neutral genetic variation in 388 Berthelot's pipits (*Anthus berthelotii*) from Tenerife.

Table S2 Summary of marker diversity at 19 microsatellites in Berthelot's pipits (*Anthus berthelotii*) from Tenerife: k = number of alleles; H_{Obs} = observed heterozygosity; H_{Exp} = expected heterozygosity.

Table S3 Summary of general linear models with malaria risk as the response variable in Berthelot's pipits (*Anthus berthelotii*) from Tenerife fitting, (a) number of MHC alleles and microsatellite diversity (HL), and (b) all MHC alleles in a multi-predictor model.

Table S4 Summary of model selection performed on multi-predictor generalized linear models (with $\Delta AIC \leq 2$ when compared to the best fit model) of MHC class I variation in Berthelot's pipits (*Anthus berthelotii*) from Tenerife.

Table S5 Summary of generalized linear models predicting presence of MHC class I aleles in Berthelot's pipits (*Anthus berthelotii*) in Tenerife, fitting single environmental predictor variables.

Fig. S1 Scores from the largest local eigenvector (Axis 129, $P = 0.055$) of spatial principal component analysis of microsatellite allele frequencies in Berthelot's pipits (*Anthus berthelotii*) from Tenerife.

Fig. S2 Scores at each sampled location from the seven PCNMs significantly associated with the spatial distribution of MHC class I alleles in Berthelot's pipits (*Anthus berthelotii*) from Tenerife, as determined by forward selection within a redundancy analysis (RDA).

Fig. S3 R^2 values from single predictor GLMs of environmental variation on presence/absence of microsatellite and MHC alleles in Berthelot's pipits (see main text for details).

Appendix S1. Methodology for environmental variable calculation and resampling.

Appendix S2. Detailed methodology for selection of spatial predictors from a set of principal components of neighbour matrices as detailed in the text.

© 2016 The Authors. *Molecular Ecology* Published by John Wiley & Sons Ltd.