Population genetic structure of serotine bats (*Eptesicus serotinus*) across Europe and implications for the potential spread of bat rabies (European Bat Lyssavirus EBLV-1).

Moussy C¹,², Atterby H², Griffiths AGF³, Allnut TR²,³,⁴, Mathews F⁵, Smith GC⁶, Aegerter JN⁶, Bearhop S¹, Hosken DJ¹

¹ Centre for Ecology and Conservation, University of Exeter, Cornwall Campus, Penryn, Cornwall, TR10 9EZ, UK
² Food and Environment Research Agency, Sand Hutton, York, YO41 1LZ, UK
³ Environment and Sustainability Institute, University of Exeter, Cornwall Campus, Penryn, Cornwall, TR10 9EZ, UK
⁴ CAFHS, CSIRO, Sneydes road, Werribee, Victoria, 3030, Victoria, Australia
⁵ Hatherly Laboratories, College of Life and Environmental Sciences, University of Exeter, Prince of Wales Road, Exeter, EX4 4PS, UK
⁶ National Wildlife Management Centre, Animal Health and Veterinary Laboratories Agency (AHVLA), Sand Hutton, York, YO41 1LZ, UK
⁷ Now at: Centre d’Ecologie et des Sciences de la Conservation (CESCO UMR7204), Sorbonne Universités, MNHN, CNRS, UPMC, CP51, 55 rue Buffon, 75005, Paris, France

Corresponding author: DJ Hosken, d.j.hosken@exeter.ac.uk; +44 (0)1326 371843

Running title: Genetic structure of *Eptesicus serotinus*

Word count: 6847
Abstract

Understanding of the movements of species at multiple scales is essential to appreciate patterns of population connectivity and in some cases, the potential for pathogen transmission. The serotine bat (*Eptesicus serotinus*) is a common and widely distributed species in Europe where it frequently harbours European Bat Lyssavirus Type 1 (EBLV-1), a virus causing rabies and transmissible to humans. In the UK it is rare, with a distribution restricted to south of the country and so far the virus has never been found. We investigated the genetic structure and gene flow of *E. serotinus* across the England and continental Europe. Greater genetic structuring was found in England compared with continental Europe. Nuclear data suggest a single population on the continent, although further work with more intensive sampling is required to confirm this, while mitochondrial sequences indicate an east-west substructure. In contrast, three distinct populations were found in England using microsatellite markers, and mitochondrial diversity was very low. Evidence of nuclear admixture indicated strong male mediated gene flow among populations. Differences in connectivity could contribute to the high viral prevalence on the continent in contrast with the UK. While the English Channel was previously thought to restrict gene flow, our data indicate relatively frequent movement from the continent to England highlighting the potential for movement of EBLV-1 into the UK.

Key words: Chiroptera; gene flow; genetic structure; philopatry; dispersal; rabies
Introduction

Bats are widely distributed and vary enormously in their ecology, sociality and behaviour. Largely because of their small size and vagility, population genetics has been widely used to infer aspects of bat biology that are difficult to observe directly. Substantial variation in genetic structuring has been found, reflecting differences in bat movement, behaviour, mating and social systems (Burland and Wilmer, 2001; Moussy et al., 2013). Additionally, molecular approaches have been used to assess many aspects of bat ecology including potential sex-biased dispersal and movements (Petit et al., 2001), mating systems (Furmankiewicz and Altringham, 2007), effect of landscape barriers (García-Mudarra et al., 2009), demography and history (Flanders et al., 2009) and sociality (Metheny et al., 2007). Gaining an understanding of these components of bat ecology and behaviour is fundamental in a conservation context (Racey and Entwistle, 2003), to improve our understanding of ecosystem services (Kunz et al., 2011) and in disease epidemiology since bats are now thought to be the wildlife reservoirs for a range of important zoonoses and emerging diseases (Kuzmin et al., 2011).

The serotine bat, *Eptesicus serotinus*, is a common species within its core range in continental Europe, but rare in the UK where it is largely restricted to southern England (Dietz et al., 2009). The species is considered sedentary in the UK with recorded seasonal flights of under 100 kilometres, although ringing data is sparse, which restricts our understanding of any seasonal movements (Hutterer et al., 2005). Adult females form maternity colonies in the summer, often in inhabited
buildings, and display some level of philopatry to their natal landscape, if not natal colony (Harbusch and Racey, 2006). Due to the relative ease of locating these maternity roosts, research has mainly focused on habitat and roost use, diet, and activity patterns of adult females or of juveniles of both sexes (Catto et al., 1996; Robinson and Stebbings, 1997). In contrast, adult males are more difficult to find and their ecology is very poorly understood. Similarly, little is known about the winter ecology of the species and its mating system. Information on the social and spatial organisation of this species, as well as on population and seasonal connectivity is required for a more complete understanding of its ecology and behaviour within its core range and at the edge of its distribution.

This basic information on serotine biology is particularly important as this bat is the main reservoir for the most common European bat lyssavirus, EBLV-1 (Harris et al., 2006). In such a context, knowledge of the host’s movement, mating patterns and social organisation is crucial to gain further understanding of viral dynamics. Currently the prevalence of EBLV-1 in continental Europe contrasts strongly with its apparent absence from England (Schatz et al. 2013). Indeed, both active and passive surveillance in the UK has failed to detect the virus. However, passive surveillance is biased to common species with only few E. serotinus submitted for testing. Furthermore, a juvenile bat was tested positive for antibodies against EBLV-1 in a maternity roost of south of England (Harris et al. 2009). This case was surprising since the bat hence had not dispersed or mated, and was therefore likely to have only been in contact with roost-mates, but no
other individuals from the same roost was found seropositive despite extensive sampling over four years.

Here, we used ten microsatellites and a portion of the hypervariable region II (HVII) of mitochondrial DNA (mtDNA) to investigate the genetic structure of *E. serotinus* at multiple spatial scales. We initially describe in detail the genetic structure of serotines in England and provide a first account of the level of genetic differentiation on the continent to compare the core and edge populations. Based on the bat’s biology, two contrasting predictions were made. First, because the English population is small and fragmented at the edge of the species range (Bat Conservation Trust 2013), movement and gene flow could be reduced there, resulting in genetic differentiation and structure. If the core continental population were also large and continuous enough to result in low structure, then these contrasting structure patterns could help explain the prevalence of EBLV-1 on the continent and its apparent absence in England – bats simply do not move enough in the UK to spread the virus. Alternatively, because the geographic area covered by the English population is small and there are few obvious barriers to movement within the UK, it could display low structure and low genetic differentiation. If this were so, then the apparent absence of EBLV-1 in the UK could simply be due limited dispersal across the English Channel. We also tested for possible sex-biased gene flow as male connectivity and female philopatry are common in temperate bats (Moussy et al., 2013).
**Methods**

**Sample collection**

A total of 593 individuals were captured at 28 maternity roosts in inhabited buildings across the English range in the summers of 2004-2006 and 2010-2011 (individuals per roosts: mean = 19.41; median = 14.50; range from 4 to 68; Table S1; Figure 1). No adult males were captured due to their absence from maternity roosts. A wing tissue sample was taken using a 3 mm biopsy punch (Stiefel Laboratories, Wooburn Green, UK) and stored in 70% ethanol at +4°C until DNA extraction. All bats were released back into the wild within a few minutes of the procedure at their location of capture. All work was done with the approval of ethical review bodies at FERA and University of Exeter and under license from the UK Home Office [Animals (Scientific Procedures) Act 1986] and Natural England (Habitats Regulations, 2010). Wing biopsies were also supplied for 106 individuals from eight locations in six countries across continental Europe (Table S1; Figure 1).

**DNA extraction and Whole Genome Amplification**

DNA was extracted following Sambrook and Russel (2001). Ammonium acetate (10 M; Sigma-Aldrich, Poole, UK) was used to precipitate proteins instead of potassium acetate. The DNA pellet was recovered in 1 x Tris–ethylenediaminetetraacetic acid buffer (TE) (Sigma-Aldrich) and stored at -20°C before further procedures. The whole genome of each individual was amplified using illustra GenomiPhi V2 DNA Amplification kit (GE Healthcare, Little Chalfont, UK). A subset of samples was genotyped for microsatellites before and after
whole genome amplification to ensure the procedure did not affect typing success or quality.

*Microsatellite genotyping and mitochondrial DNA sequencing*

Samples were genotyped using a panel of 11 microsatellite markers (Smith *et al.*, 2011) as described in Table S2. PCR was carried out in 15 μl, with 7.5 μl 2x PCR mastermix (ABgene, Epsom, UK) containing 1.5 mM MgCl₂, 1 μM of each primer (Sigma-Aldrich/Applied Biosystems, Foster City, CA, USA), and 1.5 μl DNA template (DNA extract or 1/10 dilution in 1 x TE of whole genome amplification product). The remaining volume was made up with molecular grade water. PCR reactions consisted of denaturation at 94°C for 2 min, followed by 35 cycles of 94°C for 1 min, 1 min at the annealing temperature (Table S2), 72°C for 1 min, and final elongation at 72°C for 5 min and 60°C for 1 h. PCR products were diluted and mixed into four sets (Table S2) and run on an ABI Prism 3130xl genetic analyser (Applied Biosystems) with Genescan Rox 500 size standard (Applied Biosystems). Microsatellite alleles were sized using GeneMapper 3.7 software (Applied Biosystems). The genotyping error rate per loci was estimated by re-amplifying and re-genotyping 11–24% of samples depending on the locus.

A 460 bp portion of the hypervariable domain II (HVII) of the mtDNA control region was amplified in a subset of samples (Table S1) using primers for conserved sequence block F on L-strand 5’-CTACCTCCGTGAAACCAGCAAC-3’ (Wilkinson and Chapman 1991) and for HVII on H-strand 5’-CGTACACGTATTGTATGTGTCT-3’ (J. Juste, pers. comm.). PCR was carried out in 20 μl, with 10 μl 2x BioMix Red.
(Bioline, London, UK) containing 1 mM MgCl₂, 0.5 μM of each primer (Sigma-Aldrich), and 1 μl of amplified genomic DNA. The remaining volume was made up with molecular grade water. PCR reactions consisted of denaturation at 94°C for 5 min, followed by 30 cycles of 94°C for 1 min, 55°C for 1.5 min, 72°C for 2 min, and final elongation at 72°C for 5 min and 60°C for 1 h. PCR products were purified by an enzymatic reaction to remove leftover primers and dNTPs. A 10 μl mix of 0.5 U Exonuclease I (Fermentas, Vilnius, Lithuania) and 0.25 U Alkaline Phosphatase (Fermentas) in molecular grade water was added to each product before incubation at 37°C for 30 min and deactivation at 85°C for 15 min. The purified DNA fragment was sequenced from both directions using the BigDye Terminator v3.1 sequencing kit (Applied Biosystems) in a 20 μl reaction volume with 1.5 μl BigDye Terminator mix, 3.5 μl 5x sequencing buffer, 3.33 pmol of primer and 1 μl DNA template. The remaining volume was made up with molecular grade water. The sequencing reaction consisted of 25 cycles of 96°C for 10 s, 50°C for 5 s, and 60°C for 4 min. The sequencing reactions were cleaned by precipitation with 100% ethanol and 20 μg glycogen before a washing step with 70% ethanol. They were then reconstituted in 11 μl HiDi formamide before being run on an ABI Prism 3130xl genetic analyser (Applied Biosystems). The resulting sequences were assembled, aligned and trimmed into GENEIOUS 5.5.7 (Drummond et al., 2011) to create a 424 bp consensus sequence for each individual.

Microsatellites analysis: Tests of assumptions, genetic diversity and basic statistics
The presence of scoring inconsistencies, null alleles, large allele dropout and stuttering were tested in MICROCHECKER (Van Oosterhout et al., 2004). Tests for departure from Hardy-Weinberg (HW) equilibrium were performed for each roost and each locus in GENEPOP 4.1.4 (Rousset, 2008) both as two-tailed tests with no assumption as the direction of the deviation, and under the hypothesis of heterozygote deficit. A sequential Bonferonni correction was applied to the p-values (Rice, 1989). Tests of linkage disequilibrium were performed in FSTAT 2.9.3.2 (Goudet, 1995, 2001) and the significance levels were adjusted using sequential Bonferonni correction. Genetic diversity indices were obtained for each roost from FSTAT and GENETIX 4.05 (Belkhir et al. 1996-2005). Estimated $F_{ST}$ averaged over loci (Weir and Cockerham, 1984) and the 95% confidence intervals (CI) after 1000 bootstrap over loci were calculated in GENETIX. Finally, a log-likelihood G-test for population differentiation (not assuming random mating within roosts) was performed (Goudet et al., 1996) in FSTAT based on 100,000 randomisations of complete multilocus genotypes.

**Nuclear population structure**

The Bayesian clustering program STRUCTURE was used to investigate genetic structure (Pritchard et al., 2000). We evaluated the number of assumed clusters K from 1 to 6 using a burn-in of $5 \times 10^6$ iterations followed by $5 \times 10^7$ MCMC iterations. The admixture model with correlated allele frequencies was run with and without population information (roosts the individuals were sampled from) as a prior. Each K value was run 12 times to ensure stability and convergence of the
chains. The optimal number of clusters K was obtained from ΔK, based on the rate of change in the log probability of data in successive K values (Evanno et al., 2005) as implemented on STRUCTURE HARVESTER (Earl and VonHoldt, 2012).

Genetic structuring was further investigated using two multivariate methods: (1) a discriminant analysis of principal components (DAPC) (Jombart et al., 2010) and (2) a spatial principal component analysis (sPCA) (Jombart et al., 2008). In contrast to Bayesian clustering, the strength of these approaches rests in their independence from population genetic model, and inferences are therefore made only on allelic similarity (Jombart et al., 2009). The two methods were implemented in the ADEGENET 1.3-4 package (Jombart, 2008) in R. DAPC has been developed to summarise the overall genetic variability of individuals within groups while optimising discrimination between groups. We used the roosts individuals were sampled from as group prior to discriminant analysis. The first 30 principal components (PC) of PCA were retained in the data transformation step, covering 90.5% of genetic variance. sPCA explicitly incorporates spatial information in the investigation of genetic variability across landscapes (Jombart et al., 2008). We applied this method at the roost level. A connection network was constructed using K-nearest neighbour (Cover and Hart, 1967), setting K to seven to reflect the number of sites sampled on the continent. This network was used for the calculation of Moran’s I (Moran, 1948, 1950). sPCA optimises the product of the variance of individual scores, based on roosts’ allelic frequencies, and of Moran’s I, in order to summarise genetic variability in a spatial context.
AMOVA and genetic differentiation among clusters

Consensus on the population structure was reached based on results from the Bayesian and multivariate analyses. Roosts were assigned to one of the clusters defined (hereafter called populations) and this genetic structure was tested in an analysis of molecular variance (AMOVA) (Excoffier et al., 1992) in ARLEQUIN. Log-likelihood G tests for differentiation were performed between and within populations using 10,000 permutations in the HIERFSTAT package implemented in R. Pairwise $F_{ST}$ between the populations were obtained from ARLEQUIN to assess the level of connectivity. Log-likelihood G tests were performed for all pairs of populations on 10,000 permutations by the HIERFSTAT package.

Isolation by distance

Mantel and partial Mantel tests were performed between matrices of linearised pairwise $F_{ST}$ (Rousset, 1997) and log-transformed geographical distances among roosts. An indicator matrix was included to account for the presence of the most obvious geographical barriers including the English Channel (isolating the UK), the Solent (between the Isle of Wight and England), the Alps (isolating the Italian roost), the Carpathians (isolating the Slovakian roost) and the Caucasus range (isolating Georgia). Tests were carried out on the full dataset, the continent only dataset and on the English-only dataset.

Contemporary gene flow
The magnitude and direction of contemporary gene flow occurring between the consensus populations were estimated using the program BAYESASS 3.0.1 (Wilson and Rannala, 2003). The simulation was run with $10^7$ iterations, with the first $25 \times 10^5$ iterations discarded as burn-in. Samples were collected every 100 iterations. Allelic frequencies and inbreeding coefficients were set at 0.2 while migration rate was left at its default value of 0.1. These mixing parameters ensured the acceptance rates were between 0.2 and 0.4. The trace file was examined in TRACER 1.5 to ensure mixing and convergence of the chains (Rambaut and Drummond, 2009).

**Mitochondrial DNA analysis: Genetic diversity and differentiation**

Haplotype ($h$) and nucleotide diversities ($\pi$), number of polymorphic sites and mean number of pairwise differences were calculated for each roost and each population in ARLEQUIN. Differentiation index for haploid data $\phi_{ST}$ (Weir and Cockerham, 1984) was estimated in ARLEQUIN and tested for significant departure from panmixia after 10,000 permutations. Differentiation test among roosts was also performed with 100,000 MC iterations after 10,000 burn-in iterations in ARLEQUIN.

**Mitochondrial population structure**

Spatial structuring of genetic variation in the mtDNA sequences was investigated by sPCA (Jombart et al., 2008). All roosts were considered as neighbours in this
analysis, but their inter-connection was weighted by the inverse Euclidian
distance separating the roosts.

Geographical structuring of mtDNA variation was also investigated in an AMOVA.
Roosts were grouped in various clusters based on the sPCA results and on the
structure observed at the nuclear level. Those structures were tested in order to
optimise the among-groups variance component. The contribution of each
hierarchical component (among groups, among roosts within groups and within
roosts) was estimated and tested for significance after 10,000 permutations.
Differentiation between identified clusters was inferred by pairwise \( \phi_{ST} \) and
tested by log-likelihood tests with 100,000 MC iterations after 10,000 burn-in
iterations. All calculations were undertaken in ARLEQUIN.

Mantel and partial Mantel tests were performed between matrices of pairwise
linearised \( \phi_{ST} \) and log-transformed geographical distances among roosts to test
for isolation by distance. The same indicator matrix used to infer presence of
geographical barriers in the microsatellite dataset was included. Tests were
carried out on the full dataset, the continent only dataset and on the English
dataset.

**Phylogenetic analysis**

Phylogenetic reconstructions were undertaken on unique haplotypes by
maximum-likelihood (ML) criteria (Beerli and Felsenstein, 2001) and a Bayesian
approach. The K80+I model (Kimura, 1980) of DNA substitution
(transition/transversion ratio = 18.374, proportion of invariable sites = 0.817) was applied to the ML and the Bayesian analyses as determined by the program jMODELTEST 0.1.1 (Wu and Eisen, 2008; Posada, 2008) using Bayesian Information Criterion (BIC). The ML analysis was undertaken in PHYML 3.0 (Wu and Eisen, 2008) and the Bayesian inference in MRBAYES 3.1 (Ronquist and Huelsenbeck, 2003). For the Bayesian analysis, four chains of $10^7$ iterations each were run, taking samples every 100 generations and the first 25% of trees were discarded as burn-in. The statistical support of the branching pattern obtained by ML approach was assessed by 1,000 bootstrap replicates. In addition, a haplotype network (Excoffier and Smouse, 1994) was constructed using ARLEQUIN and visualised in HAPSTAR (Teacher and Griffiths, 2011). Demographic history was inferred from Fu’s $F_S$, a measure of marker neutrality sensitive to changes in population sizes, and mismatch distributions. Populations that have undergone recent expansion display significantly negative $F_S$ value (Fu and Li, 1993) and a smooth or unimodal distribution of pairwise differences among haplotypes (Schneider and Excoffier, 1999). These analyses were run in ARLEQUIN with 10,000 bootstrap replicates to generate the simulated mismatch distribution in a population expansion model and 10,000 randomised samples to test for $F_S$ significance.

**Results**

**Microsatellites**

**Tests of assumption, genetic diversity and basic statistics**
Of the 11 microsatellites genotyped, one (AF141650) demonstrated high level of estimated null alleles, genotyping inconsistencies and significant deviation from the assumption of HW equilibrium in a large number of sampled roosts. It was thus removed from subsequent analyses. Of the remaining loci, three deviated from HW equilibrium in one to four roosts out of 36 and six showed low to moderate levels of estimated null alleles. Deviation from HW expectations was due to heterozygote deficiency and since it was noticed in only a few roosts, it could be caused by factors other than by selection on the markers. No significant linkage disequilibrium between loci was detected after sequential Bonferonni correction. The ten remaining microsatellites were thus used for subsequent analyses. The mean genotyping error rate for the ten markers was 4.21% (± 2.50%). Mean observed and expected heterozygosity and within roost gene diversity were moderate and consistent across all roosts (H_o = 0.540 ± 0.06; H_e = 0.572 ± 0.04; H_s = 0.604 ± 0.04; Table S3), and allelic richness was low but still similar across all roosts (A_r = 2.9 ± 0.1; Table S3). The continent had significantly higher allelic richness than England (p < 0.001), but none of the other genetic diversity measures differed significantly between the two (H_o: p = 0.594; H_e : p = 0.196, H_s: p = 0.127, F_{IS}: p = 0.721)

Estimated F_{ST} averaged over loci was low at 0.048 but significant (0.039 – 0.059 95% confidence interval). Roosts were found to be significantly genetically differentiated (randomization test: p<0.001), which in combination with the significant structure, suggests the roosts sampled were not part of a single panmictic population.
Bayesian modelling

Despite high levels of admixture, clustering was consistent with the sampling geography. At \( K = 2 \), the sites on the continent and eastern England formed one group and those in western England and Isle of Wight (IOW) another. At \( K = 3 \), the continent was discriminated from eastern England. At \( K = 4 \) and \( 5 \), the IOW was distinct (Figure 2 and S1). The standardised second order rate of change \( \Delta K \) indicated that the most likely number of clusters describing the data was three (Figure S1). Because this method only distinguished higher structuring level, each cluster was then run separately with the same settings. No obvious geographical sub-structure was supported for the continent or for eastern England. In contrast, sub-structure was revealed in the third cluster, with roosts on the IOW segregating from western England.

Multivariate analysis

The discrimination between roosts using DAPC was low, reflecting high within roost genetic variability and high admixture levels (91.72% of individuals have membership probabilities to the sampling roost lower than 90%, Jombart 2011). Despite this admixture, the first two PCs captured 42% of the genetic structure. PC-1 mostly delineated roosts from the continent and the IOW from those on mainland England, whilst PC-2 discriminated roosts from west of England and the IOW from those on the continent and in east of England, a pattern of clustering similar to the STRUCTURE analysis (Figure 2).
Population structure was also inferred by sPCA. The three first axes had the highest eigenvalues and were therefore retained. PC-1 mostly segregated England from the continent. PC-2 differentiated eastern England from western England and western continental sites from Eastern Europe. The continental sPCA scores were low compared to English ones, suggesting that the continental structuring was weak compared to that across England. PC-3 retrieved a final structure by clustering sites in France with the IOW and Central England (Figure S2). Thus, the UK was clearly differentiated from the continent and subdivided into three groups representing the east, the west and the IOW (Figure 3).

AMOVA and genetic differentiation among clusters

Four populations were retained through consensus from the previous analyses - continental Europe, eastern England, western England and the IOW (Table S2). Differentiation tests were significant both between and within populations ($F_{CT} = 0.034$, $p = 0.0001$; $F_{SC} = 0.025$, $p = 0.0001$), suggesting both the stratification of the genetic variation and the importance of roosts as structuring units within those clusters. Pairwise $F_{ST}$ suggested a closer relationship between the continent and eastern England than with the other clusters, a very close relationship between eastern and western England, and a stronger connection between the IOW and western England than to the other clusters. The IOW was also more connected to the continent than to eastern England (Table 1). Differentiation was significant for all pairwise connections with the exception of east-west England ($p = 0.510$) and continent-IOW ($p = 0.652$).
Isolation by distance

Genetic and geographical distances were significantly correlated when controlling for the effects of physical barriers on the full dataset ($r = 0.214$, $p = 0.021$), on the continental dataset ($r = 0.463$, $p = 0.024$) and on the English dataset ($r = 0.302$, $p < 0.001$) (Figure S3). The physical barriers tested also had a significant effect on genetic differentiation when controlling for geographical distance at all scales (Europe: $r = 0.373$, $p < 0.001$; Continent: $r = 0.511$, $p = 0.024$; England: $r = 0.560$, $p < 0.001$).

Contemporary gene flow

Roosts were pooled into the four populations retained from previous analyses, and recent migration rates were estimated as the mean proportion of individuals moving between populations (Table 2). The continental population displayed the highest emigration rate and the lowest immigration rate, indicating that contemporary gene flow over the Channel is mostly biased toward England. Interestingly, the IOW seems to play an important role in mediating gene flow with immigration coming mostly from the continent (5.16%). Within England, most gene flow was to the west (from east 4.55% and from IOW 3.66%), whereas there was little flow to the east or to the IOW (Table 2). It is however worth noting that England has been more intensively sampled than the continent and the full extent of genetic diversity on mainland Europe has not been sampled. As a result, some finer sub-structuring might not have been recovered. This could affect the estimation of gene flow between the UK and mainland Europe and the
values quoted are therefore to be taken as a relative indication of the magnitude of gene movement. Furthermore, since most continental samples are from central Europe rather than bordering the English Channel, gene flow is likely to be underestimated.

*Mitochondrial genetic diversity and differentiation*

A total of 26 haplotypes were identified among the 441 samples from 35 roosts. Thirty polymorphic sites (7.79%) were recorded with a transition/transversion ratio of 18.37. Higher diversity was observed in continental sites than in English roosts (mean pairwise differences: \( p < 0.01 \); haplotype diversity: \( p < 0.05 \); nucleotide diversity: \( p < 0.01 \); Table S2). Thus, all continental sites sequenced for more than one individual displayed several haplotypes. No continental haplotypes were shared among roosts, the exception being the French samples that shared all five haplotypes with several samples from eastern England and the IOW (Figure 4). In contrast, in England, one haplotype was shared by 322 out of the 409 English samples, and all but one English roost contained this haplotype.

Furthermore, 13 out of the 28 English roosts were fixed for this haplotype. More diversity was found in the roosts of eastern England and the IOW than in western England, which displayed only six haplotypes, two of them being shared with the rest of England. Estimated \( \Phi_{ST} \) was high at 0.405 and significant, indicating strong differentiation among roosts at the mtDNA level.
Mitochondrial population structure

The sPCA investigated sequence variability in a spatial context. The two first axes had the highest eigenvalues and were therefore retained. PC-1 separated France and England from the rest of the continent, although the boundary was not sharp and a slight cline across the whole of Europe could be noticed. PC-2 created two clusters on the continent corresponding to eastern sites (Slovakia, Poland) and western sites (Italy, Switzerland, France). It also segregated the IOW from the rest of England, and eastern England from western England, although these last scores were very low, suggesting this structure is weak (Figures 3 and S4).

AMOVA produced a simpler structure with the among groups variance component optimised for three groups ($\phi_{CT} = 0.481$): eastern continent, western continent and England (Table 3). Pairwise $\phi_{ST}$ between the three groups (Table 4) indicated strong differentiation between England and both the eastern and western continental populations, and much weaker differentiation between the two continental populations. All pairwise differentiation tests were significant ($p = 0.008$). The Mantel test revealed a significant correlation between genetic and geographical distances when controlling for the potential effects of geographical barriers ($r = 0.281$, $p < 0.05$; Figure S3). However, physical barriers did not contribute in structuring genetic variation when controlling for geographical distances ($r = 0.109$, $p = 0.170$). When splitting the full dataset into a continental one and an English one, no cline in genetic differentiation or effect of geographical barriers were detected on the continent or in England (continent:
geographical distance: $r = 0.206$, $p = 0.209$, barriers: $r = 0.295$, $p = 0.074$; UK: geographical distance: $r = 0.079$, $p = 0.118$, barriers: $r = 0.019$, $p = 0.175$.

Phylogenetic analysis

The haplotype network displayed a highly connected star-shaped structure, indicative of rapid population expansion (Figure 4). Some division within England was observed, although the two haplotypes responsible differed by only one base-pair, suggestive of the weakness of this structure. The topology of the ML and Bayesian analyses demonstrated a similar pattern of shallow genealogy among haplotypes (Figure 4). Overall there was a lack of obvious clustering or clear geographical patterns in the haplotype genealogies suggesting that all the haplotypes belong to the same clade and indicating rapid population expansion following a demographic bottleneck. This was confirmed by the mismatch analysis run on the three populations defined by the AMOVA and by $F_S$ values. The English population displayed significantly negative $F_S$ and non-significant sum of squared deviation (SSD) and raggedness index, indicating no significant difference from a model of sudden expansion (Table S4 and Figure S4). The signal of population expansion was less clear for eastern continent and western continent as the mismatch analysis also yielded non-significant SSD and raggedness indices (Table S4), despite mismatch distributions more consistent with those from stable populations (Figure S4). Furthermore, their $F_S$ values, while negative, were non-significant (Table S4). Together those results indicate that the continental populations are likely stable while the English one recently expanded.
Discussion

Our analysis of the genetic structure of *E. serotinus* across Europe and within England revealed low but significant population genetic structure at the largest geographical scale, with English serotines being somewhat distinct from their European peers. There were different patterns of genetic partitioning within these populations, with stronger genetic homogeneity on the continent than in England, consistent with the hypothesis of reduced movement at edge population compared to core population. Despite this structuring, high admixture and assessment of gene flow among the differentiated populations indicated westward movement within England and a lesser role of the English Channel as physical barrier than initially thought. Stronger differentiation in mtDNA was detected, suggesting females are more philopatric than males, as we expected from a bat, and therefore gene movement is mostly mediated by male connectivity. We discuss each finding in turn, as well as the limitations of our study and direction for further research where appropriate.

Genetic differentiation among roosts across Europe was low (*F*\(_{ST}\) = 0.048) but significantly different from panmixia, with a pattern of isolation by distance indicating some limit to gene flow. Bats tend to be relatively mobile and levels of population structuring are often low (Moussy *et al*., 2013). For example, the migratory bat *Nyctalus noctula* displays very low levels of genetic structuring across Europe (*F*\(_{ST}\) = 0.006) (Petit and Mayer, 1999). The fixation index we estimated for *E. serotinus*, while low, is relatively high for a bat and reflects values
reported for non-migratory species (Moussy et al., 2013). Nonetheless, the current evidence indicates *E. serotinus* is relatively sedentary.

Consistent with the departure from panmixia, bayesian and multivariate analyses of nuclear data recovered four genetic populations, geographically represented as continental Europe, east of England, west of England and the Isle of Wight. Some evidence for physical barriers to gene flow was also found. Hence, while there was weak evidence that mountains act as barriers, the sea appears to restrict gene flow to a degree, even between England and the Isle of Wight - 6km. This distance is less than the mean commuting-distance/night recorded for *E. serotinus* (Catto *et al.*, 1996). However, open water restricts movement for many bats (García-Mudarra *et al.*, 2009; Moussy *et al.*, 2013).

Despite the large geographical scale (> 3,000 km) and several mountain ranges, individual-based analyses of the nuclear markers failed to detect any strong sub-structuring within continental Europe. Similar findings have been reported in other taxa (Demont *et al.* 2008) and suggest genetic homogenisation over a large geographical scale and possibly high gene flow across the continent. Lack of information on the species’ mating system and movements restricts our understanding of the mechanisms driving this gene flow. However, it is worth noting that the sampling design to assess continental genetic diversity and distribution thereof was not ideal since only eight locations across the continent were sampled. The full extent of genetic diversity could thus not be recovered. Finer genetic sub-structuring could have be missed and further work with more
intensive sampling across more locations is required to fully understand patterns of genetic partitioning across the continent.

Within England, the fixation index was as high as within continental Europe, in spite of the much smaller geographical scale, and smaller effective population sizes could contribute to the high fixation index there. All analyses of nuclear markers identified three UK populations: eastern England, western England and the Isle of Wight. This pattern of stronger genetic structure in England than on the continent had previously been reported in other bat species (Flanders et al., 2009). The drivers for this relatively strong structuring are unknown, but the UK represents the most northern limit of many bat species, including E. serotinus, and this could affect individual behaviour (e.g. dispersal) and population dynamics, and influencing genetic structure. Lower dispersal at range edges has indeed been shown in other taxa (Travis and Dytham, 1999), primarily due to increased dispersal costs and reduced habitat density (Dytham, 2009). The identification of an Isle of Wight population can be explained by the sea restricting gene flow between the island and mainland England. However, no obvious physical or behavioural barriers, habitat fragmentation or lack of connectivity seem to explain what appears to be two partially isolated populations on the English mainland, although greater urbanisation in eastern England could potentially limit movement. An alternative explanation is that the East/West structure in England represents two catchments of gene circulation, possibly driven by limited and geographically distinct swarming/hibernation sites with ongoing gene flow through male, and to a lesser extent female, dispersal. Our finding of structure in
England starkly contrasts with a recent study of the genetic structure of *E. serotinus* in Poland (Bogdanowicz et al., 2013). The geographical scale of our study and the Polish one are similar, and four out of the seven microsatellites they used were used here. However, Bogdanowicz et al. failed to detect any nuclear structure, recovering only one genetic population and very low differentiation among roosts ($F_{ST} = 0.01$). While these results differ from our English findings, they are consistent with our continental results, despite the sampling limits of our study.

Historical events, like the post-glacial recolonisation of Europe, influence the partitioning of genetic variation for some bats (Flanders et al., 2009), and the significantly differentiated eastern and western continental clusters we detected at the mitochondrial level could be due to post-glacial recolonisation from separate refugia, with the low structure in the nuclear markers reflecting high contemporary gene flow. Additionally, lower mitochondrial diversity was found in England compared to mainland Europe, which, together with a unimodal mismatch distribution, could indicate past population bottlenecks before rapid population expansion in the UK. However, again, caution needs to be used in this interpretation due to the low continental sampling. Increasing the sample size and coverage on the continent and at the putative contact zone would be required to reconstruct the post-glacial history of the species and its contemporary influence on the distribution of genetic diversity.
While significant genetic structure was found for *E. serotinus*, there was nonetheless high admixture, as indicated by Bayesian and multivariate analyses of nuclear markers and the low $F_{ST}$. This indicates strong ongoing gene flow among populations rather than lack of drift because of recent population divergence, which in turn indicates regular movement of individuals between populations, even across the English Channel. Specifically immigration from continental Europe to the species’ northern English range was inferred from nuclear data. Immigrants appear to be going mostly to the Isle of Wight and eastern England. Whilst eastern England is geographically close to continental Europe, the shortest distance to the Isle of Wight is ca. 150 km, so it is less clear why it appears so connected to mainland Europe. However interpretation here has to be tempered by the low number and the location of the continental samples and unfortunately no samples could be obtained from the continental borders of the English Channel. Gene flow analysis therefore only estimated movements from the overall continental population and might not reflect the true extent of gene movement, which are probably underestimated.

Within England the low biparental differentiation between populations and the recovery of eastern mitochondrial haplotypes in the west indicates most bat movements are from east to west. Similarly, movements between the Isle of Wight and mainland England are mostly directed toward western England. This could indicate a westward expansion of the species range, consistent with the lower haplotypic and nucleotide diversity in western England. *E. serotinus* has indeed been mostly recorded in south eastern England and its presumed absence
from far west has never been fully explained since habitat and roosting
opportunities seem favourable. However, since 2004, the National Bat Monitoring
Programme (NBMP) has reported a significant regional changes in bat
distributions, with a steep increases in the southwest (Bat Conservation Trust,
2013). This is consistent with reports from local bat workers and volunteers
registering *E. serotinus* calls in western counties on a more frequent basis,
prompting them to suspect the establishment of new communities in the west
(Marshall, pers. comm.).

Significant differentiation in the nuclear and the mitochondrial data occurred
among sites at the European scale. However, estimates of genetic differentiation
revealed that the level of population structure displayed by mtDNA haplotypes
was ten times higher than that inferred by nuclear markers. While the comparison
of differentiation between different markers is difficult due to different rates and
modes of mutations, with high migration, as expected for vagile species like bats,
motion is likely to contribute little to differentiation (Balloux and Lugon-Moulin,
2002). Therefore the observed difference between biparental and maternal
fixation indices is probably the result of differences in effective population size for
the markers and sex-biased gene flow. Conservatively assuming that sexual
selection in *E. serotinus* is weak, the contrast in differentiation measured between
the two types of markers exceeds their expected four-fold difference in effective
population size. This suggests that gene flow in *E. serotinus* is mostly mediated by
males, while higher differentiation among roosts for mtDNA indicates female
philopatry (assuming the demography of this marker is a true reflection of
population demography (Balloux, 2010)), which correlates with the qualitative experience of workers studying *E. serotinus* across Europe (Hutson, pers. comm.; Harbusch and Racey 2006). The suggestion of male-biased gene movement and female philopatry we infer is widely observed in temperate bats (Moussy *et al.*, 2013), but further work analysing adult male movements is required to confirm this pattern.

Nonetheless, some female dispersal occurs in *E. serotinus* as several mitochondrial haplotypes are shared among roosts. For example, all haplotypes found in French specimens are shared with roosts in eastern England and the Isle of Wight. The coexistence of these haplotypes within roosts could be the results of past colonisation events, and possibly the contemporary recruitment of dispersing females. Similarly, one haplotype common to eastern England was also identified in a roost in the west, consistent with female dispersal and recruitment. However, the pattern of isolation by distance found over Europe suggests limits to female dispersal, although this could also reflect local selection for different haplotypes (Balloux 2010; Arnqvist *et al*. 2010). Furthermore, no haplotypes are shared among continental sites. The geographical scale involved there is much larger than within England, and it therefore appears that female dispersal occurs over sub-continental spatial scales (e.g. <500km). This is further supported by the fact that with the exception of the main English haplotype that is widely distributed, most shared haplotypes in the UK are found in neighbouring roosts. Hence, whilst female dispersal can occur over longer distances, most females apparently remain philopatric to their natal landscape if not to their natal roost.
Conclusion

Our findings indicated considerable gene flow on a large geographical scale in continental Europe despite the presence of physical barriers that could contribute to some level of sub-structuring, but further work with better sample coverage is required to confirm these patterns. Nonetheless putative movement could contribute to large scale circulation of EBLV-1 in Europe, maintaining the high viral prevalence recorded there (Schatz et al., 2013). This also seems to be concordant with the relatively low genetic diversity of EBLV-1 across Europe possibly promoted by its host’s movement (McElhinney et al. 2013). We also suggest that the English Channel, while apparently restricting some gene flow, does not appear to be a substantial barrier, and similar inferences have been drawn for M. daubentonii (Atterby et al., 2010), which is the host for EBLV-2, a virus found both on the continent and in the UK (Harris et al., 2006). The relative fragmentation of serotine populations in England may sufficiently alter host-virus epidemiology to explain the apparent absence of EBLV-1 in England. This explanation is consistent with a previous small-scale comparison of the genetic structure of the EBLV-1 (E. serotinus) and EBLV-2 (M. daubentonii) hosts that revealed more structure within England in E. serotinus than in M. daubentonii (Smith et al., 2011). The cause(s) of the reduced movement of serotines within the UK remains to be established.

Acknowledgements

The authors acknowledge DEFRA and University of Exeter for funding this study. We are grateful to the local bat groups for their help in locating and arranging
access to roosts, to all the home-owners for their hospitality and tolerance during field work, and to the volunteers that assisted in the capture and sampling. We also thank Laurent Arthur, Philippe Christé, Adrino Martinoli, Martina Spada, Martin Celuch, Wieslaw Bogdanowicz, and Loseb Natradze for providing the continental wing biopsies. Javier Juste shared his experience in sequencing serotine mitochondrial DNA and provided primers sequences and conditions. We also thank the Editor and Referee’s for constructive comments which greatly improved the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Data archiving

Mitochondrial haplotypes will be deposited in Genbank

Microsatellite genotypes and sample location records will be deposited in Dryad

Supplementary information is available at Heredity’s website

(http://www.nature.com/hdy)

References


http://beast.bio.ed.ac.uk/software/tracer/


