

Disentangling the complex evolutionary history of the Western Palearctic blue tits (*Cyanistes* spp.) – phylogenomic analyses suggest radiation by multiple colonization events and subsequent isolation

MARTIN STERVANDER,* JUAN CARLOS ILLERA,† LAURA KVIST,‡ PEDRO BARBOSA,* NAOMI P. KEEHNEN,*¹ PETER PRUISSCHER,*¹ STAFFAN BENSCH* and BENGT HANSSON*

*Molecular Ecology and Evolution Lab, Department of Biology, Lund University, Ecology Building, 223 62 Lund, Sweden,

†Research Unit of Biodiversity (UO-CSIC-PA), Oviedo University, Campus of Mieres, Research Building, 5th Floor. Calle

Gonzalo Gutiérrez Quirós, s/n, 33600 Mieres, Asturias, Spain, ‡Department of Biology, University of Oulu, POB 3000, 90014 Oulu, Finland

Abstract

Isolated islands and their often unique biota continue to play key roles for understanding the importance of drift, genetic variation and adaptation in the process of population differentiation and speciation. One island system that has inspired and intrigued evolutionary biologists is the blue tit complex (*Cyanistes* spp.) in Europe and Africa, in particular the complex evolutionary history of the multiple genetically distinct taxa of the Canary Islands. Understanding Afrocanarian colonization events is of particular importance because of recent unconventional suggestions that these island populations acted as source of the widespread population in mainland Africa. We investigated the relationship between mainland and island blue tits using a combination of Sanger sequencing at a population level (20 loci; 12 500 nucleotides) and next-generation sequencing of single population representatives (>3 200 000 nucleotides), analysed in coalescence and phylogenetic frameworks. We found (i) that Afrocanarian blue tits are monophyletic and represent four major clades, (ii) that the blue tit complex has a continental origin and that the Canary Islands were colonized three times, (iii) that all island populations have low genetic variation, indicating low long-term effective population sizes and (iv) that populations on La Palma and in Libya represent relicts of an ancestral North African population. Further, demographic reconstructions revealed (v) that the Canary Islands, conforming to traditional views, hold sink populations, which have not served as source for back colonization of the African mainland. Our study demonstrates the importance of complete taxon sampling and an extensive multimarker study design to obtain robust phylogeographical inferences.

Keywords: back colonization, Canary Islands, coalescence, colonization, isolation, phylogeography, RAD sequencing, radiation

Received 3 November 2014; revision received 20 February 2015; accepted 25 February 2015

Introduction

Isolated islands are often inhabited by plants and animals found nowhere else. Famous examples include

Caribbean anoles lizards (Losos 2011), the Galapagos finches (Grant & Grant 2011), Hawaiian spiders (Gillespie *et al.* 1994) and Lord Howe palms (Savolainen *et al.* 2006). It is natural to ask how these islands became colonized: which were the mainland ancestors, and when, why and how did the diversity arise? Indeed, island biota have inspired generations of researchers and have played a key role for the foundation of modern

Correspondence: Martin Stervander, Fax: +46 46 222 4716;

E-mail: martin.stervander@biol.lu.se

¹Present address: Population genetics, Department of Zoology, Stockholm University, 106 91 Stockholm, Sweden

evolutionary and ecological theory (Warren *et al.* 2015). Yet today, isolated islands attract attention from evolutionary biologists and after entering the genomics era,

there is an increasing optimism to uncover patterns and processes of colonization, island diversification and relationships to mainland relatives.

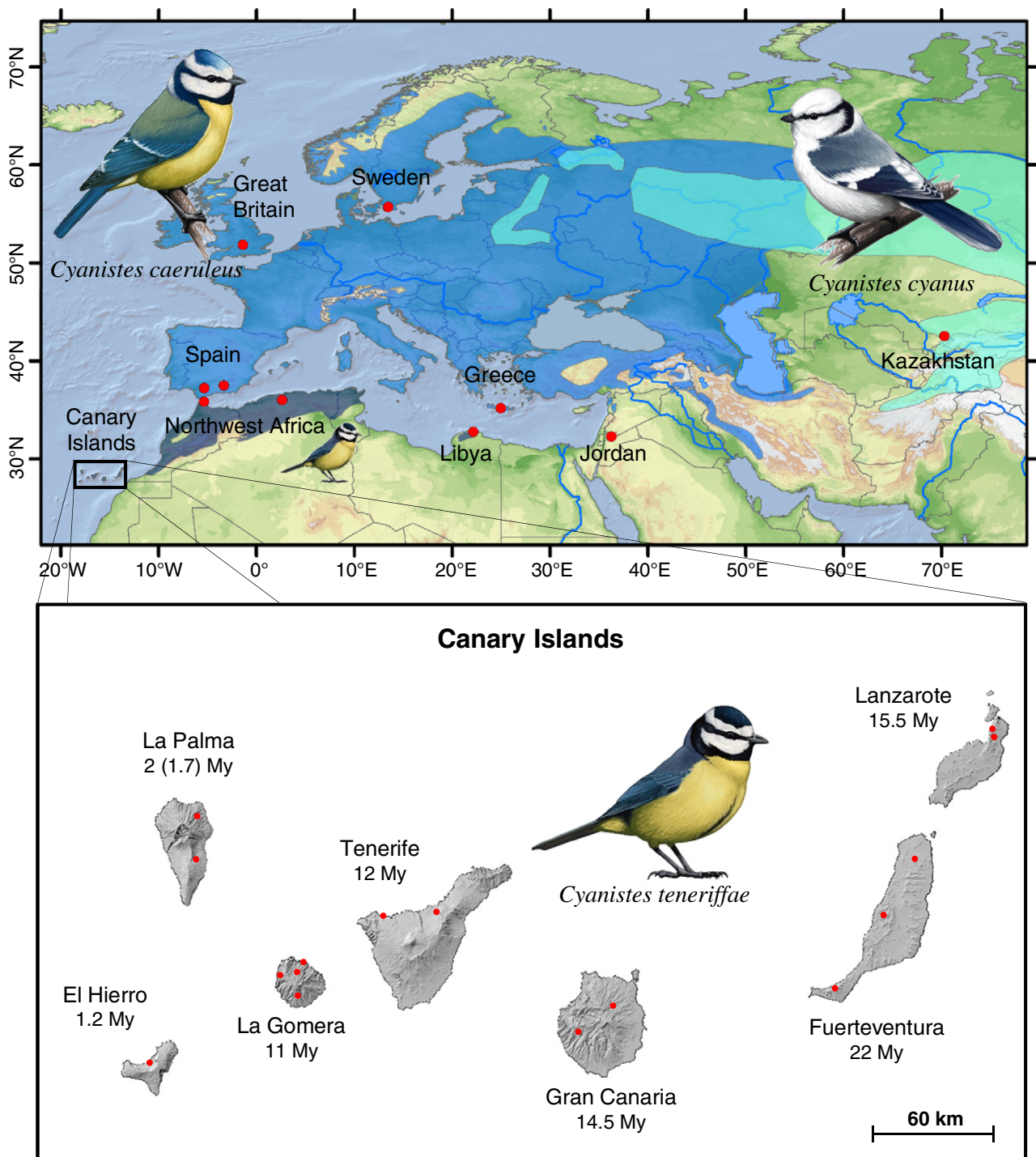


Fig. 1 Sampling locations (red circles) and distribution of African blue tit *Cyanistes teneriffae* (dark cobalt blue), Eurasian blue tit *C. caeruleus* (medium blue) and azure tit *C. cyanus* (light cyan blue). The inset shows the Canary Islands and their estimated geological ages. Illustrations by Martí Franch. Distribution data from BirdLife International and NatureServe (2014) modified for the western part of azure tit distribution according to Gosler & Clement (2007); geological ages from Carracedo & Day (2002) and alternative age for La Palma in parentheses, following Guillou *et al.* (2001).

Island populations are often considered marginal to the mainland distribution and implicitly assumed to result from colonization from a mainland source. Indeed, there are ample examples of this, such as the colonization of Darwin's finches on Galapagos Islands (Grant & Grant 2011), white-eyes in the Gulf of Guinea (Melo *et al.* 2011) and finch-like weavers on Mauritius and Réunion (Warren *et al.* 2012). However, the general view of islands as sinks has been challenged lately, and there are now examples where island populations have been acting as a source for reverse mainland colonization (Bellemain & Ricklefs 2008). For example, the spread of monarch flycatchers (family Monarchidae) from the Papuan archipelago includes colonizations from the archipelago to South-East Asia (Filardi & Moyle 2005), and similar examples come from the Corvoidea (Jonsson *et al.* 2011) and other taxa (Balke *et al.* 2009; Jonsson *et al.* 2010). Those recently documented examples of back colonizations, from islands to mainland, have led to a reassessment of the evolutionary and biogeographical roles of islands in a broader geographical context.

The Canary Islands, located approximately 100 km off the coast of north-west Africa, are a part of the Mediterranean basin biodiversity hotspot (Medail & Quezel 1999) with many endemic species of plants and animals, including five endemic bird species and several subspecies (Juan *et al.* 2000; Illera *et al.* 2012). The islands span over 500 km from east to west in a gradient of decreasing age from 23 million years (My) to 1 My (Fig. 1; Carracedo & Day 2002). Of the bird populations in the Canary Islands community, the African blue tit *Cyanistes teneriffae* is the most widespread and diverse taxon with five subspecies recognized (according to the International Ornithological Union; Gill & Donsker 2014): *C. t. palmensis* on La Palma, *C. t. ombriosus* on El Hierro, *C. t. teneriffae* on La Gomera and Tenerife, *C. t. hedwigae* on Gran Canaria, and *C. t. degener* on Fuerteventura and Lanzarote. In addition, *C. t. ultramarinus* occurs throughout Morocco, Algeria, Tunisia and Pantelleria (a small Italian island located in the strait of Sicily), and an isolated population, *C. t. cyrenaicae*, occurs in north-eastern Libya. Outside the Afrocanarian region, there are two more species in the genus: Eurasian blue tit *C. caeruleus*, occurring in Europe and western Asia with nine recognized subspecies, and Azure Tit *C. cyanus*, occurring in Asia from the Pacific coast through to eastern Europe with eight subspecies (Gill & Donsker 2014). For further details on the distribution of taxa, see Fig. 1.

The African blue tit populations of the Canary Islands have received plenty of interest for their differences in plumage (Martin 1991; Cramp & Perrins 1993), structural morphology and ecology (Partridge & Pringmill 1977; Grant 1979; Carrascal *et al.* 1994) and vocalization (Schottler 1993, 1995). As the different taxa are allopat-

ric, any direct test of the propensity to interbreed is difficult, but song playback experiments by Schottler (1995) showed that the vocal differentiation between some island taxa has progressed thus far that song from other populations may not elicit any response, and Sangster (2006) argued that they would probably best be warranted species status.

Although the population genetic structure of the Canary Island blue tits is well studied, the phylogeographical history of the Afrocanarian blue tits is highly complex and has been difficult to resolve, given that every single study has suggested widely to moderately different scenarios (Salzburger *et al.* 2002; Kvist *et al.* 2005; Dietzen *et al.* 2008; Illera *et al.* 2011; Päckert *et al.* 2013; Hansson *et al.* 2014; Gohli *et al.* 2015). Thus, the number of colonization events (1–3), the colonization pathways within and between the Canary Islands and the mainland (westward or eastward stepping-stone colonization, or from centre and out) and the origin of the genetically divergent population on La Palma (from Europe or Africa) remains debated. The latter has been grouped with other Afrocanarian populations (Kvist *et al.* 2005; Päckert *et al.* 2013; Gohli *et al.* 2015), but also with Eurasian population owing to sequence similarity and homologous deletions (Dietzen *et al.* 2008).

Furthermore, the Canary Islands' role as a colonization sink by mainland sources or as source for mainland colonization remains controversial. At least to some extent, this uncertainly may be explained by the occurrence of strong bottlenecks and incomplete lineage sorting in the Canary Islands clade, which implies that different loci often have different evolutionary histories. Unfortunately, most studies of the Afrocanarian blue tits have been based on single mitochondrial markers (Salzburger *et al.* 2002; Kvist *et al.* 2005; Dietzen *et al.* 2008), one or few additional nuclear markers (Illera *et al.* 2011; Päckert *et al.* 2013), and/or suffer from incomplete or inadequate sampling of relevant taxa, particularly the Libyan subspecies *C. t. cyrenaicae* (Salzburger *et al.* 2002; Kvist *et al.* 2005; Dietzen *et al.* 2008; Illera *et al.* 2011; Hansson *et al.* 2014; Gohli *et al.* 2015).

The recent development in sequencing technology and bioinformatics provide dramatically larger genomic data sets and promise novel insights into disputed biogeographical and evolutionary concepts (Davey *et al.* 2011; Ekblom & Galindo 2011). These technological advances also call for new types of analyses, and single-gene phylogenies are replaced by phylogenetic analyses of concatenated sequence matrices, or coalescence-based species tree inference (see e.g. Patel *et al.* 2013). In a recent study, Gohli *et al.* (2015) did employ massively parallel sequencing, but they only had access to degraded museum samples of the Libyan subspecies,

which produced a phylogeny with low nodal support of the main clades. In addition, their main coalescence-based species tree and demography analyses (excluding the Libyan population) lack strong statistical support (i.e. species tree analyses seem to converge poorly and most effective sample sizes are much too low), rendering published evolutionary interpretations uncertain.

In this study, we investigate the relationship between mainland and island Afrocanarian blue tit taxa, and their relationship to Eurasian *Cyanistes* taxa, using a combination of traditional Sanger multigene sequencing at a population level, with next-generation sequencing of single population representatives, analysed in coalescence, phylogenetic and structure assignment frameworks. With the combined strength of these different frameworks, we aim to understand the evolutionary history of the Afrocanarian blue tits, their interrelationships and their relationships to Eurasian lineages. Specifically, we address the number of colonization events to the Canary Islands, the colonization pathway(s), and we estimate the timing of these events. Finally, we aim to support or dispute the role of the Canary Island blue tits as founders of the widespread populations on the African mainland.

Materials and Methods

Feather or blood samples were collected nondestructively from mist-netted wild birds in populations of African blue tit on all seven Canary Islands, in north-west Africa (Ceuta, hereafter referred to as Morocco, and Algeria), in Libya, and three populations of Eurasian blue tits in Spain, Great Britain and Sweden (Table S1, Supporting information, Fig. 1). From each population, 4–5 individuals were sampled for Sanger sequencing, as were 1–2 individuals from Greece and Jordan, as well as azure tit and great tit *Parus major*. From all populations, a single representative was selected for restriction-site-associated DNA (RAD) sequencing, together with additional outgroup populations such as *Periparus* spp. and *Poecile* spp. (for details, see Table S1, Supporting information).

Blood samples were preserved in ethanol or SET buffer. DNA was extracted using a standard phenol–chloroform protocol for blood (Sambrook & Russel 2001) or standard sodium acetate method for feathers. DNA extracted from feather samples were whole-genome amplified using the REPLI-g kit (Qiagen) according to the manufacturer's instructions.

Sanger sequencing of population samples

Amplification and Sanger sequencing. We amplified DNA using Qiagen Multiplex PCR Kit in 10 µl reactions, for two mitochondrial and 18 nuclear markers. For details

on primers and PCR profiles, see Table S2 (Supporting information). PCR products were precipitated (NH₄Ac and ethanol) and then dissolved in 10–40 µl of water depending on amplification success, of which 2 µl was used for 10 µl sequencing reactions (BigDye sequencing kit; Applied Biosystems; Table S2, Supporting information) in 35 PCR cycles at 96 °C for 10 s, 47 °C for 5 s and 60 °C for 4 min. The products were run on an ABI Prism 3100 capillary sequencer (Applied Biosystems).

Sequence editing, alignment and haplotype inference. All Sanger sequences were edited with Geneious 6.1 (Biomatters). We trimmed sequences and searched for double peaks in the electropherograms using the find heterozygotes plug-in. Regions of sequences with too low quality causing ambiguities were coded as missing data. We aligned the sequences using the MAFFT (Katoh & Standley 2013) multiple alignment plugin.

To obtain phased data from the Sanger sequencing data set, we inferred haplotypes with the PHASE v. 2.1 algorithm (Stephens *et al.* 2001; Stephens & Donnelly 2003) available in DNASP v. 5.10.01 (Librado & Rozas 2009). For further details on haplotype inference, see the Supporting information.

Substitution model selection and implementation of a molecular clock. Substitution models were evaluated with JMODELTEST v. 2.1.4 (Guindon & Gascuel 2003; Darriba *et al.* 2012), selecting from 88 available models allowing for rate heterogeneity according to four gamma categories and for a proportion of invariable sites. Model selection was performed according to the Bayes Information Criterion (BIC; Schwarz 1978).

We used molecular clock settings as follows: for cytochrome b, 0.0105 ± 0.0005 (95% CI) substitution/site/Ma (Weir & Schluter 2008); for the control region, the *Cyanistes*-specific estimate 0.02 ± 0.01 substitution/site/Ma (Päckert *et al.* 2007); and for nuclear markers, 0.00135 ± 0.00045 substitution/site/Ma (Ellegren 2007). The bounds of 95% CI were used as bounds for uniform priors, and, unless stated differently, the molecular clock was modelled as strict.

Coalescent species tree analyses. *BEAST infers a species tree simultaneously with marker-specific tree reconstruction (Heled & Drummond 2010). Input was prepared in BEAUti, part of the BEAST v. 2.0.2 suite (Bouckaert *et al.* 2014) along with *BEAST. Estimates by JMODELTEST of transition/transversion ratio, proportion of invariable sites and gamma shape parameter were implemented, whereas base frequencies were left to be estimated empirically (except for K80 models). Clock models were linked for all nuclear markers, and tree models were linked for the two mitochondrial markers, whereas other tree and substitu-

tion models were left unlinked. We set the species tree prior to follow the Yule process and the population size model to follow the piecewise linear with constant root.

In *BEAST*, we ran two independent MCMC runs with a chain length of >850M, sampling every 100K. The output was inspected with *TRACER* v. 1.5 (Rambaut & Drummond 2007), to ensure that likelihood scores were stationary and that effective sample sizes (ESS) were adequate, setting burnin to 10%.

An additional **BEAST* analysis was performed according to above, excluding the two mitochondrial markers and thus only including nuclear markers.

Analyses of genetic variation and demographic reconstruction. Within-population nucleotide diversity, haplotype diversity and θ were calculated using *DNASP* v. 5.10.01 (Librado & Rozas 2009).

We reconstructed historical demography using extended Bayesian skyline plot analyses (Heled & Drummond 2008) in *BEAST* v. 1.8.0 (Drummond *et al.* 2012). The analyses were based on variable loci (see Table S3, Supporting information) and applied clock and tree priors as specified for the **BEAST* analyses above, with some modification of operator weighting. Analyses were run for 500M generations, sampling every 40K generation. Results were drawn using the *GGPLOT2* v. 1.0 package (Wickham 2009) in *R* v. 3.0.3 (R Core Team 2013).

RAD sequencing of individual representatives

We analysed a genomewide representation (RAD markers) of single representatives per population, in accord with the notion that optimization of resolution and precision in phylogenetic inference is best achieved by increasing the number of markers rather than the number of individuals (Corl & Ellegren 2013).

Library preparation and sequencing. We prepared paired-end RAD libraries, for one individual per population and additional outgroups of the genera *Periparus* and *Poecile*, following the protocols of Baird *et al.* (2008) and Etter *et al.* (2011) with modifications according to the Supporting information. Six samples yielded lower sequencing coverage (in particular *C. t. degener* from Lanzarote and *P. major*) and were therefore included in a subsequent library. For these samples, the DNA aliquots from the first library, which were already digested and ligated with P1 adapter, were religated with adapter P1 and included in the remainder of library preparation. For RAD sequencing and data filtering results, see Table S4 (Supporting information).

Bioinformatic pipeline. The paired raw reads were first filtered using the *process_radtags* and *clone_filter* compo-

nents of *STACKS* v. 1.17 (Catchen *et al.* 2013). Reads of samples, which had been included in two libraries, were then merged per sample. We mapped the remaining reads to the Zebra Finch *Taeniopygia guttata* genome (assembly *taeGut3.2.4*, downloaded from the UCSC genome website <http://genome.ucsc.edu/>) with *BOWTIE* v. 2.2.2 (Langmead & Salzberg 2012). Given the interspecific mapping, we set maximum and minimum mismatch penalties (*-mp*) to 2 and 1, respectively. Further, we only accepted paired reads mapping in the expected relative direction within 150–1000 base pairs (bp) from each other (*-no-discordant*, *-no-mixed*). From the resulting SAM files, we removed read 2 and used read 1 as input to *STACKS*, following the *ref_map* track for phylogenetic and coalescent analyses. We used default parameters except for setting a higher minimum sequencing coverage to call a locus in a given sample (*m* = 4). The sample of *C. t. degener* from Lanzarote as well as the *P. major* sample was of substantially lower coverage, and further treatment of data was carried out in parallel for one data set in which they were retained, and one in which they were removed.

We filtered the output so that it only contained loci (i) in which all samples had been assigned 1–2 alleles; (ii) which contained 1–10 single-nucleotide polymorphisms (SNPs); (iii) which comprised ≤ 12 alleles. The rationale for these filters was to eliminate any loci in which individuals had been assigned more than two alleles and/or which were remarkably variable (cut-off values, however, assigned arbitrarily after inspecting data). By doing so, we may have traded off erroneous exclusion of highly variable loci not to include misassembled loci.

The *STACKS* output was converted to biallelic sequences for all samples and loci, using custom scripts, and concatenated based on sample (but with random allele pairing) to long super-matrices. Finally, we defined one data set in which all taxa were included, and no missing data at any locus were tolerated: super-matrix *STRICT*, comprising 11 426 loci with total length of 994 062 bp. In another data set, we removed the two samples with lowest coverage and allowed up to 48% missing data among the remaining samples at any given locus: super-matrix *RELAXED*, comprising 37 292 loci with a total length of 3 244 404 bp.

Phylogenomic analyses. Super-matrices *STRICT* and *RELAXED* were analysed in *RAXML* v8.0.26 (Stamatakis 2014), using the *GTRGAMMA* substitution model (general time-reversible model allowing for rate heterogeneity according to a gamma distribution). The best maximum-likelihood (ML) tree from 1000 searches was selected, and 1000 bootstrap replicates for the super-matrix *STRICT*, and 200 replicates for super-matrix *RELAXED*, were performed. Convergence was verified running a *posteriori* bootstrap

Table 1 Mean estimated node ages in million years (95% highest posterior density) for three different calibration sets and four subsets of 14-Kbp genomic RAD sequence data; three autosomal (presented here with average values) and one from the Z chromosome. For additional calibration sets and taxon sets, see Table S5 (Supporting information)

Clade		Genomic subset	Calibration set*		
			A (BER rel., MSC rel.)	B (BER, MSC)	C (BER, MSC, CAN)
AB	<i>Cyanistes</i>	aut. (mean)	4.4 (2.4–6.6)	5.4 (3.9–7.0)	3.9 (3.1–4.8)
		Z chrom.	4.9 (2.9–7.0)	5.7 (4.2–7.2)	4.5 (3.6–5.5)
A	<i>C. caeruleus</i> + <i>C. cyanus</i>	aut. (mean)	3.0 (1.6–4.6)	3.7 (2.6–4.9)	2.9 (2.1–3.6)
		Z chrom.	3.1 (1.7–4.5)	3.6 (2.6–4.8)	3.0 (2.2–3.8)
A1	<i>C. caeruleus</i>	aut. (mean)	2.8 (1.5–4.1)	3.3 (2.3–4.4)	2.6 (1.9–3.3)
		Z chrom.	2.4 (1.3–3.6)	2.8 (1.9–3.7)	2.3 (1.6–3.1)
B	<i>C. teneriffae</i>	aut. (mean)	3.9 (2.2–5.8)	4.8 (3.4–6.2)	3.2 (2.8–3.5)
		Z chrom.	3.8 (2.2–5.5)	4.4 (3.3–5.7)	3.2 (2.8–3.5)
B1†	<i>C. t. palmensis</i> + <i>C. t. cyrenaicae</i>	aut. (mean)	3.1 (1.6–4.7)	3.7 (2.5–5.0)	2.7 (2.1–3.3)
		Z chrom.	3.4 (1.9–5.0)	3.9 (2.8–5.2)	3.0 (2.4–3.5)
B1b	<i>C. teneriffae</i> excl. <i>C. t. palmensis</i>	aut. (mean)	3.7 (2.0–5.5)	4.4 (3.1–5.8)	3.0 (2.5–3.4)
		Z chrom.	3.5 (2.0–5.1)	4.0 (2.9–5.2)	2.9 (2.4–3.4)
B2	<i>C. teneriffae</i> , clades B3 + B4	aut. (mean)	2.8 (1.5–4.3)	3.4 (2.4–4.6)	2.3 (1.8–2.9)
		Z chrom.	2.1 (1.2–3.1)	2.5 (1.7–3.2)	1.8 (1.3–2.2)
B3	<i>C. teneriffae</i> , central Canary Is.	aut. (mean)	1.8 (0.9–2.8)	2.1 (1.3–3.0)	1.1 (0.9–1.2)
		Z chrom.	1.7 (0.9–2.5)	1.9 (1.3–2.6)	1.1 (1.0–1.2)
B4	<i>C. teneriffae</i> , E. Canary Is., NW. Afr.	aut. (mean)	2.0 (1.0–3.1)	2.5 (1.6–3.3)	1.7 (1.3–2.3)
		Z chrom.	1.5 (0.8–2.2)	1.7 (1.1–2.3)	1.3 (0.9–1.7)

rel., relaxed; aut. (mean), autosomal (average of three independent genomic subsets); Z chrom., Z chromosome;

*Calibration points were as follows: BER, Beringia: Nearctic/Holarctic split within *Poecile* at 5.4–5.5 My or relaxed at 4.8–14.0 My; MSC, Messinian Salinity Crisis: African/European split within *Periparus ater* at 5.32–5.33 My or relaxed at 0–5.96 My; CAN, Canary Islands volcanic age at El Hierro 0–1.2 My and La Palma 0–3.5 My.

†*C. t. palmensis* (La Palma) and *C. t. cyrenaicae* (Libya) enforced as a monophyletic clade according to species tree analyses (see text).

tests using the autoMRE function (Pattengale *et al.* 2009).

Cluster analyses. A multidimensional scaling analysis of the unfiltered SNP data set of 3670 spaced SNPs was performed in PLINK v. 1.9beta2c (Purcell & Chang 2014) to detect similarities or dissimilarities among and within the populations studied.

Dating. For computational limitation reasons, we selected subsets comprising 14 Kbp (160 loci) each, from super-matrix STRICT, which we used for phylogenomic dating. We selected three subsets of autosomal loci at random, using function sample{base} in R (R Core Team 2013) and made one subset of the loci which had mapped to the sex chromosome Z.

Calibration priors were put on the split between Nearctic and Holarctic taxa of *Poecile*, separated before or at the last opening of the Bering Strait (5.5–5.4 million years ago (Mya) according to Marinovich & Gladenkov (1999) and Gladenkov *et al.* (2002)) and the split between North African and European populations of coal tit *Periparus ater*, corresponding to the Messinian salinity crisis (full range covering 5.96–5.32 Mya according to

Krijgsman *et al.* (1999)). Following the argument by Päckert *et al.* (2011) that the faunal interchange between the Holarctic and the Nearctic probably pre-dated the last opening of the Bering Strait, and comparable to Illera *et al.* (2011) and Päckert *et al.* (2013), we employed a more relaxed calibration set (A) where a uniform prior of 14–4.8 Mya was applied for the Nearctic–Holarctic split and a uniform prior of 5.96–0 Mya applied for the Messinian salinity crisis. We also employed a number of alternative calibration sets, including more conservative settings for the calibration points above, for example setting the prior for the split between North African and European populations of coal tit to the period of desiccation of the Mediterranean Sea (calibration sets B, D–E; Table 1, Table S5, Supporting information).

Finally, two further constraints according to the geological ages of islands were added (calibration set C; Tables 1 and S5, Supporting information) in accord with Päckert *et al.* (2013): La Palma (clade B or B1) set at a uniform prior of 3.5–0 Mya and El Hierro (clade B3) at 1.2–0 Mya (following Fernández-Palacios *et al.* 2011).

All datings were performed both for the full taxon set and for a partial taxon set, from which the samples from Lanzarote and Tenerife were removed, owing to

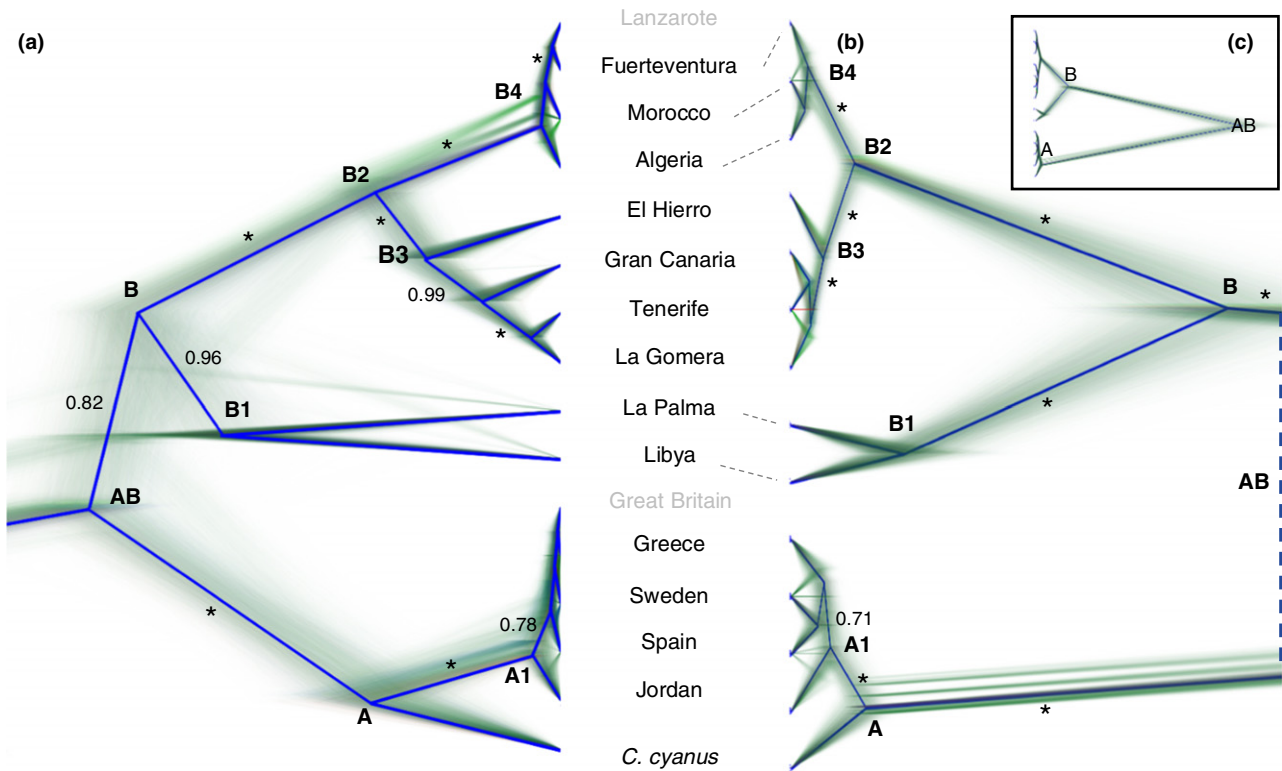


Fig. 2 Coalescence-based species trees of blue tits *Cyanistes* spp. Trees are drawn in DensiTree with individual trees represented by very thin lines, making up tree densities, and main topology highlighted with a distinct line. Posterior probabilities higher than 0.6 are stated along the branches (* = 1.0). The species trees are based on (a) population samples of two mitochondrial and 18 nuclear Sanger sequence markers, estimated with *BEAST and (b, c) single-nucleotide polymorphism (SNP) data from restriction-site-associated DNA (RAD) sequencing of single individuals per lineage, estimated using SNAPP. Note that the species-specific branches are systematically shortened in the RAD-based tree due to methodology (see text); therefore, the terminal end of the RAD-based tree is zoomed in (b), while the full topology is indicated in an inset (c). Clade names are referred to in the text. Populations of *C. teneriffae* and *C. caeruleus* are referred to by their geographical origin, and labels are aligned with the tips of sequence-based tree (a). The RAD-based tree (b) does not include samples from Lanzarote or Great Britain (marked grey), and dashed lines are indicating alignment with population labels when unaligned.

much higher within-individual genetic variation than other samples.

The ML tree from super-matrix STRICT (pruned for the partial taxon set) was transformed in R package APE v. 3.1.4 (Paradis *et al.* 2004) to fit the calibration constraints in units of My, and these trees were used as starting trees in dating analyses performed in BEAST v. 2.1.3 (Bouckaert *et al.* 2014). To accommodate for the major incongruence between ML and coalescence species trees, replicates of all dating runs were performed with starting trees in which the populations of La Palma and Libya were enforced as a monophyletic clade (cf. clade B1, Fig. 2).

The concatenated sequences were not partitioned, and an uncorrelated lognormal relaxed clock was implemented (Drummond *et al.* 2006). We employed a Yule speciation model, setting a diffuse gamma prior on the birth rate. Major clades were constrained as

monophyletic according to the topology of the starting tree, which was rooted according to the intergeneric relationships in Johansson *et al.* (2013). We ran the MCMC chain 100M generations, sampling every 10K generation, and inspected the results with TRACER v. 1.5 (Rambaut & Drummond 2007), after which 10% of the trees were discarded as burn-in, and maximum credibility clade trees were calculated in TREEANNOTATOR (Bouckaert *et al.* 2014). For 3 of the 80 analyses, stationarity was not reached within 10M generations, and effective sample size was not $\gg 200$, and these were then run another 10–25M generations, after which burn-in was discarded so that the final number of trees analysed corresponded to all other dating analyses.

Coalescent species tree analyses of SNPs. Using VCFTOOLS v. 0.1.12a (Danecek *et al.* 2011), we extracted the in-group only from super-matrix STRICT and selected

effectively nonlinked SNPs by allowing a maximum of one SNP per RAD locus. This unfiltered SNP set contained 3670 spaced SNPs. For computational purposes, autapomorphies were filtered out from the SNP set by requiring a minor allele count of ≥ 3 , resulting in retention of 3421 SNPs. The filtered data set was imported into BEAUTI (Bouckaert *et al.* 2014), where priors for forward (u) and reverse (v) mutation rates were initially set to $1/(2 \times \text{allele frequency})$ for reference (u) and alternative (v). However, preliminary runs rendered improved efficiency when u and v parameters were co-estimated during the runs, thus co-estimation was implemented in final runs. Remaining parameters were left at default values, for exam-

ple species divergence rate $\lambda = 0.00765$, and θ defined by a γ prior with shape parameter $\alpha = 11.750$ and scale parameter $\beta = 109.73$. We performed two independent runs for ≥ 8.4 M generations, sampling every 1K generation using the SNAPP v. 1.1.16 plugin (Bryant *et al.* 2012) in BEAST v. 2.1.3 (Bouckaert *et al.* 2014). Output was inspected with TRACER v1.5 (Rambaut & Drummond 2007), and the converging runs were combined after excluding 10% as burnin. Effective sample sizes of prior, posterior, likelihood and tree height were $\gg 200$ (whereas estimations of θ for ancient populations, as expected, did not reach sufficient effective sample sizes based on one individual per population).

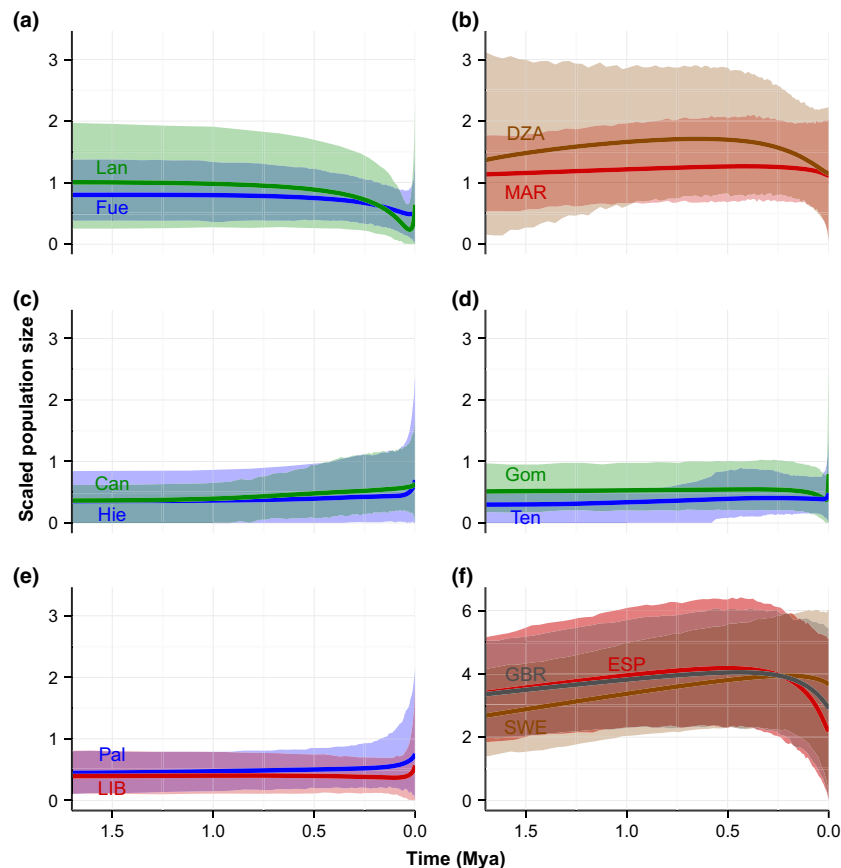


Fig. 3 Demographic reconstruction of African blue tit *Cyanistes teneriffae* (a–e) and Eurasian blue tit *C. caeruleus* (f) populations using extended Bayesian Skyline Plots based on multilocus Sangers sequence data. Solid lines represent mean estimated scaled population size (effective population size $\times$ generation time in My), and shaded areas the 95% highest posterior density. Populations are indicated with abbreviations and island taxa are drawn in blue/green whereas mainland populations are drawn in red/brown/grey. (a) Eastern Canary Islands: Fuerteventura (Fue; blue) and Lanzarote (Lan; green), (b) north-west Africa: Morocco (MAR; red) and Algeria (DZA; brown), (c) central Canary Islands: Gran Canaria (Can; green) and El Hierro (Hie; blue), (d) central Canary Islands: La Gomera (Gom; green) and Tenerife (Ten; blue), (e) La Palma (Pal; blue) and Libya (LIB; red), (f) European populations: Spain (ESP; red), Great Britain (GBR; grey) and Sweden (SWE; brown). Note that the last panel (f) has a different scale of the y-axis. Alternative versions of this figure, zooming out to cover the full reconstructions, and zooming in on the most recent period, respectively, are available as Figs S3 and S4 (Supporting information).

Results

Sanger sequencing of population samples

The Sanger sequencing data included 4–5 individuals per population and covered a total of 12 498 bp distributed over 18 nuclear and two mitochondrial markers. For the results of substitution model selection for the Sanger sequencing markers, see Table S6 (Supporting information).

Coalescence species tree analyses. The two sequence-based species tree runs, based on mitochondrial and nuclear markers, in *BEAST converged, and resulting data were combined, producing a final set of 15 656 trees. The emerging species tree reflects a fully supported (posterior probability (PP) = 1.0) Eurasian clade (clade A) with Eurasian blue tits (clade A1) and azure tit and a fully supported Afrocanarian clade (clade B), to which the populations from Libya and La Palma are assigned with PP = 0.82 (Fig. 2a). Within the Afrocanarian clade, a basal branch contains the Libyan and La Palma populations, grouped as sister taxa with PP = 0.96 (clade B1). The remaining taxa diverge into an unresolved and shallow group of north-west Africa and the eastern Canary Islands (i.e. Fuerteventura and Lanzarote; clade B4) and a group containing the populations of the central Canary Islands (clade B3), both fully supported. In the latter, El Hierro is basal, followed in branching

sequence by Gran Canaria and then La Gomera/Tenerife (PP ≥ 0.99 for all nodes; Fig. 2a).

The marker-specific trees vary in topology, but the fast-evolving mitochondrial markers recovered the general species tree topology including all major clades (see, e.g. tree for the control region, Fig. S1, Supporting information). In these trees, all populations within clades B1 and B3 turn out reciprocally monophyletic, whereas there is paraphyly with respect to populations within clades A1 and B4 (Fig. S1, Supporting information).

The run excluding mitochondrial markers resulted in a species tree with fully supported clades B3 and B4 (Fig. S2, Supporting information), identical in topology to the species tree including mitochondrial markers (Fig. 2a). However, resolution at larger evolutionary depths is poorer: the basal split of *Cyanistes* divides the population of La Palma from all other populations, next followed by European populations (clade A) branching off, then the Libyan population and finally the Canary Islands clades B3 and B4 (Fig. S2, Supporting information). The posterior probability for the nodes comprising all populations except La Palma is 0.70 and for the node comprising clades B3, B4 and Libya is 0.74. Furthermore, in clade A, *C. cyanus* is nested within *C. caeruleus* populations.

Genetic variation and demographic history. The genetic diversity is generally low among the Canary Island populations, in particular in the populations of El Hierro and Lanzarote (Tables S3, S7 and S8, Supporting

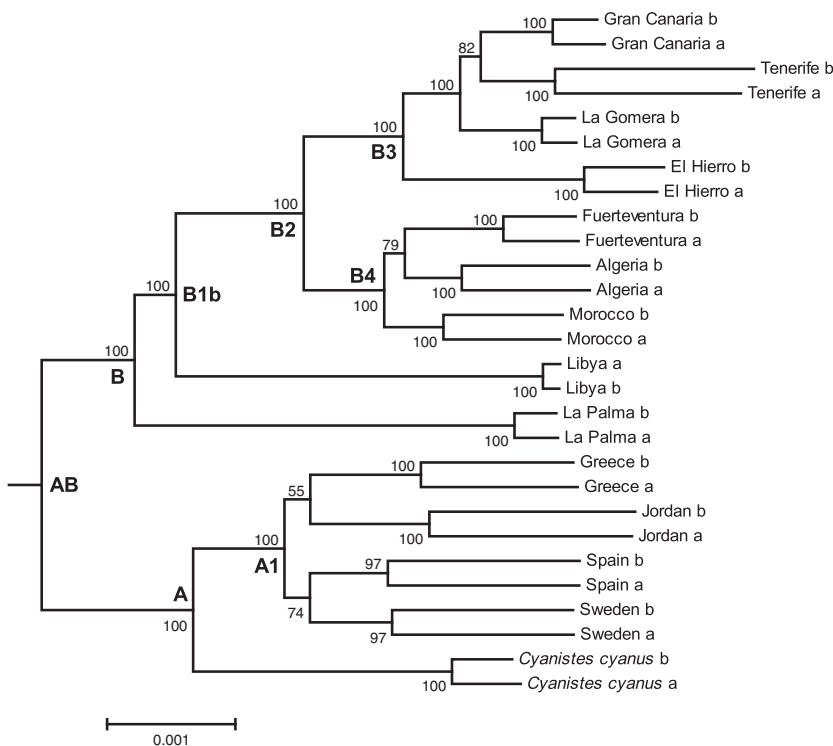


Fig. 4 Phylogenomic maximum-likelihood tree calculated in RAXML of supermatrix RELAXED, comprising 37 292 restriction-site-associated DNA (RAD) sequences (3.244 Mbp) from all individuals except low-coverage samples (i.e. Lanzarote and *P. major*) and allowing for up to 48% missing data at any locus. In the sequence names, 'a' and 'b' signify the two 'alleles' produced by randomly combining alleles from each individual RAD locus. Bootstrap support >50 displayed at nodes; clades referred to in the text; outgroups not shown.

information). Among mainland African populations, the Libyan one displays low genetic variation in par with Lanzarote. The Moroccan and, especially, Algerian population displays intermediate diversity and θ values which are higher than those for the Canary Islands and Libya, but lower than those of European populations, which, in turn, show relatively high diversity and θ values (Tables S3, S7 and S8, Supporting information).

The reconstruction of demographic history with Bayesian skyline plots show generally higher average population sizes for the mainland populations, except Libya, than for the island populations, although confidence intervals are large (Figs 3, S3 and S4, Supporting information). Most island populations show signs of a recent expansion, and in the populations of Lanzarote, La Gomera and to some extent Fuerteventura, this is preceded by bottlenecks (Figs 3a,c–e and S4a,c–e, Supporting information). On the contrary, the mainland populations of north-west Africa, as well as the European populations, display no signs of bottlenecks, and the population size trends span over longer time (Figs 3b,f and S3b,f, Supporting information).

RAD sequencing of individual representatives

Single individuals per population, along with additional outgroups, were subject to RAD sequencing, resulting in reduced representation genomic data (Table S4, Supporting information). Different filtering of the data rendered two main data sets of concatenated RAD loci: super-matrix STRICT that included all taxa and allowed for no missing data (11 426 loci containing 994 Kbp) and super-matrix RELAXED that excluded two low-coverage samples and allowed for up to 48% missing data (37 292 loci containing 3.24 Mbp; see Material and Methods).

Phylogenomic analyses of concatenated sequence data. The phylogenomic trees inferred by RAXML differ very little between super-matrix STRICT (Fig. S5, Supporting information) and super-matrix RELAXED (Fig. 4). The overall topology in the concatenation trees is similar to the coalescent species trees, with a Eurasian and an Afrocanarian clade. There are, however, some differences: The populations of La Palma and Libya are not inferred as sister taxa bifurcating from a long branch. Instead, La Palma branches off first, after which Libya shortly thereafter branches off from the ancestor of the remaining Afrocanarian populations (Figs 4 and S5, Supporting information). For super-matrix STRICT (Fig. S5, Supporting information), the central Canary Islands clade (clade B3) matches the topology of the species trees (cf. Fig. 2), whereas for super-matrix RELAXED, El Hierro is identified

as basal, after which La Gomera branches off, and finally Gran Canaria/Tenerife (the latter, however, not with full bootstrap support; Fig. 4).

In super-matrix STRICT, clade B4 is resolved with Algeria basal, followed in branching sequence by Morocco, and Fuerteventura/Lanzarote (Fig. S5, Supporting information), whereas for super-matrix RELAXED, the branching sequence mixes continental and insular populations (Fig. 4). Most nodes are supported with high bootstrap values (Figs 4 and S5, Supporting information).

For the RAD sequencing data, the within-species (within-sample) genetic variation for *C. t. degener* from Lanzarote and *C. t. teneriffae* from Tenerife is remarkably higher than for other samples (Fig. S5, Supporting information).

Cluster analyses. The multidimensional scaling analysis of the unfiltered SNP data set produced six clusters, corresponding to population groups from Eurasia (clade A), the central Canary Islands (clade B3, with Tenerife somewhat deviating), north-west Africa and eastern Canary Islands (except Lanzarote, which deviates from remaining clade B4), La Palma and Libya (Fig. S6, Supporting information).

Dating. Autosomal sequence dating based on relaxed settings for two calibration points (calibration set A), the intercontinental split between Holarctic and Nearctic *Poecile* species and the African/European split within *Periparus ater* following the Messinian salinity crisis, rendered the following mean age estimates for a partial taxon set excluding samples from Lanzarote and Tenerife: Clade AB (all *Cyanistes*) 4.4 My, clade B (*C. teneriffae*) 3.9 My, clade B1b (*C. teneriffae* except La Palma) 3.7 My or clade B1 (La Palma and Libya) 3.1 My, clade B3 (central Canary Islands) 1.8 My, clade B4 (eastern Canary Islands and north-west Africa) 2.0 My and clade A (*C. caeruleus* and *C. cyanus*) 3.0 My (Table 1). The full results of all replicate analyses are available in Table S5 (Supporting information).

There were some effects of genomic subset: In relation to autosomal subset 1, autosomal subsets 2 and 3 arrived at on average 13% and 6% higher node age estimates, respectively, whereas the Z chromosome subset produced on average 2% higher node ages, for the full taxon set. These differences were more pronounced for the partial taxon set: In relation to autosomal subset 1, autosomal subsets 2 and 3 arrived at on average 28% and 17% higher node age estimates, respectively, whereas the Z chromosome subset produced on average 7% higher node ages.

The effect of excluding two samples, owing to unexpectedly high within-individual genetic variation, had a larger effect on the estimated node ages (Table S9, Supporting information). The effect size is consistent over

all calibration sets, except that incorporating constraints of Canary Islands geological ages (A, B, D and E), but varies markedly between specific nodes. For nodes A and A1 (Eurasian *C. caeruleus* and *C. cyanus*), the effect is on average 10%; for nodes AB, B and B1/B1b 24%; 43% for node B2; 55% for node B3; and 72% for node B4 (Table S9, Supporting information).

The configuration of calibration priors affected the dating (of the partial taxon set) as follows: setting the calibration point of the Messinian salinity crisis to include only the period of actual desiccation of the Mediterranean basin (MSC strict), rather than relaxing the prior by only setting an upper time limit, increases the estimated node ages uniformly by on average 22% (Table S10, Supporting information). However, the difference in estimated node ages differs across nodes only by 2% between relaxed and strict priors for the trans-Beringian faunal interchange (Table S10, Supporting information). When employing a strict prior of the Messinian salinity crisis, adding a trans-Beringian prior has no effect at all (Table S10, Supporting information). Adding constraints according to the geological ages of El Hierro and La Palma rendered estimated node ages that are 21–31% lower, except for the central Canary Islands clade B3, which was 45% lower.

Finally, enforcing clade B1 according to the coalescence species tree topology had no effect on the estimated node ages of any other clade (data not shown).

Coalescent species tree analyses of SNPs. The two SNP-based species tree runs in SNAPP converged, and resulting data were combined, producing a final set of 15 639 trees. The coalescent analyses of the RAD SNP data set produced species trees which largely match the topology of the *BEAST species tree of the Sanger data (see Fig. 2b). The same major clades are recovered with high support, although within the clade B3 El Hierro is identified as basal but the following order of divergence is unresolved, and the divergence within clade B4 is unresolved. Species-specific branch lengths are short, owing to the removal of most autapomorphies.

Discussion

Irrespective of method and data set, our data show that the Afrocanarian blue tits are monophyletic and represent four major clades. The Afrocanarian complex has a continental origin, and the Canary Islands have been colonized three times (Fig. 2). Populations on La Palma and in Libya represent relicts of an ancestral North African population, either sharing a common ancestor separated from the other Afrocanarian lineages (Fig. 2) or diverging in rapid sequence from the African ancestor (Figs 4 and S5, Supporting information).

There has been much debate on the practice of concatenation into super-matrices vs coalescence-based species tree methods, which integrate the species tree over marker trees (see e.g. Gatesy & Springer 2013; McVay & Carstens 2013). Here, we performed both approaches, as well as coalescence-based species tree integration without inference of marker trees (Bryant *et al.* 2012). Results obtained with the different types of data sets (i.e. fewer long Sanger sequences vs many short RAD sequences, and SNPs derived thereof) produced consistent results with only slight differences (mainly regarding the exact relationship between the populations of La Palma and Libya; Figs 2 and 3).

The use of RAD sequencing for interspecific phylogenomic studies is still rare, and typically either phylogenetic analyses of sequence super-matrices (Eaton & Ree 2013; Wagner *et al.* 2013; Cruaud *et al.* 2014), SNP super-matrices (Jones *et al.* 2013) or coalescence species SNP trees (Gohli *et al.* 2015) have been performed, but with no evaluation across methods. Cruaud *et al.* (2014) also investigated the differences and similarities compared to traditional Sanger sequencing in beetles. Recent simulations of RAD sequence data show that if one concatenates markers, allowing for a substantial proportion of missing data – and thereby increase the number of loci included – it results in more accurate trees (Huang & Knowles 2014). However, we see little difference between data sets with and without tolerance for missing data (Figs 4 and S5, Supporting information). For the differences that exist, the super-matrix RELAXED which includes missing data (Fig. 4), deviates more from the coalescent species trees (Fig. 2) than does super-matrix STRICT, which includes no missing data.

Within the Afrocanarian blue tits, there is a strongly and uniformly supported clade comprising the north-west African populations and all Canary Islands, except La Palma (Figs 2 and 4), in agreement with Dietzen *et al.* (2008), Illera *et al.* (2011) and Päckert *et al.* (2013). This further splits into a shallow clade comprising the eastern Canary Islands and north-west Africa (clade B4), and a clade comprising the central Canary Islands (clade B3). The short branch lengths obtained within clade B4 suggests there is little differentiation among these populations, and indeed, there is not full reciprocal monophyly between the different populations in a mitochondrial marker tree (Fig. S1, Supporting information). This corresponds to the little plumage differentiation between the populations of Lanzarote and Fuerteventura (ssp. *degener*) compared to north-west African populations (ssp. *ultramarinus*) (Martin 1991); however, notable structural differences with shorter wings and larger bills and legs exist (Grant 1979). Clade B3 displays marked differentiation, with long branches, basal and unequivocal position of the El Hierro popula-

tion (ssp. *ombriosus*), which is also distinguished by its greenish rather than greyish back (Martin 1991), followed by sequential bifurcations (Figs 2 and 4). Most analyses support the bifurcation sequence El Hierro–Gran Canaria–La Gomera/Tenerife (Figs 2 and 4), but some conflicts were revealed in the concatenation analysis of super-matrix RELAXED, which included missing data (Fig. 4). This sequence is in accordance with current taxonomy with ssp. *hedwigae* on Gran Canaria, and ssp. *teneriffae* on Tenerife and La Gomera (Dietzen *et al.* 2008; Gill & Donsker 2014).

Basal to the main clade of continental African and insular populations were the populations from La Palma (ssp. *palmensis*), which is distinguished by whitish rather than yellowish belly (Martin 1991), and narrower bill than other populations on the Canary Islands (Grant 1979), and Libya (ssp. *cyrenaicae*). These were placed either as sister taxa (coalescence-based species trees; Fig. 2), as found by Päckert *et al.* (2013), or branching off independently in rapid sequence from the common ancestor of Afrocanarian blue tits (concatenation-based phylogenetic analyses; Figs 4 and S5, Supporting information). Gohli *et al.* (2015) recommended that the population of La Palma should be warranted species status, based on differentiation in sperm length (Gohli *et al.* 2015) in combination with previously established long-term isolation (Päckert *et al.* 2013), and morphological (Grant 1979; Martin 1991) and vocal (Schottler 1993, 1995) differentiation from other island populations. While this may be justified, we refrain from suggesting taxonomic revision of single populations, and recommend comprehensive analyses of all populations using Bayesian species delimitation methods (Jones *et al.* 2014; Yang & Rannala 2014).

Colonization of the Canary Islands

In general, passerines in the Canary Islands are considered sedentary, that is there are no dispersal events recorded among islands (Martín & Lorenzo 2001), and genetic data support such statement (Illera *et al.* 2012). In fact, for bird species occurring on multiple islands, there are several examples representing multiple independent colonization events from the continental mainland, for example *Erithacus rubecula* (Dietzen *et al.* 2003), *Regulus* spp. (Päckert *et al.* 2006), *Lanius meridionalis* (Padilla *et al.* 2015) and *Fringilla* spp. (Marshall & Baker 1999).

Based on morphology, Grant (1979) suggested a single colonization from the mainland, followed by a westward stepping-stone scenario. Kvist *et al.* (2005) and Dietzen *et al.* (2008) proposed an initial colonization of Tenerife, from which other islands were colonized (except La Palma, which Dietzen *et al.* (2008) suggested, was colonized separately from Europe). Illera *et al.*

(2011) and Päckert *et al.* (2013) recognized La Palma as a unique colonization event, but conclude that all other island population originate from a single colonization event to the central islands, followed by an eastward stepping-stone colonization.

Our tree topologies refute the above models, and suggest, in accord with Gohli *et al.* (2015), no less than three colonization events. The first one represents the population on La Palma, which – just as the Libyan population – splits from an ancestral Afrocanarian blue tit. The distribution of forested vegetation has varied considerably with the North African monsoon cycles over the past 8 My (Larrasoana *et al.* 2003, 2013). In cycles of 50–400 Ky intervals, ‘green Sahara periods’ typically developed over 2–3 Ky, lasted 4–8 Ky and diminished over 2–3 Ky, during which the vegetation was rapidly fragmented (Larrasoana *et al.* 2003, 2013). Such changes make plausible that African blue tits have periodically been widely distributed in North Africa, but retracted during dry periods in parallel with diminished forest cover. After such a retraction, the Libyan population remained isolated as a relict. Isolation following climate-driven forest fragmentation has also been suggested as the main driver for diversification in African forest robins (within Muscicapidae) during late Pliocene (Voelker *et al.* 2010).

The most ancient colonization event took place on La Palma, with an estimated age of 1.7–2.0 My (Guillou *et al.* 2001; Carracedo & Day 2002), which could imply an upper limit of the time of divergence (cf. dating with calibration set C, in which the La Palma population was constrained to a maximum age of 3.5 My, in accord with Päckert *et al.* (2013); Table 1). However, the genetic differentiation between the Afrocanarian populations is extensive, and dating, which does not incorporate island ages, estimates node ages considerably older than this limit (3.7–4.3 My for relaxed calibration set A and 4.3–5.4 My for strict calibration set B; Tables 1 and S5, Supporting information; cf. Päckert *et al.* (2013)). This could suggest a much accelerated molecular rate. Alternatively, the first colonization event initially reached older island(s) than La Palma (for instance, La Gomera and Tenerife are more than 10 My old). After its emergence, La Palma was then colonized from one of these older islands, and the population remained there, while populations on other island(s) went extinct or were outcompeted by a second colonization from the mainland. Yet another alternative could be that the differentiation of the La Palma lineage took place on the mainland before the emergence of La Palma. Subsequently, La Palma would have been colonized, whereas the source population would have gone extinct.

Our data infer that a second colonization (or third; see the discussion below about clade ages) occurred from

mainland Africa to the central Canary Islands. The population on El Hierro constitutes the basal single branch, but we argue that El Hierro is probably not the target of the primary colonization and a source of subsequent eastward stepping-stone colonization of the remaining central islands. We base this argumentation on three circumstances: (i) The age of the island is estimated to only 1.2 My (Carracedo & Day 2002), to compare with an estimated basal node (B3) age of 1.5–2.2 My for relaxed calibration set A and 1.8–2.7 My for strict calibration set B (Tables 1 and S5, Supporting information). (ii) The population on El Hierro displays very low current genetic diversity, which is four- to sixfold lower than that of the populations on the other central islands (Table S5, Supporting information; see also Illera *et al.* 2011). (iii) The low genetic diversity does not seem to be derived of any recent bottleneck, as the demographic reconstruction for El Hierro shows a long-term low population size, which has increased somewhat recently (Figs 3c and S3c, Supporting information).

A simple stepping-stone pathway from El Hierro is thus not plausible. We can only speculate about the exact pathway, but our results would favour a scenario where an initial colonization of one of the larger islands, for example Gran Canaria, was followed by independent colonizations from there to El Hierro (possibly via another island, from which the population later became extinct or outcompeted) and Tenerife, and from Tenerife further to La Gomera. Alternatively, analogous to parts of the suggestions by Kvist *et al.* (2005), Tenerife could have been the first island to be colonized, from which all other central islands were colonized: first El Hierro (possibly via e.g. La Gomera, from which the population later became extinct or outcompeted), second Gran Canaria and last La Gomera (a (re) colonization that would replace any previous remnant population). Such scenarios are, counter to common intuitive interpretation, entirely compatible with the species tree topology (Fig. 2), and we would contend that they represent parsimonious interpretations based on tree topology, genetic diversity (Tables S3, S7 and S8, Supporting information), demographic histories (Fig. 3) and geological history of the islands.

The third colonization (or second; see the discussion below about clade ages) reached the eastern islands of Fuerteventura and Lanzarote. There is a notable difference between species tree methods, in which clade B4 is very shallow (103 Ky, 95% HPD 66–149 Ky; Fig. 2a), and RAD-based concatenation analyses, in which the clade is dated to 1.3–2.9 My (Tables 1 and S5, Supporting information). The detection of bottlenecks in the eastern Canary Island populations, compatible with classical founder events, is also estimated to be very recent (27.5 Kya; Fig. 3a), more in concert with the species tree analyses.

Thus, coalescence analyses based on Sanger data and employed with a strict molecular clock arrive at a recent clade B4, while calibration point-based phylogenetic analyses of concatenated RAD sequence data, using a relaxed molecular clock, produce a considerably older estimated age. There are indications that this difference between coalescence species tree methods and concatenation analyses may be replicated: judging from relative node ages, Gohli *et al.* (2015) also arrived at a recent clade B4 based on coalescence species tree analyses of SNPs, whereas phylogenetic dating of concatenated sequence data in Päckert *et al.* (2013) rendered node ages of c. 1.4–2.7 My (depending on whether a Mediterranean island population is included). In fact, when the volcanic ages of El Hierro and La Palma are not employed as constraints, clade B4 is estimated marginally older than clade B3 (Tables 1 and S5, Supporting information).

It seems as though the reason for the drastically lower RAD sequencing success of the samples from Lanzarote and Tenerife (Table S4, Supporting information) may be owing to generally lower DNA quality. The comparably very large within-population (within-sample) genetic variation in Lanzarote and Tenerife (Fig. S5, Supporting information) does not reflect the very low genetic diversity in those populations (Tables S3, S7 and S8, Supporting information); likewise, the separation of the Lanzarote sample from remaining clade B4 and the Tenerife sample from clade B3 (Figs S5 and S6, Supporting information) deviates from the population-based results (Fig. 2a). This suggests that the lower DNA quality is manifested in the sequencing results as false genetic diversity. Although we are uncertain of by what mechanism sequencing errors may have been introduced, we regard the dating results following exclusion of those samples as more reliable.

In general, calibration without good fossil information introduces much uncertainty, and the variation in our dating results is in par with recent dated phylogenies (Päckert *et al.* 2013). It is worth pointing out that 95% highest posterior density intervals are large and overlap between different settings for the dating analyses (Table S5, Supporting information). The best practice in transforming geological events into reliable calibration points may also be debated. In the case of the Messinian salinity crisis (MSC), Päckert *et al.* (2013) used the time of desiccation of the Mediterranean basin as a maximum bound for the divergence between European and North African lineages of coal tit (but without any lower bound). While we agree that it is not unlikely that there may have been limited gene flow between the lineages even after the Mediterranean basin was refilled, the main splitting event should have co-occurred with the MSC. We therefore included dating analyses in which a strict prior, covering only the period of desiccation, was employed for

the MSC (calibration set B, and also D and E, which are similar; Tables 1 and S5, Supporting information).

Which are the source and sink populations?

Island populations are often considered marginal to the mainland distribution and implicitly assumed to be the result of colonization from a mainland source. Although this is probably true in the majority of cases, this view has been challenged lately, and there are now examples where island populations may have been acting as a source for reverse mainland colonization (Filardi & Moyle 2005; Bellemain *et al.* 2008; Jonsson *et al.* 2011). Although perhaps rare or at least infrequent, such back colonization events, from island to mainland, are important to document as they would support a potential role of islands as founders/sources of mainland systems as opposed to the traditional view of islands being single-ended evolutionary tracks (sinks). Indeed, back colonization demography underlies both the peripatric and the founder-flush speciation model (Mayr 1963; Powell 1978; Slatkin 1996), models that have been criticized as highly unlikely (Charlesworth & Smith 1982; Barton & Charlesworth 1984; Charlesworth 1997).

In the case of the Afrocanarian blue tit, the results in Illera *et al.* (2011) and Päckert *et al.* (2013) suggest such a back colonization event from the small eastern island populations to the large mainland populations in Morocco and Algeria, but this is not supported by our data. Instead we found (i) low levels of genetic variation in the island populations, contrasting the high variation found on mainland Africa (Tables S3, S7 and S8, Supporting information); this is also observed in a recent microsatellite data set (Hansson *et al.* 2014). Illera *et al.* (2011) report relatively higher genetic diversity in the Canary Island populations than in North Africa, but this seems to have a strong mitochondrial bias, as the one nuclear gene shows the opposite pattern. Further (ii), our tree-based analyses place the mainland populations basal in relation to the populations of the eastern Canary Islands (Figs 2, 4 and S5, Supporting information), in a sister clade to that of the central Canary Islands. This is in accord with Gohli *et al.* (2015), but differs from the terminal position of North African and the eastern islands, nested within the same clade as the central islands, reported by (Päckert *et al.* 2013). Finally (iii), our demographic reconstructions show clear signatures of very recent strong bottlenecks for the populations of Fuerteventura and Lanzarote (Fig. 3b), that is a sign compatible with a recent founder event, while no such patterns can be discerned for the mainland populations (Fig. 3a). Thus, our data do not support that the currently relatively large mainland populations have expanded rapidly after having been founded from the eastern Canary Islands.

Conclusion

The evolutionary history of the Afrocanarian blue tits has remained disputed despite intensive efforts to resolve it (Salzburger *et al.* 2002; Kvist *et al.* 2005; Dietzen *et al.* 2008; Illera *et al.* 2011; Päckert *et al.* 2013; Hansson *et al.* 2014; Gohli *et al.* 2015). Single-locus analyses, and multilocus concatenation analyses based on only few genes, with a bias towards mitochondrial markers, have produced conflicting results. In addition, a recent SNP-based study (Gohli *et al.* 2015) was restricted by deviating sampling of a key population, described relationships among groups as well as population demography with low effective sample sizes, and lack nodal support for the deeper nodes, making the evolutionary interpretations uncertain. Here, we combined extensive traditional sequencing and next-generation sequencing and demonstrate that the Canary Islands have been colonized from the African mainland at three separate occasions, after which substantial differentiation has occurred. The two peripheral populations on La Palma and in Libya, respectively, constitute relicts of a population ancestral to all Afrocanarian taxa. Further, our study strongly contradicts back colonization from the Canary Islands to the African mainland and instead conforms to the traditional view of islands most often acting as sinks. Based on the tree topology recovered by Illera *et al.* (2011) and Päckert *et al.* (2013), their interpretation that north-west Africa was back colonized from the Canary Islands may be justified. However, we argue that their analyses were flawed because they included very few nuclear loci, and because incomplete lineage sorting may have obscured true patterns. For resolving such complex evolutionary history, we argue that high genomic coverage as well as dense and comparable taxon sampling is necessary and that coalescence-based methods for species tree inference, which are most appropriate for population samples of multiple individuals, should routinely be explored. Therefore, reduced representation next-generation sequencing should forthcoming play an important role in phylogeography.

Acknowledgements

We thank Jens Hering and Martin Päckert, Senckenberg Natural History Collections, for providing samples of *C. t. cyrenaica* from Libya and *Periparus* spp.; Andrey Gavrilov, Chokpak Ornithological Station, and Mats Waern for providing *C. cyanus* samples from Kazakhstan; José Luis Copete and Ferran López for providing samples of *P. a. atlas* from Morocco; CHAGRA ringing group for providing us with *C. teneriffae* samples from Ceuta; Barboutis Christos for providing *C. caeruleus* samples from Crete; Matt Wood and the Edward Grey Institute at Oxford for providing samples of British *C. c. obscurus*; Olof

Hellgren and Kvismare Bird Observatory for providing samples of *Parus major* and *Poecile palustris*. Martí Franch provided the *Cyanistes* illustrations. The Regional Government of the Canary Islands, Andalusia and Madrid in Spain and the Government of Morocco gave permission to conduct fieldwork in these countries. We thank two anonymous reviewers for valuable comments on the manuscript. M.S. wishes to thank the Geneco Research School for training and support. This work was supported by the Swedish Research Council (621-2007-5381 and 621-2009-4945), the Oscar and Lili Lamm Foundation, the Craford Foundation, EU-FP7 (project Avian Genomics) (all to B.H.) and the Royal Physiographic Society (to M.S.).

References

- Baird NA, Etter PD, Atwood TS, *et al.* (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, **3**, e3376.
- Balke M, Ribera I, Hendrich L *et al.* (2009) New Guinea highland origin of a widespread arthropod supertramp. *Proceedings of the Royal Society B-Biological Sciences*, **276**, 2359–2367.
- Barton NH, Charlesworth B (1984) Genetic revolutions, founder effects, and speciation. *Annual Review of Ecology and Systematics*, **15**, 133–164.
- Bellemain E, Ricklefs RE (2008) Are islands the end of the colonization road? *Trends in Ecology & Evolution*, **23**, 461–468.
- Bellemain E, Bermingham E, Ricklefs RE (2008) The dynamic evolutionary history of the bananaquit (*Coereba flaveola*) in the Caribbean revealed by a multigene analysis. *BMC Evolutionary Biology*, **8**, 240.
- BirdLife International, NatureServe (2014) *Bird Species Distribution Maps of the World*. BirdLife International, Cambridge.
- Bouckaert R, Heled J, Kuhnert D *et al.* (2014) BEAST 2: a software platform for bayesian evolutionary analysis. *Plos Computational Biology*, **10**, e1003537.
- Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A (2012) Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, **29**, 1917–1932.
- Carracedo JC, Day S (2002) *Canary Islands*. Terra Publishing, Hertfordshire, London.
- Carrascal LM, Moreno E, Valido A (1994) Morphological evolution and changes in foraging behavior of island and mainland populations of Blue Tit (*Parus-Caeruleus*) - a test of convergence and ecomorphological hypotheses. *Evolutionary Ecology*, **8**, 25–35.
- Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) STACKS: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124–3140.
- Charlesworth B (1997) Is founder-flush speciation defensible? *American Naturalist*, **149**, 600–603.
- Charlesworth B, Smith DB (1982) A computer-model of speciation by founder effects. *Genetical Research*, **39**, 227–236.
- Corl A, Ellegren H (2013) Sampling strategies for species trees: the effects on phylogenetic inference of the number of genes, number of individuals, and whether loci are mitochondrial, sex-linked, or autosomal. *Molecular Phylogenetics and Evolution*, **67**, 358–366.
- Cramp S, Perrins C (1993) *Handbook of the Birds of Europe, the Middle East, and North Africa: The Birds of the Western Palearctic*. Oxford University Press, Oxford.
- Cruaud A, Gautier M, Galan M *et al.* (2014) Empirical assessment of RAD sequencing for interspecific phylogeny. *Molecular Biology and Evolution*, **31**, 1272–1274.
- Danecek P, Auton A, Abecasis G *et al.* (2011) The variant call format and VCFtools. *Bioinformatics*, **27**, 2156–2158.
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772.
- Davey JW, Hohenlohe PA, Etter PD, *et al.* (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, **12**, 499–510.
- Dietzen C, Witt H-H, Wink M (2003) The phylogeographic differentiation of the European robin *Erithacus rubecula* on the Canary Islands revealed by mitochondrial DNA sequence data and morphometrics: evidence for a new robin taxon on Gran Canaria? *Avian Science*, **3**, 115–131.
- Dietzen C, Garcia-del-Rey E, Castro GD, Wink M (2008) Phylogeography of the blue tit (*Parus teneriffae*-group) on the Canary Islands based on mitochondrial DNA sequence data and morphometrics. *Journal of Ornithology*, **149**, 1–12.
- Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *Plos Biology*, **4**, 699–710.
- Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.
- Eaton DAR, Ree RH (2013) Inferring phylogeny and introgression using RADseq Data: an example from flowering plants (*Pedicularis*: *Orobanchaceae*). *Systematic Biology*, **62**, 689–706.
- Eklom J, Galindo J (2011) Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity (Edinb)*, **107**, 1–15.
- Ellegren H (2007) Molecular evolutionary genomics of birds. *Cytogenetic and Genome Research*, **117**, 120–130.
- Etter PD, Bassham S, Hohenlohe PA, Johnson EA, Cresko WA (2011) SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Methods in Molecular Biology*, **772**, 157–178.
- Fernández-Palacios JM, de Nascimento L, Otto R *et al.* (2011) A reconstruction of Palaeo-Macaronesia, with particular reference to the long-term biogeography of the Atlantic island laurel forests. *Journal of Biogeography*, **38**, 226–246.
- Filardi CE, Moyle RG (2005) Single origin of a pan-Pacific bird group and upstream colonization of Australasia. *Nature*, **438**, 216–219.
- Gatesy J, Springer MS (2013) Concatenation versus coalescence versus “concordance”. *Proceedings of the National Academy of Sciences, USA*, **110**, E1179.
- Gill FB, Donsker D (2014) IOC World Bird List (v 4.3). Available from: <http://www.worldbirdnames.org/>.
- Gillespie RG, Croom HB, Palumbi SR (1994) Multiple origins of a spider radiation in Hawaii. *Proceedings of the National Academy of Sciences, USA*, **91**, 2290–2294.
- Gladnikov AY, Oleinik AE, Marincovich L, Barinov KB (2002) A refined age for the earliest opening of Bering Strait. *Palaeogeography Palaeoclimatology Palaeoecology*, **183**, 321–328.

- Gohli J, Leder EH, Garcia-del-Rey E *et al.* (2015) The evolutionary history of Afrocanarian blue tits inferred from genome-wide SNPs. *Molecular Ecology*, **24**, 180–191.
- Gosler A, Clement P (2007) Family Paridae (tits and chickadees). In: *Handbook of the Birds of the World: Vol. 12: Picathartes to Tits and Chickadees* (eds del Hoyo J, Elliott A, Christie DA), pp. 662–750. Lynx Edicions, Barcelona.
- Grant PR (1979) Ecological and morphological variation of Canary Island blue tits, *Parus-Caeruleus* (Aves, Paridae). *Biological Journal of the Linnean Society*, **11**, 103–129.
- Grant PR, Grant BR (2011) *How and Why Species Multiply: The Radiation of Darwin's Finches*. Princeton University Press, Princeton, New Jersey.
- Guillou H, Carracedo JC, Duncan RA (2001) K-Ar, Ar-40-Ar-39 ages and magnetostratigraphy of Brunhes and Matuyama lava sequences from La Palma Island. *Journal of Volcanology and Geothermal Research*, **106**, 175–194.
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.
- Hansson B, Ljungqvist M, Illera JC, Kvist L (2014) Pronounced fixation, strong population differentiation and complex population history in the Canary Islands blue tit subspecies complex. *PLoS ONE*, **9**, e90186.
- Heled J, Drummond AJ (2008) Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology*, **8**, 289.
- Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570–580.
- Huang H, Knowles LL (2014) Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Systematic Biology*. doi: 10.1093/sysbio/syu046.
- Illera JC, Koivula K, Broggi J *et al.* (2011) A multi-gene approach reveals a complex evolutionary history in the Cyanistes species group. *Molecular Ecology*, **20**, 4123–4139.
- Illera JC, Rando JC, Richardson DS, Emerson BC (2012) Age, origins and extinctions of the avifauna of Macaronesia: a synthesis of phylogenetic and fossil information. *Quaternary Science Reviews*, **50**, 14–22.
- Johansson US, Ekman J, Bowie RC *et al.* (2013) A complete multilocus species phylogeny of the tits and chickadees (Aves: Paridae). *Molecular Phylogenetics and Evolution*, **69**, 852–860.
- Jones JC, Fan SH, Franchini P, Scharl M, Meyer A (2013) The evolutionary history of Xiphophorus fish and their sexually selected sword: a genome-wide approach using restriction site-associated DNA sequencing. *Molecular Ecology*, **22**, 2986–3001.
- Jones G, Aydin Z, Oxelman B (2014) DISSECT: an assignment-free Bayesian discovery method for species delimitation under the multispecies coalescent. *Bioinformatics*. doi: 10.1093/bioinformatics/btu770.
- Jonsson KA, Bowie RCK, Moyle RG *et al.* (2010) Historical biogeography of an Indo-Pacific passerine bird family (Pachycephalidae): different colonization patterns in the Indonesian and Melanesian archipelagos. *Journal of Biogeography*, **37**, 245–257.
- Jonsson KA, Fabre PH, Ricklefs RE, Fjeldsa J (2011) Major global radiation of corvid birds originated in the proto-Papuan archipelago. *Proceedings of the National Academy of Sciences, USA*, **108**, 2328–2333.
- Juan C, Emerson BC, Oromi P, Hewitt GM (2000) Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology & Evolution*, **15**, 104–109.
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**, 772–780.
- Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS (1999) Chronology, causes and progression of the Messinian salinity crisis. *Nature*, **400**, 652–655.
- Kvist L, Broggi J, Illera JC, Koivula K (2005) Colonisation and diversification of the blue tits (*Parus caeruleus teneriffae*-group) in the Canary Islands. *Molecular Phylogenetics and Evolution*, **34**, 501–511.
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**, 357–359.
- Larrasoana JC, Roberts AP, Rohling EJ, Winkhofer M, Wehausen R (2003) Three million years of monsoon variability over the northern Sahara. *Climate Dynamics*, **21**, 689–698.
- Larrasoana JC, Roberts AP, Rohling EJ (2013) Dynamics of green Sahara periods and their role in hominin evolution. *PLoS ONE*, **8**, e76514.
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.
- Losos JB (2011) *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles*. University of California Press, Oakland, California.
- Marincovich L, Gladenkov AY (1999) Evidence for an early opening of the Bering Strait. *Nature*, **397**, 149–151.
- Marshall HD, Baker AJ (1999) Colonization history of Atlantic island common chaffinches (*Fringilla coelebs*) revealed by mitochondrial DNA. *Molecular Phylogenetics and Evolution*, **11**, 201–212.
- Martin JL (1991) Patterns and significance of geographical variation in the Blue Tit (*Parus caeruleus*). *Auk*, **108**, 820–832.
- Martín A, Lorenzo JA (2001) *Aves del archipiélago canario*. Francisco Lemus Editor, San Cristobal De La Laguna.
- Mayr E (1963) *Animal Species and Evolution*. Harvard University Press, Cambridge, Massachusetts.
- McVay JD, Carstens BC (2013) Phylogenetic model choice: justifying a species tree or concatenation analysis. *Journal of Phylogenetics and Evolutionary Biology*, **1**, 114.
- Medail F, Quezel P (1999) Biodiversity hotspots in the Mediterranean basin: setting global conservation priorities. *Conservation Biology*, **13**, 1510–1513.
- Melo M, Warren BH, Jones PJ (2011) Rapid parallel evolution of aberrant traits in the diversification of the Gulf of Guinea white-eyes (Aves, Zosteropidae). *Molecular Ecology*, **20**, 4953–4967.
- Päckert M, Dietzen C, Martens J, Wink M, Kvist L (2006) Radiation of Atlantic goldcrests *Regulus regulus* spp.: evidence of a new taxon from the Canary Islands. *Journal of Avian Biology*, **37**, 364–380.
- Päckert M, Martens J, Tietze DT *et al.* (2007) Calibration of a molecular clock in tits (Paridae)—do nucleotide substitution rates of mitochondrial genes deviate from the 2% rule? *Molecular Phylogenetics and Evolution*, **44**, 1–14.
- Päckert M, Martens J, Sun YH *et al.* (2011) Horizontal and elevational phylogeographic patterns of Himalayan and South-

- east 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 DP, 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.
- Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, **20**, 289–290.
- Partridge L, Pringmill F (1977) Canary Island Blue Tits and English coal tits – convergent evolution. *Evolution*, **31**, 657–665.
- Patel S, Kimball RT, Braun EL (2013) Error in phylogenetic estimation for bushes in the tree of life. *Journal of Phylogenetics and Evolutionary Biology*, **1**, 110.
- Pattengale ND, Alipour M, Bininda-Emonds ORP, Moret BME, Stamatakis A (2009) How many bootstrap replicates are necessary? *Research in Computational Molecular Biology, Proceedings*, **5541**, 184–200.
- Powell JR (1978) The founder-flush speciation theory: an experimental approach. *Evolution*, **32**, 465–474.
- Purcell S, Chang C (2014) *Plink 1.9beta2c*, Available from: <https://www.cog-genomics.org/plink2>.
- R Core Team (2013) *R: A Language and Environment for Statistical Computing*. Available from: <http://www.R-project.org>
- Rambaut A, Drummond AJ (2007) *Tracer v1*. Available from: <http://tree.bio.ed.ac.uk/software/tracer/>.
- Salzburger W, Martens J, Sturmbauer C (2002) Paraphyly of the Blue Tit (*Parus caeruleus*) suggested from cytochrome b sequences. *Molecular Phylogenetics and Evolution*, **24**, 19–25.
- Sambrook J, Russel DW (2001) *Molecular Cloning A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Sangster G (2006) The taxonomic status of ‘phylogroups’ in the *Parus teneriffae* complex (Aves): comments on the paper by Kvist et al. (2005). *Molecular Phylogenetics and Evolution*, **38**, 288–289; author reply 290.
- Savolainen V, Anstett MC, Lexer C et al. (2006) Sympatric speciation in palms on an oceanic island. *Nature*, **441**, 210–213.
- Schottler B (1993) Canary Islands Blue Tits (*Parus caeruleus* ssp) – differences and variation in territorial song – preliminary results. *Boletim do Museu Municipal do Funchal, (Suppl 2)*, 273–277.
- Schottler B (1995) Songs of blue tits *Parus caeruleus* palmensis from La Palma (Canary Islands): a test of hypotheses. *Bioacoustics*, **6**, 135–152.
- Schwarz G (1978) Estimating dimension of a model. *Annals of Statistics*, **6**, 461–464.
- Slatkin M (1996) In defense of founder-flush theories of speciation. *American Naturalist*, **147**, 493–505.
- Stamatakis A (2014) RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**, 1312–1313.
- Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, **73**, 1162–1169.
- Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978–989.
- Voelker G, Outlaw RK, Bowie RCK (2010) Pliocene forest dynamics as a primary driver of African bird speciation. *Global Ecology and Biogeography*, **19**, 111–121.
- Wagner CE, Keller I, Wittwer S et al. (2013) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, **22**, 787–798.
- Warren BH, Bermingham E, Bourgeois Y et al. (2012) Hybridization and Barriers to Gene Flow in an Island Bird Radiation. *Evolution*, **66**, 1490–1505.
- Warren BH, Simberloff D, Ricklefs RE et al. (2015) Islands as model systems in ecology and evolution: prospects fifty years after MacArthur-Wilson. *Ecology Letters*, **18**, 200–217.
- Weir JT, Schluter D (2008) Calibrating the avian molecular clock. *Molecular Ecology*, **17**, 2321–2328.
- Wickham H (2009) *ggplot2: Elegant Graphics for Data Analysis*. Springer, New York.
- Yang Z, Rannala B (2014) Unguided species delimitation using DNA sequence data from multiple loci. *Molecular Biology and Evolution*, **31**, 3125–3135.

M.S. and B.H. conceived and designed the study with input from J.C.I. and L.K.; J.C.I. carried out the major parts of the field work and sample collection, with additions from L.K.; M.S. performed all laboratory work, with assistance from P.B. (Sanger sequencing) and N.K. and P.P. (RAD library setup); M.S. performed all analyses with input from B.H., S.B. and J.C.I.; M.S., B.H. and J.C.I. wrote the manuscript with input from the other authors, all of which have approved of the final version.

Data accessibility

Sanger sequences have been deposited at GenBank with accession nos KP757995–KP759278 (except 24 sequences <200 bp, which are not accepted by GenBank; these are available from the alignment files), and RAD sequencing data (quality filtered, aggregated per individual) have been deposited at the NCBI Sequence Read Archive with accession nos ranging SRX879656–SRX883383 under BioProject 275690, for details see Table S1 (Supporting information). Data sets (the filtered SNP data set, aligned Sanger sequence data sets, and RAD sequence super-matrices), tree files (from *BEAST Sanger sequence analyses, SNAPP analyses of RAD sequencing SNPs and RAXML analyses of concatenated RAD sequences), an analysis log, and custom scripts are available at DataDryad, doi:10.5061/dryad.67dn4.

Supporting information

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

Appendix S1 Methods.

Table S1 Sample details and accession nos.

Table S2 Marker and primer details, PCR settings etc.

Table S3 Nucleotide diversity within populations.

Table S4 Summary of results from the restriction-site-associated (RAD) sequencing.

Table S5 Full dating results from five calibration sets, two taxon sets, and four genomic subsets.

Table S6 Substitution models for the 20 Sanger sequencing markers.

Table S7 Haplotype diversity within populations.

Table S8 Theta within populations.

Table S9 Summary of the dating analyses, highlighting the difference in estimated node ages between the partial and the full taxon set.

Table S10 Summary of the dating analyses, highlighting the difference in estimated node ages between the five different calibration sets.

Fig. S1 Phylogenetic tree for the mitochondrial control region.

Fig. S2 Coalescence based species trees based on analyses of population samples of 18 nuclear Sanger sequence markers (but no mitochondrial markers).

Fig. S3 Bayesian Skyline Plots based on multilocus Sanger sequence data, covering the full time period of demographic reconstructions.

Fig. S4 Bayesian Skyline Plots based on multilocus Sanger sequence data, covering the last 500 000 years.

Fig. S5 Phylogenomic maximum-likelihood tree of super-matrix *STRICT*.

Fig. S6 Cluster assignment using multi-dimensional scaling.

Disentangling the complex evolutionary history of the Western Palearctic blue tits (*Cyanistes* spp.) – phylogenomic analyses suggest radiation by multiple colonization events and subsequent isolation

published in *Molecular Ecology*, 2015
doi: 10.1111/mec.13145

MARTIN STERVANDER,* JUAN CARLOS ILLERA,† LAURA KVIST,‡ PEDRO BARBOSA,* NAOMI P. KEEHNEN,*¹ PETER PRUISSCHER,*¹ STAFFAN BENSCH* and BENGT HANSSON*

**Molecular Ecology and Evolution Lab, Dept of Biology, Lund University, Ecology Building, 223 62 Lund, Sweden, †Research Unit of Biodiversity (UO-CSIC-PA), Oviedo University, Campus of Mieres, Research Building, 5th Floor. C/ Gonzalo Gutiérrez Quirós, s/n, 33600 Mieres, Asturias, Spain, ‡Department of Biology, POB 3000, University of Oulu, 90014 Oulu, Finland, ¹Present affiliation: Population genetics, Dept of Zoology, Stockholm University, 106 91 Stockholm, Sweden*

Correspondence: Martin Stervander, Fax: +46 46 222 4716; E-mail: martin.stervander@biol.lu.se

Appendix S1 – Supplementary methods

Phasing

We initially chose to infer haplotypes for all *Cyanistes* taxa as one dataset, despite these being more or less separated populations of varying size and demographic history, while haplotypes of the outgroup *Parus major* were inferred separately. However, a comparison gave at hand that inference of the whole dataset increased phase probabilities for both ingroup and outgroup. Most markers included individuals with partial sequences, and for computational reasons the alignments were then divided in blocks of full, or almost full, sequences only, and the part of the full alignment containing partial sequences, respectively.

We ran the phasing algorithm at least 1–3 times, with 100 burn-in iterations and 100 iterations (thinning interval 1), starting with randomized seeds. We used a model that allowed for recombination, using default values for the initial estimate of the recombination parameter ρ to 0.0004 (see Li and Stephens (2003)). Phases were considered true if the overall phase probability was >0.95 .

RAD library preparation

We prepared paired-end restriction site-associated DNA (RAD) libraries following Baird *et al.* (2008) and Etter *et al.* (2011) with the following modifications. We considered no RNase A treatment necessary; DNA fragments were sheared with a Bioruptor Standard UCD-200 (Diagenode) and selected to a size range comprising 150–750 bp (length excluding adapters); PCR reactions were set up at 50 μ l with 16 μ l RAD library template running 16 cycles. Such protocol, with more library template and fewer cycles, leads to more P2–P2 fragments but fewer PCR clones. We used the restriction enzyme *SbfI* and multiplexed 16 samples per library; each library was then sequenced on one HiSeq 2000 (Illumina) channel at BGI, Hong Kong.

Supplementary results

- Table S1: Sample details and accession numbers.
- Table S2: Marker and primer details, PCR settings etc.
- Table S3: Nucleotide diversity within populations.
- Table S4: Summary of results from the Restriction site Associated DNA (RAD) sequencing.
- Table S5: Full dating results from five calibration sets, two taxon sets, and four genomic subsets.
- Table S6: Substitution models for the 20 Sanger sequencing markers.
- Table S7: Haplotype diversity within populations.
- Table S8: Theta within populations.

- Table S9: Summary of the dating analyses, highlighting the difference in estimated node ages between the partial and the full taxon set.
- Table S10: Summary of the dating analyses, highlighting the difference in estimated node ages between the five different calibration sets.
- Figure S1: Phylogenetic tree for the mitochondrial control region.
- Figure S2: Coalescence based species trees based on analyses of population samples of 18 nuclear Sanger sequence markers (but no mitochondrial markers).
- Figure S3: Bayesian Skyline Plots based on multi-locus Sangers sequence data, covering the full time period of demographic reconstructions.
- Figure S4: Bayesian Skyline Plots based on multi-locus Sangers sequence data, covering the last 500,000 years.
- Figure S5: Phylogenomic maximum likelihood tree of super-matrix STRICT.
- Figure S6: Cluster assignment using multi-dimensional scaling.

Table S1. Details and accession numbers for core samples. "short" means GenBank does not accept deposition since the sequence is <200 bp.

Sample	Taxon	Region/ country	Island/site	Year	Identifier	Ring no	Mitochondrial		ATP6AP2	BTF3L4	DNAJA2	ENSGALT 01183	ENSGALT 01658
							Cyt b	control region					
CLP1	<i>C. teneriffae palmensis</i>	Canary Is	La Palma	2003	ST336	839579	KP759217	KP759244	KP758459	KP758077	KP758804	KP758915	KP759059
CLP2	<i>C. teneriffae palmensis</i>	Canary Is	La Palma	2003	ST338	839581	KP759226	KP759181	KP758882	KP758682	KP758977	KP759131	KP758504
CLP3	<i>C. teneriffae palmensis</i>	Canary Is	La Palma	2003	ST340	839583	KP759206	KP759193	KP758742	KP758125	KP758466	KP758750	KP758237
CLP4	<i>C. teneriffae palmensis</i>	Canary Is	La Palma	2009	ST1818	N636281	KP759260	KP759257	KP758636	KP758167	KP758629	KP758940	KP759032
CLP5	<i>C. teneriffae palmensis</i>	Canary Is	La Palma	2009	ST1819	N636280	KP759186	KP759183	KP758757	KP758644	KP758441	KP758869	KP759065
CEH1	<i>C. teneriffae ombriosus</i>	Canary Is	El Hierro	2003	ST347	839590	KP759167	KP759196	KP758482	KP758592	KP759066	KP758645	KP758744
CEH2	<i>C. teneriffae ombriosus</i>	Canary Is	El Hierro	2003	ST348	839591	KP759201	KP759224	KP758656	KP758032	KP759130	KP758427	KP758994
CEH3	<i>C. teneriffae ombriosus</i>	Canary Is	El Hierro	2003	ST349	839592	KP759241	KP759245	KP758943	KP758642	KP758322	KP759126	KP758557
CEH4	<i>C. teneriffae ombriosus</i>	Canary Is	El Hierro	2003	ST351	839594	KP759267	KP759238	KP758044	KP759091	KP759120	KP758194	KP758357
CEH5	<i>C. teneriffae ombriosus</i>	Canary Is	El Hierro	2003	ST352	839595	KP759194	KP759163	KP758641		KP758925	KP758558	KP758924
CLG1	<i>C. teneriffae teneriffae</i>	Canary Is	La Gomera	2004	ST699	L254946	KP759216	KP759175	KP758971	KP758318	KP759031	KP758881	KP758754
CLG2	<i>C. teneriffae teneriffae</i>	Canary Is	La Gomera	2004	ST701	L254945	KP759166	KP759239	KP758153	KP758688	KP758817	KP759081	KP758985
CLG3	<i>C. teneriffae teneriffae</i>	Canary Is	La Gomera	2007	ST1698	EX8256	KP759171		KP758917	KP758040	KP758823	KP758740	KP758843
CLG4b	<i>C. teneriffae teneriffae</i>	Canary Is	La Gomera	2009	ST1815	N861959	KP759165		KP758098		KP758991	KP758684	KP758548
CLG5	<i>C. teneriffae teneriffae</i>	Canary Is	La Gomera	2009	ST1817	N861962	KP759202	KP759255	KP758474	KP758732	KP758008	KP759023	KP758515
CTE1	<i>C. teneriffae teneriffae</i>	Canary Is	Tenerife	2003	ST329	839572	KP759247	KP759187	KP759141	KP758447	KP758529	KP758161	KP758711
CTE2	<i>C. teneriffae teneriffae</i>	Canary Is	Tenerife	2003	ST331	839574	KP759210	KP759243	KP758317	KP758326	KP759107	KP758303	KP758879
CTE3	<i>C. teneriffae teneriffae</i>	Canary Is	Tenerife	2003	ST332	839575	KP759192	KP759182	KP758330	KP758708	KP759044	KP758468	KP758987
CTE4	<i>C. teneriffae teneriffae</i>	Canary Is	Tenerife	2003	ST333	839576	KP759263	KP759266	KP758778	KP758533	KP758371	KP758597	KP758503
CTE5	<i>C. teneriffae teneriffae</i>	Canary Is	Tenerife	2008	ST1616	CN7596	KP759259	KP759261	KP758927	KP758846	KP758524	KP758234	KP759136
CGC1	<i>C. teneriffae hedwigae</i>	Canary Is	Gran Canaria	2003	ST514	L254921	KP759233	KP759168	KP759155	KP758417	KP758671	KP758856	KP758131
CGC2	<i>C. teneriffae hedwigae</i>	Canary Is	Gran Canaria	2003	ST516	AZ0640	KP759250	KP759212	KP758564	KP758051	KP758062	KP758471	KP758438
CGC3	<i>C. teneriffae hedwigae</i>	Canary Is	Gran Canaria	2003	ST517	L254923	KP759205	KP759219	KP758329	KP758462	KP758521	KP758964	KP758488
CGC4	<i>C. teneriffae hedwigae</i>	Canary Is	Gran Canaria	2004	ST697	L254932	KP759269	KP759240	KP758246	KP758298	KP758498	KP758345	KP758382
CGC5	<i>C. teneriffae hedwigae</i>	Canary Is	Gran Canaria	2004	ST698	L254931	KP759253	KP759248	KP758405	KP758132	KP758007	KP758965	KP758810
CFU1	<i>C. teneriffae degener</i>	Canary Is	Fuerteventura	2003	ST417	L254913	KP759251	KP759184	KP758526	KP758354	KP758109	KP758523	KP758277
CFU2	<i>C. teneriffae degener</i>	Canary Is	Fuerteventura	2003	ST420	L254917	KP759220	KP759215	KP758388	KP758978	KP758563	KP758745	KP758848
CFU3	<i>C. teneriffae degener</i>	Canary Is	Fuerteventura	2004	ST693	L254942	KP759176	KP759178	KP758393	KP758082	KP759012	KP758195	KP758842
CFU4	<i>C. teneriffae degener</i>	Canary Is	Fuerteventura	2004	ST695	L254937	KP759221	KP759211	KP758169	KP758934	short	KP759038	KP758497
CFU5	<i>C. teneriffae degener</i>	Canary Is	Fuerteventura	2004	ST696	L254940	KP759164	KP759225	KP758036	KP759092	KP758600	KP758385	KP758574
CLA1	<i>C. teneriffae degener</i>	Canary Is	Lanzarote	2004	ST688	L254943	KP759188	KP759169	KP758709	KP758480	KP758384	KP758190	KP758069
CLA2	<i>C. teneriffae degener</i>	Canary Is	Lanzarote	2004	ST689	L254936	KP759162	KP759265	KP759005	KP757997	KP758116	KP759115	KP758163
CLA3	<i>C. teneriffae degener</i>	Canary Is	Lanzarote	2004	ST690	334875	KP759179	KP759234	KP758803	KP758244	KP758650	KP758344	KP758055
CLA4	<i>C. teneriffae degener</i>	Canary Is	Lanzarote	2008	ST1649	CH7249	KP759262	KP759161	KP759085	KP758599	KP758391	KP758990	KP758643
CLA5	<i>C. teneriffae degener</i>	Canary Is	Lanzarote	2008	ST1651	CH7253	KP759275	KP759200	KP758707	KP758832	KP758324	KP759009	KP758675
MCE1	<i>C. teneriffae ultramarinus</i>	North Africa	Ceuta	2004	ST898		KP759249	KP759246	KP759102		KP759007	KP758587	KP759040
MCE2	<i>C. teneriffae ultramarinus</i>	North Africa	Ceuta	2004	ST899		KP759276	KP759190	KP758652	KP758455	short	KP758073	KP758387
MCE3	<i>C. teneriffae ultramarinus</i>	North Africa	Ceuta	2004	ST902		KP759271	KP759277	KP758958	KP758173	KP758895	KP758904	KP758293
MCE4	<i>C. teneriffae ultramarinus</i>	North Africa	Ceuta	2004	ST905		KP759231	KP759274	KP758364	KP758225	KP758765	KP758527	KP759061
MCE5	<i>C. teneriffae ultramarinus</i>	North Africa	Ceuta	2004	ST906		KP759170	KP759242	KP758872	KP758271	KP758402	KP758939	KP758920

Table S1, continued.

Sample	Nuclear Sanger markers													RAD data (clean)
	ENSGALT 05087	ENSGALT 11836	ENSGALT 15724	ENSGALT 16264	ENSGALT 25924	LDH-B	OGDHL	PDCD6	RPL7	SPAG9	TGFb2	UBE3A	WDR12	
CLP1	KP758009	KP759055	KP758594	KP758808	KP758813	KP758035	KP758429	KP758432	KP758118	KP758028	KP758415	KP758539	KP758269	SRX879656
CLP2	KP758204	KP758502	KP758795	KP758933	KP758530	KP758894	KP758544	KP758460	KP758660	KP758436	KP758694	KP758634	KP758258	
CLP3	KP758369	KP758245	KP758383	KP758079	KP758519	KP758615	KP758198	KP758284	KP758786	KP758905	KP758721	KP758452	KP758486	
CLP4	KP758473	KP758885	KP758751	KP758160	KP758247	KP758581	KP758433	KP758179	KP758734	KP758945	KP759071	KP759150	KP758792	
CLP5	KP758332	KP758886	KP758372	KP758550	KP758265		KP758451	KP758796	KP758260	KP758240	KP758761	KP758833	short	
CEH1	KP758847	KP758360	KP758970	KP758076	KP758982	KP758508	KP758998	KP758936	KP758755	KP758252	KP758921	KP759068	KP758617	SRX879658
CEH2	KP758716	KP758674	KP758413	KP759101	KP758370	KP758825	KP758389	KP758431	KP758214	KP758321	KP758487	KP759037	KP758249	
CEH3	KP759058	KP758580	KP758352	KP758510	KP758864	KP758760	KP758865	KP759127	KP758023	KP758386	KP758845	KP758821	KP758454	
CEH4	KP758238	KP758787	KP758220	KP758710	KP758232	KP758094	KP758011	KP758589	KP758835	KP758702	KP758986	KP758791	KP758701	
CEH5	KP758513	KP758483	KP758178	KP758334	KP758426	KP758481	KP758492	KP758627	KP758010	KP759027	KP759087	KP758756	KP759073	
CLG1	KP758203	KP758876	KP759028	KP758610	KP758562	KP758399	KP758561	KP758443	KP758714	KP758794	KP759021	KP759074	KP758071	SRX879659
CLG2	KP758880	KP758603	KP758122	KP758059	KP758632	KP759036	KP759056	KP758829	KP758718	KP758511	KP758255	KP758297	KP758262	
CLG3	KP758449	KP758816	KP758748	KP758918	KP758661	KP758199	KP758719	KP759079	KP758067	KP758900	KP758693	KP758091	KP758568	
CLG4b	KP758543	KP758027	KP758646	KP758651	KP758573	KP759051	KP758836	KP758068	KP758229	KP758884	KP758572	KP758570	KP758053	
CLG5	KP758640	KP759097	KP758528	KP758221	KP758000	KP758928	KP758045	KP759019	KP759129	KP758175	KP758937	KP759143	KP758003	
CTE1	KP758948	KP759029	KP758026	KP758100	KP757999	KP758689	KP758086	KP758283	KP758299	KP759152	KP759033	KP758659	KP758654	SRX879660
CTE2	KP758039	KP758584	KP759117	KP758875	KP758090	KP758667	KP758478	KP758087	KP758147	KP758901	KP758030	KP758509	KP758211	
CTE3	KP758815	KP758309	KP758522	KP758565	KP759057	KP758896	KP758085	KP758538	KP759002	KP758259	KP758070	KP759062	KP758282	
CTE4	KP758281	KP758205	KP758713	KP758909	KP758859	KP758108	KP758961	KP758726	KP758499	KP758619	KP759053	KP758278	KP758806	
CTE5	KP758111	KP758690	KP758892	KP758637	KP758601	KP758374	KP759132	KP758477	KP759000	KP758738	KP759022	KP758285	KP758942	
CGC1	KP758906	KP758608	KP758144	KP758788	KP759017	KP758569	KP758362		KP759063	KP759083	KP758453	KP758780	KP758061	SRX883380
CGC2	KP758741	KP758668	KP758931	KP758239	KP759148	KP759010	KP758653	KP758470	KP758762	KP758546	KP758192	KP758434	KP758554	
CGC3	KP758117	KP758975	KP758731	KP759157	KP758311	KP758021	KP758423	KP758866	KP758450	KP758551	KP758871	KP758712	KP759145	
CGC4	KP758253	KP758197	KP758403	KP758201	KP758962	KP758595	KP758950	KP758250	KP758319	KP758749	KP758868	KP758542	KP758103	
CGC5	KP759125	KP758997	KP758302	KP758495	KP758578	KP758607	KP758893	KP758401	KP758410	KP758110	KP759043	KP758802	KP758294	
CFU1	KP758242	KP758141	KP758941	KP759100	KP759149	KP758049	KP758897	KP758857	KP758849	KP758902	KP759113	KP758621	KP758723	SRX879662
CFU2	KP759123	KP758752	KP758930	KP759144	short	KP758809	KP758696	KP759111	KP758138	KP758156	KP758692	KP759069	KP759088	
CFU3	KP758016	KP758703	KP758729	KP758805	KP758270	KP758938	KP759138	KP758912	KP758807	KP758789	KP758724	KP759114	KP759020	
CFU4	KP758922	KP758852	KP758394	KP758669	KP758365	KP759076	KP758764	KP758566	KP758162	KP758699	KP759142	KP758323	KP759004	
CFU5	KP758839	KP758022	KP758944	KP758779	KP758625	KP758968	KP758626	KP758926	KP758313	KP758257	KP758952	KP758844	KP758411	
CLA1	KP758469	KP758442	KP759078	KP758812	KP758491	KP758614	KP758670	KP759108	KP758899	KP758157	KP758379	KP758560	KP758514	SRX879663
CLA2	KP758200	KP758101	KP758874	KP758500	KP758146	KP758490	KP758308	KP758300	KP758736	KP758272	KP759039	KP758150	KP758730	
CLA3	KP758720	KP758191	KP758687	KP758630	KP758054	KP758209	KP758456	KP758484	KP758377	KP758923	KP759098	KP758678	KP758043	
CLA4	KP759016	KP758217	KP758623	KP758166	KP758149	KP758602	KP759049	KP758056	KP758215	KP758081	KP758695	KP758029	KP758342	
CLA5	KP758375	KP758525	KP758579	KP758919	KP759124	KP759015	KP758686	KP758112	KP758505	KP758339	KP759146	KP758152	KP758556	
MCE1	KP758819	KP758133	KP758288	KP758398	KP758673	KP758180	KP758628	KP758955	KP758210	KP758631	KP758559	KP758031	KP758867	SRX879664
MCE2	KP758341	KP758883	KP758773	KP757996	KP758840	KP758889	KP758691	KP758066	KP758353	KP759080	short	KP759140	KP758664	
MCE3	KP758328	KP758799	KP758392	KP758467	KP758037	KP758911	KP758213	KP758775	KP758251	KP758873	KP758350	KP758320	KP758181	
MCE4	KP758254	KP758414	KP758959	KP759110	KP758737	KP758611	KP758494	KP758493	KP758208	KP758706	KP758635	KP759060	KP758910	
MCE5	KP758126	KP758983		KP758768	KP758665	KP758018	KP758378	KP758767	KP758002	KP759147	KP758184	KP758605	KP759013	

Table S1, continued.

Sample	Taxon	Region/count	Island/site	Identifier	Ring no	Cyt b	control region	ATP6AP2	BTF3L4	DNAJA2	ENSGALT 01183	ENSGALT 01658
ALG1	<i>C. teneriffae ultramarinus</i>	Algeria	S of Algiers	2004 ST611		KP759209	JF755317	KP758182	KP758418	KP758333	KP758286	KP758107
ALG2	<i>C. teneriffae ultramarinus</i>	Algeria	S of Algiers	2004 ST614		KP759254	JF755318	KP758827	KP758227	KP759077	KP758783	KP759112
ALG3	<i>C. teneriffae ultramarinus</i>	Algeria	S of Algiers	2004 ST616		KP759237	JF755319	KP758770	KP758355	KP758075	KP758596	KP758416
ALG4	<i>C. teneriffae ultramarinus</i>	Algeria	S of Algiers	2004 ST618		KP759268	JF755320	KP758638	KP758207	KP759137	KP758170	KP758464
ALG5	<i>C. teneriffae ultramarinus</i>	Algeria	S of Algiers	2004 ST613		KP759232	KP759173	KP758363	KP758412	KP759054	KP758390	KP758722
LIB1	<i>C. teneriffae cyrenaicae</i>	Libya	Close to Cyrene	2010 MAR7992		KP759204		KP758189	KP758759	KP758376	KP758743	KP758143
LIB2	<i>C. teneriffae cyrenaicae</i>	Libya	Close to Cyrene	2010 MAR7991		KP759258		KP758908	KP758739	KP758336	KP758033	KP758006
LIB3	<i>C. teneriffae cyrenaicae</i>	Libya	Close to Cyrene	2010 MAR7993		KP759256		KP758233	KP758083	KP758226	KP758531	KP758612
LIB4	<i>C. teneriffae cyrenaicae</i>	Libya	Close to Cyrene	2010 MAR8040		KC202349		KP759052	KP759047	KP758074	KP758421	KP758622
JDI1	<i>C. caeruleus satunini</i>	Jordan	Dibbeen	2005 ST936		KP759203	EF523808	KP758092	KP758358	KP758537	KP758877	KP758063
JDI2	<i>C. caeruleus satunini</i>	Jordan	Dibbeen	2005 ST938								
GRE1	<i>C. caeruleus calamensis</i>	Greece	Crete	2009 AI65608	AI65608	KP759197	KP759159	KP758222	KP759133	KP759134	KP759095	KP758001
GRE2	<i>C. caeruleus calamensis</i>	Greece	Crete	2009 090925								
ECA1	<i>C. caeruleus oligastrae</i>	Spain	Cádiz	2008 ST1716	EX8364		KP759227					
ECA1b	<i>C. caeruleus oligastrae</i>	Spain	Granada	2008 ST1714	EX8366	KP759223		KP758850	KP758980	KP758231	KP758428	KP758325
ECA2	<i>C. caeruleus oligastrae</i>	Spain	Cádiz	2008 ST1719	EX8351		KP759207					
ECA2b	<i>C. caeruleus oligastrae</i>	Spain	Granada	2008 ST1717	EX8369	KP759270		KP759070	KP758828	KP758105	KP758733	
ECA3	<i>C. caeruleus oligastrae</i>	Spain	Cádiz	2008 ST1720	EX8357	KP759230	KP759185				KP758095	
ECA3b	<i>C. caeruleus oligastrae</i>	Spain	Granada	2008 ST1718	EX8370	KP759213		KP758851	KP758404	KP758834	KP758798	KP758154
ECA4	<i>C. caeruleus oligastrae</i>	Spain	Cádiz	2008 ST1721	EX8356		KP759264				KP758890	
ECA4b	<i>C. caeruleus oligastrae</i>	Spain	Granada	2008 ST1722	EX8371	KP759191		KP758772	KP759006	KP758988	KP758046	KP758128
ECA5	<i>C. caeruleus oligastrae</i>	Spain	Sevilla	2008 N989540	N989540							
GB1	<i>C. caeruleus obscurus</i>	Great Britain	Oxford	2008 P806943	P806943	KP759272		KP759156	KP758292	KP758588	KP758327	KP758439
GB2	<i>C. caeruleus obscurus</i>	Great Britain	Oxford	2008 R639492	R639492	KP759198	KP759218	KP758148	KP759034	KP758995	KP758863	KP758060
GB3	<i>C. caeruleus obscurus</i>	Great Britain	Oxford	2008 T177704	T177704	KP759229	KP759199	KP758088	KP758853	KP758340	KP758230	KP758746
GB4	<i>C. caeruleus obscurus</i>	Great Britain	Oxford	2008 V729932	V729932	KP759273	KP759214	KP758139	KP758979	KP758367	KP758609	
SRE1	<i>C. caeruleus caeruleus</i>	Sweden	Revinge	2006 B1078	1EN80119	KP759228		KP758586	KP758898	KP758273	KP758381	KP758475
SRE2	<i>C. caeruleus caeruleus</i>	Sweden	Revinge	2006 B1092	1EN80134	KP759235	KP759160	short	KP758946	KP758976	KP759116	KP758820
SRE3	<i>C. caeruleus caeruleus</i>	Sweden	Revinge	2006 B1100	1EN80158	KP759222	KP759189	short	KP759128	KP758811	KP758261	KP758186
SRE4	<i>C. caeruleus caeruleus</i>	Sweden	Revinge	2006 B1140	1EN80198	KP759208	KP759180	short	KP758193	KP758598	KP758624	KP758861
Ccya	<i>C. cyanus flavipectus</i>	Kazakhstan	Chokpak	2012 766		KP759278	KP759172	KP758337	KP758140	KP758291	KP758174	KP758516
Pmaj	<i>Parus m. major</i>	Sweden	Lake Kvismaren	2004 261	2KH17891		KP759236	KP758818	KP758567	KP758065	short	KP758248
Pmaj2	<i>Parus m. major</i>	Sweden	Lake Kvismaren	2004 312	2KH17842	KP759174	KP759195	KP758771	KP759067	KP758647	short	KP758657
Ppal	<i>Poecile p. palustris</i>	Sweden	Lake Kvismaren	2004 255	BZ 01510							
Phud	<i>Poecile h. hudsonicus</i>	Canada	Gogama	1995 HU1	1B-2348							
Pate	<i>Periparus ater ater</i>	Russia	Vladivostok	1994 1444								
Paab	<i>Periparus ater abietum</i>	Spain	Sierra Nevada	2012 PN373	PA7602							
Pasa	<i>Periparus ater sardus</i>	Corsica	C. de Bonifato	2006 5567								
Patl	<i>Periparus ater atlas</i>	Morocco	Ifrane	2014 Perate3								
Pruf	<i>Periparus rufonuchalis</i>	Nepal	Purano Marpha	1995 90130								

Table S1, continued.

Sample	ENSGALT 05087	ENSGALT 11836	ENSGALT 15724	ENSGALT 16264	ENSGALT 25924	LDH-B	OGDHL	PDCD6	RPL7	SPAG9	TGFb2	UBE3A	WDR12	
ALG1	KP759106	KP758159	KP758446	KP758089	KP758776	KP758576	KP758295	KP759025	KP758331	KP758777	KP758753	KP758407	KP758534	SRX879665
ALG2	KP758397	KP759151	KP758241	KP758102	KP758801	KP758981	KP758316	KP759026	KP758855	KP758679	KP758676	KP758093	KP758064	
ALG3	KP758974	KP758448	KP758124	KP758130	KP758038	KP758793	KP758346	KP758096	KP758536	KP758935	KP758380	KP758014	KP758972	
ALG4	KP759030	KP758571	KP758590	KP758395	KP758223	KP758017	KP758967	KP758685	KP758315	KP759011	KP758781	KP758348	KP758052	
ALG5	KP758727	KP758177	KP758347	KP758437	KP758575	KP758993	KP758790	KP758219	KP759084	KP758266	KP758472	KP758465	KP758176	SRX879668
LIB1	KP758134	KP758814	KP758989	KP758774	KP759001	KP759050	KP758545	KP758279	KP758479	KP758243	KP759135	KP759153	KP758461	
LIB2	KP758171	KP758725	KP758860	KP758440		KP758050		KP758747	KP758305	KP758947	KP758120	KP758264	KP758024	
LIB3	KP758034	KP758704	KP758954	KP758267	KP758216	KP758953	KP758633	KP759082	KP759099	KP758822	KP758280	KP758072	KP759064	
LIB4	KP758583	KP758137			KP758555	KP758992		KP758728	KP758826		KP759072	KP758276	KP758408	SRX879669
JDI1	KP758830	KP758858	KP758769	KP758903	KP758155	KP758501	KP758097	KP758758	KP758275	KP758891	KP759089	KP758419	KP758966	
JDI2														
GRE1	KP758996	KP758104	KP758041	KP758457	KP758929	KP758078	KP758127	KP759048	KP758582	KP759096	KP758114	KP758119	KP758476	
GRE2														SRX879670
ECA1	KP758957	KP758400								KP758553				
ECA1b		KP758356		KP758218	KP758135	KP759154	KP758677	KP759118	KP758914	KP758907	KP758984	KP758206	KP758463	
ECA2		KP758585												
ECA2b		KP758020	KP758681	KP758717	KP759090	KP758084	KP758359	KP758591	KP758658	KP758797	short	KP758662	short	SRX883381
ECA3	KP758435	KP759035		KP759105	KP759042					KP759041		KP758145	KP759104	
ECA3b		KP758235	KP758224	KP759075	KP758445		KP758351	KP758314	KP759008	KP758618	KP758496	KP758196	KP758485	
ECA4	KP758212	KP758547		KP759158	KP758420					KP758396				
ECA4b		KP758960	KP759094	KP758115	KP758672	KP758489	KP759139	KP758616	KP759119	KP758304	KP759122	KP758541	KP758057	SRX883381
ECA5														
GB1	KP758785	KP758705	KP758307	KP758649	KP758406	KP758188	KP758296	KP758202	KP758256	KP758916	short	KP758158	KP758366	
GB2	KP758168	KP758005	KP758361	KP759018	KP758949	KP758552	KP758549	KP758854	KP758512	KP758025	KP758349	KP758424	KP758306	
GB3	KP758172	KP758683	KP758268	KP758606	KP758164	KP758042	KP758763	KP758784	KP758430	KP758312	short	KP758700	short	SRX879680
GB4	KP758666	KP758099	KP758080	KP758715	KP758187	KP758185	KP758289	KP758301		KP758129	short	KP758999	KP758228	
SRE1	KP758887	KP758517	KP759093	KP758648	KP758620	KP758577	KP758800	KP758236	KP758013	short	short	KP758837	KP758963	
SRE2	KP758838	KP758106	KP758518	KP757995	KP759103		KP758520	KP758824	short	KP758604	short	KP758655	KP758951	
SRE3	KP758507	KP758310	KP758540	KP758663	KP759024	KP758136	KP758142	KP758593	KP758287	KP758735	short	KP758290	KP758613	SRX879680
SRE4	KP758831	KP758165	KP758019	KP759121	KP759045	KP758015	KP759003	KP758888	KP758878	KP758048	KP759014	KP758870	short	
Ccya	KP758183	KP758535	KP758373	KP758506	KP758338	KP758458	KP758698	KP758121	KP758151	KP759046	KP758956	KP758123	KP758932	SRX879681
Pmaj	KP758422	KP758058	KP758368	KP758425	KP759086	KP758782	KP758343	KP758004		short	KP758532	KP758335	KP758047	SRX879682
Pmaj2	KP758444	KP758274	KP758409	KP759109	KP757998	KP758766	KP758862	KP758639		short	KP758012	KP758969	KP758113	SRX879683
Ppal														
Phud														
Pate														
Paab														SRX879685
Pasa														SRX879686
Patl														SRX879687
Pruf														SRX879688

Table S2. Details of the genetic markers, primers, and PCR conditions used in this study.

Locus	Marker	Gene position	Chrom. (Tgu)	Location (bp)	Primer	Primer sequence (5'–3')	Reference	Anneal. temp.	PCR cycles	Target length (intron, cds etc.)	Seq. primer
<i>Nuclear markers</i>											
<i>ATP6AP2</i>	ENSGALT 26187	intron 7	1	6882519	F	GGTGGGAATGCAGTAGTAGA	Backström <i>et al.</i> 2008	63–53	20+20	721	F
				6883313	R	CATGTTCCAGAGGTTGTAGG					
<i>UBE3A</i>	UBE3A	intron 12	1	31837212	F	AGAAACTACAGAATATGATGGTGGC	this study	60.5–55.5	10+25	468	R
				31837682	R	CTGTCTGTGCCTGTTGTAAACTG					
<i>LDHB</i>	LDH-B	intron 3	1A	65040952	F	GGAAGACAACTAAAAGGAGAAATG	Friesen <i>et al.</i> 1999	70–65	10+25	477	F
				65041439	R	ATGGA TTCCTCTGAAGCAGGTTGAGACGAC					
<i>PDCD6</i>	PDCD6	intron 4	2	94138715	F	TGGAARTACATCACAGAYTGGC	this study	61–56	10+25	681	R
				94139392	R	ACACAGCACTGAATAAARTCATCAA					
<i>RPL7</i>	RPL7	intron 3	2	124033847	F	TACCATAAGGAGTACAGRCACATG	this study	60–55	10+25	645	F
				124034480	R	TAGGGYTCAACAATCCGCA					
<i>TGFB2</i>	TGFb2	intron 6	3	10256872	F	GAAGCGTGCTCTAGATGCTG	Primmer <i>et al.</i> 2002	66–61	10+25	547	F
				10257446	R	AGGCAGCAATTATCCTGCAC					
<i>CRIP1</i>	ENSGALT 16264	intron 3	3	18103287	F	GCATTGACCTCAAAGAAGGC	Backström <i>et al.</i> 2008	63–53	20+20	705	R
				18103966	R	TTTATAGGCACATCCTTGAC					
<i>PRADC1</i>	ENSGALT 25924	intron 3	4_rand	4029445	F	TTCATCCAGGACCAGATCGC	Backström <i>et al.</i> 2008	60–50	20+20	618	R
				4030096	R	GTGCTGTCCTGGATCATCTC					
<i>OGDHL</i>	OGDHL	intron 21	6	8571613	F	GCAGTTAGAATGTCTTTCCCTCA	this study	60.5–55.5	10+25	720	R
				8572304	R	GGTGTGAGGAGGARCACAA					
<i>ACADSB</i>	ENSGALT 15724	intron 9	6	32012042	F	AAATGTATTGAATGGATGGG	Backström <i>et al.</i> 2008	60–50	20+20	674	F
				32012734	R	CAGCTGGATATTTGAAGTTC					
<i>WDR12</i>	WDR12	intron 9	7	21530614	F	TGACTTCTCACACAGGCTGG	this study	60.5–55.5	10+25	557	F
				21531143	R	CATGAGCAGCCAAGTCATACA					
<i>BTF3L4</i>	BTF3L4	intron 6	8	21805788	F	GGAYAGTAAAGCACCAAAATCTGAA	this study	60–55	10+25	867	F
				21806642	R	GTCCATGCCAGCTTCCAG					
<i>DNAJA2</i>	DNAJA2	intron 3	11	13887105	F	ACAGTATCGCAATCCTTTTGAGA	this study	61–56	10+25	499	F
				13887564	R	TCTGGCCACCACCTGAAC					
<i>PSMD6</i>	ENSGALT 11836	intron part 5	12	14512888	F	GAAAAGATCTTGGGGGAGAG	Backström <i>et al.</i> 2008	63–53	20+20	517	F
				14513460	R	CTTGCGGAATGCAGTCAGAG					
<i>SPAG9</i>	SPAG9	intron 35	18	9602824	F	AYGCACAGCTYTGCTTCC	this study	61–56	10+25	333	F
				9603142	R	CCRCTTATACCAACATTGAT					
<i>SBDS</i>	ENSGALT 01658	intron 2	19	770102	F	ATTGAACGTGCTCACATGAG	Backström <i>et al.</i> 2008	63–53	20+20	629	R
				770728	R	TTCCTCCACATCTTTCAGAC					
<i>POLDIP2</i>	ENSGALT 05087	intron 5	19	6981967	F	CTGGCAAATCATGATGACAG	Backström <i>et al.</i> 2008	63–53	20+20	591	R
				6982485	R	CTGGATGGGCACTTGATCAG					
<i>OAZ1</i>	ENSGALT 01183	intron 3	28	1722106	F	CAGTCAAGGCTTACAGATGC	Backström <i>et al.</i> 2008	63–53	20+20	529	R
				1722624	R	ATCTCTGTTCTTGTTGAAAC					

Table S2, continued.

Locus	Marker	Gene position	Chrom. (Tgu)	Location (bp)	Primer	Primer sequence (5'-3')	Reference	Anneal. temp.	PCR cycles	Target length (intron, cds etc.)	Seq. primer
<i>Mitochondrial markers, amplification</i>											
<i>cytB</i>		full cds	mt	L14838 H16057	ND5-Syl mtFNP	GGCCTAATCAARGCCTACYTAGG CCRAATGTTCTGGTTACAAA	A. J. Helbig unpubl. Fregin <i>et al.</i> 2009	55–45	20+20	1143	see below
<i>cytB</i>		partial cds	mt	L14995 H15917	mt-A1 mt-Fs-H	GCCCCATCCAACATCTCAGCATGAT GAAACTTCCG TAGTTGGCCAATGATGATGAATGGG TGTTCTACTGGTT	Dietzen <i>et al.</i> 2007 Dietzen <i>et al.</i> 2007	53	35	920	see below
<i>cytB</i>		partial cds	mt	L14990 H16065	L14841 H16065	AAAAAGCTTCCATCCAACATCTCAGC ATGATGAAA TCATCTCCGGTTTACAAGAC	Kocher <i>et al.</i> 1989 Helm-Bychowski & Cracraft 1993	52.5–47.5	10+25	1045	see below
control region		control region	mt	L16700 H590	L16700 H590	GAGATGAGGAGTATTCAACCGAC	Kvist <i>et al.</i> 1999 Kvist <i>et al.</i> 1999	53	30	577	H590
<i>Mitochondrial markers, sequencing</i>											
<i>cytB</i>				L15320	L15320	TAYGTCCTACCATGAGGACAAATATC ATTCTAGG	Prager <i>et al.</i> 1993	45	30		
				H15541	Cytb_seq_ H15541	CCCACCTTACCYTAAAACAG	this study	45	30		

Chrom. (Tgu), chromosome number in Zebra Finch *Taeniopygia guttata* ; anneal. temp., annealing temperature (°C); seq. Primer, sequencing primer.

Table S3. Nucleotide diversity within populations, measured in 18 nuclear introns. The colour scheme pronounces differences by highlighting low values in blue, and high in red. n/a = measurement not possible to calculate.

Nucleotide diversity (π)	2 n	ENSGALT																		
	avg (range)	Average	ATP6AP2	BTF3L4	DNAJA2	01183	01658	05087	11836	15724	16264	25924	LDH-B	OGDHL	PDCD6	RPL7	SPAG9	TGFB2	UBE3A	WDR12
Libya (N. Africa)	3,8 (3-4)	0.00046	0.00035	0.00200	0.00000	0.00048	0.00000	0.00000	0.00000	0.00000	0.00131	0.00000	0.00000	n/a	n/a	0.00160	0.00160	0.00000	0.00000	0.00000
La Palma (W. Canary Is)	4,9 (4-5)	0.00079	0.00160	0.00000	0.00178	0.00000	0.00085	0.00034	0.00091	0.00000	0.00000	0.00000	0.00000	0.00065	0.00403	0.00238	0.00000	0.00000	0.00076	0.00097
Central Canary Is	19,9 (18-21)	0.00193	0.00025	0.00451	0.00077	0.00034	0.00201	0.00586	0.00267	0.00008	0.00096	0.00333	0.00043	0.00053	0.00191	0.00102	0.00646	0.00097	0.00161	0.00096
El Hierro	4,9 (4-5)	0.00017	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00033	0.00000	0.00000	0.00083	0.00197	0.00000	0.00000	0.00000	0.00000
La Gomera	5,1 (3-6)	0.00098	0.00047	0.00000	0.00040	0.00000	0.00523	0.00028	0.00094	0.00000	0.00000	0.00190	0.00043	0.00065	0.00458	0.00000	0.00000	0.00000	0.00204	0.00064
Tenerife	5 (5-5)	0.00107	0.00028	0.00124	0.00187	0.00038	0.00000	0.00068	0.00078	0.00030	0.00028	0.00119	0.00100	0.00050	0.00318	0.00128	0.00487	0.00000	0.00114	0.00036
Gran Canaria	4,9 (3-5)	0.00077	0.00030	0.00109	0.00040	0.00090	0.00000	0.00224	0.00000	0.00000	0.00123	0.00000	0.00000	0.00084	0.00162	0.00264	0.00000	0.00102	0.00161	0.00000
E. Canary Is, NW Africa	19,6 (18-20)	0.00207	0.00194	0.00072	0.00052	0.00145	0.00235	0.00296	0.00103	0.00507	0.00213	0.00283	0.00193	0.00167	0.00140	0.00205	0.00373	0.00389	0.00061	0.00104
Fuerteventura	4,8 (4-5)	0.00154	0.00132	0.00000	0.00000	0.00081	0.00122	0.00095	0.00039	0.00415	0.00186	0.00081	0.00242	0.00256	0.00000	0.00335	0.00440	0.00219	0.00000	0.00120
Lanzarote	5 (5-5)	0.00058	0.00226	0.00000	0.00000	0.00000	0.00000	0.00000	0.00091	0.00208	0.00167	0.00000	0.00000	0.00028	0.00053	0.00000	0.00000	0.00228	0.00043	0.00000
Morocco	4,8 (4-5)	0.00220	0.00163	0.00000	0.00000	0.00000	0.00194	0.00330	0.00138	0.00416	0.00189	0.00285	0.00271	0.00274	0.00427	0.00173	0.00414	0.00451	0.00100	0.00132
Algeria	5 (5-5)	0.00237	0.00336	0.00289	0.00147	0.00342	0.00318	0.00516	0.00069	0.00481	0.00224	0.00235	0.00161	0.00050	0.00083	0.00263	0.00200	0.00382	0.00085	0.00084
Europe	12,6 (7-16)	0.00505	0.00771	0.00448	0.00174	0.00770	0.00455	0.00602	0.00430	0.00670	0.00131	0.00910	0.00278	0.00243	0.00879	0.00393	0.00842	0.00397	0.00316	0.00373
Sweden	3,6 (3-5)	0.00465	n/a	0.00508	0.00100	0.00790	0.00473	0.00519	0.00361	0.00415	0.00144	0.00694	0.00328	0.00165	0.01288	0.00248	0.00801	n/a	0.00397	0.00203
Great Britain	3,6 (3-4)	0.00519	0.00691	0.00630	0.00236	0.00798	0.00339	0.00581	0.00777	0.00948	0.00092	n/a	0.00254	0.00220	0.00650	0.00163	0.01001	n/a	0.00397	0.00527
Spain	4,4 (3-7)	0.00478	n/a	0.00257	0.00136	0.00844	0.00455	0.00849	0.00385	n/a	0.00082	0.00889	0.00301	0.00301	0.00776	0.00433	0.00703	0.00504	0.00275	0.00455

Table S4. Results from the Restriction site Associated DNA (RAD) sequencing for core samples. "Cleaned read pairs" refers to number of read pairs retained after de-multiplexing and filtering Stacks component process_radtags. "Clone-filtered read pairs" refers to number of read pairs retained after further removing PCR clones (within-individual copies in which both read1 and read2 are identical) using Stacks component clone_filter, thus representing the true sequencing depth. For detailed sample information, see Table S1.

Sample	Site	Barcode	Cleaned read pairs		Clone-filtered read pairs		
			Library 1	Library 2*	Library 1	Library 2*	Combined
CEH2	CI, El Hierro	AACATGC	2797738	3499897	1548312	2712229	4260541
CFU4	CI, Fuerteventura	AATGCCT	7771797		4313960		
CGC2	CI, Gran Canaria	AGAGTCG	5325038		3020224		
CLG2	CI, La Gomera	CAATGAC	9312551		5153234		
CLP3	CI, La Palma	CAGACAT	5543324		3045934		
CLA1	CI, Lanzarote	CTCAGCT	638647	1759253	361432	1164323	1525755
CTE2	CI, Tenerife	CTTCTAG	2174946	3796468	1220375	2733561	3953936
ALG3	Algeria	GAGTGGA	5422017		3084908		
MCE2	Morocco, Ceuta	GCGGATA	7115603		4046892		
SRE4	Sweden, Revinge	GCTTGAT	6807039		3796007		
GRE1	Greece, Crete	GTTTCAGC	9513442		5357445		
JDI2	Jordan, Dibbeen	TGGAGTT	3501104	4158679	1956034	3124029	5080063
Ccya	Kazakhstan, Chokpak	TGGTCAA	6506916		3613155		
Ppal	Sweden, Lake Kvismaren	TTATCCG	9814060		5378238		
Phud	Canada, Ontario	TTCCTTC	3669415	4929142	2062304	3762516	5824820
Pmaj	Sweden, Lake Kvismaren	TTCGAAG	814614	1441286	453824	880345	1334169
Pate	Russia, Vladivostok	AATGCCT		9381958		7446333	
Paab	Spain	AGAGTCG		7088604		5562194	
Pasa	Corsica, Cirque de Bonifato	CAATGAC		13474915		10415080	
Patl*	Morocco, Ifrane	AACATGC		7615064		6064212	
Pruf	Nepal, Purano Marpha	GTTTCAGC		5419947		4218345	

*Sample Patl was run in a separate library (Library 3).

Table S5. Mean estimated node ages in million years (95% highest posterior density) for five different calibration sets and four subsets of 14 Kbp genomic RAD sequence data; three autosomal and one from the Z chromosome. The results are presented for analyses of the full taxon sets and partial taxon sets, from which samples of dubious genetic variation were removed.

Clade		Genomic subset	Calibration set ^a									
			A (BER rel., MSC rel.)		B (BER, MSC)		C (BER, MSC, CAN)		D (BER rel., MSC)		E (MSC)	
			full tax.set	partial tax.set ^b	full tax.set	partial tax.set ^b	full tax.set	partial tax.set ^b	full tax.set	partial tax.set ^b	full tax.set	partial tax.set ^b
AB	<i>Cyanistes</i>	aut. 1	5.4 (2.9–8.0)	4.2 (2.4–6.1)	6.3 (4.3–6.4)	4.9 (3.3–6.5)	3.7 (2.8–4.8)	3.8 (2.9–4.7)	6.2 (4.1–8.3)	4.8 (3.4–6.3)	6.0 (3.9–8.1)	4.7 (3.3–6.2)
		aut. 2	5.4 (2.8–8.3)	4.6 (2.4–6.9)	6.7 (4.8–8.7)	5.7 (4.2–7.5)	3.5 (2.8–4.3)	3.6 (3.0–4.4)	6.9 (4.9–9.3)	5.7 (4.2–7.3)	6.9 (5.1–9.1)	5.9 (4.4–7.5)
		aut. 3	5.3 (2.7–8.0)	4.5 (2.4–6.7)	6.4 (4.5–8.5)	5.5 (4.1–7.1)	4.0 (3.1–5.2)	4.2 (3.4–5.2)	6.5 (4.6–8.7)	5.7 (4.3–7.2)	6.4 (4.5–8.5)	5.7 (4.3–7.2)
		Z	6.2 (3.1–9.2)	4.9 (2.9–7.0)	7.1 (4.6–9.6)	5.7 (4.2–7.2)	4.4 (3.2–5.8)	4.5 (3.6–5.5)	7.1 (4.8–9.7)	6.0 (4.6–7.5)	7.1 (4.6–9.8)	6.0 (4.7–7.6)
A	<i>C. caeruleus</i> + <i>C. cyanus</i>	aut. 1	3.2 (1.6–5.0)	2.8 (1.6–4.1)	3.8 (2.4–5.4)	3.3 (2.3–4.5)	2.5 (1.7–3.5)	2.7 (2.0–3.6)	3.7 (2.3–5.3)	3.2 (2.2–4.3)	3.6 (2.2–5.1)	3.1 (2.1–4.2)
		aut. 2	3.0 (1.4–4.7)	2.8 (1.4–4.3)	3.7 (2.5–5.2)	3.5 (2.4–4.7)	2.3 (1.6–3.1)	2.5 (1.8–3.1)	3.8 (2.5–5.3)	3.5 (2.5–4.7)	3.8 (2.5–5.1)	3.6 (2.6–4.7)
		aut. 3	3.9 (2.0–6.0)	3.5 (1.9–5.3)	4.7 (3.2–6.5)	4.3 (3.1–5.5)	3.1 (2.2–4.0)	3.4 (2.6–4.2)	4.8 (3.2–6.6)	4.4 (3.3–5.7)	4.8 (3.2–6.5)	4.4 (3.3–5.6)
		Z	3.6 (1.6–5.8)	3.1 (1.7–4.5)	4.2 (2.5–6.3)	3.6 (2.6–4.8)	2.8 (1.7–4.0)	3.0 (2.2–3.8)	4.2 (2.5–6.2)	3.7 (2.7–4.8)	4.2 (2.4–6.3)	3.8 (2.7–4.9)
A1	<i>C. caeruleus</i>	aut. 1	2.8 (1.5–4.3)	2.6 (1.4–3.7)	3.3 (2.0–4.7)	3.0 (2.0–4.0)	2.2 (1.4–3.1)	2.4 (1.8–3.1)	3.3 (2.0–4.7)	2.9 (2.0–3.9)	3.2 (1.9–4.5)	2.8 (1.9–3.8)
		aut. 2	2.8 (1.4–4.4)	2.6 (1.3–4.0)	3.5 (2.3–4.8)	3.3 (2.3–4.4)	2.1 (1.4–2.9)	2.3 (1.7–3.0)	3.5 (2.4–4.9)	3.3 (2.3–4.3)	3.5 (2.3–4.7)	3.4 (2.4–4.4)
		aut. 3	3.3 (1.7–5.1)	3.1 (1.7–4.7)	4.1 (2.8–5.6)	3.7 (2.6–4.8)	2.7 (1.9–3.6)	3.0 (2.3–3.8)	4.2 (2.7–5.7)	3.9 (2.9–5.0)	4.1 (2.7–5.6)	3.9 (2.9–5.0)
		Z	2.7 (1.2–4.4)	2.4 (1.3–3.6)	3.2 (1.8–4.7)	2.8 (1.9–3.7)	2.2 (1.2–3.1)	2.3 (1.6–3.1)	3.2 (1.8–4.7)	2.9 (2.1–3.8)	3.2 (1.7–4.8)	2.9 (2.1–3.8)
B	<i>C. teneriffae</i>	aut. 1	4.8 (2.6–7.1)	3.7 (2.1–5.3)	5.6 (3.9–7.5)	4.3 (3.0–5.8)	3.1 (2.6–3.5)	3.1 (2.6–3.5)	5.5 (3.8–7.5)	4.1 (3.0–5.5)	5.3 (3.6–7.3)	4.1 (2.9–5.5)
		aut. 2	5.2 (2.7–7.9)	4.3 (2.3–6.5)	6.3 (4.6–8.3)	5.4 (4.0–7.0)	3.2 (2.8–3.5)	3.2 (2.8–3.5)	6.6 (4.6–8.8)	5.4 (4.0–7.0)	6.5 (4.7–8.5)	5.5 (4.1–7.1)
		aut. 3	4.6 (2.4–7.0)	3.8 (2.1–5.7)	5.6 (3.9–7.3)	4.6 (3.3–5.9)	3.2 (2.8–3.5)	3.3 (2.9–3.5)	5.7 (4.1–7.7)	4.8 (3.6–6.2)	5.6 (3.9–7.5)	4.8 (3.6–6.2)
		Z	4.9 (2.6–7.5)	3.8 (2.2–5.5)	5.7 (3.8–7.8)	4.4 (3.3–5.7)	3.1 (2.6–3.5)	3.2 (2.8–3.5)	5.7 (3.8–7.7)	4.7 (3.4–5.8)	5.7 (3.7–7.9)	4.7 (3.5–5.9)
B1 ^c	<i>C. t. palmensis</i> + <i>C. t. cyrenaicae</i>	aut. 1	3.8 (1.9–5.9)	3.0 (1.6–4.6)	4.4 (2.5–6.4)	3.5 (2.2–4.9)	2.8 (1.9–3.5)	2.8 (2.1–3.5)	4.4 (2.6–6.5)	3.5 (2.3–4.9)	4.3 (2.3–6.3)	3.4 (2.2–4.6)
		aut. 2	4.8 (2.4–7.3)	4.0 (2.1–6.1)	5.8 (3.8–7.7)	4.9 (3.5–6.5)	3.1 (2.4–3.5)	3.1 (2.6–3.5)	6.0 (4.2–7.8)	5.0 (3.6–6.6)	5.9 (4.1–7.9)	5.1 (3.6–6.8)
		aut. 3	3.2 (1.6–4.9)	2.2 (1.1–3.4)	3.9 (2.7–5.2)	2.7 (1.9–3.7)	2.7 (1.8–3.6)	2.3 (1.6–3.0)	4.0 (2.8–5.3)	2.8 (2.0–3.7)	4.0 (2.7–5.3)	2.8 (2.0–3.8)
		Z	4.0 (1.9–6.4)	3.4 (1.9–5.0)	4.7 (2.8–6.8)	3.9 (2.8–5.2)	2.8 (1.9–3.5)	3.0 (2.4–3.5)	4.8 (2.7–6.9)	4.1 (3.0–5.3)	4.6 (2.9–6.7)	4.1 (2.9–5.3)
B1b	<i>C. teneriffae</i> excl. <i>C. t. palmensis</i>	aut. 1	4.4 (2.4–6.5)	3.2 (1.8–4.7)	5.1 (3.5–6.7)	3.7 (2.5–5.0)	2.8 (2.2–3.4)	2.7 (2.2–3.3)	5.0 (3.5–6.8)	3.6 (2.6–4.8)	4.8 (3.3–6.7)	3.6 (2.5–4.8)
		aut. 2	5.0 (2.6–7.6)	4.2 (2.2–6.3)	6.1 (4.4–8.0)	5.2 (3.7–6.8)	3.1 (2.6–3.5)	3.1 (2.7–3.5)	6.4 (4.5–8.6)	5.2 (3.9–6.7)	6.3 (4.5–8.2)	5.3 (3.9–6.8)
		aut. 3	4.3 (2.2–6.6)	3.6 (2.0–5.4)	5.2 (3.6–6.9)	4.3 (3.1–5.6)	3.0 (2.5–3.5)	3.1 (2.6–3.4)	5.3 (3.8–7.2)	4.5 (3.3–5.7)	5.3 (3.7–7)	4.5 (3.3–5.8)
		Z	4.5 (2.3–6.9)	3.5 (2.0–5.1)	5.2 (3.4–7.1)	4.0 (2.9–5.2)	2.9 (2.3–3.4)	2.9 (2.4–3.4)	5.3 (3.5–7.1)	4.2 (3.2–5.4)	5.3 (3.5–7.4)	4.3 (3.1–5.4)

B2	<i>C. teneriffae</i> , clades B3 + B4 (see below)	aut. 1	3.7 (2.0–5.5)	2.4 (1.4–3.5)	4.4 (3.1–5.9)	2.8 (1.9–3.8)	2.4 (1.8–2.9)	2.0 (1.5–2.6)	4.3 (3.0–5.9)	2.7 (1.8–3.6)	4.2 (2.8–5.7)	2.7 (1.8–3.6)
		aut. 2	4.2 (2.2–6.5)	3.3 (1.7–5.0)	5.1 (3.7–6.8)	4.1 (2.8–5.4)	2.6 (2.1–3.2)	2.5 (1.9–3.0)	5.4 (3.7–7.1)	4.1 (3.0–5.3)	5.3 (3.8–7.0)	4.1 (3.0–5.5)
		aut. 3	3.6 (1.8–5.4)	2.8 (1.5–4.3)	4.4 (3.1–5.8)	3.4 (2.4–4.5)	2.5 (2.0–3.0)	2.5 (2.0–3.0)	4.5 (3.1–6.1)	3.5 (2.6–4.6)	4.4 (3.1–5.9)	3.5 (2.6–4.6)
		Z	3.4 (1.8–5.2)	2.1 (1.2–3.1)	4.0 (2.7–5.5)	2.5 (1.7–3.2)	2.2 (1.6–2.8)	1.8 (1.3–2.2)	4.0 (2.7–5.4)	2.6 (1.9–3.3)	4.0 (2.6–5.6)	2.6 (1.9–3.3)
B3	<i>C. teneriffae</i> , central Canary Islands	aut. 1	2.7 (1.4–4.2)	1.6 (0.8–2.4)	3.2 (2.0–4.5)	1.8 (1.1–2.6)	1.1 (1.0–1.2)	1.1 (0.9–1.2)	3.2 (1.9–4.4)	1.8 (1.1–2.5)	3.1 (1.9–4.4)	1.8 (1.1–2.5)
		aut. 2	3.0 (1.5–4.8)	2.2 (1.1–3.5)	3.7 (2.5–5.0)	2.7 (1.7–3.9)	1.1 (1.0–1.2)	1.1 (1.0–1.2)	3.9 (2.5–5.2)	2.8 (1.8–3.8)	3.8 (2.6–5.2)	2.8 (1.9–3.8)
		aut. 3	2.4 (1.2–3.8)	1.5 (0.7–2.4)	2.9 (1.8–4.0)	1.8 (1.1–2.6)	1.1 (1.0–1.2)	1.1 (0.9–1.2)	3.0 (1.8–4.1)	1.9 (1.2–2.7)	2.9 (1.9–4.1)	1.9 (1.2–2.7)
		Z	2.6 (1.3–4.0)	1.7 (0.9–2.5)	3.0 (1.9–4.3)	1.9 (1.3–2.6)	1.1 (1.0–1.2)	1.1 (1.0–1.2)	3.1 (2.0–4.3)	2.0 (1.4–2.7)	3.1 (1.9–4.3)	2.0 (1.5–2.7)
B4	<i>C. teneriffae</i> , eastern Canary Islands, north- western Africa	aut. 1	3.3 (1.7–4.8)	1.6 (0.8–2.4)	3.8 (2.5–5.2)	1.9 (1.2–2.6)	2.1 (1.5–2.7)	1.4 (1.0–1.9)	3.8 (2.4–5.2)	1.8 (1.2–2.5)	3.6 (2.3–5.1)	1.8 (1.1–2.5)
		aut. 2	3.5 (1.8–5.5)	2.3 (1.1–3.6)	4.3 (3.0–5.7)	2.9 (1.9–3.9)	2.2 (1.7–2.8)	1.8 (1.3–2.4)	4.5 (3.1–6.0)	2.8 (2.0–3.9)	4.4 (3.1–5.9)	2.9 (2.0–3.9)
		aut. 3	3.2 (1.7–4.9)	2.2 (1.1–3.4)	3.8 (2.7–5.2)	2.6 (1.8–3.5)	2.3 (1.7–2.8)	2.0 (1.5–2.5)	3.9 (2.7–5.4)	2.8 (2.0–3.7)	3.9 (2.6–5.1)	2.7 (2.0–3.7)
		Z	2.8 (1.4–4.3)	1.5 (0.8–2.2)	3.2 (2.1–4.5)	1.7 (1.1–2.3)	1.8 (1.3–2.4)	1.3 (0.9–1.7)	3.3 (2.2–4.5)	1.8 (1.2–2.4)	3.3 (2.1–4.6)	1.8 (1.3–2.4)

NOTE.— rel. = relaxed; aut. = autosomal, Z = Z chromosome, tax.set = taxon set

^a Calibration points were:

BER = Beringia: Nearctic/Holarctic split within *Poecile* at 5.4–5.5 My or rel. = relaxed at 4.8–14.0 My;

MSC = Messinian Salinity Crisis: African/European split within *Periparus ater* at 5.32–5.33 My or rel. = relaxed at 0–5.96 My;

CAN = Canary Islands volcanic age at El Hierro 0–1.2 My and La Palma 0–3.5 My.

^b The partial taxon set excludes individuals from Lanzarote and Tenerife, owing to high within-individual genetic variation (see text)

^c *C. t. palmensis* (La Palma) and *C. t. cyrenaicae* (Libya) enforced as a monophyletic clade according to species tree analyses (see text).

Table S6. Substitution models selected according to the Bayesian Information Criterion, using jModelTest.

Marker	Substitution model
<i>Nuclear</i>	
ATP6AP2	HKY+Invariant
BTF3L4	HKY+Invariant
DNAJA2	HKY
ENSGALT00000001183	HKY+Invariant
ENSGALT00000001658	HKY+Gamma
ENSGALT00000005087	K80+Invariant
ENSGALT00000011836	HKY
ENSGALT00000015724	HKY+Invariant+Gamma
ENSGALT00000016264	HKY+Invariant
ENSGALT00000025924	K80+Invariant
LDH-B	HKY
OGDHL	HKY+Invariant
PDCD6	HKY
RPL7	HKY+Invariant
SPAG9	F81+Invariant
TGFb2	HKY+Invariant
UBE3A	HKY
WDR12	HKY
<i>Mitochondrial</i>	
Control region	HKY+Invariant
Cytochrome b	HKY+Invariant+Gamma

Table S7. Haplotype diversity within populations, measured in 18 nuclear introns. The colour scheme pronounces differences by highlighting low values in blue, and high in red. n/a = measurement not possible to calculate.

Haplotype diversity (Hd)	2 n	ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT ENSGALT																		
	avg (range)	Average	ATP6AP2	BTF3L4	DNAJA2	01183	01658	05087	11836	15724	16264	25924	LDH-B	OGDHL	PDCD6	RPL7	SPAG9	TGFb2	UBE3A	WDR12
Libya (N. Africa)	3,8 (3-4)	0.175	0.250	0.429	0.000	0.250	0.000	0.000	0.000	0.000	0.733	0.000	0.000	n/a	n/a	0.607	0.533	0.000	0.000	0.000
La Palma (W. Canary Is)	4,9 (4-5)	0.299	0.511	0.000	0.711	0.000	0.533	0.200	0.467	0.000	0.000	0.000	0.000	0.467	0.778	0.822	0.000	0.000	0.356	0.536
Central Canary Is	19,9 (18-21)	0.440	0.162	0.763	0.233	0.177	0.388	0.769	0.559	0.056	0.556	0.717	0.203	0.351	0.354	0.312	0.749	0.491	0.595	0.491
El Hierro	4,9 (4-5)	0.075	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.200	0.000	0.000	0.511	0.644	0.000	0.000	0.000	0.000
La Gomera	5,1 (3-6)	0.287	0.333	0.000	0.200	0.000	0.733	0.167	0.485	0.000	0.000	0.742	0.200	0.467	0.800	0.000	0.000	0.000	0.682	0.356
Tenerife	5 (5-5)	0.373	0.200	0.711	0.467	0.200	0.000	0.378	0.378	0.200	0.200	0.644	0.467	0.356	0.356	0.689	0.733	0.000	0.533	0.200
Gran Canaria	4,9 (3-5)	0.315	0.200	0.467	0.200	0.467	0.000	0.644	0.000	0.000	0.644	0.000	0.000	0.533	0.733	0.733	0.000	0.511	0.533	0.000
E. Canary Is, NW Africa	19,6 (18-20)	0.550	0.690	0.243	0.252	0.317	0.684	0.679	0.497	0.848	0.796	0.738	0.510	0.528	0.235	0.460	0.763	0.885	0.273	0.500
Fuerteventura	4,8 (4-5)	0.403	0.622	0.000	0.000	0.200	0.250	0.511	0.200	0.467	0.644	0.250	0.600	0.778	0.000	0.644	0.689	0.800	0.000	0.600
Lanzarote	5 (5-5)	0.172	0.533	0.000	0.000	0.000	0.000	0.000	0.467	0.200	0.511	0.000	0.000	0.200	0.356	0.000	0.000	0.622	0.200	0.000
Morocco	4,8 (4-5)	0.571	0.644	0.000	0.000	0.000	0.644	0.733	0.622	0.750	0.844	0.778	0.511	0.644	0.733	0.511	0.889	0.857	0.467	0.644
Algeria	5 (5-5)	0.663	0.844	0.911	0.644	0.800	0.778	0.956	0.356	0.844	0.844	0.867	0.511	0.356	0.511	0.600	0.511	0.756	0.378	0.467
Europe	12,6 (7-16)	0.820	0.956	0.746	0.587	0.935	0.859	0.818	0.747	0.939	0.598	0.965	0.757	0.820	0.954	0.692	0.943	0.802	0.782	0.859
Sweden	3,6 (3-5)	0.809	n/a	0.786	0.464	1.000	0.929	0.929	0.857	0.733	0.733	1.000	0.800	0.643	0.978	0.533	0.933	n/a	0.893	0.733
Great Britain	3,6 (3-4)	0.836	0.929	0.929	0.643	0.964	0.733	0.679	0.867	0.893	0.600	n/a	0.786	0.857	0.964	0.600	1.000	n/a	0.929	1.000
Spain	4,4 (3-7)	0.803	n/a	0.464	0.607	0.955	0.933	1.000	0.682	n/a	0.536	0.909	0.867	0.929	0.964	0.643	0.868	0.867	0.733	0.893

Table S8. Theta within populations, measured in 18 nuclear introns. The colour scheme pronounces differences by highlighting low values in blue, and high in red. n/a = measurement not possible to calculate.

θ	2 n		ENSGALT																	
	avg (range)	Average	ATP6AP2	BTF3L4	DNAJA2	01183	01658	05087	11836	15724	16264	25924	LDH-B	OGDHL	PDCD6	RPL7	SPAG9	TGFB2	UBE3A	WDR12
Libya (N. Africa)	3,8 (3-4)	0.00117	0.00054	0.00180	---	0.00074	---	---	---	---	0.00132	---	---	n/a	n/a	0.00128	0.00132	---	---	---
La Palma (W. Canary Is)	4,9 (4-5)	0.00128	0.00150	---	0.00142	---	0.00056	0.00060	0.00069	---	---	---	---	0.00050	0.00393	0.00219	---	---	0.00076	0.00069
<i>Central Canary Is</i>	19,9 (18-21)	0.00192	0.00108	0.00281	0.00188	0.00045	0.00224	0.00437	0.00226	0.00036	0.00167	0.00302	0.00052	0.00132	0.00202	0.00146	0.00349	0.00129	0.00298	0.00127
El Hierro	4,9 (4-5)	0.00109	---	---	---	---	---	---	---	---	---	0.00057	---	---	0.00105	0.00165	---	---	---	---
La Gomera	5,1 (3-6)	0.00131	0.00062	---	0.00071	---	0.00337	0.00057	0.00064	---	---	0.00215	0.00076	0.00050	0.00303	---	---	---	0.00142	0.00063
Tenerife	5 (5-5)	0.00114	0.00050	0.00082	0.00142	0.00068	---	0.00121	0.00137	0.00052	0.00050	0.00115	0.00076	0.00050	0.00316	0.00110	0.00318	---	0.00076	0.00063
Gran Canaria	4,9 (3-5)	0.00133	0.00052	0.00082	0.00071	0.00068	---	0.00181	---	---	0.00150	---	---	0.00149	0.00158	0.00200	---	0.00129	0.00227	---
<i>E. Canary Is, NW Africa</i>	19,6 (18-20)	0.00220	0.00209	0.00129	0.00097	0.00190	0.00231	0.00394	0.00091	0.00303	0.00167	0.00310	0.00301	0.00198	0.00201	0.00186	0.00282	0.00392	0.00151	0.00127
Fuerteventura	4,8 (4-5)	0.00184	0.00105	---	---	0.00143	0.00188	0.00121	0.00069	0.00315	0.00150	0.00125	0.00302	0.00198	---	0.00220	0.00318	0.00194	---	0.00127
Lanzarote	5 (5-5)	0.00139	0.00150	---	---	---	---	---	0.00069	0.00367	0.00150	---	---	0.00050	0.00053	---	---	0.00194	0.00076	---
Morocco	4,8 (4-5)	0.00268	0.00250	---	---	---	0.00225	0.00338	0.00137	0.00492	0.00201	0.00287	0.00302	0.00248	0.00302	0.00184	0.00425	0.00423	0.00076	0.00127
Algeria	5 (5-5)	0.00229	0.00300	0.00246	0.00142	0.00272	0.00337	0.00592	0.00069	0.00367	0.00251	0.00230	0.00227	0.00050	0.00105	0.00186	0.00212	0.00323	0.00151	0.00064
<i>Europe</i>	12,6 (7-16)	0.00580	0.00755	0.00572	0.00412	0.00765	0.00406	0.00521	0.00686	0.00600	0.00206	0.00847	0.00328	0.00396	0.01454	0.00528	0.00682	0.00387	0.00324	0.00577
Sweden	3,6 (3-5)	0.00482	n/a	0.00594	0.00155	0.00742	0.00432	0.00659	0.00449	0.00390	0.00135	0.00712	0.00374	0.00216	0.01430	0.00204	0.00658	n/a	0.00412	0.00157
Great Britain	3,6 (3-4)	0.00513	0.00599	0.00660	0.00309	0.00892	0.00348	0.00461	0.00680	0.00801	0.00067	n/a	0.00284	0.00270	0.00743	0.00215	0.00921	n/a	0.00412	0.00550
Spain	4,4 (3-7)	0.00503	n/a	0.00396	0.00155	0.00766	0.00498	0.00930	0.00514	n/a	0.00059	0.00808	0.00283	0.00325	0.00814	0.00543	0.00755	0.00485	0.00227	0.00485

Table S9. Average difference, based on four genomic subsets, in estimated node ages for the partial taxon set (excluding individuals from Lanzarote and Tenerife, owing to high within-individual genetic variation) compared to the full taxon set. The colour scheme pronounces differences by highlighting low values in blue, and high in red.

Clade	Calibration set ^a				
	A	B	C	D	E
A	12.7%	12.5%	-6.9%	11.5%	10.6%
A1	9.4%	9.6%	-8.4%	9.0%	7.9%
AB	21.9%	21.7%	-3.3%	20.5%	19.1%
B	25.2%	24.2%	-0.9%	24.0%	21.7%
B1 ^b	27.6%	26.9%	2.2%	26.4%	23.7%
B1b	27.2%	27.0%	0.2%	26.1%	23.5%
B2	43.4%	43.2%	13.1%	43.1%	41.3%
B3	54.5%	56.7%	1.0%	55.1%	52.4%
B4	74.3%	71.2%	32.4%	74.5%	69.0%
Overall	32.9%	32.6%	3.3%	32.3%	29.9%

^a Calibration sets were A (BER rel., MSC rel.); B (BER, MSC); C (BER, MSC, CAN); D (BER rel., MSC); E (MSC), where calibration points were:

BER = Beringia: Nearctic/Holarctic split within *Poecile* at 5.4–5.5 My or rel. = relaxed at 4.8–14.0 My;

MSC = Messinian Salinity Crisis: African/European split within *Periparus ater* at 5.32–5.33 My or rel. = relaxed at 0–5.96 My;

CAN = Canary Islands volcanic age at El Hierro 0–1.2 My and La Palma 0–3.5 My.

^b *C. t. palmensis* (La Palma) and *C. t. cyrenaicae* (Libya) enforced as a monophyletic clade according to species tree analyses (see text).

Table S10. Average difference, based on four genomic subsets, in estimated node ages between different calibration sets, calculated for the partial taxon set (excluding individuals from Lanzarote and Tenerife, owing to high within-individual genetic variation). The colour scheme pronounces differences by highlighting low values in blue, and high in red.

Clade	Average difference between calibration sets ^a (My)									
	B vs. A	C vs. A	D vs. A	E vs. A	C vs. B	D vs. B	E vs. B	D vs. C	E vs. C	E vs. D
AB	20.2%	-5.4%	21.9%	21.7%	-21.2%	1.5%	1.2%	29.1%	28.9%	-0.2%
A	19.9%	-6.3%	21.6%	21.4%	-21.6%	1.5%	1.2%	29.9%	29.8%	-0.2%
A1	19.4%	-11.6%	21.7%	21.9%	-25.9%	1.9%	2.1%	38.5%	38.9%	0.2%
B	19.4%	-18.1%	21.2%	21.6%	-31.3%	1.5%	1.8%	48.6%	49.2%	0.3%
B1 ^b	18.7%	-18.1%	21.1%	21.5%	-30.8%	2.1%	2.4%	48.5%	49.0%	0.3%
B1b	19.0%	-9.5%	22.0%	21.9%	-23.9%	2.6%	2.4%	36.3%	36.3%	-0.1%
B2	19.4%	-17.6%	21.3%	21.0%	-30.9%	1.6%	1.4%	47.6%	47.3%	-0.2%
B3	17.7%	-35.9%	22.0%	21.7%	-45.3%	3.6%	3.4%	94.6%	94.3%	-0.2%
B4	20.4%	-14.0%	20.7%	21.6%	-28.4%	0.3%	1.0%	40.7%	41.9%	0.7%
Overall	19.3%	-15.2%	21.5%	21.6%	-28.8%	1.8%	1.9%	46.0%	46.2%	0.1%

^a Calibration sets were A (BER rel., MSC rel.); B (BER, MSC); C (BER, MSC, CAN); D (BER rel., MSC); E (MSC), where calibration points were:

BER = Beringia: Nearctic/Holarctic split within *Poecile* at 5.4–5.5 My or rel. = relaxed at 4.8–14.0 My;

MSC = Messinian Salinity Crisis: African/European split within *Periparus ater* at 5.32–5.33 My or rel. = relaxed at 0–5.96 My;

CAN = Canary Islands volcanic age at El Hierro 0–1.2 My and La Palma 0–3.5 My.

^b *C. t. palmensis* (La Palma) and *C. t. cyrenaicae* (Libya) enforced as a monophyletic clade according to species tree analyses (see text).

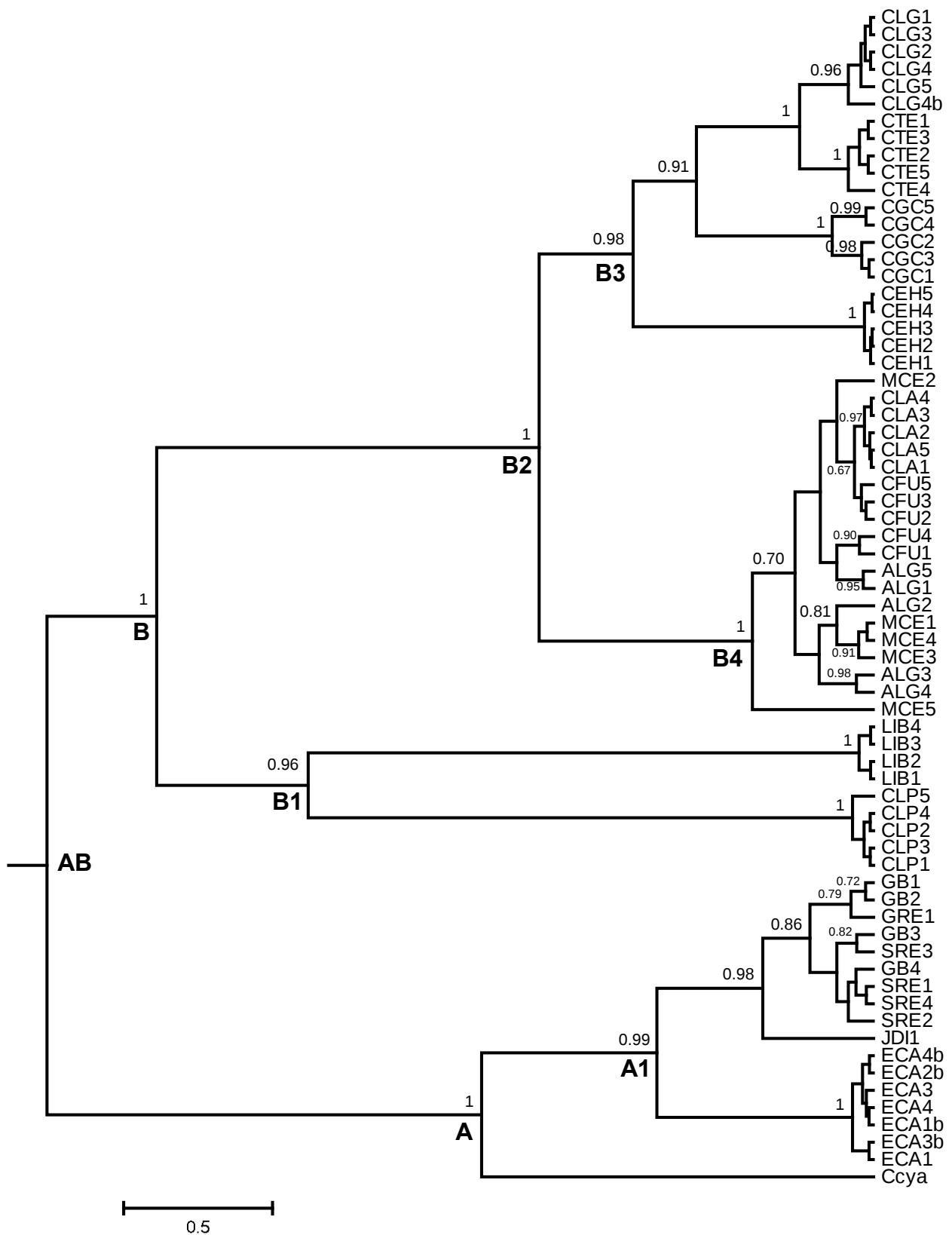


Figure S1. Phylogenetic tree for the mitochondrial control region, co-estimated with the species tree (Figure 2a) using *BEAST. Posterior probability >0.6 displayed at nodes; clades referred to in the text; sample names listed in Table S1. Populations: CLG = La Gomera, CTE = Tenerife, CGC = Gran Canaria, CEH = El Hierro, CLA = Lanzarote, CFU = Fuerteventura, MCE = Morocco, ALG = Algeria, LIB = Libya, CLP = La

Palma, GB = Great Britain, GRE = Greece, SRE = Sweden, ECA = Spain, JDI = Jordan, Ccya = *Cyanistes cyanus*.

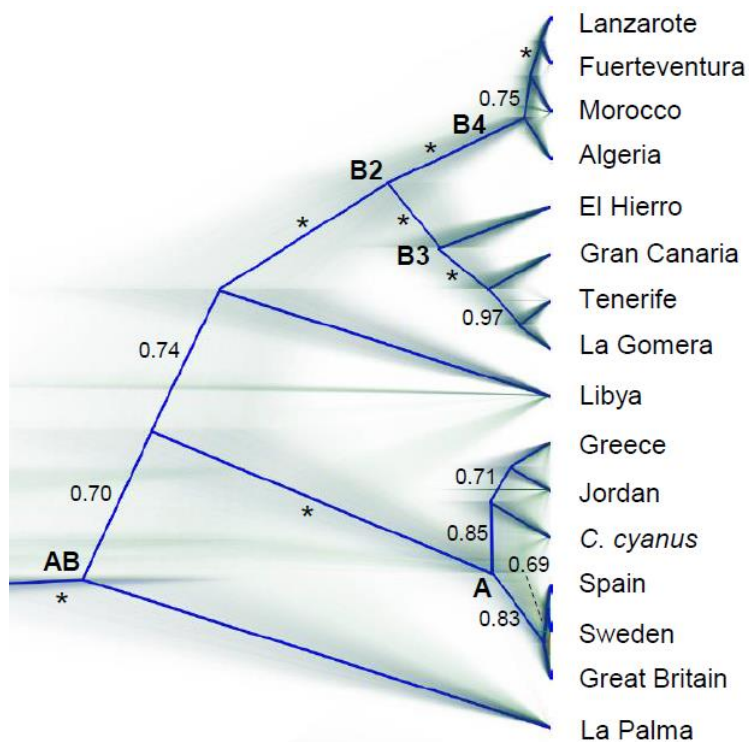
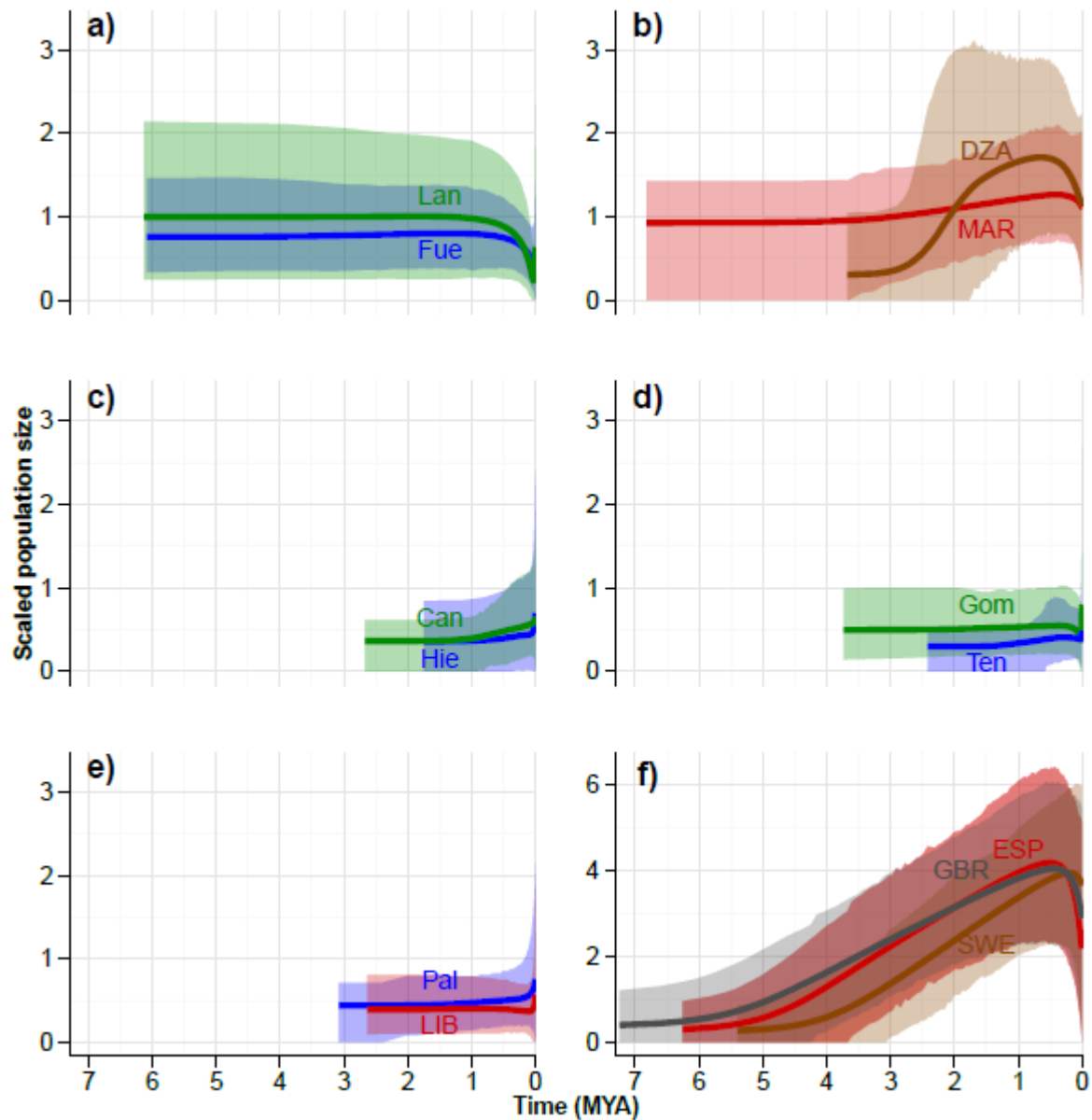


Figure S2. Coalescence based species trees of blue tits *Cyanistes* sp. based on analyses of population samples of 18 nuclear Sanger sequence markers (but no mitochondrial markers), estimated with *BEAST. Trees are drawn in DensiTree with individual trees represented by very thin lines, making up tree densities, and main topology highlighted with a distinct blue line. Posterior probabilities higher than 0.6 are stated along the branches (* = 1.0). Clade names are referred to in the text. Populations of *C. teneriffae* and *C. caeruleus* are referred to by their geographical origin.



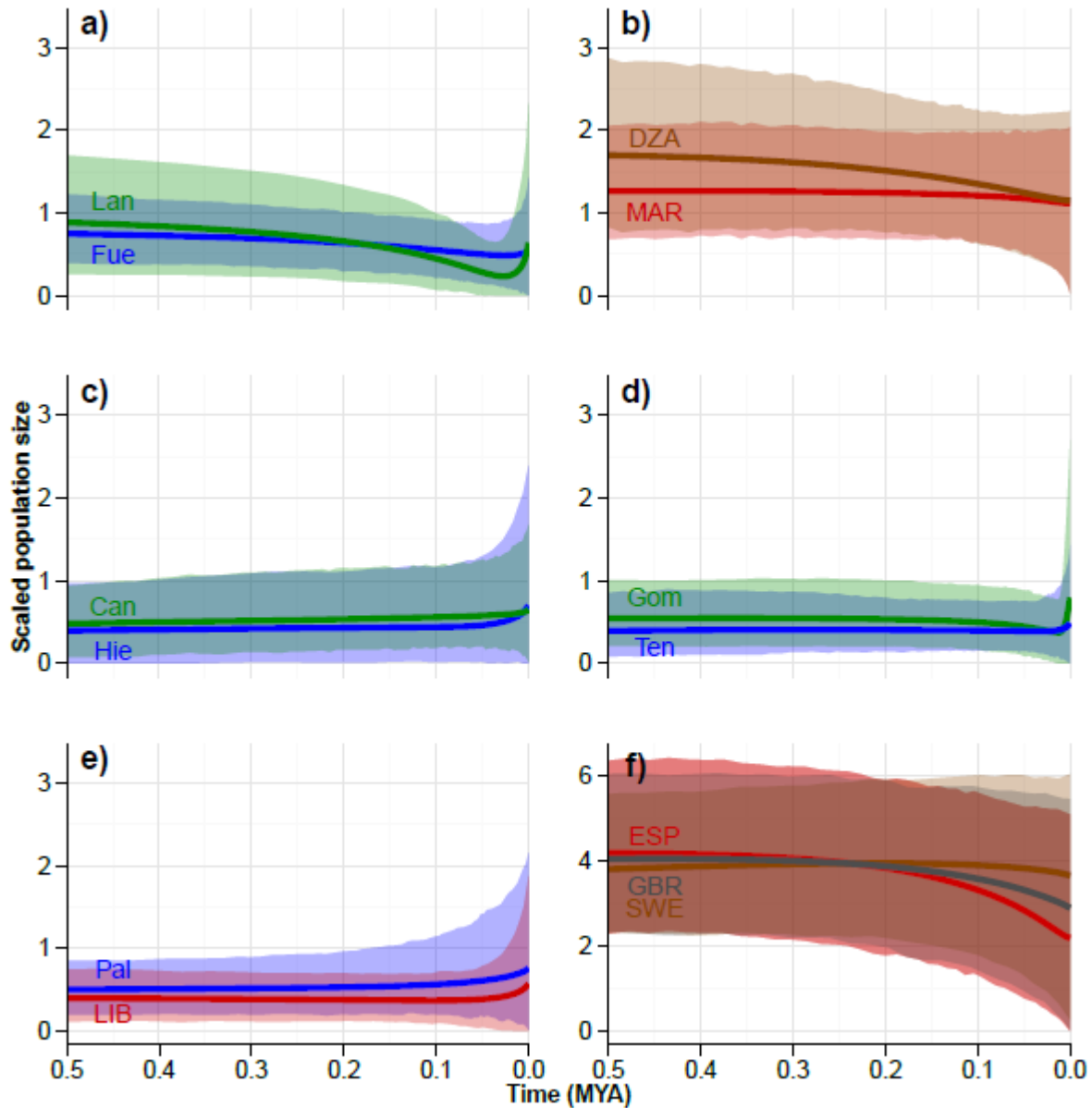


Figure S4. Demographic reconstruction of African Blue Tit *Cyanistes teneriffae* populations using Extended Bayesian Skyline Plots based on multi-locus Sangers sequence data. Solid lines represent mean estimated scaled population size (effective population size x generation time in My), and shaded areas the 95 % highest posterior density. Populations are indicated with abbreviations and island taxa are drawn in blue/green whereas mainland populations are drawn in red/brown/grey. (a) Eastern Canary Islands: Fuerteventura (Fue; blue) and Lanzarote (Lan; green), (b) Northwest Africa: Morocco (MAR; red) and Algeria (DZA; brown), (c) central Canary Islands: Gran Canaria (Can; green) and El Hierro (Hie; blue), (d) central Canary Islands: La Gomera (Gom; green) and Tenerife (Ten; blue), (e) La Palma (Pal; blue) and Libya (LIB; red), (f) European populations: Spain (ESP; red), Great Britain (GBR; grey), and Sweden (SWE; brown). Note that the last panel (f) has a different scale of the y axis. This figure is based on the same analyses as Figure 3, but is zoomed in to make clear the demographic developments during relatively recent times.

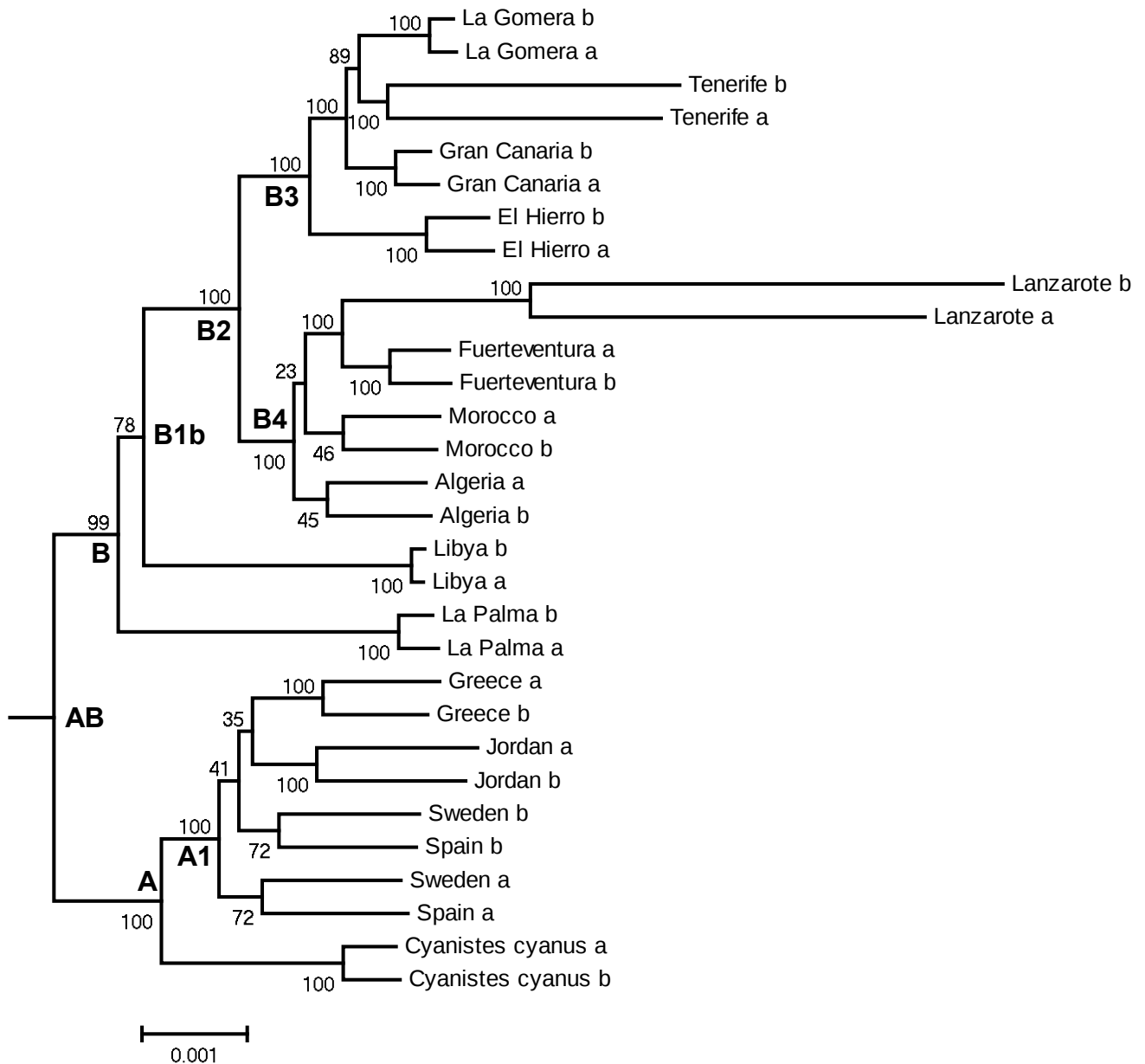


Figure S5. Phylogenomic maximum likelihood tree of blue tits *Cyanistes* sp. calculated in RAxML of super-matrix STRICT, comprising 11,426 restriction-site associated DNA (RAD) sequences (994 Kbp) from all individuals, with no tolerance for missing data. In the sequence names, “a” and “b” signify the two “alleles” produced by randomly combining alleles from each individual RAD locus. Bootstrap support >50 displayed at nodes; clades referred to in the text; outgroups not shown. Note the within-sample genetic variation, translated as long branches, for samples from Lanzarote and Tenerife, believed to be artefactual (see text).

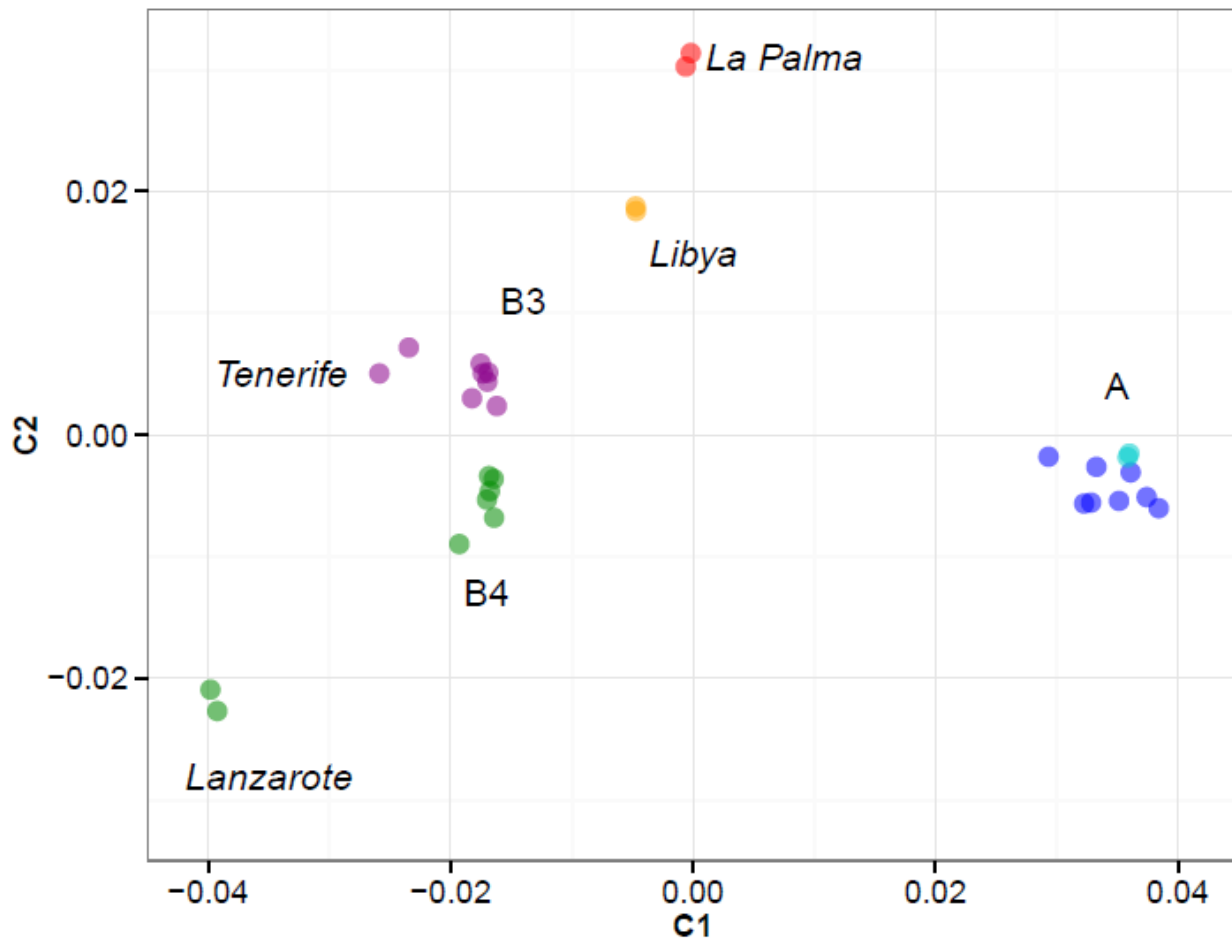


Figure S6. Cluster assignment using multi-dimensional scaling (MDS). Main clades (see text) and some single populations (*italic*) are labeled. Clade A holds European populations of *Cyanistes caeruleus* (dark blue) and Asian *C. cyanus* (cyan blue); clade B3 contains *C. teneriffae* of the central Canary Islands (purple); clade B4 comprises the eastern Canary Islands (including Lanzarote) and northwest Africa (green); Libya is drawn in orange; La Palma in red.

Supplementary references

- Backstrom N, Fagerberg S, Ellegren H (2008) Genomics of natural bird populations: a gene-based set of reference markers evenly spread across the avian genome. *Mol Ecol* **17**, 964-980.
- Baird NA, Etter PD, Atwood TS, *et al.* (2008) Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS One* **3**.
- Dietzen C, Garcia-del-Rey E, Castro GD, Wink M (2008) Phylogeography of the blue tit (*Parus teneriffae*-group) on the Canary Islands based on mitochondrial DNA sequence data and morphometrics. *Journal of Ornithology* **149**, 1-12.
- Etter PD, Bassham S, Hohenlohe PA, Johnson EA, Cresko WA (2011) SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Methods Mol Biol* **772**, 157-178.

- Fregin S, Haase M, Olsson U, Alstrom P (2009) Multi-locus phylogeny of the family Acrocephalidae (Aves: Passeriformes) - The traditional taxonomy overthrown. *Mol Phylogenet Evol* **52**, 866-878.
- Friesen V, Congdon B, Kidd M, Birt T (1999) Polymerase chain reaction (PCR) primers for the amplification of five nuclear introns in vertebrates. *Mol Ecol* **8**, 2147-2149.
- Helm-Bychowski K, Cracraft J (1993) Recovering Phylogenetic Signal from DNA-Sequences - Relationships within the Corvine Assemblage (Class Aves) as Inferred from Complete Sequences of the Mitochondrial-DNA Cytochrome-B Gene. *Mol Biol Evol* **10**, 1196-1214.
- Kocher TD, Thomas WK, Meyer A, *et al.* (1989) Dynamics of Mitochondrial-DNA Evolution in Animals - Amplification and Sequencing with Conserved Primers. *Proc Natl Acad Sci U S A* **86**, 6196-6200.
- Kvist L, Ruokonen M, Lumme J, Orell M (1999) Different population structures in northern and southern populations of the European blue tit. *Journal of Evolutionary Biology* **12**, 798-805.
- Li N, Stephens M (2003) Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics* **165**, 2213-2233.
- Prager EM, Sage RD, Gyllenstein U, *et al.* (1993) Mitochondrial-DNA Sequence Diversity and the Colonization of Scandinavia by House Mice from East Holsten. *Biological Journal of the Linnean Society* **50**, 85-122.
- Primmer C, Borge T, Lindell J, Saetre G (2002) Single-nucleotide polymorphism characterization in species with limited available sequence information: High nucleotide diversity revealed in the avian genome. *Mol Ecol* **11**, 603-612.