THE LINNEAN SOCIETY of London BIOLOGICAL Journal of the Linnean Society

Gene flow in European coal tits (Aves, Passeriformes, Periparus ater): low among Mediterranean populations but high in a continental contact zone

Journal:	*Biological Journal of the Linnean Society*
Manuscript ID	BJLS-5388.R1
Manuscript Type:	Original article
Date Submitted by the Author:	n/a
Complete List of Authors:	Tritsch, Christian; Senckenberg Naturhistorische Sammlungen Dresden, Museum of Zoology Stuckas, Heiko; Senckenberg Naturhistorsiche sammlungen Dresden, population genetics Martens, Jochen; Institut für Zoologie, ; Pentzold, Stefan; Max Planck Institute for Chemical Ecology, Department of Bioorganic Chemistry Kvist, Laura; Uni, Department of Biology LoValvo, Mario; Universita degli Studi di Palermo, Dipartimento di Scienze e Tecnologie biologiche, chimiche e farmaceutiche Giacalone, Gabriele; Cooperativa Silene Tietze, DT; Natural History Museum Basel Nazarenko, Alexander; Institute of Biology and Soil Sciences FEB RAS, Lab. of Ornithology Päckert, Martin; Senckenberg Natural History Collections, Museum of Zoology
Keywords:	phylogeography, island populations, microsatellites, subspecies, glacial refugia

Gene flow in European coal tits (Aves, Passeriformes, *Periparus ater*): low among Mediterranean populations but high in a continental contact zone

Christian Tritsch[1,2], Heiko Stuckas[1], Jochen Martens[3], Stefan Pentzold[2,4], Laura Kvist[5], Mario LoValvo[6], Gabriele Giacalone[7], Dieter Thomas Tietze[8], Alexander A. Nazarenko[9], Martin Päckert[1]

[1)]Senckenberg Natural History Collections, Königsbrücker Landstraße 159, 01109 Dresden, Germany

[2)]Institute of Biology, Molecular Evolution & Animal Systematics, University of Leipzig, 04103 Leipzig, Germany

[3)]Institute of Organismic and Molecular Evolution, Johannes Gutenberg University, 55099 Mainz, Germany

[4)]Max Planck Institute for Chemical Ecology, 07745 Jena, Germany

[5)]Department of Ecology and Genetics, University of Oulu, POB 3000, 90014 Oulu, Finland

[6)]Dipartimento di Scienze e Tecnologie biologiche, chimiche e farmaceutiche, Via Archirafi 18, 90123 Palermo, Italy

[7)]Cooperativa Silene, Via Dondes Regio, 8/a, 90127 Palermo, Italy

[8)]Natural History Museum Basel, Augustinergasse 2, 4001 Basel, Switzerland

[9)]Federal Scientific Center of the East Asia Terrestrial Biodiversity, Far East Branch of the Russian Academy of Sciences, Vladivostok 690022, Russia

Corresponding author: Martin Päckert; martin.paeckert@senckenberg.de; phone: +49-351-7958414344; fax: +49-351-7958414344

Abstract

Extant phylogeographic patterns of Palearctic terrestrial vertebrates are generally believed to have originated from glacial range fragmentation. Post-Pleistocene range expansions have led to the formation of secondary contact zones among genetically distinct taxa. For coal tits (*Periparus ater*), such a contact zone has been localized in Germany. In this study, we quantified gene flow between Fennoscandian and southern European coal tits using a set of 13 microsatellite loci. STRUCTURE analysis revealed four genetic clusters two of these on Mediterranean islands. German populations were genetically admixed but introgression of southern alleles was evident for Fennoscandian populations. In the South, we found negligible introgression of northern alleles (and haplotypes) but slight admixture of two southern genetic clusters in the Pyrenees and on the Balkan Peninsulae and near complete sorting of these two allelic lineages on the islands of Corsica and Sardinia. Genetic distinctiveness of the Mediterranean island populations reflects general patterns of endemism in the Corso-Sardinian fauna and the Cypriot fauna. Wide-range gene flow in Central Europe suggests a broad zone of intergradation between subspecies of the coal tit rather than a narrow contact zone. This is in accordance with low morphological and bioacoustic differentiation of European coal tit populations.

Key words: phylogeography – island populations – microsatellites – subspecies – glacial refugia

Introduction

Evolutionary biologists widely agree that glacial impact considerably shaped phylogeographic patterns and speciation of terrestrial vertebrates in the Palearctic (Avise & Walker, 1998; Hewitt, 2000, 2004; Lovette, 2005; Zink *et al.*, 2008; Stewart *et al.*, 2010). Pleistocene separation of Eastern and Western Palearctic populations led to divergence of gene pools among distant refugia, in a few extreme cases across a large extant distributional gap, such as seen in the marsh tit, *Poecile palustris* (Tritsch *et al.*, 2017) or in the azure-winged magpie, *Cyanopica cyanus* (Zhang *et al.*, 2012). Other East-West lineage splits dating back to Pleistocene events were reconstructed for example in corvids (Haring *et al.*, 2007, 2012) and tits (Kvist *et al.* 2003; Päckert *et al.*, 2005; Kvist & Rytkönen, 2006). One noticeable result from Holocene range expansion is the spatial overlap of genetically distinct populations that is manifested in secondary contact zones of a highly variable extent (Woodruff, 1973; Haffer, 1989; Aliabadian *et al.*, 2005). In Western Europe, the apparent spatial clustering of secondary contact zones among terrestrial vertebrate sister taxa was the result of postglacial expansion from southern glacial refugia (Hewitt, 2000; Schmitt, 2007). Parapatry along sharp and narrow hybrid zones is typically found at geographic barriers, such as the European mountain systems that separate two larger glacial refugia from the continent: i) the Iberian Peninsula in the Pyrenees (Fig. 1A; birds: Helbig *et al.*, 2001; Pons *et al.*, 2011; Backström *et al.*, 2013; Kuhn *et al.*, 2013; reptiles: Mila *et al.*, 2013; insects: Vasquez *et al.*, 1994; Shuker *et al.*, 2005; Bella *et al.*, 2009); ii) the Italian Peninsula in the Alps (Fig. 1A) II; birds: Hermansen *et al.*, 2011; rodents: Sutter *et al.*, 2013; Gimenéz *et al.*, 2017; insects: Flanagan *et al.*, 1999). Apart from parapatry across mountain ranges, narrow hybrid zones of a wide latitudinal extent exist in Central Europe (Fig. 1A), the best-studied examples being those of crows (*Corvus c. corone*, *C. c. cornix*: Haas *et al.*, 2009, 2010; Wolf *et al.*, 2010; Poelstra *et al.*, 2014a, b; other birds: Secondi *et al.*, 2011), the house mouse (*Mus m. musculus*, *M. m. domesticus*: Macholán *et al.*, 2008; Gimenéz *et al.*, 2017) and hedgehogs (*Erinaceus*

europaeus, *E. roumanicus*: Berggren *et al*., 2005; Bolfíková & Hulva, 2012; Waters et al. [2013: Fig. 1B]; Pfäffle *et al*., 2014).

In several bird species pairs, zones of secondary contact and hybridization are not restricted to a narrow band, but extend along a wide longitudinal range into Eastern Europe (Fig. 1B), such as found in flycatchers (*Ficedula*: Sætre *et al*., 2001; Hogner *et al*., 2012a), reed warblers (*Acrocephalus*: Reifová *et al*., 2016), nightingales (*Luscinia*: Vokurkova *et al*., 2013), tits (*Cyanistes*: Woodruff, 1973; Stervander *et al*., 2015) and Old World buntings (*Emberiza*: Irwin *et al*., 2009).

In a few other examples, pre-mating barriers were either not established during a short separation time in refuge areas or they simply broke down in secondary contact, which had led to merging of divergent genetic lineages. The signal from mitochondrial markers might then remain the only testimony of past (Pleistocene) lineage separation, presently contrasted by narrow or wide-range gene flow (Fig. 1C), as suggested for some passerine bird species (Zink *et al*., 2008; Päckert *et al*., 2010; Hogner *et al*. 2012b; Block *et al*., 2015). Interbreeding and merging of gene pools between cryptic genetic lineages of a phenotypically uniform species has been sometimes termed "speciation in reverse" (e.g. in birds Webb *et al*., 2011). However, in the strict sense, reverse speciation is more appropriately applied to those examples where gene pools become largely absorbed due to hybridization with one of the two parental species running the risk of going extinct (in fishes: Seehausen *et al*., 2008; Taylor *et al*., 2006; Hudson *et al*., 2013; Bath *et al*., 2014; in Darwin's finches Kleindorfer *et al*., 2014).

Our study focuses on the recent evidence of secondary range overlap in Western Europe among a north-eastern and a south-western mitochondrial lineage of the coal tit, *Periparus ater* (Pentzold *et al*., 2013). In this study, we aim at verifying the extent and degree of nuclear gene flow among the two coal tit lineages using nuclear markers (microsatellites). We expect significant gene flow at least in the region of considerable mtDNA lineage overlap at a contact zone extending throughout Germany (Fig. 2). We also expect nuclear gene flow

to extend across a wider range than mitochondrial introgression, as was shown for another parid hybrid zone (*Parus major*/ *P. minor*: Kvist & Rytkönen, 2006).

Material and methods

Study species

Eight divergent mitochondrial lineages are presently known in the coal tit. These lineages are distributed across large parts of the Palearctic, the mountain forests of China, the Himalayas, Karakoram and Hindu Kush as well as on Taiwan (Tietze *et al.*, 2011). Across the Western Palearctic, four distinct mtDNA lineages of the coal tit occur (Fig. 2): i) the north-eastern Palearctic (*ater* subspecies group; distributed from Northern Europe across the Eurasian continent to the Pacific coast and Japan), ii) Central and Southern Europe (*abietum* subspecies group) including the British Isles and the islands of Corsica and Sardinia, iii) North Africa and iv) Cyprus (Martens *et al.*, 2006, Tietze *et al.*, 2011, Pentzold *et al.*, 2013). Range overlap of the south-western *abietum* and the north-eastern *ater* lineages could so far be restricted to the German populations only (Fig. 2). Due to a lack of reliable morphological and bioacoustic distinctiveness of north-eastern versus south-western Palearctic coal tits, the spatial dimension of the contact zone cannot be delimited by geographical variation of phenotypes or song types (Tietze *et al.*, 2011; Pentzold *et al.*, 2013, 2016).

Sampling and multilocus genotyping

We sampled 166 birds from Russia, Kyrgyzstan, Kazakhstan, Finland, Norway, Germany, French Pyrenees, Corsica and Sardinia. DNA preparation was conducted either using the innuPREP DNA Mini Kit (muscle tissue, Analytik Jena AG, Germany) or the PEQLAB GOLD Blood DNA Mini Kit (blood samples, PEQLAB Biotechnologie GmbH, Germany), following the manufacturers' advice.

New microsatellite loci for *P. ater* were identified by Ecogenics GmbH (Zürich, Switzerland) based on an enriched DNA library. Size selected genomic DNA was ligated into SNXforward/SNX reverse-linker (Hamilton *et al.*, 1999) and enriched by magnetic bead selection with biotin-labelled oligonucleotide repeats ($(CT)_{13}$, $(GT)_{13}$, $(GTAT)_7$, $(GATA)_7$; Gautschi *et al.*, 2000 a,b). A total number of 528 recombinant colonies were screened and 415 gave a positive signal after dot-blot hybridization. Plasmids from 48 positive clones were Sanger sequenced and primers were designed for 16 microsatellite inserts of which nine (listed in Table 1) were finally used for amplification of polymorphic microsatellite loci using the protocol described by Schuelke (2000). For this protocol a M13(-21) tail (18 bp) was adhered to the 5' end of the forward primer. The reverse primer remained unmodified. In addition to the two regular PCR primers, a fluorescent-labelled universal M13(-21) primer was added to the reaction mixture. The reaction contained 10 to 40 ng of template DNA, 0.04 µM of the M13-forward primer, 0.16 µM of the reverse primer and the labelled M13 primer, 0.2 mM of each dNTP, 1 µL of 10x PCR reaction buffer "complete" and 0.5 units of DFS-Taq DNA polymerase (Bioron GmbH, Germany) in a total volume of 10 µL. The thermo-treatment consisted of two successive steps: a) amplification of the microsatellite fragment with M13 fusion, b) labelling of the fragment with the fluorescent dye. The PCR program was 95 °C for 10 min followed by 30 cycles of 30 s of 95 °C, 45 s of 50 °C (Parate8) or rather 56 °C (all other Parate loci) and 45 s of 72 °C (step a), followed by eight cycles of 30 s of 95 °C, 45 s of 53 °C and 45 s of 72 °C (step b) and a final elongation at 72 °C for 30 min.

Four additional primer pairs targeting microsatellite loci were obtained from previous studies on *Poecile atricapillus* (Table 1). We tested for cross amplification with *P. ater* samples in a total volume of 10 µL containing 10 to 40 ng of template DNA, 0.3 µM of each primer, 0.2 mM of each dNTP, 1 µL of 10x PCR reaction buffer "complete" and 0.5 units of DFS-Taq DNA polymerase (Bioron GmbH, Germany). The thermo-cycling protocol was as

follows: 94 °C for 5 min followed by 35 cycles of 94 °C for 30 s, 60 °C (Pat2-43) or 57 °C (PmaC25, PmaTGAn33, Pma69) for 30 s, 72 °C for 45 s and a final elongation at 72 °C for 5 min.

Specimens were genotyped at 13 loci (Table 1) using the reaction conditions outlined above and labelled PCR fragments were run on a 16-column ABI 3130*xl* capillary sequencer (Applied Biosystems). The alleles were scored using the STRand Analysis Software vers. 2.4 (UC Davis, veterinary genetics lab, http://www.vgl.ucdavis.edu/STRand; Toonen and Hughes, 2001).

The software package MICROCHECKER 2.2.3 (van Oosterhout *et al.*, 2004) was used to test the probability that experimental errors occurred during microsatellite genotyping, i.e. large allelic dropout, scoring errors due to misinterpretation of stutter bands and null alleles.

Diversity and Divergence

Summaries of allele sizes and the existence and frequencies of population specific alleles (private alleles) were calculated using the program CONVERT vers. 1.31 (Glaubitz, 2004), which was also employed to generate input files for various software packages. Linkage between loci was determined using ARLEQUIN vers 3.5.1.3 (Excoffier *et al.*, 2005). The same software package was used to calculate locus-specific observed and expected heterozygosities (H_O, H_E) for each sample population and to test for locus-specific deviations from Hardy Weinberg expectations (HWE). Population specific deviations from HWE (excess or deficiency of heterozygosity) across all loci were explored using inbreeding coefficients (F_{IS}), calculated with the software FStat (ver. 2.9.3.2; Goudet 1995) using a randomisation test (3600 randomisations) to test for significance. The same software was used to estimate the mean number of alleles per locus and populations as well as the mean allelic richness (AR) per population across all loci. For these, we analysed a reduced dataset

consisting of 145 specimens that belong to 14 distinct populations with a minimum of five specimens per population (Table 2, Table S1).

Divergence between populations was estimated using F-statistics (inferred from microsatellite allele frequencies) and Φ-statistics (inferred from mitochondrial nucleotide sequences) by pairwise F_{ST} and Φ_{ST} values as well as by non-hierarchical and hierarchical locus-by-locus analysis of molecular variance (AMOVA with F_{CT} and Φ_{CT} values as a measure of divergence among groups) using ARLEQUIN. 20000 permutations were performed to test for significance of these values. All p-values obtained from tests implementing multiple comparisons (i.e. test for deviation from HWE expectation, test for linkage between loci) were Bonferroni corrected to adjust the significance threshold (Rice, 1989). In order to depict divergence between populations, pairwise F'_{ST} values were used in a distance matrix to construct a UPGMA phenogram with MEGA v.6 (Tamura *et al.*, 2013).

Bayesian inference of the population structure

Non-spatial Bayesian inference of population structure was performed using the software package STRUCTURE vers. 2.3.3. (Pritchard *et al.*, 2000; Falush *et al.*, 2003). STRUCTURE runs were performed i) under the *a priori* assumption of genetic admixture and correlated allele frequencies and ii) under a LOCPRIOR model that allows for classification of the individuals into groups, which are given to the algorithm as an *a priori* parameter (Hubisz *et al.*, 2009). The model was run under two different LOCPRIOR settings: i) by classifying the individuals of the complete data set (n = 166) according to their assignment to mitochondrial lineages (inferred from the data set by Pentzold *et al.*, 2013) and ii) by assigning the individuals to 14 local populations of $n \geq 5$ (total sampling n = 145, Table 2).

All STRUCTURE runs were conducted for 1–10 putative genetic clusters (K) with ten replicates for each value of K. The number of MCMC runs was 10^5 with a burn-in period of 25 000 throughout all model runs. For further processing of the STRUCTURE output,

STRUCTURE HARVESTER (Earl & von Holdt, 2012) was used. In order to select the most likely number of genetic clusters (K), we used the approach by Evanno *et al.* (2005). STRUCTURE analysis was also used to estimate the extent of genetic admixture in different populations according to the method described by Randi (2008). Accordingly, we used a threshold $q > 0.80$ for the assignment of individuals to a cluster or we classified individuals as admixed individuals, if the proportion of membership was $q < 0.80$ (see Randi, 2008).

Spatial Bayesian clustering was performed using the software packages TESS vers. 2.3.1 (Chen *et al.*, 2007) and GENELAND vers. 4.03 (Guillot *et al.*, 2005b). Unlike the non-spatial models of STRUCTURE, the spatially explicit models implemented in TESS and GENELAND consider the geographic coordinates of the samples, but do not consider affiliation as a model parameter. Data exploration under different models is recommendable, because explanatory power of spatial *versus* non-spatial models depends on demographic scenarios to be tested and clustering output from one model-based method might reveal a finer scaled spatial structure that other models fail to detect (François & Durand 2010). Both admixture models in TESS (BYM and CAR model, Durand *et al.*, 2009a, b) were run with the complete dataset of n = 166 individuals (MCMC iterations: 10^5, burn-in period: 20 000, Kmax: 2–10, five replicates for each Kmax for both models). Unlike STRUCTURE or TESS the standard models of GENELAND do not account for admixture, but assign posterior probabilities of cluster membership to the single individuals. A benefit of GENELAND is that it can correct for the occurrence of null alleles. Although from a biological perspective it seems obvious to assume the allele frequencies to be correlated between populations, the respective model was assessed to systematically overestimate the number of clusters (Guillot *et al.*, 2005a). Hence according to the authors' recommendations the analysis was conducted in two steps: first, resolving the number of populations (K) using the D model (*frequency model* = uncorrelated; K: 1–10); second, deriving the correct population assignment by applying the F model (*frequency model* = correlated) with a fixed number of K (which was

determined in the first step, Guillot *et al*., 2005a, GENELAND documentation). The model parameters were: 10^6 MCMC iterations, Thinning = 1000, Null allele model = TRUE and ten replicates per analysis step. The outputs of STRUCTURE and TESS both were further processed with CLUMPP (Jakobsson & Rosenberg, 2007). For visualization, DISTRUCT (Rosenberg, 2004) was used.

Admixture rate in the hybrid zone

For estimating admixture rate in the hybrid zone we applied demographic modelling based on Approximate Bayesian Computation (ABC) using the program DIYABC v. 2.0.4 (Cornuet *et al*., 2010). We first used both microsatellite and mitochondrial control-region sequences and included individuals from which both data were available. We chose samples from Norway and Finland (n = 18) with q-values from STRUCTURE above 0.8 to their cluster '*ater*' to represent the 'northern' parental population and samples from French Pyrenees with q-values from STRUCTURE above 0.8 to their cluster '*abietum*' to represent the 'southern' (n = 8) parental population. This was based on further evidence that these populations represented only one mitochondrial lineage each and were clearly separated from, but still related to the admixture populations in the STRUCTURE analysis from the microsatellite data. In the central European admixture population, we included samples from Schleswig Holstein, Harz, Saxony, Palatine Forest and Black Forest (n = 47, all Germany).

We started by constructing four historical models (Fig. 3): 1) parental 'northern' population and 'southern' population were split from each other at time t2 and come into a contact at time t1 to form the admixture population, 2) the parental populations split first from each other and the central European population was split later from the northern population, 3) the northern and central populations split first from each other and the southern population was split later from the northern population and 4) the parental populations split first from each other and the central European population was split later from the southern population.

Mutation rate for the microsatellite data was set to 10^{-4}–10^{-3} as was done with another tit species, the blue tit *Cyanistes caeruleus* (Hansson *et al.*, 2014). For the mitochondrial sequences, we applied the substitution rate 1.156 * 10^{-8} calibrated for coal tit control region in Pentzold et al. (2013), HKY+Gamma model with gamma = 0.09 (as suggested by the test for the substitution model implemented in MEGA v. 6.06; Tamura *et al.*, 2013). As the fit of observed data with the simulated data was poor, we next performed the same analysis separately for the microsatellite data and the mitochondrial data. For the microsatellite data, uniform prior distributions for effective population sizes were set to 10–10 000 and for coalescence times, t1 was set to 10–1000 and t2 to 10–4000. The uniform prior distribution for the admixture rate was 0.001–0.999. The priors for effective population sizes and coalescence times were changed for mitochondrial analyses to N_e = 1000–1 000 000 for 'northern', 'southern' and ancient populations, N_e = 1000–2 000 000 for the population at the contact zone, t1 = 10–20 000 and t2 = 1000–2 000 000, as the fit of the observed and simulated data was poor when using the same priors as for the microsatellite data. Altogether, 4 000 000 data sets were simulated for both microsatellite and mitochondrial data.

For calculation of time of divergence and time since admixture we assumed a mean generation time of approximately two years in tits (as applied by Hansson *et al.*, 2014; compare 1.5 years for the great tit, *Parus major*, in Qu *et al.* (2015) and 2.26 years for the willow tit, *Poecile montanus*, in Kvist *et al.* (2001)). We expect time estimates inferred from mitochondrial DNA to correspond with the onset of lineage splitting and admixture caused by paleoclimatic events more accurately, because nuclear loci will generally reach coalescence slower and at a later stage of evolution (Palumbi *et al.*, 2001). Therefore, time estimates inferred from microsatellite data are generally supposed to post-date paleoclimatic events that triggered lineage divergence.

Furthermore, the ratio between male to female gene flow (m_m/m_f) was calculated according to the equations in Hedrick *et al.* (2013). The main assumptions implemented in

this approach are that populations can be described according to the island model and that populations are in migration-drift equilibrium. The approach uses divergence levels caused by female gene flow as $F_{ST(f)}$ values derived from mtDNA and estimates divergence levels caused by male gene flow $F_{ST(m)}$ from microsatellites (Eqn. 7a; Hedrick *et al.*, 2013). Both divergence levels were used to calculate the m_m/m_f ratio (Eqn. 7b; Hedrick *et al.*, 2013). Given the island model and migration drift equilibrium as basic assumptions, the estimates were alternatively performed i) for the total set of populations and ii) under exclusion of Mediterranean island populations for all continental Eurasian populations.

Results

Microsatellite genotyping

Deviations from HWE were predominately found at loci Parate 6 (six populations) and Parate 8 (seven populations), but also in single populations at loci PmaC25, Parate 15, Parate 16, Parate 3 and Parate 2. The occurrence of HWE deviations at these loci was associated with the presence of null alleles (Table S1). Most HWE deviations were found in the population from Schleswig-Holstein, which also had the highest positive F_{IS} value (Table 2). Furthermore, in this population we found evidence of linkage disequilibrium among alleles at six loci (Table S1). None of the other study populations showed a signal of linkage disequilibrium except two island populations from Cyprus and Corsica (at two loci each; Table S1). Our analyses did not provide evidence of genotyping artefacts due to large allele dropout and misscoring of genotypes due to stuttering was almost absent except at locus Parate 8 in a single population (Cyprus). These analyses suggested that loci Parate 6 and Parate 8 should be treated with a precaution in subsequent analyses (see below).

Diversity and Divergence

Given the null allele bias at loci Parate 6 and Parate 8, we excluded these data from estimating diversity and divergence for 14 populations (n = 145). The mean allele number over loci varied between 9.9 (Schleswig-Holstein) to 4.5 (Central Asia; Table 2) due to variation in sample sizes. However, allelic richness that is corrected for differences in sample sizes, ranged at similar values across populations (4.1–5.2), except for Sardinia, where it was considerably lower than in other populations (3.5; Table 2). Based on the expected heterozygosity (H_E), the genetic variation appeared relatively high and fairly constant between the samples (0.75–0.86), excluding again Sardinia that had the lowest H_E (0.63; Table 2). Most population samples showed moderate deficit of heterozygotes as indicated by positive inbreeding coefficients (Table 2). This was the most pronounced in the three German populations from Schleswig Holstein, Harz and Saxony, but also in Norway.

Non-hierarchical AMOVA indicated divergence between populations (F'_{ST} = 0.065, P < 0.0001). Pairwise F_{ST} values were calculated (Table S2) and depicted as an UPGMA phenogram reflecting the existence of four clusters of populations (Fig. 4). There is a significantly high amount of genetic divergence between these four groups as indicated by a hierarchical AMOVA (microsatellites: F_{CT} = 0.076, P < 0.001; mtDNA: Φ_{CT} = 0.662, P < 0.001). Two of the Mediterranean island populations exhibited the strongest divergence (Cyprus *vs.* all. F_{ST} = 0.13–0.28, Sardinia *vs.* all, F_{ST} = 0.10–0.26; Table S2). For Cyprus, this was apparent not only in terms of high F_{ST} values, but also by a considerably high accumulation of private alleles despite the small sample size (Table 2). In comparison, Φ_{ST} values inferred from the mtDNA data set (control-region sequences) were much higher than F_{ST} values from the microsatellite data, but likewise indicated the strongest divergence between island and continental populations (Cyprus vs. all. Φ_{ST} = 0.58–0.99, Sardinia vs. all, Φ_{ST} = 0.41–1.00).

The continental populations can be considered as two population clusters (Fig. 4). North-eastern Eurasian populations (Russia and Fennoscandia) are divergent from the central

and south-western European populations (Fig. 4). To give a general picture, pairwise F_{ST} values within these two groups were lower than values between populations of both groups (Table S2). Maximum divergence between the two continental groups was observed between southern European populations (Pyrenees, Greece) and north-eastern Palearctic populations (Russian Far East, Central Asia; F_{ST} = 0.075–0.106; Table S2). Lowest divergence was observed between German populations from the zone of overlap and all other continental Eurasian populations (F_{ST} = 0.00–0.088; Table S2).

For the entire set of populations, the ratio between male to female gene flow (m_m/m_f) was estimated based on AMOVA performed for microsatellite data (F_{ST} = 0.065) and mtDNA ($F_{ST(f)}$ = 0.667). These estimates suggest that male gene flow contributes less to the total divergence ($F_{ST(m)}$ = 0.13) than female gene flow and m_m/m_f was 13.41 When we limited the analysis to Eurasian continental populations (under exclusion of island populations) the ratio slightly increased to 19.96 with F_{ST} = 0.029 (microsatellites), divergence caused by female gene flow $F_{ST(f)}$ = 0.556 (mtDNA) and divergence caused by male gene flow $F_{ST(m)}$ = 0.059 (microsatellites).

Bayesian inference of population structure

For the complete data set (n = 166) under the admixture – frequency-correlated model, Evanno's ΔK separated two large clusters (K = 2) as the most plausible population structure (Fig. S1). There was a high level of admixture between these two groups all across Europe and only few populations appeared to be pure representatives of either of the two clusters: i) Central Asian and Far East Russian populations of the north-eastern cluster and ii) the island population from Sardinia of the south-western cluster. All continental European study populations included a high number of genetically admixed individuals. Using LOCPRIOR (classification according to mitochondrial lineages), a population structure with four genetic groups (K = 4) resulted as the most plausible situation (Fig. S1). Likewise for the reduced

population dataset (n = 145 for 14 local populations) four genetic clusters (K = 4) were identified by the ΔK method to be the most plausible scenario, independent of the model applied (with and without LOCPRIOR; Figs 5, S1). The spatial differentiation pattern under K = 4 was as follows: In the West Palearctic, four groups were distinguished (Fig. 5): i) The north-eastern Palearctic cluster including Central Asian, Far East Russian and Fennoscandian populations, ii) the south-western Palearctic cluster including all central and southern continental European populations, iii) the Mediterranean cluster comprising of populations from Corsica and Sardinia (Fig. 5) and iv) the Cypriot population. No signs of genetic admixture were found in Far East Russia, on Sardinia and on Cyprus (Fig. 5). Local genetic admixture was found all across Europe: i) introgression of southern European alleles into Fennoscandian populations, ii) introgression of north-eastern Palearctic alleles into German populations, but (near) absence of north-eastern alleles in the Pyrenean and Greek populations (Fig. 5), iii) wide admixture of two southern clusters (continental: orange; Mediterranean islands: yellow) and introgression of continental European alleles into the Corsican island population (but not into the Sardinian population; Fig. 5).

Spatial clustering studied with GENELAND identified three clusters which exactly matched the phylogeographic pattern of three mitochondrial lineages: a north-eastern Palearctic cluster (Russian Far East, Central Asia, Fennoscandia), a south-western Palearctic cluster (central and southern Europe, including Corsica and Sardinia) and Cyprus (Fig. 6). This pattern was identical in nine out of ten model runs. In a single run, Corsica and Sardinia together represented one separate group whereas the north-eastern and south-western coal tits were united in a second cluster (the number of possible clusters in the analysis was fixed at K = 3). Both of the two spatially explicit admixture models that were run in TESS reflected the same population subdivision inferred by GENELAND (Fig. 6, S2). Neither TESS nor GENELAND separated populations from Corsica and Sardinia from the south-western mainland group.

Admixture rate in the hybrid zone

For both the mitochondrial and the microsatellite data set the best fit was for the admixture model (scenario 1, Fig. 2). However, parameter estimates inferred from microsatellite data do not seem reasonably interpretable for yielding extremely recent and unreliable estimates for times of divergence and time since admixture (during the last six centuries). For the mitochondrial data, scenario 1 was the best model with high support (posterior probabilities 0.9740 and 0.9983 for the direct and logistic regression approaches, respectively), also supported by all the used 40 summary statistics. Type I and II errors were small; type I errors were 0.006 and 0.192 and type II errors were 0.011 and 0.049 for the direct and logistic approaches, respectively. Modes of the effective population sizes for Northern Europe (Norway and Finland) were 280 000 (95 % HPD = 105 000–940 000), for the hybrid zone 9 820 000 (95 % HPD = 3 120 000–9 910 000) and for the Southern Europe 595 000 (95 % HPD = 222 000–981 000). The admixture was estimated to have occurred 57 600 generations ago (95 % HPD = 12 400–95 900) with an admixture rate of 0.216 (95 % HPD = 0.079–0.422) relative to the northern population and divergence of southern and northern lineages 1 260 000 (95 % HPD = 521 000–5 710 000) generations ago. Applying a mean generation time of two years, our estimates correspond to a mean time of divergence of 2.52 mya (95% HPD 1.04–11.42 mya) and a mean time since admixture of 0.114 mya (95% HPD 0.024–0.192 mya).

Discussion

Patterns of gene flow in Europe

Despite considerable genetic differences, the European range of secondary overlap among south-western and north-eastern coal tit lineages does not match the general pattern of a narrow and geographically restricted secondary contact zone (Fig. 1A; compare Haffer, 1989;

Aliabadian, 2005). All German populations appeared to be genetically strongly admixed and introgression of south-western alleles extended northward into Fennoscandian populations, where corresponding (south-western) mtDNA haplotypes were absent in our study (but see Johnsen *et al*., 2010). A moderate deficit of heterozygotes and linkage disequilibrium in three German populations (strongest in Schleswig-Holstein) are indicative of local admixture of two diverged genetic lineages and are typically found in populations from the centre of a hybrid zone (Jiggins & Mallet, 2000; Alexandrino *et al*., 2005; Brelsford & Iwrin, 2009). The actual eastward and southward extent of the contact zone is far from being fully described, because to a lesser degree the northwestern alleles and haplotypes were present in southern European populations (Greece).

Despite all limitations of our sampling, northward allelic introgression and strong differences between pairwise Φ_{ST} values and F_{ST} values hint to wider spatial extent of nuclear gene flow as compared to mtDNA introgression. Such mito-nuclear discordance, can arise from selection against hybrids and/or sterility of the F_1 heterogametic sex (Haldane's rule: Davies & Pomiankowski, 1995; Wu *et al*., 1996) as suggested to be the case in other passerine hybrid zones (European *Ficedula* flycatchers: Tegelström & Gelter, 1990; Sætre *et al*., 2001; Qvarnström *et al*., 2010; Far East Russian great tits: Kvist & Rytkönen, 2006). However, in coal tits there was no evidence of hybrid sterility or selection against hybrids from cross-fostering experiments with individuals from European and Afghan populations (Löhrl, 1994). Therefore, with given certainty we can rule out selection against hybrids in admixed European coal tit populations, too.

Secondly, sex-biased dispersal is considered as another possible cause of mito-nuclear discordance (reviewed by Prugnolle & de Meus, 2002; in birds: Kvist & Rytkönen, 2006; Illera *et al*., 2011; Lin *et al*., 2011). Though the common paradigm of female-biased dispersal in birds (Clarke *et al*., 1997; Petit & Excoffier, 2009) has recently been challenged (Li & Merilä, 2010; Both *et al*., 2012; Dobson, 2013), there is only very scarce information on sex-

specific dispersal distances for many bird species, including the coal tit (except Dietrich *et al.*, 2003). Because of such data deficiency further evidence from field studies is required to substantiate assumptions on any putative correlation between sex-biased dispersal and mito-nuclear discordance in coal tits.

Thirdly and lastly, extreme ratios between male to female gene flow might arise from stochastic effects when comparing different levels of genetic diversity, e.g. high allelic variation of microsatellite loci and sequence variation between deeply divergent lineages (Karl *et al.*, 2014; Putman & Carbone, 2014). Due to relatively long coalescence times, incomplete lineage sorting of nuclear markers might blur spatial patterns of genetic variation. In the coal tit, this is reflected by strong admixture of two southern European allelic clusters (yellow and orange for K = 4) in continental populations on the one hand and a near-complete allelic lineage sorting in island populations of Corsica and Sardinia on the other hand. This is in accordance with low parameters of genetic variation on these islands and with the general assumption that density-dependent processes, such as founder effects and genetic drift, are most effective in island populations (Waters et al. 2013; birds: Padilla *et al.*, 2015). Even during short evolutionary time spans, fast lineage sorting derived from ancestral polymorphisms in founder populations can occur in organisms with high dispersal ability, as inferred from a comparison of historical and extant Mediterranean populations of hawkmoths (*Hyles*; Mende & Hundsdörfer, 2013).

Genetic admixture on the European continent

Extant phylogeographic patterns and lineage diversification in the coal tit are likely to have emerged from glacial range fragmentation (Martens *et al.*, 2006; Pentzold *et al.*, 2013) as suggested for other tit species (Kvist *et al.*, 2003, 2005; Päckert *et al.*, 2013; Stervander *et al.*, 2015; Tritsch *et al.*, 2017). Our time estimates inferred from the mitochondrial data set support a scenario of lineage divergence close to the Pliocene-Pleistocene boundary at a mean

time of divergence of 2.5 mya (in accordance with Päckert *et al.*, 2012). Our mean coalescence-based estimate for time since admixture of 0.114 mya pre-dates a Holocene post-glacial expansion and thus suggest that admixture of north-eastern and south-western gene pools could have already started in southern refuges during the late Pleistocene. This assumption is supported by sound evidence that northward dispersal of forest birds from Mediterranean refuges already started before the onset of the Holocene, because fossil remains of forest birds from interglacial periods have been found north of the Alps across Central Europe up to a latitude of 50° N (Holm & Svenning, 2014). Furthermore, mean coalescence time estimates are rather rough estimates, because there is no reliable empirical value of coal-tit generation time and several authors have applied shorter generation times for tits (Garant *et al.*, 2005; Qu *et al.*, 2015) that would shift our time since admixture estimates closer to a Holocene expansion scenario.

The wide range of mitochondrial introgression and nuclear gene flow in central European coal tits is indicative of a partial reversal of Pleistocene divergence patterns (for a similar case in North American chickadees compare Manthey et al. 2012). The phylogeographic pattern in continental European coal tits matches a broad trans-European zone of intergradation at the subspecies level similarly to e.g. in Eurasian nuthatches, *Sitta europaea* (Red'kin & Konovalova, 2006). Unlike in the latter species, phenotypical variation of continental European coal tits is very subtle and body size parameters and plumage coloration vary along a pan-European cline with phenotypical extreme forms *vieirae* and *abietum* in the South and *ater* in the North (Wolters, 1968; Niethammer, 1943; Glutz von Blotzheim & Bauer, 1993; Martens, 2012). Furthermore, the vocal repertoire of coal tits is remarkably uniform throughout continental Eurasia and seems to provide a less effective premating barrier compared to songs of other tit species (Thielcke, 1973; Tietze *et al.*, 2011; Pentzold *et al.*, 2016). In contrast, in many cases of asymmetric gene flow across narrow hybrid zones among Holarctic bird taxa (regardless of their taxonomic rank), assortative

mating seems to be associated with strong divergence of vocal repertoires (Haavie *et al.*, 2004; Helbig *et al.*, 2001; Päckert *et al.*, 2005; Kvist & Rytkönen, 2006; Sattler *et al.*, 2007; Vorkurková *et al.*, 2013; Shipilina *et al.*, 2017). Generally, separation of gene pools is strongly enhanced by the variation of morphological and behavioural traits that play a key role for species recognition, e.g. in *Ficedula* flycatchers (Sætre *et al.*, 2003; Backström *et al.*, 2013; Ellegren *et al.*, 2012). In contrast, it seems that phenotypical and behavioural differentiation between northern and southern European coal tits is too subtle to provide an effective premating barrier.

The same holds true for potential segregation of ecological or climatic niches in secondary contact (in tits and chickadees: Päckert *et al.*, 2005; Zhao *et al.*, 2012; Taylor *et al.*, 2014). Webb *et al.* (2011) pointed out that merging of genetic lineages might be more likely to occur in generalist species having a lower probability of evolving unique adaptations. This argument may apply to the coal tits as well, because despite a strong adaptation to coniferous forests, they exploit a great variety of food resources. In those regions where the species has adapted to deciduous forests, coal tits use a broader range of the tree's canopy and trunk than many other parid species do (Glutz von Blotzheim & Bauer, 1993; Gosler & Clement, 2007). Habitat structure might also have a considerable effect on local population structure, because, in mixed conifer-broadleaved forests of Ussuriland (Far Eastern Russia), population densities of coal tits were estimated 2.5 to 3 times higher compared to pure spruce-fir taiga forests of the upper mountain-forest belt (Nazarenko 1984).

Third and last, there is in fact recent evidence of spatial variation in an adaptive trait of European coal tits. Schmoll & Kleven (2011) found differences in sperm size between coal tits from Norway and Germany, as was reported among European and Afro-Canarian blue tits (*Cyanistes caeruleus* and *C. teneriffae*; Gohli *et al.*, 2014). Whether in blue tits these differences would constitute an effective post-mating barrier, cannot be judged due to a lack of range overlap in the field and missing evidence from experimental studies. In European

coal tits, intraspecific differences in sperm morphology do not seem to effectively prevent gene flow across the European contact zone of coal tits.

Allopatric differentiation on Mediterranean Islands

Typically, in the central areas of a species' range, the degree of gene flow is often high, whereas it is low at the range margins (Kvist *et al*., 2007; Lehtonen *et al*., 2009; Küpper *et al*., 2012; Päckert *et al*., 2013). In widespread Palearctic bird species greatest phylogeographic structure is often observed at the southwestern range margins, e.g. in the Mediterranean and on the southern European peninsulae (revision in Steward *et al*., 2010; birds: Tietze et al. 2011; Brambilla *et al*., 2008). Since genetic drift and lineage sorting are most effective in small isolated populations, genetic distinctiveness of island populations is a common phylogeographic pattern.

At the global scale levels of vertebrate endemism are significantly higher on islands when compared to the same ecoregions on the adjacent mainland (Fa & Funk, 2007; Kier *et al*., 2009). In the coal tit, the population from Cyprus (*cypriotes*) stands out as a genetically and phenotypically distinctive form that dates back to a more ancient (though still Pleistocene) colonization (Pentzold *et al*., 2013). Phylogenetic studies have revealed complex circum-Mediterranean phylogeographic patterns including distinct island lineages on Cyprus (Voelker & Light, 2011) and highly distinctive populations and even endemic species or subspecies. Apart from the famous examples of the extinct megafauna from Cyprus (Hadjisterkotis & Masala, 1995) weak insular endemism has also been postulated for the extant Cypriote herpetofauna (Böhme & Wiedel 1994) and the Cypriote avifauna (Förschler & Randler, 2009; Randler et al., 2012).

Genetic distinctiveness of Corsican and Sardinian coal tit populations was less manifest than that of *P. a. cypriotes*. The shallow genetic divergence of *P. a. sardus* from its continental relatives (see also Tietze *et al*., 2011; Pentzold *et al*., 2013) contrasts the long

evolutionary histories of some Corso-Sardinian faunal elements (reviewed in Ketmaier & Caccone, 2013). Accordingly, rather ancient (pre-Pleistocene) Corso-Sardinian species-level lineages have been found in amphibians and reptiles (Rodríguez *et al*., 2017; Salvi *et al*., 2010, 2017; Fritz *et al*., 2012). In a Corsican endemic frog species, *Discoglossus montalentii*, phylogeographic structure in microallopatry was found even within the island (Bisconti *et al*., 2013). But also in highly mobile vertebrates, such as birds, several endemic species occur on these islands, such as the Corsican nuthatch, *Sitta whiteheadi* (Pasquet *et al*., 2013) and the Corsican finch, *Carduelis corsicana* (Förschler *et al*., 2009), which also breeds on the Balearic Islands and on a few smaller neighbouring islands. In addition, there are distinct genetic lineages at the subspecies level in other bird species (Pons *et al*., 2016).

Subtle genetic admixture of the Corsican population might imply that this island does or did receive more influx from continental populations than the Sardinian population, e.g. due to its closer proximity to the mainland and along a North-South migratory pathway of migrants and/or dispersers. However, a greater number of local samplings from both islands would be needed to reliably confirm this hypothesis. Moreover, dispersal behaviour of coal tits is quite variable (Löhrl, 1974; Glutz von Blotzheim & Bauer, 1993; Gosler & Clement, 2007) and seems to depend on the availability of food resources (Löhrl, 1974; Harrap & Quinn, 1996). On Corsica, the breeding phenology of coal tits has strongly adapted to local food peaks (Blondel *et al*., 1988) and such adaptive processes might effectively have contributed to the fixation of genetic lineages on islands e.g. in Corsican blue tits (*Cyanistes caeruleus ogliastrae*; Porlier *et al*., 2012). Generally, the genetic composition of the Corsican coal tit population might be the result of both incomplete lineage sorting during a short separation time (Pentzold et al., 2013) and recent gene flow from irregular influx of continental vagrant individuals and/or dispersers.

These examples demonstrate that the circum-Mediterranean phylogeographic pattern in the coal tit is partly or often paralleled in other island endemics of the Corso-Sardinian fauna.

Traditionally, phenotypical distinctiveness has been a crucial factor for species delimitation and in fact, distinct genetic lineages of the coal tit in North Africa (*atlas* subspecies group) and on Cyprus (ssp. *cypriotes*) are corroborated by the differences in the plumage coloration (Harrap & Quinn, 1996; Gosler & Clement, 2007) and partially by subtle differences in song (Tietze *et al.*, 2011; Pentzold *et al.*, 2016). A deeper understanding of the range-wide intraspecific differentiation in the coal tit will therefore benefit (i) from an integrative taxonomic approach and (ii) from broad population sampling across gradients of genetic introgression (e.g. in narrow hybrid zones that exist for example in the Himalayas).

Acknowledgements: We are particularly grateful to all colleagues who provided samples for molecular analyses: T. Dolich, M. Fischer (Naturkundemuseum Erfurt, Germany), A. Johnsen (Natural History Museum Oslo, Norway), T. Lubjuhn, P. Lymberakis (Natural History Museum of Crete), S. Martens, A. Ostastshenko, F. Schramm and H. Zang. We are particularly grateful to F. Spina for providing collection permits in Italy. B. Martens helped on various field trips. Travel and research funds for field trips by J.M. were granted by the Deutsche Ornithologen-Gesellschaft and the Gesellschaft für Tropenornithologie, by Feldbausch-Stiftung and Wagner-Stiftung (both University of Mainz, Germany) (various grants from all four). This project was substantially funded by the State Ministry of Science and Arts of Saxony (Staatsministerium für Wissenschaft und Kunst Sachsen, AZ 4-7531.50-02-621-09/1). We would like to thank two anonymous reviewers for their constructive criticism and helpful comments on the manuscript.

References

Alexandrino J, Baird SJE, Lawson L, Macey JR, Moritz C, Wake DB. 2005. Strong selection against hybrids at a hybrid zone in the *Ensatina* ring species complex and its evolutionary implications. *Evolution* **59**: 1334–1347.

Aliabadian M, Roselaar CS, Nijman V, Sluys R, Vences M. 2005. Identifying contact zone hotspots of passerine birds in the Palaearctic region. *Biology Letters* **22**: 21–23.

Avise JC, Walker D. 1998. Pleistocene phylogeographic effects on avian populations and the speciation process. *Proceedings of the Royal Society B: Biological Sciences* **22**: 457–463.

Backström N, Sætre GP, Ellegren H. 2013. Inferring the demographic history of European *Ficedula* flycatcher populations. *BMC Evolutionary Biology* **13**: 2.

Bath S, Amundsen PA, Knudsen R, Gejlland KØ, Fevolden SE, Bernatchez L, Præbel K. 2014. Speciation reversal in European whitefish (*Coregonus lavaretus* (L.)) caused by competitor invasion. *PLoS ONE* **9**: e91208.

Bella JL, Serrano L, Orellana J, Mason PL. 2007. The origin of the *Chorthippus parallelus* hybrid zone: chromosomal evidence of multiple refugia for Iberian populations. *Journal of Evolutionary Biology* **20**: 568–576.

BirdLife International. 2017. Species factsheet: *Periparus ater*. Downloaded from http://www.birdlife.org on 07/04/2017.

Bisconti R, Canestrelli D, Nascetti G. 2013. Has living on islands been so simple? Insights from the insular endemic frog *Discoglossus montalentii*. *PLoS ONE* **8**: e55735.

Block NL, Goodman SM, Hackett SJ, Bates JM, Raherilalao MJ. 2015. Potential merger of ancient lineages in a passerine bird discovered based on evidence from host-specific ectoparasites. *Ecology and Evolution* **5**: 3743–3755.

Blondel J, Chessel D, Frochot B. 1988. Bird species impoverishment, niche expansion and density inflation in Mediterranean island habitats. *Ecology* **69**: 1899–1917.

Böhme W, Wiedel H. 1994. Status and zoogeography of the herpetofauna of Cyprus, with taxonomic and natural history notes on selected species (genera *Rana*, *Coluber*, *Natrix*, *Vipera*). *Zoology in the Middle East* **10**: 31–52.

Bolfíková B, Hulva P. 2012. Microevolution of sympatry: landscape genetics of hedgehogs *Erinaceus europaeus* and *E. roumanicus* in Central Europe. *Heredity* **108**: 248–255.

Both C, Robinson RA,van der Jeugd HP. 2012. Long-distance dispersal in migratory pied flycatchers *Ficedula hypoleuca* is relatively common between the UK and the Netherlands. *Journal of Avian Biology* **43**: 193–197.

Brambilla M, Vitulano S, Spina F, Baccetti N, Gargallo G, Fabbri E, Guidali F, Randi E. 2008. A molecular phylogeny of the *Sylvia cantillans* complex: cryptic species within the Mediterranean basin. *Molecular Phylogenetics and Evolution* **48**: 461–472.

Brelsford A, Irwin DE. 2009. Incipient speciation despite little assortative mating: the yellow-rumped wabler hybrid zone. *Evolution* **63**: 3050–3060.

Chen C, Durand E, Jay F, Francois O. 2007. Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Molecular Ecology Notes* **7**: 747–756.

Clarke A L, Sæther BE, Roskaft E. 1997. Sex biases in avian dispersal: a reappraisal. *Oikos* **79**: 429–438.

Cornuet JM, Ravigné V, Estoup A. 2010. Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). BMC *Bioinformatics* **11**: 401.

Davies N, Pomiankowski A. 1995. Haldane's rule: old theories are the best. *Trend in Ecology and Evolution* **10**: 350–351

Dietrich VCJ, Schmoll T, Winkel W, Lubjuhn T. 2003. Survival to first breeding is not sex-specific in the Coal Tit (*Parus ater*). *Journal für Ornithologie* **144**: 148–156.

Dobson FS. 2013. The enduring question of sex-biased dispersal: Paul J. Greenwood's (1980) seminal contribution. *Animal Behaviour* **85**: 299–304.

Durand E, Jay F, Gaggiotti OE, Francois O. 2009a. Spatial Inference of Admixture Proportions and Secondary Contact Zones. *Molecular Biology and Evolution* **26**: 1963–1973.

Durand E, Chen C, François O. 2009b. Tess version 2.3 - Reference Manual August 2009 (http://membres-timc.imag.fr/Olivier.Francois/manual.pdf)

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

Ellegren H, Smeds L, Burri R, Olason PI, Backström N, Kawakami T, Künster, A, Mäkinen H, Nadachowska-Brzyska K, Qvarnström A, Uebbing S, Wolf JBW. 2012. The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature* **491**: 756–760.

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

Excoffier G, Laval L, Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**: 47–50.

Fa JE, Funk SM. 2007. Global endemicity centres for terrestrial vertebrates: an ecoregions approach. *Endangered Species Research* **3**: 31-42.

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

Falush D, Stephens M, Pritchard JK. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* **7**: 574–578.

Flanagan NS, Mason PL, Gosálvez, Hewitt GM. 1999. Chromosomal differentiation through an alpine hybrid zone in the grasshopper *Chorthippus parallelus*. *Journal of Evolutionary Biology* **12**: 577–585.

Förschler MI, Randler C. 2009. A song analysis of the insular Cyprus Short-toed reecreeper, *Certhia brachydactyla dorotheae*, supports its subspecies status. *Zoology in the Middle East* **46**: 37–40.

Förschler MI, Senar JC, Perret P, Björklund M. 2009. The species status of the Corsican finch *Carduelis corsicana* assessed by three genetic markers with different rates of evolution. *Molecular Phylogenetics and Evolution* **52**: 234–240.

François O, Durand E. 2010. Spatially explicit Bayesian clustering models in population. *Molecular Ecology Resources* **10**: 773–784.

Fritz U, Corti C, Päckert M. 2012. Mitochondrial DNA sequences suggest unexpected phylogenetic position of Corso-Sardinian grass snakes (*Natrix cetti*) and do not support their species status, with notes on phylogeography and subspecies delineation of grass snakes. *Organisms Diversity and Evolution* **12**: 71–80.

Garant D, Kruuk LEB, Wilkin TA, McCleery RH, Sheldon BC. 2005. Evolution driven by differential dispersal within a wild bird population. *Nature* **433**: 60–65.

Gautschi B, Widmer A, Koella J. 2000a. Isolation and characterization of microsatellite loci in the dice snake (*Natrix tessellata*). *Molecular Ecology* **9**: 2192–2193.

Gautschi B, Tenzer I, Müller JP, Schmid B. 2000b. Isolation and characterization of microsatellite loci in the bearded vulture (*Gypaetus barbatus*) and cross-amplification in three old world vulture species. *Molecular Ecology* **9**: 2193–2195.

Glaubitz JC. 2004. CONVERT: A user-friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Molecular Ecology Notes* **4**: 309–310.

Glutz von Blotzheim UN, Bauer KM. 1993. Handbuch der Vögel Mitteleuropas, Band 13/1, Passeriformes (4. Teil) Muscicapidae–Paridae, Aula-Verlag, Wiesbaden

GENELAND documentation: the Geneland development group, August 2012. Population genetic and morphometric data analysis using R and the Geneland program (http://www2.imm.dtu.dk/~gigu/Geneland/Geneland-Doc.pdf)

Giménez MD, Förster DW, Jones EP, Jóhannesdóttir F, Gabriel SI, Panithanarak T, Scascitelli M, Merico V, Garagna S, Searle JB, Hauffe HC. 2017. A half-century of studies on a chromosomal hybrid zone of the house mouse. *Journal of Heredity* **108**: 25–35.

Gohli J, Leder EH, Garcia-Del-Rey E, Johannessen LE, Johnsen A, Laskemoen T, Popp M, Lifjeld JT. 2015. The evolutionary history of Afrocanarian blue tits inferred from genomewide SNPs. *Molecular Ecology* **24**: 180–191.

Gosler AG, Clement P. 2007. Family Paridae (Tits and Chickadees). In del Hoyo, J., Elliott, A. & Christie, D.A. (eds) Handbook of the Birds of the World: Vol. 12: 662–750. Barcelona: Lynx Edicions.

Goudet J. 1995. FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. *Journal of Heredity* **86**: 485–486.

Guillot G, Estoup A, Mortier F, Cosson JF. 2005a. A Spatial Statistical Model for Landscape Genetics. *Genetics* **170**: 1261–1280.

Guillot G, Mortier F, Estoup A. 2005b. GENELAND: a computer package for landscape genetics. *Molecular Ecology Notes* **5**: 712–715.

Guo S W, Thompson EA. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**: 361–372.

668 **Haas F, Pointer MA, Saino N, Brodin A, Mundy NI, Hansson B. 2009.** An analysis of
669 population genetic differentiation and genotype-phenotype association across the
670 hybrid zone of carrion and hooded crows using microsatellites and MC1R. *Molecular*
671 *Ecology* **18**: 294–305.
672 **Haas F, Knape J, Brodin A. 2010.** Habitat preferences and positive assortative mating in an
673 avian hybrid zone. *Journal of Avian Biology* **41**: 237–247.
674 **Haavie J, Sætre GP, Moum T. 2000.** Discrepancies in population differentiation at
675 microsatellites, mitochondrial DNA and plumage colour in the pied flycatcher–
676 inferring evolutionary processes. *Molecular Ecology* **9**: 1137–1148.
677 **Haavie J, Borge T, Bures S, Garamszegi LZ, Lampe HM, Moreno J, Qvarnström A,**
678 **Török J, Sætre GP. 2004.** Flycatcher song in allopatry and sympatry–convergence,
679 divergence and reinforcement. *Journal of Evolutionary Biology* **17**: 227–237.
680 **Haffer J. 1989.** Parapatrische Vogelarten der palaearktischen Region. *Journal für*
681 *Ornithologie* **130**: 475–512.
682 **Hadjisterkotis E, Masala B. 1995.** Vertebrate extinction in Mediterranean islets: an example
683 from Cyprus. *Biogeographia* **18**: 691–699.
684 **Hamilton MB, Pincus EL, Di Fiore A, Fleischer RC. 1999.** Universal linker and ligation
685 procedures for construction of genomic DNA libraries enriched for microsatellites.
686 *BioTechniques* **27**: 500–507.
687 **Hansson B, Ljungqvist ML, Illera J-C, Kvist L. 2014.** Pronounced fixation, strong
688 population differentiation and complex population history in the Canary Islands blue
689 tit subspecies complex. *PLoS ONE* **9**: e90186.
690 **Haring E, Gamauf A, Kryukov A. 2007.** Phylogeographic patterns in widespread corvid
691 birds. *Molecular Phylogenetics and Evolution* **45**: 840–862.
692 **Haring E, Däubl B, Pinsker W, Kryukov A, Gamauf A. 2012.** Genetic divergences and
693 intraspecific variation in corvids of the genus *Corvus* (Aves: Passeriformes: Corvidae)

– a first survey based on museum specimens. *Journal of Zoological Systematics and Evolutionary Research* **50**: 230–246.

Harrap S, Quinn D. 1996. Tits, Nuthatches and Treecreepers. A and C Black, London.

Hedrick PW, Allendorf FW, Baker CS. 2013. Estimation of male gene flow from measures of nuclear and female genetic differentiation. *Journal of Heredity* **104**: 713–717.

Helbig AJ, Salomon M, Bensch S, Seibold I. 2001. Male-biased gene-flow across an avian hybrid zone: evidence from mitochondrial and microsatellite DNA. *Journal of Evolutionary Biology* **14**: 277–287.

Hermansen JS, Sæther SA, Elgvin TO, Borge T, Hjelle E., Sætre GP. 2011. Hybrid speciation in sparrows I: phenotypic intermediacy, genetic admixture and barriers to gene flow. *Molecular Ecology* **20**: 3812–3822.

Hewitt GM. 2000. The genetic legacy of the Quarternary Ice Ages. *Nature* **405**: 907-913.

Hewitt GM. 2004. The structure of biodiversity–insights from molecular phylogeography. *Frontiers in Zoology* **1**: 4.

Hogner S, Sæther SA, Borge T, Bruvik T, Johnsen A, Sætre GP. 2012a. Increased divergence but reduced variation on the Z chromosome relative to autosomes in *Ficedula* flycatchers: differential introgression or the faster-Z effect? *Ecology and Evolution* **2**: 379–396.

Hogner S, Laskemoen T, Lifjeld JT, Porkert J, Kleven O, Albayrak T, Kabasakal T, Johnsen A. 2012b. Deep sympatric mitochondrial divergence without reproductive isolation in the common redstart *Phoenicurus phoenicurus*. *Ecology and Evolution* **2**: 2974–2988.

Holm SR, Svenning J-C. 2014. 180,000 years of climate change in Europe: Avifaunal responses and vegetation implications. *PLoS ONE* **9**: e94021.

718 **Hubisz M, Falush D, Stephens M, Pritchard JK. 2009.** Inferring weak population structure
719 with the assistance of sample group information. *Molecular Ecology Research* **9**:
720 1322–1332.
721 **Hudson AG, Vonlanthen P, Bezault E, Seehausen O. 2013.** Genomic signatures of relaxed
722 disruptive selection associated with speciation reversal in whitefish. *BMC*
723 *Evolutionary Biology* **13**: 108.
724 **Illera JC, Koivula K, Broggi J, Päckert M, Martens J, Kvist L. 2011.** A multi-gene
725 approach reveals a complex evolutionary history in the *Cyanistes* species group.
726 *Molecular Ecology* **20**: 4123–4139.
727 **Irwin DE, Rubtsov AS, Panov EN. 2009.** Mitochondrial introgression and replacement
728 between yellowhammers (*Emberiza citrinella*) and pine buntings (*E. leucocephalos*;
729 Aves, Passeriformes). *Biological Journal of the Linnean Society* **98**: 422–438.
730 **Jakobsson M, Rosenberg NA. 2007.** CLUMPP: a cluster matching and permutation program
731 for dealing with label switching and multimodality in population structure.
732 *Bioinformatics* **23**: 1801–1806.
733 **Jiggins CD, Mallet J. 2000.** Bimodal hybrid zones and speciation. *Trends in Ecology and*
734 *Evolution* **15**: 250–255.
735 **Johnsen A, Rindal E, Ericson PGP, Zuccon D, Kerr KCR, Stoeckle MY, Lifjeld JT.**
736 **2010.** DNA barcoding of Scandinavian birds reveals divergent lineages in trans-
737 Atlantic species. *Journal of Ornithology* **151**: 565–578.
738 **Karl SA, Toonen RJ, Grant WS, Bowen BW. 2012.** Common misconceptions in molecular
739 ecology: echoes of the modern synthesis. *Molecular Ecology* **21**: 4171–4189.
740 **Kawano KM. 2003.** Isolation of polymorphic microsatellite markers in the great tit (*Parus*
741 *major minor*). *Molecular Ecology Notes* **3**: 314–315.
742 **Ketmaier V, Caccone A. 2013.** Twenty Years of Molecular Biogeography in the West
743 Mediterranean Islands of Corsica and Sardinia: Lessons Learnt and Future Prospects.

In Silva-Opps (ed.) Current Progress in Biological Research. InTech, DOI: 10.5772/55458.

Kier G, Kreft H, Lee TM, Jetz W, Ibisch PL, Nowicki C, Mutke J, Barthlott W. 2009. A global assessment of endemism and species richness across island and mainland regions. *Proceedings of the National Academy of Sciences USA* **106**: 9322–9327.

Kleindorfer S. O'Connor JA, Dudaniec RY, Myers SA, Robertson J, Sulloway FJ. 2014. Species Collapse via Hybridization in Darwin's Tree Finches. *The American Naturalist* **183**: 325–341.

Kuhn K, Schwenk K, Both C, Canal D, Johansson U, Töpfer T, Päckert M. 2013. Differentiation in neutral genes and a candidate gene in the pied flycatcher using biological archives to track global climate change. *Ecology and Evolution* **3**: 4799–4814.

Küpper C, Edwards SV, Kosztolányi A, Alrashidi M, Burke T, Herrmann P, Argüelles-Tico A, Amat JA, Amezian M, Rocha A, Hötker H, Ivanov A, Chernicko J, Székely T. 2012. High gene flow on a continental scale in the polyandrous Kentish plover *Charadrius alexandrinus*. *Molecular Ecology* **21**: 5864–5879.

Kvist L, Rytkönen S. 2006. Characterization of a secondary contact zone of the great tit *Parus major* and the Japanese tit *P. minor* (Aves: Passeriformes) in Far Eastern Siberia with DNA markers. *Zootaxa* **1325**: 55–73.

Kvist L, Martens J, Ahola A, Orell M. 2001. Phylogeography of a Palaearctic sedentary passerine, the willow tit (*Parus montanus*). *Journal of Evolutionary Biology* **14**: 930–941.

Kvist L, Martens J, Higuchi H, Nazarenko AA, Valchuk OP, Orell M. 2003. Evolution and genetic structure of the great tit (*Parus major*) complex. *Proceedings of the Royal Society B: Biological Sciences* **270**: 1447–1454.

Kvist L, Arbabi T, Päckert M, Orell M, Martens J. 2007. Population differentiation in the marginal populations of the great tit (Paridae: *Parus major*). *Biological Journal of the Linnean Society* **90**: 201–210.

Lehtonen PK, Laaksonen T, Artemyev AV, Belskii E, BothC, Bures S., Bushuev AV, Krams I, Moreno J, Mägi M, Nord A, Potti J, Ravussin PA, Sirkiä PM, Sætre GP, Primmer CR. 2009. Geographic patterns of genetic differentiation and plumage colour variation are different in the pied flycatcher (*Ficedula hypoleuca*). *Molecular Ecology* **18**: 4463–4476.

Li MH, Merilä J. 2010. Genetic evidence for male-biased dispersal in the Siberian jay (*Perisoreus infaustus*) based on autosomal and Z-chromosomal markers. *Molecular Ecology* **19**: 5281–5295.

Lin F, Jiang P, Ding P. 2011. Genetic evidence for male-biased dispersal in Elliot's Pheasant (*Syrmaticus ellioti*) in China. *Chinese Birds* **2**:72–78.

Löhrl H. 1974. Die Tannenmeise. Die Neue Brehm-Bücherei 472. Ziemsen, Wittenberg.

Löhrl H. 1994. Beitrag zum noch offenen Fragenkomplex der Hybridisation von Tannen- und Schopfmeise (*Parus ater*, *P. melanolophus*) (Aves: Passeriformes: Paridae). *Zoologische Abhandlungen des Museums für Tierkunde Dresden* **48**: 149–157.

Lovette IJ. 2005. Glacial cycles and the tempo of avian speciation. *TRENDS in Ecology and Evolution* **20**: 57–59.

Macholán M, Baird SJE, Munclinger P, Dufková P, Bímová B, Piálek J. 2008. Genetic conflict outweighs heterogametic incompatibility in the mouse hybrid zone? *BMC Evolutionary Biology* **8**: 271.

Manthey JD, Klicka J, Spellman GM. 2012. Is Gene Flow Promoting the Reversal of Pleistocene Divergence in the Mountain Chickadee (*Poecile gambeli*)? *PLoS ONE* **7**: e49218.

Martens J. 2012. Arten und Unterarten im Spannungsfeld aktueller ornithologischer Systematik. *Anzeiger des Vereins Thüringer Ornithologen* **7**: 135–170.

Martens J, Tietze DT, Sun YH. 2006. Molecular phylogeny of *Parus* (*Periparus*), a Eurasian radiation of tits (Aves: Passeriformes: Paridae). *Zoologische Abhandlungen des Museums für Tierkunde Dresden* **55**: 103–120.

Mende M, Hundsdörfer AK. 2013. Mitochondrial lineage sorting in action – historical biogeography of the *Hyles euphorbiae* complex (Sphingidae, Lepidoptera) in Italy. *BMC Evolutionary Biology* **13**: 83.

Milá B, Surget-Groba Y, Heulin B, Gosá A, Fitze PS. 2013. Multilocus phylogeography of the common lizard *Zootoca vivipara* at the Ibero-Pyrenean suture zone reveals lowland barriers and high-elevation introgression. *BMC Evolutionary Biology* **13**: 192.

Nazarenko AA. 1984. Bird populations in the mixed conifer-broadleaved, and in the dark conifer (fir and spruce) forests of South Primorye Territory, 1962-1971. In: Kostenko VA, Nazarenko AA, eds. *Faunistics and Biology of Birds in the South Far East*. Vladivostok , Far East Sci. Centre, Acad. of Sci. of the USSR, 60–70. (In Russian, with English summary).

Niethammer G. 1943. Beiträge zur Kenntnis der Brutvögel des Peloponnes. *Journal für Ornithologie* **91**: 167–238.

Päckert M, Martens J, Nazarenko AA, Valchuk O, Petri B, Eck S, Veith M. 2005. The great tit, *Parus major*, a misclassified ring species. *Biological Journal of the Linnean Society* **86**: 153–174.

Päckert M, Martens J, Sun YH. 2010. Phylogeny of long-tailed tits and allies inferred from mitochondrial and nuclear markers (Aves: Passeriformes, Aegithalidae). *Molecular Phylogenetics and Evolution* **55**: 952-967.

Päckert M, Martens J, Sun YH, Severinghaus LL, Nazarenko AA, Ting J, Töpfer T, Tietze DT. 2012. Horizontal and elevational phylogeographic patterns of Himalayan

and Southeast Asian forest passerines (Aves: Passeriformes). *Journal of Biogeography* **39**: 556–573.

Päckert M, Martens J, Hering J, Kvist L, Illera JC. 2013. Return flight to the Canary Islands – the key role of peripheral populations of Afrocanarian blue tits (Aves: *Cyanistes teneriffae*) in multi-gene reconstructions of colonization pathways. *Molecular Phylogenetics and Evolution* **67**: 458–467.

Padilla DT, Spurgin LG, Fairfield EA, Illera JC, Richardson DS. 2015. Population history, gene flow, and bottlenecks in island populations of a secondary seed disperser, the southern grey shrike (*Lanius meridionalis koenigi*). *Ecology and Evolution* **5**: 36–45.

Palumbi SR, Cipriano F, Hare MP. 2001. Predicting nuclear gene coalescence from mitochondrial data: the three-times-rule. *Evolution* **55**: 859-868.

Pasquet E, Barker FK, Martens J, Tillier A, Cruaud C, Cibois A. 2014. Evolution within the nuthatches (Sittidae: Aves, Passeriformes): molecular phylogeny, biogeography, and ecological perspectives. *Journal of Ornithology* **155**: 755–765.

Pentzold S, Tritsch C, Martens J, Tietze DT, Giacalone G, LoValvo M, Nazarenko AA, Kvist L, Päckert M. 2013. Where is the line? Phylogeography and secondary contact of western Palearctic coal tits (*Periparus ater*: Aves, Passeriformes, Paridae). *Zoologischer Anzeiger* **252**: 367–382.

Pentzold S, Förschler MI, Tietze DT, Randler C, Martens J, Päckert M. 2016. Geographic variation in coal tit song across continents and reduced species recognition between Central European and Mediterranean populations. *Vertebrate Zoology* **66**: 191–199.

Petit RJ, Excoffier L. 2009. Gene flow and species delimitation. *Trends in Ecology & Evolution* **24**: 386–393.

Pfäffle M, Bolfíková BČ, Hulva P, Petney T. 2014. Different parasite faunas in sympatric populations of sister hedgehog species in a secondary contact zone. *PLoS ONE* **9**: e114030.

Poelstra JW, Vijay N, Bossu CM, Lantz H, Ryll B, Müller I, Baglione V, Unneberg P, Wikelski M, Grabherr MG, Wolf JB. 2014a. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* **20**: 1410–1414.

Poelstra JW, Ellegren H, Wolf JBW. 2014b. An extensive candidate gene approach to speciation: diversity, divergence and linkage disequilibrium in candidate pigmentation genes across the European crow hybrid zone. *Heredity* **111**: 467–473.

Pons JM, Olioso G, Cruaud C, Fuchs J. 2011. Phylogeography of the Eurasian green woodpecker (*Picus viridis*). *Journal of Biogeography* **38**: 311–328.

Pons JM, Thibault JC, Aymí R, Grussu M, Muntaner J, Olioso G, Sunyer JR, Touihri & Fuchs J. 2016. The role western Mediterranean islands in the evolutionary diversifications of the spotted flycatcher *Muscicapa striata*, a long-distance migratory passerine species. *Journal of Avian Biology* **47**: 386–398.

Porlier M, Garant D, Perret P, Charmantier A. 2012. Habitat-linked population genetic differentiation in the blue tit *Cyanistes caeruleus*. *Journal of Heredity* **103**: 781–791.

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

Pritchard J K, Wen X, Falush D. 2009. Documentation for *structure* software: Version 2.3

Prugnolle F, de Meeus T. 2002. Inferring sex-biased dispersal from population genetic tools: a review. *Heredity* **88**: 161–165.

Putman A I, Carbone I. 2014. Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecology and Evolution* **4**: 4399–4428.

QuY, Tian S, Han N, Zhao H, Gao B, Fu J, Cheng Y, Song G, Ericson PGP, Zhang YE, Wang D, Quan Q, Jiang Z, Li R, Lei F. 2015. Genetic responses to seasonal

variation in altitudinal stress: whole-genome resequencing of great tit in eastern Himalayas. *Scientific Reports* **5**: 14256.

Qvarnström A, Rice AM, Ellegren H. 2010. Speciation in *Ficedula* flycatchers. *Philosophical Transactions of the Royal Society B* **365**: 1841–1852.

Randi E. 2008. Detecting hybridization and their domesticated relatives. *Molecular Ecology* **17**: 285–293.

Randler C, Förschler MI, Aliabadian M, Gonzalez J, Bairlein F, Wink M. 2012. Phylogeography, taxonomic status and pre-zygotic isolation in the endemic Cyprus wheatear *Oenanthe cypriaca*. *Journal of Ornithology* **153**: 303–312.

Raymond M, Rousset F. 1995. GENEPOP Version 1.2: population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**: 248–249.

Red'kin Y, Konovalova M. 2006. Systematic notes on Asian birds. 63. The eastern Asiatic races of *Sitta europaea*. *Zoologische Medelingen Leiden* **80**: 241–261.

Reifová R, Majerová V, Reif J, Ahola M, Lindholm A, Procházka P. 2016. Patterns of gene flow and selection across multiple species of *Acrocephalus* warblers: footprints of parallel selection on the Z chromosome. *BMC Evolutionary Biology* **16**: 130.

Rice WR. 1989. Analyzing tables of statistical tests. *Evolution* **43**: 223–225.

Rodríguez V, Buades JM, Brown RP, Terrasa B, Pérez-Mellado V, Corti C, Delaugerre M, Castro JA,Picornell A, Ramon MM. 2017. Evolutionary history of *Podarcis tiliguerta* on Corsica and Sardinia. *BMC Evolutionary Biology* **17**: 27.

Rosenberg NA. 2004. DISTRUCT: a program for graphical display of population structure. *Molecular Ecology Notes* **4**: 137–138.

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

Sætre GP, Borge T, Haavie J, Johnsen A. 2001. Speciation, introgressive hybridization and nonlinear rate of molecular evolution in flycatchers. *Molecular Ecology* **10**:737–749.

SaladinV, BonfilsD., BinzT, RichnerH. 2003. Isolation and characterization of 16 microsatellite loci in the Great Tit *Parus major*. *Molecular Ecology Notes* **3**: 520–522.

Salvi D, Harris DJ, Bombi P, Carretero MA, Bologna MA. 2010. Mitochondrial phylogeography of the Bedriaga's rock lizard, *Archaeolacerta bedriagae* (Reptilia: Lacertidae) endemic to Corsica and Sardinia. *Molecular Phylogenetics and Evolution* **56**: 690-697.

Salvi D, Pinho C, Harris DJ. 2017. Digging up the roots of an insular hotspot of genetic diversity: decoupled mito-nuclear histories in the evolution of the Corsican-Sardinian endemic lizard *Podarcis tiliguerta*. *BMC Evolutionary Biology* **17**: 63.

Sattler GD, Sawaya P, Braun MJ. 2007. An assessment of song admixture as an indicator of hybridization in black-capped chickadees (*Poecile atricapillus*) and Carolina chickadees (*P. carolinensis*). *The Auk* **124**: 926–944.

Schmoll T, Kleven, O. 2011. Sperm dimensions differ between two Coal Tit *Periparus ater* populations. *Journal of Ornithology* **152**: 515–520.

Schmitt T.2007. Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology* **4**: 11.

Schuelke M. 2000. An economic method for the fluorescent labelling of PCR fragments. *Nature Biotechnology* **18**: 233–234.

Secondi J, Bordas P, Hipsley CA, Bensch S. 2011. Bilateral song convergence in a passerine hybrid zone: genetics contribute in one species only. *Journal of Evolutionary Biology* **38**: 441–452.

Seehausen O, Takimoto G, Roy D, Jokela J. 2008. Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Molecular Ecology* **17**: 30–44.

Shipilina D, Serbyn M, Ivanitskii, Marova I, Backström N. 2017. Patterns of genetic, phenotypic, and acoustic variation across a chiffchaff (*Phylloscopus collybita abietinus/tristis*) hybrid zone. *Ecology and Evolution* **7**: 2169–2180.

Shuker DM, King T, Bella JL, Butlin RK. 2005. The genetic basis of speciation in a grasshopper hybrid zone. In: Fellowes MDE, Holloway GJ, Rolff J. (eds) Insect evolutionary ecology: Proceedings of the Royal Entomological Society's 22nd Symposium. CABI Publishing.

Stervander M, Illera JC, Kvist L, Barbosa P, Keehnen NP, Pruisscher P, Bensch S, Hansson B. 2015. Disentangling the complex evolutionary history of the Western Palearctic blue tits (*Cyanistes* spp.)–phylogenomic analyses suggest radiation by multiple colonization events and subsequent isolation. *Molecular Ecology* **24**: 2477–2494.

Stewart JR, Lister AM, Barnes I, Dale L. 2010. Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B* **277**: 661–671.

Sutter A, Beysard M, Heckel G. 2013. Sex-specific clines support incipient speciation in a common European mammal. *Heredity* **110**: 398–404.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEAG6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30**: 2725–2729.

Taylor EB, Boughman JW, Groenenboom M, Sniatynski M, Schluter D, Gow JL. 2006. Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Molecular Ecology* **15**: 343–55.

Taylor SA, White TA, Hochachka WM, Ferretti V,Curry RL, Lovette I. 2014. Climate-Mediated Movement of an Avian Hybrid Zone. *Current Biology* **24**: 671–676.

Tegelström H, Gelter HP. 1990. Haldane's rule and sex-biased gene flow between two hybridizing flycatcher species (*Ficedula albicollis* and *F. hypoleuca*, Aves: Muscicapidae). *Evolution* **44**: 2012–2021.

Tietze DT, Martens J, Sun YH, Severinghaus LL, Päckert M. 2011. Song evolution in the coal tit (*Parus ater*). *Journal of Avian Biology* **42**: 214–230.

Thielcke G. 1973. Uniformierung des Gesangs der Tannenmeise (*Parus ater*) durch Lernen. *Journal für Ornithologie* **114**: 443–454.

Toonen RJ, Hughes S. 2001. Increased Throughput for Fragment Analysis on ABI Prism 377 Automated Sequencer Using a Membrane Comb and STRand Software, *Biotechniques* **31**: 1320–1324.

Tritsch C, Martens J, Sun YH, HeimW, Strutzenberger P, Päckert M. 2017. Improved sampling at the subspecies level solves a taxonomic dilemma–A case study of two enigmatic Chinese tit species (Aves, Passeriformes, Paridae, *Poecile*). *Molecular Phylogenetics and Evolution* **107**: 538–550.

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

Vasquez P, Cooper SJ, Gosalvez, Hewitt GM. 1994. Nuclear DNA introgression across a Pyrenean hybrid zone between parapatric subspecies of the grasshopper *Chorthippus parallelus*. *Heredity* **73**:436–443.

Voelker G, Light JE. 2011. Palaeoclimatic events, dispersal and migratory losses along the Afro-European axis as drivers of biogeographic distribution in *Sylvia* warblers. *BMC Evolutionary Biology* **11**: 163.

Vokurková J, Petrusková T, Reifová R, Kozman A, Mořkovský L, Kipper S, Weiss M, Reif J, Dolata PT, Petrusek A. 2013. The causes and evolutionary consequences of mixed singing in two hybridizing songbird species (*Luscinia* spp.). *PLoS ONE* **8**: e60172.

Waters JM, Fraser CI, Hewitt GM. 2013. Founder takes all: density-dependent processes structure biodiversity. *Trends in Ecology and Evolution* **28**: 78–85.

Webb WC, Marzluff JM, Omland KE. 2011. Random interbreeding between cryptic lineages of the Common Raven: evidence for speciation in reverse. *Molecular Ecology* **20**: 2390–2402.

Wolf JB, Bayer T, Haubold B, Schilhabel M, Rosenstiel P, Tautz D. 2010. Nucleotide divergence vs. gene expression differentiation: comparative transcriptome sequencing in natural isolates from the carrion crow and its hybrid zone with the hooded crow. *Molecular Ecology* **19**: 162–175.

Wolters HE. 1968. Aus der ornithologischen Sammlung des Museums AlexanderKoenig I. *Bonner Zoologische Beiträge* **19**: 157–164.

Woodruf DS. 1973. Natural hybridization and hybrid zones. *Systematic Biology* **22**: 213-218.

Wu CI, Johnson NA, Palopoli MF. 1996. Haldane's rule and its legacy: why are there so many sterile males? *Trends in Ecology and Evolution* **11**: 281–284.

Zhang R, Song G, Qua Y, Alström P, Ramos R, Xing X, Ericson PGP, Fjeldså J, Wang H, Yang X, Kristin A, Shestopalov AM, Choe JC, Lei F. 2012. Comparative phylogeography of two widespread magpies: Importance of habitat preference and breeding behavior on genetic structure in China. *Molecular Phylogenetics and Evolution* **65**: 562–572.

Zhao N, Dai C, Wang W, Zhang R, Qu Y, Song G, Chen K, Yang X, Zou F, Lei F. 2012. Pleistocene climate changes shaped the divergence and demography of Asian populations of the great tit *Parus major*: evidence from phylogeographic analysis and ecological niche models. *Journal of Avian Biology* **43**: 297–310.

Zink RM, Pavlova A, Drovetski SV, Rohwer S. 2008. Mitochondrial phylogeographies of five widespread Eurasian bird species. *Journal of Ornithology* **149**: 399–413.

Figure captions

Figure 1: Spatial patterns of secondary contact and hybrid zones among divergent genetic lineages (schematized haplotype networks; right) of terrestrial vertebrates in the Western Palearctic; A) narrow contact zone along geographic barriers (Pyrenees, Alps) or of a wide latitudinal North-South extent; B) broad intergradation zone often along phenotypical clines; C) range wide merger and local co-occurrence of distinct genetic lineages (reversal of past lineage divergence); Europe outline map inferred from www.freeworldmaps.net.

Figure 2: Distribution range and phylogeographic pattern of the coal tit, *Periparus ater* (modified from Pentzold et al. 2013; sampling sites of mtDNA data indicated by black dots; range boundaries in light brown according to BirdLife International 2017); pie charts indicate percentages of haplotypes belonging to four different clusters (Scandinavia/Russia, W and SW Europe, North Africa and Cyprus; subspecies included in the mtDNA dataset listed at the corresponding clusters) indicated by different colours; study populations (abbreviations): BF = Black Forest, Cors = Corsica, Cyp = Cyprus, Fin = Finland, Grec = Greece, Mor = Morocco, Nor = Norway, PF = Palatine Forest, Pyr = French Pyrenees, Sard = Sardinia, Sax = Saxony, SH = Schleswig-Holstein (strongly divergent North African subspecies *atlas* not included in our population genetic study); coal tit drawing: K. Rehbinder, University of Mainz.

Figure 3: Historical models used for DIYABC analyses; pop1 = northeastern populations from Finland and Norway; pop2 = populations from the German zone of overlap (Pentzold et al. 2012); pop3 = southwestern population from the French Pyrenees.

Figure 4: UPGMA phenogram inferred from pairwise F_{ST} values (computed with MEGA v.6, and listed in Table S2).

Figure 5: Genetic variation of 14 Western European and Mediterranean coal tit populations (n = 145) based on 13 microsatellite loci; STRUCTURE analysis under the admixture – frequency-correlated model without locpriors a priori defined, STRUCTURE plots for K = 2 to K = 4 (left); threshold q > 0.8 for assignment of individuals to genetic clusters (according to Randi 2008) indicated for the most plausible scenario of K = 4; coloured bars above the plots indicate individual assignment to three mitochondrial lineages (control region; data from Pentzold et al. 2013); grey bars above indicate regional origin of samples; right: a) estimate of most plausible K = 4 according to Evanno et al. (2005; ΔK plotted against the number of modelled genetic clusters) and b) according to L(K) (Prichard2000); abbreviations of populations: BF = Black Forest, Cor = Corsica, Cyp = Cyprus, PF = Palatine Forest, Sard = Sardinia, Sax = Saxony, Schl. Holstein = Schleswig-Holstein,.

Figure 6: Spatial clustering of coal tit populations as inferred from GENELAND analysis; assignment probabilities of the individuals to spatial clusters identified by GENELAND displayed in a contour map for a) the northwestern, b) the south-western and c) the Cypriot cluster respectively. The spatial membership probability is visualized by colour: bright yellow indicates a high-, dark red a low assignment probability; black dots: sampling localities.

Figure S1: Genetic variation in the complete data set of the Western European and Mediterranean coal tits (n = 166) based on 13 microsatellite loci; STRUCTURE analysis under the admixture – frequency-correlated model with locpriors *a priori* defined (assignment according to mtDNA lineages), STRUCTURE plots for K = 2 to K = 4 (left); right: estimate of most plausible K according to Evanno et al. (2005; ΔK plotted against the number of

modelled genetic clusters) a) under the admixture – frequency-correlated model without locpriors *a priori* defined (K = 2) and b) with locpriors *a priori* defined (K= 4); assignment according to mtDNA lineages, control region: coloured bars above the plots.

Figure S2: Spatial clusters as inferred by the spatial explicit CAR admixture model of TESS (Kmax = 3). The TESS admixture models did not distinguish more than three units even if the number of possible clusters in model is higher (Kmax > 3). The DIC criterion (arithmetic mean of ten replicates) confirmed three spatial clusters.

Tables

Table 1: Characteristics and variation of nine newly identified microsatellite loci for population genetic analysis of coal tits. The Parate loci were amplified using the protocol described by Schuelke (2000). For this protocol an 18-bp long M13(-21) tail (5'-TGTAAAACGACGGCCAGT-3') was adhered to the 5' end of the forward primer (= Label); * based on sequence clone; ** based on experimental optimization during microsatellite primer design; *** minimum and maximum sizes based on analyses of 14 populations.

GenBank accession numbers to be provided upon acceptance.

newly designed primers					
Locus	Primer sequence 5'-3'	GenBank accession no	Repeat Motif*	T_a (°C)**	Allele Size (bp)***
Parate 01	F: (Label) TCCTGGAGCACATTATGTCTATG R: AATCTGCTGCTCCATACTTGG	MG970333	$(TAGA)_{14}$ $(CAGA)_4$	56	209 - 271
Parate 02	F: (Label)-AGGGACAGAATTGTGCAAGG R: TGCATTCATGCATACATAGACAC	MG970334	$(ATCT)_{14}$	56	189 - 373
Parate 03	F: Label-TGTTGTCTGCAAAAGGCAAG R: CAAAGCCTTCATCTGCTTGG	MG970335	$(ATAG)_{14}$	56	117 - 177
Parate 06	F: (Label)-TTCAGTGCAGGTGCATAATTG R: GGCCAAGAGAAGTAGGGTGTAG	MG970336	$(CTAT)_{15}$	56	219 - 355
Parate 07	F: (Label)-CTCCCAAGAGAGTCTGTGTCG R: AAGGCTTTTGAAACAGGAGAAG	MG970337	$(CTAT)_{12}$	56	166 - 199
Parate 08	F: (Label)-TTGTAACGACCTTGCACCTC R: AGGCAGTAAAACCCTCATGG	MG970338	$(CA)_{20}$	50	90 - 153
Parate 09	F: (Label)-GGCACAGATGCATATTTTGTTTAC R: TGCACAATCATGCTTAATCCTC	MG970339	$(GT)_{13}$	56	122 - 136
Parate 15	F: (Label)-TCACAAAAAGGCATTTGCAG R: GGAGACAGGAGAGCAGCAAC	MG970340	$(TC)_{12}(C)_4(TC)_7CC(TC)_4$	56	129 - 204
Parate 16	F. (Label)-CTTTCTTGAATGCTCAGATTGC	MG970342	$(CT)_{27}$	56	166 - 263

	R: CAAGCCCATGTTCAAGGTTC				
primers from previous studies					
Pat2-43	F: ACAGGTAGTCAGAAATGGAAAG R: (Label)-GTATCCAGAGTCTTTGCTGATG	- Otter *et al.* (1998)	(CT)n	60	126 - 213
PmaTGAn 33	F: (Label)-TTCCCCAAGTATCCTGCATC R: AAACCATATCACCCAGTGCC	AY260539 Saladin *et al.* (2003)	(GATA)14GAT(GATA)8	57	258 - 398
Pma69u	F: (Label)-CCCAGACAAAGCATCACTGG R: GACAGTTCACATAGCCCTGG	AB094107 Kawano (2003)	(TG)6	57	214 – 222
PmaC25	F: CGTCCTGCTGTTTGTATTTCTG R: (Label)-CCATGAACCATTTTTAGGGTG	AY260526 (Saladin *et al.*, 2003)	(GAT)11	57	313 – 349

Table 2: Diversity parameters and inbreeding coefficient (F_{IS}) for the surveyed population samples of in total 145 individuals. Loci Parate 06 and Parate 08 were excluded due to the presence of null alleles. # = population number, n = number of sampled indifiduals, HO = observed heterozygosity, HE = expected heterozygosity, * Far East = pooled samples from Far East Russia (n = 11) and Japan (n = 2); ** Schleswig-Holstein = pooled samples from Itzehoe (n = 15), Amrum (n = 3) and Brillit (n = 1).

#	Population	n	Mean number of alleles per locus	Mean allelic richness	Private alleles	Mean HO	Mean HE	F_{IS}	p-value (F_{IS})
1	Central Asia	5	4.455	4.089	1	0.777	0.753	-0.036	0.692
2	Far East*	13	7.818	4.717	5	0.825	0.814	-0.015	0.688
3	Finland	11	8.455	4.993	10	0.807	0.832	0.031	0.204
4	Norway	13	8.545	4.992	7	0.769	0.839	0.087	0.016
5	Schleswig-Holstein**	19	9.909	5.026	7	0.740	0.833	0.115	0.0003
6	Harz	12	8.273	4.886	2	0.758	0.822	0.082	0.018
7	Saxony	5	5.364	4.770	4	0.755	0.824	0.094	0.080
8	Palatine Forest	7	7.364	5.237	3	0.831	0.848	0.022	0.373
9	Black Forest	8	7.364	5.175	3	0.849	0.860	0.014	0.336
10	Pyrenees	13	8.636	4.803	4	0.816	0.819	0.003	0.473
11	Greece	12	7.909	4.718	14	0.760	0.807	0.062	0.084
12	Corsica	10	6.273	4.249	6	0.755	0.765	0.015	0.414
13	Sardinia	10	4.818	3.462	1	0.605	0.630	0.042	0.240
14	Cyprus	7	5.182	4.059	13	0.786	0.771	-0.020	0.683

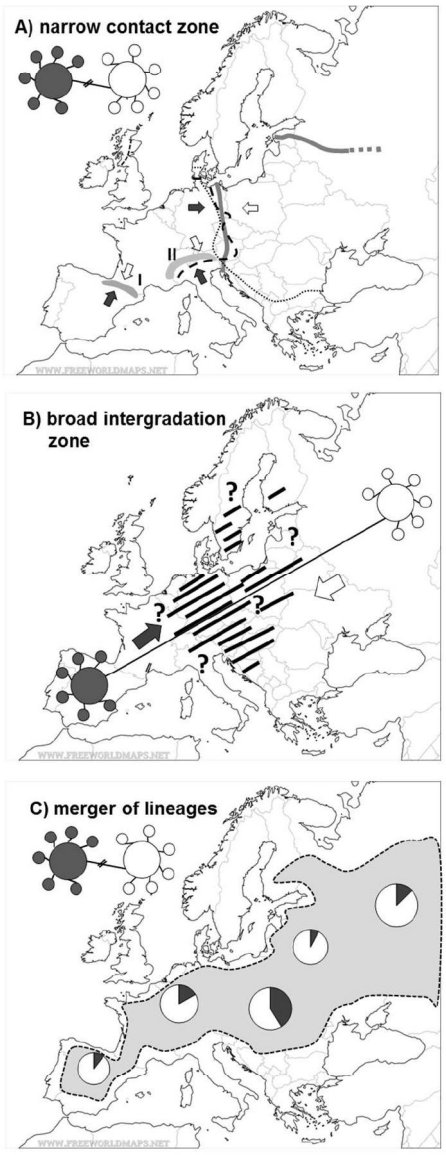


Figure 1

120x268mm (150 x 150 DPI)

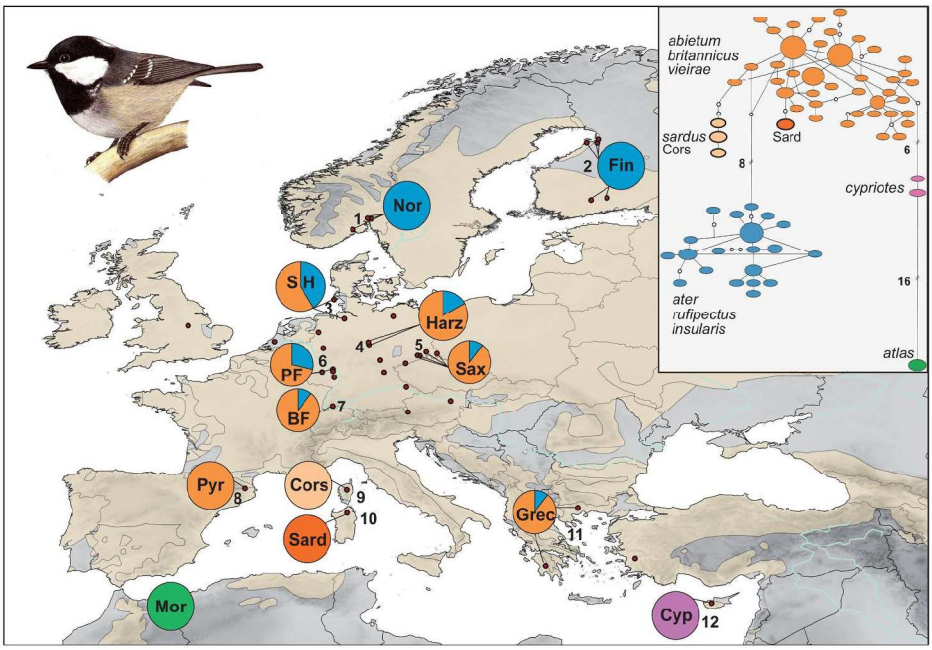
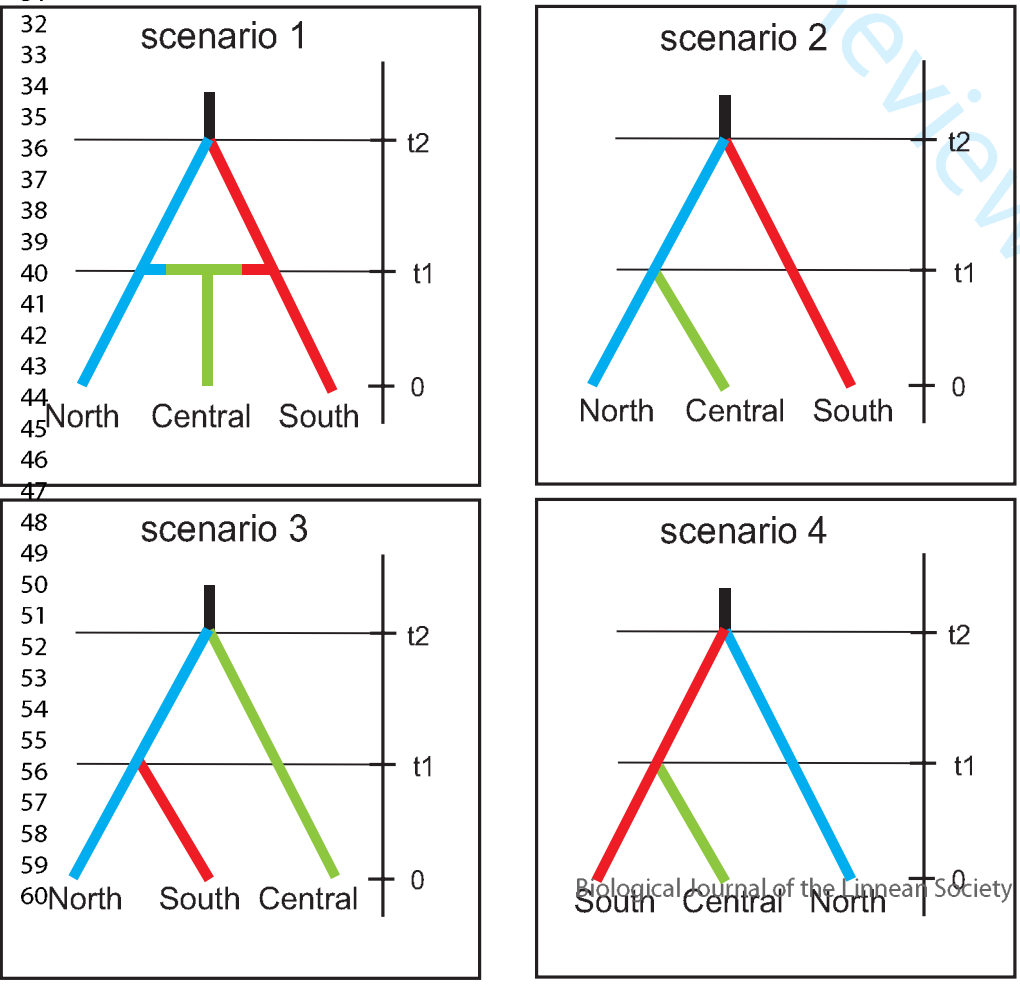
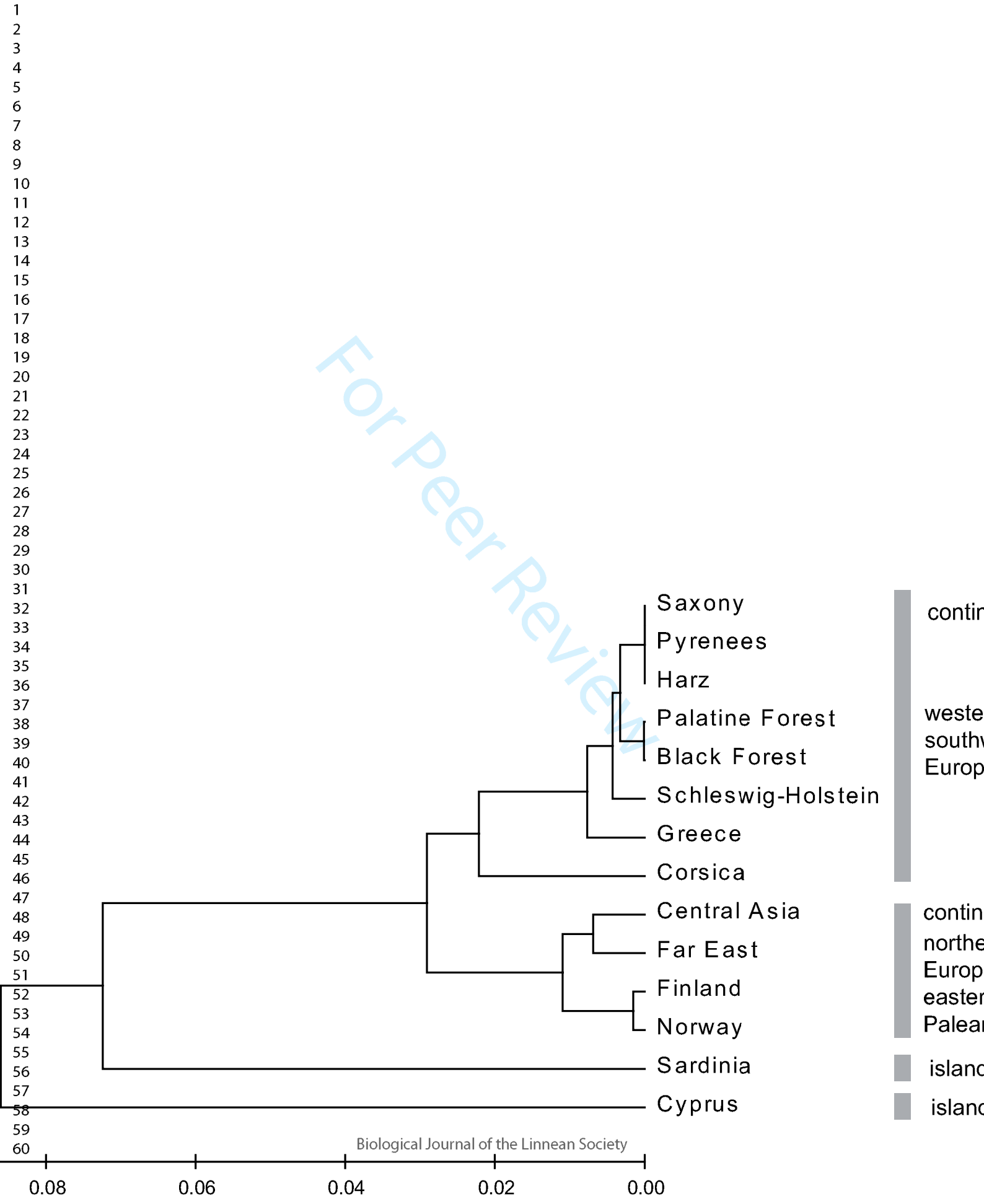
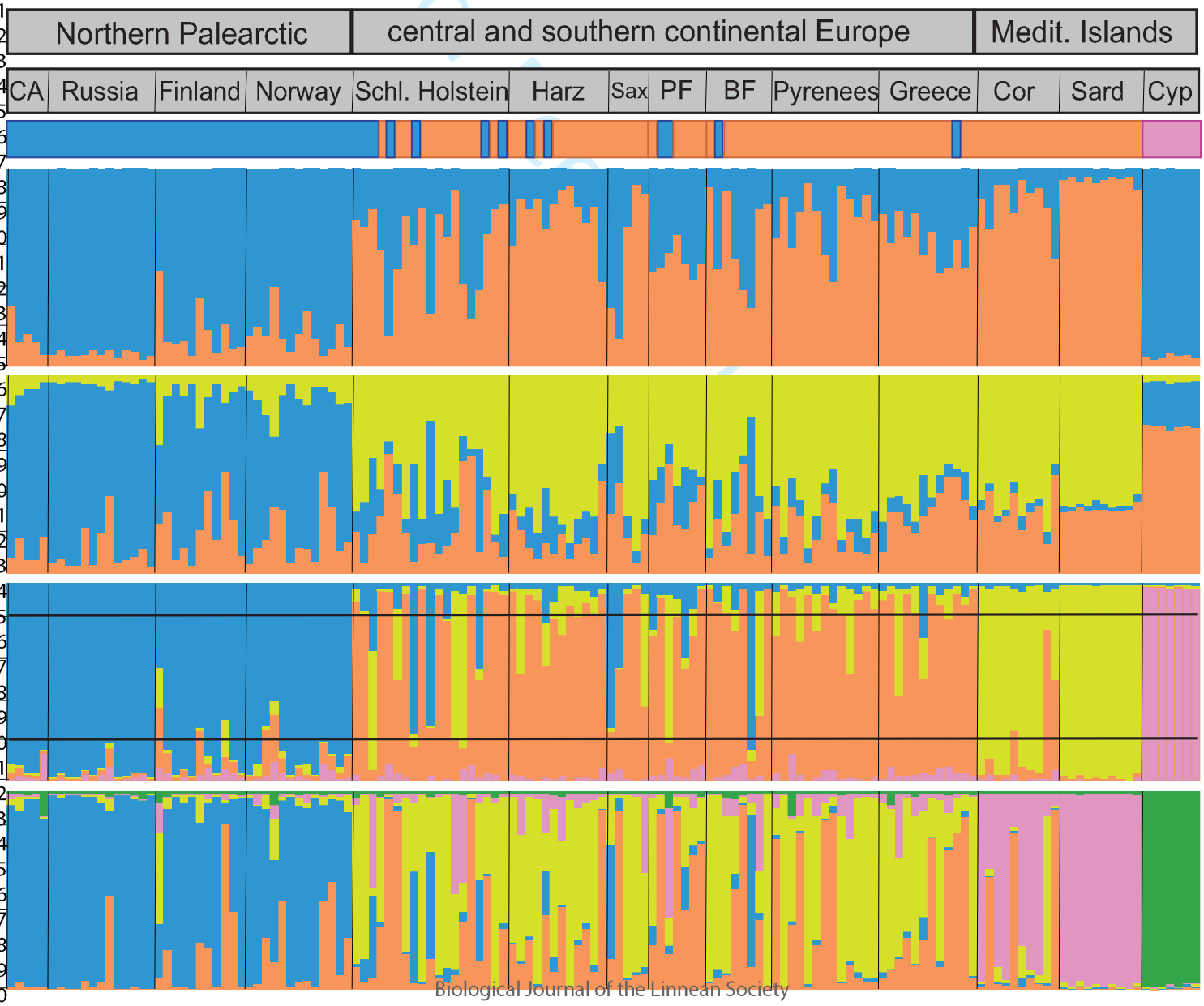


Figure 2

197x139mm (300 x 300 DPI)






Northern Palearctic
central and southern continental Europe
Medit. Islands
CA
Russia
Finland
Norway
Schl. Holstein
Harz
Sax
PF
BF
Pyrenees
Greece
Cor
Sard
Cyp
pop
mtl
K =
K =
K =
K =

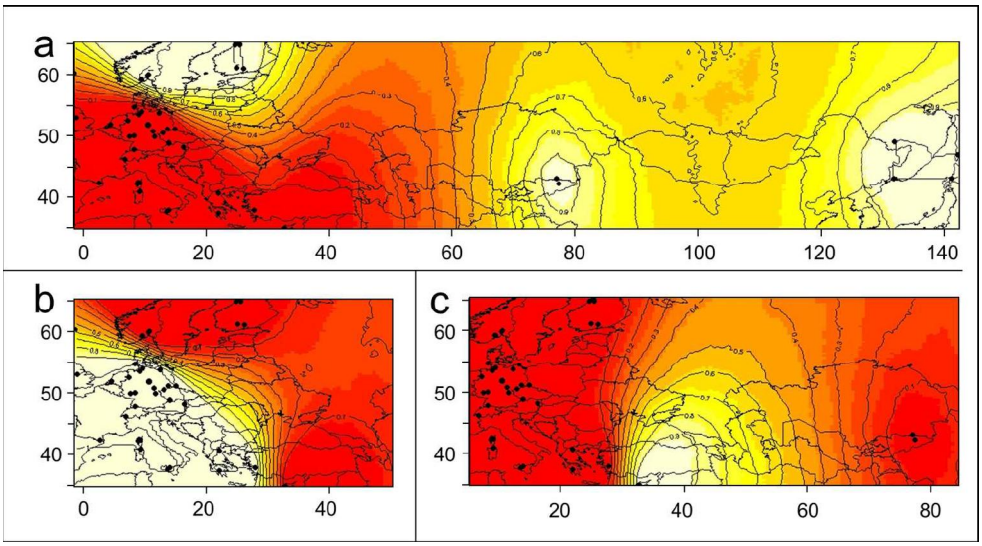
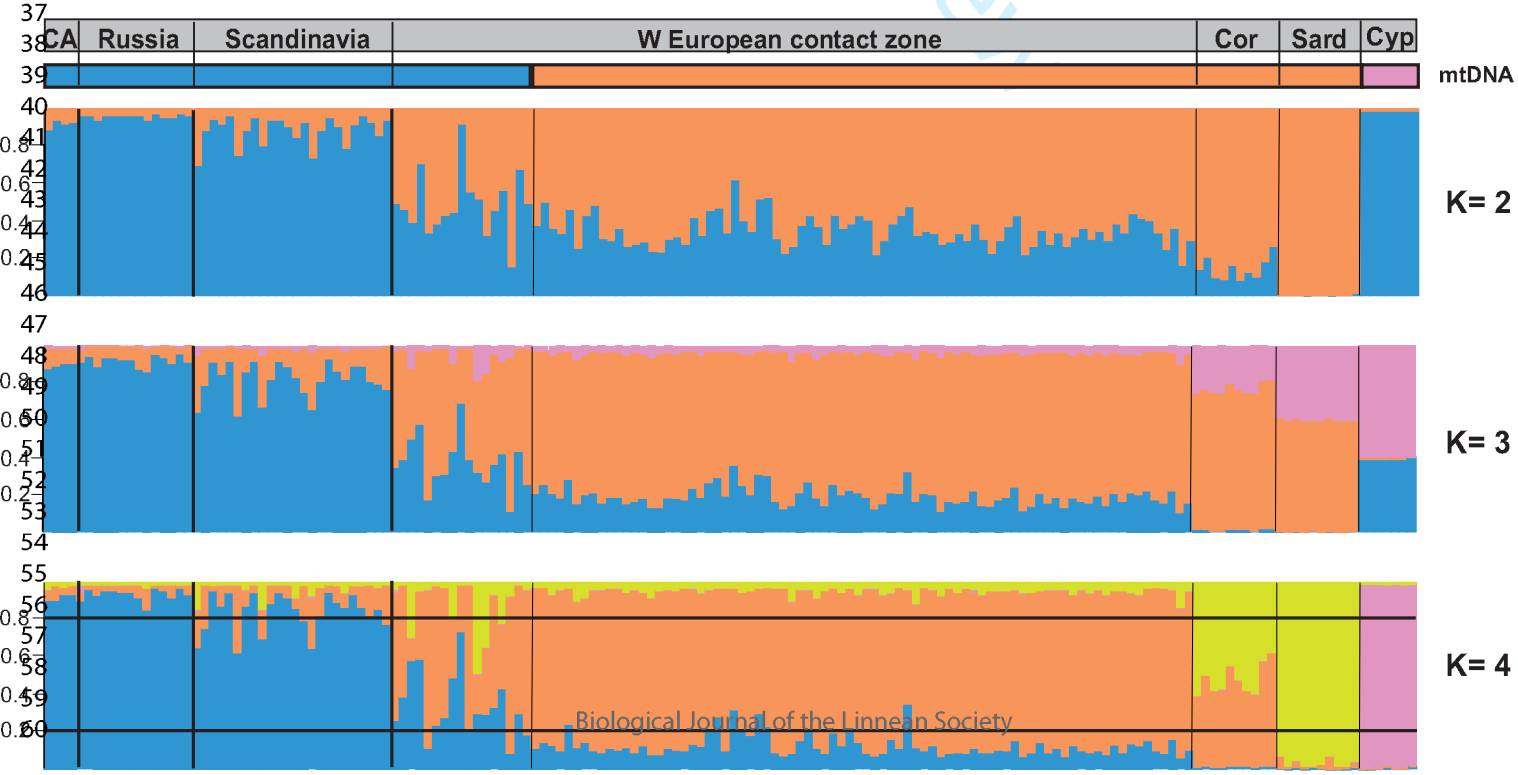
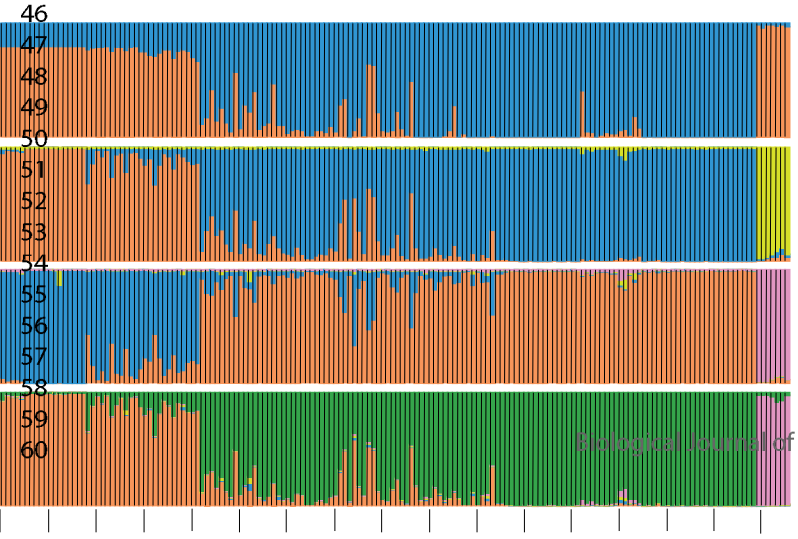


Figure 6

273x150mm (150 x 150 DPI)





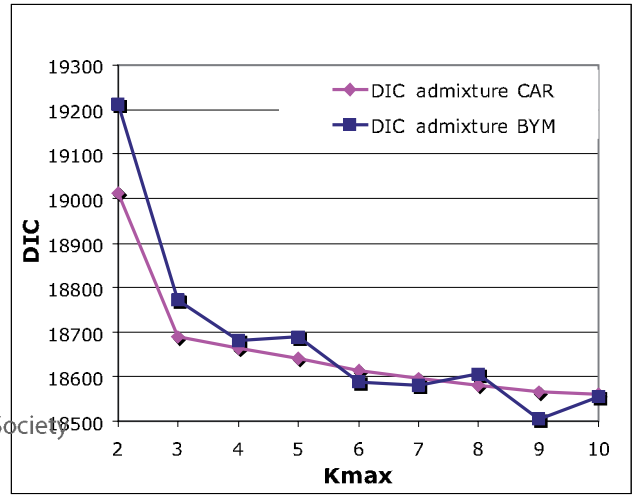


Table S1: Observed and expected heterozygosity, departure from Hardy-Weinberg equilibria, evidence of null alleles and linkage disequilibrium for each locus and population

Bonferroni corrected p-value for HWE = 0.05 / 13 = **0.0038;** Bonferroni corrected p-value for Linkage Disequilibrium = 0.05/78 = **0.00064**

Locus	**Parate 01**	**Parate 02**	**Parate 03**	**Parate 06**	**Parate 07**	**Parate 08**	**Parate 09**	**Parate 15**	**Parate 16**	**Pat2-43**	**PmaTGAn33**	**Pma69**	**PmaC25**
Central Asia (N = 5)													
H_O	1.000	0.800	0.800	0.460	1.000	0.750	0.200	0.800	0.800	0.800	0.750	0.800	0.800
H_E	0,844	0.844	0.822	0.755	0.844	0.893	0.467	0.600	0.867	0.733	0.857	0.622	0.777
HWE p-value	1,000	0.779	0.890	0.028	1.000	0.428	0.333	1.000	0.734	1.000	0.699	1.000	1.000
Null allele bias	no	no	no	no	no	no	no	no	no	no	no	no	no
Linkage disequilibrium	no												
Far East (N = 13)													
H_O	0.692	1.000	0.923	0.461	0.846	0.385	0.769	0.615	0.923	0.923	0.769	0.692	0.923
H_E	0.778	0.932	0.840	0.791	0.849	0.898	0.775	0.757	0.895	0.834	0.834	0.665	0.791
HWE p-value	0.641	0.4998	0.929	**0.0004**	0.2795	**< 0.001**	0.436	0.298	0.126	0.705	0.087	1.000	1.000
Null allele bias	no	no	no	**yes**	no	**yes**	no	no	no	no	no	no	no
Linkage disequilibrium	no												

Locus	Parate 01	Parate 02	Parate 03	Parate 06	Parate 07	Parate 08	Parate 09	Parate 15	Parate 16	Pat2-43	PmaTG An33	Pma69	PmaC25
Finland (N = 11)													
H_O	0.909	0.818	0.909	0.455	1.000	0.545	0.636	0.636	0.800	0.800	0.909	0.636	0.818
H_E	0.896	0.931	0.866	0.857	0.823	0.922	0.645	0.757	0.821	0.884	0.913	0.688	0.922
HWE p-value	0.079	0.491	0.896	**0.002**	0.597	**< 0.001**	0.333	0.2795	0.412	0.525	0.901	0.342	0.532
Null allele bias	no	no	no	**yes**	no	**yes**	no	no	no	no	no	no	no
Linkage disequilibrium	no												
Norway (N = 13)													
H_O	0.769	0.846	0.846	0.769	0.923	0.846	0.692	0.692	0.769	0.923	0.769	0.462	0.769
H_E	0.775	0.898	0.853	0.745	0.843	0.935	0.751	0.877	0.902	0.902	0.883	0.649	0.898
HWE p-value	0.875	0.177	0.171	0.152	0.594	0.057	0.466	0.189	0.009	0.697	0.1397	0.358	**< 0.001**
Null allele bias	no	no	no	no	no	no	no	no	no	no	no	no	no
Linkage disequilibrium	no												

Locus	Parate 01	Parate 02	Parate 03	Parate 06	Parate 07	Parate 08	Parate 09	Parate 15	Parate 16	Pat2-43	PmaTG An33	Pma69	PmaC25
Schleswig-Holstein (N = 19)													
H_O	0.789	0.857	0.722	0.500	0.556	0.500	0.526	0.722	0.583	1.000	0.944	0.667	0.778
H_E	0.771	0.939	0.8396	0.846	0.806	0.867	0.644	0.900	0.873	0.924	0.878	0.722	0.867
HWE P-value	0.617	0.104	0.160	**0.0025**	0.144	0.055	0.012	**<0.001**	**<0.001**	0.969	0.991	0.822	0.4395
Null allele bias	no	no	no	**yes**	**yes**	**yes**	no	**yes**	**yes**	no	no	no	no
Linkage disequilibrium	Parate01, Parate 03, Parate06, Parate 07, Parate12, Parate15												
Harz (N = 12)													
H_O	0.750	0.583	0.917	0.500	0.917	0.583	0.667	0.750	0.667	1.000	0.750	0.583	0.750
H_E	0.786	0.942	0.902	0.833	0.819	0.895	0.746	0.812	0.819	0.948	0.8695	0.670	0.728
HWE p-value	0.665	**< 0.001**	0.212	0.093	0.753	**< 0.001**	0.339	0.498	0.077	1.000	0.189	0.347	0.615
Null allele bias	no	**yes**	no	**yes**	no	**yes**	no	no	no	no	no	no	no
Linkage disequilibrium	no												

Locus	Parate 01	Parate 02	Parate 03	Parate 06	Parate 07	Parate 08	Parate 09	Parate 15	Parate 16	Pat2-43	PmaTG An33	Pma69	PmaC25
Saxony (N = 5)													
H_O	0.800	0.800	1.000	0.800	0.600	0.800	0.400	0.800	0.800	1.00	0.750	0.750	0.600
H_E	0.756	0.933	0.844	0.933	0.800	0.889	0.711	0.844	0.867	0.889	0.857	0.714	0.844
HWE p-value	1.000	0.227	1.000	0.256	0.4998	0.57989	0.144	0.362	0.621	1.000	0.657	1.000	0.386
Null allele bias	no	no	no	no	no	no	no	no	no	no	no	no	no
Linkage equilibrium	no												

Locus	Parate 01	Parate 02	Parate 03	Parate 06	Parate 07	Parate 08	Parate 09	Parate 15	Parate 16	Pat2-43	PmaTG An33	Pma69	PmaC25
Palatine Forest (N = 7)													
H_O	0.857	0.857	0.857	0.286	0.714	0.429	0.714	1.000	0.857	0.857	0.857	0.857	0.714
H_E	0.857	0.923	0.868	0.846	0.659	0.945	0.703	0.890	0.923	0.923	0.956	0.780	0.846
HWE p-value	0.928	0.513	0.829	**0.001**	0.838	**< 0.001**	0.623	0.793	0.592	0.5099	0.224	1.000	0.6099
Null allele bias	no	no	no	**yes**	no	**yes**	no	no	no	no	no	no	no
Linkage disequilibrium	no												
Black Forest (N = 8)													
H_O	1.000	1.000	1.000	0.500	1.000	0.625	1.000	0.625	0.750	0.714	0.750	0.875	0.625
H_E	0.875	0.942	0.875	0.833	0.833	0.883	0.775	0.842	0.892	0.923	0.875	0.792	0.842
HWE p-value	1.000	1.000	0.102	0.051	0.887	0.102	0.366	0.089	0.350	0.089	0.583	1.000	0.083
Null allele bias	no	no	no	**yes**	no	no	no	no	no	no	no	no	no
Linkage disequilibrium	no												

Locus	Parate 01	Parate 02	Parate 03	Parate 06	Parate 07	Parate 08	Parate 09	Parate 15	Parate 16	Pat2-43	PmaTG An33	Pma69	PmaC25
Pyrenees (N = 13)													
H_O	0.846	0.923	0.692	0.538	0.846	0.538	0.846	0.833	0.615	1.000	0.833	0.846	0.692
H_E	0.840	0.957	0.908	0.815	0.769	0.892	0.785	0.743	0.837	0.938	0.855	0.702	0.671
HWE p-value	0.660	0.359	0.157	0.064	0.973	**< 0.001**	0.646	0.713	0.157	1.000	0.413	0.942	0.067
Null allele bias	no	no	**yes**	**yes**	no	**yes**	no	no	no	no	no	no	no
Linkage disequilibrium	no												
Greece (N = 12)													
H_O	0.833	0.818	0.583	0.750	0.583	0.750	1.000	0.818	0.600	0.909	0.912	0.545	0.750
H_E	0.884	0.9697	0.833	0.743	0.692	0.917	0.747	0.848	0.733	0.896	0.906	0.610	0.754
HWE p-value	0.305	0.145	0.230	0.331	0.264	0.006	0.042	0.811	1.000	0.156	0.637	0.762	0.897
Null allele bias	no	no	no	no	no	no	no	no	no	no	no	no	no
Linkage disequilibrium	no												

Locus	Parate 01	Parate 02	Parate 03	Parate 06	Parate 07	Parate 08	Parate 09	Parate 15	Parate 16	Pat2-43	PmaTG An33	Pma69	PmaC25
Corsica (N = 10)													
H_O	0.700	0.900	0.900	0.400	0.500	0.300	0.800	0.600	0.900	0.900	0.900	0.400	0.800
H_E	0.795	0.821	0.832	0.868	0.768	0.884	0.679	0.742	0.926	0.811	0.884	0.584	0.574
HWE p-value	0.305	0.239	0.311	**< 0.001**	0.068	**< 0.001**	0.813	0.320	0.794	0.984	0.984	0.141	0.198
Null allele bias	no	no	no	**yes**	no	**yes**	no	no	no	no	no	no	no
Linkage disequilibrium	Parate 03, Parate 07												
Sardinia (N = 10)													
H_O	0.800	0.900	0.600	0.800	0.778	0.600	0.200	0.333	0.800	0.556	0.889	0.500	0.300
H_E	0.900	0.795	0.684	0.816	0.752	0.737	0.189	0.582	0.621	0.712	0.791	0.426	0.479
HWE p-value	0.667	0.623	0.341	0.751	0.255	0.165	1.000	0.137	1.000	0.203	0.010	1.000	0.4799
Null allele bias	no	no	no	no	no	no	no	no	no	no	no	no	no
Linkage disequilibrium	no												

Locus	Parate 01	Parate 02	Parate 03	Parate 06	Parate 07	Parate 08	Parate 09	Parate 15	Parate 16	Pat2-43	PmaTG An33	Pma69	PmaC25
Cyprus (N = 7)													
H_O	1.000	1.000	0.714	0.429	0.714	0.000	0.571	0.429	0.714	1.000	0.857	mono.*	0.857
H_E	0.846	0.934	0.879	0.901	0.802	0.615	0.4395	0.582	0.714	0.934	0.912	mono.*	0.670
HWE p-value	0.707	1.000	0.504	**0.0003**	0.4099	**0.002**	1.000	0.637	1.000	1.000	0.636	-	0.412
Null allele bias	no	no	no	**yes**	no	**yes**	no	no	no	no	no	no	no
Linkage disequilibrium	Parate 01, Pat2-43												

*: mono.= monomorphic

Table S2: Matrix of pairwise F_{ST} values between all local populations (computed with Arlequin 3.1). Bold: significant F_{ST} values after Bonferroni correction for the number of comparisons at the level p = 0.05. (Bonferroni correction alpha (α) = 0.05 / for 91 comparisons α = 0.0005).

	C Asia	Russia	Finland	Norway	SH	Harz	Saxony	P Forest	B Forest	Pyrenees	Greece	Corsica	Sardinia
Central Asia	-												
Russian Far East	0.014	-											
Finland	0.010	0.010	-										
Norway	0.046	0.021	0.003	-									
Schleswig-Holstein	**0.088**	**0.059**	**0.038**	0.018	-								
Harz	0.072	0.064	0.036	0.034	0.010	-							
Saxony	0.067	0.044	0.021	0.034	0.004	-0.007	-						
Palatine Forest	0.039	0.043	0.024	0.024	0.010	0.007	-0.014	-					
Black Forest	0.067	**0.042**	0.028	0.019	0.009	0.001	-0.003	0.0002	-				
Pyrenees	**0.095**	**0.075**	**0.050**	**0.047**	0.010	-0.004	-0.004	0.009	0.022	-			
Greece	**0.106**	**0.082**	**0.059**	**0.037**	0.026	0.011	0.017	0.011	0.014	0.013	-		
Corsica	**0.137**	**0.129**	**0.095**	**0.090**	**0.048**	0.036	0.033	0.051	**0.056**	0.033	**0.055**	-	
Sardinia	**0.256**	**0.215**	**0.179**	**0.155**	**0.134**	**0.102**	**0.156**	**0.148**	**0.122**	**0.101**	**0.106**	**0.064**	-
Cyprus	0.172	**0.143**	**0.130**	**0.116**	**0.172**	**0.176**	0.187	0.156	**0.145**	**0.179**	**0.158**	**0.218**	**0.284**