Pronounced Fixation, Strong Population Differentiation and Complex Population History in the Canary Islands Blue Tit Subspecies Complex

Bengt Hansson1, Marcus Ljungqvist1, Juan-Carlos Illera2,3, Laura Kvist4

1 Department of Biology, Lund University, Lund, Sweden, 2 Island Ecology and Evolution Research Group, La Laguna, Tenerife, Spain, 3 Research Unit of Biodiversity, Oviedo University, Oviedo, Spain, 4 Department of Biology, University of Oulu, Oulu, Finland

Abstract

Evolutionary molecular studies of island radiations may lead to insights in the role of vicariance, founder events, population size and drift in the processes of population differentiation. We evaluate the degree of population genetic differentiation and fixation of the Canary Islands blue tit subspecies complex using microsatellite markers and aim to get insights in the population history using coalescence based methods. The Canary Island populations were strongly genetically differentiated and had reduced diversity with pronounced fixation including many private alleles. In population structure models, the relationship between the central island populations (La Gomera, Tenerife and Gran Canaria) and El Hierro was difficult to disentangle whereas the two European populations showed consistent clustering, the two eastern islands (Fuerteventura and Lanzarote) and Morocco weak clustering, and La Palma a consistent unique lineage. Coalescence based models suggested that the European mainland forms an outgroup to the Afrocanarian population, a split between the western island group (La Palma and El Hierro) and the central island group, and recent splits between the three central islands, and between the two eastern islands and Morocco, respectively. It is clear that strong genetic drift and low level of concurrent gene flow among populations have shaped complex allelic patterns of fixation and skewed frequencies over the archipelago. However, understanding the population history remains challenging; in particular, the pattern of extreme divergence with low genetic diversity and yet unique genetic material in the Canary Island system requires an explanation. A potential scenario is population contractions of a historically large and genetically variable Afrocanarian population, with vicariance and drift following in the wake. The suggestion from sequence-based analyses of a Pleistocene extinction of a substantial part of North Africa and a Pleistocene/Holocene eastward re-colonisation of western North Africa from the Canaries remains possible.

Introduction

Population divergence and speciation have fascinated evolutionary biologists ever since Darwin [1]. Divergence in allopatry, perhaps the most accepted mode of divergence, may occur over time in subdivided populations. Climatic cycles cause repeated range expansions and contractions of most populations, with vicariance, founder events as well as admixture following in the wake [2]. In birds, the genetic distances between newly diverged sister taxa translate to divergence times in the Pleistocene approximately 0.01–3.0 MYA [3,4]. The Pleistocene speciation hypothesis proposes that these speciation events occurred in isolated refugia over one to several full glacial cycles [5,6]. A good model system to study genetic effects of isolation in the context of Pleistocene speciation model would be a system where several subpopulations or subspecies are available. Several such study systems are found on the Canary Islands [7,8]. The Canary archipelago consists of seven large islands – from west to east: El Hierro, La Palma, La Gomera, Tenerife, Gran Canaria, Fuerteventura and Lanzarote – and are situated in the Atlantic, 100–500 km off the northwest coast of Africa (Figure 1). The islands are of increasing age from west to east (less than one up to twenty million years old) [9,10]. The volcanic origin, the geographical situation (isolated but still relatively close to the mainland), the altitude (the highest peak is 3,718 m above sea level) and the absence of any land bridge connecting the archipelago with the continent, has led to a unique flora and fauna with a high degree of endemism [11,12].

The blue tit (Cyanistes caeruleus) is found all over Europe, western Asia and northern Africa, including the Canary Islands [13]. Blue tits on the Canaries and North Africa are morphologically distinct from the European populations, and two subspecies complexes are distinguished: the European continental blue tits (the caeruleus group) and the Afrocanarian blue tit complex (the teneriffae group) [14–19]. Birds in the teneriffae group have dark blue or almost black cap, blue back and some populations lack the white wing pattern that is seen in the caeruleus group [20]. Further, the Afrocanarian complex have longer and thinner bills than the European...
subspecies (La Palma is an exception to this), shorter wings and longer tarsi [21] and their song differs to a great extent, even between the islands [22,23]. Suggestions for giving caeruleus and teneriffae full species status have been put forward [14–19].

Studies of the blue tit radiation on the Canary Islands could potentially lead to new insights in the role of colonization, vicariance, populations size and drift in population divergence (cf. [24]). However, to be able to do so the historical colonization patterns and admixture events need to be understood. The closest mainland sources from where colonizing birds could come are north-western Africa (100 km; C. c. ultramarinus) and the southern Iberian peninsula (950 km; C. c. ogliastrae). Based on morphological data, Grant [21] proposed a historical demographic scenario following a classical stepping stone model, where the old eastern islands (Lanzarote and Fuerteventura) were colonized first from the African continent, then followed by the colonisation of the central islands (Gran Canaria, Tenerife, and La Gomera) and from there the western islands (El Hierro and La Palma). This scenario also included that the original populations on Lanzarote and Fuerteventura went extinct, and became recolonized from one of the central Islands, namely Tenerife [21]. Molecular studies of Canary Islands blue tits, analysing mitochondrial DNA (mtDNA) sequence data [15,17–19] and recently also data from nuclear DNA (nDNA) [18,19], interpret the results according to three different scenarios. The first interpretation follows Grant’s [21] original model and proposes that one of the central islands (Tenerife) was the first to be colonized from Africa, perhaps via Lanzarote and Fuerteventura (whose original populations went extinct), and from there all other islands were colonized, including Lanzarote and Fuerteventura [15], or alternatively these two populations were recolonized from Africa [17]. The second model suggests that the Canary Islands were colonized by blue tits from the African continent, but also from Europe [15,17]. A main reason for this interpretation is that some island populations, in particular La Palma, share some specific genetic features with the European population (including a 12 bp mtDNA fragment) which are difficult to explain from the perspective of pure African colonisation [15]. However, recent work including additional North African populations does not support this model [18]. Finally, the third suggestion is that the Canaries represent an ancestral colonization from North Africa, and that these island populations later recolonized continental Africa after a large part of its original population had gone extinct [15,18,19]. This suggestion is particularly interesting since it would support a prominent role for small, peripheral populations as a source for wide-spread mainland populations, as opposed to the traditional view of such island populations forming population sinks (cf. [25,26]). In line with this reasoning, other recent avian phylogenetic studies have also described systems where island populations may have acted as sources and contributed to the diversity of adjacent mainland populations [26–28]. Apart from these three main suggestions several other, not necessarily mutually exclusive, scenarios can be suggested, including that the Afrocanarian population was historically very large, holding high levels of...
genetic variation with at least occasional dispersal between islands and between islands and the mainland, and that relatively recent (Holocene) population declines and vicariance events have shaped the complex pattern of differentiation observed in sequence data of contemporary populations.

In the present study, we evaluate the degree of population genetic variation, fixation, differentiation and structure of the Canary Islands and North African blue tit subspecies complex using microsatellite markers and aim to get insights in the population history using coalescence based methods. In blue tits, an extensive number of nuclear microsatellites have been developed [29] and a first generation linkage map has been constructed [30]. This enabled us chose a set of markers with known linkage map position for our analyses. We analyse our genetic data using traditional methods based on predefined populations [31], as well as Bayesian clustering models [32,33] and coalescence-based models [34], to explore the population history of Africanarian blue tits. These models allows testing different scenarios of population divergence, admixture and population size changes [34], and we were interested in evaluating the split between the European and the Afrocanarian lineages, the split between different Canary Islands, and, in particular, the split between the Canaries and North Africa.

**Materials and Methods**

**Study Populations**

This study included samples from the seven Canary Islands populations, as well as from North Africa (Morocco and Ceuta), and the mainland of Spain and Sweden (see Figure 1). In categorizing the island populations into subspecies, we followed Kvist et al. [15], Kvist [16] and Dietzen et al. [17]. According to Kvist [16] and Dietzen et al. [17] there are five subspecies breeding on the Canary Islands: *umbrosus* on El Hierro, *palmenis* on La Palma, *tenerifae* on La Gomera and Tenerife, *hedvigi* on Gran Canaria, and *ultramarius* on Fuerteventura and Lanzarote. The subspecies *ultramarius* is also found in North Africa, whereas *ogliastrae* occurs in Spain and *caeruleus* in Sweden. In Kvist et al. [15], the birds on Fuerteventura and Lanzarote were separated from North Africa and put into a separate subspecies, *degneri*, but the newer data indicate that this may not be the case [16,17]. The subspecies categorisation of Dietzen et al. [17] is shown on the map in Figure 1.

The birds were captured with mist-nets, measured and weighted, and a blood sample (≥50 μl) was taken by puncturing the brachial vein or with a syringe from the jugular vein. The blood samples were stored in a SET buffer containing 0.15 M NaCl, 0.05 M Tris and 0.001 M EDTA with a pH of 8.0, or in pure ethanol. The field work has been described elsewhere [15,19]. The study and the protocols for handling and examining the birds were approved by the County Administrative Board and the Lund/Malmö Animal Review Board in Sweden, the Ministerio de Agricultura, Alimentacion y Medio Ambiente in Spain and the Haut Commissariat aux Eaux et Forêts et a la Hutte Contrue la Desertification in Morocco. After capture and examination the birds were immediately released into the wild.

**Molecular Markers and Genotyping**

We selected a set of 21 microsatellite markers of which some were linked (located on the same chromosome) and some unlinked (situated on unique chromosomes) (Table S1) [38,42], and genotyped a total of 206 blue tits. Summary statistics for all loci were linked (located on the same chromosome) and some unlinked (situated on unique chromosomes) (Table S1) [38,42], and genotyped a total of 206 blue tits. Summary statistics for all loci in Table S1. All birds were molecularly sexed by amplifying a Z- and W-linked locus, TGu9-002 (D. Dawson, University of Sheffield, unpublished), and this information was used for interpreting the genotypes of the Z-linked microsatellites.

All loci were PCR-amplified in three different multiplexes using QIAGEN Multiplex PCR kit (Qiagen, ltd.). Primer sequences and annealing temperatures for the microsatellites are given in Table S1 (see also Olano-Marin et al. [29] and Hansson et al. [30]). The PCR-products were separated and visualized using an ABI 3730 capillary sequencer (Applied Biosystems), and the genotypes were scored with Genemapper 4.0 (Applied Biosystems).

**Analyses of Population Fixation, Differentiation and Structure**

To evaluate the genetic variability in each populations, we calculated number of alleles, allele richness, observed and expected heterozygosity [36] and *F*<sub>ST</sub> [37] in FSTAT ver. 2.9.3.2 [38]. We used this program also to evaluate deviations from Hardy–Weinberg equilibrium, calculated *F*<sub>ST</sub> between all pairs of populations and tested these statistically (10,000 permutations; using a nominal level of 0.001).

In addition to *F*<sub>ST</sub>, Jost’s *D*<sub>st</sub> was used as a measure of genetic differentiation between populations [31] and calculated for each population pair using the web based resource SMOGD v. 1.2.5 [39]. *D*<sub>st</sub> is a relative measure of differentiation, which ranges from zero (no differentiation) to one (complete differentiation), and simulations have shown that it is an unbiased estimator of differentiation, and outperforms *F*<sub>ST</sub> over a range of sample sizes and for markers with different numbers of alleles (including highly variable microsatellite loci) [40]. We used 1,000 bootstrap replicates and the harmonic mean of *D*<sub>st</sub> across loci.

We evaluated population structure with the Bayesian clustering processes and MCMC simulations implemented in the program STRUCTURE [32]. By exploring a parameter space consisting of multi-locus allele frequencies of genetic clusters, the STRUCTURE algorithms search for clusters that maximize the likelihood that the observed individual genotypes belong to them. When there is strong population structure, the clusters that have the highest likelihood will coincide with a structure that minimizes deviations from Hardy–Weinberg equilibrium (single-locus measure) and linkage equilibrium (multi-locus measure) within the clusters. The most common and perhaps most biological meaningful way of running the STRUCTURE models is to use ‘admixture models’, which allow (but do not force) individuals to have a genetic origin from more than one genetic cluster [32]. For unlinked loci, any linkage disequilibrium in a data set is attributed to presently occurring substructuring. However, after an admixture event, linked loci will persist in linkage disequilibrium within a population for a time period that is inversely related to the rate of recombination between loci. When the genetic distance between loci is known, one can model linkage disequilibrium due to substructuring as well as due to linkage, and ‘linkage models’ that take both these types of linkage disequilibria into account have been implemented in later versions of STRUCTURE [33]. The linkage model can potentially provide additional information about the ancestry of individuals and may improve understanding complex relationships between populations [33]. We performed admixture and linkage models with STRUCTURE ver. 2.3.3 [32,33].

The genetic distances between loci were provided (and came from Hardy–Weinberg equilibrium (single-locus measure) and linkage equilibrium (multi-locus measure) within the clusters. The most common and perhaps most biological meaningful way of running the STRUCTURE models is to use ‘admixture models’, which allow (but do not force) individuals to have a genetic origin from more than one genetic cluster [32]. For unlinked loci, any linkage disequilibrium in a data set is attributed to presently occurring substructuring. However, after an admixture event, linked loci will persist in linkage disequilibrium within a population for a time period that is inversely related to the rate of recombination between loci. When the genetic distance between loci is known, one can model linkage disequilibrium due to substructuring as well as due to linkage, and ‘linkage models’ that take both these types of linkage disequilibria into account have been implemented in later versions of STRUCTURE [33]. The linkage model can potentially provide additional information about the ancestry of individuals and may improve understanding complex relationships between populations [33]. We performed admixture and linkage models with STRUCTURE ver. 2.3.3 [32,33].

The genetic distances between loci were provided (and came from the blue tit linkage map [30]) and allele frequencies were allowed to be correlated between populations in the models. For the full data set including 19 loci (i.e. excluding TGu9 and Pca8; see below), we started each run with a burn-in period of 50,000 replicates, followed by a sampling period of 50,000 replicates. We
also conducted separate analyses for the different linkage groups (LG1b, four loci; LG2, five loci; LGZ, five loci; Table S1) using a burn-in of 10,000 and a sampling period of 20,000 replicates. For each data set and model, we set the number of clusters (K) from 1 to 10, and used 20 iterations.

The most likely K was evaluated with the ΔK-method [41], where the change in log probability of data between two Ks, and the variation in probability within a K, are used instead of the K-specific probability per se to evaluate which K has the strongest support from the data. ΔK was calculated with STRUCTURE HARVESTER (http://taylor02biology.ucla.edu/structure_harvester/).

The posterior probability of the admixture proportions of each individual from Structure was visualised using DISTRUCT (N.A. Rosenberg, University of Michigan) and CLUMPP [42].

Finally, we analyzed our data using the mtDNA subspecies classification as a prior. We performed these analyses following the subspecies structure suggested by Kvist et al. [15] and Dietzen et al. [17], respectively. We used admixture models with correlated allele frequencies, burn-in periods of 10,000, sampling periods of 20,000 replicates and 20 iterations. K was set to 7 (i.e. the number of subspecies). This procedure was used in order to evaluate whether the clustering process may be facilitated by such priors, and thus be able to support any of the suggested subspecies classifications.

### Analyses of Population History with Coalescence Models

We used the program DIYABC v. 1.0.4.45 [34] to explore the population history of blue tits. DIYABC is a coalescence-based program that infers the population history by looking backwards in time to examine genealogy of alleles until reaching the most recent common ancestor using approximate Bayesian computation algorithm. We chose 14 loci based on map distances at least 10 cM apart for diminishing any potential effect of linkage. The loci used were both autosomal (Cdi31-ZFM, CcaTgu19, CcaTgu21, Ase18, Pdoq5, ApCo46-ZEST, LS2, PmaTGAn42, TG02-088, Titgata02) and sex-linked (Ase46-ZFM, CcaTgu31, TGZ-040, Phn3). Based on previous studies [17], we started by building three different combinations of splitting between European, African and Canary Island populations, assuming no admixture after splitting events. We used the same mean mutation rate for all the loci, 10^{-1}-10^{-3} [43,44]. Priors for all three effective population sizes were 100–10,000,000 generations, again depending on the population. We simulated 4,000,000 data sets for scenario 1, 5,000,000 for scenario 2 and 6,000,000 for scenarios 3. Priors for effective population sizes were 100–10,000 individuals and for splitting times 100–1,000,000 generations, again depending on the population. The best scenarios from these runs (see below) included simultaneous splits within the eastern, central and western islands and the European population as an outgroup. These were included in the final run, where we built six scenarios differing in relation to when the splits occurred and from where the branch leading to eastern islands originate. All islands were treated as distinct populations and 6,000,000 simulations were run.

### Results

#### Genetic Variation, Fixation and Differentiation

Two of the microsatellite loci included in the present study did not amplify in at least one population: Tgu09 did not amplify in samples from the Canary Islands and North Africa, and PcaB did not amplify in the samples from La Palma (Table S1). The reason for this is most likely that there are substantial primer site mutations leading to amplification failure in these populations. These two loci were excluded from further analyses, but we note that the pattern observed for Tgu09 supports a relationship between populations on the Canaries and North Africa.

For the remaining 19 loci, the gene diversity, allele richness, and number of alleles were highest for the three mainland populations, Sweden, Spain and North Africa (Table 1). Observed and expected heterozygosities were similar, FIS was low in all cases, and there were no statistical deviations from Hardy–Weinberg equilibrium within populations (Table 1).

There was strong fixation and differentiation between populations, with FST-values ranging between 0.022 for Sweden and Spain and 0.636 for El Hierro and La Palma, and with Dst values ranging between 0.015 for Sweden and Spain and 0.793 for La Palma and Tenerife (Table 2). The pair-wise FST- and Dst-values were highly correlated (Pearson’s correlation: r = 0.667). All FST-values were significantly different from zero (p<0.001).

#### Population Structure using STRUCTURE

STRUCTURE analyses supported the presence of strong population differentiation (Figure 2). Admixture models showed high membership assignment of specific individuals and populations independently of which K is modelled, but weak consistency in clustering of some populations within and between Ks. This has probably to do with the fact that there is a high degree of fixation of unique alleles in the populations on the Canary Islands (Table 1), and that some populations are sharing skewed allele frequencies with a few other populations with different combinations of populations for different loci. The relationship between the central island populations (La Gomera, Tenerife and Gran Canaria) and El Hierro was difficult to disentangle with different combinations of populations clustering in different iterations. Moreover, the two eastern islands (Fuerteventura and Lanzarote) and Morocco showed weak clustering with partly consistent clustering (in particular at low Ks), whereas the two European populations showed highly consistent clustering, and La Palma a consistent unique lineage in the majority of iterations (Figure 2).

The ΔK-method suggested 6 clusters as the most likely population structure (Figure 3), but the low ΔK for K = 7 was mainly an effect of a single run with very low posterior probability.
The STRUCTURE analyses using linkage models, where the genetic distances between linked loci were taken into account, gave very similar results compared to the admixture analyses. This was true for the full data set of 19 loci as well as for the three linkage groups separately. Thus, modelling linkage did not provide any additional resolution of the genetic relationship between populations (data not shown).

We also performed admixture models using the subspecies classifications suggested by Kvist et al. [25] and Dietzen et al. [27] as priors. \( K \) was set to 7 (i.e. the number of suggested subspecies) and 20 iterations were performed. These analyses were made because we were interested in evaluating whether detecting a structure corresponding to the subspecies categorisations would be facilitated by using such priors. However, the substructure groupings suggested by Kvist et al. [15] and by Dietzen et al. [17] were only partly supported: most strikingly teneriffae on La Gomera and teneriffae on Tenerife did not cluster together more often than in runs without priors, teneriffae (hedwigii) on Gran Canaria did not cluster with teneriffae on Tenerife, and ultramarinus/degener in North Africa, Lanzarote and Fuerteventura did not cluster in these runs (Figure S3).

**Table 1.** Genetic characteristics of blue tit populations at 19 microsatellite loci including number of genotyped individuals, number of alleles, number of monomorphic loci, number of unique alleles, allelic richness, observed heterozygosity (\( H_o \)), expected heterozygosity (\( H_e \)) and \( F_{IS} \).

<table>
<thead>
<tr>
<th>Population</th>
<th>Sample size</th>
<th>Mean number of alleles</th>
<th>Number of monomorphic loci</th>
<th>Number of unique alleles</th>
<th>Allele richness</th>
<th>( H_o )</th>
<th>( H_e )</th>
<th>( F_{IS} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
<td>20</td>
<td>7.5</td>
<td>1</td>
<td>22</td>
<td>5</td>
<td>0.65</td>
<td>0.63</td>
<td>−0.03</td>
</tr>
<tr>
<td>Spain</td>
<td>22</td>
<td>7.3</td>
<td>0</td>
<td>16</td>
<td>5</td>
<td>0.61</td>
<td>0.61</td>
<td>−0.01</td>
</tr>
<tr>
<td>El Hierro</td>
<td>13</td>
<td>2.5</td>
<td>9</td>
<td>3</td>
<td>2</td>
<td>0.26</td>
<td>0.28</td>
<td>0.02</td>
</tr>
<tr>
<td>La Palma</td>
<td>24</td>
<td>3.4</td>
<td>6</td>
<td>16</td>
<td>3</td>
<td>0.31</td>
<td>0.33</td>
<td>0.05</td>
</tr>
<tr>
<td>La Gomera</td>
<td>21</td>
<td>3.5</td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>0.34</td>
<td>0.34</td>
<td>−0.01</td>
</tr>
<tr>
<td>Tenerife</td>
<td>25</td>
<td>5.1</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>0.45</td>
<td>0.46</td>
<td>0.01</td>
</tr>
<tr>
<td>Gran Canaria</td>
<td>22</td>
<td>4.8</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>0.48</td>
<td>0.46</td>
<td>−0.05</td>
</tr>
<tr>
<td>Fuerteventura</td>
<td>20</td>
<td>3.8</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>0.31</td>
<td>0.36</td>
<td>0.13</td>
</tr>
<tr>
<td>Lanzarote</td>
<td>17</td>
<td>1.9</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>0.27</td>
<td>0.25</td>
<td>−0.05</td>
</tr>
<tr>
<td>North Africa</td>
<td>22</td>
<td>5.8</td>
<td>1</td>
<td>11</td>
<td>4</td>
<td>0.58</td>
<td>0.54</td>
<td>−0.05</td>
</tr>
</tbody>
</table>

**Table 2.** \( F_{ST} \)-values (below the diagonal) and \( D_{est} \)-values (above diagonal) between blue tit populations based on genotypic data from 19 loci.

<table>
<thead>
<tr>
<th></th>
<th>Swe</th>
<th>Spa</th>
<th>El Hie</th>
<th>La Pal</th>
<th>La Gom</th>
<th>Ten</th>
<th>Gr Can</th>
<th>Fue</th>
<th>Lan</th>
<th>N Afr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
<td>*</td>
<td>0.015</td>
<td>0.604</td>
<td>0.667</td>
<td>0.684</td>
<td>0.504</td>
<td>0.504</td>
<td>0.625</td>
<td>0.600</td>
<td>0.425</td>
</tr>
<tr>
<td>Spain</td>
<td>0.022</td>
<td>*</td>
<td>0.557</td>
<td>0.711</td>
<td>0.633</td>
<td>0.473</td>
<td>0.518</td>
<td>0.642</td>
<td>0.581</td>
<td>0.459</td>
</tr>
<tr>
<td>El Hierro</td>
<td>0.421</td>
<td>0.415</td>
<td>*</td>
<td>0.725</td>
<td>0.232</td>
<td>0.194</td>
<td>0.354</td>
<td>0.300</td>
<td>0.324</td>
<td>0.305</td>
</tr>
<tr>
<td>La Palma</td>
<td>0.454</td>
<td>0.469</td>
<td>0.636</td>
<td>*</td>
<td>0.712</td>
<td>0.793</td>
<td>0.737</td>
<td>0.622</td>
<td>0.619</td>
<td>0.567</td>
</tr>
<tr>
<td>La Gomera</td>
<td>0.445</td>
<td>0.435</td>
<td>0.453</td>
<td>0.614</td>
<td>*</td>
<td>0.250</td>
<td>0.440</td>
<td>0.387</td>
<td>0.374</td>
<td>0.337</td>
</tr>
<tr>
<td>Tenerife</td>
<td>0.325</td>
<td>0.321</td>
<td>0.325</td>
<td>0.571</td>
<td>0.341</td>
<td>*</td>
<td>0.205</td>
<td>0.371</td>
<td>0.456</td>
<td>0.184</td>
</tr>
<tr>
<td>Gran Canaria</td>
<td>0.322</td>
<td>0.333</td>
<td>0.428</td>
<td>0.559</td>
<td>0.439</td>
<td>0.252</td>
<td>*</td>
<td>0.547</td>
<td>0.574</td>
<td>0.479</td>
</tr>
<tr>
<td>Fuerteventura</td>
<td>0.406</td>
<td>0.410</td>
<td>0.444</td>
<td>0.577</td>
<td>0.481</td>
<td>0.389</td>
<td>0.447</td>
<td>*</td>
<td>0.167</td>
<td>0.180</td>
</tr>
<tr>
<td>Lanzarote</td>
<td>0.442</td>
<td>0.431</td>
<td>0.552</td>
<td>0.635</td>
<td>0.538</td>
<td>0.467</td>
<td>0.515</td>
<td>0.374</td>
<td>*</td>
<td>0.277</td>
</tr>
<tr>
<td>North Africa</td>
<td>0.245</td>
<td>0.263</td>
<td>0.350</td>
<td>0.469</td>
<td>0.366</td>
<td>0.234</td>
<td>0.331</td>
<td>0.243</td>
<td>0.319</td>
<td>*</td>
</tr>
</tbody>
</table>
within these groups (scenario 4 in Figure S1; scenario 5 in Figure S2), with no admixture after the divergence (posterior probabilities 0.963–1.000). The direct approach, on the other hand, did not give strong support for any of the scenarios (posterior probabilities very even for all scenarios). Divergence times for a split between Fuerteventura, Lanzarote and Morocco were estimated to be only 120 generations ago (95% HPD 107–1,610 generations), between La Gomera, Tenerife and Gran Canaria 100 generations ago (95% HPD 100–270 generations) and between La Palma, El Hierro and Central Islands 1,710 generations ago (95% HPD 465–44,100 generations).

The final run included these simultaneous divergence events within island groups, but with different branching orders. Best support with the direct approach was given to a model where all
Canary Island and North African populations diverged simultaneously (scenario 6 in Figure 5, posterior probability 0.296). Logistic regression method gave the best support (posterior probability 0.261, scenario 2 in Figure 5) for the scenario including the most recent split between the eastern Islands and Morocco and a previous split between the central islands Tenerife, La Gomera and Gran Canaria from a common branch that had separated from a branch leading to La Palma and El Hierro, i.e. the eastern group and central group share a common ancestor. Divergence of Central Islands occurred after divergence of western islands La Palma and El Hierro. However, the support for these scenarios was only marginally stronger than for the other scenarios (posterior probabilities for the different scenarios varied between 0.091 and 0.296).

Discussion

Our results complement the findings from previous mtDNA and nDNA studies of the Canary Islands blue tit radiation [15,17–19] by showing that there is low genetic variation and strong fixation within island populations and high differentiation between populations at a larger set of nuclear microsatellite loci. The western and central Canary Islands populations, El Hierro, La Palma, La Gomera, Tenerife and Gran Canaria, stand out as very different genetically, and assessing how they are related is problematic. The eastern islands, Lanzarote and Fuerteventura, cluster with North Africa in admixture models with lower number of genetic clusters. These results only partly reflect the results from studies applying other loci [15,17–19]. That different sets of markers and loci are showing partly different patterns of population relationships is in line with expectations from drift scenarios with incomplete lineage sorting [45,46].
In an aim to achieve increased resolution of potential admixture events, we applied linkage models in STRUCTURE. The linkage model accounts for allelic correlations between linked loci that arise both due to population structure and due to admixture, i.e. ‘admixture linkage disequilibrium’ [4]. Admixture linkage disequilibrium will persist in a population for some time after an admixture event due to the reduced recombination between linked loci. Applying such models when using linked markers could improve the clustering process [33]. However, this was not the case in our analyses; the results from the admixture and the linkage models were similar. Perhaps our markers were not linked sufficiently close (map distance between adjacent markers ranged between 2 and 50 cM; Table 1), implying that any admixture linkage disequilibrium would have been broken down quite rapidly after an admixture event, or perhaps any admixture events took place a long time ago. Similar outcomes for admixture and linkage models could also arise if the differentiation that we observed today are mainly caused by drift and fixation in isolation e.g. due to vicariance, rather than by colonisation and admixture.

The Canary Islands were formed 1 to 20 MYA [20,21], and the phylogeography reconstructions by Illera et al. [19] and Packert et al. [18] suggest that the blue tit colonisation dates back 3 to 4 MYA. Our reconstructions of the population history in DIYABC suggest that the European population is an outgroup to the Afrocanarian population, and that the populations within the western, central and eastern groups split simultaneously. This result differs to some extent from those of the phylogeographical reconstructions, which show ancient splits between some islands – in particular, between La Palma and the other islands [18,19]. Our models suggest that most of the splits were comparably recent; assuming a generation time of c. 2 years, our divergence time estimates ranged from c. 38,000 (14,000–188,000) years between the European and Afrocanarian clades to c. 240 (200–3,200) years for the different central islands and for the eastern islands and North Africa. These timing estimates are based on rates of $10^{-4}$–$10^{-5}$ mutations per generation, i.e. commonly employed microsatellite mutation rates [43,44]. However, our timing estimates are much more recent than those being estimated with phylogenetic reconstructions [18,19], which indicates that the microsatellite loci we have screened have lower mutation rates than commonly assumed.

The pronounced divergence with low genetic diversity and yet unique genetic material in the Canary Island system requires an explanation. The population at La Palma is particularly intriguing since it shows highly divergent patterns on our sets of markers as well as on mtDNA and nDNA sequence data [15,17–19]. La Palma (i) shares a unique 12 bp mtDNA (control region) fragment with the European population which is not present in the other Afrocanarian populations (which indicates a Eurasian ancestry at this locus), (ii) differs with 34 nucleotides to the Gran Canaria population at the cytochrome b (1,005 bp) [17] and at 26 nucleotides to Tenerife at the control region (539 bp) [15], (iii) has six loci with fixed alleles and 16 unique alleles at the 19 loci we have studied (and at locus Pca8 it is the only population with null alleles), and (iv) shares null-alleles at the Z-linked marker Tgu09 with all other Afrocanarian populations (which indicates Afrocanarian ancestry). Interestingly, a recent phylogenetic study that included a previously unstudied Libyan population of blue tits (subspecies cyanaceae in the teneriffae group) in the eastern part of the species’ Afrocanarian range, clustered the Libyan population as a basal lineage of C. teneriffae together with La Palma [18]. These two eastern and western peripheral populations at the Afrocanarian range shared all major indels at the studied sequences and differed in this respect from the other taxa in the Afrocanarian group. Such large genetic distances over a widespread area suggests an old

history of isolation. The remote location of La Palma northwest of the other islands has implied only sporadic inflow of immigrants. Also the Libyan population shows distinct allopatry, being located east of the nearest African population in Tunisia (see Figure 1; [18]). It is worth noticing that La Palma has low but not the lowest degree of microsatellite genetic variation, which may suggest a relatively large long-term population size.

In conclusion, most mtDNA and nDNA haplotypes [15,18,19] and several microsatellite alleles [19] [this study] are unique to the Canaries and are not found on the mainland populations. A phylogeographical scenario that could explain the substantial amount of genetic variation that is observed today among all the different populations in the teneriffae complex is the presence of a historically much larger Afrocanarian population in which genetic variation was generated and maintained (cf. [47]). Such a population could have been distributed over a substantial part of North Africa and the Canaries. Indeed, North Africa has been going through several cycles of climatic shifts, including a major transition from wet to dry habitat after the last glaciation [48], with possible strong impact on the amount of suitable habitat for a forest species like the blue tit. Historical dispersal events within and between the Afrocanarian region and a relatively recent population contraction in North Africa resulting in vicariance, drift, lineage sorting and extinctions and recolonizations could have resulted in the complex population structures that we observe today. In line with recent promising undertakings by Illera et al. [19] and Packert et al. [18], further phylogenetic and coalescent reconstructions using sequence data from a larger set of nuclear loci and populations will be decisive to understand the processes that shaped the blue tit diversity in the region. Such information could confirm the interesting possibility that some of the Canary Islands populations have acted as a source for the current populations on the western parts of North African. This would be of particular interest since it would highlight the importance of small, island populations as sources for mainland colonisations, as opposed to the traditional view of such peripheral populations acting as population sinks (cf. [18,19,25,26]).

Supporting Information

**Figure S1** Graphical illustration of five historical events tested with DIYABC. Pop 1 = Europe (Spain and Sweden), Pop 2 = La Palma, Pop 3 = El Hierro, Pop 4 = Central Islands (La Gomera, Tenerife, Gran Canaria).

TIFF

**Figure S2** Graphical illustration of six historical events tested with DIYABC. Test of the eastern Islands and North Africa: Pop 1 = Fuerteventura, Pop 2 = Lanzarote, and Pop 3 = North Africa (Morocco). Test of the Central Islands: Pop 1 = Tenerife, Pop 2 = Gran Canaria, and Pop 3 = La Gomera.

TIFF

**Figure S3** Proportion of membership of individual blue tits of different populations to $K = 7$ clusters corresponding to the subspecies categorisation of (a) Kvist et al. (2005) and (b) Dietzen et al. (2008). All 20 iterations for $K = 7$ are shown (sorted according to declining posterior probability from top to bottom). Results are from admixture models using the full set of 19 loci. Populations are, from left to right: Sweden (Sw), Spain (Sp), El Hierro (Hi), La Palma (Pa), La Gomera (Go), Tenerife (Te), Gran Canaria (GC), Fuerteventura (Fu), Lanzarote (La) and Morocco (Mo).

EPS
Table S1 Genomic location of 21 microsatellite loci on the blue tit linkage map and the zebra finch genome assembly. Total number of alleles, and average observed heterozygosity (Ho), expected heterozygosity (He) and FIS, in the populations are given, followed by primer sequences and annealing temperatures (Ta; TD indicates touch-down PCR).

<table>
<thead>
<tr>
<th>Locus</th>
<th>Chromosome</th>
<th>Position (cM)</th>
<th>Ho</th>
<th>He</th>
<th>FIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tss1</td>
<td>1</td>
<td>123456</td>
<td>1</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>Tss2</td>
<td>2</td>
<td>789012</td>
<td>3</td>
<td>4</td>
<td>0.6</td>
</tr>
</tbody>
</table>

References


Author Contributions

Conceived and designed the experiments: BH ML, JCI, LK. Analyzed the data: BH ML, LK. Contributed reagents/materials/analysis tools: JCI. Wrote the paper: BH ML, JCI, LK.