



FINE-SCALE GENETIC STRUCTURE IN A WILD BIRD POPULATION: THE ROLE OF LIMITED DISPERSAL AND ENVIRONMENTALLY BASED SELECTION AS CAUSAL FACTORS

Colin J. Garroway,^{1,2,*} Reinder Radersma,^{1,3,*} Irem Sepil,¹ Anna W. Santure,⁴ Isabelle De Cauwer,^{4,5} Jon Slate,⁴ and Ben C. Sheldon¹

¹Edward Grey Institute, Department of Zoology, University of Oxford, Oxford, OX1 3PS, United Kingdom

²E-mail: colin.garroway@zoo.ox.ac.uk

³E-mail: reinder.radersma@zoo.ox.ac.uk

⁴Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, United Kingdom

⁵Laboratoire de Génétique et Evolution des Populations Végétales UMR CNRS 8198, Bâtiment SN2, Université des Sciences et Technologies de Lille—Lille 1 F-59655, Villeneuve d'Ascq Cedex, France

Received September 4, 2012

Accepted March 24, 2013

Data Archived: Dryad doi:10.5061/dryad.4f46b

Individuals are typically not randomly distributed in space; consequently ecological and evolutionary theory depends heavily on understanding the spatial structure of populations. The central challenge of landscape genetics is therefore to link spatial heterogeneity of environments to population genetic structure. Here, we employ multivariate spatial analyses to identify environmentally induced genetic structures in a single breeding population of 1174 great tits *Parus major* genotyped at 4701 single-nucleotide polymorphism (SNP) loci. Despite the small spatial scale of the study relative to natal dispersal, we found multiple axes of genetic structure. We built distance-based Moran's eigenvector maps to identify axes of pure spatial variation, which we used for spatial correction of regressions between SNPs and various external traits known to be related to fitness components (avian malaria infection risk, local density of conspecifics, oak tree density, and altitude). We found clear evidence of fine-scale genetic structure, with 21, seven, and nine significant SNPs, respectively, associated with infection risk by two species of avian malaria (*Plasmodium circumflexum* and *P. relictum*) and local conspecific density. Such fine-scale genetic structure relative to dispersal capabilities suggests ecological and evolutionary mechanisms maintain within-population genetic diversity in this population with the potential to drive microevolutionary change.

KEY WORDS: Genome-environment association study, landscape genetics, Moran's eigenvector maps, *Parus major*, *plasmodium*, SNPs.

Individuals are typically not randomly distributed in space; rather, they aggregate in patches or groups, or are distributed along gradients in spatially variable physical and social environments. This spatial structure is key to much of ecological and evolutionary theory including social evolution, mate choice, epidemi-

ology, population ecology, the maintenance of genetic diversity, adaptation, and speciation (Gaston 2003; Coyne and Orr 2004; Fortin and Dale 2005; Bourke 2011; Nosil 2012). Individuals at different resource patches are less likely to meet and thus unlikely to be selected as mates or to transfer disease or parasites, whereas individuals within patches, or in closer spatial proximity, are more likely to interact with each other, and hence influence

*These authors contributed equally to this work.

selection on social properties. Divergent selection and nonrandom gene flow in space lead to population differentiation, can result in local adaptation (Bohonak 1999; Lenormand 2002; Coyne and Orr 2004), and in a subset of cases can lead to speciation (Bohonak 1999; Coyne and Orr 2004). Clearly, spatial heterogeneity of environments plays a general and fundamental role in many ecological and evolutionary processes. Within this context, the central challenge of landscape genetics is to address the role of spatial heterogeneity of environments as contributing factors to neutral and adaptive components of population genetic structure from evolutionary and ecological perspectives.

Within populations, divergent selection in space may contribute to the maintenance of genetic diversity (Hargeby et al. 2004; Postma and van Noordwijk 2005) and is crucial for adaptive responses to environmental changes. Divergent selection can even lead to reproductive barriers if immigration to new environments is selected against (Nosil et al. 2005). When selection pressures are strong enough, divergent selection should be apparent from correlations between genetic polymorphisms and the environment. One particular difficulty in identifying these correlations is teasing apart and controlling for the relative roles of various sources of spatial variation.

Spatial genetic structures may be generated by two nonexclusive types of process. The first is *spatial autocorrelation* (sensu stricto; Sokal and Oden 1978; Fortin and Dale 2005; Legendre and Legendre 2012). In this case, spatial structure emerges as a result of endogenous biotic processes, such as genetic drift, inherent dispersal tendencies (i.e., isolation-by-distance), kin structure, and shared population histories. The second involves *induced spatial dependence* (Fortin and Dale 2005; Legendre and Legendre 2012). Here, spatial genetic structure is induced by relationships with external spatially structured explanatory variables; for example, when a population responds to a spatially structured environmental process, either by selection, local adaptation, or differential dispersal. Landscape genetics studies have typically been most interested in understanding the consequences, in terms of microevolutionary processes and gene flow, of species interactions with their environment (i.e., induced spatial dependence), while controlling for spatial autocorrelation as a nuisance variable. We suggest that very often both processes leading to spatial dependence are biologically interesting and our approach here models both explicitly.

In this study, we aimed to identify sources of environmentally based selection causing spatial genetic structure at a fine-spatial scale relative to dispersal. Specifically, we test for evidence of microevolutionary responses to environmental traits which are known to affect fitness and which vary over space within an intensively studied population of great tits (*Parus major*). The environmental traits we analyze are the risk of infection with malaria, local density of conspecifics, the local density of oak

trees *Quercus* spp. (a key food plant for larvae that are an important component of the diet of nestling great tits; Gosler 1993), and altitude. In each case, there is evidence that fitness is correlated with these environmental traits (see section Materials and Methods) and hence that the environmental trait has the potential to act as a spatially variable source of selection.

Numerous approaches for controlling for spatial autocorrelation have been proposed for landscape genetics. These include (1) partialling out the effects of isolation-by-distance and assessing whether the remaining variation in the data are related to landscape features (e.g., Cushman et al. 2006); (2) incorporating isolation-by-distance into measures of landscape resistance and proceeding with standard model fitting (e.g., Garroway et al. 2011); (3) using mixed effects models to account for lack of independence (e.g., van Strien et al. 2012); (4) fitting gravity models to networks of population genetic connectivity (e.g., Murphy et al. 2010); and (5) fitting models with multiple regression on matrices (e.g., Dyer et al. 2010). Here, we take a different approach (see Manel et al. 2010, 2012; Lasky et al. 2012). Rather than treating space as a nuisance variable to be controlled for, we treat spatial patterns in the data as indicative of often unmeasured processes of dispersal, drift, and historic influences on population structure and incorporate space explicitly into our models (McIntire and Fajardo 2009; Manel et al. 2010, 2012; Dray et al. 2012). To do this we calculate Moran's eigenvector maps, which are eigenvectors from a truncated spatial distance network that is built from the spatial coordinates of sample sites (Borcard and Legendre 2002; Dray et al. 2006). Moran's eigenvector maps summarize spatial patterns across all scales as orthogonal vectors of a distance matrix. They can be incorporated into regression analyses as spatially structured proxy variables accounting for spatial autocorrelation-related processes, such as isolation-by-distance, shared histories, and kin structure (Manel et al. 2010, 2012; Lasky et al. 2012).

Thus, our goal was to identify instances of within-population selection on genetic variants (in this case single-nucleotide polymorphisms [SNPs], derived from a large-scale genotyping project) associated with spatially structured environmental variables of differing spatial and temporal predictability. We expected that the effects of environmentally induced habitat or natural selection on spatial genetic structure would be most apparent for those variables that are spatially predictable over time.

Materials and Methods

STUDY SYSTEM

This study was conducted in the great tit population in Wytham Woods, a 385 ha mixed deciduous woodland close to Oxford, U.K. (51°46'N, 1°20'W). The study site contains 1020 nest boxes of which 250–450 are used annually for breeding by great tits.

As part of a long-term monitoring project, all birds are ringed for individual identification and breeding performance has been recorded systematically since the early 1960s (Lack 1964). In this study we analyzed the relationship between spatially variable environmental gradients and genetic variation in 1174 individual birds that bred in the study population in 2008 or 2009. Because about half of the breeding birds survive until the next year (Clobert et al. 1988; Bouwhuis et al. 2012), many individuals were present in both years. As breeding location, we took the earliest recorded breeding location for those individuals that were observed in >1 year. This is a good proxy for all years because, after natal dispersal, individuals tend to breed close to the location they bred before (e.g., 61% of breeding birds move <100 m between years, 87% move <200 m and 94% move <300 m; Harvey et al. 1979). We analyzed only breeding adults because these birds had a well-defined spatial location, and have potentially undergone extensive selection by environmental variables.

EXTERNAL VARIABLES

We studied the effects of three different environmental variables: (1) malaria infection risk; (2) the density of oak trees *Quercus* spp.; (3) altitude; and (4) an endogenous property of the population, the local density of conspecifics.

Avian malaria in tits

The spatial distribution, epidemiology, and effects of avian malaria (*Plasmodium* spp.) on great and blue tits (*Cyanistes caeruleus*) have been well studied in this population (Wood et al. 2007; Knowles et al. 2010a, b, 2011; Lachish et al. 2011a, b, 2013). Two *Plasmodium* species are common in the study area (*P. circumflexum* and *P. relictum*) and differ in important aspects with regard to their spatial epidemiology and effects on hosts. *Plasmodium circumflexum* has pronounced, temporally stable, fine-scale spatial structure, with more than 10-fold relative increase in infection risk over space in the study population, whereas *P. relictum* is much less predictably distributed in space and time and with less variation in relative infection risk in space (Knowles et al. 2011; Lachish et al. 2013). *Plasmodium circumflexum* reduces survival in blue tits to a much greater extent than *P. relictum* (Lachish et al. 2011a) suggesting, together with pronounced spatial structure, the opportunity for parasite-mediated selection driven by *P. circumflexum*, but less so with *P. relictum* (Lachish et al. 2013). Similarly, reduced survival is observed in great tits infected by *P. circumflexum* but not *P. relictum* (I. Sepil et al., unpubl. ms.). In 2008 and 2009, 85% of breeding great tits (2008: 643 of 756; 2009: 472 of 556) were screened for malaria infection (*P. circumflexum* and *P. relictum*) as described in Lachish et al. (2013). We calculated infection risk for individuals for each malaria species as the proportion of infected conspecifics within a 500 m radius (500 m gave qualitatively similar results to 100 m

and 250 m radii) around breeding location for both years combined (Lachish et al. 2013).

Oak tree density

Oak tree (*Quercus robur* and *Q. petraea*) density is an important indicator of breeding season territory quality for great tits in this population (Wilkin et al. 2009; Wilkin and Sheldon 2009). Caterpillars (e.g., winter moth *Operophtera brumata*, and green tortrix *Tortrix viridana*) are a major food resource fed to nestlings (Gosler 1993) and are most abundant on freshly emerged oak leaves. The density of oak trees is temporally predictable and so, we might expect the opportunity for selection associated with exploiting this resource. For the density of the oak trees, we used the number of oak trees within a radius of 50 m (approximate breeding territory size) of every breeding location as described and calculated by Wilkin et al. (2007), but based on mapping of oak tree locations in 2010.

Local conspecific density

Nestbox spacing in the Wytham Woods study varies between different sections of the woodland, with a range from <1 ha⁻¹ to >6 ha⁻¹; this variable spacing results in variation in local conspecific density. Clutch size, fledging mass, and the number of offspring recruited to the population per breeding attempt are all positively related to territory size, a measure of local conspecific density, in the study population (Wilkin et al. 2006; Wilkin and Sheldon 2009). Although population size fluctuates annually, it has tended to increase over the past three decades (e.g., Garant et al. 2004). However, relative local conspecific density has been consistent over the longer term, owing to the fixed position of nestboxes. Local conspecific density was calculated for each individual as the number of conspecific breeding pairs per hectare of forest within a 500 m radius (for consistency with the malaria infection risk measures) of every breeding location. Local conspecific density was calculated for 10 consecutive years (2001–2011) and averaged over those years.

Altitude

Altitude varies by 106 m within the study site, and while this difference may seem relatively small, it is associated with predictable differences in vegetation phenology. Probably as a consequence, clutches are laid later at higher altitudes. In addition, fledglings are lighter as altitude increases in the study area (Wilkin and Sheldon 2009). A further strong environmental effect related to altitude concerns soil calcium concentration: the study site consists of two Corallian limestone hilltops surrounded by layers of sand and clay. As a consequence, calcium concentration in the soil varies 300-fold along the altitudinal gradient, and there is evidence that at low calcium concentrations, calcium may be limiting with respect to reproductive output (Wilkin et al. 2009). Altitude

and the environmental conditions associated with it are spatially structured and predictable. Altitude for every breeding location was calculated from a 50-m resolution Land Form PROFILE Digital-Terrain-Model dataset provided by Ordnance Survey as described and calculated by Wilkin et al. (2007).

GENOTYPES

A SNP chip with 9193 markers was developed based on transcriptome sequencing of great tits from the Wytham Woods population and genomic sequencing of great tits from populations in the Netherlands (see van Bers et al. 2010; Santure et al. 2011; van Bers et al. 2012 for details). Of those 9193 markers, 7032 passed quality control (using the criteria genotyping frequency >95%, minor allele frequency >0.05, and Hardy–Weinberg equilibrium $P > 0.001$, calculated using PLINK v1.06; Purcell et al. 2007) and 4878 were incorporated into the linkage map for our population. Because we can expect, given sex-biased dispersal distances, divergent gene flow for alleles located on the Z-chromosome relative to alleles on the autosomes (Mank et al. 2010), we only analyzed the 4701 SNPs located on the autosomes. These markers were used to genotype 2652 great tits (van Bers et al. 2012), of which 1174 formed the large majority of the breeding population in 2008–2009 and thus formed the basis of this study.

SPATIAL ANALYSES

We examined the spatial structure of our environmental data with Moran's I correlograms and genetic data with a Mantel correlogram. We used spatial principal components analysis (sPCA), a spatially explicit multivariate method that describes allele frequency variation, to search for discrete genetic clusters, allele frequency gradients, or both (Jombart et al. 2008). sPCA optimizes the product of the variance in the data and Moran's I (Moran 1948, 1950) to build synthetic components that summarize spatial patterns of allele frequencies (Jombart et al. 2008). Global (positive eigenvalues) and local (negative eigenvalues) eigenvectors are built. Global scores can be used to identify distinct genetic clusters and spatial clines. Local scores can be used to detect differentiation between neighboring sites. To calculate Moran's I, we needed to define neighboring nest sites. We did this by building a neighbor network that linked all nestboxes within the minimum distance that would keep the network connected (i.e., the longest edge in a minimum spanning tree); that is, all nestboxes within 296 m of each other were linked in our case (mean number of neighbors was 61). We tested the null hypothesis that the raw data were distributed randomly on the neighbor network to indicate which structures, global or local, should be interpreted (Jombart et al. 2008). Nonrandom distributions indicate that the scores display some spatial structure, but does not suggest how many axes are important; the null distributions were determined via permutation ($n = 999$). Local structures are more likely to be related to

behavioral processes (e.g., territoriality, or spatial disassortment via social interactions) than spatially predictable environmental variation, so here we focus on global structures. sPCA makes no assumptions regarding Hardy–Weinberg or linkage equilibrium. sPCA analysis was calculated using the adegenet package (Jombart 2008) and visualized using the ade4 (Dray and Dufour 2007) package in R (The R Core Development Team 2012).

Population structure can be a result of individual responses to spatially structured environmental variables (induced spatial dependence), population dynamics and history (spatial autocorrelation), or both. Several methods have been developed to incorporate space as a predictor in multivariate analyses (see Legendre and Legendre 2012 for an overview). Prominent among these are (1) the inclusion of spatial coordinates directly into models (trend surface analyses) and (2) the construction of distance-based Moran's eigenvector maps (dbMEMs), which model spatial structure at multiple scales independently based upon eigenvector decomposition of distance matrices. In trend surface analysis, polynomial transformations of x - y coordinates of increasing power increase the local flexibility in the spatial trends that can be modeled; however, they also increase the number of parameters needed to model a spatial trend, and so in practice, trend surface analysis is constrained to identify only broad spatial patterns, and parameters are highly correlated. In contrast, dbMEMs are variables that represent spatial structures by drawing eigenvectors through a distance matrix calculated from the spatial configuration of samples. They summarize spatial structures at decreasing scales. Thus, dbMEMs serve to identify and quantify spatial variation in response data due to spatial autocorrelation (Legendre and Legendre 2012). dbMEMs (Dray et al. 2006) were first developed as principal coordinates (PCoA) of neighbor matrices (Borcard and Legendre 2002) and generalized by Dray et al. (2006). They are called Moran's eigenvector maps because the eigenvalues are equal to Moran's I coefficients of the neighbor network multiplied by a constant (Dray et al. 2006). Moran's I is directly analogous to Malecot's estimator of spatial correlations among gene frequencies (Malecot 1955; Epperson 2005) and has been demonstrated to accurately describe processes leading to neutral variation in gene flow and frequencies (e.g., Sokal and Oden 1978; Epperson 2005). We used dbMEMs to generate independent spatial variables for inclusion in a genome–environment association study (GEAS) aiming to identify relationships between alleles, space, and the environment. Conceptually, these dbMEM variables can be thought of as independent vectors summarizing the spatial structure associated with the neighborhood network (the distance matrix) across scales. They are used as estimates of the variation in spatial structure of the SNPs due entirely to spatial autocorrelation-related processes (Peres-Neto and Legendre 2010).

To calculate dbMEMs we (1) computed a distance matrix from the spatial coordinates of all genotyped individuals, (2) chose

a threshold distance to truncate the geographic distances, and (3) computed PCoA of the truncated distance matrix. The resulting dbMEMs were models of distance relationships among breeding location and were used as explanatory variables to model spatial processes. We used the same neighbor network as we used in the sPCA as our truncated distance matrix. Positive eigenvalues for dbMEMs correspond to the Euclidean representation of the neighbor network, are orthogonal by definition, and thus can be included in regression analyses as spatial predictors (Borcard and Legendre 2002; Dray et al. 2006). As with the sPCA, positive eigenvalues describe global structures and negative eigenvalues describe local structures. Here too it seems unlikely that local structures will be related to environmental processes, so we only examined global dbMEMs. Not all dbMEMs will be related to genetic structure; we identified the important dbMEMs related to genetic structure following Blanchet et al. (2008). We first tested for a relationship between the allele frequency data and all global dbMEMs with a redundancy analysis. If there was a relationship, we used forward selection with a double stopping criterion such that variables that increased the α above 0.05 (arbitrarily chosen) or raised the adjusted R^2 above that of the global model were not retained. We differed from Blanchet et al. (2008), in that we used χ^2 -tests rather than F -tests. Simulations have shown that this approach corrects for increased type 1 errors and inflated R^2 values typical of forward selection based only upon an alpha stopping rule (Blanchet et al. 2008). dbMEM eigenvalues describe spatial variation from broad (largest eigenvalues) to fine (smallest eigenvalue) scales. dbMEMs were calculated with the PCNM package (Legendre et al. 2012) for R (R Core Development Team 2012). Detailed descriptions of this approach can be found in Legendre and Legendre (2012) and example R code in Borcard et al. (2011).

GEAS

We compared the SNPs with the environmental traits by performing SNP by SNP generalized linear regressions on each trait. In the nonspatial GEAS, we used each SNP as a response variable and environmental traits and local conspecific density as explanatory variables. In the spatial GEAS, we added significant dbMEM spatial variables as covariates to the nonspatial GEAS. *Plasmodium circumflexum* and *P. relictum* infection risk were normalized by logit transformation and local conspecific density was normalized by a logarithmic transformation. All regressions had a Gaussian error distribution except for the oak tree density that had a Poisson error distribution and log link. To correct for multiple testing, we used the Bonferroni correction that set the significance level α at $0.05/4701 = 0.00001063$, so $-\log_{10}(\alpha) = 4.973$. All calculations were performed in R 2.15.0 (R Core Development Team 2012).

We compared the results of the GEAS to more standard approaches in population genetics in which we used the social

pedigree to correct for relatedness (Thompson and Shaw 1990; Amin et al. 2007; Aulchenko et al. 2007a) and genomic control to correct for population stratification (Amin et al. 2007) by using GenABEL (Aulchenko et al. 2007b)—see Figure S1. Here, to be consistent with these other approaches, we treated the environmental variables and population density as response variables and SNP genotypes as explanatory variables. Conceptually, this differs in that it suggests that genotypes select the environment. The results of the pedigree models did not differ from the GEAS, but when genomic control was introduced to the models all significant SNPs disappeared. Genomic control corrects for population structure regardless of its origin (Devlin and Roeder 1999), and hence will have the effect of removing population stratification originating from spatial structure, which is the key phenomenon of interest in this study. In the main body of the MS, we will only present the results of the GEAS.

Results

General details of the SNP genotyping results can be found in van Bers et al. (2012). There were varying degrees of spatial correlation in the molecular and environmental data (Fig. 1). Allele frequencies were positively correlated below distances of approximately 700 m and became negatively correlated at distances over about 1500 m, suggesting that individuals living less than 700 m apart were genetically more similar than others, whereas individuals living more than 1500 m apart were genetically more dissimilar. Both strains of malaria and oak tree density were also spatially correlated across the study area, but there were no strong patterns of spatial autocorrelation for local conspecific density and altitude (Fig. 1).

Spatial principal components analysis (sPCA) suggested that global axes should be further interpreted (global $P < 0.01$; Figs. 2, S2). The first global axis showed a clear cline in allele frequencies from southeast to northwest (Fig. 2B). The second global axis showed a less pronounced north–south cline, but differentiated a southern “peninsula” part of the population from the rest (Fig. 2C); the third global axis differentiated the western-most portion of Wytham Woods from the rest of the wood (Fig. 2D). There was no evidence for discrete genetic clusters.

We identified 201 positive dbMEM variables. Of these, 169 had significant Moran’s I values ranging from fine to broad spatial scales suggesting that the redundancy analysis could model all spatial scales (e.g., from local family structure and shared ancestry to broad-scale patterns, such as dispersal). The redundancy analysis suggested that there was a weak relationship between the allele frequencies and these dbMEM variables (adj $R^2 = 0.057$, $P < 0.005$) and there was no evidence of residual spatial variation (Mantel test: $nperm = 999$, $r = -0.025$, $P = 0.888$); thus, we

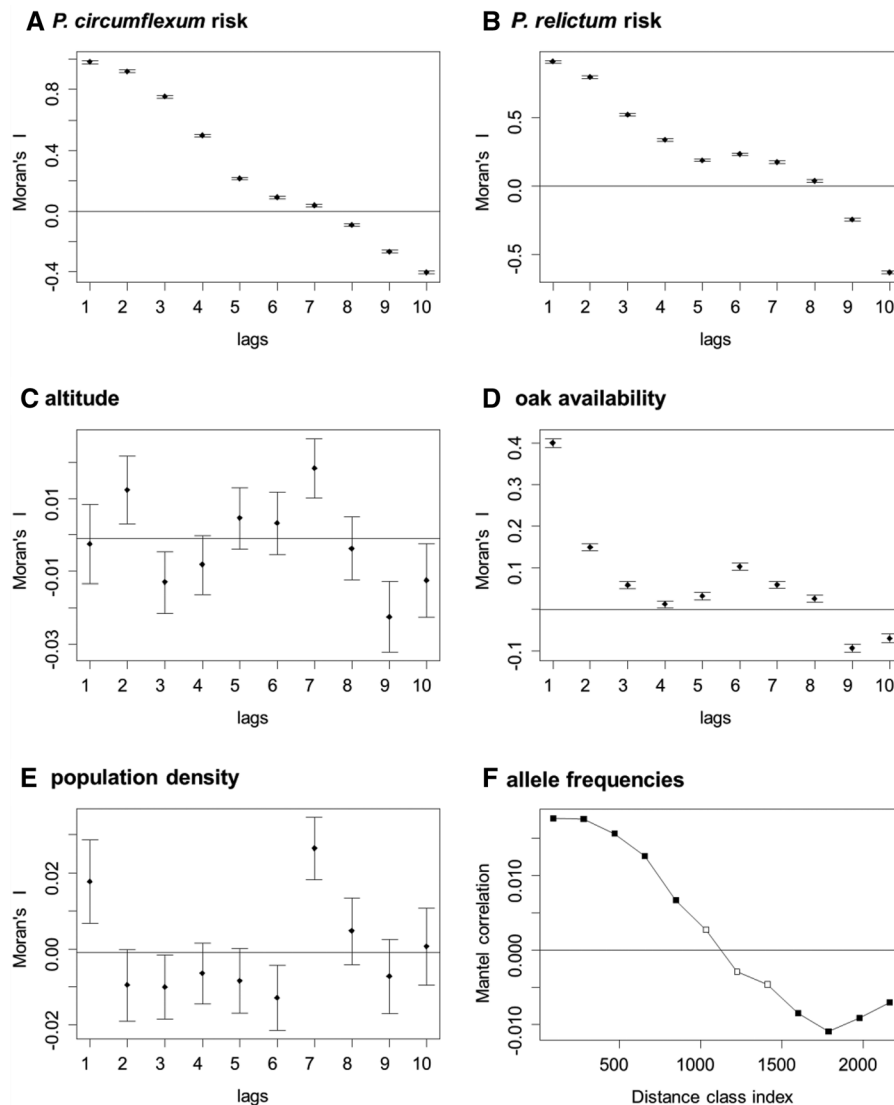


Figure 1. Moran's I correlograms illustrating the spatial patterns in variables for (A) *Plasmodium circumflexum* risk, (B) *P. relictum* risk, (C) altitude, (D) oak tree density at nest sites, (E) local conspecific density and (F) a Mantel correlograms of allele frequencies of great tits in Wytham Woods, Oxford, U.K. Morans's I standard errors are calculated by randomization. Significant correlation at distance classes for Mantel correlograms is indicated by filled squares.

continued with the forward variable selection. We retained the first, second, third, fifth, and sixth dbMEMs ($\text{adj } R^2 = 0.041$, $P < 0.005$) for inclusion as spatial predictors in the GEAS. That these are five of the six largest eigenvalues indicates that allele frequencies are structured at the broadest spatial scales of our study site with no detectable signatures at mid to fine scales apparent in the data. By including these variables in our models, we estimate and account for the effects of spatial autocorrelation-related processes (e.g., kin structure and isolation-by-distance). Combinations of dbMEMs were related to each of the environmental variables as follows: (1) *P. circumflexum* infection risk was positively related to dbMEMs 1, 2, and 6 and negatively related to dbMEM 5; (2) *P. relictum* infection risk was positively related

to dbMEM 5 and negatively related to dbMEM 6; (3) altitude was negatively related to dbMEM 6; (4) local conspecific density was negatively related to dbMEM 1 and positively related to dbMEM 5; (5) oak tree density was negatively related to dbMEM 5 and positively related to dbMEM 6. dbMEM 3 was not related to any of our chosen environmental variables and so is likely to relate to some other unmeasured spatial environmental variables that influence genetic structure and was included in the spatial GEAS.

The spatially explicit GEAS suggested that *P. circumflexum* infection risk, *P. relictum* infection risk, and local conspecific density were significantly explained by allelic variation at 21, 7, and 9 SNPs, respectively (Fig. 3; see Fig. S3 for the locations



Figure 2. The (A) eigenvalues of spatial principal components analysis of allele frequencies of great tits breeding in Wytham Woods, Oxford, U.K. and (B) scores from the first, (C) second, and (D) third global eigenvectors plotted at nest box sites throughout the wood visualizing the spatial genetic structure of breeding birds in the population.

of those SNPs on the genome); nonspatial GEAS suggested 24, 10, and 11 significant SNPs, respectively (Fig. 3). All important dbMEMs for explaining spatial variation in environmental variables remained significant when the individual SNPs were added to the model in the spatial GEAS; however, after the spatial dbMEM variables were included eight SNPs lost significance. Those SNPs that lost significance in the presence of dbMEMs are likely to have been significant due to spatial autocorrelation, and not through environmentally induced spatial dependence. As both the SNPs identified as being putatively under selection in the spatial GEAS and the dbMEMs maintained significance, their spatial structure is explained by a combination of environmentally induced spatial structure and spatial autocorrelation.

There were strong and significant correlations between the parameter estimates of the spatial GEAS across all SNPs for *P. circumflexum* and *P. relictum* infection risk ($r = -0.89$, $N = 4701$, $P < 0.001$), *P. circumflexum* infection risk and local conspecific density ($r = -0.66$, $N = 4701$, $P < 0.001$), *P. relictum* infection risk and local conspecific density ($r = 0.61$, $N = 4701$, $P < 0.001$), and local conspecific density and altitude ($r = -0.66$, $N = 4701$, $P < 0.001$), which indicates that SNPs correlating strongly with one trait tended to correlate strongly with both other traits as well (see Fig. S4, all other correlations were significant as well, but $r < 0.5$). This might originate from the correlations between some of the environmental traits (*P. circumflexum* and *P. relictum* infection risk, $r = -0.82$, $N = 1174$, $P < 0.001$; *P. circumflexum* infection risk and local conspecific density, $r = -0.45$, $N = 1174$, $P < 0.001$; *P. relictum* infec-

tion risk and local conspecific density, $r = 0.39$, $N = 1174$, $P < 0.001$; and local conspecific density and altitude $r = -0.60$, $N = 1174$, $P < 0.001$; see Fig. S4). It is possible that space is an important underlying factor contributing to the correlations. If so, we may have been able to account for and remove the spatial component of the correlations in a similar manner to the spatial GEAS with dbMEMs. Thus, as a post hoc analysis, we identified the dbMEMs important for explaining the spatial structure of these three variables. There were 29 important dbMEMs for *P. relictum* ($R^2 = 0.15$), 28 for *P. circumflexum* ($R^2 = 0.16$), and 30 for local conspecific density ($R^2 = 0.14$). The malaria strains shared 20 dbMEMs and nine of these were common to all three variables suggesting that shared spatial processes explained some of the correlation among environmental variables. We regressed these nine shared dbMEMs on the environmental variables and local conspecific density independently to account for the shared spatial structure and investigated correlations among residuals. In each case considerable correlations remained suggesting that common underlying spatial patterns were not solely responsible for the correlation among these variables. We therefore cannot draw firm conclusions regarding which of these three variables affected the genetic structure solely from these results.

Discussion

We used a high-density SNP genotype dataset, applied to a single population of passerine birds, to explore associations between

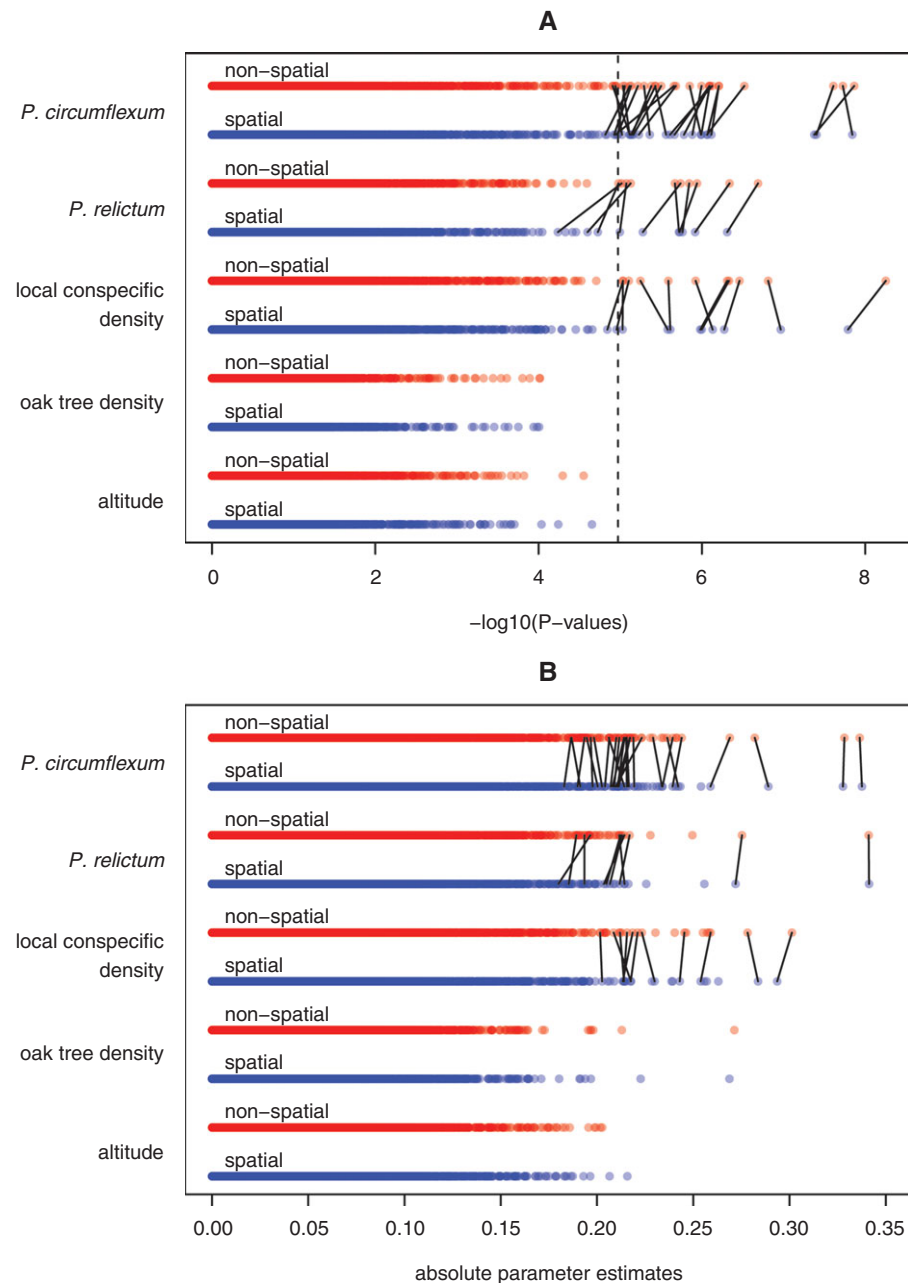


Figure 3. The (A) *P* values and (B) parameter estimates for the genome-environment association studies (GEASs) between single-nucleotide polymorphisms (SNPs) and various environmental traits for great tits breeding in Wytham Woods, Oxford, U.K. Red dots are correlations that were not corrected for space, whereas blue dots are spatial corrected correlations. SNPs that were significant (either for the nonspatial or the spatial models, or both) are indicated by a line connecting the dots representing the same SNP in both the nonspatial and the spatial models. The dashed line in (A) represents the Bonferroni-corrected significance level. In (B) all but the parameter estimations for the oak tree density were standardized for graphical purposes.

genotypes, space, and the environment. Mantel correlograms and sPCA analysis suggested that there was some spatial genetic structure within the study population of great tits. Allele frequencies were weakly but positively correlated at distances up to approximately 700 m and weakly negatively correlated at distances greater than 1500 m. This pattern is most likely to result

from restricted natal dispersal of birds that recruit to the breeding population at this site, which have a median natal dispersal distance of 528 m and 788 m for males and females, respectively (Szulkin and Sheldon 2008). sPCA suggested that there were important spatial clines in allele frequencies, the most dominant of which was aligned on a southeast–northwest axis. Accounting for

spatial structure with dbMEMs, we found evidence for fine-scale environmentally induced genetic structure with respect to the risk of infection by two malaria species (*P. circumflexum* and *P. relictum*) and local conspecific density, but not for oak tree density at breeding sites or altitude. SNPs at 21 (0.004% of loci tested), seven (0.001% of loci), and nine (0.002% of loci) loci were putatively of ecological and evolutionary relevance for *P. circumflexum*, *P. relictum*, and local conspecific density, respectively.

Studies relating environmental variables to molecular markers in a spatial context at the population level are becoming increasingly common and have successfully identified a number of putative instances of environmentally induced selection (Eckert et al. 2010; Manel et al. 2010, 2012; Hancock et al. 2011). Divergent selection in sympatry has most often been studied within the context of ecotypes (Schluter 2000; Nosil 2012), perhaps due to their conspicuousness, and is thought to generally be related to selection for resource exploitation leading to reduced gene flow (Schluter 2000; Nosil 2012). Two previous studies using long-term data and pedigree-based quantitative genetic approaches have suggested fine-scale adaptive divergence within great tit populations (Garant et al. 2005; Postma and van Noordwijk 2005). In both cases, it appeared that nonrandom dispersal was key to maintaining structure. Postma and van Noordwijk (2005) found that microgeographic variation in clutch size could be explained by the number and genotypes of immigrants into different sections of their study population, allowing some regions to attain an optimal clutch size where others were swamped by maladaptive genotypes. Garant et al. (2005), also working on the Wytham Woods population, found that fledgling mass differed between ecologically distinct sections of the wood and suggested that this differentiation was maintained by phenotype-dependent dispersal of heavy (putatively high quality) birds into high-quality habitat.

Finding genetic structure in such a small area relative to dispersal capabilities was initially a surprising result given, first, that natal dispersal of individuals within the population would seem likely to reshuffle genotypes in space quite quickly, and second, that approximately half of the breeding birds in any year are immigrants to the population. Immigrants are defined as birds that appear in the breeding population and that are unringed, and are therefore of unknown origin. Although it is possible that a small proportion of these originate from breeding attempts in natural cavities, the majority originate externally to the study population. Because the study site is a block of forest, surrounded by agricultural land and hedgerows, the majority of immigrants have dispersed from natal sites >3 km from the study site (Verhulst et al. 1997). The pattern of fine-scale environmentally induced genetic differentiation found here, despite high levels of gene flow, suggests to us that either differential survival postdispersal, or differential habitat selection in relation to genotypes (cf. Postma and van Noordwijk 2005; Garant et al. 2005), maintain fitness-

related genetic variation within this population. Classically, the more dispersal there is within a system, the shallower a genetic cline is predicted to be (Slatkin 1973; Lenormand 2002). Models of the process of cline formation typically assume that a fixed fraction of individuals disperse each generation. Armsworth and Roughgarden (2008) reexamined the emergence of genetic clines with fitness-dependent models that first allowed individuals to gather information about habitat quality prior to settling and second allowed individuals to base the decision to leave on local habitat quality but with no directional movement bias once the decision to leave is made. They found that fitness-dependent dispersal, in response to environmental conditions, produced steeper and more responsive clines than classical models and perhaps that processes partially explain the strong genetic gradient found at such a fine-scale in our study.

Avian malaria is a globally distributed vector-transmitted disease infecting many bird species, but substantial mortality caused by malaria is generally only reported for newly established host-parasite interactions, such as isolated island populations where the parasite, vector or both are recently introduced (Bennett et al. 1993; LaPointe et al. 2012). However, malaria-induced mortality is most frequently found after initial infection, during the acute phase of the infection, a stage which is brief and very difficult to study in wild birds; hence, the effect of endemic malaria may be underestimated in wild populations. Both *P. circumflexum* and *P. relictum* have a global distribution (Bennett et al. 1993) and are endemic to our study site (Lachish et al. 2011a). Of the two, only *P. circumflexum* is known to increase mortality in tits, based on work on the closely related blue tit, where uninfected birds showed reduced survival in areas of high disease risk (Lachish et al. 2011a); similar results have been obtained for the great tit (I. Sepil et al., unpubl. ms.).

Genetic effects on resistance to malaria in humans provide some of the best-known cases of genetic polymorphisms under selection, with a range of different types of gene (associated with genetic hemoglobin disorders, erythrocyte polymorphisms, enzymopathies and immunogenic variants) having been found to affect resistance to malaria infections (Lopez et al. 2010). Much less is known about the genetics of resistance to malaria in natural nonhuman systems, with the majority of work on passerine birds having investigated the link between major histocompatibility complex (MHC) variation and malaria resistance. There are several indications from such work of genetic effects on resistance to avian malaria (Bonneaud et al. 2006; Westerdahl 2006; Foster et al. 2007; see also I. Sepil et al., unpubl. ms. for evidence of links between MHC class I variants and *P. circumflexum* resistance in this population) and this could lead to selection for particular genotypes under the presence of avian malaria. An interesting difference between the spatial distributions of infections with *P. circumflexum* and *P. relictum* has

been documented for both great and blue tits in our study area. Although the spatial distribution of *P. circumflexum* was very stable over multiple years, the distribution of *P. relictum* varied between years (Lachish et al. 2013). Further, the relative risk of infection with *P. circumflexum* was considerably more variable in space than that of *P. relictum*. Because breeding dispersal is limited in great tits (Greenwood et al. 1979), this would give scope for consistent population stratification caused by *P. circumflexum*, but much less so with *P. relictum*; further, we might expect stronger effects owing to the greater spatial variability in infection risk in *P. circumflexum*. Overall, the fact that the greatest number of SNPs showed spatial associations with this feature of the environment is consistent with this scenario.

Besides associations with avian malaria, both the spatial and nonspatial GEAS found several significant SNPs correlating with local conspecific density. Although this might be the result of correlations between *P. circumflexum*, *P. relictum*, and local conspecific density, we might also expect local conspecific density to play a role in the infection risk of one or both *Plasmodium* spp. Density dependence of *Plasmodium* infection has been described in blue tits (Lachish et al. 2013), and while vector-transmitted diseases are expected to be frequency-dependent rather than density-dependent, it is possible that indirect effects of host density (e.g., physiological stress due to over-crowding) might increase the risk of infection (see Lachish et al. 2013 for more discussion). However, there are a number of other routes by which variable local conspecific density could lead to selection on underlying genetic variants. For example, increased local conspecific density is likely to be associated with increased rates of social interaction, which might select for individuals resilient to social stress. Alternatively, density-dependent life-history decisions (e.g., Wilkin et al. 2006; Wilkin and Sheldon 2009) may select for particular suites of life-history traits in areas of different local conspecific density (e.g., individuals that prioritize offspring quality over number, and future reproduction over current reproduction, at high local conspecific density). Clearly, more work needs to be done to understand the potential selective mechanisms that might generate selection due to local conspecific density.

Although both oak tree density and altitude correlate with fitness and are spatially structured (Wilkin et al. 2009a, b; Wilkin and Sheldon 2009), we did not find any significant correlations with SNPs. This suggests, given that our SNP set sufficiently covers the genome, that no detectable genetic structure was correlated with these environmental features. This could be either because the spatial scale of our study was too small or selection was too weak given high gene flow (Lenormand 2002). It is also possible that other selection patterns interfere: for instance, in the case of territory quality there is scope for sexually antagonistic selection. Because lifetime reproductive success of females depends largely on the quality of the breeding locations, whereas

the lifetime reproductive success of males seems to depend more on their natal locations (Wilkin and Sheldon 2009), sexual antagonistic selection might reduce selection for traits associated with territory quality.

Understanding the mechanisms responsible for maintaining fine-scale genetic structure and driving microevolutionary change within populations is important for evolutionary and ecological theory, conservation, and management. Many theoretical models in ecology and evolution implicitly or explicitly assume that individuals are close enough to interact in predictable ways in space and time. Discontinuities among interacting individuals are often of research interest (i.e., individuals, families, extended social groups, demes, subpopulations, and populations). Empirical work toward understanding how the genetic make-up of these units interacts with the spatial environmental processes leading to the heterogeneity of interest is thus vital. For example, the origin and maintenance of adaptive variation is often considered to be a trade-off between the swamping effects of gene flow and the strength of local selection (Lenormand 2002). Here, we provide evidence that adaptive genetic variation can be maintained despite high levels of gene flow. Evolutionary change over short time scales or at fine spatial scales is rarely investigated; however, a small but growing body of work (Garant et al. 2005; Postma et al. 2005; Tonniss et al. 2005; Porlier et al. 2012) show that it is important to do so. The mechanisms of maintenance, whether due to nonrandom dispersal or fitness- and survival-related costs associated with genotypes, remain to be tested. From a conservation and management perspective, the importance of protecting locally adapted populations has been a long recognized, if elusive, goal (Lande and Barrowclough 1987; Funk et al. 2012). However, increasingly available genomic data will change how important conservation units are delineated (Funk et al. 2012). For example, for great tits there is little genetic differentiation over broad spatial scales (e.g., United Kingdom and The Netherlands: van Bers et al. 2012). However, the potential for local adaptation despite little overall genetic differentiation remains. Although great tits are not in need of management, our results are illustrative of the potential for landscape genetic approaches for understanding local adaptation to contribute importantly to conservation.

ACKNOWLEDGMENTS

We are grateful to the many people who assisted in collection of data and material from the long-term great tit study, particularly to R. Alves, A. Davies, M. Robinson, and S. Lee for laboratory assistance. We thank K. Petren, T. Jombart, and an anonymous referee for comments on the manuscript. We thank C. Blancher and G. Buck at the Wellcome Trust Centre for Human Genetics in Oxford for genotyping. We would like to acknowledge the use of the Oxford Supercomputing Centre (OSC) in carrying out part of this work. The work was in part funded by Natural Environment Research Council grants (NER/A/S/2002/00877 and NE/F005725/1) to BCS, and by an European Research Council (ERC)

Advanced Investigator Award to BCS (AdG 250164) and ERC grant, project Avian-EGG, number 202487, awarded to JS.

LITERATURE CITED

- Amin, N., C. M. van Duijn, and Y. S. Aulchenko. 2007. A genomic background based method for association analysis in related individuals. *PLoS ONE* 2:e1274.
- Armstrong, P. R. and J. E. Roughgarden. 2008. The structure of clines with fitness-dependent dispersal. *Am. Nat.* 172:648–657.
- Aulchenko, Y. S., D. J. de Koning, and C. Haley. 2007a. Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. *Genetics* 177:577–585.
- Aulchenko, Y. S., S. Ripke, A. Isaacs, and C. M. Van Duijn. 2007b. GenABEL: an R library for genome-wide association analysis. *Bioinformatics* 23:1294–1296.
- Bennett, G. F., M. A. Peirce, and R. W. Ashford. 1993. Avian haematocrit: mortality and pathogenicity. *J. Nat. Hist.* 27:993–1001.
- Blanchet, F., P. Legendre, and D. Borcard. 2008. Forward selection of explanatory variables. *Ecology* 89:2623–2655.
- Bohonak, A. 1999. Dispersal, gene flow, and population structure. *Q. Rev. Biol.* 74:21–66.
- Bonneaud, C., J. Perez-Tris, P. Federici, O. Chastel, and G. Sorci. 2006. Major histocompatibility alleles associated with local resistance to malaria in a passerine. *Evolution* 60:383–389.
- Borcard, D., and P. Legendre. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecol. Model.* 153:51–69.
- Borcard, D., F. Gillet, and P. Legendre. 2011. Numerical ecology with R. Springer, London.
- Bourke, A. F. G. 2011. Principles of social evolution. Oxford Univ. Press, Oxford, U.K.
- Bouwhuis, S., R. Choquet, B. C. Sheldon, and S. Verhulst. 2012. The forms and fitness cost of senescence: age-specific recapture, survival, reproduction and reproductive value in a wild bird population. *Am. Nat.* 179:E15–E27.
- Clobert, J., C. M. Perrins, R. H. McCleery, and A. G. Gosler. 1988. Survival rate in the great tit *Parus major* in relation to sex, age, and immigration status. *J. Anim. Ecol.* 57:287–306.
- Coyne, J. A. and H. A. Orr. 2004. Speciation. Sinauer Associates, Sunderland, MA.
- Cushman, S., K. McKelvey, J. Hayden, and M. Schwartz. 2006. Gene flow in complex landscapes: testing multiple hypotheses with causal modeling. *Am. Nat.* 168:486–585.
- Devlin, B., and K. Roeder. 1999. Genomic control for association studies. *Biometrics* 55:997–1004.
- Dray, S., and A. Dufour. 2007. The ade4 package: implementing the duality diagram for ecologists. *J. Stat. Softw.* 22:1–20.
- Dray, S., P. Legendre, and P. Peres-Neto. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecol. Model.* 196:483–493.
- Dray, S., R. Pélissier, P. Couteron, M. J. Fortin, P. Legendre, P. Peres-Neto, P. Bellier, R. Bivand, F. G. Blanchet, M. De Caceres, et al. 2012. Community ecology in the age of multivariate multiscale spatial analysis. *Ecol. Monogr.* 82:257–275.
- Dyer, R. J., J. D. Nason, and R. C. Garrick. 2010. Landscape modeling of gene flow: improved power using conditional genetic distance derived from the topology of population networks. *Mol. Ecol.* 19:3746–3759.
- Eckert, A., A. Bower, S. González-Martínez, J. Węgrzyn, G. Coop, and D. Neale. 2010. Back to nature: ecological genomics of loblolly pine (*Pinus taeda*, Pinaceae). *Mol. Ecol.* 19:3789–4594.
- Epperson, B. K. 2005. Estimating dispersal from short distance spatial autocorrelation. *Heredity* 95:7–15.
- Fortin, M. J., and M. R. T. Dale. 2005. Spatial analysis: a guide for ecologists. Cambridge Univ. Press, Cambridge, U.K.
- Foster, J. T., B. L. Woodworth, L. E. Eggert, P. J. Hart, D. Palmer, D. C. Duffy, and R. C. Fleischer. 2007. Genetic structure and evolved malaria resistance in Hawaiian honeycreepers. *Mol. Ecol.* 16:4738–4746.
- Funk, W., J. McKay, P. Hohenlohe, and F. Allendorf. 2012. Harnessing genomics for delineating conservation units. *Trends Ecol. Evol.* 27:489–585.
- Garant, D., L. E. B. Kruuk, R. H. McCleery and B. C. Sheldon. 2004. Evolution in a changing environment: a case study with great tit fledging mass. *Am. Nat.* 164:E115–E129.
- Garant, D., L. E. B. Kruuk, T. Wilkin, R. McCleery, and B. C. Sheldon. 2005. Evolution driven by differential dispersal within a wild bird population. *Nature* 433:60–65.
- Garraway, C. J., J. Bowman, and P. J. Wilson. 2011. Using a genetic network to parameterize a landscape resistance surface for fishers, *Martes pennanti*. *Mol. Ecol.* 20:3978–3988.
- Gaston, K. J. 2003. The structure and dynamics of geographic ranges. Oxford Univ. Press, Oxford, U.K.
- Gosler, A. 1993. The Great Tit. Hamlyn, London.
- Greenwood, P. J., P. H. Harvey, and C. M. Perrins. 1979. Role of dispersal in the great tit (*Parus Major*)—causes, consequences and heritability of natal dispersal. *J. Anim. Ecol.* 48:123–142.
- Hancock, A. M., D. B. Witonsky, G. Alkorta-Aranburu, C. M. Beall, A. Gebremedhin, R. Sukernik, G. Utermann, J. K. Pritchard, G. Coop, and A. Di Rienzo. 2011. Adaptations to climate-mediated selective pressures in humans. *PLoS Genet.* 7:e1001375.
- Hargeby, A., J. Johansson, and J. Ahnesjö. 2004. Habitat-specific pigmentation in a freshwater isopod: adaptive evolution over a small spatiotemporal scale. *Evolution* 58:81–175.
- Harvey, P. H., P. J. Greenwood, and C. M. Perrins. 1979. Breeding area fidelity of great tits (*Parus major*). *J. Anim. Ecol.* 48:703–711.
- Jombart, T. 2008. Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405.
- Jombart, T., S. Devillard, A. B. Dufour, and D. Pontier. 2008. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity* 101:92–103.
- Knowles, S. C. L., V. Palinauskas, and B. C. Sheldon. 2010a. Chronic malaria infections increase family inequalities and reduce parental fitness: experimental evidence from a wild bird population. *J. Evol. Biol.* 23:557–569.
- Knowles, S. C. L., M. J. Wood, and B. C. Sheldon. 2010b. Context-dependent effects of parental effort on malaria infection in a wild bird population, and their role in reproductive trade-offs. *Oecologia* 164:87–97.
- Knowles, S. C., M. J. Wood, R. Alves, T. A. Wilkin, S. Bensch, and B. C. Sheldon. 2011. Molecular epidemiology of malaria prevalence and parasitaemia in a wild bird population. *Mol. Ecol.* 20:1062–1076.
- Lachish, S., S. C. Knowles, R. Alves, M. J. Wood, and B. C. Sheldon. 2011a. Fitness effects of endemic malaria infections in a wild bird population: the importance of ecological structure. *J. Anim. Ecol.* 80:1196–1206.
- . 2011b. Infection dynamics of endemic malaria in a wild bird population: parasite species-dependent drivers of spatial and temporal variation in transmission rates. *J. Anim. Ecol.* 80:1207–1216.
- Lachish, S., S. C. Knowles, R. Alves, I. Sepil, A. Davies, S. Lee, M. J. Wood, and B. C. Sheldon. 2013. Spatial determinants of infection risk in a

- multi-species avian malaria system. *Ecography. In press*. doi: 10.1111/j.1600-0587.2012.07801.x.
- Lack, D. 1964. A long-term study of the great tit (*Parus major*). *J. Anim. Ecol.* 33:159–173.
- Lande, R., and G. F. Barrowclough. 1987. Effective population size, genetic variation, and their use in population management. Pp. 87–124 in M. E. Soulé, ed. *Viable populations for conservation*. Cambridge Univ. Press, Cambridge, U.K.
- LaPointe, D. A., C. T. Atkinson, and M. D. Samuel. 2012. Ecology and conservation biology of avian malaria. *Ann. N. Y. Acad. Sci.* 1249:211–226.
- Lasky, J. R., D. L. Des Marais, J. K. MC KAY, J. H. Richards, T. E. Juenger, and T. H. Keitt. 2012. Characterizing genomic variation of *Arabidopsis thaliana*: the roles of geography and climate. *Mol. Ecol.* 21:5512–5529.
- Legendre, P., and L. Legendre. 2012. *Numerical ecology*. Elsevier, Amsterdam, The Netherlands.
- Legendre, P., D. Borcard, F. Blanchet, and S. Dray. 2012. PCNM: MEM spatial eigenfunction and principal coordinate analyses. Available via http://r-forge.r-project.org/R/?group_id=195.
- Lenormand, T. 2002. Gene flow and the limits to natural selection. *Trends Ecol. Evol.* 17:183–372.
- Lopez, C., C. Saravia, A. Gomez, J. Hoebeke, and M. A. Patarroyo. 2010. Mechanisms of genetically-based resistance to malaria. *Gene* 467:1–12.
- Malecot, G. 1955. Remarks on the decrease of relationship with distance. Following paper by M. Kimura. *Cold Spring Harb. Symp.* 20:52–53.
- Manel, S., B. Poncet, P. Legendre, F. Gugerli, and R. Holderegger. 2010. Common factors drive adaptive genetic variation at different spatial scales in *Arabis alpina*. *Mol. Ecol.* 19:3824–3859.
- Manel, S., F. Gugerli, W. Thuiller, N. Alvarez, P. Legendre, R. Holderegger, L. Gielly, P. Taberlet, and I. Consortium. 2012. Broad-scale adaptive genetic variation in alpine plants is driven by temperature and precipitation. *Mol. Ecol.* 21:3729–3738.
- Mank, J. E., K. Nam, and H. Ellegren. 2010. Faster-Z evolution is predominantly due to genetic drift. *Mol. Biol. Evol.* 27:661–670.
- McIntire, E., and A. Fajardo. 2009. Beyond description: the active and effective way to infer processes from spatial patterns. *Ecology* 90:46–56.
- Moran, P. A. 1948. Rank correlation and product-moment correlation. *Biometrika* 35:203–206.
- . 1950. Notes on continuous stochastic phenomena. *Biometrika* 37: 17–23.
- Murphy, M. A., R. Dezzani, D. S. Pilliod, and A. Storfer. 2010. Landscape genetics of high mountain frog metapopulations. *Mol. Ecol.* 19:3634–3649.
- Nosil, P. 2012. *Ecological speciation*. Oxford Univ. Press, Oxford, U.K.
- Nosil, P., T. Vines, and D. Funk. 2005. Perspective: reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution* 59:705–724.
- Peres-Neto, P. R., and P. Legendre. 2010. Estimating and controlling for spatial structure in the study of ecological communities. *Global Ecol. Biogeogr.* 19:174–184.
- Porlier, M., D. Garant, P. Perret, and A. Charmantier. 2012. Habitat-linked population genetic differentiation in the blue tit *Cyanistes caeruleus*. *J. Hered.* 103:781–791.
- Postma, E., and A. van Noordwijk. 2005. Gene flow maintains a large genetic difference in clutch size at a small spatial scale. *Nature* 433:65–73.
- Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. de Bakker, M. J. Daly, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81:559–575.
- Santure, A. W., J. Gratten, J. A. Mossman, B. C. Sheldon, and J. Slate. 2011. Characterisation of the transcriptome of a wild great tit *Parus major* population by next generation sequencing. *BMC Genomics* 12:283.
- Schluter, D. 2000. *The ecology of adaptive radiation*. Oxford Univ. Press, Oxford, U.K.
- Slatkin, M. 1973. Gene flow and selection in a cline. *Genetics* 75:733–756.
- Sokal, R. R., and N. L. Oden. 1978. Spatial Autocorrelation in biology. 1. Methodology. *Biol. J. Linn. Soc.* 10:199–228.
- Szulkin, M., and B. C. Sheldon. 2008. Dispersal as a means of inbreeding avoidance in a wild bird population. *Proc. R. Soc. B* 275: 703–711.
- The R Core Team. 2012. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Available via <http://www.r-project.org/>
- Thompson, E. A., and R. G. Shaw. 1990. Pedigree analysis for quantitative traits—variance-components without matrix-inversion. *Biometrics* 46:399–413.
- Tonniss, B., P. R. Grant, B. R. Grant, and K. Petren. 2005. Habitat selection and ecological speciation in Galápagos warbler finches (*Certhidea olivacea* and *Certhidea fusca*). *Proc. R. Soc. Lond. B* 272:819–826.
- van Bers, N. E. M., K. van Oers, H. H. D. Kerstens, B. W. Dibbitts, R. P. M. A. Crooijmans, M. E. Visser, and M. A. M. Groenen. 2010. Genome-wide SNP detection in the great tit *Parus major* using high throughput sequencing. *Mol. Ecol.* 19:89–99.
- van Bers, N. E., A. W. Santure, K. van Oers, I. De Cauwer, B. W. Dibbitts, C. Mateman, R. P. Crooijmans, B. C. Sheldon, M. E. Visser, M. A. Groenen, et al. 2012. The design and cross-population application of a genome-wide SNP chip for the great tit *Parus major*. *Mol. Ecol. Res.* 12:753–770.
- van Strien, M. J., D. Keller, and R. Holderegger. 2012. A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. *Mol. Ecol.* 21:4010–4023.
- Verhulst, S., C. M. Perrins, and R. Riddington. 1997. Natal dispersal of great tits in a patchy environment. *Ecology* 78:864–872.
- Westerdahl, H. 2006. Passerine MHC; genetic variation and disease resistance in the wild. *J. Ornithol.* 147:48–48.
- Wilkin, T. A., and B. C. Sheldon. 2009. Sex differences in the persistence of natal environmental effects on life histories. *Curr. Biol.* 19: 1998–2002.
- Wilkin, T. A., D. Garant, A. G. Gosler, and B. C. Sheldon. 2006. Density effects on life-history traits in a wild population of the great tit *Parus major*: analyses of long-term data with GIS techniques. *J. Anim. Ecol.* 75:604–615.
- Wilkin, T. A., C. M. Perrins, and B. C. Sheldon. 2007. The use of GIS in estimating spatial variation in habitat quality: a case study of lay-date in the Great Tit *Parus major*. *Ibis* 149:110–118.
- Wilkin, T. A., L. E. King, and B. C. Sheldon. 2009. Habitat quality, nestling diet, and provisioning behaviour in great tits *Parus major*. *J. Avian Biol.* 40:135–145.
- Wood, M. J., C. L. Cosgrove, T. A. Wilkin, S. C. Knowles, K. P. Day, and B. C. Sheldon. 2007. Within-population variation in prevalence and lineage distribution of avian malaria in blue tits, *Cyanistes caeruleus*. *Mol. Ecol.* 16:3263–3273.

Associate Editor: K. Petren

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Figure S1. Comparison of different methods to calculate significance of correlations between single-nucleotide polymorphisms (SNPs) and various environmental traits for great tits breeding in Wytham Woods, Oxford, U.K.

Figure S2. A spatial principal components analysis (sPCA) screeplot (Jombart et al. 2008) displaying the decomposition of eigenvalues into variance and a spatial autocorrelation (Moran's I) components.

Figure S3. Locations of the significant single-nucleotide polymorphisms (SNPs) for malaria infection risk and local conspecific density on the Great tit genome.

Figure S4. Correlations between (a, c) the parameter estimates of the spatial genomic regressions and (b, d) the environmental trait values themselves.